

Assimilation of hyperspectral data into crop growth models

Precision farming application for maize in Catalonia

Franz Kurz

Institute of Remote Sensing Technology
DLR – German Aerospace Center
Oberpfaffenhofen, Germany
Franz.Kurz@dlr.de

Lucas Martínez, Vicenç Palà, Roman Arbiol

Remote Sensing Department
ICC – Institut Cartogràfic de Catalunya
Barcelona, Spain
lmartinez@icc.es, vicencp@icc.es, arbiol@icc.es

Pere Villar, Ferran Calvet

LAF – Laboratori d'Anàlisi i Fertilitat dels Sòls
Sidamon, Spain
pvillar@lafsols.com, fcavet@lafsols.com

Abstract—We propose a general framework for the derivation of crop information from hyperspectral airborne CASI data using physical radiative-transfer models with an assimilation link to existing crop growth models. This framework is designed for precision farming and crop monitoring applications, which take into account the agricultural and climatic conditions in Catalonia. The derivation of crop parameters is based on physical radiative-transfer models (*PROSPECT* and *ACRM*), which are numerically inverted using optimization algorithms (Simplex, Quasi-Newton, and Levenberg-Marquardt). Additionally, we propose a method for the inclusion of a priori information in order to solve the ill-posed problem. In a further processing step, existing crop growth models (ISBA and STICS) assimilate the derived crop interface parameters to benefit from the area wide and spatial high resolution information.

Several measurement campaigns were conducted to calibrate and validate the physical radiative-transfer and crop growth models. In them, CASI airborne hyperspectral imagery and simultaneously in-situ measurements of crop and soil parameters were acquired on six different dates during the crop growth period in 2004. Six test sites with two maize hybrids (Oropesa and Eleonora) were selected for the in-situ measurements. The validation shows good correspondence of the derived canopy structure parameters with the in-situ measured values, whereas the derived leaf chlorophyll and nitrogen contents is not validated. Further investigations will show the potential of different assimilation techniques to improve the output of the crop growth models.

Keywords—Inverse problem; Physical model; Precision farming; Hyperspectral data; Data assimilation; Crop growth models.

I. INTRODUCTION

The agriculture in Catalonia (Spain) suffers from the excessive use of fertilization and the waste of irrigation water.

Over-application of pig manure and flood irrigations can cause severe environmental damage. Thus, development of improved cultivation techniques, stricter controls, and gathering of more detailed information about the current crop state are necessary to solve these environmental problems. Recent developments in the agriculture tend to dose the amount of fertilizer according to the needs of the plants and the capacity of the soil. This has been made possible through a better understanding of complex mechanisms in the soil and their interactions with the plants. Besides, the remote sensing techniques have reached a development state which allows the derivation of important agricultural information with a reasonable temporal and spatial resolution in order to allow the control of the cultivation practice and to support the agriculturist decisions.

The goal is the development of a system based on hyperspectral remote sensing data which provides agricultural relevant information about the crop fields at a reasonable level of costs. The information will be offered to farmers or governmental institutions interested in the monitoring of the cultivation to improve the environmental conditions.

The paper is organized as follows: Chapter II describes the proposed method to derive agricultural relevant parameters from hyperspectral CASI data based on physical models. In Chapter III, the concept of data assimilation is explained in more detail. Chapter IV describes the measurement and flight campaigns, which took place at maize test sites in Catalonia. In chapter V, the results of calibration and validation of the proposed method and the crop growth models are described and discussed.

II. DERIVATION OF CROP PARAMETERS FROM HYPERSPECTRAL DATA

A. Radiative transfer models

The radiative transfer plays an important role in modern remote sensing techniques, as optical remote sensing sensors measure the earth surface reflected solar irradiation. For this project, we choose the *ACRM* canopy spectral model (A two-layer Canopy Reflectance Model) [5] for the calculation of the radiative transfer within vegetation canopies. It is based on the homogeneous multispectral model *MSRM* [6] and the *SAIL* model, which models the diffuse radiative transfer. The *ACRM* model accounts for non-lambertian soil reflectance, specular reflection of direct sun rays on leaves, the hot spot effect, a two-parameter leaf angle distribution and the spectral effect of cultivation rows. For the spectral properties of leaves, we choose the *PROSPECT* model [4].

B. Inverse problem

The derivation of crop parameters from hyperspectral remote sensing data can be regarded as an optimization problem, i.e. the goal is to search the set of crop and soil parameters, which are “best” suited to the observations. One possibility to solve this problem is to search the most probable set of parameters, which is the basis of the maximum likelihood methods. In our case, the goal is to find a set of unknown crop and soil parameters, for which the a posteriori probability of the parameters $\hat{\mathbf{V}}$ with measured reflectances \mathbf{r}^* is maximal.

A basic problem is that the number of parameters describing the scene exceeds the number of the independent observations available. The reasons for this are the complexity of the vegetated earth surface and the influences on the measured reflectances. Therefore, it is impossible to find unique estimators only by relying on the measured reflectances, because the inverse function is ill-posed. A first step to solve this problem is to reduce the number of variable parameters, i.e. the input parameters of the physical models are divided in variable \mathbf{V} and constant \mathbf{c} .

C. Introduction of a priori information

The introduction of *a priori* information is another possibility to solve this ill-posed problem. In general, there are several possibilities to introduce *a priori* information into the inversion process due to the variety of available “knowledge”:

- In-situ measurements of parameters at defined sites
- Knowledge about parameters based on assumptions
- Knowledge about the parameters’ range based on assumptions or measurements

To introduce this knowledge into the derivation process, we use *a priori* values \mathbf{V}^* and probabilities $\mathbf{P}_{\mathbf{V}\mathbf{V}}$ and apply penalties \mathbf{p} to maintain the parameters within the definition range. The final cost function C , which has to be minimized, is described in equation (1). It contains the difference between observed \mathbf{r}^* and physically modeled reflectances $\mathbf{r} = f_{ACRM}(\mathbf{V}, \mathbf{c})$,

the stochastic model for the observations $\mathbf{P}_{\mathbf{b}\mathbf{b}}$, the a priori information, and the penalties \mathbf{p} (see [2]).

$$C = (\mathbf{r}^* - \mathbf{r})\mathbf{P}_{\mathbf{b}\mathbf{b}}(\mathbf{r}^* - \mathbf{r}) + \dots + (\mathbf{V} - \mathbf{V}^*)\mathbf{P}_{\mathbf{V}\mathbf{V}}(\mathbf{V} - \mathbf{V}^*) + \mathbf{p} \quad (1)$$

The stochastic model for the *a priori* information $\mathbf{P}_{\mathbf{V}\mathbf{V}}$ contains no covariances, and the variances are defined as *a priori* weights according to the definition range and the natural variability of each parameter. For example, the variance of the leaf area index (LAI) is defined as high as the whole definition range, because this parameter show high variability, whereas the specific leaf weight is supposed to remain constant within the fields and thus, the variance is defined as a fraction of the definition range. The *a priori* values themselves were set to in-situ measured values according to the specific hybrid or are set to standard values in case of parameters with high variability.

D. Model fitting

Due to the complexity of the physical models, an analytical minimization of C is not possible. Thus, the model inversion is conducted numerically using common mathematical methods. The following numerical optimization methods are applied: Simplex algorithm (Simulated annealing) SA, Quasi-Newton QN, and Levenberg-Marquardt LM.

III. ASSIMILATION OF DERIVED PARAMETERS INTO CROP GROWTH MODELS

For precision farming purposes, crop growth models are necessary, as the information derived only from remote sensing data is not sufficient for the agronomists. Remote sensing techniques provide information about the plant’s present state, but not about the growth and development processes. Therefore, the assimilation of remote sensing data into crop growth models could improve the prediction of later crop development. Besides, the use of crop growth models over large areas is limited by our ability to provide them with required input parameters. It is impossible to obtain area wide plant and soil characteristics directly from in-situ observations. Thus, remote sensing data are acquired to provide information about crop processes over large areas.

We select two crop growth models for the data assimilation, the *STICS* model [1] and the *ISBA* model [6] supported by simple SVAT (Soil Vegetation Atmosphere Transfer) models. Figure 1 gives an overview about the proposed assimilation of hyperspectral remote sensing data into the applied crop growth models. Crop and soil parameters are derived from the *CASI* reflectances using radiative transfer models like the *ACRM* model. Part of the derived parameters act as interface parameters, which are used to correct the continuous crop growth simulation of the *STICS* and *ISBA* models at the image acquisition dates. For this, various techniques can be applied: forcing (continuous), sequential assimilation (non continuous), and variational assimilation (adjustment). We propose a sequential assimilation of the derived parameters to improve

the prognosis of water and fertilization use as well as crop yield.

First, we calibrate the *STICS* and *ISBA* models using in-situ measurements (not in this paper) to fit the hybrid and cultivation practices, then we validate the results of these models with non stressed crops (v. chapter V). Further investigations will be conducted to test the proposed data assimilation concept.

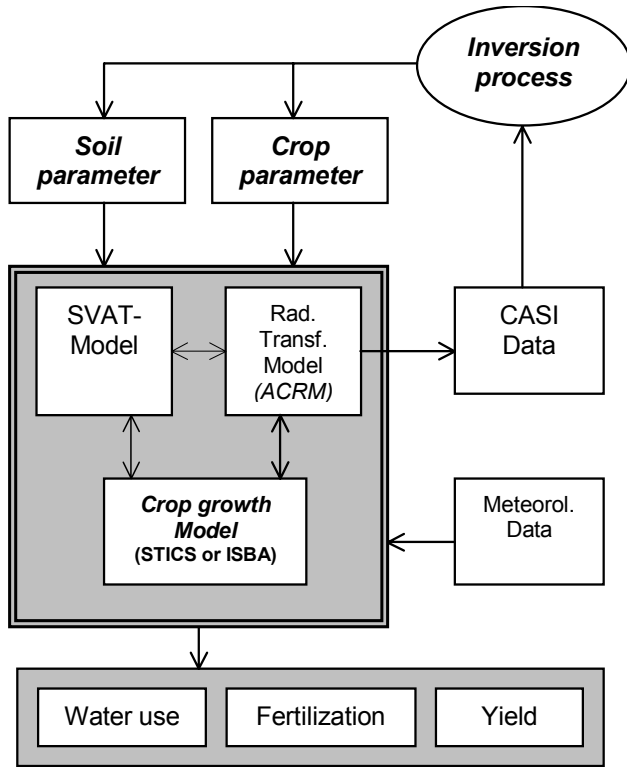


Figure 1. Illustration of the data assimilation process

IV. DATABASE

From May to August 2004, several measurement and flight campaigns were performed at six test sites near *Lleida*, Catalunya ($0^{\circ}57' E$, $41^{\circ}40' N$). The test sites were planted with two different maize hybrids: *Maize Oropesa* and *Maize Eleonora*. In order to get different nitrogen amounts, selected parts of the test sites were left without nitrogen fertilization, whereas in the rest of the field nitrogen fertilization was applied. Simultaneously to the image acquisition dates, several cultivation, crop, and soil parameters were measured. Additionally, half-hour sampled meteorological parameters were obtained from the automatic weather stations.

A. Hyperspectral imagery

The *CASI* sensor was configured in the spatial mode with eleven spectral channels and a ground pixel size of about 10 meters. After the data acquisition, the hyperspectral images were calibrated and geometrically corrected. Then an atmospheric correction was done. For the physical atmosphere

correction radiative transfer models based on the 6S code [8] were applied.

B. In-situ measurements

Simultaneously to the *CASI* image acquisition, crop and soil parameters were measured in-situ at all test sites taking into account the regions with different fertilization amounts. Thus, for each site and date, two plant and soil samples were taken and analyzed. Structural and chemical plant parameters were measured, the leaf area index, the leaf water content, nitrogen and chlorophyll content (SPAD readings), the specific leaf weight, soil water and nitrate content, etc. Besides, cultivation data (fertilization and irrigation date and amount, ploughing details, seeding date and density), constant soil parameters (soil structure, organic material, and soil chemical compounds), and selected meteorological data were recorded.

V. RESULTS

A. Model and algorithm calibration

The goal of the calibration process is to find optimal sets of configuration parameters for all applied models and inversion algorithms. Optimal configuration parameters are determined by comparing the *RMSE* of the derived parameters, the convergence, the amount of *a priori* information used, and the processing time. Table I compares the convergence and the processing time of model inversions for different configurations using synthetically generated *CASI* data. Best results are obtained by applying the LM algorithm with up to six variable parameters, which are selected according to the spectral sensitivity (sensitivity analysis not shown). In general, the introduction of *a priori* information improves the convergence and accuracy of the derived parameters, even in cases where the *a priori* information is slightly erroneous.

TABLE I. PERFORMANCE OF MODEL INVERSION

Experiment			Inversion algorithm					
Noise	A priori info.	Parameters	<i>SA</i>		<i>LM</i>		<i>QN</i>	
No	No	5	99%	6.5s	99%	0.5s	96%	1.5s
No	No	6	94%	10.0s	98%	1.0s	91%	1.5s
No	Yes	6	99%	10.0s	99%	1.0s	99%	1.5s
Yes	No	5	65%	10.5s	75%	1.0s	72%	1.5s
Yes	Yes	5	99%	10.5s	93%	1.0s	99%	1.5s
Yes	No	6	45%	20.0s	53%	1.6s	59%	2.5s
Yes	Yes	6	97%	20.0s	<u>99%</u>	<u>1.6s</u>	99%	2.5s
Yes	Yes	13	-	-	73%	3.0s	78%	6.6s

Next, in-situ measurements are used to set the *a priori* values and weights for the selected variable parameters. Model inversions with measured *CASI* reflectances are conducted to assess the *a priori* configuration with minimal *RMSE*. In table II the final *a priori* values and weights are listed. For the

parameters with high variability, low *a priori* weights are proposed (see also [2]). Standard *a priori* values for these parameters are determined, which partly have to be adjusted to the growth stage, e.g. for early development stages the *a priori* value of the leaf area index is 1.0.

In general, the above model and algorithm parameters are determined for the particular *CASI* configuration and the two maize hybrids Eleonora and Oropesa. For other configurations and hybrids, the calibration should be repeated.

TABLE II. CONFIGURATION OF THE PHYSICAL MODELS

	<i>Parameter</i>	<i>A priori value</i>	<i>A priori weight</i>
Variable parameter	Leaf area index	1.0, 3.0 ^a	0.25
	Leaf chlorophyll a content	0.55 g/m ²	0.25
	Leaf water content	15.0 mg/cm ²	0.25
	Leaf structure	0.04, 0.01 ^a	0.25
	Soil reflectance coefficients	0.01	0.25
	Modal leaf angle	60	1.0
	Eccentricity	2.5	1.0

a. depending on the actual growth stage of the crop

B. Model validation

The model validation is conducted by comparing the in-situ measurements (meas), the derived parameters (*CASI*) and the simulations (*STICS*, *ISBA*) of the parameters *LAI*, dry matter (*DM*), leaf nitrogen (*N*), and leaf chlorophyll (*Chl*). Table III lists the linear regression coefficients. The comparisons of crop structure parameters *LAI* and dry matter show high correlation, whereas the derived chlorophyll values are not validated. The reasons for the bad correlation probably is bad georeferencing of the in-situ measurements, different scales (10m pixel vs. single leaves), and low accuracy in atmospheric correction of the *CASI* reflectances.

TABLE III. VALIDATION SUMMARY

	<i>LAI_{STICS}</i>	<i>LAI_{ISBA}</i>	<i>LAI_{CASI}</i>	<i>DM_{STICS}</i>	<i>Chl_{CASI}</i>
<i>LAI_{meas}</i>	0.97 ^b	0.99 ^b	0.98		
<i>LAI_{CASI}</i>	0.97 ^b	0.87 ^b	-		
<i>Dm_{meas}</i>				0.95	
<i>N_{meas}</i>					× ^a
<i>Chl_{meas}</i>					× ^a

a. ×: No significant correlation

b. Only test sites without nitrogen- or water stress

VI. SUMMARY

A method to derive crop parameters from hyperspectral remote sensing data was designed, calibrated and validated with real measured data. For this, in-situ measurements and *CASI* imagery of test sites with two maize hybrids were acquired during six dates in 2004. The proposed derivation method showed good results in particular for crop structure parameters, *LAI* and the dry matter. Moreover, the inclusion of *a priori* information increased the convergence of the inversion process and accuracies of derived parameters.

A concept for the assimilation of remote sensing data into crop growth models was proposed. For this, two crop growth models were calibrated and validated using the in-situ measurements. Further investigations will be the sequential assimilation of the derived crop parameters to improve the model prediction of the crop growth models.

ACKNOWLEDGMENT

The authors would like to thank Dr. Andres Kuusk affiliated with the Tartu Observatory in Tõravere Estonia and Dr. Jaan Praks affiliated with the University of Technology, Laboratory of Space Technology in Helsinki, for providing the code of the *ACRM* model including a module which deals with row structured vegetation surfaces, as well as Florence Habets affiliated with the Meteo-France Toulouse in France for providing the code of the *ISBA* crop growth model.

REFERENCES

- [1] N. Brisson et al. (1998) "STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn" *Agronomie* 18 311-346.
- [2] B. Combal et al. (2002) "Retrieval of canopy biophysical variables from bi-directional reflectance using prior information to solve the ill-posed inverse problem" *Remote Sensing of Environment* 84: 1-15
- [3] Habets, F. and J. Noilhan, 1997: Resultats des simulations d'ISBA dans la phase PILPS2c: bilan hydrique du bassin de l'Arkansas, Note Trav. 50, GMMÉ/Météo-France, 59 pp
- [4] S. Jacquemoud, S.L. Ustin, J. Verdebout, G. Schmuck, G. Andreoli, and B. Hosgood (1996) "Estimating leaf biochemistry using the PROSPECT leaf optical properties model" *Remote Sensing of Environment* 56: 194-202.
- [5] Kuusk A. (2001) "A two-layer canopy reflectance model" *Journal of Quantitative Spectroscopy & Radiative Transfer* 71 1-9.
- [6] Kuusk A. (1994) "A multispectral canopy reflectance model" *Remote Sensing of Environment* 1994 50:75-82.
- [7] Noilhan J. and Mahfouf J.-F. (1996) "The ISBA land surface parameterization scheme" *Global and Plan. Change* 13 145-159.
- [8] E. Vermonte, D. Tanré, J.L. Deuzé, M. Herman, and J.J. Morcrette (1997) "Second simulation of the satellite signal in the solar spectrum, 6S: an overview" *IEEE Transactions on Geoscience and Remote Sensing*, 35: 675-686.