Assessment of Performance of Tree-Based Algorithms to Reduce Errors of Omisssion and Commission in Change Detection

Peter Rasmussen
Discovery Lab
Accenture Federal Services
Arlington, VA, United States
peter.g.rasmussen@afs.com

Jenna Abrahamson
Center for Geospatial Analytics
North Carolina State University
Raleigh, United States
jnabraha@ncsu.edu

Xiaojing Tang
School of Integrated Sciences
James Madison University
Harrisonburg, United States
tang3xx@jmu.edu

Owen Smith
Center for Geospatial Analytics
North Carolina State University
Raleigh, United States
osmith2@ncsu.edu

Josh Gray
Center for Geospatial Analytics
Forestry and Environmental Resources
Raleigh, United States
josh_gray@ncsu.edu

Curtis Woodcock
Department of Earth and Environment
Boston University
Boston, United States
curtis@bu.edu

Marc Bosch
Accenture Federal Services
Arlington, VA, United States
marc.bosch.ruiz@afs.com

***Abstract*—The ability to detect land use and land cover change quickly and accurately is crucial for earth system modeling, policy making, and sustainable land management. Remote sensing has been widely used to map and monitor land use and land cover change over very large areas. Many change detection algorithms (CDAs) have been developed with promising accuracy. However, accuracy of detecting specific types of change using these algorithms is often not satisfactory owing to errors of commission. We present a novel pixel-based broad area search (BAS) approach that detects and classifies heavy construction, which is an important indicator of human development and of interest to the intelligence community. The BAS system combines an online CDA, roboBayes, with a supervised tree-based classifier that removes the CDA's errors of commission. To assess the performance of the classifier, we examined three tree-based algorithms – decision tree, random forest, and LightGBM – trained on roboBayes model parameters, tuning the models using a leave-one-region-out cross-validation strategy. We compared the performance of the tree-based classifiers against a baseline of filters created by the authors. Performance was evaluated at the pixel-level using precision, recall, and F1-score, which are analogues of commission error, omission error, and accuracy, respectively. The BAS system with optimized tree-based filters performed nearly 80% better than the BAS system without any filters and more than 50% better than the authors' filters.**

Keywords*—*change detection, broad area search, remote sensing, filtering, error of commission, random forest

I. Introduction

Over the past 60 years, human activities in the form of land use and cover change have affected nearly one third global land [1]. Land use and land cover change is a large source of anthropogenic carbon emissions [2] and affects the ecosystem services, biodiversity, surface energy balance, water cycles, and atmospheric composition [3]. The ability to detect change quickly and accurately is crucial for earth system modeling, policy-making, and sustainable land management [4]. Remote sensing has been widely used to map and monitor land use and land cover change over very large areas [5, 6, 7, 8, 9]. The availability of free remote sensing data, especially after the opening of the Landsat archive in 2008 [10], has sparked the use of dense time series analysis of remote sensing data for land change detection [11, 12]. Many novel time-series based change detection algorithms have been developed with promising accuracy [13, 14, 15]. However, the accuracy of detecting specific types of change – including heavy construction – using these algorithms is often not satisfactory [16]. We focus on the detection and monitoring of heavy construction activity because it has a significant impact on urban climate, is an important indicator of human development, and is of interest to the intelligence community [17, 18].

Most of the mapping efforts of construction activities using remote sensing have been focused on impervious surfaces [19] and urbanization [20]. These types of studies track the expansion of human development and often fail to capture redevelopment projects, where new construction occurs in already developed areas. There have also been studies that use very high-resolution imagery [21] and data collected by UAV [22] to monitor construction activities. Cost considerations limit these studies to localized, smaller areas and target a very specific type of construction. What is lacking here is an approach that can perform a broad area search (BAS) to detect and monitor new heavy construction activities, including redevelopment projects, over large spatial scales. This paper presents one such approach that combines an unsupervised change detection algorithm (CDA), Robust Online Bayesian Monitoring (roboBayes) [23], with a supervised classifier that removes the CDA's errors of commission.

II. METHODS

A. *BAS pipeline*

The methods developed in this study automate the BAS of dense time series of satellite imagery to detect and monitor anthropogenic change with a focus on heavy construction activities. Our approach achieves this in two steps: 1) roboBayes [23] is employed as a general change detector capable of flagging both natural and anthropogenic change in time series of imagery, and 2) resulting candidate change pixels are filtered using either a) an expert system or b) a machine learning (ML) model to obtain change pixels indicative of heavy construction. Fig. 1 depicts the BAS pipeline we used to generate inputs for the experiment. In this version, we left out the postprocessing task, which is executed after filtering and applies spatial filtering and converts pixel clusters to polygons, so that we could focus on the effect filtering alone has on performance.

B. *Data*

The imagery used in this work consists of all Landsat-8 imagery collected between 2014-2021, atmospherically corrected, and at its native 30 m spatial resolution. Input spectral bands include Red, Blue, Green, NIR, SWIR1, and SWIR2. The images were first transformed using Linear Spectral Mixture Analysis (LSMA) [27] into fractions of four Endmembers (High Albedo Surface, Low Albedo Surface, Vegetation, and Soil), and NDVI was also calculated [28]. Candidate change pixels were generated by applying roboBayes across sixteen different geographic regions, which we labeled by biome zone and humidity level, as shown in Table 1. Annotations were provided by [18].

C. *Change detection: roboBayes*

roboBayes is a multivariate, Bayesian change detection algorithm that runs online and extracts changepoints by fusing information from multiple remotely sensed data streams to calculate the probability of recent change. The hyperparameters governing the prior distributions for running roboBayes are estimated by fitting multivariate linear regressions to a sample of historical time series from the geographical region of interest. This helps to lower latency in change detections when applying the algorithm in an online monitoring mode. roboBayes was applied to every region using the default parameters outlined in [29] to detect changes across broad spatial scales. roboBayes is sensitive to and detects any anthropogenic or natural landcover changes. For this reason, a filtering system is needed to recover the desired change class(es).

TABLE I. DISTRIBUTION OF REGIONS USED FOR MODELING

Zone	**Humidity Level**		
	Arid	***Semi-Arid***	***Humid***
Temperate	3 regions	1 region	8 regions
Tropical	0 regions	0 regions	4 regions

D. *ML Filter System*

After candidate change pixels were generated, two different filtering sequences were separately tested to classify pixels as positive (heavy construction) or negative (not heavy construction). First, an expert system, devised by the authors as a baseline methodology and defined as a suite of human-created IF-THEN rules with set thresholds and based on the difference in mean values of the fitted time series harmonic models before and after a change is detected. The premise is that expert systems draw on human perceptual abilities combined with domain knowledge to better assign each pixel as positive or negative. For example, if the model observes a sharp increase in NDVI after a change detection, this could be filtered out due to the change likely being associated with phenological greening or reforestation rather than heavy construction. The second filtering sequence – illustrated by Fig. 2 – is a ML-based approach that tested the decision tree, random forest, and LightGBM algorithms to remove error-of-commission pixels.

The ML models were trained using roboBayes model coefficients including the mean, variance, and amplitude of the harmonic models fit for each signal before and after a detected change point. Spatial information was encoded into features by computing the number of contiguous and adjacent change pixels for each candidate pixel. A biome map obtained from [30] was included to enable tree splitting based on characteristics unique to each biome. We compared the performance of each ML model in addition to ML versus expert-system filters using 16 regions, wherein each region served as a test region and the other 16 were used for training and hyperparameter tuning. We also assessed the performance of applying principal component analysis (PCA) [31] on roboBayes model coefficients. Performance was evaluated at the pixel level using recall (a measure of omission error, precision (a measure of commission error), and F1 score (a measure of overall accuracy).

III. RESULTS

A. *Feature and Model Selection*

We conducted leave-on-region-out cross validation across all 16 regions and aggregated true positives, false positives, and

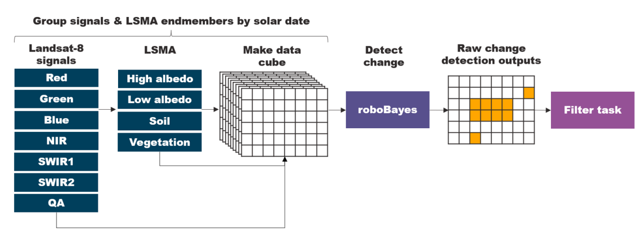

Fig. 1. BAS pipeline.

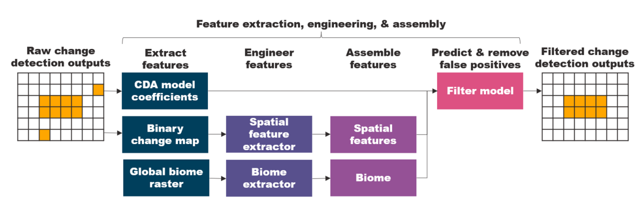

Fig. 2. ML filter system.

false negatives to compute precision, recall, and F1 at the algorithm level. Table II summarizes each algorithm's test-on-test performance across all feature permutations (e.g., PCA / no PCA) and all regions. Based on F1, LightGBM and random forest outperformed the simpler decision tree algorithm.

Based on the results of Table 2, we selected LightGBM and random forest for subsequent analysis. Table 3 provides the test-on-test results of LightGBM versus random forest across all feature permutations included in the analysis. Both LightGBM and random forest performed better without PCA, with spatial features, and with biome features. Based on these results, we selected a configuration that did not use PCA, and which included spatial features and biome data. We selected random forest, which outperformed LightGBM for this feature permutation, as our filtering algorithm.

B. *Performance of Random Forest versus Baselines*

We then compared the selected ML configuration's performance to that of the expert filters and no filtering. For each region, labeled by biome zone and humidity level, the random forest outperformed the alternatives.

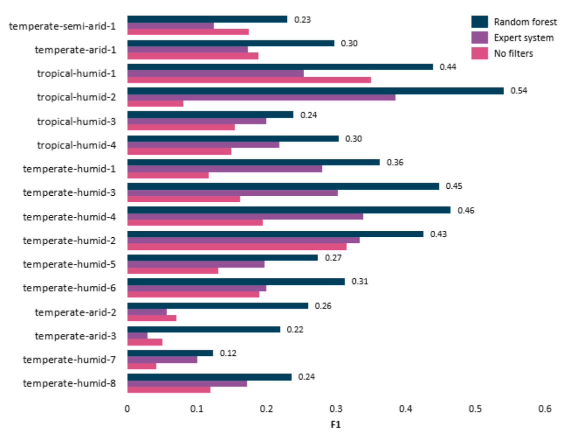

Fig. 3. Random forest performance across regions.

TABLE II. ALGORITHM PERFORMANCE

Algorithm	Precision	Recall	F1
LightGBM	0.256	0.459	0.287
Random forest	0.242	0.487	0.283
Decision tree	0.177	0.438	0.221

TABLE III. ALGORITHM PERFORMANCE BY FEATURE SET

Feature set			Algorithm	
PCA	***Spatial Features***	***Biome***	***LightGBM***	***Random Forest***
No	No	No	0.290	0.291
No	No	Yes	0.292	0.293
No	Yes	No	0.288	0.289
No	**Yes**	**Yes**	0.298	**0.304**
Yes	No	No	0.272	0.265
Yes	No	Yes	0.283	0.256
Yes	Yes	No	0.282	0.279
Yes	Yes	Yes	0.288	0.282

C. *Effect of Adding Regions to the Training Set*

We explored the effect of adding more regions to the training. To do this, we let the temperate-humid-6 region be the holdout and trained separate models on an increasing number of the other regions. We generated 30 orderings of the remaining 15 regions and aggregated by number of training regions to obtain the learning curve depicted by Fig. 3. Performance increases from zero training regions – obtained by using the unfiltered outputs – through 15 regions, albeit at a decreasing rate. Versus the baseline of unfiltered results, we see that reasonable improvement in F1 is reached after the model has seen only a few regions.

D. *On the relatively low F1 scores*

Ideally, BAS in the context of change detection is used to narrow the search space so that more powerful yet computationally expensive algorithms (e.g., deep neural networks) can be employed to more accurately determine whether the targeted change has occurred. Consequently, BAS F1 scores may seem low.

IV. CONCLUSIONS

Our BAS system combined an unsupervised CDA with a supervised tree-based classifier that outperformed the CDA alone and the CDA with the expert filters. An upside of this approach is that one can simply retrain the classifier on another set of annotations to detect a different change class. Another benefit is that the classifier's predictive power improves as more annotations are added to the training set. The reliance on training data is the key downside of our approach, but this can be mitigated through the adoption of a semi-supervised approach, whereby the CDA is used to detect potential change candidates and the researcher can decide whether they should be included in the annotation set or not.

Fig. 4. Learning curve for increasing training regions.

Acknowledgment

This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA) via 2021-20111.000006. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S Government. The U.S Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

References

[1] K. Winkler, R. Fuchs, M. Rounsevell, and M. Herold, “Global land use changes are four times greater than previously estimated,” Nat Commun, vol. 12, no. 1, p. 2501, May 2021, doi: 10.1038/s41467-021-22702-2.

[2] R. A. Houghton et al., “Carbon emissions from land use and land-cover change,” Biogeosciences, vol. 9, no. 12, pp. 5125–5142, Dec. 2012, doi: 10.5194/bg-9-5125-2012.

[3] E. F. Lambin, H. Geist, R. R. Rindfuss, “Introduction: Local Processes with Global Impacts,” in Land-Use and Land-Cover Change, E.F. Lambin and H. Geist, Eds., Heidelberg, Germany: Springer Berlin, 2006, ch. 1, pp. 1–8.

[4] J. J. Feddema et al., “The Importance of Land-Cover Change in Simulating Future Climates,” Science, vol. 310, no. 5754, pp. 1674–1678, Dec. 2005, doi: 10.1126/science.1118160.

[5] M. C. Hansen et al., “High-Resolution Global Maps of 21st-Century Forest Cover Change,” Science, vol. 342, no. 6160, pp. 850–853, Nov. 2013, doi: 10.1126/science.1244693.

[6] J.-F. Pekel, A. Cottam, N. Gorelick, and A. S. Belward, “High-resolution mapping of global surface water and its long-term changes,” Nature, vol. 540, no. 7633, pp. 418–422, Dec. 2016, doi: 10.1038/nature20584.

[7] X.-P. Song et al., “Global land change from 1982 to 2016,” Nature, vol. 560, no. 7720, pp. 639–643, Aug. 2018, doi: 10.1038/s41586-018-0411-9.

[8] X. Zhan et al., “The 250 m global land cover change product from the Moderate Resolution Imaging Spectroradiometer of NASA’s Earth Observing System,” International Journal of Remote Sensing, vol. 21, no. 6–7, pp. 1433–1460, Jan. 2000, doi: 10.1080/014311600210254.

[9] Z. Zhu, S. Qiu, and S. Ye, “Remote sensing of land change: A multifaceted perspective,” Remote Sensing of Environment, vol. 282, p. 113266, Dec. 2022, doi: 10.1016/j.rse.2022.113266.

[10] C. E. Woodcock et al., “Free Access to Landsat Imagery,” Science, vol. 320, no. 5879, pp. 1011–1011, May 2008, doi: 10.1126/science.320.5879.1011a.

[11] M. A. Wulder et al., “The global Landsat archive: Status, consolidation, and direction,” Remote Sensing of Environment, vol. 185, pp. 271–283, Nov. 2016, doi: 10.1016/j.rse.2015.11.032.

[12] C. E. Woodcock, T. R. Loveland, M. Herold, and M. E. Bauer, “Transitioning from change detection to monitoring with remote sensing: A paradigm shift,” Remote Sensing of Environment, vol. 238, p. 111558, Mar. 2020, doi: 10.1016/j.rse.2019.111558.

[13] Z. Zhu and C. E. Woodcock, “Continuous change detection and classification of land cover using all available Landsat data,” Remote Sensing of Environment, vol. 144, pp. 152–171, Mar. 2014, doi: 10.1016/j.rse.2014.01.011.

[14] R. E. Kennedy, Z. Yang, and W. B. Cohen, “Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms,” Remote Sensing of Environment, vol. 114, no. 12, pp. 2897–2910, Dec. 2010, doi: 10.1016/j.rse.2010.07.008.

[15] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, “Detecting trend and seasonal changes in satellite image time series,” Remote Sensing of Environment, vol. 114, no. 1, pp. 106–115, Jan. 2010, doi: 10.1016/j.rse.2009.08.014.

[16] S. V. Stehman, B. W. Pengra, J. A. Horton, and D. F. Wellington, “Validation of the U.S. Geological Survey’s Land Change Monitoring, Assessment and Projection (LCMAP) Collection 1.0 annual land cover products 1985–2017,” Remote Sensing of Environment, vol. 265, p. 112646, Nov. 2021, doi: 10.1016/j.rse.2021.112646.

[17] D. Li, E. Bou-Zeid, and M. Oppenheimer, “The effectiveness of cool and green roofs as urban heat island mitigation strategies,” Environ. Res. Lett., vol. 9, no. 5, p. 055002, May 2014, doi: 10.1088/1748-9326/9/5/055002.

[18] H. Goldberg, C. Ratto, A. Banerjee, M. Kelbaugh, M. Giglio, E. Vermote, “Automated global-scale detection and characterization of anthropogenic activity using multi-source satellite-based remote sensing imagery,” presented at Geospatial Informatics XIII, Orlando, FL, United States, May 2023, no. 12525-1.

[19] Q. Weng, “Remote sensing of impervious surfaces in the urban areas: Requirements, methods, and trends,” Remote Sensing of Environment, vol. 117, pp. 34–49, Feb. 2012, doi: 10.1016/j.rse.2011.02.030.

[20] Q. Zhang and K. C. Seto, “Mapping urbanization dynamics at regional and global scales using multi-temporal DMSP/OLS nighttime light data,” Remote Sensing of Environment, vol. 115, no. 9, pp. 2320–2329, Sep. 2011, doi: 10.1016/j.rse.2011.04.032.

[21] D. C. Wickramasinghe, T. T. Vu, and T. Maul, “Satellite remote-sensing monitoring of a railway construction project,” International Journal of Remote Sensing, vol. 39, no. 6, pp. 1754–1769, Mar. 2018, doi: 10.1080/01431161.2017.1415481.

[22] K. Kumarapu, M. Shashi, and V. R. Keesara, “UAV in Construction Site Monitoring and Concrete Strength Estimation,” J Indian Soc Remote Sens, vol. 49, no. 3, pp. 619–627, Mar. 2021, doi: 10.1007/s12524-020-01246-w.

[23] L. J. Wendelberger, J. M. Gray, B. J. Reich, and A. G. Wilson, “Monitoring Deforestation Using Multivariate Bayesian Online Changepoint Detection with Outliers,” 2021, doi: 10.48550/ARXIV.2112.12899.

[24] L. Breiman, J. H. Friedman, R. A. Olshen, Classification and regression trees, Monterey, CA, United States: Wadsworth & Brooks/Cole Advanced Books & Software, 1984.

[25] L. Breiman, “Random Forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, 2001, doi: 10.1023/A:1010933404324.

[26] G. Ke et al., "Lightgbm: A Highly Efficient Gradient Boosting Decision Tree," in Advances in neural information processing systems 30: 31st Annual Conference on Neural Information Processing Systems (NIPS 2017): Long Beach, California, USA, 4-9 December 2017. U. von Luxburg et al., Eds., Red Hook, NY: Curran Associates, Inc, 2018.

[27] C. Wu and A. T. Murray, “Estimating impervious surface distribution by spectral mixture analysis,” Remote Sensing of Environment, vol. 84, no. 4, pp. 493–505, Apr. 2003, doi: 10.1016/S0034-4257(02)00136-0.

[28] C. J. Tucker, “Red and photographic infrared linear combinations for monitoring vegetation,” Remote Sensing of Environment, vol. 8, no. 2, pp. 127–150, May 1979, doi: 10.1016/0034-4257(79)90013-0.

[29] L. Wendelberger, J. Gray, B. Reich, A. Wilson, A. Holloway. roboBayes: Robust Online Bayesian Monitoring, 1.1 ed. (2022). Accessed: May 29, 2023. [Online]. Available: https://cran.r-project.org/web/packages/roboBayes/index.html.

[30] P. Reich and H. Eswaran, “Biomes,” in Terrestrial Ecosystems and Biodiversity, Y. Wang, Ed., 2nd ed. Boca Raton, FL, United States: CRC Press, 2023, pp. 75-80.

[31] A. C. Rencher and W. F. Christensen, “Principal Component Analysis,” in Methods of Multivariate Analysis: Rencher/Methods. in Wiley Series in Probability and Statistics. Hoboken, NJ, USA: John Wiley & Sons, Inc., pp. 405–433, doi: 10.1002/9781118391686.