# Using data fusion to update built-up areas of the 2012 European High-Resolution Layer Imperviousness

Antoine Lefebvre[1], Nicolas Beaugendre[1], Alexandre Pennec[1], Christophe Sannier[1] and Thomas Corpetti[2]

[1] *SIRS, Villeneuve d'Ascq, France; antoine.lefebvre@sirs-fr.com*
[2] *CNRS, Rennes, France; tcorpetti@gmail.com*

**Abstract.** One of the Copernicus (ex-GMES) Initial Operations 2011–2013 services consists in updating 5 seamless European coverage of a 20m-resolution land cover. Among those, the High Resolution Layer Imperviousness requires a specific processing. This paper focuses on the update of built-up areas and more especially, on a methodology dedicated to the manipulation of massive data. The main goals are to provide a tool, which is (1) operational from one end of Europe to the other and (2) robust enough to the fluctuant data availability. The approach is based on the ESA Data WareHouse (DWH), which provides a large choice of images including IRS LISS-3, RapidEye, SPOT-5 and IRS AWiFS time series. However, even if it contains numerous sources, it remains that these latter are heterogeneous and particularly complex to process all in one: they have different spectral/spatial resolution and they have been acquired at different time, in different conditions. The originality of the approach relies on (1) a separate image classification and (2) a combination of each classification probabilities using a data fusion technique. The classification step is based on a neural network algorithm and the fusion step is performed with the Dempster-Shafer Theory. Results performed on 7 countries and present an overall accuracy around 90. It brings out the efficiency of the method to process large project such as GIO HR Layer and to benefit from the large diversity and availability of the ESA DWH products.

**Keywords.** Urban areas, Data fusion, High Resolution Layer, Copernicus, GMES, ESA Data WareHouse.

## 1. Introduction

The High Resolution Layer (HR Layer) Imperviousness consists in a 20x20 m grid covering all Corpernicus member countries and aims to represent a degree of imperviousness from 0 to 100%. This layer was already produced in 2006 and 2009 and the 2012 update remains a difficult task, urban change is a small and irregular phenomenon. In this context, service provider companies like SIRS, which are in charge of the 2012 HR layer production, need to develop innovative and efficient method to manage and produce such large dataset.

In the framework of the HR Layer project, ESA Data WareHouse (DWH) provides numerous sources of satellite images. The 2012 HR Layer imperviousness is based on at least 3 sensors: RapidEye, IRS LISS-3 and IRS AWiFS where a full coverage on the Copernicus member countries should be available. However, image acquisition is, in practice, more complex and can be delayed due to high cloud coverage. Some SPOT-5 images are then also available on the ESA DWH to fill inherent coverage gaps. It results 4 data sources, which have many different parameters (Table 1): (1) different spectral and spatial resolution, (2) different radiometric information (images were acquired at a different time period and their calibration are slightly different), (3) sensors have different path and different swath and each scenes cover a different area.

**Table 1.** Data available on ESA Data WareHouse for High Resolution Imperviousness production

| Sensor | Resolution | Swath | Bands | Availability / comments |
|---|---|---|---|---|
| RapidEye | 5x5 m | 77 km | RGB-RED-NIR | Full coverage |
| IRS LISS-3 | 20x20m | 140 km | RG-NIR-SWIR | Full coverage |
| IRS AWiFS | 56x56m | 740 km | RG-NIR-SWIR | Time series |
| SPOT-5 | 2.5x2.5m | 60 km | GB-NIR | Gap filling |

This paper aims to propose a method relying on the diversity of the DWH products and providing a homogeneous built-up mask covering all Copernicus member countries. In this context, RapidEye and SPOT-5 images are high spatial resolution images and can distinguish sharply the built-up areas. AWiFS time series are suitable to monitor phenological cycles of agricultural areas. It then reveals a useful data to delete bared soil, which could be previously detected as imperviousness areas on RapidEye or IRS LISS-3 scenes.

A previous work on the HR Layer imperviousness production has been proposed to update 2006 Soil Sealing [1]. It consisted on a change detection method, which compares bi-temporal (2006 and 2009) NDVI images. The main drawback of this approach is the strong dependency to the properties and quality of the input data. It is sensitive to the time acquisition window of the images and also to the cloud cover and haze. Moreover, it is not suitable to swap the original data source to another one.

In this paper, we suggest a generic and robust approach able to combine all DWH products. It relies on the Dempster-Shafer Theory (DST), which has already proven is efficiency on remote sensing data [2;3]. This technique relies on evidence theory, it combines multiple classifiers dealing with the imprecision and uncertainty [4].

## 2. Methods

The method can be divided in two main steps: (1) single image classification and (2) data fusion.

### *2.1. Separate image classification*

#### *2.1.1. IRS LISS-3 and RapidEye images*

A supervised classification based on neural networks algorithm was used for the IRS LISS-3 and RapidEye images. This classifier is particularly suitable when a large quantity of samples is available [5]. For each scene, calibration samples were defined randomly from the 2009 HR Layer imperviousness and split in 2 classes (built-up and non built-up areas).

#### *2.1.2. IRS AWiFS time series smoothing*

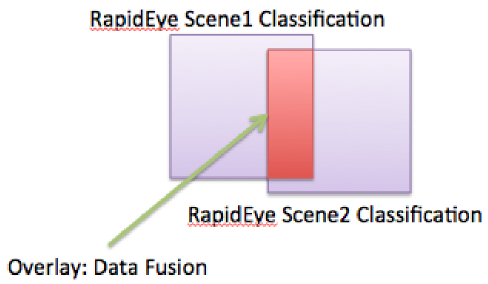
Times series are particularly noisy: (1) images have no radiometric correction and are not directly comparable from one scene to another one; (2) the cloud coverage can be really heavy or scattered. AWiFS time series were then smoothed in the time with a central moving average process [6]. The process is simple, fast on large dataset and also efficient to highlight impervious areas, which are spectrally homogeneous in the time.

#### *2.1.3. Classification uncertainty*

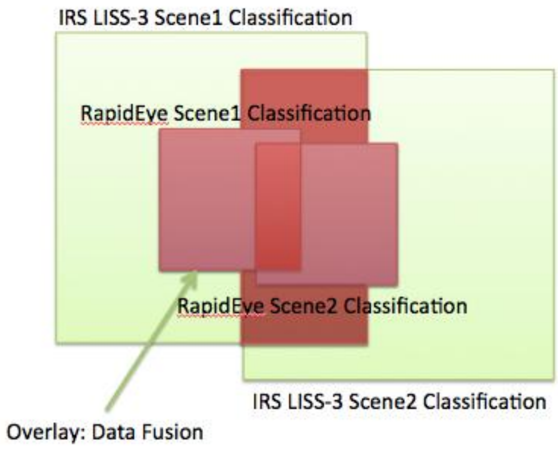
Uncertainty index is calculated for each scene classification. This step is important for the data fusion; it will prioritize the classifications. Uncertainty index is based on a classification validation performed with samples from the 2009 HR Layer imperviousness. A Kappa index (K) is computed on each validation and uncertainty is established as follow: $\Delta = 1\text{-}K$

*2.2. Data fusion*

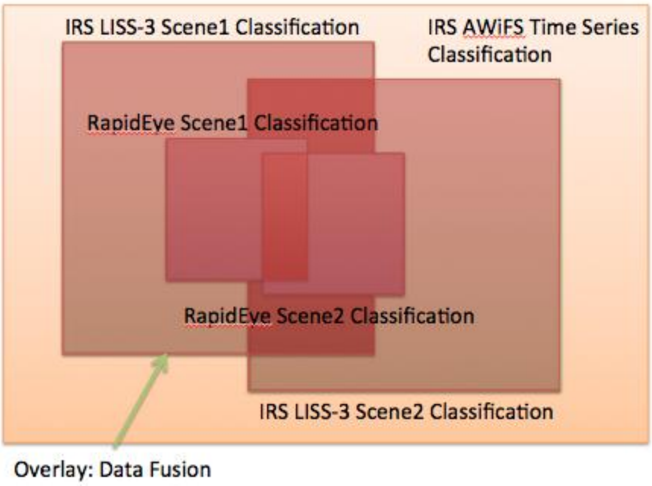
Among the data fusion techniques, we used the Dempster Shafer Theory (DST). In practice, classifications are fused when their extent has an overlap, scene classifications from a same sensor are fused first (Figure 1.a) and then, it follows by classifications from a different sensor (Figure 1.b). It results progressively a classification covering the overall area and providing a more accurate result (**Figure 1.c**).




*a. Fusion of classifications from a same sensor (RapidEye)*




*b. Fusion of classifications from a different sensor (RapidEye and IRS LISS-3)*




*c. Fusion progressively cover the overall area and accumulated overlaps become more accurate*

**Figure 1.** Process of the multi-classifications fusion

## 3. Results

This method was applied on 7 countries (over 1’000’000 km2) and validated with a manual imperviousness evaluation inside 100x100 m grid cells. Grid cells were defined randomly with a ratio of 1/1’000 km2. It results from this experiment that the method was validated at around 90 % in each country (**Table 1**Table 2).

| Country | Global Precision |
|---|---|
| Bulgaria | 98,9 |
| Czech Republic | 94,2 |
| France | 90 |
| Hungary | 89,9 |
| Poland | 91,5 |
| Slovenia | 91,4 |
| Slovakia | 93,8 |

**Table 2.** Global precision

Selected examples from the Strasbourg Large Urban Zone (Figure 2) present results in different context: rural area (Figure 3), urban fringe (Figure 4) and 2012 new built-up (Figure 5). The first example represents some villages surrounded by vineyards landscapes and the 3 separate classifications show some confusion with agricultural areas. The data fusion of these latter provides more high probabilities on the urban areas whereas the probabilities of the agricultural areas turned lower.

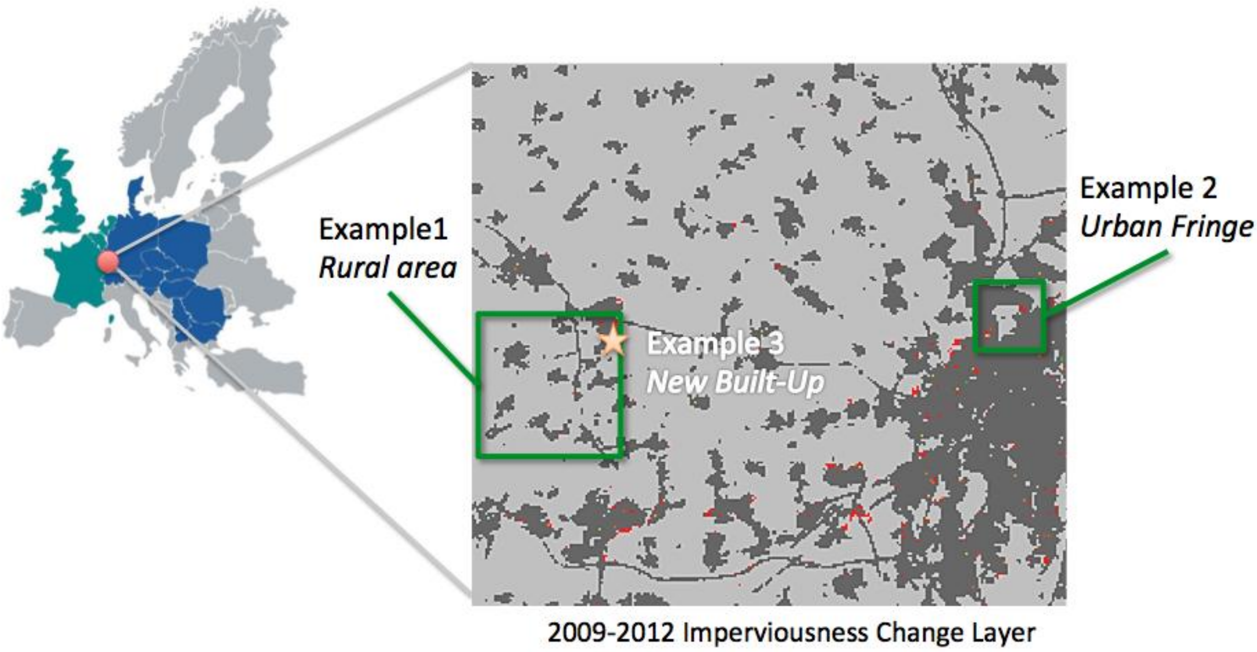
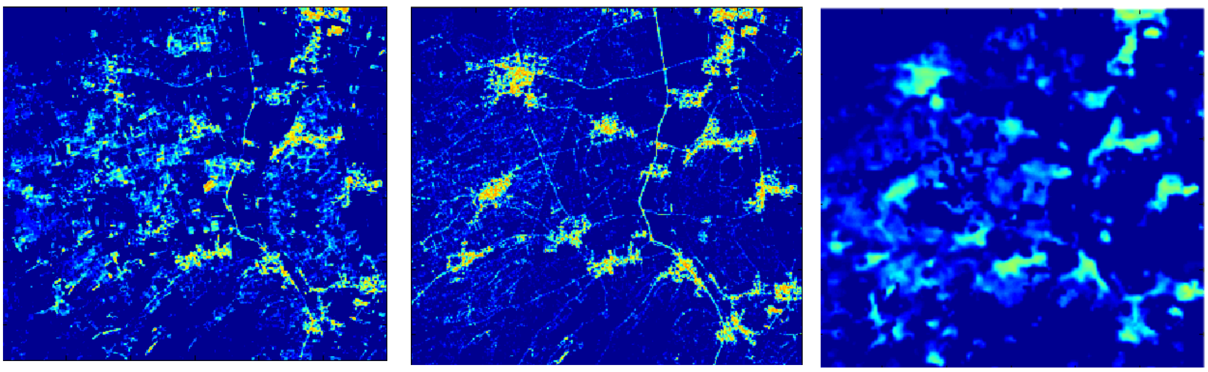
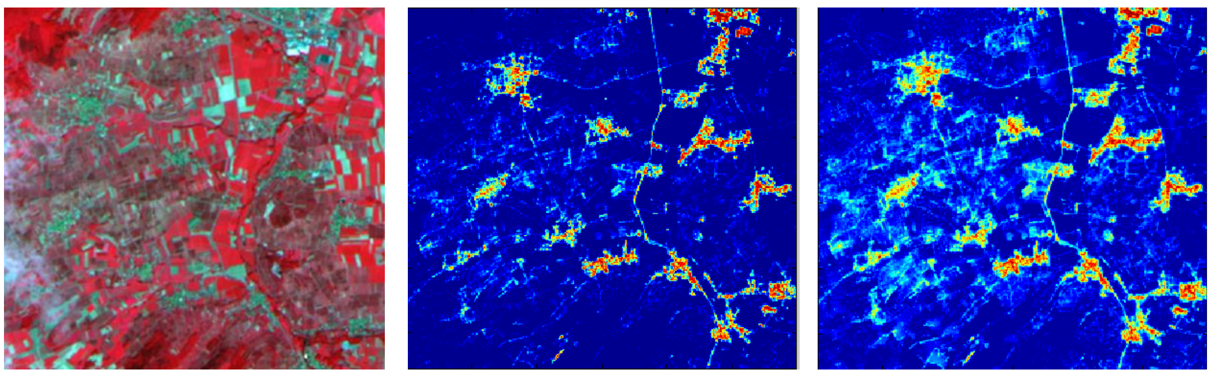



**Figure 2.** Presentation of the selected example



*a. IRS LISS-3 Classification* *b. RapidEye Classification* *c. AWiFS Classification*

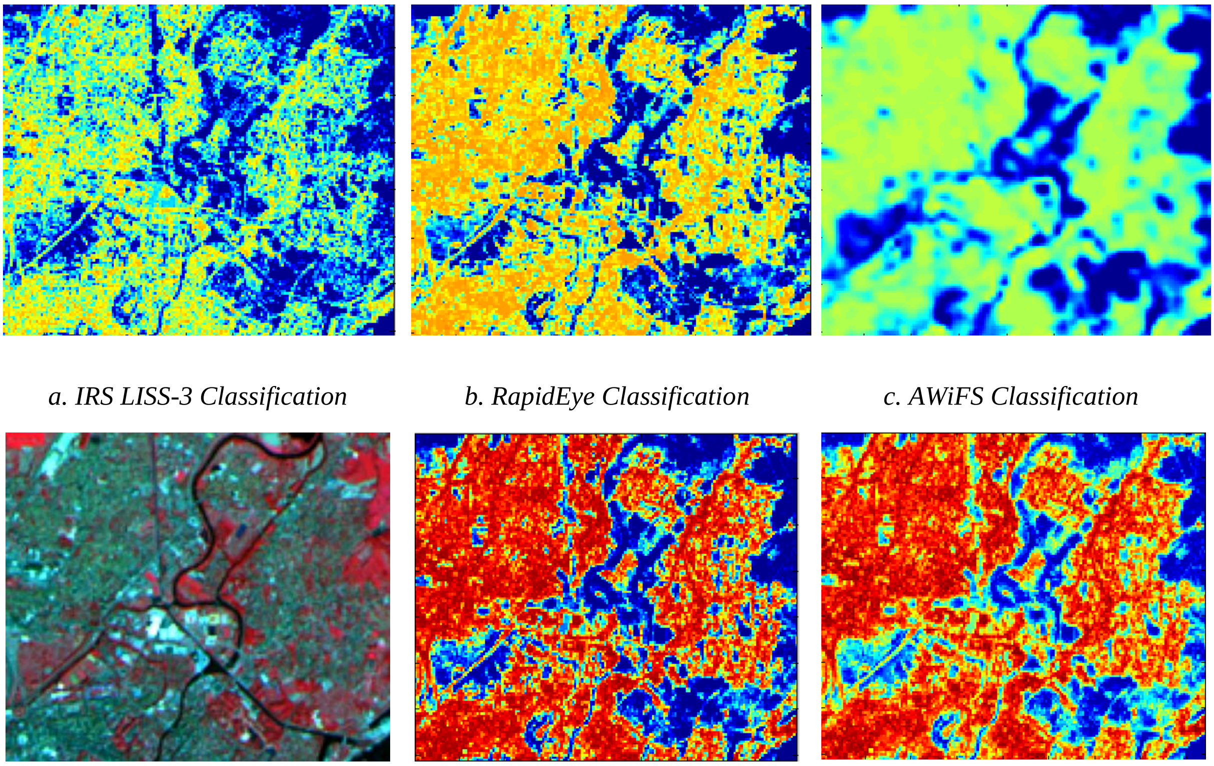


*d. IRS image* *e. Classifications Fusion* *f. Classifications Mean*

**Figure 3.** Single image classifications and data fusion in rural areas

The second example is located at the urban fringe (Figure 4) and represents a heterogeneous pattern of urban areas and semi-natural areas. The IRS AWiFS classification is coarse and boundaries of the urban areas are not clear. On the contrary, the RapidEye classification delineated precisely the border between the built-up and non built-up areas but it also seems to be noisier in agricultural areas. Compared to the mean of the 3 separate classifications, data fusion provides more sharpened edges on the urban border and is less noisy.

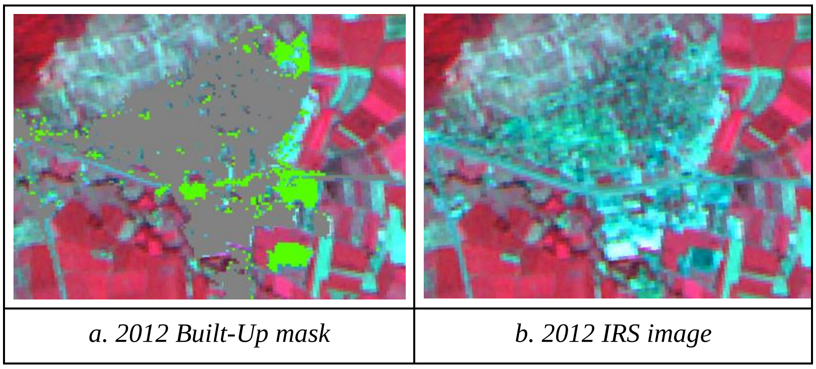
The new built-up example (Figure 5) shows a new housing lot at the North, new commercial buildings at the South and an enlargement of the road crossing the urban areas from West to East. The result shows these change areas are detected precisely (in green on Figure 5). Even if 2009 HR Layer imperviousness for the calibration, results are not over fitted and detect the new urban areas



*a. IRS LISS-3 Classification* *b. RapidEye Classification* *c. AWiFS Classification*

*d. IRS image* *e. Classifications Fusion* *f. Classifications Mean*

**Figure 4.** Single image classifications and data fusion in urban fringe



*a. 2012 Built-Up mask* *b. 2012 IRS image*

**Figure 5.** Detection of the new built-up

## 4. Conclusions

This paper presents a methodology dedicated to the manipulation of massive data to detect and update new built-up areas. The approach is based on the ESA Data WareHouse (DWH), which provides a large choice of images including medium, high and very high spatial resolution and time series. The approach relies on a separate image classification and a combination of each classification probabilities using evidence theory. Results performed 7 countries present an overall accuracy around 90%. Selected examples show that the method remains suitable in complex landscapes such as rural areas and urban fringe. It brings out the efficiency of the method to process large diversity and availability of the ESA DWH products and manage large projects such as GIO HR Layer.

## References

[1] Gangkofner, U., Weichselbaum, J., Kuntz, S., Brodsky, L., Larsson, K., & De Pasquale, V. (2009). Update of the European High-resolution Layer of Built-up Areas and Soil Sealing 2006 with Image2009 Data. 30th EARSeL Symposium: Remote Sensing for Science, Education and Culture.

[2] Lefebvre, A., Corpetti, T., & Hubert-Moy, L. (2010). Segmentation of very high spatial resolution panchromatic images based on wavelets and evidence theory. *SPIE Remote Sensing Conference, 7830-13* (pp. 1–13). Toulouse.

[3] Lu, Y. H., Trinder, J. C., & Kubik, K. (2006). Automatic Building Detection Using the Dempster-Shafer Algorithm. *Photogrammetric Engineering & Remote Sensing*, *72*(4), 395–403.

[4] Shafer, G. (1976). *Mathematical Theory of Evidence*. (Princeton, Ed.) (Princeton.). Princeton: Princeton University Press.

[5] Benediktsson, J., Swain, P. H., & Ersoy, O. K. (1990). Neural network approaches versus statistical methods in classification of multisource remote sensing data. *IEEE Transactions on Geoscience and Remote Sensing*, *28*(4), 540–552.

[6] Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, *5*(5), 703–714.