



Monitoring land-cover changes by combining a detection step with a classification step

Item Type	Conference Paper
Authors	Harrou, Fouzi;Zerrouki, Nabil;Sun, Ying;Hocini, Lotfi
Citation	Harrou F, Zerrouki N, Sun Y, Hocini L (2018) Monitoring land-cover changes by combining a detection step with a classification step. 2018 IEEE Symposium Series on Computational Intelligence (SSCI). Available: http://dx.doi.org/10.1109/SSCI.2018.8628774 .
Eprint version	Post-print
DOI	10.1109/SSCI.2018.8628774
Publisher	Institute of Electrical and Electronics Engineers (IEEE)
Journal	2018 IEEE Symposium Series on Computational Intelligence (SSCI)
Rights	Archived with thanks to 2018 IEEE Symposium Series on Computational Intelligence (SSCI)
Download date	2024-03-29 12:44:01
Link to Item	http://hdl.handle.net/10754/631693

Monitoring land-cover changes by combining a detection step with a classification step

Fouzi Harrou¹, *Member IEEE*, Nabil Zerrouki^{2,3}, Ying Sun¹, Lotfi Hocini^{3,4}

¹King Abdullah University of Science and Technology,

CEMSE Division, Thuwal, 23955-6900, Saudi Arabia Email: fouzi.harrou@kaust.edu.sa

²University of Sciences and Technology Houari Boumédiène, LCPTS

Faculty of Electronics and Computer Science, Algiers, Algeria

³ DIIM Laboratory, Center for Development of Advanced Technology (CDTA),
Baba-Hassen, Algiers, Algeria.

⁴ Mouloud Mammeri University (UMMTO), Departement of Electronics,

Faculty of Electronic Engineering and Computer Science,

Laboratory of Analysis and Modeling of the Random Phenomena, BP 17 RP, Tizi-Ouzou, Algeria.

Abstract— An approach merging the Hotelling T^2 control scheme with weighted random forest classifier is proposed and used in the context of detecting land cover changes via remote sensing and radiometric measurements. Hotelling T^2 procedure is introduced to identify features corresponding to changed areas. However, T^2 scheme is not able to separate real from false changes. To tackle this limitation, the weighted random forest algorithm, which is an efficient classification technique for unbalanced problems, has been successfully applied on features of the detected pixels to recognize the type of change. The performance of the algorithm is evaluated using SZTAKI AirChange benchmark data, results show that the proposed detection scheme succeeds to appropriately identify changes to land cover. Also, we compared the proposed approach to that of the conventional algorithms (i.e., neural network, random forest, support vector machine and k -nearest neighbors) and found improved performance.

Index Terms—Land cover change detection, multi-spectral sensors, multi-date measurements; remote sensing, multivariate statistical approach, Random Forest classification.

I. INTRODUCTION

With its development over few years, remote sensing is increasingly used for environmental protection, monitoring, security and soil mapping. Compared to conventional mapping techniques, like photo-interpretation, spatial remote sensing (passive/ active) has incontestable advantages in terms of cost, speed, and the covered area [1], [2]. In addition to soil mapping, land cover change detection (LCCD) allows an automatic measure of change variations in a specific region. In fact, LCCD can be used for various purposes including monitoring and management of pollution [3], desertification [4], and deforestation [5]. Recently, several LCCD techniques have been proposed in the literature [6]–[9]. In [6], Zhu et al. proposed a change detection and classification algorithm based on threshold derived from all spectral bands of the Landsat data. In such approach, pixels are assigned to the changed class when the difference between observed and predicted images exceeds a threshold value. Fu et al. proposed a change detection system for urban growing monitoring using thermal

infrared data on board Landsat sensors [7]. In [8], Mayes et al. introduced a linear spectral mixture analysis approach with Landsat data applied for forest changes assessment. In [9], Espinoza-Molina et al. used local features extracted from spectral indices using a clustering algorithm. The LCCD is then performed by counting the frequency of the assignments of a changed pixel to a specific class along the time series. It is worth understanding that numerous change detection systems have been proposed in the literature. However, several proposed systems in the literature recognize false changes as changed regions. Indeed, many phenomena such as weather and vegetation growth can lead to results with false changes. Furthermore, many proposed systems do not consider that change detection is an imbalanced problem, where changed class constitutes only a reduced number of samples compared to unchanged class

An effective methodology to monitor land cover changes using radiometric features is presented. The intensity of radiometric data is obtained from passive sensors (images acquired in visible wavelength). It should be noticed that LCCD based on remote sensing data remains a challenging problematic where there are different phenomena like shadows of clouds which can change values of radiometric data and consequently deliver poor or misleading results. Therefore, it is crucial to design an effective approach to appropriately identify land-cover changes. Towards this end, we propose a framework merging the benefits of the Hotelling T^2 monitoring test and those of the weighted random forest algorithm (WRF) classifier to suitably detect and identify land cover changes. The advantage of the T^2 scheme is its ability to monitor multivariate data simultaneously and its implementation is an easy task [10]. Unfortunately, real and false (reversible) changes are not separable by using the T^2 scheme. This means that these two types of changes may have relatively similar features when the T^2 is used. WRF algorithm, which is an efficient classifier to deal with imbalanced data classification, is used to discover the dissimilarity between real changes and

false changes. Specifically, only features of selected pixels during detection phase via T^2 scheme are used by WRF. In other words, the role of detection phase is to remove features without change, hence the size of the training dataset in WRF is reduced. Data from SZTAKI Air Change benchmark is used to evaluate the proposed strategy. The results indicate that the integrated T^2 -WRF strategy is found to be suitable in detecting real changes with a reduced error rate. Section II presents radiometric feature extraction. T^2 -based LCCD and WRF classification are described in Sections III and IV respectively. In Section V we assess the efficiency of the developed approach, and conclusions are presented in Section VI.

II. FEATURES EXTRACTION FOR LCCD

Extracting relevant features from aerial or remote sensing images is not an obvious task due to the presence of regional characteristics like build regions and large areas with vegetation. In traditional change detection technologies, feature generation is accomplished via a software, like ENVI or eCognition2. The key idea behind the use of radiometric values as an indicator to identify changed regions is to associate to the same pixel at two time instants t_1 and t_2 a sequence of radiometric values (extracted from three spectral bands).

With the presence of potential changes in remote sensing measurements, the main target is to classify real and false changes. To do that, we propose the use of two stages procedure namely detection and classification of changes. On one hand, the detection phase consists of the application of the T^2 technique on radiometric values to separate unchanged from changed regions. On the other hand, the WRF-based classification focuses only on pixels previously detected by T^2 to separate real changes from false changes. The main advantage of using two-stages-based approach (detection followed by classification) is to reduce the number of training samples used by the WRF algorithm compared to classification without a detection phase. Unchanged pixels, which are identified by the T^2 scheme, have been excluded from the classification stage.

III. HOTELLING T^2 MONITORING SCHEME

Consider a data matrix $\mathbf{X} \in R^{n \times m}$ with n measurements and m process variables, $\mathbf{X} = [\mathbf{x}_1^T, \dots, \mathbf{x}_n^T]^T$. Assume \mathbf{x}_t , $t = 1, 2, \dots, n$ is normally distributed with mean μ and covariance matrix Σ . A multivariate Shewhart chart, also known as a T^2 chart or a χ^2 chart [11], [12], to monitor the process mean is based on the decision statistic:

$$T_t^2 = \left[(\mathbf{x}_t - \mu)^T \Sigma^{-1} (\mathbf{x}_t - \mu) \right] \quad (1)$$

where \mathbf{x}_t is a vector of m variables, μ is a vector of in-control means of each variable, and Σ is the covariance matrix of \mathbf{X} . This chart is the most common multivariate process monitoring technique used to monitor the mean of a normally distributed process. The upper limit for this control chart is: $UCL = \chi_{\alpha, m}^2$. We claim that the process is abnormal when the T_t^2 overpass the threshold value, $\chi_{\alpha, m}^2$ [13].

IV. LAND-COVER CLASSIFICATION PHASE VIA WRF

Land-cover classification aims at separating real from false (reversible) changes for the detected pixels. The use of the detection phase permits to reduce the number of pixels to be classified, where unchanged pixels are excluded from the classification phase. The key concept in classification is to define the model's parameters in the training phase. After the model is determined, then it will be employed to classify new data. Here, we choose the weighted random forest procedure as classifier because of its capability to operate efficiently with large and imbalanced datasets. Furthermore, it can handle an important number of features without overfitting (by its capacity to manage an important number of trees).

Compared to the classical random forest algorithm which tends to be biased towards the majority class (unchanged pixels), WRF algorithm attributes a weight to each class individually, with the minority class given larger weight (i.e., higher misclassification cost for changed pixels).

It should be pointed out that Random forest classifier can reach a high accuracy than many conventional classification algorithms, where it minimizes the overall classification error rate and has the capacity to merge many classifiers [14]. However, in the presence of imbalanced data RF ability is slightly weakened, it is because that each classifier has the same weight before combination (voting phase) [15]. Hence, to address this problem, a methodology of merging classification algorithm with different weights has been developed. In WRF approach, each individual classifier is pondered by a weight (representing the individual confidence of its decisions). Class weights are used to permit a robust separation between the majority and the minority classes.

The key concept of the WRF classifier is to compute misclassification rate for each tree Out-Of-Bag (OOB), and the class prediction for WRF is then given by combining the weighted vote from each individual tree. Instead algorithms based on error minimization like Support Vector Machine or neural network, Random Forest has an effective methodology for estimating classes when proportion of the data is missing (unbalanced problems).

Figure 1 illustrates the main principles of the weighted random forest. Initially, the algorithm assigns a weight to each class, where the minority class attributed higher weight. The class weights, $CW(j)$, are introduced in two places of the algorithm namely (i) in the tree induction procedure to find tree's splits, (ii) and in the terminal nodes, $Fn(i)$, of each tree whose impact the guidance of the voting process.

Initially, a set of new training subsets $S(i)$ are generated randomly from original training set using bootstrap sampling technique. The same strategy is applied to each subset (bootstrapping), where only \sqrt{M} features are picked from the extracted features (M represents the total number of features). The new generated features are then predicted separately using different trees for each subsample. The remaining samples after each bootstrap sampling are used for tree's evaluation procedure. Only tree corresponding to the best accuracy is selected. In fact, the class weights are considered in the tree selection

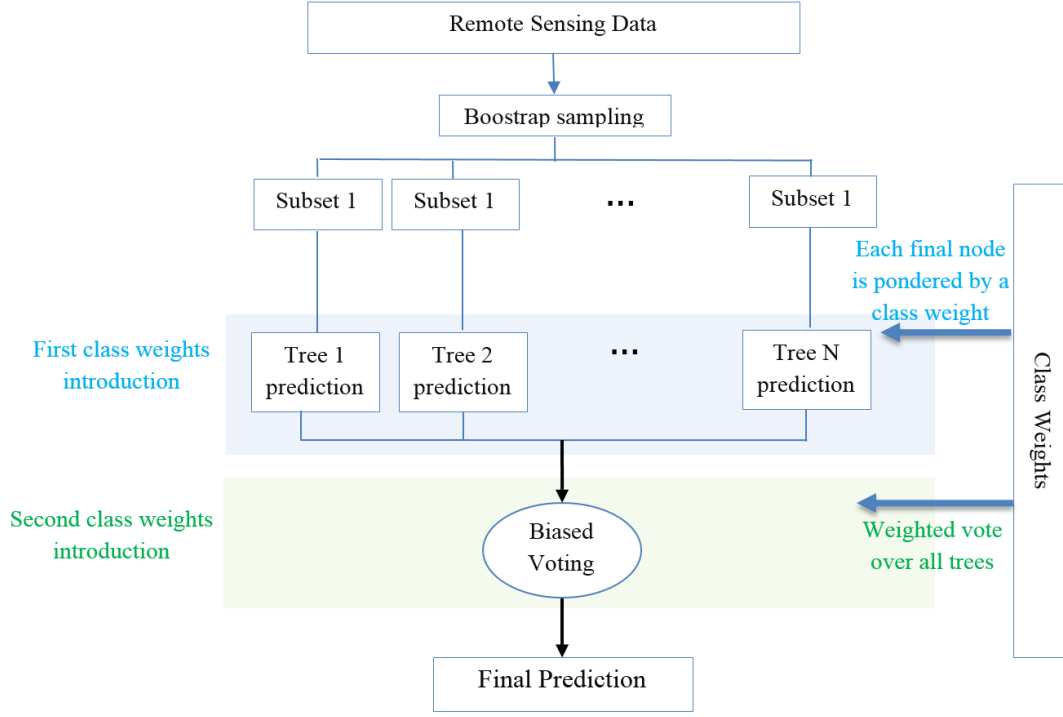


Fig. 1. A diagrammatic representation of the WRF classifier.

during separate prediction phase. The next step consists of aggregating the results of all the trees. This procedure is based on weighted voting procedure $V(j)$ instead of the majority voting to allow an adequate separation between the majority and the minority classes.

Algorithm 1: Pseudo-Code of WRF algorithm

Input: X Original data (feature vector)

- Initializing class weights $CW(j)$, $j \in [1 \dots K]$; where K =Number of classes

For $i \in [1 \dots N]$, where N is the number of trees

- Bootstrap sample $S(i)$ from original data
- Construct tree $Tr(i)$ for $S(i)$
- Assign class weight $CW(k)$ to $Fn(i)$ if $Y(i) = k$ and $k \in [1 \dots K]$ where $Y(i)$ is the predicted class label by the tree $Tr(i)$

End for

For $j \in [1 \dots K]$

- $V(j) = \sum_{i=1}^N \frac{Fn(i)}{\sum_{k=1}^K CW(k)}$ if $Y(i) = k$

End for

- $Y_{Final} = \text{Argmax}_j(V(j))$

Output: Y_{Final} (Class label)

V. RESULTS AND DISCUSSION

A. Data collection

This part is dedicated to an evaluation of the T^2 -WRF approach in detecting LCCs by using real multi-date-images. Data from SZTAKI AirChange Benchmark set [16] has been used to investigate the capacity of the proposed approach. This dataset is gathered by the Hungarian Institute of Geodesy Cartography and Remote Sensing (HIGCRS), and consisting of several optic aerial image pairs with a resolution of 1.5 meters of 952×640 pixels [16]. The multi-date images are associated with the ground of truth masks. They are acquired

with different time intervals to include different changes such as deforestation assessment and urban growing planning. Figure 2(a-c) depicts an example of the image pairs at two time instant ($t1$ and $t2$) the corresponding binary mask, highlighting changed areas with white dots.



Fig. 2. Sample of an image at $t1$ (a), same image at $t2$ (b), and the ground truth image (c).

SZTAKI AirChange Benchmark, which is rich data, covers the whole thematic classes and contains different regional physiognomies like vegetation (agricultural areas known by homogenous regions) urban (generally cities and towns known by heterogeneous areas). This variety of data makes change detection a challenging task, in particular in the ground truth, reversible changes are labeled as unchanged pixels. To address this difficulty, three main categories of pixels are considered here: unchanged, real changes and false changes. The first class presents irreversible changes like desertification and natural phenomena (e.g., forest fires). The unchanged class contains pixels that keep the same thematic classes across the time interval, such as water areas. The third class, false change is for pixels that have undergone a change in radiometric values but not in the thematic classes such as variation in

the degree of greenery for grass areas or snowfall. This type of change is not labeled as changed because it is seasonal.

B. Evaluating LCCD Results

To evaluate the performance of T^2 monitoring scheme, we first used change-free data to compute the T^2 decision threshold. The training data consists of the variation between radiometric values in different bands collected at two different instants. The data consist of three columns corresponding to the number of bands and 500 rows (representing the number of pixels). Based on the training data, the threshold is set to be $UCL = 21.03$ with the predefined probability of false alarm of 5%. Here, the training data has been chosen from different areas including bare soil, vegetation, and urban regions, to consider the variety of land cover. To illustrate the detection capability of the T^2 approach, we used testing data of 500 pixels comprising real changes. Results of the T^2 approach is presented in Figure 3(a), which indicates the presence of possible LCC. T^2 approach based on testing data with false changes is displayed in Figure 3(b). We conclude that this multivariate approach is able to detect changes but it cannot differentiate between true and false changes.

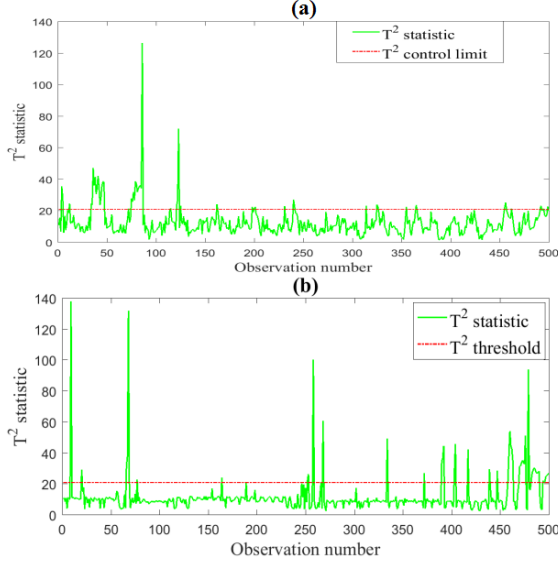


Fig. 3. T^2 approach: (a) real and (b) false changes.

The WRF classifier is used to separate real from false changes. The WRF is applied only to the features of the flagged pixels by T^2 approach. The proposed T^2 -WRF strategy is compared with some common machine learning schemes namely neural network, Random Forest, K-NN, and SVM without a detection step.

The parameters of each algorithm, which provide the highest accuracy, have been determined in the training stage. The NN classifier with one hidden layer and nine neurons has been selected. The number of neighbors in KNN is set to at 3. In SVM, radial basis function (RBF) kernel with the width of the Gaussian kernel $\sigma = 0.125$, and the parameter for the soft margin cost function $C = 128$, is used. For WRF

classifier, after testing different parameter settings, we fix the number of trees in the forest to 10 and weights to 15. Since change detection presents a case of imbalanced data, so, to accord more interest to the rare class (changed) two accuracies corresponding to changed and unchanged classes are separately computed for different algorithms.

TABLE I
COMPARISON OF THE T^2 -WRF PROCEDURE WITH CONVENTIONAL MACHINE LEARNING APPROACHES.

	Accuracy of changed class	Accuracy of unchanged class	Kappa
Neural Network	66.53	90.92	0.76
KNN	41.21	86.66	0.72
SVM	69.59	91.05	0.76
RF	51.20	93.18	0.78
WRF	85.30	95.20	0.87

Table I presents the classification accuracy of the T^2 -WRF strategy and the conventional procedures when applied to Sztaki Airchange dataset. Results in Table I point out that the proposed approach is found to be better (highest accuracy) in identifying LCCs in comparison to the conventional learning algorithm. It can be concluded that the LCC detection is improved by integrating T^2 -based detection approach with WRF classification. Also, the conventional procedures (i.e., NN, k -NN, and SVM) resulted in many misclassifications (Table I). This is because pixels with false changes are misclassified as "changed". Furthermore, some algorithms such as Support Vector Machine or neural network where classification is based on overall error minimizations (well classifying majority class and completely misclassifying the minority class) which are not suitable in such situations (imbalanced data). Another reason that the T^2 -WRF approach surpassed the other algorithms is that weighted random forest provides high accuracy over the minority class while keeping acceptable accuracy for the majority class.

VI. CONCLUSION

In this paper, we introduced an effective land-cover change detection (LCCD) technique that utilizes variation in radiometric values of aerial images. In this approach, T^2 monitoring scheme has been utilized for LCCD and the WRF classifier has been applied to distinguish the type of change. Using SZTAKI Air Change benchmark dataset, we show the efficiency of T^2 -WRF in detection land-cover changes. The results indicate that the integrated T^2 -WRF is found to be better in LCCD in comparison to other commonly used algorithms.

ACKNOWLEDGMENT

This publication is based upon work supported by the King Abdullah University of Science and Technology (KAUST) Office of Sponsored Research (OSR) under Award No: OSR-2015-CRG4-2582. The authors (Nabil Zerrouki and Lotfi H Hocini) would like to thank the DIIM laboratory, Centre de Developpement des Technologies Avancees (CDTA) for the continued support during the research.

REFERENCES

- [1] S. Jin, L. Yang, P. Danielson, C. Homer, J. Fry, and G. Xian, "A comprehensive change detection method for updating the national land cover database to circa 2011," *Remote Sensing of Environment*, vol. 132, pp. 159–175, 2013.
- [2] N. Zerrouki, F. Harrou, and Y. Sun, "Statistical monitoring of changes to land cover," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 12, pp. 1–5, 2018.
- [3] P. Zhang, Z. Lv, and W. Shi, "Local spectrum-trend similarity approach for detecting land-cover change by using SPOT-5 satellite images," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 4, pp. 738–742, 2014.
- [4] F. Ling, W. Li, Y. Du, and X. Li, "Land cover change mapping at the subpixel scale with different spatial-resolution remotely sensed imagery," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 1, pp. 182–186, 2011.
- [5] J. Inglada and G. Mercier, "A new statistical similarity measure for change detection in multitemporal sar images and its extension to multiscale change analysis," *IEEE transactions on geoscience and remote sensing*, vol. 45, no. 5, pp. 1432–1445, 2007.
- [6] 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.
- [7] P. Fu and Q. Weng, "A time series analysis of urbanization induced land use and land cover change and its impact on land surface temperature with landsat imagery," *Remote Sensing of Environment*, vol. 175, pp. 205–214, 2016.
- [8] M. T. Mayes, J. F. Mustard, and J. M. Melillo, "Forest cover change in miombo woodlands: modeling land cover of african dry tropical forests with linear spectral mixture analysis," *Remote sensing of environment*, vol. 165, pp. 203–215, 2015.
- [9] D. Espinoza-Molina, R. Bahmanyar, R. Díaz-Delgado, J. Bustamante, and M. Datcu, "Land-cover change detection using local feature descriptors extracted from spectral indices," in *Geoscience and Remote Sensing Symposium (IGARSS), 2017 IEEE International*. IEEE, 2017, pp. 1938–1941.
- [10] F. Harrou, Y. Sun, M. Madakyaru, and B. Bouyedou, "An improved multivariate chart using partial least squares with continuous ranked probability score," *IEEE Sensors Journal*, vol. 18, no. 16, pp. 6715–6726, 2018.
- [11] H. Hotelling, "Multivariate quality control illustrated by the air testing of sample bomb sights, techniques of statistical analysis, ch. ii," 1947.
- [12] M. Madakyaru, F. Harrou, and Y. Sun, "Improved data-based fault detection strategy and application to distillation columns," *Process Safety and Environmental Protection*, vol. 107, pp. 22–34, 2017.
- [13] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Journal of Educational Psychology*, vol. 24, pp. 417–441, 1933.
- [14] Q. Wu, Y. Ye, Y. Liu, and M. K. Ng, "SNP selection and classification of genome-wide SNP data using stratified sampling random forests," *IEEE transactions on nanobioscience*, vol. 11, no. 3, pp. 216–227, 2012.
- [15] B. Xia, H. Jiang, H. Liu, and D. Yi, "A novel hepatocellular carcinoma image classification method based on voting ranking random forests," *Computational and mathematical methods in medicine*, vol. 2016, 2016.
- [16] C. Benedek and T. Szirányi, "Change detection in optical aerial images by a multilayer conditional mixed markov model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 10, pp. 3416–3430, 2009.