STUDY ON ENVISAT ASAR DATA ASSIMILATION IN RICE GROWTH MODEL FOR YIELD ESTIMATION

S. H. Shen [a], S. B. Yang [a], B. B. Li [b], X. Y. Zhao[a], B. X. Tan [c], Z. Y. Li [c], T. Le Toan [d]

[a] College of Applied Meteorology, Nanjing University of Information Science and Technology, Nanjing 210044 - China – jaasyang@163.com
[b] Institute of Agricultural Resources and Environment, Jiangsu Academy of Agricultural Sciences, Nanjing 210014 - China – bbli88@163.com
[c] Institute of Forest Resources Information Technique, Chinese Academy of Forestry, Beijing 100091, China - tan@caf.ac.cn
[d] CESBIO, 18 Avenue Edouard Belin, 31401 Toulouse Cedex 9, France - Thuy.Letoan@cesbio.cnes.fr

Commission VIII, WG VIII/10

KEY WORDS: Rice Yield Map, Crop Model, Radar Model, Data Assimilation, Optimization Algorithm, Classification

ABSTRACT:

In this paper, a practical scheme for assimilation of multi-temporal and multi-polarization ENVISAT ASAR data in rice crop model to map rice yield has been presented. To achieve this, rice distribution information should be obtained first by rice mapping method to retrieve rice fields from ASAR images, and then an assimilation method is applied to use the temporal single-polarized rice backscattering coefficients which are grouped for each rice pixel to re-initialize ORYZA2000. The assimilation method consists in re-initializing the model with optimal input parameters allowing a better temporal agreement between the rice backscattering coefficients retrieved from ASAR data and the rice backscattering coefficients simulated by a coupled model, i.e. the combination of ORYZA2000 and a semi-empirical rice backscatter model through LAI. The SCE-UA optimization algorithm is employed to determine the optimal set of input parameters. After the re-initialization, rice yield for each rice pixel is calculated, and the yield map over the area of interest is produced finally. The scheme was applied over Xinghua study area located in the middle of Jiangsu Province of China by using the data set of an experimental campaign carried out during the 2006 rice season. The result shows that the obtained rice yield map generally overestimates the actual rice production situation, with an accuracy of 1133 kg/ha on validation sites, but the tendency of rice growth status and spatial variation of the rice yield are well predicted and highly consistent with the actual production variation.

1. INTRODUCTION

Since the launch of ERS satellites and Radarsat, a considerable amount of programs has been set up to investigate the capability and efficiency of radar data in agricultural monitoring. For rice crop, SAR data has been successfully applied for rice mapping and growth parameters inversion (Le Toan et al., 1997; Ribbes & Le Toan, 1999; Shao et al., 2001; Dong et al., 2006; Tan et al., 2006; Yang et al., 2008a; Yang et al., 2008b). Physical models have also been developed to study the interaction between backscatter and rice canopy (Le Toan et al., 1997; Dong et al., 2006; Koay et al., 2007). The results prove that the rice crop is in favor of radar monitoring. However, among the aforementioned research activities, the potential of SAR data in rice yield prediction has not been fully investigated.

In this paper, a practical scheme for rice yield estimation and mapping rice yield based on ASAR (Advanced Synthetic Aperture Radar) data is presented, in which the assimilation strategy (Moulin et al., 1998) is adopted, and validated using the data acquired in 2006 over the Xinghua study area located in the middle of Jiangsu Province of China.

2. MATERIALS AND METHODS

2.1 Scheme for mapping rice yield based on ASAR data

The scheme for rice yield estimation is based on multi-temporal and multi-polarization ASAR data (Figure 1). It consists of two parts for realizing the whole process. In the first part, ASAR data is employed with rice mapping method to obtain the rice distribution map over the area of interest. The rice map is used to mask all the ASAR images to select only rice fields and retrieve rice backscattering coefficients. The temporal rice backscattering coefficients corresponding to each rice pixel of the rice map are then grouped for each polarization respectively. It should be noted that the accuracy of rice yield estimation is somehow influenced by the rice mapping accuracy. The key to rice mapping is to find proper data acquisition dates and/or data of proper polarizations to maximize the distinction between rice and other land cover objects (Le Toan et al., 1997; Ribbes & Le Toan 1999). Therefore, multi-temporal and multi-polarization radar data are recommended for rice mapping, because higher rice mapping accuracy of more than 80% has been reported by several studies with the threshold or supervised classification method (Chen et al., 2007; Yang et al., 2008a; Yang et al., 2008b).

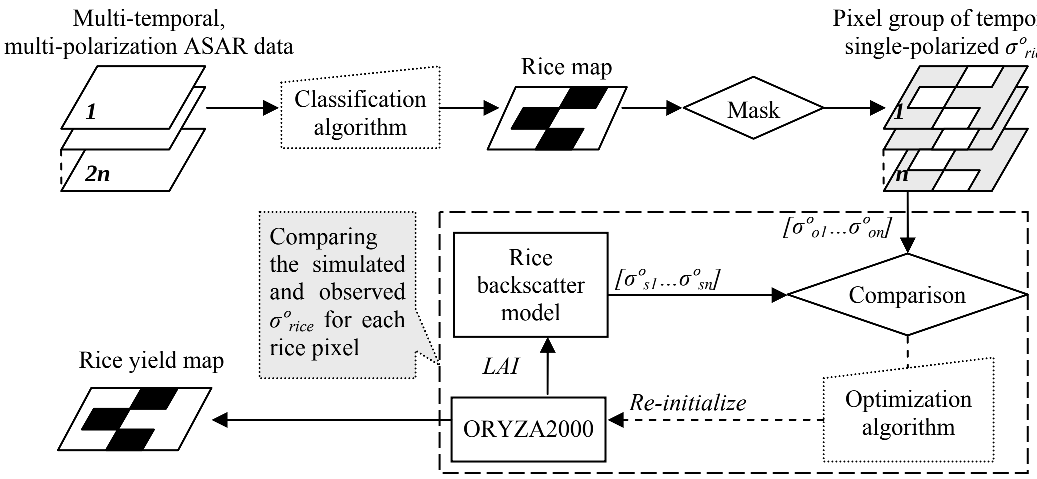

Figure 1. Scheme for mapping rice yield based on ASAR data

Acquisition	Day of year (DoY)	Polarization	Orbit	Spatial resolution	Incidence angle	Rice phenological stage
06/30/2006	181	HH/VV	Descending	30 m	19.2°-26.7°	Tillering stage
08/04/2006	216	HH/VV	Descending	30 m	19.2°-26.7°	Jointing stage
08/19/2006	231	HH/VV	Ascending	30 m	19.2°-26.7°	Booting stage
09/23/2006	266	HH/VV	Ascending	30 m	19.2°-26.7°	Grain filling stage

Table 1. An overview of the acquired ASAR data and the phenological stage of rice at each of the acquisition date

The second part in this scheme is mainly shown in the dashed box in Figure 1, where an assimilation method is adopted to calculate the rice yield for each rice pixel. The assimilation method is the direct use of observed rice backscattering coefficients to re-initialize the rice growth model ORYZA2000 (Bouman et al., 2001). During the assimilation, ORYZA2000 is coupled with a rice backscatter model using LAI as an essential link to simulate rice backscattering coefficients. Here, a semi-empirical rice backscatter model, based on the Cloud model (Attema & Ulaby, 1978), is used for simplicity and practicability. But, before it can be used, the model should be calibrated to determine the polarization, in which the temporal rice backscatter can be simulated with a higher accuracy. The assimilation method re-initializes the ORYZA2000 model with optimal input parameters allowing a better temporal agreement between the rice backscattering coefficients simulated and those observed. The global optimization algorithm SCE-UA (Duan et al., 1992; Duan et al., 1993) is applied to determine the optimal set of input parameters. After the re-initialization, rice yield corresponding to each rice pixel is calculated by ORYZA2000, and finally, the rice yield map of the area of interest is produced.

2.2 Study area and data description

The Study area is a flat agricultural area located in the middle of Jiangsu Province of China, approximately 32º51′N-32º58′N and 120º00′E-120º06′E. The crop system here is a two-crop rotation system, with wheat in winter and rice in summer. During the rice season, more than 95% of rice is direct-seeded. The dominant rice species is japonica rice with a life span of about 135 days.

In 2006, four rice growth monitoring plots A, B, C and D with size about 10 ha each were established, and ground truth data were collected within the plots at 10-day intervals from June 25 to October 20. Within each plot, 5 fixed observation sites with area of 1 m^2 for each were monitored for rice growth stages, plant density, plant height and general information about field management. The aboveground biomass was measured separately for stems, green and dead leaves, and ears by collecting 25 samples randomly within each plot. LAI was also measured. During the rice harvest, detailed information on actual rice production was collected. Meanwhile, boundaries of the monitoring plots and other land surface objects were recorded using DGPS devices. Daily meteorological data were obtained from local meteorological bureau.

During the rice season of 2006, four ASAR APMode products were acquired over the study area (Table 1). The calibration was carried out using the *Basic ENVISAT SAR Toolbox* software provided by European Space Agency (ESA) to extract backscattering coefficients. In order to reduce the speckle noise, two SAR image filters, the multichannel filter (Quegan & Yu, 2001) and Gamma Map filter were performed successively to the previously calibrated images. Then, images were geo-referenced using a topographic map with a root means square error of the control points about 22 m.

3. MODELS AND OPTIMIZATION METHOD

3.1 ORYZA2000

ORYZA2000 is a dynamic, eco-physiological rice crop model to simulate the growth, development and water balance of lowland rice in situations of potential production, water limitations and nitrogen limitations (Bouman et al., 2001). The model follows the daily calculation scheme for the rates of dry matter production of the plant organs and the rate of phenological development from emergence until harvest. By integrating these rates over time, dry matter production of rice is simulated throughout the growing season and final yield is calculated.

In this study, rice growth is assumed under the condition of potential production, and ORYZA2000 is calibrated by DRATES and PARAM to estimate the variety-specific parameters which consist of development rate, partitioning factors, relative leaf growth rate, specific leaf area, leaf death rate and fraction of stem reserves using the experimental data collected over the plots A and B. The experimental data of the plots C and D are kept for the validation.

After the calibration, the predicted final yield and the modeled temporal course of dry aboveground biomass were compared with the experimental data by computing the simulation error, i.e. the mean absolute difference between the estimated and the observed data normalized by the measured value and expressed in percentage. As a result, the final yield was predicted with a simulation error less than 18%, and the total dry aboveground biomass was predicted with a simulation error of 17.4%. The correlation coefficient of simulated and measured LAI reaches 96%.

However, some parameters can be largely variable over the study area and are rarely available through field measurements. Since the rice is direct-seeded, the difference between plant densities from fields to fields reaches more than 22% on average and 69% to the maximum showing that the plant density varies remarkably over the study area. Therefore, a sensitivity analysis was carried out to determine whether or not there exists a relatively small subset of model inputs affecting, more than others, the temporal behavior of the state variable of interest for the assimilation. The inputs taken into account concerned emergence date (EMD), plant density (NPLDS) and maximum relative growth rate of leaves (RGRLMX), other parameters related to the nitrogen fertilization and water irrigation are not included for the potential production is adopted in this study. The result shows that the predicted rice yield and LAI are mainly sensitive to the changes of emergence date and plant density which constitute the set of model inputs which were involved in the assimilation and re-initialization process.

3.2 Cloud

The Cloud model (Attema & Ulaby, 1978) is a semi-empirical radar model to simulate the volume scattering for vegetation. It assumes the vegetation consists of a collection of water droplets, which are represented as small identical particles.

The simple model has been applied successfully for a range of crop types and conditions (Champion et al., 2000; Inoue et al., 2002). For paddy rice, we simply assume the scattering from the paddy background (water surface) is constant before the ripening stage and the temporal variation of rice backscatter is mainly attributed to the change of canopy size (e.g., stem height, leaf size), canopy water content and canopy biomass.

The inputs W and h in the Cloud model (Attema & Ulaby, 1978) are rice parameters refer to the canopy water content and canopy size. In order to link the simple model with ORYZA2000 by the LAI, regression analysis is applied to investigate the relationship between $W \cdot h$ and LAI, in which the $W \cdot h$ represents the canopy water content per unit soil surface (kg/m^2). As a result, a significant relationship is found as the following rational equation:

$$LAI = (10.19 \cdot W \cdot h - 0.1534) / (W \cdot h + 1.511) \quad (1)$$

Finally, the Cloud model for the rice paddy can be expressed (in dB) as follows:

$$\sigma^o = 10 \cdot \log10[\alpha \cdot cos\theta \cdot (1 - k^2) + k^2 \cdot \sigma^o_{BG}] \quad (2)$$

$$k^2 = \exp\{2 \cdot \beta \cdot (1.511 \cdot LAI + 0.1534) / [(LAI - 10.19) \cdot cos\theta]\} \quad (3)$$

where σ^o_{BG} = constant backscattering from canopy background (m^2/m^2)

Here, three parameters α, β and σ^o_{BG} should be estimated by fitting the model to the observed rice backscatter data. In this paper, the global optimization method SCE-UA was applied to estimate the optimal values of the three parameters, and tested for HH and VV polarization separately. The brief description of SCE-UA method and optimization configurations can be found in the next part. The results in terms of the α (in dB), β, and σ^o_{BG} (in dB) with some statistics indicate that the Cloud model calibrated in HH polarization has a better performance to simulate the rice backscattering coefficients.

3.3 SCE-UA Method

SCE-UA (Shuffled Complex Evolution) method (Duan et al., 1993), developed at the University of Arizona in 1992 is not problem specific and is easy to handle, which has been widely used in various fields for nonlinear optimization problems and reported exact results (Duan et al., 1994; Yan et al., 2006).

The SCE-UA algorithm contains many parameters that control the probabilistic and deterministic components of the method. These parameters should be carefully selected for the optimal performance of the algorithm (Duan et al., 1993). Here, we only present the optimization configuration for the model input parameters of the rice backscatter model and ORYZA2000, with the system settings of the SCE-UA method for each of them (see Table 2 and 3). Least-squares function is used as the objective function. The optimization process is terminated if one of the following criteria is satisfied: (1) the algorithm is unable to improve 0.0001 percent of the value of the objective function over five iterations; (2) the algorithm is unable to change the parameter values and simultaneously improve the function value over five iterations; (3) the maximum number of iterations (10000) is exceeded.

Configuration	Rice backscatter model			ORYZA2000	
	α (m^2/m^2)	β (m^2/kg)	σ^o_{BG} (m^2/m^2)	EMD (DoY)	NPLDS (plants/m^2)
Initial value	0.143	8.2	0.083	166	230
Sample interval	(0, 1)	(0,10)	(0, 1)	[145, 175]	[100, 300]

Table 2. Optimization configuration for the model input parameters of the rice backscatter model and ORYZA2000

Models	Number of Complexes	Number of points in each complex	Number of points in a sub-complex	Number of evolution steps	Minimum number of complexes	Trials
Cloud model	4	7	4	7	3	10
ORYZA2000	3	5	3	5	2	10

Table 3. System settings of SCE-UA method for the rice backscatter model and ORYZA2000

4. RESULTS

To achieve the presented scheme for estimating rice production and mapping rice yield, rice map should be produced at first as pointed out in Section 2. Fortunately, the rice map of the study area has been retrieved in our previous study using threshold classification method (Yang et al., 2008a) with the rice mapping accuracy of 84.36%.

The semi-empirical rice backscatter model was calibrated and coupled with ORYZA2000 to simulate HH-polarized rice backscattering coefficients. During the process, the optimization stopped successfully in criteria (1), and recorded the optimal values of EMD and NPLDS for further analysis. Figure 2 shows the obtained distribution maps of EMD and NPLDS, and the map of rice yield, with the spatial resolution of 30 m. For each map, different colors are assigned to the pixels according to their values. The probability density plots of the retrieved maps were also displayed in Figure 3. It shows that the optimal parameters vary mainly in the following ranges: the EMD between DoY 160 and DoY 175, and the NPLDS between 150 and 300 plant/m^2. According to our field survey, the variation of the estimated rice emergence date and rice plant density is realistic for the study area, except that the average rice plant density is slightly underestimated. The rice yield in Figure 2(c) varies between 9000 and 11200 kg/ha. The mean estimated is about 10422 kg/ha, higher than the average observed about 1200 kg/ha.

The red color region in the EMD map shown in Figure 2(a) is conspicuous. The pixel value corresponding to the red area is between 171 and 175 with the average of 172.3. It indicates that the rice grown in this area have a late emergence date compared to the average rice emergence date observed (DoY 166). In addition, the red color region is well corresponding to the light blue area in the NPLDS map and the blue area in the yield map shown in Figure 2(b) and (c) respectively. The light blue area in the NPLDS map has the pixel value ranging from 100 to 150 with the average of 118.9, which shows the rice plant density of the area is rather low. As a result, the blue area in the yield map has a relatively low rice production ranges from 9000 to 9800 kg/ha.

For a quantitative evaluation of the reliability of the produced rice yield map, it was compared to the in-situ data collected over the 10 monitored fields. As shown in Figure 4, it was found that the estimated rice yield is generally higher than the observed with a root mean square error of approximately 1133 kg/ha (i.e. approximately 12.2%). The overestimation of the rice yield was due to the simulation under the condition of potential production used in this study. In fact, rice disease and pests are severe during the rice growth period of 2006, which causes the universal reduction of the rice yield. However, the tendency of rice growth status and final yield are well predicted according to the field survey during the rice harvest.

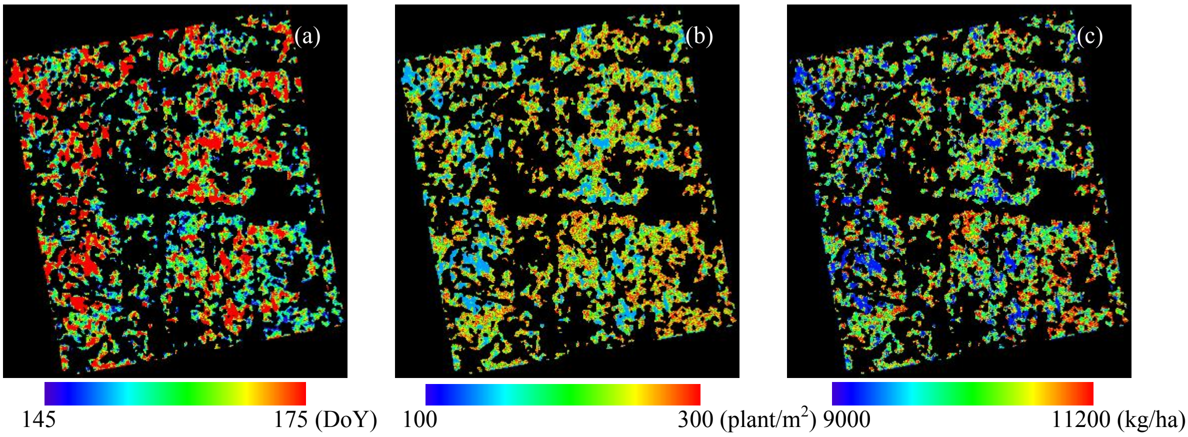

Figure 2. Distribution maps of rice emergence date (a), rice plant density (b) and final rice production (c)

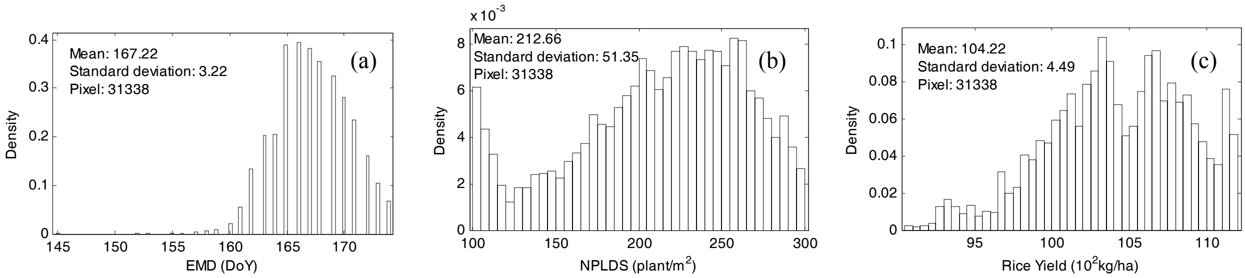

Figure 3. Probability density of pixel values for maps of rice emergence date (a), rice plant density (b) and final production (c)

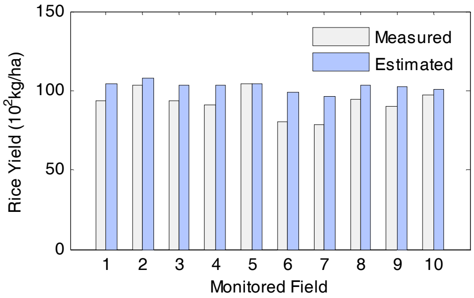

Figure 4. Rice yield estimated after the assimilation of ASAR data (blue) compared to the yield measured over 10 fields (grey)

5. DISCUSSION AND CONCLUSIONS

In this study, a practical scheme for mapping rice yield based on ASAR data has been presented and was applied over Xinghua study area of China and validated with experimental data collected in 2006. The ORYZA2000 model predicts a rice yield map with a spatial resolution of 30 m. The estimated rice yield is generally higher than the observed with an accuracy of approximately 1133 kg/ha. This is due to the potential rice growth condition were assumed in this study. But the tendency of rice growth status and final yield are well predicted and the spatial variation of the rice yield is highly consistent with the actual rice production situation.

In conclusion, the scheme described in this study is a promising technique to apply radar data for regional rice production estimation, when no accurate in-situ information is available and/or optical data are hampered by heavy clouds during the rice season. However, further validation of the presented scheme at different rice planting areas and with different radar configurations is needed. Therefore, future work will be dedicated to improve the models and assess the scheme over longer series of data and over different sites.

REFERENCES

Attema, E. P. W., Ulaby, F. T., 1978. Vegetation modeled as a water cloud. Radio Science, 13(2), pp. 357-364.

Bouman, B. A. M., van Kraalingen, D. W. G., Stol, W., and van Leeuwen, H. J. C., 1999. An agroecological modeling approach to explain ERS SAR radar backscatter of agricultural crops. Remote Sensing of Environment, 67(2), pp. 137-146.

Bouman, B. A. M., Kropff, M. J., Tuong, T. P., Wopereis, M. C. S., ten Berge, H. F. M., and van Larr, H. H., 2001. ORYZA2000: modeling lowland rice. International Rice Research Institute, Los Banos.

Champion, I., Prevot, L., and Guyot, G., 2000. Generalized semi-empirical modeling of wheat radar response. International Journal of Remote Sensing, 21, pp. 1945– 1951.

Chen, J. S., Hui, L., and Pei, Z. Y., 2007. Application of ENVISAT ASAR data in mapping rice crop growth in southern China. IEEE Transactions on Geoscience and Remote Sensing Letters, 4(3), pp. 431-435.

Dong, Y. F., Sun, G. Q., and Pang, Y., 2006. Monitoring of rice crop using ENVISAT ASAR data. Science in China: Series D Earth Sciences, 49(7), pp. 755-763.

Duan, Q. Y., Sorooshian, S., Gupta, V. K., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resource Research, 28(4), pp. 1015 - 1031.

Duan, Q. Y., Gupta, V. K., and Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. Journal of Optimization Theory and Its Applications, 76(3), pp. 501-521.

Duan, Q. Y., Sorooshian, S., Gupta, V. K., 1994. Optimal use of the SCE-UA global optimization method for calibrating Watershed Models. Journal of Hydrology, 158, pp. 265-284.

Fang, H., Wu, B., Liu, H., and Xuan, H., 1998. Using NOAA AVHRR and Landsat TM to estimate rice area year-by-year. International Journal of Remote Sensing, 3, pp. 521-525.

Inoue Y., Kurosu T. Maeno H., Uratsuka S., Kozu T., Dabrowska-Zielinska K., and Qi J., 2002. Season-long daily measurements of multifrequency (Ka, Ku, X, C, and L) and full-polarization backscatter signatures over paddy rice field and their relationship with biological variables. Remote Sensing of Environment, 81, pp. 194-204.

Koay, J. Y., Tan, C. P., Lim, K. S., bin Abu Baka, S. B., Ewe, H. T., Chuah, H. T., and Kong, J. A., 2007. Paddy fields as electrically dense media: theoretical modeling and measurement comparisons. IEEE Transactions on Geoscience and Remote Sensing, 45(9), pp. 2837-2849.

Le Toan, T., Ribbes, F., Wang, L. F., Floury, N., Ding, K. H., Kong, J. A., Fujita, M., and Kurosu, T., 1997. Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. IEEE Transactions on Geoscience and Remote Sensing, 35(1), pp. 41-56.

Moulin, S., Bondeau, A., and Delécolle, R., 1998. Combining agricultural crop models and satellite observations : From field to regional scales. International Journal of Remote Sensing, 19(6), pp. 1021-1036.

Quegan, S., Yu, J. J., 2001. Filtering of multichannel SAR images. IEEE Transactions of Geoscience and Remote Sensing, 39(11), pp. 2373-2379.

Ribbes, F., and Le Toan, T., 1999. Rice field mapping and monitoring with RADARSAT data. International Journal of Remote Sensing, 20(4), pp. 41-56.

Shao, Y., Fan, X., Liu, H., Xiao, J., Ross, S., Brisco, B., Brown, R., and Staples, G., 2001. Rice monitoring and production estimation using multitemporal RADARSAT. Remote Sensing of Environment, 76(3), pp. 310-325.

Tan, B. X., Li, Z. Y., Li, B. B., and Zhang, P. P., 2006. Rice field mapping and monitoring using single-temporal and dual polarized ENVISAT ASAR data. Transactions of the CSAE, 22(12), pp. 121-127.

Yan, Y., Liu, Q. H., Liu, Q., Li, J., Chen, L. F., 2006. Methodology of winter wheat yield prediction based on assimilation of remote sensing data with crop growth model. Journal of Remote Sensing, 10(5), pp. 804-811.

Yang, S. B., Shen, S. H., Li, B. B., Le Toan, T., and He, W., 2008a. Rice mapping and monitoring using ENVISAT ASAR data. IEEE Geoscience and Remote Sensing Letters, 5(1), pp. 108-112.

Yang, S. B., Li, B. B., Shen, S. H., Tan, B. X., and He, W., 2008b. Rice mapping research based on multi-temporal, multi-polarization backscattering differences. Journal of Remote Sensing, 13(3), pp. 138-144.

ACKNOWLEDGEMENTS

This work has been partly supported by the ESA-NRSCC Dragon Cooperation Program (http://earth.esa.int/dragon/) and partly by the Project of Jiangsu Graduate for Science Research and Innovation (No. CX07B_048z). The authors are grateful to ESA for supplying the ASAR images, and would like to thank Prof. B.A.M. Bouman of the International Rice Research Institute (IRRI) for permitting us to use the ORYZ2000 model, and thank Dr. Q.Y. Duan of the Lawrence Livermore National Laboratory, for providing the program of the SCE-UA algorithm.