Research activities in regional crop modelling and yield forecasting

Allard de Wit, Kees van Diepen and Hendrik Boogaard

Wageningen UR, Alterra-CGI, P.O. Box 47, 6700 AA Wageningen, the Netherlands. Tel. +31-(0)317-481850. Email: rob.lokers@wur.nl

Background

CGMS is being applied successfully within the MARS Crop Yield Forecasting System for qualitative monitoring of the growing season and for making quantitative crop yield forecasts. Nevertheless, there are large uncertainties related to applying crop growth models over large areas. Examples of these uncertainties are the generally unknown within-season sowing dates, the uncertainty in the effect of drought due to limited weather station density and poorly known soil parameters, the lack of information about irrigation and the weighting of individual simulation results to administrative regions.

For regional applications, particularly the uncertainty in spatial and temporal distribution of rainfall has a large impact on the results of crop models (De Wit et al. 2005). This is caused by the erratic behaviour of rainfall in space and time and the strong non-linear response of crop models to rainfall. Consequently, when the model simulation results are used for regional crop yield forecasting the forecast accuracy is limited by this uncertainty.

Within Alterra-CGI, we have put considerable research effort on ways to quantify and reduce uncertainty in CGMS. Similar to meteorological application such as weather forecasting we are looking at probabilistic modelling and assimilation of satellite observations in order to improve CGMS.

Probabilistic modelling

Due to the complexity of crop models it is (in practice at least) n`a is more difficult because these values have a spatial and temporal structure which should be preserved in the ensemble of input data. Ideally, even correlations between meteorological variables (for example rainfall and radiation) should be preserved in the ensemble.

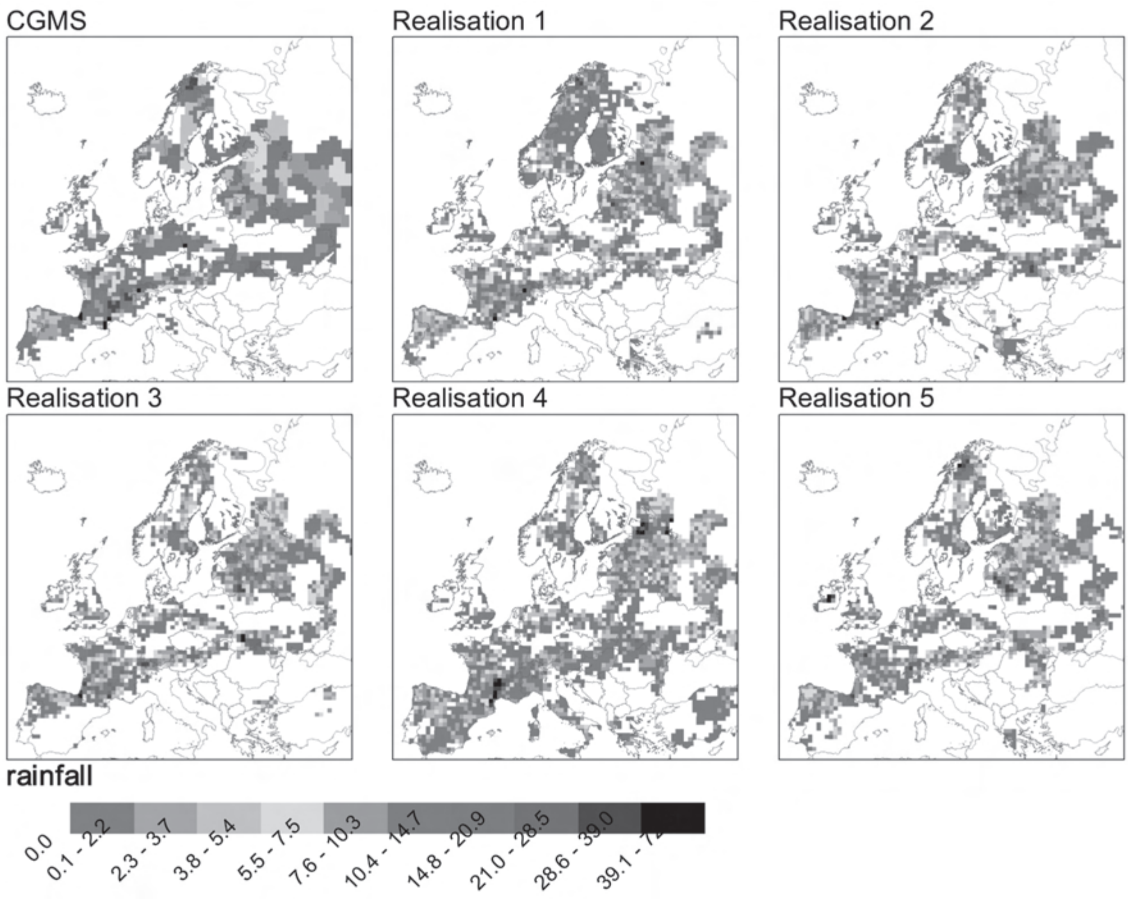

Figure 1. Rainfall product from CGMS (top left) and five realisations of this rainfall field for a single day showing the uncertainty in the rainfall field (values in mm).

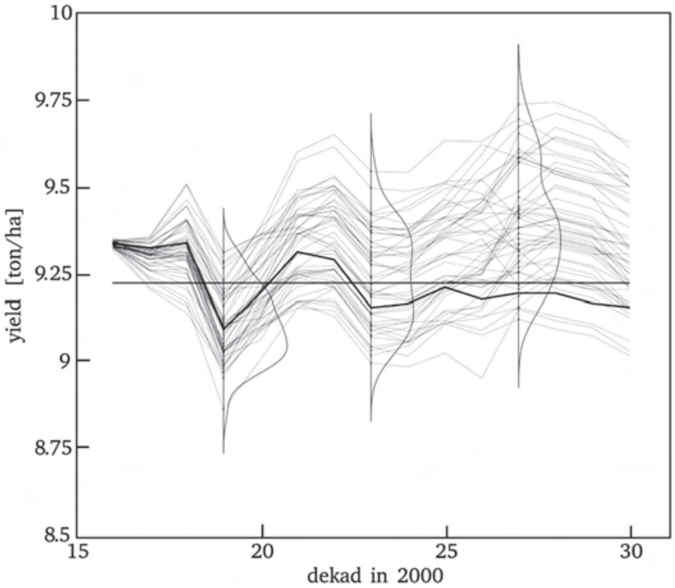

Figure 1. Rainfall product from CGMS (top left) and five realisations of this rainfall field for a single day showing the uncertainty in the rainfall field (values in mm).

Given that rainfall is the most uncertain and influential meteorologic variable, we have focussed on generating an ensemble rainfall product based on the CGMS gridded precipitation product (De Wit et al. 2008). We used an error model fitted to a highly accurate precipitation dataset which was available for the year 2000. The error model consisted of two components. The first is an additive component generating precipitation residues over the entire spatial domain. The residues are generated by quantile-based back transformation of standard Gaussian fields using a set of histograms for different CGMS precipitation bins. The second component is multiplicative and generates binary rain/no-rain events on locations where the CGMS precipitation records report no precipitation. The error model was used to generate an ensemble of 50 realizations of daily precipitation over the period 1990-2009.

Figure 1 shows an example of such a rainfall ensemble product. The top left image shows the original rainfall product from the CGMS database, the other images are ensemble realisations rainfall field. It can be clearly seen that the realisations are perturbations of the original rainfall fields which have roughly the same structure but the rainfall estimates vary for each realisation.

We used the generated ensemble of rainfall products as input in CGMS and the ensemble of simulated biomass values was used to create a probabilistic crop yield forecast for grain maize for a province in South-France (Figure 2.) The figure demonstrates that the uncertainty in precipitation has a profound influence on the value of the yield forecast during the growing season. Compared to the deterministic yield forecast (blue line), the probabilistic yield forecasts shows a diverging ensemble of yield forecasts which keeps diverging almost up to end of growing season with maximum spread in the yield forecast of around 0.65 ton/ha. Given that EUROSTAT uses a tolerance of 0.2 ton/ha as an acceptable accuracy for yield forecasts (G. Genovese, pers. comm.), this is a significant deviation.

Satellite data assimilation

The use of satellite observations to improve crop model simulation and/or their direct use as crop yield predictors has been demonstrated at local level (e.g. a field or small region), but has proven difficult to implement over Europe. There are several reasons:

- Land cover in Europe is highly fragmented and the inter-

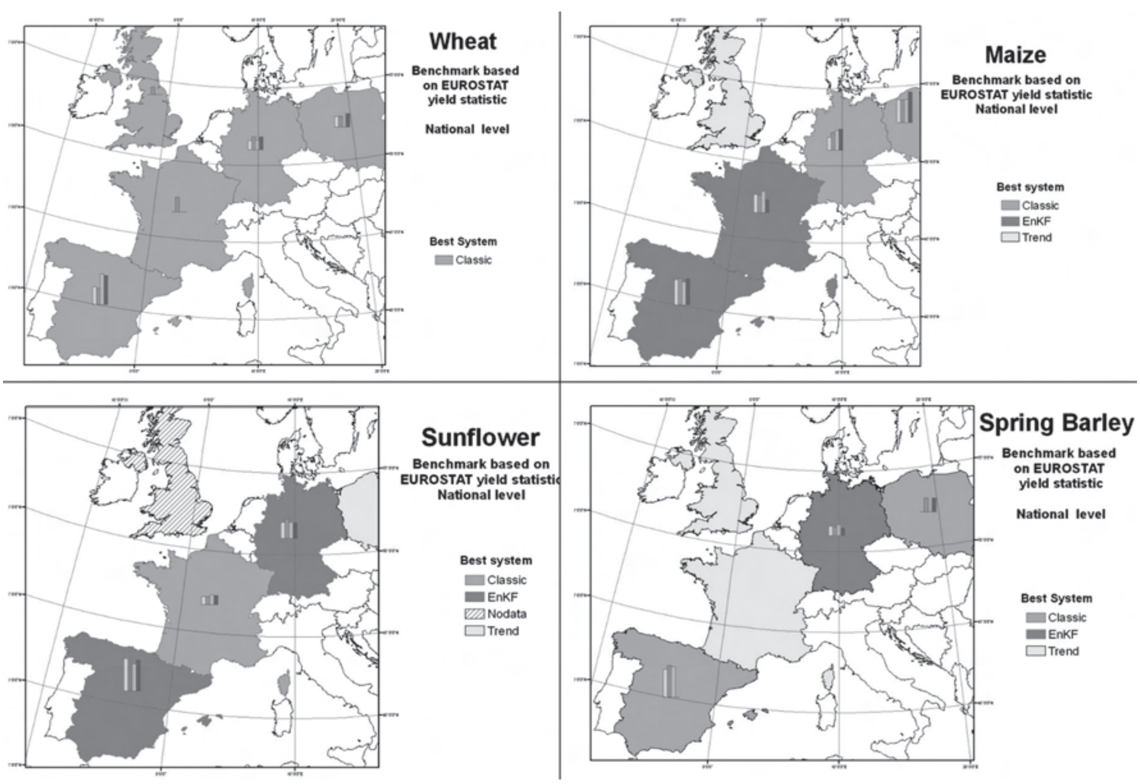

Figure 2. Ensemble crop yield forecast during the growing season in 2000. The thick black line is the official reported yield by EUROSTAT, the blue line represents the deterministic yield forecast and the dotted lines represent the influence of uncertainty in precipitation on the yield forecast. Density plots indicate the shape of the yield forecast ensemble at three moments during the growing season.

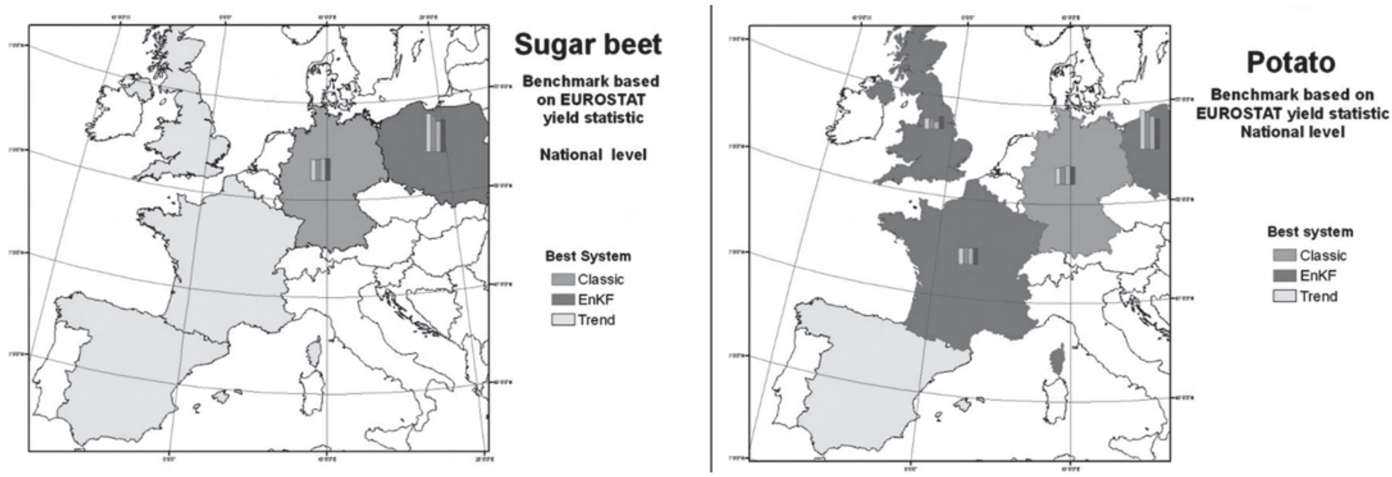

Figure 3. Performance of assimilating SWI for various crop types at national level. Bars show the relative error of the yield prediction

pretation of the low resolution optical satellite observations (250m to 1km pixels) is ambiguous because it represents most often a mixture of several land cover types;

- Lack of consistent time-series of remote sensing data due to persistent cloud cover, sensor calibration problems or satellite mission continuity;
- Lack of sensitivity of commonly used remote sensing indicators (e.g. NDVI) in much of Europe due to the high crop production levels and the relatively small year-to-year variability.

For assimilating satellite observations in CGMS we have therefore used a totally different satellite-derived product: the Soil Water Index (SWI) derived from scatterometer-based soil moisture estimates. This product has the advantage that it is operational and its availability is guaranteed for the coming 15 years. Moreover, an archive spanning 1992-2006 exists that can be used to implement and validate the approach.

The assimilation itself was carried out using an Ensemble Kalman Filter (EnKF) which uses an ensemble of models to estimate the variance of the modelled soil moisture. The variance on the model soil moisture relative to the variance in the satellite observation is then used as a measure to adjust the soil moisture in the model. The simulation results at regional scale are then compared with EUROSTAT crop yield statistics in order to determine if assimilation of SWI leads to improved relationships between simulation results and reported EUROSTAT yields.

Figure 3 shows some results at national level for Spain, France, Germany, Poland and the UK for several crops. Countries that are coloured orange have no benefit from the data assimilation, countries coloured green have improved while for countries that are coloured grey no conclusions can be drawn. The results demonstrate that the assimilation of SWI is not necessarily beneficial. Results for spring barley and particularly winter-wheat have deteriorated, while maize, potato and sunflower do benefit from assimilation of SWI for some countries. For sugar beet there are only results for Germany and Poland with contrasting results.

Further analyses of these results indicated that the yield forecasting performance is probably linked to the performance of the Ensemble Kalman filter, because crops with a poor performance in yield forecasting also showed unfavourable Kalman filter statistics. This may provide a handle to improve on the data assimilation, or to discourage the assimilation of SWI for particular crops..