POLARIMETRY-BASED LAND COVER CLASSIFICATION WITH SENTINEL-1 DATA

Xavier Banqué[(1)], Juan M Lopez-Sanchez[(2)], Daniel Monells[(1)], David Ballester[(2)], Javier Duro[(1)], Fifame Koudogbo[(1)]

[(1)]Altamira Information S.L., Barcelona, Spain, Email: xavier.banque@altamira-information.com
[(2)]Universidad de Alicante, Alicante, Spain

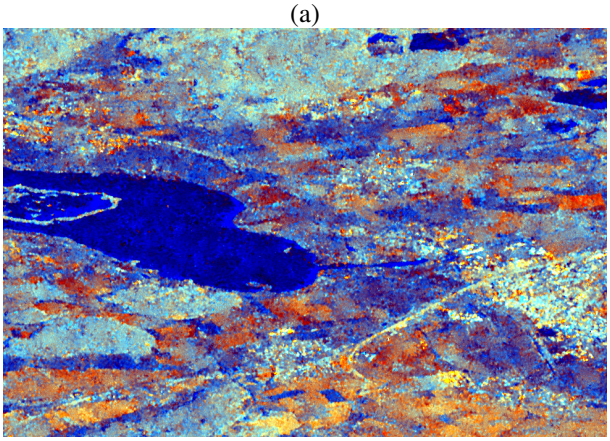
ABSTRACT

The presented research focuses on the assessment of the exploitation of the Sentinel-1 dual polarization data for land cover classification. In order to take advantage of massive data availability produced by Sentinel-1, data used in this research work is Interferometric Wide Swath mode, acquired over the Altmühlsee, Weißenburg-Gunzenhausen, Germany during November 2014. The developed preliminary classifier is based on the interpretation of several polarimetric figures as well as the Dual Polarization Entropy/Alpha Decomposition. Specifically, the following polarimetric indicators will be assessed: the channels cross-correlation, the cross and co-polar channels ratio and both cross and co-polar backscattering coefficients. The work carried out concentrates on the joint interpretation of the backscattering response of the co-pol and cross-pol channels for four or five different distributed targets that set the basis for an unsupervised simple land cover classifier. The developed research targets a preliminary unsupervised classifier able to differentiate between four or five terrain classes, including water, urban, forest and bare soil. Obtained results pave the way for the development of a Sentinel-1 based land classifier.

1. INTRODUCTION

With the arrival of Sentinel-1, SAR based Earth Observation applications will benefit from a new open source of dual polarization coherent data with a short revisit time that will boost the spectrum of remote sensing applications. Among them, land cover classification is a baseline technology upon which a myriad of higher level applications such as emergency management or urban planning can be built. The availability of coherent dual polarization products, comprising complex values and inter-channel phase information, will involve a major advantage in land cover classification with respect to former GMES SAR missions. The possibility of working with both co-polar and cross-polar channels provides two of the three components of the lexicographic feature vector in monostatic radar, theoretically allowing, for example, the detection of volume scattering produced by the forest canopy. The work described in this paper intends to prove the expected major step forward in the classification capabilities of Sentinel-1 data with respect to single pol space borne SARs.

(a)



(b)

Figure 1 (a) Optical image from Google Earth and (b) RGB composite of the Area of interest, from the VV, the VH and VH/VV descriptors.

2. METHODOLOGY

The test site selected for this work is located over the Altmühlsee, Weißenburg-Gunzenhausen, in Germany. It features many land cover types, including inland waterbodies, forest, bare soil, agricultural fields and urban areas, which represent a fair training ground for classification. The image was acquired by Sentinel-1 in 15 November 2014, using the Interferometric Wide Swath (IW) mode. This work is based only on the first swath (IW1) of the three available, covering a total area of 180 km in range and 80 km in azimuth. Figure 1

shows an optical image of the area and the corresponding RGB composite, created from the co-polar and the cross-polar channels and the power ratio between them. Table 1 describes the principal characteristics of the acquired data.

TABLE 1
SENTINEL-1 IW DATA CHARACTERISTICS

SYSTEM PARAMETER	VALUE
Coordinate System	Slant Range
Pixel Value	Complex
Bits Per Pixel	16 I and 16 Q
Polarization	Dual (VV+VH)
Ground Range Coverage [km]	251.8 (~84 per swath)
Slant Range Resolution [m]	2.7 (IW1)
Azimuth Resolution [m]	21.7
Slant Range Pixel Spacing [m]	2.3
Azimuth Pixel Spacing [m]	17.4
Incidence Angle [º]	32.9
NESZ [dB]	< -23.7

Since the acquisition is in TOPS configuration, a *debursting process* [1] must be applied to the Single Look Complex (SLC) product delivered. First, the composition of bursts to obtain the global image is done. This process also adjusts the azimuth time of each burst, in order to obtain a continuous reference.

The next step is to perform the radiometric calibration in order to obtain the sigma naught parameters for each polarimetric channel. This process has been done with the L1 Calibration Annotation Data Set, according to Sentinel-1 documentation [2].

One of the most important steps for SAR data based classification applications is the de-Speckle filtering. The following approaches have been tested: Boxcar filter, Refined Polarimetric Lee filter [3], and NL-means filter [4]. The later was finally selected, since it provided the best compromise between filtering and resolution preservation.

Regarding the classification step, several strategies have also been considered. First, some supervised classifiers have been analysed. Classical options such as Maximum Likelihood and Minimum Distance have been tested, to obtain a quick overview of the Sentinel-1 data capabilities for classification. A more robust approach, the Wishart classifier, has been used to improve the supervised classification.

Finally, this work proposes a model based classifier, based on the polarimetric features statistics, presented in section 3.

All classification strategies use as input information channels the polarimetric features available: sigma nought parameters, power ratio, span, normalized correlation, entropy and dominant alpha. The later parameter refers to the alpha angle of the dominant scattering mechanism.

Five different target classes for classification have been chosen, labelled *water*, *forest*, *bare soil*, *crops* and *urban*. For each class, a particular region of interest (ROI) over the SAR acquisition interest zone has been selected. The purpose of the ROIs is twofold. First, it provides a training area to distinguish each class using the supervised classifiers. Second, it defines the target areas where the polarimetric feature statistics are analysed for each class, when defining the model based classifier. In addition, within the present research work, these ROIs are further used for performance evaluation.

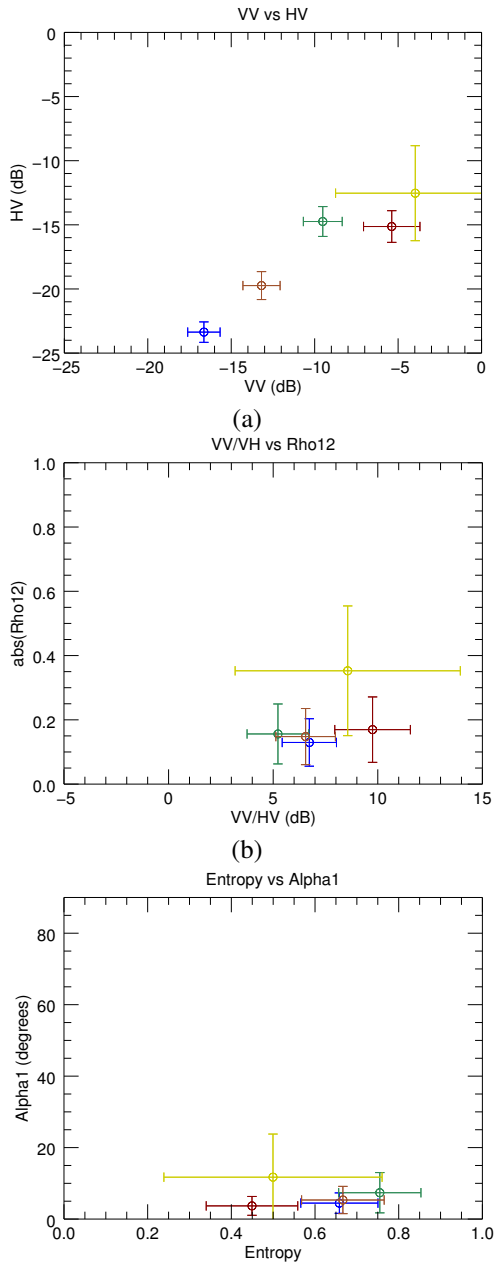


Figure 2: Distribution of the classes defined, using the mean and standard deviation in each ROI, by pairs of polarimetric features: (a) VV vs VH, (b) Power ratio vs Normalized correlation and (c) alpha1 vs entropy. The color code is Water (blue), Bare Soil (brown), Forest (green), Urban (yellow) and Crops (red)

3. LAND COVER CLASSIFIER

This section describes the model based classifier proposed, in order to improve the classification given by the supervised classifiers and work towards an unsupervised classifier.
This approach exploits the different statistical behaviour of the polarimetric features for each defined class, due to the different scattering mechanisms of the chosen classes. Figure 2 shows the statistical distribution of each class, by pairs of polarimetric features. The information in these graphs in combination with the polarimetric interpretation of the scattering mechanism of the different classes lead to several assumptions for each class:

a) *Water*: Both the co-polar and the cross-polar channels are expected to be very low due to waterbodies low backscattering.

b) *Bare Soil:* Both the co-polar and the cross-polar channels are expected to be low due to general low backscattering, but above water backscattering power. By using these two information channels seems feasible to identify this class from the rest.

c) *Forest*: Volumetric backscattering mechanism suggests that the cross-polar channel has more relevance than for other classes with respect to the total span. Moreover, entropy [5] is expected to be high due to the clear presence of more than one backscattering mechanism in the scatterers. The total span has to be moderate.

d) *Urban*: In general it shows the highest backscatter values in both channels, due to the man-made structures (buildings, bridges, etc.) but there is also a noticeable heterogeneity (texture). The most specific feature of this class is the absence of reflection symmetry, hence the correlation between VV and HV is not null (as it happened for the rest of natural covers). Urban, with respect to crops, presents more heterogeneity, no reflection symmetry, and high normalized correlation. In addition, dominant alpha has been observed to be higher than in the case of crops.

e) *Crops*: We have named this cover type as "crops" to highlight that it corresponds to agricultural fields. Its main characteristic is that it presents a high VV backscatter and a much lower cross-polar channel, which involves a very low dominant alpha value. This is in contrast to the "bare soil" class, for which both channels are low. Moreover, crops present a low normalized correlation value.

Figure 3 summarizes the previous rationale in a flow chart, being in fact the decision tree diagram for the model based classifier.

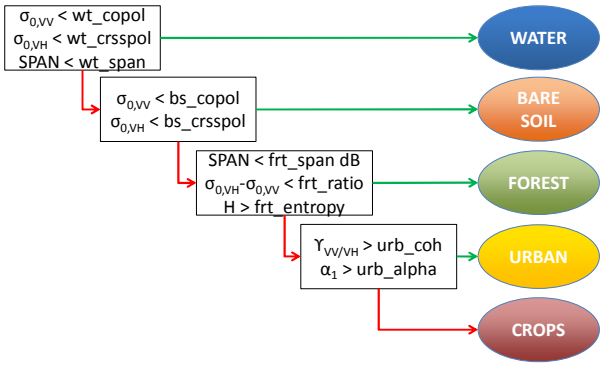


Figure 3 Model based classifier decision tree.

4. RESULTS

This section shows the classification results obtained with the different classifiers described in sections 2 and 3. Figure 5 depicts the different classification results obtained for each approach. They are focused in a small area near the Altmühlsee Lake, which contains a fair collection of the different land cover classes selected.
It has been noticed that the supervised classifiers used are especially sensitive to the radiometric parameters among the set of polarimetric indicators which are fed with. This fact was somehow expected since for the differentiation of most of the classes, the distance between classes is maximized using the co-polar and cross-polar sigma naught values. Nonetheless, for the differentiation between classes that show very similar co-pol and cross-pol radiometric values, it is important to rely on other polarimetric parameters such as normalized correlation or dominant alpha. This is exactly the goal of the proposed model based classifier, since some of the classes are hardly differentiated according to their backscattering power.
Comparing the results shown in Figure 5 as well as in the confusion matrices in Figure 4, some general observations arise.
The detection of water bodies and bare soil is almost identical in all the cases. This is expected due to the wide backscattered power difference with other classes shown in Figure 2a, which entails a high probability for the correct classification with respect to the rest of the classes. This is also reflected in the confusion matrices shown in Figure 4, where the water and bare soil classes (elements 1 and 5 of the diagonal respectively) show in general the highest agreement rates.
The main differences in the classifiers performance are found in the rest of the classes. Both the Minimum Distance and the Wishart approaches tend to overestimate the forest, mostly in detriment of the crops class.
The urban class is sub-estimated in all the cases. In fact, it shows the poorest agreement index in all the confusion matrices. The main reason for this result is the heterogeneity found in urban areas. The current definition of the ROI used may be ambiguous, since it

contains a mix of several kinds of land cover. Due to backscattering morphology, the highest weight in the definition of the class urban is the high reflectivity of man-made targets, which may represent only a small portion of the land cover considered as urban. Thus the results for the urban classification seem to be the worst since they are obtained from the evaluation of such a heterogeneous ROI, containing not only urban targets but others. This effect can be especially observed in the model based classifier results (see confusion matrices in Figure 4), since this classifier has been developed to select as urban only man-made targets, and in the evaluated urban ROI there is a myriad of kinds of targets.

Finally, the crops class is underestimated in the Minimum distance and the Wishart classifications, while the model based classifier seems to preserve it better. This is likely to be due to the fact that the model based classifier uses the normalized correlation between polarimetric channels and the dominant alpha in order to differentiate between urban and crops, instead of the level of backscattering which is comparable for both classes. See an example in the white circles in Figure 5.

Some outlier effects are also observed: it is hard to distinguish water bodies from other very low backscattering areas, such as airport runways or shadowing areas. Airports in the image are in general misclassified by all the candidate classifiers.

In summary, the results of all the evaluated classifiers confirm the classification potential of Sentinel-1 data. The proposed model based classifier can the improved to use the optimal polarimetric parameters in the detection of each land cover class, providing a better detection confidence potential. The urban class scope definition has to be refined since depending on what is to be classified as urban, the model based classifier has to be redefined for this class. Currently, the proposed preliminary version targets specifically man-made structures for the urban class, and not heterogeneous areas with man-made scatters among a variety of targets.

5. CONCLUSIONS

The main conclusions out the presented research work are described as follows. Sentinel-1 SLC data present good potential for land cover classification. Preliminary obtained results with the proposed model based classifier present a good classification performance (over approximately 95% for all the classes except for urban, which is the real challenge of the classifier). In the rest of the classes, different evaluated classifiers present similar results. A clear advantage of the model based classifier is its potential to become non-supervised with respect to other classifiers evaluated, despite a prior adjustment of the tree decision thresholds is needed during the classifier development.

Concerning further work, the identified next steps are the following. Evaluate results over an area with a good land-cover map (ground truth for performance evaluation). Assess the influence of the incidence angle and the impact of meteorological conditions during the data acquisition in the classification performance, which has not been taken into account in the presented work. Finally, the potential for crop type mapping evaluation will be as well studied, based on the morphology of the different crops effects in the scattering mechanism.

99,97	0,00	0,00	0,00	0,03
0,00	**99,61**	0,29	0,03	0,07
0,00	3,89	**84,72**	11,39	0,00
0,00	19,06	19,56	**61,15**	0,22
0,31	0,18	0,00	0,00	**99,51**

(a)

99,87	0,00	0,00	0,00	0,08
0,00	**97,66**	0,31	4,93	0,07
0,00	2,34	**97,44**	33,49	0,00
0,00	0,00	2,25	**61,09**	0,00
0,13	0,00	0,00	0,49	**99,86**

(b)

99,83	0,00	0,00	0,00	0,17
0,00	**99,03**	0,91	0,05	0,00
0,00	3,39	**96,50**	0,11	0,00
0,00	20,99	55,83	**23,04**	0,14
0,78	0,61	0,08	0,00	**98,53**

(c)

Figure 4 Confusion matrices for all classifiers. Each column is related to a class, ordered as: water, forest, crops, urban and bare soil. (a) Minimum distance classifier. (b) Wishart Distance classifier. (c) Model Based classifier.

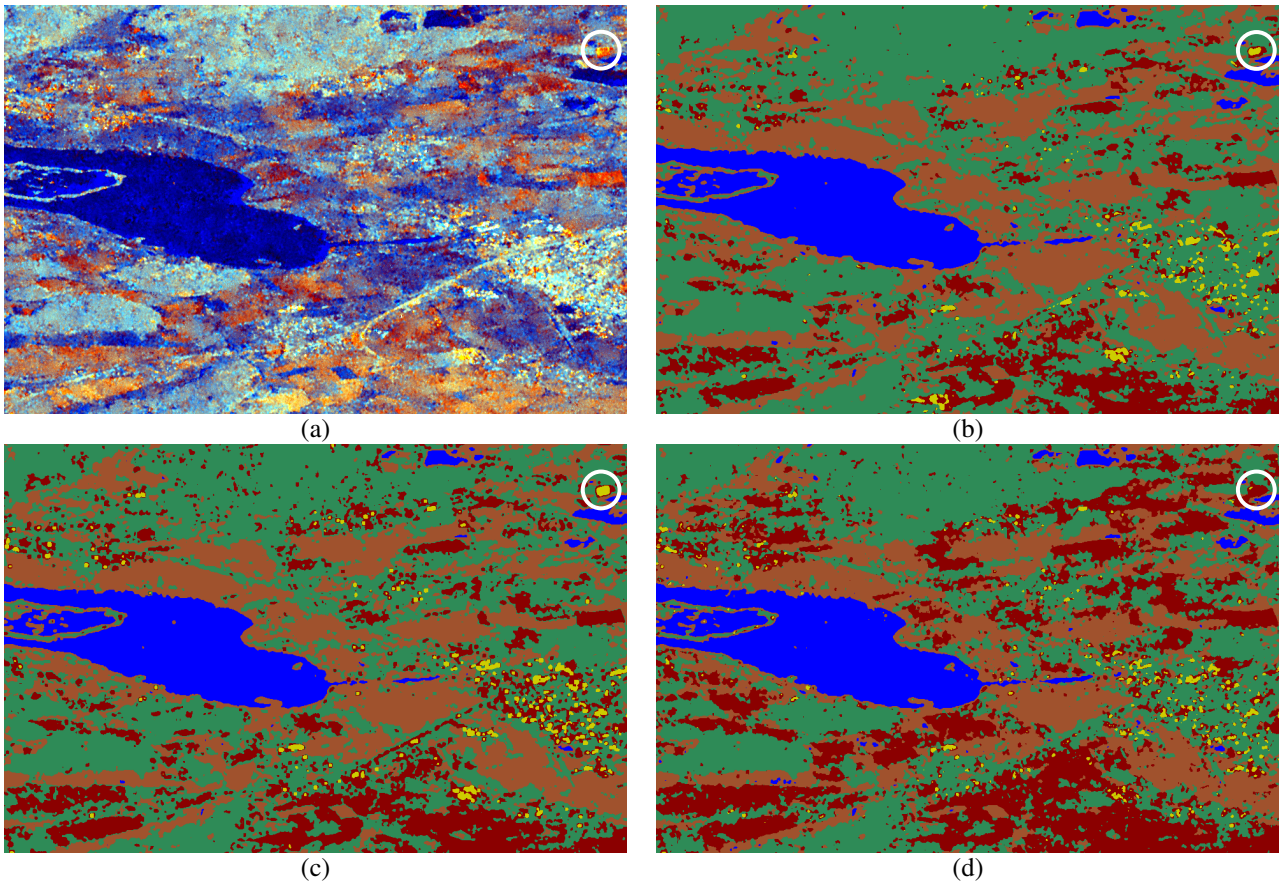


Figure 5 Classification results near the Altmühlsee Lake. (a) RGB composite.(b) Minimum distance classifier results. (c) Wishart Distance classifier results. (d) Model Based classifier results. The color code is Water (blue), Bare Soil (brown), Forest (green), Urban (yellow) and Crops (red)

6. REFERENCES

1. Prats-Iraola, Pau; et al. *TOPS Interferometry With TerraSAR-X,* IEEE TGRS Vol. 50. NO. 8. August 2012

2. MDA, *Sentinel-1 Product Specification document*, July 2013

3. Lee, Jong-Sen; Pottier, Eric, *Polarimetric Radar Imaging: From Basics to Applications* ,CRC Press, February 2009

4. Deledalle, Charles-Alban; et al., *NL-SAR: a unified Non-Local framework for resolution-preserving (Pol)(In)SAR denoising,* IEEE Trans. on Geoscience and Remote Sensing, vol. 53, no. 4, 2015

5. S. R. Cloude, *The Dual Polarization H2alpha Decomposition: a PALSAR Case Study*", Proceedings of POLinSAR 2007.