Estimation of rice yield by SIMRIW-RS, a model that integrates remote sensing data into a crop growth model

Masayasu MAKI a,†, Kosuke SEKIGUCHI a, Koki HOMMA b, Yoshihiro HIROOKA b and Kazuo OKI c

a Graduate School of Engineering, Kyoto University, Kyoto, 615-8540, Japan
b Graduate School of Agriculture, Kyoto University, Kitashirakawa Oiwake-cho, Sakyo-ku, Kyoto, 606-8502, Japan
c Institute of Industrial Science, the University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan

Abstract

Food security has become a serious concern recently in Southeast Asia. The reduction of agricultural land because of economic development is decreasing the food supply. Simultaneously, due to rapid population growth, the food demand is increasing. Therefore, to ensure a stable food supply, it is important to estimate the supply capability of rice, which is the staple food in most Asian countries. In this study, a crop model (SIMRIW-RS) that can combine remote sensing data with a crop model (SIMRIW) was used to estimate rice yield at a regional scale. This model was applied to the estimation of rice yield in paddy fields located in the suburbs of Vientiane, Laos. Satellite (COSMO-SkyMed)-derived data for leaf area index (LAI) were integrated into SIMRIW-RS, and the transplanting date detected by COSMO-SkyMed was used to set the starting date of the simulation. Results were verified by surveying farmers. Transplanting dates were detected with high accuracy in all but a few fields. On the basis of the results of regression analysis between actual LAIs and the corresponding backscatter coefficients of COSMO-SkyMed, we suggest that COSMO-SkyMed can estimate LAIs at early growth stages when LAI is small. The results of yield estimation after integrating the LAIs derived from COSMO-SkyMed data into SIMRIW-RS indicated that the estimation accuracy of the rice yield was improved compared with the estimation result without adjusting parameters in the model, and this held so long as LAI was retrieved with high accuracy by satellite data. However, when LAI could not be estimated accurately, integration has the potential to worsen the model’s accuracy compared with the estimation result without any such readjustment. This study therefore indicates that SIMRIW-RS has the potential to estimate rice yield accurately when the LAI of rice is estimated with high accuracy from satellite data.

Key words: Assimilation, Crop growth model, Remote sensing, Yield estimation.

1. Introduction

Recently, food security has become a serious concern for countries in Southeast Asia. The reduction of agricultural land as a result of recent economic development has decreased food supply, and due to rapid population growth, food demand is increasing. It is therefore crucial to estimate the supply capability of rice, which is the staple food in most Asian countries, at the regional scale to ensure a stable food supply. Crop models have been used worldwide for yield estimation, and these models are broadly divided into two types: those for crop growth and yield estimation at a global scale and those for cultivation management at the field scale. For estimation by global-scale crop models, such as the Simulated Model for Rice–Weather relations (SIMRIW) (Horie et al., 1995), only cultivar information and meteorological data are required to simulate yield. However, this type of model does not simulate water and nitrogen stresses; therefore, the results from this type of simulation model indicate the potential yield under ideal field conditions. In contrast, models for cultivation management at the field scale, as typified by ORYZA2000 (Bouman et al., 2001), can simulate yield while considering field management effects such as water management and fertilizer application, and such models therefore give more realistic yield predictions. Because this type of model requires several sets of information in addition to cultivar and meteorological data, it is difficult to apply this type of model to estimation of crop growth at a regional scale.

A separate technique for crop growth estimation at a regional scale, used worldwide, relies on satellite remote sensing. However, satellite remote sensing data contains the growing condition of the crop at the observation time only, so it is difficult to directly estimate crop yield from satellite data unless the crop condition during the harvesting period has been already observed. In recent years, integration of remote sensing data into crop models has been investigated as a way to estimate crop yield at regional and global scales, and this method has yielded successful results (Curnel et al., 2011; Zhao et al., 2013). Our research group has developed one such crop model, called SIMRIW-RS (Homma et al., submitted), that can simulate development stage, growth condition, and yield while taking water and nitrogen stresses into consideration. The model is designed on the assumption that remote sensing data are integrated into the model. Therefore, the number of readjusted parameters for optimal yield estimation with using SIMRIW-RS is lower than with using other crop models that can simulate crop yield while considering water and nitrogen stresses. For regional estimation of rice yield by using SIMRIW-RS, the location and transplanting date of each paddy field must
be considered in the model, in addition to the cultivar and field parameters. Previous studies have successfully retrieved this information from satellite remote sensing data. In particular, optical remote sensing data have allowed some successful results (Xiao et al., 2002; Xiao et al., 2005). However, it may be difficult to apply such approaches in the monsoon regions of Southeast Asian during the wet season because the ground surface is blocked from satellites by cloud cover. However, synthetic aperture radar (SAR) can monitor the ground surface under any weather condition, and SAR data can thereby extract information on which paddy fields have been placed with rice (Shao et al., 2001; Zhang et al., 2009). However, it had been difficult to detect accurate transplanting dates from SAR data because the data with high spatial and temporal resolution have not been available until the past few years. In recent years, high temporal- and spatial-resolution SAR sensors have been launched, allowing access to these SAR data.

In this study, to achieve an accurate estimate of rice yield at the regional scale in monsoon regions, the use of high temporal- and spatial-resolution SAR data for detecting the transplanting date in rice paddy fields and retrieving the crop physical factors used to adjusting the field parameters of the model was investigated. The effect of integrating these remote sensing data into the SIMRIW-RS crop model was then examined.

2. Materials and methods

This study was carried out during the wet season of 2013 at paddy fields located in a suburban area near Vientiane, Laos. This area is within the Asian monsoon region and it is therefore difficult to monitor the ground surface with optical satellites during the wet season due to cloud cover. X-band SAR images from the Constellation of Small Satellites for the Mediterranean Basin Observation (COSMO-SkyMed system) were used in this study to acquire high temporal- and spatial-resolution SAR data. COSMO-SkyMed can supply high temporal- and spatial-resolution data by operating 4 satellites on the same orbit. All COSMO-SkyMed images used in this study were acquired during ascending orbit. The incidence angle for data acquisition was approximately 30°. Spatial resolution was adjusted to 30 m by multi-look processing and spatial filtering to reduce speckle noise. The 3 × 3 pixel Lee filter (Lee, 1980) was applied to the images used in this study, and the backscattering coefficients of each image were normalized by the values for deep bodies of water such as lakes. This technique has been reported to be useful for creating datasets for rice growth monitoring from time series SAR imagery (Miyaoka et al., 2013). Table 1 shows the acquisition dates of the COSMO-SkyMed data used in this study.

There are several techniques to integrate remote sensing data into a crop model. The main techniques are as follows: remote sensing information is directly inputted to drive the crop model; physical variables, such as leaf area index (LAI), derived from remote sensing data are used to update the relevant variables in the crop model; remote sensing information is used to verify the prediction result of the crop model; and parameters in a crop model are readjusted by comparing physical variables, such as LAI, derived from both the crop model and remote sensing data. This last technique was employed to integrate the data in this study, and LAI was selected as the physical variable used to readjust model parameters. The readjusted model parameters were the maximum amount of nitrogen stored in soil (Nstore_max), nitrogen supply level (N_supply), and nitrogen loss (N_loss); these are each related to field characteristics used in SIMRIW-RS. Initial values of these parameters were obtained from field experiments at the Kyoto University Experimental Farm (Kyoto, Japan). The cultivar parameters required to drive SIMRIW-RS were also obtained from the results of field experiments at Kyoto University Experimental Farm, and those values were fixed in the model for this study.

Figure 1 shows a flowchart detailing the integration of remote sensing data into the crop model used in this study. As shown in this figure, location and transplanting date of each rice paddy field are first detected by satellite image. Second, transplanting date and areal distribution of water bodies, including flooded paddy fields, were extracted according to the threshold derived from time series SAR imagery. This technique has been reported to be useful for creating datasets for rice growth monitoring from time series SAR imagery (Miyaoka et al., 2013). Table 1 shows the acquisition dates of the COSMO-SkyMed data used in this study.

![Flowchart](image-url)

**Table 1. Acquisition dates of COSMO-SkyMed data used in this study.**

<table>
<thead>
<tr>
<th>Month</th>
<th>Day</th>
</tr>
</thead>
<tbody>
<tr>
<td>June</td>
<td>28</td>
</tr>
<tr>
<td>July</td>
<td>2, 6, 14, 18, 22, 30</td>
</tr>
<tr>
<td>August</td>
<td>7, 15, 23, 31</td>
</tr>
<tr>
<td>September</td>
<td>16, 24</td>
</tr>
<tr>
<td>October</td>
<td>2</td>
</tr>
</tbody>
</table>

† LAI measurements were conducted within a few days of the dates in bold.
as according to the land-use map. Terra/ASTER images archived on 23 January and 11 November in 2004, 9 January in 2005, and 6 March in 2008 were used for classifying land use. The land-use map consisted of 8 categories (1: paddy field, 2: dry field, 3: plantation, 4: forest, 5: build-up area, 6: bare land, 7: water body, 8: cloud and shadow) and spatial resolution is 15 meters. The behavior of time-series normalized backscattering coefficients at each paddy field was then analyzed to detect transplanting dates. In this study, it was assumed that the value of each pixel classified paddy field by land-use map represented pure value of paddy field because almost all of paddy fields in this study area were larger than a pixel size of COSMO-SkyMed data used in this study.

In this study, the estimated equation for LAI was derived from the relationship between actual LAI and normalized backscattering coefficients calculated from COSMO-SkyMed data. To obtain actual LAI, LAI measurements were taken with a LAI-2200 plant canopy analyzer (Li-Cor, Inc., Lincoln, NE, USA) several times within a few days of the dates observed by COSMO-SkyMed. Measurements were conducted within a few days of the dates highlighted in boldface type in Table 1. The locations of LAI measurements and detection of transplanting dates (19 locations) are shown in Fig. 2. LAI measurements were conducted 5 times at each location on each date. The average of LAI measurements at each location on each date was taken as the actual LAI for each location and date. For validation of each transplanting date, interviews were conducted with the farmers who managed each of the paddy fields where LAI measurements were taken. For validation of the yield simulated by SIMRIW-RS, the rice weights after harvesting at each of the same locations were measured.

Fig. 1. Flowchart of the integration of remote sensing data into the crop model.

Fig. 2. The locations of LAI measurements and detection of transplanting dates.
3. Results and discussion

To prepare for detecting the transplanting date, backscattering coefficients of 3,000 pixels of the Nam Ngum Dam reservoir in each image were used for normalizing the remaining pixels. The threshold for extracting the water bodies from each image was then decided by Otsu’s method (Otsu, 1979), using a histogram consisting of the major land use in all images. For this study, the threshold value was 12.58. Finally, transplanting dates of rice paddy fields were detected by analyzing the behavior of time-series normalized backscattering coefficients at each paddy field extracted by the land-use map. Paddy fields where the normalized backscattering coefficient had continuously increased over the course of 2 weeks to 1 month, ultimately exceeding the threshold, after which the field had been detected as a water body, were considered to be fields with rice planted, and the date that the field was detected as water body was identified as the transplanting date because transplanting was generally performed within approximately 1 week after flooding in the study area. These conditions were later empirically confirmed by interview to consider rice growth after transplanting in the study area. Table 2 shows comparisons between transplanting dates detected by use of COSMO-SkyMed data and actual transplanting dates as obtained by interviews with farmers. This revealed that the transplanting date of each paddy field was successfully detected by COSMO-SkyMed data, except for a few points where the actual date was different. Therefore, it is considered that the proposed method for detecting the transplanting date is useful for this study area.

Figure 3 shows the relationship between actual LAI and the corresponding normalized backscattering coefficients of the COSMO-SkyMed data. The estimated equation for LAI using normalized backscattering coefficients of the COSMO-SkyMed data is given in the figure. This estimated equation was derived from the data obtained through all field experiments and has a coefficient of determination \( r^2 \) of 0.80. Several measured data points were excluded from calculation of the estimated equation because LAI were not accurately measured due to rain just before field measurements. Consequently, the estimated equation was derived from 47 points of measured data. This result suggests that COSMO-SkyMed data can estimate LAI with relatively high accuracy. Furthermore, the estimated equation derived from this relationship may be able to successfully predict the result if the LAI is smaller, because of correspondingly smaller dispersion of values in that region. This means that this estimated equation may be more suitable for the early growing period than for the heading and maturing periods.

Figure 4 indicates the time-series LAI change simulated by SIMRIW-RS without readjustment of field parameters in the model between transplanting and harvesting. Transplanting date was that detected by COSMO-SkyMed data. Table 3 shows actual and simulated yields. Errors in this table were calculated by the following equation,

\[
\text{Error(\%)} = \left( \frac{\text{Yield}_{\text{Actual}} - \text{Yield}_{\text{Model}}}{\text{Yield}_{\text{Actual}}} \right) \times 100
\]

where \( \text{Yield}_{\text{Actual}} \) and \( \text{Yield}_{\text{Model}} \) indicate field-measured yield and yield predicted by SIMRIW-RS, respectively. On the basis of this figure and case 1 in Table 3, we considered that SIMRIW-RS can simulate rice growth and yield accurately without integration when the parameters set in the model are suitable for the target field. Similar results were obtained at 4 locations. In this study, error of less than 10 % was defined high accuracy. On the other hand, as shown in Fig. 5 and case 2 in Table 3, it is also possible that SIMRIW-RS cannot accurately simulate rice growth and

<table>
<thead>
<tr>
<th>Difference between actual and estimated dates</th>
<th>Number of paddy field</th>
</tr>
</thead>
<tbody>
<tr>
<td>Within 1 week</td>
<td>15</td>
</tr>
<tr>
<td>1 week to 2 weeks</td>
<td>3</td>
</tr>
<tr>
<td>2 weeks to 3 weeks</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Comparisons of transplanting dates detected by COSMO-SkyMed data and actual dates.

Fig. 3. The relationship between actual LAI and the corresponding normalized backscattering coefficients of COSMO-SkyMed data.
yield, and this is likely if the setting of parameter values in the model is unsuitable. However, even when a successful result is not obtained by SIMRIW-RS, the accuracy of yield estimation can be improved by integrating COSMO-SkyMed data into SIMRIW-RS if LAI is accurately estimated by COSMO-SkyMed data (see Fig. 5). All LAI derived from COSMO-SkyMed data shown in Fig. 5 were used for readjustment of field parameters in the model. In this study, field parameters ($N_{\text{storemax}}$, $N_{\text{supply}}$, and $N_{\text{loss}}$) were selected as readjusted parameters, signifying that readjustment of those parameters by the LAI derived from COSMO-SkyMed data is available for improving yield estimation accuracy so long as LAI is accurately estimated from COSMO-SkyMed data. Similar results were obtained at 7 locations.

Unfortunately, LAI cannot always be derived from the estimated equation accurately because, in this study, the coefficient of determination of the equation is not very high. The estimation accuracy of larger LAI may even be lower because, as Fig. 3 shows, the dispersion of LAI becomes wider as the target LAI become larger. Figure 6 gives the result obtained from the readjusted model by using inaccurate LAI derived from COSMO-SkyMed data. All LAI derived from COSMO-SkyMed data shown in Fig. 6 were used to readjust the field parameters in the

<table>
<thead>
<tr>
<th>Case</th>
<th>Actual</th>
<th>Default model (Error: %)</th>
<th>Readjusted model (Error: %)</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1(Fig. 4)</td>
<td>278.3 (8.2)</td>
<td>-</td>
<td>-</td>
<td>303.2</td>
</tr>
<tr>
<td>Case 2(Fig. 5)</td>
<td>292.1 (19.4)</td>
<td>375.0 (3.5)</td>
<td>362.4</td>
<td></td>
</tr>
<tr>
<td>Case 3(Fig. 6)</td>
<td>288.4 (110.5)</td>
<td>303.3 (121.4)</td>
<td>137.0</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 4. Successful time-series LAI simulation by SIMRIW-RS without readjustment of field parameters.

Fig. 5. Time-series LAI simulation by SIMRIW-RS before and after readjustment of field parameters, using accurate estimated LAIs from COSMO-SkyMed data. CSK means COSMO-SkyMed.
model. As a consequence, although optimized values of field parameters could not be obtained because the convergence condition for Equation (1) could not be satisfied, it seems that yield estimation accuracy was lower than the accuracy of the simulation without readjustment (see case3 in Table 3). This result was due to the use of inaccurate LAIs derived from the equation estimated by using COSMO-SkyMed data. Similar results were obtained at 8 locations.

As mentioned above, although SIMRIW-RS itself is a reliable crop model if field parameters in the model are suitable for the target area, the accuracy of yield estimation with using integrated SIMRIW-RS varies with the estimation accuracy of LAI from COSMO-SkyMed data, at least when readjustment of field parameters is required. Therefore, to estimate rice yield precisely by using SIMRIW-RS across the study area, it is necessary to accurately estimate LAIs at each rice paddy field from COSMO-SkyMed data.

4. Conclusion

In this study, integration of remote sensing data into a crop model was investigated. SIMRIW-RS, which was developed by our research group, was selected as a crop model, and COSMO-SkyMed images were used as remote sensing data. Detection methods for transplanting date, which is required to drive SIMRIW-RS, were investigated. The results clearly indicated that transplanting date was accurately detected by applying the threshold determined by Otsu’s method to time-series COSMO-SkyMed data and analyzing the behavior of time-series normalized backscattering coefficients at each pixel. Moreover, estimation of LAI used for readjustment of field parameters in the model was conducted by using COSMO-SkyMed data, and actual LAIs were measured by field survey. The estimated equation was derived from the relationship between normalized backscattering coefficients of COSMO-SkyMed and LAI measured at each field within a few days of the dates observed by COSMO-SkyMed. The results suggested that the estimated equation could estimate LAI relatively accurately because the coefficient of determination of the equation was 0.80. Rice growth and yield were first simulated by SIMRIW-RS without readjustment of field parameters in the model. This result indicates that SIMRIW-RS itself has an ability to accurately simulate rice growth and yield without readjustment of field parameters when the parameter values set in the model are suitable for the target field. Even when a successful result was not obtained by SIMRIW-RS itself, the accuracy of yield estimation was improved by readjustment of field parameters in the model according to the LAIs derived from COSMO-SkyMed data, so long as those LAIs were accurately obtained by the estimated equation. However, when the LAIs were inaccurately derived from COSMO-SkyMed data, the accuracy of the prediction by the readjusted SIMRIW-RS about yield may be lower than that of the prediction by SIMRIW-RS without readjustment.

Although rice yield simulation by integrating COSMO-SkyMed data into SIMRIW-RS was not performed across the entire study area, this study is evidence that SIMRIW-RS itself is a reliable crop model if field parameters in the model are suitable for the target area, and that even when a successful result is not obtained by SIMRIW-RS itself, the accuracy of yield estimation is improved by readjustment of field parameters in the model using LAIs obtained from COSMO-SkyMed data, so long as LAIs are accurately derived by the estimated equation.

To accurately estimate rice yield across a wide area, we consider the improvement of the estimated equation for LAI from COSMO-SkyMed data to be the most important factor. Therefore, further study about the estimation of LAIs by using COSMO-SkyMed data is necessary. Moreover, the combined use of C-band and L-band SAR data with X-band data taken by COSMO-SkyMed should be investigated in the near future.

Acknowledgements

This research was supported by the Environment Research and Technology Development Fund (E-1104: Development and Practice of Advanced Basin Model in Asia).
References