Crop Monitoring as an
E-agricultural tool in
Developing Countries

EVALUATION REPORT ON THE INTEGRATION OF RS DATA IN BIOMA MODELS

Reference: *E-AGRI_D35.3_ Evaluation Report on the integration of RS data in BioMA models*

Author(s): Simone Bregaglio, Valentina Pagani, Fabien Ramos, Roberto Confalonieri

Version: 1.0
Date: 2/01/2014

DOCUMENT CONTROL

Signatures

Author(s) : Simone Bregaglio
Valentina Pagani,
Fabien Ramos,
Roberto Confalonieri

Reviewer(s) :

Approver(s) :

Issuing authority :

Change record

Release	Date	Pages	Description	Editor(s)/Reviewer(s)
1.0	02/01/2014	48	D35.3 Report	

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

This report presents the results of the evaluation of the BioMA model WOFOST for wheat yield forecasting in Morocco when remote sensing data are integrated in the simulation environment. Remote sensing data used refer to the maximum Normalized Difference Vegetation Index (NDVI), in turn used to update the simulated state variable leaf area index (LAI). WOFOST was parameterized for soft and durum wheat within the activities reported in the E-AGRI tasks D34.X.

The simulated outputs in the period 1998-2011 were aggregated at regional and national level and then processed via the CGMS Statistical Tool Box to be compared with official statistics. Four empirical relationships to derive wheat LAI values from NDVI data were tested, and simulated values were compared to those obtained without the use of exogenous data. Forecasting was carried out at day of the year 60, 90, 120 and 150 in order to assess the impact of the forcing events during the crop cycle.

In general, the assimilation of NDVI data into the BioMA-WOFOST model determined an increase in the accuracy of wheat forecasting in Morocco, especially when it was carried out before maturity. However, differences in the performances obtained across regions and with different empirical models to derive LAI from NDVI suggest further studies before transferring these techniques into operational contexts.

1. Introduction

The spatially distributed application of crop simulation models is affected by a large uncertainty in the agro-meteorological, pedological and management inputs used to reproduce the spatial heterogeneity of the agricultural systems (Aggarwal, 1995). These information are usually stored into agrometeorological databases after the interpolation of scattered point measurements (e.g., weather data from meteorological stations, soil properties from field level analyses) into grids of different spatial resolution, representing the model simulation unit (Confalonieri et al., 2013). The integration of remotely sensed (RS) information into crop simulators, even if recent, is becoming a standard methodology to reduce the uncertainty of the model inputs (Launay and Guèrif, 2005), thanks to the availability of satellite data at high temporal resolution (e.g., NOAA AVHRR, TERRA/AQUA MODIS, METEOSAT sensors).

Three main strategies were developed to integrate RS information into crop simulators (Dorigo et al., 2007), which are calibration, forcing and updating. The calibration method consists in the adjustment of the model parameters after the minimizations of the errors between the simulation outputs and the observed canopy state variables (e.g., Bouman et al., 1995); the forcing technique is used to directly replace simulated state variables using RS data at each model time step: since many crop simulators work at a daily time step, different interpolation methods were developed to dispose of continuous series of RS indices (Roerink et al., 2000); the updating method, also known as sequential data assimilation, is used to update the model outputs only when the RS observation is available, via the application of algorithms to convert observations into simulated variables (Mc Laughlin, 2002).

Many canopy state variables can be retrieved by remote sensing and are currently used to improve crop model predictions. Leaf area index (LAI, m^2 m^{-2}), fractional cover (fCOVER) and the photosynthetically active radiation absorbed by the canopy (fAPAR) are the most used in the available studies (e.g., Gobron et al., 2000; Mo et al., 2005), other than evapotranspiration (e.g., Batiaansen and Ali, 2003) and crop phenological information (e.g., Xin et al., 2002). LAI, defined as the one sided green leaf area per unit ground area, is a crucial variable simulated by the crop models, since it determines the amount of intercepted radiation for the photosynthesic activity and influences the actual crop evapotranspiration (Liu et al., 1997). The normalized difference vegetation index (NDVI), derived by the values of reflectance in red and near infrared wavebands, is the most commonly used RS spectral index to derive LAI values (e.g., Wang et al., 2007; Chattaraj et al., 2013). Since the quality of NDVI data is deeply decreased by background noise such as soil water content, chemical use, and tillage conditions (Koller and Upadhyaya, 2001), many empirical relationships with LAI were developed, according to the specific agroenvironmental conditions explored (e.g., Cihlar et al., 2002; Duchemin et al., 2006; Song et al., 2008; Chaurasia Sasmita et al., 2011; Chattaraj et al., 2013).

In this report, we evaluate the accuracy of the BioMA-WOFOST model in reproducing official statistics of soft and durum wheat in Morocco using NDVI data to force the model. Simulation outputs were processed via the CGMS Statistical Tool Box and they were compared at national and regional level versus official yields.

2. Materials and methods

2.1. NDVI-LAI methods

Four empirical relationships between wheat LAI values and NDVI data (f_{NDVI_LAI}) retrieved in literature were tested.

Duchemin et al. (2006) developed a f_{NDVI_LAI} function(R^2=0.92) basing on a field dataset collected in an irrigated area of the Haouz plain (region of Marrakesh, Central Morocco) during the 2002–2003 agricultural season (Equation 1).

$$NDVI = NDVI_{\inf} - \left(NDVI_{soil} - NDVI_{\inf} \right) e^{-0.54 LAI} \quad [1]$$

where $NDVI_{\inf}$ is the NDVI of an infinitely-dense canopy and $NDVI_{soil}$ is the NDVI of soil.

Song et al. (2008) developed the f_{NDVI_LAI} function (R^2=0.62) on experimental winter wheat field trials performed in the Beijing province in China in 2001, 2003 and 2004 cropping seasons (Equation 2).

$$NDVI = 0.83\left(1 - e^{\frac{LAI}{1.155}} \right) \quad [2]$$

Chaurasia Sasmita et al. (2011) derived the f_{NDVI_LAI} function (R^2=0.75) using field data collected across five different agroclimatic regions in India in 2006-2007, with wheat patches of more than 90,000 m^2 for in situ measurements in each region (Equation 3).

$$NDVI = 0.114 e^{4.906 LAI} \quad [3]$$

Chattaraj et al. (2013) found the f_{NDVI_LAI} function using experimental field data gathered in two cropping seasons for cultivars grown under different irrigation regimes in the Indo-Gangetic plains (Equation 4).

$$NDVI = 0.2128 \ln(LAI) + 0.6 \quad [4]$$

The four alternative f_{NDVI_LAI} functions were then inverted to derive LAI as a function of NDVI. They are plotted in Figure 1.

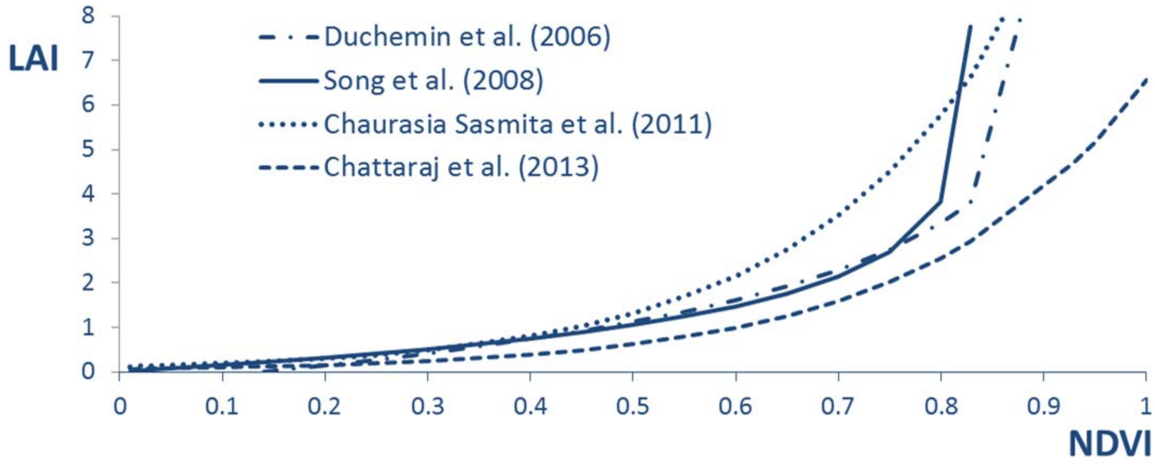

Figure 1: Leaf area index (LAI, m^2 m^{-2}) values computed as a function of NDVI with the four alternative methods tested.

The four alternative methods present an overall similarity at low NDVI values, whereas NDVI values higher than 0.4 determine large differences in derived LAI. Figure 2 presents the average value and the standard deviation of LAI values computed by the four approaches:

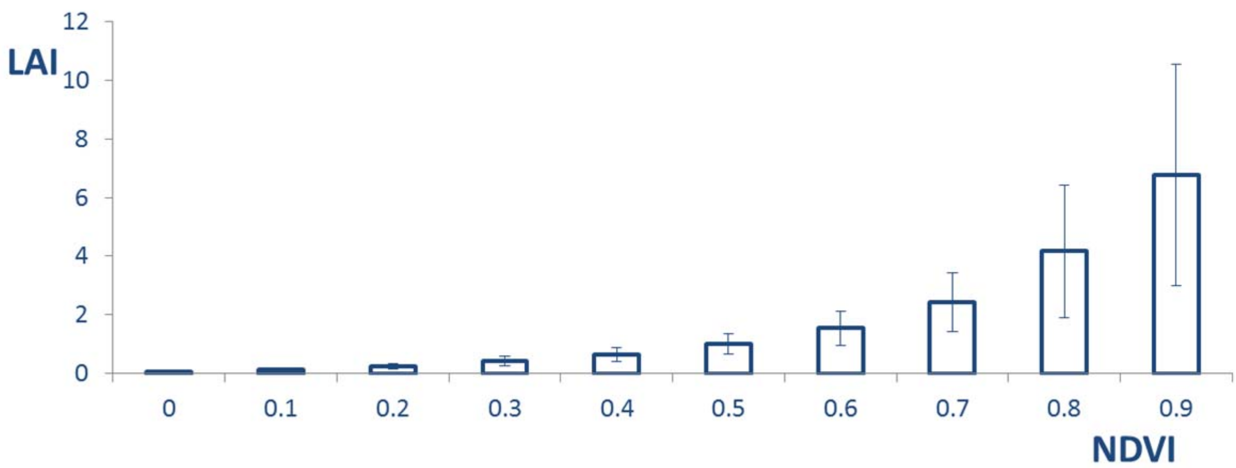

Figure 2: Average and standard deviation of leaf area index values computed by the four approaches tested to force the BioMA model WOFOST.

Standard deviation of LAI values increase in accordance with NDVI, thus indicating large differences in the LAI values computed by the four approaches even at low-medium NDVI values (0.4 ÷ 0.7).

2.2. The UNIMI.Forcing component

The four approaches to derive wheat LAI values from NDVI were implemented in the UNIMI.Forcing component, which in turn was linked to the BioMA WOFOST model calibrated to simulate durum and soft wheat in Morocco (E-AGRI report D34.3).
UNIMI.Forcing is a software component developed to force a crop model using exogenous LAI values periodically available. In case vegetation indices data are available instead of LAI values, UNIMI.Forcing has routines for the estimation of LAI from NDVI. The component is autonomous in selecting the most reliable source of information among those available (e.g., LAI, NDVI) and can be coupled with all the crop models implementing a dynamic approach for daily partitioning of assimilates into leaves, stems and storage organs, as the BioMA model WOFOST. The strategy diagram of the component is presented in Figure 3.

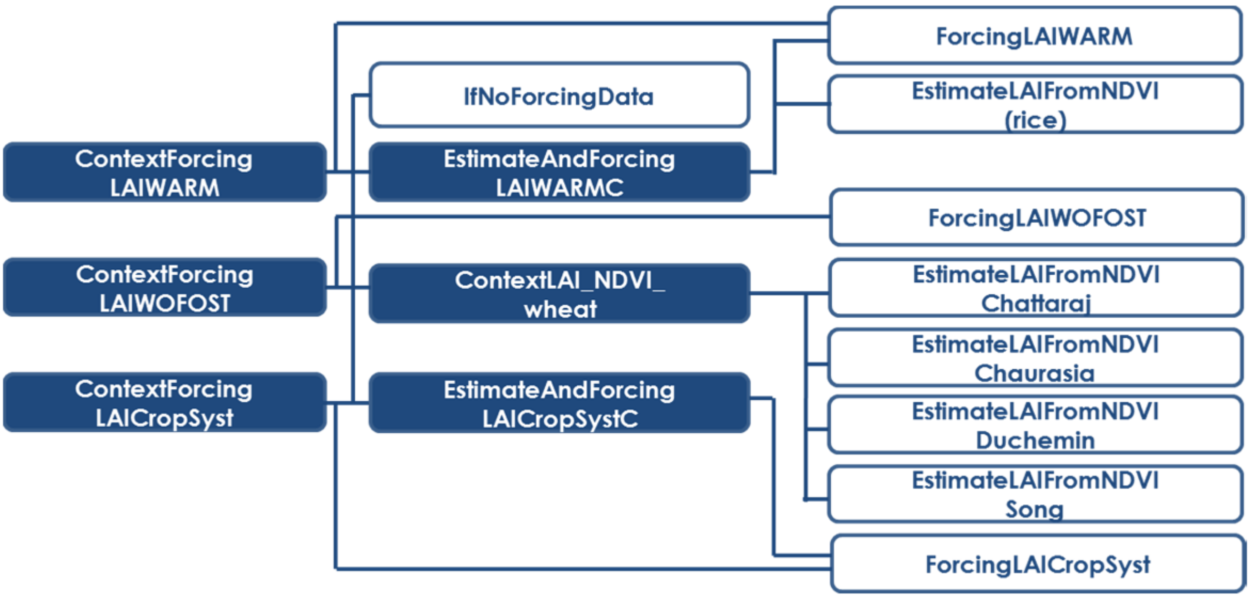

Figure 3: Strategy diagram of the UNIMI.Forcing component. The four approaches to derive LAI from NDVI for wheat are implemented as discrete units.

The help and the code documentation files of the component are available at http://agsys.cra-cin.it/tools/. The component design allows for extensions by the users without requiring the re-compilation.

2.3. NDVI data used to derive wheat LAI in Morocco

The Moroccan NDVI data were delivered by the Joint Research Centre of the European Commission. They refer to the grid cells of the MARS weather database in which wheat is cultivated in the period 1998-2011. The time resolution of the NDVI data is the decade.

The approach used to force the BioMA model WOFOST was the updating one: for each grid cell and cropping season, the maximum NDVI value ($NDVI_{max}$) and the corresponding day of year were extracted from the NDVI time series. Then, the $NDVI_{max}$ was converted into LAI via the f_{NDVI_LAI} functions tested and it was given as input to the BioMA model WOFOST. After the forcing event and within the same model time step, the biomass of the plant organs (i.e., leaves, stems and storage organs) was derived from the new LAI value and from the updated specific leaf area (SLA, m^2 kg^{-1}). Figure 4 presents a graphical representation of the methodology adopted to force the BioMA-WOFOST model with the $NDVI_{max}$ value, and the impact on the simulated LAI.

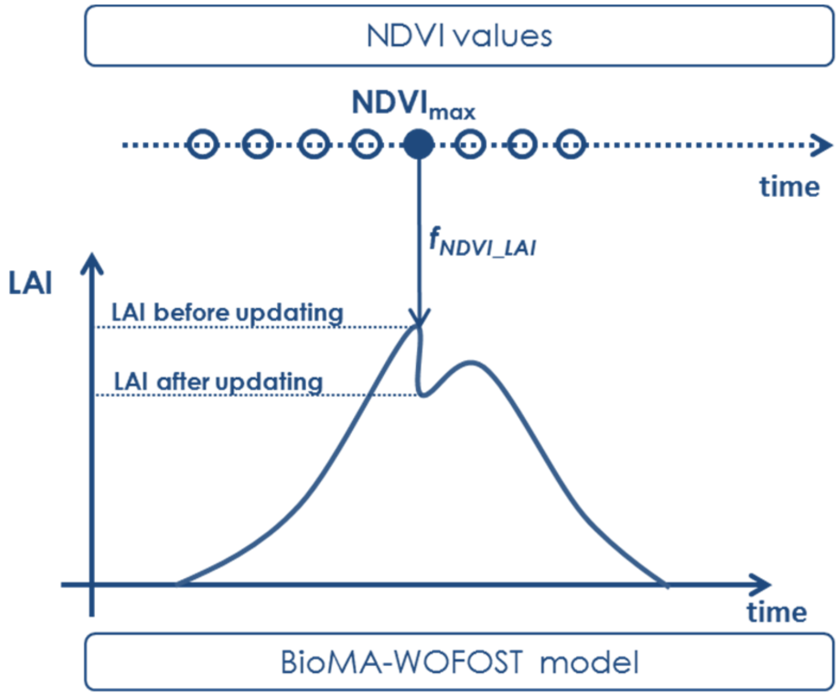

Figure 4: Schematic representation of the methodology adopted to force the BioMA-WOFOST model with LAI derived by NDVImax data with the alternative f_{NDVI_LAI} functions.

The total number of the $NDVI_{max}$ data used to force the WOFOST model was 2795. The box plots presented in Figure 5 and Figure 6 display the distributions of the $NDVI_{max}$ values and the relative days of the year in the period 1998-2011. They refer to the whole Moroccan wheat area.

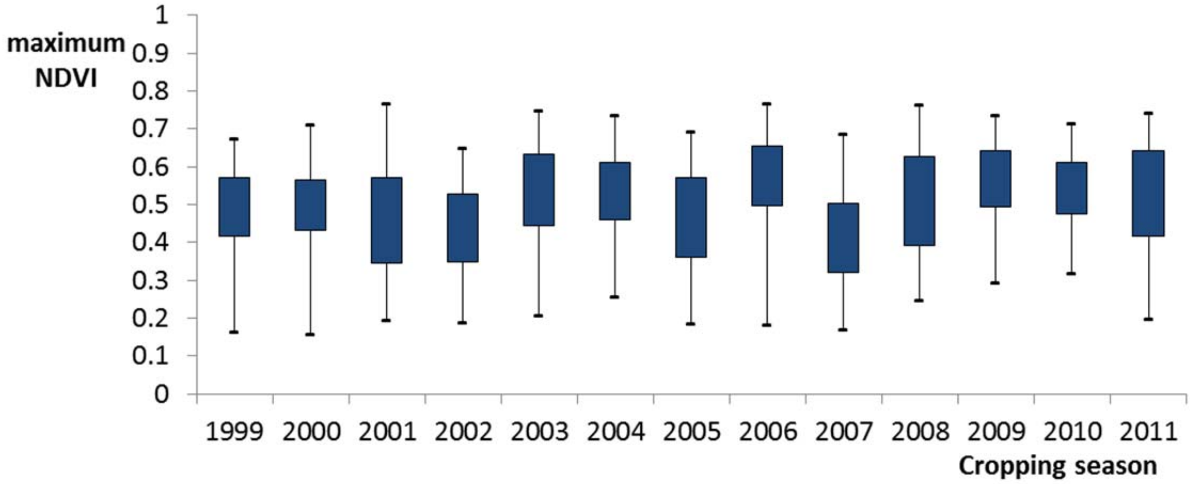

Figure 5: Boxplot showing the distributions of the maximum NDVI in the Moroccan wheat area in the period 1999-2011.

The analysis of the distribution of the $NDVI_{max}$ values in the 1999-2011 cropping seasons highlighted a variable pattern across the years, characterized by mainly right skewed boxplots, thus indicating more outliers right of median. The minimum value reached by $NDVI_{max}$ in the whole series is 0.154, whereas the maximum is 0.764.

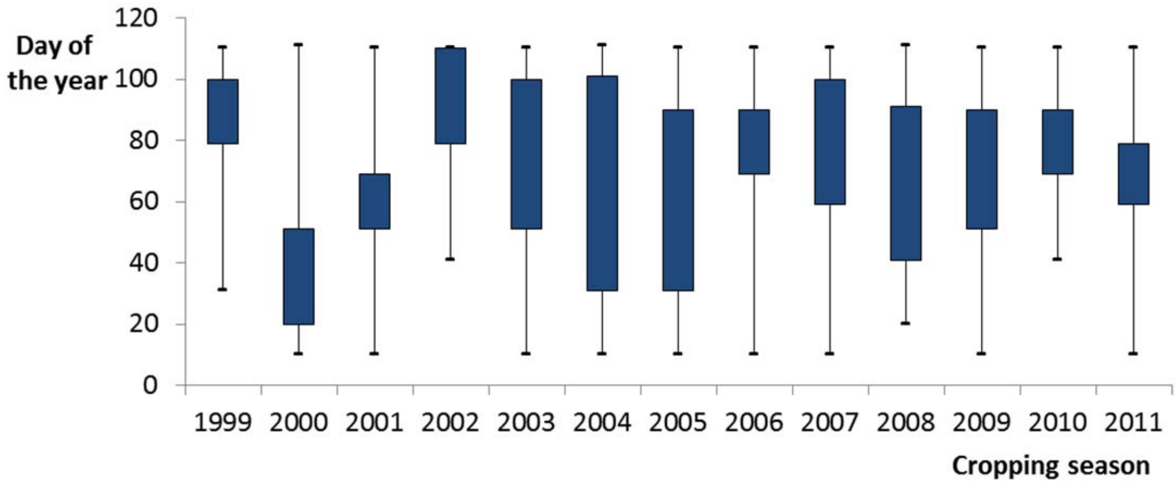

Figure 6: Boxplot showing the distributions of the day of the year in which maximum NDVI was reached in the Moroccan wheat area in the period 1999-2011.

The analysis of the distribution of the day of year in which $NDVI_{max}$ was observed in the 1999-2011 cropping seasons highlighted a decided variable pattern across the years. For example, the 2001 cropping season was characterized by very similar $NDVI_{max}$ values, with limited spread of data. On the contrary, 2004 cropping season was characterized by a large variability of the day of year in which $NDVI_{max}$ was observed.

2.4. Evaluation of the impact of forcing WOFOST model in reproducing official statistics

The simulation experiment design aimed at evaluating the impact of forcing the BioMA model WOFOST with RS exogenous data is presented in Figure 7.

The simulated outputs produced by the BioMA model WOFOST in each combination cropping season × grid cell were aggregated into decades to Moroccan administrative regions (i.e., national and regional level) using the CGMS database and then processed into the CGMS Statistical Tool Box, in order to be compared with official statistics at national and regional level. Five simulation runs were performed for each calibrated crop (durum wheat, soft wheat high potential, soft wheat low potential, see E-AGRI reports D34.4). A reference simulation was performed without updating the BioMA model WOFOST with RS data, and four simulation runs were carried out forcing the WOFOST model with $NDVI_{max}$ values (see paragraph 2.1).

The CGMS Statistical Tool Box is able to autonoumously select up to four model outputs (i.e., the indicators) to develop a multiple regression model using yearly official yields as dependent variable in different periods of the year (i.e., decades). The best regressive model (i.e., the one with the highest correlation index R^2 between simulated and official yields) obtained by each simulation run (i.e., forced and not forced) was chosen for the comparison. For each crop (i.e., durum and soft wheat) and for each Moroccan region, the forecasting was carried out at DOY 60, 90, 120 and 150, to evaluate the impact of forcing during the cropping season. The comparison of official and simulated yields was performed at National level and for the Centre, Sud, Centre Nord, Tensift, Nord Ouest, Centre Sud and Oriental regions.

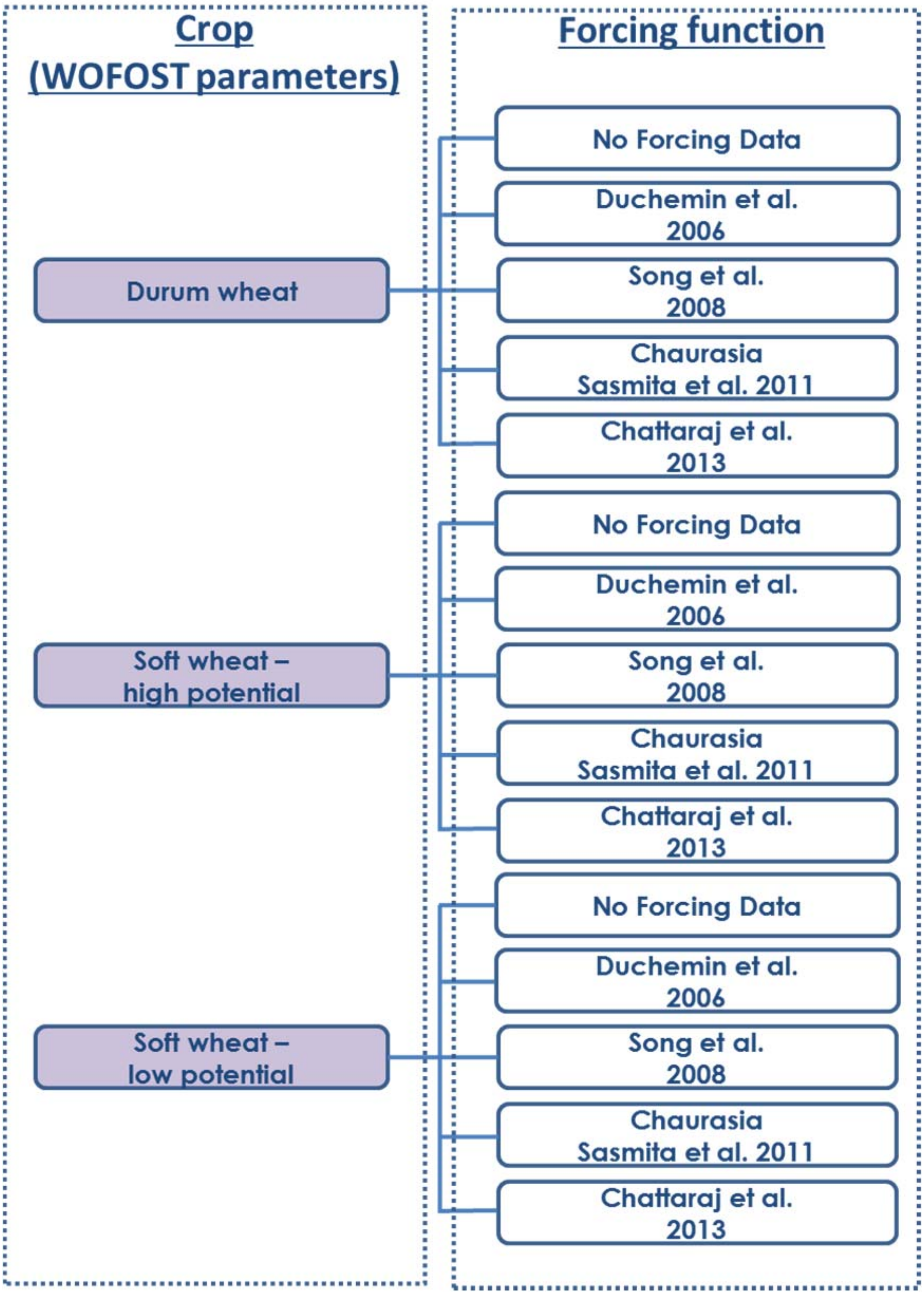

Figure 7: Simulation experiment design to evaluate the impact of forcing the BioMA-WOFOST model with LAI derived by exogenous NDVI data.

3. Results and Discussion

3.1. Analysis of the impact of forcing on time series of simulated outputs

The updating of the BioMA model WOFOST with RS data has different impacts on the simulated LAI time evolution, depending both on the f_{NDVI_LAI} used and on $NDVI_{max}$ value. Two contrasting examples are shown in Figures 8-11.

Figure 8 shows the LAI trend without forcing (red line) in the CGMS grid 25041 and year 2006. The low value of $NDVI_{max}$ (0.186) determined lower LAI values than the simulated one according to all the f_{NDVI_LAI} functions tested. It can be observed a large variability in the resulting LAI values: the f_{NDVI_LAI} developed by Chattaraj et al. (2013) determined the lowest LAI (less than 1 $m^2\ m^{-2}$) and the f_{NDVI_LAI} developed by Chaurasia Sasmita et al. (2011) derived the highest one (more than 2 $m^2\ m^{-2}$). This has a strong impact on the interception of photosynthethic active radiation, in turns reflecting in the variability of simulated trend of aboveground biomass accumulation (Figure 9).

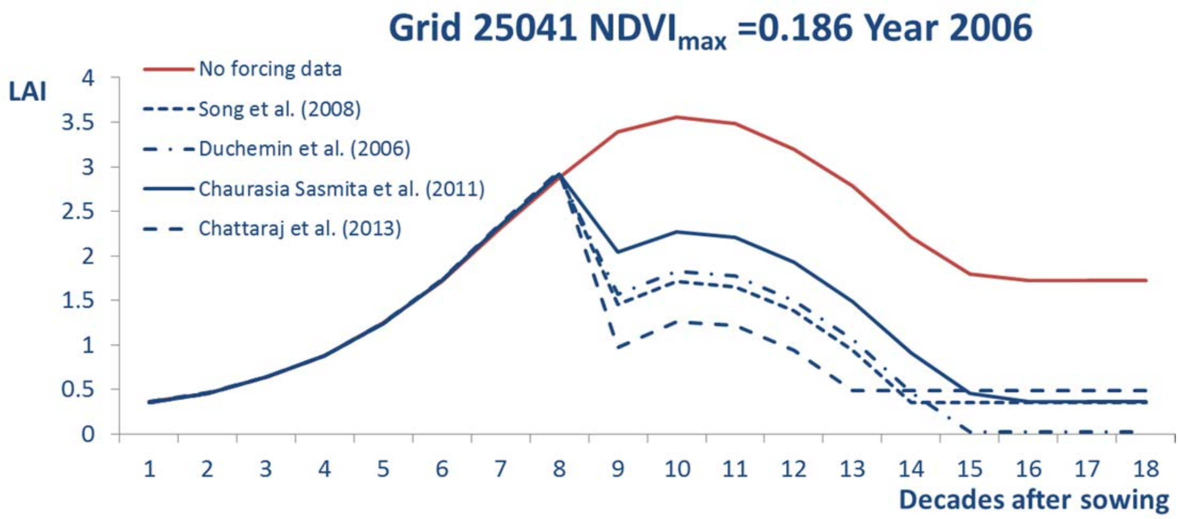

Figure 8: LAI simulations performed with the BioMA WOFOST model in the CGMS grid cell 25041 and year 2006. Comparison of the LAI evolution trends without the updating of remotely sensed data and with the assimilation of maximum NDVI using the four functions tested.

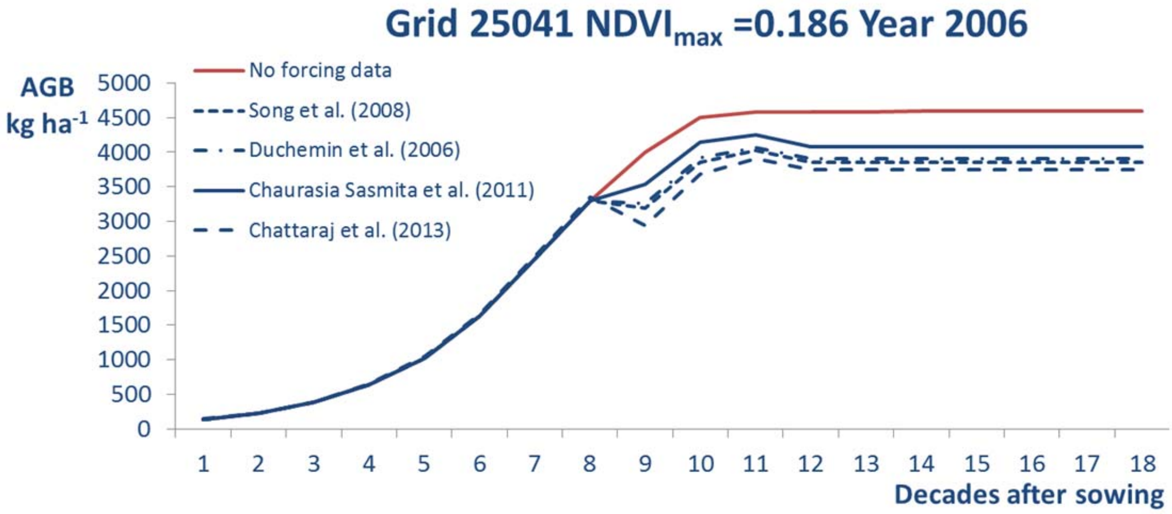

Figure 9: Aboveground biomass evolution simulated with the BioMA WOFOST model in the CGMS grid cell 25041 and year 2006. Comparison of the time trends without the updating of remotely sensed data and with the assimilation of maximum NDVI using the four functions tested.

Figure 10 shows the LAI trend without forcing (red line) in the CGMS grid 31044 and year 2005. In this case, the high value of $NDVI_{max}$ (0.746) determined higher LAI values than the simulated one for all the f_{NDVI_LAI} functions, except for Song et al. (2008). The f_{NDVI_LAI} developed by Chaurasia Sasmita et al. (2011) computed the highest LAI value (more than 3 m^2 m^{-2}). This implies an earlier reaching of close canopy stage, in turns causing a higher rate of accumulation of aboveground biomass immediately after the forcing event (Figure 11).

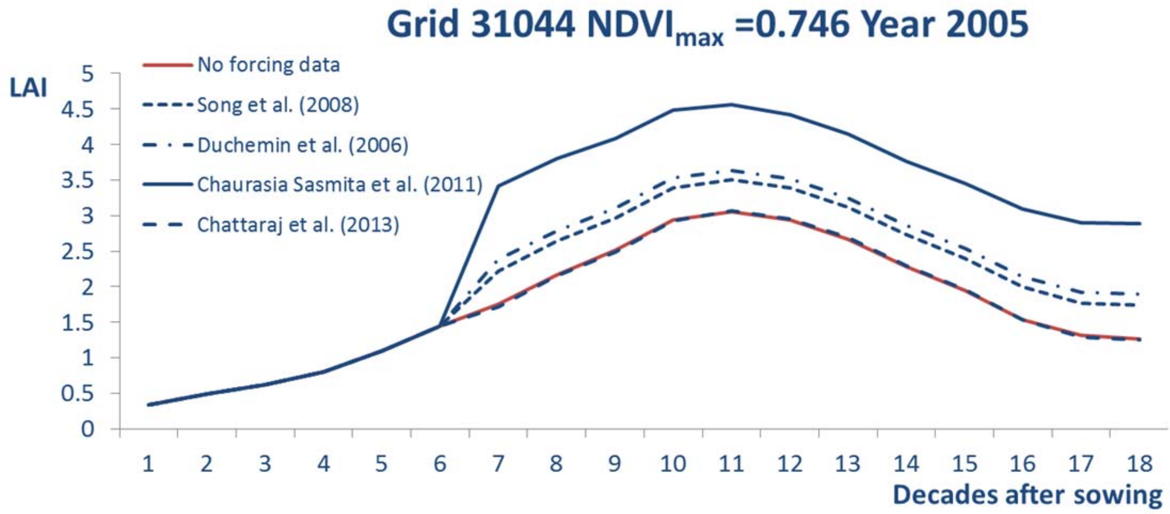

Figure 10: LAI simulations performed with the BioMA WOFOST model in the CGMS grid cell 31044 and year 2005. Comparison of the LAI trends without the updating of remotely sensed data and with the assimilation of maximum NDVI using the four functions tested.

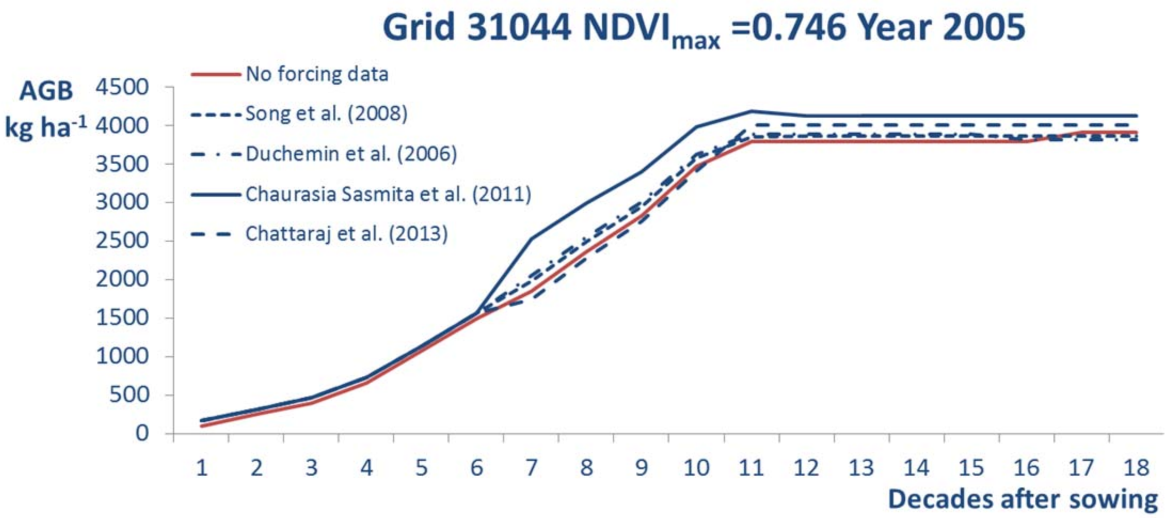

Figure 11: Aboveground biomass evolution simulated with the BioMA WOFOST model in the CGMS grid cel 31044 and year 2005. Comparison of the time trends without the updating of remotely sensed data and with the assimilation of maximum NDVI using the four functions tested.

3.2. Evaluation of the impact of the f_{NDVI_LAI} functions on the forecasting of official statistics

The official yields statistics at national level and in a subset of the Moroccan regions (i.e., Centre Nord, Sud, Tensift) were used to evaluate the impact of the four f_{NDVI_LAI} functions on the forecasting ability of the BioMA model WOFOST. Figure 12 reports the average results for durum wheat crop referred to the forecasting performed at DOY 60, 90, 120 and 150.

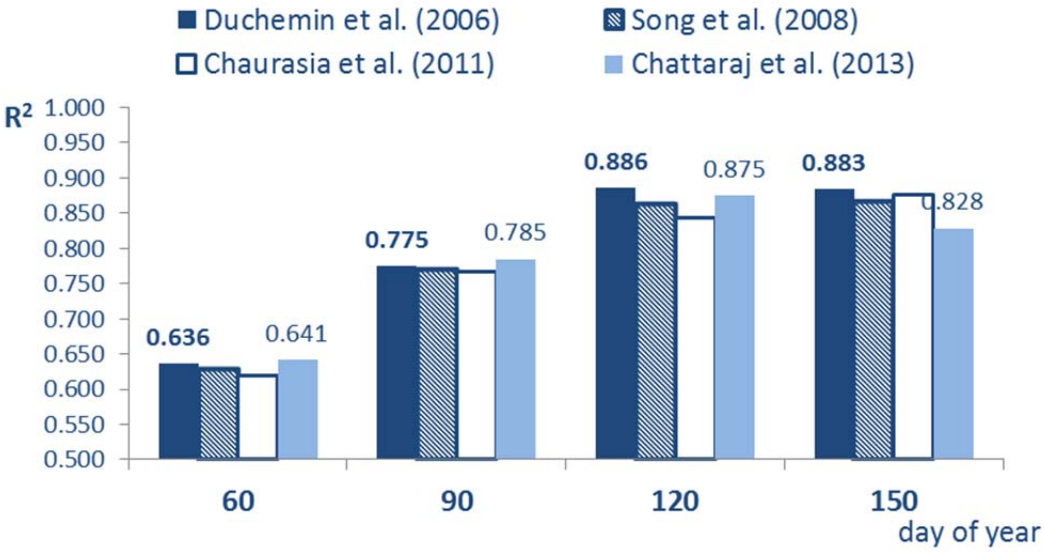

Figure 12: Average R^2 values obtained by the multiple regression model built with outputs simulated by the BioMA-WOFOST model forced with the four f_{NDVI_LAI} functions. Forecasting of the official statistics refers to a subset of Moroccan regions in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year.

The correlation values (R^2)obtained by the forecasting carried out with the four f_{NDVI_LAI} functions are reported in Table 3.
The two f_{NDVI_LAI} developed by Duchemin et al. (2006) and Chattaraj et al. (2013) obtained the best R^2 in 7 out of 20 cases, respectively. On average, the f_{NDVI_LAI} function by Duchemin et al. (2006) obtained the highest correlation value (R^2=0.795) among the four function tested, therefore it was used to the extensive comparison with no forcing simulations.

Table 1: Correlation (R^2) values obtained by the multiple regression models built with outputs simulated by the BioMA-WOFOST model forced with the four f_{NDVI_LAI} functions. Forecasting of the official statistics was carried ou on a subset of Moroccan regions in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year

f_{NDVI_LAI}	DOY	National	Centre Nord	Sud	Tensift	Average
Duchemin et al. (2006)	60	0.515	0.437	**0.840**	0.753	0.636
	90	0.773	**0.761**	**0.936**	0.630	0.775
	120	0.893	**0.903**	**0.928**	0.819	**0.886**
	150	**0.943**	**0.832**	0.931	**0.827**	**0.883**
Song et al. (2008)	60	0.512	0.425	0.837	0.744	0.629
	90	0.772	0.759	0.925	0.625	0.770
	120	0.894	0.822	0.917	0.819	0.863
	150	0.928	0.804	0.932	0.807	0.868
Chaurasia et al. (2011)	60	0.514	0.341	0.838	**0.781**	0.618
	90	0.751	0.744	0.935	0.639	0.767
	120	0.853	0.791	0.921	0.807	0.843
	150	0.919	0.830	**0.942**	0.813	0.876
Chattaraj et al. (2013)	60	**0.531**	**0.448**	0.834	0.754	**0.641**
	90	**0.777**	0.732	0.921	**0.709**	**0.785**
	120	**0.929**	0.830	0.900	**0.840**	0.875
	150	0.928	0.807	0.910	0.668	0.828

3.3. Comparison with no forced simulations

3.3.1. Average comparison

The comparison of the accuracy of forced and not forced simulations in forecasting soft and durum wheat official yields in Morocco is shown in Figure 13. Average correlation values (R^2) and standard deviations obtained by the best regressive model (i.e., the one including potential and water limited LAI) at national and regional level are reported. The assimilation of exogenous RS data determined an increase in the correlation of the simulated and official yields in all the combinations crop (soft and durum wheat) × DOY (60, 90, 120, 150) tested.

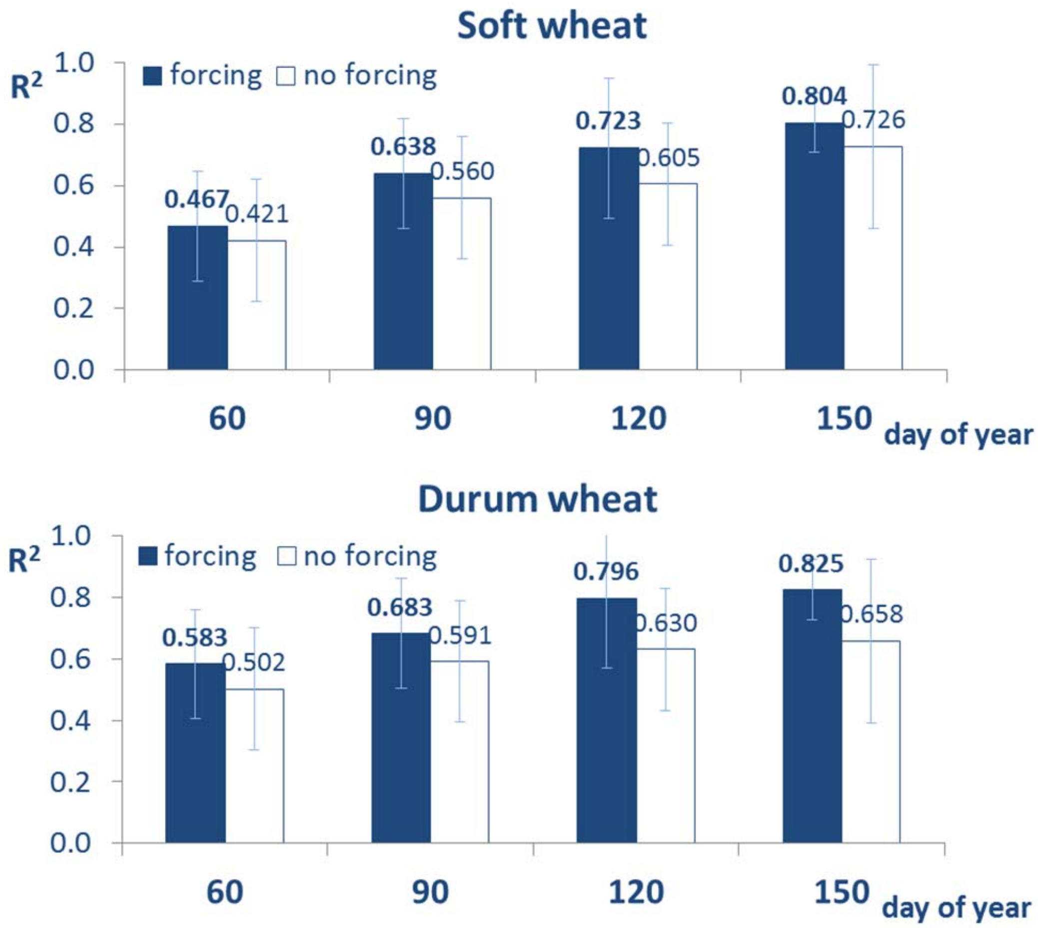

Figure 13: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the official statistics refers to all the Moroccan wheat regions in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year.

In general, the benefit due to forcing is more evident in the forecasting at DOY 90 and 120, which correspond to a deep increase of the percentage of cells in which $NDVI_{max}$ is observed (Figure 14). Durum and soft wheat crops showed a similar trend in the average improvement due to the forcing event. The standard deviation bars indicate a marked spread of the data about the mean value.

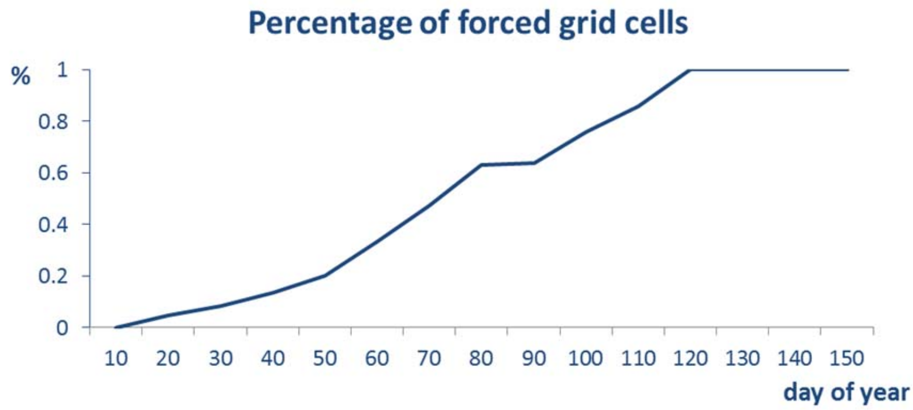

Figure 14: Percentage of Moroccan grid cells in which $NDVI_{max}$ was reached and the forcing event was implemented as a function of time in the period 1998-2011 (source CGMS data).

3.3.2. National level

At national level, the improvement due to forcing is evident both for durum and soft wheat, especially when the forecasting was carried out at DOY 90 and DOY 120. The two methodologies (i.e., forced and not forced simulation) showed very similar accuracy in earlier stages of the growth cycle (DOY 60) or close to the maturity stage (DOY 150). The regressive model built with forced indicators reaches very high R^2 values (0.943 for durum wheat at DOY 150 and 0.956 for soft wheat at DOY 120), presenting an increasing trend over time.

By analysing the graphs related to the simulated and official yields (Figure 15 and 16), it emerged that forced and not forced regressive models are concordant and not very accurate at DOY 60. The correlation with official statistics improved at DOY 90 and 120, with forced simulations producing outputs very close to the official ones. At DOY 150, the forced and not forced predicred yields almost match completely. Since no trend was computed by the CGMS Statistical Tool Box in the official yields at National level, these results refer only to the BIOMA model WOFOST predictive ability.

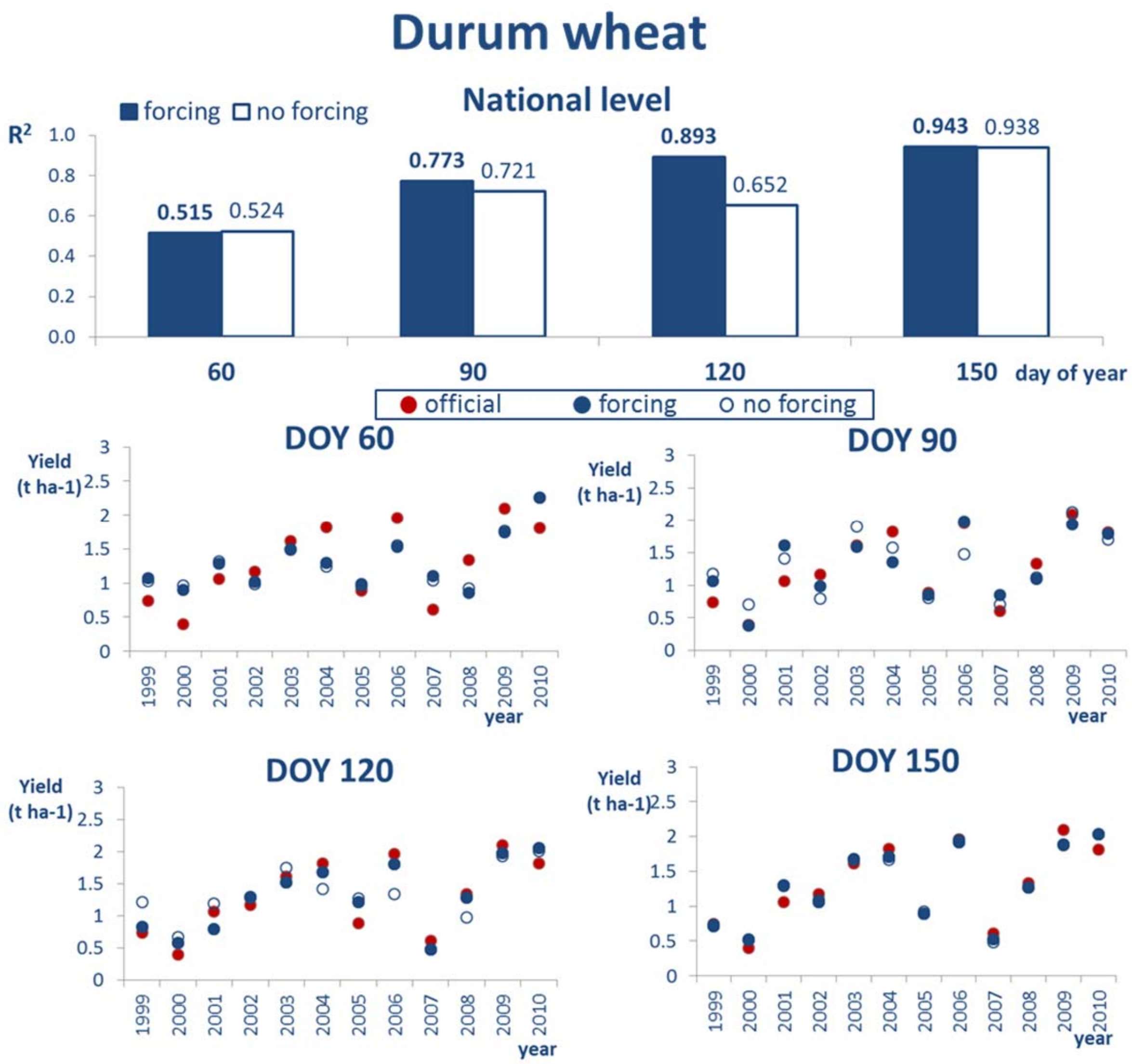

Figure 15: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics refers to the Moroccan national level in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

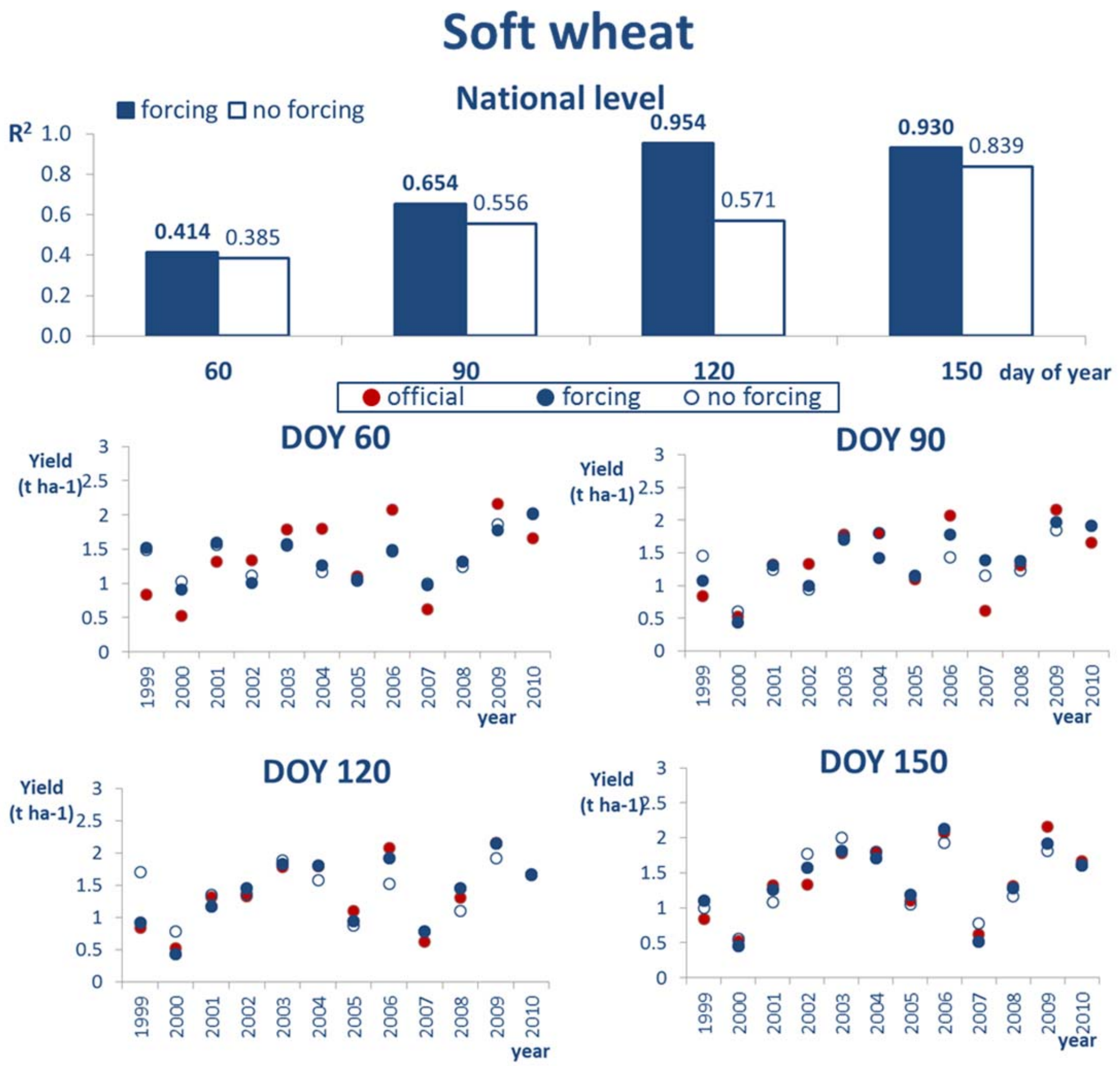

Figure 16: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan national level in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.3. Centre region

No improvements due to forcing can be observed in the Centre region when forecasting was performed at DOY 60 and 90, since the regressive model including not forced indicators is slightly more correlated with official yields. This situation is common both for durum and soft wheat. An improvement due to forcing can be noticed for durum wheat when forecasting is made at DOY 120 (R^2=0.812 versus R^2=0.786) and 150 (R^2=0.784 versus R^2=0.595), whereas for soft wheat crop the correlation of the forced indicators is always equal or below the one of the not forced regressive model. In general forced indicators obtained a high correlation with official yields (R^2=0.812 for durum wheat at DOY 120 and R^2=0.655 for soft wheat at DOY 150). By analysing the graphs related to the simulated and official yields (Figure 17 and 18), it emerged that the forced and the not forced regressive models predict very similar values at DOY 120, whereas they differentiate in all the other time periods (DOY 60, 90 and 150). Since no trend was computed by the CGMS Statistical Tool Box in the official yields in the Centre Region, these results refer only to the BioMA model WOFOST predictive ability.

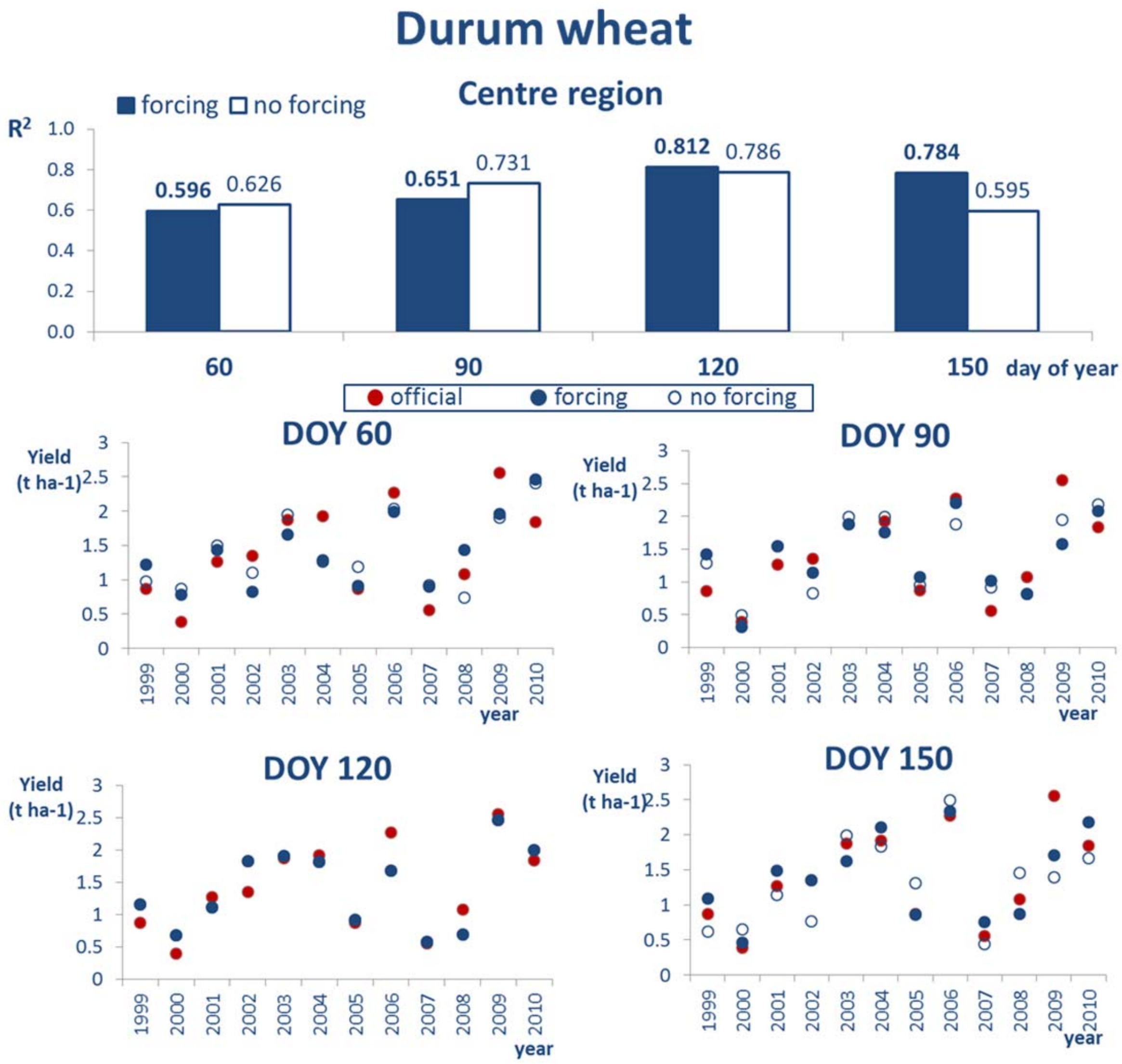

Figure 17: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics refers to the Moroccan Centre region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

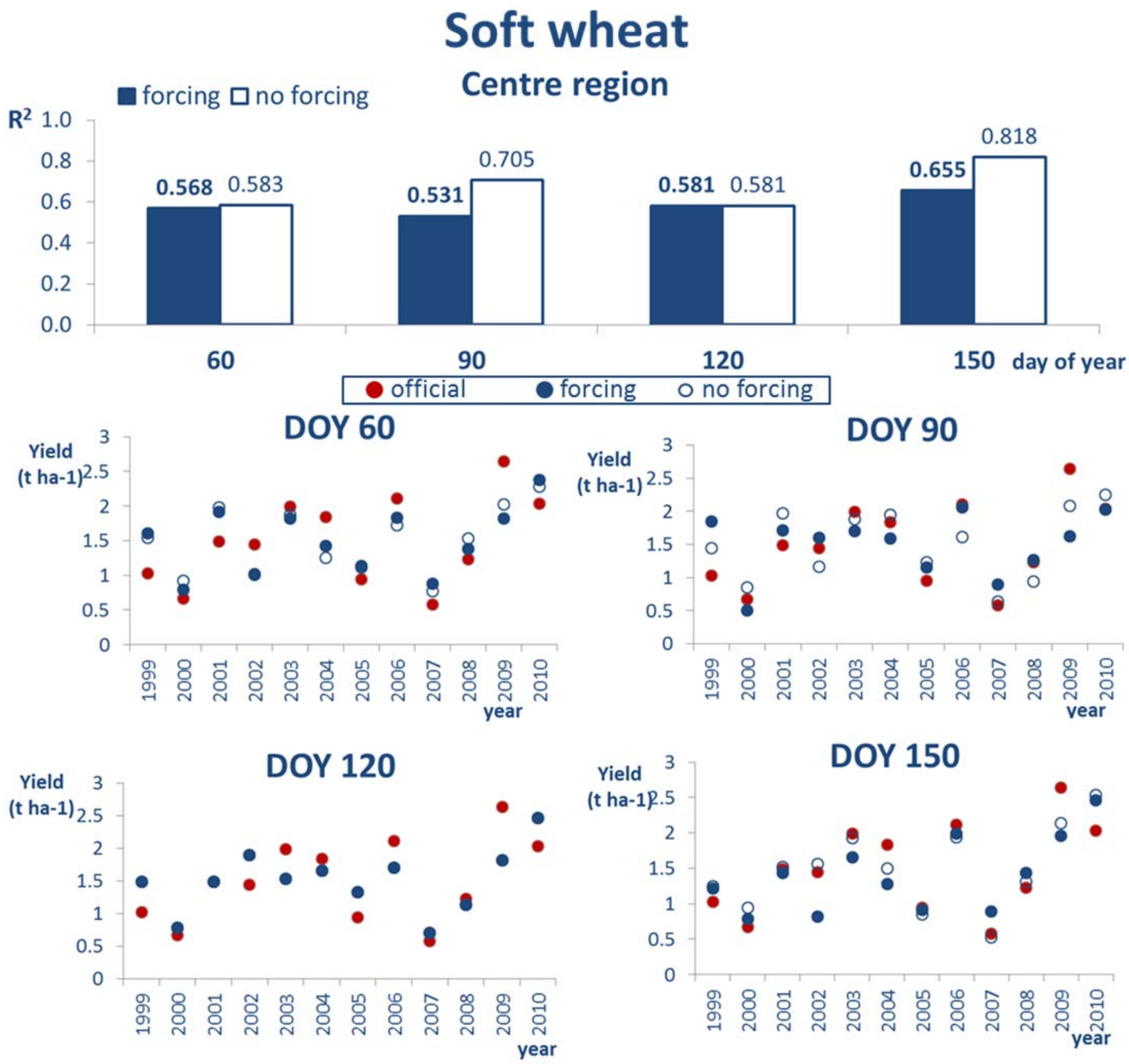

Figure 18: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Centre region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.4. Sud region

It can be observed a high correlation of simulated and official durum wheat yields in the Sud Region, obtained by not forced and forced indicators, and according to the different periods in which the forecasting was performed. A marked improvement due to forcing can be observed for soft wheat, especially at DOY 90 (R^2=0.938 versus R^2=0.695). No increasing trend can be observed in the predictive ability of the regressive models during the crop cycle.

By analysing the graphs related to the simulated and official yields (Figure 19 and 20), it emerged a marked trend in the official yields (R^2=78.1).

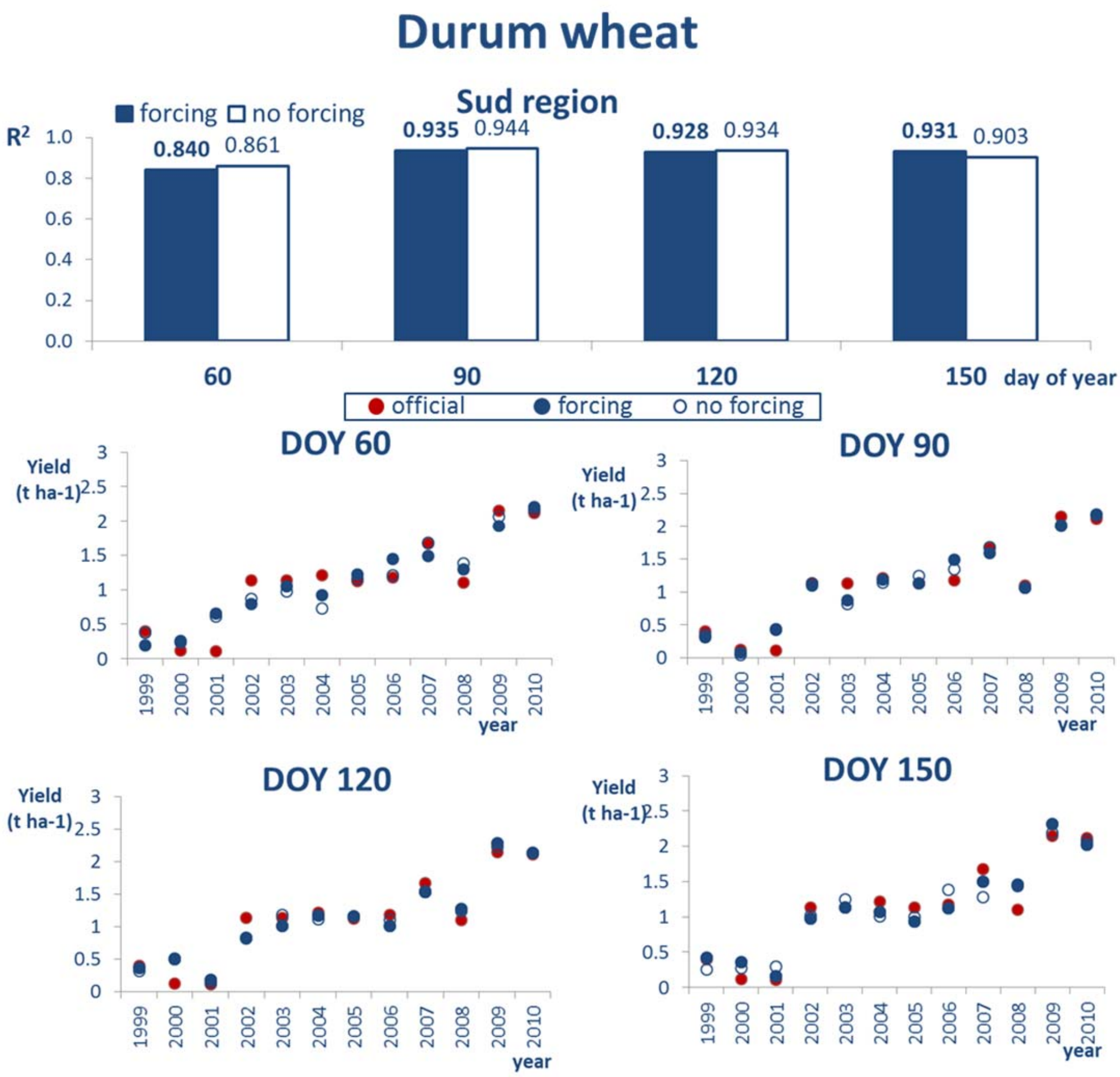

Figure 19: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics

refers to the Moroccan Sud region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

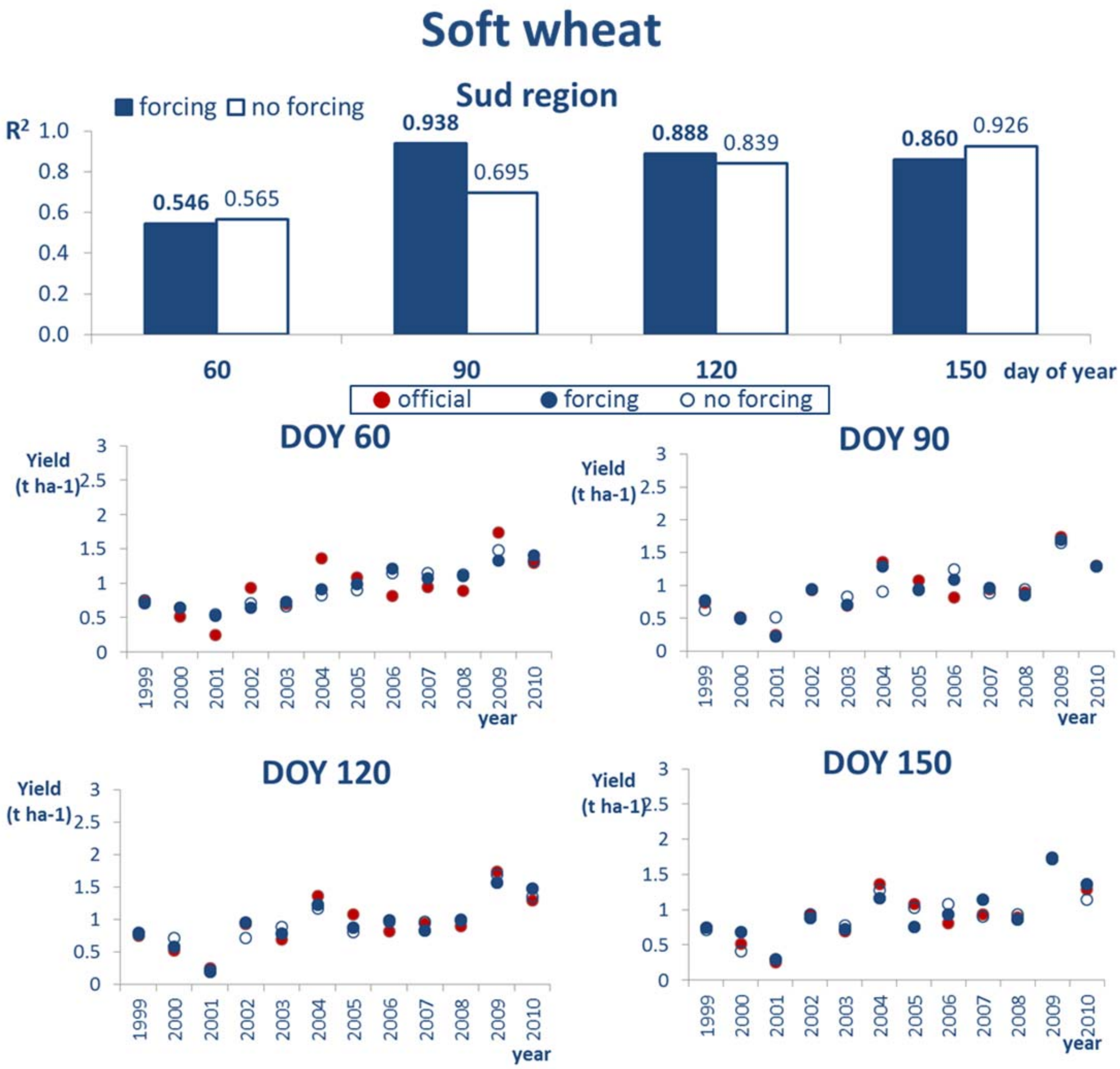

Figure 20: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Sud region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.5. Centre Nord region

A marked and constant improvement due to forcing can be observed in the Centre Nord region, both for durum and soft wheat. This is more pronounced at DOY 90 (e.g., R^2=0.761 versus R^2=0.559 for durum wheat) and especially at DOY 120 (e.g., R^2=0.903 versus R^2=0.548 for soft wheat). Results obtained in this region are very important since the regressive models including not forced indicators did not improve their accuracy according to the evolution of the crop cycle, thus indicating a real improvement attainable by the assimilation of RS data in the simulation. When the forecasting was made at DOY 150, the two methodologies allowed to obtain very similar results.
By analysing the graphs related to the simulated and official yields (Figure 19 and 20), it emerged that the forced and not forced regressive models are not concordant since DOY 120 both for soft and durum wheat. At DOY 150 the two methodologies agree in predicting very similar outputs, close to the official yields. Since there is no trend in the official durum and soft yields in the Centre Nord Region, these results refer only to the BioMA model WOFOST predictive ability.

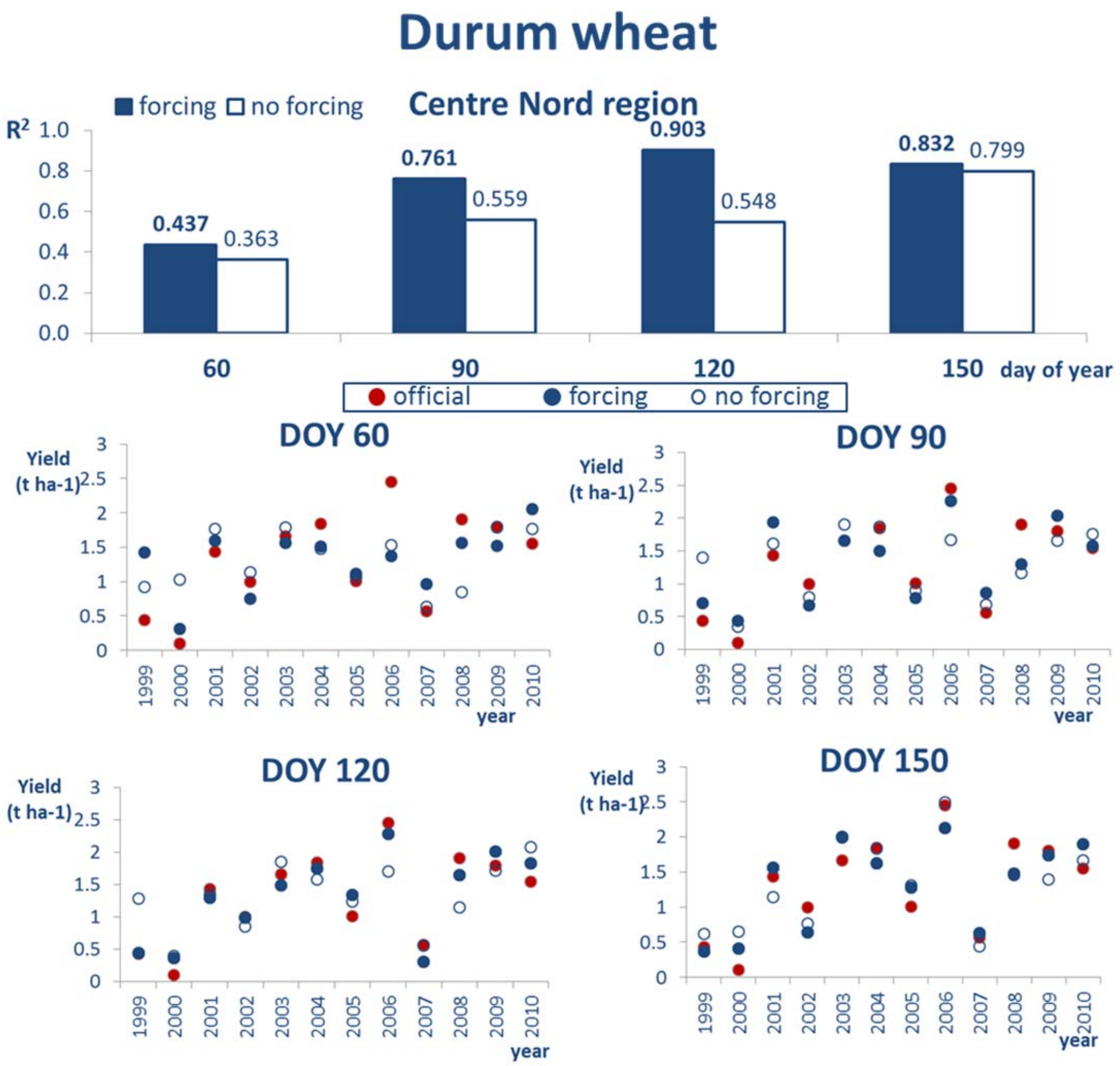

Figure 21: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics refers to the Moroccan Centre Nord region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

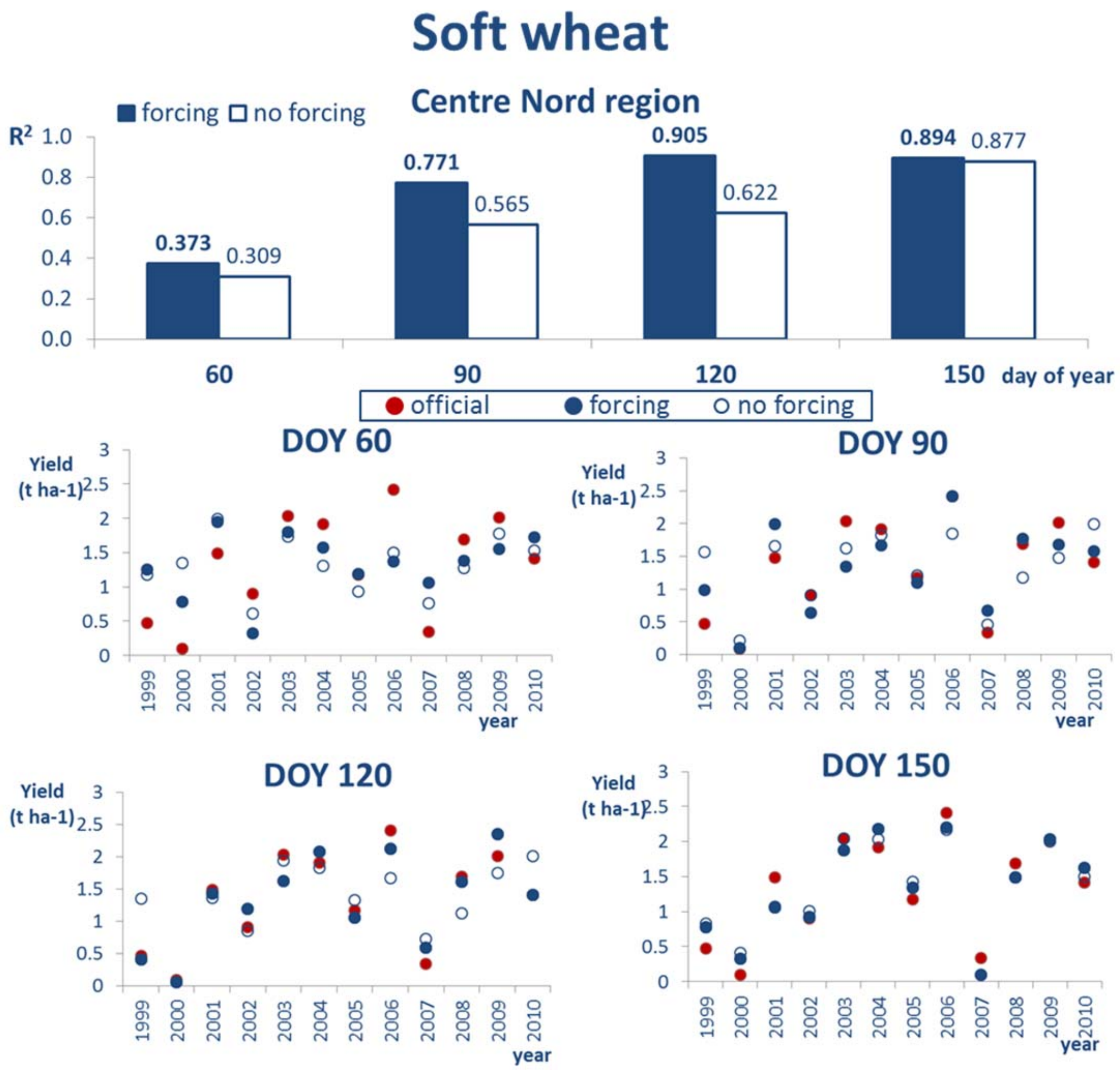

Figure 22: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Centre Nord region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.6. Tensift region

A slight improvement due to forcing can be observed in the Tensift region when the forecasting of official durum wheat yields was performed at DOY 60 (R^2=0.753 versus R^2=0.697), 90 (R^2=0.630 versus R^2=0.596) and 120 (R^2=0.819 versus R^2=0.753), whereas at DOY 150 the regressive model including not forced indicators performed slightly better

(R^2=0.827 versus R^2=0.849). A different situation is depicted for soft wheat, where forcing allowed to markedly improve the accuracy of yield forecasting at DOY 120 (R^2=0.800 versus R^2=0.607), whereas it was less correlated with official yields in the other time windows. In general the R^2 values reached by forced indicators highlight a good correlation with official yields (R^2=0.827 for durum wheat at DOY 150 and R^2=0.833 for soft wheat at DOY 150), even if lower than the maximum R^2 value obtained by not forced indicators for soft wheat when forecasting was performed at DOY 150. By analysing the graphs related to the simulated and official yields (Figure 23 and 24), it emerged that the forced and not forced regressive models are more concordant for durum wheat than for soft wheat. Since there is no trend in the official durum and soft yields in the Tensift, these results refer only to the BioMA model WOFOST predictive ability.

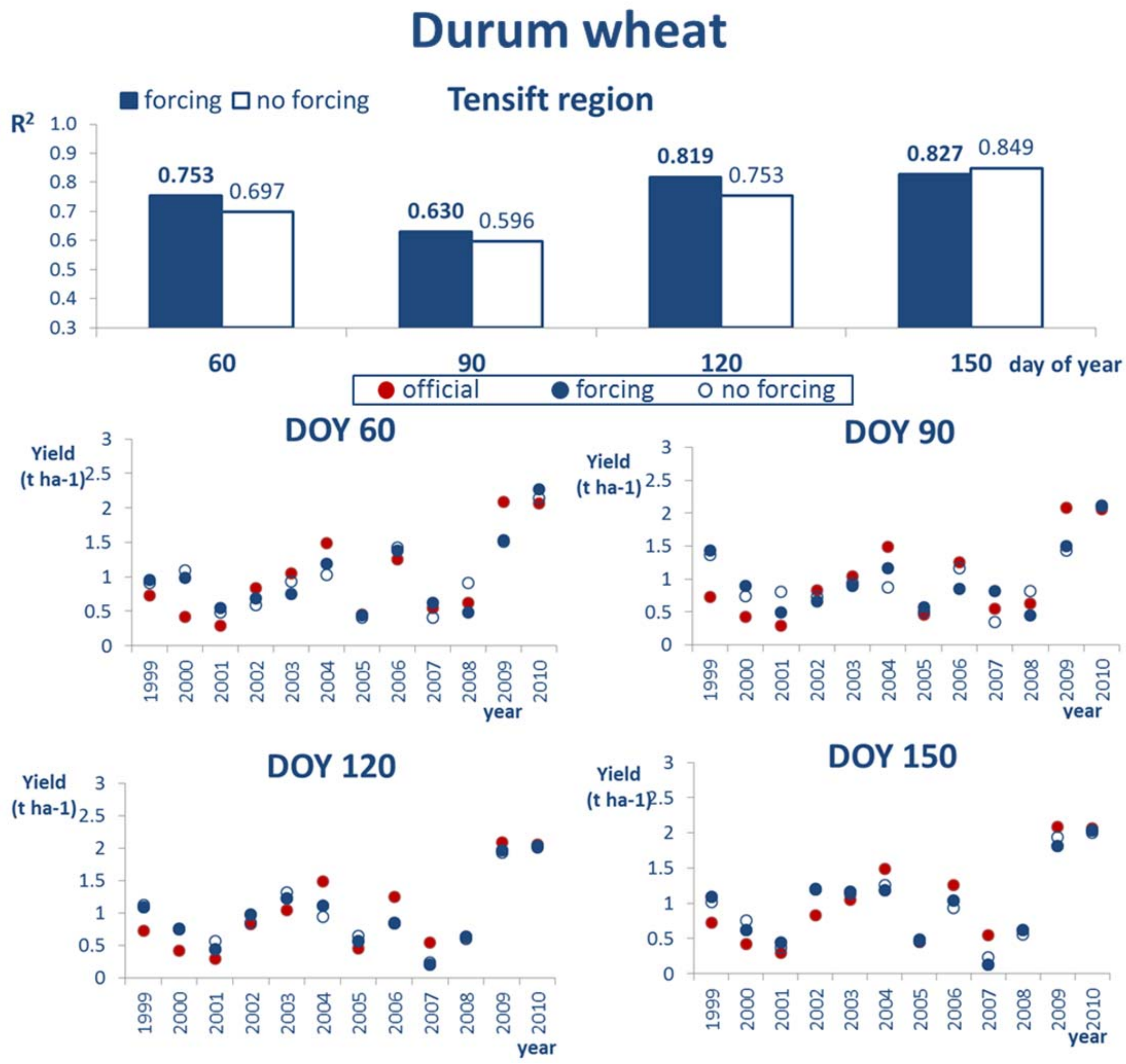

Figure 23: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Tensift region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

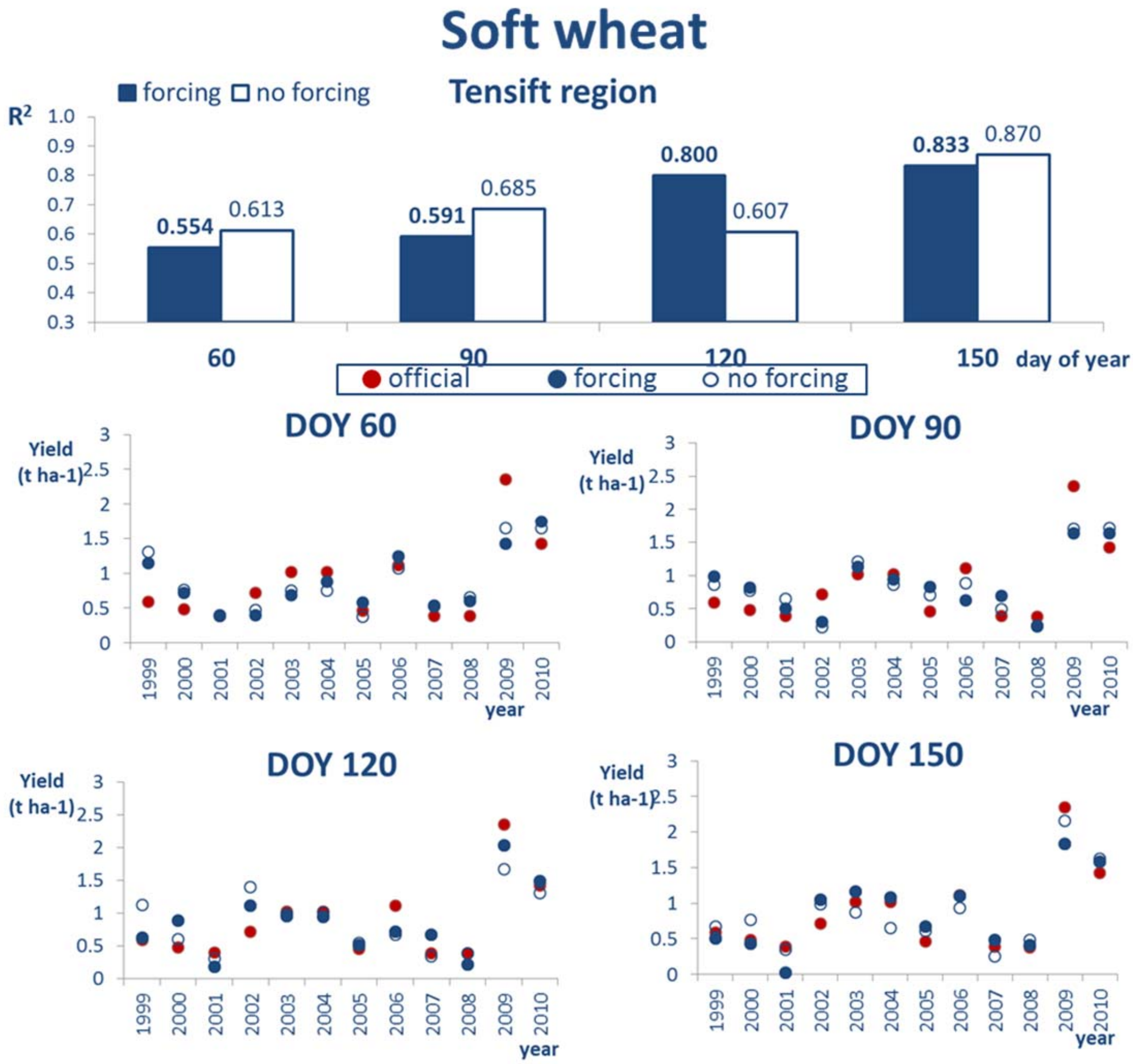

Figure 24: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Tensift region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.7. Nord Ouest region

No decided improvements due to forcing and an overall lower performance of the regressive models can be observed in the Nord Ouest region than in the other Moroccan wheat regions, especially for soft wheat. A sligh improvement can be observed for durum wheat at DOY 120 (R^2=0.630 versus R^2=0.533) at DOY 150 (e.g., R^2=0.640 versus R^2=0.550) for durum wheat.

By analysing the graphs related to the simulated and official yields (Figure 25 and 26), it emerged that the forced and not forced regressive models are very concordant in all the combinations DOY×crop tested. Since there is no trend in the official durum and soft yields in the Nord Ouest Region, these results refer only to the BioMA model WOFOST predictive ability.

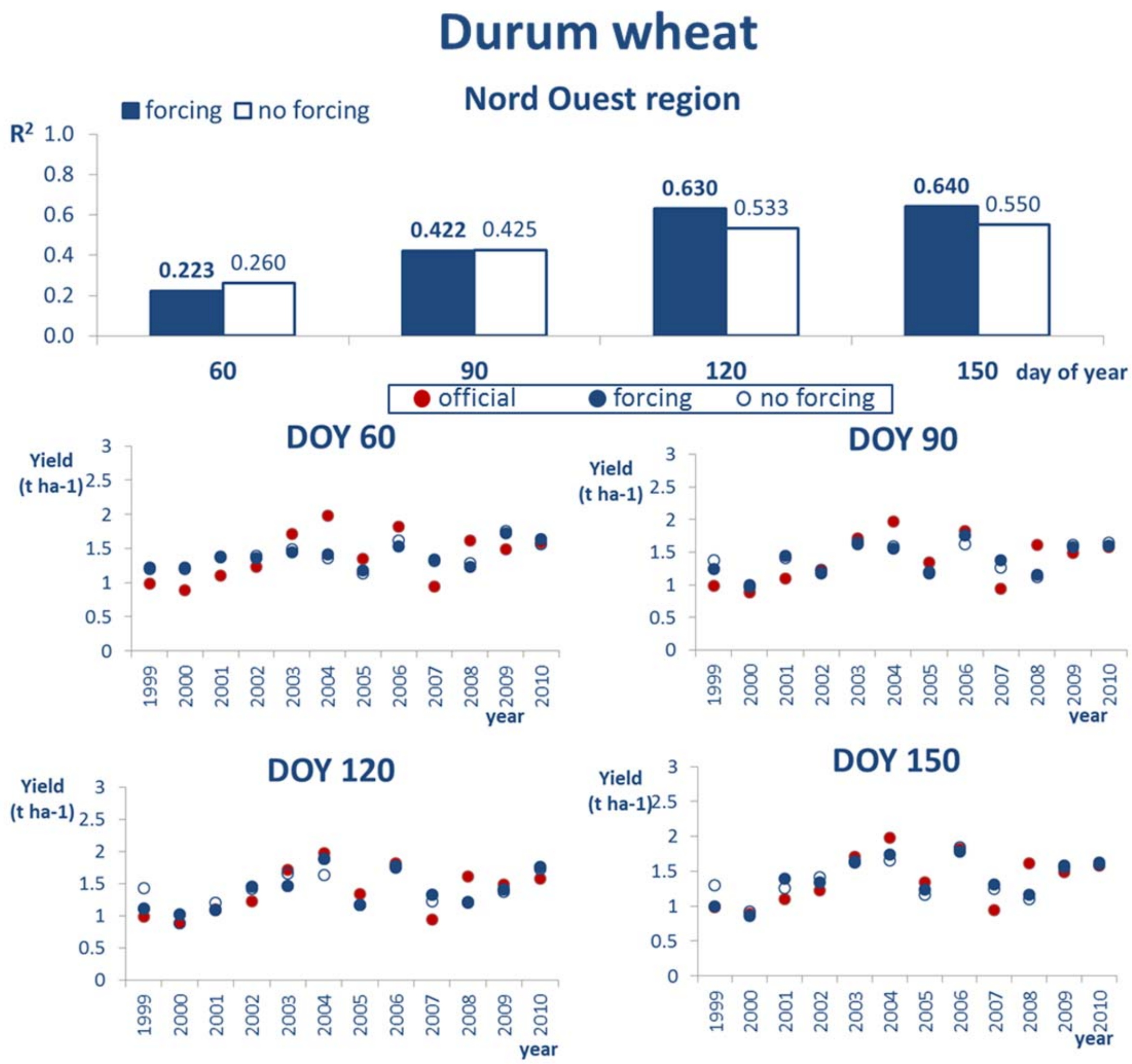

Figure 25: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics refers to the Moroccan Nord Ouest region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

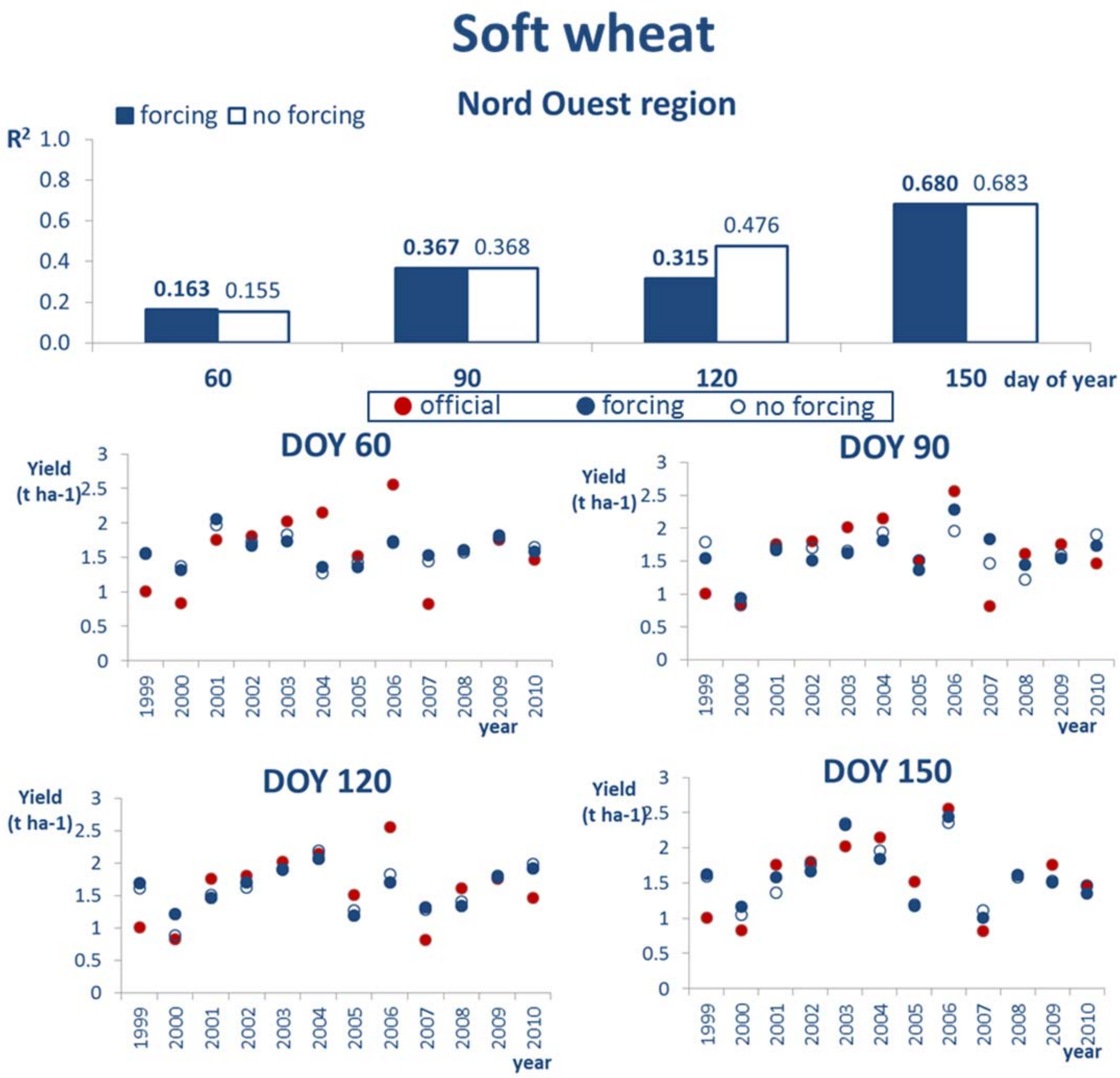

Figure 26: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Nord Ouest region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.8. Centre Sud region

The assimilation of RS data determined improvements in the Centre Sud region only when forecasting is performed at DOY 150 both for durum (R^2=0.760 versus R^2=0.534) and soft (R^2=0.796 versus R^2=0.727) wheat. In all the other cases, the regressive model built with not forced indicators performed shlightly better for the two crops. In general the R^2 values reached by forced indicators show a discrete correlation with official yields (maximum R^2=0.760 for durum wheat at DOY 150 and R^2=0.796 for soft wheat at DOY 150), in line with the ones obtained by not forced indicators. By analysing the graphs related to the simulated and official yields (Figure 27 and 28), it emerged that the forced and not forced regressive models present different predicted values in all the DOY of forecasting. The trend computed in the Centre Sud by the CGMS statistical tool box for durum wheat yields was 26%, whereas no trend was computed for soft wheat yields.

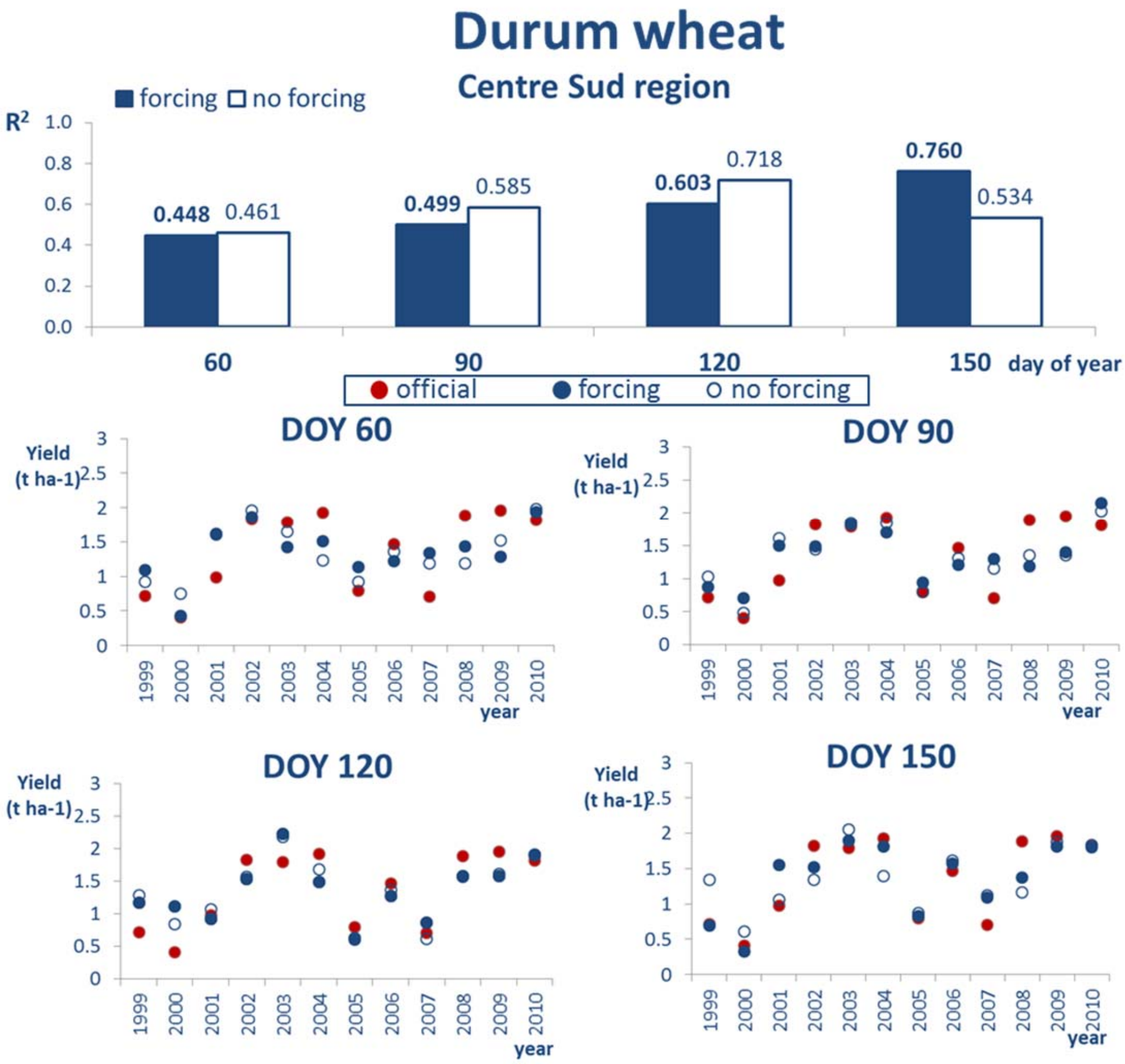

Figure 27: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics

refers to the Moroccan Centre Sud region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

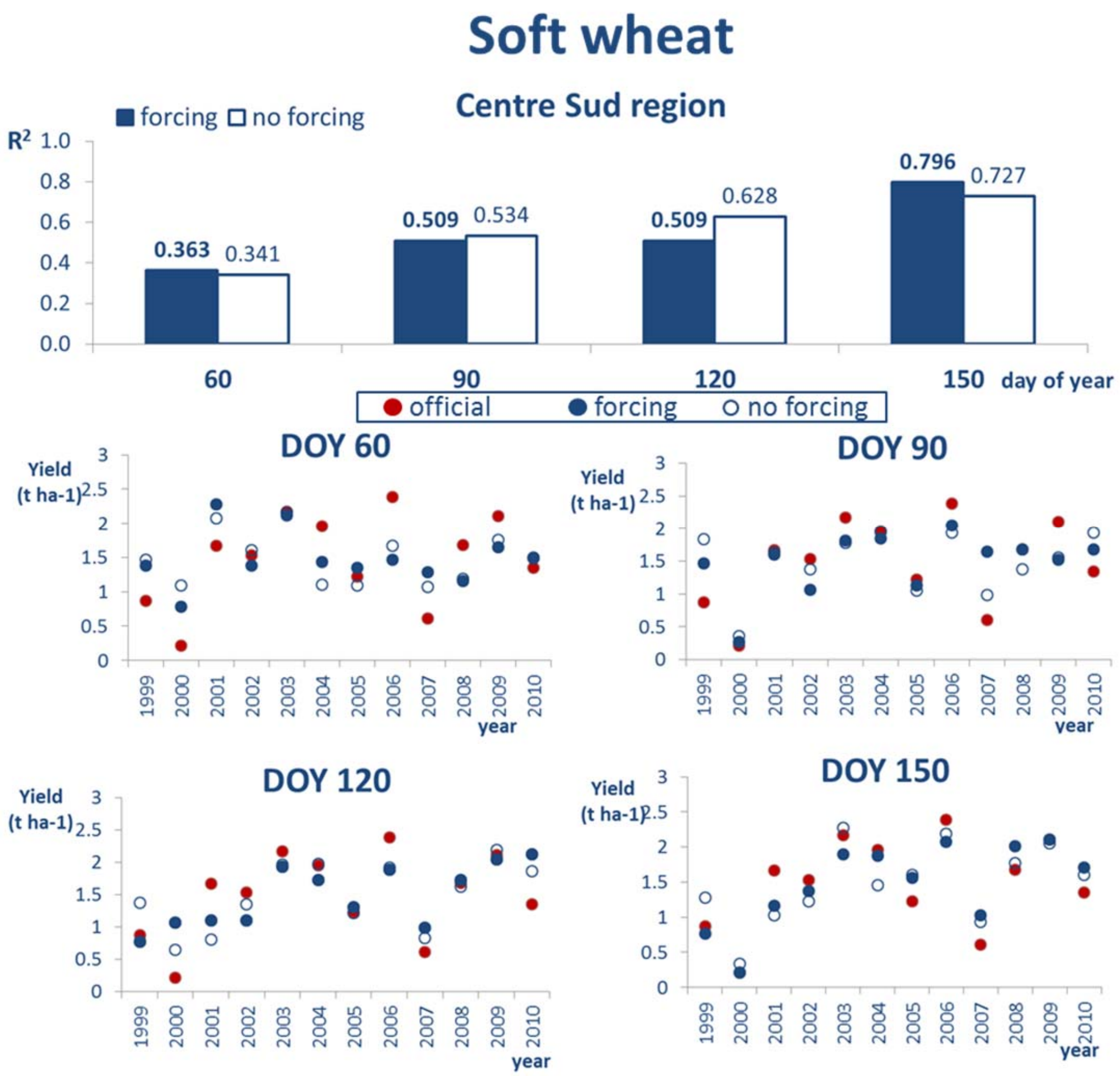

Figure 28: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Centre Sud region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

3.3.9. Oriental region

The analysis of the predictive ability of the regressive models in the Oriental region highlighted a very similar situation than the Sud region one, with a high correlation of simulated and official durum wheat yields in the different periods of forecasting. No clear benefits due to forcing can be observed in the Oriental region. In general the R^2 values reached by forced indicators show a very good correlation with official yields (maximum R^2=0.880 for durum wheat at DOY 150 and R^2=0.828 for soft wheat at DOY 120), in line with the ones obtained by not forced indicators

No increasing trend can be observed in the predictive ability of the forced and not forced regressive models during the crop cycle. By analysing the graphs related to the simulated and official yields (Figure 19 and 20), it emerged that very similar yields were predicted by the forced and not forced regressive models, with no trend computed by the CGMS statistical tool box for the two crops.

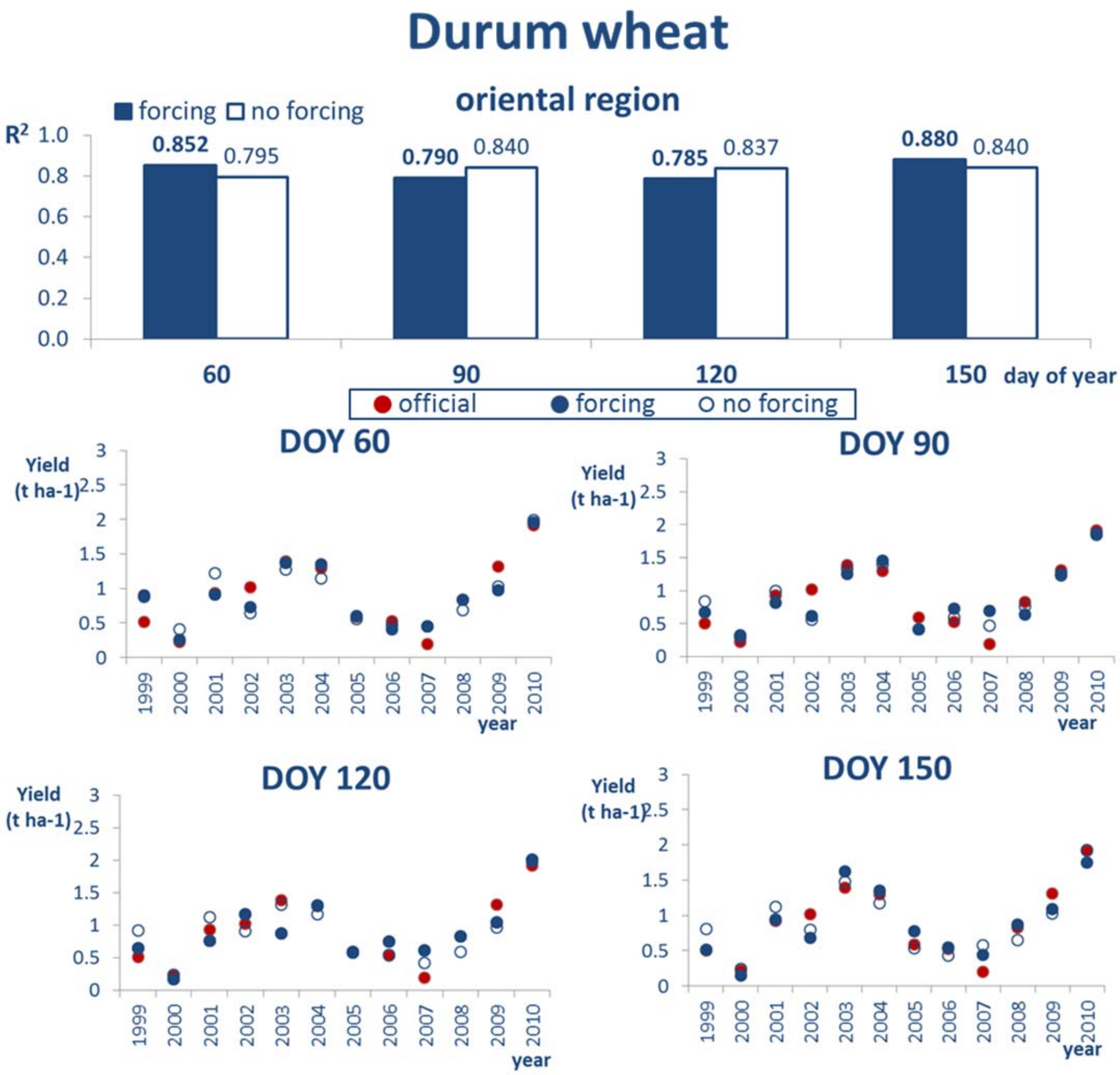

Figure 29: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the durum wheat official yield statistics refers to the Moroccan Oriental region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

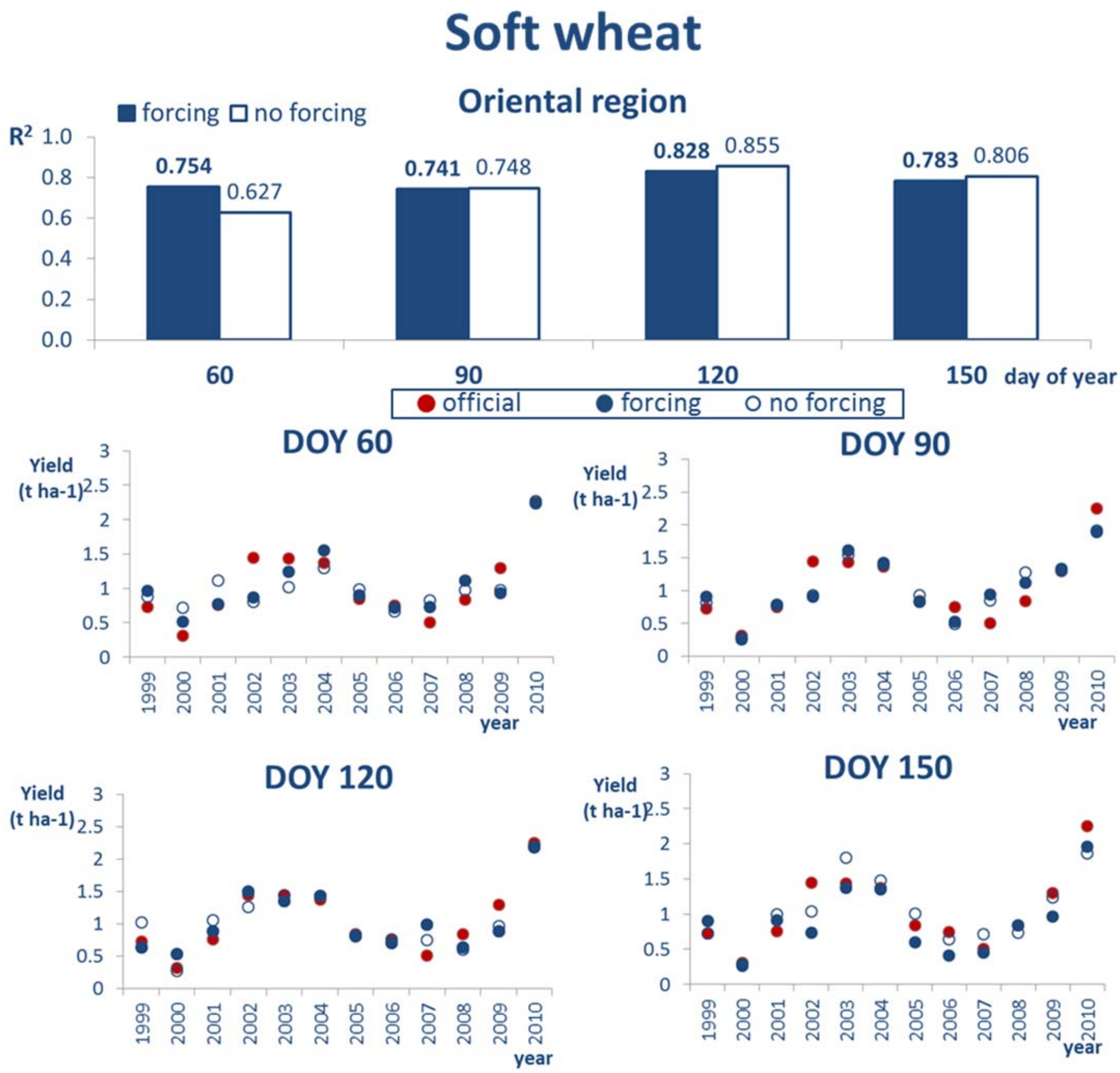

Figure 30: Average R^2 values obtained by the multiple regression models built with forced and no forced simulation outputs. Forecasting of the soft wheat official yield statistics refers to the Moroccan Oriental region in the period 1998-2011 and was performed at 60, 90, 120 and 150 day of the year. The official and the forecasted simulated yields 60, 90, 120 and 150 day of year are plotted.

4. Conclusions

The assimilation of RS data into the BioMA-WOFOST model determined an overall improvement of the accuracy of wheat forecasting in Morocco. The enhancement of the correlation between simulated and official yields is maximized when the forecasting is made from March (DOY 90) to April (DOY 120), and tend to decrease when the crop is reaching the maturity stage (DOY 150). The comparison of four methods to derive LAI values from NDVI data allowed to highlight their dissimilarities and to identify the best approach in the tested conditions. The analysis of the performance of the regressive models in the different Moroccan wheat regions presented a marked heterogeneity of the impact of forcing on the predictive ability of the BioMA model WOFOST, thus suggesting a case by case assessment and a further analysis before its operative implementation in the forecasting activities.

This study encourages further activities aimed at assessing assimilation techniques within forecasting systems, although marked among empirical methods used to derive LAI and differences across regions suggest to avoid – at the moment – the use of such technique within operational contexts.

5. References

Aggarwal, P.K., 1995. Uncertainties in crop, soil and weather inputs used in growth models: implication for simulated outputs and their application. Agricultural Systems 48, 361-384.

Bastiaanssen, W.G.M., Ali, S., 2003. A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan. Agriculture Ecosystem Environmen 94, 321-340.

Bouman, B.A.M., 1995. Crop modeling and remote-sensing for yield prediction. Netherlands Journal of Agricultural Science 43, 143-161.

Chattaraj, S., Chakraborty, D., Garga, R.N., Singh, G.P., Gupta, V.K., Singha, S., Singh, R., 2013. Hyperspectral remote sensing for growth-stage-specific water use in wheat. Field Crops Research 144, 179-191.

Chaurasia Sasmita, Nigam R, Bhattacharya BK, Sridhar VN, Mallick K, Vyas SP, Patel NK, Mukherjee J, Chander Shekhar, Kumar D, Singh KRP, Bairagi GD, Purohit NL, Parihar JS, 2011. Development of regional wheat VI-LAI models using Resourcesat-1 AWiFS data. J. Earth Syst. Sci. 120, 1113–1125

Cihlar, J., Chen, J., Li, Z., Latifovic, R., Fedosejevs, G., Adair, M., Park, W., Raser, R., Trishchenko, A., Guindon, B., Stanley, D., Morse, D., 2002. GeoComp-n, an advanced system for the processing of coarse and medium resolution satellite data. Part: Biophysical products for northern ecosystems. Canadian Journal of Remote Sensing 28, 21-44.

Confalonieri, R., Bregaglio, S., Cappelli, G., Francone, C., Carpani, M., Acutis, M., El Aydam, M., Niemeyer, S., Balaghi, R., Dong, Q., 2013. Wheat modeling in Morocco unexpectedly reveals predominance of photosynthesis versus leaf area expansion plant traits. Agronomy for Sustainable Development 33, 393-403.

Dorigo, W.A., Zurita-Milla, R., de Wit, A.J.W., Brazile, J., Singh, R., Schaepman, M.E., 2007. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling. International Journal of Applied Earth Observations and Geoinformation 9, 165-193.

Duchemin, B., Hadria, R., Erraki, S., Boulet, G., Maisongrande, P., Chehbouni, A., Escadafal, R., Ezzahar, J., Hoedjes, J.C.B., Kharrou, M.H., Khabba, S., Mougenot, B., Olioso, A., Rodriguez, J.C., Simonneaux, V., 2006. Monitoring wheat phenology and irrigation in Central Morocco: On the use of relationships between evapotranspiration, crops coefficients, leaf area index and remotely-sensed vegetation indices. Development of regional wheat VI-LAI models using Resourcesat-1 AWiFS data. Agricultural Water Management 79, 1-27.

Gobron, N., Pinty, B., Verstraete, M.M., Widlowski, J.L., 2000. Advanced vegetation indices optimized for up-coming sensors— design, performance, and applications. IEEE Transactions on Geoscience and Remote Sensing 38, 2489-2505.

Launay, M., Guèrif, M., 2005. Assimilating remote sensing data into a crop model to improve predictive performance for spatial applications. Agriculture, Ecosystem Environment 111, 321–339.

Liu, W.T., Kogan, F., 2002. Monitoring Brazilian soybean production using NOAA/AVHRR based vegetation condition indices. International Journal of Remote Sensing 2, 1161-1179.

Mo, X., Liu, S., Lin, Z., Xu, Y., Xiang, Y., McVicar, T.R., 2005. Prediction of crop yield, water consumption and water use efficiency with a SVAT-crop growth model using remotely sensed data on the North China Plain. Ecological Modelling 183, 301-322.

McLaughlin, D., 2002. An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering. Advances in Water Resources 25, 1275-1286.

Roerink, G.J., Menenti, M., Verhoef,W., 2000. Reconstructing cloudfree NDVI composites using Fourier analysis of time series. International Journal of Remote Sensing 21, 1911-1917.

Song, J., Wang, J., Xiao, Y., Shuai, Y., Huawei, W., 2008. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B1. Beijing.

Van Keulen, H., Wolf, J., 1986. Modelling of agricultural production: weather soils and crops. Simulation Monographs. Pudoc, Wageningen, The Netherlands, pp. 479.

Xin, J., Zhenrong, Y., van Leeuwenb, L., Driessen, P.M., 2002. Mapping crop key phenological stages in the North China plain using NOAA time series images. International Journal of Applied Earth Observation and Geoinformation 4, 109-117.

Wang, F.M., Huang, J.F., Tang, Y.L., Wang, X.Z., 2007. New vegetation index and its application in estimating leaf area index of rice. Rice Science 14, 195-203.