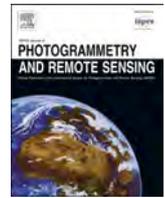


Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

Mapping fine-scale human disturbances in a working landscape with Landsat time series on Google Earth Engine

Tongxi Hu^a, Elizabeth Myers Toman^{a,*}, Gang Chen^b, Gang Shao^c, Yuyu Zhou^d, Yang Li^a, Kaiguang Zhao^{a,*}, Yanan Feng^a

^a Environmental Science Graduate Program, School of Environment and Natural Resources, The Ohio State University, Columbus, OH 43210, USA

^b Department of Geography and Earth Sciences, University of North Carolina at Charlotte, Charlotte, NC 28223, USA

^c Libraries and School of Information Studies, Purdue University, West Lafayette, IN 47907, USA

^d Department of Geological and Atmospheric Sciences, Iowa State University, Ames, IA 50011, USA

ARTICLE INFO

Keywords:

Google Earth Engine
Working landscape
Ensemble learning
Change detection
Hydraulic fracturing
BEAST
Land cover change
Sub-pixel

ABSTRACT

Large fractions of human-altered lands are working landscapes where people and nature interact to balance social, economic, and ecological needs. Achieving these sustainability goals requires tracking human footprints and landscape disturbance at fine scales over time—an effort facilitated by remote sensing but still under development. Here, we report a satellite time-series analysis approach to detecting fine-scale human disturbances in an Ohio watershed dominated by forests and pastures but with diverse small-scale industrial activities such as hydraulic fracturing (HF) and surface mining. We leveraged Google Earth Engine to stack decades of Landsat images and explored the effectiveness of a fuzzy change detection algorithm called the Bayesian Estimator of Abrupt change, Seasonality, and Trend (BEAST) to capture fine-scale disturbances. BEAST is an ensemble method, capable of estimating changepoints probabilistically and identifying sub-pixel disturbances. We found the algorithm can successfully capture the patterns and timings of small-scale disturbances, such as grazing, agriculture management, coal mining, HF, and right-of-ways for gas and power lines, many of which were not captured in the annual land cover maps from Cropland Data Layers—one of the most widely used classification-based land dynamics products in the US. For example, BEAST could detect the initial HF wellpad construction within 60 days of the registered drilling dates on 88.2% of the sites. The wellpad footprints were small, disturbing only 0.24% of the watershed in area, which was dwarfed by other activities (e.g., right-of-ways of utility transmission lines). Together, these known activities have disturbed 9.7% of the watershed from the year 2000 to 2017 with evergreen forests being the most affected land cover. This study provides empirical evidence on the effectiveness and reliability of BEAST for changepoint detection as well as its capability to detect disturbances from satellite images at sub-pixel levels and also documents the value of Google Earth Engine and satellite time-series imaging for monitoring human activities in complex working landscapes.

1. Introduction

Humans have been and will be altering Earth's natural landscapes through myriads of activities (Chen et al., 2015; Tilman et al., 2019). In the US, for example, 44% of the land has been converted from native vegetation to croplands and currently, the largest driver for land use change is energy development (Trainor et al., 2016). A large proportion of human-modified lands across the globe are working landscapes—an umbrella term coined to balance social, economic, and ecological needs within a landscape and highlight the importance of understanding and

observing the interactions between people and nature (Eastburn et al., 2017). Critical to the efforts of this kind for sustainability goals is the establishment of capabilities for reliably tracking and monitoring human footprints and disturbances (Miller and Zégre, 2016; Ma et al., 2018). Traditional field-based ways of surveying disturbances are expensive and spatially limited. In contrast, remote sensing provides alternative observations over large areas at lower costs and has increasingly become a popular method for detecting disturbances at various spatial and temporal scales.

Our capabilities of remotely observing landscape disturbances, being

* Corresponding author.

E-mail address: zhao.1423@osu.edu (K. Zhao).

<https://doi.org/10.1016/j.isprsjprs.2021.04.008>

Received 26 August 2019; Received in revised form 7 April 2021; Accepted 15 April 2021

Available online 17 May 2021

0924-2716/© 2021 International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Published by Elsevier B.V. All rights reserved.

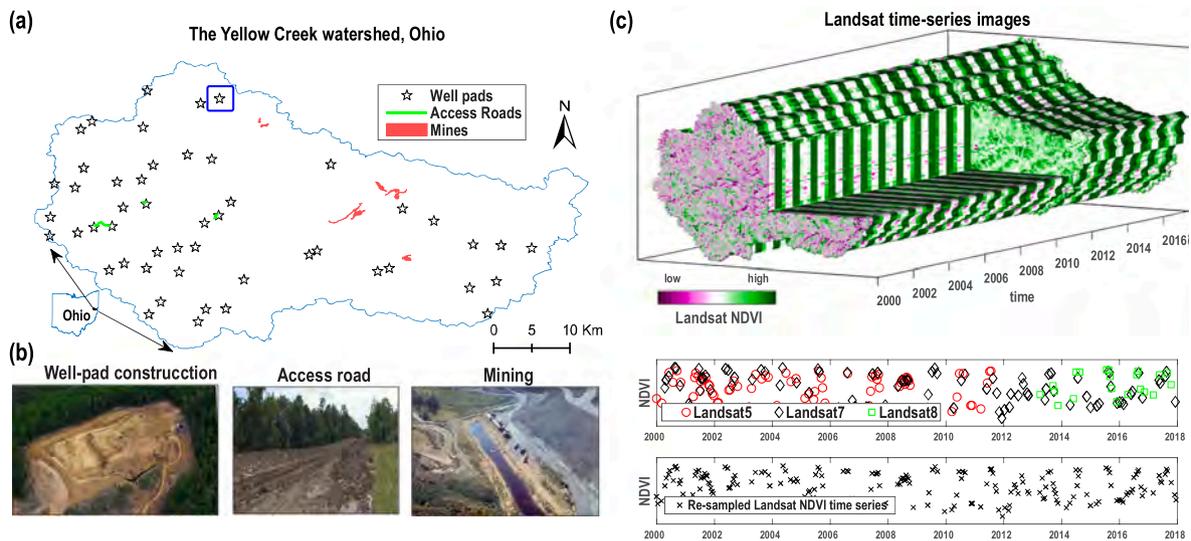


Fig. 1. Study area and Landsat time-series images: (a) The Yellow Creek watershed is a forest-dominated landscape with frequent disturbances from industrial and energy development activities. (b) Examples of the disturbances include hydraulic fracturing (e.g., construction of wellpads and access road) and coal and minerals mining. (c) High-density Landsat time-series NDVI images, with a specific example of the NDVI trajectory for a given sample. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

naturally occurring or anthropogenically driven, have been constantly improving, especially made possible by recent advances in data analytics and computational capacity (Hu et al., 2017, 2019; Wulder et al., 2016; Zhao et al., 2018). A milestone is the arrival of Google Earth Engine (GEE). GEE is a multi-petabyte catalog of satellite imagery collected during the past several decades, open to the general research communities (Gorelick et al., 2017). It enables fast and efficient processing of multiple years of satellite images via a user-friendly online cloud computing platform. GEE has become an increasingly popular platform for a wide spectrum of remote sensing applications (Ghazaryan et al., 2018). In particular, because of the opening of global Landsat archive data and the immediate availability of the Landsat data as high-density time-series images on GEE, great progress has been made in developing remote sensing approaches to monitor landscape dynamics and human-induced disturbances over time. Examples include the mapping of wildfire, forest logging, glacier melting, insect infestation, flooding, mining, and industrial activities (Watts and Laffan, 2014; Yu et al., 2018; He et al., 2021).

Concomitant with the advances and progresses in mapping disturbances are some practical challenges in characterizing human activities of diverse forms. This is especially true due to the vagaries of the ways and scales at which humans modify the landscapes. In the US, for example, the majority of human-dominated lands are working landscapes where the rural–urban continuum is frequently managed or altered to meet multiple competing needs. These activities occur often at small scales and overlap with each other over time. One example is the recent boom in unconventional shale gas extraction and hydraulic fracturing (HF). The vegetation disturbances caused by HF from the construction of wellpads and access roads are local in extent. An individual HF wellpad at its full operation is normally 1 to 2 ha (Ohio Department Natural Resources) and sometimes up to 4 ha for a super-sized pad. These areas are often too small to be detected by a moderate-resolution sensor (e.g., MODIS). Even with high-resolution sensors such as Landsat, these areas can be represented by only 10 to 45 pixels at 30-m resolutions; the issue of mixing pixels is much severer than characterizing large-scale disturbance such as logging and forest fire. The majority of existing remote sensing studies for mapping such activities took classification approaches based on paired pre- and post-disturbance images (Slonecker and Milheim, 2015). These traditional classification approaches are not optimal to tackle the small and time-varying nature

of the disturbance; for example, training sample sizes can be highly skewed and the disturbances occurred not before a fixed point of time but spread across many years continuously. The importance and challenges in mapping these small-scale activities have been well recognized (Pickell et al., 2014; Powers et al., 2015; Wasson and Franklin, 2018) but remain under-explored, with satellite time-series analysis being the most promising solution.

Numerous time-series algorithms have been developed in the passing decade to leverage the time dimension of satellite data—still being an active area of research. The advantages of time-series analysis are self-evident for mapping landscape dynamics and disturbances, but the availability of many alternative algorithms highlights a potential problem: no single algorithm is always applicable to all scenarios. The problem is recently stressed in a study evaluating seven common methods and finding that the agreement of the detected disturbances among the algorithms is close to nil at the pixel levels (Cohen et al., 2017). A belief in alleviating this dilemma is to discard the single-best-algorithm paradigm and switch to ensemble modeling. One such ensemble algorithm is the Bayesian Estimator of Abrupt change, Seasonality, and Trend (BEAST) originally reported in Zhao et al. (2019). The first evaluation of this algorithm is a case study monitoring wetland vegetation dynamics (Cai et al., 2020), and the practical performances of BEAST for detecting many other types of disturbances over different regions are yet to be done.

This study aims to design and evaluate a satellite time-series approach to detecting human-induced disturbances at fine scales, especially those associated with energy development. We first gathered decades of Landsat time series data via Google Earth Engine and then applied the BEAST algorithm to decompose the stacked high-density Landsat time series into individual components: seasonality, trend, and changepoints. Finally, we assessed the results and the effectiveness of BEAST with different ancillary data. More specifically, we seek to answer three questions related to landscape disturbance: (1) Is BEAST able to capture land cover disturbance at fine scales? (2) What is the magnitude of the detected disturbances or vegetation loss caused by human activities in a typical working landscape in the middle US? (3) How large are the disturbances from hydrofracking activities relative to other disturbances and how do they alter the landscape together?

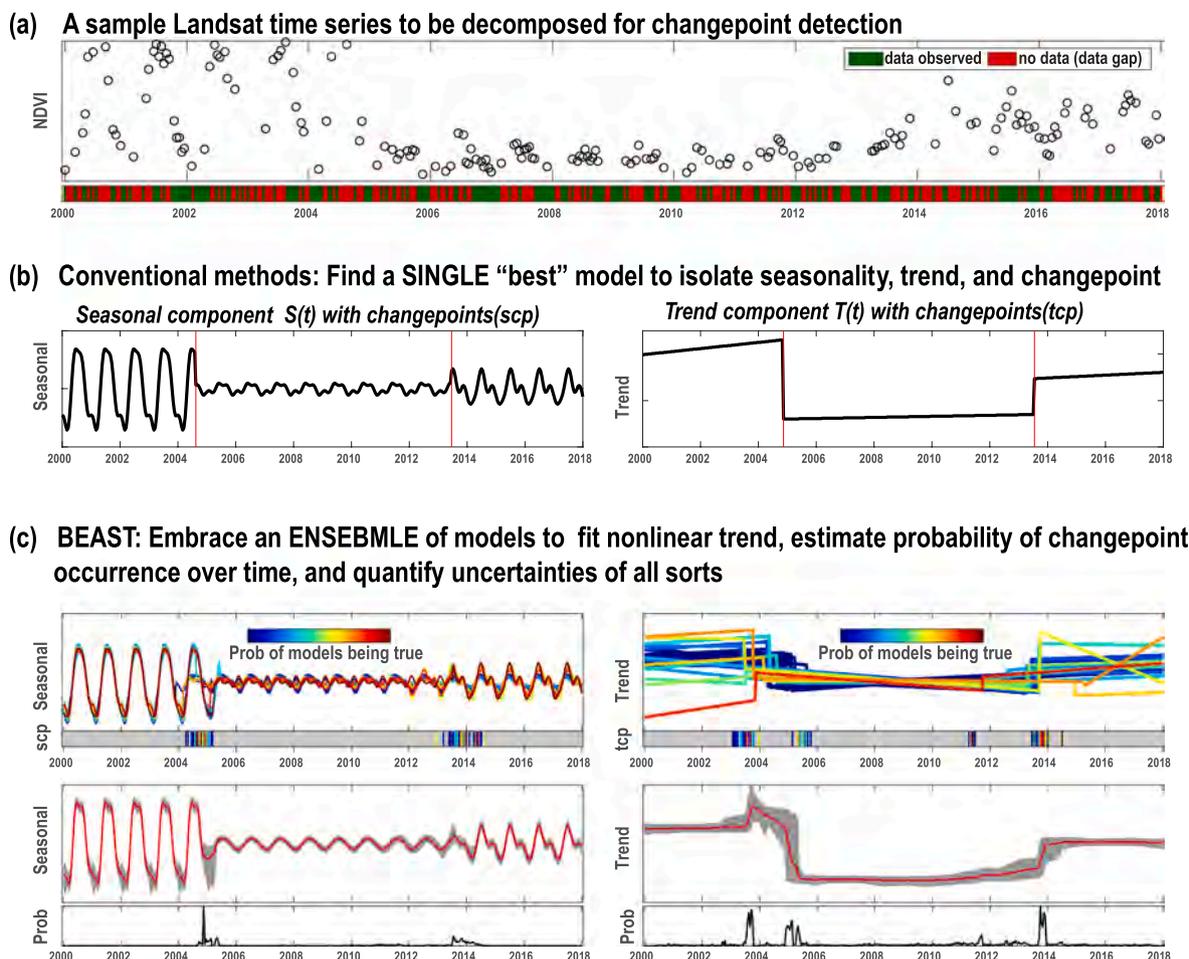


Fig. 2. Illustration of how BEAST works: BEAST is an ensemble algorithm seeking to decompose a time series into trend, seasonality, and changepoints. The interest of this study is mainly in the changepoint component. Unlike conventional methods that search only for a so-called best model, BEAST includes all the models into the inferences and synthesizes them into an average model, allowing explicitly quantification of model uncertainties.

2. Material and methods

2.1. Study area

Our study area is the Yellow Creek Watershed in Eastern Ohio, USA (Fig. 1a). Its size is 380 km² and its land cover is dominated by forest (72.5%), followed by grassland or pasture (17.3%), urban (5.4%), and cropland (2.3%). We selected this area because it is typical of the Appalachian region that has been disturbed by various human activities, such as coal and minerals mining, development of utility transmission lines, and unconventional shale oil and gas extraction (Fig. 1(a) and (b)). For example, from 2008 to 2017, 51 wellpads were constructed and 185 horizontal wells were drilled. Detailed information about the locations of the wellpads and land cover characterization for this watershed is supplied in the [Supplementary Materials](#).

2.2. Satellite time-series data

We processed 18 years' worth of Landsat images using Google Earth Engine from 2000 to 2017. In our study watershed, there are a total of 6575 scenes being pooled from GEE's Tier-1 atmospherically corrected surface reflectance data acquired by Landsat 5, 7, and 8. All the data preparation and processing were performed in a batch mode on GEE's cloud platform. In particular, we applied multiple filters to the time-series images using the quality control flag and the snow/cloud masks (e.g., ≥ 5 in quality control flag); the filtered images were then converted into NDVI. Because the time-series algorithm to be used (i.e., BEAST)

was formulated to handle regularly-spaced time series data only, we further resampled the NDVI time series at evenly-spaced points of time using a time interval of 15 days—a period commensurate with Landsat's revisit period (i.e., 16 days). The resampling resulted in 443 values per time series. Still, missing data were prevalent across the region and over time. The overall miss rate was 56.1%, but we kept the missing values without further special processing because the BEAST algorithm can handle missing data on its own.

2.3. BEAST: A Bayesian ensemble change-detection algorithm

BEAST is a Bayesian statistical algorithm to detect seasonality, trend, and abrupt changes in time series (Fig. 2). It differs from conventional changepoint detection algorithms mainly in two aspects. First, BEAST does not rely on any single model for decomposing the time series but rather combines numerous models into an average model using a technique termed Bayesian model averaging, that is, combining many weak models into a stronger model. Second, as a Bayesian algorithm, BEAST treats all the unknowns as random and therefore characterizes uncertainties of all sorts explicitly. Unlike the existing changepoint algorithms that report only whether there is a disturbance or not (i.e., a binary result), BEAST is a fuzzy detector that estimates the probability that a disturbance occurs for any given point in time (Fig. 2).

Mathematically speaking, BEAST breaks a time series $Y(t)$ into four components: trend (T), seasonal variability (S), abrupt changes (θ , and θ_s), and noise (ϵ). These four components are combined additively to model the time series:

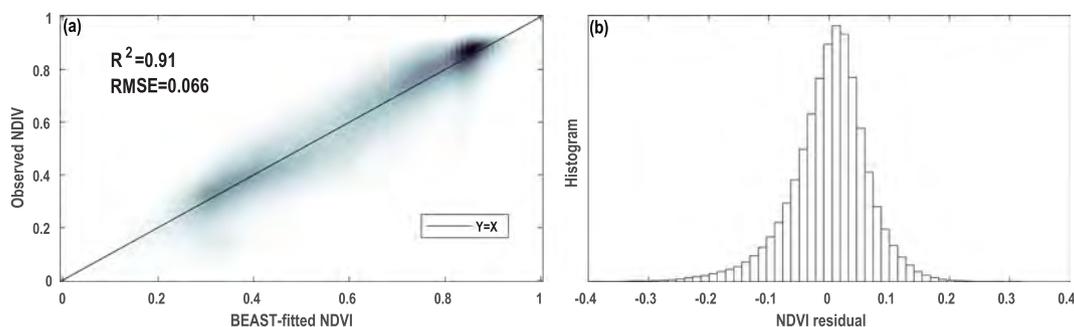


Fig. 3. Diagnostic statistics to evaluate BEAST as a regression tool: (a) the observed vs BEAST-fitted NDVI values; (b) the distribution of residuals for all pixels.

$$Y(t) = T(\theta_t) + S(\theta_s) + \varepsilon$$

where ε is the Gaussian random error term $N(0, \delta^2)$ with an unknown variance δ^2 ; T and S are the basic terms for the trend and seasonality components where abrupt changes are implicitly encoded in the parameters θ_t and θ_s . More specifically, θ_t and θ_s represent the numbers and locations of changepoints in the trend and seasonal components. Let us denote $\{\theta_t, \theta_s, \delta^2\}$ by M . The unknown parameters $M = \{\theta_t, \theta_s, \delta^2\}$ are formulated as a posterior probability distribution according to Bayes' theorem:

$$f(M|Y) \propto f(Y|M) * f(M)$$

The exact forms of the three terms are detailed in Zhao et al. (2019). The posterior probability $f(M|Y)$ encodes all the information on the time series decomposition, including the numbers and locations of changepoints in the trend and seasonal components. But $f(M|Y)$ is analytically intractable and it needs to be stimulated via the Markov Chain Monte Carlo (MCMC) sampling. In our analysis, we chose five parallel MCMC chains, each with 50,000 iterations with the first 5000 discarded as burn-in samples.

2.4. Algorithm evaluation and ancillary data

To test the effectiveness of BEAST, we primarily used two approaches for algorithm assessments: one focusing on the regression aspect of the algorithm, and another on the accuracies of changepoint detection. We first evaluated the performance of BEAST as a regression model in fitting the Landsat NDVI time series. Common diagnostic measures, such as R^2 and RMSE, were computed to evaluate the goodness-of-fit of the algorithm. These measures are applicable here because, in essence, BEAST is a general linear regression model. The premise for this assessment is that a time-series algorithm is less unlikely to capture the true underlying dynamics (e.g., seasonality, trend, and abrupt changes) if the algorithm

fits the time-series curve poorly. We also calculated a confusion matrix of disturbance/non-disturbance with 254 randomly selected pixels in the watershed as a complementary test of the algorithm (Supplementary Materials).

Second, we gathered ancillary data from multiple sources as independent references to contextualize and validate the disturbance results associated with BEAST-detected changepoints from the Landsat time series. The first ancillary dataset we considered is the annual land cover/land use maps from USDA's Cropland Data Layer (CDL). CDL is recognized as the best land use product with a resolution of ~ 30 m for the US; it is available annually for our study area starting from the year 2006. The CDL dataset was used to identify changes in land cover types before and after disturbance activities occurred. The second ancillary datasets we chose are GIS layers of human activities from multiple public sources. These include a layer of mine sites from the Ohio Department Natural Resources (ODNR) Division of Mineral Resources (<https://ohiodnr.gov>), a layer of wellpads from the ODNR Division of Oil and Gas, a layer of electric transmission lines from Homeland Infrastructure Foundation (<https://hifld-geoplatform.opendata.arcgis.com>), a layer of gas transmission pipelines from the U.S. Energy Information Administration (<https://www.eia.gov/state/maps.php>), and layers of hazardous liquid pipelines and hydrocarbon liquid gas pipelines from the National Pipeline Mapping System (<https://www.npms.phmsa.dot.gov>). All the layers contain information about the locations and extents of the industrial projects and sometimes the timelines of the projects (e.g., drilling and spud dates for the wellpads).

Other ancillary data sources we considered include a variety of high-resolution aerial photos and satellite imagery over time that helped us to visually and manually identify how and when the landscape was disturbed over selected sites. We particularly leveraged historical high-resolution images in Google Earth and Google Earth Engine. Additionally, our time-series analysis derives all types of abrupt changes in NDVI, regardless of their exact drivers. Many of the abrupt changes, especially

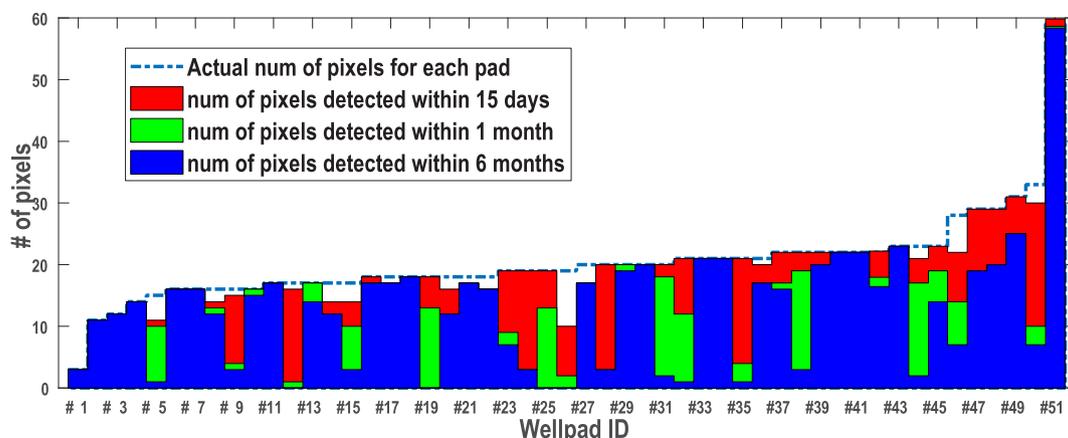


Fig. 4. Numbers of disturbed pixels detected by BEAST over individual wellpads within different time windows around the actual disturbance dates.

Probabilities of detecting seasonal and trend changepoints (scp and tcp) in each year

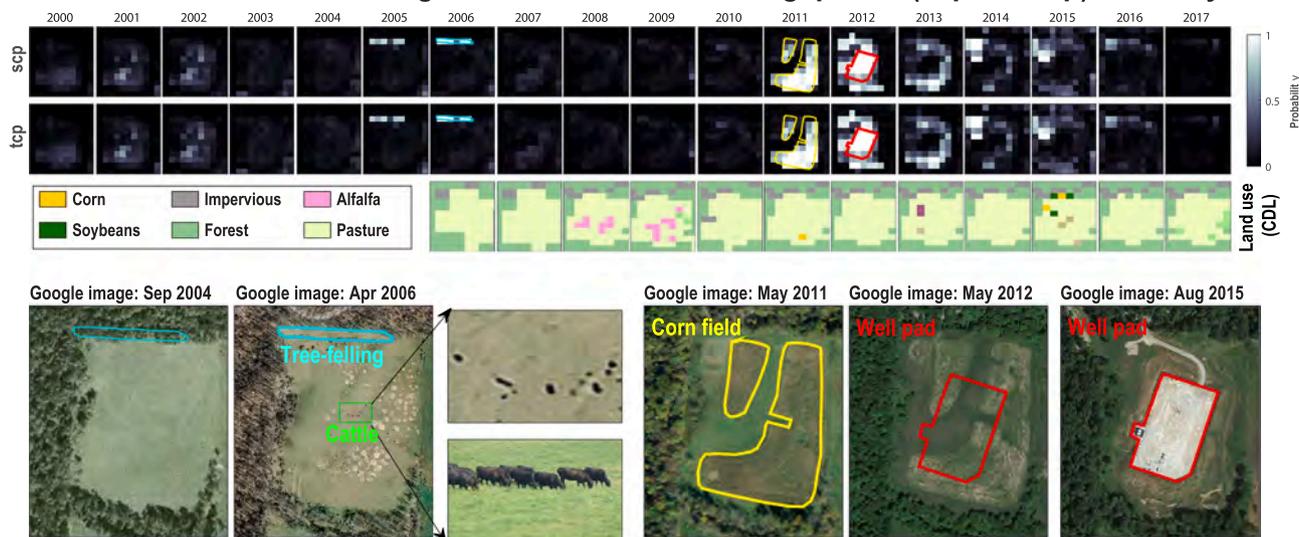


Fig. 5. Spatiotemporal dynamics of disturbances detected by BEAST over a selected wellpad site. The top two rows are the BEAST-estimated probabilities of seasonal and trend changepoints (i.e., scp and tcp) from the year 2000 to 2017; the third row is the concomitant annual Cropland Data Layer (CDL) land cover products, available only after 2006. Google Earth images in the bottom panel are used to illustrate land cover changes that have been captured by BEAST but not CDL.

those related to subtle variation in surface greenness, were attributed not to human activities but to natural factors (e.g., extreme weather). Therefore, we also resorted to rainfall and crop survey data to interpret our results (<https://www.nass.usda.gov>).

3. Results and discussion

3.1. Evaluation of BEAST

As a statistical model, BEAST fits the observed data well with an R^2 of 0.91 and RMSE of 0.066, according to a total of 100,000 randomly

sampled time series across the watershed. Furthermore, residuals errors in the fitted NDVI were found to closely followed a normal distribution (Fig. 3(b)), as tested by the Shapiro-Wilk normality statistic (p-value = 0.38). The test result shows that the residuals of the BEAST model did not violate the normality assumption and also confirms that our filtering and screening in the preparation of the Landsat images removed the majority of outliers (e.g., those associated with snow, clouds, and shadows), which otherwise would have skewed the residual distribution. In addition, a disturbance/non-disturbance confusion matrix indicated that the overall accuracy of BEAST to detect disturbed pixels was 77.2%.

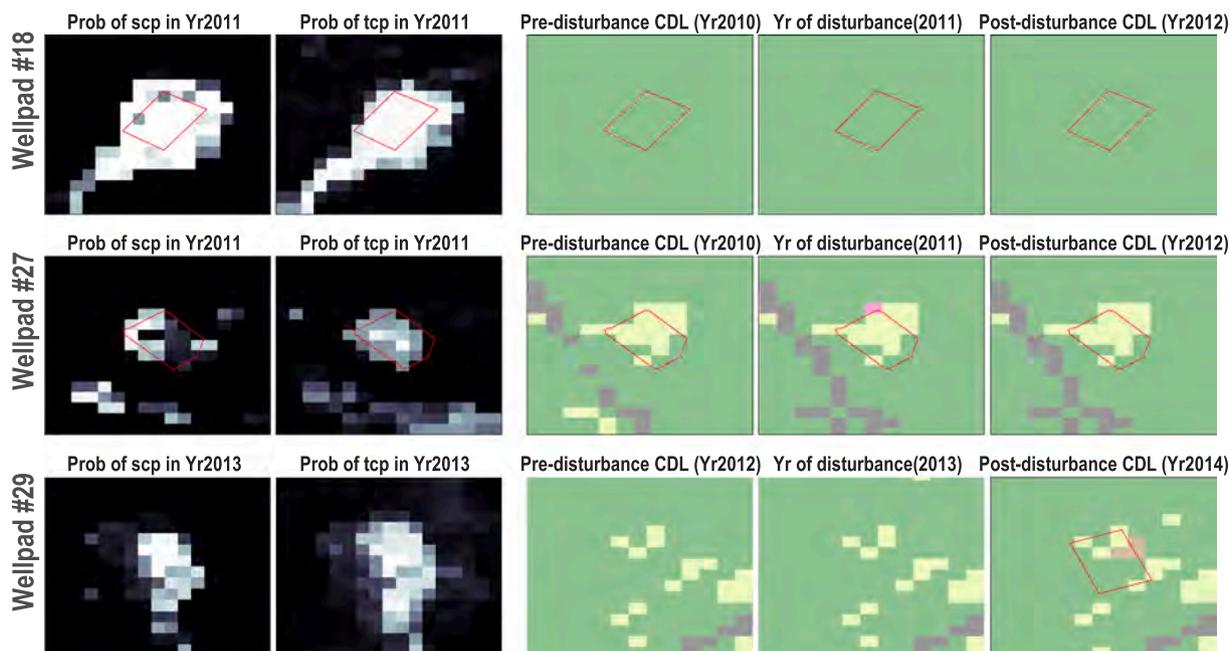


Fig. 6. Three more hydraulic fracturing wellpads as examples to further highlight the difference between land disturbances detected by BEAST (left) and the Cropland Data Layer (CDL) annual land cover dynamics products (right): Brighter pixels indicate higher occurrence probabilities of seasonal or trend changepoints (the first two columns on the left), as detected by BEAST. The CDL maps on the last three columns wrongly showed essentially no land cover changes before and after the wellpad onconstructions.

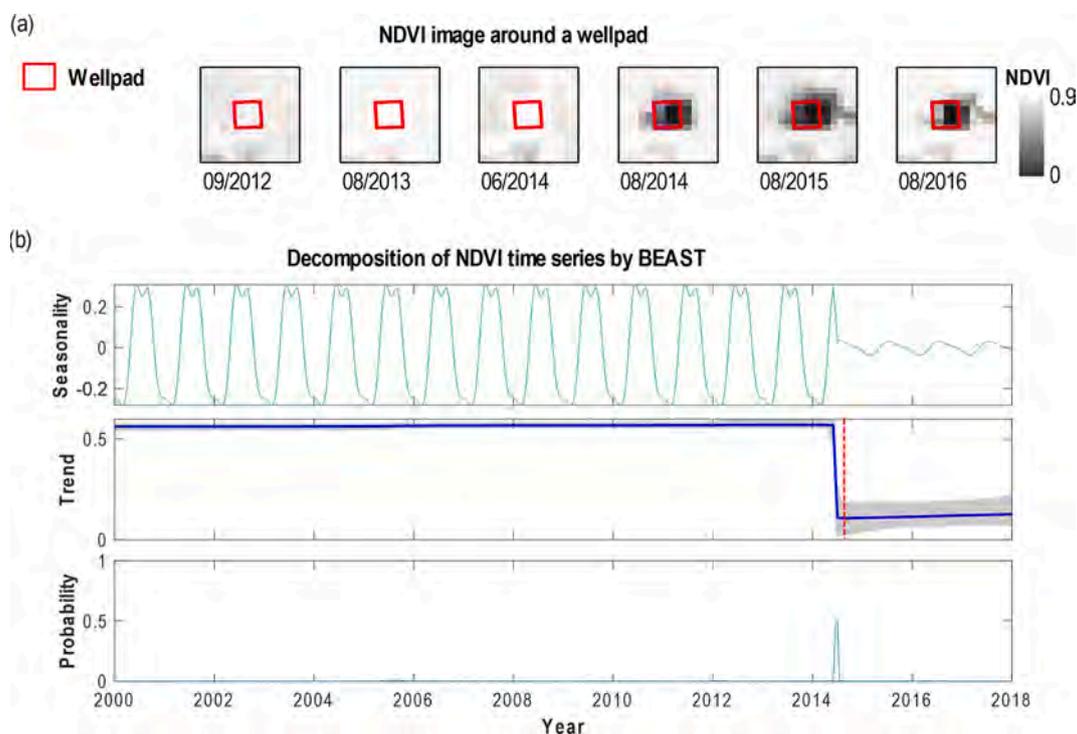


Fig. 7. Season and trend trajectories of NDVI on a selected wellpad to illustrate the BEAST time series decomposition. (a) NDVI time-series images around the wellpad. (b) Results of decomposed NDVI time series by BEAST include response curves of seasonality (top panel) and trends (middle panel), and abrupt changes with probabilities (bottom panel). The blue line in the middle panel represents an average trend of all candidate models, accompanied by a credible interval (lightly shaded area) at the 95% level. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

BEAST was also able to capture the disturbances caused by the constructions of most wellpads when comparing with ground reference data—ONDR’s fracking wellpad layer, which contains details about 51 wellpads such as geographical boundaries and operation and construction dates. In particular, BEAST captured abrupt changes within 60 days around the initial development of the wellpads on 45 of the 51 wellpad sites (i.e., a rate of 88%), over which pixels the vegetation was partially or fully removed. Such results are comparable to a conventional trajectory-based change-detection algorithm that was verified for forest disturbances at larger scales using both areas where trees were totally cut (90%) and partially cut (77%) (Kennedy et al., 2007).

On average, the wellpads for HF activities in the watershed cover 21 pixels, but only 13 of them (i.e., 60.5%) were identified to have changepoints within the 2-month window of the reported initial development dates. This result shouldn’t be interpreted as the ineffectiveness of BEAST. Rather, two major reasons explain the mismatching. First, the data missing rates of the NDVI time series were large; for more than 55% of the cases over the wellpad pixels, there were no Landsat observations available within 60 days around the construction dates—an inherent data problem independent of the algorithm. Indeed, the missing data problem became less of an issue when we increased the size of the time windows around the construction dates and found more pixels labeled as changepoints by BEAST (Fig. 4). Second, the construction dates reported by ODNDR are just a point in time, but the real construction and disturbances are a prolonged process that started from a point in space and spanned multiple years in time. The spatiotemporal dynamics of such disturbances were captured by BEAST in the estimated changepoint probabilities over time (Fig. 5).

BEAST captured land dynamics that were real but missed in conventional classification-based land-use products (Figs. 5 and 6). For example, only 4 of the 51 wellpads were found in the annual CDL maps to have land cover types altered before and after the construction years over large fractions of the pads; for all the other wellpads, the land use types roughly remained the same before and after the wellpad

constructions, although there might have some minor changes in land covers attributable mostly to random classification errors. Using the sample site in Fig. 5 as an example, the disturbances from the fracking activities starting from the year 2012 were evident in the BEAST-estimated probabilities for seasonal changepoints or trend changepoints (top, Fig. 5), but the wellpad footprint was not revealed at all in the annual CDL land cover maps and was classified mostly as pastures instead of open land/impervious surfaces. In 2015, part of this fracking site was even misclassified as corn and soybean. Three other examples are also given in Fig. 6 to illustrate similar observations.

Despite our emphasis being on industrial activities, we also found that BEAST can detect other types of changepoints, such as tree-felling, and agriculture management practices, whereas these changes were not captured by the traditional classification-based approaches (e.g., the CDL products). For example, before the year 2011, the site in Fig. 5 was a pasture field and in 2011 large portions of it were converted to two irregularly-shaped corn fields—a pattern detected by BEAST but not shown in the CDL annual land dynamics products. Sometime between 2005 and 2006, a narrow strip of trees north of the site was removed. The average width of this strip was 21 m, less than the 30-m resolution of Landsat. This sub-pixel disturbance was detected by BEAST; in particular, the pixels involved were estimated to have abrupt changes in seasonality and trend with an average probability of 84%.

3.2. Time series decomposition by BEAST

The BEAST algorithm decomposed the NDVI time series of a pixel into seasonality and trend signals, embedded with changepoints. We used a sample site in Fig. 1 to show more details about the dynamic nature of the BEAST results. The pixels that cover this wellpad experienced vegetation loss between June and August in 2014, according to the Landsat time-series imagery (Fig. 7(a)). The loss was captured by BEAST with an abrupt change in both the seasonality and trend response curves (Fig. 7(b)). The seasonality curve explicitly demonstrated

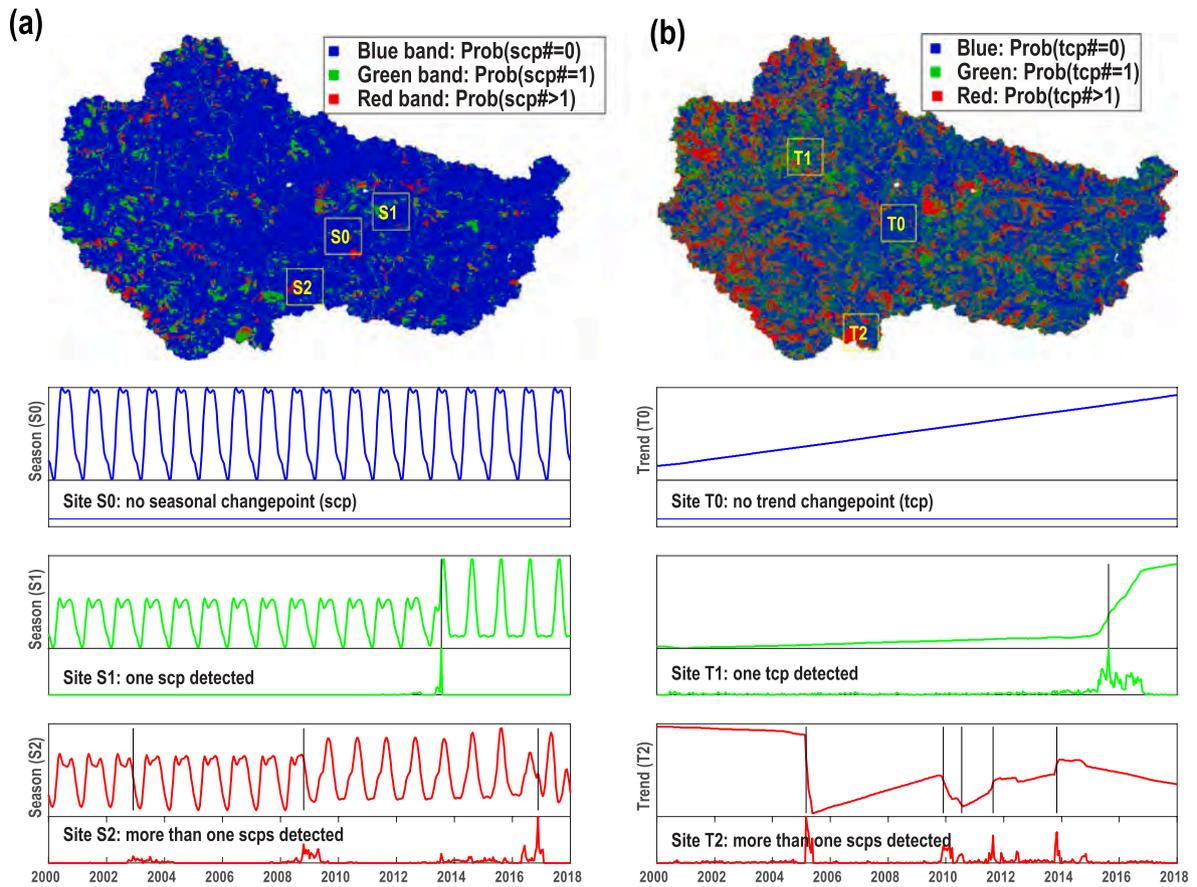


Fig. 8. Mapping the number of changepoints detected by BEAST in both (a) seasonal and (b) trend signals for each pixel: In the color composites used, bluish areas were less frequently disturbed (e.g., high probabilities of detecting no changepoints as exemplified at Site S0 and T0) whereas reddish areas were more frequently disturbed (e.g., Site S2 and T2).

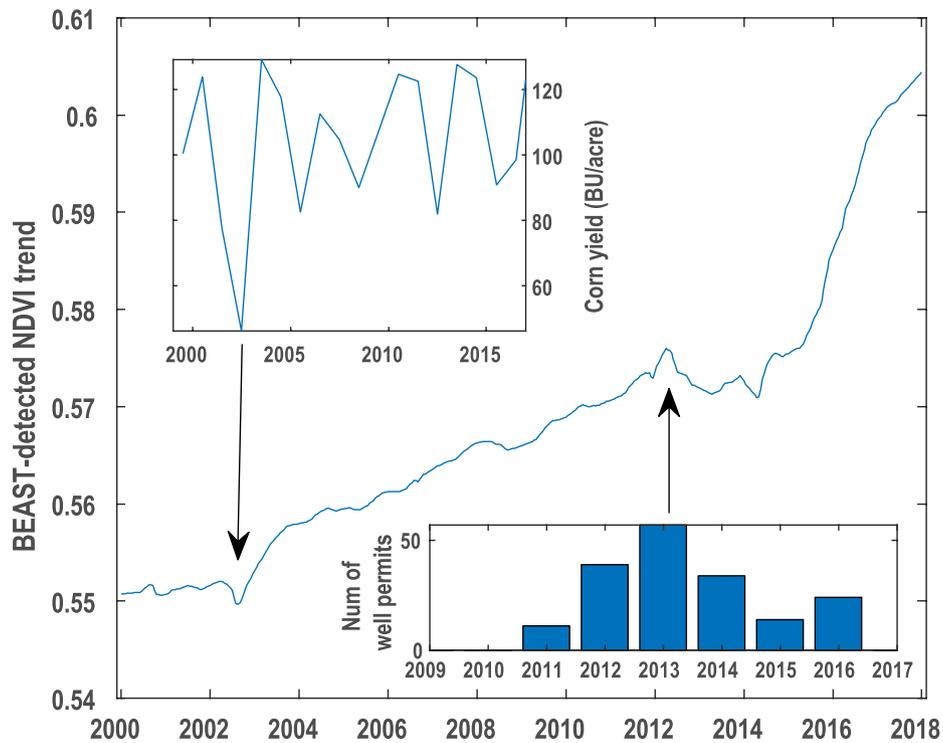


Fig. 9. The overall BEAST-detected trend in NDVI (i.e., surface greenness) averaged over the whole watershed. Around 2003 and 2013, a larger number of pixels were detected with abrupt changes (i.e., the drops in the trend curve marked by the two arrows), and these changes coincided with a drought-infllicted crop loss in 2003 and the peak in hydraulic fracturing activities in 2013, respectively.

declines in the magnitude of the NDVI time series after well pads were constructed. In the trend curve, a sudden drop was identified before the construction date (red line) of the wellpad. This is consistent with the practice of clearing vegetation before well pad development. The bottom panel in Fig. 7(b) depicts a unique feature of the BEAST algorithm—the probability curve of changepoint occurrence over time that gives the likelihood of a changepoint occurring for any given point in time.

Changepoints could appear in both seasonal and trend signals, and they do not necessarily occur at the same time. Seasonal changepoints normally indicate some phenological shifts whereas trend changepoints mean a shift from one regime to another regime in land surface dynamics. The numbers of seasonal changepoints (i.e., scp) are not necessarily the same as those of trend changepoints (i.e., tcp). We mapped the numbers of scp and tcp separately by dividing them into three categories: zero, one, and more than one changepoint (Fig. 8). The number of changepoints is an unknown parameter automatically estimated by BEAST. In our results, 24% of the watershed showed at least one seasonal changepoint and 41% showed at least one trend changepoint, with a probability of 0.5. These areas where scp and tcp co-occurred correspond to regions with large or dramatic disturbances (e.g., altered land use, stand-replacing forest clearing, and development of wellpads).

We noted that BEAST only estimated when and how many abrupt changes occurred in a time series but couldn't determine what caused the changes. The specific drivers have to be identified by referring to additional reference information. All the known human activities in our reference data were successfully detected by BEAST (more results reported in Section 3.3). For example, a large portion of changepoints was detected around 2013, as also suggested by the large drop in the BEAST-

estimated NDVI trend averaged over the watershed (Fig. 9); the timings of these changepoints coincide with the peak in the number of oil gas wells permits granted. Beside these known industrial activities, we suspect that the many other changepoints, especially tcp, are associated with short-time anomalies in surface greenness due to natural drivers such as extreme climate events. For example, the year 2003 saw the second largest number of changepoints detected, the majority of which were subtle variations with low magnitudes (e.g., the slight dip in the averaged NDVI trend in Fig. 9). The changepoints largely occurred over agricultural fields and the region experienced unusual drought in 2003; therefore, the occurrences of these changepoints were associated with drought-induced crops, which were further confirmed by the dip in the reported crop yield by NASS.

3.3. Spatiotemporal patterns of disturbances

We further mapped the maximum probability of disturbances in the probability curves (Fig. 10) and their associated dates (Fig. 11) for each pixel to show spatiotemporal patterns of disturbances in the watershed. In the map of maximum probabilities, spatial patterns of disturbances could provide insight into what activities caused them. For example, line-shaped patterns might indicate the development of right-of-ways for utility transmission lines whereas patch-shaped patterns could identify mining activities. Some disturbances with these two patterns (highlighted using red rectangles in Fig. 10) were confirmed to correspond to known industrial activities related to the development of power lines, gas pipelines, as well as surface coal and minerals mining. We also confirmed the patterns by visually comparing the BEAST results with Google Earth images; the two were consistent with each other, despite

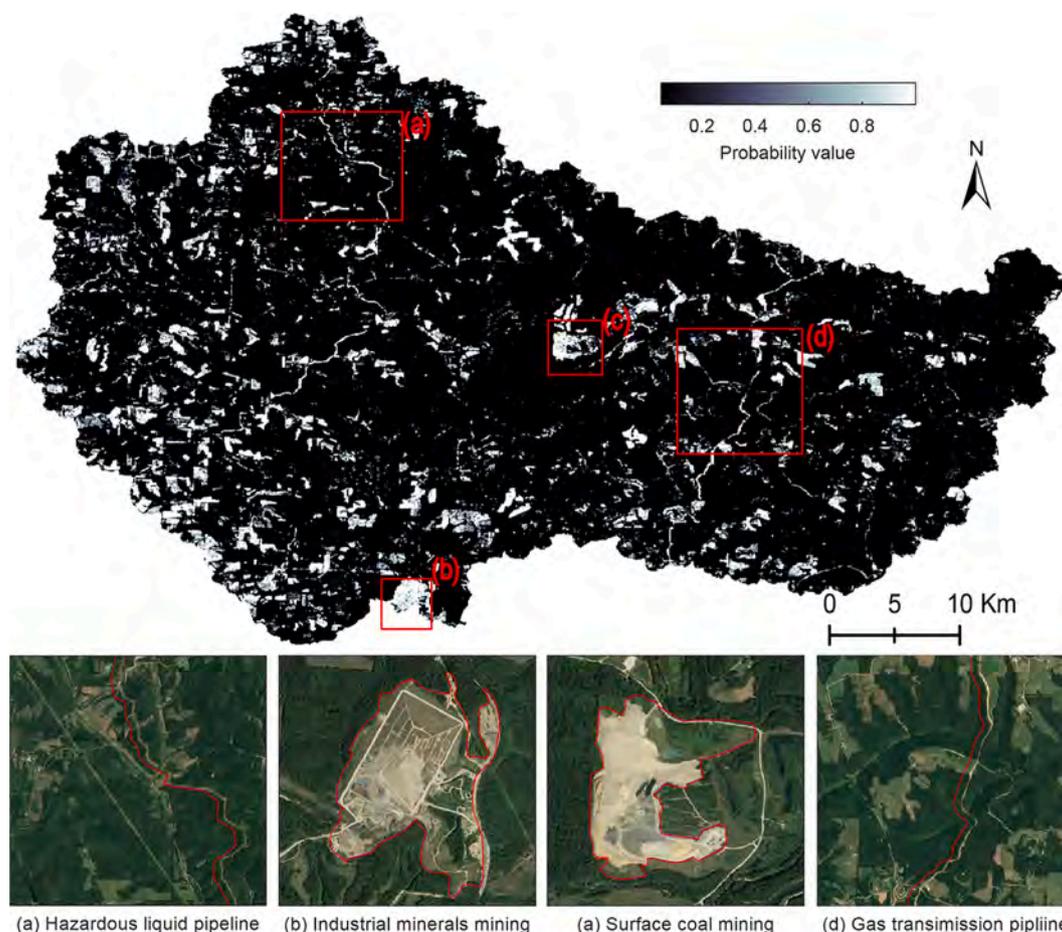


Fig. 10. Spatial patterns in the maximum probability of land disturbances from the year 2000 to 2017, together with examples to show the matching between the BEAST-detected patterns and high-resolution Google Earth images.

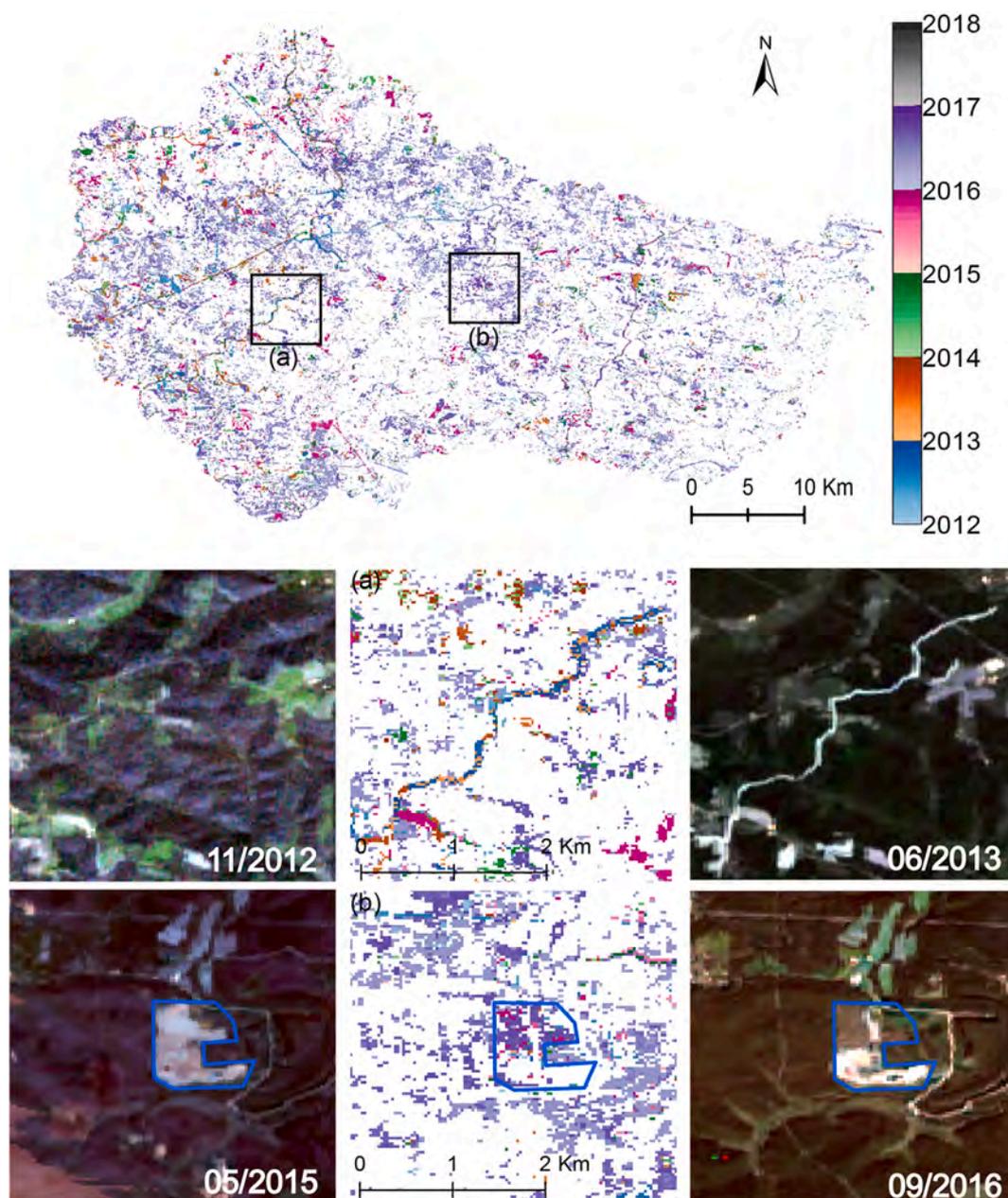


Fig. 11. Estimated dates of disturbances with the maximum probability that occurred between 2012 and 2017 in the watershed. Examples of (a) line- and (b) patch-shaped disturbances were highlighted using black squares and the corresponding areas were enlarged in the lower panels. The left and right insets in the lower panels are from Landsat 7 imagery (RGB = 3-2-1) for (a) and Landsat 8 imagery (RGB = 4-3-2) for (b).

the discrepancy in spatial resolutions.

The BEAST-detected dates and locations of disturbances closely resembled those revealed by Landsat images when evaluated at an annual level. We chose this temporal resolution due to a lack of information about the actual dates of all these disturbances. In particular, two areas with black squares were highlighted and enlarged to show how estimates of dates by BEAST were consistent with remote sensing observations (Fig. 11). The first area has a line-shaped disturbance by a hazardous liquid pipeline which was estimated to happen in 2012 and 2013 by BEAST. This result was confirmed by two Landsat images in November 2012 and June 2013. The second one is an abandoned surface mining site where disturbances were estimated by BEAST to happen primarily in 2016. These disturbances come from vegetation recovery based on the Landsat-8 images in May 2015 and September 2016.

3.4. Human-introduced disturbances in working landscapes

We additionally analyzed the relative areas of disturbances or altered land use associated with the known human activities. For mining activities, only active ones were considered because inactive sites were developed before 2000, earlier than our time series data. The classification of land cover that was most altered by mining activities was deciduous forest with a total of 316.4 ha (Table 1). While for line-shaped activities, the most affected land cover was evergreen forest (838.7 ha in total) that was impacted mainly by electric transmission lines and gas pipelines (406.4 and 311.7 ha). Overall, the top three land cover categories that had been altered in the watershed were forest, pasture/hay, and farmland, with an estimated area of 1192.5, 240.0, and 133.4 ha, respectively (Table 1).

Human activities have altered and fragmented the landscape of the

Table 1
Areas (ha) of land cover types altered by human activities.

Land Cover	Activities							
	HF	SCM	SIMM	ETL	GTP	HLP	HGLP	Total Area
Open Water	0	7.4	1.1	1.9	0.3	0	0.2	10.9
Developed Areas	0.3	9.8	0.3	36.5	15.8	11.5	3.5	77.7
Deciduous Forest	21	194.9	121.5	7.5	7.6	1.4	0	353.9
Evergreen Forest	0.3	3.6	6.8	406.4	311.7	72.2	37.7	838.7
Mixed Forest	0	3.2	7.5	0	0	0	0	10.7
Grassland	4.8	1.2	0	17.7	8.8	1	1.2	34.7
Pasture/Hay	45.6	25.1	10.9	76.1	62.4	8.6	11.3	240.0
Crops	19.8	25.5	25.6	25.8	27.4	5.1	4.2	133.4
Wetland	0	1.4	0.5	0	0	0	0	1.9
Total Area	91.8	272.1	174.2	571.9	434	99.8	58.1	1701.9

Note: HF – Hydraulic Fracturing; SCM – Surface Coal Mine (active); SIMM – Surface Industrial Minerals Mine (active); ETL – Electric Transmission Line; GTP – Gas Transmission Pipeline.
HLP – Hazardous Liquid Pipeline; HGLP – Hydrocarbon Gas Liquid Pipeline. Numbers in red color indicate the most impacted land cover area for each type of human activity.

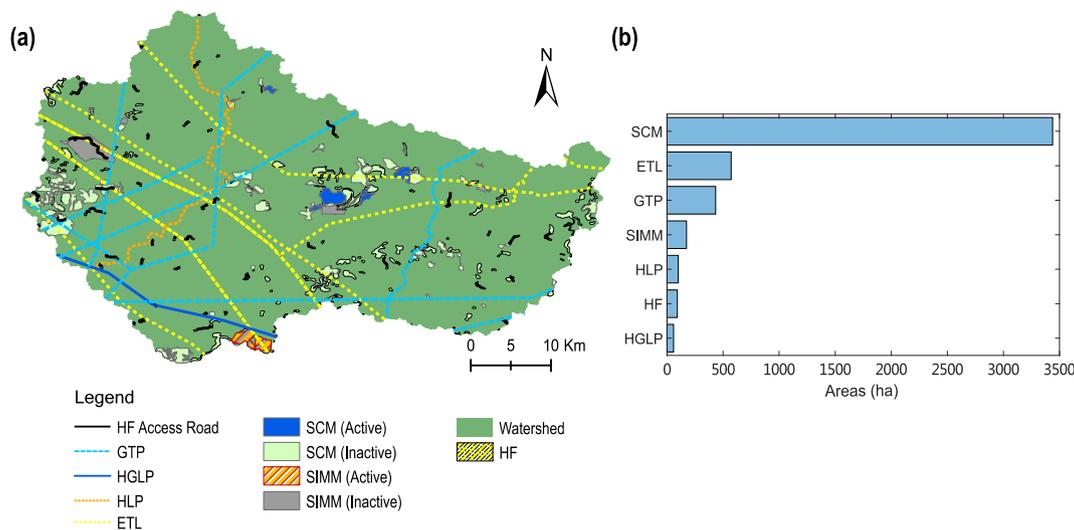


Fig. 12. Fragmentation of the Yellow Creek watershed and comparison of vegetated areas disturbed by different human activities. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Yellow Creek watershed frequently and sometimes, severely. The activities addressed in this research, for example, have disturbed 9.7% of the watershed over 18 years (Fig. 12(a)). The greatest loss of vegetated area was due to surface mining activities (i.e., SCM) with nearly 3479 ha when considering both inactive and active sites, followed by the development of utility pipelines (ETL and GTP) with about 1100 ha (Fig. 12(b)). Although the disturbed area seems relatively small, the disturbances are discontinuous with various patterns (e.g., lines, patches) and not concentrated in particular areas but spread over the entire watershed. The environmental and ecological consequences of these changes are far-reaching, raising many local management issues, especially related to impacts on water quality (Jackson et al., 2013) and habitat.

We further examined the lands disturbed by HF activities in detail. Thirty-three well pads were initially covered with grassland/pasture and

14 with deciduous forest. Four wellpads were located on agricultural land that grew corn, soybean, or alfalfa before they were developed. Shale gas extraction using HF is still at the center of considerable debate, mostly focused on environmental quality instead of landscape disturbance. Not surprisingly, our results confirmed that the fraction of landscapes disturbed by HF was relatively small, which was only 91.8 ha (about 2.7% of the land areas disturbed by surface mining) and the most impacted land cover was pasture/hay (45.6 ha).

HF-caused disturbances have not been analyzed in most studies that examine the environmental impacts of HF (Burton et al., 2014). One reason could be that disturbance caused by HF is possibly overshadowed by other environmental impacts of HF activities (e.g., drinking water contamination). But as shown in Fig. 12(a), HF activities (construction and development of well pads and access roads) have a higher density than mines, and the density is increasing. Indeed, access roads and well

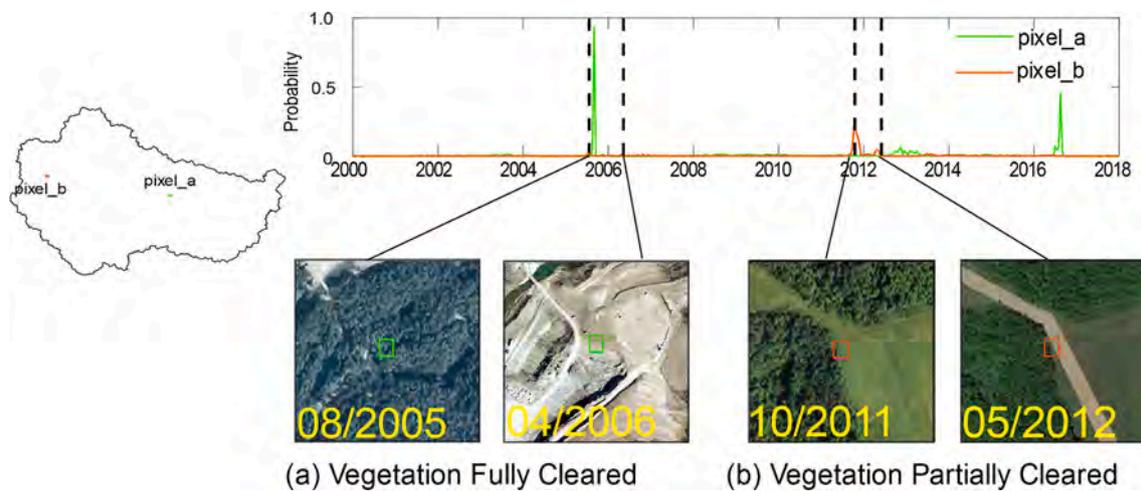


Fig. 13. Probabilities of abrupt changes determined by BEAST at two pixels with vegetation fully and partially cleared.

pads create a network of bare ground or non-vegetated land over the watershed that can break apart continuous forest habitat and create barriers for wildlife movement (Bischof et al., 2017; Proctor et al., 2012) or corridors for invasive plant species (Lemke et al., 2019). Their potential impacts on land fragmentation should not be overlooked (Kiviat, 2013).

3.5. Change detection at sub-pixel levels

The BEAST algorithm was shown to effectively capture fine-scale land disturbances. Not only because its performance is comparable with conventional algorithms (Section 3.1), but also because it could describe abrupt changes more flexibly than conventional algorithms. The biggest difference between BEAST and conventional change-detection algorithms is that conventional algorithms rely on a “best-fit” model to make inference while BEAST embraces an ensemble of candidate models and quantifies their relative usefulness. The use of ensemble modeling allows BEAST to deal with complicated cases such as disturbance at sub-pixel levels. Given an abrupt change that only covers half or one-third of a pixel (e.g., partial clear of vegetation), conventional change-detection algorithms give a response of either “Yes” or “No” (Wang et al., 2019; Zhao et al., 2019). While BEAST describes such sub-pixel changes with probability—a ratio between the number of candidate models that detect the change and the total number of candidate models (bottom, Fig. 7). The magnitude of the probability could tell how many candidate models respond to such a change and how many do not (Fig. 13).

The description of sub-pixel changes with probability curves makes BEAST suitable to detect fine-scale human disturbances since capturing the partially disturbed pixels is necessary especially when they are the majority type of pixels representing disturbances in remote sensing imagery. For example, the well pad highlighted in Fig. 1 is represented by 25 pixels in Landsat images and 64% of them are with vegetation partially removed. There are chances that some of these partially disturbed pixels are missed by conventional algorithms with a response of “No” but could be captured by BEAST with a low probability value.

4. Conclusions

We investigated disturbances to landscape resulted from various human activities (e.g., HF, gas pipelines, and mining activities) as identified by NDVI trajectories in a watershed in Ohio using an ensemble change detection algorithm—BEAST. The BEAST algorithm detected abrupt changes in land cover over 88% of the wellpads. A complementary test using the confusion matrix of disturbance to non-disturbance

demonstrated an overall accuracy of 77.2%. Such results indicated that the BEAST algorithm is a reliable tool for change detection. The algorithm could also capture disturbances at sub-pixel levels, which is highly needed for monitoring fine-scale disturbances.

Due to the small spatial extents, fine-scale human disturbances, for example, those introduced by HF activities, are difficult to capture and thus their impacts on the landscape were overlooked. BEAST provides the possibility to detect these fine-scale landscape disturbances and quantify their disturbed areas. Activities related to the development of HF were confirmed to cause disturbances at the local scale, but the total area of vegetated land cover lost to HF activities was minor (about 98 ha) compared with other activities. Such loss accounts for only 2.7% of the land disturbance caused by surface mining.

Detecting abrupt changes within time series is the key to understanding the acting processes and drivers of disturbances and their impacts on the environment. This paper will prompt further research regarding the ecological impacts of fine-scale land cover disturbances with a better understanding of the spatial and temporal patterns of disturbance at the finer resolution. For example, research addressing the environmental impacts of local land cover changes on the landscape in regards to habitat fragmentation, surface runoff and stream water quality, and invasive species movement would benefit from the advanced information of detected disturbance using BEAST or other change detection algorithms.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We gratefully acknowledge the support from the following programs: the USGS 104(b) program (Ohio Water Resources Center-2018OH567B); Dean’s Emerging Faculty Leaders award from the College of Liberal Arts and Science at the Iowa State University; the Open Research Fund from the State Key Laboratory of Digital Earth Science, Chinese Academy of Sciences, China; and Microsoft Azure for Research (CRM0518513). We also thank Lee Sorrell from the Ohio Department of Natural Resources for his help with data collection and interpretation. Our special thanks go to Dr. Oni Mutanga and four anonymous reviewers for their insightful comments. The software used and tested in this study is hosted via the Comprehensive R Archive Network at <https://CRAN.R-project.org/package=Rbeast>.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.isprsjprs.2021.04.008>.

References

- Bischof, R., Steyaert, S.M.J.G., Kindberg, J., 2017. Caught in the mesh: roads and their network-scale impediment to animal movement. *Ecography* 40 (12), 1369–1380. <https://doi.org/10.1111/ecog.02801>.
- Burton, G.A., Basu, N., Ellis, B.R., Kapo, K.E., Entekhabi, S., Nadelhoffer, K., 2014. Hydraulic “Fracking”: Are surface water impacts an ecological concern? *Environ. Toxicol. Chem.* 33 (8), 1679–1689. <https://doi.org/10.1002/etc.2619>.
- Cai, Y., Liu, S., Lin, H., 2020. Monitoring the vegetation dynamics in the dongting lake wetland from 2000 to 2019 using the BEAST algorithm based on dense landsat time series. *Appl. Sci.* 10 (12), 4209. <https://doi.org/10.3390/app10124209>.
- Chen, G., Powers, R.P., de Carvalho, L.M.T., Mora, B., 2015. Spatiotemporal patterns of tropical deforestation and forest degradation in response to the operation of the Tucuruí hydroelectric dam in the Amazon basin. *Appl. Geogr.* 63, 1–8. <https://doi.org/10.1016/j.apgeog.2015.06.001>.
- Cohen, W.B., Healey, S.P., Yang, Z., Stehman, S.V., Brewer, C.K., Brooks, E.B., et al., 2017. How similar are forest disturbance maps derived from different Landsat time series algorithms? *Forests* 8, 98, 8, 98. <https://doi.org/10.3390/f8040098>.
- Eastburn, D.J., O’Geen, A.T., Tate, K.W., Roche, L.M., 2017. Multiple ecosystem services in a working landscape. *PLoS ONE* 12 (3), e0166595. <https://doi.org/10.1371/journal.pone.0166595>.
- Ghazaryan, G., Dubovyk, O., Graw, V., Schellberg, J., 2018. Vegetation monitoring with satellite time series: An integrated approach for user-oriented knowledge extraction. In: Neale, C.M.U., Maltese, A. (Eds.), *Remote Sensing for Agriculture, Ecosystems, and Hydrology Xx*, vol. 10783. *Spie-Int Soc Optical Engineering*, Bellingham p. UNSP 107830W.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- He, Y., Chen, G., Cobb, R.C., Zhao, K., Meentemeyer, R.K., 2021. Forest landscape patterns shaped by interactions between wildfire and sudden oak death disease. *For. Ecol. Manage.* 486, 118987. <https://doi.org/10.1016/j.foreco.2021.118987>.
- Hu, T., Zhao, T., Shi, J., Wu, S., Liu, D., Qin, H., Zhao, K., 2017. High-Resolution Mapping of Freeze/Thaw Status in China via Fusion of MODIS and AMSR2 Data. *Remote Sensing* 9 (12), 1339. <https://doi.org/10.3390/rs9121339>.
- Hu, T., Zhao, T., Zhao, K., Shi, J., 2019. A continuous global record of near-surface soil freeze/thaw status from AMSR-E and AMSR2 data. *Int. J. Remote Sens.* 1–24. <https://doi.org/10.1080/01431161.2019.1597307>.
- Jackson, R.B., Vengosh, A., Darrah, T.H., Warner, N.R., Down, A., Poreda, R.J., et al., 2013. Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc. Natl. Acad. Sci.* 110 (28), 11250–11255. <https://doi.org/10.1073/pnas.1221635110>.
- Kennedy, R.E., Cohen, W.B., Schroeder, T.A., 2007. Trajectory-based change detection for automated characterization of forest disturbance dynamics. *Remote Sens. Environ.* 110 (3), 370–386. <https://doi.org/10.1016/j.rse.2007.03.010>.
- Kiviat, E., 2013. Risks to biodiversity from hydraulic fracturing for natural gas in the Marcellus and Utica shales. *Ann. N. Y. Acad. Sci.* 1286, 1–14. <https://doi.org/10.1111/nyas.12146>.
- Lemke, A., Kowarik, I., von der Lippe, M., 2019. How traffic facilitates population expansion of invasive species along roads: The case of common ragweed in Germany. *J. Appl. Ecol.* 56 (2), 413–422. <https://doi.org/10.1111/1365-2664.13287>.
- Ma, Q., He, C., Fang, X., 2018. A rapid method for quantifying landscape-scale vegetation disturbances by surface coal mining in arid and semiarid regions. *Landscape Ecol.* 33 (12), 2061–2070. <https://doi.org/10.1007/s10980-018-0726-9>.
- Miller, A.J., Zégre, N., 2016. Landscape-Scale Disturbance: Insights into the Complexity of Catchment Hydrology in the Mountaintop Removal Mining Region of the Eastern United States. *Land* 5 (3), 22. <https://doi.org/10.3390/land5030022>.
- Pickell, P.D., Hermsilla, T., Coops, N.C., Masek, J.G., Franks, S., Huang, C., 2014. Monitoring anthropogenic disturbance trends in an industrialized boreal forest with Landsat time series. *Remote Sensing Lett.* 5 (9), 783–792. <https://doi.org/10.1080/2150704X.2014.967881>.
- Powers, R.P., Hermsilla, T., Coops, N.C., Chen, G., 2015. Remote sensing and object-based techniques for mapping fine-scale industrial disturbances. *Int. J. Appl. Earth Obs. Geoinf.* 34, 51–57. <https://doi.org/10.1016/j.jag.2014.06.015>.
- Proctor, M.F., Paetkau, D., McLellan, B.N., Stenhouse, G.B., Kendall, K.C., Mace, R.D., et al., 2012. Population fragmentation and inter-ecosystem movements of grizzly bears in western Canada and the northern United States. *Wildlife Monographs* 180 (1), 1–46. <https://doi.org/10.1002/wmon.6>.
- Slonecker, E.T., Milheim, L.E., 2015. Landscape Disturbance from Unconventional and Conventional Oil and Gas Development in the Marcellus Shale Region of Pennsylvania, USA. *Environments* 2 (2), 200–220. <https://doi.org/10.3390/environments2020200>.
- Tilman, D., Hartline, N., Clark, M.A., 2019. Saving biodiversity in the era of human-dominated ecosystems. *Biodiversity Climate Change: Transform. Biosphere* 356–365.
- Trainor, Anne M., McDonald, Robert I., Fargione, Joseph, 2016. Energy Sprawl Is the Largest Driver of Land Use Change in United States. *PLoS ONE* 11 (9), e0162269. <https://doi.org/10.1371/journal.pone.0162269>.
- Wang, Y., Ziv, G., Adami, M., Mitchard, E., Batterman, S.A., Buermann, W., et al., 2019. Mapping tropical disturbed forests using multi-decadal 30 m optical satellite imagery. *Remote Sens. Environ.* 221, 474–488. <https://doi.org/10.1016/j.rse.2018.11.028>.
- Wasson, R.J., Franklin, S.E., 2018. Detection accuracy of new well sites using Landsat time series data: a case study in the Alberta Oil Sands Region. *Remote Sens. Lett.* 9 (2), 160–169. <https://doi.org/10.1080/2150704X.2017.1410293>.
- Watts, L.M., Laffan, S.W., 2014. Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region. *Remote Sens. Environ.* 154, 234–245. <https://doi.org/10.1016/j.rse.2014.08.023>.
- Wulder, M.A., White, J.C., Loveland, T.R., Woodcock, C.E., Belward, A.S., Cohen, W.B., et al., 2016. The global Landsat archive: Status, consolidation, and direction. *Remote Sens. Environ.* 185, 271–283. <https://doi.org/10.1016/j.rse.2015.11.032>.
- Yu, L., Xu, Y., Xue, Y., Li, X., Cheng, Y., Liu, X., et al., 2018. Monitoring surface mining belts using multiple remote sensing datasets: A global perspective. *Ore Geol. Rev.* 101, 675–687. <https://doi.org/10.1016/j.oregeorev.2018.08.019>.
- Zhao, K., Suarez, J.C., Garcia, M., Hu, T., Wang, C., Londo, A., 2018. Utility of multitemporal lidar for forest and carbon monitoring: Tree growth, biomass dynamics, and carbon flux. *Remote Sens. Environ.* 204, 883–897. <https://doi.org/10.1016/j.rse.2017.09.007>.
- Zhao, K., Wulder, M.A., Hu, T., Bright, R., Wu, Q., Qin, H., et al., 2019. Detecting change-point, trend, and seasonality in satellite time series data to track abrupt changes and nonlinear dynamics: A Bayesian ensemble algorithm. *Remote Sens. Environ.* 232, 111181. <https://doi.org/10.1016/j.rse.2019.04.034>.