LAND-COVER MAPS USING MULTIPLE CLASSIFIER SYSTEM FOR POST-DISASTER LANDSCAPE MONITORING

Hidetake Hirayama [1]*, Mizuki Tomita[2], Ram C. Sharma[2], Keitarou Hara[2]

[1] Graduate School of Tokyo University of Information Sciences, 4-1Onaridai, Wakaba-ku, Chiba, 265-8501 Japan –
h17002hh@edu.tuis.ac.jp
[2] Tokyo University of Information Sciences, 4-1Onaridai, Wakaba-ku, Chiba, 265-8501 Japan

ABSTRACT:

Recently, land cover maps created from high resolution satellite images have been used for landscape analysis, in order to understand the impact of natural disasters on biodiversity and ecosystems. Conventional land cover classification methods, however, suffer from problems with isolated pixels (salt and pepper effect). Filtering can remove the isolated pixels, but can also result in loss of accurate information. The purpose of this study is to create a land cover map for landscape analysis of large-scale disturbances caused by the Great East Japan Earthquake of 2011, utilizing a Multiple Classifier System (MCS), which allows for reduction of isolated pixels while maintaining classification accuracy. RapidEye satellite images covering the Pacific Ocean side of the Tohoku district damaged by the earthquake and subsequent tsunami were obtained for 2010, 2011, 2012 and 2016, and land cover classification was implemented using individual classifiers and the MCS method. The results showed that the MCS land cover map was able to reduce the number of isolated pixels significantly (61-71%) compared with the individual classifiers, while maintaining very high accuracy (0.976-0.986) for all four years. These results indicate that MCS land cover maps have a great potential for analyzing disturbances following infrequent large-scale natural disasters such as earthquakes and tsunami, and for monitoring the process of recovery afterwards. We expect that the results of this research will be useful in managing the recovery process in the region disturbed by the Great Eastern Japan Earthquake and Tsunami of 2011, and also for developing future Ecosystem-based Disaster Risk Reduction programs for the region.

1. INTRODUCTION

The occurrence of a natural disasters is increasing year by year, causing serious damage worldwide (Ritchie and Roser, 2019). In the case of Japan, earthquake-related disasters occur frequently, due to high levels of volcanic and seismic activity. For example, the Great East Japan Earthquake (magnitude 9.0) and subsequent tsunami occurred on March 11, 2011, resulting in 21,000 dead and missing persons (Cabinet Office Govermant of Japan, 2015). This earthquake and tsunami can be regarded as a typical example of a "Large, Infrequent Disturbance" (Tomita et al., 2014), which not only results in loss of lives, but also results in dramatic changes in land cover, and has a major impact on biodiversity and ecosystems (Hara, 2014; Hara et al., 2016). Satellite remote sensing technology is considered to be an effective tool for monitoring natural disasters (Römer et al., 2012). Analysis of satellite imagery has been utilized to analyze changes in land cover before and after the Great East Japan Earthquake (Harada et al., 2015; Ishihara and Tadono, 2017); and studies using high-resolution satellite images have shown that 90% (4.2 km^2 to 0.5 km^2) of coastal forests in the Sendai Bay coastal area were destroyed by the tsunami (Zhao et al., 2013). As the region slowly recovers, continuous monitoring of the changing landscape structure will be essential for conserving biodiversity and ecosystems.

In recent years, machine learning methods have made it possible to generate highly accurate land cover maps. These maps, however, suffers from a problem known as the salt and pepper effect (Blaschke et al., 2000; Zhai et al., 2017). This problem involves occurrence of isolated pixels with high local spatial heterogeneity between neighboring pixels in the classification results. Isolated pixels reduce visibility of the results, and can lead to overestimation of analysis results. From the standpoint of landscape ecology, this problem reduces the ability to evaluate the degree to which the landscape facilitates or impedes

movement of plants and animals among resource patches (Taylor et al., 1993).

Filter processing, performed before or after the classification process, has been utilized to minimize the effect of isolated pixels (Eastman, 2003; Zhu, 2013). This method, however, also involves a risk of deleting correct information along with the unwanted isolated pixels (Eastman, 2003). Segment-based analysis methods have also been proposed as a means for preventing isolated pixels (Blaschke et al., 2000). In this approach, however, parameter selection is qualitative and complex (Neubert et al., 2006). The Multiple Classifier System (MCS), has also been utilized for pixel-based land cover classification (Du et al., 2012; Kittler et al., 1998). Our previous research (Hirayama et al., 2018), has shown that the MCS method is effective for reducing the number of isolated pixels that occur on land cover maps while retaining classification accuracy.

The purpose of this study is to utilize the MCS method for creating a land cover map for landscape analysis of large-scale disturbances caused by the Great East Japan Earthquake of 2011. RapidEye satellite images covering the Pacific Ocean side of the Tohoku district damaged by the earthquake and subsequent tsunami were obtained for 2010 (before the disaster), 2011, 2012 and 2016, and land cover classification was implemented using five individual classifiers and the MCS method. The results for the five individual classifiers and the MCS method were compared in terms of classification accuracy (Kappa coefficient), and occurrence of isolated pixels.

2. DATA AND METHODS

2.1 Research area

The target area, shown in Figure 1, covered an area along the Pacific Ocean side of Miyagi Prefecture, in the Tohoku Region of northern Honshu Island. This area was heavily damaged by the

---
* Corresponding author

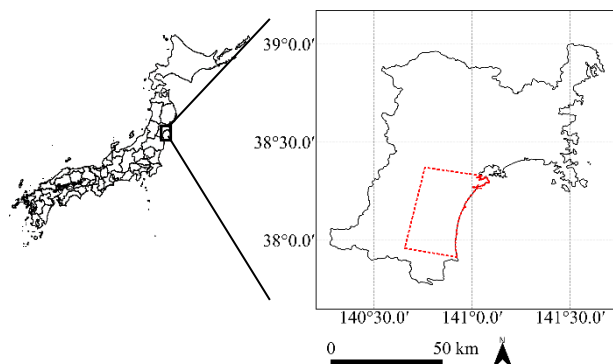earthquake and subsequent tsunami, and it was placed in the red dotted box (Figure 1).

Figure 1. Location of study area; The red dotted line shows the area for land cover classification.

2.2 Data

2.2.1 Training and test data

Ground truth data were prepared based on field survey information and visual interpretation using Google Earth images. Six land cover classes (forest, shrub/grassland, cropland, urban area, waterbody and bare ground) were established. Ground truth data were collected at 2,800 points for each class. Of the 2,800 ground truth points, 2,100 points (75%) were used as training data and 700 points (25%) as test data.

2.2.2 Satellite and elevation data

RapidEye satellite data with spatial resolution of 6.5 m, were obtained for 4 April 2010, 13 April 2011, 10 April 2011 and 5 April 2016. The data consisted of atmosphere reflectance in five channels (Blue, Green, Red, Red edge, and Near-infrared). Five spectral indices, Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), NDVI red edge (NDVIre), Red edge NDVI (ReNDVI), and Normalized Difference Water Index (NDWI), were calculated pixel by pixel, employing the respective equations noted in Table 1. In addition, the research also employed 5.0 m spatial resolution Digital Elevation Model (DEM) data, available from the Ministry of Land, Infrastructure, Transport and Tourism of Japan, as well as land surface slope (in degrees) data, generated from the DEM data.

Index	Equation	Reference
NDVI	$\dfrac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}}$	Rouse et al., 1974
EVI	$2.5 \dfrac{\rho_{nir} - \rho_{red}}{\rho_{nir} + 6.0\rho_{red} - 7.5\rho_{blue} + 1.0}$	Huete et al., 2002
NDVIre	$\dfrac{\rho_{nir} - \rho_{rededge}}{\rho_{nir} + \rho_{rededge}}$	Gitelson and Merzlyak, 1994
ReNDVI	$\dfrac{\rho_{rededge} - \rho_{red}}{\rho_{rededge} + \rho_{red}}$	Sims and Gamon, 2002
NDWI	$\dfrac{\rho_{green} - \rho_{nir}}{\rho_{green} + \rho_{nir}}$	Mcfeeters, 1996

Table 1. Equations for calculating spectral indices. ρ_{blue}, ρ_{green}, ρ_{red}, $\rho_{rededge}$ and ρ_{nir} refer to the reflectance obtained from Blue, Green, Red, Red edge, and Near infrared channels of the RapidEye image.

2.3 Generation of land cover maps

A total of five recently available machine learning classifiers (Random forest (RF); (Breiman, 2001), Bagging (BAG);

(Breiman, 1996), XGBoost (XGB); (Chen and Guestrin, 2016), Support vector machine (SVM); (Mountrakis et al., 2011), K-nearest neighbor (KNN); (Beckmann et al., 2015)) ware used as individual classifiers. The parameters for each classifier were selected by the highest kappa coefficient value, using a trial and error method. For the MCS method, the classification results obtained from the five individual machine learning classifiers were aggregated by calculating the mode value from the result of each classifier (Figure 2). The classification accuracies for the individual classifiers and MCS method were evaluated using Kappa coefficient; and the numbers of isolated pixels were counted to evaluate the reduction of the salt and pepper effect.
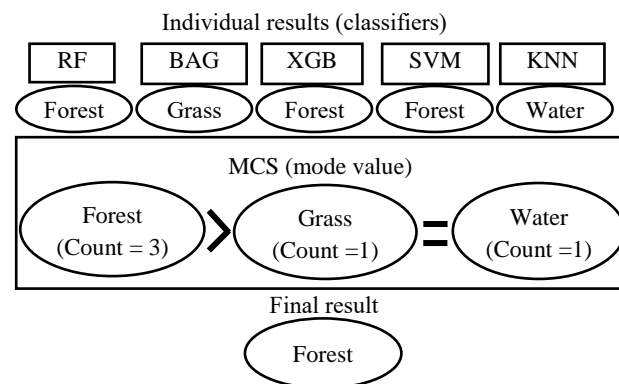

Figure 2. Flow chart of the classification.

3. RESULTS

3.1 Accuracy assessment

Kappa coefficients for the five individual classifiers and the MCS method are shown in Table 2. The results show that for all four years the classification accuracy was high for all the individual classifiers, especially for RF, BAG, and XBG, with values ranging from 0.969 to 0.986 in all years; The SVM and KNN classifiers were slightly lower, but still above 0.8 for all years. The accuracy of the MCS, 0.976-0.986, compared favorably with the best of the individual classification results.

		Kappa coefficient			
		2010	2011	2012	2016
Classifier	RF	0.982	0.979	0.970	0.977
	BAG	0.976	0.977	0.965	0.969
	XGB	0.986	0.984	0.975	0.980
	SVM	0.895	0.912	0.872	0.860
	KNN	0.863	0.890	0.840	0.834
	MCS	0.986	0.984	0.976	0.984

Table 2. Kappa coefficient for five individual classifiers and MCS aggregation.

3.2 Generation of isolated pixels

The number of isolated pixels for each of five classifiers and the MCS method are shown in Figure 4. Among the isolated classifiers, KNN generated the highest number of isolated pixels, followed by SVM, BAG, XGB, and RF. The number of isolated pixels MCS generated by the MCS method amounted to 61-71% of the mode value of individual classifiers for each year.
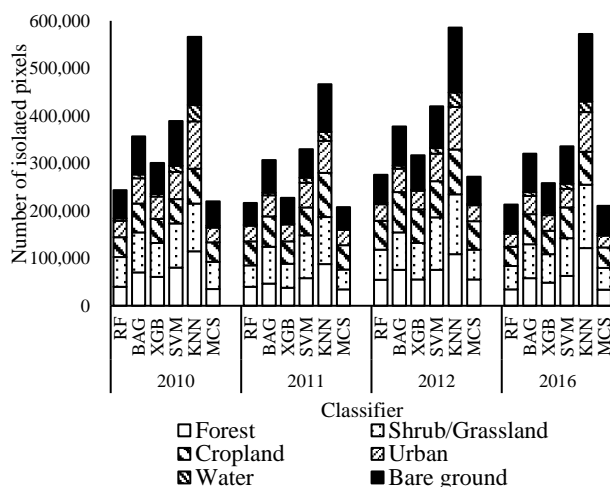
Figure 3. Number of isolated pixels by land cover class for five individual classifiers and MCS aggregation.

3.3 Broad scale land cover maps

The land cover maps generated by the MCS classification are shown in Figure 4. These maps folow the changes in land cover from before the earthquake and tsunami (2010) to immediately after (2011) to 6 years later (2016). Evem at this broad scale, a comparison of the 2010 (4a) and 2011 (4b) maps demonstrates that most of the forests stretching along the coast from north to south were nearly completely destroyed by the tsunami. In addition, vast areas of cropland on the alluvial plain inland from the coast changed to bare ground. The two subsequent maps show that some of the cropland areas ware already recoverd by 2012 (4c), and that this recovery was almost complete by 2016 (4d).
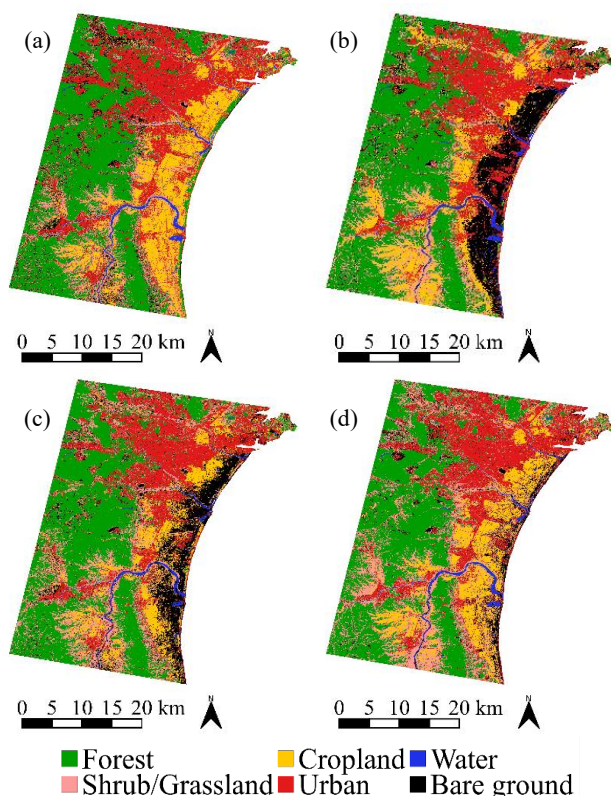

Figure 4. Land cover maps obtained with MCS aggregation: (a) 2010, (b) 2011, (c) 2012, (d) 2016.

3.4 Changes in Coastal Forest

In order to take a closer look at the disturbance to the coastal forest ecosystem caused by the earthquake and tsunami, the MCS land cover maps for each of the four years were compared to contemporary Google Earth images. The results are shown in Figure 5.
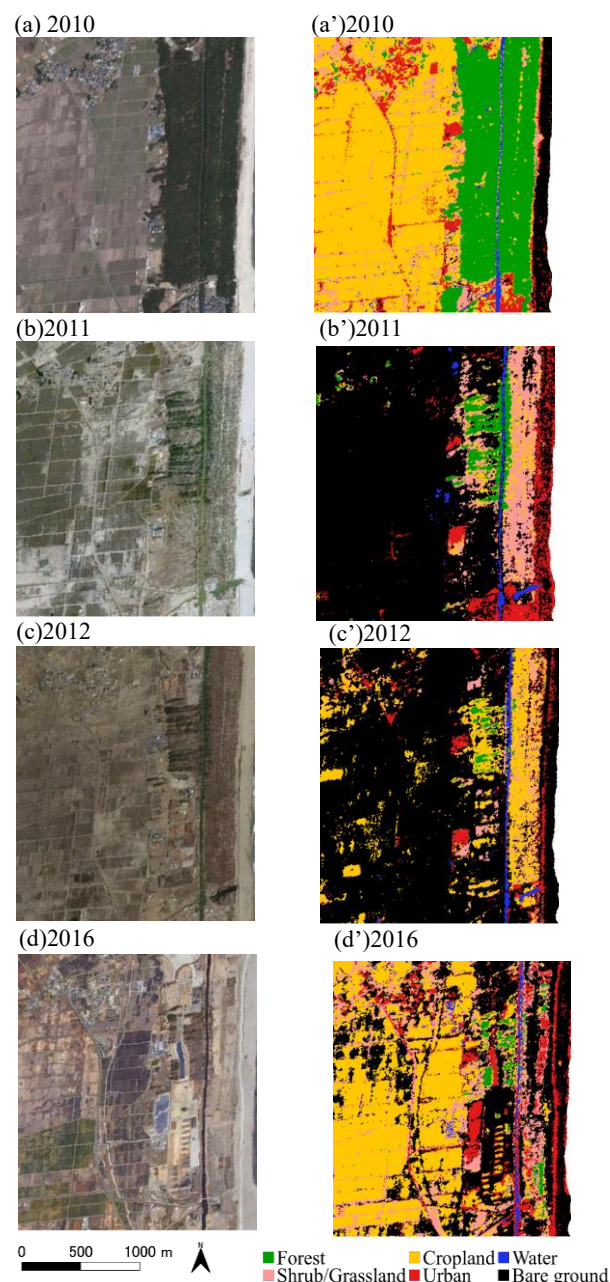

Figure 5. Detailed map of isolated pixels produced by MCS aggregation: (a) to (d) are Google Earth images. (a') to (d') are landcover maps.

The maps for 2010 (5a, 5a'), before the disaster, show a wide belt of pine forest running parallel to the coast. These pines had been planted along the raised banks of a central canal. Immediately after the disaster (5b, 5b'), almost all the pines have disappeared, with the exception of several narrow strips on the inland side of the canal. These strips of residual forest remain, although slightly diminished, in both the 2012(5c, 5c') and 2016

(5d, 5d') maps. In addition, the 2016 maps show some recovery of forest on the seaward side of the bank. Small patches of forest, representing planted groves surrounding farmsteads on the alluvial plain, can also be seen in the upper left corner of the pre-disaster land cover map and Google image. The subsequent maps and images indicate that these groves survived the immediate disaster, but were later reduced or disappeared.

4. CONCLUSIONS

The MCS aggregation resulted in a significant reduction (61-71%) in isolated pixels when compared to the mean of the five individual classifiers, while at the same time producing excellent kappa coefficients (0.98 to 0.99) that exceeded those of all the individual classifiers. These results clearly show that the MCS aggregation method is capable of reducing the effect of isolated pixels, while at the same time retaining a high classification accuracy. In addition, a detailed comparison of land cover maps generated by the MCS method with ground truth Google Earth images showed that the maps accurately portray changes in land cover over a six year period.

5. DISCUSSION

Classification of high spatial resolution satellite data is a promising tool for evaluating and monitoring changes in land cover due to large, infrequent disaster such as earthquakes and tsunami. Such classifications, however, suffer from problems with isolated pixels, which compromise the effectiveness of the system for evaluating important ecological variables such as connectivity between patches. Extant methods for reducing isolated pixels, such as filtering, can result in loss of classification accuracy. This study tested the MCS method, which aggregates the results of multiple individual classifiers, as a tool for reducing isolated pixels. The results showed that the MCS is capable of reducing the effect of isolated pixels while maintaining high classification accuracy. The results also demonstrated the MCS method's ability to accurately monitor changes in land cover in a target area that was heavily damaged by the Great East Japan Earthquake and Tsunami of 2011. Hopefully the results of this study will prove useful in monitoring future changes in the region, and also in designing recovery and restoration projects that incorporate habitat connectivity and other ecological concerns.

ACKNOWLEDGEMENTS

REFERENCES

Beckmann, M., Ebecken, N.F.F., Lima, B.S.L., Lima, P. De, 2015. A KNN Undersampling Approach for Data Balancing. J. Intell. Learn. Syst. Appl. 7, 104–116. https://doi.org/10.4236/jilsa.2015.74010

Blaschke, T., Lang, S., Lorup, E., Strobl, J., Zeil, P., 2000. Object-Oriented Image Processing in an Integrated GIS / Remote Sensing Environment and Perspectives for Environmental Applications. Environ. Inf. Planning, Polit. Public 555–570.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., 1996. Bagging Predictors. Mach. Learn. 24, 123–140. https://doi.org/10.1023/A:1018054314350

Cabinet Office Govermant of Japan, 2015. Disaster Mnagement in Japan.

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System. Kdd 1–10. https://doi.org/10.1145/2939672.2939785

Du, P., Xia, J., Zhang, W., Tan, K., Liu, Y., Liu, S., 2012. Multiple classifier system for remote sensing image classification: A review. Sensors 12, 4764–4792. https://doi.org/10.3390/s120404764

Eastman, J.R., 2003. IDRISI Kilimanjaro: guide to GIS and image processing. Worcester, MA: Clark Labs, Clark University.

Gitelson, A., Merzlyak, M.N., 1994. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus-hippocastanum L. and Acer-platanoides L. Leaves - Spectral Features and Relation to Chlorophyll Estimation. J. Plant Physiol. 143, 286–292. https://doi.org/10.1016/S0176-1617(11)81633-0

Hara, K., 2014. Damage to coastal vegetation due to the 2011 tsunami in northeast Japan and subsequent restoration process: analyses using remotely sensed data. Glob. Environ. Res. 18, 27–34.

Hara, K., Zhao, Y., Tomita, M., Kamagata, N., Li, Y., 2016. Impact of the Great East Japan Earthquake and Tsunami on Coastal Vegetation and Landscapes in Northeast Japan: Findings Based on Remotely Sensed Data Analysis, in: Urabe, J., Nakashizuka, T. (Eds.), . Springer Japan, Tokyo, pp. 253–269. https://doi.org/10.1007/978-4-431-56448-5_16

Harada, I., Hara, K., Park, J., Asanuma, I., Tomita, M., Hasegawa, D., Short, K., Fujihara, M., 2015. Monitoring of rapid land cover changes in eastern Japan using Terra/MODIS data. ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XL-7/W3, 403–408. https://doi.org/10.5194/isprsarchives-XL-7-W3-403-2015

Hirayama, H., Sharma, R.C., Tomita, M., Hara, K., 2018. Evaluating multiple classifier system for the reduction of salt-and-pepper noise in the classification of very-high-resolution satellite images. Int. J. Remote Sens. 40, 2542–2557. https://doi.org/10.1080/01431161.2018.1528400

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213.

Ishihara, M., Tadono, T., 2017. Land cover changes induced by the great east Japan earthquake in 2011. Sci. Rep. 7, 45769. https://doi.org/10.1038/srep45769

Kittler, J., Hatef, M., Duin, R.P.W., Matas, J., 1998. On Combining Classifiers. IEEE Trans. Pattern Anal. Mach. Intell. 20, 226–239. https://doi.org/10.1109/34.667881

Mcfeeters, S.K., 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. 17, 1425–1432. https://doi.org/10.1080/01431169608948714

Mountrakis, G., Im, J., Ogole, C., 2011. Support vector machines in remote sensing: A review. ISPRS J. Photogramm. Remote Sens. 66, 247–259.

Neubert, M., Herold, H., Meinel, G., 2006. Evaluation of Remote Sensing Image Segmentation Quality – Further Results and Concepts. 1st Int. Conf. Objectbased Image Anal. OBIA06 1–6.

Ritchie, H., Roser, M., 2019. Natural Disasters [WWW Document]. OurWorldInData.org. URL https://ourworldindata.org/natural-disasters (accessed 6.4.19).

Römer, H., Willroth, P., Kaiser, G., Vafeidis, A.T., Ludwig, R., Sterr, H., Revilla Diez, J., 2012. Potential of remote sensing techniques for tsunami hazard and vulnerability analysis -a case study from Phang-Nga province, Thailand. Nat. Hazards Earth Syst. Sci. 12, 2103–2126. https://doi.org/10.5194/nhess-12-2103-2012

Rouse Jr, J., Haas, R.H., Schell, J.A., Deering, D.W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. NASA Spec. Publ. 351, 309.

Sims, D.A., Gamon, J.A., 2002. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. Remote Sens. Environ. 81, 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X

Taylor, P.D., Fahrig, L., Henein, K., Merriam, G., 1993. Connectivity is a vital element of landscape structure. Oikos 68, 571–573. https://doi.org/10.2307/3544927

Tomita, M., Hirabuki, Y., Kanno, H., Hara, K., 2014. Influences of Large, Infrequent Disturbance Caused by Tsunami on Coastal Forest Communities. Japanese J. Conserv. Ecol. 19, 163–176. https://doi.org/10.1007/978-4-431-56448-5_22

Zhai, H., Zhang, H., Zhang, L., Li, P., Plaza, A., 2017. A New Sparse Subspace Clustering Algorithm for Hyperspectral Remote Sensing Imagery. IEEE Geosci. Remote Sens. Lett. 14, 43–47.

Zhao, Y., Tomita, M., Hara, K., 2013. Landscape change analysis before and after the earthquake disaster in Sendai coastal area by using SPOT satellite data. J Nat Restor Conserv 6, 43–49.

Zhu, X., 2013. Land cover classification using moderate resolution satellite imagery and random forests with post-hoc smoothing. J. Spat. Sci. 8596, 1–15. https://doi.org/10.1080/14498596.2013.819600