



Challenges in Grassland Mowing Event Detection with Multimodal Sentinel Images

Anatol Garioud, Sébastien Giordano, Silvia Valero, Clément Mallet

► To cite this version:

Anatol Garioud, Sébastien Giordano, Silvia Valero, Clément Mallet. Challenges in Grassland Mowing Event Detection with Multimodal Sentinel Images. 2019 10th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp), Aug 2019, Shanghai, France. pp.1-4, 10.1109/Multi-Temp.2019.8866914 . hal-02387167

HAL Id: hal-02387167

<https://hal.science/hal-02387167>

Submitted on 29 Nov 2019

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.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/336557273>

Challenges in Grassland Mowing Event Detection with Multimodal Sentinel Images

Conference Paper · August 2019

DOI: 10.1109/Multi-Temp.2019.8866914

CITATIONS

0

READS

15

4 authors, including:



Anatol Garioud

Institut national de l'information géographique et forestière

8 PUBLICATIONS 2 CITATIONS

[SEE PROFILE](#)



Sébastien Giordano

Institut national de l'information géographique et forestière

34 PUBLICATIONS 77 CITATIONS

[SEE PROFILE](#)



Clément Mallet

Institut national de l'information géographique et forestière

142 PUBLICATIONS 3,069 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Radar Polarimetry [View project](#)



Trajectory analysis of multi-temporal/multi-modal satellite images for monitoring grassland dynamics [View project](#)

CHALLENGES IN GRASSLAND MOWING EVENT DETECTION WITH MULTIMODAL SENTINEL IMAGES

Anatol Garioud^{1,2}, Sébastien Giordano¹, Silvia Valero², Clément Mallet¹

¹ Univ. Paris Est, IGN-ENSG, LaSTIG/Strudel, Saint-Mande, FRANCE

² CESBIO, Université de Toulouse, CNES/CNRS/IRD/UPS, Toulouse, FRANCE

Abstract—Permanent Grasslands (PG) are heterogeneous environments with high spatial and temporal dynamics, subject to increasing environmental challenges. This study aims to identify requirements, key constraining factors and solutions for robust and complete detection of Mowing Events. Remote sensing is a powerful tool to monitor and investigate Near-Real-Time and seasonally PG cover. Here, pros and cons of Sentinel-2 (S2) and Sentinel-1 (S1) time series exploitation for Mowing Events (MowEve) detection are analysed. A deep-based approach is proposed to obtain consistent and homogeneous biophysical parameter times series for MowEve detection. Recurrent Neural Networks are proposed as regression strategy allowing the synergistic integration of optical and Synthetic Aperture Radar data to reconstruct dense NDVI times series. Experimental results corroborates the interest of deriving consistent and homogeneous series of biophysical parameters for subsequent MowEve detection.

Index Terms—Permanent Grasslands; Mowing Events; Time Series; Sentinel; NDVI; Regression; Recurrent Neural Networks

I. INTRODUCTION AND MOTIVATION

Permanent Grasslands (PG) are defined as areas used permanently (at least 5 years) for cultivated, natural herbaceous forage crops or for growing grasses. Grasslands correspond to 37% of the agricultural areas in Europe. They are major ecosystems supporting a large number of functions, such as carbon storage, biodiversity conservation, renewable energy production, or soil erosion. They provide a considerable number of what are referred to as *ecosystem services*. Mainly, grasslands behave as a significant carbon sink. Strong links between Mowing Events (MowEve), act of levelling or cutting down grass, and soil carbon storage have for instance been demonstrated. Accurate monitoring is relevant since recent decades have witnessed significant modifications such as abandonment and conversion to arable land. MowEve are highly variable in time and space: the farmer's decision to mow depends on many factors. Both yearly and Near Real-Time (NRT) inspections are mandatory to capture the large diversity of behaviours.

Remote sensing allows a synoptic and regular monitoring through systematic acquisitions of Earth Observation imagery. Sentinels of the European Space Agency (ESA) are combining high temporal and fine spatial resolutions. Their free and easily accessible policy makes it possible to envisage regular monitoring of grassland cover [1]. A majority of the literature focuses on discriminating grasslands from other land-cover types [2]. Only few studies address management practises detection. Optical time series of images are privileged [3], even

if combined exploitation of optical and Synthetic Aperture Radar (SAR) data has proven its potential in distinguishing management practices [4]–[6]. Besides focusing on restricted areas [6]–[10], optical and SAR data sources still remain very often independently exploited. Such approaches do not face operational constraints (automation, scalability), and the high heterogeneity of landscapes.

The aim of this paper is threefold: (i) to review main strengths and weaknesses of S2 and S1 data for PG monitoring; (ii) to propose a methodological framework for low-level S2/S1 fusion for dense biophysical feature retrieval; (iii) to assess the proposed strategy over a test site.

II. PROBLEM ANALYSIS

A. Thematic data analysis

Optical Sentinel-2 images: 13 spectral bands are acquired every 5 days. Visible and Near-Infra-Red channels are available at 10m resolution, and strongly correlate to the vegetation cover. The Normalized Difference Vegetation Index (NDVI) measures the vegetation health and is highly relevant [1] for monitoring and visualizing grassland phenology. NDVI time series varies with phenological stages and highlights the different agronomic types and management methods.

The most important limitation in optical time series is the recurrent presence of cloud cover. Cloud-masking techniques are now effective but not faultless (e.g., non-detected thin clouds). Yet, a large number of missing data exists and significantly decreases S2 revisit rate. For instance, the average revisit rate obtained by studying 200 PG parcels comes down to 13 days (ranging from 2 days up to 120 days). Unfortunately, this actual temporal resolution would not allow exhaustive MowEve detection. They are abrupt phenomena with very different natures and duration (Fig. 1). When rapid re-growth of grass or partial mowing are occurring, sparse S2 measurements would thus fail effective monitoring.

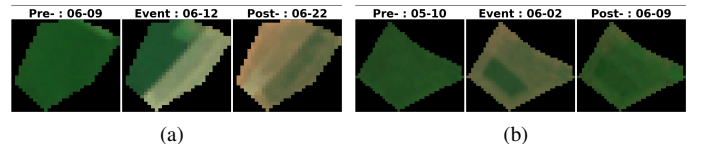


Fig. 1. Several stages of MowEve spanning over S2 acquisitions (Pre-Event, Event, Post-Event). A large diversity of behaviours exists. For example grass cover fast regrowth (b), here within 7 days.

When optical data is not cloud-affected, MowEve are well spotted (Fig. 1) by a significant change in the surface color. Fig. 2 shows another example validating how optical data can

provide key information for MowEve detection. Vertical red lines correspond to MowEve on a NDVI time series. The slope breaks confirms how simple spectral indices such as NDVI are effective for MowEve detection. However, a high temporal resolution is essential. Indeed, sudden movements in time series are more likely to be due to anthropogenic impacts.

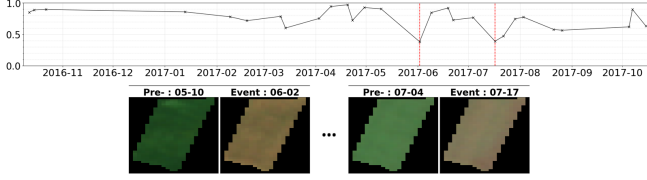


Fig. 2. PG MowEve (red line) example.

SAR Sentinel-1 imagery: Active SAR sensors allows acquiring more dense time series. S1 dual-polarized (VV & VH) SAR systems is independent of illumination conditions or cloud cover. It permits a 3 day revisit rate alternating ascending/descending orbit acquisitions. The measure describing the backscattering coefficient per unit area is σ_0 . It is based on the intensity of emitted signal returning directly to the antenna. It therefore varies accordingly to the reflective strength of the target. Compared with optical NDVI ones, σ_0 time series appear more persistent in time. However, phenological stages are contrarily less clearly discernible. Recently, coherence has appeared to be more suitable for MowEve [4], [6]. Coherence measures changes between pair of images and ranges from 0 to 1. When scatterers remain geometrically and physically stable between two images, coherence estimation tends to be high. Conversely, it decreases with ground-based changes. Therefore, MowEves following pair of images will have a different response and lead to shifts in the time series.

Surface roughness, local slope and orientation, soil moisture, precipitations or frost events will affect the amplitude and phase recorded by the sensor. They have a strong impact on σ_0 and coherence values. Fig. 3 visually correlates daily precipitations (bars) with S1 σ_0 measurements (black lines) over one PG parcel. Abrupt variations in the S1 σ_0 signal are observed after rainy events. In addition, S1 images are also affected by shadow, layover, and foreshortening (occurring differently between ascending and descending orbits) with no possibilities to offset these effects.

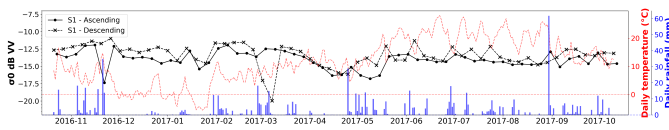


Fig. 3. S1 $\sigma_{0,vv}$ (dB) fluctuates jointly with daily precipitation and temperature (source: Météo France). Strong drops in σ_0 affecting both polarization (end of November, April) coincides with heavy precipitations and probably open water on the parcel.

Mitigating external factors is generally a major challenge in the exploitation of SAR signal due to its high sensitivity. Auxiliary data can partly improve time series interpretation by decreasing the variations of σ_0 and coherence values.

Lack of reference data: despite a high environmental importance, there is yet no national or pan-European initiative that reports on the number of mowing performed on grassland parcels during a season. Field campaigns can help. But they are tedious, costly, and limited in view of the spatial and temporal scales involved. Therefore, any MowEve methodology should sparsely rely on reference data or should be designed in an unsupervised way.

B. Methodological analysis: related work

Both time series of optical and SAR data have been used for MowEve detection. Optical images are traditionally privileged in remote sensing. Classification of PG management practices have been proposed using several SPOT and Quickbird images in [8]. RapidEye time series to classify PG management practices intensity are processed in [7]. S1 and S2 datasets are proposed for exploring correlation between NDVI and σ_0 over PG in [5], and in [4] for monitoring MowEve and grazing occurrences. Some studies integrate optical multi-sensor data to increase the temporal resolution. Radiometric and geometric discontinuities between modalities have then to be considered [2]. To address low optical temporal resolution, exponential amount of work is being done to investigate SAR time series for PG monitoring and MowEve assessment as in [9] with COSMO-SkyMed data, and in [10] for PG classification from RapidEye augmented with TerraSAR-X data. The potential of dual-pol capability of S1 to discriminate MowEve and grazing is discussed in [5]. The interest of S1 coherence variation for MowEve detection with 12-days pairs of images over PG in Estonia is also explored in [6]. It is highlighted that in-situ data is required to interpret coherence behaviour in regard to climate effects.

A vast majority of MowEve detection analysis is based on predefined threshold(s) [4], [8] or conventional regression analysis on handcrafted features (spectral indices). Thresholds are usually locally and manually determined (in space and time). Thus, they do not handle the vast temporal and structural variations of MowEve on larger geographical areas. Furthermore, they usually only address the offline detection case (i.e., retrospective detection), through the full scanning of complete time series.

The instability of thresholding methods can be easily demonstrated by experimentally studying the NDVI time series of 200 contiguous PG plots. It is considered that a MowEve is detected when the difference of two consecutive NDVI values is higher than a predefined value. Table I shows the wide variation in results based on several arbitrarily defined thresholds.

Threshold Value	0.32	0.3	0.28	0.26	0.24	0.22	0.2	0.18
Mean number of MowEve per parcel	1,27	1,42	1,6	1,88	2,3	2,65	3,3	4

TABLE I
MEAN NUMBER (PER PARCEL) OF MOWEVE DETECTED THOROUGH ONE YEAR OF S2 NON-CLOUDY NDVI TIME SERIES AMONG 200 PG.

The problem analysis corroborates how the ability of optical imaging to detect changes in vegetation cover and the consistency of SAR acquisitions must be jointly exploited.

Considering complementary strengths and weaknesses of S1 and S2, methodology must be guided by these requirements:

- High temporal sampling over a full year;
- Both NRT and retrospective detection should be possible;
- Limited supervision: to cope with sparse reference data;
- Interpretability of the time series to end-users;
- Operational constraints: automation & scalability

III. PROPOSED APPROACH

Our approach consists in using the complete S1 time series to derive dense biophysical optical features (here, NDVI) allowing the detection of MowEwe. The new dense temporal super-resolution will allow to cope with missing data and account for fluctuating weather and topographic conditions.

In the literature, regression methods and other statistical parametric modeling approaches are applied, in order to increase temporal resolution. Unfortunately, they do not fulfill all requirements (multimodal, unsupervised, and NRT). Autoregressive ARIMA [11] or MASD [3] models are mishandling sparse datasets [11]. Alternatively, decomposition-based models can be used (e.g., BFAST [12]), but they require explicit analytic functions and known distributions.

Although some improvements have been proposed (multimodality, climate variations), non-parametric Machine Learning approaches have been privileged [13]. Random Forest, Support Vector Machines [4], [10], Gaussian Processes [14] or Neural Networks [15] are part of the solutions used. Recently, Deep Learning methods have proved to successfully generate nonlinear estimation models. Specially, Recurrent Neural Networks (RNNs) [16] allow the use of multimodal datasets and exploit the temporal dimension of time series. They can regress values based on a temporal memory of past observations. Thus, exploration of the potential and flexibility of RNN model is proposed here.

To derive yearly real-time synthetic S2 NDVI ($S2_{NDVI}^{\hat{}}$) time series, the use of a trained RNN model is selected. The training model uses (dense) S1 as input data and (sparse) non-cloudy S2 observations as target. Inputs and targets of training samples corresponds to Sentinels time series pixels values describing PG parcels. The French agricultural farmer declaration (LPIS) allows to extract enough samples to permit supervised training. Our proposed RNN architecture can be fed with external data through attention mechanisms [17]. The resulting trained model allows to estimate NDVI values for cloudy observations by only considering S1 information.

IV. EXPERIMENTS

A. Test site and Dataset

Agricultural season from October 2016 to October 2017 is observed. The test site is located North-East of Lyon in France over a S2 orbit overlap area. This 5,267 km² area offers a high geographical diversity (valleys, hillsides, mountains). The LPIS is used to retrieve the 23,846 PG parcels over the test site. The average size of parcels is 5.1 ha (ranging 1 - 88 ha).

The slope is in-between 0 and 46°, with notable steep areas in the Eastern and Western parts of a river catchment.

56 Level-2A S2 A&B images were downloaded from the French Theia web-portal. They were processed with the MAJA chain, which performs atmospheric correction and provides cloud and shadow masks. 60 ascending and 60 descending S1 Ground Range Detected (GRD) IW dual-pol images were downloaded from ESA's Copernicus Open Access Hub, and pre-processed in order to compute σ_0 time series. In addition, InSAR coherence estimation (in a averaging window size of azimuth×range: 3×9) images with 6 days temporal gap were computed from S1 Single Look Complex (SLC) IW dual-pol images. Produced σ_0 and coherence time series are stacked, but separately for ascending and descending orbits. This avoids viewing angle correction.

B. First results

The proposed method requires the use of a common temporal grid for all satellite data. Therefore interpolation by the nearest neighbour method is done following S1 descending acquisitions. Non cloudy observations are used for S2. This grid consists of 60 dates with a fixed 6-day interval. The dataset is splitted in train and test sets with a 80/20 rule resulting in 19076 samples for training and 4770 samples for validation. Fig. 4 shows the proposed RNN architecture which is composed of two main components. First, a Gated Recurrent Unit (GRU) as a encoder exploiting the temporal dimension of the inputs sequences. Secondly, a Multi Layer Perceptron (MLP) is used as a classifier.

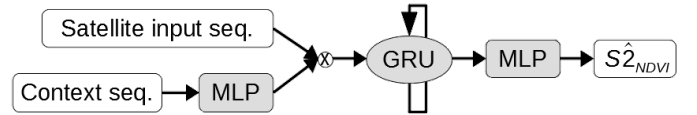


Fig. 4. RNN workflow architecture used for regression.

As shown in Fig. 4, two data sequences are used as input data. All the data is standardized. The Sentinel's sequence is composed by the information of the 60 temporal S1 and S2 acquisitions. Concerning S2, a 60 × 1 vector is composed of NDVI values. In contrast, each S1 acquisitions contains 30 features by computing the means, median and standard deviation for each polarization and orbit (for σ_0 and coherence). Besides, in the case of σ_0 , the three measures are also computed for the VV/VH ratio. A second vector, the context sequence describes temporal information (e.g. day of the year, month, season) for both S1 and S2 acquisitions. This vector is passed through a MLP so that the output of this context sequence is of same size of the Sentinel's sequence. The resulting vector is used as *attention vector* for the S1 input sequence through element wise multiplication in order to temporally weight their importance. Hyperparameters of the network have been chosen empirically. The recurrent GRU cell has a hidden state size of 128. The last hidden state of the GRU network is used as embedding to a final MLP classifier which uses batch normalization, dropout, and ReLu activation on each hidden layer. The last layer of this MLP includes sigmoid

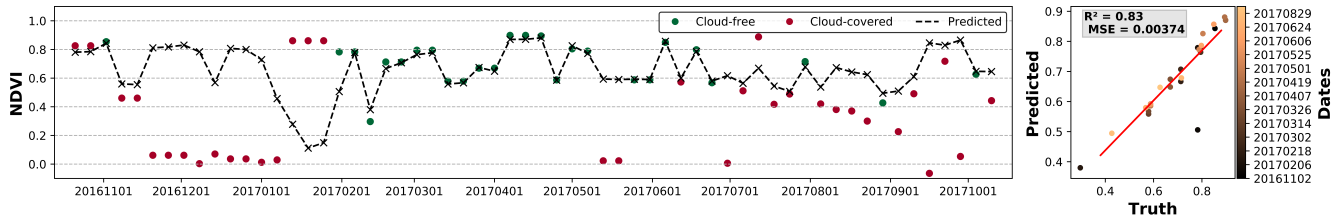


Fig. 5. *Left*: 6-day dense estimation of NDVI values over a year for one parcel. Dots: S2 acquisitions (green: valid and selected for training, red: invalid and removed). Black dashed line: resulting time series from model inference. *Right*: R^2 and MSE score for this parcel. Dots indicate NDVI values along regression line. They are colored with respect to the position in the temporal sequence (dark $\rightarrow$ bright: early $\rightarrow$ late).

activation function. The model is trained for 200 epochs with Adam optimizer set with a batch size of 32 and a learning rate of 10^{-3} . Accuracy during training is measured through Mean Squared Error (MSE) for each epoch. Training the model with a NVIDIA Geforce RTX 2070 GPU takes about 15 minutes. Inference takes less than 10 seconds.

Fig. 5 shows an example of resulting $S2_{NDVI}$ over a PG parcel. The black curve contains 60 values of NDVI over a one-year period, regressed from S1 features. The variations of $S2_{NDVI}$ approximate well the phenological cycle of a PG parcel. On Fig. 5 right, regression line shows a general good correlation between non-cloudy S2 NDVI and matching $S2_{NDVI}$. Although the two worst predicted results are associated to the NDVI values at the beginning of the season, temporal analysis denotes similar results almost regardless of the position in the time series. Model performances have been evaluated on the 4770 validation samples representing each of them a PG parcel. The original non cloudy S2 NDVI values are compared with the $S2_{NDVI}$ values predicted by the trained model. The mean R^2 for all the samples and dates obtains **0.805** and Mean Standard Deviation of **0.158**. Regarding MSE, Mean MSE is **0.0071** but with a same order for Mean Standard Deviation MSE that is **0.0077** expressing the difficulty of the model for some outlying observations.

V. CONCLUSION & FURTHER PERSPECTIVES

Optical time series has proven to be effective for detecting MowEve phenomenon. Unfortunately, temporal resolution of S2 time series is affected by cloud coverage greatly reducing the number of exploitable observations. To address it, a RNN architecture has been proposed to densify available S2 time series by exploiting S1 acquisitions.

The results obtained are promising by predicting $S2_{NDVI}$ values using only S1 SAR features. The reconstruction of dense biophysical features as NDVI can be used as input data for a change detection algorithm. Mainly, $S2_{NDVI}$ obtained in this paper has been with few supervision data and allows to comply with operational constraints. The result is easily interpreted by end users being expressed as a biophysical parameter. Once the model has been trained over a limited time series (e.g. a full year), it will be possible to obtain a $S2_{NDVI}$ value from S1 acquisitions. Our architecture will be improved, in particular with dynamic weighting derived from neighborhood or topological information. The reconstructed

signal will be used to detect MowEve under different geographical conditions.

ACKNOWLEDGMENT

This research has been funded by the Agence pour le Développement Et la Maîtrise de l'Énergie (ADEME) and the Centre National d'Études Spatiales (CNES).

REFERENCES

- [1] I. Ali et al. Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, (6):649–671, 2016.
- [2] P. Griffiths, C. Nendel, and P. Hostert. Intra-annual reflectance composites from Sentinel-2 and Landsat for national-scale crop and land cover mapping. *Remote Sensing of Environment*, 220:135–151, 2019.
- [3] J. Franke, V. Keuck, and F. Siegert. Assessment of grassland use intensity by remote sensing to support conservation schemes. *Journal for Nature Conservation*, 20(3):125–134, 2012.
- [4] R. d'Andrimont, G. Lemoine, and M. van der Velde. Targeted Grassland Monitoring at Parcel Level Using Sentinels, Street-Level Images and Field Observations. *Remote Sensing*, 10(8), 2018.
- [5] L. Stendardi et al. Exploiting Time Series of Sentinel-1 and Sentinel-2 Imagery to Detect Meadow Phenology in Mountain Regions.. *Remote Sensing*, 11(5), 2019.
- [6] T. Tamm et al. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. *Remote Sensing*, 8(10):802, 2016.
- [7] M.-G. Giménez et al. Determination of grassland use intensity based on multi-temporal remote sensing data and ecological indicators. *Remote Sensing of Environment*, 198:126–139, 2017.
- [8] P. Dusseux et al. Identification of grassland management practices from leaf area index time series. *Journal of Applied Remote Sensing*, 8, 2004.
- [9] R. Siegmund et al. Satellite-based monitoring of grassland: assessment of harvest dates and frequency using SAR. *Remote Sensing for Agriculture, Ecosystems, and Hydrology XVIII*, 9998, 2016.
- [10] C. Schuster et al. Grassland habitat mapping by intra-annual time series analysis. comparison of RapidEye and TerraSAR-X satellite data. *International Journal of Applied Earth Observation and Geoinformation*, 34:25–34, 2015.
- [11] J. Haworth and T. Cheng. Non-parametric regression for spacetime forecasting under missing data. *Computers, Environment and Urban Systems*, 36(6):538–550, 2012.
- [12] L. Dutrieux et al. Monitoring forest cover loss using multiple data streams, a case study of a tropical dry forest in Bolivia. *ISPRS Journal of Photogrammetry and Remote Sensing*, 107:112–125, 2015.
- [13] M. Reichstein et al. Deep learning and process understanding for data-driven earth system science. *Nature*, 566(7743):195–204, 2019.
- [14] G. Camps-Valls et al. A survey on gaussian processes for earth-observation data analysis: A comprehensive investigation. *IEEE Geoscience and Remote Sensing Magazine*, 4(2):58–78, 2016.
- [15] A. Taravat, M.-P. Wagner, and N. Oppelt. Automatic Grassland Cutting Status Detection in the Context of Spatiotemporal Sentinel-1 Imagery Analysis and Artificial Neural Networks. *Remote Sensing*, 11(6), 2019.
- [16] E. Ndikumana et al. Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing*, 10(8), 2018.
- [17] X. Yang et al. Deep Multimodal Representation Learning from Temporal Data. *CoRR*, abs/1704.03152, 2017.