Exploring BFAST to detect forest changes in Portugal

Hugo Costa[*a,b], Anny Giraldo[c], Mário Caetanoa,[b]

[a]Direção-Geral do Território, 1099-052 Lisboa, Portugal;

[b]NOVA Information Management School (NOVA IMS), Universidade Nova Lisboa, Campus de Campolide, 1070-312 Lisbon, Portugal;

[c]Hochschule für Technik Stuttgart, 70174 Stuttgart, Germany

Exploring BFAST to detect forest changes in Portugal

Hugo Costa*[a,b], Anny Giraldo[c], Mário Caetano[a,b]

[a]Direção-Geral do Território, 1099-052 Lisboa, Portugal; [b]NOVA Information Management School
(NOVA IMS), Universidade Nova Lisboa, Campus de Campolide, 1070-312 Lisbon, Portugal;
[c]Hochschule für Technik Stuttgart, 70174 Stuttgart, Germany

ABSTRACT

Landsat 8 data and Breaks For Additive Season and Trend (BFAST) were used in a region of central Portugal to detect
forest clear-cuts and burnt areas. A total of 79 Landsat 8 images from 2013 to 2019 were downloaded for path/row
204/032, and the NDVI was calculated. The same data processing was done for path/row 203/032 to create a denser time
series in the overlapping area, which increased to 124 images. The output of the analysis is a binary map of change (i.e.,
forest loss) and no-change. A probabilistic accuracy assessment based on random stratified sampling was implemented
with 100 random points per stratum. Each point was interpreted as being either "no-change", "clear-cut" or "burnt area"
based on reference data. Furthermore, the date of change (if any) was defined. Results show an overall accuracy of
0.85±0.02 for the binary classification with omission and commission errors of class "Change" of 0.30±0.02 and
0.19±0.02. Moreover, it is estimated that 32% of the forested area in path/row 204/032 went through at least one episode
of clear-cut or fire in the period analyzed. The time lag between the date of change and detection was about 2.5 months
on average, which decreased to 1.5 months in the regions of the denser time series. The results are promising but BFAST
is somewhat slow and hence some concerns remain about its efficiency in operation use.

Keywords: Landsat, NDVI time series, Forest clear cuts, Burnt areas, change detection

1. INTRODUCTION

Most of the Portuguese mainland has been historically covered by forest and shrublands, but large of this area suffers
regular change within a cycle interchanging between vegetation loss and growth. The dynamics of forested areas and
shrublands is mainly driven by clear-cuts and wildfires, and there is great interest in monitoring these phenomena for
multiple purposes, such as landscape planning, wildfire monitoring and greenhouse gas emission reporting. Traditional
methods to collect data on forest and shrublands exist, such as the national forest inventory, but there are limitations.
Existing methods are typically expensive and slow, founded on sampling (whereas there is interest in full coverage, wall-
to-wall products), and do not address all the needs. For example, the Portuguese *Instituto da Conservação da Natureza e
das Florestas* (ICNF) produces an official annual map of burnt areas, but clear-cuts are not included. To address these
issues, the Portuguese *Direção-Geral do Território* (DGT) is doing research to develop new products and services to
provide the society with updated and wall-to-wall information on the status and change of forest and shrublands.

This paper describes results obtained with Landsat 8 time series and Breaks For Additive Season and Trend (BFAST)[1] in
a region of central Portugal. BFAST has been used to analyze Normalized Difference Vegetation Index (NDVI) time
series and other vegetation indices to detect deforestation in a range of biomes[2,3], but rarely in Mediterranean and
Atlantic biomes (but see [4]). BFAST decomposes a time series into trend and seasonal components with methods for
detecting abrupt changes within the two components. The breaks detected in the pixels denote a potential change in
vegetation cover with a date associated, thus enabling vegetation loss to be detected in space and time.

BFAST was used to detect forest clear-cuts and burnt areas, the two most relevant events impacting forest in Central
Portugal. The goal was to evaluate the accuracy of BFAST in analyzing historical NDVI data of Landsat 8. Special focus
was given to the performance of BFAST as a function of data density to understand the potential benefit of using denser
time series acquired by Sentinel-2 satellites when they become longer in the future.

*hcosta@dgterritorio.pt; phone +351 21 381 96 00;

2. STUDY AREA AND DATA

This study covers an area of more than 2 million hectare in the centre of Portugal (Figure 1). This corresponds to path/row 204/032 of Landsat, which covers a dense forested area in Portugal, in which clear-cuts and wildfires are abundant and most prominent. The forest found in this path/row includes 44% of maritime pine, 27% of eucalyptus, 18% of shrublands, 6% of broadleaves, and 5% of oaks[5].
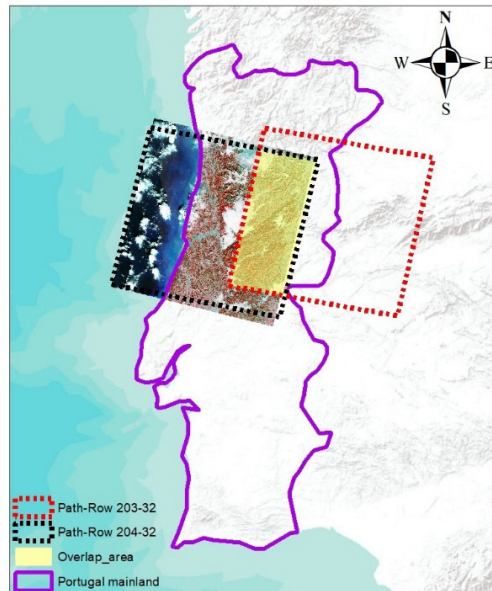

Figure 1. Study area and footprint of Landsat path/row 204/32 and 203/33.

The archive of Landsat 8 was downloaded and a total of 79 images (level 2, surface reflectance) from 17 April 2013 to 25 September 2019 with less than 30% cloud cover were used. Data for the neighboring path/row 203/032 was also processed to build a denser time series in the region overlapping with the study area (yellow stripe Figure 1). This increased to data density to 124 images in the overlapping region. Figure 2 shows the number of observations of the time series. The goal was to inspect the performance of BFAST as a function of the amount of data.
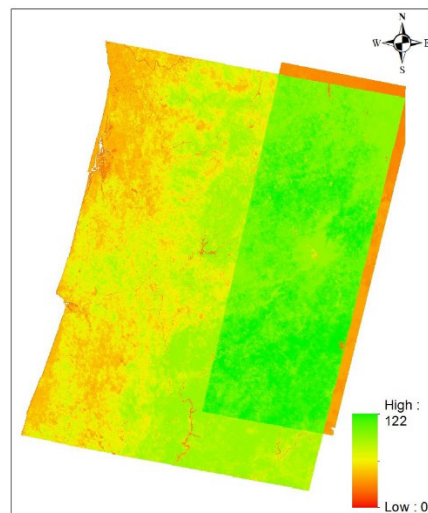

Figure 2. Number of valid observations per pixel in the NDVI Landsat time series.

The NDVI was calculated from each data acquisition to be analyzed with BFAST. Also, data acquisitions were masked to remove pixels related to non-forest. This was based on the Portuguese national land cover and land use map (COS) of 2010. Although COS is available for the reference years of 2015 and 2018, the year of 2010 was used as it refers to a date before the time series starts[5].

The removal of clouds and shadows and the denser time series in the area overlapping the neighboring path/row produced an irregular time series. As BFAST requires a regular time series[1], that is, all pixels must hold a valid value in all dates, linear interpolation was used to fill all the gaps. Finally, the time series under analysis was a series of NDVI observed or interpolated values for pixels of forest and shrublands.

3. METHODS

BFAST[1] is a method that analyses a univariate time series and decomposes it into three components: trend (T_t), seasonal (S_t) and noise (e_t). The general model of BFAST is an addictive decomposition of the tree components in the form of

$$Y_t = T_t + S_t + e_t \quad (t = 1, 2, \ldots, n) \tag{1}$$

in which Y_t is the observed data at time t. The model iteratively fits a piecewise linear trend and seasonal model and the intercept and slope of consecutive linear models define breaks in the signal and can be used to derive the magnitude and direction of abrupt changes. Magnitude is the difference between the trend components of consecutive linear models. The magnitude and its direction are one of the main outputs of BFAST and can be used to detect clear-cuts and burnt areas in space and time. BFAST is available from the Comprehensive R Archive Network (CRAN), but a faster version available on githug (https://github.com/bfast2/bfast) was used. The code was run with a computer with 36 logical processors and 64 GB RAM, which proved to be 500% faster than the original code.

BFAST requires a set of parameters to run and the defaults were used[6], except parameter h. Parameter h is the minimal segment size between potentially detected breaks in the trend model. This parameter can be used to force the breakpoints to be apart a minimum period like one year and hence ensure that enough observations are used to fit robust linear models between potential breaks. However, large h values can prevent the detection of real changes that occur within short periods. Thus, a sensitivity analysis of h was conducted as follows.

Small subsets of the study area were defined with a size of 17 x 17 pixels (circa 26 hectares). Their location was randomly defined and a sample of five subsets was selected with characteristics representative of the whole area. The selected subsets (Figure 3) are located to include intact forest (subsets 5 and 7), clear-cuts (subset 14 and 15), and burnt areas (subset 2) in the non-overlapping and overlapping region (yellow area in Figure 3). Subset 2 is partially inside the overlapping region.
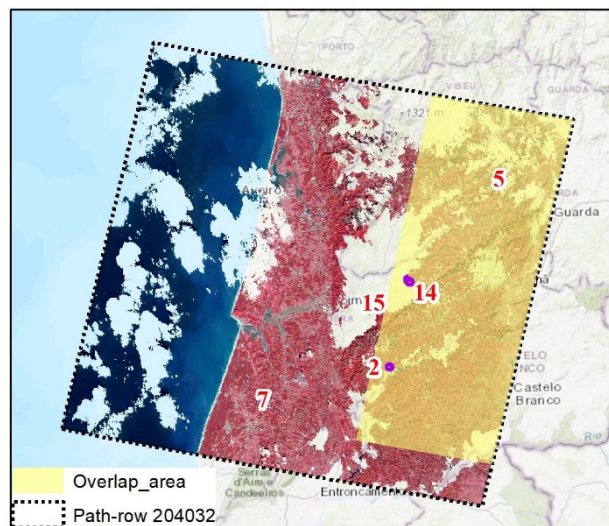

Figure 3. Location of the small data subsets used to conduct a sensitivity analysis of parameter h.

Following previous publications[2,6], the following values were tested: 1/2, 1/3, 1/5, and 1/10. Figure 4 shows an example of how BFAST behaves with different h values. In general, large values like 1/2 failed to detect clear-cuts and fires with accuracy, especially if they occurred at the extremes of the time series. On the other hand, very small h values like 1/10 allowed BFAST to fit to noise and irrelevant fluctuations of the signal and thus detected breaks not related to abrupt forest disturbances. Nevertheless, apparently the magnitude of change in the trend component always reveal the severity of the event with larger magnitudes being associated with abrupt changes such as clear-cuts and fires. The h value considered as optimal was 1/5 and used in all subsequent analyses.
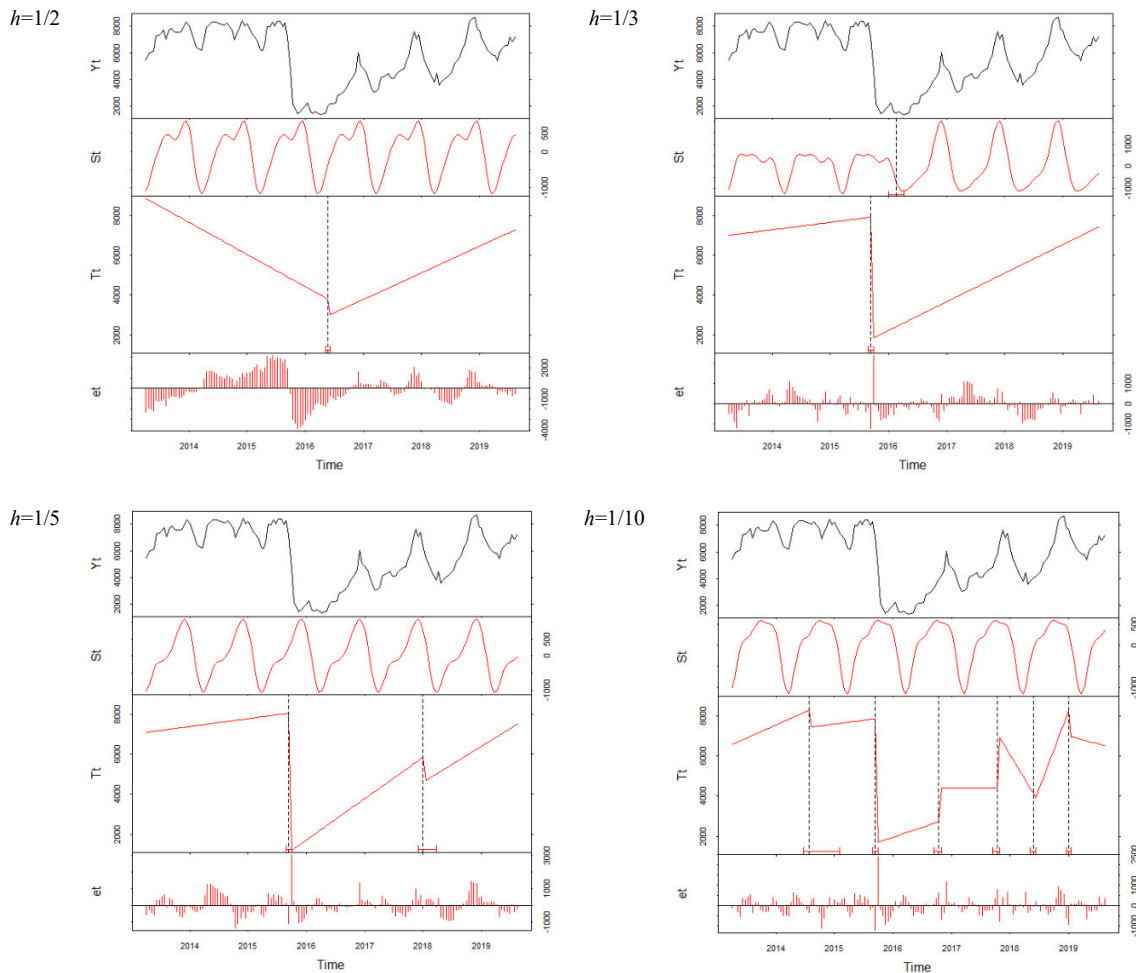

Figure 4. Results of BFAST as a function of parameter h for a pixel in subset 14 (see Figure 3) in which a clear-cut occurred in 2016. In each plot, the black line at the top represents the NDVI time series (Y_t), followed by components seasonal (S_t), trend (T_t) and noise (e_t) in red. The breaks detected by BFAST are represented by dashed vertical black lines.

Note that the analysis of non-disturbed forest also detected breaks in the trend component regardless of h value (Figure 5). However, the magnitude of change tends to be relatively small as compared to disturbed forest. Therefore, the magnitude of the breaks detected is key to distinguish real abrupt changes from other disturbances of the signal in the time series.

An additional analysis was needed to define a threshold of magnitude at which a decrease of NDVI should be considered as clear-cut or burnt area. A sample of pixels of known land cover (intact forest, clear-cut, and burnt forest) used in the sensitivity analysis above were manually classified as no-change (value 0) or change (value 1) depending on whether a clear-cut or fire occurred during the period analyzed. These set of points were compared to the magnitude of change in the trend component retrieved by BFAST (Figure 6), which makes clear that forest disturbance is associated with large magnitude values of NDVI change. Using a binomial logistic model, a threshold was set at -0.12, and all pixels with magnitude below -0.12 were considered as forest loss. The *h* value and threshold found as optimal were applied to the entire data set in the study area. The output was a binary map of change (i.e., forest loss) and no-change.
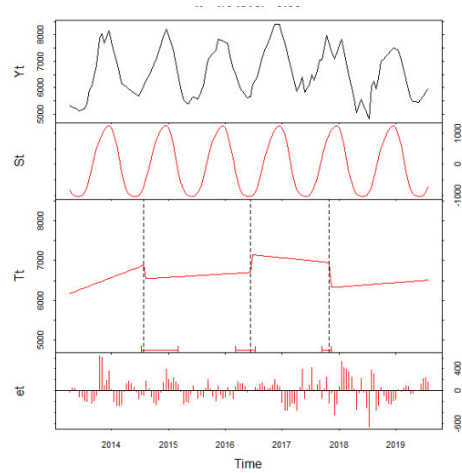

Figure 5. Results of BFAST with *h*=1/5 for a pixel of undisturbed forest in subset 7 (see Figure 3).
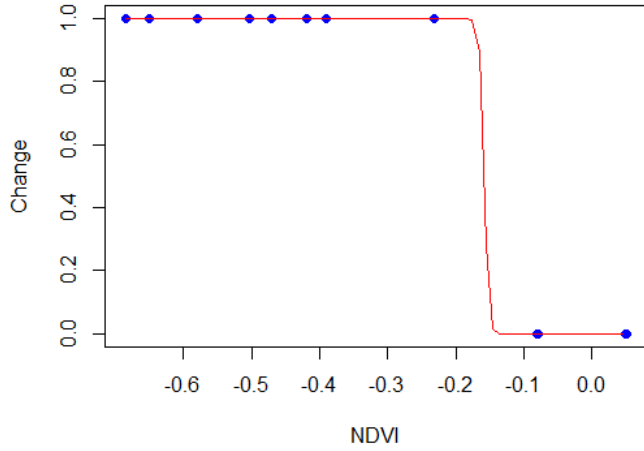

Figure 6. Relationship between the magnitude of change in the BFAST trend component and forest loss.

A probabilistic accuracy assessment of the binary map was implemented following good practices for accuracy and area estimation[7]. Four strata were defined to conduct random stratified sampling. Stratification was defined based on the threshold of change defined above (-0.12) as it is expected that errors are likely to occur in magnitude values close to the threshold. Therefore, two strata are defined to include all pixels with a magnitude of change nearby the threshold, namely below [-0.24; -0.12] and above the threshold [-0.12; 0]; In addition, pixels with magnitude values below -0.24 and above 0 formed another two strata. However, points were prevented to fall in the burnt area of 2017. This is because the extreme event of 2017 destroyed a very large area of forest (>500 000 ha)[8], mostly in the study area, and its inclusion in the assessment would impact negatively on the results as most of the points associated with change would certainly be

caused by this extreme event. However, the goal was to evaluate BFAST on average conditions, including a diversity of clear-cuts and fires occurred over time, and not only during a specific extreme event. Thus, the results of the accuracy assessment express the performance of BFAST in the study area except the burnt area of 2017. A total of 100 random points per stratum were collected. Each point was interpreted as being either "no-change" or "change" (with the additional labels "clear-cut" or "burnt area") based on reference data, including official orthophotos of 2015 and 2018 available for Portugal, Google Earth imagery, and the Landsat 8 dataset. Furthermore, the date of change (if any) was defined based on the same reference data.

4. RESULTS AND DISCUSSION

The predicted magnitude of change in the trend component and estimated date of change given by BFAST are shown in Figure 7. The extreme fires of 2017 are evident along with fires of 2016 and other years. Clear-cuts typically cover smaller areas and are not apparent in the figure at the scale used.
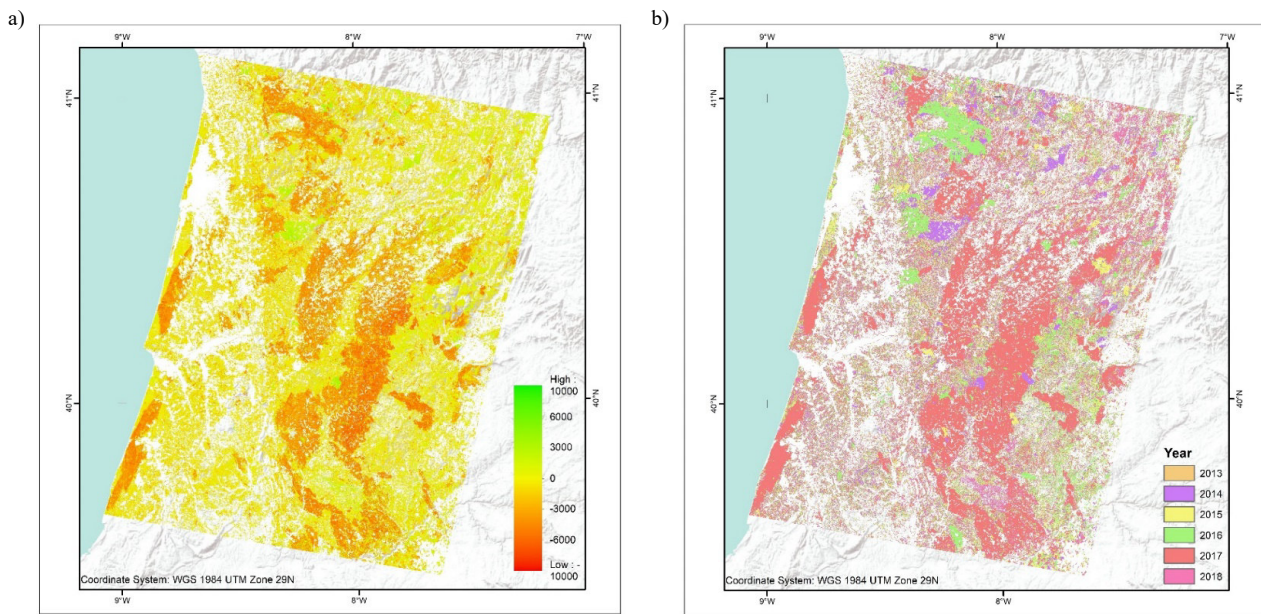

Figure 7. (a) Magnitude of change in the trend component given by BFAST; reddish values represent drastic negative changes of NDVI typically associated with vegetation loss; magnitude values are multiplied by 10000. (b) Estimated year of change.

Although most of the clear-cuts and fires are associated with a negative magnitude of change (decrease of NDVI), some cases are associated with a positive change (e.g. some fires of 2014). This is caused by the fact that BFAST is unable to detect events at the very beginning of the time-series, depending of parameter h.

Therefore, BFAST can detect recovery of vegetation that follows an early event, which is associated with a substantial increase of NDVI. This may be unexpected while monitoring vegetation loss, but can also be positively used with adequate adjustments. However, BFAST is also unable to detect events at the end of the time-series and such cases are missed completely as the time series includes no data on subsequent vegetation recovery.

The map of Figure 7(a) was reclassified to represent two classes, "change" and "no-change" (Figure 8). The accuracy assessment estimates an overall accuracy of 0.85±0.02 at the 95% confidence level, which is similar to some previous research[3]. Omission and commission errors of class "change" are 0.30±0.02 and 0.19±0.02. Table 1 shows the estimated proportion of each class in the study area and respective accuracy. Importantly, it is estimated that 32% of the forested area in path/row 204/032 went through at least one episode of forest loss (clear-cut or fires) in the period analyzed. Specifically, the estimated area of change is 333 886 ± 673 ha. Note that this excludes the 2017 fires, which alone damaged a larger area (> 500 000 ha) as mentioned above.

Another important result is the accuracy of the estimated date of the events. This is clearly a function of the density of the time series and this research defined a region inside the study area with a larger amount of data from a neighboring Landsat path/row (Figure 2). While the time lag between the real date of change and the estimated date of change in the non-overlapping area was about 2.5 months on average, for the overlapping area the time lag was cut to 1.5 months. This result is very good comparing to other studies in the tropics affected by frequent cloud cover[9]. For other sensors such as Sentinel-2 imagery, with a revisit time of 5 days, time lag can be even shorter.

Table 1. Error matrix of estimated area proportions. Map classes are in rows and reference classes in columns.

	No-change	Change	Total	User's accuracy	Producer's accuracy
No-change	0.6264	0.0978	0.7242	0.87±0.24	0.92±0.3
Change	0.0538	0.2220	0.2758	0.81±0.28	0.70±0.4
Total	0.6802	0.3198	1		

Another interesting aspect is that BFAST appears to detect burnt areas with more accuracy than clear-cuts. This may be partially related to the larger spatial extent typical of burnt areas, possibly something necessary for an event to be clearly visible at the resolution of Landsat, but possibly the spectral traits of fires also play a role here. This needs further investigation however, along with potential benefits of using spectral indices other than the NDVI[9].

The results obtained with Landsat 8 indicate a potential benefit of using Sentinel-2 data, which will become more relevant as data acquisitions accumulates in the future and longer time series become available. However, there are some concerns about the efficiency of the method because the data set analyzed needed around eight days to process, even with the faster version of BFAST available on github. Denser time series with the larger spatial resolution will require longer processing times to analyze equivalent spatial extents, which can correspond to several weeks of processing to cover a relatively small country like Portugal. Although BFAST is not a near-real time analysis tool (the near-real time version of BFAST[10] was not tested), very long processing times are not desirable.
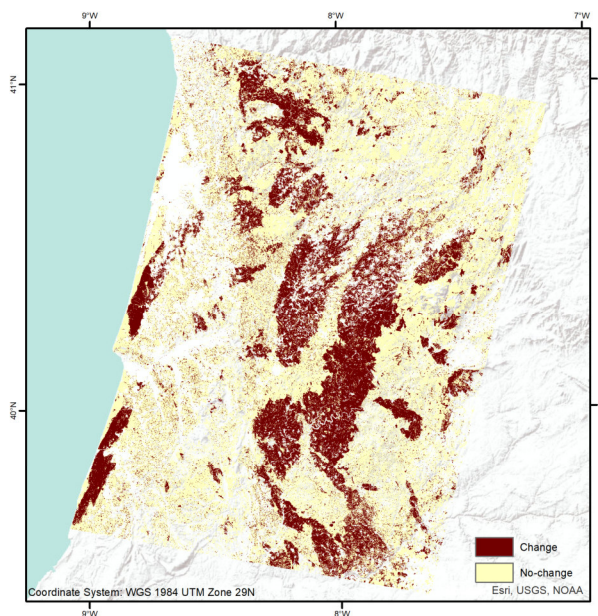

Figure 8. Map of change/no-change (vegetation loss).

ACKNOWLEDGEMENTS

REFERENCES

[1]   Verbesselt, J., Hyndman, R., Newnham, G. and Culvenor, D., "Detecting trend and seasonal changes in satellite image time series," Remote Sens. Environ. **114**(1), 106–115 (2010).

[2]   Watts, L. M. and Laffan, S. W., "Effectiveness of the BFAST algorithm for detecting vegetation response patterns in a semi-arid region," Remote Sens. Environ. **154**(1), 234–245 (2014).

[3]   Wu, L., Li, Z., Liu, X., Zhu, L., Tang, Y., Zhang, B., Xu, B., Liu, M., Meng, Y. and Liu, B., "Multi-type forest change detection using BFAST and monthly landsat time series for monitoring spatiotemporal dynamics of forests in subtropical wetland," Remote Sens. **12**(2) (2020).

[4]   Lambert, J., Denux, J. P., Verbesselt, J., Balent, G. and Cheret, V., "Detecting clear-cuts and decreases in forest vitality using MODIS NDVI time series," Remote Sens. **7**(4), 3588–3612 (2015).

[5]   Caetano, M., Igreja, C., Marcelino, F. and Costa, H., "Estatísticas e dinâmicas territoriais multiescala de Portugal Continental 1995-2007-2010 com base na Carta de Uso e Ocupação do Solo (COS)" (2017).

[6]   Verbesselt, J., Hyndman, R., Zeileis, A. and Culvenor, D., "Phenological change detection while accounting for abrupt and gradual trends in satellite image time series," Remote Sens. Environ. **114**(12), 2970–2980 (2010).

[7]   Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E. and Wulder, M. A., "Good practices for estimating area and assessing accuracy of land change," Remote Sens. Environ. **148**, 42–57 (2014).

[8]   San-Miguel-Ayanz, J., Durrant, T., Boca, R., Libertà, G., Branco, A., de Rigo, D., Ferrari, D., Maianti, P., Artés Vivancos, T., Costa, H., Lana, F., Löffler, P., Nuijten, D., Christofer Ahlgren, A. and Leray, T., [Forest fires in Europe, Middle East and North Africa 2017], Publications Office of the European Union (2018).

[9]   Smith, V., Portillo-Quintero, C., Sanchez-Azofeifa, A. and Hernandez-Stefanoni, J. L., "Assessing the accuracy of detected breaks in Landsat time series as predictors of small scale deforestation in tropical dry forests of Mexico and Costa Rica," Remote Sens. Environ. **221**(December 2018), 707–721 (2019).

[10]   Verbesselt, J., Zeileis, A. and Herold, M., "Near real-time disturbance detection using satellite image time series," Remote Sens. Environ. **123**, 98–108 (2012).