Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 14, No. 1, 2025

Rapid and Accurate Mapping of Winter Wheat Distribution in Jiaozuo City Using Sentinel-2 and Google Earth Engine

Haofeng Luo

School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China

Abstract: Rapid and accurate mapping using remote sensing and cloud computing is crucial for food security. This study employs Google Earth Engine (GEE) and Sentinel-2A multispectral imagery to analyze the distribution of winter wheat in Jiaozuo City from 2019 to 2021 using the one-class support vector machine (OCSVM) classification method. The results show high classification accuracy, with an overall accuracy of 0.952, a Kappa coefficient of 0.952, a producer's accuracy of 0.953, and a consumer's accuracy of 0.978, validating the classification effectiveness. This method is efficient, easy to implement, and suitable for large-scale crop mapping, providing valuable support for sustainable agriculture and policy-making.

Keywords: Winter wheat distribution; Sentinel-2; Remote sensing cloud computing; Accuracy; Classifier.

1. Introduction

Food security has always been a fundamental national policy in China, reflecting the ancient wisdom that "food is the first necessity of the people." Ensuring food security is not only the foundation of China's social development but also provides valuable insights for addressing the global food crisis [1]. While China adheres to the strict policy of maintaining 1.8 billion mu (120 million hectares) of arable land, accurately determining the specific cultivated area each year remains a significant challenge [2]. Reliable data on crop types and distribution are essential for securing food production [3].

Wheat is one of the most widely cultivated crops worldwide, and accurately assessing its planting area is crucial for ensuring food security. In the past, traditional survey methods such as field investigations and on-site measurements were commonly used. Although these methods provide high accuracy, they require substantial human, financial, and material resources, often taking a long time to complete. A major drawback is that by the time a survey is completed, the initial planting conditions may have already changed, reducing the timeliness and effectiveness of the data.

With the continuous advancement of technology and the rapid development of remote sensing cloud computing platforms, significant progress has been made in monitoring crop cultivation areas. Winter wheat, a major crop in Jiaozuo City, plays a vital role in regional food balance and agricultural resource planning. Accurate data on its planting area and spatial distribution are essential for optimizing agricultural production and ensuring sustainable economic development. By leveraging remote sensing cloud computing platforms [4], it is possible to obtain real-time and accurate winter wheat distribution data, facilitating dynamic monitoring[5]. This approach is particularly valuable for enhancing agricultural management strategies and improving overall efficiency. The monitoring process for winter wheat distribution involves several key steps, as illustrated in Figure 1:

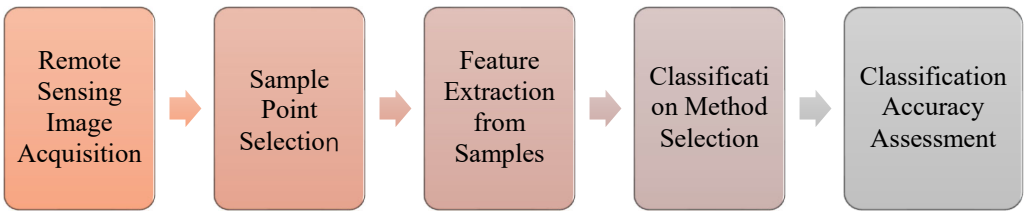

Figure 1. Monitoring of Wheat Distribution

When selecting classification features, it is essential to consider the spectral characteristics of different crops. Since each crop exhibits unique spectral indices, distinguishing them requires careful analysis of their spectral differences. The spectral time-series analysis of wheat and other major crops in the study area reveals that wheat has distinct spectral characteristics, which form the foundation for classification.

Throughout the entire growing period, wheat can be roughly divided into four stages. During the sowing stage, seeds are newly planted, and the cultivated land appears as bare soil, resulting in a high Bare Soil Index (BSI). As the seeds germinate and seedlings emerge, the Normalized Difference Vegetation Index (NDVI) begins to rise. In the overwintering stage, wheat seedlings enter dormancy, slowing their above-ground growth. This process leads to a significant decrease in NDVI, while the Normalized Difference Phenology Index (NDPI) exhibits an upward trend, reflecting physiological changes in the plants. In the growth

stage, wheat experiences rapid vegetative expansion, causing NDVI to continue increasing, accompanied by a significant upward trend in the Enhanced Vegetation Index (EVI), which indicates increasing chlorophyll content and canopy density. As wheat enters the maturity stage, chlorophyll content declines, causing NDVI to drop sharply, whereas the Plant Senescence Reflectance Index (PSRI) increases, signaling leaf senescence and grain ripening.

By analyzing spectral characteristics across these four stages, it is evident that wheat can be effectively distinguished from most other crops. However, a challenge encountered in the study is that garlic exhibits highly similar spectral features to wheat, leading to potential misclassification where garlic fields are mistakenly identified as wheat. Although solutions to this issue are still being explored, it is worth noting that garlic cultivation in Jiaozuo City is relatively limited, making such classification errors negligible. To further enhance classification accuracy, future research may incorporate additional spectral indices, texture features, and advanced machine learning algorithms. The integration of time-series analysis and multi-temporal image fusion techniques may also improve the ability to differentiate wheat from spectrally similar crops, further refining monitoring precision.

2. Materials and Methods

2.1. Study Area Overview

The study area selected for this research is Jiaozuo City, Henan Province, China. Jiaozuo is located in the northwest of Henan Province, between 35°10′- 35°21′N latitude and 113°4′-113°26′E longitude. The city is bordered by the Taihang Mountains to the north and the Yellow River to the south. It consists of 11 counties and districts, covering a total area of 4,071 square kilometers, with a population of 3.58 million, of which 825,000 reside in the urban area.

The topography of Jiaozuo City slopes from north to south, exhibiting significant elevation variations. The plains below 200 meters in elevation cover approximately 2,959 square kilometers, accounting for 73% of the city's total area, while hills and mountainous regions above 200 meters cover about 1,066 square kilometers, representing 27% of the total land area (Figure 2). The city falls within the temperate zone, characterized by a mild climate, distinct seasons, hot and rainy summers, and cold winters with little snowfall. These natural conditions have made wheat, corn, and rice the primary crops cultivated in Jiaozuo.

Figure 2. Overview of Jiaozuo Study Area

2.2. Data Acquisition and Processing

2.2.1. Data Processing Steps

The first step in data processing involved downloading the vector boundary map of the study area from the Google Earth Engine (GEE) official website and importing it into the platform. Next, appropriate remote sensing imagery was selected. GEE provides access to several widely used remote sensing datasets, including Sentinel, Landsat, and MODIS. For this study, Sentinel-2A data were chosen due to their high resolution and suitability for assessing vegetation, soil, and water conditions [6]. The surface reflectance dataset used in this study underwent Level-2A orthorectification[7], ensuring precise georeferencing and radiometric corrections before being loaded into the software for analysis.

Due to climatic conditions, acquiring completely cloud-free remote sensing imagery can be challenging. Most images contain some level of cloud cover, cloud shadows, or aerosol interference, which can impact classification accuracy. To mitigate these effects, a cloud-removal process was applied to the imagery, incorporating spectral feature adjustments [8,9] (Figure 3).

```
// Cloud Masking Function
function maskS2clouds(image) {}

function CalculateBrightnessIndex(image){
  var bsi = image.expression(
    '((SWIR+Red)-(NIR+Blue))/((SWIR+Red)+(NIR+Blue))', {
    'SWIR':image.select('Band11'),
    'Red':image.select('Band4'),
    'NIR':image.select('Band8'),
    'Blue':image.select('Band2'),
  }).rename("BrightnessIndex");
  return image.addBands([bsi]);
}

// Load Spectral Features
function CalculateVegetationIndex(image){
  var ndvi =
image.normalizedDifference(["Band8","Band4"]).rename("VegetationIndex");
  return image.addBands([ndvi]);
}
```

```
// Load Spectral Features
function CalculateVegetationIndex(image){
  var ndvi =
image.normalizedDifference(["Band8","Band4"]).rename("VegetationIndex");
  return image.addBands([ndvi]);
}

function CalculateGreenVegetationIndex(image){
  var gndvi = image.expression(
    '(NIR-Green)/(NIR+Green)', {
    'NIR':image.select('Band8'),
    'Green':image.select('Band3'),
  }).rename("GreenVegetationIndex");
  return image.addBands([gndvi]);
}
```

Figure 3. Image Preprocessing Operations

After preprocessing, an RGB composite image was generated to visually inspect the study area (Figure 4).

Figure 4. Color Composite Image of the Study Area

From Figures 3 and 4, it is evident that residual cloud cover and atmospheric effects may still interfere with classification accuracy. To eliminate these potential sources of error, all Sentinel-2 images from 2019 to 2021 covering Jiaozuo were filtered based on cloud percentage. A cloud-free composite image was then generated, preserving rich phenological information for winter wheat monitoring. After applying the cloud-percentage filter, a total of 561 images were selected for further processing.

Based on data provided by the GEE platform and previous classification studies, sample points were manually selected using visual interpretation methods. These selected sample

points were then overlaid onto the satellite imagery to generate a reference dataset

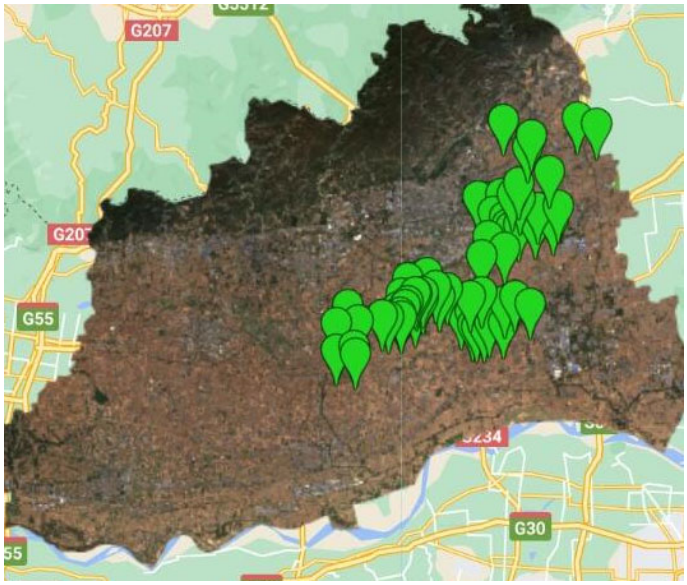

Figure 5. Distribution of Sample Points in the Study Area

By analyzing the spectral characteristics of these sample points, the most suitable classification features were identified. A classification model was then developed based on these features. Finally, a support vector machine (SVM) classifier was applied to classify winter wheat, and the classification results were validated by comparing the predicted wheat distribution with sample-based verification data.

2.2.2. Theoretical Basis of the Method

Supervised classification methods are commonly used in remote sensing but often suffer from small-sample problems, where the limited number of training samples cannot represent all land cover types. This limitation can lead to misclassification or omission errors. Traditional supervised classification methods such as Maximum Likelihood Classification (MLC) and Minimum Distance Classification (MDC) are highly dependent on large training datasets. When applied to small-sample scenarios, these methods often produce unstable classification results due to algorithmic constraints.

To address these challenges, this study adopts the One-Class Support Vector Machine (OCSVM) classification method, a widely used approach in remote sensing classification. Support Vector Machines (SVMs) are binary classification models that define a maximum-margin linear classifier in a feature space. SVMs also incorporate kernel techniques, allowing them to function as nonlinear classifiers. The core learning strategy of SVMs is margin maximization, which can be mathematically formulated as a convex quadratic programming problem, equivalent to minimizing a regularized hinge loss function. The SVM learning algorithm is thus an optimization process for solving convex quadratic programming problems.[10-12]

3. Results and Analysis

3.1. Selection of Sample Points

The selection of sample points is a critical factor determining the success of this study. Since both classification feature selection and classification model development rely on these reference points, the accuracy of the final classification results is directly influenced by the quality and representativeness of the sample data.

Due to the COVID-19 pandemic, field surveys could not be conducted extensively. As a result, this study could not fully rely on on-site investigations to collect sample points. Instead, a combination of limited field surveys, high-resolution imagery from Google Earth, and visual interpretation methods was used to obtain reliable training samples. A total of 120 accurately identified sample points were manually selected based on remote sensing data analysis.

The distribution of these sample points was analyzed using the Google Earth Engine (GEE) remote sensing cloud platform, ensuring a diverse spatial representation of different land cover types within the study area. The spatial distribution of the selected sample points

is illustrated in Figure 6.

```
var withRandom = traindata.randomColumn('random');
var split = 0.5;
var trainingPartition = withRandom.filter(ee.Filter.lt('random',
var testingPartition = withRandom.filter(ee.Filter.gte('random',
```

Figure 6. Distribution of Sample Points in the Study Area

To ensure the robustness of the classification model and the reliability of validation, 50% of the sample points were used for model training, while the remaining 50% were used for accuracy assessment. This balanced approach enhances the classification reliability, reducing the risk of overfitting and improving the generalizability of the model for large-scale agricultural monitoring applications.

3.2. Selection of Classification Features

The classification of winter wheat is based on its unique

spectral characteristics throughout its growth cycle. To enhance the accuracy of classification, this study divides the entire growth cycle of winter wheat into four distinct phenological stages: sowing stage, overwintering stage, growth stage, and maturity stage. For each stage, appropriate spectral indices are selected as classification features to maximize differentiation from other crops.

The selection of classification features is guided by the spectral-temporal variation of winter wheat in comparison to other local crops. Through time-series spectral analysis, it is evident that winter wheat exhibits distinct spectral characteristics at different growth stages, forming the basis for effective classification [13-16].

The detailed classification features corresponding to each phenological stage are presented in Table 1 below.

Table 1. Classification Features for Different Phenological Stages of Winter Wheat

Spectral Feature	Calculation Formula	Phenological Stage
BSI	(SWIR+R)-(NIR+B))/((SWIR+R)+(NIR+B)	Sowing Stage
NDVI	(NIR-R)/(NIR+R)	Growth Stage
GNDVI	(NIR-G)/(NIR+G)	Growth Stage
$NDVI_6$	(6*NIR-R)/(NIR+6*R)	Growth Stage
EVI	(2.5*NIR-R)/(NIR+6*R-7.5*B+1)	Growth Stage
NDPI	(NIR-(0.74*R+0.26*SWIR))/(NIR+(0.74*R+0.26*SWIR))	Overwintering Stage
PSRI	(R-B)/R2	Maturity Stage

3.3. Selection of Classifier

Since this study focuses solely on extracting winter wheat, a single-stage classifier is employed instead of the traditional multi-stage classifier to enhance efficiency. The One-Class Support Vector Machine (OCSVM) is a variation of the Support Vector Machine (SVM) that is particularly well-suited for cases where only one class of data is available. By training on normal data samples, OCSVM constructs an optimal hyperplane in the feature space and separates the training samples from the origin based on the principle of maximum margin.

A significant advantage of OCSVM is that it requires only one class of training samples, significantly reducing the time and human effort required for sample collection. Previous studies have demonstrated that OCSVM can achieve high extraction accuracy when applied to winter wheat mapping. In this study, new spectral features of winter wheat were incorporated into the OCSVM model to improve classification performance. After multiple experimental iterations, the optimal parameters for the classifier were determined, as shown in Figure 7.

```
var classifier = ee.Classifier.libsvm({
  svmType: 'ONE_CLASS',
  kernelType: 'RBF',
  gamma: 5.0,
  nu:0.1
```

Figure 7 Experimental Parameters for Winter Wheat Extraction

Based on the classification results generated by the OCSVM classifier, the distribution of winter wheat planting areas in Jiaozuo City was calculated using the remote sensing cloud platform. The final mapping results illustrate the spatial distribution of winter wheat across the study area, providing valuable insights for agricultural monitoring and decision-making (Figure 8).

```
var mianji = ee.Image.pixelArea( ).multiply(wheat.neq(1)).multi
    reducer:ee.Reducer.sum( ),
    geometry: roi,
    scale:100,
    tileScale:16,
    maxPixels:1e15
});
var area = mianji.getNumber('area').divide(10000);
print(area);
```

Figure 8. Remote Sensing Cloud Computing Operation Diagram

The classification process was implemented using Google Earth Engine (GEE), and the final classification results were generated after processing and analysis.

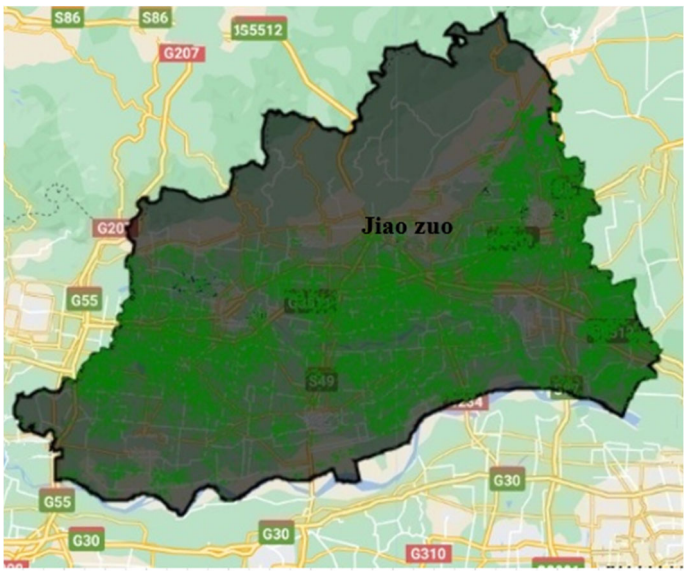

Figure 9. Distribution of Winter Wheat Planting Areas in Jiaozuo City

The final classification results show the spatial distribution of winter wheat planting areas in Jiaozuo City, providing a valuable reference for agricultural monitoring, land-use planning, and food security assessment.

3.4. Accuracy Assessment

To validate the classification method, 50% of the 120 selected sample points were used for classification, while the remaining 50% were used for accuracy verification. The accuracy assessment was performed using four key evaluation metrics, including Overall Accuracy (OA), Kappa Coefficient, Producer's Accuracy (PA), and User's Accuracy (UA). The Overall Accuracy represents the proportion of correctly classified pixels to the total number of pixels, with correctly classified pixels aligning along the diagonal of the confusion matrix. The Kappa Coefficient measures classification agreement by considering both observed and expected accuracy, calculated by subtracting the expected accuracy from the observed accuracy and dividing it by the maximum possible accuracy.

The Producer's Accuracy reflects the probability that a given land cover type is correctly classified on the map, indicating how effectively the classification process identifies a specific category without omission errors, and is calculated as PA = 100% - Omission Error. Similarly, User's Accuracy represents the likelihood that a pixel classified as a specific land cover type truly belongs to that category in the reference data, providing a measure of classification reliability from the user's perspective, and is calculated as UA = 100% - Commission Error. Using this method, the classification accuracy of winter wheat distribution in Jiaozuo City for the 2019-2021 period was evaluated, as shown in Table 2.

Table 2. Accuracy of the extracted distribution area of winter wheat

Area (10,000 mu)	Overall Accuracy	Producer's Accuracy	Kappa Coefficient	User's Accuracy
45.303	0.952	0.952	0.952	0.985

Based on the classification results and accuracy assessment, the extracted winter wheat distribution area in Jiaozuo City was evaluated for precision. As shown in Table 2, the total extracted winter wheat planting area was 45.303 million mu. The classification achieved a high Overall Accuracy of 95.2%, indicating that most pixels were correctly classified. The Producer's Accuracy, also at 95.2%, suggests that the classification model effectively identified winter wheat areas with minimal omission errors. The Kappa Coefficient of 0.952 further confirms the strong agreement between the classification results and reference data, highlighting the reliability of the classification model [17,18].

Moreover, the User's Accuracy reached 98.5%, demonstrating that the classified winter wheat pixels closely corresponded to actual wheat-growing regions, with very few commission errors. These results indicate that the proposed method provides an effective and accurate approach for monitoring winter wheat distribution using remote sensing and cloud computing technologies.

4. Discussion and Conclusion

This study employs a widely used method for winter wheat identification and mapping, incorporating novel features as input data for the classifier. By integrating four key phenological stages of winter wheat and computing seven spectral indices for each stage, the method effectively enhances the spectral separability of winter wheat from other crops. Since data acquisition, processing, and analysis were conducted on the Google Earth Engine (GEE) cloud platform, this approach offers strong applicability for monitoring and analysis. The cloud-based platform allows for rapid adaptation to other regions or crop species, providing greater potential for application compared to traditional local computing methods.

Through this experiment, the estimated total winter wheat planting area in Jiaozuo City was approximately 1.45 million mu, which significantly deviates from the values reported in statistical yearbooks. However, the overall accuracy and Kappa coefficient were both relatively high, suggesting that classification errors might have resulted from the selection of

sample points, ultimately affecting classification precision [19,20].

Additionally, while this method demonstrates promising results in winter wheat extraction, it also presents notable limitations. One major drawback is its lack of timeliness, as the study was conducted after wheat maturity, making it unsuitable for real-time monitoring of annual wheat cultivation. Furthermore, the method is highly dependent on the selection of sample points. Since wheat planting patterns continuously evolve, sample points cannot be reliably used over extended periods. Therefore, there remains significant room for improvement in this approach.

References

[1] FAO. The future of food and agriculture – Trends and challenges. Rome: Food and Agriculture Organization of the United Nations, 2017.

[2] Gorelick, N., Hancher, M., Dixon, M., et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 2017, 202: 18-27.

[3] Cai Y, Guan K, Peng J, et al. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. Remote Sensing of Environment, 2018, 210: 35-47.

[4] Zhu, Z., Wulder, M.A., Roy, D.P., et al. Benefits of the free and open Landsat data policy. Remote Sensing of Environment, 2019, 224: 382-385.

[5] Yin, Jie, et al. "A comparative study on wheat identification and growth monitoring based on multi-source remote sensing data."Remote Sensing Technology and Application36.2 (2021): 332-341.

[6] Liang, Z. H. A. O., et al. "Large-scale winter wheat extraction based on GEE platform and automatic statistical allocation algorithm." HUBEI AGRICULTURAL SCIENCES 61.19 (2022): 132.

[7] Cai, Wenxin, Jinyan Tian, Xiaojuan Li, Lin Zhu, and Beibei Chen. 2022. "A New Multiple Phenological Spectral Feature for Mapping Winter Wheat" *Remote Sensing*14, no. 18: 4529.

[8] Zhang, H.; Du, H.; Zhang, C.; Zhang, L. An automated early-season method to map winter wheat using time-series Sentinel-2 data: A case study of Shandong, China.Comput. Electron. Agric.2021,182, 105962.

[9] Thenkabail, P.S., et al. An Automated Methodology for Mapping Global Croplands and Cropland Subtypes Using Landsat and MODIS Data. Remote Sensing of Environment, 2015, 162: 25-45.

[10] Xiao, X., et al. A Time-Series MODIS-NDVI Approach for Mapping Winter Wheat in China. Remote Sensing of Environment, 2005, 98: 441-457.

[11] Zhang, M., et al. Mapping Winter Wheat in China Using MODIS Time-Series Data. Remote Sensing of Environment, 2008, 112: 135-148.

[12] Atzberger, C. Advances in Remote Sensing of Agriculture: Context Description and Classification. Remote Sensing of Environment, 2013, 130: 85-110.

[13] Huete, A.R., et al. Overview of the Radiometric and Biophysical Performance of the MODIS Vegetation Indices. Remote Sensing of Environment, 2002, 83: 195-213.

[14] Gitelson, A.A., et al. Vegetation Index-Based Assessment of Winter Wheat Canopy Chlorophyll Content at the Plot Scale. Remote Sensing of Environment, 2002, 82: 70-85.

[15] Tucker, C.J. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. Remote Sensing of Environment, 1979, 8: 127-150.

[16] Huete, A.R., et al. A Soil-Adjusted Vegetation Index (SAVI) to Eliminate the Need for a Bare Soil Line. Remote Sensing of Environment, 1985, 18: 29-50.

[17] Jiang, Z., et al. Development of the Enhanced Vegetation Index. Remote Sensing of Environment, 2000, 73: 164-177.

[18] Meroni, M., et al. A Critical Analysis of the Use of Normalized Difference Vegetation Index (NDVI) for Crop Monitoring. Remote Sensing of Environment, 2004, 92: 429-441.

[19] Qi, J., et al. A Vegetation Index for the Remote Assessment of Plant Canopy Chlorophyll Content. Remote Sensing of Environment, 1994, 48: 43-54.

[20] Serra, P., et al. Mapping Winter Wheat in Spain Using MODIS Time-Series Data. Remote Sensing of Environment, 2007, 107: 345-358.