Implementation of Bayesian Online Changepoint Detection Algorithm to Monitor Deforestation

1st Peter Rasmussen
Whiting School of Engineering
Johns Hopkins University
Baltimore, United States
prasmus3@jh.edu

***Abstract*—Land cover change detection and monitoring techniques use remote sensing data to find and monitor a variety of natural and anthropogenic events and activities. Detecting change at scale - at the city or region level - is computationally burdensome. To reduce runtime and memory usage, lower resolution satellite data - on the order of 10m to 30m - can be used for broad area search change detection. A desirable broad area search change detection algorithm would be online and have runtime and memory complexities comparable to the state-of-the-art. This paper proposes the following. First, we aim to implement an online change detection algorithm, roboBayes, in Python to make it more widely available to the remote sensing community. Second, we will explore the feasibility of using a variation of roboBayes, MR roboBayes, that employs wavelet decomposition on lower resolution satellite imagery. Third, we plan on validating the Python implementation of roboBayes using annotated benchmark deforestation data.**

***Index Terms*—change detection, remote sensing, deforestation, Bayesian**

I. Introduction

Remote sensing data - which comprise visible and non-visible spectra captured by orbiting satellites - is the essential input for landcover change detection algorithms. For broad area search, lower resolution, publicly available data is suitable. The Landsat program collects spectral data using sensors deployed across multiple satellites operating in sun-synchronous orbit, which provides users with regular-interval time series imagery data [1]. Multiple established landcover change detection (LCCD) algorithms use Landsat data, including Continuous Change Detection and Classification (CCDC) and Robust Online Bayesian Monitoring (roboBayes).

II. Related Work

CCDC fits multivariate linear models with seasonality terms to pixel-level data. The algorithm produces segments delimited by change. Change is detected if the current model error exceeds a threshold; if so, the change point is recorded and a new model is fit. CCDC is widely used and an implementation exists in Google Earth Engine. As it is an offline model, it does not meet this paper's aim of implementing an online change detection algorithm.

Online change detection algorithms are better suited for monitoring as they update given new data. Reference [3] developed one such algorithm, Bayesian Online Changepoint Detection (BOCD), which predicts the "probability distribution of the current 'run', or time since the last changepoint, using a simple message-passing algorithm".

Reference [4] adapted BOCD to landcover change detection by adding a "multivariate linear regression framework that supports seasonal trends" to successfully detect deforestation in Myanmar. This algorithm, now known as Robust Online Bayesian Monitoring (roboBayes), predicts change at the pixel-level. roboBayes has a runtime complexity of $O(tn^2)$, where n is the number of sensor signals and t is the number of time steps. The n^2 complexity is due to the generation of a covariance matrix at each time step.

A more recent iteration of the algorithm, multiresolution roboBayes (MR roboBayes), uses wavelet decomposition on high-resolution, gapless 3m satellite data to find changes across pixels [5]. To perform the wavelet decomposition, gapless time series data is required. Therefore, clouds and missing data must be imputed to perform the decomposition.

III. Proposed Methodology

A. Datasets

This analysis will use level 2 Landsat data over the period 2014 through 2021. All available spectra will be employed to perform the change detection. Principal components analysis (PCA) may be used to reduce dimensionality to decrease runtime and memory usage. We will use publicly available annotated deforestation data in one to two regions; if none is available we will self-annotate.

B. Pipeline

A data pipeline will handle data download, preprocessing, model execution, post-processing, and visualization. Code will be version controlled in GitHub and an executable Docker image will be available to enable reproduction of the results.

Preprocessing will consist of organizing each scene - a multi-band raster pertaining to one time step - into a spatiotemporal data cube. In the case of the MR roboBayes experiment, missing data will be interpolated.

A Python implementation of roboBayes will ingest the data cube and predict change at the pixel level. In the MR roboBayes case, pixel clusters will undergo wavelet decomposition.

Post-processing will organize clusters of change pixels into segments, similar to the output CCDC produces. Clusterized pixels will be vectorized and visualized in a simple slippy map hosted locally.

IV. Conclusion

This effort will create the first Python implementation of roboBayes. In addition, we aim to extend MR roboBayes to predict on lower resolution data, something which the MR roboBayes creators have not done. Finally, we will validate the Python implementation using publicly available annotations or self-made annotations.

References

[1] U.S. Geological Survey, "Landsat-Earth observation satellites," USGS Publications Warehouse, version 1.4, 2022.

[2] 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, 2014.

[3] R. P. Adams and D. J. MacKay, "Bayesian online changepoint detection," arXiv e-prints, arXiv-0710, 2007.

[4] L. J. Wendelberger, J. M. Gray, B. J. Reich, A. G. Wilson, "Monitoring deforestation using multivariate Bayesian online changepoint detection with outliers," arXiv preprint, arXiv:2112.12899, 2021.

[5] L. J. Wendelberger, J. M. Gray, A. G. Wilson, R. Houborg, B. J. Reich, "Multiresolution broad area search: Monitoring spatial characteristics of gapless remote sensing data," Journal of Data Science, 20.4, 2022.