Hydrol. Earth Syst. Sci. Discuss., 12, 1329–1353, 2015
www.hydrol-earth-syst-sci-discuss.net/12/1329/2015/
doi:10.5194/hessd-12-1329-2015

Hydrology and Earth System Sciences
Discussions
Open Access

SACRA – global data sets of satellite-derived crop calendars for agricultural simulations: an estimation of a high-resolution crop calendar using satellite-sensed NDVI

S. Kotsuki[1] **and K. Tanaka**[2]

[1]Advanced Institute for computational Science, RIKEN, Japan
[2]Disaster Prevention Research Institute, Kyoto University, Japan

Received: 22 December 2014 – Accepted: 21 January 2015 – Published: 29 January 2015

Correspondence to: S. Kotsuki (shunji.kotsuki@riken.jp)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

To date, many studies have performed numerical estimations of food production and agricultural water demand to understand the present and future supply–demand relationship. A crop calendar (CC) is an essential input datum to estimate food production and agricultural water demand accurately with the numerical estimations. CC defines the date or month when farmers plant and harvest in cropland. This study aims to develop a new global data set of a satellite-derived crop calendar for agricultural simulations (SACRA) and reveal advantages and disadvantages of the satellite-derived CC compared to other global products. We estimate global CC at a spatial resolution of 5 min (≈ 10 km) using the satellite-sensed NDVI data, which corresponds well to vegetation growth and death on the land surface. We first demonstrate that SACRA shows similar spatial pattern in planting date compared to a census-based product. Moreover, SACRA reflects a variety of CC in the same administrative unit, since it uses high-resolution satellite data. However, a disadvantage is that the mixture of several crops in a grid is not considered in SACRA. We also address that the cultivation period of SACRA clearly corresponds to the time series of NDVI. Therefore, accuracy of SACRA depends on the accuracy of NDVI used for the CC estimation. Although SACRA shows different CC from a census-based product in some regions, multiple usages of the two products are useful to take into consideration the uncertainty of the CC. An advantage of SACRA compared to the census-based products is that SACRA provides not only planting/harvesting dates but also a peak date from the time series of NDVI data.

1 Introduction

Recent population growth has increased food demand significantly, and humans have expanded cropland globally. Agriculture occupies more than 70 % of world water usage and has a large impact on the water cycle (Rost et al., 2008). Consequently, simulations of food production and agricultural water demand are necessary to understand the

HESSD
12, 1329–1353, 2015

present and future supply–demand relationship. To date, many studies have estimated food production (Fischer et al., 2000; Tan and Shibasaki, 2003; Stehfest et al., 2007) and agriculture water demand (Döll et al., 2002; Hanasaki et al., 2008; Rockström et al., 2009; Siebert and Döll, 2009; Pokrel et al., 2011). Those studies estimated food production and agricultural water demand with numerical models using meteorological forcing data and land surface parameters. A crop calendar (CC) is an essential input datum to estimate food production and agricultural water demand accurately with those hydrological models. CCs define the date or month when farmers plant and harvest in cropland. There are three major approaches to develop CC data sets: a census-based method, modeling method, and satellite-based method.

The first approach, the census-based method, estimates CCs by collecting and integrating agricultural census data provided by international and national organizations such as the Food and Agriculture Organization (FAO) and US Department of Agriculture (USDA). The census-based CC products are represented by MIRCA2000 (Portmann et al., 2010; Monthly Irrigated and Rain-fed Crop Areas around the year 2000) and Sacks et al. (2010). The census-based products have an advantage of high reliability in regions that have sufficient census data. However, they also have disadvantages of low reliability in regions that have no census data. Additionally, spatial resolution of the census-based products is limited because of the sampling scheme (Portmann et al., 2010). Because only one CC is defined in an administrative unit for each crop, the differences of CCs within the same administrative unit are not considered.

The second approach, the modeling method, generates CC using crop growth models. These models simulate crop growth with meteorological and agricultural forcing data such as temperature, solar radiation, and soil moisture. Especially, accumulated temperature is widely used to indicate crop growing stage. Hanasaki et al. (2008) estimated global CC for several crops using the soil and water integrated model (SWIM; Krysnova et al., 2000). The crop growth models have an advantage of accurate crop-growth simulation with well-calibrated parameters. However, it is difficult to calibrate parameters properly in areas where observation data are insufficient. Additionally, the

HESSD
12, 1329–1353, 2015

crop growth model has difficulty in identifying the planting dates because the cultivation period is heavily affected by human activities. Normally, modelling methods use the census data to identify the planting dates or determine CC to maximize crop yields.

The third satellite-based studies estimate CC using time series of satellite observation data. Time series of vegetation indexes (VIs) corresponds well to vegetation growth and death on the land surface. In this context, the satellite-derived VIs have been widely used to classify crop type and monitor crop growth on a regional scale (Mingwei et al., 2008; Sakamoto et al., 2005, 2010; Wardlow and Egbert, 2008; Wardlow et al., 2007). An advantage of satellite-derived data is its spatial resolution (less than 1 km). However, few studies have estimated global CC with satellite-derived data. Yorozu et al. (2005) estimated global CC using the normalized difference vegetation index (NDVI), but they performed no comparison with other global CC data sets.

Here, we present a new global data set of a SAtellite-derived CRop calendar for Agricultural simulations (SACRA). Using the satellite-sensed NDVI data, we estimate global CC at a spatial resolution of 5 min ($\approx$ 10 km). This study aims to develop a high-resolution and high-accurate CC product by using both the satellite-derived NDVI and a census-based product. We also aim to reveal advantages and disadvantages of our satellite-derived CC compared to existing census-based and model-derived products.

Section 2 describes input data and methods to estimate CC, and Sect. 3 describes the validation and discussion through comparisons with other CC data sets. Finally, Sect. 4 provides the summary.

2 Data and methods

This section describes the methods to produce SACRA according to data processing scheme (Fig. 1). SACRA is produced from four different data sets: time series of NDVI, land cover data, reanalysis temperature, and monthly agricultural data (Table 1). We identify planting and harvesting dates (t_{pl} and t_{hv}) using two CC parameters ($\mathrm{nNDVI_{pl}}$ and $\mathrm{nNDVI_{hv}}$) and a time series of NDVI data (Fig. 2a). The planting and harvesting

HESSD
12, 1329–1353, 2015

dates are determined by the following:

$$t = \begin{pmatrix} t_{\text{pl}} \\ t_{\text{hv}} \end{pmatrix} \text{ when } \begin{pmatrix} t \le t_{\text{pk}}, & \text{nNDVI}(t) > \text{nNDVI}_{\text{pl}} \\ t \ge t_{\text{pk}}, & \text{nNDVI}(t) < \text{nNDVI}_{\text{hv}} \end{pmatrix}, \tag{1}$$

where nNDVI represent normalized NDVI. The subscripts pl, pk, and hv denote plant, peak, and harvest, respectively. The following subsections describe vegetation indexes, such as NDVI and nNDVI (Sect. 2.1), typical crop and cropping pattern (Sect. 2.2), and the two CC parameters (Sect. 2.3) and the produced SACRA data sets (Sect. 2.4).

2.1 Vegetation index

2.1.1 SPOT-Vegetation NDVI data

The vegetation indexes are simple, graphic indicators to assess whether the targeting area contains live, green vegetation or not. In this study, we use NDVI defined by the following:

$$\text{NDVI} = \frac{\text{NIR} - \text{VIS}}{\text{NIR} + \text{VIS}}, \tag{2}$$

where VIS and NIR indicate the spectral reflectance in the visible and near-infrared bands. Leaves absorb visible light for photosynthesis and reflect near-infrared light. The time series of NDVI corresponds to vegetation growth. Over the cropland, the time series of NDVI represents crop growth. The time series of satellite-sensed NDVI at a double-cropping grid in China is shown in Fig. 3a. As shown in Fig. 3a, peak dates can be clearly identified from the time series of NDVI. NDVI becomes relatively large in the cultivation seasons and decreases after peaks. In this study, we use a 10 day composite NDVI provided by SPOT-Vegetation (Maisongrande et al., 2004). To remove the effect of clouds, the best index slope extraction (BISE) method (Viovy et al., 1992) is applied to the time series of NDVI (Fig. 3a). We use the time series of NDVI averaged over three years (2004–2006). Hereafter, this averaged NDVI is indicated by SPOT-NDVI in this manuscript.

HESSD
12, 1329–1353, 2015

2.1.2 Aggregation and normalization of NDVI

Our method needs time series of NDVI, which represents crop growth on the cropland. Two NDVI data sets (NDVI-Crop and NDVI-Filled; 5 min resolution) are aggregated from original SPOT-NDVI (1 km resolution) using two land-cover data sets: global land cover characterization version 2.0 (GLCC) and Ecoclimap version 2.0. The GLCC and Ecoclimap data are provided by the US Geological Survey and Meteo France, respectively. The NDVI-Crop and NDVI-Filled are generated by averaging the SPOT-NDVI grids where both the GLCC and Ecoclimap denote the cropland. If NDVI-Crop and NDVI-Filled are not determined, only NDVI-Filled is generated by averaging the SPOT-NDVI grids where the GLCC or Ecoclimap denotes the cropland. The NDVI-Crop is used to identify the two CC parameters (Sect. 2.3). The NDVI-Filled is used to produce the global CC in Sect. 2.4.

In this study, planting and harvesting dates are determined by time series of vegetation index and two CC parameters. These two CC parameters are defined for each crop type. We apply the same CC parameters to global cropland. Peak value of NDVI ($t = t_{pk}$), however, differs depending on climate condition and density of crops. Therefore, NDVI data are normalized to have an identical maximum value:

$$\mathrm{nNDVI}(t) = \frac{\mathrm{NDVI}(t) - \mathrm{NDVI}_{bas}}{\mathrm{NDVI}_{pk} - \mathrm{NDVI}_{bas}}, \tag{3}$$

$$\mathrm{NDVI}_{bas} = \max(\mathrm{NDVI}_{btm}, \mathrm{NDVI}_{param}), \tag{4}$$

where nNDVI represents normalized NDVI. The subscripts btm and bas denote bottom and base, respectively. We have 36 NDVI data per year from 10 day composite data of the SPOT-NDVI. The NDVI_{pk} and NDVI_{btm} are determined in the following manner: first, the number of cultivations is determined to be equal to the number of peaks of NDVI up to twice per year. If NDVI at peak is less than 0.40, the peak is not counted as cultivation. Second, the minimum NDVI between the peaks are determined to be the NDVI_{btm}. If the number of cultivations is once in a year, the NDVI_{btm} equals the

HESSD
12, 1329–1353, 2015

minimum NDVI over the year. The $NDVI_{param}$ is a parameter to avoid remnant irregular NDVI caused by cloud and snow cover reflection. The $NDVI_{param}$ is set to be 0.10 in this study. As shown in Fig. 2b, the normalized NDVI (hereafter, nNDVI) always have same maximum nNDVI = 1.0. We apply this normalization for both NDVI-Crop and NDVI-Filled.

2.2 Typical crop and cropping pattern

This subsection describes the method to determine a typical crop and cropping pattern using MIRCA2000, reanalysis temperature, and time series of NDVI-Filled (Fig. 1). MIRCA2000 provides irrigated and rain-fed crop areas of 26 crop types at a spatial resolution of 5 min. Their crop calendars are also defined in MIRCA2000. We define the typical crop that has largest cultivation area in a grid of the 26 crop types. Since our method cannot consider the mixture of several crops in a grid, we assume that the NDVI-Crop and NDVI-Filled represent the phenology of the typical crop in each grid. The typical crop in the second cultivation season is also identified in the same manner.

Our method is unsuitable for estimation of planting dates of winter crops because it assumes consistent growth in NDVI from planting to peak. However, NDVI decreases if the surface is covered by snow (Fig. 2b). Therefore, we consider summer crops and winter crops separately. If the minimum monthly-averaged temperature during the cultivation period is below 5.0 °C, the typical crop is categorized as a winter crop. In this categorization, we use cultivating periods from MIRCA2000 and reanalysis temperature (Hirabayashi et al., 2008). The global distribution of the typical crop is shown in Fig. 4a. The cropping pattern is shown in Fig. 4b, which is a combination of cultivation times and summer/winter crop discrimination. The minimum monthly-averaged temperature during the cultivation period of the typical crop is shown in Fig. 4c. Because we define winter crops by temperature, cultivated wheat in Australia and Northern India is defined as spring wheat. Regions having the minimum monthly-averaged temperature below 5.0 °C in Fig. 4c are categorized as winter wheat or fodder (permanent crop) in Fig. 4a.

HESSD
12, 1329–1353, 2015

2.3 Crop calendar parameter

This subsection describes a method to identify the two CC parameters, $nNDVI_{pl}$ and $nNDVI_{hv}$, which are produced from time series of NDVI-Crop, crop calendar of MIRCA2000, typical crop, and cropping pattern (Fig. 1). The two CC parameters are determined for each crop type. In this study, we consider five widely cultivated crops: spring wheat, maize, rice, soybean, and cotton. To remove the noise of the time series of NDVI data as much as possible, we use limited grids (hereafter, pure grids) to estimate the CC parameters. The pure grids satisfy the following conditions: (1) single summer crop defined by cropping pattern, (2) cropland occupying more than 25 % of the grid area, and (3) up to five grids from the same administrative unit of MIRCA2000. Once the parameters $nNDVI_{pl}$ and $nNDVI_{hv}$ are determined, planting and harvesting dates are calculated automatically using Eq. (1). The values of two CC parameters $nNDVI_{pl}$ and $nNDVI_{hv}$ are determined for each crop to minimize the errors between planting/harvesting dates of determined and MIRCA2000 over pure grids. Table 2 shows the number of pure grids, the two CC parameters and averaged errors in planting/harvesting dates between determined and MIRCA2000 over pure grids of the five crop types. As shown in Table 2, averaged differences between planting/harvesting dates of determined and MIRCA2000 over pure grids are less than 15 days. The CC of the winter wheat is substituted for by the spring wheat because our method is unsuitable for the winter crops.

2.4 SACRA and SACRA-filled data sets

The global planting and harvesting dates are determined by Eq. (1) using time series of nNDVI-Filled (Sect. 2.1), typical crop, cropping pattern (Sect. 2.2), and the two CC parameters (Sect. 2.3). Our method has the advantage of capturing the cultivation season properly using time series of the satellite-sensed NDVI. However, our algorithm has the possibility of overestimating or underestimating cultivation periods. Therefore, we adjusted the cultivation period of SACRA to that of MIRCA2000. For the summer

HESSD
12, 1329–1353, 2015

crops, planting and harvesting dates are advanced/postponed the same days to adjust to the cultivation period of MIRCA2000. For the winter crops, only the planting date is advanced for the adjustment because our method is unsuitable to estimate the planting date of the winter crops. Although we adjust cultivation period, SACRA's cultivation season is governed by the time series of NDVI. SACRA data sets contain planting and harvesting dates of the global cropland after the adjustment.

Our algorithm cannot compute crop calendar at grids where no cropland is categorized by GLCC and Ecoclimap. We also produce SACRA-Filled data sets by substituting CC of MIRCA2000 in grids where SACRA contains no CC data. We discuss the results using the produced SACRA data set in Sect. 3.

3 Results and discussion

This section provides validations and discussion on the produced SACRA data set. Spatial distributions of the planting dates of the typical crop in SACRA and MIRCA2000 are shown in Fig. 5. Although SACRA (Fig. 5a1 and a2) and MIRCA2000 (Fig. 5b1 and b2) are produced from different methods, they have similar spatial patterns. Since SACRA uses high-resolution satellite data, it reflects a variety of planting dates in the same administrative unit, as shown in Fig. 5a2 and b2. The satellite-based methods can produce the CC with a finer resolution than the census-based methods. Crop calendars of SACRA are compared with those of MIRCA2000 in Sect. 3.1, and advantages and disadvantages of SACRA compared to other CCs are discussed in Sect. 3.2.

3.1 Comparison with MIRCA2000

The difference in planting dates of a typical crop between SACRA and MIRCA2000 is shown in Fig. 6a. MIRCA2000 has one crop calendar in the same administrative unit. As SACRA can have several crop calendars in the same administrative unit, an averaged crop calendar is used for comparison. If the typical crop has no crop calen-

HESSD
12, 1329–1353, 2015

dar (e.g., a permanent crop) in the region, the region is drawn in white. SACRA and MIRCA2000 have the same cultivation period because the cultivation period of SACRA is adjusted to that of MIRCA2000. As shown in Fig. 6a, the two data sets show small differences in many regions (regions with gray and purple colors in Fig. 6a). In those regions, the SACRA and MIRCA results suggest the accuracy of each other. However, large differences are observed in several regions (regions with red color in Fig. 6a). Figure 6b compares cultivation seasons of the two products in nine regions, having a large difference in planting month in Fig. 6a. The lower panels of Fig. 6b show the cultivating periods of the two products. The upper panels of Fig. 6b show the time series of NDVI. The upper panels also show histograms of the planting dates of SACRA because SACRA can have several CCs in the same administrative unit. NDVI shows averaged NIDV-Filled over the corresponding regions. Because SACRA is produced from the time series of NDVI, cultivation seasons of SACRA should exist over peaks of NDVI.

A possible question is that if SACRA and MIRCA have different CCs, which one is more reliable? The accuracy of SACRA is affected by the accuracy of the land cover data sets. It is known that the 1 km land cover data sets contain uncertainties (Herold et al., 2008; Nakaegawa, 2011). For example, forests can be classified as croplands in the 1 km land cover data sets. If the land cover data sets are inaccurate in terms of the regions, SACRA can be different from MIRCA2000. However, SACRA can be used to validate MIRCA2000 data sets. For example, MIRCA2000 predicts winter wheat in Mongolia (Fig. 6b3) because the cultivation period is over the winter. Several reports indicate that spring wheat is the typical crop in North China (e.g., US Department of Agriculture, 1994). Spring wheat can be a typical crop in Mongolia because Mongolia also has a cold winter (see also Fig. 4c). The cultivation period of SACRA in Mongolia is reasonable for spring wheat. It is difficult to determine whether one product is accurate or inaccurate through the comparisons. However, multiple usages of the two products are useful to take into consideration the uncertainty of the CC.

HESSD
12, 1329–1353, 2015

Two peaks are captured by the time series of NDVI in Arizona and Uzbekistan (Fig. 6b1 and b2). Because NDVI of the double-cropping grid is not included in this average NDVI, the two peaks suggest that there are two phenologies in Arizona and Uzbekistan. The histograms of the planting month indicate that there are several cultivation periods in Arizona, and mainly two phenologies in Uzbekistan. MIRCA2000 may represent one of the phenologies. SACRA has the advantage of considering several CC in the same administrative unit depending on the VI phenology, as shown in Arizona and Uzbekistan.

3.2 Advantages and disadvantages of SACRA

This subsection discusses the advantages and disadvantages of SACRA compared to two other global crop calendars: census-based methods (e.g., Portmann et al., 2010; Sacks et al., 2010) and modeling methods (e.g., Hanasaki et al., 2008). Additionally, this subsection also discusses possible improvements of SACRA. Table 3 summarizes the advantages and disadvantages of the census-based methods, modeling methods, and SACRA.

First, an advantage of SACRA is its fine resolution compared to the other two data sets. Therefore, different CCs in the same administrative unit are considered in SACRA. However, our method has a disadvantage in that the mixture of several crops in a grid is not considered. We assume that the NDVI-Crop and NDVI-Filled represent the phenology of the typical crop at each grid. The spatial resolution of SACRA is equal to the maximum resolution of the satellite-sensed NDVI and the crop classification map. At present, NDVI from the moderate-resolution imaging spectroradiometer (MODIS) is available at a spatial resolution of 250 m (e.g., Zhang et al., 2006). However, present studies provide a global crop classification map at a spatial resolution of 10 km (e.g., Monfreda et al., 2008; Potmann et al., 2010). Present land cover data sets, such as GLCC and Ecoclimap, contain a few agricultural land classes. At the regional scale, many studies have been performed to classify the crops species using satellite-sensed data (e.g., Mingwei et al., 2008; Wardlow and Egbert, 2008; Wardlow et al., 2007). In

HESSD
12, 1329–1353, 2015

this study, we use the crop classification map from MIRCA2000 at a spatial resolution of 10 km. If high-resolution crop classification maps become available in the future, SACRA can be downscaled into the fine resolution.

Second, an advantage of SACRA is easy identification of cultivation using peak NDVI. Because agriculture is controlled by the human activities, it is difficult to estimate whether or not farmers perform cultivation from the census-based and modeling methods. Additionally, agriculture is also affected by disasters, such as droughts, inundations, heat waves, and cool summer damages. The satellite-sensed NDVI are observed data corresponding to crop growth. It is also possible to identify double-cropping by NDVI, as shown in Fig. 3. In this study, our method was applied to the average NDVI over a 3 year period. Application of our method to the annual NDVI data would facilitate estimation of the CC for each corresponding year.

Third, SACRA has a disadvantage in that it is inapplicable to future simulations such as impact assessments of climate change because SACRA is produced using past observational data. Future changes in agricultural water demand and food production are major issues in assessment studies of climate change (Hanjra and Qureshi, 2010). SACRA can be used to calibrate parameters of the crop growth model. An advantage of SACRA compared to MIRCA2000 is that SACRA provides not only planting/harvesting dates but also peak date from the time series of NDVI data. SACRA can contribute to future assessment studies indirectly by being utilized to calibrate the model parameters.

Finally, accuracy of SACRA depends on the accuracy of NDVI and land cover data sets. In this study, we use the two land cover data sets to reduce the uncertainty of the land cover data. The land cover data sets, however, contain uncertainties (Herold et al., 2008; Nakaegawa, 2011). The land cover data sets could be improved by altering the algorithms, increasing the amount of supervised data, and utilizing multi-spectrum information. Further improvements of the land cover data sets would lead to improvement of SACRA.

HESSD
12, 1329–1353, 2015

4 Summary

This study aimed at producing a new crop calendar, SACRA, using satellite-sensed NDVI. This paper describes the methods to produce SACRA from the following four data sets: time series of NDVI, land cover data sets, reanalysis temperature, and census monthly agricultural data. The resulting SACRA data set included three products of subset at a spatial resolution of 10 km: (1) the spatial distribution of typical crop and cropping patterns, (2) time series of NDVI of the cropland, (3) planting and harvesting dates of the typical crop. The advantages and disadvantages of SACRA compared to other global crop calendars are summarized as follows.

First, an advantage of SACRA is its finer spatial resolution than other existing global crop calendars. However, a disadvantage is that the mixture of several crops in a grid is not considered in SACRA. Second, the cultivation period of SACRA is identified from the time series of NDVI, which clearly corresponds to crop growth. Therefore, we can consider the effects of human activities and natural disasters. Satellite-sensed NDVI data enable the identification of whether or not famers perform cultivation. It is possible to generate annual crop calendars by applying our method to annual NDVI data. Finally, SACRA is inapplicable to future simulations because SACRA is produced by *past* satellite-sensed data. However, we can use SACRA to calibrate the parameters of crop growth models. An advantage of SACRA compared to the census-based products is that SACRA provides not only planting/harvesting dates but also a peak date from the time series of NDVI data.

Many improvements to SACRA are possible. We plan to make SACRA data sets available on our web page free of charge. We encourage researchers to utilize our data and provide feedback on errors or possible improvements.

Acknowledgement. S. Kotsuki and K. Tanaka thank the members of the Water Resources Research Center, Kyoto University, for valuable suggestions and discussions. This study was supported in part by the Japan Society for Promotion Science (JSPS) KAKENHI grants 12J03816, 22246066 and the SOUSEI Program of the Ministry of Education, Culture, Sports, Science, and Technology.

References

Döll, P. and Siebert, S.: Global modeling of irrigation water requirements, Water Resour. Res., 38, 1037, doi:10.1029/2001WR000355, 2002.

Fischer, G., Velthuizen, H., and Nachtergaele, F.: Global Agro-Ecological Zones Assessment: Methodology and Results, Interim report, International Institute for Systems Analysis (IIASA), Luxemburg, Austria, and FAO, Rome, 2000.

Hanasaki, N., Kanae, S., Oki, T., Masuda, K., Motoya, K., Shirakawa, N., Shen, Y., and Tanaka, K.: An integrated model for the assessment of global water resources – Part 2: Applications and assessments, Hydrol. Earth Syst. Sci., 12, 1027–1037, doi:10.5194/hess-12-1027-2008, 2008.

Hanjra, M. A. and Qureshi, M. E. Global water crisis and future food security in an era of climate change, Food Pol., 35, 365–377, doi:10.1016/j.foodpol.2010.05.006, 2010.

Herold, M., Mayaux, P., Woodcock, C. E., Baccini, A., and Schmullius, C.: Some challenges in global land cover mapping?: an assessment of agreement and accuracy in existing 1 km datasets, Remote Sens. Environ., 112, 2538–2556, doi:10.1016/j.rse.2007.11.013, 2008.

Hirabayashi, Y., Kanae, S., Motoya, K., Masuda, K., and Petra, D.: A 59-year (1948–2006) global near-surface meteorological data set for land surface models. Part I: Development of daily forcing and assessment of precipitation intensity, Hydrol. Res. Lett., 2, 36–40, doi:10.3178/HRL.2.36, 2008.

Krysanova, V., Wechsung, F., Arnold, J., Srinivasan, R., and Williams, J.: SWIM (Soil and Water Integrated Model) User Manual, Potsdam Institute for Climate Impact Research, Potsdam, Germany, 2000.

Maisongrande, P., Duchemin, B., and Dedieu, G.: VEGETATION/SPOT: an operational mission for the Earth monitoring; presentation of new standard products, Int. J. Remote Sens., 25, 9–14, doi:10.1080/0143116031000115265, 2004.

Mingwei, Z., Qingbo, Z., Zhongxin, C., Jia, L., Yong, Z., and Chongfa, C.: Crop discrimination in Northern China with double cropping systems using Fourier analysis of time-series MODIS data, Int. J. Appl. Earth Obs. Geoinf., 10, 476–485, doi:10.1016/j.jag.2007.11.002, 2008.

Monfreda, C., Ramankutty, N., and Foley, J.: Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000, Global Biogeochem. Cy., 22, 1–19, doi:10.1029/2007GB002947, 2008.

HESSD
12, 1329–1353, 2015

Nakaegawa, T.: Comparison of water-related land cover types in six 1-km global land cover datasets, J. Hydrometeorol., 13, 649–664, doi:10.1175/JHM-D-10-05036.1, 2012.
Pokhrel, Y., Hanasaki, N., Koirala, S., Cho, J., Yeh, P., Kim, H., Kanae, S., and Oki, T.: Incorporating anthropogenic water regulation modules into a land surface model, J. Hydrometeorol., 13, 255–269, doi:10.1175/JHM-D-11-013.1, 2012.
Portmann, F., Siebert, S., and Döll, P.: MIRCA2000 – Global monthly irrigated and rainfed crop areas around the year 2000: a new high-resolution data set for agricultural and hydrological, Global Biogeochem. Cy., 24, 1–24, doi:10.1029/2008GB003435, 2010.
Rockström, J., Falkenmark, M., Karlberg, L., Hoff, H., Rost, S., and Gerten, D.: Future water availability for global food production: the potential of green water for increasing resilience to global change, Water Resour. Res., 45, W00A12, doi:10.1029/2007WR006767, 2009.
Rost, S., Gerten, D., Bondeau, A., Lucht, W., Rohwer, J., and Schaphoff, S.: Agricultural green and blue water consumption and its influence on the global water system, Water Resour. Res., 44, W09405, doi:10.1029/2007WR006331, 2008.
Sacks, W. J., Deryng, D., Foley, J. A., and Ramankutty, N.: Crop planting dates: an analysis of global patterns, Global Ecol. Biogeogr., 19, 607–620, doi:10.1111/j.1466-8238.2010.00551.x, 2010.
Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., and Ohno, H.: A crop phenology detection method using time-series MODIS data, Remote Sens. Environ., 96, 366–374, doi:10.1016/j.rse.2005.03.008, 2005.
Sakamoto, T., Wardlow, B. D., Gitelson, A., Verma, S. B., Suyker, A. E., and Arkebauer, T. J.: A Two-Step Filtering approach for detecting maize and soybean phenology with time-series MODIS data, Remote Sens. Environ., 114, 2146–2159, doi:10.1016/j.rse.2010.04.019, 2010.
Siebert, S. and Döll, P.: Quantifying blue and green virtual water contents in global crop production as well as potential production losses without irrigation, J. Hydrol., 384, 198–217, doi:10.1016/j.jhydrol.2009.07.031, 2010.
Stehfest, E., Heistermann, M., Priess, J. A., Ojima, D. S., and Alcamo, J.: Simulation of global crop production with the ecosystem model DayCent, Ecol. Modelling, 209, 203–219, doi:10.1016/j.ecolmodel.2007.06.028, 2007.
Tan, G. and Shibasaki, R.: Global estimation of crop productivity and the impacts of global warming by GIS and EPIC integration, Ecol. Modelling, 168, 357–370, doi:10.1016/S0304-3800(03)00146-7, 2003.

HESSD
12, 1329–1353, 2015

An estimation of a high-resolution crop calendar using satellite-sensed NDVI

S. Kotsuki and K. Tanaka

Viovy, N., Arino, O., and Belward, A. S.: The Best Index Slope Extraction (BISE): a method for reducing noise in NDVI time-series, Int. J. Remote Sens., 13, 1585–1590, doi:10.1080/01431169208904212, 1992

Wardlow, B., Egbert, S., and Kastens, J.: Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains, Remote Sens. Environ., 108, 290–310, doi:10.1016/j.rse.2006.11.021, 2007.

World Agricultural Outlook Board: Major World Crop Areas and Climate Profiles. US Department of Agriculture, Washington DC, US, 1994.

Yorozu, K., Tanaka, K., Ikebuchi, S.: Creating a global 1-degree dataset of crop type and cropping calendar through timeseries analysis of NDVI for GSWP simulation considering irrigation effect, Proc. of 85th AMS Annual Meeting, J6.8, San Diego, USA, 8–14 January 2005.

Zhang, X., Friedl, M., and Schaaf, C. B.: Global vegetation phenology from Moderate Resolution Imaging Spectroradiometer (MODIS): evaluation of global patterns and comparison with in situ measurements, J. Geophys. Res., 111, G04017, doi:10.1029/2006JG000217, 2006.

Table 1. Characteristics and sources of the four input data used to develop SACRA.

Data	Characteristic and resolution	Source
NDVI	10 days composite/1 km/averaged over 2004–2006	SPOT-Vegetation
Land cover	30 s (≈ 1 km)	GLCC and Ecoclimap
Monthly agricultural data	5 min (≈ 10 km), crop classification map and crop calendar	MIRCA2000 (Portmann et al., 2010)
Temperature	0.5°/averaged over 2004–2006	Hirabayashi et al. (2008)

HESSD
12, 1329–1353, 2015

Table 2. The number of pure grids, two crop calendar parameters ($nNDVI_{pl}$ and $nNDVI_{hv}$), and averaged errors in planting/harvesting dates between determined and MIRCA2000 over pure grids of the five crop types.

	Spring Wheat	Maize	Rice	Soybean	Cotton
Num. of grid	100	136	97	98	35
$nNDVI_{pl}$	0.19	0.18	0.21	0.16	0.18
$nNDVI_{hv}$	0.36	0.50	0.70	0.32	0.25
Ave. err. Plant	15	6	10	7	14
Ave. err. Harvest	12	4	4	8	4

HESSD
12, 1329–1353, 2015

Table 3. Advantages and disadvantages of the three global crop calendars: census-based methods, modeling method studies, and this study (SACRA).

	Census-based methods	Modelling methods	This study (SACRA)
Main inputs	Census data	Forcing data	Satellite-sensed NDVI
Resolution	Country/state scale	Equal to forcing data (dozens of km)	10 km
Mixture of several crops in a grid	possible	possible	impossible
Judgement of cultivation	hard	hard	easy
Different CC in a same admin. unit	impossible	possible	possible
Applications to future simulations	to assume same CC	applicable	to assume same CC

HESSD
12, 1329–1353, 2015

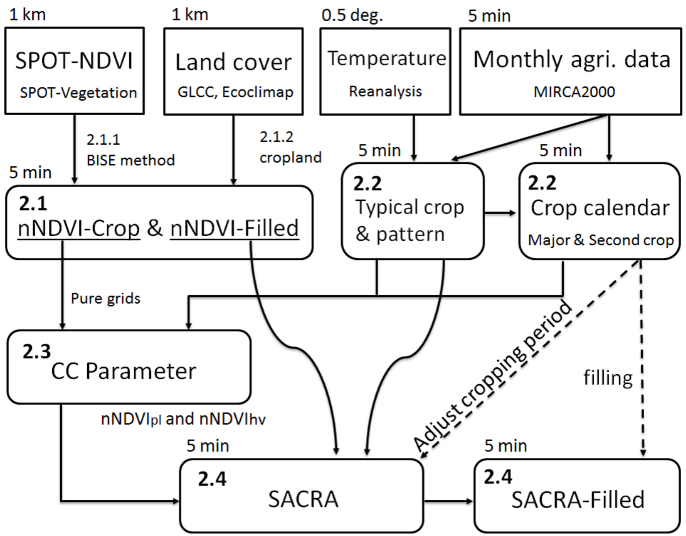

Figure 1. Data processing scheme to produce global data sets of satellite-derived crop calendars.

HESSD
12, 1329–1353, 2015

An estimation of a high-resolution crop calendar using satellite-sensed NDVI

S. Kotsuki and K. Tanaka

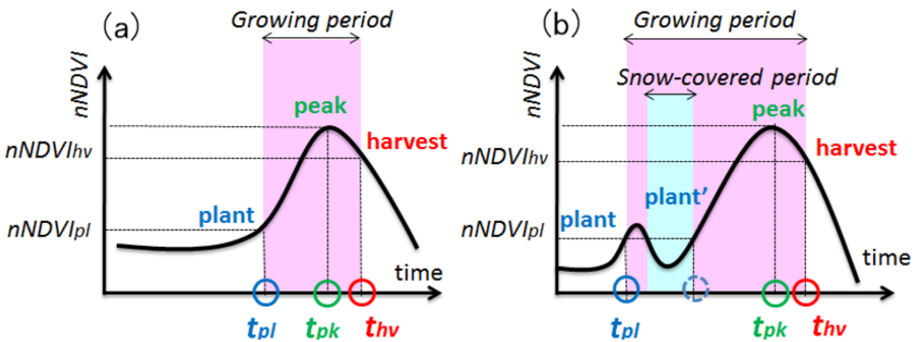

Figure 2. Schematic of identification of planting and harvesting dates in this study. Planting and harvesting dates (t_{pl} and t_{hv}) are identified together with a vegetation index time series and two crop calendar (CC) parameters: $\mathrm{nNDVI}_{\mathrm{pl}}$ and $\mathrm{nNDVI}_{\mathrm{hv}}$. Figures **(a)** and **(b)** indicate summer and winter crops, respectively. The two CC parameters are defined for each crop type.

HESSD
12, 1329–1353, 2015

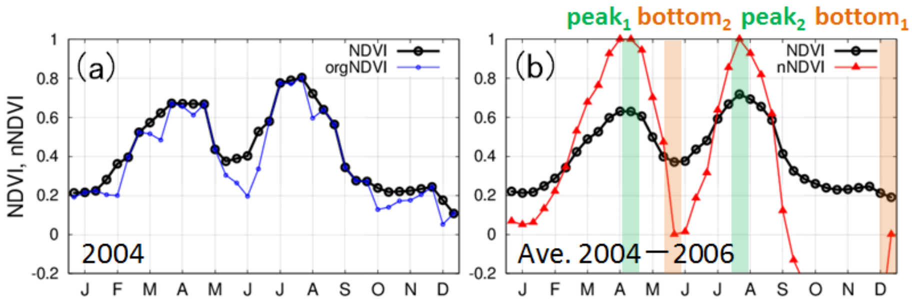

Figure 3. Time series of NDVI at a double-cropping grid in China (116.76° E, 32.60° N). Figure **(a)** represents the original NDVI (blue line) and NDVI with the BISE method (black line) in 2004. Figure **(b)** represents the NDVI (black line) and normalized NDVI (nNDVI; red line) average from 2004 to 2006.

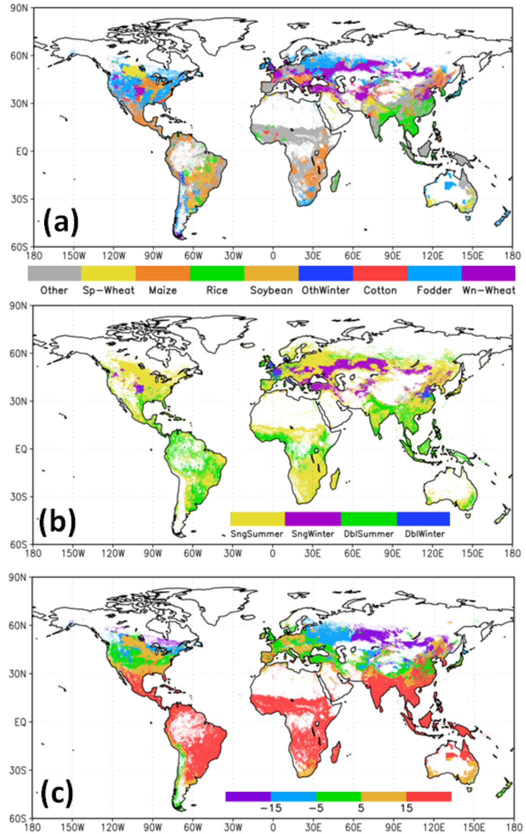

Figure 4. Global distribution of **(a)** a typical crop, **(b)** cropping pattern, and **(c)** minimum monthly-averaged temperature (°C) during cultivation period of the typical crop. The cropping pattern is a combination of cultivation times and summer/winter crop discrimination. SngSummer, SngWinter, and DblSummer in **(b)** represent single summer crop, single winter crop, and double summer crop, respectively. DblWinter represents double cropping with summer and winter crops.

HESSD
12, 1329–1353, 2015

An estimation of a high-resolution crop calendar using satellite-sensed NDVI

S. Kotsuki and K. Tanaka

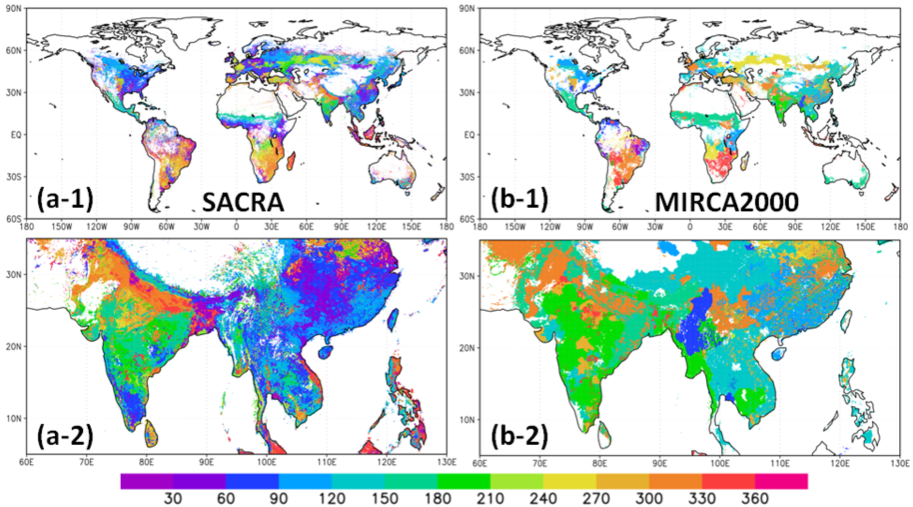

Figure 5. Planting dates of a typical crop in **(a)** SACRA and **(b)** MIRCA2000 (unit: day of year). Figures **(a-1/b-1)** and **(a-2/b-2)** show global and South Asian maps, respectively.

HESSD
12, 1329–1353, 2015

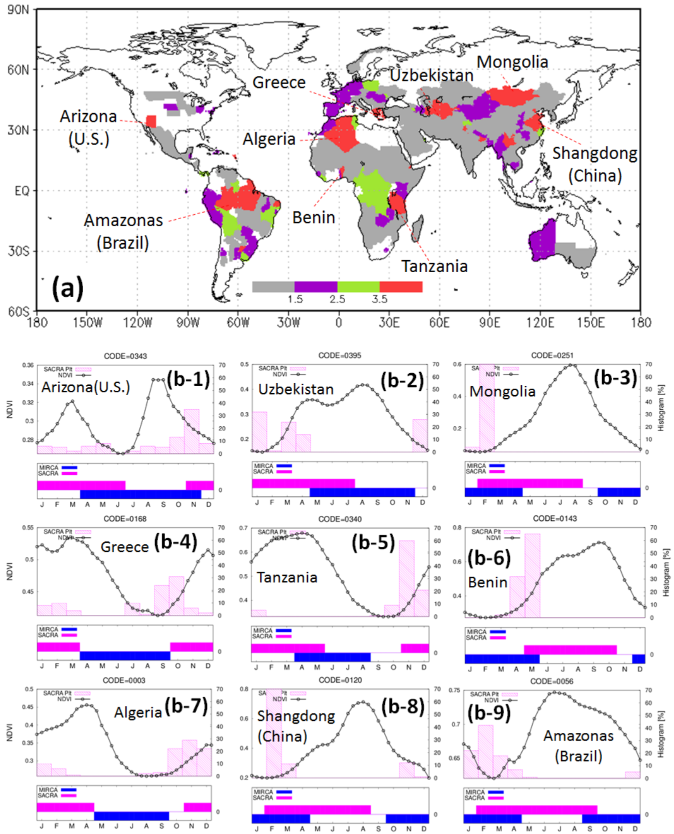

Figure 6. (a) Difference in planting month of the typical crop between MIRCA2000 and SACRA, and **(b)** cultivation seasons of the two products in nine regions (from **b-1** to **b-9**; Arizona in the United States, Uzbekistan, Mongolia, Greece, Tanzania, Benin, Algeria, Shandong in China, and Amazonas in Brazil). Lower panels in **(b)** compare the cultivation periods of SACRA (magenta bands) and MIRCA2000 (blue bands). Upper panels in **(b)** show time series of NDVI (black line) and a histogram of the planting month of SACRA (magenta bars).

HESSD
12, 1329–1353, 2015

An estimation of a high-resolution crop calendar using satellite-sensed NDVI

S. Kotsuki and K. Tanaka