THEME ENV.2012.6.1-1

EUPORIAS

(Grant Agreement 308291)

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

Deliverable D23.5

Report summarising predictability of impact parameters for the agriculture sector

Deliverable Title	Report summarising predictability of impact parameters for the agriculture sector	
Brief Description	This report summarizes general progress made in WP23 relevant to the agricultural sector, and describes the development, application and assessment of crop models for use with seasonal forecasts.	
WP number	23	
Lead Beneficiary	Pete Falloon, Met Office	
Contributors	Met Office: Karina Williams, Pete Falloon WUR: Ronald Hutjes, Iwan Supit, Wietse Franssen, Hester Biemans, Geoffrey Ogutu University of Leeds: Kathryn Nicklin, Andy Challinor	
Creation Date	13/04/2015	
Version Number	1.7	
Version Date	29/04/2015	
Deliverable Due Date	31/04/2015	
Actual Delivery Date		
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**
0.9	15/04/2015	P Falloon	Initial draft, awaiting input from WUR and additional input from University of Leeds.
1.0	16/04/2015	P Falloon	Initial draft including WUR input
1.1	16/04/2015	P Falloon	Initial draft for circulation and comment
1.2	20/05/2015	P Falloon/K Nicklin	Incorporating comments and revisions from University of Leeds
1.3	22/05/2015	P Falloon	Further edits on v1.2
1.4	22/05/2005	P Falloon	Second version for circulation
1.5	23/05/2015	P Falloon	Updated references/papers
1.6	27/04/2015	P Falloon/K Nicklin	Incorporating comments from University of Leeds
1.7	29/04/2015	P Falloon/R Hutjes/G Ogutu/K Williams	Added comments from WUR, final Met Office comments.

EUPORIAS

Table of Contents

List of Tables

List of Figures

1. Executive Summary

Climate variability and extremes can have significant impacts on crops, so the ability to translate seasonal forecasts of meteorological variables such as temperature and rainfall into crop yield forecasts on a seasonal timescale has significant economic and humanitarian benefits. With this in mind, key tasks of EUPORIAS work package 23 have included the development of impacts models, their application to seasonal forecasts, and assessment of skill with respect to their accuracy for key sector specific variables such as crop yields. This report briefly summarizes general progress made in the work package relevant to the agricultural sector, and next describes the development, application and assessment of crop models for use with seasonal forecasts.

Previous findings of work package 23 relevant to this deliverable

Previous reports produced by work package 23 have focused on two aspects of relevance to this deliverable. Firstly, the work package has agreed a modelling strategy and workflow, highlighting 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 to assess the effects of different climate model forcing data, and the effect of impact model initialisation uncertainties
- a common climatology (WFDEI) would be used for reference forcing and general initialisation

Secondly, the work package has found that the impact of initialisation on model outputs differs between types of impact models, differs between seasons and regions, with lead time and with variable; in general the importance of initialisation is greatest for hydrological models, and least for crop models. A range of crop models have been developed and applied to different regions in this work package, including the GLAM crop model (East and West Africa), the JULES-crop model (Global, East Africa, France), and the WOFOST and LPJml models (East Africa).

Key model developments achieved

Several key model developments have been achieved in the work package, with a view to improving their performance for use with seasonal forecasts in the case study regions. For example, a new version of GLAM (version 3) has been produced, incorporating both improved processes (e.g. drainage) and new processes including:

- the ability to run the soil water balance to ‘spin-up’ the soil moisture before planting;
- a soil moisture requirement for successful emergence;
- temperature dependence of leaf growth,
- And a simplified method of simulating maize development.

The default GLAM-maize parameter set was also reviewed and updated. Defining planting dates and crop development parameters is a particular challenge in East Africa due to a lack of data and the spatial variation in climate and cropping seasons. New methods have therefore been developed to define more realistic planting windows and thermal time requirements for local maize varieties.

The initial development of the new JULES-crop model has been completed, and the model structure and global performance is reported in Osborne et al. (2015). Additional developments to the JULES model have included:

- a weather data disaggregator (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),
- enabling spatially varying heights of weather/climate forcing data in the model,
- And an irrigation supply (crop requirement) scheme.

A new WOFOST version has been developed which includes, the effects of fertilisation (N) on crop production in an explicit way. This is essential for East Africa where nutrients are often as limiting as water. Also, nutrient limitations suppress the sensitivity of crop growth to meteorological variability. A set of parameters for nine maize varieties have been optimised for applying the WOFOST crop model to East Africa, reflecting the difference in growing season length across the region.

For LPJml a recoding now allows easier regional applications, still in parallel mode, where previously only global simulations were practically feasible. This greatly facilitates S2D applications that require relatively large ensembles to be run and analysed.

Model assessment and application to seasonal forecasts

Seasonal hindcast datasets were obtained for use as inputs to the crop models, and Wageningen University will make a re-gridded, bias-corrected dataset of the 15-member ECMWF System4 hindcasts available to all partners. An assessment of the impact of the bias corrections applied showed that temperature biases are strongly related to topography effects and that they are perfectly corrected, while precipitation biases are due to a combination of topography and regional effects and the bias correction removes most but not 100% of the biases (notably for drizzle days which are under-corrected).

An investigation of the sources of inter-annual variability in modelled maize yield has been performed with the JULES-crop model, using global runs driven by reanalysis data, with a view to understanding the impact of various approximations in the driving data and initialisation for applications using seasonal hindcasts. These included a) use of driving data at daily rather than sub-daily resolution, and disaggregated internally to the model time-step; b) use of a subset of daily driving data and setting the rest to a daily climatology, and c) initialisation with climatology on the crop sowing date. Each of these approximations could significantly simplify the use of JULES-crop for seasonal crop yield forecasts, due to the reduction in required driving and initialisation data. Using daily forcing data and disaggregating performed well, although care should be taken if modelling the Amazon basin, where the precipitation disaggregation parameters may have been tuned to compensate for biases in JULES.

In most regions outside South-east Asia, Central Brazil, the northern part of the Amazon basin and Bangladesh/East India, the inter-annual variability of the yield from a JULES-crop run in the control configuration is mainly driven by precipitation, which affects the crop via water availability from the soil. As a result, in these regions, it is a good approximation to drive the model with forecast precipitation and leave the other driving data at their climatological values for each day of year. Driving the model with both precipitation and temperature improves the performance in areas with high soil moisture and some further improvement in these areas can be obtained from the addition of downward shortwave radiation.

Perhaps the most important approximation considered for JULES-crop is initialising with climatology on the sowing date, since obtaining accurate initialisation data on the timescales needed for seasonal forecast runs is a particularly significant practical challenge. Building on the earlier findings of the work package, we have further confirmed that this approximation performs well across the majority of maize-growing regions and identified areas where the approximation breaks down. For example, for Maize in France, there was 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. On the other hand, the model performed poorly for maize in Ethiopia even when driven by the full CRU-NCEP simulation, possibly due to the dependency of annual variability in Ethiopian yield on a wider range of factors than meteorological drivers, inaccuracies in the FAO observations for Ethiopia, 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). Nevertheless, assuming the model is driven by seasonal forecasts which have been bias corrected to WFDEI, this study implies that JULES-crop maize simulations runs for EUPORIAS can be reliably initialised using WFDEI climatology.

Taken together, these approximations allow JULES-crop to be driven by seasonal meteorological forecasts where ensembles of bias corrected daily precipitation and daily temperature (and possibly downward short-wave radiation) are available. The reference dataset used for the bias correction can be used to generate the climatology of the initialisation variables and the other driving variables. Since this data is widely available, this provides a practical methodology by which to obtain seasonal crop forecasts with JULES-crop.

Similar to the JULES-crop results, application of the GLAM model to simulate maize yields over East Africa using the WFDEI forcing dataset has found generally poor correlations between simulated and observed yields, with some variation in performance across countries. Further work is needed to understand the reasons for this poor model performance before applying the model to seasonal hindcasts. Additional work could therefore include an in-depth evaluation of GLAM-maize using trial data to make sure that all the key processes affecting maize in East Africa are properly represented, and comparison of simulated yields to sub-national observed statistics to assess whether model performance varies spatially, and if so find out why. It is also possible that weather is not the dominant driver of maize yields in East Africa, or that there are issues with the quality of observed maize yield data.

Forcing GLAM with GloSea5 hindcast data for groundnut over West Africa indicated 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.

Two models that simulate agricultural production over East Africa (LPJml and WOFOST) were used, the latter being specifically implemented and optimised for EUPORIAS. For both models three sets of runs have been performed, using baseline/reference forcing (WFDEI), and using either raw or bias-corrected System4 forcing. In comparing model outputs to observations, several issues were encountered with the observed statistics:

- calendar issues (different databases use different calendars, causing shifts in the series)
- aggregation issues (reported production at different administrative levels is not always consistent due in part to administrative reorganizations)
- only part of inter-annual variability is climate related (check for major alternative causes, e.g. political upheaval)
- records at all levels exhibit considerable gaps.

In addition, the preliminary work illustrated the strong link between both observed and modelled yields and precipitation changes, supporting the argument to not need long spin-up periods for model initialisation for crop models in East Africa proposed by the University of Leeds.

For LPJml, the observed inter-annual variability in crop yields was much higher than modelled especially for Tanzania and Ethiopia, while for Kenya it was more realistic. Validation against sub-national yield data awaits consolidation of these data as crops statistics are not consistent between the various administrative levels. For LPJml driven by System4 hindcasts, compared to simulations driven with WFDEI reference forcing (i.e. not a direct comparison to observed yields), the forecast skill of net primary productivity (NPP) for the first rain (i.e. growing season, March-May) shows considerable forecast resolution, i.e. ability to discriminate between events and non-events, for above and below normal NPP up to lead times of about 3 months and slightly higher resolution for below-normal than for above-normal vegetation productivity. For lead month 1, large areas of the study region show good skill and again, below-normal events have slightly better skill than above-normal events. The skill deteriorates up to lead month 4, though certain areas retain considerable skill (e.g. western Tanzania). Surprisingly, lead month 5 appears to have slightly improved skill, for which presently we have no explanation.

A first evaluation of WOFOST runs forced by the SYS4 hindcasts shows that for crop simulation the use of bias corrected forcing data is essential, since using non-bias corrected hindcasts generally leads to an underestimation of crop yields.

Implications of findings, and recommendations for future work

The work performed under this deliverable has several implications for future work under EUPORIAS:

1. For most of the crop models used here (and most regions relevant to EUPORIAS), precipitation is the dominant driver of inter-annual variability in modeled crop yields. This implies that the models may be more confidently applied with seasonal forecasts in regions where precipitation predictability is strong.
2. The importance of precipitation in driving modeled yields suggests that lengthy spin-up periods are not required, and that in some cases models can be reliably initialized

with short spin-up periods, climatological forcing data on the sowing date, confirming previous work package findings.

3. A number of approximations may be made when driving the JULES-crop model for maize with WFDEI data without significant impacts on model performance in many regions, which are relevant to use with seasonal hindcasts, and likely also relevant to other crop models:
   a. Daily weather data may be used disaggregated to sub-daily values, if necessary
   b. Climatological weather data may be used, apart from for precipitation (which requires ‘actual’ values)
   c. Confirming 2 above, the model may be reliably initialized using WFDEI climatology on the sowing date
4. Bias-correction of seasonal hindcasts to WFDEI climatology is important for crop modelling applications: as crop varieties are optimized to relatively narrow Thermal Time requirements, relatively small temperature biases may result in significant yield biases.
5. The performance of the different crop models applied here varies with region, crop, forcing data and lead-time. For example:
   a. JULES_crop and GLAM perform generally poorly over East Africa for maize when driven by observed climate data, possibly because weather may not be the dominant driver of yield variability, due to inadequate parameter sets to represent local crop types, or due to missing processes (e.g. heat/water stress). Hence further model development and assessment is needed for confident application to seasonal hindcasts.
   b. Overall, model performance was generally best for maize in Kenya, and poorest for Ethiopia
   c. By comparison to simulations driven with reference forcing (WFDEI), LPJml showed generally good performance for maize when driven by seasonal hindcasts for the first rain crop, showing good ability to distinguish high, and particularly low productivity events, and better skill for shorter (1 month) lead times relative to longer lead times (3 months).
6. Availability of high quality, detailed observational datasets of crop yield and development is a significant challenge for both assessing and developing crop models for application to seasonal forecasts, particularly in East Africa. Given the potential value of seasonal crop productivity outlooks to the region, further effort should be invested in obtaining suitable datasets.

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.	X	
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 agricultural 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). 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 how model skill affects 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 including agriculture (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 or grassland 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).

3.1 Approaches for estimating agricultural 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 relevant to agricultural impacts (e.g. extreme temperature and precipitation, storm tracks and cyclones, soil moisture and runoff; 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 or rainfall and crop yield); 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 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 method 3, particularly on complex (offline) impact models.

3.2 General methodological issues for estimating skill in S2D agricultural 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. anticipated crop yield amount and harvest date)
- The probability of particular events (in both categories above) occurring

Furthermore, agricultural 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. This is because 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).

3.2 Producing seasonal hindcasts of crop yield

The ability to forecast crop yield on a seasonal timescale has significant economic and humanitarian benefits (Hansen et al, 2006; Iizumi et al, 2014; Mishra et al,2008). Climate variability and extremes can have significant impacts on crops (e.g. Challinor et al (2014)) and improvements in the seasonal forecast of meteorological variables such as temperature and rainfall (Molteni et al, 2011; MacLachlan et al, 2014; Manzanas et al, 2014) therefore have the potential to improve yield forecasts. However, existing studies of crop model performance focused on seasonal forecast applications show considerable variation depending on the region, scale, processes and crops involved (Hansen et al, 2011; Dessai and Bruno Soares, 2013; Falloon et al, 2013).

Crop model simulations driven by statistically downscaled seasonal hindcasts for European wheat (Palmer et al, 2004; Cantelaube and Terres, 2005), and specifically for wheat in Italy (Marletto et al, 2007) showed that reliable crop yield predictions could be produced using an ensemble multi-model approach and the WOFOST crop model, for instance accurately estimating the probability of positive (1996) and negative (1998) events over the UK. Similarly, Coelho and Costa (2010) used an ensemble of bias-corrected and disaggregated seasonal forecasts to simulate maize yields over Southern Brazil, with the GLAM crop model. The model showed generally good agreement with observed yields, with observed yields within the 95% forecast interval for most years. Using a statistical approach to assess the reliability of global-scale seasonal crop failure hindcasts, Iizumi et al (2013) found that within-season hindcasts generally reproduced inter-annual variability in observed yields in major wheat exporting countries (r^2 =0.56-0.61) better than pre-season hindcasts (r^2 =0.43-0.59). Iizumi et al (2014) modelled global yields of major crops by combining satellite derived NPP data and global agricultural datasets for crop calendar and harvested area and country yield statistics. This statistical model mostly performed well compared to observations, with modelled yields explaining 45-81% of the spatial variation of observed yields in 2000, and correlation coefficients between modelled yield time series and sub-national yield statistics for 1982-2006 in major crop-producing regions generally greater than 0.8. Nicklin et al (2013) found some positive skill in reproducing crop failure of groundnut in West Africa with GLAM driven by seasonal forecast data, and that these results were relatively independent of assumptions on the varieties of groundnut modelled. Mishra et al (2008) ran the SARRA-H crop model at five locations in Burkina Faso and found that, in most cases, incorporating seasonal rainfall forecasts improved sorghum yield predictions made early in the season.

Palmer et al (2004) and Cantelaube and Terres (2005) also found that downscaling seasonal hindcasts improved crop model performance – for example the r^2 value of simulated biomass for the whole of Europe increased from 0.62 to 0.69 with greater regional improvements. On the other hand Challinor et al (2005) found that bias correction of GCM-derived seasonal hindcast data had generally small effects for simulation of groundnut yields in India. Watson and Challinor (2013) found that errors in rainfall data had the largest impact on crop model skill for groundnut in India, mainly because the study region was rainfall limited, while generally the largest yield errors were caused by errors in inter-annual variability in temperature and precipitation. In contrast, for French maize, temperature errors had a stronger influence on yield estimates from both a statistical model and a process-based model than precipitation (Watson et al, 2014).

The ability of crop models to represent inter-annual effects of climate variables also varies depending on the processes represented in the models (Falloon et al. 2014b). For example, high temperature stress around anthesis (the onset of flowering) can have strong impacts on crop yields but not all models include this effect, and responses vary across models that do (Asseng et al, 2013). In general, there is little information of the role of initial conditions in crop model performance on seasonal timescales (Falloon et al, 2013); although hydrological studies have shown that different spin-up approaches may be needed for different impacts (Cosgrove et al, 2003) and different regions.

4. Summary of progress under work package 23

4.1 Implications of previous studies and deliverables D23.1/D23.2/D23.3 for D23.5

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, D23.2 D23.3: Gornall et al. 2013; Hutjes et al. 2013; Falloon et al. 2013), 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, have downloaded the available seasonal hindcasts from the ECOMS portal; and

- Used these models and datasets to assess and improve the predictive skill of impact models by analysing simulations driven by 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 agricultural impact models being used in Work Package 23 (WP23).

Table 1: Agricultural 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
	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(EI)	Global	0.5 degrees	Crop Yield
	WOFOST *WU*	WFD(EI)	Regional	5km	Crop yield

Following on from the workshop on model initialisation (and in D23.1 and D23.2; Gornall et al. 2013; Hutjes et al. 2013) 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. It was assumed that such initialisation would be less important for models of annual vegetation and crops, compared to models of perennial vegetation (e.g. forestry) and hydrological models. This assumption was confirmed by the workshop participants based on their expert judgement and existing literature, and therefore it was agreed (and noted in D23.1/D23.2: Gornall et al. 2013; Hutjes et al. 2013) that sensitivity experiments would only be performed for those models where initialisation is considered critical. Following on from this, D23.3 (Falloon et al. 2014) described recent literature and EUPORIAS studies to assess the importance of initial conditions, with the following key findings:

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.

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 (Hutjes et al. 2013) 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, 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 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.

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.

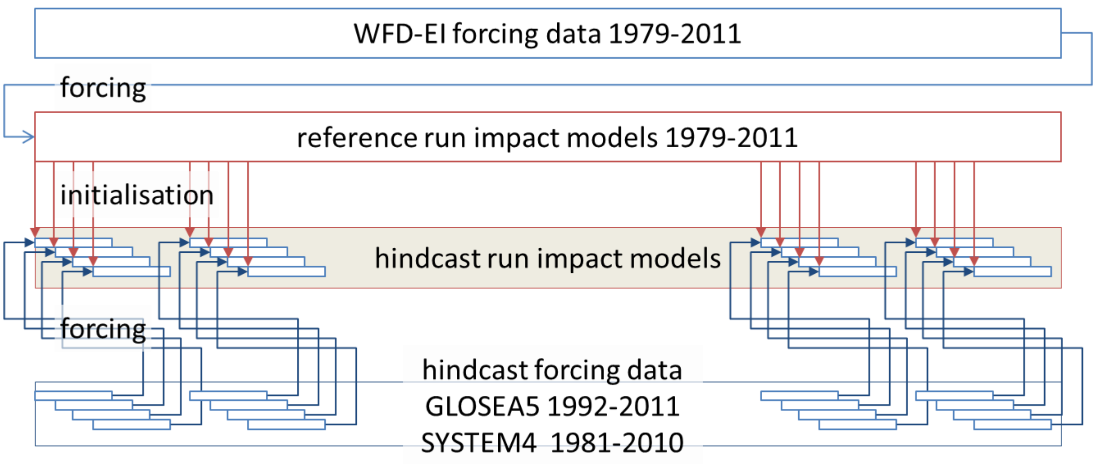

Figure 1: Schematic representation of modelling approach in EUPORIAS WP23

In summary, the EUPORIAS WP23 and WP31 partners 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 Using the GLAM crop model for East Africa (University of Leeds)

The University of Leeds has set up the GLAM-maize crop model for East Africa. The model was then run with Watch Forcing Data Era-Interim (WFDEI) and the simulated yields compared to country-scale observed yields for Ethiopia, Kenya and Tanzania.

A new version of GLAM (version 3) has been produced, in which several processes have been updated (e.g. drainage) and a number of additional processes have been added. Four of these additional processes were considered particularly important for the simulation of maize in East Africa:

1) the ability to run the soil water balance to 'spin-up' the soil moisture before planting,
2) soil moisture requirement for successful emergence,
3) temperature dependence of leaf growth,
4) A simplified method of simulating maize development.

The default GLAM-maize parameter set was also reviewed and updated.

The input data required by GLAM consists of daily weather data, soil hydrological parameters and planting dates. In addition, location-specific parameter values, such as crop development parameters and initial soil moisture, need to be determined. In line with the work package 23 protocol, the WFDEI was downloaded from the ECOMS-UDG and used to produce GLAM weather files. The soil hydrological parameters required are the wilting point, the drained upper limit and the saturation limit. These parameters were determined using information on soil texture from the Global Soil Dataset for Earth System Modelling (GSDE, Shangguan et al 2014).

Defining planting dates and crop development parameters is a particular challenge in East Africa due to a lack of data and the spatial variation in climate and cropping seasons. Most of the rainfall in the region is brought by the Inter Tropical Convergence Zone (ITCZ), which moves over the region from south to north during the first half of the year and from north to south during the second half of the year. This results in a single rainy season in most of Ethiopia and Tanzania, and two rainy seasons in southern Ethiopia, Kenya and northernmost Tanzania. In general, the rainy season associated with the northward movement of the ITCZ is longer and more reliable (Camberlin and Phillippon 2002). It is therefore more important agriculturally and is the initial focus of this study.

The Sacks Crop Calendar dataset (Sacks et al 2010) is a widely used global dataset of planting and harvesting windows for 19 different crops, including maize. However, the spatial resolution of the data is often country scale, which is inadequate for East Africa due to the within-country variation in rainfall regimes and therefore cropping seasons (e.g. Figure 2). In order to define more realistic planting windows, the study area (Ethiopia, Kenya and Tanzania) was divided into 8 separate regions depending on the rainfall regime. Information from the FAO crop calendar data set and from The Livelihoods Atlas for Ethiopia (Ministry of Agriculture, Ethiopia), was used to define an updated planting window for each region (see Table 2 and Figure 2). The WFDEI was used to check that these planting windows are consistent with the start of the rainy season in each location.

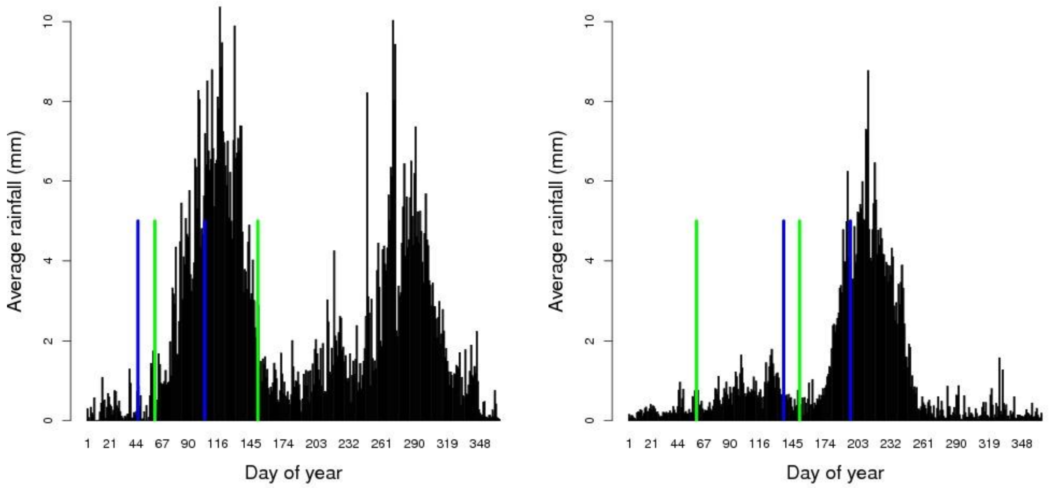

Figure 2: Daily rainfall climatology for two grid cells in Ethiopia created using WFDEI 1979-2012. Green lines - Sacks planting window for Ethiopia. Blue lines – updated planting window.

Table 2: Updated planting window for each region in East Africa. The regions were defined according to the rainfall regime. The first (second) season refers to the first (second) season within the calendar year (Jan-Dec).

Region	Rainfall regime	Updated planting window
North west Ethiopia	Unimodal or second season dominant	Start April to end June
North east Ethiopia	Unimodal or second season dominant	Mid May to mid July
Southern Ethiopia	First season dominant	Mid Feb to mid April
Central Kenya	First season dominant, inland	Mid Jan to end March
Eastern Kenya	First season dominant, coastal	Mid Feb to end April
Western Kenya	Unimodal	Start March to end April
Northern Tanzania	First season dominant	Mid Jan to mid March
Southern Tanzania	Unimodal	Mid Nov to mid Jan

For each grid cell, the updated planting window was divided into a number of 10 day planting windows. A 10 day planting window was excluded if it resulted in an unrealistically short or long cropping season (see Figure 3). Maize yields were simulated for each remaining 10 day planting window to account for variability in the planting date. If a 10 day planting window is unrealistic because there is not enough soil moisture, the crop will fail to emerge. The seasonal yield is the average yield over 10 day planting windows, excluding the crops that fail to emerge.

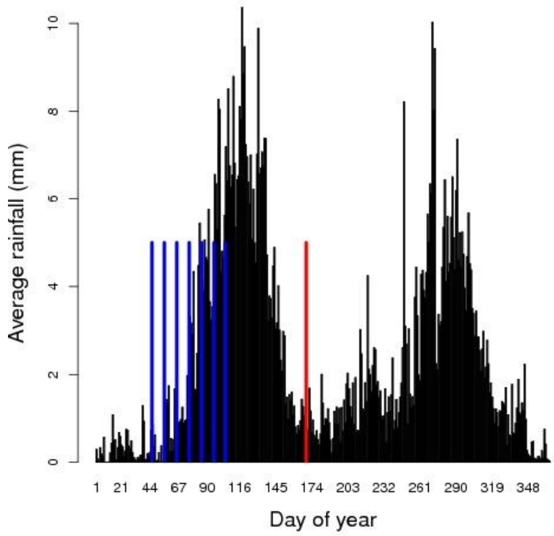

Figure 3: Daily rainfall climatology for a grid cell in Ethiopia. The blue lines show the 10 day planting windows and the red line shows the average end of the (first) rainy season. 10 day planting windows were excluded if there is less than 60 days or greater than 200 days between the planting window and the average end of the rainy season (Nigussie et al 2002, Brink and Belay 2006).

In GLAM-maize, crop development can be simulated by specifying the amount of thermal time that must be accumulated between emergence and flowering (TSUM1) and between flowering and maturity (TSUM2). This simple approach was taken in East Africa due to the limited information on crop development parameters and on how these parameters vary spatially. Values of TSUM1 and TSUM2 for 10 East African maize varieties were kindly provided by Dr. Iwan Supit, Wageningen University. A further 6 varieties were added, resulting in 16 varieties that cover the full range of possible thermal time requirements. It was assumed that farmers will select varieties that mature at the average end of the rainy season. Therefore, for each grid cell and 10 day planting window, the average amount of thermal time between emergence and the average end of the rainy season was calculated and the variety with the closest total thermal time requirement was selected (Figure 4).

The average date on which the rainy season ends was calculated for each grid cell using the WFDEI. In a given year, the end of the rainy season was defined as a dry spell of at least 15 days within the end of season window, with the criteria being relaxed to a dry spell of at least 5 days during the last month of the end of season window (Segele and Lamb 2005). A day was considered dry if the rainfall was less than 1mm. Table 3 gives the end of season window used for each of the 8 regions in East Africa. Figure 5 shows the date calculated as the average end of the rainy season for each grid cell.

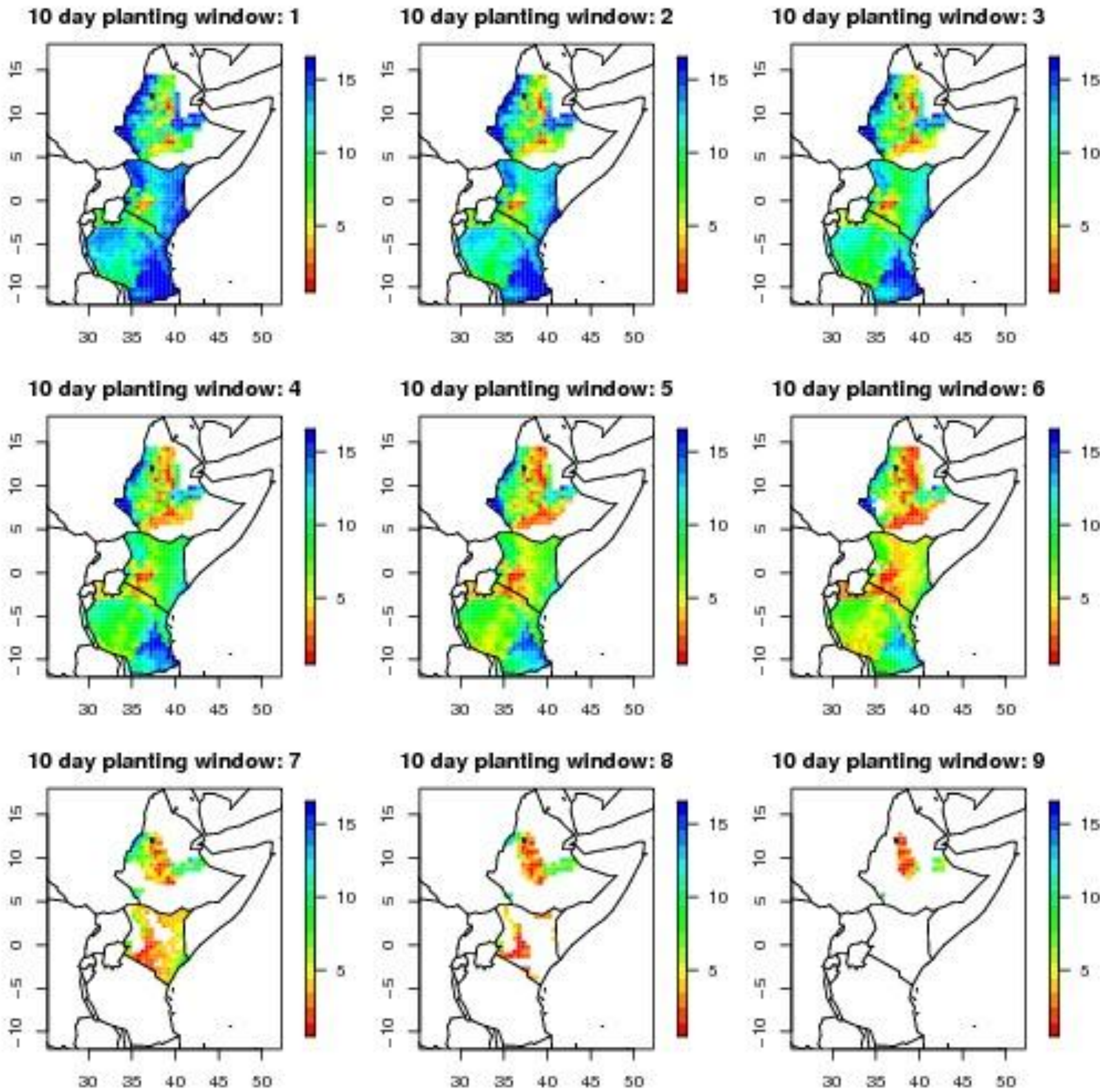

Figure 4: Selected maize variety for each grid cell and 10 day planting window. The total thermal time required from emergence to maturity increases from variety 1 to variety 16. The number of 10 day planting windows varies spatially with a maximum of 9.

Finally, GLAM requires the soil moisture at the start of the planting window to be specified. However, in East Africa, this soil moisture will vary spatially and from one year to the next. Therefore, the GLAM code was modified to allow the soil water balance to be run for a set number of days before the start of the planting window in order to 'spin-up' the soil moisture. The soil moisture at the start of this spin-up period was set to zero and the length of the spin-up period was set to 30 days. As noted in D23.1, D23.2 and D23.3 (Gornall et al. 2013; Hutjes et al. 2013; Falloon et al. 2014), soil moisture initialisation is not likely to have a significant impact in the GLAM model for crop production in East Africa.

Table 3: End of season window for each region in East Africa.

Region	End of season window
North west Ethiopia	Mid August to end Nov
North east Ethiopia	Mid August to end Nov
Southern Ethiopia	Start May to end July
Central Kenya	Start May to end July
Eastern Kenya	Start June to mid Aug
Western Kenya	Start June to end Sept
Northern Tanzania	Mid April to mid July
Southern Tanzania	Start April to end June

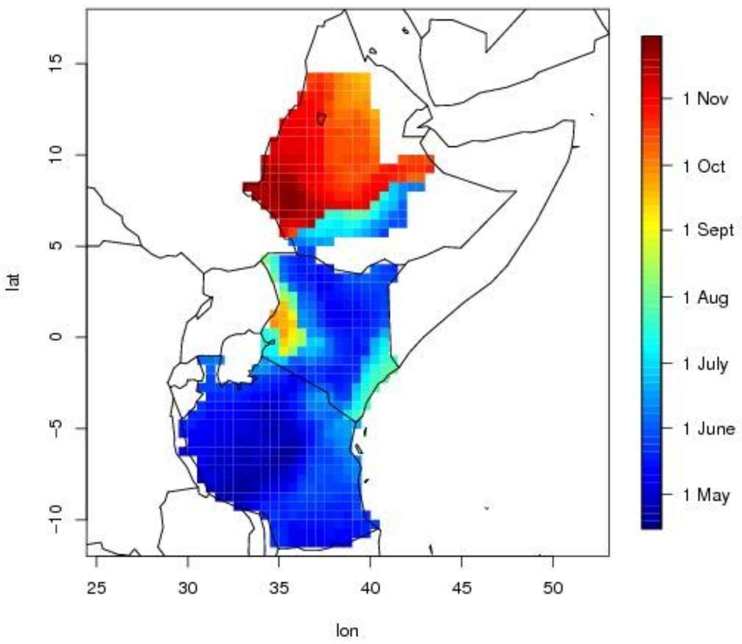

Figure 5: Average end of the rainy season calculated using the WFDEI (1979-2012).

GLAM-maize was used to simulate crop yields on a 0.5° grid across East Africa using the WFDEI. The simulated yields were scaled up to country-level taking into account the maize cultivated area in each grid cell (Monfreda et al 2008). Figure 6 shows the simulated and observed yield time series for each country. The agreement between simulated and observed yields is poor, with correlations between simulated and detrended observed yields of 0.004, 0.14 and 0.12 for Ethiopia, Kenya and Tanzania respectively. Work is currently being done to understand and address the reasons for this poor agreement before GLAM-maize is run with seasonal hindcasts.

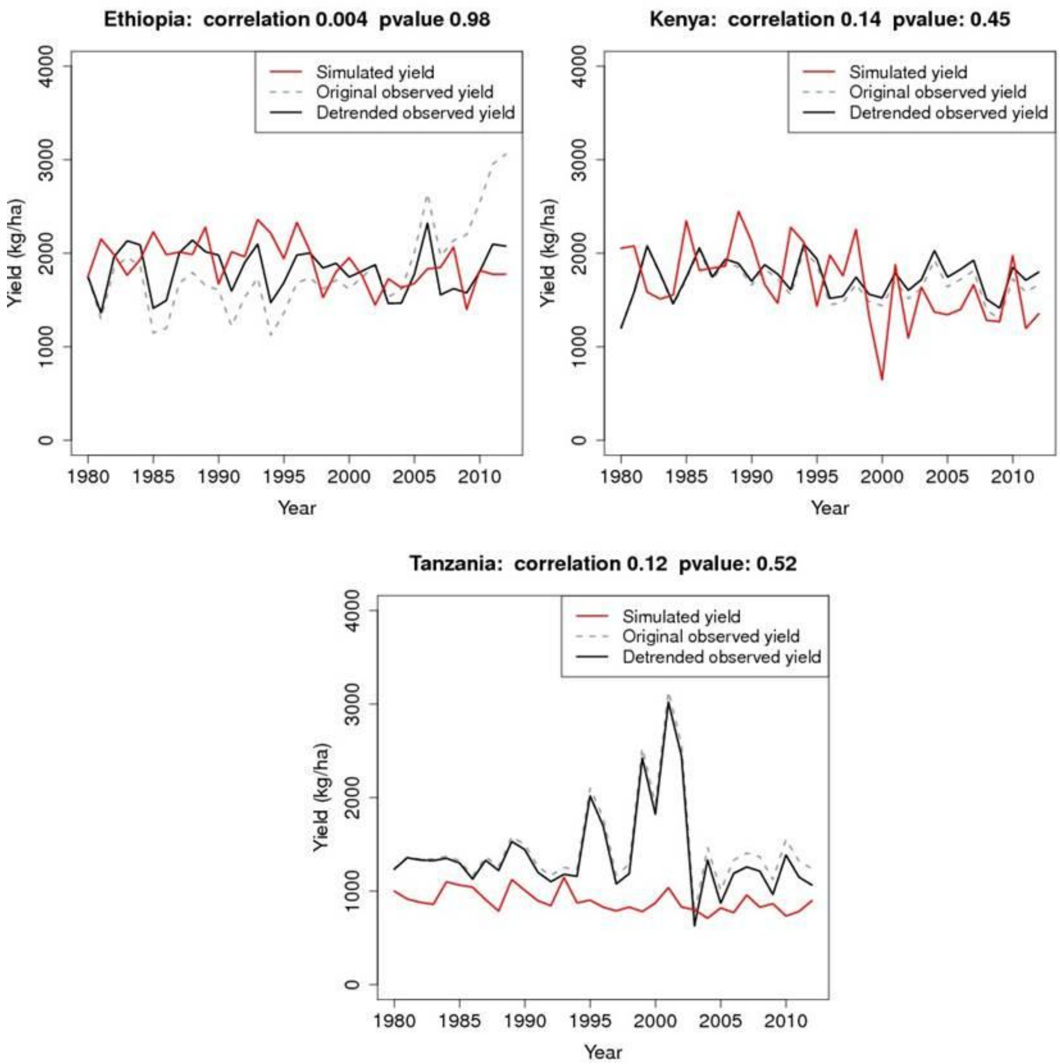

Figure 6: Yields simulated using GLAM-maize and WFDEI for Ethiopia, Kenya and Tanzania (solid red lines). Also shown are the original (grey dashed lines) and detrended (black solid lines) observed yields from FAOstat.

4.2 Development and application of the JULES-crop model (Met Office)

The Met Office contribution to WP23 has included the following activities:

1. development of the JULES-crop crop model, and publication of the model structure and initial performance
2. development of related components of the JULES model system including a weather data disaggregator, enabling spatially varying heights of forcing data in the model, and an irrigation supply (crop requirement) scheme
3. assessment of the impact of weather data disaggregation on model results (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)
4. investigation of the sources of inter-annual variability in modelled maize yield, using global runs driven by reanalysis data, with a view to understanding the impact of various approximations in the driving data and initialisation for applications using seasonal hindcasts
5. collaboration with the University of Leeds on forcing the GLAM crop model with GloSea5 hindcast data for groundnut over West Africa

Here, we focus particularly on activities 1.3, and 4, whilst noting implications of the other activities for the work reported here.

Improved assessment of the impacts of climate change on food and water security requires the development and use of models not only representing each sector (e.g. water, agriculture) but also their interactions. To meet this requirement the Joint UK Land Environment Simulator (JULES) land surface model has been modified to include a generic The JULES-crop model (Osborne et al, 2015; Figure 7) was developed with the dual aim of being able to simulate the impact of weather and climate on crop productivity and the impact that crop-lands have on weather and climate. It is a component of the Joint UK Land Environment Simulator (JULES) (Best et al, 2011; Clark et al, 2011), which is a community land surface model that can be used both online as part of the Met Office Unified Modelling system and offline for impacts studies.

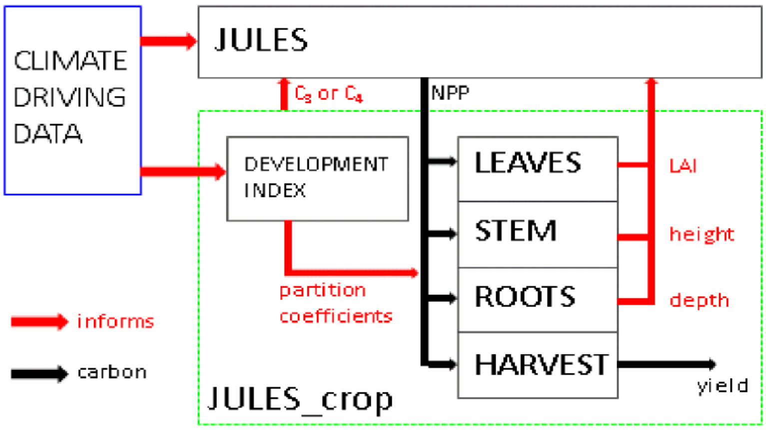

Figure 7: structure of the JULES-crop model.

The new model, JULES-crop, is fully described by Osborne et al. (2015) and has been evaluated at global and site levels for the four globally important crops; wheat, soy bean, maize and rice, driven by NCEP-CRU forcing data (Figure 8). JULES-crop demonstrates skill in simulating the inter-annual variations of yield for maize and soy bean at the global level, and for wheat for major spring wheat producing countries.

The impact of the new parameterisation, compared to the standard configuration, on the simulation of surface heat fluxes is largely an alteration of the partitioning between latent and sensible heat fluxes during the later part of the growing season. Further evaluation at the site level shows the model captures the seasonality of leaf area index and canopy height better than in standard JULES. However, this does not lead to an improvement in the simulation of sensible and latent heat fluxes. The performance of JULES-crop from both an earth system and crop yield model perspective is encouraging; however more effort is needed to develop the parameterisation of the model for specific applications. Key future model developments identified in Osborne et al. (2015) include the specification of the yield gap to enable better representation of the spatial variability in yield.

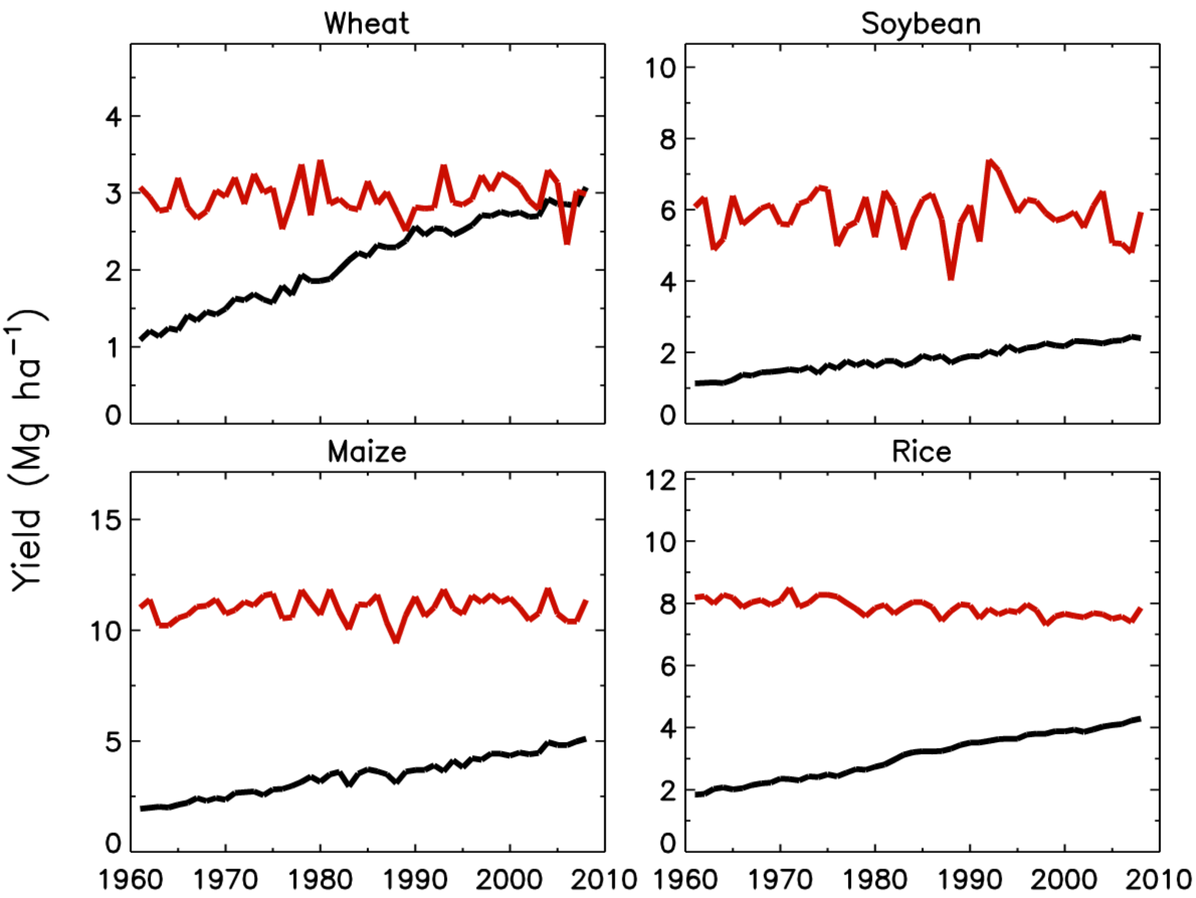

Figure 8. JULES-crop simulated (red) and observed (black) global yield of wheat, soy bean, maize and rice between 1961–2008.

Further assessment of the global-scale performance JULES-crop model has been performed, driving the model with both CRU-NCEP and WFDEI forcing data, as fully reported in D23.3 (Falloon et al. 2014), with a particular focus on the implications of initialising the model with different forcing datasets and differently time averaged data for model results. For Maize in France, there was 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).

The study reported in D23.3 (Falloon et al. 2014) suggests 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.

As reported in D23.3 (Falloon et al. 2014), the Met Office and the University of Leeds also experimented with forcing GLAM with GloSea5 hindcast data for groundnut over West Africa, using the procedure formulated in Nicklin (2013). Preliminary evaluation indicated 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.

Further work, submitted for publication (Williams & Falloon, 2015), involved investigation of the sources of inter-annual variability in modelled maize yield, using global runs driven by reanalysis data, with a view to understanding the impact of various approximations in the driving data and initialisation for applications using seasonal hindcasts. Using JULES-crop on a seasonal timescale introduces a number of technical and scientific issues. Many of these are centred around the availability of data. JULES is driven by a combination of meteorological variables describing air temperature, precipitation, radiation, wind speed, humidity and pressure for each grid box in the model domain, ideally at sub-daily resolution. Output in this format for each ensemble member requires a large amount of storage space and is typically not made externally available by seasonal forecast centres. It is therefore useful to investigate whether the yield variability can be modelled sufficiently well if only a subset of the forcing variables are taken from the seasonal forecast and the others set to climatology, or if the model is forced with daily meteorological data and disaggregated internally to the model time step. To gain a better understanding of the dependence of the yield on the different forcing variables, we investigated the effect of removing water stress and the correlation of the yield with the total grid box precipitation during the crop growing season.

The second data availability issue concerns the variables required to initialise the JULES-crop runs, such as the moisture content of each soil layer (as a fraction of the water content at saturation). Obtaining accurate values for these variables on the start date of the seasonal forecast runs would present a significant practical challenge, as recent observations would be required to estimate these values directly or as input to a reanalysis run. Therefore, we investigated the loss in predictability of yield if the run is started on the sowing date of the crop in that grid box and initialised by the climatological values for that date. This set-up would be simple to reproduce with seasonal forecast forcing that has been bias corrected to a reanalysis dataset, such as those available as part of EUPORIAS, since JULES-crop can be run with this reanalysis dataset to produce a climatology of the initialisation variables. Starting the run before or on the sowing date means that the initialisation of crop variables (e.g. height) is trivial. It has also been suggested that the initialisation of impact model runs driven by seasonal forecasts is more critical for some impacts and regions than others, for example, it may be more critical for water resources in cold regions where snow stores are important than for to dry land cropping (Falloon et al, 2014a).

A number of simulations were performed using JULES-crop as follows. The experimental set-up for the control run follows the global set-up in Osborne et al (2015). The control run (control) was forced by 6 hourly CRU-NCEPv4 climate data (extended to include 2012) as used by the Global Carbon Project (Le Quere et al, 2014), re-gridded to a n96 grid (i.e. grid boxes are 1.875 degrees by 1.25 degrees). The run was from 1960 to 2009 and spun up by repeating the first 10 years five times, before starting the main run. Wheat, soybean, maize and rice were modelled, with the crop parameters listed in Osborne et al (2015). A multi-layer canopy radiation scheme was used, which accounts for direct/diffuse radiation components including sun-flecks (can ran mod=5). The crop sowing dates were taken from Sacks et al. (2010) and extended using nearest neighbour interpolation. The crop tile fractions were taken from Monfreda et al (2008) and other ancillaries taken from HadGEM2-ES (Collins et al, 2011; Jones et al, 2011). Irrigation was not switched on.

We repeated the control run with irrigation demand switched on (irrig), such that, when one of the crops on the grid box had a crop development index (DVI) >-1, water was added to the top two soil levels until the critical soil moisture content θ^c was reached, so that the soil water factor β was 1. Each run repeated the first 10 years five times, to spin up, before starting the main run.

A fully disaggregated run (disagg) was performed as follows. We created daily means and daily temperature ranges from the CRU-NCEPv4 driving data, and used this to drive a JULES run. The internal JULES disaggregator (described in Williams and Clark (2014)) was used to disaggregate this forcing data to the model time step. The run was initialised with the dump file from the beginning of the control main run. All other settings were the same as the control run.

Next, disaggregated runs with some forcing from climatology (sens-*) were produced. In order to investigate the sensitivity to variability in different parts of the driving data, we created daily climatologies of each driving data variable in the full disaggregated run. We then repeated the runs with climatological driving data for all variables apart from certain combinations. The combinations we refer to here are shown in Table 4. Each run had 50 years of spin up (first 10 years five times) before starting the main run (this was particularly important for the sens-T run).

name	sens-T	sens-P	sens-TP	sens-TPR	sens-TPW
mean temperature (T)	x		x	x	x
precipitation (P)		x	x	x	x
downward short-wave radiation (R)				x	
wind speed (W)					x

Table 4. Combinations of driving variables that are allowed to vary in the sens-* runs. Each column is a separate run. All driving variables not marked with an 'x' are set to their daily climatology.

Runs initialised from climatology (init) were created as follows. We created a climatology for the initialisation variables for each day of the year, using daily means outputted from the control run. The model domain was split by sowing date and we performed a separate run for each sowing date for each crop for each year, initialised by the climatology for that sowing date. For example, for Maize, we modelled 77 different sowing dates across the globe for 48 years, which involved 77×48 individual JULES runs. Each run lasted 1 year and the full 6 hourly driving data was used.

Global time series for each crop were constructed from the model output by first masking any grid boxes which had one or more years in which the crop did not reach a DVI of 1.5 or greater or had a yield less than the seed carbon 0.01 kg C m^{-2} (which we assumed was due to a failure on the part of the model or model settings to represent the crops in this grid box) and then weighting according to grid box size and crop tile fraction. Note that the year is defined as January 1st to December 31st (i.e. the model year). In a small fraction of the grid boxes with harvest dates around the end of December/beginning of January, this definition caused issues, as two harvests could fall in one year and none in the next. These points were masked out, as the zero yield appears as a model failure.

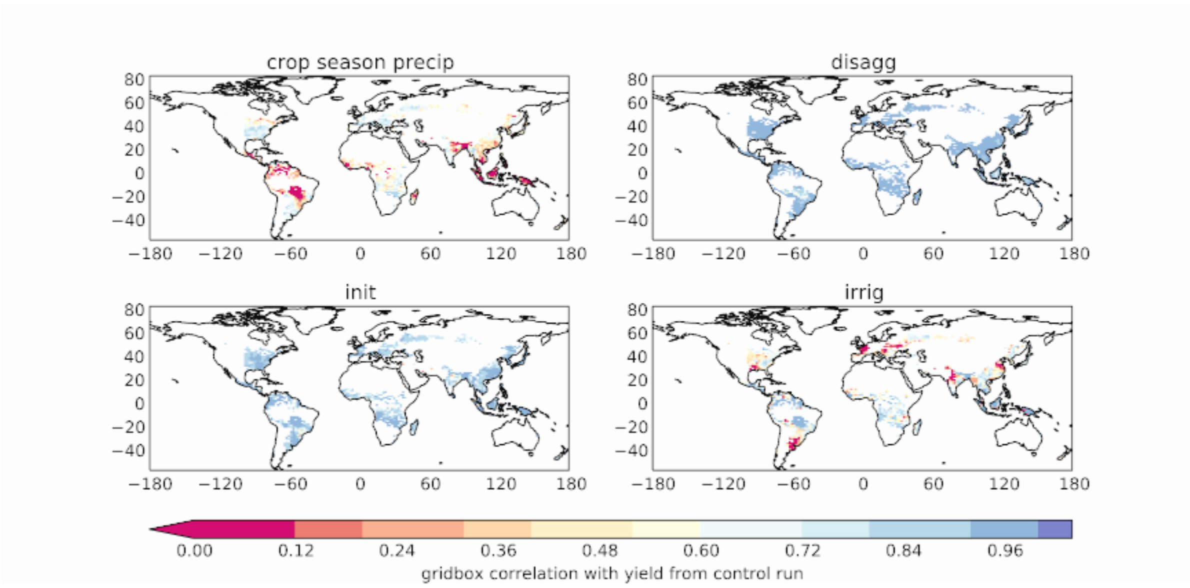

Figure 9. All plots show the correlations with the annual maize yield in the control run for each grid box. Top left: the correlation between yield in control run and crop season precipitation.

Top right, bottom left and bottom right: the correlation between yield in control run and yield in the disagg, init and irrig runs, respectively.

Osborne et al (2015) found that maize yield in the control run had the highest correlation with de-trended global FAO yield observations out of the four crop types modelled; therefore we will present the results for maize only, although we have confirmed that our overall conclusions apply to each of the four crops individually.

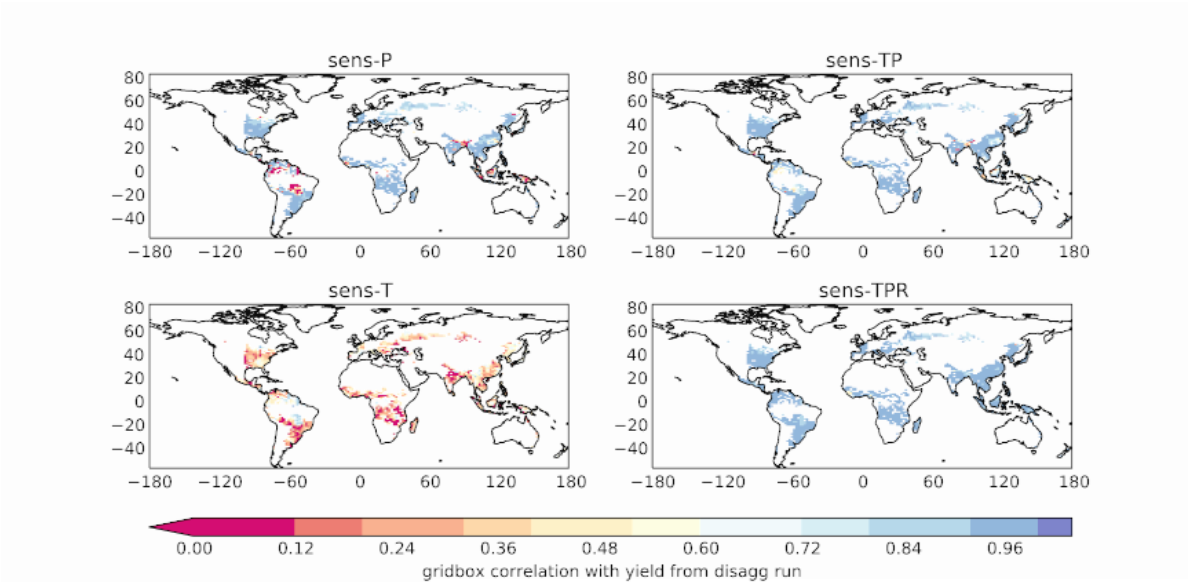

Figure 10. The correlations between the annual maize yield in the control run and the annual maize yield from the sens-P (top left), sens-TP (top right),sens-T (bottom left) and sens-TPR (bottom right) runs for each grid box.

name	mean	standard deviation	global corr with control	global corr with disagg
control	10.6	0.55		
irrig	16.2	0.18	0.48	
init	10.3	0.48	0.91	
disagg	10.2	0.53	0.98	
sens-T	10.7	0.23		0.23
sens-P	10.9	0.42		0.87
sens-TP	11.1	0.51		0.92
sens-TPR	11.1	0.50		0.92
sens-TPW	10.3	0.52		0.96

Table 5. Results from the global JULES-crop runs. First column is the run name, second is the mean maize yield in Mg ha^{-1}, third is the standard deviation of the annual global yield time series in Mg ha^{-1}. The fourth column gives the Pearson correlation coefficient with the global yield in the control run and the fifth column gives the Pearson correlation coefficient with the global yield in the disagg run. All results have been weighted.

Using daily forcing data and disaggregating rather than the full six hourly data results in a slightly lower mean global yield (10.2 Mg ha^{-1} for the disaggregated run, compared to 10.6 Mg ha^{-1}, see Table 5). The global yield time series from the disaggregated run correlates very well with the global yield time series from the control run: the Pearson correlation coefficient is 0.98. Figure 9 (top right) shows the correlation for each grid box, 94% of which are greater than 0.85 (note that there will be spatial correlation between grid boxes and autocorrelation in the time series for each grid box. Also the Pearson correlation coefficient is not resistant to outliers). It is interesting to note that many of the grid boxes with low correlations are in Brazil, a region where the disaggregator has been seen previously to reproduce the climatology of key variables such as evaporation better than runs driven with three hourly data (Williams and Clark, 2014). As discussed in Williams and Clark (2014), since the three hourly data is more representative of the underlying driving data than the disaggregated data, this apparent 'improvement' with the disaggregator is likely to be result of the extra parameters involved in the disaggregation being tuned to compensate for a bias elsewhere in the model. As a result, the maize yield from the disaggregated run actually has a higher correlation with FAO country yield data than the control run for Brazil (not shown here). We can therefore conclude that using daily forcing data and disaggregating is a very good approximation to the control run, for the purposes of looking at variability in the maize yield.

Comparing the control run with the fully irrigated run allows us to determine how much of the modelled yield variability is driven by soil moisture variability. Removing the effect of soil moisture stress increases global NPP as expected, which results in considerably higher global mean yields: maize yield rises from 10.6 to 16.2 Mg ha^{-1} (Table 5). This increase in NPP also has the effect of increasing the number of grid boxes which contribute to the global yield time series, since fewer grid boxes have crops that are harvested prematurely in the model due to lack of growth. Removing soil moisture stress also significantly decreases the (year-to year) standard deviation for maize yield, which has a global standard deviation of 0.55 Mg ha^{-1} in the control run and 0.18 Mg ha^{-1} in the irrigation run.

We also calculated the Pearson correlation coefficient between the control run yield and irrigated run yield for each grid box (Figure 9, bottom right). There was a high correlation coefficient between the two runs in areas with high rainfall during the model maize growing season such as South-east Asia, Central Brazil, the northern part of the Amazon basin and Bangladesh/East India, where we would not expect soil moisture to be a limiting factor in crop growth, even with no irrigation. However, in drier regions, these correlations were much lower, as expected. The percentage of unmasked grid boxes with correlations above 0.85 was just 20% for maize, showing that in most regions, soil moisture variability is an import contribution to the yield variability in the control run.

Moving on from soil moisture to precipitation, we constructed a time series for the crop season precipitation by integrating the rainfall between the sowing and harvesting dates for each crop in each grid box. In many regions, this crop season precipitation index correlates reasonably well with the crop yield for the unmasked grid boxes, particularly outside of South-east Asia, Central Brazil, the northern part of the Amazon basin and Bangladesh/East India, where, as we have already identified, the modelled yield variability does not follow soil moisture variability. Comparing the correlation between modelled yield and detrended observed yields (FAO country scale data) to the correlation between crop season precipitation (between sowing and harvest date) and detrended observed yields (Figure 11) confirms that JULES-crop tends to perform well where there is a strong correlation between the total precipitation between sowing and harvesting date and the (detrended) yield observations.

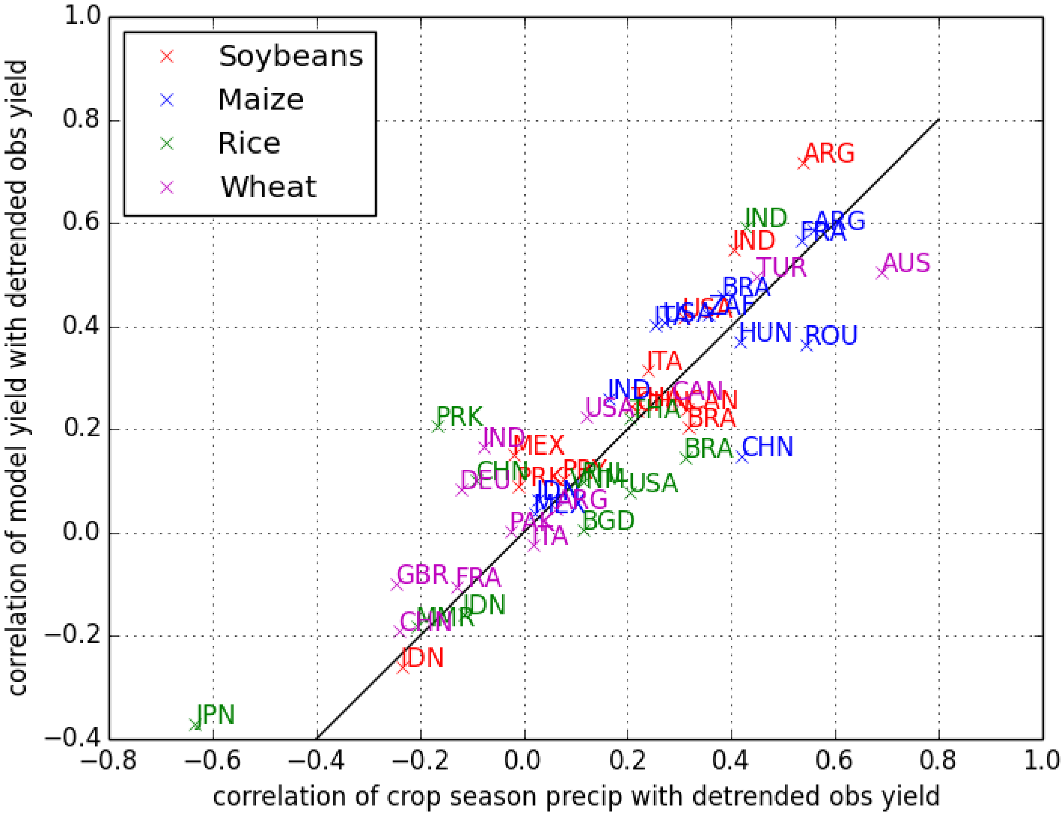

Figure 11, comparing the correlation between modelled yield and detrended observed yields (FAO country scale data) to the correlation between crop season precipitation (between sowing and harvest date) and detrended observed yields.

It is therefore interesting to look at how much of the modelled yield variability can be reproduced if the daily precipitation is used to drive the model, while keeping all other variables at their climatological value for each day of the year (sens-P). A priori we can not assume this will be a good approximation to using the full daily driving data (disagg) from the result for the crop season precipitation index above, since, in the control run, the precipitation is not independent of the other driving data. However, Figure 10 (top left) shows that the sens-P run does indeed correlate well with the disagg run in areas outside of South-east Asia, Central Brazil, the northern part of the Amazon basin and Bangladesh/East India. 74% of the grid boxes shown in Figure 10 have a correlation of 0.85 or more. The correlation between the global yield time series from the sens-P run and the disagg run is 0.87. The sens-P run does have a slightly higher mean global maize yield than the disagg run: 10.9 Mg ha^{-1} as compared to 10.2 Mg ha^{-1}.

If both daily precipitation and daily mean temperature are allowed to vary (run sens-PT), the grid box correlations with the disagg run are much more spatially uniform than for the sens-P run: 81% of the grid boxes have a correlation higher than 0.85 (Figure 9, top right). We can see in Figure 10 (bottom left) that, as expected, the sens-T run does not correlate well with the disagg run in the areas where the sens-P run had a higher correlation. Many of the areas with low correlations in the sens-P run are much higher in the sens-PT run, such as parts of Brazil, Columbia, Bangladesh and Southeast Asia, although these still remain lower than surrounding regions. The correlation between the global maize yield time series in the sens-TP run and the disagg run is 0.92. In general, therefore, driving the model with daily precipitation and mean temperature and using climatology for all other driving variables is a good approximation to make when looking at the inter-annual yield variability across the majority of global maize-growing regions.

In order to improve the approximation further, it may be desirable to additionally allow downward shortwave radiation to vary (sens-PTR) or additionally allow wind speed to vary (sens-PTW). Allowing downward shortwave radiation to vary improves performance (i.e. grid box correlations with the disagg run) in the areas which still have relatively low performance in the sens-PT run i.e. Brazil, Columbia, Bangladesh and Southeast Asia (Figure 10, bottom right). Alternatively, allowing wind speed to vary results in a mean global yield that is closer to the mean global of the disagg run (Table 5).

The final remaining question concerns the model initialisation. The set of runs that are initialised on each sowing date with climatology (init) in general reproduce the spatial distribution of yield from the control run. The global yield is generally lower than in the control run, which results in slightly lower mean global yield (10.3 Mg ha^{-1}) compared to the control run (10.6 Mg ha−1). The correlation between the global maize yield in the init run and the control run is 0.91 and 70% of individual grid boxes have a correlation above 0.85 (Figure 9, bottom left). The correlations are relatively poor in some parts of India, the Congo basin and South/Southeastern Brazil. However, outside these areas, initialising on the sowing date has the potential to be a very useful approximation.

In conclusion, we have investigated a number of possible approximations that could be made when running JULES-crop:

- Use driving data at daily rather than sub daily resolution, and disaggregate internally to the model time step
- Use a subset of daily driving data and set the rest to a daily climatology
- Initialise with climatology on the crop sowing date

Each of these approximations significantly simplifies the use of JULES-crop for seasonal crop yield forecasts, due to the reduction in required driving and initialisation data. With this usage in mind, we have concentrated on the effect of these approximations on the inter-annual variability of the modelled yield.

Using daily forcing data and disaggregating performs the best out of these approximations, although care should be taken if modelling the Amazon basin, where the precipitation disaggregation parameters may have been tuned to compensate for biases in JULES.

We have shown that, in most regions outside South-east Asia, Central Brazil, the northern part of the Amazon basin and Bangladesh/East India, the inter-annual variability of the yield from a JULES-crop run in the control configuration is mainly driven by precipitation, which affects the crop via water availability from the soil. As a result, in these regions, it is a good approximation to drive the model with forecast precipitation and leave the other driving data at their climatological values for each day of year. Driving the model with both precipitation and temperature improves the performance in areas with high soil moisture and some further improvement in these areas can be obtained from the addition of downward shortwave radiation.

Perhaps the most important approximation considered here is initialising with climatology on the sowing date, since obtaining accurate initialisation data on the timescales needed for seasonal forecast runs is a particularly significant practical challenge. We have confirmed that this approximation performs well across the majority of maize-growing regions and identified areas where the approximation breaks down.

Taken together, these approximations allow JULES-crop to be driven by seasonal meteorological forecasts where ensembles of bias corrected daily precipitation and daily temperature (and possibly downward short-wave radiation) are available. The reference dataset used for the bias correction can be used to generate the climatology of the initialisation variables and the other driving variables. Since this data is widely available, this provides a practical methodology by which to obtain seasonal crop forecasts with JULES-crop.

Additional work by the Met Office has performed WFDEI-CRU forced runs of JULES-crop over East Africa. These simulations applied the first and second sowing dates from the Sacks dataset in separate runs. The WFDEI data were applied in three different ways: a) as full three hourly driving data, b) using daily driving data with disaggregation applied, and c) as daily mean temperature and precipitation data with disaggregation applied. One run was also performed with a seasonal hindcast (System 4, 1 year, 1 start date, 1 ensemble member, using bias corrected mean temperature). The model performance when driven by WFDEI-CRU data was generally poor, so results are not shown here and further runs with seasonal hindcasts were not performed since these would likely result in even poorer overall performance. Model yield results were also compared to sub-national data from Geoffrey Ogutu (Wageningen University) for: Tanzania (Mbeya, Rukwa, Mwanza, Shinyanga, Tabora, Kagera, Kigoma, Morogoro, Dodoma, Iringa, Lindi, Mtwara, Ruvuma, Singida, Arusha, Manyara, Kilimanjaro, Mara, Tanga) Kenya (Central, Eastern, Rift Valley, Western, Nyanza, Coast, North-Eastern). The work performed so far has indicated that further development of the JULES-crop model is still required for beneficial application to seasonal hindcasts of maize in East Africa, so ongoing work at the Met Office is now focusing on improving the representation of maize in the model. Literature values and detailed datasets from the Mead FLUXNET sites are being used to improving the current model parameter sets.

4.3 Development and application of the WOFOST and LPJml models (Wageningen University)

Wageningen University has worked on setting up and applying the WOFOST and LPJml models for agricultural forecasts in East Africa (with potential to perform later simulations for Europe). An initial task was to obtain and prepare bias corrected seasonal hindcast datasets for use as inputs to the crop models.

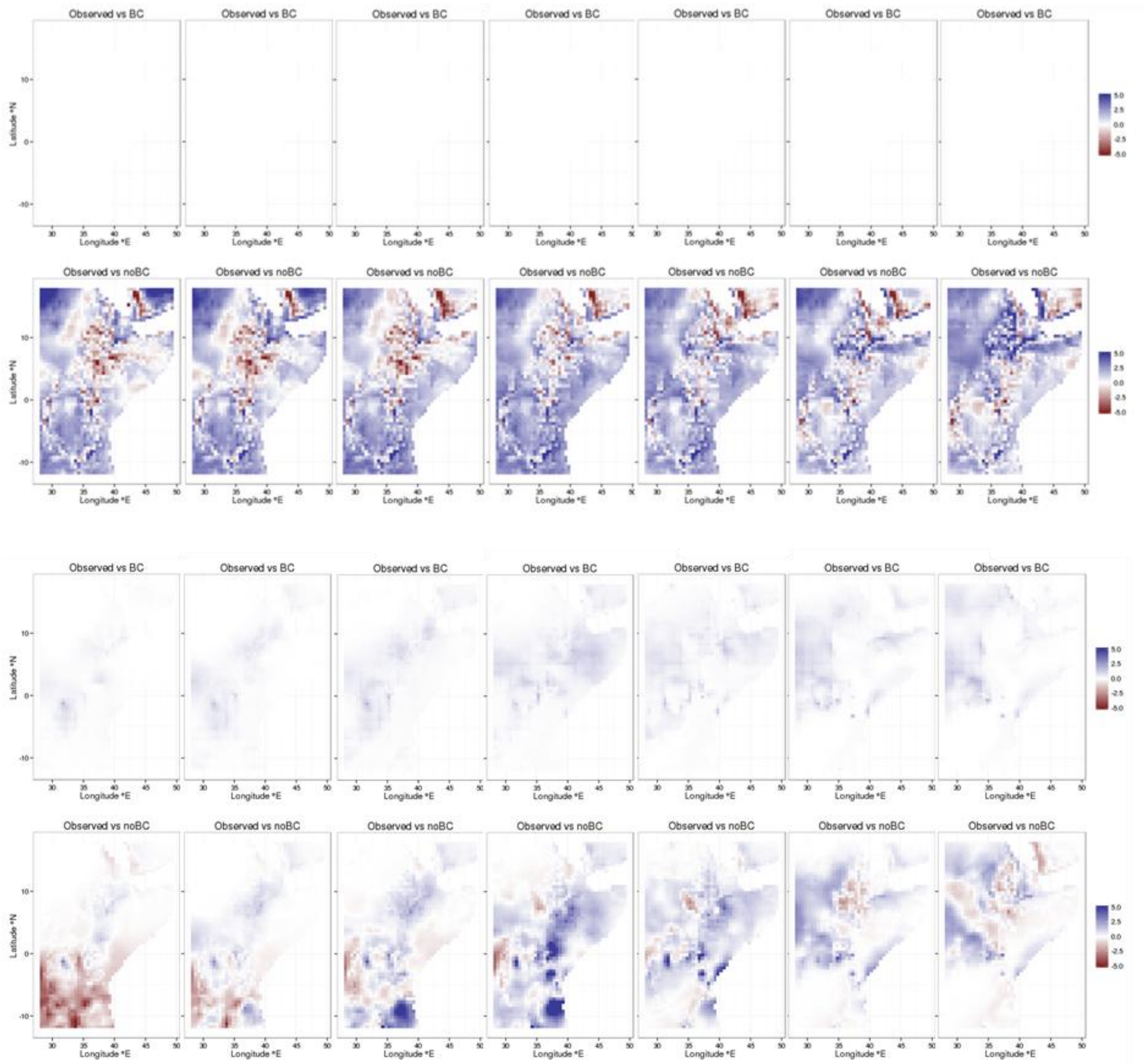

Figure 12: maps of differences between WFDEI, System4 with bias correction (SYS4-BC) and System 4 without bias correction (SYS4-noBC) for temperature (top 2 rows) and precipitation (bottom two rows) for the seven forecast months (left to right) starting in January.

Wageningen University downloaded all required daily variables for the hydrological and agricultural impact models from the ECOMS-UDG data portal for two domains Europe and East Africa for all 15 available members of the ECMWF System4 hindcasts (SYS4). These were bias corrected against the Watch Forcing Data Era Interim (WFDEI), using Quantile-Quantile (QQ) mapping based on the scripts developed in collaboration with the University of Cantabria, using one common set of correction parameters for all 15 members. Then these were interpolated from the native grid (0.75°) to the WFDEI grid (0.5°) land masked and rearranged into one NETCDF file containing all variables for each hind cast run of 7 months. In total 2regions x 2resolutions x 30yrs x 12mo x 15 members = 21600 files. An NCDUMP is given in Appendix 1. These were shared with all EUPORIAS wp23/31 partners. A bias and skill assessment has been performed for selected grid boxes across East Africa (see WP31/32 reports).

The effect of the bias corrections is shown in the example Figure 12. This shows maps of differences between WFDEI, System4 with bias correction (SYS4-BC) and System 4 without bias correction (SYS4-noBC) for temperature (top 2 rows) and precipitation (bottom two rows) for the seven forecast months (left to right) starting in January. It clearly shows that temperature biases are strongly related to topography effects (2nd row) and that these are perfectly corrected (1st row). Figure 13 further illustrates the perfect correlation with temperature, since the red and green bars coincide. Precipitation biases are due to a combination of topography and regional effects and that the bias correction removes most but not 100% of the biases. In particular, the drizzle days are not corrected enough, due to the implementation of the Quantile-Quantile mapping that uses a threshold for corrections of 1mm.

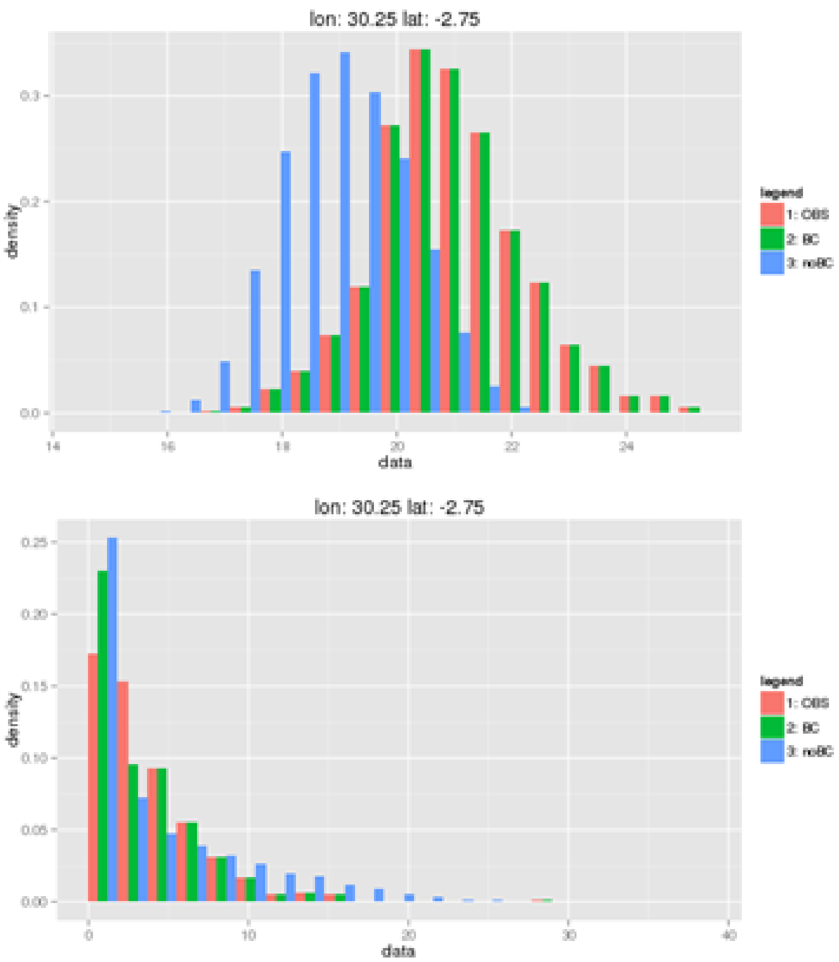

Figure 13: histograms for January/lead month 1, for one particular grid box of temperature (top) and rainfall (bottom) for the observed (red bars), bias-corrected (green) and non-bias corrected (blue) data respectively.

Building on the forcing data preparation described above, Wageningen University has been working with two models that simulate agricultural production: LPJml and WOFOST. The first is a generic global model which in the present context will be analysed for Eastern Africa only, with a single generic maize crop. WOFOST is a specifically for EUPORIAS implemented version for Eastern Africa, also parameterised for maize but using 9 varieties optimised for the different growing season lengths as they vary over the region. A new WOFOST version has been developed which includes, the effects of fertilisation (N) on crop production in an explicit way. This is essential for East Africa where nutrients are often as limiting as water. Also, nutrient limitations suppress the sensitivity of crop growth to meteorological variability. Both presently simulate only a single crop rotation per year for the entire region.

For both models three sets of runs have been performed, all input and output is in netcdf files available:

- baseline/reference run forced by WFDEI
- hindcast with SYS4 forcing, **not** bias corrected (30yrs x 12mo x 15 members = 5400 runs of 7 months each)
- hindcast with SYS4 forcing, **bias** corrected (idem 5400 runs of 7 months each)

For LPJml a recoding now allows easier regional applications, still in parallel mode, where previously only global simulations were practically feasible. This greatly facilitates S2D applications that require relatively large ensembles to be run and analyzed. Previously LPJml was mainly used to study the effects of climate change on crop production, generally requiring (after spin up) a few, century-long runs (for different CMIP models). The 5400 runs listed above are equivalent to more than 31 centuries, justifying the effort to increase computational efficiency.

Initial work reported in D23.3 (Falloon et al. 2014) and the first periodic report performed an initial comparison of model results to national statistics from the FAO only. Several issues were encountered with the observed statistics: 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), only part of inter-annual variability is climate related (check for major alternative causes, e.g. political upheaval), records at all levels exhibit considerable gaps. In addition, the preliminary work illustrated the strong link between both observed and modelled yields and precipitation changes, supporting the argument to not need long initialisation for crop models in East Africa proposed by the University of Leeds.

The full set of runs described above is presently under evaluation. The base line run can be used for an evaluation of average model performance, preferably against observed data, to establish e.g. crop model biases, etc.

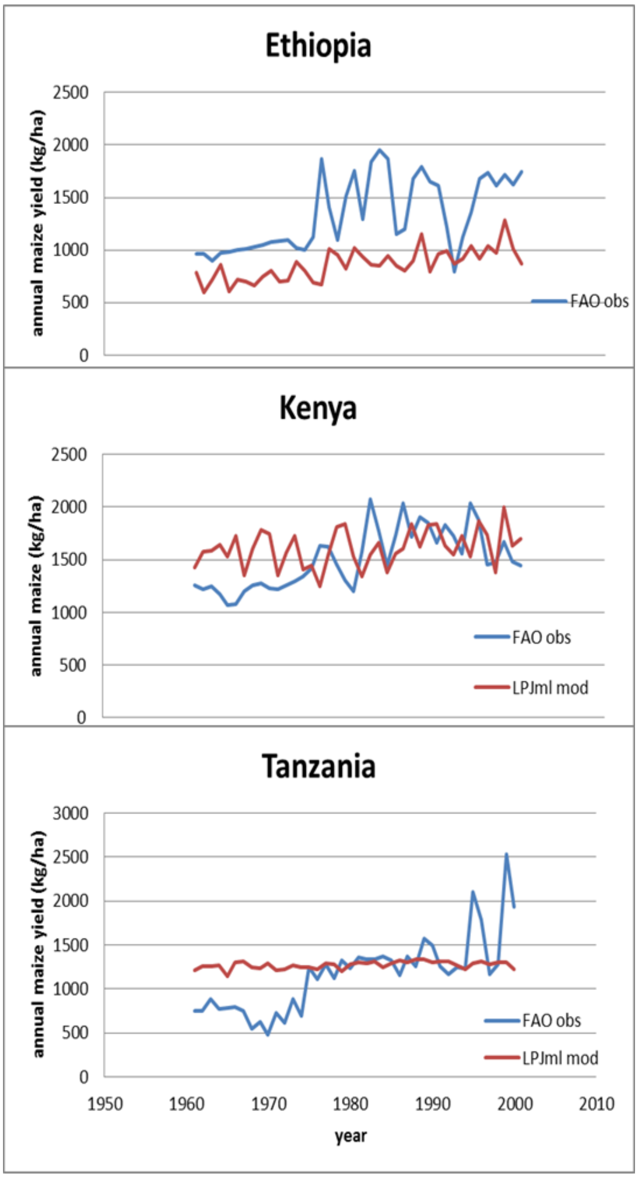

Figure 14: An example evaluation of the LPJml model at country level for East Africa, showing observed and modelled yields for Ethiopia, Kenya and Tanzania.

Figure 14 shows an example evaluation for the LPJml model at country level for East Africa. Since the model uses the same crop parameters throughout the run it does not reflect any yield improvements obtained through the use of better varieties as introduced during the green revolution starting in the seventies of last century. Also the observed inter-annual variability in realized crop yields is much higher than modelled especially for Tanzania (detrended standard deviation 258 versus 39 kg ha^{-1}, respectively for observed and modelled) and Ethiopia (270 versus 103 kg ha^{-1}, respectively) and while for Kenya it is more realistic (206 versus 165 kg ha^{-1}, respectively). Correlations between modelled and observed yields (Table 6) were generally low, especially for Kenya, and were significantly different from zero at the 5% level (n=40) when r>0.312, so those for Ethiopia and Tanzania are correlated, for Kenya were not. Similarly, the Kuiper skill score (KSS) for above normal and below normal yield events showed that for Ethiopia and Tanzania there was some skill, mostly for below normal yields.

Table 6: Model performance statistics for LPJml simulations driven by WFDEI forcing (Figure 14) compared to FAO yield statistics, showing correlation (r) and Kuiper Skill Score (KSS) for above normal (AN) and below normal (BN) yield events. KSS<0.5 indicates no skill, and KSS=1 perfect skill.

Country	r	KSS	
		AN	BN
Kenya	0.155	0.54	0.54
Ethiopia	0.525	0.60	0.72
Tanzania	0.361	0.60	0.66

Validation against sub-national data awaits consolidation of these data as crops statistics are not consistent between the various administrative levels, as shown in Figure 15. It shows for Kenya the yield statistics for maize from county level (NUTS 2) to provincial (NUTS1) to national level (NUTS 0) to FAO reported each time aggregated to the highest level. Especially, prior to about 1985 big discrepancies between yields reported at different levels exist in terms of average yields. Since 1985 this improved, but differences in inter-annual variability still remain. Data consolidation is underway and for case studies we will use the more recent years (since about 1995) to ensure higher quality of observed sub national data.

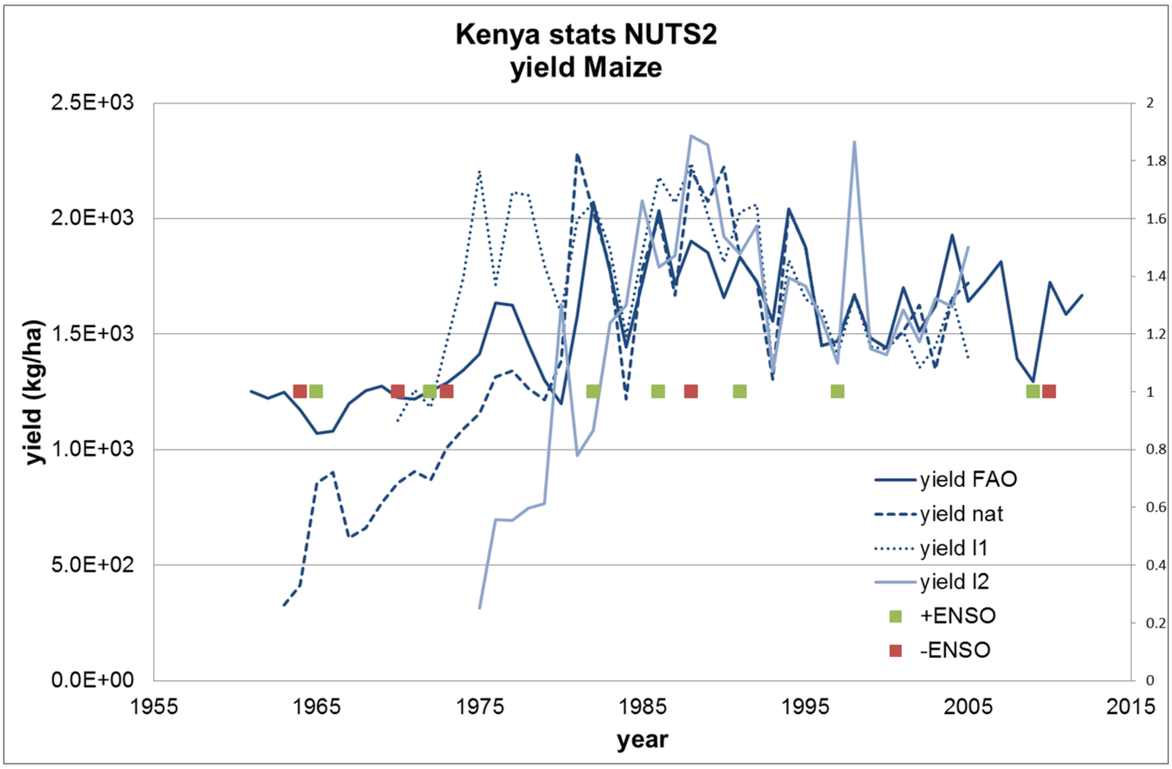

Figure 15: showing for Kenya the yield statistics for maize from county level (' yield l2' = NUTS 2) to provincial ('yield l1' = NUTS1) to national level ('yield nat' = NUTS 0) to FAO (yield FAO) reported each time aggregated to the highest level. Positive (+ENSO) and negative ENSO (-ENSO) phases are also marked on the figure.

The WOFOST models runs at high spatial resolution (i.e. 5km cells) which next are aggregated to either an 0.5 degree grid or to administrative regions (NUTS level 2), as shown is the next few maps. We define 10 varieties for maize with different TT parameters for germination, flowering, etc. For each variety we simulated 12 different sowing dates around the dates given by SAGE and FAO respectively. For each cell we then select that variety/sowing date combination that produced the highest yield for the WFDEI forced run, as shown in Figure 16. The average yield and its inter-annual variability (standard deviation over 32 years) are shown in Figure 17.

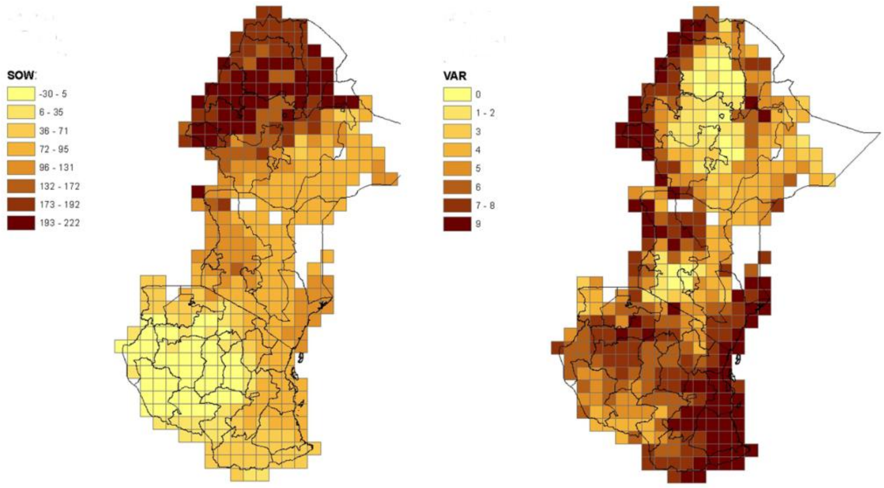

Figure 16: The variety/sowing date combination that produced the highest yield in WOFOST simulations with the WFDEI forced run for each grid cell.

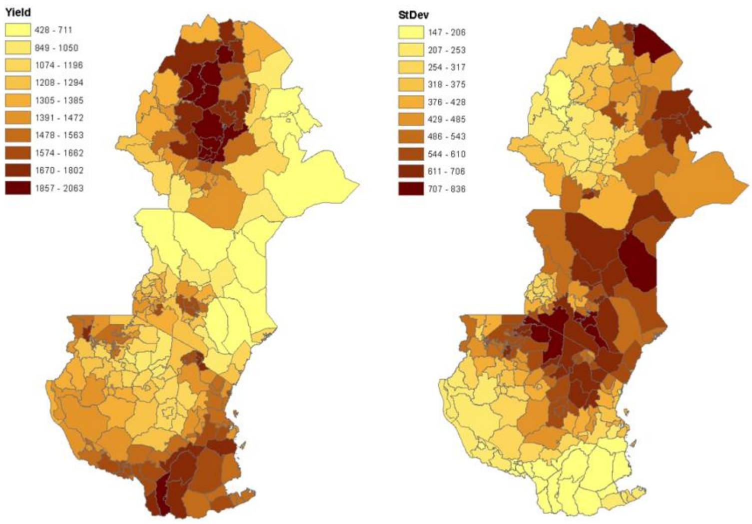

Figure 17: Average yield and its inter-annual variability (standard deviation over 32 years) simulated by WOFOST forced by WFDEI.

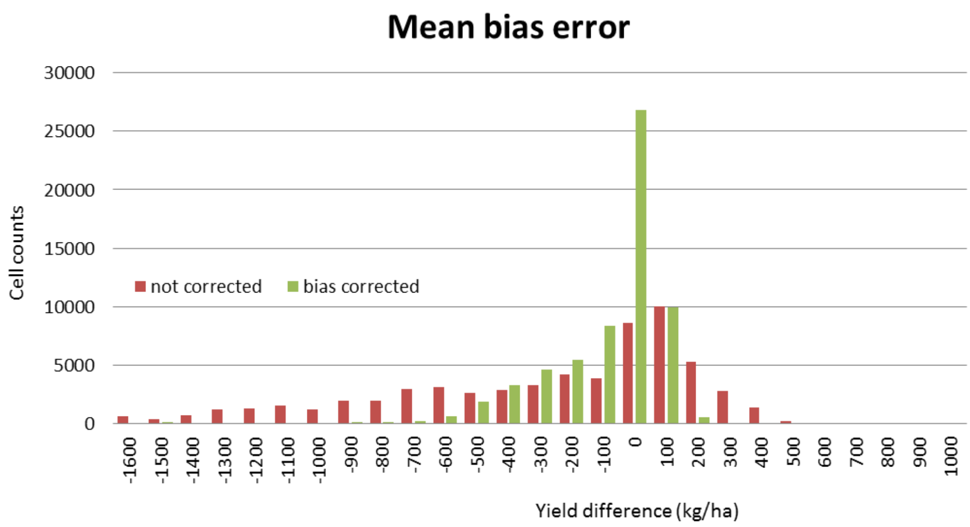

Figure 18: Histograms of mean yield differences over all 65000 WOFOST cells using bias corrected or non bias corrected SYS4 forcing data as compared to the WFDEI forced yields.

A first evaluation of runs forced by the SYS4 hindcasts shows that for crop simulation the use of bias corrected forcing data is essential. Figure 18 shows the histograms of mean yield differences over all 65000 cells using bias corrected or non bias corrected SYS4 forcing data as compared to the WFDEI forced yields. It clearly shows that using non-bias corrected SYS4 data generally leads to an underestimation of crop yields as simulated by WOFOST. The reason is that crop varieties are optimized to relatively narrow Thermal Time requirements for different phenological stages. Thus, relatively small temperature biases may result in significant yield biases.

With the hindcast runs we can do a probabilistic evaluation of seasonal forecast skills. Here we will present only a few tentative results for randomly selected points.

For LPJmL we arbitrarily compare the hindcast SYS4 simulated NPP (net primary productivity) and runoff to that of the base run WFDEI forced NPP and runoff, using some typical probabilistic skill score metrics. Scoring against real observations will follow later as especially sub-national crop yield statistics need further quality assessment and consolidation before use. A further limitation of even good real data is that for crops we can validate only one number (i.e. yield) per year for some administrative units, of the discharge of a few stations only. Validation against the baseline run allows evaluation of more variables at pixel level. Such an analysis may teach us more about the why of any (lack of) forecast skills. Performance against real data will probably be lower than those reported next.

Figure 19 shows ROC and attribution diagram for a pixel in the lower left of our domain. The forecast skill of NPP for the first rain -, i.e. growing season (MAM) is presented as forecasted at the start of that season in March (top) down to a forecast lead time of five months, i.e. a forecast started in November the previous year (bottom). The left column shows the ROC plots and AROC scores (in the respective legends) for above normal (AN, black) near normal (NN, blue) or below normal (BN, red) vegetation productivity. It shows considerable forecast resolution, i.e. ability to discriminate between events and non-events, for above and below normal NPP up to lead times of about 3 months and slightly higher resolution for BN than for AN vegetation productivity. This is confirmed in the other two columns showing attribute diagrams for AN (middle) and BN (right), i.e. how well do the predicted probabilities of an event correspond to their observed frequencies. Points within the grey area are considered Ok and the closer to the diagonal the better. This suggests that especially the BN events are well attributed and show some reliability for lead times up to three months.

These metrics can be combined, following the Murphy ‘73 decomposition, in the Brier Score (BS) which can then be mapped. Lower scores are better here. An example is given on the last page, again for NPP. The left column are maps for above normal (AN) events and the right columns for below normal (BN) events. For lead month 1 (top) it shows large areas with skills below 0.2 which can be considered good skill; BN events have slightly better skill than AN events. This gradually deteriorates up to lead month 4, though certain areas retain considerable skill, like e.g. western Tanzania. Surprisingly, lead month 5 seems slightly better again, for which presently we have no explanation.

As noted before these is just a very brief tentative first probabilistic evaluation of forecast skills. The scatter in all the graphs and maps is still a bit high. Or phrased alternatively, the statistics are a bit poor for 10 bins for a single pixel of a 15 member ensemble. Pooling more pixels (e.g. in to NUTS regions or into climatically more homogeneous zones) and/or reducing the number of bins will reduce the scatter and may thus produce more informative results. Also comparing these results to the skill of just using a climatological average as forecast, or to statistical forecasts, will better show the added value of the computationally very expensive modelling chain used here. The Brier Skill Score provides such a relative metric ($BSS = 1 - BS_{model}/BS_{climatology}$), as illustrated in Figure 20.

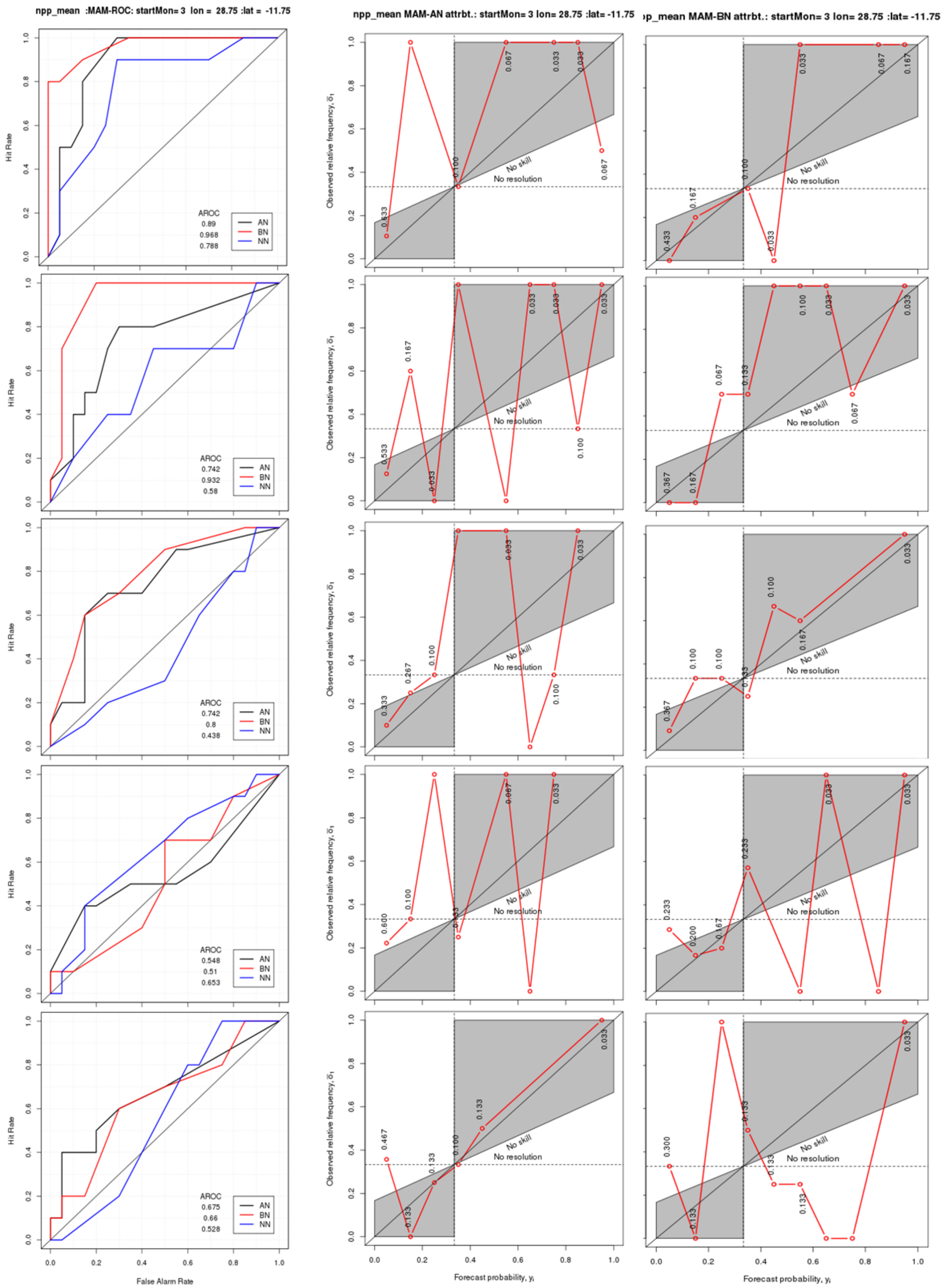

Figure 19: ROC and attribution diagram of SYS4 forced LPJml simulations for a pixel in the lower left of our domain. The forecast skill of crop NPP for the first rain -, i.e. growing season (MAM) is presented as forecasted at the start of that season in March (top) down to a forecast lead time of five months, i.e. a forecast started in November the previous year (bottom). The left column shows the ROC plots and AROC scores (in the respective legends) for above normal (AN, black) near normal (NN, blue) or below normal (BN, red) vegetation productivity.

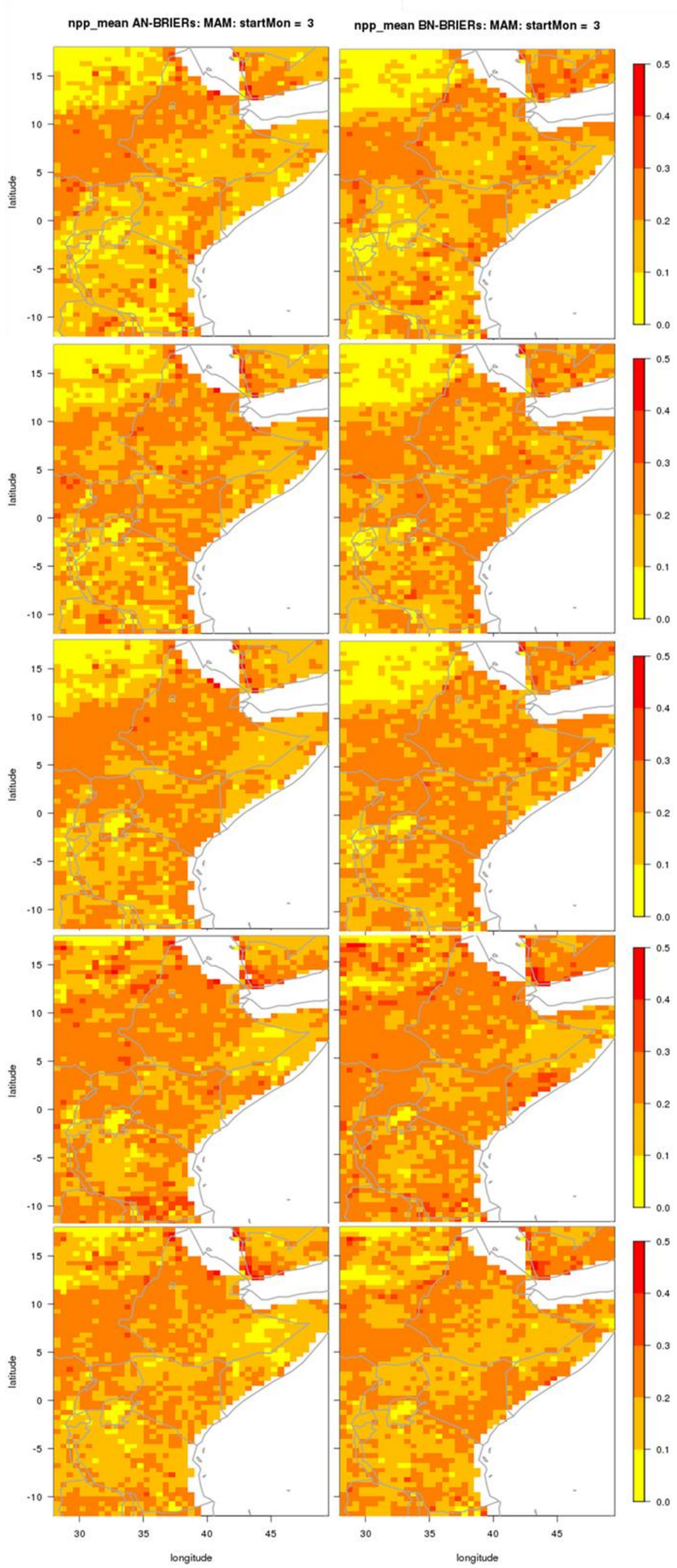

Figure 20: The Brier Skill Score (BSS = 1 – $BS_{model}/BS_{climatology}$) of SYS4 forced LPJmL crop yield simulations, comparing model results to the skill of just using a climatological average as forecast.

1.3 List of publications

Pete Falloon, David Fereday, Nicky Stringer, Karina Williams, Jemma Gornall, Emily Wallace, Rosie Eade, Anca Brookshaw, Joanne Camp, Richard Betts, Rutger Dankers, Kathryn Nicklin, Michael Vellinga, Richard Graham, Alberto Arribas and Craig MacLachlan (2013) Assessing Skill for Impacts in Seasonal to Decadal Climate Forecasts (editorial), J Geol Geosci 2: e111, doi: 10.4172/2329-6755.1000e111.

Osborne, T., Gornall, J., Hooker, J., Williams, K., Wiltshire, A., Betts, R., and Wheeler, T.: JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator, Geosci. Model Dev., 8, 1139-1155, doi:10.5194/gmd-8-1139-2015, 2015.

Karina Williams and Pete Falloon (2015) Sources of inter-annual yield variability in JULES-crop and implications for forcing with seasonal weather forecasts. Geoscientific Model Development Discussions (submitted).

Williams K, Clark D (2014) Disaggregation of daily data in JULES. Tech. Rep. 96, Met Office Hadley Centre, URL http://www.metoffice.gov.uk/media/pdf/2/j/HCTN96.pdf

5. Lessons Learnt

The assessment of crop model performance with WFDEI datasets has highlighted several additional areas for further model assessment and development, as noted in detail in the report. This has been an important precursor to driving the models with seasonal hindcasts, since if the models perform poorly when driven by “best estimates” of observed climate data, they would be expected to perform even less well when driven by seasonal hindcasts. The ability of the different crop models used here to represent observed variability in yields varies across models, regions, crops and lead-times. In some cases, the availability and quality of observational data has proven a challenge to making robust assessments of model performance – particularly the case for East Africa where access to quality controlled, detailed datasets of crop yield and crop development and growth is difficult.

6. Links Built

The work presented here as strong links to the ECOMS (EUPORIAS-SPECS) data portal as a source of driving data for the crop models, and has benefitted from the active involvement of partners at the University of Cantabria. There are also links with the prototypes in Work package 42, most notably the LEAP food security outlook for East Africa, and the land management prototype for the UK. Ongoing work related to this deliverable will be carried out under the strongly related Work package 31.

References

Asseng S, Ewert F, Rosenzweig C, Jones JW, Hatfield JL, Ruane AC, Boote KJ, Thorburn PJ, Rotter RP, Cammarano D, Brisson N, Basso B, Martre P, Aggarwal PK, Angulo C, Bertuzzi P, Biernath C, Challinor AJ, Doltra J, Gayler S, Goldberg R, Grant R, Heng L, Hooker J, Hunt LA, Ingwersen J, Izaurralde RC, Kersebaum KC, Muller C, Naresh Kumar S, Nendel C, O'Leary G, Olesen JE, Osborne TM, Palosuo T, Priesack E, Ripoche D, Semenov MA, Shcherbak I, Steduto P, Stockle C, Stratonovitch P, Streck T, Supit I, Tao F, Travasso M, Waha K, Wallach D, White JW, Williams JR, Wolf J (2013) Uncertainty in simulating wheat yields under climate change. Nature Clim Change 3(9):827–832, DOI 10.1038/nclimate1916, URL http://dx.doi.org/10.1038/nclimate1916

Best MJ, Pryor M, Clark DB, Rooney GG, Essery, M´enard CB, Edwards JM, Hendry MA, Porson A, Gedney N, Mercado LM, Sitch S, Blyth E, Boucher O, Cox PM, Grimmond CSB, Harding RJ (2011) The Joint UK Land Environment Simulator (jules), model description Part 1: Energy and water fluxes. Geoscientific Model Development 4(3):677–699, DOI 10.5194/gmd-4-677-2011, URL http://dx.doi.org/10.5194/gmd-4-677-2011

Camberlin P and Phillippon N (2002) The East African March-May Rainy Season: Associated Atmospheric Dynamics and Predictability over the 1968-1997 Period. Journal of Climate 15:1002-1019

Cantelaube P, Terres jM (2005) Seasonal weather forecasts for crop yield modelling in europe. Tellus A 57(3):476–487, DOI 10.1111/j.1600-0870.2005.00125.x, URL http://dx.doi.org/10.1111/j.1600-0870.2005.00125.x

Challinor AJ, Watson J, Lobell DB, Howden SM, Smith DR, Chhetri N (2014) A meta-analysis of crop yield under climate change and adaptation. Nature Clim Change 4(4):287–291, DOI 10.1038/nclimate2153, URL http://dx.doi.org/10.1038/nclimate2153

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(3):498–512, DOI 10.1111/j.1600-0870.2005.00126.x, URL http://dx.doi.org/10.1111/j.1600-0870.2005.00126.x

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.

Clark DB, Mercado LM, Sitch S, Jones CD, Gedney N, Best MJ, Pryor M, Rooney GG, Essery RLH, Blyth E, Boucher O, Harding RJ, Huntingford C, Cox PM (2011) The Joint uk Land Environment Simulator (jules), model description Part 2: Carbon fluxes and vegetation dynamics. Geoscientific Model Development 4(3):701–722, DOI 10.5194/gmd-4-701-2011, URL http://dx.doi.org/10.5194/gmd-4-701-2011

Coelho CAS, Costa SMS (2010) Challenges for integrating seasonal climate forecasts in user applications. Current Opinion in Environmental Sustainability 2(5-6):317–325, DOI 10.1016/j.cosust.2010.09.002, URL http://dx.doi.org/10.1016/j.cosust.2010.09.002

Collins WJ, Bellouin N, Doutriaux-Boucher M, Gedney N, Halloran P, Hinton T, Hughes J, Jones CD, Joshi M, Liddicoat S, Martin G, O'Connor F, Rae J, Senior C, Sitch S, Totterdell I, Wiltshire A, Woodward S (2011) Development and evaluation of an Earth-system model hadgem2. Geoscientific Model Development 4(4):1051–1075, DOI 10.5194/gmd-4-1051-2011, URL http://dx.doi.org/10.5194/gmd-4-1051-2011

Cosgrove BA, Lohmann D, Mitchell KE, Houser PR, Wood EF, Schaake JC, Robock A, Sheffield J, Duan Q, Luo L, Higgins RW, Pinker RT, Tarpley JD (2003) Land surface model spin-up behavior in the north american land data assimilation system (NLDAS). J Geophys Res 108(D22):8845+, DOI 10.1029/2002jd003316, URL http://dx.doi.org/10.1029/2002jd003316

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.

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

Dessai S, Bruno Soares M (2013) Systematic literature review on the use of seasonal to decadal climate and climate impacts predictions across European sectors. Tech. Rep. D12.1, University of Leeds

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, Challinor A, Dessai S, Hoang L, Johnson J, Koehler AK (2014b) Ensembles and uncertainty in climate change impacts. Frontiers in Environmental Science 2:33+, DOI 10.3389/fenvs.2014.00033, URL http://dx.doi.org/10.3389/fenvs.2014.00033

Falloon P, Fereday D, Stringer N, Williams K, Gornall J, Wallace E, Eade R, Brookshaw A, Camp J, Betts R, Dankers R, Nicklin K, Vellinga M, Graham R, Arribas A, MacLachlan C (2013) Assessing skill for impacts in seasonal to decadal climate forecasts. Geology & Geosciences 2(3), URL http://www.omicsgroup.org/journals/assessing-skill-for-impacts-inseasonal-to-decadal-climate-forecasts-2329-6755-2-1000e111.php?aid=18791

Falloon P, Buontempo C, Ceron JP, Dubus L, Greuell W, Hutjes R, Jonsson AM, Lowe R, Manzanas RG, San Martin MT, Newton P, Nicklin K, Pechlivanidis I, Pouget L, Supit I, Williams K (2014a) Report on initialisation of impacts models for seasonal predictions. Tech. Rep. D23.3, EUPORIAS, URL http://www.euporias.eu/system/files/D23.3 Final.pdf

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.

Falloon P, Williams K, Buontempo C, Newton P, Hutjes R, Greuell W, Supit I, Nicklin K, Pechiavandis I, García-Manzanas R, Jönsson A-M, Lowe R, Pouget L, Ceron J-P, Dubus L, Tuni San Martin M (2014) EUPORIAS (Grant agreement 308291) Deliverable 23.3, Report on initialisation of impacts models for seasonal predictions, Met Office Hadley Centre, Exeter, UK.

Gornall J, Williams K, Falloon P, Buontempo C, Dankers R, Arribas A, Wiltshire A, Newton P, Hutjes R, Nicklin K, Pechlivanidis I, García-Manzanas R, Jönsson A-M, Lowe R, Pouget L (2013) EUPORIAS (Grant agreement 308291) Deliverable 23.1, Workshop on initialisation of impacts models for seasonal predictions, Met Office Hadley Centre, Exeter, UK

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

Hansen JW, Challinor A, Ines A, Wheeler T, Moron V (2006) Translating climate forecasts into agricultural terms: advances and challenges. Climate Research33:27–41, DOI 10.3354/cr033027, URL http://dx.doi.org/10.3354/cr033027

Hansen JW, Mason SJ, Sun L, Tall A (2011) Review of seasonal climate forecasting for agriculture in Sub-Saharan Africa. Experimental Agriculture 47:205–240, DOI 10.1017/s0014479710000876, URL http://dx.doi.org/10.1017/s0014479710000876

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.

Hewitt C, Buontempo C, Newton P (2013) Using climate predictions to better serve society's needs. Eos Trans AGU 94(11):105–107, DOI 10.1002/2013eo110002, URL http://dx.doi.org/10.1002/2013eo110002

Hutjes R, Gornall J, Williams K, Falloon P, Buontempo C, Dankers R, Arribas A, Wiltshire A, Newton P, Nicklin K, Pechlivanidis I, García-Manzanas R, Jönsson A-M, Lowe R, Pouget L (2013) EUPORIAS (Grant agreement 308291) Deliverable 23.2, Report of the workshop on initialisation of impacts models for seasonal predictions, Wageningen University, Wageningen, Netherlands.

Iizumi T, Sakuma H, Yokozawa M, Luo JJ, Challinor AJ, Brown ME, Sakurai G, Yamagata T (2013) Prediction of seasonal climate-induced variations in global food production. Nature Clim Change 3(10):904–908, DOI 10.1038/nclimate1945, URL http://dx.doi.org/10.1038/nclimate1945

Iizumi T, Yokozawa M, Sakurai G, Travasso MI, Romanenkov V, Oettli P, Newby T, Ishigooka Y, Furuya J (2014) Historical changes in global yields: major cereal and legume crops from 1982 to 2006. Global Ecology and Biogeography 23(3):346–357, DOI 10.1111/geb.12120, URL http://dx.doi.org/10.1111/geb.12120

Jones CD, Hughes JK, Bellouin N, Hardiman SC, Jones GS, Knight J, Liddicoat S, O'Connor FM, Andres RJ, Bell C, Boo KO, Bozzo A, Butchart N, Cadule P, Corbin KD, Doutriaux-Boucher M, Friedlingstein P, Gornall J, Gray L, Halloran PR, Hurtt G, Ingram WJ, Lamarque JF, Law RM, Meinshausen M, Osprey S, Palin EJ, Chini LP, Raddatz T, Sanderson MG, Sellar AA, Schurer A, Valdes P, Wood N, Woodward S, Yoshioka M, Zerroukat M (2011) The Hadgem2-es implementation of CMIP5 centennial simulations. Geoscientific Model Development 4(3):543–570, DOI 10.5194/gmd-4-543-2011, URL http://dx.doi.org/10.5194/gmd-4-543-2011

Le Qu´er´e C, Peters GP, Andres RJ, Andrew RM, Boden TA, Ciais P, Friedlingstein P, Houghton RA, Marland G, Moriarty R, Sitch S, Tans P, Arneth A, Arvanitis A, Bakker DCE, Bopp L, Canadell JG, Chini LP, Doney SC, Harper A, Harris I, House JI, Jain AK, Jones SD, Kato E, Keeling RF, Klein Goldewijk K, Kortzinger A, Koven C, Lef`evre N, Maignan F, Omar A, Ono T, Park GH, Pfeil B, Poulter B, Raupach MR, Regnier P, R¨odenbeck C, Saito S, Schwinger J, Segschneider J, Stocker BD, Takahashi T, Tilbrook B, van Heuven S, Viovy N, Wanninkhof R, Wiltshire A, Zaehle S (2014) Global carbon budget 2013. Earth System Science Data 6:235–263, DOI doi:10.5194/essd-6-235-2014, URL http://www.earth-syst-sci-data.net/6/235/2014/essd-6-235-2014.html

MacLachlan C, Arribas A, Peterson KA, Maidens A, Fereday D, Scaife AA, Gordon M, Vellinga M, Williams A, Comer RE, Camp J, Xavier P, Madec G (2014) Global seasonal forecast system version 5 (GloSea5): a high-resolution seasonal forecast system. QJR Meteorol Soc p n/a, DOI 10.1002/qj.2396, URL http://dx.doi.org/10.1002/qj.2396

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

Manzanas R, Fr´ıas MD, Cofin˜o AS, Guti´errez JM (2014) Validation of 40 year multimodel seasonal precipitation forecasts: The role of ENSO on the global skill. J Geophys Res Atmos 119(4):1708–1719, DOI 10.1002/2013jd020680, URL http://dx.doi.org/10.1002/2013jd020680

Marletto V, Ventura F, Fontana G, Tomei F (2007) Wheat growth simulation and yield prediction with seasonal forecasts and a numerical model. Agricultural and Forest Meteorology 147(1-2):71–79, DOI 10.1016/j.agrformet.2007.07.003, URL http://dx.doi.org/10.1016/j.agrformet.2007.07.003

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.

Mishra A, Hansen JW, Dingkuhn M, Baron C, Traor´e SB, Ndiaye O, Ward MN (2008) Sorghum yield prediction from seasonal rainfall forecasts in burkina faso. Agricultural and Forest Meteorology 148(11):1798–1814, DOI 10.1016/j.agrformet.2008.06.007, URL http://dx.doi.org/10.1016/j.agrformet.2008.06.007

Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnusson L, Mogensen K, Palmer T, Vitart F (2011) The new ECMWF seasonalforecast system (system 4). Tech. Rep. 656, ECMWF, URL http://old.ecmwf.int/publications/library/ecpublications/ pdf/tm/601-700/tm656.pdf

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem Cycles 22(1):GB1022+, DOI 10.1029/2007gb002947, URL http://dx.doi.org/10.1029/2007gb002947

Nicklin KJ (2013) Seasonal crop yield forecasting in semi-arid West Africa. PhD Thesis. The University of Leeds, School of Earth and Environment.

Osborne, T., Gornall, J., Hooker, J., Williams, K., Wiltshire, A., Betts, R., and Wheeler, T.: JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator, Geosci. Model Dev., 8, 1139-1155, doi:10.5194/gmdd-7-6773-2014, URL http://dx.doi.org/10.5194/gmdd-7-6773-2014gmd-8-1139-2015, 2015.

Palmer TN, Doblas-Reyes FJ, Hagedorn R, Alessandri A, Gualdi S, Andersen U, Feddersen H, Cantelaube P, Terres JM, Davey M, Graham R, D´el´ecluse P, Lazar A, D´equ´e M, Gu´er´emy JF, D´ıez E, Orfila B, Hoshen M, Morse AP, Keenlyside N, Latif M, Maisonnave E, Rogel P, Marletto V, Thomson MC (2004) Development of a european multimodel ensemble system for seasonal-to-inter-annual prediction (DEMETER). Bull Amer Meteor Soc 85(6):853–872, DOI 10.1175/bams-85-6-853, URL http://dx.doi.org/10.1175/bams-85-6-853

Sacks WJ, Deryng D, Foley JA, Ramankutty N (2010) Crop planting dates: an analysis of global patterns. Global Ecology and Biogeography 19(5):607–620, DOI 10.1111/j.1466-8238.2010.00551.x, URL http://dx.doi.org/10.1111/j.1466-8238.2010.00551.x

Sacks WJ, Deryng D, Foley JA and Ramankutty N (2010) Crop planting dates: An analysis of global patterns. Global Ecology and Biogeography 19:607-620

Segele ZT and Lamb PJ (2005) Characterization and variability of Kiremt rainy season over Ethiopia. Meteorology and Atmospheric Physics 89:153-180

Shangguan W, Dai Y, Duan Q, Liu B and Yuan H (2014) A Global Soil Data Set for Earth System Modeling. Journal of Advances in Modeling Earth Systems 6:249-263

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.

Watson J, Challinor A (2013) The relative importance of rainfall, temperature and yield data for a regional-scale crop model. Agricultural and Forest Meteorology 170:47–57, DOI 10.1016/j.agrformet.2012.08.001, URL http://dx.doi.org/10.1016/j.agrformet.2012.08.001

Watson J, Challinor A, Fricker T, Ferro C (2014) Comparing the effects of calibration and climate errors on a statistical crop model and a process-based crop model. Climatic Change, pp 1–17, DOI 10.1007/s10584-014-1264-3, URL http://dx.doi.org/10.1007/s10584-014-1264-3

Williams K, Clark D (2014) Disaggregation of daily data in JULES. Tech. Rep. 96, Met Office Hadley Centre, URL http://www.metoffice.gov.uk/media/pdf/2/j/HCTN96.pdf

Williams K, Falloon (2015) Sources of inter-annual yield variability in JULES-crop and implications for forcing with seasonal weather forecasts, submitted to Geoscientific Model Development Discussions.

Appendix 1: NCDUMP of one of the forcing files created by WU for the Eastern African domain for each of the 5400 hindcasts

file naming convention:

forcing_seas15	refers to the ECMWF System 4, 15 member, 7 months (~212 days) seasonal forecasts
GHA	refers to Greater Horn of Africa, i.e. Eastern Africa fully comprising Ethiopia, Kenya and Tanzania plus parts of the surrounding area; **EU** for European domain
BC	refers to Bias Corrected data; **noBC** for non corrected data
E01	refers to ensemble member 1 (out of 15)
1981_01	refers to start month of forecast
.nc4	refers to NETCDF v4 compressed format

File "forcing_seas15_GHA_BC_E01_1981_01.nc4"

Dataset type: Hierarchical Data Format, version 5

```
netcdf file:/D:/bigdata_werk/EUPORIAS/S4_GHA/forcing_seas15_GHA_BC_E01_1981_01.nc4 {
 dimensions:
  lon = 43;
  lat = 60;
  time = UNLIMITED;  // (212 currently
 variables:
  double lon(lon=43);
   :standard_name = "longitude";
   :long_name = "Longitude";
   :units = "degrees_east";
   :axis = "X";

  double lat(lat=60);
   :standard_name = "latitude";
   :long_name = "Latitude";
   :units = "degrees_north";
   :axis = "Y";

  double time(time=212);
   :standard_name = "time";
   :long_name = "time";
   :units = "days since 1981-01-01 00:00:00";
   :calendar = "standard";
   :_ChunkSize = 1; // int

  float tas(time=212, lat=60, lon=43);
   :standard_name = "air_temperature";
   :long_name = "Near-Surface Air Temperature";
   :units = "K";
   :_FillValue = 1.0E20f; // float
   :missing_value = 1.0E20f; // float
   :_ChunkSize = 1, 60, 43; // int

  float tasmax(time=212, lat=60, lon=43);
   :standard_name = "air_temperature";
   :long_name = "Daily Maximum Near-Surface Air Temperature";
   :units = "K";
   :_FillValue = 1.0E20f; // float
   :missing_value = 1.0E20f; // float
   :_ChunkSize = 1, 60, 43; // int

  float tasmin(time=212, lat=60, lon=43);
   :standard_name = "air_temperature";
   :long_name = "Daily Minimum Near-Surface Air Temperature";
   :units = "K";
   :_FillValue = 1.0E20f; // float
   :missing_value = 1.0E20f; // float
   :_ChunkSize = 1, 60, 43; // int

  float pr(time=212, lat=60, lon=43);
   :standard_name = "precipitation_flux";
   :long_name = "Precipitation";
   :units = "kg m-2 s-1";
   :_FillValue = 1.0E20f; // float
   :missing_value = 1.0E20f; // float
   :_ChunkSize = 1, 60, 43; // int

  float rsds(time=212, lat=60, lon=43);
```

```
    :standard_name = "surface_downwelling_shortwave_flux_in_air";
    :long_name = "Surface Downwelling Shortwave Radiation";
    :units = "W m-2";
    :_FillValue = 1.0E20f; // float
    :missing_value = 1.0E20f; // float
    :_ChunkSize = 1, 60, 43; // int

   float rlds(time=212, lat=60, lon=43);
    :standard_name = "surface_downwelling_longwave_flux_in_air";
    :long_name = "Surface Downwelling Longwave Radiation";
    :_FillValue = 1.0E20f; // float
    :missing_value = 1.0E20f; // float
    :units = "W m-2";
    :_ChunkSize = 1, 60, 43; // int

   float huss(time=212, lat=60, lon=43);
    :standard_name = "specific_humidity";
    :long_name = "Near-Surface Specific Humidity";
    :units = "1";
    :_FillValue = 1.0E20f; // float
    :missing_value = 1.0E20f; // float
    :_ChunkSize = 1, 60, 43; // int

   float sfcWind(time=212, lat=60, lon=43);
    :standard_name = "wind_speed";
    :long_name = "Near-Surface Wind Speed";
    :units = "m s-1";
    :_FillValue = 1.0E20f; // float
    :missing_value = 1.0E20f; // float
    :_ChunkSize = 1, 60, 43; // int

 // global attributes:
 :institution = "Wageningen University and Research centre (WUR)";
 :source1 = "distributed by ECOMS User Data Gateway (ECOMS-UDG), partially funded by the European
Union FP under Grant Agreement no 308291 (EUPORIAS), http://meteo.unican.es/ecoms-udg";
 :source2 = "IPR remains with European Centre For Medium-Range Weather Forecasts (ECMWF)";
 :source3 = "ECMWF System4 hind cast, seasonal-15: 12 Runtimes per year (the 1st of J,F,etc) running
for 7 months. Ensemble of 15 members for 1981-2010";
 :source4 = "variables renamed, domain subset, masking, reorganised, interpolated to 0.5 degree by
Wageningen University, Earth System Sciences";
 :source5 = "variables naming convention: CMIP5 (http://cmip-
pcmdi.llnl.gov/cmip5/docs/standard_output.pdf)";
 :domain = "East Africa";
 :member = 1; // int
 :comment1 = "Processed and re-gridded by W.Franssen";
}
```