Large Scale Mowing Event Detection on Dense Time Series Data Using Deep Learning Methods and Knowledge Distillation

Tilemachos Moumouris[1], Vasileios Tsironis[1], Athena Psalta[1], Konstantinos Karantzalos[1]

[1] NTUA, Remote Sensing Lab, 15772 Zografou Athens, Greece
tmoumouris@mail.ntua.gr
{tsironisbi, psaltaath, karank}@central.ntua.gr

Keywords: Large Scale Monitoring, Mowing Event Detection, Land Use, Knowledge Distillation, Time series Analysis.

Abstract

The intensity of agricultural land use is a critical factor for food security and biodiversity preservation, necessitating effective and scalable monitoring techniques. This study presents a novel approach for large-scale mowing event frequency detection using dense time series data and deep learning (DL) methods. Leveraging Sentinel-2 and Landsat data, we developed a benchmark dataset of over 1,600 annotated parcels in Greece, capturing mowing events through photo-interpretation and Enhanced Vegetation Index (EVI) analysis. Four DL architectures were evaluated, including MLP, ResNet18, MLP+Transformer, and Conv+Transformer, with additional handcrafted features incorporated to assess their impact on performance. Our results demonstrate that the Conv+Transformer architecture achieved the highest improvement when enriched with additional features, while ResNet18 showed a decline in performance under similar conditions. To address data scarcity, we employed knowledge distillation, pre-training models on pseudo-labeled data derived from a dataset in Germany. This process significantly enhanced model performance, with fine-tuned ResNet18 and Conv+Transformer architectures achieving significant performance improvements. This study highlights the importance of architecture selection, feature engineering, and pre-training strategies in time series classification for agricultural monitoring. The proposed methods provide a scalable, non-invasive solution for monitoring mowing events, supporting sustainable land management and compliance with agricultural policies. Future work will explore multimodal data integration and advanced training techniques to further enhance detection accuracy.

1. Introduction

The intensity of land use in agricultural regions is directly related to food security, and biodiversity preservation (Klein et al., 2020). In addition, the Common Agricultural Policy (CAP) requires all European Union (EU) member states to implement a more stringent monitoring program in order to preserve the sustainability of natural resources while maximizing crop production yields. Every growing season, mowing events in cultivated areas serve as a reliable indicator of parcel management; therefore, effective monitoring techniques that enable the effective implementation of national programs must be established. These techniques are required to be not only able to monitor large areas, but also to be non-invasive in terms of not requiring in-situ observation of parcel management intensity. Considering the requirements of parcel monitoring, Earth Observation (EO) data is essential in order to be able to monitor vast areas over an extended period of time (KARAKIZI et al., 2024) and provide decision-makers with the necessary tools.

DL models, as well as machine learning algorithms in general, can decisively support such tasks given their efficiency and effectiveness. In recent years, DL-based methods have been predominantly used for mowing event detection. Previous studies on large-scale areas tend to focus on selecting the optimal EO data source combination for the task. EO data has been widely used for monitoring and managing agricultural areas such as grasslands using different methods of machine learning (Ali et al., 2016). Additionally, referring to the use of satellite data for practice monitoring, Vegetation Indexes and specifically of Normalized Difference Vegetation Index (NDVI) was highlighted by (Ottosen et al., 2019).

The study of (Lange et al., 2022) have shown that Sentinel-2 (S2) data could be utilized for land-use intensity (LUI) paired with DL methods such as Convolutional Neural Networks (CNN) that resulted in accuracy of 68% for mowing detection. (Andreatta et al., 2022) also studied the use of S2 data for the task of mowing frequency detection, showing that spatial resolution plays a significant role in the prediction of their proposed algorithm. (Watzig et al., 2023) also proposed a method that, although not based on DL, gives better insight into data usage (S2) and small-scale parcels due to geography. In terms of data quality, inaccuracies in cloud masking could significantly lower the ability of these methods to detect events accurately, while in the meantime, excluding these areas could also lead to big fluctuations in NDVI values, leading to poor results (Kolecka et al., 2018). Additionally, (Halabuk et al., 2015) used MODIS-derived NDVI and EVI along with a simple CART classifier, obtaining an accuracy of 85% in the case of NDVI. LandSat8 (L8), along with S2 data, were also used by (Schwieder et al., 2022) to map mowing events at a national scale, using a rule based (non-ML) method that best describes the events under study. A significant study by (Griffiths et al., 2020) utilized imagery collected by both S2 and L8 in the form of timeseries to map mowing events in Germany, both in terms of frequency and timing, using machine learning methods. Similarly, (Liu et al., 2020) developed algorithms that take advantage of L8 and S2 time series to map mowing event intensities in China, showing that climate change alters parcel activities at a global scale.

The use of Synthetic Aperture Radar (SAR) data derived from Sentinel-1 (S1) was studied by (Taravat et al., 2019) to extract the status of grassland cutting using artificial neural networks, particularly multilayer perceptron (MLP), achieving an accuracy of 85.71%.

Combination of S1 and S2 data as well as field evaluation of results was done by (De Vroey et al., 2022), yielding a F1-score of 79% for a specific type of crop through an object-based approach. (Lobert et al., 2021) studied the combination of different EO data modalities such as optical (S2, L8) and radar (S1), proving that using NDVI along with SAR data, yielded the better results compared to utilizing solely NDVI. Apart from optical or radar data (Holtgrave et al., 2023) included additional data sources such as weather data, showing that more data sources do not implicitly lead to better results, with optical data proving to be the most crucial EO data source for the task.

In this work, we propose a novel approach for large-scale mowing event detection by leveraging dense time series data, deep learning architectures, and knowledge distillation techniques to enhance model performance in data-scarce environments. In addition, we introduce a benchmark dataset of over 1,600 annotated parcels from three agriculturally significant regions in Greece, capturing mowing events through photo-interpretation and Enhanced Vegetation Index (EVI) analysis. Dataset as well as code implementation for this work are available under the MIT license at https://github.com/rslab-ntua/mowing-event-detection.

2. Datasets

2.1 Novel Mowing-Event detection dataset in Greece

As a first step in our work, we defined the Regions of Interest (ROIs) in Greece as shown in Figure 1, for which data are collected and annotated, taking into account the intensity of agricultural activity. Moreover, three different areas were chosen to further capture any spectral differences present. Specifically these areas are located in Central and Southern-Western Greece, Thessaly and Peloponnese; areas which host a significant part of agricultural activity. The data source for this study was NASA's Harmonized LandSat Sentinel Dataset (Claverie et al., 2018) for a period spanning from April to November 2020. During dataset creation, pixels with high cloud coverage were excluded. The EVI was calculated for each time step of the time series. During this phase, each parcel was assigned a label corresponding to the number of mowing events that occurred during the study period. Events were identified via photo-interpretation from optical data in synergy with the computed EVI.

Annotation process posed several challenges in terms of most suitable parcels selection. At first, cloud masking created gaps in the timeseries that could result in false positives. Another issue we encountered was the small scale of a significant number of parcels leading to low-quality observations, especially in parcel borders. This issue is common due to the way parcels are arranged in agricultural areas in Greece, posing challenges for small-parcel monitoring using HLS imagery.

After annotation, we impose a temporal pre-processing to derive the median value for each timestep, resulting in a single vector representation for each parcel containing 45 timesteps.

Observations containing missing values were preserved as different approaches, e.g. data imputation, could be tested. Finally, the dataset we introduce with this study consists of more than 1600 parcels across the three selected geographical areas. This dataset contains timeseries for 5 different classes, 0 to 4 events during the season. Furthermore, this dataset can be used for classification for low intensity as well as heavily managed parcels. An 80-10-10 (train-validation-test) split was utilized for this study after a random split of the whole dataset.

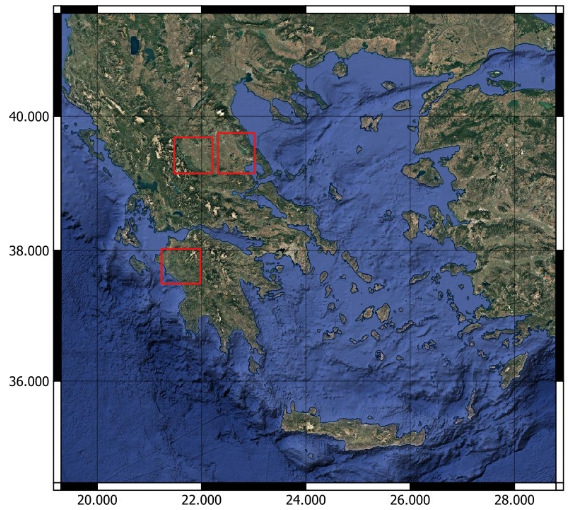


Figure 1. Different ROIs in Greece, where data were collected and annotated.

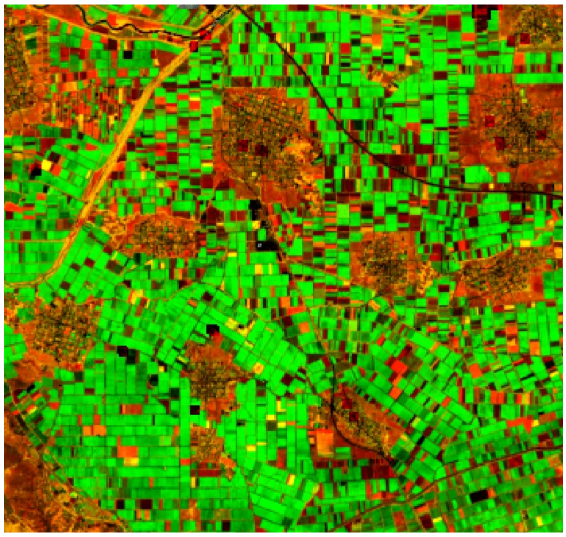


Figure 2. Computed EVI in central Greece during annotation.

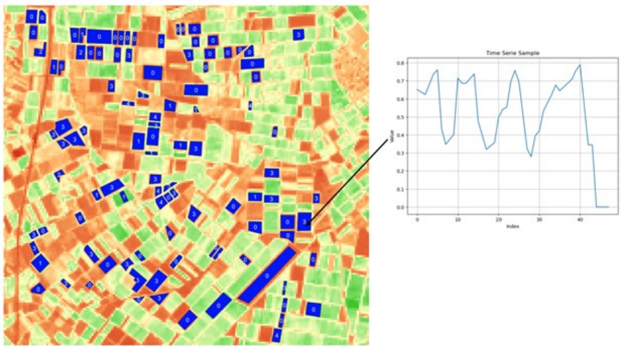


Figure 3. Gap filled timeseries with added padding after dataset creation.

2.2 Additional datasets

To perform knowledge distillation on our selected models, we utilized a dataset comprised of predictions rather than actual ground-truth labels, as described in (Schwieder et al., 2022). This dataset includes pseudo-labels generated for over 50,000 samples in West Germany. To leverage this dataset effectively, we formulated a knowledge distillation process that incorporates these pseudo-labels, aligning them with our benchmark dataset through HLS-based pre-processing to derive time series features for pre-training.

3. Methodology

3.1 Model selection

In this study we focused on different DL-based architectures to model relations between time-steps in the most efficient way, while keeping the computational cost low.

Multi-Layer Perceptron (MLP) (Goodfellow et al., 2016): This was the initial architecture used in our study in order to set the baseline results for the next experiments. This simple yet efficient architecture was built of 2 hidden layers, followed by the network output. Initial Learning Rate was set to 1×10^{-3} and the ReLU activation function. During the second round of experiments, where handcrafted features were used, the input was flattened to be utilized by the network.

ResNet 1D: It is a well-studied architecture (Kiranyaz et al., 2021) and was selected due to its ability to process multiple features per step as well as intra-step time relations (Di Mauro et al., 2017). It is similar to that of (Zhang et al., 2021) used for Electrocardiogram (ECG) diagnosis. In our implementation ReLU activation function was used, as well as Batch Normalization. Stride was set to 1 for the start of the experiments, kernel size was set to 7, padding used was **same**, and the Learning Rate value was 1×10^{-3}.

MLP followed by a Transformer architecture (MLP + Transformer): Taking advantage of the simplicity the MLP offers and the abilities of the Transformer Encoder architecture, this model was utilized to enhance the classification performance, leveraging the cooperation of both basic architectures. The MLP part was built using 2 hidden layers that encoded each step in order to be used later. The initial Learning Rate was set to 1×10^{-4} paired with Dropout value of 0.1. Additionally, a standard Feed-Forward Layer of size 2048 was utilized for the transformer.

1D CNN followed by a Transformer architecture (Conv + Transformer): Finally, another hybrid network was tested similar to that of (Safari et al., 2020), that is composed of 1D CNN and Transformer Encoder one on top of another. In this architecture similar values for the hyperparameters were used as a starting point regarding the Transformer encoder part. The convolution-encoder part was constructed with two layers. Both layers' stride was set to 1 and a same kernel size (3). The output channels from the convolutional part were 64 and used as input to the transformer encoder layers.

Regarding the two transformer-based architectures MLP and CNN networks respectively act as an embedding layer prior to Transformer input. The latter, in particular, leverages an additional attention pooling layer at the final part of the network.

3.2 Feature Selection

During our initial experiments, only the EVI time series features were used without any additional features to establish an early baseline prior to further experimentation, while any temporal gaps were filled using interpolation. In succession, we enriched our benchmark dataset by incorporating a series of several handcrafted features calculated over the dataset's reference period. The final list of features included in our dataset is the following:

1. EVI: The basis of our selected input features.
2. Difference to the mean for each time step: Aim to better combat sudden but small drops in EVI that could lead to False Positives.
3. Difference to the maximum for each time step: This feature was used as it helped better model drops in the index that are not caused by unmasked clouds.
4. Difference to the minimum for each time step: In this case, minimum was selected to further enhance the networks ability to reject False Positives.
5. Difference to the next for each time step: Finally, this feature was utilized to better describe temporal relations in timesteps.

These features selected proved to better model the relations between subsequent time-steps while their on-the-fly computation minimizes computational and storage needs.

3.3 Knowledge Distillation

To address the challenge of limited data availability we are introducing a model pre-training step through knowledge distillation in a typical teacher-student setup. This approach allows us to leverage additional information from pseudo-labeled data, enhancing the performance of our models in a data-scarce environment. As detailed in subsection **2.2**, we leveraged external datasets for pretraining the two best-performing models, ResNet and Conv+Transformer. The pseudo-labels, generated during the inference step of the algorithm proposed in (Schwieder et al., 2022), were employed as a teacher model to facilitate the transfer of knowledge to our student models.

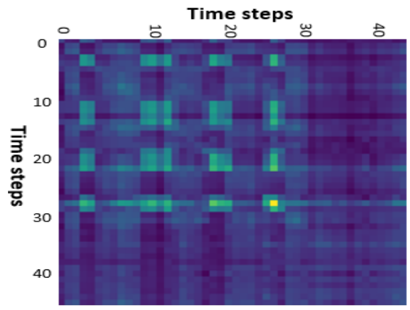
The straightforward teacher-student training scheme allows the pre-training of our models on the pseudo-labels derived from the teacher model in (Schwieder et al., 2022). This step enabled us to expose our models to a broader range of data representations, even in the absence of additional rigorously labeled data.

Finally, a fine-tuning step was conducted on our dataset to further refine the models and evaluate the performance gains achieved through the knowledge distillation pre-training. This final step allowed us to quantify the improvements in model performance and validate the effectiveness of the knowledge transfer process.

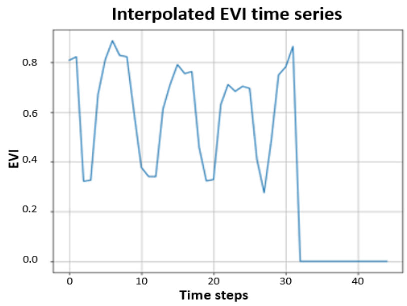
3.4 Activation Maps

For further insights into the learned representations by our networks, activation and attention maps were extracted for the various architectures validated in this study. This information aims to gain a deeper understanding relevant to the temporal regions each architecture is more likely to focus on. Resulting maps reveal that our models' activations correlate to the peaks of the EVI index.

This behavior can yield satisfactory results but also exhibits weaknesses, particularly in cases where peaks are not directly associated to mowing events. Specifically, peaks may additionally be caused by weather phenomena or inaccuracies in dataset masking that fail to eliminate shadows or other effects influencing the index in the examined parcel.



(a)



(b)

Figure 4. (a) Attention map extracted from Transformer Encoder architecture (MLP encoder), (b) corresponding input EVI time series.

3.5 Training Configuration

The training process was designed to solve a classification problem, where the goal was to predict the number of mowing events (classes) for each parcel. To achieve this, we employed the categorical cross-entropy (CCE) loss function, which is well-suited for multi-class classification tasks.

Initially, all models were trained and evaluated using the full dataset without incorporating any additional features. This configuration served as a baseline, providing a reference point for subsequent experiments. In the next configuration, additional handcrafted features were introduced into the training process to assess their impact on model performance. This allowed us to evaluate whether the inclusion of these features contributed to any measurable improvements in classification accuracy. Finally, we focused on refining the two best-performing models from the previous configurations. This refinement included leveraging external pseudo-labeled datasets through a knowledge distillation process to enhance the models' performance under data-scarce conditions.

All models were implemented using PyTorch, and the experiments were conducted on an NVIDIA 3060 GPU with 6GB of memory, which was sufficient to meet the computational requirements of the study.

4. Results

Following the training process, each model was evaluated on the validation set. Table 1 presents a performance comparison, detailing the F1-score, Precision, and Recall metrics achieved by each architecture, and indicates whether additional features were incorporated during training.

4.1 MLP Architecture

To establish a baseline for time series classification, a basic MLP model was employed. During the initial experiments, this architecture achieved a satisfactory classification performance, yielding an F1 score of 60.1%. When additional handcrafted features were incorporated into the input, a significant improvement of +12.7% in the F1 score was observed. Such significant improvement highlights the value of engineered features for enhancing the performance of simpler architectures like MLP, even in basic setups.

4.2 1D ResNet18

The next architecture that produced noteworthy results was the 1D ResNet18. In the initial training iteration, without the inclusion of any extra handcrafted features, this model achieved an impressive F1 score of 80.6%. However, when the input was enriched with additional features, resulting in a total of five input channels, a decrease in performance was observed. The observed findings suggest that the extra handcrafted features may not align well with the model's internal feature extraction mechanisms, potentially introducing noise or redundancy that hindered its performance in this specific case.

4.3 MLP+Transformer

The combination of an MLP followed by a Transformer Encoder proved particularly effective for encoding the temporal steps of the time series data. This architecture demonstrated an improved F1 score compared to the baseline MLP model, showcasing its ability to model inter-time relationships more efficiently. The inclusion of the MLP layer before the Transformer Encoder appears to have facilitated the extraction of meaningful representations, which were then further refined by the Transformer to capture temporal dependencies.

4.4 Conv+Transformer

Without additional handcrafted features, the Conv+Transformer architecture achieved an F1 score of 67.8%. In contrast to the 1D ResNet18 (which saw decreased performance with added features), the Conv+Transformer benefited substantially from incorporating the selected handcrafted features. This resulted in a significant performance improvement, increasing the F1 score by +12.8% to reach 80.6%. This finding underscores the adaptability of the Conv+Transformer architecture, highlighting its capacity to effectively leverage engineered features, likely due to its ability to integrate both spatial and temporal information.

4.5 Knowledge Distillation

The knowledge distillation pre-training process demonstrated its potential to enhance model performance, particularly in scenarios with limited labeled data. Without fine-tuning, the results indicated that it is feasible to learn meaningful and transferable features during the pre-training phase, leveraging the knowledge distillation process. After fine-tuning and with the

Architecture		Extra Features	F1	Precision	Recall
Our Dataset	**MLP**	✗	60.1	60.0	63.2
	MLP	✓	72.8	72.2	75.0
	MLP+Transformer	✓	75.6	75.1	79.4
	ResNet18 1D	✗	80.6	79.5	82.7
	ResNet18 1D	✓	79.1	79.2	80.7
	Conv+Transformer	✗	67.8	69.7	75.3
	Conv+Transformer	✓	80.6	80.7	81.5
Pre-trained w/ fine-tuning	**ResNet Pre-trained**	✓	82.3	81.2	84.4
	Conv+Transformer Pre-trained	✓	82.3	81.8	83.7

Table 1. Performance comparison of various models on our benchmark dataset, including the impact of fine-tuning and extra features.

inclusion of additional features in the input, both the ResNet18 and Conv+Transformer architectures exhibited further improvements in performance, with F1-score increases of +3.2% and +1.7%, respectively. These findings highlight the effectiveness of knowledge distillation as a pre-training strategy, enabling models to better utilize available data and improve their overall classification accuracy.

4.6 Attention-Activation Maps

Basic model behavior was studied by extracting activation maps and correlation matrices from ResNet and Transformer based models respectively. Experiments showed that models tend to attend to spikes in timeseries resulting in low performance, especially when those sudden spikes are not followed by an event that indicated a mowed parcel. Despite the weaknesses models performed well in cases when the mowing intensity was heavier, which in some cases results in cleaner patterns of drops and spikes in EVI.

5. Conclusions and Future Work

In our study, we experimented with popular architectures to study their performance in classification of time series data. MLP was the first kind of model to be trained and showed that even without any additional data as input, it was able to yield Recall of 63.2%. Further experimentation with the same network but with extra features as input, resulted in higher Recall that reached 75.0% in the test set. After performing our study with the process of encoding initial timesteps into embeddings, our results showed that long temporal relations between timesteps could be detected not only with the use of an MLP network as an encoder but also by a 1D CNN, with each of the previous encoder networks followed by a Transformer Encoder. The final architecture was that of a 1D ResNet 18, which revealed that extra features acted as noise, deteriorating overall performance and future studies might consider avoiding adding such features. A significant part of our study was dedicated to studying whether the pre-training of various architectures has an impact on the final performance of these models. To further investigate their behavior we chose ResNet 1D and the Conv+Transformer architectures and trained them with a considerably larger dataset which was produced by the model developed in (Schwieder et al., 2022). Data used in that phase, contains a large amount of labels that do not correspond to the actual mowing events that have occurred in a parcel or larger area and thus are considered pseudo- labels. Results showed that this simple method of knowledge distillation has a significant impact on model performance after fine-tuning on our benchmark dataset.

Although significant progress has been made regarding the task of mowing event detection, there are potential areas that could be explored or improved as a part of a future study. Firstly, the developed dataset must be enriched with more samples that cover even more agricultural areas in Greece to better combat any spectral profile differences between ROIs. Secondly, DL models are promising in finding patterns in large quantities of data. This could be exploited in order to integrate different data sources in our workflow such as radar data, where the synergy with modern architectures could lead to better results.

As a conclusion, we demonstrated the importance of architecture selection and the significance of data availability for mowing event detection. In addition, we establish a method that takes advantage of pseudo-labels for efficient pre-training of models used with limited access to ground truth training data. Next steps should include the use of multimodal data and different training techniques to further improve detection accuracy. Concluding, we hope our study will support the use of deep learning methods in the field of large scale monitoring with the maximum use of open-access data.

6. Acknowledgments

Part of this research was supported by the research project Bi-CUBES “Analysis-Ready Geospatial Big Data Cubes and Cloud-based Analytics for Monitoring Efficiently our Land & Water” funded by HFRI (grant: 03943).

References

Ali, I., Cawkwell, F., Dwyer, E., Barrett, B., Green, S., 2016. Satellite remote sensing of grasslands: from observation to management. *Journal of Plant Ecology*, 9(6), 649-671. https://doi.org/10.1093/jpe/rtw005.

Andreatta, D., Gianelle, D., Scotton, M., Vescovo, L., and, M. D., 2022. Detection of grassland mowing frequency using time series of vegetation indices from Sentinel-2 imagery. *GIScience & Remote Sensing*, 59(1), 481–500. https://doi.org/10.1080/15481603.2022.2036055.

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.

De Vroey, M., de Vendictis, L., Zavagli, M., Bontemps, S., Heymans, D., Radoux, J., Koetz, B., Defourny, P., 2022. Mowing detection using Sentinel-1 and Sentinel-2 time series for large scale grassland monitoring. *Remote Sensing of Environment*, 280, 113145.

Di Mauro, N., Vergari, A., Basile, T. M. A., Ventola, F. G., Esposito, F. et al., 2017. End-to-end learning of deep spatio-temporal representations for satellite image time series classification. *DC@ PKDD/ECML*.

Goodfellow, I., Bengio, Y., Courville, A., Bengio, Y., 2016. *Deep learning*. 1Number 2, MIT press Cambridge.

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.

Halabuk, A., Mojses, M., Halabuk, M., David, S., 2015. Towards Detection of Cutting in Hay Meadows by Using of NDVI and EVI Time Series. *Remote Sensing*, 7(5), 6107–6132. https://www.mdpi.com/2072-4292/7/5/6107.

Holtgrave, A.-K., Lobert, F., Erasmi, S., Röder, N., Kleinschmit, B., 2023. Grassland mowing event detection using combined optical, SAR, and weather time series. *Remote Sensing of Environment*, 295, 113680.

KARAKIZI, C., KARANTZALOS, K., KANDYLAKIS, Z., 2024. Exploiting Multitemporal Multispectral High-resolution Satellite Data toward Annual Land Cover and Crop Type Mapping: A Case Study in Greece. *Multitemporal Earth Observation Image Analysis: Remote Sensing Image Sequences*, 81–122.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M., Inman, D. J., 2021. 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151, 107398.

Klein, N., Theux, C., Arlettaz, R., Jacot, A., Pradervand, J.-N., 2020. Modeling the effects of grassland management intensity on biodiversity. *Ecology and Evolution*, 10(23), 13518-13529. https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.6957.

Kolecka, N., Ginzler, C., Pazur, R., Price, B., Verburg, P. H., 2018. Regional Scale Mapping of Grassland Mowing Frequency with Sentinel-2 Time Series. *Remote Sensing*, 10(8). https://www.mdpi.com/2072-4292/10/8/1221.

Lange, M., Feilhauer, H., Kühn, I., Doktor, D., 2022. Mapping land-use intensity of grasslands in Germany with machine learning and Sentinel-2 time series. *Remote Sensing of Environment*, 277, 112888.

Liu, L., Xiao, X., Qin, Y., Wang, J., Xu, X., Hu, Y., Qiao, Z., 2020. Mapping cropping intensity in China using time series Landsat and Sentinel-2 images and Google Earth Engine. *Remote Sensing of Environment*, 239, 111624.

Lobert, F., Holtgrave, A.-K., Schwieder, M., Pause, M., Vogt, J., Gocht, A., Erasmi, S., 2021. Mowing event detection in permanent grasslands: Systematic evaluation of input features from Sentinel-1, Sentinel-2, and Landsat 8 time series. *Remote Sensing of Environment*, 267, 112751.

Ottosen, T.-B., Lommen, S. T., Skjøth, C. A., 2019. Remote sensing of cropping practice in Northern Italy using time-series from Sentinel-2. *Computers and Electronics in Agriculture*, 157, 232-238.

Safari, P., India, M., Hernando, J., 2020. Self-attention encoding and pooling for speaker recognition.

Schwieder, M., Wesemeyer, M., Frantz, D., Pfoch, K., Erasmi, S., Pickert, J., Nendel, C., Hostert, P., 2022. Mapping grassland mowing events across Germany based on combined Sentinel-2 and Landsat 8 time series. *Remote Sensing of Environment*, 269, 112795.

Taravat, A., Wagner, M. P., Oppelt, N., 2019. Automatic Grassland Cutting Status Detection in the Context of Spatiotemporal Sentinel-1 Imagery Analysis and Artificial Neural Networks. *Remote Sensing*, 11(6).

Watzig, C., Schaumberger, A., Klingler, A., Dujakovic, A., Atzberger, C., Vuolo, F., 2023. Grassland cut detection based on Sentinel-2 time series to respond to the environmental and technical challenges of the Austrian fodder production for livestock feeding. *Remote Sensing of Environment*, 292, 113577.

Zhang, D., Yang, S., Yuan, X., Zhang, P., 2021. Interpretable deep learning for automatic diagnosis of 12-lead electrocardiogram. *iScience*, 24(4), 102373.