



**HAL**  
open science

## Recurrent-based regression of Sentinel time series for continuous vegetation monitoring

Anatol Garioud, Silvia Valero, Sébastien Giordano, Clément Mallet

### ► To cite this version:

Anatol Garioud, Silvia Valero, Sébastien Giordano, Clément Mallet. Recurrent-based regression of Sentinel time series for continuous vegetation monitoring. *Remote Sensing of Environment*, 2021, 263, pp.112419. 10.1016/j.rse.2021.112419 . hal-03274870

**HAL Id: hal-03274870**

**<https://hal.science/hal-03274870>**

Submitted on 30 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Recurrent-based regression of Sentinel time series for continuous vegetation monitoring

Anatol Garioud<sup>a,b,\*</sup>, Silvia Valero<sup>b</sup>, Sébastien Giordano<sup>a</sup>, Clément Mallet<sup>a</sup>

<sup>a</sup>Univ. Gustave Eiffel, IGN-ENSG, LaSTIG, Saint-Mande, France

<sup>b</sup>CESBIO, Université de Toulouse, CNES/CNRS/INRAE/IRD/UPS, Toulouse, France

---

## Abstract

Dense time series of optical satellite imagery describing vegetation activity provide essential information for the efficient and regular monitoring of vegetation. Nevertheless, the temporal resolution of optical sensors is strongly affected by cloud cover, resulting in significant missing information. The use of complementary acquisitions, such as Synthetic Aperture Radar (SAR) data, opens the door to the development of new multi-sensor methodologies aiming at the reconstruction of missing information. However, the joint exploitation of new radar and optical missions, such as the Sentinel, raises new challenges given the different nature and response of the two data sources. In this work, the SenRVM methodology is proposed as a new multi-sensor approach to regress SAR time series towards Normalized Difference Vegetation Index (NDVI). A deep Recurrent Neural Network architecture which integrates SAR acquisitions and ancillary data is adopted. The regression task permits a continuous optical temporal resolution of 6 days. Multiple experiments are carried out to assess the SenRVM framework by studying two large-scale areas in France. Through an extensive interpretation of the results, SenRVM is evaluated on three main vegetation types (grasslands, crops, and forests). High accurate results ( $R^2 > 0.83$  and  $MAE < 0.05$ ) over more than 140,000 vegetation polygons are obtained, both for multi-class and class-specific models. The importance of the SenRVM input features is discussed through an ablation study, highlighting that the relevance of the features differs for the various classes. The good performances of SenRVM are assessed by evaluating the quality of the reconstruction over short- and long-term data gaps. Results are compared to five state-of-the-art methodologies. Finally, some preliminary experiments are first carried out to show how SenRVM results could be used to improve the existing cloud & shadow masks. The recovering of time series breaks caused by vegetation cover changes is also illustrated.

**Keywords:** Time series, Recurrent Neural Network (RNN), Regression, Missing Data, Sentinel-2, Sentinel-1, Data fusion, NDVI, Vegetation, Monitoring

---

## 1. Introduction

A growing body of evidence shows the essential benefits of vegetation cover for ecosystem services. Its capacity for carbon storage, quality food production, or as a biodiversity reservoir (Hufkens et al., 2016; Yang et al., 2019) are increasingly threatened by current climate changes and human interventions. Large vegetation areas across the globe can be divided into three major types: *grasslands*, *crops* and *forests*. *Forests* and *grasslands* are respectively the world's first and second carbon sinks for climate regulation (Dass et al., 2018). Both are home to the most important terrestrial biodiversity and preserve soil from erosion or impoverishment. *Crops* play a key role in many sustainable agricultural practices aiming a healthy environment, economic profitability and social and economic equity (Massawe et al., 2015). Monitoring the status and evolution of vegetation surfaces at a global scale provides essential information for climate modeling or public policy implementation (Bengtsson et al., 2019; Buchmann, Nina and Fuchs, Kathrin and Feigenwinter, Iris and Gilgen, Anna K., 2019). Vegetation types exhibit multiple modes of growth and senescence within their different phenology cycles, which vary under human interventions and climate influences. Regular vegetation monitoring over the globe thus requires frequent wide-area measurements. Large scale remote sensing time series is a well-known tool meeting the global coverage and repeated measurement requirements (Chuvieco, 2008; Li and Roy, 2017).

The relevance of remote sensing imagery for vegetation monitoring has been proven since the launch of the first Earth Observation (EO) satellites (Wulder et al., 2012). Vegetation dynamics can be recognized from optical measurement given their spectral reflectance properties. Consequently, numerous studies exploit optical time series for

vegetation related applications. Mapping the extent of vegetation areas (Fensholt and Proud, 2012; Zhang et al., 2017), estimating yields and management practices, monitoring the impact of climatic events and diseases (Liu et al., 2020) are some classical examples. The literature commonly adopts vegetation indices derived from satellite data for applications involving vegetation monitoring. The most prominent vegetation index is the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979), which is still extensively used (Ali et al., 2016; Belgiu and Csillik, 2018; Misra et al., 2020; Yang et al., 2020). Besides its simple computation, the popular success is explained by its efficiency to retrieve vegetation phenology information and health conditions (Reinermann et al., 2020; Zeng et al., 2020).

Temporal NDVI profiles provide efficient information for vegetation condition monitoring. Vegetation phenological changes can be accurately captured when optical satellite observations are acquired with sufficient repetitivity. Unfortunately, the temporal resolution of optical satellites is significantly affected by weather conditions. Cloud coverage causes considerable data gaps in both spatial and temporal domains. As a consequence, the temporal sampling of optical time series is irregular. It is among others, constrained by the geographical location of the area under study (Sudmanns et al., 2020). Large temporal data gaps can last from weeks to months or occur at key dates involving important vegetation changes (e.g., growth or senescence periods or even harvest time). To address the data loss, gap filling methods are commonly proposed to increase the time series temporal resolution. Most of the existing methodologies exploit past and future observations of the same sensor to recover missing data. This classical strategy can suffer from important limitations in the presence of large data gaps (Moreno-Martínez et al., 2020). These drawbacks become more apparent for non-stationary vegetation time series having seasonal gradual or abrupt breaks. To address such shortcomings, the integration of external complementary data in the gap filling process is proposed by some recent strategies.

---

\*Corresponding author

Email address: anatol.garioud@ign.fr (Anatol Garioud)

In this era of new satellite missions, the opportunity of fusing information from different sensors opens the door to the development of new methodologies aiming at reconstructing vegetation index time series with high temporal sampling. One of the most promising complementary satellite data sources is the one acquired by the Sentinels missions. The trade-off between spatial, spectral, and temporal resolutions is significantly improved by the Sentinels satellites. The constellation provides free and open access data covering the globe which are acquired by different high resolution sensors. Sentinel-2 A&B (Gascon et al., 2014) are carrying multi-spectral imaging sensors while Synthetic Aperture Radar (SAR) instruments are on-board Sentinel-1 A&B (Potin et al., 2019) satellites.

Given the particular nature of optical and radar signals, each of the sensors has its pros/cons and fields of application. Sentinel-2 brings new opportunities for regular vegetation monitoring, however, its high temporal resolution cannot be guaranteed due to cloud coverage. Weather independent Sentinel-1 acquisitions can allow constant vegetation monitoring (Schmitt and Zhu, 2016). The benefit of SAR data exploitation for vegetation studies involving *forests* (Frison et al., 2018; Sr. et al., 2018), *crops* (Veloso et al., 2017; Kumar et al., 2018; Hosseini et al., 2019; Ouadi et al., 2020) or *grasslands* (Tamm et al., 2016; Fauvel et al., 2020) is highlighted by numerous studies. Nevertheless, the extraction of interpretable descriptors is not straightforward. High statistical fluctuations of SAR signals bring strong inconsistencies in their time series. SAR acquisitions are significantly influenced by exogenous factors such as climate and relative humidity effects (Vreugdenhil et al., 2018) or geometric effects led by the side-looking nature of SAR sensors (Nasrallah et al., 2019). Another inherent problem of SAR systems is the speckle noise, which is due to the negative or positive interferences of all elementary targets within a pixel. As a consequence, the use of SAR sensors for monitoring changes in vegetation is more restricted in the literature.

In this context, this work proposes a multi-source methodology exploiting the capabilities of both satellite data sources. The joint complementarity between SAR and optical Sentinel imagery is used to reconstruct dense temporal NDVI time series. A new recurrent deep-based architecture is designed to explore the high dimensionality of the data and their complex relationships. The proposed architecture, namely SenRVM, targets to regress multivariate SAR time series and proposes the incorporation of contextual knowledge to reduce the impact of exogenous factors leading to SAR data variability. The contributions of the proposed object-based methodology are:

- (i) a recurrent neural network (RNN) approach that jointly integrates multivariate SAR and optical time series with ancillary data. The proposed approach permits to densely estimate NDVI at a 6 days temporal rate,
- (ii) an extensive evaluation of the results which are obtained over large scale geographical areas for the next three vegetation classes: *grasslands*, *crops* and *forests*,
- (iii) a comparison with existing methodologies both in terms of prediction accuracy and ability to recover time series breaks.

The content of this work is organised as follows. A review of prominent gap filling and deep-based data fusion techniques is first proposed in Sec. 2. In a second step, Sec. 3 details the study areas, data sources, and the compulsory pre-processing steps, both for satellite and ancillary data. Sec. 4 explains the proposed architecture and experimental results are presented in Sec. 5. Finally, conclusions and future perspectives are drawn in Sec. 6 and 7.

## 2. Related work

A large range of research efforts has been devoted to develop non-parametric methodologies for recovering time series of optical-derived vegetation indices with high temporal sampling (Verrelst et al., 2015; Cai et al., 2017; Belda et al., 2020a). Two main categories of methods are found in the literature that target to increase the temporal resolution of the time series: standard interpolation methods and machine learning (ML) regression algorithms.

Standard interpolation methods can be considered as gap filling reconstruction strategies recovering missing information. These methodologies can be divided into different categories (Shen et al., 2015; Desai and Ganatra, 2012; Yin et al., 2017; Lepot et al., 2017; Gerber et al., 2018; Moreno-Martínez et al., 2020). Traditionally, these mono-sensor approaches exploit past and future observations acquired by the same sensor to estimate missing data. Despite numerous relevant spatial and spatio-temporal approaches (Kang et al., 2005; Zhang et al., 2007; Das and Ghosh, 2016; Ding et al., 2017; Vuolo et al., 2017; Moreno-Martínez et al., 2020), local and global temporal interpolation approaches remain the most prevalent methods.

Local temporal methods exploit the temporal evolution of the time series by using a sliding temporal window. Among these approaches, the classical linear interpolation method is the most well-known and used (Inglada et al., 2017; Defourny et al., 2019; Hubert-Moy et al., 2019; Bolton et al., 2020; Kamir et al., 2020). Polynomial-based strategies have been extensively proposed as alternatives. Some examples are spline interpolation methods (Xu et al., 2017; Meng and Li, 2019), Savitzky-Golay filter-based methodologies (Chen et al., 2004; Jönsson and Eklundh, 2004; Kandasamy et al., 2013; Julien and Sobrino, 2019) or locally weighted scatterplot smoothing (Moreno et al., 2014) method. The accuracies of these methods are directly influenced by the sliding window's size, which is a predefined parameter related to the gap length.

Instead of working at the local scale, global temporal methods propose to recover missing information by fitting the data to predefined functions. For instance, the widely used Whittaker smoother fits the time series by minimizing penalized weighted spline regression squared errors (Atkinson et al., 2012; Kandasamy et al., 2013). Further global approaches include asymmetric Gaussian function fitting (Jonsson and Eklundh, 2002; Beck et al., 2006) or Fourier-based harmonic analysis (de Wit and Su, 2005; Zhou et al., 2015; Julien and Sobrino, 2019; Solano-Correa et al., 2020). One of the main limitations of global strategies is that they generally assume that the data follows some *a priori* distribution shape. They are therefore class-specific, resulting in a lack of flexibility in the presence of non-stationary data (Chen et al., 2004; Moreno-Martínez et al., 2020).

One of the main weaknesses of the standard interpolation methods is their poor effectiveness when large data gaps are occurring. In these situations, these methods fail in reconstructing temporal trajectories describing high frequency variations. It can result in missing crucial information about vegetation changes described by such variations. The effectiveness of standard interpolation approaches directly depends on the valid number of observations acquired by the used sensor. Consequently, these methods have been mostly applied to dense optical time series (e.g., MODIS or SPOT-VEGETATION) with coarse spatial resolutions or for tasks requiring limited temporal information, e.g., yearly land-cover classification (Cai et al., 2017; Sun et al., 2021). The temporal resolution of time series acquired at high spatial resolution (e.g., Landsat or Sentinel) is usually less dense. In this case, missing data periods can range from weeks to months (Roy et al., 2008). To address such limitations, some interpolation methods consider the fusion of complementary optical data (Gao et al., 2017; Claverie et al., 2018; Dwyer et al., 2018; Dong et al., 2020; Griffiths et al., 2020; Moreno-Martínez et al., 2020). Unfortunately, these multi-sensor methods require important corrections to homogenize the different spatial (Zhu et al., 2016) and spectral (Barsi et al., 2018; Bolton et al., 2020) resolutions. Additionally, complementary optical data is also affected by cloud coverage and cannot guarantee to provide a high number of supplementary valid observations.

The exploitation of multi-sensor observations for recovering time series of optical-derived vegetation indices has been emphasized by ML regression methodologies (Kamilaris and Prenafeta-Boldú, 2018; Reichstein et al., 2019). An increasing number of works has proposed the use of optical and weather-independent SAR time series (Schmitt and Zhu, 2016). The availability of complementary optical and SAR satellite missions (e.g., Sentinels) has supported their joint exploitation. Three categories of ML regression algorithms approaches can be found in the literature, exploiting Sentinel images: classical ML approaches, Gaussian processes, and deep learning methods.

Classical ML approaches such as Support Vector Machines (SVM) or Random Forests (RF) have been commonly adopted. For instance, the work in (Wang et al., 2019) has proposed to apply SVM and RF algorithms on Sentinel-1, Sentinel-2 and Landsat 8 data to predict frequent Leaf Area Index (LAI) estimations. RF and Support Vector Regression have been used in (Mohite et al., 2020) to generate dense NDVI time series. A six-month time period has been investigated over five different *crop* types. Despite the good results obtained by SVM and RF approaches, it must be noticed that they have been mostly validated on small agricultural datasets composed almost only of *crops* samples. It is therefore difficult to assess whether these techniques could be accurate over large areas with different vegetation covers. In addition, these techniques require a feature extraction step and do not exploit the temporal trajectory of the input time series.

Gaussian process (GP) is another supervised regression method which has been increasingly exploited in several works. Pipia et al. (2019) have proposed a multi-output GP methodology to fill gaps in LAI time series derived from the joint exploitation of Sentinel-2 and Sentinel-1 observations. Besides time series reconstruction, the GP performances have been also corroborated by other regression tasks involving vegetation monitoring. An example is found in (Mercier et al., 2020) where biophysical parameters have been extracted from wheat and rapeseed parcels by exploiting Sentinel-1&2 time series. Although neglecting Sentinel images, the use of GP for *crop* yield estimation is also studied in (Martinez-Ferrer et al., 2020), combining MODIS and SMAP datasets. Despite previous works have shown a satisfactory ability of GP for regression tasks, the scalability of these methods can be challenging. Long training times and significant computational resources are required. The tuning of GP is also complex and very sensitive to the choice of the kernel. Such a choice questions its generalization capacity in the presence of very heterogeneous covers. Finally, likewise classical ML approaches, GP does not exploit the temporal order of the input data used for regression.

In the last decade, deep neural networks have attracted a lot of attention of the scientific community for solving regression problems Lathuilière et al. (2020). The increasing computer capabilities and the availability of large datasets have greatly supported their expansion. Modern neural network architectures have proven to be efficient for time series data mining where there is limited knowledge about the underlying processes. The use of large training datasets describing high variability further improves their generalization ability on unseen data. This capability is essential for handling large geographical areas.

Deep-based SAR-to-optical regression architectures proposed for Sentinel data have first been devoted to exploit the spatial dimension of images acquired at a single date (He and Yokoya, 2018; Cresson et al., 2019; Gao et al., 2020; Meraner et al., 2020). Convolutional Neural Networks (CNN) or Generative Adversarial Networks architectures have been proposed as a single-date regression solution without exploiting the temporal information of time series. The main objective of this work is the regression of SAR data to optical raw spectral bands to fill the missing data. Only a few deep-based architectures have been proposed to exploit the temporal information of high resolution remote sensing time series and almost exclusively for classification tasks (land-cover mapping). Scarpa et al. (2018) have proposed the use of CNNs to estimate NDVI from Sentinel data for dates between May and November. Several scenarios have been investigated to exploit optical and SAR data either separately or jointly. The images preceding or following the date to be estimated have been used to integrate temporal information. While the scenario involving both optical and SAR data as input to their architecture allowed a satisfying estimation of NDVI, in the case of SAR data only, the results were less accurate. The need for non-cloudy optical data can nevertheless be a problem depending on the persistence of the cloud cover. The satisfactory results also highlight the interest of deep learning approaches compared to standard interpolation methods. High regression scores are obtained, albeit the very limited time series used.

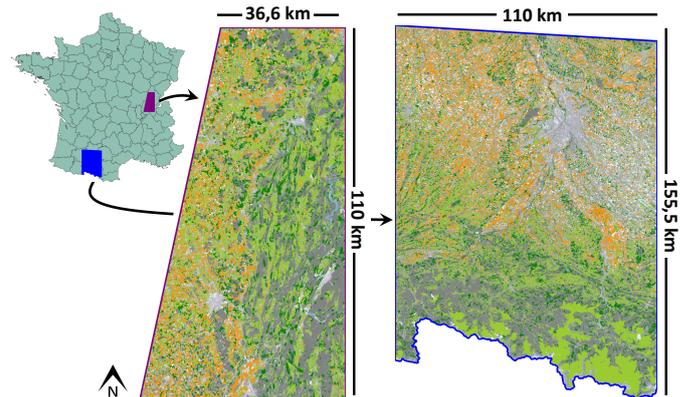
To the best of our knowledge, only Zhao et al. (2020) have proposed a deep-based regression framework based on yearly SAR time series to retrieve optical-derived vegetation indices. The authors combine CNNs and RNNs in a reconstruction framework aiming at predicting gap-free optical NDVI time series. 1D-CNNs have been used to extract

information from SAR time series separately for VV and VH polarizations. RNNs were then exploited to integrate the temporal evolution of the time series. Promising results have been presented by this latter approach. Nevertheless, its generalization capability over heterogeneous vegetation types and large areas can be questioned. Indeed, the work proposed in Zhao et al. (2020) has solely focused on *crops* and a limited spatial extent. *Crops* are known to have a clearly defined phenological cycle. This is especially true on restricted areas, where agricultural practices take place at the same time and phenological cycles are similar. Consequently, the limited geographical area and the lack of class variability do not allow to conclude on the potential accuracy in the context of larger areas and more complex vegetation types. Furthermore, high calculation costs and long training times, which may be challenging in an operational context have been reported, due to the pixel-wise approach. Finally, the input data in Zhao et al. (2020) seems restricted: only descending orbit Sentinel-1 acquisitions have been used, discarding the ascending orbit which could provide additional information to the regression process. Concerning Sentinel-2, cloud & shadow masks have been used to select only non- or partially-cloudy images for both training and validation. The predictions are thus made on dates that have been extensively seen during training. This strategy does not allow to assess the approach neither on dates strongly affected by clouds nor on unseen dates. As for the feature used, only SAR backscatter coefficient information is selected. Additional features such as coherence information, which embeds temporal information and has demonstrated its usefulness for the study of vegetation could have been considered. Furthermore, Zhao et al. (2020) have reported inferior results when important changes in scattering mechanism (e.g., from diffuse soil scattering to vegetation volume scattering) are occurring. The introduction of ancillary information to support the contextualisation of SAR measurements could improve the results. Moreover, this inclusion could allow to improve the generalization of the regression over larger geographical areas and heterogeneous vegetation covers.

### 3. Study areas and data

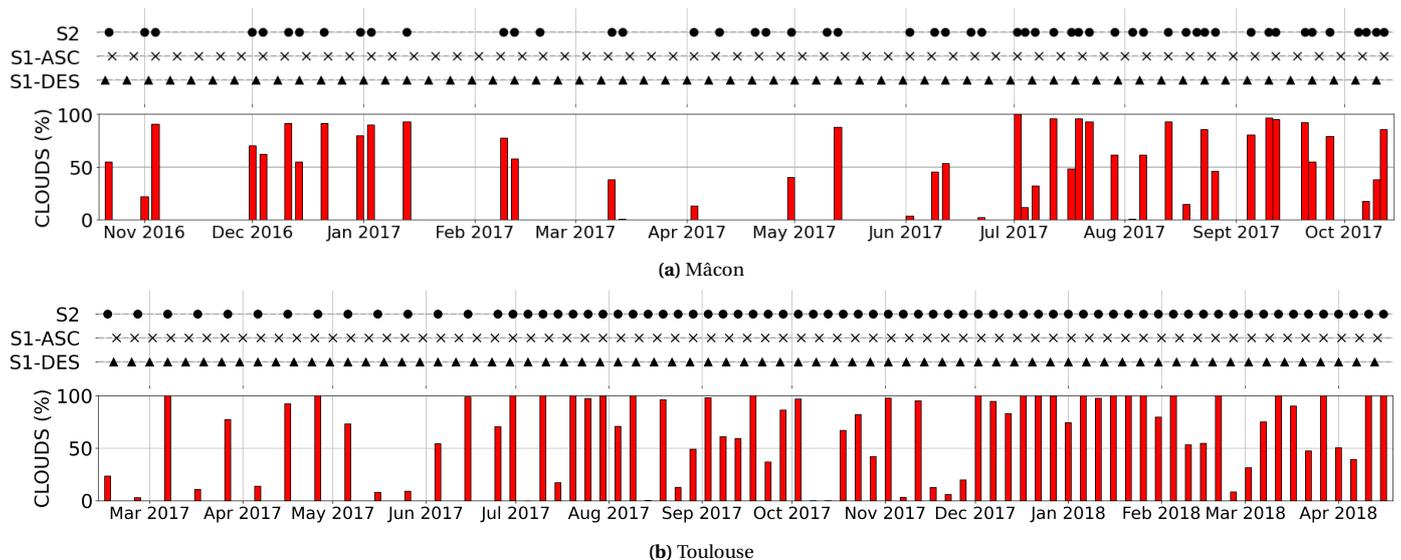
#### 3.1. Study areas

The study is carried out on two geographical areas located in France, as shown in Fig. 1. The study of each area is performed over two distinct time periods.



**Fig. 1** The two study areas (left: *Mâcon*, right: *Toulouse*) with a true-color composition of Sentinel-2 image overlaid by *grasslands* (light-green), *crops* (orange), and *forests* (dark-green).

The first area, denoted as *Mâcon*, covers the Sentinel-2 T31TFM MGRS tile. Agricultural season from October 2016 to October 2017 is studied. Consequently, only Sentinel-2A data is considered for the first half of the studied period, until Sentinel-2B was available (see Fig. 2). The study site is thus defined over overlapping swaths of two adjacent orbits. It allows a temporal resolution of 6 days despite the momentary



**Fig. 2** Temporal distribution of satellite acquisitions for the two study areas. Sentinel-2 (dots), Sentinel-1 ascending (crosses), and descending (triangles) orbits are shown in the upper part of each plot. Percentage of polygons affected by clouds (at least one pixel) for each Sentinel-2 acquisition is indicated by red bars.

availability of only Sentinel-2A satellite. This area of 5,328 km<sup>2</sup> is located in Eastern France and is depicted by a varied topography. A river valley, hills, and plateaus characterise the area.

The second area, denoted as *Toulouse*, is located in the South-West of France and straddles two Sentinel-2 MGRS tiles (T31TCJ and T31TCH). The area covers 15,120 km<sup>2</sup> and it is characterized by different landscapes. Plains regions are located in the North whereas the Pyrenees mountains cover the southern part. The time period under consideration is from February 2017 to April 2018. Sentinel-2A&B full revisit capacity is therefore available from July 2017.

### 3.2. Vegetation classes and reference data

Reference data is composed by a large set of polygons which are obtained from two governmental databases. Three vegetation classes are considered:

- *Grasslands* are composed of 19 and 21 subclasses for the respective *Mâcon* and *Toulouse* areas. This class is composed by some sub-classes such as permanent meadows or pastures, fodder legumes, or temporary herbaceous areas.
- *Crops* are composed of three major cereals cultivated throughout both areas: maize, winter wheat and winter barley.
- *Forests* include closed-canopy deciduous and coniferous forests which belong to different species. The resulting datasets contain 11 subclasses for both studied areas.

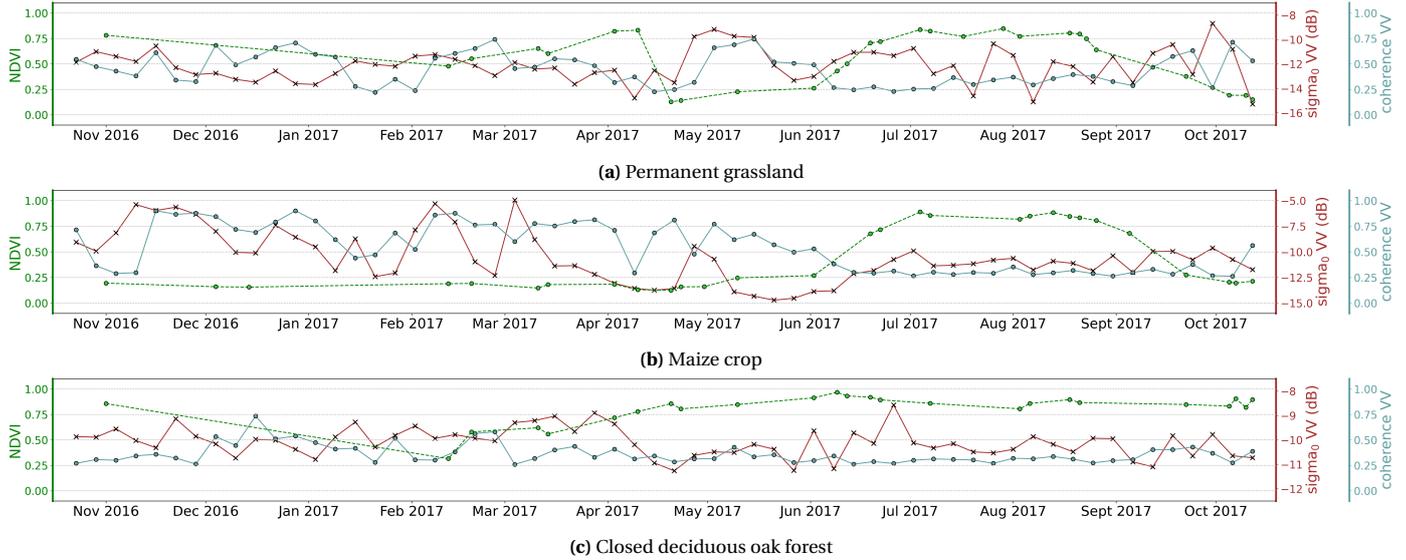
The first database used is the yearly updated French Land Parcel Identification System (LPIS). For each parcel, this database describes the main cultivated species according to its farmer's declaration. The LPIS is obtained under the framework of the Common Agricultural Policy (Campos-Taberner et al., 2019; Sitokonstantinou et al., 2020). *Grasslands* and *crops* polygons, corresponding to parcels, are selected from the LPIS for both areas. The resulting polygons are filtered to remove non-agricultural surfaces within their boundaries. This filtering task removes pixels identified for example as isolated trees, ponds or ditches. The second database is the BD FORET (produced by the French Mapping Agency, IGN), which delineates *forest* areas and provides semantic information on the dominant species. For this database, only polygons having sizes ranging from 4 to 40 hectares are considered. This consideration permits to balance the size of the *forest* polygons with respect to the *grassland* and *crop* ones. The specific details of the reference data are described in Tab. 1.

**Tab. 1** Spatial statistics for *grassland*, *crop*, and *forest* polygons, which are obtained for the two *Mâcon* and *Toulouse* study areas.

Area	Class	Polygons	Area size (km <sup>2</sup> )	Total polygons
<i>Mâcon</i>	<i>Grasslands</i>	27,832	1274,9	46,001
	<i>Crops</i>	12,557	594,9	
	<i>Forests</i>	5,612	579,2	
<i>Toulouse</i>	<i>Grasslands</i>	50,103	2758	98,203
	<i>Crops</i>	34,504	1870,4	
	<i>Forests</i>	13,596	1177,4	

### 3.3. Satellite and ancillary data

**Satellite data** includes optical Sentinel-2A&B and SAR Sentinel-1A&B images. The temporal distribution of their acquisitions is shown in Fig. 2, where information about cloud coverage is also displayed. The cloud cover percentage of the reference polygons illustrates how optical time series may contain many missing values. Optical data was retrieved from the THEIA web portal (<https://theia.cnes.fr/>, accessed: 2020-11-23). Level-2 (Bottom-Of-Atmosphere) data are produced by the MACCS ATCOR Joint Algorithm processor (MAJA, (Baetens et al., 2019)) developed jointly by the French and German space agencies (CNES and DLR, respectively). The resulting products contain a cloud and shadow mask, generated for each acquisition through a multi-temporal approach. Between October 2016 and October 2017, 53 Sentinel-2 images were recovered for the *Mâcon* area. For *Toulouse*, 73 images were downloaded between February 2017 and April 2018. The images have 13 spectral bands with three different spatial resolutions. In this study, only the 10×10 m red and near-infrared bands are kept to compute NDVI time series profiles. SAR data are obtained from the CNES PEPS platform (<https://peps.cnes.fr/>, accessed: 2020-11-26). Time series of radar backscatter coefficients, denoted  $\sigma_0$ , are computed from Ground-Range-Detected (GRD) products. For the *Mâcon* area, 60 ascending orbit and 60 descending orbit GRD products are collected. GRD products gathered on the *Toulouse* area consists of 71 ascending orbit and 71 descending orbit acquisitions. A cross-ratio band between VV and VH polarizations is additionally calculated for each date. The interest of this cross-ratio band is its good correlation with vegetation biophysical parameters (Veloso et al., 2017; Vreugdenhil



**Fig. 3** Yearly time series of optical and SAR features retrieved from Sentinel’s for the three studied vegetation classes. Green dashed-line corresponds to the cloudless NDVI time series, brown line to  $\sigma_0$  VV polarization in dB, and light-blue line to 6-days coherence VV polarization time series.

et al., 2020). All  $\sigma_0$  bands are converted from linear scale to logarithmic decibels (dB). For the same dates, interferometric coherence time series were computed from Single-Look-Complex (SLC) products. Coherence measurements are obtained every 6 days calculated from following pairs of images, independently for ascending and descending orbits. As the SLC products are delivered with a pixel spacing of  $2.3 \times 14.1$  m (range and azimuth, respectively), a  $9 \times 3$  window (range and azimuth, respectively) is used for the coherence calculation. The processing of the GRD products to  $\sigma_0$  and SLC products to coherence images is done using the freely available SNAP software (<http://step.esa.int>, accessed: 2020-11-23). All images are then orthorectified using Range Doppler Terrain Correction algorithm and the Shuttle Radar Topographic Mission (SRTM) 1 Arc-Second corresponding elevation data. Resulting images have an output spatial resolution of  $10 \times 10$  m matching the optical ones.

**Ancillary data** integrates climatic and topographic data. Climate data are extracted from the Météo France SAFRAN dataset (Quintana-Seguí et al., 2008; Vidal et al., 2010). This dataset gathers measurements from all climate stations of the French national meteorological service (more than 500 stations across the country). Data are provided daily and measure 25 variables such as solid and liquid precipitation, average daily temperature, radiation, evapotranspiration, soil moisture index, or drainage. A point grid evenly spaced 8 km apart is used for the spatial sampling of the data. Each point gathers data from the nearest meteorological station. The large number of climate stations

allows to ensure data variability despite the low spatial resolution of the grid.

Topographic data are retrieved from  $5 \times 5$  m spatial resolution Digital Terrain Models (DTMs) acquired by IGN. Terrain height, slope, and exposure bands are computed from the DTMs. Reference data and the DTMs are perfectly co-registered with the Sentinel images. IGN-France is indeed responsible for providing the georeferencing grid of Sentinels over Europe.

### 3.4. Satellite data pre-processing

The prediction of NDVI measurements from SAR time series is framed here as a supervised learning problem. Thus, a common temporal grid is defined for the optical and SAR datasets. This target grid is the ascending or descending SAR time series (e.g., crosses or triangles in Fig. 2) having the lowest average time lag (in days) with respect to optical acquisitions. Optical and remaining SAR data are then re-sampled to this temporal grid using the nearest neighbour technique. The resulting optical and SAR time series contain 60 dates for *Mâcon* and 71 dates for *Toulouse*. Polygons described in Tab. 1 are then eroded with an internal buffer of 10 m to remove border pixels. The satellite information of the resulting polygons is extracted at the object level. The object-oriented strategy was chosen for three reasons: 1) the reference data permit the delineation of relatively homogeneous clusters (parcels) which reduces possible inconsistencies in measurements (Atzberger, 2004), 2) pixel-wise analysis would require further

**Tab. 2** Summary of features computed at polygon level for Sentinel-1 and ancillary data.

Sentinel-1 features computed for descending and ascending orbits		
(1)	Sigma <sub>0</sub> dB	mean, median and standard deviation for VV, VH and VV/VH bands from Sentinel-1 GRD
(2)	Coherence	mean, median, standard deviation for VV and VH bands from Sentinel-1 SLC
(3)	Derivative features	first order derivative computed for the previous (1) and (2) mean features
(4)	Neighborhood features	features in (3) are averaged on the polygon neighborhood within a specific radius
Features computed from ancillary data		
(5)	Topographic features	mean and standard deviation for height, slope, exposure bands. Area, perimeter, and size are also provided
(6)	Climatic features	day of SAR acquisition and day before from the nearest weather station
(7)	Metadata features	temporal distribution of satellite acquisition and LPIS subclass

SAR processing as it suffers from the speckle noise, 3) the computational and storage challenges associated to the high data volume (Atzberger, 2013; Inglada et al., 2017; Mallet and Le Bris, 2020) can be avoided.

From Sentinel-2 data, the information extracted from each polygon is the average NDVI. These values are used to train and validate the proposed supervised deep learning methodology. Each NDVI value has its corresponding validity flag  $\in [0, 1]$  which is computed by considering MAJA’s cloud & shadow mask information. The cloud & shadow mask information is used in the most restrictive way: if one pixel of the polygon is affected by a cloud, the polygon is considered as cloudy.

From Sentinel-1 data, the two categories of features summarized in Tab. 2 are computed for each polygon. The first category describes the statistics computed on the different polarization bands of Sentinel-1 dataset and is denoted as (1) and (2). The statistics correspond to the mean, median, and standard deviation values calculated on the processed GRD and SLC images. For each date, the resulting datasets contain 9 features describing the statistics of  $\sigma_0$  bands and 6 features from the coherence bands. The second category of Sentinel-1 features corresponds to the datasets (3) and (4) of the Tab. 2. These features provide information about the first order derivatives computed on the time series considered in (1) and (2). The features in (3) describe the statistics previously computed on (1) and (2) for the first order derivative between date  $t$  and date  $t_{-1}$ . In order to incorporate information about the polygon neighbourhood, the set of features in (4) is also proposed. The dataset (4) contains then the average of features computed on (3) on a specific neighbourhood. The polygon neighbourhood is defined by all the polygons belonging to the same vegetation class inside a given pre-defined radius of 2 km. The goal of features in (4) is to highlight if a specific polygon has a diverging behaviour compared to its neighbourhood (Ding et al., 2017).

The above described features derived from satellite data are computed for the different studied classes. Fig. 3 illustrates three examples of feature-based time series derived from satellite data. For each example, NDVI,  $\sigma_0$ , and coherence (VV polarization and ascending orbit) are shown. Different phenology dynamics are observed for the three vegetation classes across the temporal dimension. For instance, *crops* NDVI response curve follows a double logistic function as described in (Jonsson and Eklundh, 2002). In contrast, no specific shape is observed for the *grasslands* NDVI response curve. Concerning *forests*, it can be observed how the temporal NDVI signal is relatively stable describing seasonal trends. Fig. 3 also shows that the interpretation of SAR time series is not straightforward given their strong fluctuations. For instance, at the beginning of the year until the sowing crop period in May,  $\sigma_0$ , and coherence time series of Fig. 3b exhibit low stabilities. In this case, the presence of bare soil leads to a high surface roughness and diffuse scattering and important response variations. Besides drawing phenological cycles, Fig. 3 also corroborates that the features time series can be used to monitor anthropic activity. For instance, a mowing act is observed in Fig. 3a at the end of April. The agricultural practice results in different responses among the features. A drastic drop is observed in NDVI response, whereas the act results in a marked increase of SAR coherence signal.

As shown in Tab. 2, object-level information is also extracted from ancillary data. Additional features (corresponding to (5) and (6) in Tab. 2) are integrated into the network architecture as an input information. For each Sentinel-1 acquisition date and the previous day, 25 climatic variables are collected. Information about the previous day of the acquisition is incorporated to take into account the morning schedule of some SAR acquisitions as well as rain accumulation. Concerning topographic information, 9 features are derived to incorporate polygon geometric characteristics alongside SAR measurements (O’Grady et al., 2013; Nasrallah et al., 2019). Working at the polygon level, mean and standard deviation values are computed from the height, slope, and exposure bands. Additionally, polygon shape features are also considered with area in hectares, perimeter, and the number of Sentinel pixels.

Finally, the last category of ancillary features comes from two types of metadata information (features (7) in Tab. 2). The first one concerns the temporal information about Sentinel’s acquisitions. The days of

the year of Sentinel-1 ascending and descending acquisitions are provided. Temporal differences in days between Sentinel-1 acquisitions and their associated Sentinel-2 NDVI observations are taken into account. This last information is considered to reduce the impact of the nearest neighbouring resampling pre-processing step used to unify the temporal grid. The subclass of each polygon described in the LPIS and BD FORET databases is proposed as second metadata information.

## 4. Method

### 4.1. Regression task and loss function

Let us denote as  $X = (x_1, x_2, \dots, x_T)$  the multivariate time series of length  $T$  containing all features derived from Sentinel-1 (see Tab. 2). For each  $t \in \{1, 2, \dots, T\}$ , SAR features derived from an image acquired at instant  $t^{th}$  are represented by  $x_t$ . In parallel,  $Z = (z_1, z_2, \dots, z_T)$  is the multivariate time series where each  $z_t$  contains the features computed from ancillary data, tailored to provide information about SAR measurements (Benninga et al., 2019). Considering these definitions,  $\{X, Z\}^{asc}$  and  $\{X, Z\}^{desc}$  correspond to the couples of features from ascending and descending orbits. SAR orbits are individually processed to avoid mixing information acquired from different viewing angles. The proposed SenRVM regression method then uses both couples to predict the time series  $\hat{Y} = (\hat{y}_1, \hat{y}_2, \dots, \hat{y}_T)$  where  $\hat{y}_t$  denotes the predicted NDVI measure at the instant  $t^{th}$ .

The proposed supervised SenRVM methodology is based on a deep learning recurrent architecture. As described in Sec. 4.3, our methodology mainly relies on two types of neural networks which are briefly introduced in Sec. 4.2.

To supervise SenRVM, NDVI time series  $Y = (y_1, y_2, \dots, y_T)$  of length  $T$  derived from Sentinel-2 acquisitions are used. The training process estimates the network parameters by minimizing a loss function  $\mathcal{J}$ . This function quantifies the error  $\mathcal{L}$  between predicted and expected NDVI values. Given  $n$  training samples,  $\mathcal{J}$  is defined as the average Mean Squared Error (MSE)  $\mathcal{L}$  and is committed during the forward training propagation step as:

$$\mathcal{J} = \frac{1}{n} \sum_{i=1}^n \mathcal{L}(\hat{Y}_i, Y_i), \quad (1)$$

where  $\mathcal{L}$  evaluates the average error between prediction  $\hat{y}_t$  and expected  $y_t$  values as instant  $t$ . The MSE is preferred to other classical regression loss functions Lathuilière et al. (2020) given its ability to converge towards the optimal solution. Note that the Mean Absolute Error outputs continuously large gradients even in the case of small errors, which can lead to convergence problems. The alternative Huber loss requires the setting of a hyper-parameter delta. The setting of this value can be data dependent and is an iterative problem.

To take into account the sensitivity of MSE to outliers, cloudy NDVI acquisitions are removed from the MSE computation. The information contained in the cloud & shadow masking vector  $M = (m_1, m_2, \dots, m_T)$  is incorporated in Eq. 1 with the validity flag  $m_t \in \{0, 1\}$  associated with each  $y_t$ :

$$\mathcal{L} = \frac{1}{T} \sum_{t=1}^T m_t (\hat{y}_t - y_t)^2. \quad (2)$$

It must be noticed that the presence of outliers in the training data cannot be entirely discarded since cloud & shadow masks can contain errors. The performance impact of such errors is discussed in Appendix A where a re-training/refinement strategy is presented to slightly improve the SenRVM results.

### 4.2. Multilayer Perceptron and Recurrent Neural Networks

A Multilayer perceptron (MLP) is an artificial neural network (Atkinson and Tatnall, 1997) based on the formalisation of the perceptron (Rosenblatt, 1958). MLPs are feedforward networks, meaning that each

layer output is fed into the next layer. Traditionally, MLPs are used to map inputs into an output representation describing the complex relations among the data, which is known as the *encoding* task. Conversely, they are used for *decoding* tasks, which intend to output the closest match from the given input to the intended output.

MLPs are composed of an input layer, one or more hidden layers, and an output layer. During the training process, the weights and biases of the network are iteratively adjusted by minimizing a cost function. To perform it, the most common technique is to use the back-propagation algorithm which calculates the gradient of the error function with respect to each weight. Stochastic gradient descent and its variations such as AdaGrad, RMSprop, or Adam are commonly used for updating weights to minimize the loss. By means of nonlinear activation functions (e.g., Sigmoid, ReLU, or Tanh), a multi-layer network introduces nonlinear relationships in the data, which is essential to effectively back-propagate the model's error in the training phase. To improve the generalization of training and prevent overfitting (*i.e.*, fitting too closely to the data used for training), regularization techniques such as dropout or normalization layers are routinely used. Dropout limits the interdependence between neurons in a layer by randomly ignoring neurons during training. Normalization helps to stabilize the training accuracy when data have different ranges.

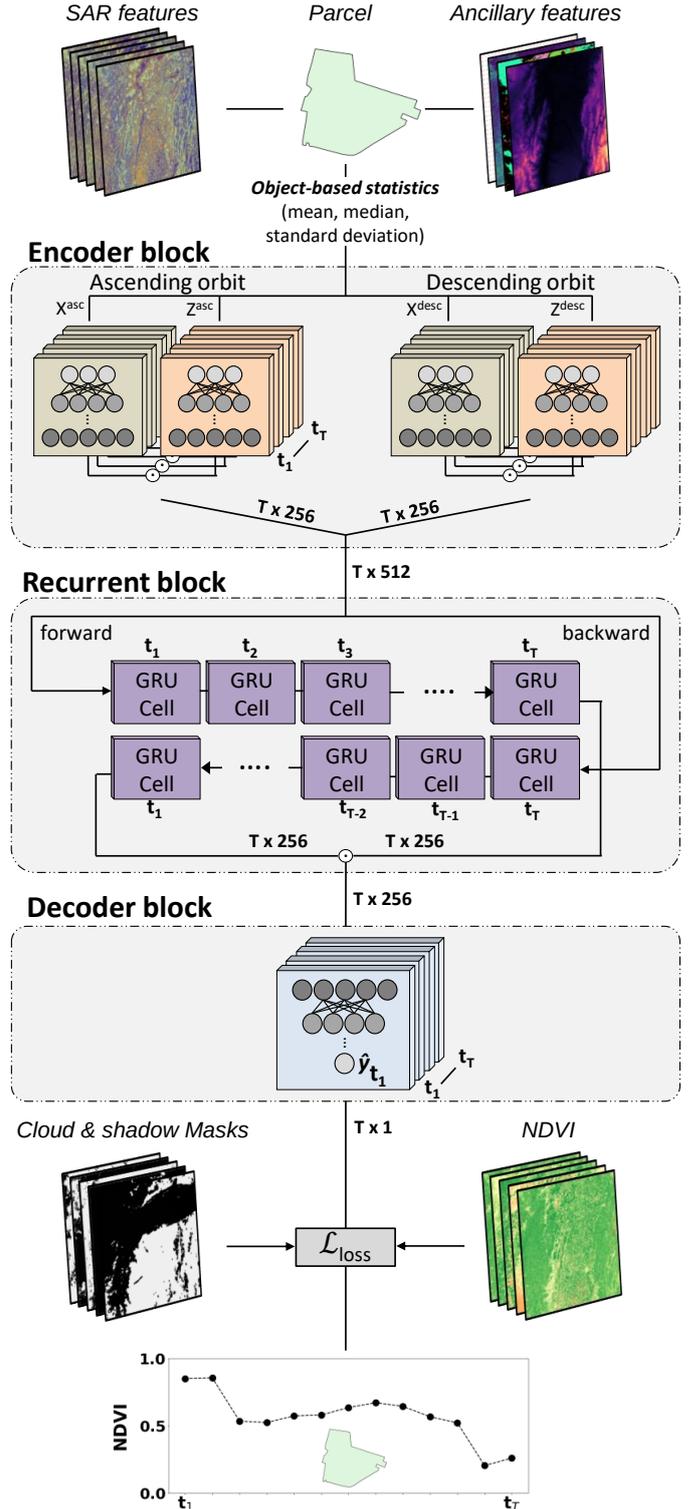
Recurrent neural networks (RNNs) are capable of learning features and long-range dependencies from sequential data (Campos-Taberner et al., 2020). Compared with MLP, RNN breaks the imposed feedforward neural network rule where information is processed unidirectionally. RNNs have recurrent layers where the neurons composing a layer are interconnected in both directions of the network. These recurrent connections are used for dynamic information processing and are naturally suited to satellite time series.

Bi-directional RNNs (BRNN) process the sequence in both directions. Typically, two separate RNNs are used: one for the forward direction and one for the reverse direction. Two common variants of RNN include Long Short-Term Memory (LSTMs, (Hochreiter and Schmidhuber, 1997) units and Gated Recurrent Units (GRUs) (Cho et al., 2014). These variants propose the use of gating mechanisms to control the memorization process along the sequence. GRUs are similar to LSTMs networks but with a simpler architecture (*e.g.*, one gate less). This reduces the network complexity and allows for significant decrease in network parameters and computing times, while providing similar results (Ndikumana et al., 2018).

### 4.3. SenRVM architecture

The Sentinels Regression for Vegetation Monitoring (SenRVM) architecture is decomposed into three blocks as depicted in Fig. 4. Firstly, the encoder block combines SAR and ancillary data to extract a joint complex representation. Secondly, the recurrent block captures the temporal dependencies among the previous representations. Finally, the decoder block translates the networks representation into the target variable.

The encoding block individually processes ascending and descending orbit datasets by two parallel branches, which are fed by  $\{X, Z\}^{asc}$  or  $\{X, Z\}^{desc}$  couples. Each branch is composed of two MLPs, which separately maps each  $\{x_t, z_t\}$  couple to an output representation of 256 features. The four MLPs composing the encoder block contain 4 fully connected (FC) layers whose output sizes are equal to 128, 128, 200 and 256. FC layers, except for the last one, are followed by batch normalization, dropout, and non-linear activation layers. As described in Sec. 4.2, these layers are used to improve the accuracy and generalization of the training. A fixed probability  $p$  for the dropout layers is set to 0.2 and a rectified linear unit (ReLU) is used as non-linear activation function. The use of MLPs allows us to obtain a feature representation with the same output size (256). This intermediate representation encodes complex relationships for each data modality and allows their fusion without prior correlation knowledge. An element-wise multiplication (Hadamard product) is proposed to combine the encoded SAR and ancillary feature representations computed for each acquisition date. Alternatives such as concatenation, summation, or subtraction are discarded because lower results were obtained for these strategies in preliminary



**Fig. 4** SenRVM model overview. It targets to regress SAR and associated ancillary data into optical-like vegetation index (here NDVI). The size of the vectors is also indicated.

experiments. Finally, the encoding block concatenates the two ascending and descending branches in a single vector of 512 features for each date  $t$ . The concatenation is chosen here to keep both orbits information.

The second block is a single layer BRNN, scanning the input in both directions. The choice of a bi-directional RNN is made to obtain robust predictions using past and future observations. A single BRNN layer is proposed as it achieves similar performances than more complex multi-layers BRNN while reducing the number of network parameters. This block takes as input the outputs of the encoder block (*i.e.*, the joint SAR and ancillary representation). The RNN is composed of GRU cells which have a memory size equal to 256. Working with short-length satellite time series, GRU has been preferred to LSTM. GRU cells carry two memories (update and reset gates) allowing the model to determine which information is transmitted or removed from hidden state  $t$  to hidden state  $t_{+1}$ . The purpose of this block is to extract the underlying temporal information contained in the time series described for each  $t$  by 512 features. The bi-directional RNN outputs two vectors corresponding to forward and backward scanning. These two vectors are element-wise multiplied to obtain a single vector of 256 values per date which is passed to the next block. The recurrent block outputs a time series of size  $T \times 256$ .

For each date  $t$ , the last SenRVM block processes its corresponding 256 features to predict  $\hat{y}$ . This last block is composed of a funnel-shaped MLP which is a succession of six FC layers whose input sizes are equal to 256, 64, 32, 16, 8, and 4. As previously proposed, FC layers are followed by batch normalization, dropout, and ReLU activation layers. Dropout probability is here initially set to 0.4 and decreases by 0.1 for the successive following layers. Finally, the last FC layer applies the sigmoid function as the final activation function, considering the dynamic range of NDVI values. The decoding block regresses the 256 features to a single NDVI value  $\hat{y}$  for all dates  $t \in T$ .

SenRVM network weights are learned through the training phase by minimizing the loss function of Eq. 1. Iterative backpropagation with Adaptive Moment Estimation (Adam) algorithm (Kingma and Ba, 2017) is used to find the optimal network weights. After each iteration, Adam algorithm updates the weights towards the global minimum of the loss function.

## 5. Results

Multiple experiments are carried out to evaluate the performances of the SenRVM architecture. The accuracy of the predicted NDVI time series at high temporal coverage is first investigated over a multi-class dataset including three types of vegetation. In a second step, single-class models are studied by analyzing the specific class results considering different spatial and temporal criteria. The importance of the different input SenRVM features is then evaluated by an ablation study. To further assess the SenRVM performances, the proposed methodology is compared with standard interpolation and ML regression methodologies. The SenRVM prediction accuracy is evaluated on small and significant data gaps.

### 5.1. Experimental setup

All experiments are carried on a Intel(R) Xeon(R) Silver 4116 CPU with 12 cores @2.10 GHz with 32 GB of RAM and a NVIDIA GeForce RTX 2080 Ti GPU with 11 GB GDDR6 memory. Reference data is randomly splitted into disjoint train, validation and test data subsets. The training dataset contains  $3/5^{th}$  of the polygons describing the complete reference dataset. The remaining polygons are divided equitably for the validation and test datasets. A 5-fold cross-validation is performed by repeating the splitting procedure five times. Hereinafter presented results are obtained by averaging the 5-fold results.

The SenRVM input features are described in Tab. 2. Each dataset is standardized by its mean and standard deviation to homogenize the different ranges. For the different experiments, the batch size (Bs) and learning rate (Lr) are empirically set. During the training process, the number of epochs is set to 150. The validation dataset is used to select the epoch having the lowest loss result.

Four classical regression metrics are used to assess the SenRVM performances. The coefficient of determination  $R^2$  is used to estimate

how strong the linear relationship is between the expected and predicted NDVI measures. Prediction errors are evaluated by the Mean Absolute Error (MAE, with Absolute Error =  $|y_t - \hat{y}_t|$ ), Mean Squared Error (MSE), and the Root Mean Squared Error (RMSE).

### 5.2. Evaluation of the multi-class SenRVM model

The performances of a multi-class SenRVM model trained on a large dataset are evaluated here. The multi-class training dataset is obtained by merging the three (*grassland*, *crop* and *forest*) vegetation training datasets (see Tab. 1). For *Mâcon*, 27, 599 polygons are used for training the models while 58, 921 are used for the *Toulouse* area. As previously described in Sec. 3.3, the three vegetation classes have distinct phenologies, which leads to a high data variability. For both study areas, the learning rate and batch size are set to 0.0005 and 512 respectively. Training times for the models are 71 minutes for the *Mâcon* area and 145 minutes for *Toulouse*.

Results obtained for the multi-class SenRVM model are shown in Tab. 3. Results are averaged over all polygons belonging to the test dataset and for all dates of the time series. Highly accurate results are obtained, with  $R^2$  above 0.86 and MAE errors below 0.042. Low standard deviations are also found across the four metrics for both areas. This shows the ability of the SenRVM network to provide good predictions despite the different phenologies of the vegetation types.

**Tab. 3** Average results and  $\pm$  standard deviation obtained by multi-class SenRVM models.

		<i>Mâcon</i>	<i>Toulouse</i>
Multi-class	$R^2$	0.8650 $\pm$ 0.0039	<b>0.8947</b> $\pm$ 0.0016
	MAE	0.0419 $\pm$ 0.0016	<b>0.0404</b> $\pm$ 0.0016
	MSE	0.0039 $\pm$ 0.0002	<b>0.0030</b> $\pm$ 0.0002
	RMSE	0.0628 $\pm$ 0.0021	<b>0.0547</b> $\pm$ 0.0017

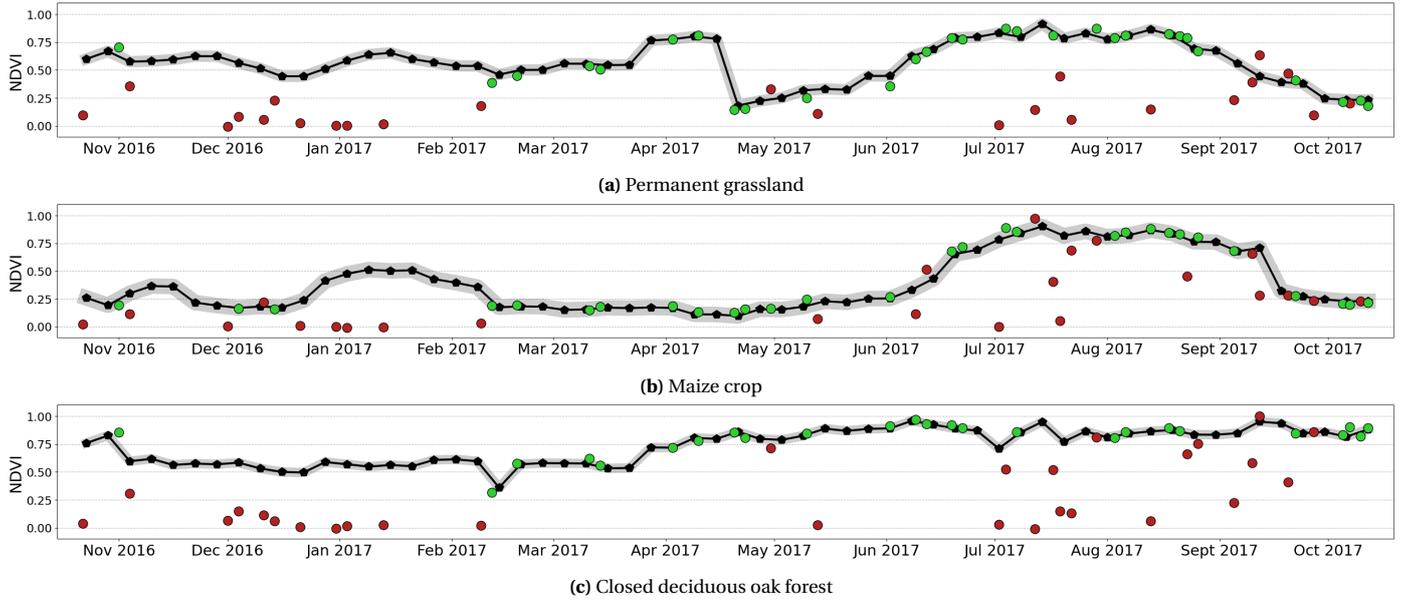
The best performances are reached on *Toulouse*. Nonetheless, the differences found between both areas are relatively small. The differences can be justified by several reasons. Firstly by a difference in landscapes between the two areas, with large parts of plains marked by a little topography for the *Toulouse* area. The *Mâcon* area, on the other hand, has a more spatially distributed topography. Because of the side-looking nature of SAR data, this can lead to data being masked by topography (see Sec. 5.3.1). The temporal distribution of satellite acquisitions, with more dates for the *Toulouse* area can also be beneficial. Lastly, the uneven number of training samples allows the models of *Toulouse* to integrate for learning about double the number of polygons of *Mâcon*.

Fig. 5 shows the predicted NDVI time series for the three validation polygons previously studied in Fig. 3. As observed, the NDVI time series obtained by SenRVM follow the 6-day temporal sampling of Sentinel-1. The figure visually assesses the results detailed in Tab. 3 and corroborate how the predicted NDVI values fit the expected NDVI measurements. Those results illustrate how SenRVM is capable of predicting dense NDVI time series describing different phenological profiles.

Fig. 5a shows how the *grassland* clear cut previously discussed in Fig. 3a is well retrieved. The interest of having dense NDVI time series for regular vegetation monitoring is also attested by Fig. 5b. In this case, the resulting NDVI allows us to observe the maize growth during the month of June and the harvest mid-September. Such interest is also visible in Fig. 5c where several NDVI measurements are recovered from a large winter data gap. Despite several months without optical acquisitions, the results corroborate how SenRVM allows to predict the stable NDVI response curve expected on this *forest* polygon.

### 5.3. Evaluation of single-class SenRVM models

This study aims to investigate if previous results could be improved by training SenRVM on single-class datasets. Hence, three SenRVM models are individually trained on *grassland*, *crop* and *forest* datasets.



**Fig. 5** Example of three reconstructed time series belonging to the testing dataset. These three examples are previously studied in Fig. 3. Dots are the actual NDVI values calculated on the Sentinel-2 images. Green dots correspond to cloudless acquisitions, whereas red dots indicate a cloudy measurement. The black line is the prediction of the multi-class SenRVM model. Grey surrounding region is the average MAE of predictions for the time series.

**Tab. 4** Training time (T) and main SenRVM hyperparameters: learning rate (Lr) and batch size (Bs).

Area	Class	Lr	Bs	T (min)
Mâcon	Grasslands	0.0005	256	57
	Crops	0.0005	256	32
	Forests	0.001	128	24
Toulouse	Grasslands	0.0001	256	96
	Crops	0.0005	512	61
	Forests	0.001	256	36

Hyperparameters are empirically tuned according to each vegetation class, which are reported in Tab. 4.

Results obtained over *Mâcon* and *Toulouse* are shown in Tab. 5. The minimum  $R^2$  and the maximum MAE are here respectively equal to 0.83 and 0.044. Results of single-class and multi-class models are of the same order of magnitude. The differences found with the multi-class model are highlighted in green (improvement) or red (decrease).

The highest accuracies are reached by the *crops* class for both study areas, which obtain similar results. Conversely, *grasslands* obtains the worst results and the highest standard deviations. This can be explained by its high intra-class variability and the important number of abrupt events impacting them. Despite some  $R^2$  differences are observed between *crops* and *grasslands*, similar results are obtained by the rest of the precision metrics. *Forests* class results obtain the lowest precision errors which are related to the class signal stability discussed in Fig. 3c. Human intervention in *forests* remains rare and their phenology response curve only varies with some specific factors such as climate. Some differences are observed by comparing the *forest* results obtained on the two study areas. The best performances are reached on *Toulouse* which can be justified by a greater number of training samples, approximately double that of the *Mâcon* area.

Even with a high heterogeneity within the training dataset (see Sec. 3.2), the results corroborate that a multi-class SenRVM model can obtain accurate predictions. Furthermore, the use of a large training dataset permits to increase the batch size, which reduces the computational times. Compared to single-class models, training a multi-

**Tab. 5** Class-specific average results  $\pm$  standard deviation, over all the predicted instant of times and for all testing polygons. Differences with results obtained with the multi-class model are marked in green (improvement) or red (decrease).

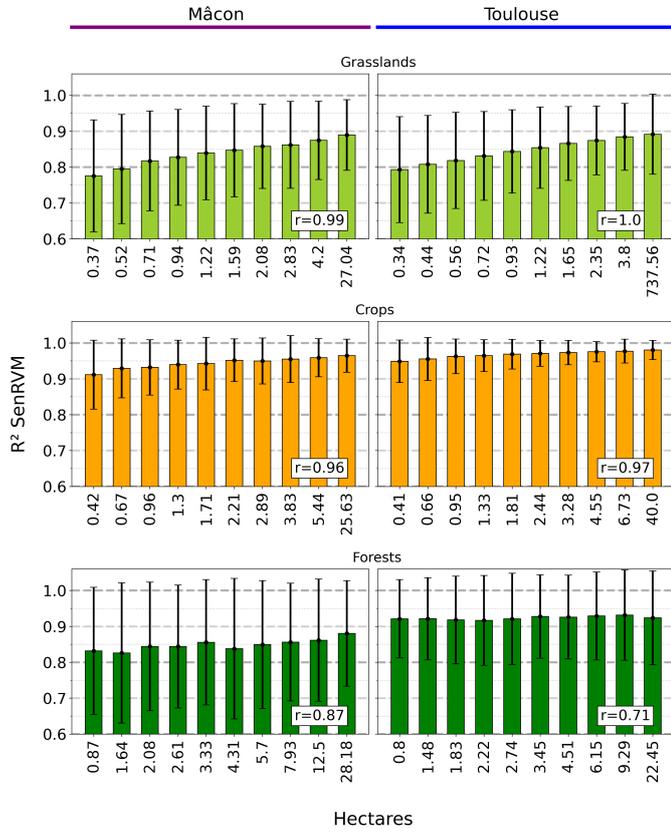
		<i>Mâcon</i>		<i>Toulouse</i>	
Grasslands	$R^2$	0.8384 (+0.0038) $\pm$ 0.0152	<b>0.8464</b> (+0.0089) $\pm$ 0.0115		
	MAE	<b>0.0418</b> (+0.0011) $\pm$ 0.0058	0.0443 (-0.0002) $\pm$ 0.0029		
	MSE	0.0040 (+0.0003) $\pm$ 0.0009	<b>0.0037</b> (-0.0002) $\pm$ 0.0006		
	RMSE	0.0629 (+0.0018) $\pm$ 0.0069	<b>0.0606</b> (-0.0011) $\pm$ 0.0049		
Crops	$R^2$	0.9433 (+0.0033) $\pm$ 0.0017	<b>0.9676</b> (+0.0001) $\pm$ 0.0037		
	MAE	0.0420 (-0.0061) $\pm$ 0.0014	<b>0.0353</b> (-0.0019) $\pm$ 0.0038		
	MSE	0.0040 (+0.0008) $\pm$ 0.0001	<b>0.0026</b> (+0.0001) $\pm$ 0.0005		
	RMSE	0.0630 (+0.0065) $\pm$ 0.0010	<b>0.0503</b> (-0.0001) $\pm$ 0.0044		
Forests	$R^2$	0.8486 (+0.0006) $\pm$ 0.0268	<b>0.9235</b> (+0.0030) $\pm$ 0.0106		
	MAE	0.0343 (-0.0001) $\pm$ 0.0066	<b>0.0318</b> (-0.0017) $\pm$ 0.0051		
	MSE	0.0032 (+0.0002) $\pm$ 0.0007	<b>0.0020</b> (-0.0001) $\pm$ 0.0005		
	RMSE	0.0562 (+0.0013) $\pm$ 0.0057	<b>0.0450</b> (-0.0013) $\pm$ 0.0049		

class SenRVM model can offer some advantages as, for instance, the reduction of the number of parameters to be learned. Besides simplifying the parameter tuning, the reference data scarcity problem which may exist for minority classes can be reduced by training a multi-class model. In addition, the variability of multi-class training data is usually increased. Therefore, multi-class SenRVM models can improve their generalization performance and potentially reduce overfitting, while preserving accurate results.

### 5.3.1. Spatial and temporal assessment

Firstly, a spatial assessment evaluates the effect of size and location of polygons on the reconstruction accuracy. In a second step, the time-dependent accuracy of SenRVM is evaluated on different dates and seasons of the year. The temporal assessment also evaluates how the number of valid Sentinel-2 acquisitions can influence the SenRVM performances.

The polygon area size distribution is studied by ten categories, where each category contains the same number of polygons. In general, the number of polygons for each category is similar for both study areas (see Fig. 6). Only some differences are found for *grasslands* since large summer pastures covering several hundred hectares are present in the Toulouse area.

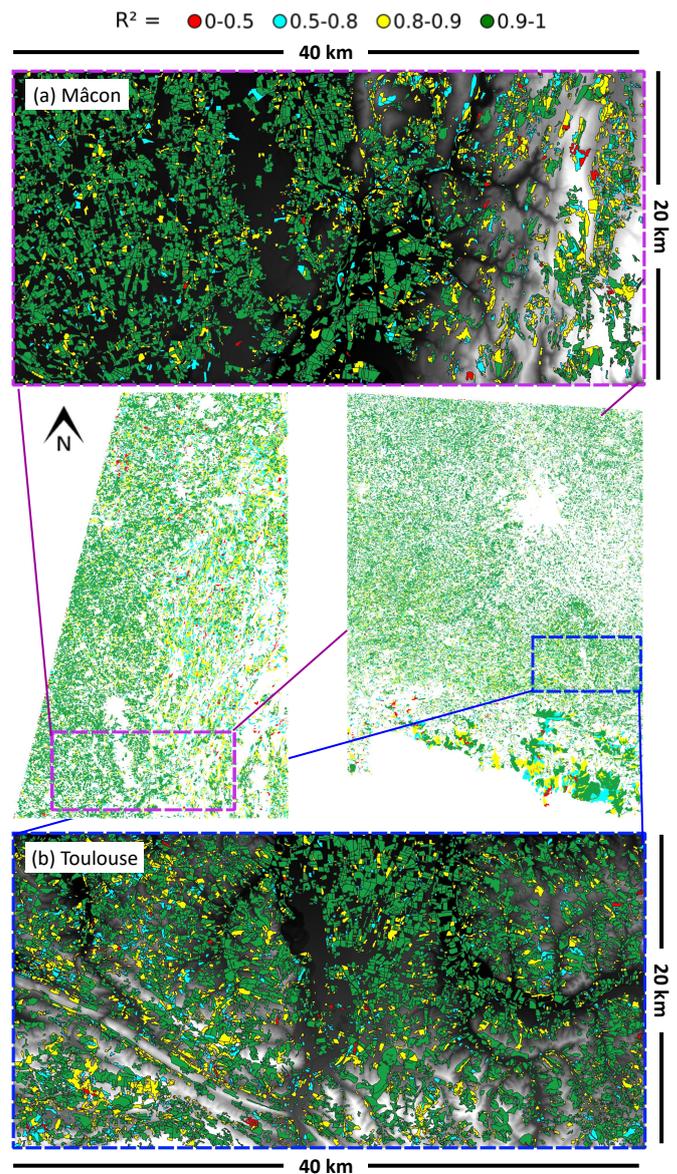


**Fig. 6** SenRVM accuracies evaluated according to the polygon size. Ten polygon sizes classes of equal population are defined for each vegetation class (rows) and both areas (columns). Average  $R^2$  of SenRVM predictions per class and their standard deviation is displayed by bars. Correlation coefficient  $r$  between SenRVM predictions and polygon sizes are displayed on the bottom right corner of each plot.

The previous results obtained at Sec. 5.2 are studied according to the ten categories. Fig. 6 shows the obtained results where for each vegetation class the correlation coefficient  $r$  between polygon size and  $R^2$  score is also displayed. The results show that a strong correlation exists between the polygon size and the prediction performances, with the highest accuracies obtained by the largest polygons. *Grasslands* reach the highest correlation coefficients, which are higher than 0.99 for both study areas. In contrast, the lowest correlation coefficients are obtained on *forests*. The strong correlation obtained on *grasslands* and *crops* can be mainly explained by three reasons. The first is that large polygons contain more pixels and the resulting statistics (*i.e.*, mean, median, and standard deviation used as features) are more stable and less sensitive to outliers. The second reason is the mechanization of farming practices on large agricultural parcels, which leads to the presence of more homogeneous and human-controlled vegetation covers. Time series describing homogeneous vegetation polygons lead to more predictable and reliable results. Lastly, large polygons are less affected by speckle noise given the proposed object-oriented approach. Although similar satisfactory results are obtained for the three classes, the effects of polygon size on prediction accuracies are different. The  $R^2$  difference between small and large polygons is minimal for *crops* and *forests*. In contrast, for *grasslands*, impact of polygon size appears greater with  $R^2$  gains of about 0.1 for larger polygons.

Following the same idea, the effects of the altitude, slope, and exposure of polygons on SenRVM results are also studied. Polygon altitude does not seem to influence the SenRVM performances with correlation coefficients close to 0. Slope and exposure exhibit a significant negative correlation ( $-0.75 < r < -0.6$ ), except for *forests*. In the case of *grasslands* and *crops*, the increase of slope and exposure decreases the prediction quality. Nevertheless, the correlations obtained for these three topographic features are not as significant as the ones presented in Fig. 6. Consequently, these results are not further explored in the following.

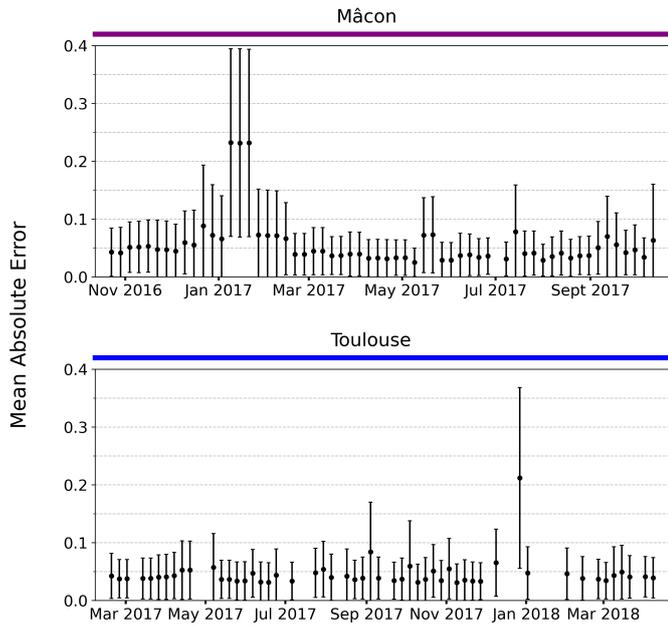
Fig. 7 shows the spatial distribution of the  $R^2$  results obtained by SenRVM for both study areas. A four-color map is used to evaluate the defined  $R^2$  scale ranges. Results show that high relief areas obtain the worst SenRVM performances. This can be observed by looking at the eastern part of *Mâcon* and the Southern part of *Toulouse*. The zoom box of each area is superimposed with a DTM. Light colours describe high altitude areas, where an accuracy decrease is observed



**Fig. 7** Spatial visualisation of the SenRVM performance obtained on all vegetation polygons. DTM is displayed as base-map in the zoom boxes. Four scale  $R^2$  ranges are used to evaluate SenRVM predictions which exhibit less accurate prediction in mountainous areas (East part of *Mâcon* and South part of *Toulouse*).

in mountainous areas. Polygons located on mountain regions can suffer from non-exploitable Sentinel-1 data (*i.e.*, due to geometric effects such as layover or foreshortening). Furthermore, as these high relief parts also have strong slopes and exposures, this visual assessment confirms the quantitative correlation results. Another interesting remark is that SenRVM errors seem not to be concentrated in specific areas but rather isolated.

The second evaluation carried out here assesses the temporal consistency of the SenRVM results. For each single date, the SenRVM performances are evaluated by computing the MAE. Only valid NDVI measurements not affected by clouds are considered. The MAEs and their respective standard deviations obtained are shown in Fig. 8. To simplify the result interpretation (and considering the similar results obtained for the three classes), the results are averaged across the three vegetation classes.

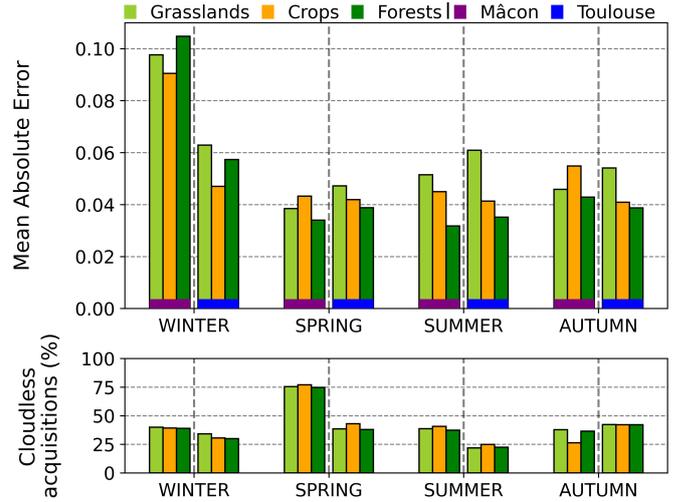


**Fig. 8** MAE obtained for each date comparing the expected non cloudy NDVI observations with their corresponding SenRVM predictions. Results are averaged for the three classes and reported for the *Mâcon* (top) and *Toulouse* (bottom) areas. Large MAE are explained by the presence of cloud & shadow mask errors.

As observed, most of the dates obtain MAE lower than 0.1, which confirms the SenRVM performances. The two study areas show similar and satisfactory results despite the timelines of image acquisitions are different. Some abnormal MAE higher than 0.2 are visible in January (*Mâcon* area) and December (*Toulouse* area). These high values are explained by the presence of cloud & shadow mask errors which can be visually confirmed. At these dates, the corresponding validity flags denote as valid numerous cloudy Sentinel-2 observations. In this situation, the NDVI values predicted by SenRVM are compared with invalid NDVI measures (see Sec. Appendix A). As a result, MAE calculated for those dates is high, however, it cannot be associated with methodological limitations among the temporal dimension.

Further analysis is carried out to evaluate if MAEs are influenced by the yearly season. Previous per-date results are averaged over the four seasons of the year as illustrated in Fig. 9. At the bottom of the figure, the percentages of cloudless NDVI measurements per season are also reported. Despite the seasons having a similar number of valid optical observations (except for spring in *Mâcon*), the winter period obtains the highest MAEs. These high values can be explained by the persistent winter cloud coverage producing large gaps without Sentinel-2 images (see Fig. 2).

Although Fig. 8 shows similar MAEs for both study areas, they are



**Fig. 9** On top, MAE computed for the four seasons of the year are shown for both study areas. The bottom chart indicates the percentage of cloudless observations available per season.

differently impacted by cloud coverage as shown in Fig. 2. Considering the  $R^2$  between predicted and expected NDVI time series of each polygon, correlation between  $R^2$  and the number of cloudless NDVI observations used is also investigated. Although 60 dates are available for the *Mâcon* area, cloudless measurements are ranging from 18 to 43 among all the polygons. Concerning *Toulouse*, the valid number of observations ranges from 7 to 35 out of the 71 dates. Significant correlation, above 0.7, is only found for *grasslands* and *crops* from the *Mâcon* area. In these cases, the performances are improved with an increased number of cloudless NDVI observations available for training. These results must be nevertheless tempered since the temporal distribution of missing data nor is their duration taken into account in this study. Moreover, the number of polygons greatly varies between classes. The results are therefore indicative and are not illustrated.

### 5.3.2. Ablation study

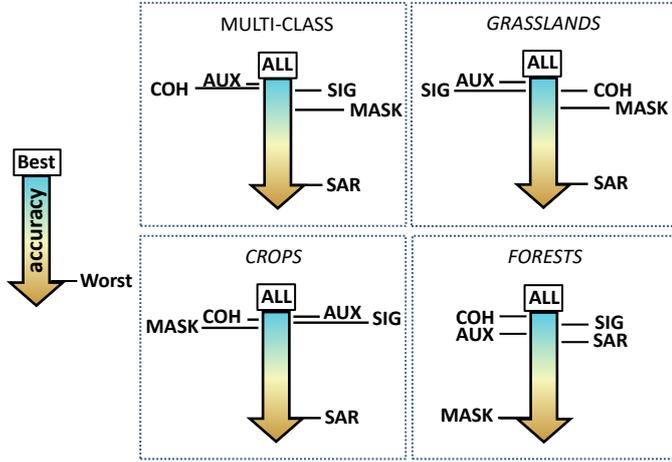
An ablation study is presented here by analyzing the SenRVM performances obtained with 6 input data scenarios. The study aims to evaluate the impact of the input features on the SenRVM predictions. The first scenario is the baseline one (denoted as ALL), in which the input SenRVM data corresponds to all features described in Tab. 2. The other 5 scenarios are constructed by removing some specific features from the baseline one.

Referring to Tab. 2, the  $\sigma_0$  (1) and coherence (2) features are removed, respectively, for the SIG and COH scenarios. In the SAR scenario, all features derived from Sentinel-1 and denoted as (1),(2),(3), and (4) are not considered. The AUX scenario studies the removal of (5), (6), and (7) ancillary features from ALL. Finally, the MASK scenario investigates how SenRVM performances differ when the validity flags provided by the cloud & shadow masks are not incorporated in the loss function. The different scenarios consider the same SenRVM parameter configurations and the same train/validation/test datasets.

The results of the ablation study are detailed in Appendix B, where the different scenarios are both evaluated with multi-class and single-class models. The four past metrics are computed to assess the results over *Mâcon* and *Toulouse*. Both study areas obtain similar results and attest that as expected, the highest accuracies are obtained by the ALL scenario. This result corroborates that none of the set of features removed from the different input data scenarios decrease the optimal ALL SenRVM accuracies. Only *forest* class models obtain different feature removal impact.

A visual interpretation of the ablation study results is illustrated in Fig. 10, where the summarized results of each specific scenario are compared with the ALL scenario. For this visual evaluation, a

global score is computed by averaging the four metrics obtained by both study areas. To take into account that the metrics have different ranges, they are normalized between 0 and 1 by considering the results of the different scenarios. The best result, obtained here in each case by ALL, is 1. Conversely, 0 indicates the worst result among the assessed scenarios. The global score is used to represent the accuracy decrease in Fig. 10. The arrow direction represents the accuracy decrease between the best (top) and worst (bottom) scenarios.



**Fig. 10** Ablation study results obtained by the multi-class and single-class SenRVM models. The accuracies of the five scenarios are compared to the baseline (ALL) by using a global score summarizing the four metrics results obtained for both study areas. The arrow direction represents the accuracy decrease between the best and worst scenarios. For further analysis, these results are detailed in Appendix B.

**SAR results :** the worst prediction accuracies are obtained by the SAR scenario where all features derived from Sentinel-1 are removed. Looking at the  $R^2$  values obtained by single-class *grassland* model, it can be observed that SAR scenario obtains low values (around 0.55) compared to the baseline ALL scenario ( $\approx 0.84$ ) for both study areas. Similar results are obtained by the multi-class model, which is expected given that multi-class data is mostly populated by *grasslands*. The  $R^2$  performance decrease is also observed on single-class *crop* results, however lesser with a decrease of 0.2. As discussed in Sec. 5.2, a high intra-class variability exists in *grasslands* and *crops* classes. Furthermore, the time series describing these classes have many abrupt breaks (*i.e.*, due to agricultural practices). Ancillary data are in these cases not sufficient to model the NDVI curves. The importance of Sentinel-1 data is here highlighted by its capacity in delivering precise and temporally close information describing the vegetation polygons. The SAR scenario thus leads to a significant performance decrease. Concerning single-class *forest* models, the removal of SAR features seems less important (decrease of about 0.02) for both study areas. This result is explained by the importance of ancillary data over *forests*, explained below.

**AUX results :** the lowest accuracy decrease ( $\approx 0.02$  of  $R^2$ ) is obtained by single-class models trained on *grasslands* and *crops* and consequently by the multi-class model. In contrast, the removal of AUX features seems to have a strong impact on the single-class *forests* model. Results obtained by the AUX scenario are, for *forests*, almost equivalent to the SAR scenario. This surprising results are explained by the low temporal variability of *forests* NDVI curves, which only have small fluctuations due to seasonal climate evolution. Information provided by the ancillary data is on one hand less prone to noise compared to satellite remote sensing features. For *forests*, ancillary data are thus valuable for obtaining satisfactory reconstruction results.

**SIG and COH results :** the results obtained by all SenRVM models for both study areas corroborate that the importance of  $\sigma_0$  and coherence features are relatively similar. The decrease of  $R^2$  observed for the two scenarios ranges from 0.01 to 0.06. Despite the similar results, some differences can be discussed between both scenarios. For

instance, the high standard deviations obtained by SIG scenario indicate a slightly superior importance of  $\sigma_0$  features. In this case, the results may be more stable as the number of inputs differs between both scenarios (*i.e.*,  $\sigma_0$  additional VV/VH polarization ratio band). Furthermore, while temporal information is directly integrated in coherence features, the use of RNNs to extract temporal dependencies akin to coherence information may explain this result. Opposite results can be observed between *Macon* and *Toulouse* areas concerning the impact of the SIG and COH scenarios. The contradictory results may be due to temporal and spatial differences existing in both areas and discussed in Sec. 5.3.1. Concerning single-class *forests* results, it can be observed that the removal of one of these two families of features does not greatly impact the performances.

**MASK results :** in general, the MASK scenario results obtained by the different SenRVM models indicate that the incorporation of the validity flags in the loss function is beneficial. Excluding the single-class *forests* results, the rest of the models exhibit that this scenario obtains the second most significant performance reduction. The performance decrease is more visible on single-class *grasslands* results with  $R^2$  of 0.78 and 0.73 for *Macon* and *Toulouse* respectively. For *forests*, the MASK scenario leads to the worst overall results. As in this case cloud & shadow mask information is not incorporated in the loss function, important label noise is affecting the training process. This noise prevents an accurate learning of the small variations observed in the *forests* time series (see Fig. 3c).

#### 5.4. Comparison with existing methodologies

Single-class SenRVM models are here evaluated against standard interpolation and ML regression methodologies. The robustness and efficiency of the methods are evaluated through their ability to reconstruct short- and long-term missing data gaps. The MAEs obtained for these two different scenarios are discussed in Sec. 5.4.1 and 5.4.2.

In these experiments, a new learning constraint is incorporated in the SenRVM training stage. The learning constraint ensures that satellite observations acquired on the dates that want to be predicted are not used to train the SenRVM models. It must be remarked that the same temporal grid is shared for the disjoint training and testing datasets used in the previous experiments. Therefore, a few non-cloudy observations acquired on the specific predicted date are most of the time considered in the training step. This is prevented here by completely masking the date during the training step.

The three ML regression algorithms described in the following are studied. For the three methods, the same SAR and ancillary input data as SenRVM is considered.

- a Random Forest Regression (RF) algorithm. This ensemble learning method is based on the construction of multiple decision tree classifiers (Belgiu and Drăguț, 2016; Li et al., 2020). The individual trees are built by applying a bagging strategy which randomly selects a subset of training samples and features. Following Pelletier et al. (2016) conclusions, the number of trees is set to 100. The maximum number of features taken into account for tree splitting is set to the square root of the number of input features and the maximum depth of a tree to 25.
- a Gaussian Processes Regression (GPR). It is a non-parametric kernel-based probabilistic regression algorithm based on a Bayesian framework. GPR is selected given the convincing results obtained in similar regression tasks (Belda et al., 2020b; Mercier et al., 2020). A squared exponential kernel function is employed.
- a deep-based method using MLPs architecture (MLP). This simplified version of the SenRVM is obtained by removing the recurrent block (*i.e.*, GRU cells) from the architecture. The training is thus performed for each date individually without handling the inherent temporal information of the time series. The MLP method integrates 4 MLPs for encoding and 1 MLP for decoding.

A fair comparison of the three previous regression algorithms with SenRVM requires the use of the same training samples. Nevertheless, substantial computational problems can occur if all polygons and their entire corresponding NDVI time series are used for training RF and GPR methods. Therefore, a temporal sampling strategy is proposed for

RF and GPR methods. The solution proposed is to discard some dates that are temporally far from the prediction date. For half of the training polygons of SenRVM, the selected training dates correspond to the nearest past cloud-free date of each polygon. For the other half, the following cloud-free date is used. It must be remarked that besides the sampling selection strategy, the space-time evolution of the cloud coverage makes it possible to include numerous dates and cover up to several months. The same training sample size as for SenRVM and MLP is thus kept for the RF method while temporal information is fed to the model. Concerning the GPR method, the number of training polygons is nevertheless limited to 10,000 polygons, due to long calculation times and memory constraints. It must be remarked that despite reducing the number of training samples, the resulting training dataset of GPR remains significant. This only impacts *grasslands* for both areas and *crops* for the *Toulouse* area.

Besides the previous ML algorithms, two mono-sensor standard interpolation methods are also considered in this study:

- a Whittaker smoother (WHIT), which is based on a penalized least-square regression algorithm combining fidelity to the data and smoothness of the filtered sequence (Vuolo et al., 2017; d’Andrimont et al., 2020). The smoothing criterion ( $\lambda$  value) is set to 1. This low value preserves the temporal variability of the original signal, describing important changes in vegetation cover. The  $d$  parameter used in the penalty calculation is set to 2.
- a weighted linear interpolation method (linear), which assigns weights to neighbouring observations (two before and after), based on the distance to the interpolated value.

Compared to the previous methods, the main difference is that WHIT and linear interpolation approaches only consider valid Sentinel-2 observations. These mono-sensor methods focus on exploiting the temporal trajectory of neighbouring NDVI observations. The Decomposition and Analysis of Time Series Software (DATimeS, Belda et al. (2020a)) is selected for the WHIT method. For the linear interpolation, the Orfeo ToolBox implementation (Grizonnet et al., 2017) is used.

#### 5.4.1. Gap filling of short-term data gaps

The short-term study consists in removing 6 individual non-consecutive acquisitions from the training datasets. The chosen dates are distributed along the complete year and have a low cloud cover rate. This implies that a high number of measurements can be used to validate this experiment. Prediction results obtained on the six reconstructed dates are evaluated for the three vegetation classes and over the two study areas.

In the case of SenRVM and the MLP method, a unique model is trained and used for the prediction of the six dates. In contrast, as a temporal sampling strategy is defined for the RF and GPR methods, six models are independently trained for the six masked dates.

Fig. 11 shows that similar satisfactory results are obtained by the different methods which most of the time achieve MAE lower than 0.15. *Grassland* and *forest* results obtained on *Mâcon* show that high errors are obtained on the second reconstructed date (i.e., February 14<sup>th</sup>). These high values are justified by the high number of cloud & shadow mask errors existing at this date.

Results show how SenRVM achieves accurate and comparable performances compared to standard interpolation methods. Furthermore, SenRVM obtains in most cases the lowest standard deviations. The interest of SenRVM is especially remarkable at the dates of June 8<sup>th</sup> over *Mâcon* and June 6<sup>th</sup> over *Toulouse*. At these dates, numerous anthropic activities exist given the agricultural calendar of *grasslands* and *crops*. These activities lead to phenology breaks in the time series (see Fig. 3) occurring during the masked dates. For the ML regression methods, based on SAR data, an accurate reconstruction of these breaks is possible whereas standard interpolation methods fail. A complementary discussion about break reconstruction can be found in Appendix C. Concerning *forests*, the stability of the NDVI temporal trajectory of this class leads to obtain similar results for all methods. For this class, SenRVM obtains the best accuracies for three dates.

The RF, GPR and MLP methods generally obtain comparable results. The highest MAEs are obtained by these methods, especially remarkable on *grasslands* and *crops*. The RF method obtains slightly better re-

sults than GPR for several dates, and notably lower standard deviations. Over 18 assessed dates for each area, RF achieves greater accuracy than SenRVM for 4 dates over *Mâcon* and 3 dates over *Toulouse*. GPR reaches lower MAE than SenRVM for 2 dates over *Mâcon* and only one date over *Toulouse*. The upper accuracy found with SenRVM can be explained by several factors. In comparison to RF and GPR methods, entire time series are fed to SenRVM. The bi-directional RNN used in SenRVM architecture can therefore extract long-term phenological stages helping accurate predictions. This long-term evolution may be neglected by the temporal sampling strategy of the RF and GPR methods. Another explanatory factor is the adoption of several encoding branches in the SenRVM and MLP methods, allowing more complex and descriptive features to be considered.

Regarding the results obtained by the MLP method, it only achieves the best results over the first date on *crops* for both areas and the first date over *forests* of the *Mâcon* area. The main explanatory factor is that MLP processed each date individually without exploiting temporal information. Comparing MLP with SenRVM results, the improvements achieved by the use of recurrent networks are highlighted. For example, *crop* results obtained by MLP show a high MAE in *Toulouse* at the date of November 27<sup>th</sup>. At this date, bare soil covers most of the *crops* parcels which leads to the presence of large fluctuations in SAR time series (see Sec. 3.4 and Fig. 3). Processing individual dates with the MLP method produces low prediction results while recurrent cells permit accurate predictions, taking advantage of the full temporal trajectory.

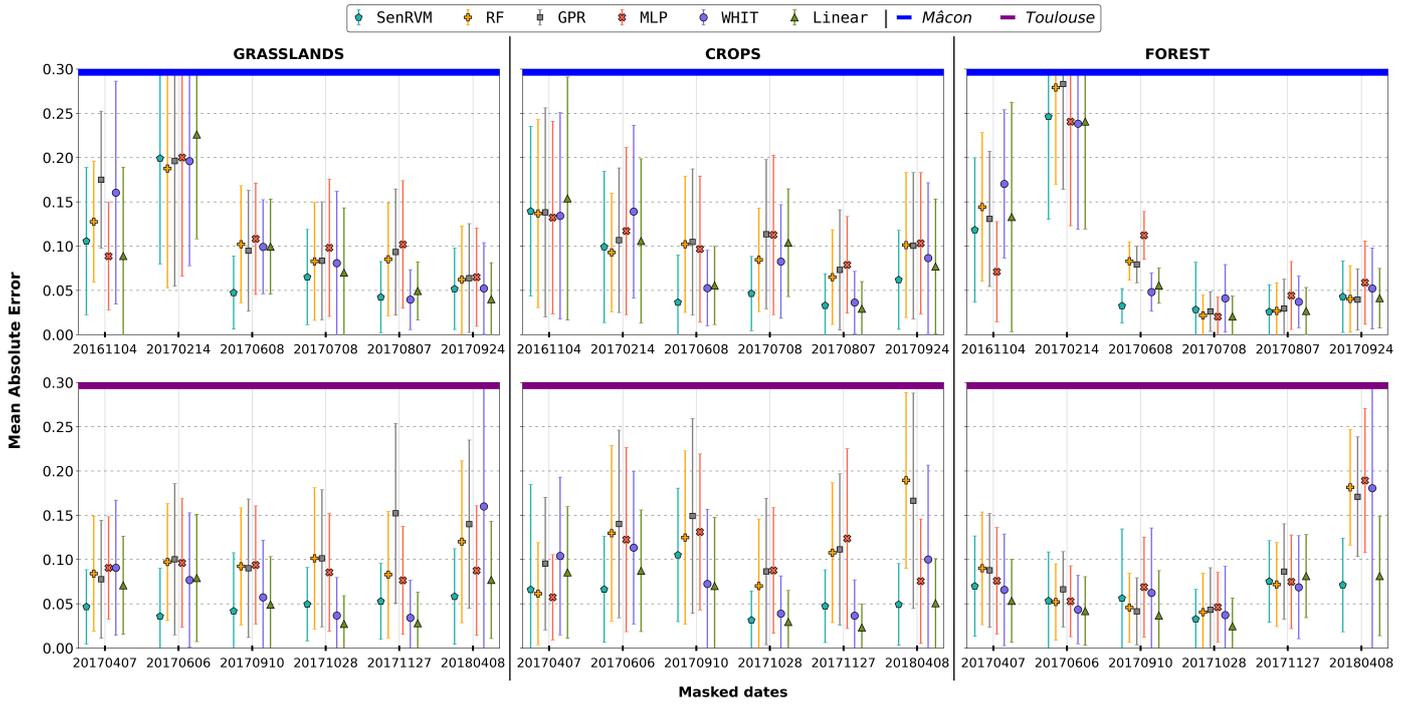
This experiment also corroborates the good results of the commonly used linear interpolation method. Over the 12 reconstructed dates, the linear interpolation reaches the highest accuracies for 5 dates on *grasslands* and *forests* and 3 dates on *crops*. The satisfactory results can be explained by the availability of cloudless neighbouring Sentinel-2 acquisitions, close to the reconstructed dates (see Fig. 2). Therefore, information describing the temporal trajectory of NDVI allows the linear method to obtain low MAE. As expected, the WHIT approach obtains results similar to the linear method. This method exhibits the highest inconsistencies that can be explained by the use of a smoothing function. This function is mostly well adapted for filtering purposes over long time series.

#### 5.4.2. Gap filling of a long-term data gap

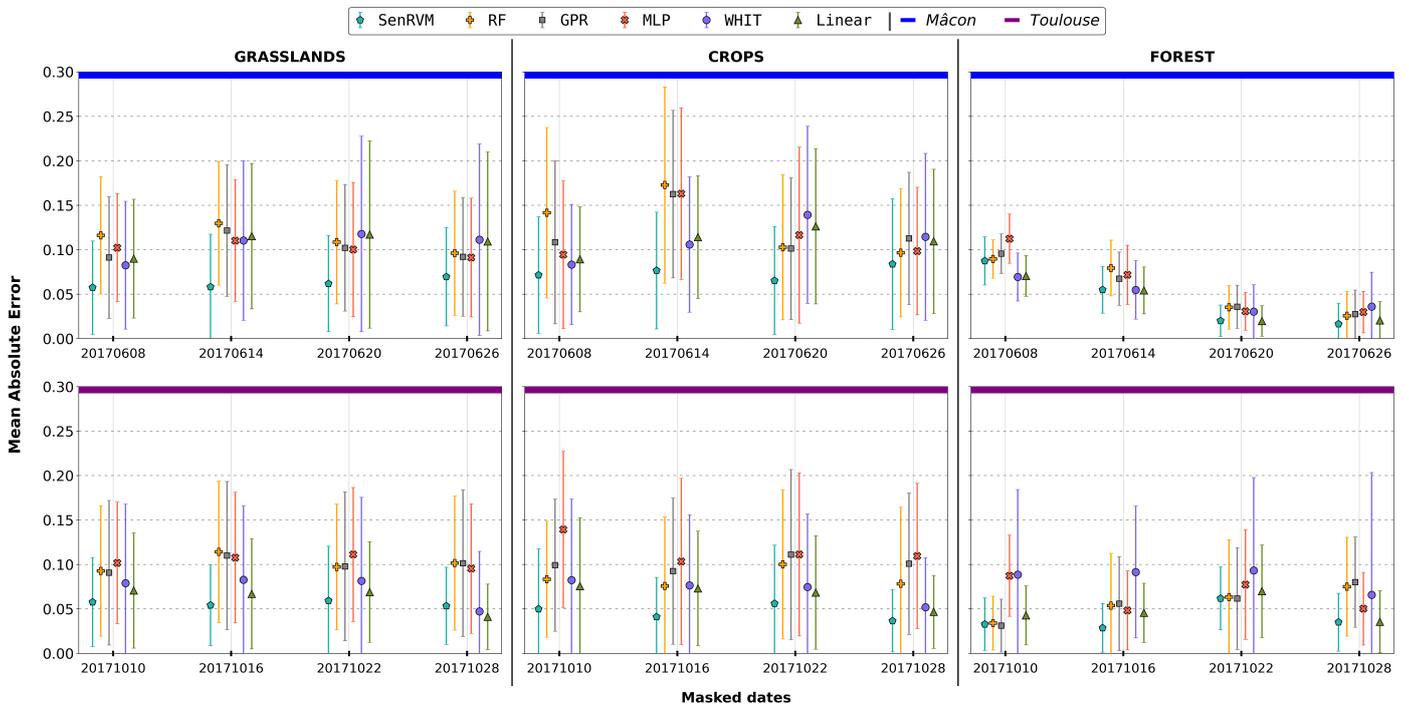
Consecutive missing values are likely to occur in optical time series. Consequently, the performances of gap filling methods can strongly decrease when the temporal frequency of exploitable observations is reduced. As previously discussed with Fig. 9, the presence of long-term gaps during winter can lead to a decrease of SenRVM performances. The robustness of the different methods in the presence of a significant data gap wants to be assessed here.

To perform this study, long-term data gaps are artificially created by removing 4 consecutive dates from the valid NDVI time series. The four consecutive dates are removed over different time periods for the two study areas. The number of valid observations as well as the agricultural practices calendars are considered as selection criteria. A long gap is then considered in June for the *Mâcon* area. Concerning *Toulouse*, October is chosen. For this experiment, as the four masked dates are consecutive, a unique model is trained for all ML regression methods.

The results obtained by the different methods can be compared by looking at Fig. 12. The reconstruction performances consolidate the previous results and further demonstrate the interest of the SenRVM approach. *Grassland* and *crop* results show how SenRVM obtains almost in all cases the lowest MAE and standard deviations. The performance improvement is more remarkable over *Mâcon*. This is explained by the frequent presence of anthropic interventions taking place during the reconstructed time period (see Appendix C). Solely the first date over *forests* for the *Mâcon* area and the last date over *grasslands* for the *Toulouse* area is better predicted by the two optical-based methods than by SenRVM. Nevertheless, the difference in MAE between SenRVM and the two methods is very low for this specific date. Concerning *forests*, similar results are obtained by all the methods. The low temporal variation of NDVI curves of *forests* again explains these similarities.



**Fig. 11** Comparison of gap filling performances for six non-consecutive dates, corresponding to **short-term** data gaps. Mean Absolute Error are assessed over *Mâcon* and *Toulouse* areas for the three vegetation classes.



**Fig. 12** Comparison of gap filling performances over four continuous dates, corresponding to **long-term** data gaps. Mean Absolute Error are assessed over *Mâcon* and *Toulouse* areas for the three vegetation classes.

Close MAEs are obtained for both areas by RF, GPR, and MLP methods for the three vegetation classes. GPR and MLP methods provide superior results than RF for the *Mâcon* area while it is the opposite for the *Toulouse* area. The exploitation of the temporal SAR trajectory and the use of multiple encoding branches explain the good SenRVM results.

Linear and WHIT methods exhibit substantially lower results than SenRVM, except for two dates. The differences are more noticeable for the *Mâcon* area. This is related to the presence of numerous agricultural practises, which are not detected by the mono-sensor interpolation methods. Consequently, differences are less marked for the *Toulouse* area, where few agricultural practices are performed.

## 6. Discussion

Sentinel family of satellite missions offers an efficient and accessible tool for continuous vegetation monitoring. Freely available optical and radar imagery can provide significant information to monitor vegetation changes over wide areas and different temporal scales. The development of new multi-source methods combining optical and radar satellite data can reduce the limitations of both sensors. To this aim, this work has proposed the SenRVM as a new joint strategy aiming at providing dense temporal resolution time series describing the phenological evolution of vegetation covers. The high temporal resolution of SAR acquisitions has been exploited to recover NDVI time series on a regular temporal grid of 6 days. The SenRVM methodology relies on a supervised recurrent deep learning architecture which integrates ancillary data to tackle the SAR signal weaknesses and its fluctuations sensitivity. The fusion of climate, topography, and SAR features has allowed to extract complex relationships. The SenRVM performances have been evaluated on a large dataset covering three vegetation classes exhibiting many agronomic subclasses over two distinct landscapes. This complete evaluation is a novelty since the majority of similar existing works have exclusively focused on *crops*. The accurate results have been obtained simultaneously on *grasslands*, *crops* and *forests*, demonstrating the relevance of the proposed SenRVM approach.

### 6.1. Vegetation monitoring with SenRVM

The quality assessment of SenRVM has been firstly evaluated by analyzing the prediction results obtained on a multi-class dataset covering the three types of vegetation. Satisfactory generalization capabilities of SenRVM have been obtained, despite the high phenological variability of the different classes. Further experiments will be necessary to study the SenRVM performances on minority classes (such as dactylis grass or common beech forests). It must be noticed that the reconstruction of minority classes with varied phenologies, such as coniferous forests or certain crop classes with less stem and leaf development could be challenging. Class-oriented results have been analyzed according to different temporal scales. The temporal stability of the results has been corroborated by the low errors. The highest errors have been observed during winter, more prone to long-term data gaps due to heavy cloud cover (Fig. 2). Besides the lack of information due to clouds, these results have been explained by the important fluctuations of SAR measurements in winter, due to the presence of bare soil or very little vegetation cover (Fig. 3). Discarding winter and focusing on a period restricted by phenological interest (e.g., from March to October), is possible for some applications (Veloso et al., 2017; Karasiak et al., 2020). To improve the current results, training on multi-annual time series could allow to gather more measurements over the winter period. RNNs could in this case extract additional recurrent phenological patterns over winter, leading to more robust SenRVM models.

A spatial evaluation of the results has been also performed to investigate if the reconstruction accuracies were impacted by the polygon sizes and their respective locations. This study has shown how a significant correlation exists between the size of the polygons (ranging in our study from 0.12 to several hundreds of hectares) and their prediction accuracies. The results have assessed that large homogeneous polygons reach the best accuracies. Among others, large polygons benefit from a strong speckle filtering, which could explain the

performance increase. Concerning the location of the polygons, a decrease in performance has been detected on high-relief geographical regions (up to 1295 *m* for the *Mâcon* area and 3127 *m* for *Toulouse*). It can be explained by the particular side-looking geometry and the strong local incidence angle effects induced by surface topography. These effects degrade the usefulness of SAR images. In some cases when steep topography is present, it may even prevent information extraction (Tamm et al., 2016). It must be remarked that the SenRVM method has proposed a solution to reduce these undesirable topographic effects through DTM-derived data. This data describing the polygons that have been integrated helped mitigate the topographic effects which nevertheless cannot be discarded.

The performances of multi-class and single-class SenRVM models have been also evaluated and compared by performing an ablation study (Fig. 10). The study has investigated the impact of different SenRVM input feature scenarios on the SenRVM performances. The obtained results have suggested different conclusions for agricultural and *forest* classes. For instance, ancillary data (e.g., climatic and topographic information) do not provide sufficient information for the prediction of NDVI time series on *grassland* and *crop* classes, driven by anthropic activities and phenological stages. Conversely, these features have seemed of utmost importance for accurate predictions on *forests*. This result has been explained by the lower impact of human activities leading *forest* time series to be defined by a more stable and seasonal evolution. The removal of the cloud & shadow mask information has allowed another distinction among vegetation classes. The use of the cloud & shadow masks have permitted a performance gain in all cases, which has been considerable on *forests*.

At last, different experiments have been carried out to evaluate SenRVM performances with respect to several existing ML regression algorithms and standard interpolation methods. Two different situations have been investigated to confirm the advantages of the SenRVM approach for regular vegetation monitoring. The reconstruction over short- and long-term data gaps has shown that the SenRVM approach obtained satisfactory results. SenRVM have reached identical absolute errors with classical mono-sensor methodologies and even the lowest in many cases. The advantages of SenRVM have been more noticeable for recovering long-term data gaps. Especially, the SenRVM performances have been remarkable when vegetation changes occur during the missing data period. In this situation, standard interpolation methods have obtained low accuracies and failed to capture changes in vegetation cover. The good performances of the deep-based SenRVM architecture have been mainly explained by two factors: its ability to extract complex features and relations between SAR and ancillary data and the efficient extraction of temporal information through recurrent cells.

### 6.2. Methodological perspectives

The combination of two standard neural networks has been proposed in the SenRVM architecture. The ability of RNNs to extract temporal information has been particularly conducive to a multimodal approach. E.g., NDVI,  $\sigma_{00}$  and coherence features time series have staggered temporal responses to vegetation changes (see Sec. 3.4). The temporal memory of the RNN has made possible the modelling of a common temporal response from the different features, resulting in stable and accurate results. Despite the SenRVM satisfactory results, the use of RNN as a key piece of the architecture could be discussed for further improvements. RNN has been traditionally used to process sequences that are longer than the exploited satellite time series. Moreover, they are relatively long to train. Alternatives can be considered for the exploitation of temporal (Pelletier et al., 2019) or spatio-temporal (Ienco et al., 2019) information. But most of these alternatives have been adopted in classification tasks. In the last years, new architectures such as the Transformers have replaced RNN, achieving remarkable results to process sequential data. Their use in remote sensing has just started (Garnot et al., 2020; Rußwurm and Körner, 2020) and needs further exploration for continuous regression problems.

Several considerations have been taken into account to avoid overfitting issues in our experiments. Nevertheless, overfitting problems could arise when a large number of parameters have to be learned.

Despite SenRVM being a supervised methodology, it must be highlighted that the lack of reference data is not an issue. The training of SenRVM models depends on Sentinel-1&2 acquisitions, which are ensured both spatially and temporally by the two constellations. Training models at different spatial scales (*e.g.*, Sentinel-2 tiles, administrative areas, climatic zones) are therefore possible. Moreover, even if RNNs have seemed relatively long to train, training times reported in Tab. 4 are reasonable. The SenRVM reconstruction task has taken about an hour for a complete Sentinel-2 tile (*i.e.*,  $1/90^{th}$  of metropolitan France). Thanks to the considerable amount of data describing the high diversity of landscapes and limited calculation times, the overfitting problem can then be minored.

Unlike the standard interpolation methods mentioned in Sec. 5.4, the temporal trajectory of Sentinel-2 NDVI time series was not exploited by the SenRVM method. Optical measurements have only been used for network supervision. The simplest solution to incorporate this optical information would be to insert valid NDVI measurements as input to the SenRVM architecture. Alongside the SAR features, these NDVI measurements might help the model to adjust its learning and improve the prediction performances. As the supervision would be this same value, there would be a significant risk that the network weights would be overwhelmingly adjusted according to it. The network could entirely discard the SAR features and completely rely on this NDVI value. Another interesting perspective for improvement would be to modify the loss function (Sec. 4), by proposing other similarity metrics measuring the trajectory similarity between predicted and expected time series. Other constraints based on first order derivatives, but computed on Sentinel-2 time series, could be explored. However, in any case, as missing data frequently occurs in optical time series, their exploitation can be tedious.

Finally, the SenRVM performances have been here only evaluated by the prediction of NDVI measurements. As previously mentioned, this choice was made given its versatility and simplicity. Nevertheless, it must be highlighted that the SenRVM methodology could be applied to different vegetation indices (*e.g.*, EVI, MSAVI, NDMI), biophysical variables (*e.g.*, LAI, fAPAR, fCOVER) or even raw optical spectral bands.

## 7. Conclusion

Prediction of dense time series describing vegetation phenology is valuable for many environmental applications. In this work, the SenRVM recurrent-based architecture has been proposed to regress SAR and ancillary features towards NDVI values. Annual NDVI time series over three prominent classes of vegetation, *grasslands*, *crops* and *forests*, have been predicted for two areas ( $> 20,000 \text{ km}^2$ ). The resulting time series with a temporal resolution of 6 days have exhibited high accurate results and limited standard deviation with respect to the expected Sentinel-2 NDVI time series.

Among the three classes and two areas,  $R^2$  above 0.83 and MAE under 0.05 have been found on a large set of vegetation polygons. Despite distinct phenologies and different observed time periods, satisfactory results have been obtained. The consistencies of the results have been assessed both spatially and temporally. The significant contribution of the SenRVM multi-sensor approach has been illustrated by its capacity to recover missing information even in the presence of breaks in the original time series (see Sec. Appendix C). Obtained results have corroborated that the reconstruction of dense NDVI time series with SenRVM can allow an accurate and regular monitoring of vegetation.

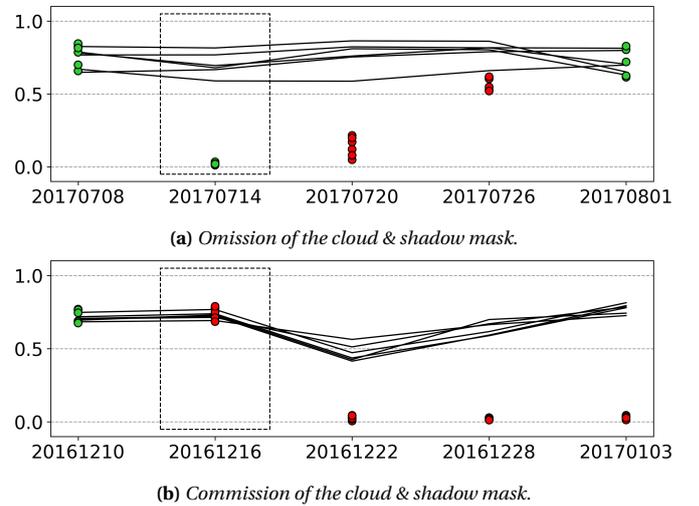
The contribution of this work has therefore opened the door to the development of SAR and optical data fusion methodologies. The use of SenRVM could be adapted for other applications, for instance, near real-time monitoring or forecasting. As such, learning a SenRVM model by integrating data acquired in one or multiple years could allow to predict NDVI measurements in near-real time, after each new SAR acquisition. Moreover, a country-wide and systematic application of the SenRVM method could be envisaged given the continuity of the Sentinel data, ensured at least until 2030. The imminent arrival of SAR satellites expanding the Sentinel-1 constellation may also let us apply the SenRVM approach with a temporal resolution higher than 6 days. Such improvements will be beneficial for regular vegetation monitoring through satellite time series.

## Acknowledgement

This work is funded by the Agence de la transition écologique (ADEME) and the Centre National d'Études Spatiales (CNES). The authors would like to thank Météo France for supplying the SAFRAN daily data and IGN-France for the preparation and provision of DTMs. Acknowledgements are due to Santiago Belda for providing the DATimes software.

## Appendix A. Towards cloud/shadow mask refinement

Accurate cloud and shadow detection remains a well-known challenge in optical remote sensing, despite a plethora of approaches (Baetens et al., 2019; Chen et al., 2019). Errors can be divided into two categories. Undetected observations flagged as valid measurements are known as *omission* errors, while *commission* errors correspond to the non-cloudy observations detected as invalid measurements. The presence of errors can occur, negatively impacting a large number of reference polygons, and subsequently the SenRVM performances (Sec 5.3.1 and 5.3.2).



**Fig. A.13** Two examples on how cloud/shadow mask errors could lead to an erroneous performance assessment. *Omission* (a) and *commission* (b) errors are shown in the black dashed boxes. As seen, accurate SenRVM results over neighboring polygons (black lines) are obtained on dates affected by cloud/shadow mask errors. Red dots are masked dates.

Fig. A.13 depicts how significant SenRVM errors could be associated to cloud *omission* and *commission* (Fig. A.13a and A.13b, respectively). SenRVM is particularly insensitive to such errors and could be used as a solution to improve the mask quality. An experimental set-up is proposed here.

To filter the errors, MAEs of SenRVM prediction are considered. This refinement is performed by applying at each date simple thresholds:

$$\tilde{m}_t = \begin{cases} 0, & \text{if } m_t = 1 \text{ and } |y_t - \hat{y}_t| > \alpha_{om} \\ 1, & \text{if } m_t = 0 \text{ and } |y_t - \hat{y}_t| < \alpha_{com} \\ m_t, & \text{otherwise.} \end{cases} \quad (\text{A.1})$$

$m_t$  is the original mask validity flag for the instant  $t$ . The expected NDVI and SenRVM predicted NDVI values are  $y_t$  and  $\hat{y}_t$ , respectively. The performance of the refinement process depends on the *commission* and *omission* thresholds, which are empirically set:  $\alpha_{om} = 0.3$  and  $\alpha_{com} = 0.02$ . Eq. A.1 is applied on the original masks of both study areas by considering the prediction results obtained by the single-class SenRVM models presented in Tab. 5. It results in removing or adding a few validation measurements: for each polygon, 1.26 and 1.82 dates

were found as *omission* and *commission*, respectively. To investigate the impact of outliers on performances, the resulting  $\tilde{m}$  masks are then used to retrain the SenRVM architecture for the three vegetation classes (the same datasets and hyperparameters as those in Tab. 5). Results obtained by using the refined mask are reported in Tab. A.6.

**Tab. A.6** SenRVM performances obtained after mask refinement. Results correspond to difference with those shown in Tab. 5, Sec.5.2. In all cases, results are improved.

		<i>Mâcon</i>	<i>Toulouse</i>
<i>Grasslands</i>	R <sup>2</sup>	+0.0334 ± 0.0043	+0.0228 ± 0.0023
	MAE	-0.0037 ± 0.0014	-0.0034 ± 0.0007
	MSE	-0.0010 ± 0.0002	-0.0008 ± 0.0001
	RMSE	-0.0083 ± 0.0023	-0.0064 ± 0.0011
<i>Crops</i>	R <sup>2</sup>	+0.0099 ± 0.0024	+0.0029 ± 0.0010
	MAE	-0.0025 ± 0.0013	-0.0026 ± 0.0008
	MSE	-0.0009 ± 0.0001	-0.0005 ± 0.0001
	RMSE	-0.0070 ± 0.0013	-0.0049 ± 0.0015
<i>Forests</i>	R <sup>2</sup>	+0.0300 ± 0.0035	+0.0092 ± 0.0044
	MAE	-0.0032 ± 0.0004	+0.0004 ± 0.0037
	MSE	-0.0008 ± 0.0001	-0.0000 ± 0.0004
	RMSE	-0.0069 ± 0.0016	-0.0003 ± 0.0038

The gain in accuracy corroborates that the removal of cloud & shadow mask errors could improve the SenRVM performances. The benefit is observed for both areas and the three vegetation classes and across the four metrics. The minimum R<sup>2</sup> of 0.83 is improved to 0.86 by using the  $\tilde{m}$  masks. Improvements are especially significant in the *Mâcon* area, given its important number of errors and over *grasslands*: the R<sup>2</sup> increase is around 0.04 and 0.02 for *Mâcon* and *Toulouse* areas, respectively. Concerning *crops*, the improvement is less noticeable given that very good scores are already obtained by using the original masks. For both *grassland* and *crop* classes, MAE errors are improved. MAE below 0.04 is achieved for both classes over the *Mâcon* area. The single-class *forests* model also improves its prediction performances (for *Mâcon*, R<sup>2</sup> goes from 0.84 to 0.87). The interest of the refinement strategy is finally reflected in the standard deviation values of Tab. A.6, which are much lower than those described in Tab. 5.

## Appendix B. Ablation study detailed results

Detailed results of the ablation study experiment conducted and analyzed in Sec. 5.3.2 are presented in Tab. B.7.

## Appendix C. Reconstruction of time series breaks

We explore the ability to reconstruct time series breaks, which are usually associated with vegetation changes over agricultural areas (*i.e.*, *grasslands* and *crops*). SenRVM is especially relevant, with respect to the other methods, for dates containing numerous breaks (June 8<sup>th</sup> and June 6<sup>th</sup> for *Mâcon* and *Toulouse* respectively, in Fig. 11). The long-term experiment (Sec. 5.4.2) shows such conclusions, SenRVM reaching the highest accuracies on agricultural classes, but without significant differences in terms of MAE. A qualitative evaluation of the results obtained in Sec. 5.4.2 is proposed. This evaluation permits to highlight the interest of the SenRVM method to recover vegetation changes.

The reconstructions of two types of breaks are considered: a drastic decrease in NDVI due to mowing or ploughing, or an increase due to vegetation growth. Fig. C.14 shows four examples over two *grassland* and two *crop* polygons. It corroborates the interest of ML regression methods: the results show that the reconstruction performances of mono-sensor methods (linear, WHIT) are strongly affected by the distance between the reconstructed date and the valid NDVI measurements used in the reconstruction. Because of the simulated data gap of

about one month, the standard interpolation methods are using temporally distant dates for interpolation. On intensively exploited and quickly evolving agricultural parcels, this interpolation significantly or even entirely obscures part of their phenological cycles. SAR-based multi-sensor solutions can integrate temporally close knowledge, and efficiently recover these cycles. Comparing RF, GPR, MLP and SenRVM, the latter exhibiting the highest accuracies obtained and more stable results.

Fig. C.14a shows the results for a permanent *grassland* parcel (a mowing followed by a vegetation regrowth). Standard interpolation approaches use the dates of June 2<sup>nd</sup> and July 26<sup>th</sup> for reconstructing the missing period. Despite the mowing occurring in-between, the vegetation has grown back on July 26<sup>th</sup>, reaching high NDVI values. The resulting reconstructed time series follow a gradual but not significant decrease in NDVI, preventing abrupt change detection.

Fig. C.14b, illustrates the capability of the methods for recovering vegetation changes on a maize parcel. As observed, the four masked dates result in a gap of two months without cloudless measurements. The vegetation growth is well captured by all methods. However, the standard interpolation methods show gradual and weak growth, whereas ML regression methods suggest that such a growth occurs mainly in June. The image acquired on June 20<sup>th</sup> confirms the accurate reconstruction of the four multi-sensor methods.

A complete phenological cycle showing the growth and harvest periods is shown in Fig. C.14c. The two dates related to both growth and mowing stages are unfortunately acquired during the data gap. As a result, the two successive dates available for the standard interpolation methods both correspond to a bare ground cover with close NDVI values. The reconstructed time series are therefore flat and do not reflect vegetation changes. In contrast, ML regression approaches accurately fit the phenology cycle of the parcel.

The harvest period of a winter wheat parcel is shown in Fig. C.14d. In this example, a single cloudless NDVI observation is masked. It results in a two-month data gap between September 10<sup>th</sup> and November 9<sup>th</sup>. The standard interpolation methods fail to reconstruct the vegetation decrease, while ML regression approaches accurately mark a clear decrease in NDVI values, suggesting that the mowing occurred between September 10<sup>th</sup> and September 16<sup>th</sup>. Even without masking the date of October 10<sup>th</sup> to the standard interpolation methods, it would have been impossible to determine the date of mowing with such precision.

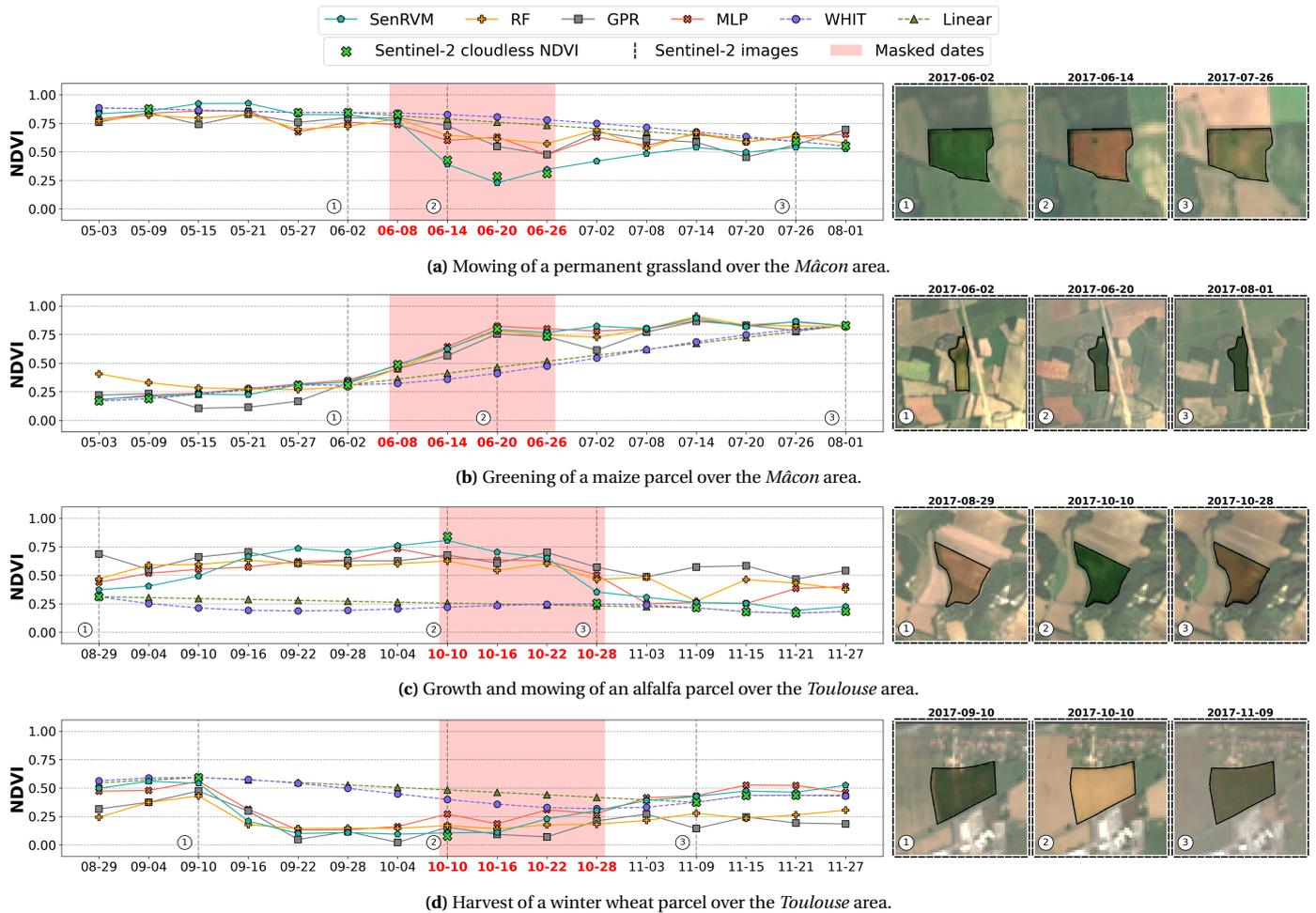
These last results further corroborate that regular vegetation monitoring needs the exploitation of multi-sensor information such as it has been proposed by the SenRVM approach.

## References

- Ali, I., Cawkwell, E., Dwyer, E., Barrett, B., Green, S., 2016. Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology* 9, 649–671.
- Atkinson, P.M., Jeganathan, C., Dash, J., Atzberger, C., 2012. Inter-comparison of four models for smoothing satellite sensor time-series data to estimate vegetation phenology. *Remote Sensing of Environment* 123, 400–417.
- Atkinson, P.M., Tatnall, A.R.L., 1997. Introduction Neural networks in remote sensing. *International Journal of Remote Sensing* 18, 699–709.
- Atzberger, C., 2004. Object-based retrieval of biophysical canopy variables using artificial neural nets and radiative transfer models. *Remote Sensing of Environment* 93, 53–67.
- Atzberger, C., 2013. Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing* 5, 949–981.
- Baetens, L., Desjardins, C., Hagolle, O., 2019. Validation of Copernicus Sentinel-2 Cloud Masks Obtained from MAJA, Sen2Cor, and FMask Processors Using Reference Cloud Masks Generated with a Supervised Active Learning Procedure. *Remote Sensing* 11.
- Barsi, J.A., Alhammoud, B., Czapla-Myers, J., Gascon, F., Haque, M.O., Kaewmanee, M., Leigh, L., Markham, B.L., 2018. Sentinel-2A MSI and Landsat-8 OLI radiometric cross comparison over desert sites. *European Journal of Remote Sensing* 51, 822–837.

**Tab. B.7** Average results and  $\pm$  standard deviation for the ablation study experiments described in Sec. 5.3.2. Five input data scenarios are compared to the baseline scenario (ALL) for which the complete dataset described in Tab. 2 is used. Four metrics are provided for the *grasslands*, *crops*, *forests* and multi-class datasets. Green color indicates the input removal having the lowest impact on performances, while the red colour denotes the highest.

		<i>Multi-class (Grasslands, Crops, Forests)</i>			
		R <sup>2</sup>	MAE	MSE	RMSE
<i>Mâcon</i>	<b>ALL</b>	<b>0.8650 <math>\pm</math> 0.0121</b>	<b>0.0419 <math>\pm</math> 0.0065</b>	<b>0.0039 <math>\pm</math> 0.0010</b>	<b>0.0624 <math>\pm</math> 0.0069</b>
	- COH	0.8293 $\pm$ 0.0097	0.0459 $\pm$ 0.0030	0.0046 $\pm$ 0.0004	0.0681 $\pm$ 0.0028
	- SIG	0.8346 $\pm$ 0.0100	0.0469 $\pm$ 0.0029	0.0048 $\pm$ 0.0004	0.0689 $\pm$ 0.0029
	- AUX	0.8458 $\pm$ 0.0145	0.0443 $\pm$ 0.0060	0.0043 $\pm$ 0.0008	0.0654 $\pm$ 0.0058
	- SAR	0.6314 $\pm$ 0.0084	0.0848 $\pm$ 0.0019	0.0148 $\pm$ 0.0004	0.1217 $\pm$ 0.0015
	- MASK	0.7854 $\pm$ 0.0152	0.0532 $\pm$ 0.0041	0.0058 $\pm$ 0.0007	0.0761 $\pm$ 0.0041
<i>Toulouse</i>	<b>ALL</b>	<b>0.8947 <math>\pm</math> 0.0056</b>	<b>0.0404 <math>\pm</math> 0.0039</b>	<b>0.0030 <math>\pm</math> 0.0005</b>	<b>0.0545 <math>\pm</math> 0.0038</b>
	- COH	0.8669 $\pm$ 0.0071	0.0452 $\pm$ 0.0024	0.0038 $\pm$ 0.0003	0.0616 $\pm$ 0.0022
	- SIG	0.8562 $\pm$ 0.0111	0.0498 $\pm$ 0.0050	0.0045 $\pm$ 0.0007	0.0671 $\pm$ 0.0049
	- AUX	0.8778 $\pm$ 0.0135	0.0435 $\pm$ 0.0055	0.0035 $\pm$ 0.0008	0.0587 $\pm$ 0.0057
	- SAR	0.5758 $\pm$ 0.0067	0.1194 $\pm$ 0.0021	0.0267 $\pm$ 0.0004	0.1634 $\pm$ 0.0012
	- MASK	0.8179 $\pm$ 0.0152	0.0531 $\pm$ 0.0044	0.0054 $\pm$ 0.0007	0.0732 $\pm$ 0.0046
		<i>Grasslands</i>			
		R <sup>2</sup>	MAE	MSE	RMSE
<i>Mâcon</i>	<b>ALL</b>	<b>0.8384 <math>\pm</math> 0.0152</b>	<b>0.0418 <math>\pm</math> 0.0058</b>	<b>0.0040 <math>\pm</math> 0.0009</b>	<b>0.0629 <math>\pm</math> 0.0069</b>
	- COH	0.7891 $\pm$ 0.0081	0.0470 $\pm$ 0.0021	0.0048 $\pm$ 0.0004	0.0691 $\pm$ 0.0027
	- SIG	0.8011 $\pm$ 0.0115	0.0470 $\pm$ 0.0038	0.0048 $\pm$ 0.0006	0.0690 $\pm$ 0.0041
	- AUX	0.8159 $\pm$ 0.0147	0.0439 $\pm$ 0.0043	0.0043 $\pm$ 0.0007	0.0657 $\pm$ 0.0048
	- SAR	0.5557 $\pm$ 0.0061	0.0791 $\pm$ 0.0019	0.0122 $\pm$ 0.0005	0.1106 $\pm$ 0.0023
	- MASK	0.7893 $\pm$ 0.0819	0.0528 $\pm$ 0.0063	0.0060 $\pm$ 0.0011	0.0772 $\pm$ 0.0063
<i>Toulouse</i>	<b>ALL</b>	<b>0.8464 <math>\pm</math> 0.0115</b>	<b>0.0443 <math>\pm</math> 0.0029</b>	<b>0.0037 <math>\pm</math> 0.0006</b>	<b>0.0606 <math>\pm</math> 0.0049</b>
	- COH	0.8076 $\pm$ 0.0106	0.0506 $\pm$ 0.0032	0.0046 $\pm$ 0.0005	0.0677 $\pm$ 0.0037
	- SIG	0.7976 $\pm$ 0.0173	0.0526 $\pm$ 0.0040	0.0050 $\pm$ 0.0007	0.0706 $\pm$ 0.0047
	- AUX	0.8256 $\pm$ 0.0114	0.0470 $\pm$ 0.0025	0.0040 $\pm$ 0.0005	0.0635 $\pm$ 0.0037
	- SAR	0.5535 $\pm$ 0.0093	0.0914 $\pm$ 0.0016	0.0151 $\pm$ 0.0004	0.1231 $\pm$ 0.0016
	- MASK	0.7332 $\pm$ 0.0173	0.0579 $\pm$ 0.0034	0.0063 $\pm$ 0.0006	0.0791 $\pm$ 0.0034
		<i>Crops</i>			
		R <sup>2</sup>	MAE	MSE	RMSE
<i>Mâcon</i>	<b>ALL</b>	<b>0.9433 <math>\pm</math> 0.0017</b>	<b>0.0420 <math>\pm</math> 0.0014</b>	<b>0.0040 <math>\pm</math> 0.0001</b>	<b>0.0630 <math>\pm</math> 0.0010</b>
	- COH	0.9270 $\pm$ 0.0070	0.0490 $\pm$ 0.0042	0.0052 $\pm$ 0.0007	0.0718 $\pm$ 0.0043
	- SIG	0.9220 $\pm$ 0.0080	0.0516 $\pm$ 0.0058	0.0057 $\pm$ 0.0010	0.0753 $\pm$ 0.0059
	- AUX	0.9358 $\pm$ 0.0081	0.0466 $\pm$ 0.0075	0.0047 $\pm$ 0.0010	0.0684 $\pm$ 0.0076
	- SAR	0.7212 $\pm$ 0.0080	0.1181 $\pm$ 0.0051	0.0257 $\pm$ 0.0011	0.1603 $\pm$ 0.0034
	- MASK	0.9176 $\pm$ 0.0100	0.0536 $\pm$ 0.0077	0.0061 $\pm$ 0.0015	0.0776 $\pm$ 0.0085
<i>Toulouse</i>	<b>ALL</b>	<b>0.9676 <math>\pm</math> 0.0037</b>	<b>0.0353 <math>\pm</math> 0.0038</b>	<b>0.0026 <math>\pm</math> 0.0005</b>	<b>0.0503 <math>\pm</math> 0.0044</b>
	- COH	0.9584 $\pm$ 0.0048	0.0389 $\pm$ 0.0040	0.0031 $\pm$ 0.0005	0.0551 $\pm$ 0.0042
	- SIG	0.9526 $\pm$ 0.0039	0.0422 $\pm$ 0.0048	0.0037 $\pm$ 0.0006	0.0608 $\pm$ 0.0045
	- AUX	0.9652 $\pm$ 0.0026	0.0363 $\pm$ 0.0031	0.0027 $\pm$ 0.0004	0.0517 $\pm$ 0.0035
	- SAR	0.7151 $\pm$ 0.0076	0.1293 $\pm$ 0.0039	0.0296 $\pm$ 0.0007	0.1720 $\pm$ 0.0021
	- MASK	0.9501 $\pm$ 0.0062	0.0440 $\pm$ 0.0063	0.0041 $\pm$ 0.0011	0.0634 $\pm$ 0.0076
		<i>Forests</i>			
		R <sup>2</sup>	MAE	MSE	RMSE
<i>Mâcon</i>	<b>ALL</b>	<b>0.8486 <math>\pm</math> 0.0268</b>	<b>0.0343 <math>\pm</math> 0.0066</b>	<b>0.0032 <math>\pm</math> 0.0007</b>	<b>0.0562 <math>\pm</math> 0.0057</b>
	- COH	0.8443 $\pm$ 0.0182	0.0352 $\pm$ 0.0042	0.0034 $\pm$ 0.0005	0.0582 $\pm$ 0.0043
	- SIG	0.8349 $\pm$ 0.0234	0.0365 $\pm$ 0.0050	0.0035 $\pm$ 0.0006	0.0587 $\pm$ 0.0044
	- AUX	0.8189 $\pm$ 0.0243	0.0373 $\pm$ 0.0050	0.0037 $\pm$ 0.0005	0.0604 $\pm$ 0.0041
	- SAR	0.8207 $\pm$ 0.0056	0.0403 $\pm$ 0.0009	0.0043 $\pm$ 0.0003	0.0653 $\pm$ 0.0021
	- MASK	0.7119 $\pm$ 0.0266	0.0514 $\pm$ 0.0073	0.0064 $\pm$ 0.0010	0.0795 $\pm$ 0.0059
<i>Toulouse</i>	<b>ALL</b>	<b>0.9235 <math>\pm</math> 0.0106</b>	<b>0.0318 <math>\pm</math> 0.0051</b>	<b>0.0020 <math>\pm</math> 0.0005</b>	<b>0.0450 <math>\pm</math> 0.0049</b>
	- COH	0.9232 $\pm$ 0.0041	0.0317 $\pm$ 0.0015	0.0021 $\pm$ 0.0001	0.0454 $\pm$ 0.0015
	- SIG	0.9152 $\pm$ 0.0076	0.0340 $\pm$ 0.0034	0.0024 $\pm$ 0.0004	0.0484 $\pm$ 0.0033
	- AUX	0.9086 $\pm$ 0.0047	0.0343 $\pm$ 0.0026	0.0024 $\pm$ 0.0002	0.0487 $\pm$ 0.0024
	- SAR	0.9058 $\pm$ 0.0050	0.0395 $\pm$ 0.0012	0.0033 $\pm$ 0.0002	0.0577 $\pm$ 0.0018
	- MASK	0.8454 $\pm$ 0.0204	0.0483 $\pm$ 0.0081	0.0049 $\pm$ 0.0011	0.0698 $\pm$ 0.0070



**Fig. C.14** Performance assessment comparing six different methods aiming at recovering vegetation changes occurring during a long-term data gap (red square). NDVI values evaluating the reconstruction results are depicted by the green crosses. Black-dashed lines and circle numbers correspond to the specific dates which are analyzed by the Sentinel-2 images shown on the right part of the figures.

Beck, P., Atzberger, C., Høgda, K., Johansen, B., Skidmore, A., 2006. Improved monitoring of vegetation dynamics at very high latitudes : a new method using MODIS NDVI. *Remote sensing of environment* 100, 321–334.

Belda, S., Pipia, L., Morcillo-Pallarés, P., Rivera-Caicedo, J.P, Amin, E., De Grave, C., Verrelst, J., 2020a. DATimeS: A machine learning time series GUI toolbox for gap-filling and vegetation phenology trends detection. *Environmental Modelling & Software* 127, 104666.

Belda, S., Pipia, L., Morcillo-Pallarés, P., Verrelst, J., 2020b. Optimizing Gaussian Process Regression for Image Time Series Gap-Filling and Crop Monitoring. *Agronomy* 10.

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment* 204, 509 – 523.

Belgiu, M., Drăguț, L., 2016. Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114, 24–31.

Bengtsson, J., et al., 2019. Grasslands—more important for ecosystem services than you might think. *Ecosphere* 10.

Benninga, H.J.F., van der Velde, R., Su, Z., 2019. Impacts of Radiometric Uncertainty and Weather-Related Surface Conditions on Soil Moisture Retrievals with Sentinel-1. *Remote Sensing* 11.

Bolton, D.K., Gray, J.M., Melaas, E.K., Moon, M., Eklundh, L., Friedl, M.A., 2020. Continental-scale land surface phenology from harmonized Landsat 8 and Sentinel-2 imagery. *Remote Sensing of Environment* 240, 111685.

Buchmann, Nina and Fuchs, Kathrin and Feigenwinter, Iris and Gilgen, Anna K., 2019. Multifunctionality of permanent grasslands: ecosystem services and resilience to climate change, in: *Improving sown grasslands through breeding and management*, European Grassland Federation EGF pp. 19 – 26.

Cai, Z., Jönsson, P., Jin, H., Eklundh, L., 2017. Performance of smoothing methods for reconstructing ndvi time-series and estimating vegetation phenology from modis data. *Remote Sensing* 9.

Campos-Taberner, M., García-Haro, F.J., Martínez, B., Izquierdo-Verdiguier, E., Atzberger, C., Camps-Valls, G., Gilabert, M.A., 2020. Understanding deep learning in land use classification based on Sentinel-2 time series. *Scientific Reports* 10, 17188.

Campos-Taberner, M., García-Haro, F.J., Martínez, B., Sánchez-Ruiz, S., Gilabert, M.A., 2019. A Copernicus Sentinel-1 and Sentinel-2 Classification Framework for the 2020+ European Common Agricultural Policy: A Case Study in València (Spain). *Agronomy* 9.

Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. *Remote Sensing of Environment* 91, 332 – 344.

Chen, Y., He, W., Yokoya, N., Huang, T.Z., 2019. Blind cloud and cloud shadow removal of multitemporal images based on total variation regularized low-rank sparsity decomposition. *ISPRS Journal of Photogrammetry and Remote Sensing* 157, 93 – 107.

Cho, K., van Merriënboer, B., Bahdanau, D., Bengio, Y., 2014. On the Properties of Neural Machine Translation: Encoder-Decoder Ap-

- proaches.
- Chuvieco, E. (Ed.), 2008. *Earth Observation of Global Change*. Springer Netherlands.
- Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.C., Skakun, S.V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment* 219, 145–161.
- Cresson, R., Ienco, D., Gaetano, R., Ose, K., Tong Minh, D.H., 2019. Optical image gap filling using deep convolutional autoencoder from optical and radar images, in: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 218–221.
- Das, M., Ghosh, S.K., 2016. A cost-efficient approach for measuring Moran's index of spatial autocorrelation in geostationary satellite data, in: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pp. 5913–5916.
- Dass, P., Houlton, B.Z., Wang, Y., Warlind, D., 2018. Grasslands may be more reliable carbon sinks than forests in California. *Environmental Research Letters* 13, 074027.
- de Wit, A., Su, B., 2005. Deriving phenological indicators from SPOT-VGT data using the HANTS algorithm, in: Veroustraete, E., Bartholomé, E., Verstraeten, W. (Eds.), *Proceedings of the 2nd international VEGETATION user conference; 1998-2004: 6 years of operational activities*, EC. pp. 195–201. 2nd international VEGETATION user conference; Antwerp (Belgium) ; Conference date: 24-03-2004 Through 26-03-2004.
- Defourny, P., Bontemps, S., Bellemans, N., Cara, C., Dedieu, G., Guzonato, E., Hagolle, O., Inglada, J., Nicola, L., Rabaute, T., Savinaud, M., Udroui, C., Valero, S., Bégué, A., Dejoux, J.F., El Harti, A., Ez-zahar, J., Kussul, N., Labbassi, K., Lebourgeois, V., Miao, Z., Newby, T., Nyamugama, A., Salh, N., Shelestov, A., Simonneaux, V., Traore, P.S., Traore, S.S., Koetz, B., 2019. Near real-time agriculture monitoring at national scale at parcel resolution: Performance assessment of the Sen2-Agri automated system in various cropping systems around the world. *Remote Sensing of Environment* 221, 551–568.
- Desai, M., Ganatra, A., 2012. Survey on Gap Filling in Satellite Images and Inpainting Algorithm. *International Journal of Computer Theory and Engineering* 4, 341–345.
- Ding, C., Liu, X., Huang, F., 2017. Temporal interpolation of satellite-derived leaf area index time series by introducing spatial-temporal constraints for heterogeneous grasslands. *Remote Sensing* 9.
- Dong, T., Liu, J., Qian, B., He, L., Liu, J., Wang, R., Jing, Q., Champagne, C., McNairn, H., Powers, J., Shi, Y., Chen, J.M., Shang, J., 2020. Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data. *ISPRS Journal of Photogrammetry and Remote Sensing* 168, 236–250.
- Dwyer, J.L., Roy, D.P., Sauer, B., Jenkerson, C.B., Zhang, H.K., Lymburner, L., 2018. *Analysis Ready Data: Enabling Analysis of the Landsat Archive*. *Remote Sensing* 10.
- d'Andrimont, R., Taymans, M., Lemoine, G., Ceglar, A., Yordanov, M., van der Velde, M., 2020. Detecting flowering phenology in oil seed rape parcels with Sentinel-1 and -2 time series. *Remote Sensing of Environment* 239, 111660.
- Fauvel, M., Lopes, M., Dubo, T., Rivers-Moore, J., Frison, P.L., Gross, N., Ouin, A., 2020. Prediction of plant diversity in grasslands using Sentinel-1 and -2 satellite image time series. *Remote Sensing of Environment* 237, 111536.
- Fensholt, R., Proud, S.R., 2012. Evaluation of Earth Observation based global long term vegetation trends — Comparing GIMMS and MODIS global NDVI time series. *Remote Sensing of Environment* 119, 131–147.
- Frison, P.L., Fruneau, B., Kmiha, S., Soudani, K., Dufrière, E., Le Toan, T., Koleček, T., Villard, L., Mougin, E., Rudant, J.P., 2018. Potential of Sentinel-1 Data for Monitoring Temperate Mixed Forest Phenology. *Remote Sensing* 10.
- Gao, F., Anderson, M.C., Zhang, X., Yang, Z., Alfieri, J.G., Kustas, W.P., Mueller, R., Johnson, D.M., Prueger, J.H., 2017. Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery. *Remote Sensing of Environment* 188, 9–25.
- Gao, J., Yuan, Q., Li, J., Zhang, H., Su, X., 2020. Cloud Removal with Fusion of High Resolution Optical and SAR Images Using Generative Adversarial Networks. *Remote Sensing* 12.
- Garnot, V.S.F., Landrieu, L., Giordano, S., Chehata, N., 2020. Satellite Image Time Series Classification with Pixel-Set Encoders and Temporal Self-Attention, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12325–12334.
- Gascon, F., Cadau, E., Colin, O., Hoersch, B., Isola, C., Fernández, B.L., Martimort, P., 2014. Copernicus Sentinel-2 Mission: Products, Algorithms and Cal/Val, in: Butler, J.J., Xiong, X.J., Gu, X. (Eds.), *Earth Observing Systems XIX, International Society for Optics and Photonics*. SPIE. pp. 455–463.
- Gerber, E., de Jong, R., Schaepman, M.E., Schaepman-Strub, G., Furrer, R., 2018. Predicting Missing Values in Spatio-Temporal Remote Sensing Data. *IEEE Transactions on Geoscience and Remote Sensing* 56, 2841–2853.
- Griffiths, P., Nendel, C., Pickert, J., Hostert, P., 2020. Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. *Remote Sensing of Environment* 238, 111124. *Time Series Analysis with High Spatial Resolution Imagery*.
- Grizonnet, M., Michel, J., Poughon, V., Inglada, J., Savinaud, M., Cresson, R., 2017. Orfeo ToolBox: Open source processing of remote sensing images. *Open Geospatial Data, Software and Standards* 2, 15.
- He, W., Yokoya, N., 2018. Multi-Temporal Sentinel-1 and -2 Data Fusion for Optical Image Simulation. *ISPRS International Journal of Geo-Information* 7.
- Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. *Neural computation* 9, 1735–1780.
- Hosseini, M., McNairn, H., Mitchell, S., Dingle Robertson, L., Davidson, A., Homayouni, S., 2019. Synthetic aperture radar and optical satellite data for estimating the biomass of corn. *International Journal of Applied Earth Observation and Geoinformation* 83, 101933.
- Hubert-Moy, L., Thibault, J., Fabre, E., Rozo, C., Arvor, D., Corpetti, T., Rapinel, S., 2019. Mapping Grassland Frequency Using Decadal MODIS 250 m Time-Series: Towards a National Inventory of Semi-Natural Grasslands. *Remote Sensing* 11.
- Hufkens, K., Keenan, T.F., Flanagan, L.B., Scott, R.L., Bernacchi, C.J., Joo, E., Brunsell, N.A., Verfaillie, J., Richardson, A.D., 2016. Productivity of North American grasslands is increased under future climate scenarios despite rising aridity. *Nature Climate Change* 6, 710–714.
- Ienco, D., Interdonato, R., Gaetano, R., Ho Tong Minh, D., 2019. Combining Sentinel-1 and Sentinel-2 Satellite Image Time Series for land cover mapping via a multi-source deep learning architecture. *ISPRS Journal of Photogrammetry and Remote Sensing* 158, 11–22.
- Inglada, J., Vincent, A., Arias, M., Tardy, B., Morin, D., Rodes, I., 2017. Operational High Resolution Land Cover Map Production at the Country Scale Using Satellite Image Time Series. *Remote Sensing* 9.
- Jonsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Transactions on Geoscience and Remote Sensing* 40, 1824–1832.
- Julien, Y., Sobrino, J.A., 2019. Optimizing and comparing gap-filling techniques using simulated NDVI time series from remotely sensed global data. *International Journal of Applied Earth Observation and Geoinformation* 76, 93–111.
- Jönsson, P., Eklundh, L., 2004. "timesat—a program for analyzing time-series of satellite sensor data". *Computers & Geosciences* 30, 833–845.
- Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture* 147, 70–90.
- Kamir, E., Waldner, F., Hochman, Z., 2020. Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing* 160, 124–135.
- Kandasamy, S., Baret, F., Verger, A., Neveux, P., Weiss, M., 2013. A comparison of methods for smoothing and gap filling time series of remote sensing observations; application to MODIS LAI products. *Biogeosciences* 10, 4055–4071.
- Kang, S., Running, S., Zhao, M., Kimball, J., Glassy, J., 2005. Improving continuity of MODIS terrestrial photosynthesis products using

- an interpolation scheme for cloudy pixels. *International Journal of Remote Sensing* 26, 1659–1676.
- Karasiak, N., Fauvel, M., Dejoux, J.F., Monteil, C., Sheeren, D., 2020. Optimal dates for deciduous tree species mapping using full years Sentinel-2 time series in South West France. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences V-3-2020*, 469–476.
- Kingma, D.P., Ba, J., 2017. Adam: A Method for Stochastic Optimization.
- Kumar, P., Prasad, R., Gupta, D.K., Mishra, V.N., Vishwakarma, A.K., Yadav, V.P., Bala, R., Choudhary, A., Avtar, R., 2018. Estimation of winter wheat crop growth parameters using time series Sentinel-1A SAR data. *Geocarto International* 33, 942–956.
- Lathuilière, S., Mesejo, P., Alameda-Pineda, X., Horaud, R., 2020. A comprehensive analysis of deep regression. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 42, 2065–2081.
- Lepot, M., Aubin, J.B., Clemens, F.H., 2017. Interpolation in Time Series: An Introductory Overview of Existing Methods, Their Performance Criteria and Uncertainty Assessment. *Water* 9.
- Li, J., Roy, D.P., 2017. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing* 9.
- Li, Y., Li, M., Li, C., Liu, Z., 2020. Forest aboveground biomass estimation using landsat 8 and sentinel-1a data with machine learning algorithms. *Scientific Reports* 10, 9952.
- Liu, L., Dong, Y., Huang, W., Du, X., Ren, B., Huang, L., Zheng, Q., Ma, H., 2020. A Disease Index for Efficiently Detecting Wheat Fusarium Head Blight Using Sentinel-2 Multispectral Imagery. *IEEE Access* 8, 52181–52191.
- Mallet, C., Le Bris, A., 2020. Current challenges in operational very high resolution land-cover mapping. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B2-2020*.
- Martínez-Ferrer, L., Piles, M., Camps-Valls, G., 2020. Crop Yield Estimation and Interpretability With Gaussian Processes. *IEEE Geoscience and Remote Sensing Letters*, 1–5.
- Massawe, E., Mayes, S., Cheng, A., Chai, H., Cleasby, P., Symonds, R., Ho, W., Siise, A., Wong, Q., Kendabie, P., Yanusa, Y., Jamalluddin, N., Singh, A., Azman, R., Azam-Ali, S., 2015. The Potential for Underutilised Crops to Improve Food Security in the Face of Climate Change. *Procedia Environmental Sciences* 29, 140 – 141. *Agriculture and Climate Change - Adapting Crops to Increased Uncertainty (AGRI 2015)*.
- Meng, C., Li, H., 2019. Assimilating Satellite Land Surface States Data from Fengyun-4A. *Scientific Reports* 9, 19567.
- Meraner, A., Ebel, P., Zhu, X.X., Schmitt, M., 2020. Cloud removal in Sentinel-2 imagery using a deep residual neural network and SAR-optical data fusion. *ISPRS Journal of Photogrammetry and Remote Sensing* 166, 333 – 346.
- Mercier, A., Betbeder, J., Rapinel, S., Jegou, N., Baudry, J., Hubert-Moy, L., 2020. Evaluation of Sentinel-1 and -2 time series for estimating LAI and biomass of wheat and rapeseed crop types. *Journal of Applied Remote Sensing* 14, 1 – 21.
- Misra, G., Cawkwell, E., Wingler, A., 2020. Status of Phenological Research Using Sentinel-2 Data: A Review. *Remote Sensing* 12.
- Mohite, J.D., Sawant, S.A., Pandit, A., Pappula, S., 2020. Investigating the performance of Random Forest and Support Vector Regression for estimation of cloud-free NDVI using Sentinel-1 SAR data. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B3-2020*, 1379–1383.
- Moreno, , García-Haro, F.J., Martínez, B., Gilabert, M.A., 2014. Noise Reduction and Gap Filling of fAPAR Time Series Using an Adapted Local Regression Filter. *Remote Sensing* 6, 8238–8260.
- Moreno-Martínez, Á., Izquierdo-Verdiguier, E., Maneta, M.P., Camps-Valls, G., Robinson, N., Muñoz-Marí, J., Sedano, F., Clinton, N., Running, S.W., 2020. Multispectral high resolution sensor fusion for smoothing and gap-filling in the cloud. *Remote sensing of environment* 247, 111901–111901.
- Nasrallah, A., Baghdadi, N., El Hajj, M., Darwish, T., Belhouchette, H., Faour, G., Darwich, S., Mhaweji, M., 2019. Sentinel-1 Data for Winter Wheat Phenology Monitoring and Mapping. *Remote Sensing* 11.
- Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., Hossard, L., 2018. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing* 10.
- Ouaadi, N., Jarlan, L., Ezzahar, J., Zribi, M., Khabba, S., Bouras, E., Bousbih, S., Frison, P.L., 2020. Monitoring of wheat crops using the backscattering coefficient and the interferometric coherence derived from Sentinel-1 in semi-arid areas. *Remote Sensing of Environment* 251, 112050.
- O’Grady, D., Leblanc, M., Gillieson, D., 2013. Relationship of local incidence angle with satellite radar backscatter for different surface conditions. *International Journal of Applied Earth Observation and Geoinformation* 24, 42 – 53.
- Pelletier, C., Valero, S., Inglada, J., Champion, N., Dedieu, G., 2016. Assessing the robustness of random forests to map land cover with high resolution satellite image time series over large areas. *Remote Sensing of Environment* 187, 156–168.
- Pelletier, C., Webb, G.I., Petitjean, F., 2019. Temporal Convolutional Neural Network for the Classification of Satellite Image Time Series. *Remote Sensing* 11.
- Pipia, L., Muñoz-Marí, J., Amin, E., Belda, S., Camps-Valls, G., Verrelst, J., 2019. Fusing optical and SAR time series for LAI gap filling with multioutput Gaussian processes. *Remote Sensing of Environment* 235, 111452.
- Potin, P., Rosich, B., Miranda, N., Grimont, P., Shurmer, I., O’Connell, A., Krassenburg, M., Gratadour, J., 2019. Copernicus Sentinel-1 Constellation Mission Operations Status, in: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, pp. 5385–5388.
- Quintana-Seguí, P., Le Moigne, P., Durand, Y., Martin, E., Habets, F., Baillon, M., Canellas, C., Franchisteguy, L., Morel, S., 2008. Analysis of Near-Surface Atmospheric Variables: Validation of the SAFRAN Analysis over France. *Journal of Applied Meteorology and Climatology* 47, 92–107.
- Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., Prabhat, 2019. Deep learning and process understanding for data-driven Earth system science. *Nature* 566, 195–204.
- Reinermann, S., Asam, S., Kuenzer, C., 2020. Remote Sensing of Grassland Production and Management—A Review. *Remote Sensing* 12.
- Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review* 65, 386–408.
- Roy, D.P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., Lindquist, E., 2008. Multi-temporal MODIS–Landsat data fusion for relative radiometric normalization, gap filling, and prediction of Landsat data. *Remote Sensing of Environment* 112, 3112 – 3130.
- Rußwurm, M., Körner, M., 2020. Self-attention for raw optical Satellite Time Series Classification. *ISPRS Journal of Photogrammetry and Remote Sensing* 169, 421 – 435.
- Scarpa, G., Gargiulo, M., Mazza, A., Gaetano, R., 2018. A CNN-Based Fusion Method for Feature Extraction from Sentinel Data. *Remote Sensing* 10.
- Schmitt, M., Zhu, X.X., 2016. Data Fusion and Remote Sensing: An ever-growing relationship. *IEEE Geoscience and Remote Sensing Magazine* 4, 6–23.
- Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L., 2015. Missing Information Reconstruction of Remote Sensing Data: A Technical Review. *IEEE Geoscience and Remote Sensing Magazine* 3, 61–85.
- Sitokontantinou, V., Koutroumpas, A., Drivas, T., Koukos, A., Karathanassi, V., Kontoes, H., Papoutsis, I., 2020. A Sentinel based agriculture monitoring scheme for the control of the CAP and food security, in: Themistocleous, K., Papadavid, G., Michaelides, S., Ambrosia, V., Hadjimitsis, D.G. (Eds.), *Eighth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2020)*, International Society for Optics and Photonics. SPIE. pp. 48 – 59.
- Solano-Correa, Y.T., Bovolo, E., Bruzzone, L., Fernández-Prieto, D., 2020. A Method for the Analysis of Small Crop Fields in Sentinel-2 Dense Time Series. *IEEE Transactions on Geoscience and Remote Sensing* 58, 2150–2164.

- Sr., S.N., Billey, A., Jr., H.T.O.A., 2018. Random forest classification using Sentinel-1 and Sentinel-2 series for vegetation monitoring in the Pays de Brest (France), in: Neale, C.M.U., Maltese, A. (Eds.), *Remote Sensing for Agriculture, Ecosystems, and Hydrology XX*, International Society for Optics and Photonics. SPIE. pp. 1 – 18.
- Sudmanns, M., Tiede, D., Augustin, H., Lang, S., 2020. Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass. *International Journal of Digital Earth* 13, 768–784.
- Sun, L., Gao, F., Xie, D., Anderson, M., Chen, R., Yang, Y., Yang, Y., Chen, Z., 2021. Reconstructing daily 30 m ndvi over complex agricultural landscapes using a crop reference curve approach. *Remote Sensing of Environment* 253, 112156.
- Tamm, T., Zalite, K., Voormansik, K., Talgre, L., 2016. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. *Remote Sensing* 8.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment* 8, 127 – 150.
- Veloso, A., Mermoz, S., Bouvet, A., Le Toan, T., Planells, M., Dejoux, J.E., Ceschia, E., 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. *Remote Sensing of Environment* 199, 415 – 426.
- Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G., Camps-Valls, G., Moreno, J., 2015. Experimental sentinel-2 lai estimation using parametric, non-parametric and physical retrieval methods – a comparison. *ISPRS Journal of Photogrammetry and Remote Sensing* 108, 260–272.
- Vidal, J.P., Martin, E., Franchistéguy, L., Baillon, M., Soubeyrou, J.M., 2010. A 50-year high-resolution atmospheric reanalysis over France with the Safran system. *International Journal of Climatology* 30, 1627–1644.
- Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., Dorigo, W., Wagner, W., 2020. Sentinel-1 Cross Ratio and Vegetation Optical Depth: A Comparison over Europe. *Remote Sensing* 12.
- Vreugdenhil, M., Wagner, W., Bauer-Marschallinger, B., Pfeil, I., Teubner, I., Rüdiger, C., Strauss, P., 2018. Sensitivity of Sentinel-1 Backscatter to Vegetation Dynamics: An Austrian Case Study. *Remote Sensing* 10.
- Vuolo, F., Ng, W.T., Atzberger, C., 2017. Smoothing and gap-filling of high resolution multi-spectral time series: Example of Landsat data. *International Journal of Applied Earth Observation and Geoinformation* 57, 202 – 213.
- Wang, J., Xiao, X., Bajgain, R., Starks, P., Steiner, J., Doughty, R.B., Chang, Q., 2019. Estimating leaf area index and aboveground biomass of grazing pastures using Sentinel-1, Sentinel-2 and Landsat images. *ISPRS Journal of Photogrammetry and Remote Sensing* 154, 189 – 201.
- Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., Woodcock, C.E., 2012. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment* 122, 2 – 10. *Landsat Legacy Special Issue*.
- Xu, H.J., Wang, X.P., bao Yang, T., 2017. Trend shifts in satellite-derived vegetation growth in Central Eurasia, 1982–2013. *Science of The Total Environment* 579, 1658 – 1674.
- Yang, Y., Tilman, D., Furey, G., Lehman, C., 2019. Soil carbon sequestration accelerated by restoration of grassland biodiversity. *Nature Communications* 10, 718.
- Yang, Y., Wu, T., Wang, S., Li, H., 2020. Fractional evergreen forest cover mapping by MODIS time-series FEVC-CV methods at sub-pixel scales. *ISPRS Journal of Photogrammetry and Remote Sensing* 163, 272 – 283.
- Yin, G., Mariethoz, G., Sun, Y., McCabe, M.F., 2017. A comparison of gap-filling approaches for Landsat-7 satellite data. *International Journal of Remote Sensing* 38, 6653–6679.
- Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multi-spectral satellite data. *Remote Sensing of Environment* 237, 111511.
- Zhang, C., Li, W., Travis, D., 2007. Gaps-fill of SLC-off Landsat ETM+ satellite image using a geostatistical approach. *International Journal of Remote Sensing* 28, 5103–5122.
- Zhang, Y., Song, C., Band, L.E., Sun, G., Li, J., 2017. Reanalysis of global terrestrial vegetation trends from MODIS products: Browning or greening? *Remote Sensing of Environment* 191, 145 – 155.
- Zhao, W., Qu, Y., Chen, J., Yuan, Z., 2020. Deeply synergistic optical and SAR time series for crop dynamic monitoring. *Remote Sensing of Environment* 247, 111952.
- Zhou, J., Jia, L., Menenti, M., 2015. Reconstruction of global MODIS NDVI time series: Performance of Harmonic ANalysis of Time Series (HANTS). *Remote Sensing of Environment* 163, 217 – 228.
- Zhu, X., Helmer, E.H., Gao, F., Liu, D., Chen, J., Lefsky, M.A., 2016. A flexible spatiotemporal method for fusing satellite images with different resolutions. *Remote Sensing of Environment* 172, 165 – 177.