Technical description of crop model (WOFOST) calibration and simulation activities for Argentina, pampas region

Allard de Wit
Diego Abelleyra
Santiago Veron
Joop Kroes
Iwan Supit
Hendrik Boogaard

June 2017

1. Sites

To define local crop varieties, sites have been selected using the GYGA-ED Climate Zones (CZ) (Wart et al, 2013). A procedure was designed to find a representative location(s) within each CZ for the targeted crop (see separate SIGMA document). Each one or site(s) then represents one GYGA-ED zone and for each site local crop data is collected to calibrate the selected crop growth model. This results in at least one variety for a GYGA-ED zone.

Five sites were selected, shown in the following Figure 1-1 and Table 1-1. Note that more than one site per zone was selected to better capture the south-north gradient of day length.

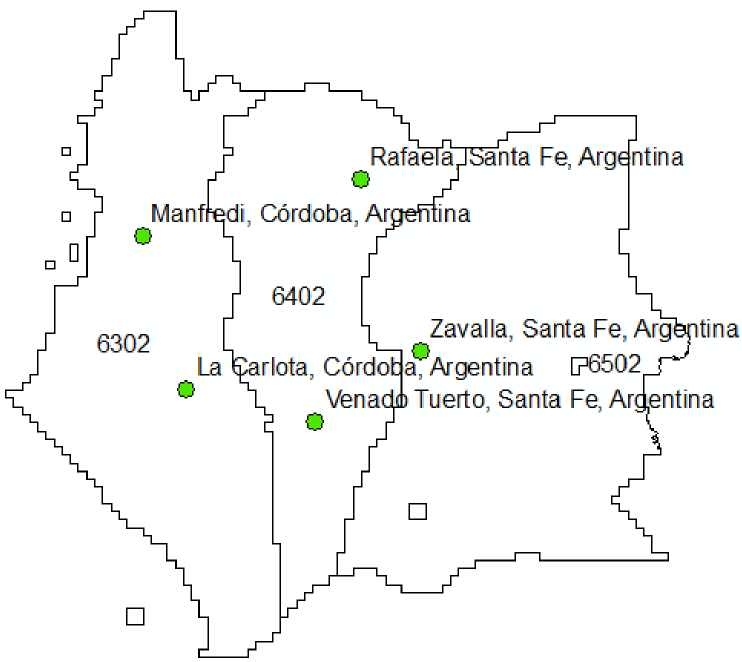


Figure 1-1 Selected sites for calibration (number refer to GYGA-ED zone)

Table 1-1 Details of selected sites and reference to collected crop data

Location	Nr	Longitude	Latitude	Data availability	Site number of field experiments
Zavalla	2	-60.883331	-33.01667	Biophysical parameters, phenology	35 (a-l)
Venado Tuerto	3	-61.965641	-33.75	yield, planting date, R1, R7/R8	31 (a-e)
Rafaela	4	-61.491642	-31.252598	yield, planting date, R1, R7/R8	24 (a-j)
Manfredi	5	-63.750001	-31.83333	yield, planting date, R1, R7/R8	15 (a-o)
La Carlota	6	-63.293568	-33.420334	yield, planting date, R7/R8	13 (a-e)

2. Data for selected sites

2.1. Weather

For the selected sites nearby weather stations were selected (see Table 2-1).

Table 2-1 Weather stations for selected sites

Location	Nr	Selected weather station	Longitude	Latitude
Zavalla	2	Zavalla UNR WS	-60.88	-33.02
Venado Tuerto	3	Venado Tuerto Aero WS	-61.95	-33.75
Rafaela	4	Rafaela INTA WS	-61.55	-31.18
Manfredi	5	Manfredi INTA WS	-63.77	-31.82
La Carlota	6	Rio Cuarto Aero WS	-64.23	-33.12

Daily weather data for these stations were provided by INTA. Table 2-2 gives an overview of the available weather elements.

Table 2-2 Available weather elements for weather stations

Element	Description	Unit	Required Unit
RG	Daily sum incoming global radiation at earth surface	MJ.m-2.d-1	KJ.m-2.d-1
TMAX	Daily maximum temperature	°C	°C
TMIN	Daily minimum temperature	°C	°C
PRECIP[1]	Daily sum precipitation	mm.d-1	mm.d-1
TMIN5	Daily minimum soil temperature at 5 cm depth	°C	Not required
HELIOF	Duration of sunlight	hours	Not required
HEREL	Duration of sunlight as percentage of day length	%	Not required
TVAP	Daily average vapour tension	hPa (=mb)	hPa
HR(%)	Daily average relative humidity	%	Not required
VV10	Wind speed at 10 m	m.s-1	m.s-1
VV2	Wind speed at 2 m	m.s-1	Not required
ETP	Evapotranspiration	mm.d-1	Not required[2]

Daily data were checked for completeness and data ranges (Table 2-3).

Table 2-3 Analysis data ranges of daily station data

Element	min	avg	max	missing values
RG	0	15.8	32.3	-99.9, -9.9
TMAX	-0.8	23.5	42.0	-99.9
TMIN	-13.0	10.8	29.4	-99.9
PRECIP	0	2.5	315	-99.9
TVAP	0	15.0	88.0	-99.9, -9.9
VV10	0	11.6	180.0	-99.9, -99

Data were completed by global gridded weather data taking the most nearby grid cell of the selected data source (see Table 2-4):

- JRC_ERA-INTERIM (see section 2.1.1)
- NASA_POWER (see section 2.1.2)
- CHG_CHIRPS (see section 2.1.3)

[1] 12 AM GMT - 12 AM GMT (Argentina weather stations)
[2] Will be calculated in PCSE (Penman-Monteith)

Elements with missing data were mainly wind speed (35%), vapour pressure (21%) and radiation (10%). Finally, radiation values were multiplied by factor 1000.

Table 2-4 Selected location (lon, lat) of gridded weather sources per site

Location	Nr	Station	JRC_ERA-INTERIM	NASA_POWER	CHG_CHIRPS
Zavalla	2	Zavalla UNR WS	-61/-33 (11.42 km)	-60.5/-33.5 (64.08 km)	-60.875/-33.125 (11.70 km)
Venado Tuerto	3	Venado Tuerto Aero WS	-62/-33.75 (4.63 km)	-61.5/-33.5 (50.15 km)	-61.875/-33.875 (15.55 km)
Rafaela	4	Rafaela INTA WS	-61.5/-31.25 (9.13 km)	-61.5/-31.5 (35.94 km)	-61.625/-31.125 (9.41 km)
Manfredi	5	Manfredi INTA WS	-63.75/-31.75 (8.02 km)	-63.5/-31.5 (43.86 km)	-63.875/-31.875 (11.67 km)
La Carlota	6	Rio Cuarto Aero WS	-64.25/-33 (13.49 km)	-64.5/-33.5 (49.20 km)	-64.125/-33.125 (9.81 km)

2.1.1. JRC_ERA-INTERIM

JRC_ERA-INTERIM data is owned by the MARS AGRI4CAST project of the Joint research Centre (JRC) of European Commission (EC). It is based on the 3-hourly ERA-Interim data obtained from the European Centre for Medium range Weather forecast (ECMWF; Berrisford et al. 2009). The ECMWF ERA-Interim data set is a reanalysis of the global atmosphere of the period 1989 to date. It has a spatial resolution of 0.75 x 0.75 degrees. Due to an improved reanalysis system, ERA-INTERIM has proved to have better performance compared to previous reanalysis data sets such as ERA-40 (ECMWF, 2007). Within the JRC MARS project the data were processed to arrive at daily or 10-daily data (aggregations and bias corrections). First 3-hourly were aggregated into daily data using indicator specific time zones. Next daily data, available at the 0.75 degree grid were downscaled (Inverse distance weight interpolation) to a regular global 0.25 degree grid and afterwards bias corrected. The bias correction between the IDW-interpolated ERA-Interim model and ECMWF operational model data available at the 0.25 degree resolution was done for temperature related elements, radiation and wind speed. The daily data of the ECMWF Operation model, available at the 0.25 degree grid for the period 2008-2010, was used as a training set to determine the bias correction. For rainfall no corrections were applied. The rainfall parameter showed less accurate results in the regression due to its intermittent nature and distribution (see for more information Hartman, 2011).

Temporal resolution: daily
Spatial resolution: 0.25 x 0.25 degree
Download: via JRC (not publicly accessible, but available for SIGMA)
Reference: http://marswiki.jrc.ec.europa.eu/agri4castwiki/index.php/Main_Page

2.1.2. NASA_POWER

The POWER project was initiated to improve upon the current Surface Meteorological and Solar Energy (SSE) project data set and to create new data sets from new satellite systems and forecast modeling data. The parameters contained in the agro-climatology archive are based primarily upon solar radiation derived from satellite observations and meteorological data from the Goddard Earth Observing System assimilation model.

Temporal resolution: daily

Spatial resolution: 1.0° resolution
Download: all products can be accessed directly through the dedicated http site: http://power.larc.nasa.gov/cgi-bin/cgiwrap/solar/agro.cgi?email=agroclim@larc.nasa.gov
Reference: NASA (http://power.larc.nasa.gov/common/php/POWER_AboutPOWER.php)

2.1.3. CHG_CHIRPS

CHIRPS v2 is a satellite-based rainfall monitoring dataset specifically designed to support trend analysis and seasonal drought monitoring around the world. It is developed as a partnership between the USGS Earth Resources Observation and Science (EROS) Center and the University of California Santa Barbara (UCSB) Climate Hazards. CHIRPS has been supported by funding from USAID, FEWS NET, NASA and NOAA. CHIRPS incorporate satellite imagery with in-situ station data to create gridded rainfall time series.

Temporal resolution: daily
Spatial resolution: 0.25° resolution
Download: all products including the daily data can be also accessed directly through the dedicated ftp site: ftp://chg-ftpout.geog.ucsb.edu/pub/org/chg/products/CHIRPS-2.0
Reference: Funk et al. 2015

2.1.4. NASA_TRMM

The TRMM Product 3B42 (V7) was selected. The Rainfall Measuring Mission (TRMM) is a joint U.S.-Japan satellite mission to monitor tropical and subtropical precipitation and to estimate its associated latent heating. The purpose of the 3B42 algorithm is to produce TRMM-adjusted merged-infrared (IR) precipitation and root-mean-square (RMS) precipitation-error estimates. The algorithm consists of two separate steps. The first step uses the TRMM VIRS and TMI orbit data (TRMM products 1B01 and 2A12) and the monthly TMI/TRMM Combined Instrument (TCI) calibration parameters (from TRMM product 3B31) to produce monthly IR calibration parameters. The second step uses these derived monthly IR calibration parameters to adjust the merged-IR precipitation data, which consists of GMS, GOES-E, GOES-W, Meteosat-7, Meteosat-5, and NOAA-12 data. The final gridded, adjusted merged-IR precipitation (mm/hr) and RMS precipitation-error estimates have a daily temporal resolution

Temporal resolution: daily (00Z UTC to 21Z UTC)
Spatial resolution: 0.25 x 0.25 degree
Download: http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&dataset=3B42:%203-Hour%200.25%20x%200.25%20degree%20merged%20TRMM%20and%20other%20satellite%20estimates&project=TRMM&dataGroup=Gridded&version=007
References: Tropical Rainfall Measuring Mission (TRMM) (2011), TRMM (TMPA) Rainfall Estimate L3 3 hour 0.25 degree x 0.25 degree V7, Greenbelt, MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: June 2015 https://disc.gsfc.nasa.gov/datacollection/TRMM_3B42_7.html

2.2. Soil data

The calibration was done for the potential production under irrigated conditions. Therefore data on soils are not required (no drought stress have to be simulated/assessed).

2.3. Crop data

Observed data for soybean were obtained from Red Nacional de Evaluación de Cultivares de Soja (RECSO) and National University of Rosario experiments (UNR). These observations consist of 1259 experiments over the period 2011-2013 (sowing year) for four different sites: Manfredi, Rafaela, Venado Tuerto and Zavalla (Figure 2-1Figure 2-1). Note that data for La Carlota and San Antonio de Areco were not available.

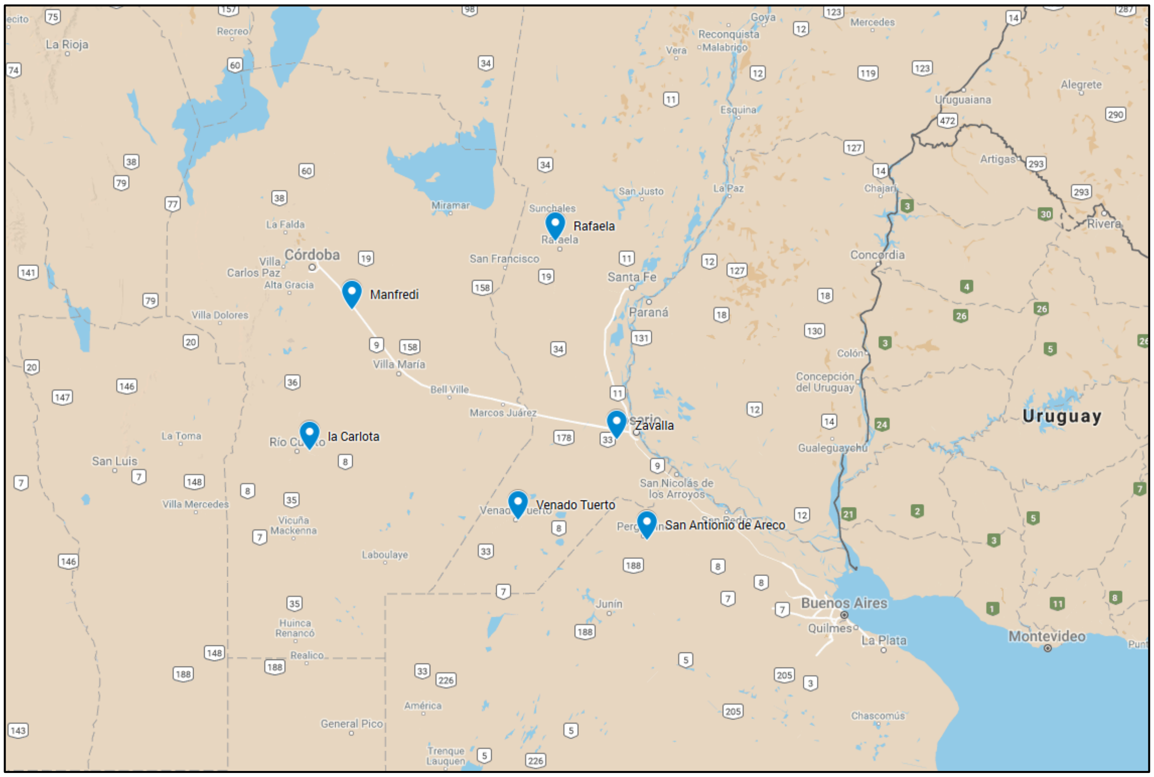


Figure 2-1 Overview of sites with experimental data for soybean (no data of La Carlota and San Antonio de Areco were used)

The number of experiments is variable over the different sites with the largest number of experiments at Rafaela and only 20 experiments at Zavalla (see Figure 2-2). Moreover, a total of 259 unique soybean cultivars were tested at those sites. Most soybean cultivars have a high sensitivity to day length and the different cultivars are categorized in so-called maturity groups (MG) that indicate the total length of the growing period. Although all different cultivars are classified in a maturity group there can be considerable variability due to differences in cultivar. Figure 2-2 (right) shows the distribution of soybean experiments over the different maturity groups. Most experiments were carried out with cultivars of maturity groups 4, 5 and 6. The extremely early cultivars (MG=3) and late cultivars (MG=7, 8) are less well represented in the experimental data. All experiments have been carried out under water-limited conditions but with proper crop management.

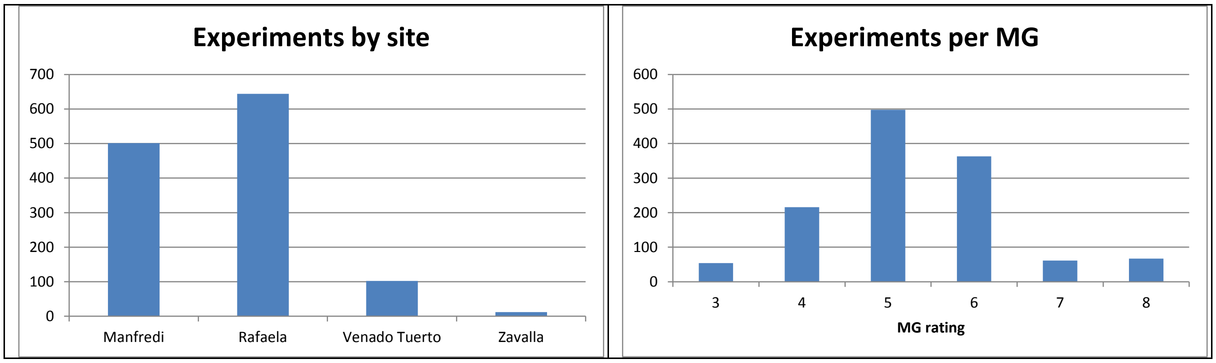


Figure 2-2 Number of soybean experiments over the different sites (left) and the distribution of experiments over the different maturity group ratings (right)

The soybean experiments provided generally contain information about crop phenology, crop height and crop yield. Observations of total crop biomass or more advanced parameters (biomass of crop organs, specific leaf area, etc.) are not available. Table 2-5 gives an overview of the available types of observations. Phenological observations of types RA, EMG, R1 and R8 are available for all experiments, while R5 and R7 are only available for 223 and 480 experiments respectively.

Table 2-5 Overview of available observations for soybean in Argentina

Parameter	*Description*	*unit*
RA	Sowing	date
EMG	Emergence: hypocotyl with cotyledons break through soil surface (“cracking stage”)	date
V0	Cotyledons completely unfolded	date
V1	2 full leaves (first leaf pair unfolded)	date
R0	Floral induction: not observable and no BBCH equivalent	date
R1	Beginning of flowering	date
R5	Beginning of seeds	date
R7	10% of pods ripe	date
R8	Fully ripe	date
Yield	Crop yield – fresh weight	kg/ha
Moisture content	Weight fraction of moisture in yield	-
Canopy height	Maximum height of crop canopy	cm

3. Model adaptation and calibration

3.1. Introduction

The WOFOST simulation model is a generic crop simulation model that can simulate the growth and development of different annual crops by applying specific parameters for each crop. Parameter files for the simulation of soybean have been available for a long time and were documented already by Van Heemst (1988). Moreover, specific parameter files for European conditions were created and documented by Boons-Prins et al. (1993) in the framework of the MARS project.

An important aspect of the simulation of crops is the phenological development of the crop. Simulation of phenological development in WOFOST 7.1 is based on the phenological development pattern of a typical cereal plant. It is defined by a dimensionless variable called the Development Stage (DVS) where DVS equals zero at crop emergence, DVS equals one at anthesis (flowering) and DVS equals two at maturity. The daily development rate from one stage to the next is calculated from the daily average temperature adjusted by a base temperature and divided by the temperature sum needed to reach the next development stage: TSUM1 for the stage from emergence to anthesis and TSUM2 for the stage from anthesis to maturity. For long-day plants such as cereals, WOFOST allows to take the impact of day length and vernalisation on phenological development into account; limiting the development rate under conditions of too short day length or when the saturated vernalisation requirement has not been reached.

A characteristic of the phenological development scheme used in WOFOST is that phenological development is essentially sequential. The plant goes through a defined set of sequences and those sequences do not overlap. This type of phenological development is typical for cereals and it is appropriate for tuber crops (potato, sugar beet) as well given that those crops have a very simple phenological development pattern. For tuber crops, the anthesis date does not correspond to flowering but to the start of tuber development.

However, in the case of soybean the sequential phenological development scheme used by WOFOST 7.1 does not describe the growth stages of soybean very well, for several reasons:

1. The phenological development of soybean is to a large extent parallel. Following the definition of soybean phenology by Fehr and Caviness (1977), the vegetative development of stems and leaves (the ‘V’ stages) runs parallel to the reproductive development of pods and seeds (the ‘R’ stages) for a considerable part of the growth cycle.
2. The temperature response function for development rate of soybean is more complicated and cannot be simulated by accumulating the daily average temperature above a base temperature.
3. In contrast to cereals, soybean is a short-day plant meaning that the phenological development rate of soybean accelerates under shorter day length. The short-day dependence of soybean cultivars is formalized in so-called “maturity groups” which indicate the critical and optimal day length for a given cultivar. The standard WOFOST 7.1 was not able to simulate this behaviour.

Given the considerations above, it was decided to develop an alternative model for phenological development of soybean.

3.2. Adaptations for WOFOST-Soybean

3.2.1. Phenology module

As described before, WOFOST has model for phenological development that is based on the principle of a sequential development stage (DVS). Although, the phenological development of soybean is partially parallel, the principle of the DVS cannot be eliminated from the model completely given that many internal variables and parameters rely on the DVS to receive an appropriate value. Therefore, we developed a hybrid phenological development model taking elements from established models for soybean phenology (SoyDev – Setiyono et al. 2007) but still applying the sequential DVS logic that is needed for WOFOST.

The new phenological development model has the following three elements. First of all, the phenological development is defined as in Table 3-1 where DVS=0 means emergence, DVS=1 is equivalent to the R1-stage (the *beginning of flowering*) and DVS=2 is equivalent to the R8 stage (*fully ripe*). There is some uncertainty on the exact interpretation of the vegetative stages. Van Heemst (1988) makes a distinction between pod wall development (starting at DVS=1.0) and seed development (starting at DVS=1.15). As there is no explicit mentioning of flowering by Van Heemst, we currently assume that DVS=1 represents both the start of flowering and pod development. This is also supported by the fact that the partitioning scheme in WOFOST only starts to partition a small fraction of assimilates to the storage organs (beans).

Second, the vegetative part from DVS=0 to 1 is driven by temperature only. It is simulated by a maximum development rate for emergence to flowering ($DVRMAX_1$) multiplied by a temperature reduction function defined by an optimal temperature where phenological development rate is maximal and the two cardinal temperatures below or above phenological development is halted (figure 3b from Setiyono et al. 2007).

$$DVR = DVRMAX_1 \cdot f(T)$$

Finally, the development rate during the reproductive part from DVS=1 to 2 is driven by both temperature and day length. Temperature is modelled using the same beta function as the vegetative stage, while the day length effect is simulated with same function as in Setiyono et al. 2007.

$$DVR = DVRMAX_2 \cdot f(T) \cdot f(P)$$

Table 3-1 ***Relation between R-stages that are common to soybean phenology data and the WOFOST DVS-phenology***

Stage	WOFOST DVS	description
RA	N/A	Sowing
EMG	DVS = 0	Emergence
R0	N/A	Floral induction: not observable and no WOFOST equivalent

R1	DVS = 1.0	Beginning of flowering, beginning of pod development
R5	DVS = 1.15	Beginning of seed development
R7	N/A	10% pods ripe, no WOFOST equivalent
R8	DVS = 2.0	Fully ripe.

3.2.2. Temperature response function

The temperature response function was taken from Setiyono et al. (2007) and is modelled as a beta function defined by an optimal temperature *Topt* where phenological development rate is maximal and the cardinal temperatures *Tmin* and *Tmax* below or above phenological development is halted:

$$f(T) = \begin{cases} 0, & T \leq T_{min}, T \geq T_{max} \\ \dfrac{2(T - T_{min})^{\propto}(T_{opt} - T_{min})^{\propto} - (T - T_{min})^{2\propto}}{(T_{opt} - T_{min})^{2\propto}}, & T_{min} < T < T_{max} \end{cases}$$

With

$$\propto = \frac{\ln(2)}{\ln\left((T_{max} - T_{min})/(T_{opt} - T_{min})\right)}$$

Values for *Tmin*, *Topt* and *Tmax* were taken from Setiyono et al. (2007) and set to 7.0, 31.0 and 40.0 degrees Celsius (see Figure 3-1).

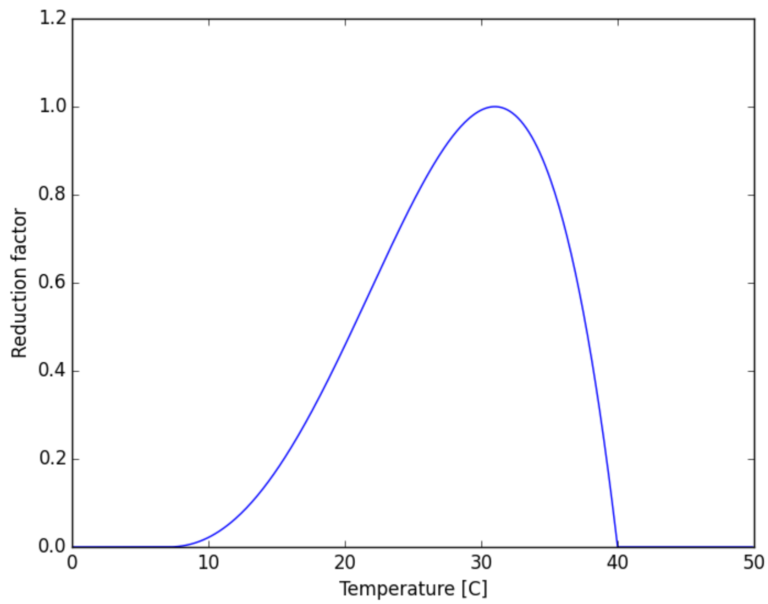


Figure 3-1 Temperature response function for phenological development of soybean (0 = complete reduction and 1 = no reduction)

3.2.3. Photoperiod response function

For the photoperiod response function the non-linear beta function proposed by Setiyono et al. (2007) was used:

$$f(P) = \left[\left(\frac{P - P_{opt}}{m} + 1\right)\left(\frac{P_{crt} - P}{P_{crt} - P_{opt}}\right)^{(P_{crt} - P_{opt})/m}\right]^{\propto}$$

With

$$\propto = \frac{\ln(2)}{\ln\left((P_{crt} - P_{opt})/m\right)}$$

This function is defined by two parameters: the optimal daylength (P_{opt}) below which there is no reduction of the development and the critical daylength (P_{crt}) above which the development rate is zero. The parameter m is a constant with the value of 3.0 h.

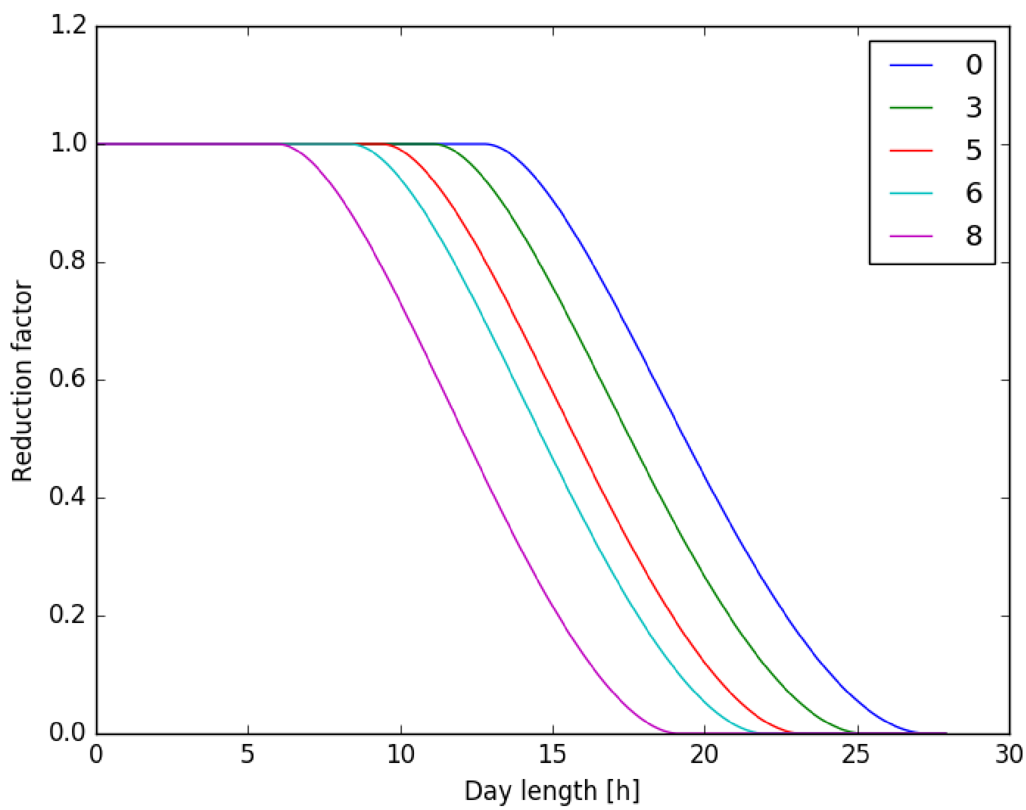


Figure 3-2 Photoperiod response function for phenological development of soybean for different maturity groups. Values for Popt and Pcrt in the figure where derived using the empirical relationship presented by Setiyono et al. (2007). See also figure 4 in Setiyono et al. 2007 – note that the *Popt* and *Pcrt* are swapped in the figure (0 = complete reduction and 1 = no reduction).

3.2.4. Estimation of parameter values

Calibration of the new WOFOST-soybean including the new hybrid phenological was carried out using observed phenological stages for soybean. Application of WOFOST-soybean required several new parameters to be calibrated. It was assumed that the temperature response for phenological development is relatively stable across cultivars and that main parameters to be estimated are those related to development rate and photoperiodicity. Therefore, the parameters that were calibrated are the maximum development rate for the vegetative and reproductive stages ($DVRMAX_1$ and $DVRMAX_2$) and the optimal and critical day length (Popt and Pcrt).

3.3. Calibration approach

For calibrating WOFOST-soybean a new generic CalibrationManager was developed that can make estimates of any set of model parameters. The main requirement is that there are observations available that are suitable to estimate the target parameters.

The CalibrationManager has several components that are combined in order to estimate the model parameters (Figure 3-3):

- Observations are retrieved from the database of observations which form the basis of the calibration;
- For each set of observations that represent a cropping season, a ModelRunner is started which consists of a WOFOST model and all data needed to execute WOFOST for the given cropping season.
- The observations and modelRunners are combined in an ObjectiveFunction-Calculator (OFC). The OFC executes all ModelRunners with a given set of parameters. Next, it retrieves for each observed value the equivalent simulated value. Finally, it computes an error value that summarizes the differences between all observed and simulated values. Several error measures can be used, but in practice the Root Weighted Mean Squared Error is used.
- The CalibrationManager integrates the NLOPT library (http://ab-initio.mit.edu/wiki/index.php/NLopt) which contains algorithms to efficiently explore the parameter space in order to find the parameter values that return the lowest error from the OFC.

Besides the components of the CalibrationManager itself, the system uses many components of PCSE (http://pcse.readthedocs.io) for retrieving weather data, parameter values and carrying out the actual crop simulation.

In case of calibration of soybean phenology the CalibrationManager uses observations of the EMG stage (emergence), R1 (beginning of flowering) and R8 (fully ripe) stages to estimate the values of the different parameters (Table 3-1).

Setiyono et al. (2007) present an empirical relationship between the maturity group rating of a soybean cultivar and the parameters related to photoperiod (Pcrt and Popt), see also Figure 3-2. This relationship was derived for soybean cultivars in the U.S. For the first calibration tests we assumed that this relationship was also valid for soybean cultivars in Argentina. However, we quickly found out that settings for Pcrt and Popt were not appropriate and did not yield appropriate results. Therefore, it was decided to calibrate the parameters as listed in Table 3-2 including the parameters related to photoperiod with the configuration settings shown in Figure 3-4.

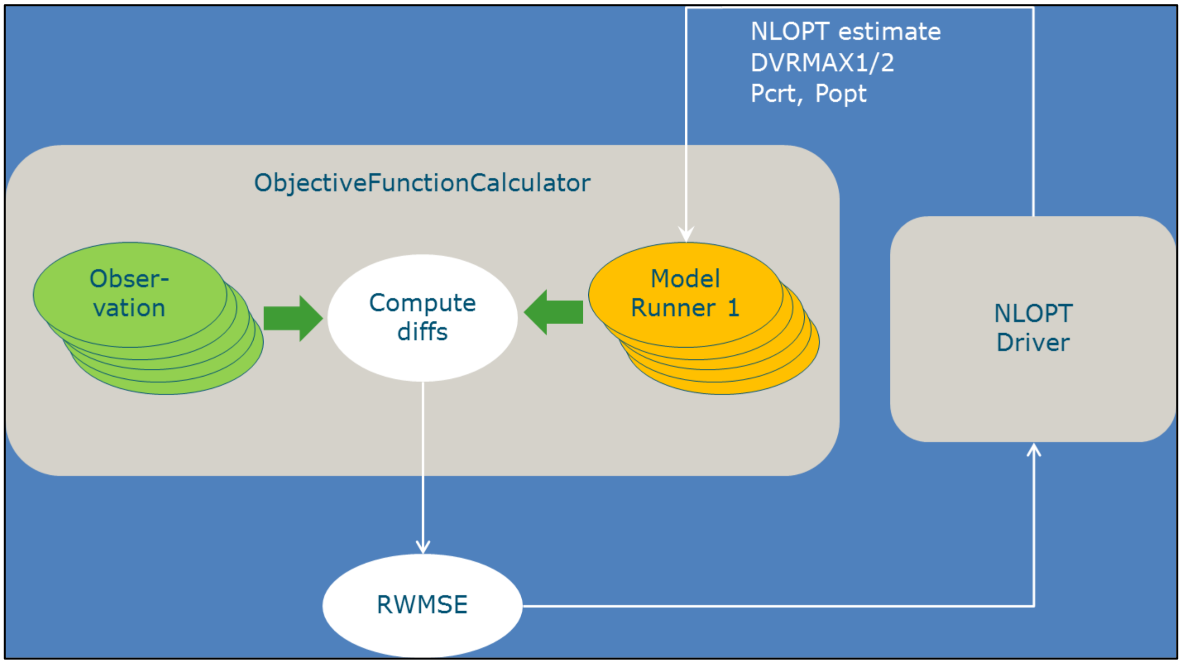


Figure 3-3 Schematic representation of the CalibrationManager procedure

Table 3-2 Overview of parameters in the soybean phenology module and the parameters that were calibrated based on observations (right column)

Parameter	Description	Unit	Calibrated
T_{min}	Minimal temperature for phenologic development	C	No
T_{opt}	Optimal temperature for phenologic development	C	No
T_{max}	Maximal temperature for phenologic development	C	No
$DVRMAX_1$	Maximum development rate for emergence to flowering	-	Yes
$DVRMAX_2$	Maximum development rate for flowering to maturity	-	Yes
P_{crt}	Critical daylength for phenologic development	h	Yes
P_{opt}	Optimal daylength for phenologic development	h	Yes
TSUMEM	Temperature sum from sowing to emergence	Cd	No

```
# target parameters for optimization and target variables for minimization
target_parameter_names = ["DVRMAX1", "DVRMAX2","Pcrt","Popt"]
parameter_lower_bounds = [0.006, 0.01, 14, 7]
parameter_upper_bounds = [0.10, 0.1, 20, 12]
parameter_default_values = [0.04, 0.054, 17, 8]
parameter_initial_step = [0.0025, 0.0025,0.025,0.025]

# Optimizer settings
max_evaluations = 200
objfunc_ftol = 0.01

# Target variables that will be used to compute the objective function
target_summary_variables = ["DOR1", "DOR8"]
```

Figure 3-4 Configuration settings for the CalibrationManager for the optimizing the phenological parameters of the soybean phenology model

3.4. Calibration results: phenology

Given the large differences in photoperiod response between maturity groups, we estimated parameter value for the different soybean maturity groups separately. The CalibrationManager was set up to select only those observations from the INTA experimental database for the given maturity group across the various sites. This implies that we assume that we can describe the phenological development of soybean for a given maturity group with a single set of parameters and thus differences in responses in phenology are caused by variations in sowing date, weather and day length at the different sites.

Table 3-3 lists the calibrated parameter values for the different maturity groups that were obtained by the CalibrationManager, including the final error value defined as the sum of the Root Mean Squared Errors of the R1 and R8 stages. In general, the error values are between 7 days (MG=8) and 21 days (MG=7). Although this is still relatively high one should realize that this is the sum of the RMSE of R1 stage and the R8 stage. Nevertheless, the results demonstrate that there is still a considerable variability in the observations that cannot be explained by the model.

The large variability is confirmed by the scatter plots showing the observed versus simulated number of days to the R1 stage (Figure 3-5) and the R8 stage (Figure 3-6) of soybean. The plot also demonstrates that within one maturity group it is often difficult to explain the variability. Nevertheless, it is promising that the model is able to explain some of the extreme cases well such as the points in the upper right corner of Figure 3-6 which refer to experiments with very early sowing.

The calibrated development rates are visualized in Figure 3-7 demonstrating that the DVRMAX1 is steadily declining with increasing MG rating, while the DVRMAX2 is generally increasing except for the highest maturity groups (7 & 8). The photoperiod parameters Pcrt and Popt are visualized in Figure 3-8 demonstrating that the optimal day length is decreasing from 10 to 7.5 hours with increasing maturity group rating, while the critical day length is relatively constant at 15 hours. If we compare these results against the empirical rating of Setiyono et al. (2007) (Figure 3-8 – dashed lines) then we see that estimates of the P_{opt} are similar while estimates of P_{crt} are vastly different. This indicates that the empirical relationship of Setiyono et al. (2007) is not applicable for soybean in Argentina.

When looking at the average length of the R1 and R8 stages (Figure 3-9), it is clear that with increasing maturity group the length of the R1 stage is increasing, while the length of the R8 stage (difference between blue and red line) is relatively constant. This observation explains the decrease in $DVRMAX_1$ as the phenological stage needs to be completed in more days and a lower development rate is required.

For the DVRMAX2, the effect is more complicated because the day length plays a role. In general higher development rates are needed for late maturing cultivars (higher MG) to complete the season before the winter. Only, the cultivars from maturity group 8 seem to break with this pattern for reasons unclear yet.

Table 3-3 Parameters of the phenology model after calibrating on observations. The number of experiments is given between brackets.

Site	*MG*	*DVRMAX₁*	*DVRMAX₂*	*P_crt*	*P_opt*	*Error[days]*
Venado Tuerto, Manfredi	2.66 (51)	0.0439	0.0544	14.78	10.23	11.22
Venado Tuerto, Manfredi	3.66 (114)	0.0398	0.0567	15.02	9.49	12.87
Zavalla, Rafaela, Manfredi	4 (100)	0.0387	0.0582	15.00	8.76	15.27
Venado Tuerto, Rafaela, Manfredi	4.66 (317)	0.0377	0.0727	15.00	7.81	17.55
Zavalla, Venado Tuerto, Rafaela, Manfredi	5 (181)	0.0328	0.0729	14.89	7.92	19.71
Rafaela, Manfredi	5.66 (123)	0.0278	0.0844	14.30	7.84	17.87
Venado Tuerto, Rafaela, Manfredi	6 (240)	0.0247	0.0994	14.11	7.66	15.31
Rafaela, Manfredi	7 (61)	0.0231	0.0920	14.00	7.78	21.40
Rafaela	8 (67)	0.0228	0.0539	15.59	7.77	7.16

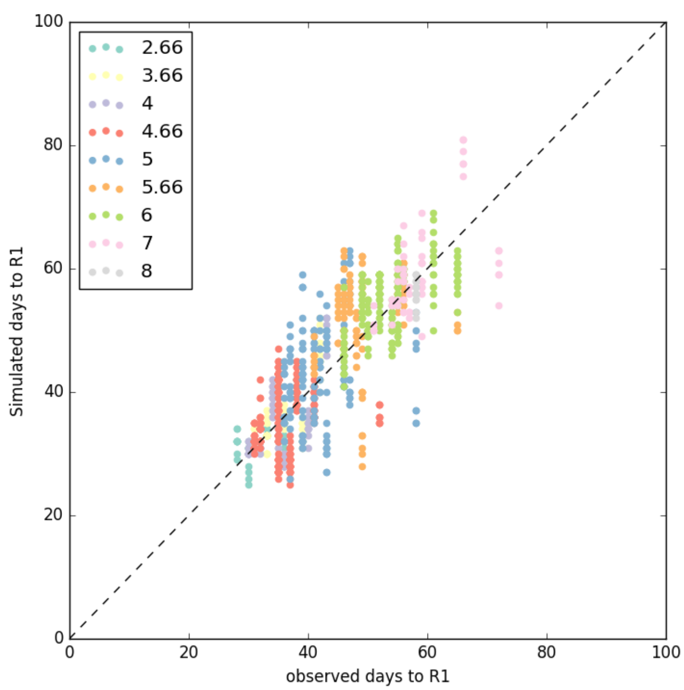


Figure 3-5 Observed versus simulated days from emergence to the R1 stage for different soybean maturity groups

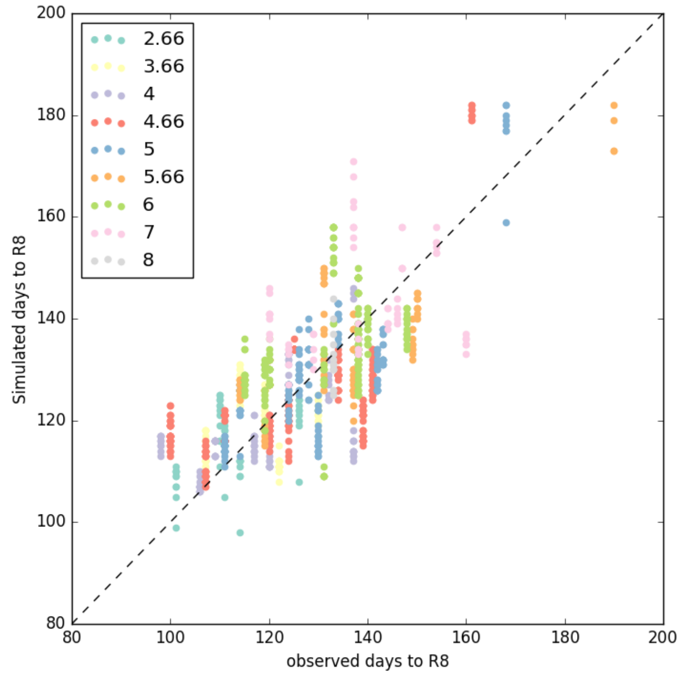


Figure 3-6 Observed versus simulated days from emergence to the R8 stage for different soybean maturity groups

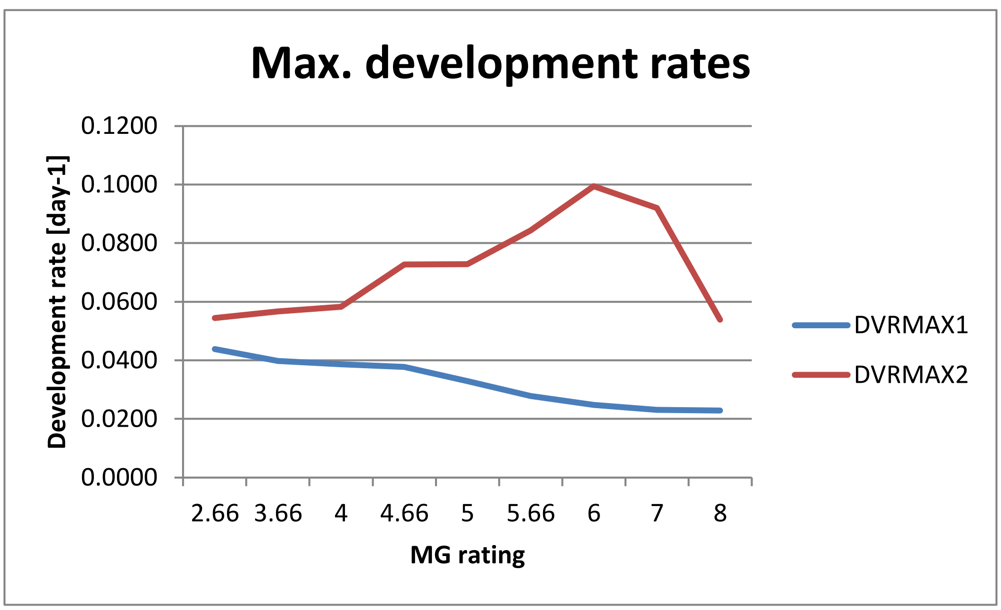


Figure 3-7 Calibrated maximum development rates for the different soybean maturity groups

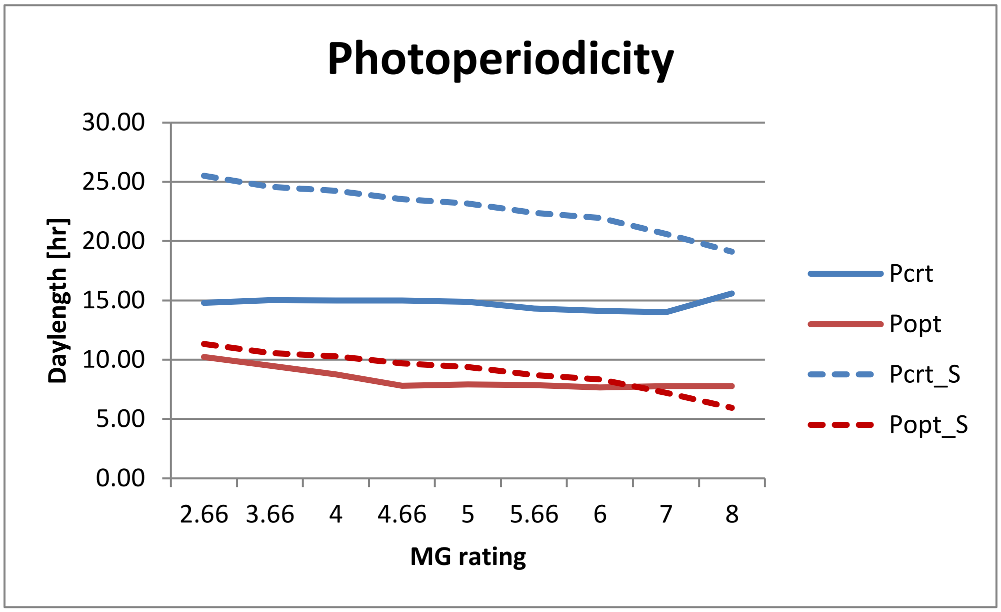


Figure 3-8 Calibrated parameter for day length sensitivity for the different soybean maturity groups (label '_S' indicates the parameters from Setiyono et al., 2007)

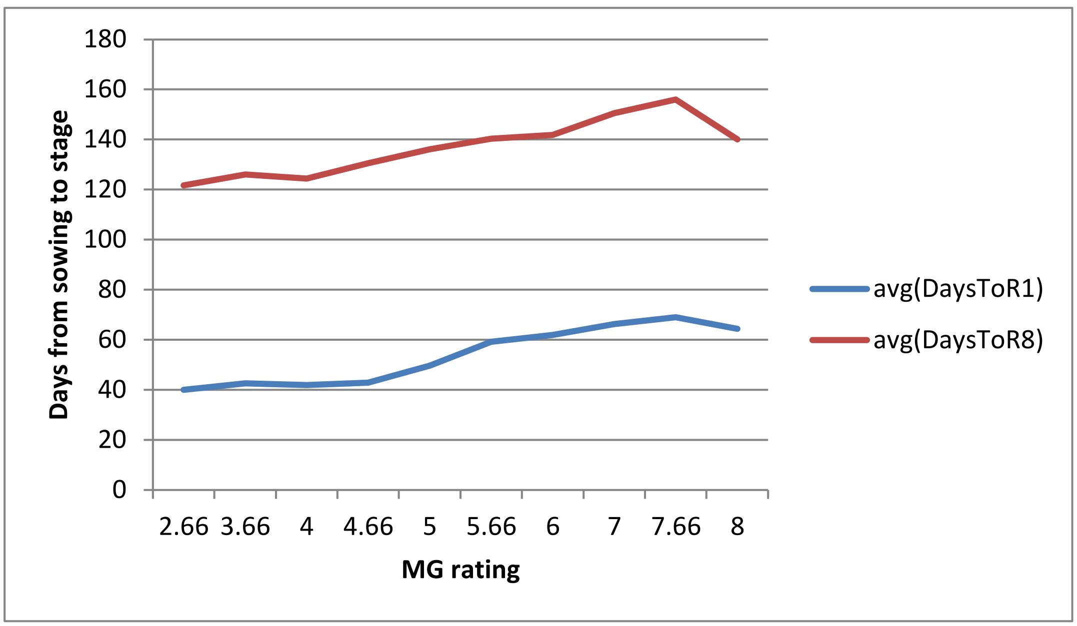


Figure 3-9 Average number of days from sowing to flowering (R1) and from sowing to maturity (R8) per maturity group

3.5. Calibration results: Yield level

3.5.1. Strategy

In order to produce maps of the yield gap of soybean the potential yield level need to be estimated using output from the WOFOST model. However, calibration of the potential yield levels is less straightforward compared to the phenological development because the field experiments were not carried out under optimal conditions and therefore are not representative of potential yield levels. Moreover, the observations do not provide observations of total crop biomass and leaf area index and thus the crop LAI development and increase in total biomass cannot be quantitatively validated.

However, for estimating potential yield levels we can apply some rules which ensure that that the simulated yield levels are plausible:

- Crop leaf area index (LAI) should be within the plausible range. Measurements of LAI at many crops demonstrate that plausible ranges of maximum LAI are generally between 4 and 7 which ensures full light interception (LAI>4) during a considerable part of the growing season.
- The harvest index (ratio between yield and total biomass) must be realistic. Literature values for soybean harvest index (HI) vary from 0.4 to 0.6 (Kakiuchi and Kobata, 2006; Bajgain et al. 2015). Kakiuchi and Kobata also indicate that the harvest index is a relatively conservative property across environments and growing conditions.
- Behaviour of total biomass, yield and harvest index should be consistent with growing season length.
- Parameter values of the model should not become unrealistic.
- Earlier results from the GYGA project and results from INTA with other crop simulation models indicated a potential yield level of 6 to 7 ton/ha (at 13% moisture level)

3.5.2. Yield levels with default parameter values

First we evaluated the results using the original parameterization, an example of these results is shown in Figure 3-10 for maturity group 5. In most cases actual yield levels are higher than the simulated yield levels. Note that yields are expressed in 0% moisture. Given the fact that those experiments were carried out under non-optimal conditions it clearly demonstrates that the potential yield levels simulated by WOFOST are too low. A detailed analysis of the results (Figure 3-11) shows that 1) the maximum leaf area index is relatively high followed by a rapid decline in LAI; 2) crop total biomass is relatively low mainly between 8 and 11 ton/ha, moreover the yield component of the crop is particularly low (3-4 ton/ha). The combination results in a harvest index between 0.25 and 0.35 which is low compared to the values in literature.

Based on the detailed results we concluded that the WOFOST parameters needed to be adapted at multiple levels:

- Crop total biomass is too low, indicating that an increase in maximum assimilation rate (AMAX) is needed. Existing crop files for WOFOST and the values tabulated by Van Heemst (1988) already indicate a considerable variability in AMAX estimates ranging from 29 to 37 kg CO_2 ha^{-1} $hour^{-1}$.
- The harvest index is too low indicating that the portion of assimilates allocated to the storage organs is too low. Given that we made major changes in the simulation of phenology, it is not surprising that changes in the allocation pattern are needed which ensure that more assimilates are allocated to the storage organs.
- Changes in the above parameters will probably require adjustment of the SPAN (life span of leaves) and the SLATB (specific leaf area) parameters.

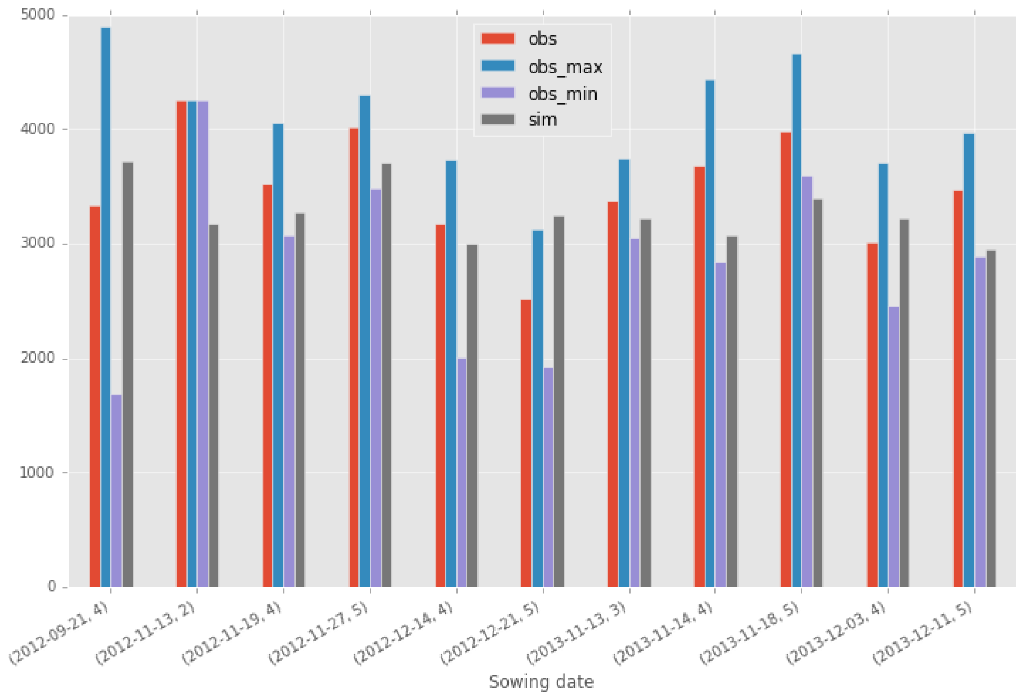


Figure 3-10 Simulated vs observed yield (expressed as 0% moisture) for cultivars in maturity group 5 using the original WOFOST parameterization for soybean. Red, blue, purple bars represent the observed mean, maximum and minimum yield. The grey bar represents the simulated yield. Different sets of bar charts represent different combination of sowing date and location.

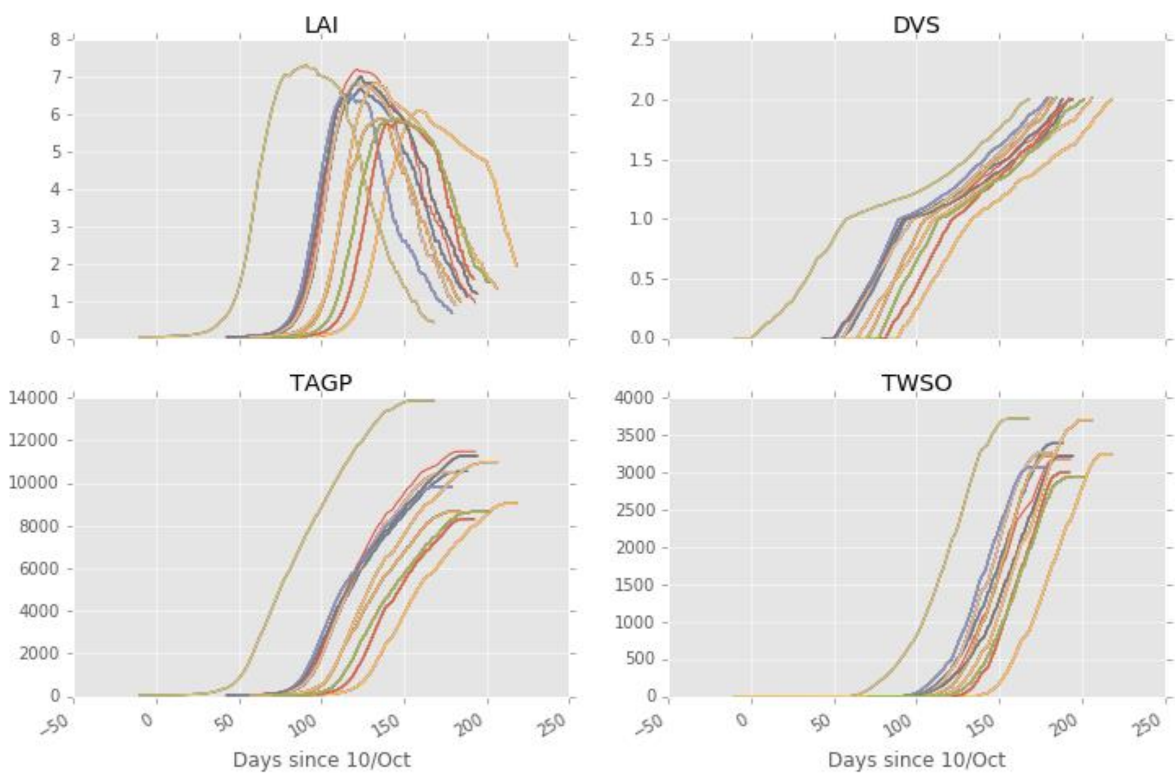


Figure 3-11 Detailed analysis of the behavior of the simulation results for maturity group 5: results shown are leaf area index (LAI), development stage (DVS), total crop biomass (TAGP) and crop yield (TWSO) (biomass and yields expressed in 0% moisture), different colors represent different experiments

3.5.3. Changes to crop parameters

Based on the conclusions above, we made several changes to crop parameters. First of all, we adjusted the default AMAX value (29 kg CO_2 ha^{-1} $hour^{-1}$) to the maximum value reported by Van Heemst (1988) (37 kg CO_2 ha^{-1} $hour^{-1}$). This change increased the total biomass produced as a result of a higher gross photosynthesis rate.

Next, the allocation pattern to storage organs was adjusted as is shown in Figure 3-12. The new allocation pattern starts to allocate assimilates to the storage organs at an earlier DVS resulting in a higher proportion of assimilates ending up in the storage organs. Also note that after DVS=1 there is a reduced development rate due to the effect of day length. The new allocation pattern takes this into account as the allocation to storage organs is increasing steeply after DVS=1 (Figure 3-12). The increased allocation to storage organs also meant that the allocation to stems and leaves was changed (decreased) as well.

The change in allocation pattern also has an impact on the LAI development during the growing season. The specific leaf area was lowered to avoid excessive high leaf area while the live span of leaves was increased in order to sustain the crop canopy a bit longer in the growing season. Finally, the maintenance respiration rate for storage organs awas decreased slightly in order to have a larger net photosynthesis.

For an overview of all crop parameters and changes, see appendix I.

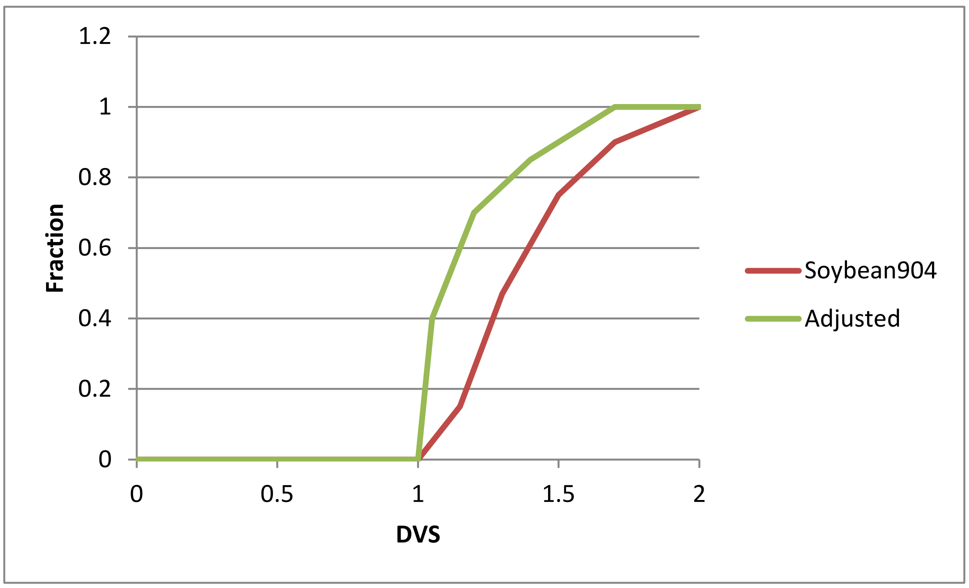


Figure 3-12 Allocation pattern of assimilates to storage organs as a function of development stage for the old parameter file (soybean904) and the adjusted version

3.5.4. Evaluation of the new crop parameters

As there are no observations available to quantitatively validate the simulation results, we therefore looked at the plausibility of the results according to the criteria that we set up in section 3.5.1.

Figure 3-13 shows the detailed simulation results with the new crop parameters. It clearly demonstrates that the model now produces more realistic output, according to the different criteria:

- Excessive high LAI has been decreased to more realistic values.
- Total crop biomass increased to range of 9 to 12 ton/ha.
- The change in allocation pattern increased the yield level (storage organs) to more realistic values between 5 and 7 ton/ha which are consistently higher than the observed yields (Figure 3-14). This figure also demonstrate the systematic higher yields (6-7 ton/ha) as results of early sowing (November) compared to late sowing in December (5-6 ton/ha).
- The harvest index is now within the realistic derived from the literature (Figure 3-15).
- Yield levels are now consistent with results from the GYGA project.

Finally, we show the distributions of the different variables (TAGP, TWSO, HI, MAXLAI) across the different maturity groups as box plots in Figure 3-16. For total biomass (TAGP) the results demonstrate that with increasing maturity group, the total biomass increases. This is the expected behavior given that the length of the growing season increases with increasing maturity group.

The average crop yield (TWSO) is nearly constant for the first 4 maturity groups although the highest yields are found for maturity groups 4.66 and 5. This is not surprising as these are the maturity groups that are most suitable for the region and which are used most commonly (see also Figure 2-2). For maturity groups higher than 5, there is a gradual decline in crop yield (TWSO) as a result of the grain filling period being delayed and the conditions for grain filling becoming less favorable (lower temperature and radiation).

The harvest index shows a consistently declining pattern which can be explained from the decreasing yield for high maturity groups. Additionally, the ratio of vegetative versus reproductive length of the season changes over the maturity groups. Figure 3-9 demonstrates that the reproductive phase is nearly equal across all maturity groups, however cultivars in a low maturity group have a shorter vegetative phase. As the harvest index under potential conditions reflects the ratio of (reproductive phase)/(vegetative + reproductive phase) it is not surprising that the harvest index decreases. For low maturity groups the harvest index currently is somewhat higher than the values found in the literature.

Finally, the maximum leaf area index shows a systematic increase with maturity group number which can be directly related to the increase in the length of the vegetative phase. We have no measurements to verify whether this pattern is realistic.

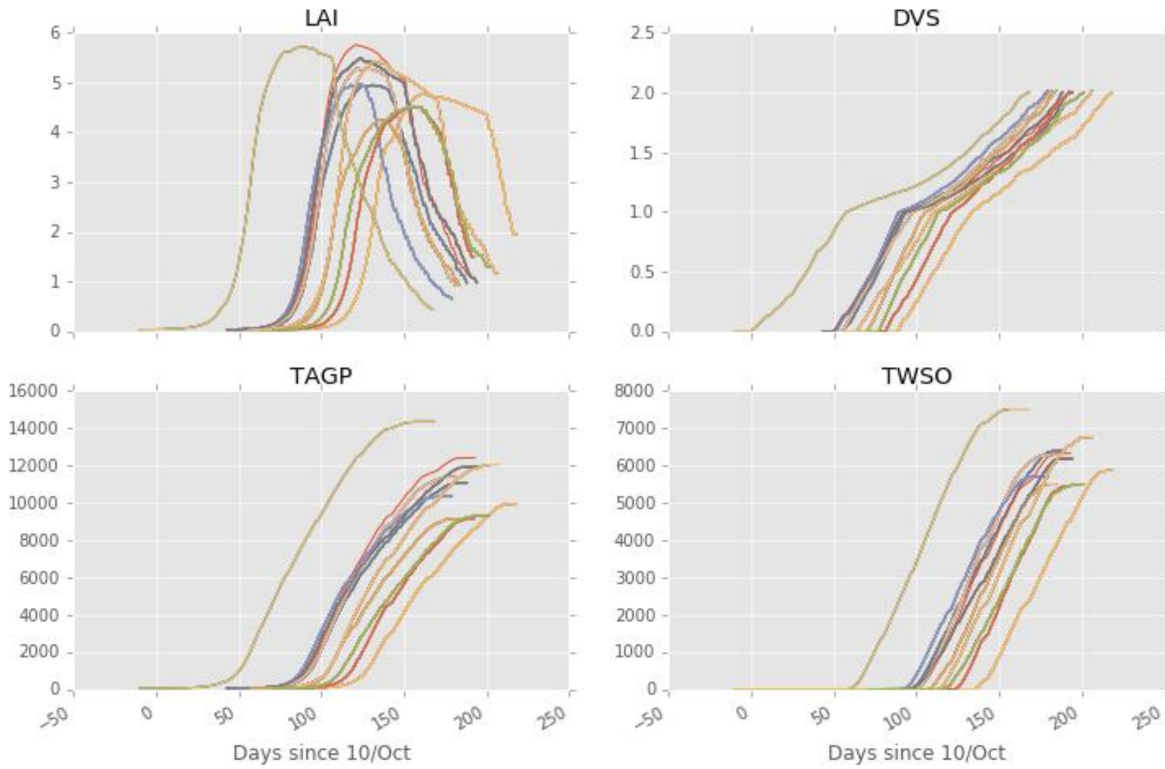


Figure 3-13 Detailed analysis of the behavior of the simulation results for maturity group 5 with new crop parameters: results shown are leaf area index (LAI), development stage (DVS), total crop biomass (TAGP) and crop yield (TWSO) (biomass and yields expressed in 0% moisture) different colors represent different experiments

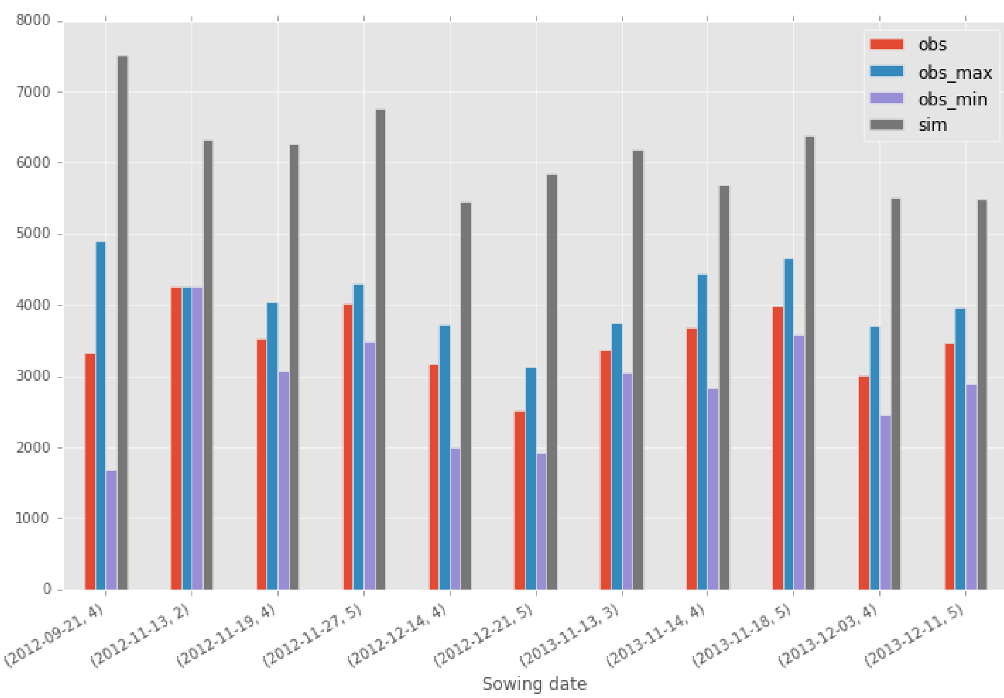


Figure 3-14 Simulated vs observed yield (expressed as 0% moisture) for cultivars in maturity group 5 using the new parameterization for soybean. Red, blue, purple bars represent the observed mean,

maximum and minimum yield. The grey bar represents the simulated yield. Different sets of bar charts represent different combination of sowing date and location.

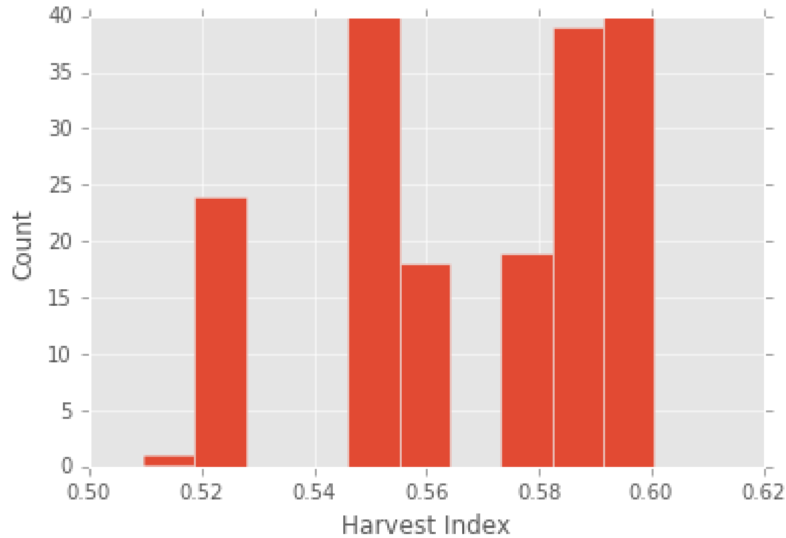


Figure 3-15 Distribution of the simulated harvest index for Maturity Group 5

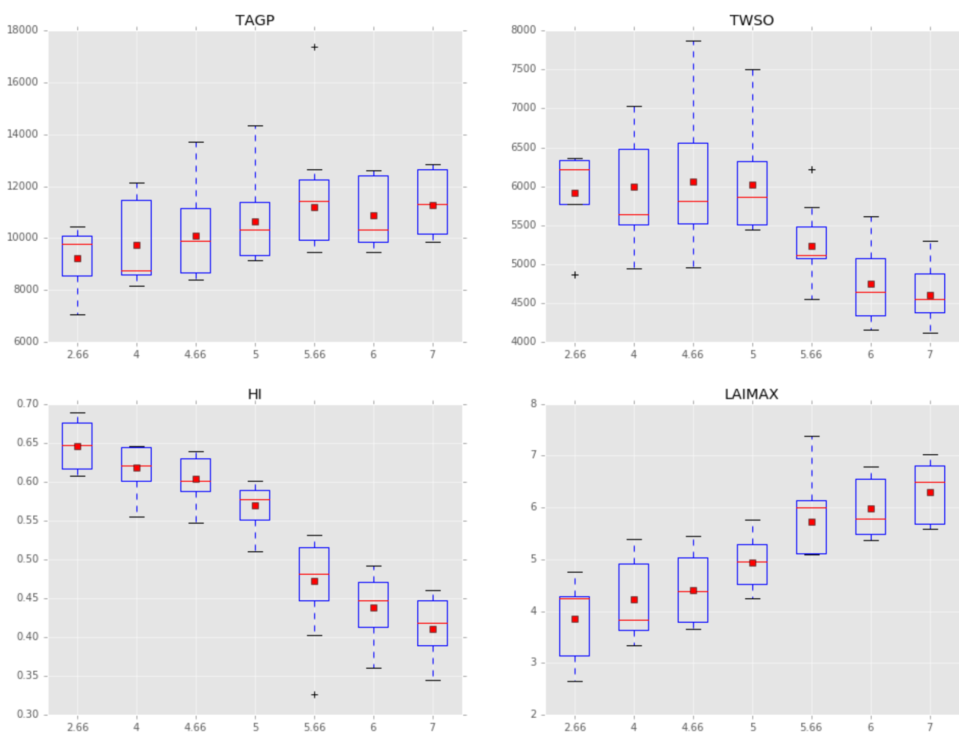


Figure 3-16 Distributions of total biomass (TAGP), yield (TWSO), harvest index (HI) and maximum LAI (LAIMAX) across the different maturity groups for soybean (biomass and yields expressed in 0% moisture)

4. Regional implementation

To arrive at spatial explicit maps of potential yield levels for soy beans a regional version of WOFOST-soybeans has been set-up. The regional implementation includes the following main components to ensure a full spatial coverage:

- Downscaled daily ECMWF ERA-Interim model weather data at a 0.25 degree grid obtained from JRC-MARS (JRC_ERA-INTERIM, see section 2.1.1)
- Daily rainfall from CHIRPS, aggregated to a 0.25 degree grid (CHG_CHIRPS, see section 2.1.3).
- Gridded soil data from WISE30SEC
- Local calibrated varieties linked to a zonation that combines the GYGA-ED zones and a south-north zonation describing different soy beans Maturity Groups.
- SIGMA specific spatial schematization combining: 1) the 0.25 degree grid; 2) the WISE30SEC soil grids and 3) the variety zonation. The daily weather is linked to the 0.25 degree grid; soil data originates from the WISE30SEC soil grids (1/120 degree); calibrated varieties are linked to the variety zonation (0.08333 degree).

The following sections give more information on the weather, crop and soil data.

4.1. Daily weather data

To obtain wall-to-wall daily weather data on a certain predefined grid there are basically two sources possible:

- Observations from weather stations that need to be interpolated to the grid
- Gridded data sets from models (weather simulation, satellite based data assimilation models etc.) that need to be scaled to the grid

Using weather stations is recommended when station density is high and the main elements are available for the desired period. If this is not the case gridded data sets can be a good alternative provided it satisfies certain criteria like a daily time step, a spatial resolution capturing the main spatial variability and a certain accuracy. In our case we had access to daily data of weather stations. However data were incomplete and some elements like radiation were missing. Instead of setting-up a procedure to process, repair and complete daily station weather, we decided to evaluate gridded data sets and select the best performing set for the regional WOFOST database. We have selected the following gridded data sets:

- JRC_ERA-INTERIM (see section 2.1.1)
- NASA_POWER (see section 2.1.2)
- NASA_TRIMM (see section 2.1.4)
- CHG_CHIRPS (see section 2.1.3)

To evaluate the accuracy an independent validation set was constructed. INTA made daily data of 178 stations available. Complicating factor is that the gridded data sources also use station observations in their assimilation schemes. However we do not exactly know which station data are used. We assumed that these products mainly use data of stations that are regularly transmitted by the national meteo office via the GTS. These data are available within the database of NOAA-GSOD. We performed a cross-check identifying those stations that are very close (< 10 km) to the stations of the NOAA-GSOD data set. The station-day combinations of those stations (see 4_Select_Similar_Stations_And_Dates.xlsx) were excluded from the validation data set.

For each of the 178 stations the most nearby grid cell was selected for each different data source. Minimum and maximum temperature, radiation and precipitation were included. In the vegetation and drought monitoring domain data are usually available at dekadal time steps. There the comparison was done at dekadal time steps and data were aggregated. For precipitation the sum was taken while for the others the average was taken. With regard to precipitation the following procedure was followed:

- Exclude all dekads where weather station data was >0 and <=5
- Assign hit (A) in case gridded data source has a value > 0 mm and the station has a value >= 5 mm
- Assign false alarm (B) in case gridded data source has a value > 0 mm and the station has a value = 0 mm
- Assign miss (C) in case gridded data source has a value = 0 mm and the station has a value >= 5 mm
- Assign correct negative (D) in case gridded data source has a value = 0 mm and the station has a value = 0 mm

Table 4-1 shows the validation statistics that were determined per GYGA-ED zone for minimum and maximum temperature, radiation and precipitation.

Table 4-1 Validation statistics

Statistic	Formula	Perfect score
Relative mean absolute error	$RMAE = \frac{\frac{1}{N} * \sum(\lvert G - O \rvert)}{\bar{O}}$	0
Bias	$Bias = \frac{\sum G}{\sum O}$	1
Probability of detection[3]	$POD = \frac{A}{(A + C)}$	1
False alarm ratio[3]	$FAR = \frac{B}{(A + B)}$	0

For each element and GYGA-ED zone the gridded data set were first ranked for each validation statistic with 1 (best) and 2 (worst) and in case of precipitation 1 (best) to 4 (worst). Finally these rankings were averaged for each gridded data set showing which gridded data performs best for the selected element and GYGA-ED zone. In case of precipitation the rankings of RMAE and BIAS received a weight of 1/3 each and the rankings of POD and FAR were weighted for 1/6 each.

The validation covered a much wider area than the regions for which WOFOST was set-up. It covered 23 GYGA-ED zones. Table 4-2,
Table 4-3 and Table 4-4 show the results for the 3 GYGA-ED zones covered by the regional implementation of WOFOST soy beans. Appendix II gives the underlying validation statistics for the gridded data sets, GYGA-ED zones and elements.

[3] Only for precipitation; based on a contingency table with A= number of hits, B = number of false alarms, C = number of misses

Table 4-2 Ranking of gridded data sources for minimum and maximum temperature, radiation and precipitation for GYGA-ED zone 6302

	Maximum temperature	Minimum temperature	Radiation	Precipitation
JRC_ERA-INTERIM	1.5	1.5	2	3.5
NASA_POWER	1.5	1.5	1	2.2
NASA_TRIMM				1.8
CHG_CHIRPS				2.5

Table 4-3 Ranking of gridded data sources for minimum and maximum temperature, radiation and precipitation for GYGA-ED zone 6402

	Maximum temperature	Minimum temperature	Radiation	Precipitation
JRC_ERA-INTERIM	1	1.5	2	3.5
NASA_POWER	2	1.5	1	2.7
NASA_TRIMM				2.2
CHG_CHIRPS				1.7

Table 4-4 Ranking of gridded data sources for minimum and maximum temperature, radiation and precipitation for GYGA-ED zone 6502

	Maximum temperature	Minimum temperature	Radiation	Precipitation
JRC_ERA-INTERIM	1	1	2	3.5
NASA_POWER	2	2	1	2.7
NASA_TRIMM				2.2
CHG_CHIRPS				1.7

Finally, it was decided to take the data from the JRC_ERA-INTERIM data source except precipitation which was taken from the CHG_CHIRPS and global radiation which was taken from NASA_POWER. Table 4-5 summarizes some validation statistics for the selected gridded data source.

Table 4-5 Summary of validation statistics for selected gridded data sources

Element	Gridded data source	Analysis
Maximum temperature	JRC_ERA-INTERIM	High correlation (97%), bias showing lower values (-1.5 to -1 degrees going from west to east)
Minimum temperature	JRC_ERA-INTERIM	High correlation (93-97%), bias showing higher values (2.5 to 1.4 degrees going from west to east)
Radiation	NASA_POWER	Moderate correlation in west (77%), high correlation in other two GYGA-ED zones (95-98%), bias showing higher values around 1.5 ($MJ.m^{-2}.d^{-1}$)
Precipitation	CHG_CHIRPS	Moderate correlation of around 71-77%, bias showing higher values between 1.6 – 3.3 (mm per dekad). Performs moderate on FAR. In fact

		NASA_TRMM performs best on FAR and JRC_ERA-INTERIM performs best on POD but both overestimates precipitation amounts more than CHG_CHIRPS

Data of CHG_CHIRPS and NASA_POWER were linked to the grid definition of JRC_ERA-INTERIM applying a nearest neighbor procedure.

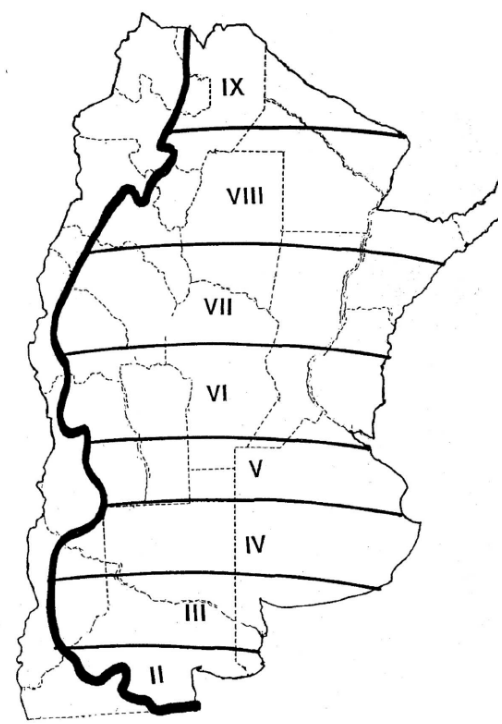
4.2. Crop data

4.2.1. Variety definition

For the calibration of WOFOST soybean, first sites were selected representing different GYGA-ED climate zones namely 6302, 6402 and 6502 following rainfall gradient from dry to wetter conditions (see Figure 1-1).

More information on the GYGA-ED zonation and coding can be found through: http://www.yieldgap.org/documents/10180/35397/The%20Global%20Yield%20Gap%20Atlas%20Extrapolation%20Domain%20documentation.pdf.

However the GYGA-ED zones in the region of Argentina do not capture the south-north gradient in Maturity Group (MG) for soy beans (see chapter 3 for more information on maturity groups). To add the MG-gradient, GYGA-ED zones were combined with a MG-zonation, provided by INTA (Figure 4-1).



Fuente: H. Baigorri. 1991 (adaptado de Remussi y Pascale. 1977)

Figure 4-1 Maturity group zonation for soy beans Argentina

It led to the following variety zonation (Figure 4-2) with two gradients:

- West – east (dry to wet)
- South – north (increasing maturity groups)

The selected sites are mainly situated in the MG zone IV. However the experiments of these sites covered different MGs:

- La Carolota (Córdoba): II-III Short, III Long, IV Short, IV Long, V Short
- Manfredi (Córdoba) and La Carlota (Córdoba): II-III Short, III Long, IV Short, IV Long, V Short, V Long, VI, VII Short
- Rafaela (Santa Fe): IV Short, IV Long, V Short, V Long, VI, VII Short, VII Long - VIII
- Venado Tuerto (Santa Fe): II-III Short, III Long, IV Long, V Short, V Long, VI
- Zavalla (Santa Fe): III Short, III Medium, III Long, IV Short, IV Medium, IV Long

The calibration (see chapter 3) resulted in soy bean parameterizations per MG. The rainfall oriented west-east gradient represented by the GYGA-ED zonation was finally not used to differentiate varieties.

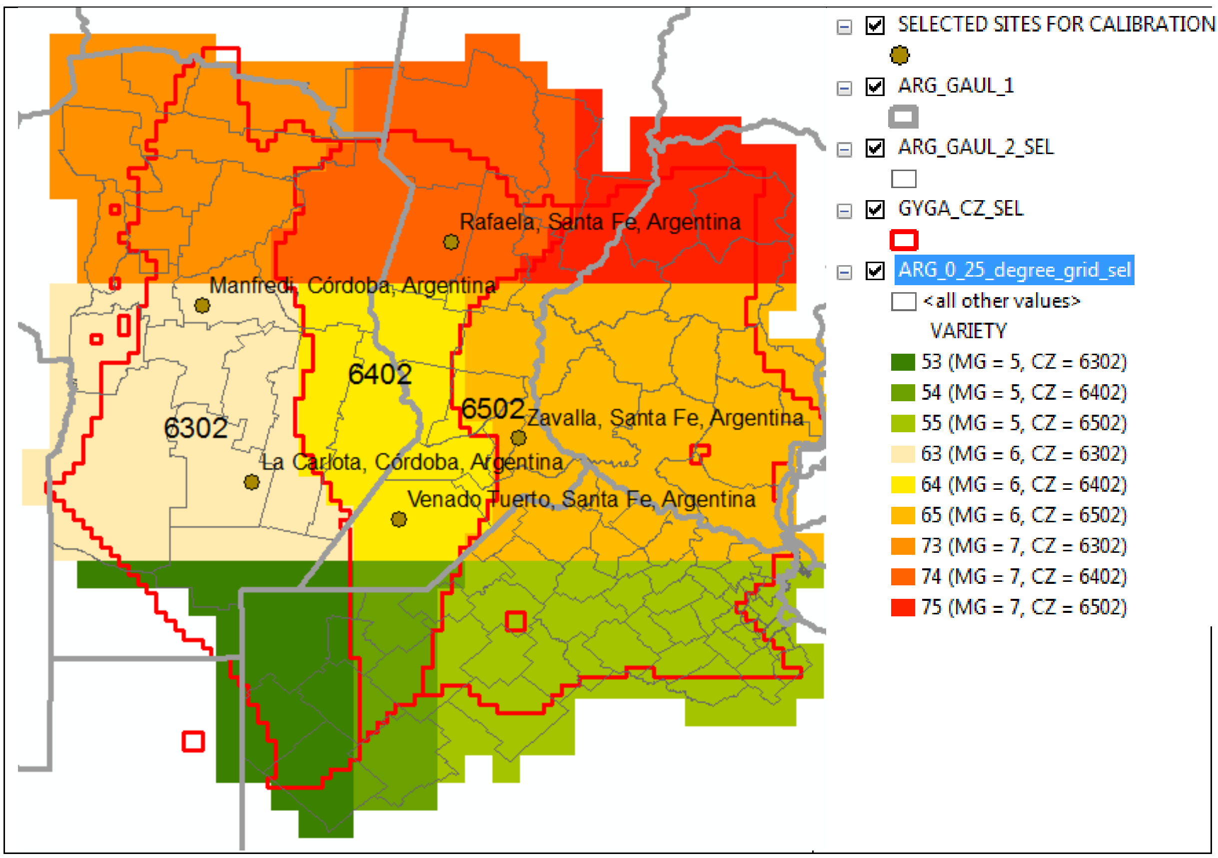


Figure 4-2 Variety zonation resulting from an overlay of the MG-zonation and the GYGA-ED zonation

During the simulation tests we found a sharp gradient in the spatial pattern of yields between the MG 5 zones and the other MG zones 6 and 7. This is caused by the allocation patterns that depend solely on phenology (DVS). When reaching flowering (R1) the development rate (DVR) vastly reduces because of day length sensitivity. In case of MG=6 the critical day length is somewhat lower than that for MG=5: 14.11 h vs 14.89 h. In case of MG=6 the DVR equals zero for some days so it keeps stuck around flowering. To avoid this gradient a more detailed calibration is needed changing Pcrt and DVRMAX2. As day length is very close to the critical day length the calibration can be rather unstable. The above illustrates the limitation of the current concept of allocation of assimilates in WOFOST solely based on DVS. Other approaches (e.g. SoySim) describe the development of pods/beans to estimate the demand of assimilates. This could be a good alternative. Considering the limited time and resources within the project we decided to use the crop parameterization of MG 5 for all zones as this is regarded as the most stable solution.

4.2.2. Crop calendar

In this study we focused on a soy bean crop within a single cropping system thus one soy bean season.

It has the following long term average emergence dates:

- South (maturity group V): 10 November
- Centre (maturity group VI): 15 November
- North (maturity group VII): 25 November

This is based on the following sources listed in Table 4-6 and Table 4-7 (locations are shown in Figure 4-3).

Table 4-6 Soy bean crop calendar (provided by INTA)

Location	**Emergence**	**Sowing**	**Estimated value emergence**
Zavalla, Santa Fe, Argentina		1-11 – 13-11	15-11
Venado Tuerto, Santa Fe, Argentina	20-11 – 13-12		1 – 12
Rafaela, Santa Fe, Argentina	21-11 – 9-12		1 – 12
Manfredi, Córdoba, Argentina	17-11 - 17-12		3- 12
La Carlota, Córdoba, Argentina	19-11		19 -11

Table 4-7 Soy bean crop calendar taken from GYGA (http://www.yieldgap.org; accessed February 2017)

Location	**Sowing**	**Estimated value emergence – single season**
Pilar	25-11 (single) / 28 – 11 (second)	2 - 12
Río Cuarto	25-10 (single) / 5-12 (second)	1 - 11
Laboulaye	25-10 (single) / 5-12 (second)	1 - 11
General Pic	5-11 (single) / 10-12 (second)	12 - 11
Pehuajo	1-11 (single) / 15-12 (second)	8 - 11
Rafaela	15-11 (single) / 20-12 (second)	22 - 11
Marcos Juárez	25-10 (single) / 5-12 (second)	1 - 11
Pergamino	1-11 (single) / 10-12 (second)	8 - 11
Paraná	15-11 (single) / 5-12 (second)	22 - 11
Gualeguaychú	1-11 (single) / 5-12 (second)	8 - 11
Mercedes	10-11 (single) / 9-12 (second)	17 - 11
Young	10-11 (single) / 2-12 (second)	17 - 11

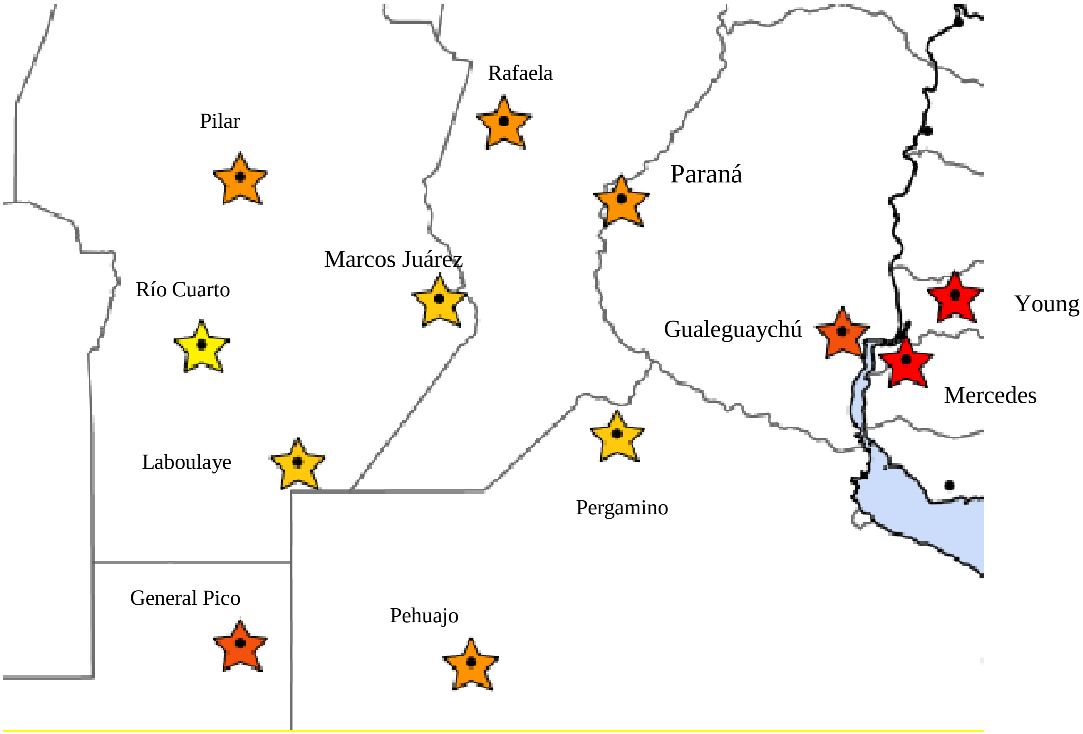


Figure 4-3 Locations for which yields levels of soybeans have been modelled in GYGA

Based on these sources the following long term average start was determined as input for the crop simulations (see Table 4-8).

Table 4-8 Emergence dates

Variety zone (see Figure 4-2)	Emergence date
53	10 November
54	10 November
55	10 November
63	15 November
64	15 November
65	15 November
73	25 November
74	25 November
75	25 November

In summary soy beans are simulated with WOFOST starting at emergence (see Table 4-8) and following a certain parameterization, calibrated for selected sites (see Chapter 3) and linked to a variety zonation (see Figure 4-2).

4.3. Soil data

The WISE30SEC version 1.0 soil database was selected (Batjes, 2015). It has some advantages compared to the previous WISE v 1.2 database. For instance the Harmonized World Soil Database (HWSD) is used instead of the Digital Soil Map of the World (DSMW). The former includes more detailed soil maps for some parts of the world like Europe and China (1 to 1 million) while the latter only includes the 1 to 5 million global soil map. Moreover soil characteristics like TAWC are based on twice as much soil profiles increasing the accuracy of these characteristics. Climate data has been used to improve the geo-mapping of soil profiles.

Data have been loaded in tables HW30S_FULL and HW30S_MAPUNIT and processed determining soil physical characteristics (rooting depth and available water capacity) for each soil type unit. This was done by views and scripts according the following steps:

1) Import tables HW30S_FULL and HW30S_MAPUNIT in ORACLE schema
2) Create views to extract a structure:

```
create or replace view soils_smu
as
select
   to_number(substr(newsuid,4)) as idsmu
from hw30s_mapunit;

create or replace view soils_stu
as
select
    to_number(substr(newsuid,4))                    idsmu
  , scid                                            idstu
  , prop                                            percentage
  , max(botdep)                                     soil_rooting_depth
  , sum(tawcc*(botdep-topdep))/sum(botdep-topdep) available_water_capacity
from
  (select
       newsuid              newsuid    -- soil map unit id
     , scid                 scid       -- soil type id
     , prop                 prop       -- proportion of soil type in map unit
     , topdep               topdep     -- top depth of layer
     , botdep               botdep     -- bottom depth of layer
     , (1-(cfrag/100))*tawc tawcc      -- tawc corrected for coarse fragments > 2mm
  from hw30s_full
  where botdep <= 100                  -- skip layers deeper than 1 meter
  and   tawc   >= 0                    -- skip records with -1, -2, -3, -4, -5, -7, -9
  and   to_number(substr(newsuid,4)) not in (0,6997,6998) -- skip records with suid 0, 6997,
6998
)
group by to_number(substr(newsuid,4)), scid, prop
order by to_number(substr(newsuid,4)), scid, prop;
```

The list of soil mapping units is stored in table SOILS_SMU while soil mapping composition and the soil physical characteristics per soil typologic unit are stored in table SOILS_STU. These two tables were used to fill the regional WOFOST database.

In the final simulation only suitable soil components are included. These are components having a rooting depth class >= 1 and an available water capacity of 0.0 or more thus effectively all soils are included.

Finally, it is assumed that all water infiltrate as long as the soil can absorb the water. Percolation rates are set to10 cm per day and a critical air content for oxygen stress of roots was set to 0.06.

Figure 4-4 demonstrates the spatial patterns of the soil associations.

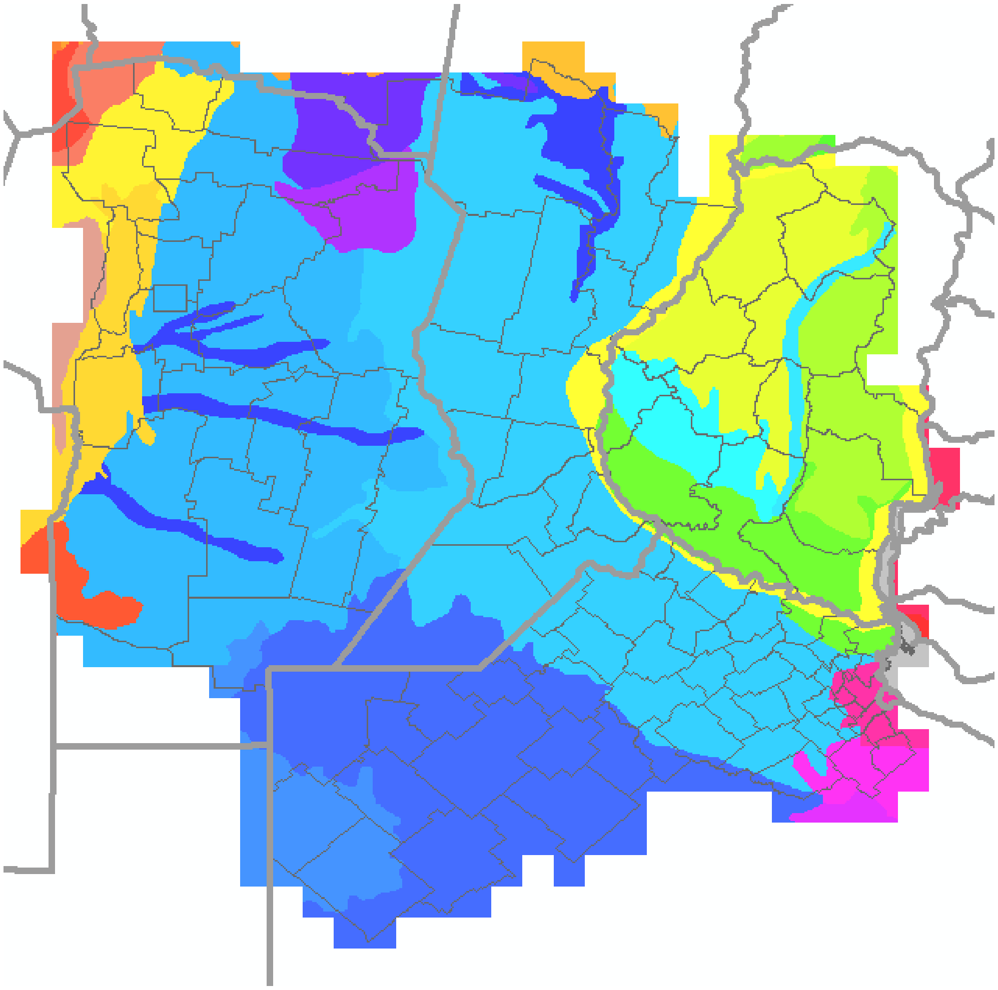


Figure 4-4 Map of soil associations from the WISE30SEC database

4.4. Regional WOFOST database

The table below gives more details on specific data.

Table	Content
CROP	Soy beans
CROP_CALENDAR	Grid and year specific definition of start (emergence), end (maturity) and a grid specific variety
CROP_GROUP	Crop group of crops
CROP_PARAMETER_VALUE	Crop parameterization of crops (see chapter 3)
DATES_75_15	Mapping between dates and dekads
EMU	Overlay between the 0.25 degree grid and WISE30SEC grid resulting in unique soil associations and their area share within a 0.25 degree grid cell
GRID	JRC-MARS global grid definition: 0.25 degree spatial resolution in WGS84 co-ordinate system
PARAMETER_DESCRIPTION	Description of crop parameters
INITIAL_SOIL_WATER	Grid and soil component specific soil water available at the start of the soil water initialization which starts 60 days before emergence. Initial soil water at 60 days before emergence is 0.5 time the Available Water Capacity
ROOTING_DEPTH	List of rooting depth classes (20, 40 and 100 cm)
SIMULATION_UNIT	Unique combinations of 0.25 degree grid cells and soil components for each crop
SITE	System wide parameters on infiltration (no surface run-off)
SMU_SUITABILITY	List of soil associations and area percentage of suitable soils based on suitable soil components
SOIL_ASSOCIATION_COMPOSITION	Mapping between soil associations and their soil components of WISE30SEC taken from table SOILS_STU
SOIL_MAPPING_UNIT	List of unique soil associations of WISE30SEC taken from table SOILS_SMU
SOIL_PHYSICAL_GROUP	List of distinct values of available water capacity of WISE30SEC taken from table SOILS_STU. Each value lead to a unique soil group number and associated soil moisture values for pF 2.5 (field capacity) and pF 4.2 (wilting point) and saturation by first defining 0.1 for wilting point, available water capacity plus 0.1 for field capacity and available water capacity plus 0.2 for saturation. Percolation rates of 10 cm per day and a critical air content of 0.06.
SOIL_TYPOLOGIC_UNIT	List of unique soil components of WISE30SEC taken from table SOILS_STU and introducing unique number of distinct rooting depth values and unique number of soil groups (see table SOIL_PHYSICAL_GROUP)

STAT_CROP	Mapping between simulation and statistical crop (in this case 1 to 1)
SUITABILITY	List of suitable soil components have a rooting depth class >= 1 and an available water capacity of 0.0 or more
SYSCON	System wide parameters to run WOFOST
VARIETY_PARAMETER_VALUE	Variety specific crop parameters (see chapter 3)
WEATHER_OBS_GRID	MARS down-scaled daily ECMWF ERA-Interim model data for years 1989-2015. Rainfall is taken from CHIRPS.
GRID25KM_GYGA_ED	Mapping between the 0.25 degree grid and the variety zonation based on an overlay between the two grids
GYGA_ED_CROP_DATA	Crop and variety zone specific start date (emergence)

4.5. Simulation runs

Runs start 60 days before emergence to run a climatic water balance and come to more realistic initial soil water.

The simulation period spans the years 1989 – 2014 (26 years). Simulation was carried out with the WOFOST implementation in PCSE:

- Under full irrigation: WofostSoybean & WaterbalancePP
- Under rain fed conditions: WofostSoybean & WaterbalanceFD

More information can be found via the following link: http://pcse.readthedocs.io/en/latest/.

Simulation results of the unique combinations of 0.25 degree grid cells and soil components are aggregated to unique combinations of 0.25 degree grid cells and soil associations, allowing spatial mapping of the results.

The following key characteristics have been derived to evaluate the simulations and to assess the potential yield levels (grain yield) under irrigated and rain-fed conditions:

- Average maximum Leaf Area Index (plausible range: 4-7)
- CV of maximum Leaf Area Index (plausible range: 0 – 0.2)
- Average final Leaf Area Index (plausible range: 0 - 1)
- CV of final Leaf Area Index (plausible range: 0 – 0.2)
- Average above ground biomass, 0% moisture (plausible range: 4000 - 13000)
- CV of above ground biomass (plausible range: 0 – 0.2)
- Average grain yield, 0% moisture (plausible range: 2000 - 6000)
- CV of grain yield (plausible range: 0 – 0.2)
- Average final development stage (plausible range: 195-200)
- Average harvest index (plausible range: 0.4 – 0.6)

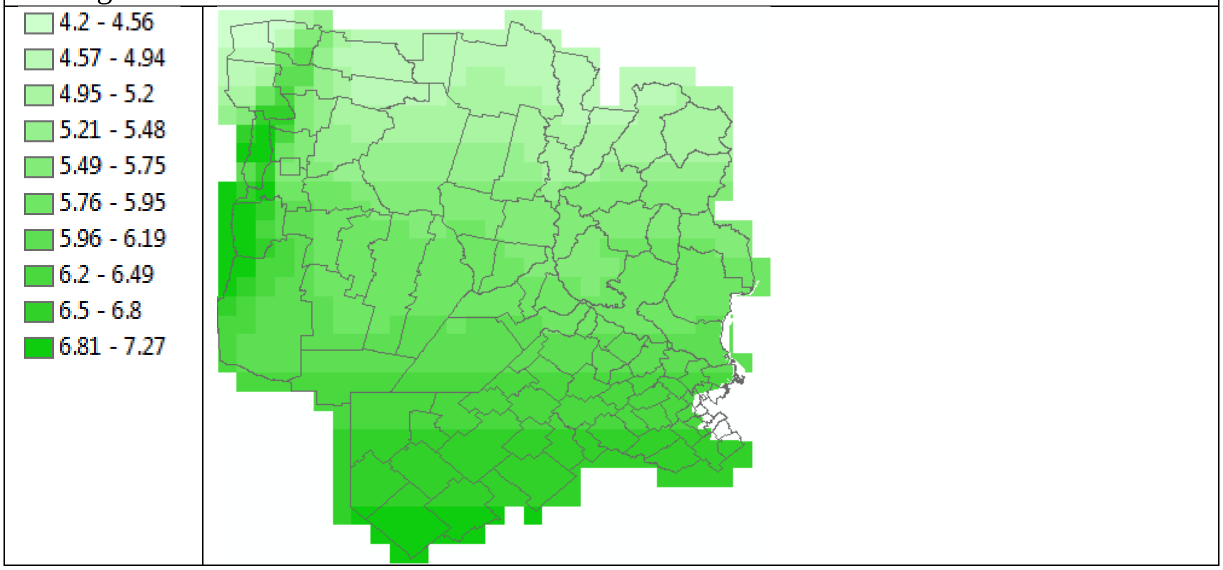
The simulated grain yield from WOFOST is expressed as kg dry weight (zero moisture) per hectare. To enable a correct comparison with downscaled official regional yield statistics, expressed in fresh weight, the modelled grain yield was converted as follows:

- modelled yield/(1-0.13)/1000 assuming 13 moisture in the dried grains

4.6. Results potential run

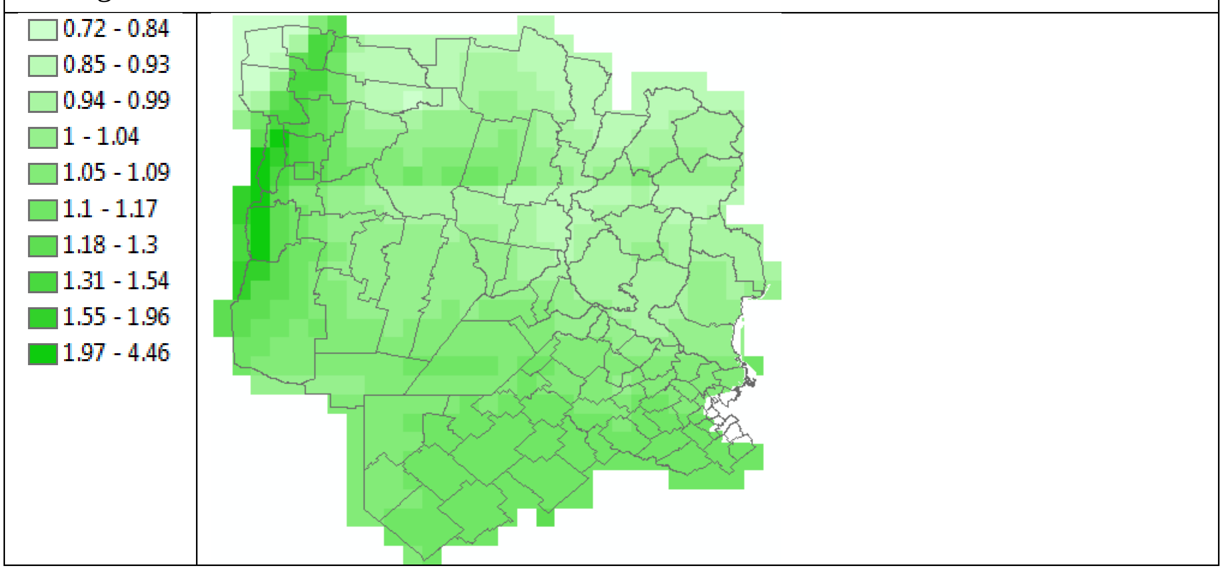
Average maximum Leaf Area Index is ok with a south-north gradient. The northern located grid cells have a lower maximum LAI due to higher temperatures and thus a little bit shorter vegetative phase. Relatively high values occur in the Sierras de Córdoba (western Cordoba province) due to the opposite reason. To improve we would need a more refined zonation and calibration. The CV (not shown) is somewhat high within a range of 20-30% indicating variation over years. Some northern located grid cells have a CV between 30 and 50%. It indicates the need for further calibration probably with respect to phenology (see also section 4.2.1)

Average maximum Leaf Area Index



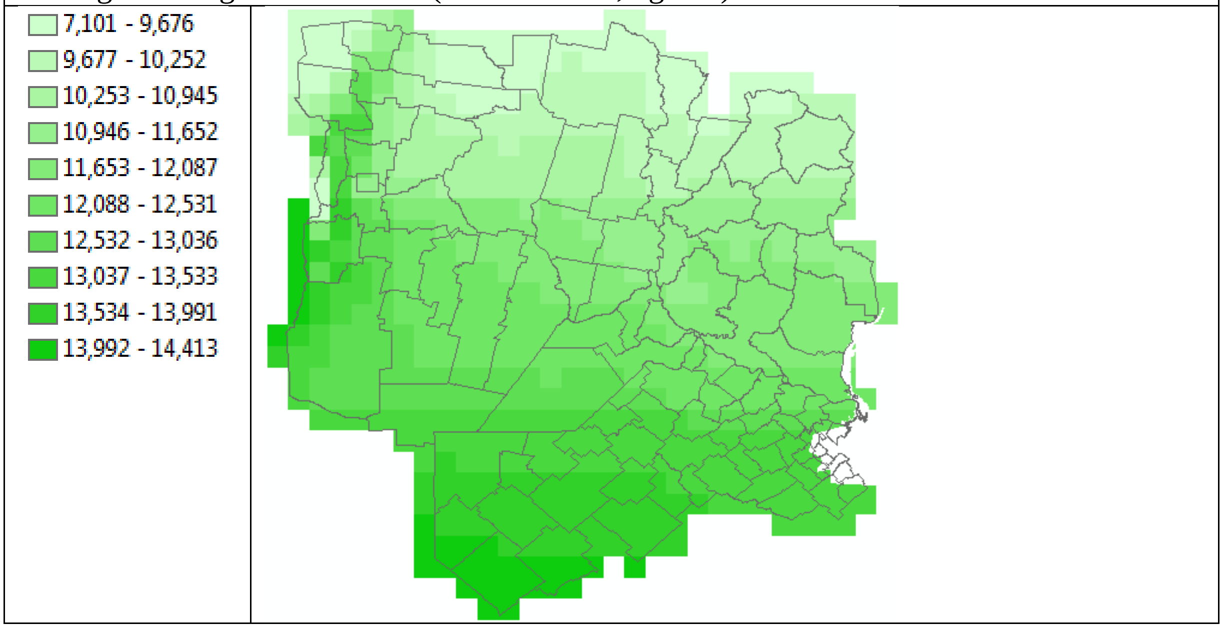
Average final Leaf Area Index is ok except some high values in the Sierras de Córdoba due to colder environments leading to elongated crop cycles. The CV (not shown) is within 20% thus relatively stable except some grid cells located in the Sierras de Córdoba and far south.

Average final Leaf Area Index



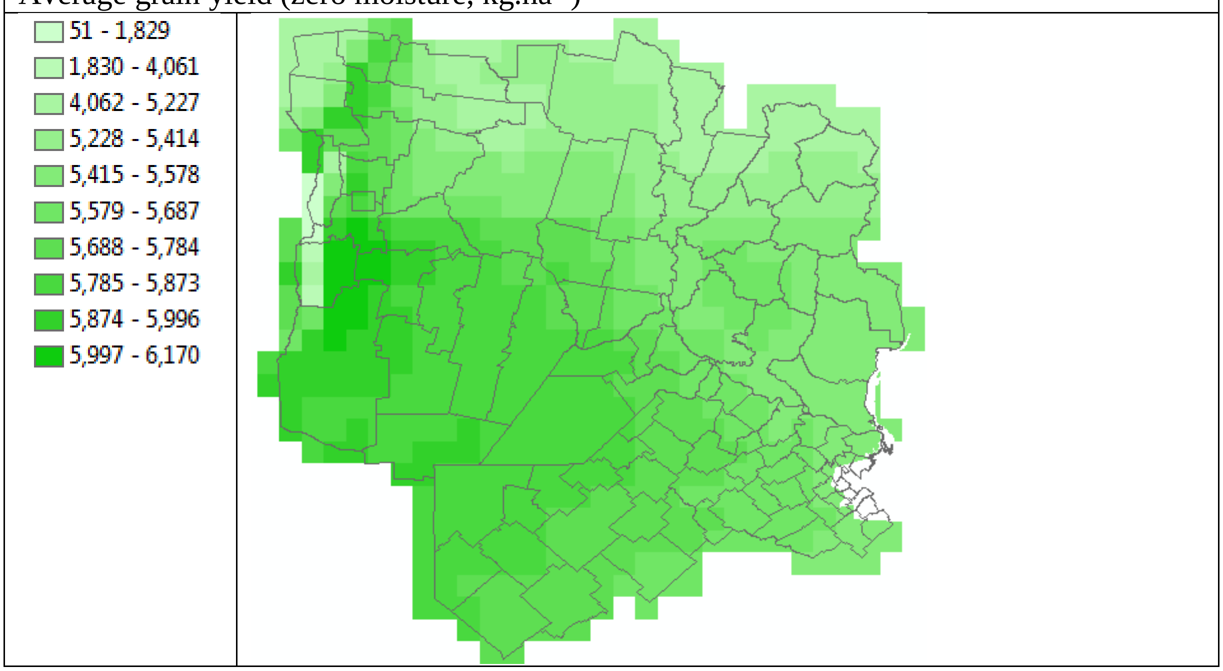
Average above ground biomass is within the indicated ranges with a south-north gradient similar like for maximum LAI. The CV (not shown) is within 20% thus stable.

Average above ground biomass (zero moisture, kg.ha^{-1})



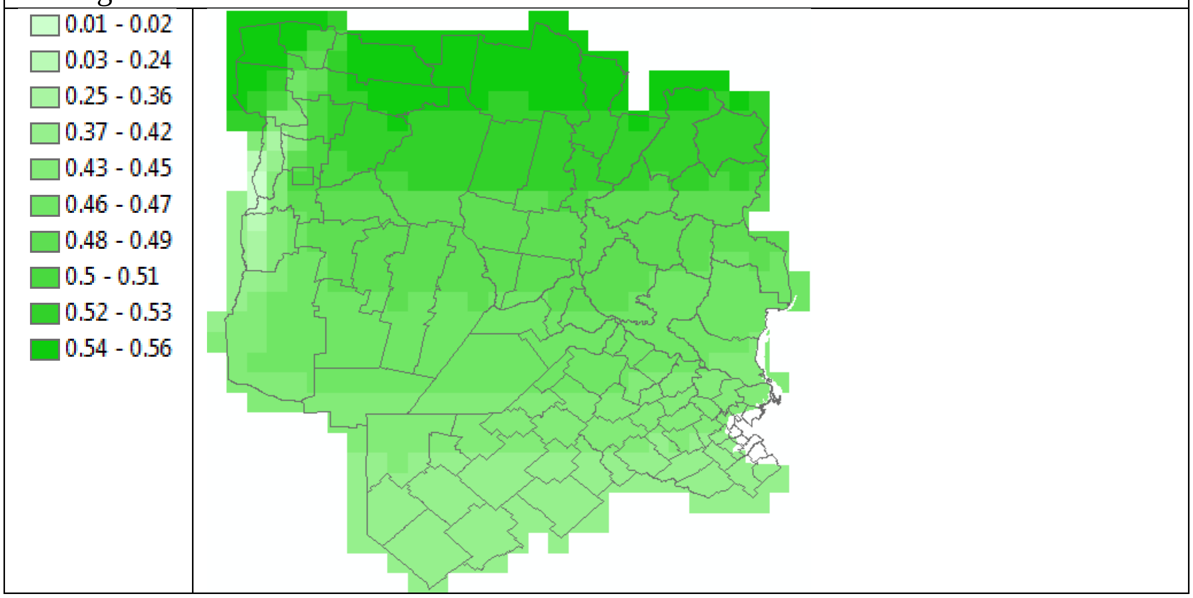
Average grain yield is within than the indicated ranges except for 5 grid cells in the Sierras de Córdoba with values between 0 and 4000 kg.ha^{-1}. However similar levels were obtained in the GYGA project. Apparently the indicated range applies more to the north. The CV (not shown) is within 15% thus stable.

Average grain yield (zero moisture, kg.ha^{-1})



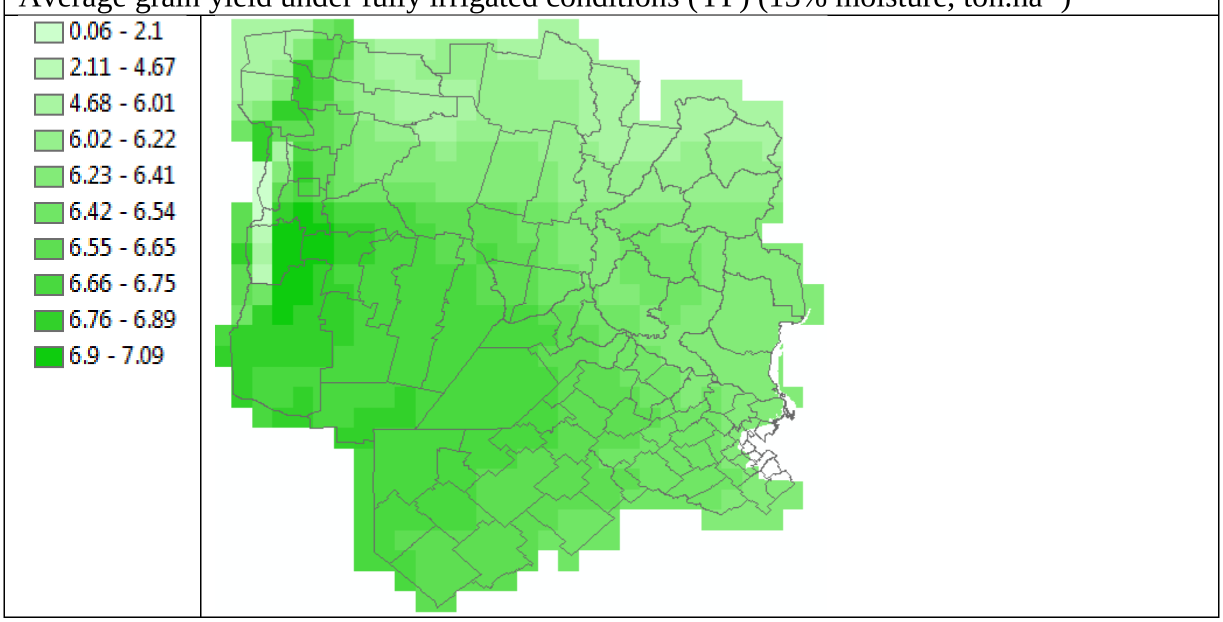
Average harvest index is within the indicated ranges except for some grid cells in the Sierras de Córdoba and a larger zone in the south with values between 0.3 and 0.4. The northern zone has a relatively high HI because of the low LAI and biomass and the relatively average grain yield. The relative short duration of the vegetative phase, due to higher temperature is followed by an average length of the grain filling phase. In the latter phase the accelerated effect of higher temperatures is offset by the slowing effect of a reduced shortening of day length compared to the southern zones. The CV (not shown) is within 20% thus relative stable except for some grid cells in the Sierras de Córdoba that have very high values.

Average harvest index



Average grain yield (13 %) between 4.1– 7.1 $ton.ha^{-1}$ (excluding some grid cells in the Sierras de Córdoba).

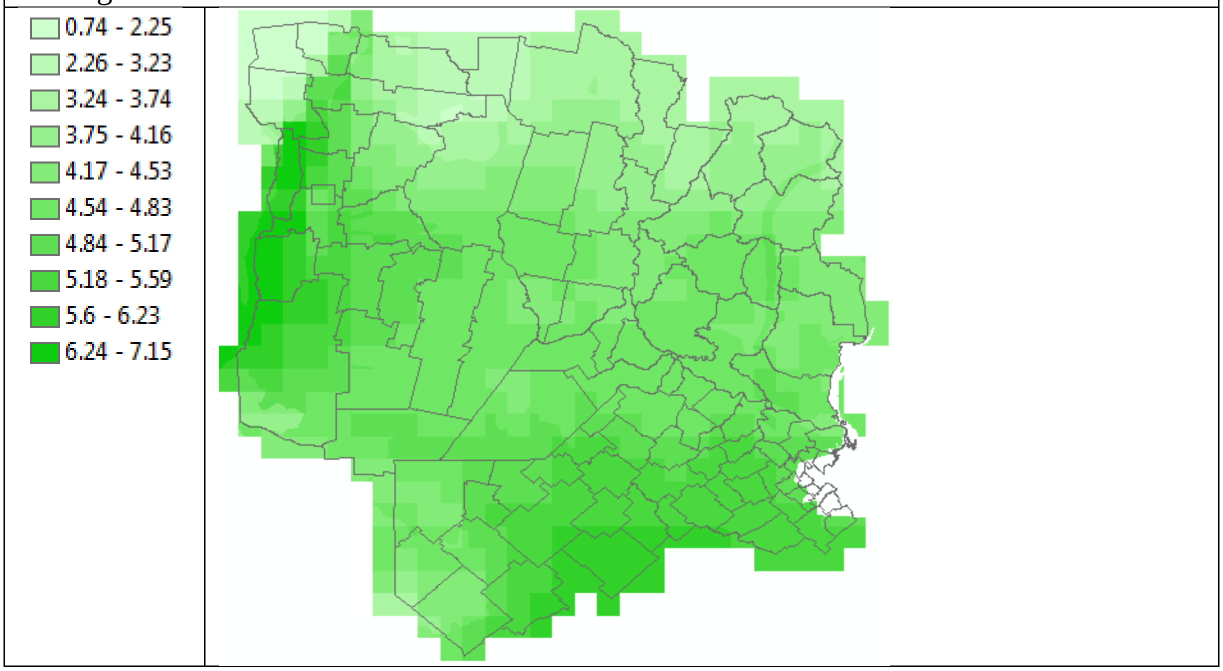
Average grain yield under fully irrigated conditions (YP) (13% moisture, $ton.ha^{-1}$)



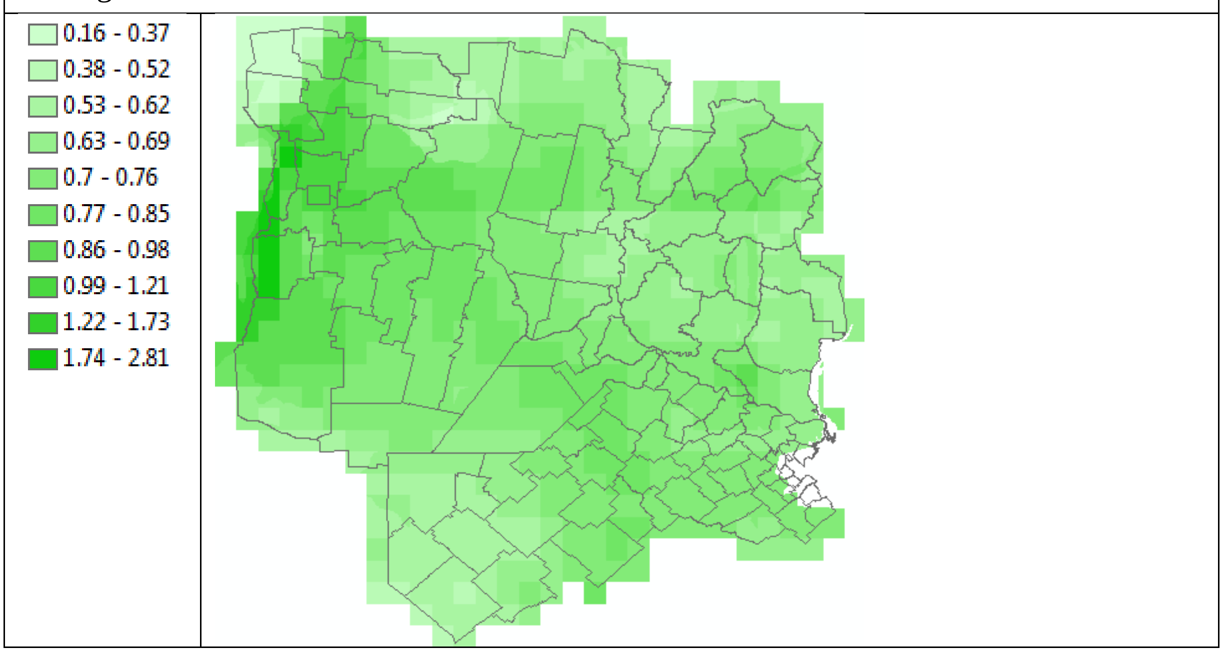
4.7. Results water-limited run

The drought effect is largest in the north-west and south-west where yield levels have been decreased by more than 30-40%. This is also clear in the figure showing the difference between the potential and the water-limited yield.

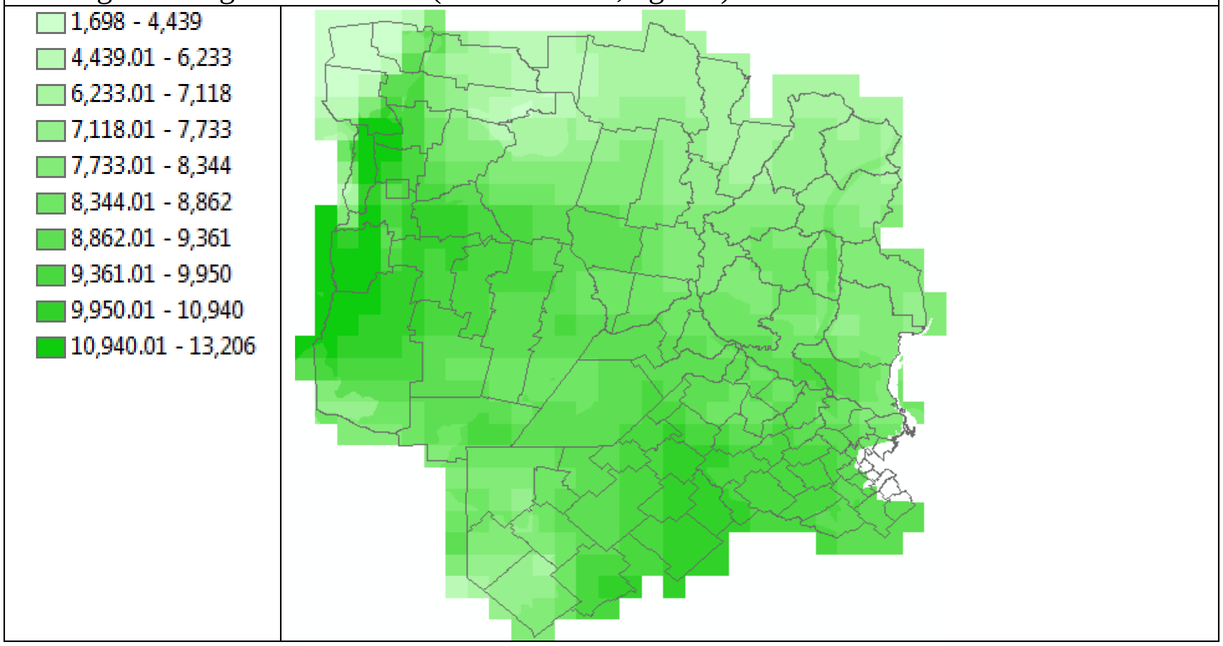
Average maximum Leaf Area Index



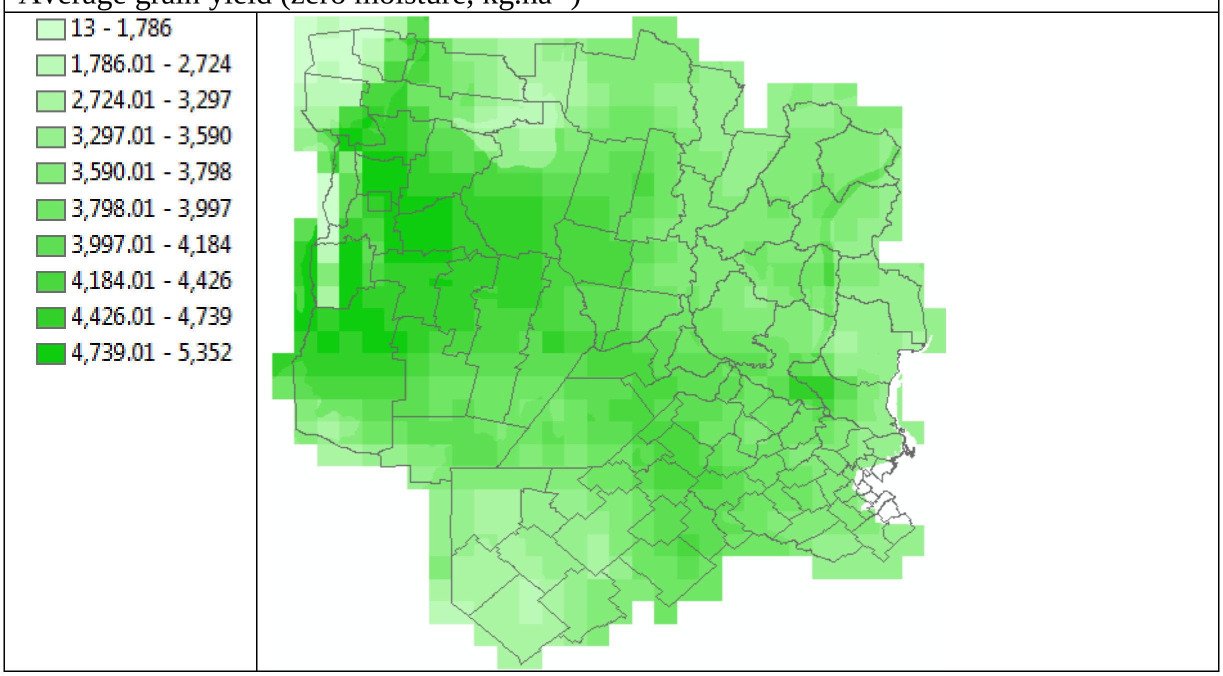
Average final Leaf Area Index



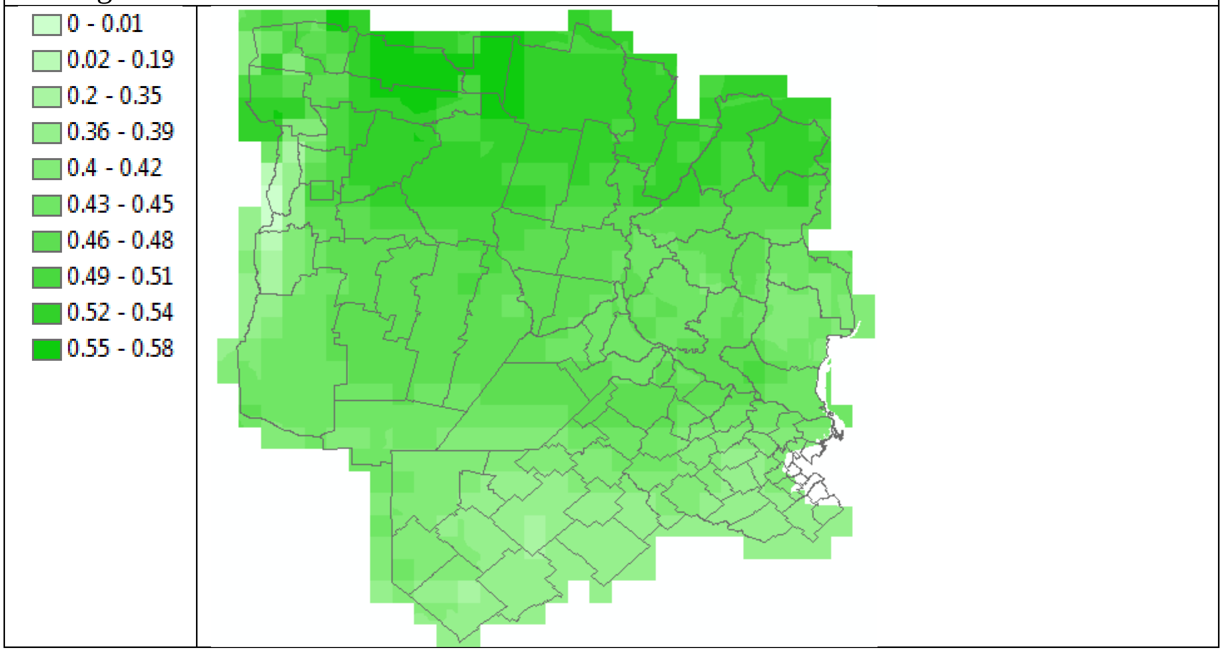
Average above ground biomass (zero moisture, kg.ha^{-1})



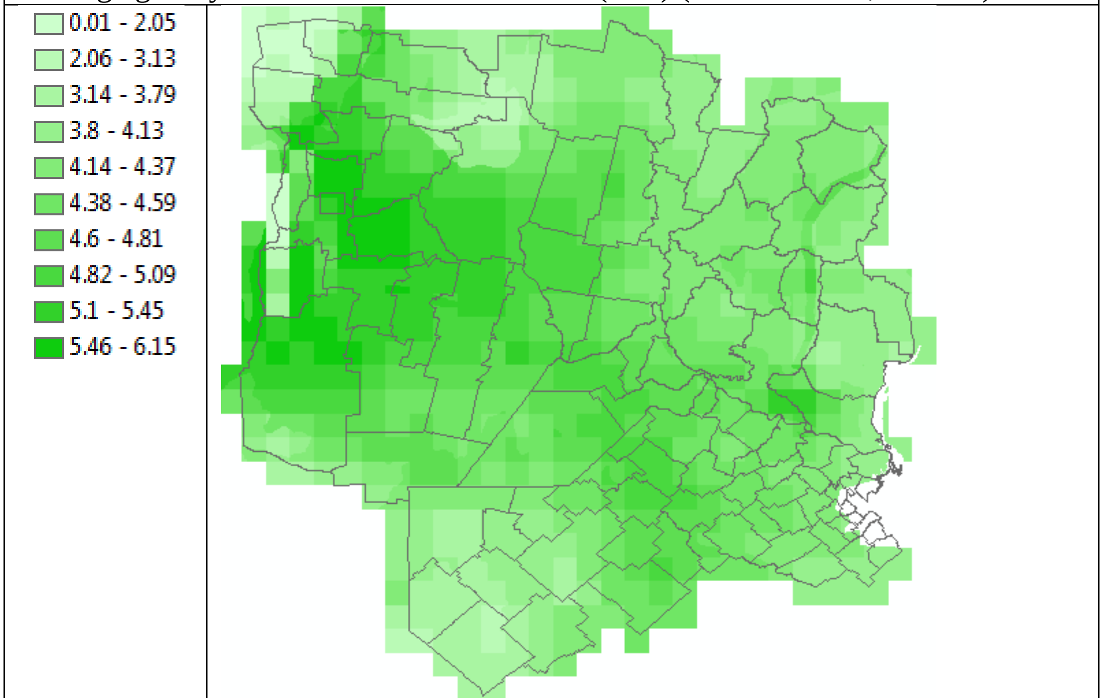
Average grain yield (zero moisture, kg.ha^{-1})



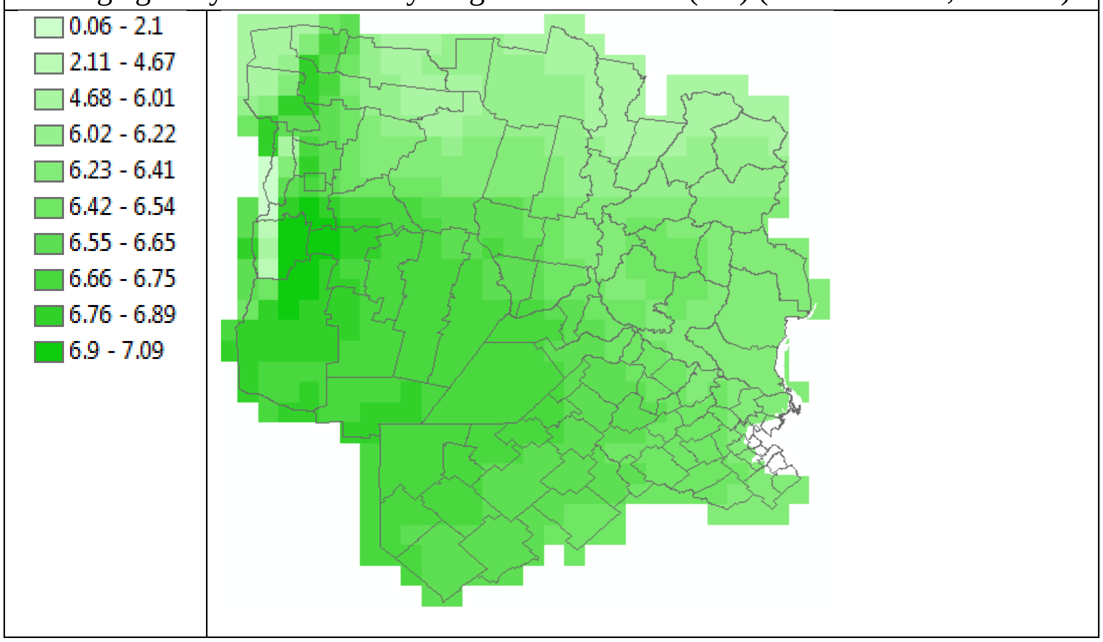
Average harvest index



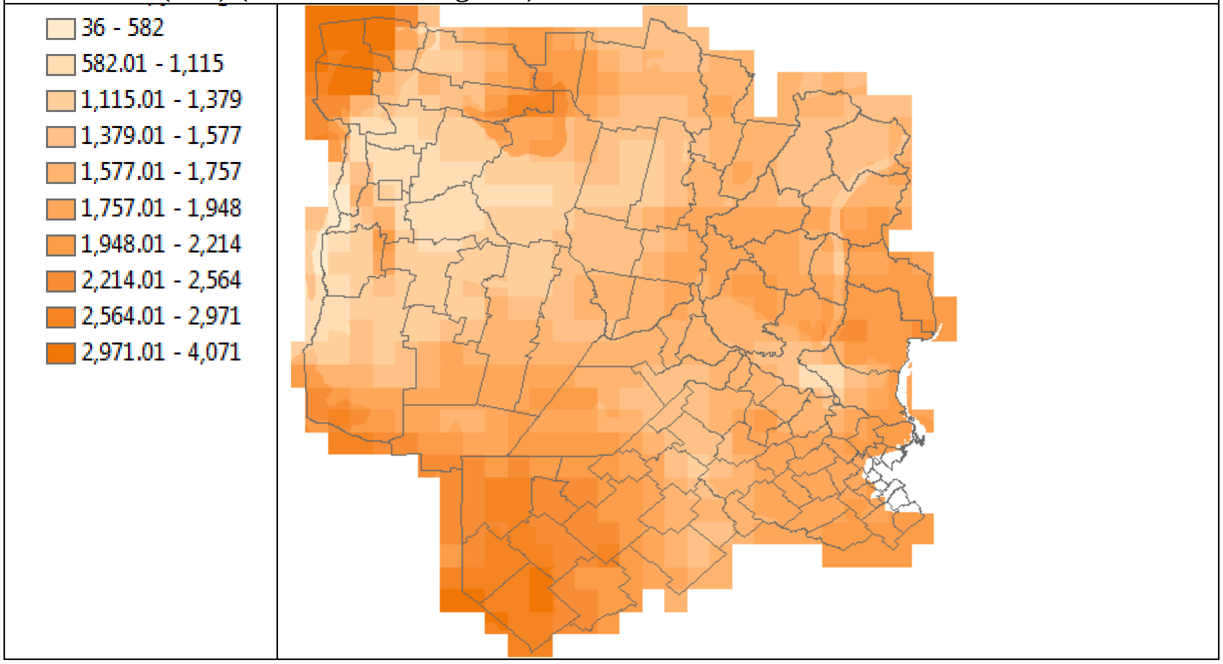
Average grain yield under rain fed condition (YW) (13% moisture, $ton.ha^{-1}$)



Average grain yield under fully irrigated conditions (YP) (13% moisture, $ton.ha^{-1}$)



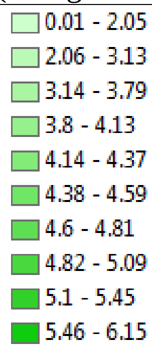
Difference between average grain yield under fully irrigated condition (YP) and rain fed conditions (WP) (zero moisture, $kg.ha^{-1}$)

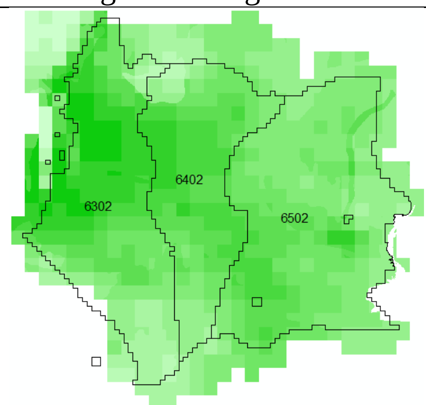


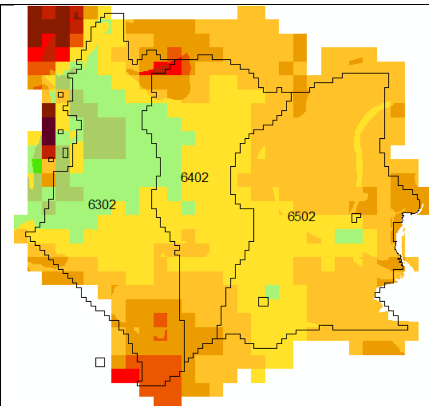
4.8. Comparison with GYGA

Water limited yield levels of SIGMA are substantially higher for the western part (GYGA-ED zone 6302). In other areas yield levels are comparable. The spatial patterns are similar: the highest yields in GYGA-ED zone 6302 and 6402 and lower values to the east and substantially lower to the south and the upper north-west.

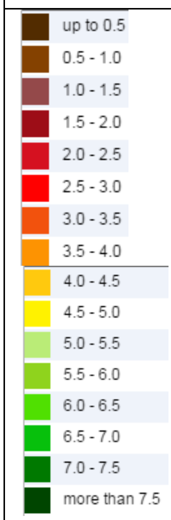
Average grain yield under rain fed conditions (YW) SIGMA (13% moisture, ton.ha^{-1}) (left: green SIGMA legend and right: GYGA legend)

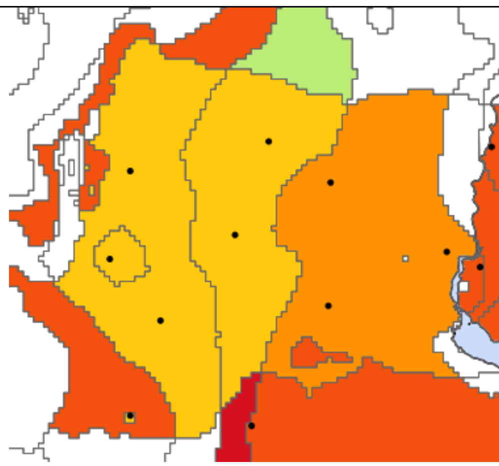


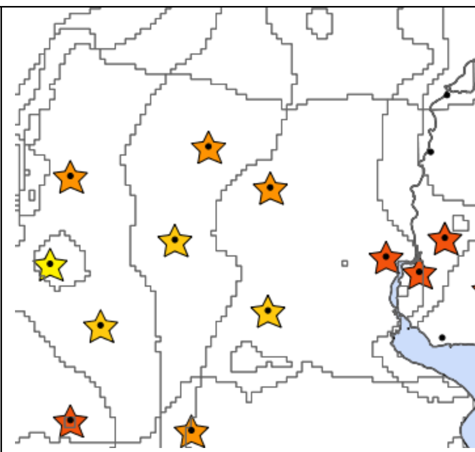




Average grain yield under rain fed conditions (YW) GYGA (13% moisture, ton.ha^{-1})





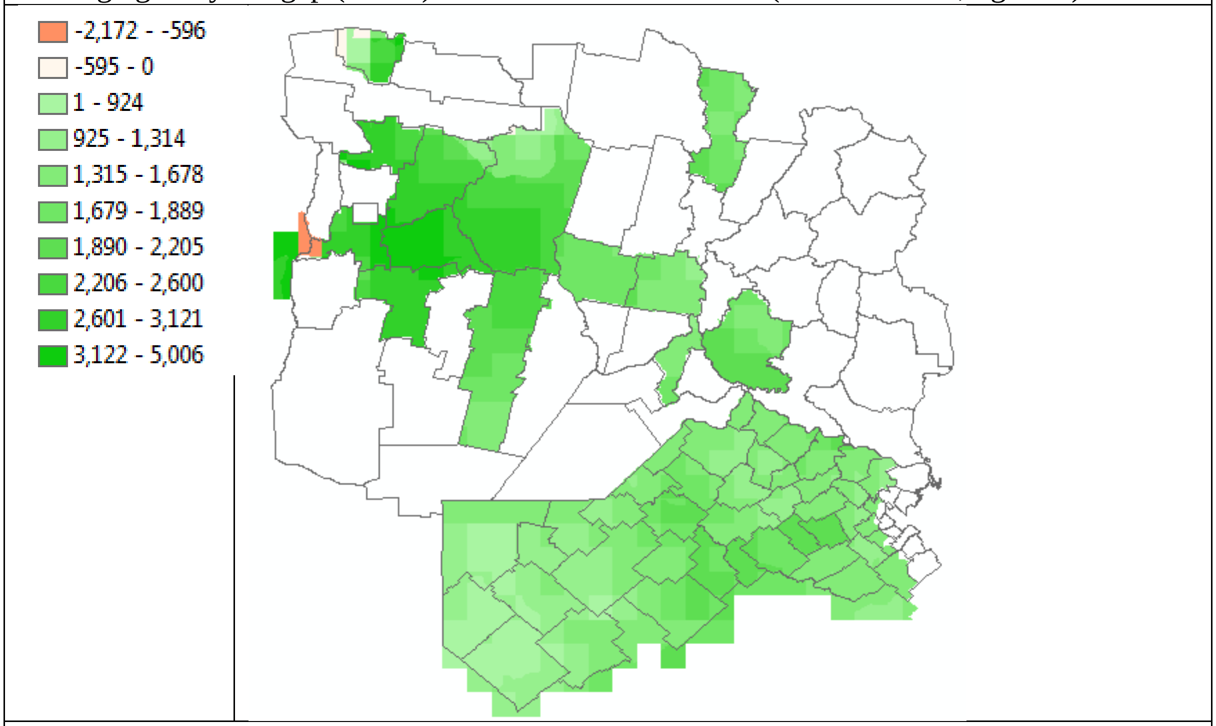


GYA-RWS	Lon	Lat	YW-GYGA	YW-SIGMA	YP-GYGA	YP-SIGMA
Pilar	-63.833	-31.667	3.86	5.68	5.90	6.74
Río Cuarto	-64.167	-33.117	4.94	5.47	7.26	6.89
Laboulaye	-63.333	-34.133	4.20	4.61	7.12	6.78
Rafaela	-61.550	-31.180	3.83	4.79	5.92	6.31
Marcos Juárez	-62.106	-32.719	4.13	4.69	6.22	6.62
Paraná	-60.530	-31.856	3.97	4.22	6.02	6.36
Pergamino	-60.569	-33.888	4.11	4.55	6.59	6.56
Gualeguaychú	-58.617	-33.000	3.16	4.01	6.34	6.33

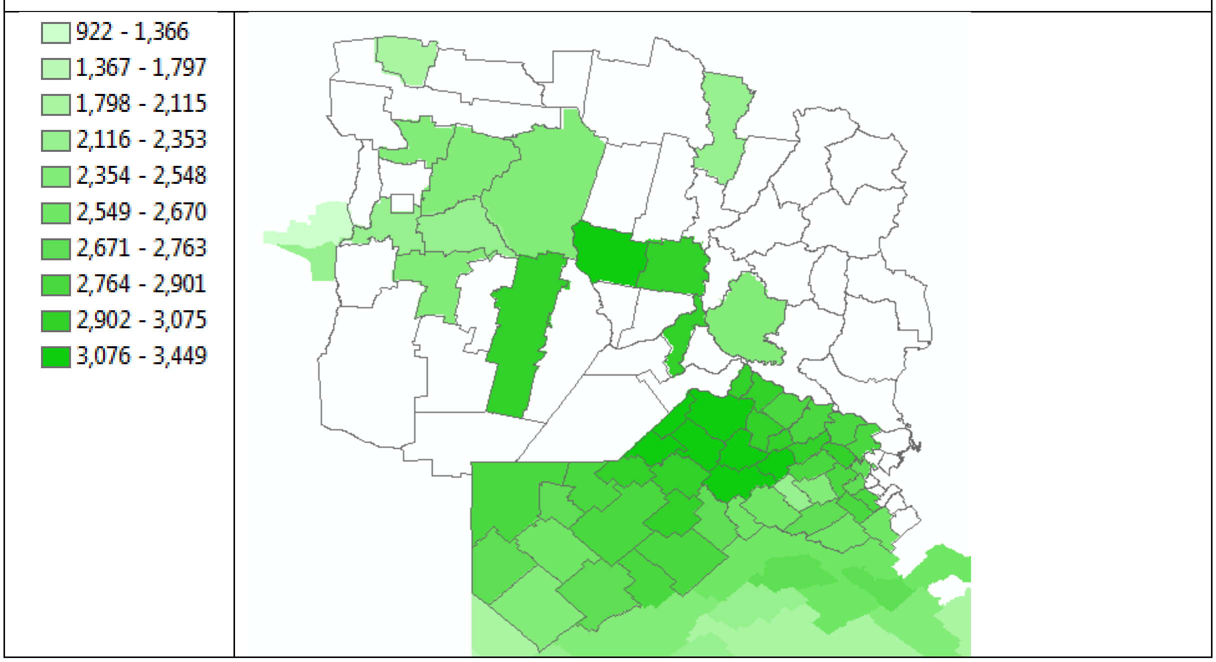
4.9. Yield gap under water-limited conditions

Actual yield levels vary between 0.9 – 3.5 ton.ha^{-1}. The resulting yield gap varies between almost 5 ton.ha^{-1} in the west to less than 1 ton.ha^{-1} in the south. A few locations in the Sierras de Córdoba (outer west) have a negative yield gap due to the very low simulated yields. For these locations a finer zonation and associated calibration is recommended or it could be decided to drop those locations in case soy beans is not an important crop. In fact an accurate arable land map or a soy bean map would help to focus the simulation effort.

Average grain yield gap (YGW) under rain fed conditions (13% moisture, kg.ha-1)



Average grain yield (YA) – actual (13% moisture, kg.ha-1) based on 2000-2013



5. References

www.wofost.wur.nl

Batjes NH 2015. World soil property estimates for broad-scale modelling (WISE30sec, ver. 1.0). Report 2015/01, ISRIC—World Soil Information, Wageningen (available via: http://www.isric.org/data/datadownload)

Berrisford, P. et al., 2009. The ERA-Interim archive. ERA Report Series. ECMWF, Shinfield Park, Reading, 16 pp.

Boons-Prins, E.R., Koning, G.H.J. de, Diepen, C.A. van, Penning de Vries, F.W.T., 1993. Crop-specific parameters for yield forecasting across the European Communtiy. Simulation Reports CABO-TT, no 32. Wageningen, The Netherlands, pp 160.

ECMWF, 2007. ERA-Interim: New ECMWF reanalysis products from 1989 onwards. ECMWF Newsletter 110. http://www.ecmwf.int/publications/newsletters/pdf/110_rev.pdf

Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Laura Harrison, Hoell, A., and Michaelson, J., 2015. The climate hazards infrared precipitation with stations - a new environmental record for monitoring extremes. Scientic Data, 2.

Fehr, W.R. and Caviness, C.E., 1977. Stages of soybean development. Spec. Rep. No. 80, Coop. Ext. Serv. Agric. and Home Econ. Expn. Stn., Iowa State Univ., Ames, IA.

Hartmann, H., 2011. Development of ERAI downscaling equations for the JRC global project. MC-MRD-CR-2011-002. Meteo Consult, Wageningen.

Kakiuchi, J. and Kobata, T., 2006. The relationship between dry matter increase of seed and shoot during the seed-filling period in three kinds of soybeans with different growth habits subjected to shading and thinning. Plant production science, 9(1), pp.20-26.

Bajgain, R., Kawasaki, Y., Akamatsu, Y., Tanaka, Y., Kawamura, H., Katsura, K. and Shiraiwa, T., 2015. Biomass production and yield of soybean grown under converted paddy fields with excess water during the early growth stage. Field Crops Research, 180, pp.221-227.

Setiyono, T.D., Weiss, A., Specht, J., Bastidas, A.M., Cassman, K.G. and Dobermann, A., 2007. Understanding and modeling the effect of temperature and daylength on soybean phenology under high-yield conditions. Field crops research, 100(2), pp.257-271.

Tote, C., Patricio, D. ,Boogaard, H., Van der Wijngaart, R., Tarnavsky, E., Funk, C., 2015. Evaluation of Satellite Rainfall Estimates for Drought and Flood Monitoring in Mozambique. Remote Sens. 7-2, 1758-1776.

Van Heemst, H.D.J., 1988. Plant data values required for simple crop growth simulation models: review and bibliography. Simulation Report CABO-TT 17, Centre for Agrobiological Research (CABO) and Department of Theoretical Production Ecology, Agricultural University Wageningen, The Netherlands.

Wart, J., Lenny G.J. van Bussel, Joost Wolf, Rachel Licker, Patricio Grassini, Andrew Nelson, Hendrik Boogaard, James Gerber, Nathaniel D. Mueller, Lieven Claessens, Martin K. van Ittersum, Kenneth G. Cassman, Use of agro-climatic zones to upscale simulated crop yield potential, Field Crops Research, Volume 143, 1 March 2013, Pages 44-55, ISSN 0378-4290.

Appendix I Overview of all crop parameters and changes

```
-- SQL*Loader control file
LOAD DATA
INFILE *
APPEND
INTO TABLE CROP_PARAMETER_VALUE
FIELDS TERMINATED BY ',' ENCLOSED BY '"'
(
CROP_NO
, PARAMETER_CODE
, PARAMETER_XVALUE
, PARAMETER_YVALUE
)
BEGINDATA
"1","AMAXTB_01","0","37"
"1","AMAXTB_02","1.7","37"
"1","AMAXTB_03","2","5"
"1","AMAXTB_04","0","0"
"1","AMAXTB_05","0","0"
"1","AMAXTB_06","0","0"
"1","AMAXTB_07","0","0"
"1","AMAXTB_08","0","0"
"1","AMAXTB_09","0","0"
"1","AMAXTB_10","0","0"
"1","CFET","1.0",""
"1","CVL","0.720",""
"1","CVO","0.480",""
"1","CVR","0.720",""
"1","CVS","0.690",""
"1","DEPNR","5",""
"1","Pcrt","14.89",""
"1","Popt","7.92",""
"1","DVSEND","2",""
"1","EFF","0.40",""
"1","FLTB_01","0","0.6"
"1","FLTB_02","1.0","0.6"
"1","FLTB_03","1.05","0.3"
"1","FLTB_04","1.2","0.2"
"1","FLTB_05","1.4","0.1"
"1","FLTB_06","1.7","0"
"1","FLTB_07","2","0"
"1","FLTB_08","0","0"
"1","FLTB_09","0","0"
"1","FLTB_10","0","0"
"1","FOTB_01","0","0"
"1","FOTB_02","1.0","0"
"1","FOTB_03","1.05","0.4"
```

"1","FOTB_04","1.2","0.7"
"1","FOTB_05","1.4","0.85"
"1","FOTB_06","1.7","1"
"1","FOTB_07","2","1"
"1","FOTB_08","0","0"
"1","FOTB_09","0","0"
"1","FOTB_10","0","0"
"1","FRTB_01","0","0.65"
"1","FRTB_02","0.75","0.35"
"1","FRTB_03","1.0","0.15"
"1","FRTB_04","1.5","0"
"1","FRTB_05","2.0","0"
"1","FRTB_06","0","0"
"1","FRTB_07","0","0"
"1","FRTB_08","0","0"
"1","FRTB_09","0","0"
"1","FRTB_10","0","0"
"1","FSTB_01","0","0.4"
"1","FSTB_02","1.0","0.4"
"1","FSTB_03","1.05","0.3"
"1","FSTB_04","1.2","0.1"
"1","FSTB_05","1.4","0.05"
"1","FSTB_06","1.7","0"
"1","FSTB_07","2","0"
"1","FSTB_08","0","0"
"1","FSTB_09","0","0"
"1","FSTB_10","0","0"
"1","IAIRDU","0",""
"1","KDIF","0.8",""
"1","LAIEM","0.0163",""
"1","PERDL","0.03",""
"1","Q10","2",""
"1","RDI","10",""
"1","RDMCR","120",""
"1","RDRRTB_01","0","0"
"1","RDRRTB_02","1.5","0"
"1","RDRRTB_03","1.5001","0.02"
"1","RDRRTB_04","2","0.02"
"1","RDRRTB_05","0","0"
"1","RDRRTB_06","0","0"
"1","RDRRTB_07","0","0"
"1","RDRRTB_08","0","0"
"1","RDRRTB_09","0","0"
"1","RDRRTB_10","0","0"
"1","RDRSTB_01","0","0"
"1","RDRSTB_02","1.5","0"
"1","RDRSTB_03","1.5001","0.02"
"1","RDRSTB_04","2","0.02"
"1","RDRSTB_05","0","0"

"1","RDRSTB_06","0","0"
"1","RDRSTB_07","0","0"
"1","RDRSTB_08","0","0"
"1","RDRSTB_09","0","0"
"1","RDRSTB_10","0","0"
"1","RFSETB_01","0","1"
"1","RFSETB_02","2","1"
"1","RFSETB_03","0","0"
"1","RFSETB_04","0","0"
"1","RFSETB_05","0","0"
"1","RFSETB_06","0","0"
"1","RFSETB_07","0","0"
"1","RFSETB_08","0","0"
"1","RFSETB_09","0","0"
"1","RFSETB_10","0","0"
"1","RGRLAI","0.01",""
"1","RML","0.03",""
"1","RMO","0.017",""
"1","RMR","0.01",""
"1","RMS","0.015",""
"1","RRI","1.2",""
"1","SLATB_01","0","0.0014"
"1","SLATB_02","0.45","0.0025"
"1","SLATB_03","0.9","0.0025"
"1","SLATB_04","2","0.001"
"1","SLATB_05","0","0"
"1","SLATB_06","0","0"
"1","SLATB_07","0","0"
"1","SLATB_08","0","0"
"1","SLATB_09","0","0"
"1","SLATB_10","0","0"
"1","SPA","0",""
"1","SPAN","30",""
"1","SSA","0",""
"1","TBASE","7",""
"1","TBASEM","7",""
"1","TDWI","120",""
"1","TEFFMX","22",""
"1","TMNFTB_01","0","0"
"1","TMNFTB_02","3","1"
"1","TMNFTB_03","0","0"
"1","TMNFTB_04","0","0"
"1","TMNFTB_05","0","0"
"1","TMNFTB_06","0","0"
"1","TMNFTB_07","0","0"
"1","TMNFTB_08","0","0"
"1","TMNFTB_09","0","0"
"1","TMNFTB_10","0","0"
"1","TMPFTB_01","0","0"

"1","TMPFTB_02","10","0"
"1","TMPFTB_03","24","1.0"
"1","TMPFTB_04","39","1.0"
"1","TMPFTB_05","42","0.63"
"1","TMPFTB_06","0","0"
"1","TMPFTB_07","0","0"
"1","TMPFTB_08","0","0"
"1","TMPFTB_09","0","0"
"1","TMPFTB_10","0","0"
"1","TSUMEM","90",""
"1","IOX","0",""
"1","DVRMAX1","0.0328",""
"1","DVRMAX2","0.0729",""
"1","MG","5",""
"1","Tmin","7",""
"1","Topt","31",""
"1","Tmax","40",""

Appendix II Validation statistics for gridded weather data sets

A number of validation statistics are given included RMAE, Bias, POD and FAR. These statistics are explained below (Tote et al., 2015).

Name	Formula	Perfect Score
Pearson correlation coefficient	$r = \frac{\sum(G-\bar{G})(S-\bar{S})}{\sqrt{\sum(G-\bar{G})^2}\sqrt{\sum(S-\bar{S})^2}}$	1
Mean Error	$ME = 1/N\sum(S-G)$	0
Relative mean absolute error	$RMAE = \left(1/N\sum\lvert(S-G)\rvert\right)/(\bar{G})$	0
Nash-Sutcliffe Efficiency coefficient	$Eff = 1-\left(\sum(S-G)^2\right)/\left(\sum(G-\bar{G})^2\right)$	1
Bias	$Bias = \sum S/\sum G$	1

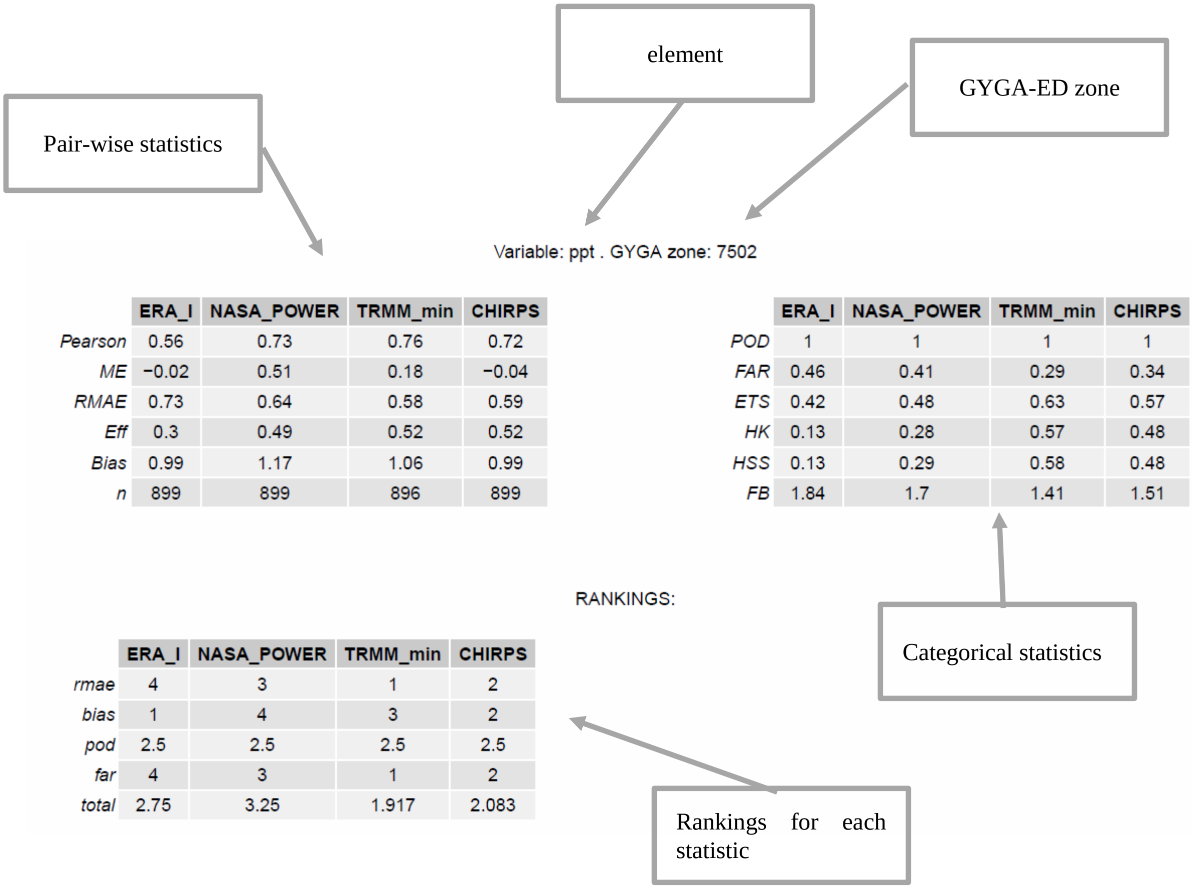
Name	Formula	Perfect Score
Probability of detection	$POD = A/(A+C)$	1
False alarm ratio	$FAR = B/(A+B)$	0
Equitable threat score	$ETS = (A-Ar)/(A+B+C-Ar)$ with hits that occur by chance: $Ar = \left((A+C)(A+B)\right)/N$	1
Hansen and Kuipers discriminate	$HK = A/(A+C) - B/(B+D)$	1
Heidke Skill Score	$HSS = \left(2(AD-BC)\right)/\left((A+C)(C+D)+(A+B)(B+D)\right)$	1
Frequency Bias	$FB = (A+B)/(A+C)$	1

The following gridded data sets are included

- JRC_ERA-INTERIM (named ERA_I)
- NASA_POWER (named NASA_POWER)
- NASA_TRIMM (named TRMM_min)
- CHG_CHIRPS (named CHIRPS)

The elements are:

- minimum temperature (named tmax)
- maximum temperature (named tmin)
- radiation (named rad)
- precipitation (named ppt)



Variable: ppt . GYGA zone: 7502

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.56	0.73	0.76	0.72
ME	−0.02	0.51	0.18	−0.04
RMAE	0.73	0.64	0.58	0.59
Eff	0.3	0.49	0.52	0.52
Bias	0.99	1.17	1.06	0.99
n	899	899	896	899

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	1	1	1	1
FAR	0.46	0.41	0.29	0.34
ETS	0.42	0.48	0.63	0.57
HK	0.13	0.28	0.57	0.48
HSS	0.13	0.29	0.58	0.48
FB	1.84	1.7	1.41	1.51

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	4	3	1	2
bias	1	4	3	2
pod	2.5	2.5	2.5	2.5
far	4	3	1	2
total	2.75	3.25	1.917	2.083

Variable: tmax . GYGA zone: 6302

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.97	0.91	NA	NA
ME	−1.6	0.47	NA	NA
RMAE	0.07	0.09	NA	NA
Eff	0.85	0.76	NA	NA
Bias	0.94	1.02	NA	NA
n	1570	1570	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	2	1	NA	NA
total	1.5	1.5	NA	NA

Variable: tmax . GYGA zone: 6402

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.97	0.93	NA	NA
ME	−1.25	1.63	NA	NA
RMAE	0.06	0.09	NA	NA
Eff	0.89	0.69	NA	NA
Bias	0.95	1.07	NA	NA
n	1547	1547	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	1	2	NA	NA
total	1	2	NA	NA

Variable: tmax . GYGA zone: 6502

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.98	0.94	NA	NA
ME	−1.05	1.92	NA	NA
RMAE	0.05	0.1	NA	NA
Eff	0.93	0.68	NA	NA
Bias	0.96	1.08	NA	NA
n	2225	2225	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	1	2	NA	NA
total	1	2	NA	NA

Variable: tmin . GYGA zone: 6302

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.93	0.89	NA	NA
ME	2.55	1.96	NA	NA
RMAE	0.27	0.28	NA	NA
Eff	0.68	0.68	NA	NA
Bias	1.25	1.19	NA	NA
n	1508	1508	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	2	1	NA	NA
total	1.5	1.5	NA	NA

Variable: tmin . GYGA zone: 6402

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.97	0.96	NA	NA
ME	2.35	2.3	NA	NA
RMAE	0.22	0.22	NA	NA
Eff	0.76	0.75	NA	NA
Bias	1.22	1.21	NA	NA
n	1580	1580	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	2	1	NA	NA
total	1.5	1.5	NA	NA

Variable: tmin . GYGA zone: 6502

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.97	0.95	NA	NA
ME	1.36	1.58	NA	NA
RMAE	0.13	0.16	NA	NA
Eff	0.87	0.8	NA	NA
Bias	1.12	1.14	NA	NA
n	2324	2324	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	1	2	NA	NA
bias	1	2	NA	NA
total	1	2	NA	NA

Variable: rad . GYGA zone: 6302

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.76	0.77	NA	NA
ME	1.75	1.57	NA	NA
RMAE	0.19	0.18	NA	NA
Eff	0.51	0.53	NA	NA
Bias	1.11	1.1	NA	NA
n	1238	1216	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	2	1	NA	NA
bias	2	1	NA	NA
total	2	1	NA	NA

Variable: rad . GYGA zone: 6402

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.97	0.98	NA	NA
ME	2.08	1.54	NA	NA
RMAE	0.13	0.1	NA	NA
Eff	0.81	0.89	NA	NA
Bias	1.13	1.09	NA	NA
n	680	670	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	2	1	NA	NA
bias	2	1	NA	NA
total	2	1	NA	NA

Variable: rad . GYGA zone: 6502

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.94	0.95	NA	NA
ME	2.12	1.4	NA	NA
RMAE	0.15	0.11	NA	NA
Eff	0.75	0.84	NA	NA
Bias	1.13	1.09	NA	NA
n	1751	1719	NA	NA

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	NA	NA	NA	NA
FAR	NA	NA	NA	NA
ETS	NA	NA	NA	NA
HK	NA	NA	NA	NA
HSS	NA	NA	NA	NA
FB	NA	NA	NA	NA

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	2	1	NA	NA
bias	2	1	NA	NA
total	2	1	NA	NA

Variable: ppt . GYGA zone: 6302

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.65	0.64	0.72	0.71
ME	10.4	1.66	2.91	2.26
RMAE	0.89	0.72	0.63	0.65
Eff	0.02	0.32	0.39	0.47
Bias	1.5	1.08	1.14	1.11
n	2009	2009	2004	2009

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	0.97	0.85	0.88	0.89
FAR	0.28	0.24	0.18	0.27
ETS	0.37	0.36	0.49	0.33
HK	0.52	0.52	0.65	0.49
HSS	0.54	0.53	0.66	0.5
FB	1.34	1.12	1.08	1.23

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	4	3	1	2
bias	4	1	3	2
pod	1	4	3	2
far	4	2	1	3
total	3.5	2.333	2	2.167

Variable: ppt . GYGA zone: 6402

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.59	0.69	0.78	0.76
ME	6.74	1.96	3.19	1.58
RMAE	0.78	0.65	0.58	0.57
Eff	0.15	0.41	0.51	0.57
Bias	1.27	1.08	1.13	1.06
n	1619	1619	1615	1619

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	0.96	0.88	0.9	0.92
FAR	0.25	0.19	0.15	0.24
ETS	0.35	0.42	0.51	0.34
HK	0.49	0.58	0.67	0.49
HSS	0.52	0.59	0.68	0.51
FB	1.29	1.08	1.05	1.21

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	4	3	2	1
bias	4	2	3	1
pod	1	4	3	2
far	4	2	1	3
total	3.5	2.667	2.333	1.5

Variable: ppt . GYGA zone: 6502

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
Pearson	0.64	0.72	0.77	0.77
ME	5.36	4.03	4.53	3.32
RMAE	0.69	0.65	0.59	0.55
Eff	0.23	0.38	0.4	0.57
Bias	1.19	1.15	1.17	1.12
n	2380	2380	2374	2380

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
POD	0.97	0.87	0.9	0.92
FAR	0.2	0.21	0.13	0.24
ETS	0.39	0.3	0.48	0.27
HK	0.51	0.44	0.64	0.39
HSS	0.56	0.46	0.65	0.43
FB	1.22	1.1	1.03	1.21

RANKINGS:

	ERA_I	NASA_POWER	TRMM_min	CHIRPS
rmae	4	3	2	1
bias	4	2	3	1
pod	1	4	3	2
far	2	3	1	4
total	3.167	2.833	2.333	1.667

Best Ranked databases in each GYGA zone