ABSTRACT

Probably a consequence of the unbalance between rice production in Asia and in Europe, satellite-based rice identification in Asia is widely discussed in scientific literature whereas SAR-based mapping of European rice paddy fields has received less attention so far. In this paper, we propose a simple methodology for identifying European rice paddy fields from time series of SAR data. Standard practices for management of water in conventional European rice paddy fields translates into a distinctive pattern of low backscatter values between April and May, typically preceded and followed by higher backscatter values due to ploughing and emergence. Our proposed method leverages such pattern to discriminate rice against other crops and in a test involving the entire Italian rice-producing region of Lombardy has achieved very good Overall Accuracy (OA) scores. This paper reports the method, our test results and draws some preliminary conclusions.

Index Terms— Crop mapping, European rice paddy fields, multitemporal analysis, Sentinel-1, time series.

1. INTRODUCTION

Rice mapping based on spaceborne remote sensing has been widely investigated in Southern Asian contexts, where most of global rice production takes place. Cultivation practices in such regions are however different from European ones for cultural, environmental and climatic reasons; simple reuse of those methods [1–3] will not lead to high levels of accuracy. This situation calls for specific research on space-based mapping of European rice, with an outlook to extracting information that may be useful for example to enhance food traceability by providing third-party (satellite-based) information on development stages, or on cross-checking compliance with production specifications. Spaceborne remote sensing data can indeed help monitoring crops at a large scale, providing precise and timely information on the phenological status and development of vegetation [4–7], detecting possible emerging problems, and also building independent records of the crop status. This may pave the way towards cross-checking of organic claims [8].

2. METHODOLOGY

Standard rice cultivation practices in Europe require flooding of paddy fields between April and May, a feature which is very specific and not generally encountered in other types of crops. Rice fields tend to display low radar backscatter values during this period compared to the rest of the year, due to the water surface covering the area. This peculiar situation can be leveraged to develop a simple yet effective rice identification strategy. As visible in Figure 1, mirror reflection from flooded fields does actually translate into significantly lower backscatter intensity in mid-to-late spring, a trend which is not observed in backscatter series from other land cover classes.

Fig. 1: Example of a typical rice field backscatter SAR time series (blue line) compared with similar series on other land cover types. The x axis reports the order number of satellite acquisitions, spanning a full calendar year; the y axis reports the recorded radar backscatter intensity in units of dB.

This suggests that a simple way to discriminate between paddy fields and other land cover classes could hinge on the identification of a clear dip in the reflectivity trend at a designated time of year. The experiments reported in this paper intend to test the feasibility of rice identification based on such feature, and to propose sensible values for classifi-
cation parameters like width, depth and temporal location of the reflectivity dip. Tests were carried out on a statistically significant area as described in the next section.

3. STUDY AREA AND DATA

3.1. Lombardy and rice production

Among all EU countries, Italy is the biggest rice producer, covering more than 53% of the entire European rice-growing area and exporting more than 45% in weight of its domestic output. In North-Western Italy, the province of Pavia in the Lombardy region (Figure 2) accounts for just above one third of the total domestic production [9] thanks to its 82,000 hectares of rice paddy fields. Lombardy was thus selected as the area of interest.

Fig. 2: Map of the Lombardy Region, Italy, and ground truth data (green polygons represent rice paddy fields).

3.2. The regional land cover map

Reference data on land cover over the Lombardy region of Italy were extracted from the geographic database named DUSAF 6.0 (“Destinazione d’uso dei suoli agricoli e forestali”, 6th version), referring to year 2018 and developed by Lombardy region using AGEA ortophotos and SPOT 6/7 satellite images, publicly available on the web Geoportal of Regione Lombardia [10].

A series of preliminary operations were carried out to adapt the database to the present study. These refined data were then split into two groups: the first one includes data of the administrative Provinces of Pavia, Milan and Lodi (PML), the largest producers of rice in the region; the second group consists of all agricultural fields in the region (AllAgr). This generates two sub-sets that are a bit better balanced, in terms of the ratio between the number of rice (R) fields and of non-rice (NR) fields, than the starting set. Plus, the two sub-sets account for 160,000 distinct polygons, which slashed the processing times with respect to the full set of 308,000 polygons.

3.3. Spaceborne data and processing tools

Cloud cover and haze are very frequent in Northern Italy in spring, when rice paddy flooding takes place. This makes spaceborne radar sensors a more reliable source of information than optical sensors. In terms of processing environment, Google Earth Engine (GEE) was selected because it offers easier access to data, processing tools and resources, than other options considered.

The C-band Sentinel-1 SAR constellation was selected as the radar data source, because of its open policy, and its features compatible with the intended application. In this work we used the standard IW acquisition mode, and the VH polarization because it offers greater sensitivity to the presence of vegetation emerging from the flooded fields. The reference year is 2018 and a time series of 60 radar acquisitions was considered for the analysis. The Sentinel-1 image collection available on GEE has already undergone a series of pre-processing steps leading to the final normalized backscatter coefficient ($\sigma^0$), ready to use.

4. TEST AND TUNING

On each polygon of the reference dataset, the spatially-averaged backscatter value was computed for every satellite acquisition, approximately every six days. The series of values was then smoothed out using a Savitzky–Golay (S-G) filter, which makes the envelope to emerge by suppressing most speckle-driven local jitter, while at the same time preserving relevant changes in time. This is of critical importance in the decision tree-based identification method proposed in this paper. Multiple tests led us to set a window length of 11, and filter order of 2 as suitable parameter values for the S-G filter.

Starting from the smoothed series, we defined our first classification approach, derived from the previous discussion. Simply, the date of the absolute minimum value is identified, and if this date falls between April 1st and June 30th, the polygon is tagged as “Rice”, whereas if it doesn’t, the polygon is tagged as “Non-Rice”. Since S-G filtering removed most of jitter, this simple criterion was sufficient to achieve a fair overall accuracy. Further research suggested the average value of the time series calculated over the entire year, and the absolute minimum value of the time series could provide additional, actionable information to improve classification results. An in-depth analysis of the rice field reference data has shown how these two parameters feature two well-defined ranges of typical values for rice paddy fields. As can be noted from Figure 4, minima and average values tend to fall in most cases around -22 dB and -18 dB respectively.
Using this information we could improve our first classifier by imposing two ranges based on these allowed values, keeping the algorithm simple with little expense in terms of omission errors, while reducing the number of false positives. Therefore, the condition on the date of the minimum is no longer sufficient to classify a polygon as “Rice”, and additional conditions were sought. Looking at statistical values for the different classes reported in Table 1, it becomes apparent that typical figures for minimum and yearly average are significantly different between rice and other land cover classes. This offered an additional, effective condition for identifying rice: the absolute minimum value must fall between -29 and -17 dB, and its average value must fall between -21 and -15 dB. This improvement proved to be more effective on the PML set, demonstrating that agricultural fields tend to be less separable than urban areas from rice paddy fields with this technique. Another remarkable point from Table 1 is that rice is more homogeneous and consistent that the other classes, since its standard deviations are lower both on minima and on yearly averages.

5. RESULTS AND DISCUSSION

Table 2 reports the accuracy results measured during the tests described in the previous section, broken down into different areas and different thematic components. On the main rows, the different geographic areas are reported. For each main row, two sub-rows report two separate discrimination experiments, i.e. rice against all other agricultural land cover classes on the upper half-row, and rice against all other general land cover classes on the lower half-row. It can be noted that accuracy values are always lower when classification is done against other agricultural land cover classes than in the general case. The corresponding $k$ values are not significant and have not been reported. It is indeed known that the polygon population is heavily unbalanced against rice, especially in the entire regional set; the overall accuracy is still deemed sufficient to guide the development of the algorithm and the tuning of its parameters. As mentioned in section 1, little previous research was done on this specific land cover discrimination problem; however, to let the reader get a sense of how rice mapping performs in other contexts, the methods proposed in [3] report Overall Accuracy levels falling between 80 and 90% in China.

6. CONCLUSIONS

In this paper, a simple yet effective method to identify rice fields from time series of spaceborne SAR backscatter intensity values has been proposed. The method focuses on rice
Table 2: Size of the statistical sample in number of polygons and overall accuracy (OA) attained on the various tests carried out on the proposed method. The first column defines the area of interest (entire Region, each single province), the second column defines against what classes rice is discriminated (all agricultural fields but rice, or all non-rice land cover classes). It can be noted that rice is more subject to be discerned against the general land cover class set than it is against other agricultural fields.

| ROI      | Consid. sets | #Polyg. | OA   |
|----------|--------------|---------|------|
| Lomb.    | R/AllAgr     | 108,171 | 91.2%|
|          | R/NR         | 85,899  | 92.38%|
| Pavia    | R/AllAgr     | 18,858  | 91.32%|
|          | R/NR         | 47,685  | 94.31%|
| Milano   | R/AllAgr     | 8,003   | 84.66%|
|          | R/NR         | 29,662  | 89.59%|
| Lodi     | R/AllAgr     | 2,770   | 90.07%|
|          | R/NR         | 8,552   | 91.38%|

Fig. 5: Flowchart summarizing the proposed algorithm. Condition on min location requires the absolute minimum to fall between April 1st and June 30th. Condition on min, avg value requires the absolute minimum value to fall between -29 and -17 dB, and the average value to fall between -21 and -15 dB.

Paddy fields in Europe, where the strong seasonality of rice growth cycle offers a means to identify them through analysis of a few temporal features. The typical dip in reflectivity associated with field flooding in April is characterized by well-defined ranges of values for a pair of relevant parameters. Specific tests assessed discriminability of rice paddy fields from the general set of land cover classes, and from other agricultural land cover classes. The proposed rule-based classifier performs well in the general case, less so when comparing rice with other agricultural fields. This may be related to the presence of farmlands that share some similarities with rice fields. Future research lines include investigation on additional shape parameters describing the spring minima, and incorporation of polarimetric information where available from Sentinel-1, and possible incorporation of crowdsourced information to build enhanced ground reference data [11].

7. ACKNOWLEDGMENTS

This work was partly funded by the European Commission through the H2020 MSCA RISE Project EOXPOSURE, Grant Agreement #734541.

8. REFERENCES

[1] G. S. Khush, “Origin, dispersal, cultivation and variation of rice,” *Plant molecular biology*, vol. 35, no. 1-2, pp. 25–34, 1997.
[2] J. M. Lopez-Sanchez, S. R. Cloude, and J. D. Ballester-Berman, “Rice phenology monitoring by means of SAR polarimetry at X-band,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 7, pp. 2695–2709, 2011.
[3] L. Xu, H. Zhang, C. Wang, B. Zhang, and M. Liu, “Crop classification based on temporal information using sentinel-1 SAR time-series data,” *Remote Sensing*, vol. 11, no. 1, p. 53, 2019.
[4] D. Zhou, S. Zhao, L. Zhang, and S. Liu, “Remotely sensed assessment of urbanization effects on vegetation phenology in China's 32 major cities,” *Remote Sensing of Environment*, vol. 176, pp. 272–281, 2016.
[5] J. Pastor-Guzman, J. Dash, and P. M. Atkinson, “Remote sensing of mangrove forest phenology and its environmental drivers,” *Remote Sensing of Environment*, vol. 205, pp. 71–84, 2018.
[6] S. Siachalou, G. Mallinis, and M. Tsakiri-Strati, “A hidden Markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data,” *Remote Sensing*, vol. 7, no. 4, pp. 3633–3650, 2015.
[7] E. Albanesi, S. Bernoldi, F. Dell'Acqua, and D. Entekhabi, “Covariation of Passive–Active Microwave Measurements over Vegetated Surfaces: Case Studies at L-Band Passive and L-, C- and X-Band Active,” *Remote Sensing*, vol. 13, no. 9, p. 1786, 2021.
[8] D. Marzi, D. De Vecchi, and G. C. Iannelli, “The ESA KSA Vialone Project: New Business from Space in Organic Farming,” *ESA Earth Observation Phi-Week, Frascati (Rome), Italy, 9th-13th September 2019*, 2019.
[9] ISTAT, “Open-data portal of the Italian National Statistics Institute,” Website, 2019, last checked: 13.11.2011. [Online]. Available: http://dati.istat.it/
[10] R. Lombardia, “Geoportal of the Lombardy region, Italy,” Website, 2018. [Online]. Available: http://www.geoportale.regione.lombardia.it/
[11] F. Dell’Acqua and D. De Vecchi, “Potentials of active and passive geospatial crowdsourcing in complementing sentinel data and supporting Copernicus service portfolio,” *Proceedings of the IEEE*, vol. 105, no. 10, pp. 1913–1925, 2017.