Evaluation of Graph-based Analysis for Forest Fire Detections

Young Gi Byun, Yong Huh, Kiyun Yu, Yong Il Kim

Abstract—Spatial outliers in remotely sensed imageries represent observed quantities showing unusual values compared to their neighbor pixel values. There have been various methods to detect the spatial outliers based on spatial autocorrelations in statistics and data mining. These methods may be applied in detecting forest fire pixels in the MODIS imageries from NASA’s AQUA satellite. This is because the forest fire detection can be referred to as finding spatial outliers using spatial variation of brightness temperature. This point is what distinguishes our approach from the traditional fire detection methods. In this paper, we propose a graph-based forest fire detection algorithm which is based on spatial outlier detection methods, and test the proposed algorithm to evaluate its applicability. For this the ordinary scatter plot and Moran’s scatter plot were used. In order to evaluate the proposed algorithm, the results were compared with the MODIS fire product provided by the NASA MODIS Science Team, which showed the possibility of the proposed algorithm in detecting the fire pixels.

Keywords—Spatial Outlier Detection, MODIS, Forest Fire

I. INTRODUCTION

Spatial outliers in remotely sensed imageries represent pixel values that are significantly different from their neighborhoods even though it is not a case in entire imagery scale [1], [2], [3]. Up to now, numerous kinds of spatial outlier detection algorithms have been developed in the fields of statistics and data mining.

A graph-based analysis is one of typical methods to define spatial outliers, which enables researchers to do visual analysis with 2-D graph based on ordinary scatter plot and Moran’s scatter plot [2]. The ordinary scatter plot method utilizes a linear regression equation of attribute values of locations and means of their neighbor values to find spatial outliers [1], [3]. In comparison, Moran’s scatter plot method uses correlations of attributive values of locations with its neighboring values and the entire data values (the total value?) [2], [4], [5], [6].

These methods may be applied in detecting the pixels indicating forest fire from the remotely sensed imageries. A breakout of forest fire in an area brings rapid change in brightness temperature of the corresponding pixels compared to its surroundings [7], [8]. Therefore, forest fire pixels can be treated as thermal anomaly in an image that represents land surface temperature [8].

With this point, forest fire detection becomes a process of finding spatial outliers which reflect local instability or extremity with respect to its neighboring pixel values. In this paper, we applied and tested the two aforementioned scatter plot methods for forest fire detection with MODIS fire detection with MODIS (Moderate Resolution Imaging Spectro-radiometer) sensor imageries from the AQUA, a satellite developed and controlled by NASA.

Traditionally, there have been three kinds of forest fire detection methods in remote sensing: the spectral method, the spatial method and the temporal method. The spectral method uses simple thresholds for single- or multi-band data to identify fire pixels [9]. The spatial method uses statistical characteristics of a local area, an average and a standard deviation of pixel values surrounding a pixel. In this case, every pixel is transformed to z-score values and a predefined threshold is used [10]. Note that the spectral method basically adopts a fixed threshold, while the spatial method employs variable thresholds [9]. The temporal method uses temperature differences between remotely sensed imageries acquired at different times [11]. At any rate, all of these methods have been modified according to their specific purposes and test regions for improved accuracy. As a noticeable case, the MODIS Science Team developed the contextual fire detection algorithm for MODIS based on the spatial method [12]. They aimed at a globally and practically applicable algorithm that takes into account relatively many factors, such as cloud effect, sun glint effect, desert boundary effect and coastal false effect, in order to prevent false detection [10]. As a result, the MODIS contextual algorithm is expected to be more stable than others despite the fact that it often misses small fires [9].

As the above literature shows, the approach proposed in this paper is far from that using the traditional methods. In this study, we wanted to know whether the graph-based analysis could be applicable to forest fire detection. Here, the ordinary scatter plot and Moran’s scatter plot method are to be employed. Consequently, in the following sections two scatter plot analyses are introduced along with the proposal and evaluation of a set of new algorithms based on these approaches. As for the
test imagery, a MODIS satellite data covering the South Korean peninsula is used, and the evaluation is carried out by comparing the results from the product provided by the NASA MODIS Science Team for the same data.

II. METHODOLOGY

All objects above absolute zero (-273°C or 0 K) emit radiant energy according to Plank’s function which explains the radiant energy by means of the wavelength observed and the object’s surface temperature [14]. Ideally, land surface temperatures of each band derived by inverse Plank’s function should be the same. However, due to many factors, such as reflected solar radiation, surface emissivity differences and water vapor attenuation, there are some differences in temperature between bands. The differences are larger when one part of a pixel is substantially warmer than the rest [12]. The hotter the object is, the more contribution in radiance is observed in the shorter wavelengths than in their longer counterparts. As a result, in the case of a physically hot phenomenon like a forest fire, surface temperatures of both bands increase, in which the surface temperature of the shorter wavelength band is dramatically higher than that of the longer wavelength band. The level of increase is proportional to the magnitude of fire [7]. Traditionally, bands of near 4 \( \mu m \) and 11 \( \mu m \) wavelengths are used in forest fire detection. Here, the former represents the shorter wavelength while the latter represents the longer wavelength band [7], [9], [12].

A. Graph-based Fire Detection

A breakout of forest fire in an area brings different changes in brightness temperature in 4 \( \mu m \) and 11 \( \mu m \) wavelengths, respectively [9]. Once a forest fire breaks out, brightness temperature in 4 \( \mu m \) changes significantly compared to the changes in 11 \( \mu m \) [7]. Considering such a phenomenon, pixels that are suspected to indicate forest fire are selected. Subsequently, the corresponding brightness temperature in 4 \( \mu m \) (T4) is used as a single spatial variable in the ordinary scatter plot and Moran’s scatter plot. Detailed results of the proposed algorithms are described in the following paragraphs.

1) Ordinary Scatter Plot

The ordinary scatter plot shows the brightness temperature T4 on the X-axis and the average brightness temperature T4 in the surrounding area on the Y-axis (Figure 1(a)). A least square regression line is used to identify abnormal pixels. In this plot, the fire pixels are marked with red points that show significant standardized errors from the least square regression line [1]. Two points are found above the line and the other is below.

Assuming that errors are normally distributed, a standardized residual

\[ s_x = \frac{(e - \mu_x)}{\sigma_x} > \theta \]

is a common test, in which \( \mu_x \) and \( \sigma_x \) represent the mean and standard deviation of the distribution of the error term \( e \), respectively. Pixels with

standardized residuals that are greater than \( \theta \) or less than \(-\theta\) are flagged as possible fire pixels. In Figure 1 (a), points (a), (b) and (c) turn out to be the farthest from the regression line and thus have a high potential to be fire pixels. As a threshold, we let \( \theta = 3 \) to detect the final fire pixels in the ordinary scatter plot method.

\[ I = \frac{n \sum \sum w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum \sum w_{ij}^2 \sum (x_i - \bar{x})^2} \]

Where, \( n \) indicates a total data count, \( \bar{x} \) the mean of total data,
x the spatial variable (T4), and \( w_{ij} \) the element of spatial weight matrix. By dividing the numerator and denominator of equation (1) with \( n^2 \), and using row-standardized spatial weight matrix, the global Moran’s I can be expressed in Z statistics as in equation (2).

\[
I = \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}^2 (\sum_{j=1}^{n} (x_j - \overline{x})^2) = \frac{1}{n} Z^2}
\]

(2)

In this study, we let \( w_{ij} = 1 \) if pixel \( i \) and pixel \( j \) were adjacent to each other; otherwise \( w_{ij} = 0 \). Therefore, \( W \) is an adjacency matrix for the observed region.

The local Moran’s I decomposes the global Moran’s I into contribution of each location. Therefore, the local Moran’s I can be used to identify fire pixels [6]. The local Moran’s I for an observation of \( i \) may be defined as

\[
I_i = z_i \times \sum_j w_{ij} z_j
\]

(3)

where the observations \( z_i, z_j \) are in deviations from the mean, and the \( w_{ij} \) is an element from a spatial weights matrix \( W \).

The Moran’s scatter plot is a plot of normalized brightness temperature value against the neighborhood of normalized brightness temperature value \((WZ)\) [6]. Where, \( W \) is the row-standardized spatial weight matrix. The row-standardized spatial weight matrix is created through the process of dividing the each element in the weight matrix by its row sum. The Moran’s scatter plot contains four quadrants. The upper left and lower right quadrants of figure 1(b) indicate a spatial association of dissimilar values. Spatial outliers can be identified form these two quadrants. However, in this paper we decided that pixels such as point (c), which is surrounded by low value neighbors in Moran’s scatter plot, are final fire pixels because forest fire areas have a higher brightness temperature than their neighbors.

### III. IMPLEMENTATION DETAILS OF ALGORITHM

We thus propose new graph-based fire detection algorithms based on the aforementioned characteristics of forest fire pixels in surface temperature images. Figure 2 presents a flowchart of this study, which can be divided into three parts, a preprocessing of MODIS L1B data, an implementation of the forest fire detection algorithm, and an evaluation of the proposed algorithm.

The pre-processing is composed of a geometric correction, subset extraction of study area, and transformation of data from a radiance unit to temperature unit. We used ENVI 3.5, a satellite image processing software for the pre-processing.

To derive temperatures of each band, offset and scale data from a metadata of MODIS L1B data set is used to convert the digital number (DN) of the sensor measurements to spectral radiance measured by satellite sensors.

Generally, a surface temperature is derived by integrating the product of Planck’s function and the relative spectral response function of a sensor within a bandwidth [7]. However, for the sake of convenience, we utilized a simple surface temperature estimation equation developed by CIMSS (Cooperative Institute for Meteorological Satellite Studies) of Wisconsin University, which provides an effective central wavelength and a slope and offset for temperature estimation according to each MODIS band.

Next, we performed a potential fire pixel test to reduce the execution time of the algorithm and to make the detection process more effective. This is done by removing obviously non-fire pixels such as clouds pixels or cold land pixels and identifying potential fire pixel which is probable to be real forest fire pixels. Then we applied the proposed fire detection algorithms to the potential fire pixels.

The non-fire pixels, such as cloud or water body pixels can cause serious problem in the contextual fire detection [15]. Thus, detections and removals of such pixels are important and necessary in the preprocessing. However, it is very difficult to identify exact cloud pixels because there exist not only thick clouds but also thin clouds as well.

In general, clouds have low radiance value in 4 \( \mu m \) and 12 \( \mu m \) band compared to land surface. From this we can detect

---

**TABLE I**

<table>
<thead>
<tr>
<th>Method</th>
<th>Ordinary scatter plot</th>
<th>Moran’s scatter plot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference function</td>
<td>( E = b x - (m_0 f(x) + b) )</td>
<td>( Z_i = \frac{I_i - \mu_i}{\sigma_i} )</td>
</tr>
<tr>
<td>Statistical test function</td>
<td>( E - \mu_i ) ( \sigma_i ) ( \theta )</td>
<td>lower right quadrants</td>
</tr>
<tr>
<td>Aggregate function</td>
<td>( \mu_i, \sigma_i, m, b )</td>
<td>( \mu_i, \sigma_i )</td>
</tr>
</tbody>
</table>
cloud pixels. In this study, we identified potential fire pixels using two criteria:

\[
\text{Criterion 1: } R_s > \max(R_{12}, R_{44}) \quad \text{and} \quad R_{12} > \max(R_{44}, R_{22})
\]

(4)

Here, \(R_s\), \(R_{12}\), \(R_{44}\), and \(R_{22}\) indicate radiances of 4 \(\mu\)m and 12 \(\mu\)m bands, respectively. On the other hand, \(\overline{R}_4\), \(\overline{R}_{12}\) indicate mean radiances of 4 \(\mu\)m and 12 \(\mu\)m bands in full image, respectively, whereas \(\overline{R}_{44}\), \(\overline{R}_{22}\) indicate mean radiances of 4 \(\mu\)m and 12 \(\mu\)m bands in local window. Such simple criteria can screen out more than half of the pixels to be processed and eliminate cloud and sun glint pixels [8].

Finally, we calculated the user accuracy and producer accuracy of the proposed algorithms by comparing their detection results with the ground data provided by the Korea Forest Service (KFS) [17]. KFS has built a forest fire database in terms of igniting and extinguishing times, locations, burnt areas, and other related information. Subsequently, the performance of the algorithms is evaluated through the comparison of the resulting accuracy with that of the MODIS fire product provided by the NASA MODIS Science Team.

IV. STUDY AREA AND MATERIALS

The MODIS L1B data used in this paper is a set of imageries covering the South Korean peninsula taken by NASA’s AQUA satellite. A metadata of these imageries is shown in Table 2. We used five imageries taken on April 5th and 11th of 2005, February 9th and 14th of 2004, and April 2nd of 2003.

![Fig. 3 MODIS RGB image of South Korea peninsula](image)

**TABLE 2**

<table>
<thead>
<tr>
<th>Satellite Name</th>
<th>AQUA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor Used</td>
<td>MODIS</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Spatial Resolution</th>
<th>1km</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Band</td>
<td>36</td>
</tr>
<tr>
<td>Area Covered</td>
<td>South Korea</td>
</tr>
<tr>
<td>Date of Acquisition</td>
<td>2003.04, 2004.02, 2005.04</td>
</tr>
</tbody>
</table>

The MODIS bands used in our study are listed in Table 3. In detecting forest fire from the MODIS imagery, band 21 as well as band 22 was used to solve the problem of sensor saturation. The above-mentioned two bands have the same band width. Sensor saturation takes place when radiance that reaches the sensor is beyond the upper limit of the sensor’s sensing range [12]. Radiance subject to sensor saturation is recorded as the upper limit. In other words, all radiances higher than the upper limit of the sensing range are recorded as the upper limit. As a result, when radiance is higher than the upper limit, accurate radiance cannot be identified. In band 22 sensor saturation takes place when brightness temperature is over 330K, and in band 21 it takes place when it is over 550K [12]. Therefore, the problem of sensor saturation was solved by using band 21 when sensor saturation took place in band 22.

**TABLE 3**

<table>
<thead>
<tr>
<th>Band number</th>
<th>Bandwidth ((\mu)m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>3.930-3.989</td>
</tr>
<tr>
<td>22</td>
<td>3.930-3.989</td>
</tr>
<tr>
<td>31</td>
<td>10.780-11.280</td>
</tr>
<tr>
<td>32</td>
<td>11.770-12.270</td>
</tr>
</tbody>
</table>

Reference data for accuracy assessment was provided by the Korea Forest Service forest fire information system. The KFS data included the original location of forest fires, the time of breakout and extinction, the size of damaged areas, and the cause of forest fires. In making use of the data, it was troublesome because the data was not referenced to geographic coordinates but addresses. Considering the spatial resolution of MODIS (1km on ground), we expected that there would be little possibility of misclassification from converting the addresses to approximate geographic coordinates. Thus, it was done as such. Regarding the time of image acquisition, the AQUA satellite scans Korea at around 13:30pm that forest fires that were running through this time was recorded. The size of damage area recorded in the KFS data can be tens to hundreds times as large as that of burnt area. Theoretically, MODIS can sense the forest fire with the temperature over at least 800~1000K and the area over 0.01ha (100m²), and the detection accuracy is about 50% [10]. Therefore, the forest fire recorded in Korea Forest Service data was selected only when the size of damaged area was over 0.1ha.
V. RESULT AND ANALYSIS

In this paper, in addition to the accuracy assessment, i.e., user accuracy and producer accuracy calculation, we compared the results with the one from the product provided by the NASA MODIS Science Team to examine applicability of the algorithm. The NASA team used the contextual algorithm to detect thermal anomalies, such as forest fires or volcanic eruptions [10], [16].

The contextual algorithm uses statistical analysis of surface temperatures in a local window centered on a pixel, in which discrimination is carried out in order to find out whether or not it is a fire pixel [10]. Instead of using fixed thresholds for determining fire pixels, the algorithm computes a mean and standard deviation of the temperature in a local window, and then thresholds are determined using equations as follow:

\[
T_4 \geq \mu_4 + \alpha \sigma_4 \\
T_{4-11} \geq \mu_{4-11} + \beta \sigma_{4-11}
\]  

Here, \( T_4 \) is the surface temperature of observing bands near 4 \( \mu m \) and \( T_{4-11} \) is temperature difference between 4 \( \mu m \) and 11 \( \mu m \) of a local window. \( \mu_4 \) and \( \mu_{4-11} \) are means of \( T_4 \) and \( T_{4-11} \), \( \sigma_4 \) and \( \sigma_{4-11} \) are standard deviations of \( T_4 \) and \( T_{4-11} \), respectively, and \( \alpha \) and \( \beta \) are constant values.

Fig. 4 Detection results using ordinary scatter plot(a), Moran’s scatter plot(b), MODIS fire product(c)

Thus, the proposed fire detection algorithms were applied on the imageries and accuracies were assessed by checking the classified fire pixels against the data from the Korea Forest Service. To understand mechanisms between the commission and omission error, both the producer accuracy and user accuracy were calculated. Then, the same imageries were used by the method provided by the NASA MODIS Science Team and both the producer and user accuracy is calculated. The test results are shown in Table 4.

<table>
<thead>
<tr>
<th>TABLE 4 ACCURACY TEST RESULTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinary Scatter Plot</td>
</tr>
<tr>
<td>Total Number of Detected Fire Pixels</td>
</tr>
<tr>
<td>Number of True Fire Pixels in Detected Pixels</td>
</tr>
<tr>
<td>Number of Omission Error</td>
</tr>
<tr>
<td>Number of Commission Error</td>
</tr>
<tr>
<td>Producer Accuracy(%)</td>
</tr>
<tr>
<td>User Accuracy(%)</td>
</tr>
</tbody>
</table>

As shown in Table 4, the proposed algorithm based on the ordinary scatter plot showed a producer accuracy which was about 16% higher than that of the MODIS fire project. On the other hand, the user accuracy turned out to be 3% higher than that of the MODIS fire product. A similar trend appears in Moran’s scatter plot case; about 6% higher in producer accuracy and about 15% higher in user accuracy than that of the MODIS fire product.

The ordinary scatter plot fire detection did not improve user accuracy considerably, whereas it reduced the commission error significantly compared to the MODIS fire product. In contrast, the Moran’s scatter plot fire detection had high user accuracy with relatively high commission error. Such results were partly due to the incineration plant and the dense industrial area. These areas were misclassified as forest fires and such misclassifications were included in the commission error.

Moreover, the ordinary scatter plot algorithm was problematic in that large forest fires with extremely high temperatures were so influential in OLS (Ordinary Least Square) estimation that they made it difficult to detect small forest fires. On the other hand, Moran’s scatter plot algorithm showed the problem of not considering the amount of relative spatial variation because it detected forest fire pixels based only on the location of the pixel in Moran’s scatter plot.
VI. CONCLUSION

This paper proposed graph-based forest fire detection algorithms which detect forest fire pixels in MODIS L1B imagery. The algorithm used spatial statistics reflecting spatial variation. After applying and testing the algorithms, the following results are acquired.

The algorithms yielded higher user and producer accuracies than those of the MODIS fire product provided by the NASA MODIS Science Team. Such a result does not mean that the proposed algorithms work better than the traditional fire detection methods developed so far. Rather, it implies the proposed algorithms have potentials to be applied in detecting the fire pixels and thus we need to put more intention on it. Looking at the algorithms, they are easy to use because they do not need the step for determining complicated thresholds despite the spatial variation. After applying and testing the algorithms, they are easy to use because they do not need the step for determining the threshold empirically according to the target regions. Thus, the algorithms have advantages over the traditional methods.

From the test, the ordinary scatter plot algorithm was proven problematic in that it was not sensitive to small forest fires, while Morán’s scatter plot was also weak for it entailed a more complicated threshold. Moreover, the ordinary scatter plot algorithm yielded higher accuracies as the MODIS fire product. An outcome does not mean that the proposed algorithms work better than the traditional fire detection methods developed so far. Rather, it implies the proposed algorithms have potentials to be applied in detecting the fire pixels and thus we need to put more intention on it. Looking at the algorithms, they are easy to use because they do not need the step for determining the threshold empirically according to the target regions. Thus, the algorithms have advantages over the traditional methods.

Thus, future research may put its focus on relieving such problems. One idea is in introducing spatial weight matrices that reflects the spatial variation of forest fire to decrease the commission error.

REFERENCES


Young Gi Byun received the BS degree in mineral and energy resources engineering from Chonnam National University and MS degree in Remote Sensing from Seoul National University in South Korea in 2004. He is now a doctoral course in Seoul National University in Korea. He is now working on forest fire detection and mapping using RS techniques. His research interests lie in the area of image matching, pattern recognition, and image processing.

Yong Huh received the BS degree in civil, urban & geosystem engineering from Seoul National University in South Korea in 2001. He is now a doctoral course in Seoul National University. He is now working on forest fire detection and mapping using RS techniques. His research interests lie in the area of digital image processing, pattern classification.

Kiyun Yu received the BS degree in civil engineering from Yonsei University in South Korea and PhD degree in GIS from University of Wisconsin-Madison in 1998. He was a director in the Ministry of Construction and Transportation until 2000. He is now an Assistant Professor of the School of Civil, Urban & Geosystem Engineering, Seoul National University in Korea.

Yong Il Kim received the BS degree and the PhD degree in Remote Sensing from Seoul National University in South Korea in 1988 and 1991, respectively. He is now an Associate Professor of the School of Civil, Urban & Geosystem Engineering, Seoul National University in Korea.