

# MAPPING FOREST DISTURBANCE USING PURE FOREST INDEX TIME SERIES AND CCDC ALGORITHM

Yaotong Cai<sup>1</sup>, Qian Shi<sup>1,\*</sup>, Xiaoping Liu<sup>1</sup>

<sup>1</sup> School of Geography and Planning, Sun Yat-Sen University, China – caiyt33@mail2.sysu.edu.cn, (shixi5, liuxp3@mail.sysu.edu.cn

Commission VI, WG III/IVa

**KEY WORDS:** Mapping, Change detection, Forest disturbance, Vegetation index, Time-series analysis, Spectral mixture analysis, Pure Forest Index (PFI)

## ABSTRACT:

Forest dynamics are closely related to climate change, natural disasters, and ecological diversity. The accumulated Landsat archive provides an unprecedented opportunity for long-term forest dynamics monitoring globally. However, using Landsat time series to detect small-scale and low-intensity disturbance events is still challenging since the moderate spatial resolution of Landsat images and the mixed pixel problem. Towards improving the ability of vegetation index (VI) in characterizing sub-pixel forest dynamics, this paper introduced the spectral mixture analysis (SMA) to develop a novel Pure Forest Index (PFI). The Continuous Change Detection and Classification (CCDC) algorithm was used to detect forest disturbance based on the PFI time series. Cross-comparison shows that PFI is far superior to other conventional VI in indicating forest conditions since it can enhance the spectral signal of the forest and suppress noises from the background. Time series analysis further demonstrates the superiority of PFI in accurately characterizing forest dynamics. The high overall accuracy of 0.96 for the forest disturbance map generated by the proposed approach was achieved. This study highlights a novel VI for accurately tracking subtle forest changes in a heterogeneous landscape.

## 1. INTRODUCTION

As the primary component of terrestrial ecosystems, forests are crucial to reaching carbon neutrality objectives (Keenan et al., 2015). However, frequent natural and human disturbances like climate change, natural disasters, urbanization, and activities related to forest management have led to extensive forest fragmentation and edge effects (Fischer et al., 2021; Bonan, 2008). These effects have been observed to contribute to forest degradation and produce substantial carbon emissions (Brinck et al., 2017), which are much larger than the direct emission caused by deforestation (Qin et al., 2021; Matricardi et al., 2020). Therefore, spatiotemporally explicit information regarding forest disturbance has become one of the key elements in evaluating the carbon dynamics of forest ecosystems (Estoque et al., 2018).

The exponential growth of remote sensing data and quickly evolving cloud computing technology provide unprecedented opportunities to obtain forest disturbance information with appropriate spatial detail and temporal frequency over long periods (Chen et al., 2021; Hansen et al., 2013). With suitable spectral, temporal, and spatial resolutions and the advantages of being freely available, Landsat satellite series imagery has emerged as one of the fundamental remote sensing data sources for extensive and long-term forest dynamic monitoring (Chen et al., 2021; Hansen et al., 2013; Senf et al., 2017).

Researchers have developed several robust change detection algorithms to promote change detection performance. Some well-known algorithms, such as Breaks for Additive Season and Trend (BFAST), Landsat-based detection of Trends in Disturbance and Recovery (Landtrendr), Continuous Change Detection and Classification (CCDC) were proposed to detect land cover change using the time series analysis. Although these methods have brought further progress, the gradual and subtle forest changes remain not to be adequately solved, in particular, in areas with high

landscape heterogeneity and forest degradation. On the other hand, due to many practical difficulties, using Landsat images is still challenging to obtain quantitative spatiotemporal information on forest disturbance at different scales. (1) Although using moderate spatial resolution Landsat imagery to monitor large-scale forest resources is advantageous, it also makes Landsat difficult to identify forest variation within a forested stand such as small-scale and low-intensity degradation events (Bullock et al., 2020a; Wulder et al., 2008; Zhang et al., 2021); (2) As the existing vegetation index (VI) is not sensitive to minor forest changes, the satellite may delay or even fail to capture these changes (Wang et al., 2021); (3) Extensive forest fragmentation and edge effects exacerbate the mixed pixel issue (Margono et al., 2014; Matricardi et al., 2020), further increasing the difficulty in characterizing the actual forest dynamics in the mixed pixel using the existing VIs. The above factors together limit the acquisition of long-term forest disturbance information.

In order to solve the above-mentioned issues, the objective of this study is to develop a novel index for improving the monitoring of forest dynamics. A comprehensive cross-comparison of the novel index with conventional VIs was conducted. The novel index has been applied to generate a forest disturbance map with the CCDC algorithm.

## 2. STUDY AREA AND DATASET

### 2.1 Study area

The study region is in the Brazilian state of Rondônia (Figure 1). The entire research region is covered by the biome of the tropical Amazon forest. Large-area deforestation in Rondônia started in the 1960s due to the development of local traffic, and the expansion of farmland and cattle pastures (Brondizio, Moran, 2012). Forest conversion has also been influenced by small-scale clearings for logging, agriculture, or raising livestock. Frequent natural disasters like fire, windfall, and floods can also damage

\* Corresponding author

trees. It is reported that about 24% of the primary forest had been changed to other land uses (Bullock et al., 2020a). In Rondônia, massive deforestation and other forest management practices have resulted in severe forest fragmentation and edge effects.

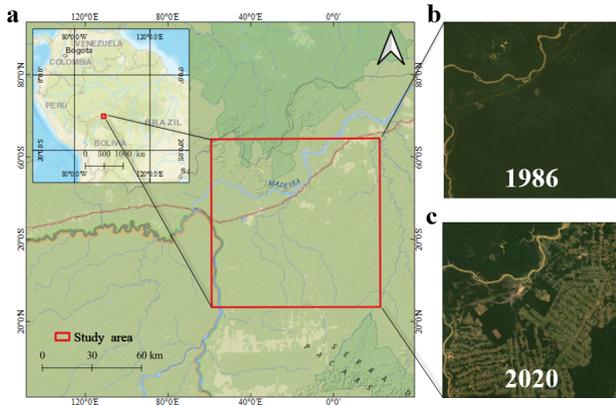


Figure 1. The study area. (a) is the location of the study area; (b) and (c) are 1986 and 2020 Landsat images in the study area

## 2.2 Landsat imagery and pre-processing

All available growing season Landsat surface reflectance over the study area from 1986 to 2020 were used to generate a forest disturbance map, including Landsat 4/5 TM, Landsat 7 ETM+, and Landsat 8 OLI. The quality assessment band was used to remove the cloud and shadow pixels. Due to the subtle but potentially significant differences between the spectral characteristics of images acquired by different Landsat sensors, spectral harmonization of TM and ETM+ to OLI spectrum was performed using a linear transformation method (Roy et al., 2016). Finally, six bands (blue, green, red, NIR, SWIR 1, and SWIR 2) were selected for subsequent analysis. All pre-processing was performed on the Google Earth Engine (GEE) platform (Gorelick et al., 2017).

## 3. METHODS

### 3.1 Pure Forest Index

Aiming to improve the ability of VI in characterizing forest dynamics over heterogeneous landscapes from moderate resolution satellite imagery, this study developed a novel Pure Forest Index (PFI) by integrating the spectral mixture analysis (SMA) and information of VI. First, the SMA was used to estimate the contribution of different endmembers (i.e., green vegetation (GV), non-photosynthetic vegetation (NPV), soil, cloud, and shade) to the spectral signal of a pixel. For each endmember, 50 pure pixels were collected by visual interpretation using the Landsat images in 2020 and the corresponding Google images at 0.5 m. The final endmembers were calculated as the average of the collected pixels. Here, we used the linear SMA approach was employed to obtain fractional abundances (Adams et al., 1986). The Root Mean Square Error (RMSE) was used to quantify the errors of SMA. Figure 2 shows an example of the SMA over a heterogeneous landscape using a linear SMA approach and the 2020 Landsat composite. A low RMSE value of 0.11 was achieved.

According to the principle of linear spectral mixture analysis, the spectral signal of a pixel is equal to the sum of the product of fractional abundance and reflectance of endmembers (Adams et al., 1986). Therefore, the PFI is designed as the transformation

of conventional VI to detangle the contribution of forest and non-forest to VI (Eq.1). The standard VI value of a non-GV endmember can be estimated from the spectral of the endmember. In order to remove the extreme value of PFI and reduce the commission errors of disturbance detection caused by the extreme index value, the Softsign function (Glorot, Bengio, 2010) was used to rescale the range of PFI into [-1, 1].

$$PFI = \text{Softsign} \left( \frac{VI - \sum_{m=1}^M VI_m A_m}{AGV} \right); \quad (1)$$

$$\text{Softsign} = \frac{X}{1 + |X|}$$

where  $VI$  = value of VI calculated from satellite imagery  
 $M$  = the number of endmembers  
 $VI_m$  = standard VI value of  $m$ -th non-GV endmember  
 $A_m$  = fractional abundance of GV endmember  
 $X$  = input of Softsign function

The principal illustration of the PFI was displayed in Figure 3. Here, we assumed that there is a mixed pixel (NDVI: 0.6) composed of three endmembers (GV, NPV, and Soil) at  $t_0$ . After deforestation/degradation ( $t_1$ ), the NDVI value of the pixel was reduced to 0.45. We can observe the change of NDVI for 0.15, which is hard to identify as an abrupt change by the change detection algorithm. By transforming NDVI to PFI, the PFI variation caused by forest change could be enlarged, which can facilitate the detection of small-scale forest change in a pixel. For PFI, the spectral component of the non-GV endmember in a pixel was estimated and eliminated (Eq. 1). Therefore, the PFI can integrated vegetation index and sub-pixel spectral information to amplify the signal of disturbances.

### 3.2 Forest disturbance mapping

This study used the CCDC algorithm (Zhu, Woodcock, 2014) and the PFI time series to generate the forest disturbance map from 1986-2020. CCDC algorithm has been widely used to detect forest change based on time-series satellite images. The CCDC algorithm uses harmonic models to fit observation data with the least absolute shrinkage and selection operator (LASSO):

$$p(i, t) = a_{0,i} + s_i t + a_{1,i} \cdot \cos \left( \frac{2\pi}{T} t \right) + b_{1,i} \cdot \sin \left( \frac{2\pi}{T} t \right) \\ + a_{2,i} \cdot \cos \left( \frac{4\pi}{T} t \right) + b_{2,i} \cdot \sin \left( \frac{4\pi}{T} t \right) \\ + a_{3,i} \cdot \cos \left( \frac{6\pi}{T} t \right) + b_{3,i} \cdot \sin \left( \frac{6\pi}{T} t \right) \quad (2)$$

where  $p(i, t)$  = the fitted value of the  $i$ -th PFI band on date  $t$   
 $T$  = the number of days per year  
 $S_i$  = the long-term trend of the  $i$ -th band  
 $a, b$  = the coefficients for COS and SIN function

When the fitted data considerably deviates from the actual observation over five times, the CCDC would consider there was an abrupt change. According to the experiments, the NBR-based PFI is superior to other VI-based PFIs in tracking forest dynamics. Thus, the NBR-based PFI time series was employed to map

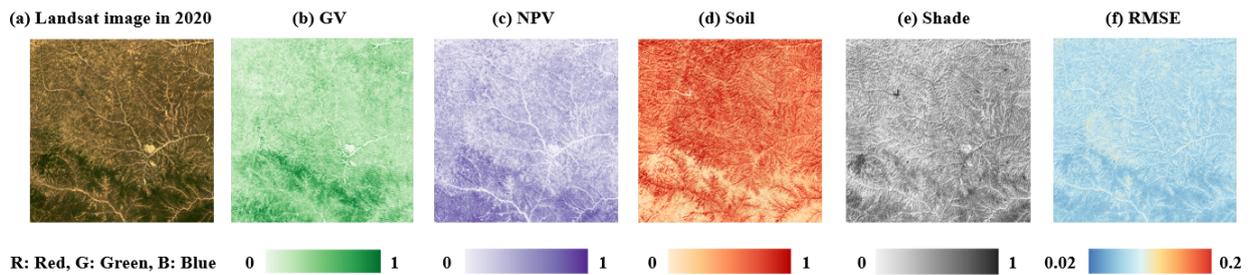


Figure 2. Example of the spectral mixture analysis. (a) is the Landsat composite in 2020, (b)-(e) are fractional abundances of GV, NPV, Soil, and Shade, respectively. (f) represents the *RMSE* of the linear SMA

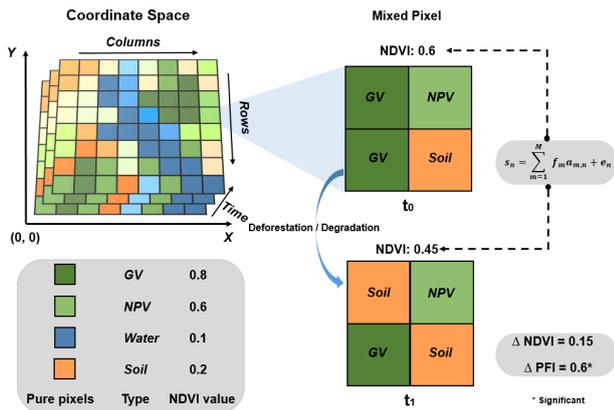


Figure 3. Illustration of the fundamental concept of PFI. The standard values of NDVI for pure pixels of different land types were shown in the bottom left panel

forest disturbance. The detection procedure was performed in the forest area. To obtain the 1986 forest mask, the Random Forest algorithm was applied to classify the land cover. A total of 327 samples, including forest, non-forest vegetation, built-up, and water, were selected for model training. The spectral bands and fractional abundances of endmembers were used as input features for the Random forest classifier to obtain the 1986 forest mask of the study area.

## 4. RESULTS

### 4.1 Comparison of PFI with other VIs in tracking forest dynamics

In this study, the NBR was used to calculate PFI. Other VIs, including Normalized Difference Fraction Index (NDFI) (Souza Jr et al., 2005), NBR, NDVI, EVI, RVI, and DVI, were used for comparison. The NDFI was developed to map canopy damage by combining spectral and spatial information. Several studies used NDFI to monitor deforestation and forest degradation with different detectors (e.g., CCDC and BFAST), which prove that NDFI is suitable to provide forest change information and improve forest monitoring (Bullock et al., 2020a; Bullock et al., 2020b; Chen et al., 2021; Souza Jr et al., 2005).

Figure 4 has displayed the VI images over forest regions. The comparison indicates that PFI can more accurately reflect the forest distribution than other VIs, which enlarges the distinctions between forest and non-forest pixels (Figure 4). Additionally, PFI can identify the proportion of forest within a pixel by integrating endmember abundance information. In general, PFI enlarged the

forest's spectral signal and suppresses background noise, which can improve the PFI's performance in capturing forest dynamics.

To intuitively understand the sensitivity difference between PFI and other VIs, we selected two distinct pixels for time series analysis. We compared the representational ability of different VIs time series in abrupt and gradual forest change (Figure 5). Compared with other vegetation indices, PFI could better describe the forest dynamic. In the region of abrupt change forest (Figure 5a), PFI is stable before forest disturbance while decreasing sharply at the time of abrupt forest change. It demonstrated that PFI is good at representing forest dynamics. Other vegetation indices are not as sensitive to forest changes as PFI, despite their ability to characterize the dynamic of forests to a certain extent (Figure 5a). In the region of the gradual change forest (Figure 5b), the performance of PFI is far better than other vegetation indices. PFI is more stable than other indices when the forest is not disturbed. In contrast, the amplitude of change of PFI is much more extensive than other indices, thus assisting CCDC in identifying forest disturbances. In the above two areas, only the forest changes were detected using CCDC-PFI, indicating that PFI is outstanding and flexible in suggesting forest changes.

### 4.2 Forest disturbance map derived from different VIs

We also generated the forest disturbance maps using the CCDC algorithm based on different VI time series (Figure 6). Note that the PFI used to generate a forest disturbance map is derived from the NBR.

The comparison of detection results derived from PFI and NDFI indicates that the PFI enables more accurate and explicit information about forest change. Due to the sensitivity of NDFI to the spectral change, the NDFI tends to produce commission errors. The NDFI is more sensitive to spectral change, leading to commission errors.

The PFI has dramatically improved the detection accuracy compared with non-SMA indices (NBR, NDVI, EVI, RVI, and DVI). The detection results were improved remarkably by PFI transformation compared with the disturbance map generated by NBR. Without the assistance of SMA, the disturbance detection using conventional time-series VIs would have failed to detect subtle spectral change. Therefore, the imperceptible forest change in a heterogeneous landscape tends to be omitted by conventional VIs since the area of forest change is too small to change the spectrum of a forested pixel. In this regard, PFI could improve the forest change detection compared with non-SMA indices by enhancing the spectral change caused by subtle forest change and reducing the commission errors induced by the misleading non-forest and noises.

In this study, we used a confusion matrix to validate the performance of the proposed method. The overall accuracy was utilized

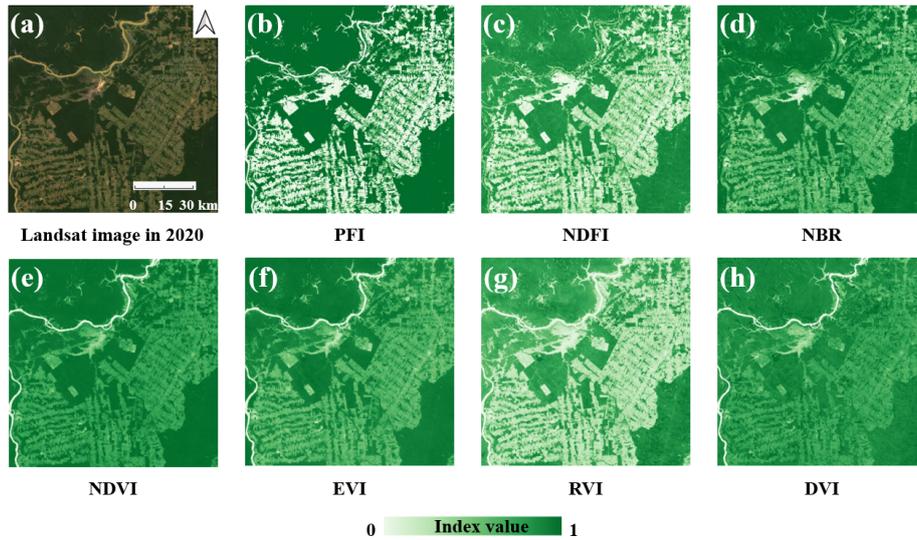


Figure 4. Landsat image and corresponding VIs. (a) is the Landsat image in 2020, (c)-(h) are corresponding images of PFI, NDFI, NBR, NDVI, EVI, RVI, and DVI, respectively

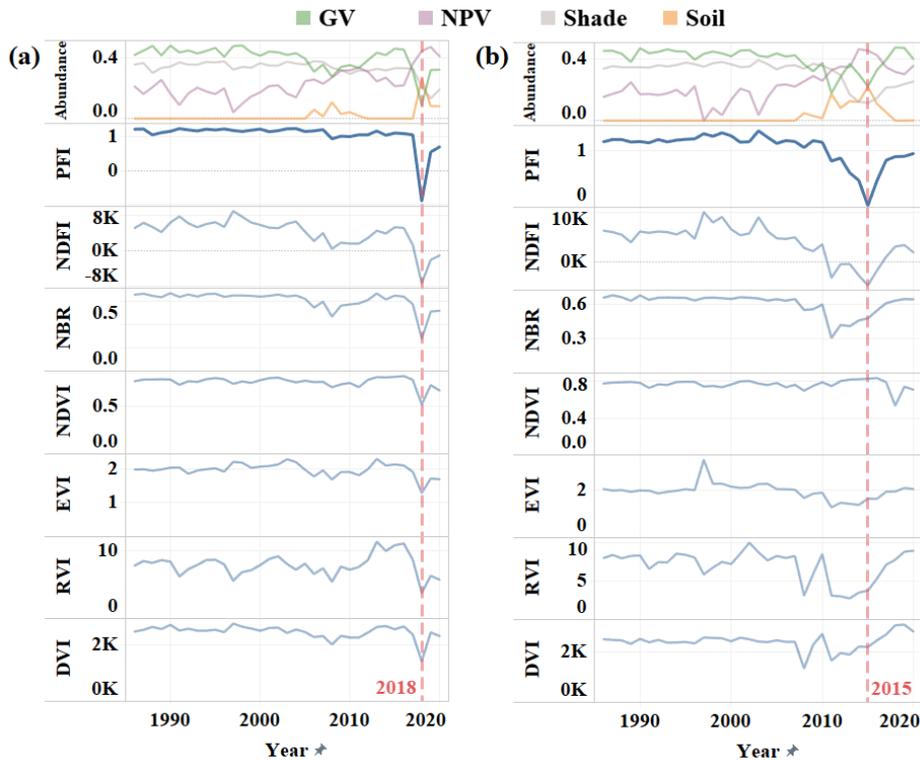


Figure 5. Analysis of fractional abundance and VIs time series of (a) a deforestation pixel and (b) a degradation pixel. The year of forest disturbance was displayed as a red dash line

to assess the detection accuracy. The accuracies of forest disturbance maps using the CCDC algorithm and different VI time series were quantified (Table 1). As shown in the Table, PFI-based mapping achieved the highest overall accuracy of 0.96, followed by NBR (0.76), DVI (0.72), NDFI (0.56), EVI (0.56), and RVI (0.32). The NBR-based PFI has greatly improved the mapping accuracy for NBR (0.96 vs. 0.76), which also demonstrates that spectral information is useful for capturing forest change.

## 5. DISCUSSION

Excellent results were achieved in forest change detection using the proposed PFI index and the CCDC-PFI method. However, changes that occurred in previous years may not be detected because starting data are required to initialize the CCDC algorithm. Therefore, the integration of PFI with other time series analysis algorithms such as BFAST and LandTrendr is worthy of in-depth and extensive research.

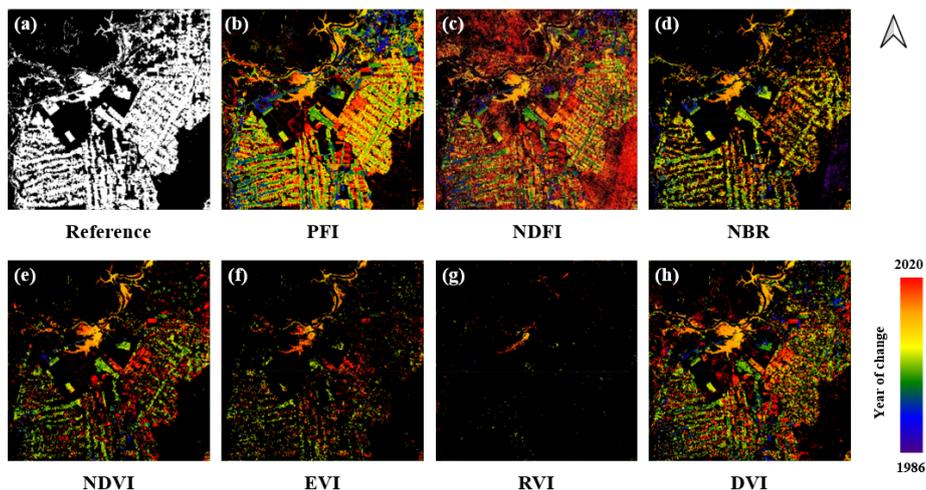


Figure 6. Forest disturbance maps generated by the CCDC algorithm based on different VI time series. White in the reference image indicates the location of forest change, while black represents non-change pixels

Vegetation index	Overall accuracy
PFI	0.96
NDFI	0.57
NBR	0.76
NDVI	0.71
EVI	0.56
RVI	0.37
DVI	0.72

Table 1. Accuracies of forest disturbance mapping using CCDC algorithm based on different VI time series

Furthermore, a lack of quality and consistent data can lead to missed detections and delays in observing forest changes in under-served areas. Long-term forest monitoring in the rainy and cloudy regions is particularly challenging due to the lack of fine-grained observations. According to recent studies, forest degradation is widespread throughout the Amazon rainforest and even on Earth, which can lead to significant carbon emissions from forest ecosystems (Bullock et al., 2020b; Qin et al., 2021; Junior et al., 2021; Matricardi et al., 2020). Therefore, the classification of degradation and deforestation is needed to elucidate and quantify their impact on the global biome. Despite small spectral differences from healthy forests, PFI is a good characterization of forest mortality from disease, insect infestation, and natural disasters. Future research should use the new index to investigate the complex spatial dynamics behind the spread of forest mortality.

## 6. CONCLUSIONS

In this study, PFI was developed to identify the specific contribution of forest change within pixels, which facilitates accurate forest change detection. The CCDC-PFI is proposed to monitor forest change accurately and has achieved excellent results in the Amazon rainforest. This study indicates that the PFI could promote detection accuracy by reducing omission and commission errors caused by data quality, spectral mixture, noise in images, and minor changes of the surface objects compared with other vegetation indices. Simultaneously, rich information regarding forest change can be extracted to support plan-making and analyse the potential drivers. Besides, the forest monitoring performance of different vegetation index-based PFIs has been compared. We found that NBR-based PFI is more remarkable for charac-

terizing forest dynamics and performs better than other vegetation indices-based PFI in forest change detection. Further study should focus on the classification of forest change types and the resulting ecological and environmental impacts.

## ACKNOWLEDGEMENTS

This research was funded by support from the National Natural Science Foundation of China under Grant 61976234.

## REFERENCES

- Adams, J. B., Smith, M. O., and Johnson, P. E. 1986. Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 Site, *Journal of Geophysical Research: Solid Earth*, 91(B8), 8098-8112.
- Bonan, G. B. 2008. Forests and climate change: forcings, feedbacks, and the climate benefits of forests, *Science*, 320(5882), 1444-1449.
- Brinck, K., Fischer, R., Groeneveld, J., Lehmann, S., De Paula, M. D., Pütz, S., Sexton, J. O., Song, D., and Huth, A. 2017. High resolution analysis of tropical forest fragmentation and its impact on the global carbon cycle, *Nature Communications*, 8(1), 1-6.
- Brondizio, E. S. and Moran, E. F. 2012. Level-dependent deforestation trajectories in the Brazilian Amazon from 1970 to 2001, *Population and Environment*, 34(1), 69-85.
- Bullock, E. L., Woodcock, C. E., and Olofsson, P. 2020a. Monitoring tropical forest degradation using spectral unmixing and Landsat time series analysis, *Remote sensing of Environment*, 238, 110968.
- Bullock, E. L., Woodcock, C. E., Souza Jr, C., and Olofsson, P. 2020b. Satellite-based estimates reveal widespread forest degradation in the Amazon, *Global Change Biology*, 26(5), 2956-2969.
- Chen, S., Woodcock, C. E., Bullock, E. L., Arévalo, P., Torchinava, P., Peng, S., and Olofsson, P. 2021. Monitoring temperate forest degradation on Google Earth Engine using Landsat time series analysis, *Remote Sensing of Environment*, 265, 112648.

- Estoque, R. C., Myint, S. W., Wang, C., Ishtiaque, A., Aung, T. T., Emerton, L., Ooba, M., Hijioka, Y., Mon, M. S., and Wang, Z. 2018. Assessing environmental impacts and change in Myanmar's mangrove ecosystem service value due to deforestation (2000–2014), *Global change biology*, 24(11), 5391-5410.
- Fischer, R., Taubert, F., Müller, M. S., Groeneveld, J., Lehmann, S., Wiegand, T., and Huth, A. 2021. Accelerated forest fragmentation leads to critical increase in tropical forest edge area, *Science advances*, 7(37), eabg7012.
- Glorot, X. and Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks, Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, *Proceedings of Machine Learning Research*, 249-256.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., and Moore, R. 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote sensing of Environment*, 202, 18-27.
- Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., and Loveland, T. R. 2013. High-resolution global maps of 21st-century forest cover change, *Science*, 342(6160), 850-853.
- Junior, C. H. S., Pessôa, A. C., Carvalho, N. S., Reis, J. B., Anderson, L. O., and Aragão, L. E. 2021. The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade, *Nature Ecology & Evolution*, 5(2), 144-145.
- Keenan, R. J., Reams, G. A., Achard, F., de Freitas, J. V., Grainger, A., and Lindquist, E. 2015. Dynamics of global forest area: Results from the FAO Global Forest Resources Assessment 2015, *Forest Ecology and Management*, 352, 9-20.
- Margono, B. A., Potapov, P. V., Turubanova, S., Stolle, F., and Hansen, M. C. 2014. Primary forest cover loss in Indonesia over 2000–2012, *Nature climate change*, 4(8), 730-735.
- Matticardi, E. A. T., Skole, D. L., Costa, O. B., Pedlowski, M. A., Samek, J. H., and Miguel, E. P. 2020. Long-term forest degradation surpasses deforestation in the Brazilian Amazon, *Science*, 369(6509), 1378-1382.
- Qin, Y., Xiao, X., Wigneron, J.-P., Ciais, P., Brandt, M., Fan, L., Li, X., Crowell, S., Wu, X., and Doughty, R. 2021. Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon, *Nature Climate Change*, 11(5), 442-448.
- Roy, D. P., Kovalskyy, V., Zhang, H. K., Vermote, E. F., Yan, L., Kumar, S. S., and Egorov, A. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity, *Remote Sensing of Environment*, 185, 57-70.
- Senf, C., Pflugmacher, D., Hostert, P., and Seidl, R. 2017. Using Landsat time series for characterizing forest disturbance dynamics in the coupled human and natural systems of Central Europe, *ISPRS Journal of Photogrammetry and Remote Sensing*, 130, 453-463.
- Souza Jr, C. M., Roberts, D. A., and Cochrane, M. A. 2005. Combining spectral and spatial information to map canopy damage from selective logging and forest fires, *Remote Sensing of Environment*, 98(2-3), 329-343.
- Wang, H., Liu, H., Huang, N., Bi, J., Ma, X., Ma, Z., Shang-guan, Z., Zhao, H., Feng, Q., and Liang, T. 2021. Satellite-derived NDVI underestimates the advancement of alpine vegetation growth over the past three decades, *Ecology*, 102(12), e03518.
- Wulder, M. A., White, J. C., Goward, S. N., Masek, J. G., Irons, J. R., Herold, M., Cohen, W. B., Loveland, T. R., and Woodcock, C. E. 2008. Landsat continuity: Issues and opportunities for land cover monitoring, *Remote Sensing of Environment*, 112(3), 955-969.
- Zhang, Y., Ling, F., Wang, X., Foody, G. M., Boyd, D. S., Li, X., Du, Y., and Atkinson, P. M. 2021. Tracking small-scale tropical forest disturbances: Fusing the Landsat and Sentinel-2 data record, *Remote Sensing of Environment*, 261, 112470.
- Zhu, Z. and Woodcock, C. E. 2014. Continuous change detection and classification of land cover using all available Landsat data, *Remote sensing of Environment*, 144, 152-171.