Modelling of agricultural SAR Time Series using Convolutional Autoencoder for the extraction of harvesting practices of rice fields

Thomas Di Martino[a,b], Élise Colin Koeniguer[b], Laetitia Thirion-Lefevre[a], and Régis Guinvarc'h[a]
[a]SONDRA, ONERA, CentraleSupélec, Université Paris-Saclay, 91190 Gif-sur-Yvette, France
[b]ONERA, Traitement de l'information et systèmes, Université Paris-Saclay, 91123 Palaiseau, France

Abstract

We apply an unsupervised learning methodology to project SAR Time Series of growing rice fields onto a 3-dimensional space, where we explicit differences between the fields. The projection method used is a Convolutional Autoencoder, trained using a reconstruction task and a mean-square cost function. The chosen embedding space is of dimension 3, to provide the possibility to visualise it spatially using an RGB false colour composite. We compare two subsets of rice fields at both embedding space and original SAR time series levels to analyze the nature of the variations between the two subsets.

1 Introduction

Since the launch of the Copernicus program and the Sentinel-1 satellites, more and more applications have used the concept of the SAR Time Series. Said applications include forest mapping [11], forest monitoring [12] or agricultural monitoring [14, 15]. Rice fields have also been receiving an increased amount of attention from the community as its agricultural practices, which include flooding, can be characterized using SAR time series [9, 10].
The increase in revisit time and depth of SAR Time Series requires the transformation of existing processing methodologies into new ones, more robust to these increased observable dimensions. Thus, a new methodology involving unsupervised learning has been introduced [6]. A deep learning architecture is used to project SAR Time-series onto a latent space where agricultural class-specific semantic is kept, while dramatically reducing the dimension of the data, allowing for the computation of clustering algorithms such as k-Means. While the original paper applies this methodology to retrieve agricultural classes and mislabels [6], we present it here to model variations in agricultural practices within rice fields.
First, we define this methodology and update it to present its potential to serve as a visualization tool of SAR Time Series. Then, we apply this methodology to Sentinel-1 Time Series of rice fields and display the resulting embedding space. Finally, we show how the differences between two subsets of rice fields in the embedding space signify differences in their original temporal profile. We then detail how variations of agricultural practices lead to these varying representations.

2 Methodology

Introduced in [6] as a mean to unsupervisingly process SAR Time Series, the convolutional autoencoder architecture (CAE) is a deep neural network architecture consisting of an encoder and a decoder. The forward pass of the model can be written as:

$$\tilde{p} = d(e(p)) \tag{1}$$

where:

- $e : \mathbb{R}^n \rightarrow \mathbb{R}^t$ is the encoder. It maps input time series p onto a representation space of lower dimension, called the embedding space, where separability is deemed superior ($n > t$).
- $d : \mathbb{R}^t \rightarrow \mathbb{R}^n$ is the decoder. It reconstructs the original input p using its latent representation $e(p)$. Its output is written $\tilde{p}$.

The weights of the encoder and the decoder are trained using backpropagation [13], with a mean-square cost function, to model the recreation error.
The first mention of such encoder-decoder architecture is mentioned in [4], as a *Nonlinear Principal Component Analysis*. We differ from this methodology with the use of convolutional layers [5], here of 1-dimension, to extract temporal patterns that acknowledge the sequential nature of time series.
Thus, rather than working with the original time-series p, we use their projected-self $e(p)$. Indeed, measures of similarity, such as distances between time-series, are more accurately calculated in the embedding space, thanks to its low dimension and its retention of semantic information from the original time series. The amount of information kept in the embedding space and the separability of that same space is conditioned on the chosen dimension, t. We find that an embedding space of size $t = 3$ is an optimal choice for easier illustration and computation.

3 Data presentation: Sector BXII, Sevilla, Spain

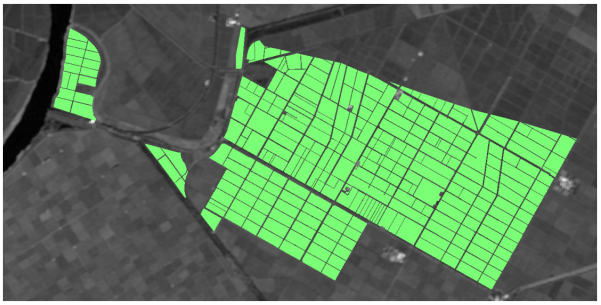


Figure 1 Illustration of agricultural fields labelled as Rice fields (in green), over a Sentinel-1 σ_0 temporal average image, in VV polarisation, of the BXII Sector (36°59 N 6°06W)

The data used in this study is a subset of the dataset introduced in [7]. It consists of a stack of 61 Sentinel-1 radiometric images, fixed on orbit 74 in Ascending mode, and all acquired during 2017. The preprocessing of the Single-Look Complex images is detailed in [7]. Each image has been manipulated and processed in dB scale, exploiting both VV and VH available polarisation.
While the original dataset contained labels for 17 classes, we focused on rice fields (**Figure 1**), thus modelling existing intra-class variance within the temporal profiles of growing rice. The presented rice fields belong to multiple farm operators. Our final trimmed dataset then consists of almost $300,000$ time series of rice fields.

The average temporal profile of rice fields (**Figure 2**)

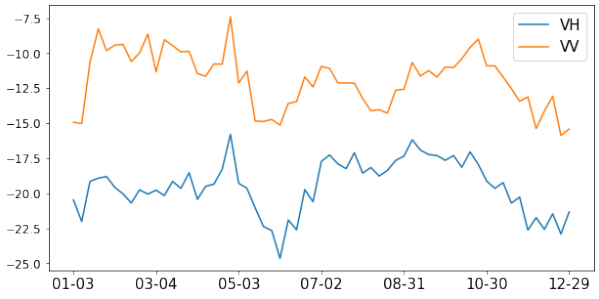


Figure 2 Average temporal profile of rice fields

shows a significant amount of temporal patterns, such as increase and decrease of backscatter, that can be correlated with the state of the field. Additionally, we observe some events appearing only in one of the two polarisations, implying that they are both crucial to characterize rice temporal radiometric profile correctly.

4 Application of the CAE to the rice-restricted BXII Sector

4.1 Training setup

On the one hand, the encoder of our CAE model is designed using 3 Stacks of 1D-Convolution, 1D-MaxPooling and ELU activation function before feeding the extracted temporal features to fully connected layers. On the other hand, our model's decoder uses stacks of fully connected layers, in-between of which we also find ELU activation functions. Our model is trained for 100 epochs, with a learning rate of 0.001, a batch size of 1024, using the ADAM optimizer [3].

4.2 Visualisation of the rice embedding space

The generated 3-dimensional embedding space can be visualised using scatter plots (**Figure 3**) and appears highly sparse. We notice the presence of compact clusters, which indicates a significant amount of sub-profiles of growing rice that have been projected onto a highly separable embedding space by the CAE model.

The rice fields embedding space can also be visualised either as 3 separate images (**Figure 4**) or as a single RGB image (**Figure 5**), where each component is the normalized embedding space. This way, we can observe the partitioning of fields into groups with supposed homogeneous temporal profiles. In addition, the variations of embedding values appear to be delimited by the crop outlines, which corroborates the assumption of the connection between the variation in rice characteristics (e.g. rice type, agricultural practice...) and the difference in time series. These division are also likely to be connected to differences in ownerships of the parcels. To further investigate the profiles available in the scene, we extract two groups of outlying colors and compare them to understand better the reason of their separation in the embedding space.

4.3 Comparison of antagonist rice temporal profiles

The two selected subsets of rice fields, with respectively light-blue and red-violet embedding space colors, are presented in **Figure 6**. Their opposite colors indicate a major difference in their average time profile, displayed in **Figure 7**.
Three key dates of 2017 acquisitions, highlighted in red in **Figure 7a** and **Figure 7b**, appear to show high differences in backscatter value between the two subsets. We will study the state of the fields at these three dates (9th of January, 21st of May, and 24th of October) using a combination of the original σ_0 images and optical Sentinel-2 images to better understand the variations in agricultural practices that led to these differences, caught on by the CAE model.
To lead our interpretation, we mostly focused on analyzing the VV signal, as we cannot explain the behavior of the VH polarization in **Figure 7a**, between January and May (slow

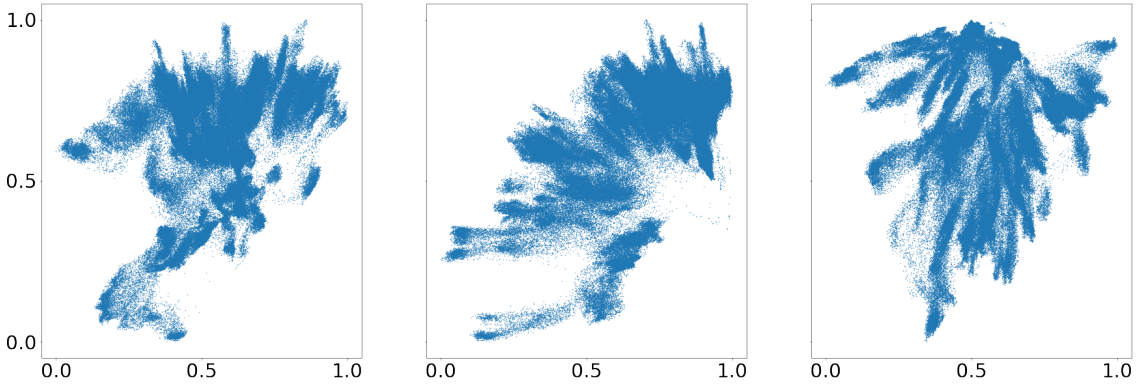
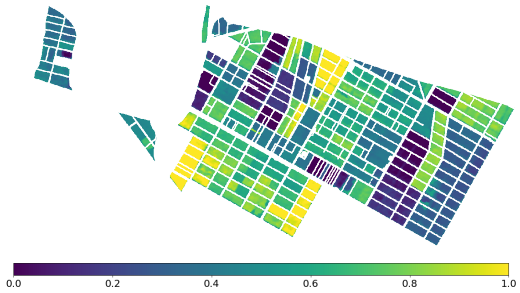
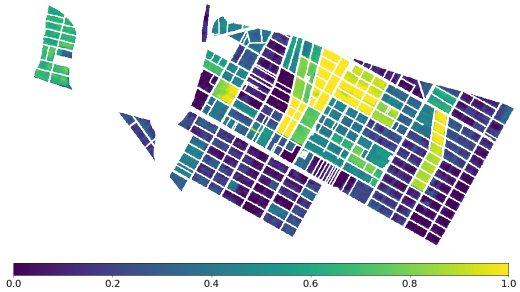


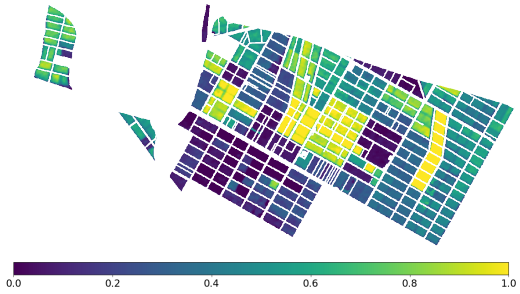
Figure 3 Illustration of rice fields embedding space. Left: 1st embedding dimension (x-axis) vs 2nd embedding dimension (y-axis), Center: 3rd embedding dimension (x-axis) vs 2nd embedding dimension (y-axis), Right: 1st embedding dimension (x-axis) vs 3rd embedding dimension (y-axis).



(a)



(b)



(c)

Figure 4 Rice fields embedding space represented as 3 separate images, where the color of each pixel corresponds to its normalized representation in the embedding space. (a) 1st normalized embedding space. (b) 2nd normalized embedding space. (c) 3rd normalized embedding space.

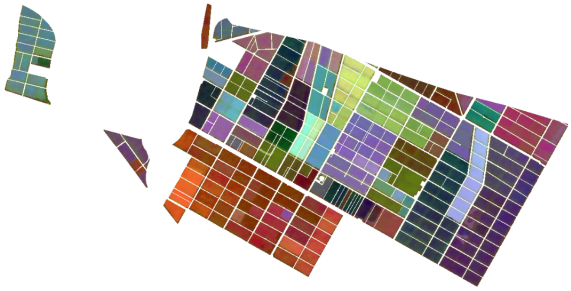
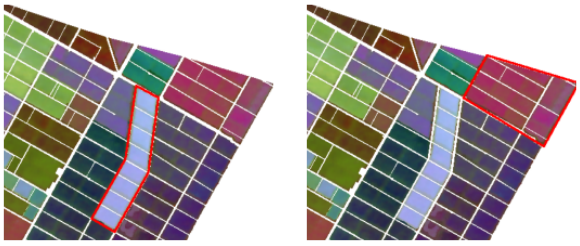


Figure 5 Rice fields embedding space represented as a single RGB image, where the Red channel corresponds to the normalized 1st embedding dimension, the Green channel corresponds to the normalized 2nd embedding dimension and the Blue channel corresponds to the normalized 3rd embedding dimension



(a) First selected rice field **(b)** Second selected rice field

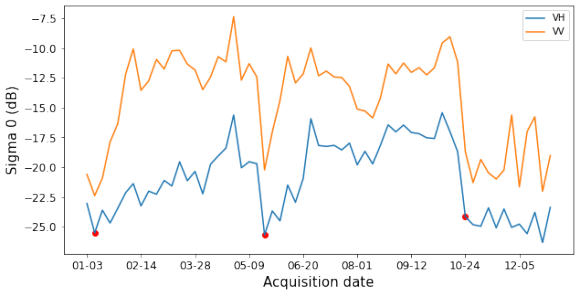
Figure 6 Select rice fields, colored using their embedding space representation

and progressive increase in VH signal against a drastic increase in VV signal).

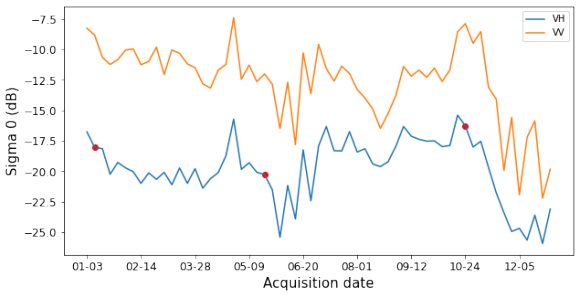
4.3.1 Analysis of the 9th of January acquisition

The typical crop calendar [1] for rice in the Sector BXII consists of:

- a seeding procedure surrounding the month of May and June.
- a growing phase during the rest of June, July, August



(a) σ_0 temporal profile of the first selected subset of rice fields (in dB)



(b) σ_0 temporal profile of the second selected subset of rice fields (in dB)

Figure 7 Comparison between the average temporal profiles of the two subset of rice fields

and September.

- and harvest, spanning over October and its nearby dates.

Thus, this first discriminatory date, the 9th of January, is within the preparation state of the field. While no Sentinel-2 images were found for this date, the Sentinel-1 results were self-explanatory enough to draw reliable conclusions. We show in **Figure 8** the clear difference in radiometry between the two selected rice subsets. It appears that the first selected field co-polarisation and cross-polarization radiometry is almost identical to the Guadalquivir river, shown on the top-left corner of the image. Oppositely, the radiometry of the second subset is high, among the highest of all 2017 acquisitions, showing the presence of vegetated structure over the crops.

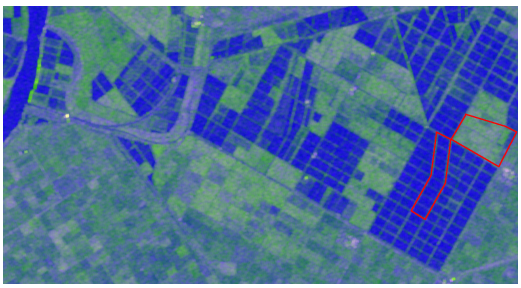


Figure 8 False color (Red: VV, Green: VH, Blue: VV-VH) σ_0 image of the 9th of January

However, the acquisition date is outside the of crop calendar of rice parcels. Thus, the observed differences are not directly linked to harvesting practices of rice. Indeed, the flooded condition of the subset 1 is likely to be connected to Spanish agri-environmental subsidies for the protection of agricultural systems of special interest for the populations of steppe birds and birds of the Andalusian rice fields [8]. Among the commitments farmers have to fullfil to receive these subsidies, there is an obligation for rice fields to be kept flooded once the productive cycle of the crop has finished, usually from November to January/February of the following year. As this practice is not mandatory but only financially encouraged, not every farm operator will flood his fields, thus resulting in differences in signals during the first weeks of the year between rice parcels.
In addition, the higher backscatter of subset 2 parcels can be linked to remaining vegetated structures, as the harvest of rice fields usually leaves 20 to 25cm-high branches.

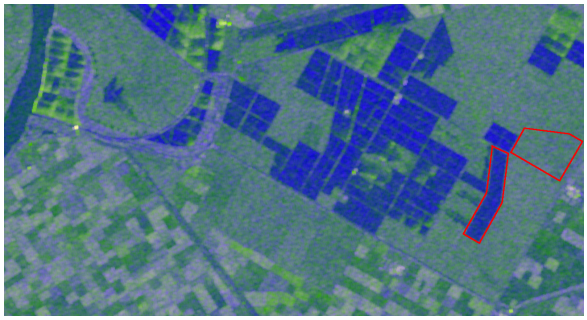
4.3.2 Analysis of the 22nd of May acquisition

We present in **Figure 9** visuals for the studied area, with Sentinel-1 false color composite image, Sentinel-2 RGB image and Sentinel-2 derived Normalized Difference Flood Index [2]. We see that the backscatter of the first selected subset is way lower than the second, in a similar manner than for the 9th of January, as mentioned in Section 4.3.1. However, the additional advantage of Sentinel-2 imagery, especially the NDFI, allows us to defend the hypothesis of a flooded state of the first rice subset, while the second remains unflooded.
To put the state of both fields in the context of rice growing and its phenology, as described by SAR Time Series, we will use the v-shaped typical temporal evolution of the Sentinel-1 VH/VV backscatter of rice fields, as presented in [9].

When plotting the VH/VV backscatter profiles of both subsets (**Figure 10**), we observe that the start of this v-shaped pattern (shown as a dashed vertical red line), an indicator of the sowing date, is offset by 24 days (or 4 Sentinel-1 acquisitions) between the two subsets, which is another difference between the agricultural practices of these two subsets of rice fields.

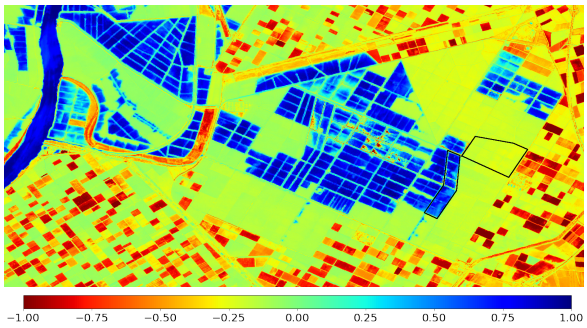
4.3.3 Analysis of the 24th of October acquisition

Now that we explicit a difference in sowing timing between the two subsets, we will further investigate the month of October, which is likely to be the month of harvest [1].
Displayed in **Figure 11**, the Sentinel-1 false colour image, the Sentinel-2 RGB and the NDFI all seem to indicate that yet again, the difference between the two subsets of rice fields is the flooded condition of the first. We can hypothesize that while the first subset of rice fields has been harvested, the second is still growing. Indeed, the decrease in signal for the second subset of rice fields only happens 4 acquisitions later (approx. 24 days), which show a temporal offset synced with the one observed for the sowing process.



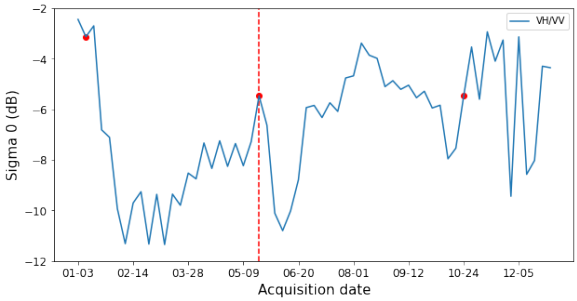
(a) False color (Red: VV, Green: VH, Blue: VV-VH) σ_0 image of the 21st of May

(b) Sentinel-2 RGB (Red: B4, Green: B3, Blue: B2) image acquired on the 22nd of May

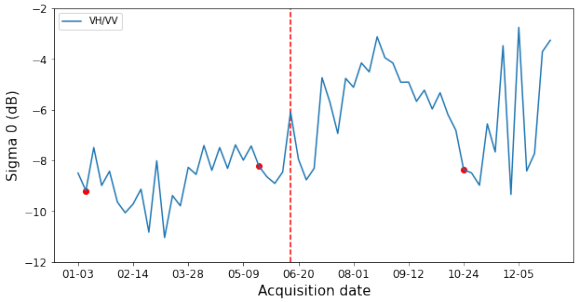


(c) Sentinel-2 derived Normalized Difference Flood Index, using B4 and B11 bands from the acquisition of the 22nd of May (Red: -1, indicates dried conditions; Blue: 1, indicates flooding conditions)

Figure 9 Study of the second key differentiating date between the two selected rice subsets, using (a) the original sigma0 Sentinel-1 imagery, (b) Sentinel-2 RGB imagery from the 22nd of May and (c) NDFI derived from the same 22nd of May Sentinel-2 acquisition



(a) VH/VV temporal profile of the first selected subset of rice fields (in dB)



(b) VH/VV temporal profile of the second selected subset of rice fields (in dB)

Figure 10 Comparison between the average temporal profiles of the VH/VV ratio of the two subsets of rice fields

5 Conclusion

In this paper, we present the application of the Convolutional Autoencoder to SAR Time Series of growing rice fields, consisting of Sentinel-1 images acquired over a year. The chosen 3-dimensional embedding space allows for an RGB visualisation of that space where the separation of rice fields in various temporal profiles is intuitively possible, without the need for a clustering methodology. Additionally, we demonstrate that the proposed methodology enables the analysis of variations in agricultural practices over the rice fields, where a temporal offset in the rice sowing, growing, and harvesting practices was unveiled.
This paper presents the potential of this method for agricultural monitoring for a large amount of data. Its high sensibility to variations in rice-growing patterns can be used to monitor, to some degrees, crop yield, crop health, as well as remotely evaluate the eventual impacts of disturbances (e.g. bacteria, fires) on agricultural growing volume.
Furthermore, its application to other fields, such as forest monitoring, are to be explored, as the behaviour of the algorithm on various agricultural plots shows excellent potential for its applicability to forest temporal modelling.

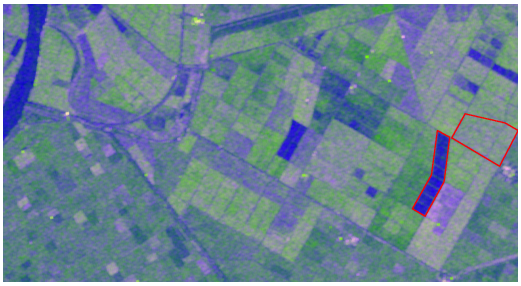
Acknowledgment

The authors would like to thank Juan M. Lopez-Sanchez for providing us with both the reference data, originating from the Regional Government of Andalucía and the Spanish Agrarian Guarantee Fund (FEGA), as well as for the preprocessed multitemporal Sentinel-1 images. We also want to thank him for the thoughtful discussions and exchanges.

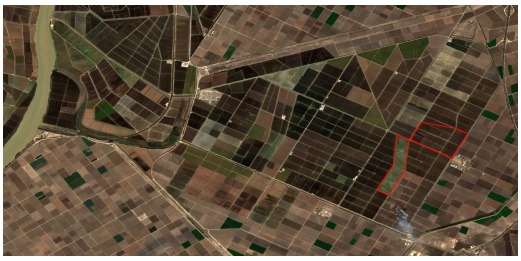
6 Literature

[1] Mario Busquier, Juan M. Lopez-Sanchez, Alejandro Mestre-Quereda, Elena Navarro, María P. González-Dugo, and Luciano Mateos. Exploring tandem-x interferometric products for crop-type mapping. *Remote Sensing*, 12(11), 2020.

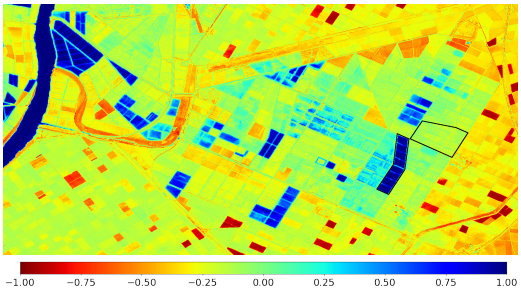
[2] Fabio Cian, Mattia Marconcini, and Pietro Ceccato. Normalized difference flood index for rapid flood



(a) False color (Red: VV, Green: VH, Blue: VV-VH) σ_0 image of the 24th of October



(b) Sentinel-2 RGB (Red: B4, Green: B3, Blue: B2) image acquired on the 24th of October



(c) Sentinel-2 derived Normalized Difference Flood Index, using B4 and B11 bands from the acquisition of the 29th of October (Red: -1, indicates dried conditions; Blue: 1, indicates flooding conditions)

Figure 11 Study of the second key differentiating date between the two selected rice subsets, using (a) the original sigma0 Sentinel-1 imagery, (b) Sentinel-2 RGB imagery from the 24th of October and (c) NDFI derived from the same 29 of October Sentinel-2 acquisition

mapping: Taking advantage of eo big data. *Remote Sensing of Environment*, 209:712–730, 2018.

[3] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2017.

[4] M. Kramer. Nonlinear principal component analysis using autoassociative neural networks. *Aiche Journal*, 37:233–243, 1991.

[5] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1(4):541–551, 1989.

[6] Thomas Di Martino, Régis Guinvarc'h, Laetitia Thirion-Lefevre, and Élise Colin Koeniguer. Beets or cotton? blind extraction of fine agricultural classes using a convolutional autoencoder applied to temporal sar signatures. *IEEE Transactions on Geoscience and Remote Sensing*, pages 1–18, 2021.

[7] Alejandro Mestre-Quereda, Juan M. Lopez-Sanchez, Fernando Vicente-Guijalba, Alexander W. Jacob, and Marcus E. Engdahl. Time-series of sentinel-1 interferometric coherence and backscatter for crop-type mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:4070–4084, 2020.

[8] Claire A. Pernollet, Anis Guelmami, Andy J. Green, Antoni Curcó Masip, Bosco Dies, Giuseppe Bogliani, Franco Tesio, Anne Brogi, Michel Gauthier-Clerc, and Matthieu Guillemain. A comparison of wintering duck numbers among european rice production areas with contrasting flooding regimes. *Biological Conservation*, 186:214–224, 2015.

[9] Hoa Phan, Thuy Le Toan, and Alexandre Bouvet. Understanding dense time series of sentinel-1 backscatter from rice fields: Case study in a province of the mekong delta, vietnam. *Remote Sensing*, 13(5), 2021.

[10] Hoa Phan, Thuy Le Toan, Alexandre Bouvet, Lam Dao Nguyen, Tien Pham Duy, and Mehrez Zribi. Mapping of rice varieties and sowing date using x-band sar data. *Sensors*, 18(1), 2018.

[11] Andrea Pulella, Rodrigo Santos, Francescopaolo Sica, Philipp Posovszky, and Paola Rizzoli. Multi-temporal sentinel-1 backscatter and coherence for rainforest mapping. *Remote Sensing*, 12:847, 03 2020.

[12] Andrea Pulella, Francescopaolo Sica, and Paola Rizzoli. Monthly deforestation monitoring with sentinel-1 multi-temporal signatures and insar coherences. *Revista de Teledeteccion*, 56:1–22, 11 2020.

[13] David E. Rumelhart and James L. McClelland. *Learning Internal Representations by Error Propagation*, pages 318–362. 1987.

[14] Tracy Whelen and Paul Siqueira. Time-series classification of sentinel-1 agricultural data over north dakota. *Remote Sensing Letters*, 9(5):411–420, 2018.

[15] Lu Xu, Hong Zhang, Chao Wang, Bo Zhang, and Meng Liu. Crop classification based on temporal information using sentinel-1 sar time-series data. *Remote Sensing*, 11(1), 2019.