24th International Congress on Modelling and Simulation, Sydney, NSW, Australia, 5 to 10 December 2021
mssanz.org.au/modsim2021

An automatic system to estimate crop phenological dates with remote sensing: A case study in the southeast Australia

Zitian Gao [a], **Danlu Guo** [a], **Dongryeol Ryu** [a] **and Andrew W Western** [a]

[a] *Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC, 3010, Australia*
Email: zitiang@student.unimelb.edu.au

Abstract: Accurate crop phenological dates over large agricultural districts are important for managing agricultural activities and resources. Compared to farmer-reported calendars, satellite-based crop phenological dates can be spatially comprehensive and cost-effective. This study developed an integrated, fully automatic system that can generate crop phenological dates over a large region across decades. The system was developed in the R programming environment and linked to Google Earth Engine (GEE). We 1) used the R package "rgee" to send commands to GEE for satellite data acquisition and processing, and 2) applied state-of-the-art statistical algorithms to extract phenology metrics that represented key phenological dates from time-series data of vegetation indices. All analyses, including satellite data acquisition and pre-processing, calculation of vegetation indices and crop phenology analyses were written in a consolidated R script, making the analyses compactly packaged for efficient handling. The system was tested on three summer crops - corn (maize), cotton and rice over 113 fields in the Coleambally Irrigation Area (CIA) in southern New South Wales (NSW), Australia. We calculated the lengths of growth stages and start and end of the season (SOS and EOS) from phenology metrics for all fields and estimated their variabilities across the district. The satellite-derived crop phenological dates and lengths of growth stage were further compared with the equivalent dates used in FAO-56 and local crop growth guidelines.

The summer crop growing season in our study area starts from Oct-Dec and ends in Mar-May. The variability in the crop development stage length is relatively low compared to mid-season and late-season lengths. Cotton has the narrowest SOS and EOS time window, spanning roughly 1-1.5 months. Rice has more uncertain SOSs and EOSs that spread across 2-3 months. Differences are observed between satellite-derived phenological dates and the equivalent dates in guidelines, with the largest difference of 38 days found in the EOS for cotton. The phenological date detection system developed in this study is highly transferable to other periods or regions to benefit the local crop management.

Keywords: *Crop phenological dates, remote sensing, Google Earth Engine, rgee package*

1. INTRODUCTION

Accurate crop phenological dates are important inputs to food production management, crop modelling and analysis of crop response to climate change (Boschetti et al., 2009; Brown et al., 2010; Studer et al., 2007). Crop phenological dates, in this study, refer to the dates of greenup, maturity, senescence and dormancy in one growing season. Conventionally, crop phenological dates in a region are obtained by interviewing farmers or using publicly available crop guidelines. Interviewing farmers can obtain accurate crop phenological dates, but the process is time consuming, leading to limited spatial coverages. Crop phenological dates from guidelines provided by regional or international agricultural organizations (e.g., Food and Agriculture Organization (FAO)) are easier to obtain, but they do not capture variability in those dates by regions or by years.

The rapid development of remote sensing technology and increasing availability of higher spatial and temporal resolution imagery provide an opportunity to estimate crop phenological dates through satellite-based phenology analysis. Time series of vegetation indices, which capture the spectral information related to the crop growth, are commonly used to estimate crop phenology metrics that can generate crop phenological dates (Araya et al., 2018; Onojeghuo et al., 2018). Vegetation indices such as the Normalised Difference Vegetation Index (NDVI) have been proved effective in detecting crop phenology in many existing studies (e.g., Boschetti et al., 2009; Pan et al., 2015; Schwartz et al., 2002). At the same time, many authors have developed robust methods to detect the key dates when crop growth moves from one stage to another, from vegetation index time series (Gocic & Trajkovic, 2013; Gu et al., 2009; Zhang et al., 2003). Most of these algorithms have been developed into software-based interfaces or packages for user-friendly applications. For example, TIMESAT, a MATLAB-based interface developed by Jönsson and Eklundh (2004), is a well-known tool to detect crop phenology. In addition, R packages, such as the phenofit (Kong et al., 2019) and CropPhenology (Araya et al., 2018) packages, have been developed recently.

While satellite-based phenology analysis has long been studied and used, improvement can be implemented to make it more efficient and practical. Firstly, most current phenology analyses are semi-automatic, requiring satellite data to be downloaded from source websites (e.g., USGS) and processed in a specific software (e.g., ENVI, ArcGIS and programming languages). The mixed use of multiple interfaces makes analyses more labour intensive, harder to manage and less efficient. Also, downloading large amounts of satellite data requires large amounts of time and storage space, which is not optimal for large-scale studies. Secondly, although many studies on detecting accurate satellite-based phenological dates exist, they seldom compare the satellite-detected phenological dates with the equivalent dates provided in the general guidelines. The guideline-based and satellite-derived phenological dates currently are applied in parallel with different practices and similarities between them are rarely discussed. Since both data are popular inputs in many crop modelling studies and analyses, it is important to understand their similarities and differences.

This study developed a fully automatic system to promote the application of satellite-based crop phenological dates identification. The system was developed in the R programming environment, and we used the "rgee" package (Aybar et al., 2020) to link to the Google Earth Engine cloud platform for satellite data acquisition and processing. The system was applied in an Australian irrigation district as a case study for three types of summer crops. Satellite-derived crop phenological dates are also compared with that in FAO-56 (Allen et al., 1998) and local crop growth guidelines (Grains Research and Development Corporation (GRDC), 2010; NSW Department of Primary Industries (DPI), 2019; NSW DPI, 2020) to understand the magnitude of difference.

2. STUDT AREA AND METHODS

The study area is the Coleambally Irrigation Area (CIA), a major irrigation district serving 79,000 hectares (ha) of irrigated land, located in south-eastern Australia (Figure 1). Crop phenological dates for three irrigated summer crops – corn, cotton and rice were analysed. Crop type information was collected in March 2021, which identified 36 corn/maize fields, 41 cotton fields and 36 rice fields to be used in this study (Figure 1). The main growing season for summer crops is from Oct to May. In the CIA, the average annual rainfall is approximately 410 mm and the annual average temperature is 24 °C (Bureau of Meteorology, 2021).

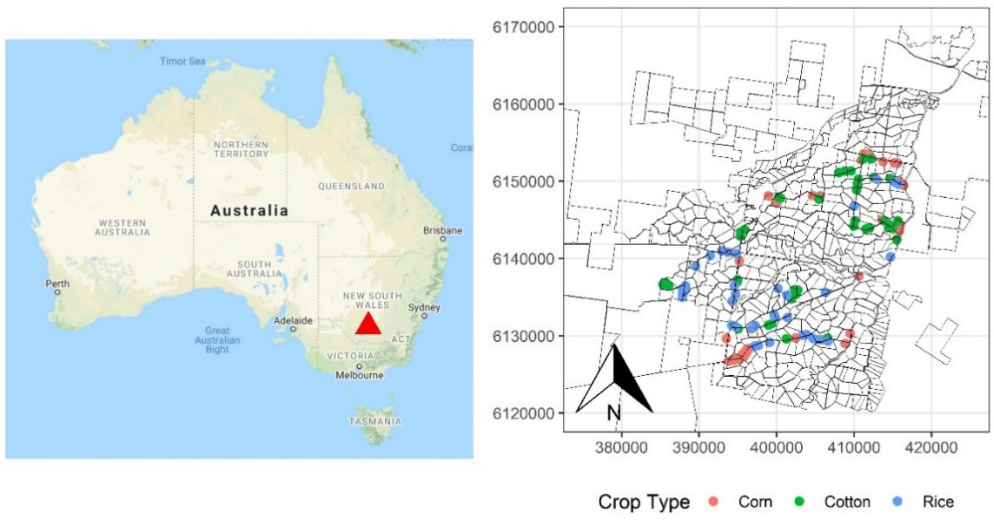

Figure 1. The left figure shows the study area in Australia and the right figure shows tested fields with black polygons indicating individual farms.

A schematic diagram of the proposed crop phenological dates identification system is presented in Figure 2. We wrote functions and used existing functions from the "rgee" package of R (Aybar et al., 2020) to send commands to Google Earth Engine (GEE) and retrieved cloud-free NDVI time series for all fields. We further used the R package "phenofit" (Kong et al., 2019) to obtain crop phenology metrics and calculated the duration of growth stages. Finally, we compared the identified lengths of growth stages and the start and end of the growing season with FAO-56 (Allen et al., 1998) and local crop growth guidelines (GRDC, 2010; DPI, 2019; DPI, 2020).

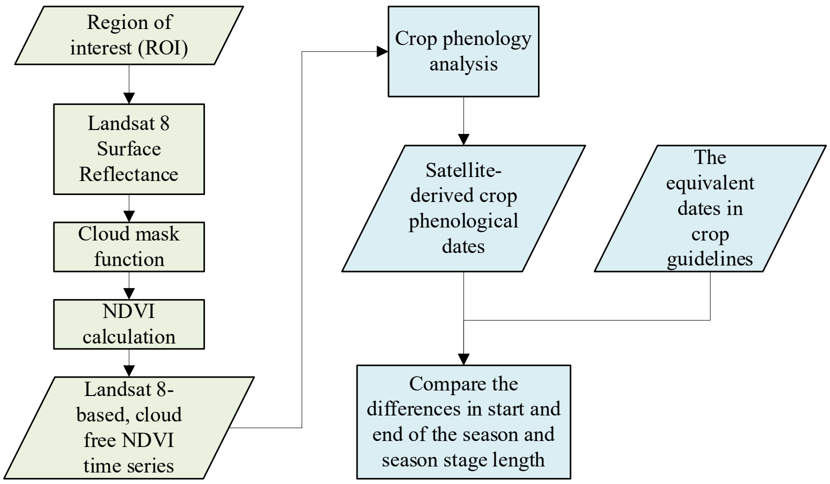

Figure 2. The schematic diagram of the proposed crop phenological dates identification system. Note: Tasks filled in green were analysed in Google Earth Engine, but commands were delivered from the R environment; Tasks filled in blue were analysed in the R environment.

2.1. Satellite data acquisition and pre-processing

Google Earth Engine (GEE) is a cloud platform that allows users to access publicly available satellite imagery easily, but it only supports JavaScript or Python. The R package "rgee" is the first bridge that links GEE with R (Aybar et al., 2020). Using this package, commands can be sent from R to GEE for satellite data acquisition and processing, which was very efficient. Landsat 8 surface reflectance data from 1 September 2020 to 30 June 2021 were extracted at the pixel level. NDVI was calculated for each pixel of each scene (NDVI time series), which was then averaged to the field level. Landsat 8 pixels were officially given the quality assessment code so we generated a cloud mask to remove pixels that are labelled as 'Cloud' or 'Cloud Shadow'. We used Landsat data for this study because CIA is located in the overlapping area of two Landsat scenes (row 92, path 84, and row 93, path 84), which effectively reduces the satellite revisit period by 50%, from 16 to 8 days. The frequent availability of satellite imagery is important for detecting accurate crop phenological dates. The resolution (30 m) of Landsat 8 data is adequate for this study given the size of fields in the CIA.

2.2. Crop phenology analysis

Crop phenology was estimated using the R package "phenofit" (Kong et al., 2019). This package contains multiple curve fitting methods and phenology metric extraction methods to extract phenology metrics from the NDVI time series. We used the algorithm developed by Zhang et al. (2003), which calculates the temporal change of the NDVI time series to detect dates for greenup, maturity, senescence and dormancy (Figure 3). The analyses included a curve fitting to the time series to reduce noise and extracted crop phenology metrics from the fitted curve. Since cloud mask was applied to data, we assumed all NDVI values in the time series were of good quality, and consequently equal weight was applied to every NDVI value for curve fitting.

2.3. Compare satellite-derived crop phenological dates with the equivalent dates in FAO-56 or local crop guidelines

We calculated durations between greenup and maturity, maturity and senescence, and senescence and dormancy from the phenology metrics of Zhang et al. (2003), and compared these three durations with the lengths of the development stage, mid-season stage and late-stage given by FAO-56 (Allen et al., 1998, Figure 4). FAO-56 suggests different lengths of these growth stages under different climatic conditions, and we used the case with the closest climate to the CIA. For example, since the CIA is in a semi-arid region, we used stage lengths for maize in the arid climate in FAO-56. We also compared the satellite-derived start of the season (SOS, same date as the greenup date) and end of the season (EOS, same date as the dormancy date) with those in the local crop guidelines (GRDC, 2010; DPI, 2019; DPI, 2020). If the crop guidelines only provide recommended sowing windows, we selected the middle date of the window as the sowing date and added 10 days backward as the guideline-based SOSs. This considers the time lag between sowing and crop emergence and makes guideline-based SOSs more equivalent to the satellited-derived SOSs. The guideline-based EOSs were calculated using SOSs and FAO-56 stage lengths (i.e., EOSs = SOSs + sum of the Dec-stage, Mid-stage and Late-stage lengths). The mode(s) (i.e., the most frequently appeared values) of satellite-derived SOSs, EOSs and stage lengths for all fields per crop type were extracted for comparison.

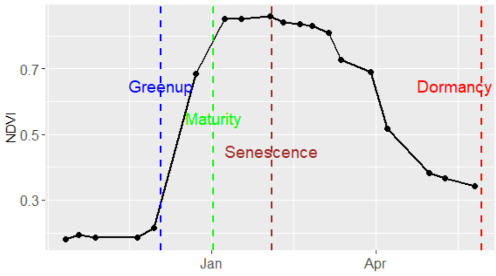

Figure 3. The growing stages in a corn (maize) field in this study (after Zhang et al., 2003)

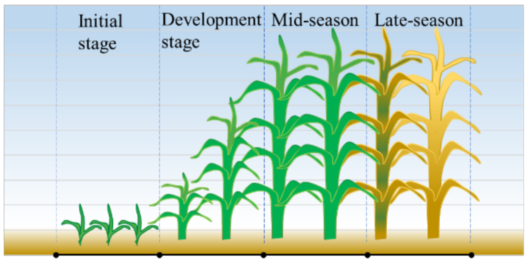

Figure 4. A schematic diagram of FAO-56 crop growth stage (after Allen et al., 1998).

3. RESULTS

Figure 5 shows the lengths of the development stage, mid-season and late-season for all crop samples derived from satellite-based crop phenology analysis. For all three crops, the distributions of the development stage length were relatively narrower, indicating that crops had consistent development lengths across the region. In addition, cotton had less variability in the length of the mid-season and late season compared to corn (maize) and rice. We then calculated the modes from distractions in Figure 5 to represent the overall stage lengths in CIA and compared them with the suggested values in FAO-56 (Table 1). Results from Table 1 showed that:

- across three stages, the estimated length of the development stage from satellite and FAO-56 has the best agreement;
- the satellite-derived method suggested longer durations for the mid-season stage compared to FAO-56 for all crops;
- for the late-season, the satellite-derived method suggested longer durations for corn (maize) and rice, but a shorter duration for cotton.

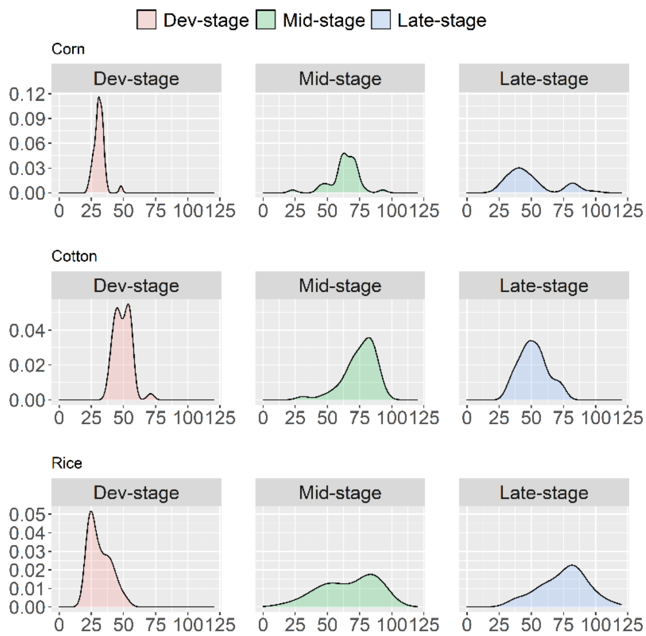

Figure 5. The stage lengths for all tested 36 corn, 41 cotton and 36 rice fields. Dev-stage = Maturity date - Greenup date; Mid-stage = Senescence date - Maturity date; Late stage = Dormancy date - Senescence date (Figure 3).

Table 1. The comparison of stage lengths from satellite-based crop phenology analysis and FAO-56 recommended values. The length difference was calculated using satellite-derived values minus FAO values, shown in days.

	Development stage			**Mid-season**			**Late season**		
	Satellite	**FAO**	**Difference**	**Satellite**	**FAO**	**Difference**	**Satellite**	**FAO**	**Difference**
Corn	31	40	-9	63	50	13	40	30	10
Cotton	54	45	9	82	60	22	50	80	-30
Rice	25	30	-5	84	55	29	82	40	42

Figure 6 shows the SOS and EOS distributions for three crops derived from the satellite-based phenology analysis method. Cotton had a narrow distribution of SOS - starting dates were concentrated within one month from mid-November to mid-December, indicating that most farmers grew cotton within a short time window in the district. This also led to a narrow harvesting time window (EOS) in May. In contrast, rice had relatively flexible SOS windows which spread across more than two months. Consequently, the distribution of its EOS was also widely spread. Table 2 suggested that satellite-derived SOS and EOS were not agreed very well with that in crop growth guidelines in NSW (GRDC, 2010; DPI, 2019; DPI, 2020). The largest difference occurred in the EOS for cotton, which is 38 days.

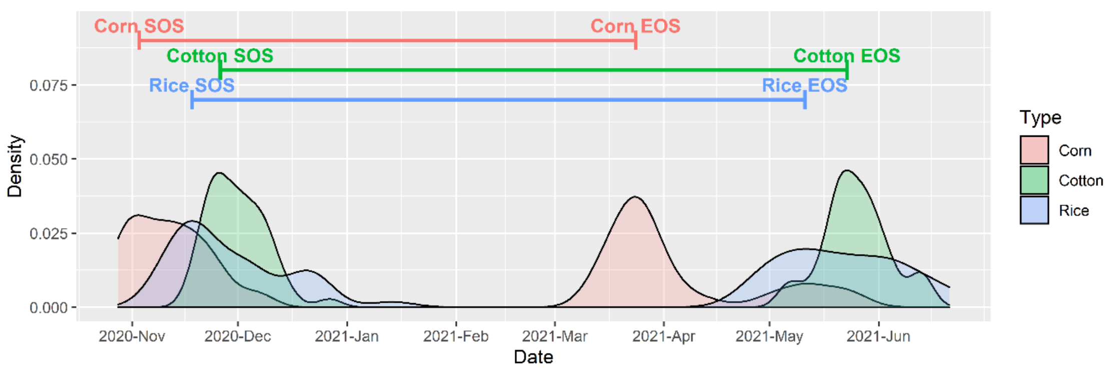

Figure 6. The distributions of SOSs and EOSs for three crops, and the district-level season lengths (EOS – SOS) which were calculated using the modes of their distributions.

Table 2. The difference in SOSs and EOSs from the satellite-derived method and the local crop guideline.

	Satellite SOS	Guideline SOS	SOS difference in days	Satellite EOS	Guideline EOS	EOS difference in days
Corn	03-Nov-20	05-Nov-20	-2	24-Mar-21	28-Feb-21	24
Cotton	26-Nov-20	01-Nov-20	25	23-May-21	15-Apr-21	38
Rice	18-Nov-20	10-Nov-20	8	11-May-21	09-Apr-21	32

4. DISCUSSION AND CONCLUSION

Crop phenological dates are important data sources in agricultural management but are often found to be poorly managed or not available in a cropping district (Whitcraft et al., 2015). The satellite-based crop phenology analysis method provides an easy way to obtain comprehensive crop phenological dates. Satellite-derived phenological dates can serve as important inputs for government/agricultural census to understand the agricultural activities at a district level and variability across fields. Compared to FAO-56 (Allen et al., 1998) or local crop guidelines (GRDC, 2010; DPI, 2019; DPI, 2020), the satellite-derived crop phenological date has the following advantages:

1) **Satellite-derived phenological dates are spatially detailed.** Satellite-derived phenological dates are derived from individual fields in the study region, so they are more representative of the local condition. In contrast, FAO-56 or local crop guidelines are more general e.g., the stage lengths for individual crops in FAO-56 come from different studies that were not necessarily conducted in south-eastern Australia. For example, the large difference in late-season length for rice (i.e., 42 days, Table 1) could be due to the difference in climate between our study region and the FAO case studies i.e., CIA has a temperate semi-arid climate, but FAO-56 only provides data for tropical and Mediterranean climate conditions. On the other hand, while guideline-based phenological dates may be informative for a large spatial extent (e.g., southern NSW), they are often too coarse for precise agriculture monitoring and modelling at the farm or field level. In contrast, satellite-derived crop phenological dates can often be derived at sub-field, field or farm scale. The satellite-derived phenological dates are spatially detailed and thus more appliable for local water and crop management.

2) **Satellite-derived phenological dates give insights into detailed within-district variability.** Understanding the variability of crop phenological dates is important for district-level water and crop production management (Zeng et al., 2020). Crop phenological dates can vary within an irrigation district in response to various farm management practices. For example, our study shows that the spatial variability of SOS/EOS can be as much as 2-3 months for the same crop type in the same season (Figure 6). This variability also changes from crop to crop. For example, we found that SOS/EOS for cotton fall within a narrower time window than for rice and corn/maize (Figure 6). More importantly, we can gain insights into the overall dynamics of calendars by looking at their SOS/EOS distributions. For example, we found that the distribution of rice SOS has two peaks, indicating that farmers concentrated on two main periods to grow rice (Figure 6). While crop guidelines may provide rough ranges of the SOS/EOS, they do not show the detailed distributions of samples within those ranges. The variability of phenological dates or patterns of variability can be further related to potential drivers such as farm characteristics, soil, fertilizer, or variety to gain more useful information for crop management.

3) **Satellite-derived phenological dates are applicable for long-term analysis.** Satellite data, such as Landsat data (30m) used in this study, are publicly available since the 1980s with global coverage. Hence, crop phenological dates using Landsat series data can be analysed over multi-decade periods. The availability of multi-decade phenological dates allows scientists to understand the variation of growth patterns under interannual variability in climatic conditions. One challenge for multi-decade studies is the efficiency of application. Large numbers of satellite imageries need to be downloaded, processed and analysed. Google Earth Engine (GEE) provides an opportunity to process satellite data on the cloud, easing the challenge of dealing with large datasets. We develop a comprehensive system that combines GEE functions with phenology analysis in one platform (the R environment).

While the satellite-derived method is convenient, efficient and appliable for large scale, it is not 100% accurate and further analysis is required to collect ground measurements over a wide range of conditions and crop types to further calibrate and validate the satellite-based crop phenology models (Zeng et al., 2020). Currently, ground-measured crop phenological dates are lacking. One option to improve ground-truth data would be a phenology observation network using digital cameras (Zeng et al., 2020).

REFERENCES

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. *Fao, Rome, 300*(9), D05109.

Araya, S., Ostendorf, B., Lyle, G., & Lewis, M. (2018). CropPhenology: An R package for extracting crop phenology from time series remotely sensed vegetation index imagery. *Ecological informatics, 46*, 45-56.

Aybar, C., Wu, Q., Bautista, L., Yali, R., & Barja, A. (2020). rgee: An R package for interacting with Google Earth Engine. *Journal of Open Source Software, 5*(51), 2272.

Boschetti, M., Stroppiana, D., Brivio, P., & Bocchi, S. (2009). Multi-year monitoring of rice crop phenology through time series analysis of MODIS images. *International Journal of Remote Sensing, 30*(18), 4643-4662.

Brown, M. E., de Beurs, K., & Vrieling, A. (2010). The response of African land surface phenology to large scale climate oscillations. *Remote Sensing of Environment, 114*(10), 2286-2296.

Bureau of Meteorology. (2021). *Historical weather observations and statistics*. Retrieved 22 September 2021, from http://www.bom.gov.au/climate/data-services/station-data.shtml.

Gocic, M., & Trajkovic, S. (2013). Analysis of changes in meteorological variables using Mann-Kendall and Sen's slope estimator statistical tests in Serbia. *Global and Planetary Change, 100*, 172-182.

Grains Research and Development Corporation. (2010). *Growing cotton in SNSW - Considerations and information*. https://grdc.com.au/resources-and-publications/grdc-update-papers/tab-content/grdc-update-papers/2010/08/growing-cotton-in-snsw-considerations-and-information.

Gu, L., Post, W. M., Baldocchi, D. D., Black, T. A., Suyker, A. E., Verma, S. B., . . . Wofsy, S. C. (2009). Characterizing the seasonal dynamics of plant community photosynthesis across a range of vegetation types. In Phenology of ecosystem processes (pp. 35-58): Springer.

Jönsson, P., & Eklundh, L. (2004). TIMESAT—a program for analyzing time-series of satellite sensor data. *Computers & geosciences, 30*(8), 833-845.

Kong, D., Xiao, M., Zhang, Y., Gu, X., Cui, J., & Kong, M. D. (2019). Package 'phenofit'.

NSW Department of Primary Industries. (2019). *Summer crop management guide 2019*. https://www.dpi.nsw.gov.au/__data/assets/pdf_file/0011/1187750/SCMG-web-FINAL-5Nov.pdf

NSW Department of Primary Industries. (2020). *Rice variety guide 2020–2021*. https://www.dpi.nsw.gov.au/__data/assets/pdf_file/0009/1248309/RVG-2020-2021-web.pdf

Onojeghuo, A. O., Blackburn, G. A., Wang, Q., Atkinson, P. M., Kindred, D., & Miao, Y. (2018). Rice crop phenology mapping at high spatial and temporal resolution using downscaled MODIS time-series. *GIScience & remote sensing, 55*(5), 659-677.

Pan, Z., Huang, J., Zhou, Q., Wang, L., Cheng, Y., Zhang, H., . . . Liu, J. (2015). Mapping crop phenology using NDVI time-series derived from HJ-1 A/B data. *International Journal of Applied Earth Observation and Geoinformation, 34*, 188-197.

Schwartz, M. D., Reed, B. C., & White, M. A. (2002). Assessing satellite‐derived start‐of‐season measures in the conterminous USA. *International Journal of Climatology: A Journal of the Royal Meteorological Society, 22*(14), 1793-1805.

Studer, S., Stöckli, R., Appenzeller, C., & Vidale, P. L. (2007). A comparative study of satellite and ground-based phenology. *International Journal of Biometeorology, 51*(5), 405-414.

Whitcraft, A. K., Becker-Reshef, I., & Justice, C. O. (2015). Agricultural growing season calendars derived from MODIS surface reflectance. *International Journal of Digital Earth, 8*(3), 173-197.

Zeng, L., Wardlow, B. D., Xiang, D., Hu, S., & Li, D. (2020). A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sensing of Environment, 237*, 111511.

Zhang, X., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C., Gao, F., . . . Huete, A. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment, 84*(3), 471-475.