Assimilation of ASAR Data for Wheat Yield Prediction: Matera Case Study

Laura Dente(1), Michele Rinaldi(2), Francesco Mattia(1), Giuseppe Satalino(1)

(1)Istituto di Studi sui Sistemi Intelligenti per l’Automazione, ISSIA-CNR
via Amendola 122/D, 70126 Bari, Italy
Email: Dente@ba.issia.cnr.it

(2)Istituto Sperimentale Agronomico, ISA
via Celso Ulpiani 5, 70125 Bari, Italy

Abstract—The objective of this work is to investigate the synergistic use of leaf area index (LAI) retrieved by ENVISAT ASAR data and crop growth models, such as CERES-Wheat, to improve the accuracy of wheat yield predictions. The estimate reliability of CERES-Wheat strongly depends on the accuracy of its numerous inputs, which are not always available or accurate. As a consequence, the model would largely benefit from using updated information on the wheat status, provided by remote sensing at field scale. This work shows that the assimilation of ENVISAT ASAR AP data into the model lead to significant improvements in the wheat dry biomass and the grain yield model predictions.

I. INTRODUCTION

The early forecast of wheat production plays a key role in crop management and marketing strategies. Crop growth models are important tools of modern agronomy, used to predict the yield before harvest and to estimate soil-crop characteristics during growing season. However a limitation of these models, which prevents to reach accurate results, is that they need numerous input parameters concerning the wheat genotype, the soil characteristics and the crop management, which are not always available in details. The average values characterising the area of interest, generally used to run the model, can lead to erroneous predictions.

In this context, this work investigates the possibility to improve the accuracy of wheat yield estimates of crop growth models, such as CERES-Wheat, by assimilating the leaf area index (LAI) estimates retrieved from polarimetric ENVISAT ASAR data.

The potential of polarimetric SAR data for the retrieval of wheat biophysical parameters has been investigated by some recent experimental studies ([1]-[3]). They have shown that the HH/VV radar backscatter ratio, acquired at high incidence angles, is extremely sensitive to wheat biomass and to LAI, while it is almost independent on soil conditions. In particular Mattia et al. [2] provided an empirical relationship between ASAR HH/VV backscatter ratio at about 40° incidence angle and wheat biomass and LAI, obtained over the Matera site in the South of Italy.

In this work a model re-initialisation strategy has been adopted to assimilate the LAI retrieved by means of ASAR AP data into CERES-Wheat and the effects on wheat yield predictions over some fields of the Matera area have been examined.

In the next section, the exploited in situ and radar data and the adopted relationship between multi-temporal HH/VV backscatter ratio and wheat LAI are described. Then, the CERES-Wheat crop growth model and the results of its calibration and validation procedures are presented. Subsequently, the method used for the assimilation of the radar estimated biophysical parameter into the crop growth model is described. In the last section a comparison between CERES-Wheat predictions before and after the assimilation of radar data is shown and discussed.

II. MATERIALS

A. Ground Data

The experimental site is located close to Matera in the Basilicata region, South of Italy. This area is mostly dedicated to the durum wheat (i.e. Triticum durum Desf.) cultivation. The data exploited in this study were collected during a measurement campaign carried out in 2004. Six wheat fields, with size ranging between 3 and 15 ha, were selected in an area of approximately 30 km². Four of them were quantitatively monitored, while in the others only qualitative information concerning vegetation and soil status, phenological stage, sowing and harvesting date were collected.

In particular gravimetric and volumetric soil moisture and above ground fresh and dry biomass of wheat were measured. LAI data were collected using a portable probe, the AccuPAR linear PAR/LAI ceptometer, which is able to measure the photosynthetically active radiation (in the 400 to 700 nanometer waveband) intercepted by the plants and to compute the LAI. Moreover a pedological campaign was carried out in order to have detailed information on the soil texture, chemical characteristics and hydrological parameters for all the soil layers from 0 to 1 m depth. At the end of the season the grain production, the seed weight and the total
above ground dry biomass were determined over all the six monitored fields. Climatic data, such as maximum and minimum temperature, precipitation, total solar radiation, wind speed and direction, air humidity, soil temperature and soil moisture, were collected during the whole growing season by an agro-meteorological station located close to the fields.

It is worth mentioning that similar campaigns were carried out in the same site also over 4 wheat fields in 2001 and over 2 wheat fields in 2003.

B. Radar Data

The in-situ measurements and in particular those concerning the agronomic parameters were collected roughly every two weeks coincidentally to ENVISAT ASAR acquisitions (further details of this campaign can be found in [2]). From tillering to harvesting stage (i.e. February to June) 7 descending acquisitions were collected by ENVISAT ASAR in alternating polarisation mode (HH and VV) at about 40° incidence angle (swath I5 and I6) over the Matera site (ENVISAT AO-662), which is the radar configuration more suitable for the wheat biophysical retrieval ([1] - [3]).

These ASAR APS 1P products have been pre-processed (radiometric calibration, ground range projection and co-registration have been performed using the GAMMA Remote Sensing Research and Consulting AG software package). Then the backscattering coefficient of the areas corresponding to the monitored fields has been extracted from all the images.

These ASAR data and the corresponding in-situ data were extensively analysed in [2], where a strong correlation between the HH/VV backscatter ratio at about 40°, \( \sigma_{HH/VV} \), and wheat LAI from tillering to the beginning of heading stage was reported. An empirical power-law relationship, with a correlation coefficient of \( R^2 = 0.82 \), was found. The rms error on the LAI estimates, obtained by means of this law, is \( \Delta_{LAI} = 0.42 \).

The LAI estimated by the 2004 ASAR data by means of this relationship has been exploited in the assimilation technique. In particular, only in three dates from tillering to beginning of heading stage of 2004 wheat season the ASAR data and the corresponding retrieved LAI information were available. The date, the orbit and the swath of ASAR acquisitions are reported in Table 1.

<table>
<thead>
<tr>
<th>Date</th>
<th>Orbit</th>
<th>Swath</th>
</tr>
</thead>
<tbody>
<tr>
<td>03/04/04</td>
<td>10940</td>
<td>6</td>
</tr>
<tr>
<td>22/04/04</td>
<td>11212</td>
<td>5</td>
</tr>
<tr>
<td>08/05/04</td>
<td>11441</td>
<td>6</td>
</tr>
</tbody>
</table>

C. Crop model

The crop growth model selected for this study is the CERES-Wheat implemented by the software DSSAT 4.0. The CERES (Crop Environment Resource Synthesis) family of crop models is used to predict the performance of several cereal crops, especially maize and wheat. A detailed description of the model can be found in [4]. CERES-Wheat estimates among the others, LAI, biomass and yield using meteorological data and detailed information on soil properties, crop genetic characteristics and crop management. The genotype input files contain numerous parameters which define the crop growth during the season, for example the duration of the crop phenological phases and the maximum percentage of biomass partitioned into the various components of the plant at each phase. These parameters depend on the wheat variety and the calibration procedure of the model consists of varying them until the model predictions approximate the measured data.

Starting from the results obtained from previous calibration and validation of CERES-Wheat for wheat variety generally sown in South of Italy [5], the model has been calibrated for the specific variety sown in the Matera site. The calibration has been carried out considering the datasets collected in two out of the four fields monitored in 2001 and in the two fields monitored in 2003. After the calibration the mean percentage absolute difference between the estimated and the observed data, normalized by the measured value, is about 24% for the yield and about 29% for the total dry biomass at harvest. These errors are acceptable in the context of a crop growth model [6].

After the calibration procedure, the model has been validated over three fields monitored in 2004, comparing the predictions of the model for these fields with the in-situ data. The mean percentage absolute difference between estimated and observed values is about 11% for the total dry biomass at harvest and about 9% for the grain yield, demonstrating that the carried out calibration has been very successful.

III. ASSIMILATION METHOD

Different techniques can be chosen for the assimilation of remote sensing data into models, as for instance described by Bach and Mauser in [7].

In this work, the time sequence of the LAI estimated by the radar has been assimilated in the crop model by means of a re-initialisation strategy, i.e. the CERES-Wheat state variable, \( \text{LAI}_{CERES} \), has been forced to fit the \( \text{LAI}_{ASAR} \) by varying a set of model initialisation parameters. The set of inputs which need to be varied to minimize the error between the radar estimated parameter and the model state variable was identified in a previous work [8], where a preliminary sensitivity analysis of CERES-Wheat was carried out and the assimilation method was successfully tested using LAI retrieved by a C-band ground-based scatterometer. In this work it was found that three model inputs mainly affect the LAI predictions: wheat sowing date, soil wilting point and field capacity. The sowing date determines in which period of the season the wheat growth occurs and then, depending on the temperature, the length of the phenological stages. The wilting point and the field capacity, which are soil hydrological properties mainly dependent on soil texture and bulk density, define the range limits and, consequently, the amount of the soil available water content for the crop.

In order to simplify the subsequent assimilation procedure, a generic curve of the LAI state variable of the model, has been
parameterised, using a fitting procedure, by a sigmoid function of time \( \Sigma(t) \):

\[
\Sigma(t) = \frac{p_1}{1 + e^{-p_2(t+p_3)}},
\]

which depends on three parameters, i.e. \( p_1, p_2 \) and \( p_3 \). This function describes quite well the behaviour of LAI from tillering to heading stage.

The assimilation method can be schematised in two steps. In the first step, the optimal curve \( \Sigma_{opt} \) (i.e. the optimal \( p_k \) parameters) which better interpolates the temporal series of radar LAI estimates (LAI\(_{ASAR,i}\)) has been determined by minimizing the following functional with a generalized reduced gradient method:

\[
J(p_1, p_2, p_3) = \frac{1}{2} \sum_{k=1}^{N} \left( \frac{p_k - p_{k0}}{\Delta^2 k0} \right)^2 + \frac{1}{2} \sum_{i=1}^{N} \left( \frac{LAI_{ASAR,i} - \Sigma_i}{\Delta^2 LAI-ASAR} \right)^2
\]

where \( N \) is the total number of radar estimates until heading, \( p_k \) are the parameters characterising the \( \Sigma \) function, \( p_{k0} \) represents the a priori information of these parameters, \( \Delta_k0 \) is the standard deviation associated with \( p_{k0} \), \( \Delta_{LAI-ASAR} \) is the standard deviation on radar estimates, \( \Sigma_i \) is \( \Sigma(t_i) \). The functional \( J \) represents a weighted sum of the square error between radar measurements and CERES-Wheat predictions and the square error between retrieved and guessed parameters. The weights are inversely proportional to measurement errors \( (\Delta_{LAI-ASAR}) \) and to a priori information error \( (\Delta_{k0}) \).

The a priori information to the optimisation procedure expressed in terms of the guessed values \( p_{k0} \) have been obtained from a curve \( \Sigma(t) \) representing an average LAI behaviour for wheat crops of the Matera area. This average curve has been derived by varying the sowing date, the wilting point and the field capacity, i.e. the three input parameters which mainly affect the CERES-Wheat LAI predictions, in the ranges that characterise the Matera site. In correspondence of each change the model has been run and a discrete set of LAI curves that the model can predict in the Matera area has been generated. Each LAI curve has been fit with a sigmoid function, determining the corresponding \( p_k \) values. Then the \( p_k \) mean and standard deviation have been computed and used respectively as guessed values of \( \Sigma(t) \) parameters, \( p_{k0} \), and associated errors, \( \Delta_{k0} \).

The second step consists of determining the CERES-Wheat configuration corresponding to the obtained \( \Sigma_{opt} \), which has to be adopted to re-initialise the model. More specifically the \( \Sigma_{opt} \) has been compared with the discrete set of LAI curves that the model can predict in the Matera area, computing the rms error. The LAI curve characterised by the lowest rms error with the \( \Sigma_{opt} \) and the corresponding CERES-Wheat configuration have been selected. The CERES-Wheat configuration found with this procedure is the one which better assimilates the information from radar data.

### IV. Results and Conclusion

The assimilation technique above described has been applied to the six wheat fields monitored in 2004, where in-situ yield analyses were carried out.

The pre-assimilation predictions, which have been used as reference to evaluate the effect of remote sensing information on the CERES-Wheat estimates, have been preliminary obtained. In case of absence of accurate information, concerning crop management and soil properties, some average values characterising the Matera area could be used. Driven by these values the crop model predicts an average wheat behaviour typical of the Matera area, which will be referred, for sake of reference, as the predictions for the “mean field”.

The experimental data collected in the six monitored wheat fields have been compared first with the model outputs obtained for the “mean field” and then with the model predictions after the assimilation of the LAI estimated by the ASAR data.

In order to explain the results it is worth mentioning that the wheat fields monitored in 2004 can be divided in two categories: three fields in which no fertilization and no herbicide treatments were applied and three fields in which these treatments were carried out. The fields of the first category were characterized by significant presence of weeds (the average percentage of weed weight with respect to the total above ground canopy weight is about 40%) and by an inhomogeneous canopy. These unfavourable conditions affected the growth of the wheat, so that the observed wheat biomass, LAI and yield were below the average values of the area. This implies that running the model for the “mean field”, in which the optimal conditions are taken into account, the predictions result considerably overestimated in comparison with the values measured in these fields, as will be shown in the following plots.

The assimilation of the radar data into the CERES-Wheat model has shown two different effects for fields with weeds and for fields without weeds.

The plots of Figure 1 compare the wheat leaf and plant dry biomass measured in one field with a significant presence of weeds (Field A) and in one field with a low percentage of weeds (Field B) with the temporal behaviour of the same parameters simulated by the model before, i.e. for the “mean field”, and after the assimilation. In Field A the estimates remain considerably overestimated after the assimilation, while in Field B the fit between the estimated curve and the observed values is significantly improved. Even though the assimilation reduces the error on the total LAI estimates in both fields, at least for Field A this is not enough to improve the predictions of wheat leaf biomass. The same comments can be extended to wheat above ground dry biomass.

The poor effect of the assimilation technique on the CERES-Wheat predictions when weeds are present in wheat fields may be due to limitations in the performance of the model and in the retrieval of wheat LAI by radar data. Due to the growth of weeds the density of the canopy is higher than the predicted one and this affects the regular growth of the wheat plants. The weed presence, indeed, reduces the solar
radiation intercepted by wheat and also the soil water resources. The model is not able to predict the weed presence and neglects the consequent effects, so that the wheat growth and the yield are overestimated.

Moreover the backscattering coefficient measured over these fields is affected by the weed presence, without the possibility to discriminate the wheat and weed contributions. This effect was taken into account by [2], in which data of both fields with and without weeds were used to obtain the empirical relationship between $(\sigma_{0}^{HH/VV})_{\text{DoY}}$ and LAI. Consequently the LAI retrieved by radar data includes the component due to the weeds and the assimilation of this information into the crop model lead to overestimated predictions.

The bar-plot in Figure 2 reports the simulation errors on the grain yield before and after the re-initialisation of the model in the three fields with low weed presence. Before the assimilation, the observed values have been compared with the grain yield predicted for the “mean field”, showing a model overestimation. After the re-initialisation of CERES-Wheat the accuracy on yield predictions has been significantly improved.

In conclusion, this study has described and validated a methodology aimed at assimilating polarimetric radar data into CERES-Wheat model. The obtained results indicate that, when the model is driven by inaccurate information about sowing date and soil properties, the assimilation of ASAR AP data acquired from tillering to heading stage leads to significant improvements in CERES-Wheat predictions of dry biomass and grain yield. However no significant effects have been observed when the same technique has been applied over fields characterised by large presence of weeds.

The encouraging results obtained in the retrieval of LAI from ASAR AP data and in the assimilation of this information into the CERES-Wheat crop model indicate the possibility to retrieve wheat LAI maps and then to early estimate wheat grain yield maps of the area.

ACKNOWLEDGMENT

This work has been partly supported by ASI under contract n. ASI/3/R/199/02 and partly by ESA-ESTEC under contract n. 17011/03/NL/JA.

REFERENCES