A thresholding process on the dissimilarities between probability models for change detection on Remote Sensing data

Luiz Gustavo Rodrigues Godoy, Rogério G. Negri, Diogo de Jesus Amore

Dear Associate Editor,

We would like to thank the reviewers for the valuable comments and the time spent on checking our manuscript. We have addressed their comments towards further improving the paper contents and presentation. Below we detail the changes and updates incorporated into the manuscript in this round of review.

Yours sincerely,

The authors.

Editor's decision

Only a few issues need to be solved as the two reviewers suggested. Please revise your manuscript.

R: We have done our best to answer all comments and included the corresponding modifications in the new version of the manuscript. To facilitate the identification of new content added in the revised manuscript, we highlight in red the text inclusions/changes.

Reviewer #1

This paper presents an application of unsupervised algorithms for change detection analysis based on stochastic distance and threshold. And the author compared it with CVA method to prove it robust and high performance. However, there are several major comments in the manuscript that must be addressed before considering its publication.

R: We thank the reviewer for all his/her contributions, technical insights, and the time spent on reviewing our manuscript.

> ### Comment 1
>
> *The method applied three different types of stochastic distances, while the difference of these distances is not explained, and the reason for different performances when they are applied to remote sensing images is not analyzed.*
>
> **R:** As a consequence of its particular formulation, each stochastic distance plays a distinct way of measuring the similarity between two probability distributions. Therefore, while two objects (i.e., regions/neighborhoods) may share high similarities according to a given stochastic distance, such objects may also demonstrate elevated divergence levels concerning other stochastic distances. Consequently, we should understand that

each stochastic distance may provide a distinct representation of the spatio-temporal dynamics. In the context of the proposed change detection framework, it is essential to assess different stochastic distance alternatives and quantify the respective performances on mapping the landscape changes.

The presented response was adapted and then inserted in the Section 2.2, as showed bellow:

> As a consequence of its particular formulation, each stochastic distance plays a distinct way of measuring the similarity between two probability distributions. Therefore, while two distributions may present a certain level of similarity according to a given stochastic distance, these distributions may also demonstrate divergences concerning other stochastic distances. Consequently, admitting that distinct stochastic distances may lead to particular conclusions, the identification of the most convenient distance for a specific application is needed.

Comment 2

Fig.1 is the key block diagram illustrating the proposed approach. Maybe I missed some key points, but I want to clarify some questions. Each pixel in the neighborhood is the same important for V? And the probability density functions fu and fv are the multivariate Gaussian distributions in experiment?

R: The reviewer's interpretation is right. All the pixels/positions s into the neighborhood $\mathcal{V}_\rho(s)$ has the same importance when modeling the probability distribution functions f_U and f_V. Such modeling is made by estimating the parameters μ and Σ from $\mathcal{V}_\rho(s)$ but considering different instants, resulting in the models f_U and f_V. In the experiments, f_U and f_V are both multivariate Gaussian distributions.

The following text was added in Section 3:

> Equivalently, it may be understood that the information from $V(\mathcal{I}_1, (i,j), \rho)$ and $V(\mathcal{I}_2, (i,j), \rho)$ are employed to estimate the parameters of $f_\mathbf{U}$ and $f_\mathbf{V}$. In this process, the radiometric values concerning the elements/coordinates into the mentioned neighborhoods, regardless of their distance from the reference position (i,j), equal importance on the distributions parameters estimation.

Comment 3

The author chose the reference samples manually. And I want to ask that what is the rule and way when you take the samples because I notice that the distribution of most samples at different times are located in similar areas.

R: The ground-reference samples were collected through a careful visual inspection process regarding pairs of instants. On certain occasions, when selecting a "non-change" sample that does not change over time, such sample tends to be reconsidered in distinct pairs of instants. In the same sense, targets/regions with frequent spectral change also may be re-employed.

An additional explanation concerning the reference samples was inserted in the experiments section:

> Under certain conditions, places of "non-change" samples that do not change over time tend to be reconsidered in distinct pairs of instants. In the same sense, regions with frequent spectral transition may be re-employed as a "change" sample.

Comment 4

The author listed some literature in the introduction. I think the authors might have missed some latest and relevant papers in this area. Please comment on the following recent papers in the revised paper. [a] Change detection using deep learning approach with object-based image analysis. [b] An unsupervised domain adaptation approach for change detection and its application to deforestation mapping in tropical biomes. [c] A deep translation (gan) based change detection network for optical and sar remote sensing images

R: Thanks for the suggested papers. We included the respective citations in the introduction section.

> Recently, diverse supervised approaches have been proposed in the literature to deal with change detection on Remote Sensing data. Ref. **?** combines object-based image analysis and convolutional neural network concepts to perform change detection. Ref. **?** employs generative adversarial nets to propose a change detection method that simultaneously analyzes optical and radar data. Ref. **?** exploits the idea of "Domain Adaptation" to train a neural network using information from one source for posterior generic use on the change detection over distinct imagery datasets.

Comment 5

The experiments need to be further analyzed. For instance, the intermediate result of Id should be displayed and explained, because it make the experiment more convincing.

R: Thanks for this suggestion.
The distance ($\mathcal{I}_d$ – SDTA) and amplitude ($\mathcal{I}_a$ – CVA) images, related to the change detection maps presented in Section 4.3, were incorporated in the manuscript. Additional discussions regarding such images were also added to the text.

> As a complementary analysis, Figures 10 to 13 show the distance ($\mathcal{I}_d$) and amplitude ($\mathcal{I}_a$) images referent to the change maps presented in Figures 6 to 9, respectively. Once the values may change as a function of the input data, distance type, and neighborhood size, the depicted maps use a quantile-based color scale. Concerning both methods, it may observe that the Kittler-Illingworth technique tends to threshold the distance/amplitude values around the second and third quantile. In consonance with the change maps (Figs.6-9), the SDTA method expresses the changes through homogeneous regions. The same is not observed for the CVA method, where isolated points of high amplitude values are spread along the study area.

Reviewer #2

General Comment

The paper proposes a change detection method for Landsat 5/8 images, referred to as Stochastic Distance Thresholding Analysis (SDTA). Please consider the following issues:

R: We thank the reviewer for all his/her contributions, technical insights and the time spent on reviewing our manuscript.

Comment 1

Line 149. Known densities. Please add details, by considering both the sensor/band/resolution characteristics and the neighborhood size.

R: Further details about the densities/neighborhoods were incorporated in Section 3, as pointed below. Details concerning the adopted sensor and neighborhood size (ρ) are found in Sections 4.1 and 4.2.

> Equivalently, it may be understood that the information from $V(\mathcal{I}_1, (i,j), \rho)$ and $V(\mathcal{I}_2, (i,j), \rho)$ are employed to estimate the parameters of $f_{\mathbf{U}}$ and $f_{\mathbf{V}}$. In this process, the radiometric values concerning the elements/coordinates into the mentioned neighborhoods, regardless of their distance from the reference position (i,j), have equal importance on the distributions parameters estimation.

Comment 2

Was pansharpening applied? If it was not, add results on pansharpened MS bands.

R: The pansharpening was not applied. One of the main reasons is the absence of a panchromatic band for images acquired in 1999, 2003, and 2008 by the TM sensor. However, according to the reviewer's comment, the conclusion section mentions the importance of future studies and analysis with higher resolution images.

Comment 3

Line 208: GitHub link is not working.

R: In fact, the repository has been created, but we have not yet provided the code. The codes will be made available immediately after the manuscript acceptance/publishing. The main reason to make it unavailable is due to possible adjusts that may be necessary through the revision process.

Comment 4 and 5

Please better explain the anomalous behavior in Fig.5b.
Fig.10b shows poor performance for low FPR with respect to CVA. Justify that from a methodological point of view or with considerations on the specificity of the 2003-2008 scenario.

R: For the sake of simplicity, we discussed comments 4 and 5 together. We included the following text in the "results and discussion" section.

> The behavior observed in Figure **?** is a consequence of the spectral behavior of the "2003 image" (Fig. **??**). External factors (like atmospheric,

climatic, or even seasonality) may impair the target distinction and the measured stochastic distance/amplitude values. It is essential to attend that the "2003 image" affects both analyzed methods (SDTA and CVA) concerning 1999–2003 and 2003–2008 comparisons in terms of results accuracy, with more influence on this latter pair. Furthermore, according to the ROC curves (Fig.14(a) and (b)), it is possible to identify that the "2003 image" demands higher False Positive Rate values (approximately 0.195 and 0.3 concerning the pairs 1999–2003 and 2003–2008, respectively) in order to deliver a True Positive Rate above 0.8. In summary, we may consider that the interference mentioned is the cause of large "blobs of changes" that overlaps the non-change reference samples and consequently drop the F1-Score values. On the other hand, since the CVA does not consider the neighborhood information, avoiding the insurgence of large "blobs of changes" yet producing very noisy results, it is still less susceptible to False Positive errors in this case. However, from a qualitative point of view, we should stress that the change maps delivered by the SDTA method are superior to those from CVA.