Fusion of Sentinel-2 and Proba-V/Sentinel-3 Images for Multi-Temporal Land-Cover Map

Jordi Inglada, Julien Osman and Tiangang Yin
CESBIO - UMR 5126, Toulouse, France; jordi.inglada@cesbio.cnes.fr

Abstract. Upcoming temporal and spatial HR satellites such as Venμs and Sentinel-2 and even the Landsat Data Continuity Mission (LDCM) will provide very valuable data for land-cover and vegetation monitoring. The 2 to 16 day revisit cycle and the 10 to 30 meter resolution is very useful. However, due to cloud cover and even to some rapid changes, a higher temporal resolution may be needed for such applications. One of the ways to improve the temporal resolution for these satellites is to merge their data with higher temporal resolution systems. For now, these other systems will fatally have a lower spatial resolution or a limited field of view. Past research works have developed fusion approaches to use the synergy between HR resolution and mid- to low-resolution images. The outcome of these works showed that the resolution ratio between the images to fuse need to be not too far apart. The increased resolution of Proba-V with respect to the Vegetation sensor seems to be well suited for this kind of applications. The goal of this work is to assess the usefulness of these techniques for the joint use of Proba-V data and Venμs/Sentinel-2/LDCM images for land-cover monitoring. As a result, one can expect an algorithm for the generation of land-cover maps and time profiles of surface reflectances with a spatial resolution of 10 to 30 m. with an update frequency of about 10 days. Particular attention is paid to the different spectral resolutions of the sensors being used.

Keywords. Proba-V, Sentinel, image fusion, land-cover maps.

1. Introduction

Upcoming temporal and spatial HR satellites such as Venμs and Sentinel-2 and even the Landsat Data Continuity Mission (LDCM) will provide very valuable data for land-cover and vegetation monitoring. The 2 to 16 day revisit cycle and the 10 to 30 meter resolution is very useful. However, due to cloud cover and even to some rapid changes, a higher temporal resolution may be needed for such applications.

One of the ways to improve the temporal resolution for these satellites is to merge their data with higher temporal resolution systems. For now, these other systems will fatally have a lower spatial resolution or a limited field of view.

Past research works have developed fusion approaches for using the synergy between HR resolution and mid- to low-resolution images. The outcome of these works showed that the resolution ratio between the images to fuse need to be not too far apart. The increased resolution of PROBA-V with respect to the Vegetation sensor seems to be well suited for this kind of applications.

The goal of this work is to assess the usefulness of these techniques for the joint use of PROBA-V data and Venμs/Sentinel-2/LDCM images for land-cover monitoring. As a result of the presented work, one can expect an algorithm for the generation of land-cover maps and time profiles of surface reflectances with a spatial resolution of 10 to 30 m., with an update frequency of about 10 days. Particular attention has to be paid to the different spectral resolutions of the sensors being used.

The main objective of this work is to propose a processing work-flow aiming at taking advantage of the very high temporal resolution of PROBA-V and the high temporal and spatial resolutions of Venµs, Sentinel-2 and other similar systems. Particular attention has been paid to the increased spatial resolution of PROBA-V with respect to the SPOT-Végétation sensor.

This general objective is described below:

1. Perform a thorough review of the literature about the available techniques for the joint use of high and mid to low resolution remote sensing images for land cover analysis.
2. Implement the most promising methods of each category and carry on a benchmark on simulated and real data.
3. Develop a fusion framework allowing us to use methods of different nature (statistical, physical) in order to benefit from the advantages of different approaches.
4. Assess the methods in the particular case of the joint use of PROBA-V data and Venµs and/or Sentinel-2 images.

Since the implementation of the afore mentioned algorithms needs a robust and modular software environment, the CNES' Open Source Software Orfeo Toolbox (OTB) has been used. As a consequence, the resulting developments were made available as a module of OTB.

2. Methods

2.1. Signal processing

2.1.1. Temporal interpolation and spatial unmixing

If we consider the high spatial resolution as starting point, the problem of increased temporal resolution poses an interpolation issue. Many approaches for temporal interpolation of time series exist, from simple linear interpolation of the target values to more sophisticated approaches using neural networks.

In the case of the joint use of LR time series with higher spatial resolution images, the LR images can be used to constrain the interpolation algorithms and therefore improve the results with respect to a classical interpolation. The way of constraining the temporal interpolation will depend on the ratio of the temporal sampling between the low and the high resolution sensors, and also on the chance of co-occurrence of the acquisitions.

For instance, in the case of several simultaneous acquisitions of the low and the high resolution sensors, one can try to derive the spectral transfer function between them (either by variational or by learning-based approaches). Also, the spatial resolution ratio between the sensors, will determine the possibility of using object-based approaches. For instance, 5 to 10 m. resolution images over agricultural areas can be segmented in order to use per field approaches instead of pixel-based ones.

2.1.2. Naive fusion

The simplest approach to the fusion of LR time series with HR imagesis is the resampling of the LR onto the geometry of the HR. Although this approach may be called *naive*, if the zoom of the LR data is done appropriately, some useful information from the LR time evolution may be incorporated to the HR images.

Of course, this approach needs a perfect knowledge of the geometric distortions between the LR and the HR data. In the simulations shown in this work, we avoid the issue by means of simulating LR data from HR data.

For the case of this simple fusion approach, we will investigate the influence of the interpolator used for the over-sampling operation. We will compare a nearest-neighbour interpolator with a bi-cubic one. We will apply this processing to NDVI values.

2.1.3. Bayesian data fusion

The problem of image pan-sharpening is similar to the one we want to address here: fusing multi-spectral images with panchromatic ones with higher spatial resolution. Most of the algorithms in the literature needs the panchromatic band to be consistent with the spectral range covered by the multi-spectral channels.

The Bayesian Data Fusion (BDF) method developed by [1] estimates the statistical link between the HR and the LR images in order to find the most probable value for a HR pixel given the LR one.

A simple sensor model for the observed pixels, $Y = g(Z) + E$, is used, where g(.) are the sensor function and E is random noise. This statistical link is represented by the conditional probability of the high resolution multi-spectral value (the fused pixel), z, when the high resolution panchromatic pixel, y_p, and the low resolution multi-spectral pixel, y_S, are known:

$$f(z \mid y_S , y_P) \sim f_Z(z) * f_{ES} * f_{EP}.$$

The authors introduce the additional possibility of giving different confidence levels to the panchromatic and the multispectral images by introducing a weight parameter. A low value for the weight gives more weight to the multispectral data.

Even though this method was developed for the fusion of synchronous LR and HR acquisitions, its usefulness for time series updating was investigated in [2].

In order to produce NDVI time series as in the 2 previous cases, we will apply the BDF to the red and near infra-red bands of the LR images and we will use a pseudo-panchromatic image produced by the sum of the red and near infra-red bands of the closest HR image in time.

Once the 2 bands are pan-sharpened, we compute the NDVI values from them.

2.1.4. STAR-FM

The Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM) successfully blends the reflectance of Landsat data (30m) with reflectance of MODIS data (500m).

The STARFM algorithm is based on the use of several high resolution (HR) and low resolution (LR) image pairs which are supposed to have been pairwise acquired at the same dates. These image pairs are used to estimate the HR reflectance (red image in the following figure) for a date where only a LR image (yellow LR1 image the figure) is available. The blending process focuses on using spectrum-similar pixels of corresponding images to predict the reflectance.

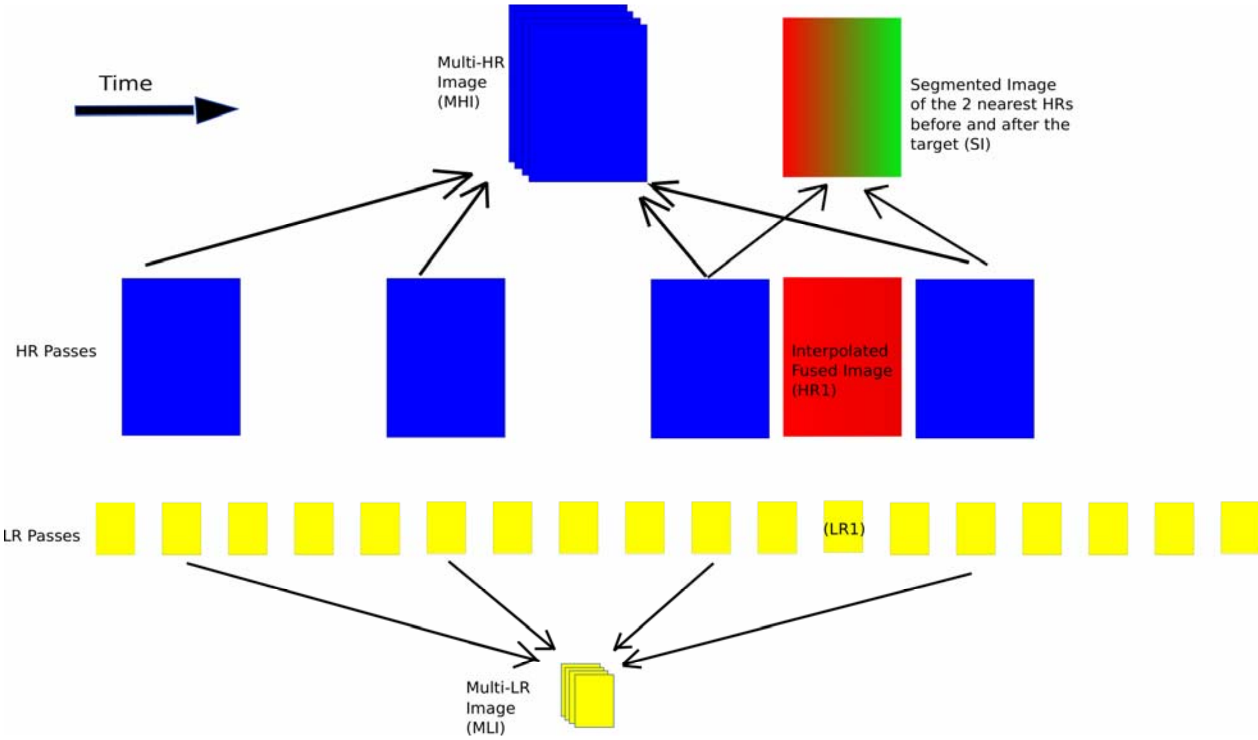

Figure 1: Illustration of STARFM input images.

Figure 1 shows an example of the work-flow used in the STARFM algorithm. In this example, we use 4 HR images and 18 LR images. The 4 HR images are combined into a multi-band image (HRI), and the LR images acquired at the corresponding times are also combined.

2.2. Model-based

2.2.1. Sigmofitting

The temporal variation in satellite derived vegetation index (VI) data for a single growth or senescence cycle can be modelled using a sum of 2 functions of the form [3]:

$$y(t) = c/(1+ e^{(a+b*t)}) +d$$

where t is the time in days, y(t) is the VI value at time t, a and b are the fitting parameters, c+d is the maximum VI value and d is the initial background value.

The algorithm has been implemented using nonlinear least squares fitting with 8 parameters to fit (4 for each sigmoid) using the Levenberg-Marquardt algorithm. Several strategies for the selection of the weights for the LR dates have been compared: constant and increasing with the distance to the HR (linear and square root).

The later seems to give the most robust results and is the one used for the algorithm benchmarks. Since cloud screening results are available in our case [4], we can use this information in order to modify the weights of every date for every pixel in the least squares fitting. The cloud screening results are available at 10 m. resolution. Our cloud screening algorithm detects different types of clouds with different reliabilities, and therefore, different weight values can be set accordingly. For the LR images, the cloud masks are used as if they were only available at the LR. We call this approach Sigmo-Cloud.

2.3. Combination of approaches

One of the goals of the work was to investigate the possibility of improving the fusion approaches by combining signal processing and model based methods.

One single model-based approach was finally available - the sigmoid fitting - with its two different versions (with and without cloud screening).

For the signal processing approaches, one possible choice was an improvement of the best algorithm, STARFM. However, this algorithm already takes into account the temporal trend of the previous and following HR images. Furthermore, the STARFM algorithm assumes that a HR image is available at the end of the time series, which can be a limitation for some applications.

On the other hand, the BDF approach does not use the temporal trend explicitly and could benefit from the combination with a phenology-based approach.

Therefore, the BDF was combined with the sigmoid fitting approach in a sequential way: first the standard BDF is applied and then the resulting time series is processed through the sigmoid fitting algorithm. The two versions of the sigmoid fitting (with and without cloud screening results) were evaluated.

3. Results

3.1. Experimental set-up

3.1.1. Data available

A very rich dataset was used for the assessment and validation tasks that were carried out in this work. The test site is located in the South-West of France, near Toulouse and covers an area of 500 km^2 presenting 24 thematic classes at the finest level of the nomenclature.

A series of 49 Formosat-2 images during the period going from February to November 2006 were available. The images have a size of 4000x4500 pixels and have been ortho-rectified with a 8 m. ground sampling distance. They have 4 spectral bands (blue, green, red and near infra-red).

A ground truth over 1764 agricultural plots was available, which adds up to about 80 km^2. These ground truth were available for most dates of the temporal series and gives the phenological state and the thematic class for each region. Only the thematic class was used for this work. We worked with a simplified nomenclature with a total of 7 classes (forest, winter crops, irrigated summer crops, non-irrigated summer crops, water, built-up and mineral surfaces).

3.1.2. Simulations

From the whole dataset, we kept only the images which had less than 20% cloud cover, which mounts to 28 images. A previous work shows that this subset gives the best performance for classification purposes [5]. This 28-image set constitutes our best scenario for comparison. This dataset will be the split in 2:

1. Five cloud-free images acquired at optimum dates during the crop season, one in March, two in June, one in July and one in September. These images are the HR time series.
2. The rest of the images (23) were used to simulate the LR time series. Using a prolate interpolator (which is a good approximation of a sensor point spread function [6], [7]), they were down-sampled to 330 m. resolution. Images at 500 m. and 1km. resolution were generated using the same approach.

The 5-image dataset can be seen as our worst case scenario, where only a few HR images are available.

Our goal was therefore to assess the usefulness of the fusion approach. In order to have a quality metric, we performed a supervised classification of the time series and used the Kappa coefficient [8] in order to rank the different methods.

The classifier algorithm used is a Support Vector Machine with a Gaussian kernel. The kernel variance and the cost parameter are set using a cross-folder optimization over 7,000 samples per class. The estimation of the Kappa coefficient is performed using 20,000 samples per class. The training and validation samples are not only different, but they come from different regions of interest, which correspond to distinct objects in the image (different fields, different forest plots, different urban areas, etc.). Fifty different runs of the supervised classification are performed for each algorithm using different random selections of pixels, which allows us to compute robust statistics on the Kappa coefficient.

3.2. Algorithm benchmark

The following 9 configurations were selected for the benchmark:

1. Naive NN
2. Naive Zoom
3. BDF
4. Sigmo: sigmoid fitting using the number of days as a weight
5. Sigmo-Cloud: as Sigmo, but the cloud screening masks are added as weights
6. Sigmo-BDF: Sigmo applied on the output of BDF
7. Sigmo-Cloud-BDF: Sigmo-Cloud applied on the output of BDF
8. STARFM: our improved version
9. HR: all high resolution images (upper bound)

We give a graphical representation of the final benchmark for which each algorithm is run 50 times over different random selections of the ground truth data. This ensures diversity of the configurations and statistical representativity. The following statistics are plotted for the Kappa values obtained:

- minimum and maximum (the ends of the segments; vertical ticks)
- 2nd and 4th quartiles (boxes)
- median (vertical line inside the boxes)

The red area gives the minimum (for all algorithms) of the median values. The green area gives the maximum (for all algorithms) of the median values. This corresponds to the use of all the HR images.

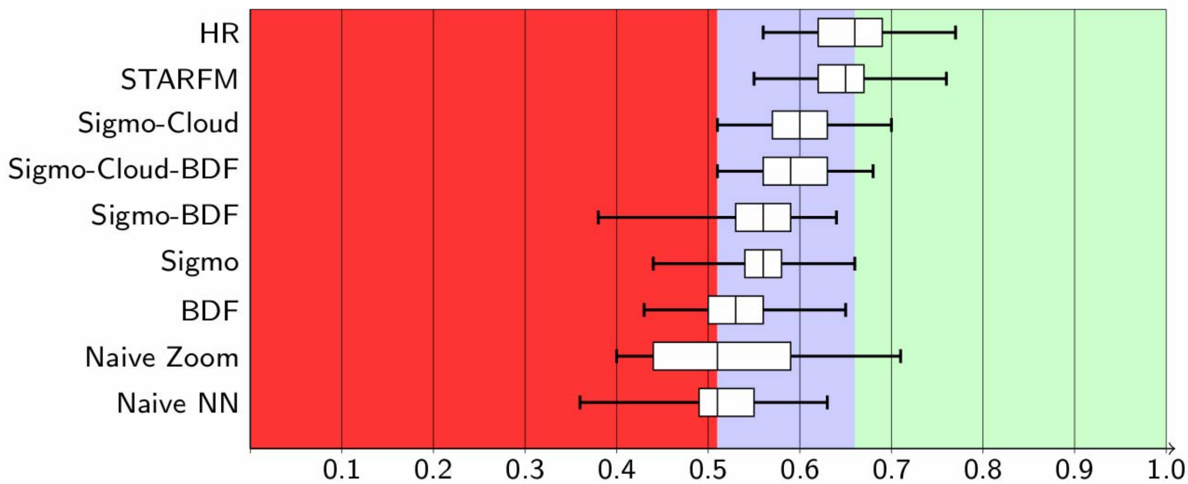

Figure 2: Comparison of the different algorithms using synthetic Sentinel2 (10m) and PROBA-V (330m) images.

Figure 2 presents the results of each algorithm for the 1/3 km. resolution for the LR images. The algorithms are ranked from top to bottom in decreasing order of the median of the Kappa values for the 50 classifications.

The following conclusions can be drawn from these results:

- STARFM may lead to classification accuracies close to those obtained with HR images (theoretical upper bound), but it needs sensors with similar spectral bands. This algorithm also assumes that there is a HR image at the end of the time series and this can be a drawback for some applications.
- Sigmoid fitting has bad performances for some classes as for instance the non vegetated classes or the vegetation classes with more than one maximum in the phenological profile.
- Bayesian data fusion has lower performances than those obtained during the first runs (smaller test areas), mainly due to textures classes as for instance urban and forests. But BDF can work on vegetation indices or reflectances.
- The combination of approaches (Sigmo+BDF) gives little improvements and is always less accurate than STARFM.
- Sigmoid fitting works well, but is designed to work on vegetation indices only and not on reflectances.
- Sigmo-Cloud: prior knowledge about cloud screening is very useful.

We have also analyzed the impact of the improved resolution of PROBA-V with respect to SPOT-Végétation. Three resolutions were compared: 1km. 1/2 km. and 1/3 km. These comparisons were carried out for the two best algorithms, i.e. STARFM and Sigmo-Cloud. Figure 3 shows the effect of the resolution of the PROBA-V images on the Kappa coefficient. One can observe that the improved resolution of PROBA-V images with respect to SPOT-Végétation allows achieving better classification results.

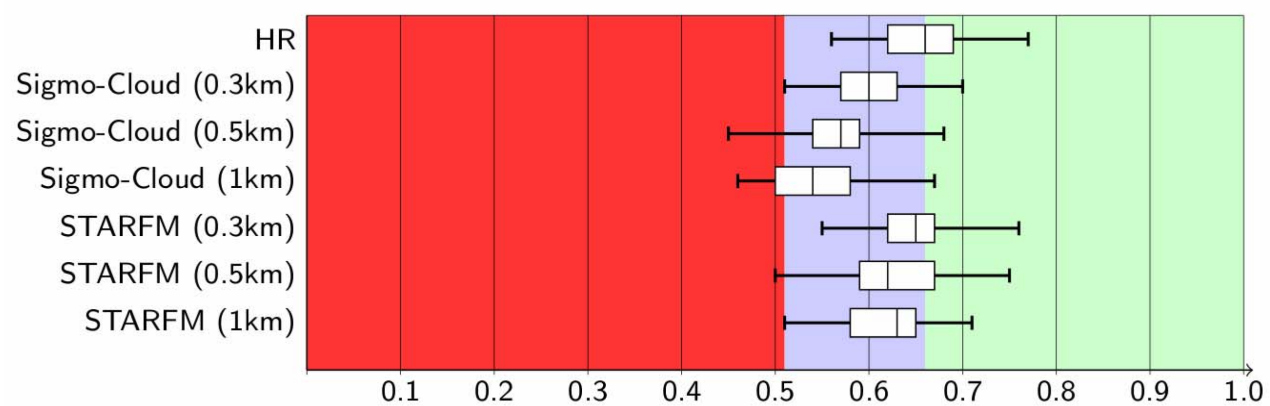

Figure 3: Comparison of the different algorithms using synthetic Sentinel2 (10m) and PROBA-V (3 resolutions) images.

4. Conclusions

In this work we have reviewed and compared methods for the joint use of Proba-V data together with high geometrical and temporal resolution sensors as Venμs and Sentinel-2. We have generated simulated datasets representative of PROBA-V over sites where other sources of imagery and ground data were available. We have also improved some of the existing methods and proposed a simple strategy for combining signal processing and rigorous models. The implementation of the selected methods and the validation of the produced software has been done using the ORFEO Toolbox and made available as an open source software package.

The benefits for the PROBA-V programme are described below:

- The demonstration of the usefulness of PROBA-V data for high geometrical resolution continental surface modelling
- A validation of the concept of PROBA-V being a precursor of Sentinel-3 by taking into account the synergy with Sentinel-2
- The assessment of the impacts of the improvement of the geometrical resolution of PROBA-V with respect to SPOT-Végétation
- The availability of open source software right from the beginning of the PROBA-V mission so that potential users can prepare the arrival of real data.
- The availability of a procedure for simulating data sets representative of both PROBA-V data and interesting vegetation related applications

Although the results of this work show that approaches like STARFM may lead to classification accuracies close to those obtained with HR images (theoretical upper bound), the following issues have to be taken into account. First of all the STARFM algorithm has 2 main limitations:

1. The need of a HR image at the end of the series
2. The need for similar spectral bands for the fusion

As a final recommendation on the choice of the algorithms, we can say that:

- for seasonal classification where all the images of the series are available, our improved version of STARFM should be used
- for on line classification (or real-time), the sigmoid fitting approach (with cloud screening if available) seems a very good candidate mainly for vegetated areas.

There are also some limitations to the representativity of the results obtained. The LR images used in this study are obtained by resolution degradation from HR images. This fact makes that:

1. The HR and LR images are perfectly registered, which may be difficult to obtain with real data
2. The spectral bands of the LR images are exactly the same that those of the HR images
3. There are no directional effects on the LR data, which is difficult to achieve with real LR sensors, since they have a very wide field of view.

The main positive conclusion of the work for the PROBA-V programme is the validation of the improvements in land-cover classification brought by the increased resolution with respect to SPOT-Végétation.

Acknowledgements

This work is funded by the Belgian Federal Science Policy in the framework of the PROBA-V Preparatory Programme under contract CB/XX/20 (PRO-Fusion Project).

References

[1] Fasbender, D., Radoux, J. and Bogaert, P., 2008. Bayesian data fusion for adaptable image pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 46(6): pp. 1847-1857.

[2] Fasbender, D., Obsomer, V., Bogaert, P. and Defourny, P., 2009. *Updating scarce high resolution images with time series of coarser images: a Bayesian data fusion solution*. In N.Milisavljević, editor, Sensor and Data Fusion, 246-261. I-Tech, Vienna (Austria).

[3] Xiaoyang, Z., Friedl, M. A., Schaaf, C. B., Strahler, A. H., Hodges, J. C. F., Gao, F., Reed, B. C. and Huete, A., 2003. Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, 84: pp. 471-475.

[4] Hagolle, O, Huc, M., Villa Pascual, D. and Dedieu, G., 2010. A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment*, 114: pp. 1747-1755.

[5] Inglada, J., Hagolle, O. and Dedieu, G., 2010. Assessment of the land cover classification accuracy of Venμs and Sentinel-2 image time series with respect to Formosat-2. *Third International Symposium on Recent Advances in Quantitative Remote Sensing}*, Torrent, València, Spain.

[6] Mathews, J., Breakall, J. and Karawas, G., 1985. The discrete prolate spheroidal filter as a digital signal processing tool. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(6): pp.1471-1478.

[7] Inglada, J., Muron, V., Pichard, D. and Feuvrier, T., 2007. Analysis of artifacts in subpixel remote sensing image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 45(1): pp. 254-264.

[8] Carletta, J., 1996. Assessing agreement on classification tasks: the kappa statistic. *Comput. Linguist.*, 22(2): pp. 249-254.