Evaluating the Potential of the PROBA-V Sensor in Estimating Forest Cover Change Over a Range of European Biogeographical Regions: The FM@PROBA-V Project

Ioannis Manakos[1], Ioannis Gitas[2], Mathias Schardt[3], Chariton Kalaitzidis[1], Eleni Dragozi[2] and Heinz Gallaun[3]

[1]International Centre for Advanced Mediterranean Agronomic Studies, Mediterranean Agronomic Institute of Chania, Department of Geoinformation in Environmental Management, Chania, Greece, manakos@maich.gr, chariton@maich.gr

[2]Aristotle University of Thessaloniki, School of Forestry and Natural Environment, Thessaloniki, Greece,
igitas@for.auth.gr, edragozi@for.auth.gr

[3]DIGITAL, Graz, Austria,
mathias.schardt@joanneum.at, heinz.gallaun@joanneum.at

Abstract. The PROBA-V mission is designed to replace the VEGETATION instruments onboard the SPOT-4 and SPOT-5 satellites and ensure data continuity for the VEGETATION users until the launch of the Sentinel-2 series, as well as compliment data provision, even after the Sentinel-2 series become fully operational. In addition, the PROBA-V mission will offer data at a higher spatial resolution, while maintaining the spectral coverage and near daily global coverage of its predecessor. Amongst the various applications that VEGETATION data have seen over the past 15 years, forest monitoring had been one of the most prominent ones. The FM@PROBA-V project, funded by the Belgian Science Policy Office (BELSPO), achieved three objectives in an effort to quantify the improvements on forest monitoring brought along with the increase of the spatial resolution by the new sensor: a) it provided an estimate of European forest cover with 300, 600, and 1000 meter resolution data, b) it evaluated a number of popular classification methods on VEGETATION and MODIS 250 meter data – serving as a substitute of PROBA-V data due to the similar spatial resolution – on five test sites representing different biogeographical regions of Europe, and c) assessed the two sensors in their capacity in mapping forest cover changes caused by disastrous events, such as forest fires and wind damage. The results suggest that the increased spatial resolution will offer an opportunity for more accurate forest monitoring at global, continental, and regional scale, particularly on highly fragmented landscapes.

Keywords. PROBA-V, forest mapping, forest cover change detection.

1. Introduction

The VEGETATION program was conceived in the early 1990s as a means to provide daily monitoring of vegetation at a regional and global scale, leading to the first payload becoming operational onboard SPOT4 on April 1998. A second payload was deployed four years later, onboard SPOT5 on May 2002. The instrument offers near daily global coverage (5 acquisitions every 6 days at the equator) due to its wide swath (2200 Km), and data is acquired in four channels (blue, red, NIR and SWIR) at a 1.15 Km nadir spatial resolution. Ten-day synthesis data, older than three months, are available free of charge to the public.

Following more than 20 years of successful operation, the VEGETATION instruments will no longer be available after the end of 2012. The upcoming Sentinel-2 series, prepared by ESA within the framework of Global Monitoring for Environment and Security (GMES), will be able to substitute the services of VEGETATION in the future. However, the Sentinel-2 series will not be available until after the discontinuation of VEGETATION, as the scheduled launch date for the first of the two Sentinel-2 is at the end of 2013, with the second satellite following a year later (ESA, 2011). In order to maintain data continuity to the VEGETATION users, the Belgian Federal Science Policy Office (BELSPO), in collaboration with ESA, took advantage of the experiences acquired through the successful PROBA mission and decided to construct the PROBA-V satellite (with “V” standing for Vegetation). The new mission, will provide data with specifications similar to the ones provided by the VEGETATION instruments and, in addition, will have improved spatial resolution (up to 1/3 of a Km in the VNIR and 2/3 of a Km in the SWIR channel), in comparison to the VEGETATION instrument, while maintaining the capability of daily global coverage. In addition to covering the data gap, the mission will supplement the Sentinel-2 satellites in the early stages of their mission.

1.1. The FM@PROBA-V project

In 2011, BELSPO launched a Preparatory Program, in order to assess the compatibility between VEGETATION and PROBA-V data on a wide range of applications served by the former instrument. The program aimed to maximize the potential benefits that can be gained from the improved characteristics of PROBA-V in various applications. Within the framework of the PROBA-V Preparatory Program, the project “Forest cover change monitoring with PROBA-V: Potentials and limitations on European terrain (FM@PROBA-V)” was realized with three particular objectives: 1) Quantification of the effects of the increased spatial resolution for the estimation of European forest cover, 2) Evaluation of forest cover classification methods at selected locations in Europe within five bio-geographical regions of Europe using data equivalent to PROBA-V, and 3) Use of data to evaluate their potential in mapping impacts of disastrous sudden events in forest canopies.

2. Datasets

For the purposes of this project, five study areas in Europe were selected, each one representing a different biogeographical region (N.W. Portugal – Atlantic, Greece – Mediterranean, Russia – Boreal, Bulgaria – Continental and Austria – Alpine). In each of those areas a recent event, causing massive change in forest cover had occurred in the recent past. In particular, for the Alpine region an event of storm damage was studied and in the case of the other four sites forest fires had burned large forested areas. For each of these areas the following datasets were acquired:

- **Landsat 5 TM data**: Two scenes were acquired from U.S. Geological Survey website, one before and one after the event.
- **Simulated PROBA-V data**: Vito, operating the VEGETATION instrument, has provided PROBA-V simulated data, based on the Landsat TM images above.
- **SPOT-VEGETATION data**: Two 10-day average (S10) products were downloaded from the VEGETATION website (http://free.vgt.vito.be/), containing the extent of the Landsat data above and including the date in which the Landsat data were acquired.
- **MODIS data**: The 8-day-average, 250-metre resolution product of MODIS, containing the Red and Near Infrared (NIR) bands were downloaded from NASA’s LAADS website. These data were used as a substitute for PROBA-V data, due to the small size of the simulated PROBA-V data, which led a) to results of low statistical significance and b) to problems with the stacking and georectification of the simulated data.
- **Ancillary data**: Digital Elevation Models, Corine Land Cover 2000, the JRC Forest Cover Map, data from the Land Parcel Identification System (LPIS), the Terra Norte

(equivalent Corine product for Russia) [1] and Google Earth images were used for the correction of the imagery and identification of forest cover and forest types.

3. Methodologies

For each of the three objectives a specific methodology was developed.

3.1. Quantification of the effects of the increased spatial resolution on the estimation of forest cover

The JRC Forest Cover Map 2000 was downloaded from the JRC website and all tiles were integrated in a mosaic with a complete European coverage. The dataset has a 25 meter resolution and discriminates between forested (stocked) and non-forested (non-stocked) land areas. The map was reclassified to three classes, Forest, Non-Forest EU land and other (including sea and non-EU land). The map was then aggregated to 300, 600 and 1000 meter resolution, and each pixel of the aggregated version was assigned a value equal to the sum of the values of the 25 meter pixels contained within. This allowed the maintenance of the information of the number of Forest, Non-Forest and Other 25 meter pixels that were included in the aggregated pixel. The process was repeated an additional three times for each simulated resolution, once with shifting of half an aggregated pixel towards the X axis, once towards the Y axis and once on both the X and Y axes. This was performed in order to assess whether the geo-location coincidence of the aggregated 300, 600 and 1000 raster grids with the 25 meter grid affected the estimated forest cover.

3.2. Evaluation of forest cover and forest type classification methods

Four classification algorithms/methods were evaluated, in terms of estimating forest cover and forest type mapping, on each of the five biogeographical regions. These were the pixel-based Maximum Likelihood, Support Vector Machines and Artificial Neural Networks, and an object-based method, using image segmentation and subsequent classification of the objects using the Nearest Neighbour algorithm. All satellite data used for forest cover and forest type mapping refer to the images acquired before the occurrence of the event. Initially, the Landsat TM data were classified in “stocked” (more than 30% forest cover) and “non-stocked” areas. These classifications served as reference layer for the classification of the coarse resolution images. A total of 400 points were selected to be used for the training of the classification algorithm and the classification results were evaluated using an additional 300 validation points. All points were selected using stratified random sampling and each point was in the middle of a 5 by 5 pixel matrix representing a homogenous “stocked” or “non-stocked” area. The training and validation points were further segregated in three forest type classes: coniferous, boreadleaved and mixed forests.

3.2.1. Forest cover mapping

The 30-metre Landsat classification products were aggregated to 250 (MODIS), 300 (PROBA-V) and 1000 (VEGETATION) metres, assigning the percentage of “stocked” area within each aggregated pixel and classifying them in five classes using six break points: 0, 10, 30, 50, 75 and 100%. A total of 400 training points and additional 300 validation points were generated through stratified random sampling for the classification and evaluation of the coarse resolution images, using the aggregated reference layer for the identification of the strata and the forest cover class represented. The four classification methods were trained and applied, using the 400 training points, on the MODIS, simulated PROBA-V and VEGETATION data and the classification products were evaluated through a confusion matrix, using the 300 validation points. The first two classes representing areas with up to 10 and up to 30 % forest cover were merged into one "non-stocked" class and the remaining three were merged into one "stocked" class.

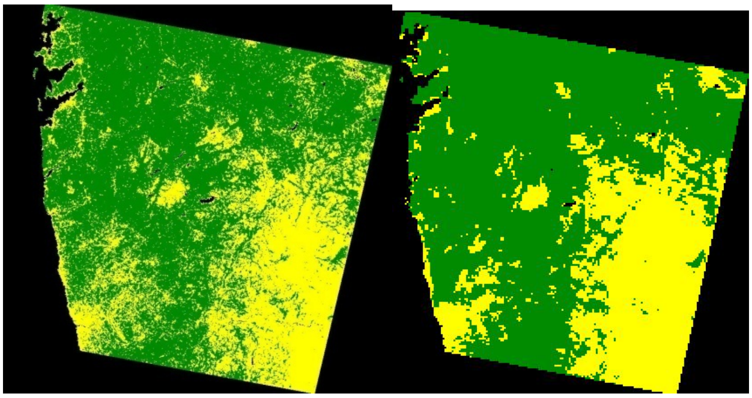

Figure 1: Forest cover maps from MODIS (left) and VEGETATION (right) imagery, using Support Vector Machine (green is "stocked" and yellow is "non-stocked" areas).

3.2.2. Forest type mapping

Using a similar method with forest cover mapping, the training and validation points representing "stocked" areas, which were used for forest cover mapping, were renamed to "coniferous", "broeadleaf" and "mixed forest", using the Corine Land Cover for the four European sites and the Terra Norte for the Russian site as reference. These points were evaluated to ensure that they represented a 3 by 3 homogenous forest type, replacing those that did not meet this criterion with additional points. Based on those training points, MODIS, simulated PROBA-V and VEGETATION data were classified and the products were evaluated through a confusion matrix, based on the renamed validation points. The evaluation of the classification methods for forest type mapping was not conducted for the Mediterranean site, since the Corine 2006 version was used as a reference and the dataset did not cover Greece for that period.

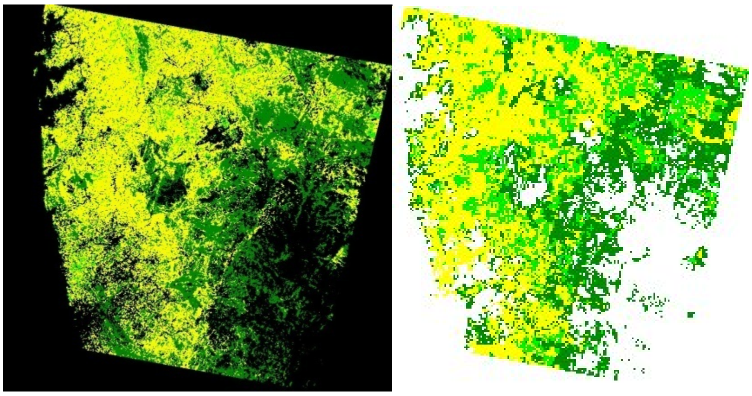

Figure 2: Forest type maps from MODIS (left) and VEGETATION (right) imagery, using Support Vector Machine (dark green is "coniferous", light green is "broadleaf" and yellow is "mixed forest").

3.3. Evaluating potential in mapping impacts of disastrous sudden events in forest canopies

In order to detect the changes from "stocked" to "non-stocked" class between the images acquired before and after the disastrous event, the Landsat TM images acquired after the event were classified and a post-classification comparison between the pre-event and post-event classifications produced the change detection reference layer. For the case of the Alpine region, the accuracy was quite high (91%), but for the remaining four regions, it was not satisfactory. In order to derive the change detection reference layer for those regions, both Landsat TM images were combined in one 12-band data set (omitting the thermal band). Principal Components Analysis was applied in those datasets, producing one or two components, which evidently portrayed the burned areas.

The training and validation points used in the classification and evaluation of the coarse resolution images in the previous classifications were also used in this task. These points were renamed based on the post-event images, with the assistance of the PCA-derived layer indicating the changes. The post-event images of MODIS and VEGETATION were classified using the algorithm that produced the most accurate forest cover classification, when the pre-event images were classified. The validation points which indicated "stocked" areas in the pre-event image and "non-stocked" in the post-event image, were assigned to a third "changed" class. A post-classification comparison of the two classifications identified the areas that changed from "stocked" to "non-stocked" due to the sudden event and were assigned to the "changed" class. The accuracy of identifying these changes was conducted using the confusion matrix deriving from the use of the validation points.

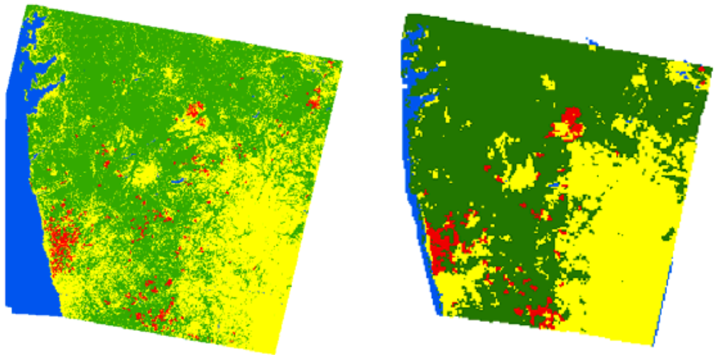

Figure 3: Change detection in MODIS (left) and VEGETATION (right) imagery, based on post-classification comparison using Support Vector Machine (light/dark green is "stocked", yellow is "non-stocked" and red is "changed").

4. Results and discussion

4.1. Results on estimated forest cover with various spatial resolutions

The results of the estimated forest cover using the JRC Forest Cover Map in four different resolutions is shown in the table below.

Table 1. Estimated forest cover for each resolution.

Resolution (m)	Average Forest Cover (%)
25	25.31
300	26.80
600	34.78
1000	36.24

The results show that the estimates at a 300 metre resolution are very close to the original 25 metre resolution JRC Forest Cover map. In contrast an increase to 600 and 1000 metre resolution overestimates the forest cover by 8 and 10% to 34.78 and 36.24% respectively. This indicates that mapping at 300 metre resolution, which is the proposed resolution of PROBA-V will be very comparable to that produced by high resolution satellites.

4.2. Results on forest cover mapping method evaluation

For the forest cover mapping accuracy on the five biogeographical regions, using the four classification algorithms/methods, the results are shown in the following table.

Table 2. Highest overall classification accuracies and the classification algorithm producing those, for each site and sensor.

	MODIS	PROBA-V	VEGETATION
Alpine	86.7 %(ML)	N/A	79.7 %(SVM)
Atlantic	83.39 %(ML/ANN/SVM)*	74.75 %(SVM)	75.79 %(SVM)
Boreal	93.6 %(SVM)	N/A	82 %(SVM)
Continental	85.71 %(SVM)	80 % (ML)	81.14 %(SVM)
Mediterranean	89.87 %(ANN/SVM/OBIA)*	71.43 %(SVM)	86.87 %(ANN/SVM)*

* All of these classification methods produced classifications with very similar accuracies.

Classification accuracies for forest cover mapping using the simulated PROBA-V data were lower than both MODIS and VEGETATION data, mainly due to the problems in georectification of the individual bands of PROBA-V. For the Alpine and Boreal region, data could not be processed and no classifications were performed for those sites. On the other hand, MODIS data produced classifications with the highest accuracies for all five biogeographical regions. Overall MODIS classifications were consistently more accurate by 3-10%. In terms of classification methods, the SVM algorithm provided the highest accuracies in most cases; however, the differences in performance were very marginal. In addition to those classifiers, the Multinomial Regression (MNR) method was applied at the Alpine and Boreal site. For the Alpine and Boreal sites, the MNR produced the best results on VEGETATION data with 81% and 83.3 % respectively.

4.3. Results on forest type mapping method evaluation

In terms of forest type mapping, the SVM classifier appears to be consistently the best performing method for forest type classification, regardless of sensor and biogeographical region. The MNR method was additionally evaluated with the Alpine and Boreal site data and produced higher accuracies than SVM for the Alpine region (72.7% and 63.5% for MODIS and VEGETATION respectively) but lower accuracies for the Boreal region.

Table 3. The highest overall classification accuracies for forest type mapping and the classification algorithm producing those, for each site and sensor.

	MODIS	VEGETATION
Alpine	70.7 %(SVM)	62.8 %(SVM)
Atlantic	75.33 %(SVM)	69.67 %(SVM)
Boreal	81.3 %(SVM)	77 % (SVM)
Continental	73.3 %(SVM)	68.3 % (SVM)

4.4. Results on comparison of change detection capability between medium and low spatial resolutions

The accuracies of detecting changes of forest cover for each sensor and each region are presented in Table 4. The results on detecting changes of forest cover consistently show that MODIS data (despite employing only two bands) are more accurate than VEGETATION data. This is true for both the accuracy of the non-changing classes (stocked and non-stocked) as well as the areas where change has occurred (from stocked to non-stocked).

It is evident that the higher resolution of MODIS data allows for a more detailed identification of the areas where sudden changes occur, as is the case with forest fires and storm damage. This suggests that the higher resolution of PROBA-V will similarly allow for such accurate mapping, in comparison with the VEGETATION sensor. The use of the SWIR channel (which was now lacking from the MODIS dataset employed here) could further support the discrimination between forested and non-forested areas.

Table 4. Change detection accuracies for MODIS and VEGETATION data for each region.

		Non-stocked Accuracy (%)	Stocked Accuracy (%)	Change Accuracy (%)	Overall Accuracy (%)	Kappa coefficient
Alpine	MODIS	95.7	80.0	60.0	86.9	0.75
	VGT	N/A	N/A	N/A	N/A	N/A
Atlantic	MODIS	86.6	88.4	53.8	86.0	0.74
	VGT	82.1	71.7	33.3	73.7	0.49
Boreal	MODIS	94.3	81.0	80.0	88.7	0.79
	VGT	92.0	61.0	80.0	80.7	0.64
Continental	MODIS	80.6	80.7	68.1	77.7	0.59
	VGT	73.8	74.1	20.0	70.3	0.46
Mediterranean	MODIS	89.4	88.8	57.1	86.7	0.75
	VGT	89.7	87.8	10.8	77.3	0.59

5. Conclusions

The main aim of this study was to investigate the capacity of PROBA-V sensor in delivering more spatially detailed and accurate forest monitoring and detecting changes in the forest cover caused by sudden events, such as forest fires and storm damage.

The initial theoretical estimation of the European forest cover, with theoretical spatial resolutions of 300, 600 and 1000 metres, showed that the estimated forest cover percentage at 300 metre resolution was very close to the one estimated by the JRC Forest Cover map with a 25 metre resolution. On the other hand, the increase of the resolution to 600 and 1000 metres, significantly overestimated the percentage of forest cover. These results are in accordance with the literature stating that overestimation of dominant land-cover is higher in the highly heterogeneous landscape mosaic with complex shapes of patches [2], [3], [4], [5].

The fact that the simulated PROBA-V data could not be processed in order to be involved in the classification processes, led to the use of 250m MODIS data, using a red and NIR bands, which has been used for forest mapping and estimating of forest cover loss in the past [6], [7]. Initial results from the stocked/non-stocked classifications of data showed that all classification methods show very comparable classification accuracies, for each sensor. As a result, no classification method can be suggested as optimum for forest cover classification. On the other hand, classifications produced using MODIS were consistently more accurate than those produced with the VEGETATION, despite the fact that VEGETATION data employed more spectral bands. It is anticipated that the availability of four bands in PROBA-V, similar to the ones available in VEGETATION, will allow for even higher mapping accuracies. In terms of forest type mapping, the SVM classification procedure provided consistently the most accurate classifications, across all sites and for both sensors. Again, the higher spatial resolution of MODIS, which is similar to the one offered by PROBA-V, supported the production of more accurate forest type classifications, in comparison with VEGETATION.

Regarding change detection, MODIS also appears to be more suitable for accurate identification of forest cover change, with the most evident differences appearing in the fragmented Atlantic, Continental and Mediterranean regions. Post-classification comparison using VEGETATION data resulted in many pixels identified as changes from stocked to non-stocked erroneously. Closer inspection revealed that in the majority of the cases, this was due to the presence of both stocked and non-stocked areas within the extent of those pixels, which lead to their misclassification as "non-stocked" during the classification of the post-fire image. Furthermore, the small size of the storm damage patches in the Alpine region did not allow their identification with the spatial resolution of VEGETATION, while MODIS showed moderate potential in detecting such changes. Even though small patches of burned forest areas are usually smaller than the pixel size of MODIS, the sensor is still capable of identifying a fire event within that pixel [8], [9], [10]. The performance of MODIS data in this application is attributed to the higher spatial resolution, in comparison to VEGETATION, which leads to the conclusion that PROBA-V will be capable of delivering classification products of equal, if not higher, accuracy.

The project employed MODIS data to indirectly assess the higher spatial resolution of PROBA-V, in comparison to the one offered by VEGETATION data. Both VEGETATION and PROBA-V are suitable for forest mapping, but on different scales. The increased spatial resolution of PROBA-V highlights the sensor as more suitable for forest cover and forest type mapping of fragmented forest landscapes of Southern Europe. In terms of forest monitoring, the maintenance of the high temporal resolution of near-daily global coverage for PROBA-V allows the generation of time series datasets and the identification of forest cover changes and forest type mapping at a finer scale, in comparison to VEGETATION.

The advent of the Sentinel-2 series of satellites and the generation of products with similar characteristics to those offered by PROBA-V, will further increase the temporal coverage at a global level and provide a higher capacity of estimating forest parameters at a regional level, with increased spatial detail, in comparison to the current capacities offered by SPOT-VEGETATION data.

Acknowledgements

We would like to thank BELSPO for funding this project, Vito for providing the simulated PROBA-V data, USGS and NASA for making available the Landsat TM and MODIS data respectively and the PROBA-V International Users Committee for their valuable feedback through the course of the project.

References

[1] Bartalev, S., Egorov, V., Loupian, E. and Uvarov, I., 2007. Multi-year circumpolar assessment of the area burnt in boreal ecosystems using SPOT-VEGETATION. *International Journal of Remote Sensing*, 28(6), pp. 1397-1404.

[2] Mayaux, P. and Lambin, E.,1995. Estimation of tropical forest area from coarse spatial resolution data: a two-step correction function for proportional errors due to spatial aggregation. *Remote Sensing of Environment*, 53(1), pp .1-15.

[3] Moody, A. and Woodcock, C., 1994. Scale-dependent errors in the estimation of land-cover proportions: Implications for global land-cover datasets, *Photogrammetric Engineering and Remote Sensing*, 60(5), pp. 585-596.

[4] Hay, G., Niernann, K. and Goodenough, D., 1997.Spatial thresholds, image-objects, and upscaling: a multiscale evaluation, *Remote Sensing of Environment*, 62(1), pp. 1-19.

[5] Lioubimtseva, E., 2003. An evaluation of VEGETATION-1 imagery for broad-scale landscape mapping of Russia: effects of resolution on landscape pattern, *Landscape and Urban Planning*, 65(4), pp. 187-200.

[6] Hansen, M., Roy, D., Lindquist, E., Adusei, B., Justice, C. and Altstatt, A., 2008.A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin, *Remote Sensing of Environment*, 112(5), pp. 2495-2513.

[7] Potapov, P., Hansen, M. C., Stehman, S. V., Loveland, T. R. and Pittman, K., 2008.Combining MODIS and Landsat imagery to estimate and map boreal forest cover loss. Remote Sensing of Environment, 112(9), 3708-3719.

[8] Achard, F., Eva, H. D., Stibig, H.-J., Mayaux, P., Gallego, J., Richards, T. & Malingreau, J-P., 2002. Determination of Deforestation Rates of the World's Humid Tropical Forests, Science, 297, 999-1002.

[9] Zhan, X., Sohlberg, R. A., Townshend, J. R. G., DiMiceli, C., Carroll, M. L., Eastman, J. C., Hansen, M. C. and DeFries, R. S., 2002. Detection of land cover changes using MODIS 250 m data. *Remote Sensing of Environment*, 83, pp. 336– 350.

[10] Jin, S. and Sader, S., 2005. MODIS time-series imagery for forest disturbance detection and quantification of patch size effects, *Remote Sensing of Environment*, 99(4), pp. 462-470.