remote sensing

MDPI

Article

Enhancing the Accuracy and Temporal Transferability of Irrigated Cropping Field Classification Using Optical Remote Sensing Imagery

Zitian Gao *, Danlu Guo, Dongryeol Ryu and Andrew W. Western

Department of Infrastructure Engineering, The University of Melbourne, Parkville, VIC 3010, Australia; danlu.guo@unimelb.edu.au (D.G.); dryu@unimelb.edu.au (D.R.); a.western@unimelb.edu.au (A.W.W.)
* Correspondence: zitiang@student.unimelb.edu.au

Section S1. Field Boundary Delineation

Field boundary delineation was completed using a very-high-resolution satellite image (one scene of a Planetscope image (3 m resolution) obtained on 30 August 2019) in the Feature Extraction (FX) module in ENVI software. In this study we used texture information to optimize the segmentation parameters. Such optimization is required because, by trial and error, we found that segmentation accuracy is mainly affected by the scale level parameter in the FX module. The challenge is that the best scale level values for individual fields are different, which means using a uniform scale level over the entire region is impractical. Besides, the best scale level values for individual fields show large spatial variability within a small region (e.g., fields within one farm may have different preferences of parameters), making it hard to divide the image into subregions through similar parameter values and apply segmentation separately.

To solve this problem we adopted an idea from Xu et al. [40], who suggested using texture information to inform segmentation parameter selection. The processes to select the best scale level values for individual fields are summarized as follows:

1. We developed four sets of parameters with scale level within a feasible range (scale level = 50, 60, 70 and 80, respectively). Another key parameter, merge level, was fixed to 90, and the rest of the parameters were fixed to the default values in the FX module. These four sets of parameters led to four different segmentation layers.
2. We calculated the texture variables—gray-level co-occurrence matrices (GLCMs)—for individual fields for four different layers separately. The GLCM contains variables such as the mean, variance, homogeneity and entropy. The calculations were completed based on the Planetscope image using the "glcm" package in R [54].
3. We selected a small set of testing samples (fields) to explore the relationship between GLCM variables and scale level. Specifically, we visually selected approximately 50 to 60 polygons with satisfactory segmented boundaries from each segmentation layer. The scale level values of these testing samples were plotted against their corresponding GLCM variables. Among all the GLCM variables, we found that GLCM variance has a strong positive correlation with scale level (Figure S1), indicating its capability to inform the best scale level values for individual fields.
4. We used the upper and lower bounds of the GLCM variance boxplot (Table S1) as thresholds to inform the best scale level for individual fields. This allows each field to obtain its segmented boundary from one of the four segmentation layers based on its field-averaged GLCM variance. The boundaries of individual fields were finally merged into a field boundary layer.

5. Although texture information helped to improve segmentation accuracy, there was still a small number of fields (approx. 10%–15%) which had boundaries that could not be delineated accurately.

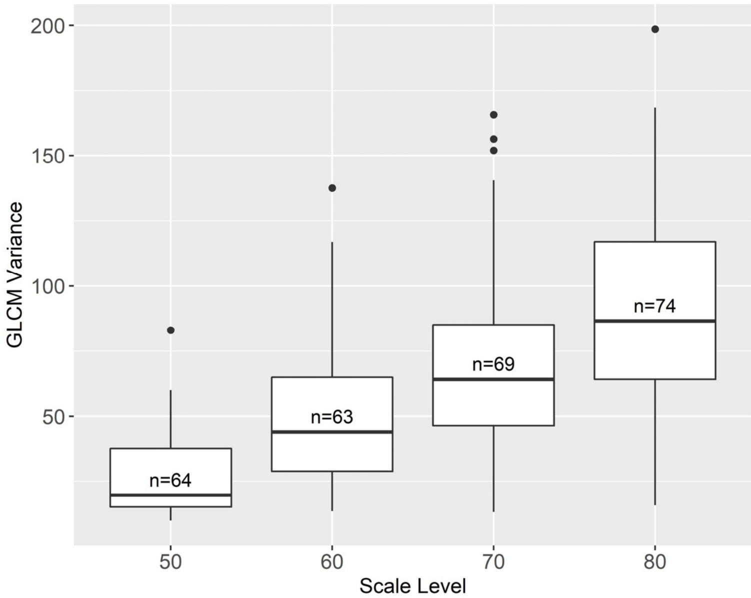

Figure S1. The boxplots of GLCM variance values for testing samples selected from four segmentation layers (scale level = 50, 60, 70 and 80, respectively).

Table S1. The rules to select the best scale level using GLCM variance.

GLCM Variance	Scale Level
<37	50
37–66	60
66–85	70
>85	80

Manual corrections were applied after automatic segmentation. The final field boundary layer for the whole district is shown in Figure S2.

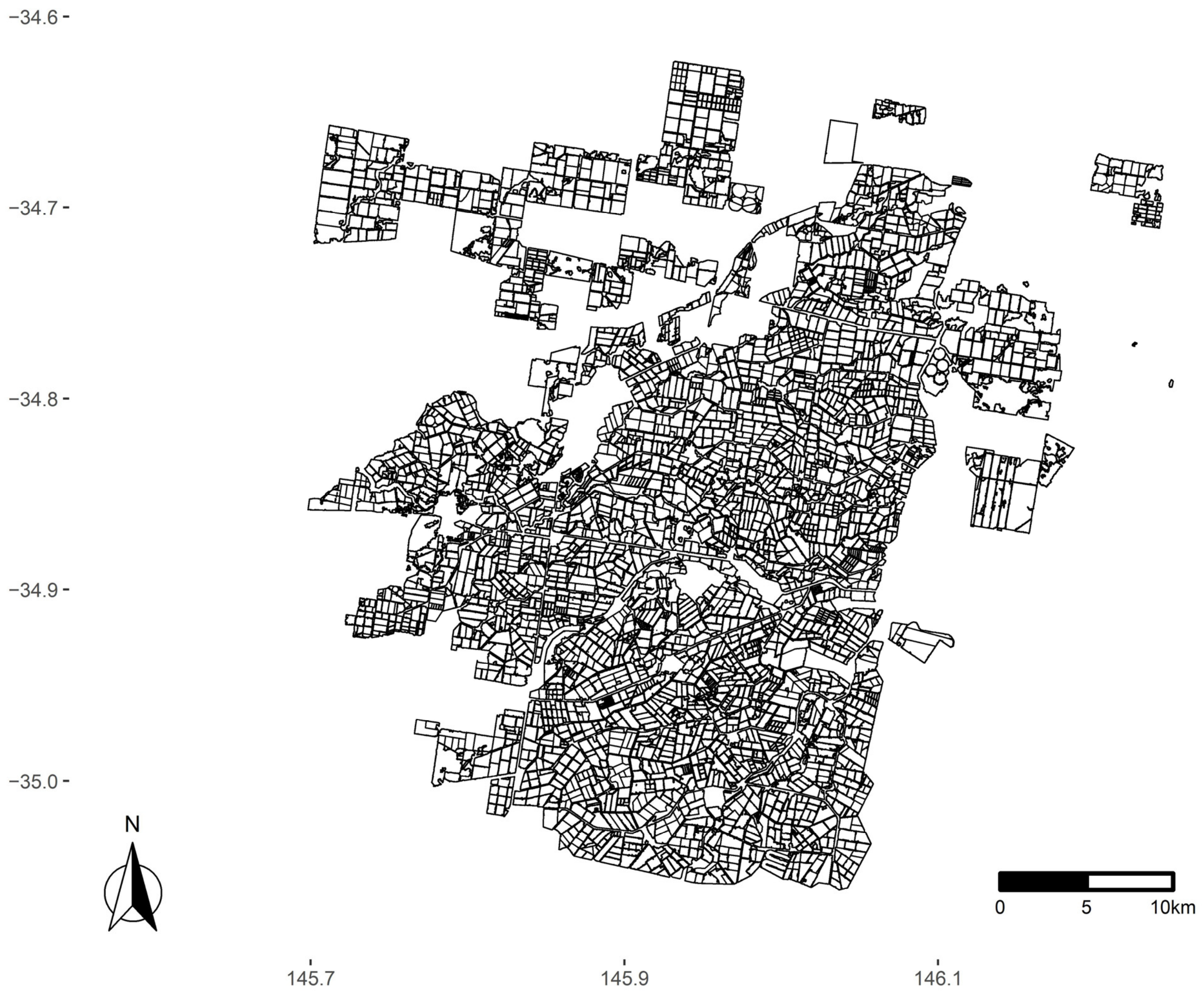

Figure S2. The final field boundary layer for the CIA.

Section S2. Ground-Truth Samples Collection and NDVI Metrics Used in *Method Two*

To visually interpret static training samples we first downloaded the monthly false-color composites (band 5-4-3) at the peak cropping season of summer and winter (January and August). We randomly selected fields that are displayed in bright red as "irrigated cropping and pasture" samples and fields that are displayed in green as "bare soil" samples. Fields with a transit color between red and green were considered as rainfed crops, some types of perennial crops or fallow land with weeds, which were marked as "unknown". NDVI time series were used to further confirm the class of marked samples. Initial "irrigated cropping and pasture" samples that maintain a high NDVI without fluctuation across the season were considered as perennial plantations; therefore, they were remarked as "perennial plantation".

Dynamic samples were drawn from Google Earth imagery according to the unique features of nonirrigated grazing land and forests (Figure S3). This was achieved by drawing polygons over homogenous grazing land and forest areas.

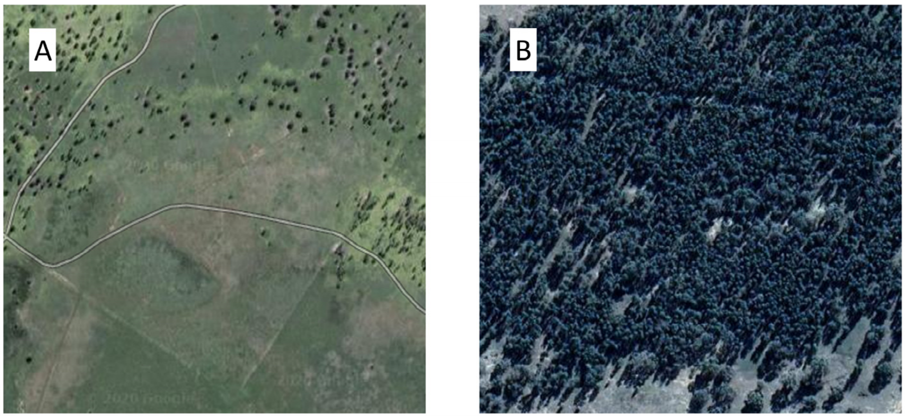

Figure S3. Landscape characteristics of nonirrigated grazing land (A) and wild forest (B) on Google Earth imagery.

Table S2 shows the number of validation samples in each class.

Table S2. The number of validation samples in each class.

		Irrigated Cropping/Pasture	Bare Soil	Nonirrigated Grazing Land	Forest	Perennial Plantations	Unknown
Summer	**2011–2012**	97	202	-	-	13	41
	2012–2013	118	194	-	-	9	23
	2013–2014	80	210	-	-	11	20
	2014–2015	98	202	-	-	9	27
	2015–2016	70	228	-	-	8	45
	2016–2017	114	209	-	-	11	30
	2017–2018	108	164	-	-	10	67
Winter	**2011–2012**	63	101	46	61	3	22
	2012–2013	111	30	46	61	1	41
	2013–2014	103	55	46	61	3	49
	2014–2015	92	29	46	61	3	39
	2015–2016	47	41	46	61	3	58
	2016–2017	110	32	46	61	3	53
	2017–2018	51	78	46	61	3	28

The detailed descriptions of NDVI metrics used in *method two* (random-forest-based method) are shown in Table S3.

Table S3. NDVI metrics used in *method two*.

Feature Name	Description
MAX	The maximum NDVI value across a season[1]
RANGE	The maximum NDVI minus the minimum NDVI. Note: if the minimum NDVI is less than 0.2, use 0.2
MIN	The minimum NDVI value across a season [1]
Monthly_NDVI	Raw values of monthly maximum NDVI
Average_max	The average of the first three max NDVI values within a season [1]

Monthly_above_thres	A series of logical numbers to describe if the monthly NDVI value is above the threshold (0.6). If the value is greater than the threshold, returns 1; otherwise, 0
Quantile_0.8_time	One value that indicates the month(s) when the NDVI is above the 0.8 quantile of the total monthly NDVI across a season. The month is described using an index (e.g., in winter classification, Mar = 1, Apr = 2… Nov = 9). If multiple months are selected, use the mean of the index
Quantile_0.2_time	One value that indicates the month(s) when the NDVI is below the 0.2 quantile of the total monthly NDVI across the season. The month is described using the index (e.g., in winter classification, Mar = 1, Apr = 2… Nov = 9). If multiple months are selected, use the mean of the index
Peak_Over_0.8	Number of months when the monthly NDVI is greater than 0.8
Peak_Over_0.7	Number of months when the monthly NDVI is greater than 0.7
Peak_Over_0.6	Number of months when the monthly NDVI is greater than 0.6
Peak_Over_0.5	Number of months when the monthly NDVI is greater than 0.5
Largest_growth_rate	The maximum NDVI growth rate, which is defined as the difference of two neighboring monthly NDVI values ($NDVI_{i+1}$-$NDVI_i$)
Diff_2_3	The difference between the second largest monthly NDVI and the third largest monthly NDVI

[1] Maximum NDVI is extracted from December to February for summer and from June to October for winter.

Section S3. Gaussian Mixture Model (GMM) Results

Figure S4 and Figure S5 show the GMM clustering results for summer and winter, respectively. The cluster number (*k*) is between 2–9. The analyses were completed using the "mclust" package in R [46].

Figure S4. The Gaussian mixture model clustering with *k* components on the seasonal maximum NDVI in summer. *k* is determined by Bayesian inference criteria (BIC).

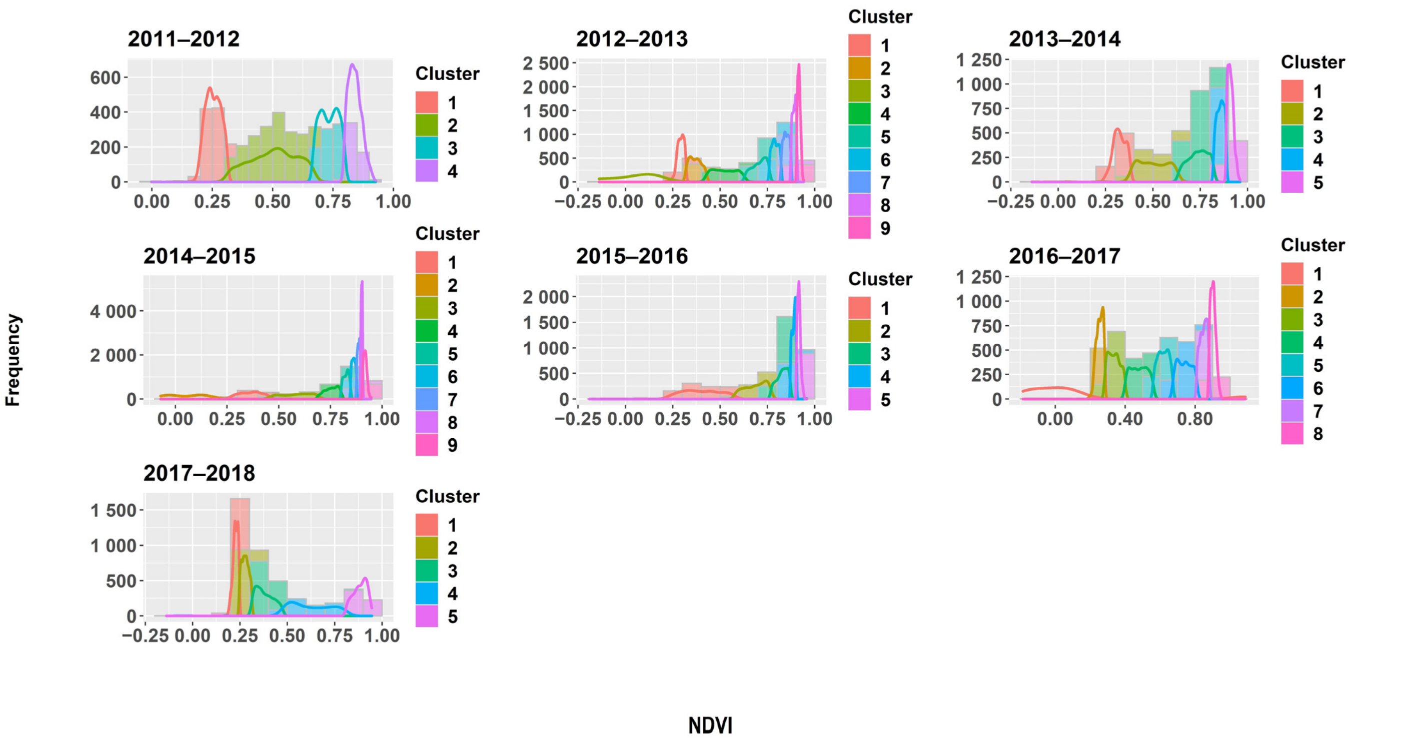

Figure S5. The Gaussian mixture model clustering with *k* components on the seasonal maximum NDVI in winter. *k* is determined by Bayesian inference criteria (BIC).

Section S4. Classification Maps

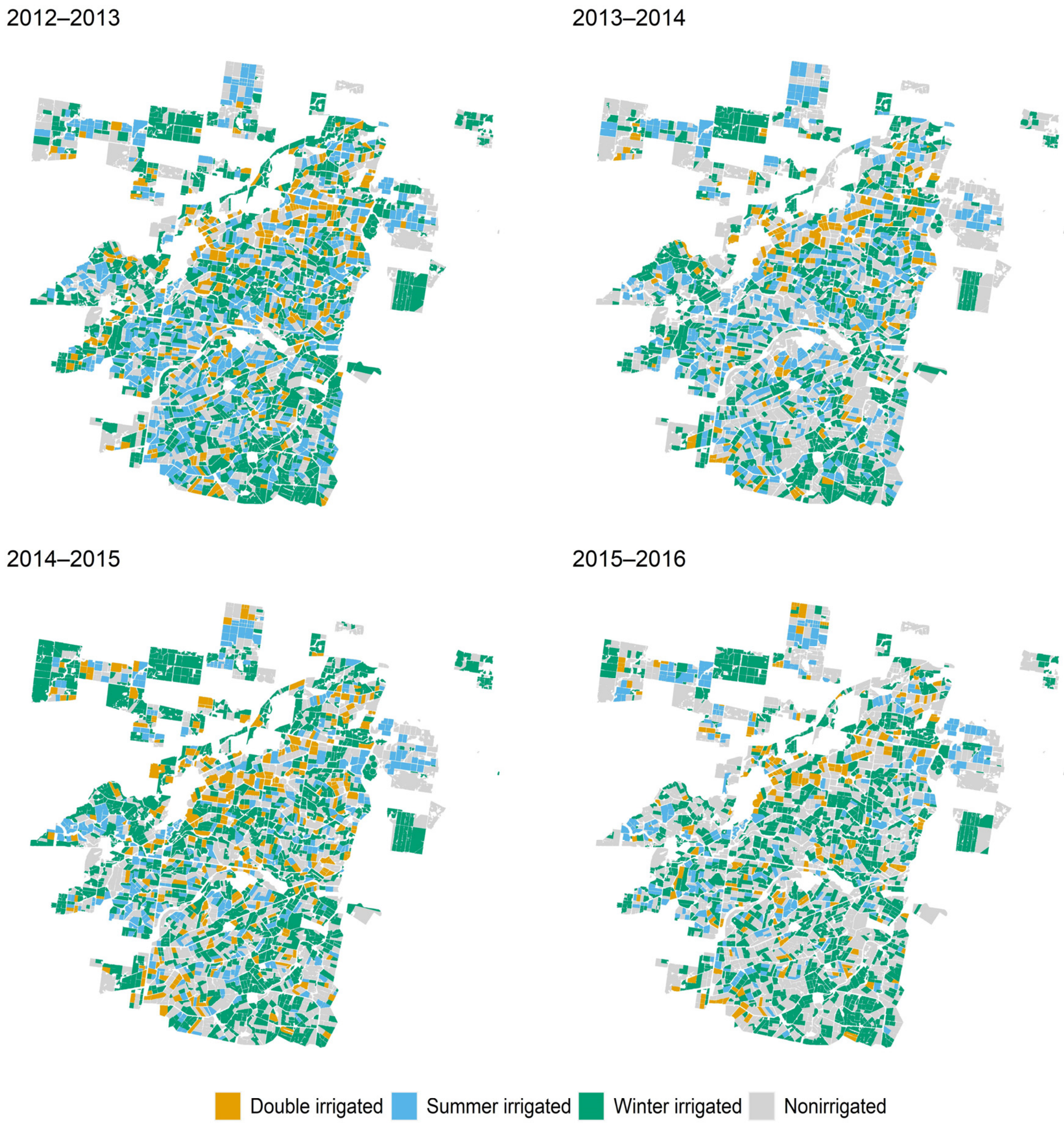

Figure S6. Classified irrigated field mapping from 2012–2013 to 2015–2016.