J. Geogr. Sci. 2018, 28(4): 429-444
DOI: https://doi.org/10.1007/s11442-018-1482-0

Spatio-temporal differences in cloud cover of Landsat-8 OLI observations across China during 2013–2016

XIAO Chiwei[1,2,*], LI Peng[1,2,3], FENG Zhiming[1,2], WU Xingyuan[3]

1. Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China
2. College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China
3. Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, Nanchang 330022, China

Abstract: Currently, the historical archive images of Landsat family sensors are probably the most effective data products for tracking global longitudinal changes since the 1970s. However, the issue of the degree and extent of cloud coverage is always a challenge and varies distinctively worldwide. So far, acquisition probability (AP) analyses of cloud cover (CC) of Landsat observations have been conducted with different sensors at regional scale. To our knowledge, CC probability analysis for the newly-launched Landsat-8 Operational Land Imager (OLI) across China is not reported. In this paper, monthly, seasonal, and annual APs for Landsat OLI (44,228 in total) images over China acquired from April 2013 to October 2016 with various CC thresholds were analyzed. The results showed that: first, the cumulative average APs of all OLI data over China at the CC thresholds ⩽30% was about 49.6% which illustrated the availability of OLI imagery across China. Second, the spatial patterns of 10%, 20%, and 30% CC thresholds of OLI observations, coincided well with the precipitation distributions separated by the respective 200 mm, 400 mm, and 800 mm isohyetal lines. Third, the APs of images with the 30% CC threshold are the highest in autumn and winter especially in October of 58.7%, while the corresponding lowest probability occurred in June of 41.0%. Finally, the spatial differences in APs of targeted images with ⩽30% CC thresholds were quite significant. At regional scales, the arid and semi-arid areas, Inland River and Songliao River basins, and northwestern side of the Hu Huanyong population line had the larger probabilities of obtaining high-quality images. Our study suggested that OLI imagery satisfy the data requirements needed for land surface monitoring, although there existed obvious spatio-temporal differences in APs over China at the 30% CC threshold.

Keywords: cloud cover (CC); spatio-temporal changes; Landsat-8 OLI; acquisition probability (AP); China

Received: 2017-06-15 **Accepted:** 2017-11-16

Foundation: National Natural Science Foundation of China, No.41430861; National Key Research and Development Program of China, No.2016YFC0503500; Opening Fund of Key Laboratory of Poyang Lake Wetland and Watershed Research, Ministry of Education, Jiangxi Normal University, No.PK2016004

Author: Xiao Chiwei (1991–), PhD Candidate, specialized in resource utilization and remote sensing of land resource. E-mail: xiaocw@igsnrr.ac.cn

[*]**Corresponding author:** Li Peng (1984–), PhD and Associate Professor, E-mail: lip@igsnrr.ac.cn. ORCID ID (https://orcid.org/0000-0002-0849-5955)

1 Introduction

China has experienced rapid land use/land cover changes (LUCC) in ecotones, fragile ecological regions (Liu *et al*., 2016; Ren and Wang, 2017), and metropolitan areas (Liu *et al*., 2014; Wu *et al*., 2016) in the past decades. Indeed, the vast territory, various climates and diverse topography in China, are conducive to a variety of LUCC. However, before detecting LUCC, and its characteristics, rates, drivers, and impacts (Hao and Ren, 2009; Liu and Deng, 2010), the fundamental issue of the "observability" of remote sensing data needs to be evaluated (Foley *et al*., 2005), as well as understanding its applicability and limitations, especially historical Landsat data.

Currently, remote sensing involving optical, microwave radar, and lidar satellites, or combinations thereof has become an effective tool to monitor land surface changes and ecosystem dynamics (Aplin, 2004; Jin *et al*., 2013b; Teillet *et al*., 2007). Of these, the freely available Landsat products have become an increasingly widespread imagery source for agricultural, ecological, and environmental monitoring (Jin *et al*., 2013b; Li *et al*., 2014; Qin *et al*., 2007), although they are usually impacted due to contamination of frequent cloud cover (CC) and cloud shadows (Dong *et al*., 2013; Hansen and Loveland, 2012; Li *et al*., 2012; Liu and Liu, 2013). Historical 30-m resolution data of Landsat-4/5 Thematic Mapper (TM, 1982–2013), Landsat-7 Enhanced Thematic Mapper Plus (ETM+, 1999–), and Landsat-8 Operational Land Imager (OLI, 2013–) are important and extensive data sources of Earth observation since the late 1980s, covering geo-information, natural vegetation, agricultural crops, and land cover types (Hansen and Loveland, 2012; Jin *et al*., 2013a; Steven *et al*., 2003). However, the most obvious challenge for their application is CC and cloud shadows which are widely reported in Landsat-derived remote sensing analyses (Asner, 2001; Ju and Roy, 2008; Hansen and Loveland, 2012; Liu *et al*., 2016; Whitcraft *et al*., 2015), particularly in the subtropics and tropics (Asner, 2001; Li *et al*., 2016). Consequently, some studies have attempted to address the issue of CC, such as the use of cloud removal algorithms to exclude CC areas or replace them with other images (Ju and Roy, 2008; Zhu and Woodcock, 2012). Nevertheless, these methods tend to result in some errors due to the variability and complexity in cloud types (Kovalskyy and Roy, 2015), and usually require time- and labor-intensive work.

Recently, a few studies have indicated that Landsat images are vulnerable to CC and cloud shadows during the dry season in many regions of tropical, sub-tropical, and temperate monsoon areas (Dong *et al*., 2013; Li *et al*., 2016). However, one awkward but unanswered question is what is the acquisition probability (AP) of usable images for the past nearly 40 years with CC less than 30% over China from the Landsat-like optical data, as China has undergone noticeable land use/cover dynamic changes since the 1980s (Liu *et al*., 2014). Unfortunately, the problem of the 'observability' of Landsat family sensors (i.e., TM, ETM+, and OLI) still unaddressed globally (Goward *et al*., 2006; Sano *et al*., 2007). Thus, quantifying CC of Landsat images is a paramount prerequisite for monitoring land surface change at various regional scales (Asner, 2001; Li *et al*., 2017), especially in the monsoon regions (e.g., China). Currently, there are only a few published studies concerning the probability analysis of CC in different Landsat sensors throughout the Amazon Basin (Asner, 2001; Sano *et al*., 2007), the availability of cloud-free images covering the conterminous

United States and globally (Ju and Roy, 2008; Kovalskyy and Roy, 2013; Kovalskyy and Roy, 2015), and the probability differences of CC in Mainland Southeast Asia (MSEA) (Laborde *et al*., 2017; Li *et al*., 2017). Unlike cloud avoidance strategy of Landsat-5/7 Long-term Acquisition Plan (LTAP) which does not collect all images outside the US (Arvidson *et al*., 2001; Asner, 2001), Landsat-8 provides global coverage of observation since 2013. Currently, APs analysis of CC of Landsat-8 observation is not fully investigated worldwide (Laborde *et al*., 2017).

To our knowledge, a comprehensive analysis of cloud coverage for the newly launched Landsat-8 OLI sensor (February 11, 2013) data over China has not been published to date. In this study, we used the CC information of all OLI historical metadata (a total of 44,228 scenes) acquired over China from January 2013 to October 2016 to evaluate the monthly, seasonal, and annual APs, and the limitation and suitability for such as terrestrial surface studies. The objectives of this study are twofold: (1) to understand and evaluate the average APs of Landsat OLI with different CC thresholds in China; and (2) to analyze the spatial differences in CC of Landsat OLI images among major river basins or grain production regions, dry and wet regions, and the two regions divided by the Hu Huanyong population line (Hu Line) in China. This study contributes to providing necessary guidance in the aspect of Landsat-8 OLI data source for monitoring land surface changes.

2 Materials and methods

2.1 Study area

China has a latitude span of 50 degrees from north to south and a longitude span of approximately 62 degrees from west to east (Figure 1) (Ge *et al*., 2016). Large spans of latitude and longitude generally lead to significant differences in precipitation, resulting in dry and wet regions (Figure 1), respectively. China has many different types of climate, such as subtropical, temperate monsoon, and temperate continental climate. There are four distinct seasons, namely spring from March to May, summer from June to August, autumn from September to November, and winter from December to February of the following year. In winter, northerly winds from high-latitude regions are cold and dry, while southerly winds from coastal areas at lower latitudes are warm and moist in summer. Among them, the cold-dry season is characterized by low CC with less precipitation per month, which facilitates obtaining high-quality images, or cloud-free or little cloudy images.

Landforms in China include large mountains and plateaus in the west and north, and low hills and alluvial plains in the southern and eastern areas. Northwestern China has very little precipitation because it is located deep inland far from the oceans. However, southwestern China has a distinction between the rainy season (from May to October) and the dry season (from November to April) due to the seasonal changes of Indian Ocean monsoon (Li *et al*., 2017). Agriculture in northwestern China is greatly determined by the accessibility of irrigation. China's northern plateau region is dominated by grasslands. However, ranchers who seek for short-term profit have caused overgrazing of the grassland. China's eastern and southern lowlands (i.e., low hills and alluvial plains) comprise the major agricultural zones and river basins (e.g., Yangtze River and Yellow River) (Jin *et al*., 2016). This region is also characterized by agriculture for major rice-producing area (Ding, 1961) with a larger popu-

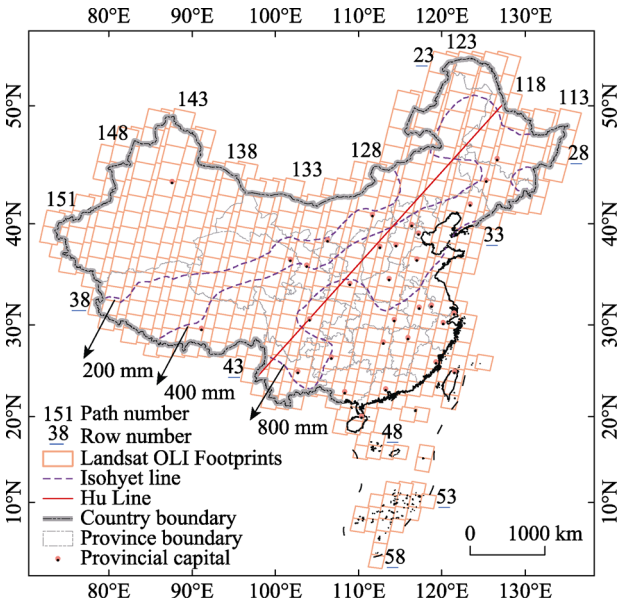

Figure 1 Map showing the location of the dry and wet areas, and the two sides divided by the Hu Line in China, and the 576 Landsat OLI footprints (path/row) over China

which is a primary form of LUCC in this region.

2.2 Landsat OLI CC data and pre-processing

There are 576 path and row (PR) coverage frames (or footprints, e.g., 121/040) of Landsat-8 OLI satellite over China according to the Worldwide References System (http://landsat.usgs.gov/tools_wrs-2_shapefile.php). In particular, all Landsat OLI footprints are required to cover China's islands (Figure 1). The PR information and all Landsat-8 OLI metadata, including cloud coverage (0–100%) and its levels (0–9), acquisition and processing date, and geographic location (e.g., scene center's latitude and longitude) were collected from the USGS Landsat inventories provided by the Landsat Bulk Metadata Service (http://landsat.usgs.gov//consumer.php). With regard to the cloud coverage, we only used the CC percentage average value in the whole Landsat OLI scene in this study. The percentage of cloud coverage of Landsat 8 cloud cover assessment (CCA) uses multiple algorithms to detect clouds in scene data (USGS, 2016), including automatic CCA (ACCA), See-5 CCA, Cirrus CCA, and Artificial Thermal-A CCA.

A total of 44,228 Landsat OLI scenes were acquired over China from April 2013 to October 2016. There were 8121, 12,671, 13,086, and 10,350 scenes in 2013, 2014, 2015, and 2016, respectively (Table 1). It should be noted that the Landsat-8 OLI satellite has acquired all the images in China, including the Diaoyu island and other surrounding islands in South China Sea (Figure 1). The thresholds of cloud coverage range from 0 to 100% in these Landsat metadata files, which were divided into ten levels (0–9) at 10% intervals by location (PR) (Asner, 2001). Note that level 0 represents the CC of all scenes equal to 10% or less (or $0\% \leqslant CC \leqslant 10\%$), and level 1 represents the range from 10% to 20% (including 20%) (or $10\% < CC \leqslant 20\%$). Following this categorization, the last level 9 represents greater than

lation density.

With rapid economic development, LUCC is currently a very common phenomenon, especially in southern and eastern China (Liu *et al.*, 2014). Several factors are driving this trend, including an enhancement of the impact of the long-standing Hu Line (Figure 1), which marks the striking difference in the distribution of China's population. The population and economy flow into the southeastern part of the Hu Line due to better biophysical and social conditions. Meanwhile, series of serious ecological problems have occurred in northern and northwestern China due to overgrazing and other unsuitable farming practice in recent decades,

90%, but less than or equal to 100% (or 90% < CC ⩽ 100%). CC thresholds are referred as 0%, 10% to 100% or level 0, level 1 to level 9 in this study. In this study, the percentage information of CC in Landsat OLI sensor data was analyzed for all path and row combinations for April 2013 to October 2016 over China.

Table 1 Annual Landsat OLI scenes acquisition statistics in China including cloud cover (CC) levels (0–9) in each footprint (path/row)

Year	Number of scenes	Proportion (%)	Number of CC level									
			0	1	2	3	4	5	6	7	8	9
2013	8121	18.4	2546	948	705	634	602	631	608	654	571	222
2014	12671	28.6	3712	1561	1185	957	868	884	915	988	1023	578
2015	13086	29.6	3655	1570	1168	962	951	906	1016	1153	1152	553
2016	10350	23.4	2712	1228	945	808	816	774	843	952	878	394

2.3 AP calculation of different CC thresholds

The monthly, seasonal and annual APs of CC thresholds ranging from level 0 (or 0% ⩽ CC ⩽ 10%) to level 9 (or 90% < CC ⩽ 100%) were computed over China at the spatial scale of each Landsat footprint (Path/Row, 576 in total) (Figure 1). Firstly, cumulative probabilities of different CC thresholds were calculated to show the appropriate CC threshold. Secondly, the seasonal cumulative probabilities of increasing CC (i.e., 10%, 20% and 30%) were calculated to display the regional and intermonth variations. Then, we further delineated monthly average probability of CC for Landsat sensors with ⩽ 30%. Lastly, the average APs of regional were calculated based on the above results. The probability of CC for a successful acquisition was calculated using formula (1) with the Microsoft Visual Basic application (Asner, 2001; Li *et al.*, 2017). The metadata archive of cloud coverage from the 44,228 Landsat OLI images and statistical results were spatially displayed along the path and row (e.g., 120/040) using the ArcGIS platform (version 10.1).

$$AP(S)_t^m = \frac{1}{N}\sum_y S_t^m \times 100\% \tag{1}$$

where S represents the probability of a certain acquisition of different CC for each year for each scene, m and y denotes each month (m = January, …) and year (i.e., between 2013 and 2016), t represents the different CC thresholds (t = 0, 10%, …, 100% or level, level 1,…, level 9, as described above), and N denotes the total number of observation scenes in month (m) in the entire Landsat archive.

3 Results and analysis

3.1 Appropriate CC threshold determination for AP analysis

In order to understand the effects of increasing CC thresholds (i.e., from 0, 10% to 100%), the average probability of all Landsat OLI observation was first evaluated. Figure 2 shows that the patterns of the average APs at the monthly (Figure 2a), seasonal, and annual scales (Figure 2b) with increasing CC thresholds for all Landsat OLI images. The results indicated that the curves of average AP with increasing CC thresholds show a “crescent” shape, which

suggests that the APs do not increase significantly over a whole year. There was a clear and rapid increase beyond level 0 CC threshold. Overall, in autumn and winter, Landsat OLI data acquisition over China had higher probabilities, while the corresponding APs remained comparatively low in spring and summer. The highest AP during the autumn and winter periods occurred in October, followed by December, January, February, November, and September, especially at the 30% CC threshold (0% ⩽ CC ⩽ 30%). Similarly, in spring and summer, June had the lowest probability, followed by any other month, August, July, May, April, and March. It should be noted that the values of average AP in March and September play a transition role for CC in response to the transformation between southeast and northwest monsoons.

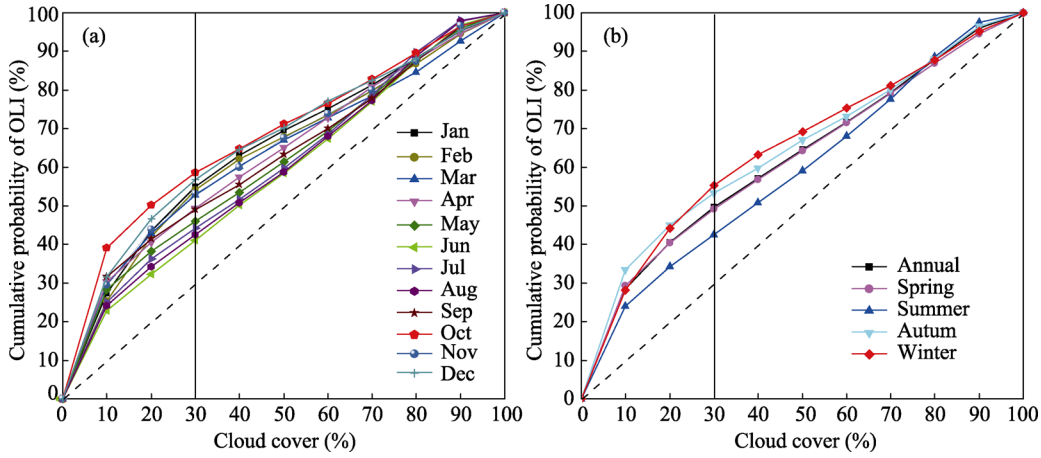

Figure 2 The cumulative average acquisition probabilities (APs) of increasing CC thresholds range from 0–100% for all OLI footprints in China: monthly (a) and annual and seasonal (b)

In general, a lower CC threshold makes an image more suitable for monitoring land surface changes. Some studies have found that the CC threshold is consistently less than or equal to 30% (Asner, 2001; Goward *et al*., 2006; Li *et al*., 2017). This CC threshold (i.e., less than or equal to 30%) represents the greatest extent possible value for land surface monitoring. Thus, the ⩽ 30% CC threshold was also selected in this study. The corresponding cumulative annual AP averages of CC for Landsat OLI images were 28.6%, 42.5%, and 49.6% for level 0, level 1 and level 2, respectively. The annual AP results with the ⩽ 30% CC in China are much larger than those of MSEA (41%) (Li *et al.*, 2017) and global estimates (37%) (Goward *et al*., 2006). This also demonstrates that the CC level 2 (i.e., less than or equal to 30%) is more appropriate for analysis of usable Landsat images (Ju and Roy, 2008; Asner, 2001), especially in China.

Next, the monthly average APs were computed for all Landsat OLI and are presented in Figure 2a, where the ⩽ 30% CC threshold is shown with a solid vertical line. It showed evident variability in the monthly average APs at the 30% CC threshold. The APs of CC at the 30% were greater than 50% over China for over six months from October to March next year. Figure 2a also shows that Landsat OLI had a relatively higher probability of successfully obtaining high-quality images from October to March during 2013–2016, with an average CC of 55.1%. Among them, the related AP reached a peak value (58.7%) in October,

followed by 56.8% in December, 55.0% in January, 54.1% in February, 53.1% in November, and 52.9% in March. It can be explained that the weakening subtropical high pressure and the beginning of cold air from Siberia in October usually leads to dry weather ("crisp autumn day") in China. This allows the acquisition of cloudless or little-cloud images in this month. On the contrary, the average monthly probability value was 45.3% between April and September. Particularly, the smallest probability for OLI acquiring targeted images occurred in June, merely 41.0%. The average monthly APs of other months are listed as follows, August (42.5%), July (44.2%), May (46.0%), April (49.3%) and September (48.9%). It can also be explained that the major river (e.g. Yangtze river and Yellow river) basins generally enter into flood season because of concentrated rainfall in this month (June), or earlier in April and May and end in September. The rainy weather hinders the observation of high-quality images.

Then, an analogous statistical analysis of seasonal AP showed that the seasonal variance with CC thresholds of $\leqslant$ 30% was distinct for the periods of autumn to winter and spring to summer. The average seasonal probability of acquiring a successful Landsat OLI image in winter was 55.3%, followed by 53.4% in autumn, 49.1% in spring, and 42.6% in summer (Figure 2b). In summary, analyses of the amount of useful Landsat OLI images at the CC thresholds $\leqslant$ 30% have great importance for monitoring land surface changes in China.

3.2 Spatio-temporal patterns of AP at the 30% CC threshold

3.2.1 Spatial comparison analysis of AP at the 10%, 20%, and 30% CC thresholds

Three thresholds of annual CC (i.e., 10%, 20% and 30%) were further selected to show the spatial differences in AP of OLI images for the period of 2013–2016 (Figure 3). Comparative results indicated that the distributions of OLI scene AP with the three CC thresholds are closely correlated with the isohyetal lines in China (Figure 3).

Specifically, the spatial patterns of AP for the three CC thresholds coincided with the 200 mm (Figure 3a), 400 mm (Figure 3b), and 800 mm (Figure 3c) isohyetal lines, respectively. Particularly, the spatial pattern of AP at 30% CC threshold coincides with the spatial distribution of the northwest and southeast, typically divided by the Qinling Mountains-Huaihe River Line (or the 800 mm isohyetal line), an important natural boundary between southern and northern China, which indicates transitions in geography, climate, and ecology. This correlation provides important insight in the influence of clouds and precipitation on optical satellite acquisition in the monsoon region. Figure 3 also indicates that the proportions of high-quality images data on the northwestern side of the line were much larger than those of the southeastern. In the following sections, the selection of the 30% CC threshold in this study was applied to further confirm the differences of AP at varied spatial and temporal scales. It should be pointed out that our analysis derived from the scene-based metadata were only for the whole Landsat OLI footprints, but did not take the overlay parts of any two adjacent footprints (or path/row) into consideration.

3.2.2 Temporal comparison analysis with the 30% CC threshold

China is situated in the transitional zone between the ocean and continent in longitude, covering the temperate, subtropical, and tropical zones. Therefore, regional differences in climate types or vegetation types are distinct. Explicit analysis of AP at the 30% CC threshold

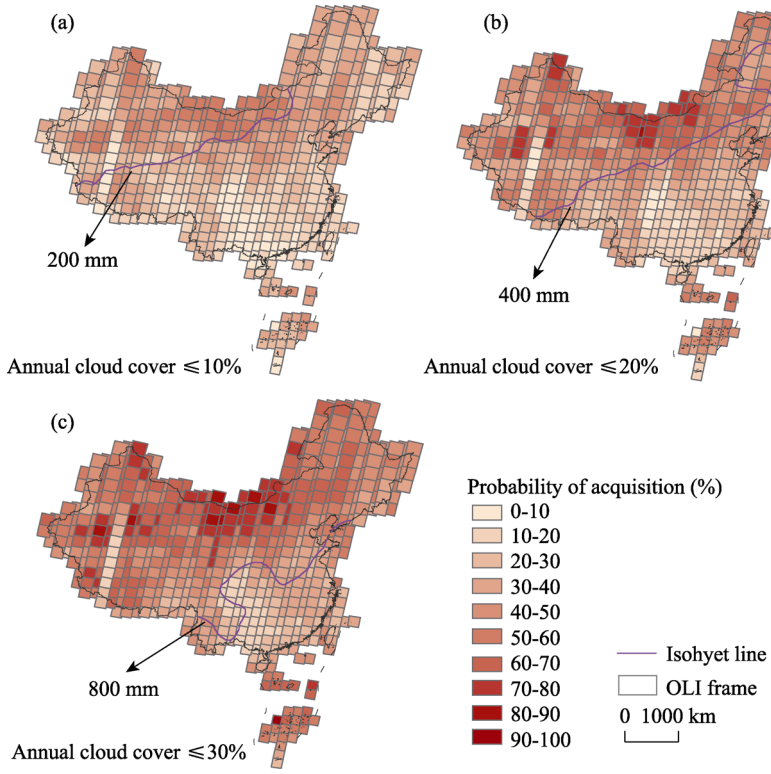

Figure 3 Annual AP with ⩽ 30% CC for Landsat OLI sensor images (44,228 scenes) between 2013 and 2016 in China. (a), (b) and (c) represent the annual CC thresholds of 10%, 20%, and 30%, respectively.

for the OLI archive products can contribute to evaluating the usefulness and applicability of OLI imagery at both spatial and temporal scales. Regional features for the 2013–2016 period have a unique spatial distribution and change in clouds as indicated by the Landsat OLI data at this threshold (Figure 4). On the whole, the probabilities of usable OLI image acquisition in autumn and winter were larger than those of the spring and summer seasons.

The monthly AP results of the historical OLI scenes showed clear regional differences (Figure 4). Among these differences, those between the northwest and the southeast regions were distinct, especially from October to March. Historical Landsat OLI observations with a threshold of AP at the 30% CC have higher chances of obtaining cloudless or little-cloud images, in most parts of northern China. This also suggests that the highest AP is in October of each year in China, while the corresponding lowest probability occurred in June. For example, in the middle reach of the Yangtze River, the observation probability remains highest in October, at the beginning of the winter season and the ending of the autumn season (Figure 4). Overall, northern and western China had AP values >90% even during the wetter spell of the year. The maximum variation in CC occurred in the northeast (especially the Sanjiang Plain), the Tibetan Plateau, and Hengduan Mountains. Among these areas, the entire Tibetan Plateau had 90%–100% chance of successful imaging from October to March. In addition, seasonal variance comparisons showed that tropical region of China (e.g., Hainan Island) had the larger APs for cloudless or little-cloud OLI images in spring and summer and lower APs in autumn and winter (Figure 4), while this relationship was slightly

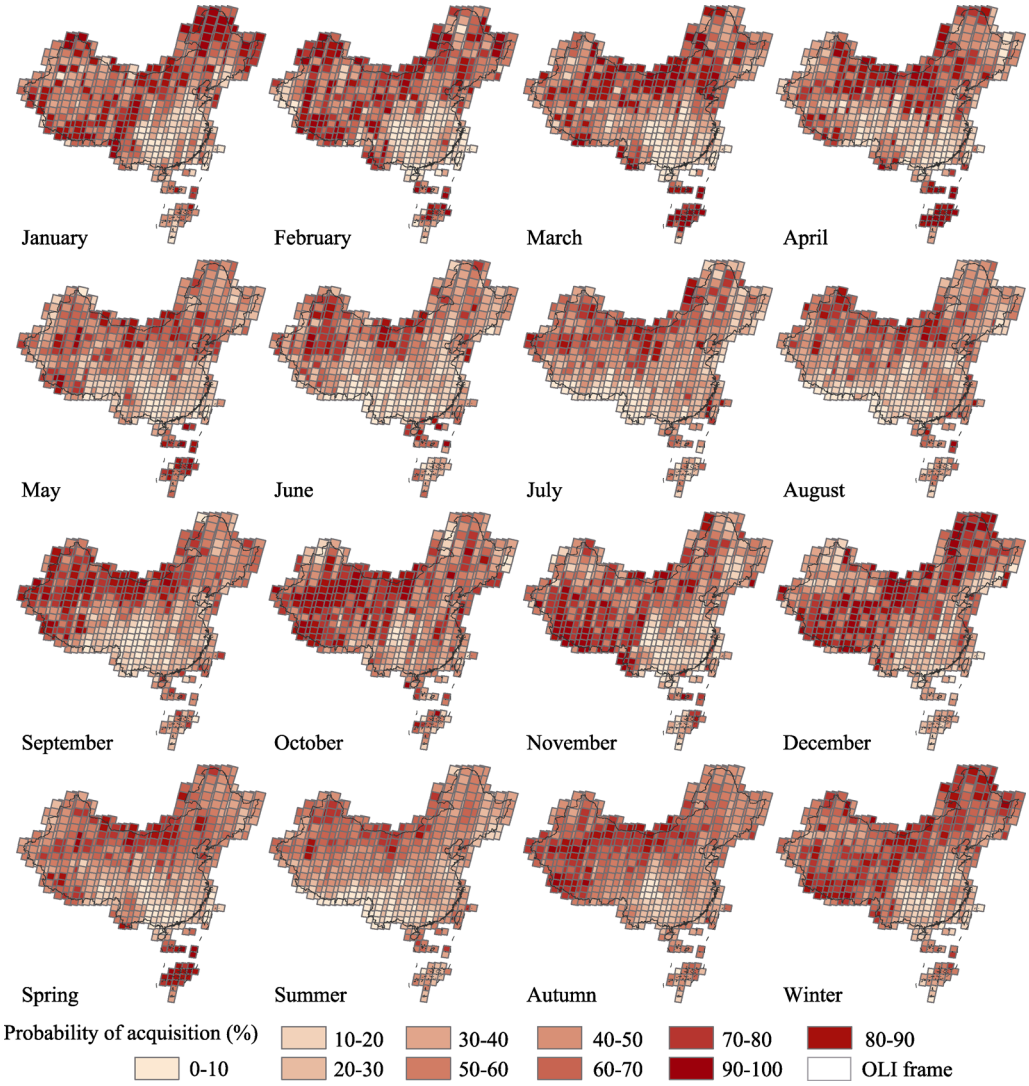

Figure 4 Monthly and seasonal AP for historical Landsat OLI scenes (42,229 scenes) with ⩽ 30% CC in China for 2013–2016

different in MSEA (Li *et al.*, 2017). The temporal changes of monsoonal wind systems may well explain the differences between them. These results confirmed the validity of the monthly result at the 30% CC threshold. In total, there is a relatively high probability of acquiring cloudless or little-cloud OLI images in autumn and winter, including the transition periods in March and September (Figure 4). These results also provide useful information for the selection of high-quality Landsat OLI imagery within proper critical window for remote sensing monitoring.

3.3 Area and distribution of AP in China based on the 30% CC threshold

The spatial distribution of probability in obtaining OLI scenes at CC level 2 (or ⩽ 30% CC) provides a chance for extracting useful surface feature information; for instance, combining image quality in different months for remote sensing monitoring. The relationships between

the AP of cloudless or little-cloud images and major geographical regions in China were further analyzed. AP values were calculated for all Landsat OLI images at the regional scale, and categorized into dry and wet regions (climate, Figure 5), river basins or grain production regions (water or land, Figure 6), and two sides of the Hu Line (population, Figure 7).

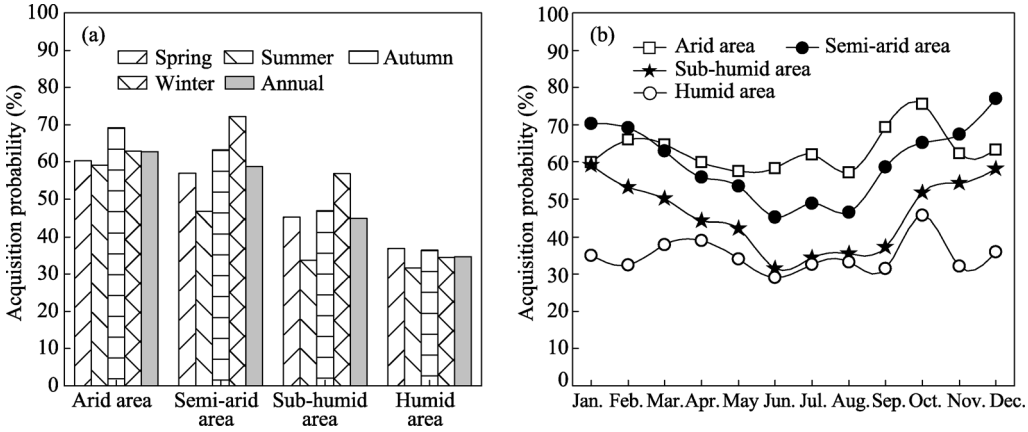

Figure 5 Regional differences in annual and seasonal (a) and monthly (b) average AP at 30% or less CC for all Landsat data. The four arid and humid areas in China are typically separated by the 200 mm, 400 mm, and 800 mm isohyetal lines.

3.3.1 Spatial differences of AP in dry and wet regions

China has the arid and semi-arid, semi-arid and sub-humid, sub-humid and humid areas, respectively (Cui *et al*., 2016). Because of rainfall differences in China, the arid area (62.7%) and semi-arid (58.8%) areas had larger annual average AP, compared to those of sub-humid (44.8%) and humid areas (34.7%).

Seasonal variance comparisons (Figure 5a) again showed that the arid area had an AP of about 70% in the autumn season, and 63.0%, 60.4%, and 59.2% for winter, spring, and summer seasons, respectively. In the semi-arid area, the AP was 72.2% in the winter season, followed by 63.3% in autumn, 57.1% in spring, and 46.8% in summer. The average probability of obtaining high-quality images in the sub-humid area in spring was 45.2%, the related probability declined in summer (33.7%), and increased in autumn (46.9%) and winter (56.9%). Only 31.7% of the images had a relatively high AP of a successful observation in summer in the humid area, as well as 36.4% in autumn, 34.5% in winter, and 36.8% in autumn. The variations in seasonal probability for the arid and humid areas were generally slight (less than 15%) in the whole year. However, there were large differences (greater than 50%) in the semi-arid and sub-humid areas between autumn to winter season and spring to summer season. Thus, the AP variations for the semi-arid and sub-humid areas were larger than those for the arid and humid areas. In addition, the chance for successful Landsat observations of arid area was about twice than that of the humid area, equal to the sub-humid area in autumn and winter, and 1.5 times than that of the sub-humid area in spring and summer. The relationship between the semi-arid area and (sub-) humid area was similar. Furthermore, in all (semi-) arid and (semi-) humid areas, the autumn and winter seasons were superior to the spring and summer seasons in obtaining cloudless or little-cloud images.

Finally, regional variations in Figure 5b demonstrate the “best” month in acquiring

high-quality images was October for the arid (75.5%) and humid (45.7%) areas. Likewise, the semi-arid (76.9%) and sub-humid areas (58.2%) had higher probabilities of acquiring Landsat OLI cloudless or little-cloud images in December. As a result, Landsat OLI imagery has proven to be more applicable for monitoring land surface changes in the arid (especially in October), semi-arid areas (particularly in December) than the sub-humid (particularly in December), and humid areas (especially in October). The results also demonstrated that CC changed due to the shift between southeast and southwest monsoonal wind systems. This distinct monthly variance of AP values will facilitate studying land surface changes comprehensively and systematically.

3.3.2 Spatial differences of AP in major river basins (or grain production regions)

At present, Landsat imagery is important and successful for evaluating and discriminating the quality of soil and water (Liu *et al.*, 2016). In China, the spatial distribution of grain production region is closely related to nine major river basins: the Songliao River (SLR), Inland River (InR), Luanhai River (LHR), Yellow River (YeR), Pearl River (PR), Yangtze River (YtR), Southwest Rivers (SWRs), Huaihe River (HR), and Southeast Rivers (SERs) basins (Jin *et al*., 2016; Liu *et al*., 2007). Spatially, regional differences in the nine river basins varied dramatically in the AP of cloudless or little-cloud OLI images (Figure 6). Among the nine major river basins, the SLR, InR, LHR, and YeR basins had annual AP greater than 50% for obtaining high-quality images, compared with the AP values for the PR (23.4%), YtR (28.7%), SWRs (32.9%), HR (36.6%), and SERs (45.8%) basins. Comparative analyses indicated that there were similar regional differences in CC over the seasonal and monthly variance. Figure 6 summarizes the average seasonal (left panel, Figure 6a) and monthly (right column, Figure 6b) AP values of obtaining cloudless or little-cloud images within a given year in the nine major river basins.

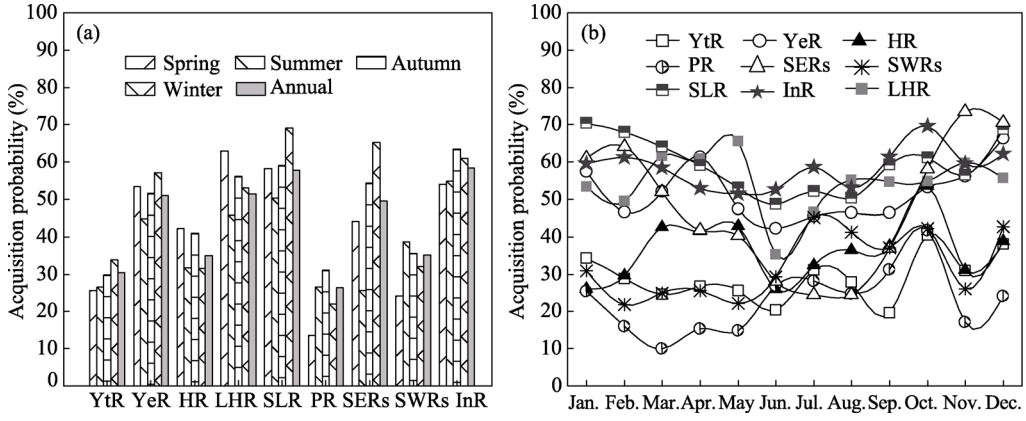

Figure 6 River basin differences in annual and seasonal (a) and monthly (b) average AP at ⩽ 30% CC for all Landsat data in China. YtR: Yangtze River; YeR: Yellow River; HR: Huaihe River; LHR: Luanhai River; SLR: Songliao River; PR: Pearl River; SERs: Southeast Rivers; SWRs: Southwest Rivers; InR: Inland River

As shown in Figure 6a, of the nine major river basins, the largest difference of AP happens in the spring season (4.6 times). The LHR basins had the largest probabilities in the spring season (62.9%) whle the corresponding lowest probabilities occurred in the PR basins (13.6%). The ratio of APs between the SERs and PR basins was about 3.1. The SERs basins

had the largest probabilities in winter (up to 69.0%) and the corresponding lowest probabilities in the PR basins (22.0%). The difference of statistical results was 2.1 in summer, with 54.8% in InR basins and 25.5% in SERs basins. A similar ratio occurred in the autumn season, the larger probabilities (63.5%) in the InR basins and the corresponding lower probabilities (13.6%). Next, the monthly average APs of Landsat OLI images showed distinct characteristics (Figure 6b). Monthly variance comparisons at a regional scale showed that the SLR basin had the highest AP in January (70.3%, Table 2), and lowest monthly average 48.7% in June. Similar analyses were applied to other basins for the highest and lowest AP values in Table 2. These results also demonstrated that Landsat OLI has higher probabilities of acquiring high-quality iamges in the autumn and winter seasons.

Table 2 Differences in monthly average AP at ≤30% CC threshold for OLI data in the nine major river basins of China

Major River Basin	AP (%)											
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Yangtze River (YtR)	34.3	28.8	24.6	26.6	25.6	20.4	31.1	27.8	19.6	40.5	31.0	38.0
Yellow River (YR)	57.4	46.5	52.2	61.4	47.5	42.2	45.5	46.5	46.5	53.2	56.2	66.3
Huaihe River (HR)	26.1	29.6	42.5	41.4	42.9	25.9	32.4	36.4	37.5	53.6	31.1	38.9
Luanhai River (LHR)	53.4	49.5	61.7	60.8	65.6	35.3	46.6	55.2	54.7	54.9	59.6	55.7
Songliao River (SLR)	70.3	67.9	64.2	59.1	53.1	48.7	52.1	50.3	59.2	61.2	56.5	68.7
Pearl River (PR)	25.4	16.0	10.0	15.3	14.9	26.9	28.1	24.7	31.2	41.7	17.2	24.0
Southeast Rivers (SERs)	60.9	64.1	52.0	41.8	40.3	27.9	24.4	24.4	37.0	58.1	73.5	70.5
Southwest Rivers (SWRs)	31.0	21.8	25.0	25.6	22.1	29.4	45.1	41.2	37.1	42.2	26.0	42.7
Inland River (InR)	59.5	61.2	58.5	52.9	51.5	52.6	58.5	53.1	61.3	69.6	59.5	62.1

3.3.3 Spatial differences of AP on the two sides of the Hu Line

The Hu Line is considered as one of the most important geographical discoveries in China (Qi *et al*., 2016). The spatial patterns of CC distribution in the two sides of the Hu Line were obviously different according to the average AP evaluation. Statistical results indicated that the northwestern side of the Hu Line (NWHL) had 56.5% annual AP for obtaining a cloudless or little-cloud Landsat 8 OLI imagery, while the southeastern side of the Hu Line (SEHL) had a lower AP of 39.9%. Generally, CC in the SEHL was much higher than that of the NWHL in any given month or season. This can be explained that SEHL has plenty of rain with humid weather, which easily leads to high CC during the revisit of Landsat sensors. However, the phenomenon was just the opposite in the NWHL. Figure 7 presents the spatial patterns of the AP of cloudless or little-cloud OLI images on both parts of the Hu Line. The results showed that the average monthly or seasonal probabilities of high-quality Landsat OLI images acquisition in the NWHL were larger than those of the SEHL throughout the year.

Seasonally, there were larger APs for cloudless or little-cloud OLI images in winter (65.4%) and autumn (61.5%) seasons in the NWHL, and the corresponding average values declined in the spring (54.7%) and summer (47.4%) seasons. The SEHL showed similar variations in the seasonal average probabilities at the threshold of level 2, with 42.5% and 41.3% in the winter and autumn seasons, respectively, and 42.4% in the spring season and

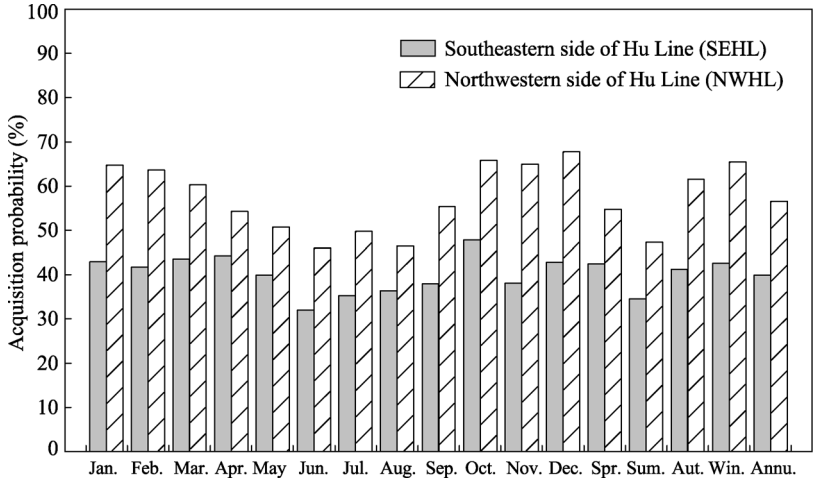

Figure 7 Regional differences in annual, seasonal and monthly average AP at ⩽ 30% CC for Landsat OLI data in China on both sides of the Hu Line

34.6% in the summer season. Monthly variances showed that the largest AP was 67.8% in December for the NWHL and 47.8% in October for the SEHL, with the difference approximately 20%. In general, the NWHL had higher (over 45%) monthly average AP (including 65.8% in October, 65.0% in November, 64.7% in January, 63.6% in February, 60.3% in March, 55.3% in September, 54.2% in April, 50.7% in May, 49.7% in July, 46.4% in August, and 45.9 in June), when compared to those of the SEHL (about 47.8%, 38.1%, 42.9%, 41.7%, 43.6%, 38.0%, 44.3%, 39.8%, 35.3%, 36.3%, and 32.0% in the corresponding month). Our CC analysis confirmed that the Hu Line acts as a geographical boundary between oceanic climate and continental climate. Therefore, differences in AP are useful for monitoring human settlements with Landsat data on both sides of the Hu Line.

4 Conclusions and discussion

4.1 Conclusions

Recently, freely available historical Landsat products have been widely used in land surface monitoring from regional to global level. Cloud cover (CC) analysis of historical Landsat data is a prerequisite for remote sensing monitoring. The acquisition probability (AP) analysis of CC helps to understand the temporal and spatial practicality of obtaining a usable OLI image. In this study, we explicitly investigated the spatio-temporal differences in AP for varied CC thresholds using 44,228 Landsat OLI images over China from April 2013 to October 2016. We then analyzed the differences in AP at the 30% or less CC threshold in the dry and wet regions, river basins or grain production regions, and two sides of the Hu Line (Hu Huanyong population line). Some main conclusions were drawn as follows:

(1) Cumulative frequency showed larger probabilities at lower CC thresholds, in contrast, smaller percentage at higher CC thresholds. The cumulative average AP of all OLI data over China at the CC thresholds ⩽ 30% was about 49.6%. This higher AP illustrated the availability of OLI imagery for monitoring land surface changes across China, in spite of obvious national differences in APs.

(2) The spatial patterns of the AP of Landsat OLI at varied CC thresholds and the major isohyetal lines of China coincided well. The spatial patterns of lower-threshold CC of OLI observations, namely 10%, 20%, and 30%, coincided well with the precipitation distributions of China divided by the isohyetal lines of 200 mm, 400 mm, and 800 mm, respectively.

(3) Temporal differences in AP for obtaining cloudless or little-cloud images were clear. China has higher probabilities of acquiring high-quality OLI images in autumn and winter especially in October of 58.7%, while the corresponding lowest probability in June merely 41.0%.

(4) Similarly, the spatial differences in APs of targeted images with $\leqslant$ 30% CC thresholds were quite significant. At regional scales, the arid and semi-arid areas, Inland River and Songliao River basins, and northwestern side of the Hu Line had the larger probabilities of obtaining high-quality images for monitoring single locations.

4.2 Discussion

In this study, we delineated the statistical features of CC of Landsat OLI imagery in China and explored the relationships between high-quality imagery and the spatial differences in water, land, and climate. The distinct monthly or seasonal AP variations in different regions at the 30% CC threshold in OLI imagery will contribute to facilitating land surface surveying. This is consistent with the observation from the Asner's study (Asner, 2001; Li *et al*., 2017). Our study suggested that the newly launched Landsat 8 OLI imagery satisfy the data requirements needed for land surface monitoring, although there existed the distinct spatio-temporal variances. To our knowledge, this study may be the first systematic analysis of CC in OLI imagery over China. The results and conclusions are of practical guidance for selecting cloudless or little-cloud Landsat OLI data for land surface changes and ecosystem dynamics research. Furthermore, the study provides useful information for future research on the availability and analysis of CC in Landsat images for a specific region, although our study mainly focused on the average value of CC percentage in the whole Landsat OLI footprints. Thus, the area of the Landsat OLI adjacent path/rows in WRS would influence the spatial analysis results in each intra-scenes scale to some extent. Following this, we will continue to analyze the CC percentage differences at the pixel level over China-Southeast Asia in the near future, especially with Landsat family sensors (TM, ETM+ and OLI).

As noted, the results show that the OLI sensor has the generally higher AP for high-quality images in China compared to reported some estimates in other studies (Goward *et al*., 2006; Li *et al.*, 2017). The scene-based metadata analysis of spatio-temporal differences of AP is a prerequisite for regional remote sensing monitoring, especially for crop mapping such as paddy rice (Li *et al*., 2016) and wheat (Feng *et al.*, 2014). It is worthwhile to mention that our study focused only on the OLI data CC, which is a global archive. Further analysis of other Landsat data (i.e., TM and ETM+) would help to comprehensively understand suitability and limitation of the historical Landsat data for land surface studies in the future.

References

Aplin P, 2004. Remote sensing: Land cover. *Progress in Physical Geography*, 28(2): 283–293.

Arvidson T, Gasch J, Goward S N, 2001. Landsat 7's long-term acquisition plan: An innovative approach to building a global imagery archive. *Remote Sensing of Environment*, 78(1/2): 13–26.

Asner, G P, 2001. Cloud cover in Landsat observations of the Brazilian Amazon. *International Journal of Remote Sensing*, 22(18): 3855–3862.

Cui Y P, Ning X J, Qin Y C *et al*., 2016. Spatio-temporal change of agricultural hydrothermal conditions in China from 1951 to 2010. *Journal of Geographical Sciences*, 26(6): 643–657.

Ding Y, 1961. Cultivation of Rice in China. Beijing: China Agriculture Press. (in Chinese)

Dong J W, Xiao X M, Chen B Q *et al*., 2013. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. *Remote Sensing of Environment*, 134: 392–402.

Feng Z M, Yang L, Yang Y Z, 2014. Temporal and spatial distribution patterns of grain crops in the West Liaohe River Basin. *Journal of Resources and Ecology*, 5(3): 244–252.

Foley J A, DeFries R, Asner G P *et al*., 2005. Global consequences of land use. *Science*, 309: 570–574.

Ge Q S, Zheng J, Hao Z X *et al*., 2016. Recent advances on reconstruction of climate and extreme events in China for the past 2000 years. *Journal of Geographical Sciences*, 26(7): 827–854.

Goward S, Arvidson T, Williams D *et al*., 2006. Historical record of Landsat global coverage: Mission operations, NSLRSDA, and international cooperator stations. *Photogrammetric Engineering & Remote Sensing*, 72(10): 1155–1169.

Hansen M C, Loveland T R, 2012. A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122: 66–74.

Hao H M, Ren Z Y, 2009. Land Use/Land Cover Change (LUCC) and eco-environment response to LUCC in farming-pastoral zone, China. *Journal of Integrative Agriculture*, 8(1): 91–97.

Jin S M, Homer C G, Yang L M *et al*., 2013a. Automated cloud and shadow detection and filling using two-date Landsat imagery in the USA. *International Journal of Remote Sensing*, 34(5): 1540–1560.

Jin S M, Yang L M, Danielson P *et al*., 2013b. A comprehensive change detection method for updating the National Land Cover Database to circa 2011. *Remote Sensing of Environment*, 132: 159–175.

Jin T, Qing X Y, Huang L Y, 2016. Changes in grain production and the optimal spatial allocation of water resources in China. *Journal of Resources and Ecology*, 7(1): 28–35.

Ju J, Roy D P, 2008. The availability of cloud-free Landsat ETM plus data over the conterminous United States and globally. *Remote Sensing of Environment*, 112(3): 1196–1211.

Kovalskyy V, Roy D P, 2013. The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for Global 30 m landsat data product generation. *Remote Sensing of Environment*, 130: 280–293.

Kovalskyy V, Roy D P, 2015. A one year Landsat 8 conterminous United States study of cirrus and non-cirrus clouds. *Remote Sensing*, 7(1): 564–578.

Laborde H, Douzal V, Ruiz Piña, H A *et al.*, 2017. Landsat-8 cloud-free observations in wet tropical areas: A case study in South East Asia. *Remote Sensing Letters*, 8(6): 537–546.

Li P, Feng Z M, Jiang L G *et al*., 2012. Changes in rice cropping systems in the Poyang Lake Region, China during 2004–2010. *Journal of Geographical Sciences*, 22(4): 653–668.

Li P, Feng Z M, Xiao C W, 2017. Acquisition probability differences in cloud coverage of the available Landsat observations over mainland Southeast Asia from 1986 to 2015. *International Journal of Digital Earth*, 10: 1–14. doi: 10.1080/17538947.2017.1327619.

Li P, Jiang L G, Feng Z M *et al*., 2016. Mapping rice cropping systems using Landsat-derived Renormalized Index of Normalized Difference Vegetation Index (RNDVI) in the Poyang Lake Region, China. *Frontiers of Earth Science*, 10(2): 303–314.

Liu J Y, Deng X Z, 2010. Progress of the research methodologies on the temporal and spatial process of LUCC. *Chinese Science Bulletin*, 55(14): 1354–1362.

Liu J Y, Kuang W H, Zhang Z X *et al*., 2014. Spatiotemporal characteristics, patterns, and causes of land-use changes in China since the late 1980s. *Journal of Geographical Sciences*, 24(2): 195–210.

Liu J Y, Shao Q Q, Yan X D. *et al*., 2016. The climatic impacts of land use and land cover change compared among countries. *Journal of Geographical Sciences*, 26(7): 889–903.

Liu R G, Liu Y, 2013. Generation of new cloud masks from MODIS land surface reflectance products. *Remote Sensing of Environment*, 133: 21–37.

Liu X N, Feng Z M, Jiang L G *et al*., 2013. Rubber plantation and its relationship with topographical factors in the border region of China, Laos and Myanmar. *Journal of Geographical Sciences*, 23(6): 1019–1040.

Liu Y J, Yang Y Z, Feng Z M, 2007. The change of the main regions for China's food grain production and its implications. *Resources Science*, 29(2): 8–14. (in Chinese)

Qi W, Liu S H, Zhao M F *et al*., 2016. China's different spatial patterns of population growth based on the "Hu Line". *Journal of Geographical Sciences*, 26(11): 1611–1625.

Qin Z, Karnieli A, Berliner P, 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. *International Journal of Remote Sensing*, 22(18): 3719–3746.

Ren Z Y, Wang L X, 2007. Spatio-temporal differentiation of landscape ecological niche in western ecological frangible region: A case study of Yan'an region in northwestern China. *Journal of Geographical Sciences*, 17(4): 479–486.

Sano E E, Ferreira L G, Asner G P *et al*., 2007. Spatial and temporal probabilities of obtaining cloud-free Landsat images over the Brazilian tropical savanna. *International Journal of Remote Sensing*, 28(12): 2739–2752.

Steven M D, Malthus T J, Baret F *et al*., 2003. Intercalibration of vegetation indices from different sensor systems. *Remote Sensing of Environment*, 88(4): 412–422.

Teillet P M, Fedosejevs G, Thome K J E *et al*., 2007 Impacts of spectral band difference effects on radiometric cross-calibration between satellite sensors in the solar-reflective spectral domain. *Remote Sensing of Environment*, 110(3): 393–409.

U.S. Geological Survey (USGS), 2016. Landsat 8 (L8) Data Users Handbook. Available online: https://landsat.usgs.gov/landsat-8-l8-data-users-handbook (accessed on 18 April 2017).

Wu Y Y, Li S Y, Yu S X, 2016. Monitoring urban expansion and its effects on land use and land cover changes in Guangzhou city, China. *Environmental Monitoring Assessment*, 188(541), 609–621.

Whitcraft A K, Vemiote E F, Becker-Reshef I *et al*., 2015. Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations. *Remote Sensing of Environment*, 156: 438–447.

Zhu Z, Woodcock C E, 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118: 83–94.