*remote sensing*

MDPI

*Article*

# Satellite-Based Detection of Farmland Manuring Using Machine Learning Approaches

**David Marzi and Fabio Dell'Acqua ***

Department of Electrical, Computer and Biomedical Engineering, University of Pavia, I-27100 Pavia, Italy; david.marzi01@universitadipavia.it
* Correspondence: fabio.dellacqua@unipv.it

**Abstract:** In agriculture, manuring offers several benefits, which include improving soil fertility, structure, water retention, and aeration; all these factors favor plant health and productivity. However, improper handling and application of manure can pose risks, such as spread of pathogens and water pollution. Mitigation of such risks requires not only proper storage and composting practices, but also compliance with correct application periods and techniques. Spaceborne Earth observation can contribute to mapping manure applications and identifying possible critical situations, yet manure detection from satellite data is still a largely open question. The aim of this research is an automated, machine learning (ML)-based approach to detecting manure application on crop fields in time sequences of spaceborne, multi-source optical Earth Observation data. In the first stage of this research, multispectral data alone was considered; a pool of different spectral indexes were analyzed to identify the ones most impacted by manure application. Increments of the selected indexes from one satellite acquisition to the next were used as features to train and test various machine learning models. Two agricultural areas—one in Spain and one in Italy—were considered. Fair levels of accuracy were achieved when training and testing were carried out in the same geographical context, whereas ML models trained on one context and tested on the other reported significantly lower—albeit still acceptable—accuracy levels. In the stage that followed, thermal data was integrated and used alongside multispectral indexes. This addition led to significant improvements in accuracy levels, despite possible thermal-to-multispectral sampling mismatch in time series. Our results appear to indicate that ML-based approaches to manuring detection from space require training on the targeted geographical context, although transfer learning can probably be leveraged and only fine-tuning training will be needed. Spaceborne thermal data, where available, should be included in the input data pool to improve the quality of the final result. The proposed method is meant as a first step towards a suite of techniques that should enable large-scale, consistent monitoring of agricultural activities to check compliance with environmental regulations and provide enhanced traceability information for food products.

**Keywords:** agriculture; agricultural operations; crop fields; earth observation; traceability; machine learning; manure application; multi-source satellite sensing

check for updates

Academic Editor: Jochem Verrelst

Received: 17 December 2024
Revised: 9 March 2025
Accepted: 11 March 2025
Published: 15 March 2025

**Citation:** Marzi, D.; Dell'Acqua, F. Satellite-Based Detection of Farmland Manuring Using Machine Learning Approaches. *Remote Sens.* **2025**, *17*, 1028. https://doi.org/10.3390/rs17061028

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

## 1. Introduction

Agriculture is vital for human survival and the global economy. In 2020, this primary sector was valued at nearly 3.8 billion USD [1]. With the world's population projected to exceed 10 billion by 2060 [2], the demand for food is set to rise, putting pressure on farmers

to increase crop yields while minimizing environmental impact. This pressure has led to widespread use of fertilizers, as well as to deforestation [3].

Fertilizers, proven beneficial in agriculture since ancient times [4], include exogenous organic matter (EOM) like minerals, compost, wood ash, and manure. These substances enhance farm productivity by enriching soil nutrients and improving its chemical and physical properties [5]. However, excessive fertilizer use can harm the environment, causing soil degradation, water pollution, and greenhouse gas emissions [6]. Nitrogen is a key factor for plant growth, but when in excess it can have detrimental effects on both humans and the environment. In Europe, agricultural nitrogen contributes significantly to water pollution, with nitrates and organic nitrogen compounds from fertilizers and manure seeping into groundwater and flowing into surface water [7].

Intense nitrate pollution can render water undrinkable and foster excessive algae growth in aquatic ecosystems, leading to eutrophication [8]. This harms biodiversity, fisheries, and recreation, emphasizing the need to combat excess nitrogen in the environment. To address this issue, the European Union (EU), for example, issued the *nitrates directive* [9] to regulate fertilizer and manure use. One proposed solution is to implement *closed periods* for organic fertilizer in nitrate-vulnerable zones [10]. Closed periods restrict fertilizer spreading to specific time windows, depending on local features and situations. Enforcement of closed periods requires consistent monitoring by designated authorities. In Europe, this is typically realized through ground-based surveys or random field assessments. This however requires investing significant time and resources, particularly in large agricultural areas, where extensive, thorough checks are simply not feasible. On-site visits, while practical in certain contexts, are insufficient to ensure full compliance across an entire country [11,12]. This is where spaceborne Earth Observation (EO) techniques offer a significant advantage. By leveraging remote sensing approaches, authorities can achieve cost-effective, large-scale monitoring of agricultural parcels. The use of free satellite data, such as those provided by Sentinel-1, Sentinel-2, and Landsat-8, makes this approach not only scalable but also economically feasible [13–18].

The capability of multispectral satellite imagery, particularly from Sentinel-2, to identify areas where livestock manure is applied has been demonstrated in several studies, as discussed in the following. A pivotal study [19] carried out a preliminary identification of spectral indexes that can support the detection of manuring events, focusing on farmlands located in France. This research demonstrated that Sentinel-2 imagery can effectively detect changes in spectral characteristics resulting from the application of exogenous organic matter (EOM) such as Green Waste Compost (GWC) and livestock manure. Specifically, the analysis revealed significant spectral differences in Sentinel-2 datasets acquired before and after the application of EOM. This points at satellite imagery as a possible means to identify when and where these amendments are applied in agricultural landscapes. The study, however, also reports limitations. For example, Sentinel-2 struggles to detect low EOM application rates, and its spatial and temporal resolution may miss rapid agricultural changes, affecting accuracy. Moreover, soil, crops, and environmental variability reduce re-usability of spectral indices across different regions. Since the research was conducted in a specific region of France, the findings may not apply to areas with different climates or agricultural practices.

Another study [20] achieved an *F-1* test score of around 89% by developing a pixel-wise manure detection method between two consecutive Sentinel-2 acquisitions. The authors suggested that specific wavelengths in the near-infrared and visible spectrum were most effective for capturing the unique reflectance characteristics of freshly manured fields. Moreover, their results highlighted the importance of timing in data acquisition. In fact, the effectiveness of remote sensing in detecting manure decreases over time as the manure

decomposes and its spectral signature fades. However, limitations include environmental variability, limited field validation, and most importantly, *a priori* knowledge of fertilization dates, which is not always available.

Another interesting study reported in [21] introduces the Manure Spread Index (MSI), a new spectral index that can effectively detect areas where livestock manure and digestate have been spread. This index is based on the spectral response of bare soil compared to that of the manure and digestate, demonstrating its potential usefulness in agricultural monitoring. The proposed spectral index achieves an accuracy of 62.53% when detecting manure spreading, which is considered a promising result. This level of accuracy, however, may not be sufficient for all applications, particularly in precision agriculture where effective decision-making requires higher levels of confidence. Moreover, the effectiveness of the spectral index relies on the availability and accuracy of ancillary datasets, such as soil moisture and precipitation. Variations or inaccuracies in these datasets could impact the reliability of the spectral index in detecting manure spreading.

The use of multispectral data alone is often the primary choice for detecting manure cover over crop fields, our literature review found. This preference stems from the ability of optical sensors to effectively capture the spectral responses of vegetation and soil to organic amendments, which are key indicators of manure application. In particular, our review also highlighted that the use of radar and thermal data for this purpose has been limited, with no more than one study integrating these data types; in [22], the authors utilized random forest models that combined C-band Synthetic Aperture Radar (SAR) data, multispectral and thermal imagery, along with other predictive variables, to monitor soil moisture conditions indicative of liquid manure applications. This analysis focused on sprayfields in eastern North Carolina, USA, where distinct patterns of manure application across various fields and timeframes were identified. By leveraging different types of remotely sensed data, the study was able to derive both the timing and location of manure applications, providing insights into the frequency and scheduling of these events.

Inspired by this approach, our study builds on these findings by applying similar techniques to detect fresh, solid cow manure applications. We considered several aspects of this prior work, particularly its use of integrated data sources, as we adapted the methodology to address the specific challenges associated with identifying solid manure applications. Our work presents an innovative method for detecting fresh manure application events through multi-time series analysis, leveraging existing indexes based on optical and thermal data and combining them in an innovative manner. To the best of our knowledge, this approach has not been previously applied in this context, representing a novel contribution to the field. Another novelty proposed in the manuscript is the identification of most suitable combinations of indexes for manuring detection. This enables the use of a smaller number of input features, making the models more efficient and easier to interpret. Integrating this method with the approach outlined in [20] may lead to detecting manure applications over various time series and determining their spatial distribution on farmland within specified time frames.

To achieve these goals, the study leverages machine learning (ML)-based classification techniques, which are widely used in Earth Observation (EO) applications. A range of ML models were implemented, trained, and evaluated using several performance metrics, including their ability to generalize across data sets [23]. This comprehensive assessment allows for a comparison of model strengths, ending up with the most effective approach for detecting and mapping manure application events over time. This paper is organized as follows: in the next section, the data and the novel proposed method are presented. In Section 3, results of the method application are presented with an extensive discussion. Section 4 draws some conclusions based on the data produced and its analysis.

## 2. Materials and Methods

This section presents the dataset utilized in our experiments and describes the proposed method with its motivations. As explained below, two case studies were defined: one in Spain, one in Italy, to gather clues on re-usability of the method in different geographical contexts.

### *2.1. Ground Truth Data*

Manuring data for crop fields located in Spain were obtained from the "Satellite imagery dataset of manure application on pasture fields" [24], which combines Sentinel-2 satellite imagery with precise on-site verification, including GPS-recorded locations and photographic evidence of manure application. The average field size exceeds 10,000 square meters, with a total of 30 fields that lie inside the region depicted in Figure 1 (top).

The Italian dataset was added to assess generalization capabilities, enabling mixed training and testing. In a first subset, manure application dates for 26 fields in Northern Italy were determined through on-site surveying. A second subset was sourced from the regional DUSAF (Destination of Use of Agricultural and Forestry Soils) [25] archive referring to year 2018 in the Lombardy region of Northern Italy. Here, agricultural parcels in use could be delineated, but manuring dates were unavailable, and only statistical evaluations could be made. Those farmland coordinates were obtained via Google Earth Engine (GEE). A total of around 300 fields—the area of interest is depicted in Figure 1 (bottom)—were included, with an average field size of almost 32,000 square meters.

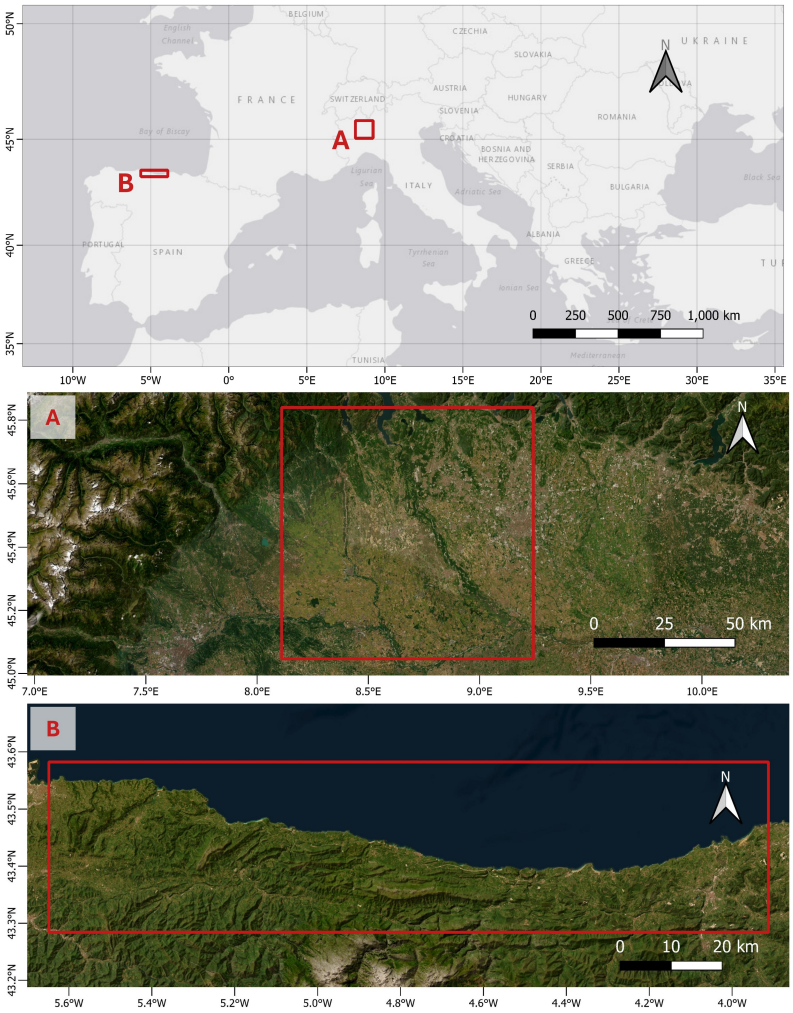



**Figure 1.** Smallest rectangles encircling all analyzed fields. Top row: Geographical context; (**A**) Plots located in northern Spain; (**B**) Plots located in north-western Italy (CRS: EPSG:4326—WGS 84).

### 2.2. Satellite Data

In this work, we used Sentinel-1, Sentinel-2, and Landsat-8 data due primarily to their open-data policy, offering free access to spaceborne sensing data. These satellites offer a comprehensive range of sensors, enabling both SAR (Sentinel-1) and optical/thermal (Sentinel-2, Landsat-8) imaging. This combination allowed us to extract essential indices for monitoring the effects of manure application.

Sentinel-1 revisits the same location on the Earth every 6 days, Sentinel-2 every 5 days, and Landsat-8 every 16 days. These revisit times ensure continuous monitoring, allowing for timely observation of changes in agricultural fields. Their long-term operation also provides consistent data streams for detecting long-term environmental trends.

Regarding spatial resolution, Sentinel-1 offers C-band radar imaging with a spatial resolution of 15 m in standard Interferometric Wide (IW) swath for both Vertical-Vertical (VV) and Vertical-Horizontal (VH) polarizations. Sentinel-2 provides multispectral imagery with resolutions ranging from 10 to 60 m depending on the band, while Landsat-8 offers 30 m resolution for multispectral data and 15 m for panchromatic imaging. These high spatial resolutions enable detailed analysis of crop health and soil conditions, supporting precise environmental monitoring.

### 2.3. Methods

The methodology illustrated in Figure 2 involves several steps aimed at analyzing time series data from agricultural fields and applying machine learning to detect the impact of manure application. First, datasets from agricultural fields located in Spain and Italy were collected. These datasets consisted of Sentinel-1 SAR data plus Sentinel-2 and Landsat-8 optical satellite data collected in a period of time. Crop fields from two different regions were selected to ensure robustness and generality of the analysis, allowing the models to perform across different geographical areas.

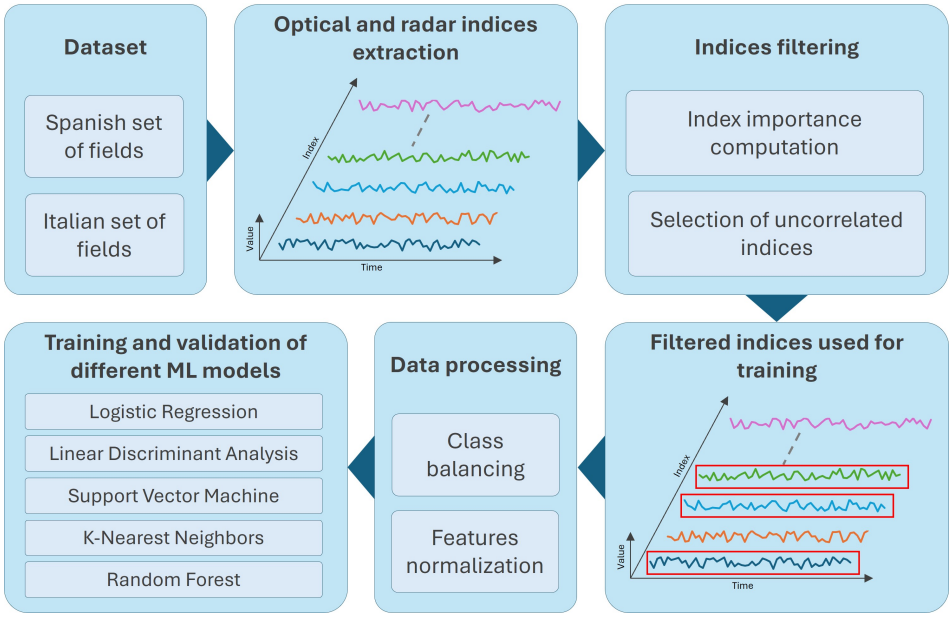



**Figure 2.** Flowchart of the overall methodology. Flowchart of the overall methodology. Different colors represent various index time series derived from optical and radar data. Red rectangles highlight the indices selected for the training process as an example.

Next, various optical and radar indices were extracted from the time series data. These indices, such as vegetation, water content, or radar-based measures, provide important

insights about the physical properties of the soil. The goal of this extraction step is to determine the most relevant indicators of how manure application influences the fields.

Although the impact of manuring on field reflectance manifests in the visible range also, the induced changes seem not to be consistent nor conclusive. Similar alterations in the RGB composites may result from other agricultural operations, such as ploughing [26]; visible bands are thus insufficient to reliably detect manure application events. Examples illustrating this observation are presented in Figure 3: manuring in some cases leads to an increase in reflectance (right column), in some others to a decrease (centre column), in others even more complex effects (left column). This highlights the importance of incorporating indices beyond the visible spectrum, such as those derived from infrared data, to enhance detection capabilities.

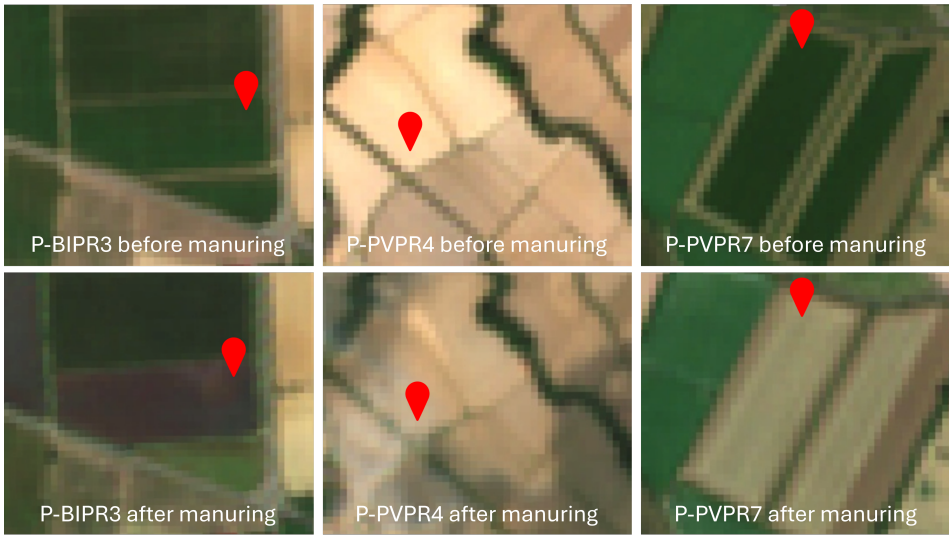



**Figure 3.** Difference in the visible spectrum considering the Sentinel-2 acquisitions (50 m scale) just before (**top**) and just after (**bottom**) manure application for three fields (namely—from left to right—P-BIPR3, P-PVPR4, P-PVPR7), identified by a geomarker.

A total of 60 indexes were computed on the entire range of visible and infrared bands of the considered datasets. Once the indices were extracted, a filtering process was applied to narrow down the dataset to the most relevant indices only.

#### 2.3.1. Correlation and Relevance Analysis

Our analysis highlighted that many of the spectral indices examined in this study display significant correlations with each other, as illustrated in Figures 4 and 5. This correlation implies redundancy if all indices are used simultaneously in the model. On the other hand, our model relies not on absolute index values but rather on the differences between consecutive satellite acquisitions for each index. By leveraging temporal changes, the model aims to capture dynamic shifts that occur following manure application, rather than static index values that might be affected by other environmental factors.

The analysis of the correlation matrix visualized in Figure 4 led to narrowing down the set of multispectral S-2-based indexes to 21 items whose mutual correlation was deemed sufficiently low. This is important because highly correlated features can easily lead to overfitting and thus loss of generalization capability.

The next step aimed at identifying indexes that are most impacted by manure application. This involved tracking feature trends throughout the year and assessing their possible changes (or lack thereof) across manure application; this helps guide feature selection for subsequent model development.

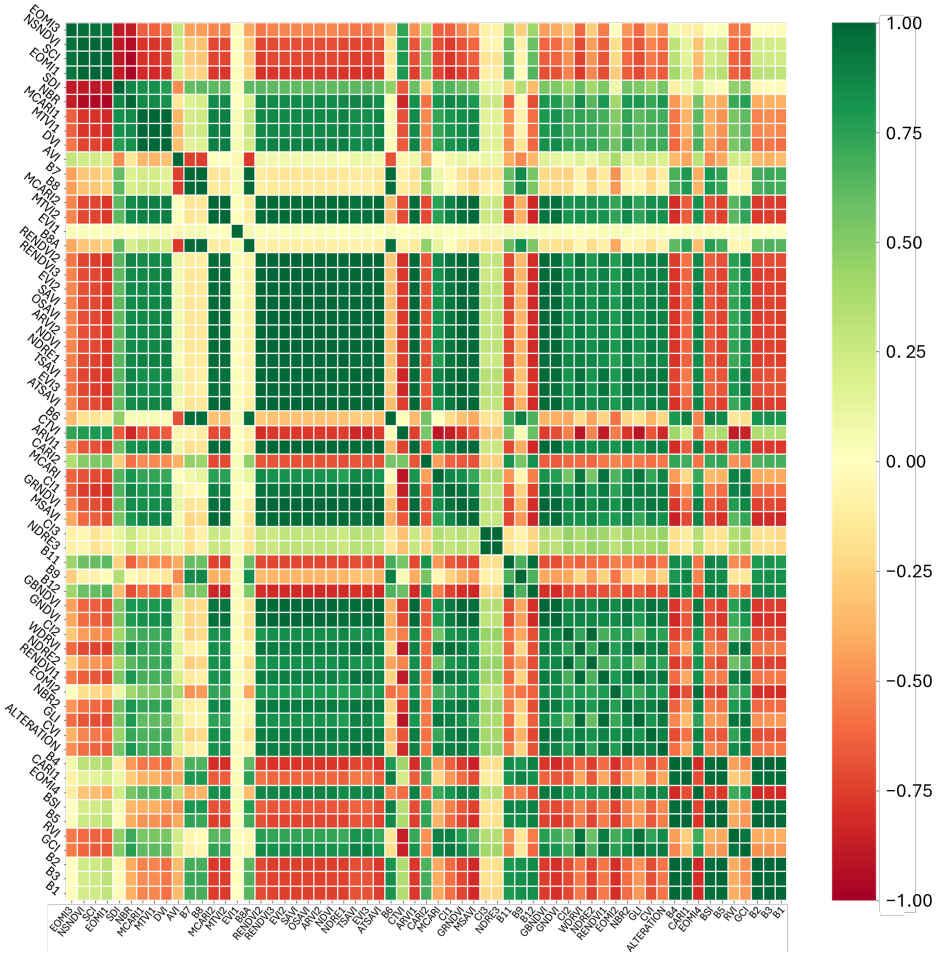



**Figure 4.** Correlation between different Sentinel-2 indexes (Spanish context).

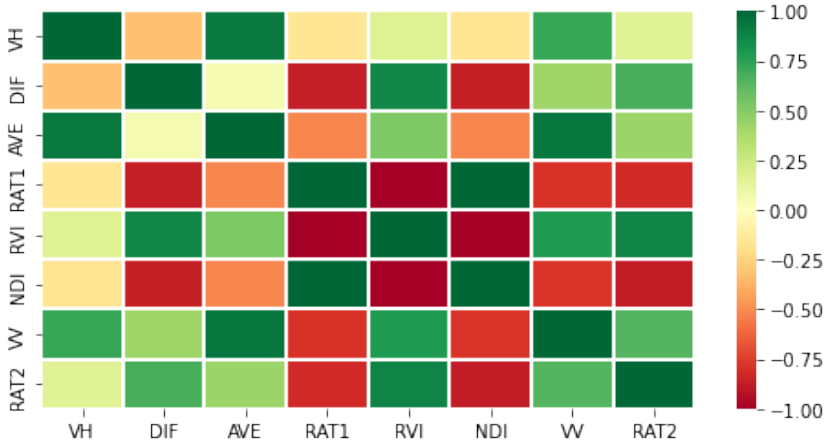



**Figure 5.** Correlation between different Sentinel-1 indexes (Spanish context).

To quantify the significance of each feature in detecting manure application events, a *feature importance parameter* was introduced. This parameter is calculated as the absolute difference in the feature value immediately after manure application (AM) and its value immediately before manure application (BM). This difference is then normalized by dividing it by the maximum absolute difference observed in the feature values between any two consecutive acquisitions during periods when no manure application occurred. By normalizing in this way, the feature importance parameter captures the relative change caused specifically by manure application, while accounting for the natural variability in

the feature over time. This approach ensures that the parameter highlights features most impacted by manure application relative to their fluctuations. The feature importance parameter's analytical expression is presented in Equation (1), where $|\cdot|$ is the absolute value, and $\neg$M indicates the "no manure application" case:

$$I = \frac{|f_{AM} - f_{BM}|}{\max|\Delta f_{\neg M}|} \quad (1)$$

Feature selection is a critical process employed to enhance model performance and prevent overfitting by identifying and retaining only the most relevant features from a dataset. By selecting a subset of informative features, it is possible to reduce the complexity of the model, thereby minimizing the risk of overfitting, which occurs when a model performs well on training data but poorly on unseen data. To assess statistical significance, a one-sample *t*-test was conducted, with the threshold set to the commonly used value of 0.05. A value of $p$ below the threshold suggests that the observed change was not the result of pure chance. If a correlation analysis is performed on the different indexes, it can be noticed that the majority of them are strongly correlated, for both the fields located in Italy and those in Spain. Therefore, only a subset of the indexes among the most mutually uncorrelated and most impacted by manure application has been considered (for optical, radar, and thermal indexes). For example, regarding multispectral optical indexes, this subset contains EOMI3, SWIR-Red edge Normalized Difference Index, EOMI1, and SDI, as illustrated in Table 1.

**Table 1.** Top 5 Sentinel-2 and Sentinel-1 indexes most affected by manure application, for crop fields located in Spain.

| Feature | Name | Sensor | Expression | *Feature Importance* Value |
|---|---|---|---|---|
| EOMI3 | Exogenous Organic Matter Index 3 | Sentinel-2 | $\frac{B12 - B4}{B12 + B4}$ | 0.720542 |
| RESNDI | SWIR-Red edge Normalized Difference Index | Sentinel-2 | $\frac{B11 - B7}{B11 + B7}$ | 0.672114 |
| SCI | Soil Composition Index | Sentinel-2 | $\frac{B11 - B8}{B11 + B8}$ | 0.661566 |
| EOMI1 | Exogenous Organic Matter Index 1 | Sentinel-2 | $\frac{B11 - B8A}{B11 + B8A}$ | 0.628979 |
| SDI | SWIR Difference Index | Sentinel-2 | $B8 - B12$ | 0.584395 |
| VH | VH polarization | Sentinel-1 | VH only | 0.366449 |
| DIF | Difference | Sentinel-1 | $VV - VH$ | 0.362870 |
| AVE | Average | Sentinel-1 | $\frac{VV \times VH}{2}$ | 0.340196 |
| RAT1 | Ratio 1 | Sentinel-1 | $\frac{VV}{VH}$ | 0.312078 |
| RVI | Radar Vegetation Index | Sentinel-1 | $\frac{VH \times 4}{VV + VH}$ | 0.305876 |

#### 2.3.2. Class Balancing and Feature Normalization

The selected indices were passed to the following steps: class balancing and feature normalization. Class balancing is used to handle imbalances in the dataset, as in our case where fields without manure application vastly outnumber manured fields. Feature normalization was applied to standardize the scale of the data. This step is crucial for certain models that are sensitive to the magnitude of input features.

Finally, a range of machine learning models was trained and validated using the processed dataset. The selected models include the following:

- Logistic Regression: a classification method that predicts the probability of an outcome using a sigmoid function. It is effective for binary classification and applicable to multi-class problems, assuming a linear relationship between input variables and log-odds.
- Linear Discriminant Analysis (LDA): a technique that finds a linear combination of features to maximize class separability by projecting data onto a lower-dimensional space. It assumes normally-distributed classes with equal covariance matrices and helps prevent overfitting.
- *K*-Nearest Neighbors (*K*-NN) Classifier: this is non-parametric classifier that assigns class labels based on the majority vote of the *K* nearest neighbors. It does not assume a specific data distribution but is sensitive to the choice of *K*, which affects performance and the risk of overfitting.
- Support Vector Classifier (SVC): a model that identifies an optimal hyperplane to separate classes by maximizing the margin between them. It uses different kernel functions to handle both linear and non-linear separability, though its performance depends on careful tuning of hyperparameters.
- Random Forest Classifier: an ensemble learning method that constructs multiple decision trees and aggregates their predictions, enhancing robustness against overfitting. It effectively handles high-dimensional data and provides insights into feature importance, but requires careful hyperparameter tuning.

Each model has different strengths, and the goal was to identify which one performed best at classifying fields based on whether manure had been applied. By comparing their performance, the most accurate and generalizable model could be selected to detect in-field changes resulting from manure application.

### *2.4. Further Experimental Details*

A Python (version 3.11.3) library was developed and made publicly available in a PyPI repository named *ee-satellites* (version 1.0.0) [27,28] to extract the identified indexes from optical, thermal, and radar satellite data. This tool simplifies working with satellite data from Sentinel-1 (Ground Range Detected) [29], Sentinel-2 (Surface Reflectance) [30], and Landsat-8 (Top of Atmosphere Reflectance) [31]. It also allowed us to automatically apply cloud masking, image compositing, and other data-cleaning steps, while exploiting multi-threading for efficient parallel processing of requests to GEE.

For each acquisition over an area of interest, the mean values of its spectral bands were calculated for that region, using the GEE `reduceRegion` function. Cloud and shadow filtering was also applied to improve data quality. Subsequently, other spectral indexes were computed using the obtained mean band values (examples are provided in Figure 6). The described extraction phase allowed us to reduce the number of requests to GEE, resulting in a significant speedup in the feature extraction procedure.

The different datasets underwent a meticulous transformation process to prepare them for ingestion in ML models. Each row was carefully modified to encapsulate the indexes' variations between consecutive satellite acquisitions, together with a flag (binary variable) that indicates whether manuring took place between the previous and the currently considered acquisitions. It is to be noted that samples with manuring flag 1 are much rarer than those with flag 0, as there may be one or two manuring episodes generating a positive sample for each observed field but several acquisition dates across an agricultural season generating negative samples.

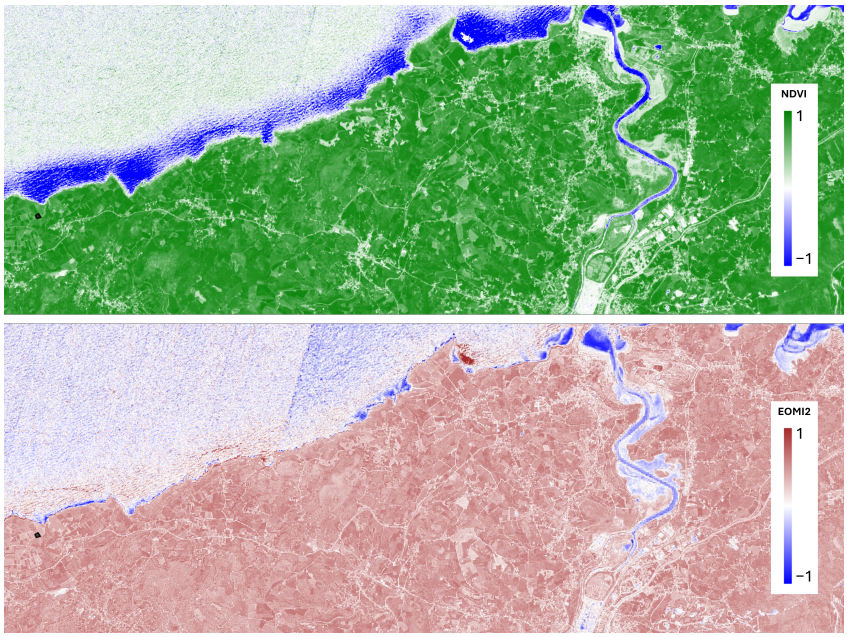



**Figure 6.** Example of (**Top**) NDVI and (**bottom**) EOMI2, at a pixel-wise level, calculated near a farmland of interest, represented in black on the leftmost side.

Thus, to mitigate class imbalance issues, both undersampling and oversampling techniques have been considered [32] as previously mentioned. More specifically, in undersampling, random samples from the majority class were picked—without replacement—while in oversampling, samples from the minority class were replicated multiple times. Feature selection played a crucial role in enhancing model performance [33]. A forward feature selection approach was employed, systematically identifying the most relevant input indexes for a range of ML classifiers. Thus, only those features that significantly improved training and testing accuracy without overfitting nor underfitting have been used. These classifiers included Logistic Regression [34], Linear Discriminant Analysis [35], Support Vector Classifier [36], K-Nearest Neighbors [37], and Random Forest Classifier [38]. For each model, hyperparameter tuning [39] has been applied, more specifically a grid-search technique. Grid search is a simple yet effective method for hyperparameter tuning that is typically used to optimize one or two hyperparameters (sometimes three). If the best combination is found at the edge of the grid, the parameter range may need to be expanded to capture the optimal value. Identifying a suitable set of potential values can be challenging, but exploratory tests can help determine an appropriate range. For hyperparameters with values spanning different orders of magnitude, using a geometric progression is advisable. The results can reveal the model's sensitivity to different hyperparameters and indicate which ones need careful tuning. As a suggestion to the readers who may want to replicate our experiments, we believe results can be improved by repeating the search with a finer grid around the initial best combination—if computational resources allow.

Furthermore, to address variations in feature values, multiple scaling techniques were considered [40], including Min-Max, Mean-Var, Max-Abs, and Robust Scaling. Finally, to assess and validate model performances comprehensively, Stratified Cross-Validation was selected to ensure more reliable and unbiased assessments of models' performances [41].

## 3. Experimental Findings and Interpretation

### *3.1. Spanish Test Case*

After identifying the spectral features most impacted by manure application, the next step involved examining the temporal trends of a few key indices for a set of randomly selected

reference crop fields, all located in Spain. This analysis aimed to observe how these indices fluctuated over time in response to manure application events and to better understand their behavior. The trends are displayed in Figure 7, which showcase the changes in selected indices across multiple acquisitions for these fields. It is important to note that the values shown in these charts are not the actual raw index values. Instead, each index value has been scaled to a standardized range of −1 to +1, for visualization purposes.

The effects of manure application on crop fields are uneven, with some fields showing more pronounced changes in specific indices than others. This variability can be attributed to differences in soil composition and conditions, and even the timing and rate of manure application, which all influence how strongly these indices respond. Furthermore, these indices naturally fluctuate over the year due to seasonal crop growth, weather patterns, and other environmental factors, meaning that changes in index values are not solely driven by manure applications.

The next phase, following this initial analysis of index importance and correlation, is to develop and train the machine learning models using these selected features. These models weredesigned to predict manure application events based on the observed changes in index values between two dates. Initially, each model was tested on Spanish crop fields, the same context from which the training data was derived. Subsequently, the model's ability to generalize would be tested in Italy, to assess whether it could detect manure application in regions with diverse environmental and agricultural conditions.

After conducting data analysis and balancing the DataFrames to ensure an even representation of manure and non-manure application events, multiple machine learning models were evaluated to determine their effectiveness. This evaluation involved a 5-fold cross-validation approach, which consists of splitting the data into five parts and training the models on four parts while validating on the fifth, rotating through different sets to ensure robust performance testing. The 5-fold validation was introduced to improve representativeness of validation accuracy as a means to determine suitability of the model for the problem at hand. Additionally, this process was repeated across different random seeds to further validate the consistency and reliability of the model results. Two main configurations of the models were compared to assess their accuracy in detecting manure applications. The first configuration focused on machine learning models trained on specific Sentinel-2 indices, selected for their sensitivity to changes caused by manure application. These indices were chosen based on initial analyses that identified which spectral bands were most impacted by manure presence. The second configuration used features derived from Sentinel-1 SAR data, allowing for a comparative analysis of optical versus radar-based manure detection capabilities. To optimize the models, index normalization was applied to the data, with various scaling techniques tested to ensure that features were appropriately scaled and comparable. Among these techniques, the Max-Abs scaler produced the best results, as determined through a forward feature selection process. This process involved iteratively adding features and assessing their impact on model performance, ultimately identifying the optimal feature set and scaling method to improve detection accuracy.

With the datasets balanced to ensure an equal representation of manure and non-manure events, model evaluation focused on general accuracy as the primary metric. Results showed a clear distinction between the effectiveness of Sentinel-2 and Sentinel-1 data for detecting manure applications. Specifically, spectral indices derived from Sentinel-2 outperformed those derived from Sentinel-1 SAR data in terms of prediction capabilities. For the models based on Sentinel-1 SAR data, accuracy scores for most classifiers were only slightly better than a random classifier, with accuracy ranging between 0.50 and 0.64. This poor performance suggests that C-band radar data did not contain meaningful indicators of manure application, possibly due to radar's limited sensitivity to the types of changes that manure induces in soil

and crop properties. In contrast, models using data from Sentinel-2, which captures reflectance in visible and infrared bands, led to far better performance levels. Among the models tested, Support Vector Classification (SVC) stood out as particularly effective, demonstrating both high accuracy and the ability to generalize well across different data samples. These findings are detailed in Table 2, which provides a comparative overview of classifier performance across different datasets and configurations.

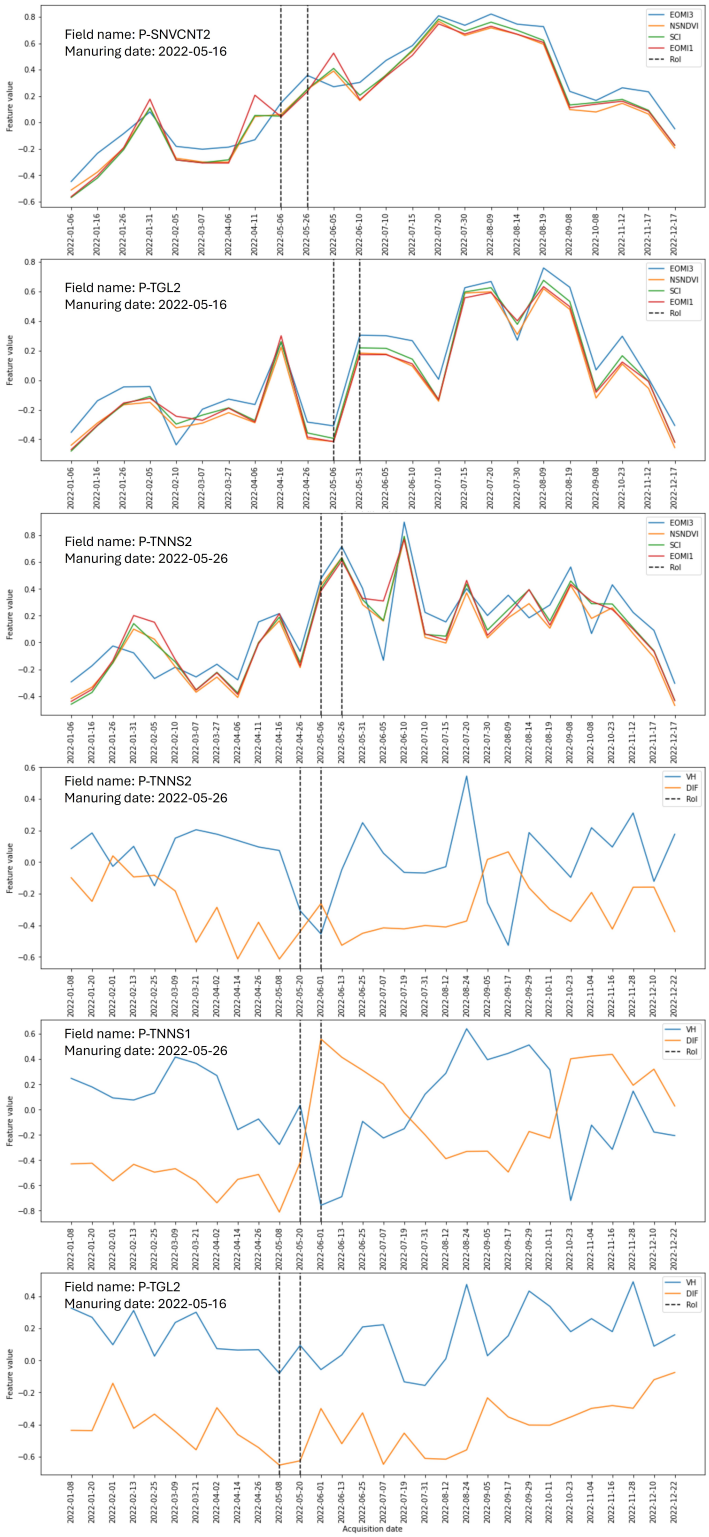



**Figure 7.** Trend of selected Sentinel-2 and Sentinel-1 indices most impacted by manure application (Spanish context).

**Table 2.** Accuracy of models using best Sentinel-1 and Sentinel-2 indexes, trained on the Spanish context.

| Sentinel-2 | | | Sentinel-1 | | |
|---|---|---|---|---|---|
| **Model** | **Training Acc.** | **Test Acc.** | **Model** | **Training Acc.** | **Test Acc.** |
| LR | 0.80 | 0.77 | LR | 0.52 | 0.52 |
| LDA | 0.86 | 0.84 | LDA | 0.51 | 0.50 |
| SVC | 0.90 | 0.88 | SVC | 0.67 | 0.64 |
| KNN | 0.82 | 0.80 | KNN | 0.67 | 0.62 |
| RFC | 0.85 | 0.83 | RFC | 0.70 | 0.52 |

### *3.2. Italian Case Study*

#### 3.2.1. Cross-Validation, Spain to Italy

The next step of the study evaluated the model's ability to generalize manure application detection across different geographic regions. Specifically, the model, which was initially trained on crop field data from Spain, was tested on crop fields located in Italy. This testing aimed to assess whether the spectral signature of manure application remained consistent in a different geographic and environmental context. Results showed a notable drop in accuracy from 90% to 83%, indicating that the model, while retaining some predictive capability, loses effectiveness when applied outside the region it was trained on. This suggests that the spectral signature of manure application, while remaining somehow similar, is impacted by regional factors, such as soil type, climate, and crop species. Further insights are provided by the confusion matrix in Table 3, detailing true and false, positive and negative classifications. The matrix highlights specific areas where the model struggled to correctly identify manure application in the Italian fields; especially of concern is the large number of false positives, leading to a disappointing precision level of 16%. This is particularly annoying considering the objective of raising a flag where manure application takes place in closed periods. Such poor results highlight the need for additional regional data or adaptations to improve cross-regional generalization.

**Table 3.** Confusion matrix of a model trained on Spanish context and tested on the Italian context. Please note that each considered field generates more than one sample, as identification is performed on each adjacent pair of acquisitions. For the "Manured" class, precision is 16% and recall is 74%.

| | | **Predicted** | |
|---|---|---|---|
| | | **Not Manured** | **Manured** |
| **Actual** | **Not manured** | 199 | 27 |
| | **Manured** | 14 | 5 |

#### 3.2.2. Training and Validating in Italy

To further our understanding of local factors, independent models were trained specifically on data from Italian farmland to assess whether region-specific training could improve the model's performance in detecting manure applications. These Italian-trained models used the same set of features, normalization methods, and cross-validation techniques as those applied to the Spanish farmland models, ensuring consistency in model configuration and evaluation. The training process utilized 5-fold Stratified Cross Validation (K), a method that divides the dataset into five equally sized, stratified subsets. In each fold, one subset is used for validation while the remaining four are used for training, rotating through each subset to ensure that each portion of the data is validated once. Stratification

was used to maintain balanced representation of manure and non-manure events across each fold, which enhances the reliability and robustness of the results. Because the Italian dataset was balanced to contain equal representations of both classes, accuracy alone was chosen as the metric for performance comparison, as it provides a straightforward measure of the models' success rate in distinguishing manure application events. Table 4 presents the accuracy results, allowing for a direct comparison with the performance of models trained on Spanish data.

**Table 4.** Accuracy of models using best Sentinel-2 extracted indexes, trained on the Italian context.

| Model | Train Acc. | Test Acc. |
|---|---|---|
| LR | 0.58 | 0.54 |
| LDA | 0.65 | 0.60 |
| SVC | 0.70 | 0.69 |
| KNN | 0.68 | 0.57 |
| RFC | 0.78 | 0.63 |

As previously discussed, the models trained on Italian farmland demonstrated a reduction in performance compared to the Spanish models when evaluated on the Italian dataset itself. However, the Italian models showed a significant improvement in their ability to generalize to other, unseen Italian farmland. This is particularly evident in the increase in recall for the "manured" class, which increased from 30% to nearly 75%. The substantial improvement in recall suggests that the Italian models were better at recognizing manure applications within the same geographic region, even when tested on different fields. This increase in performance supports the hypothesis that the spectral signature of manure application is regionally consistent, meaning that the model could better detect manure application events within the familiar environmental context of Italian farmland. This consistency may arise from similar soil types, crop varieties, climate conditions, and agricultural practices across the region, which likely result in more uniform spectral responses to manure application. In contrast, the Spanish-trained models performed worse on Italian data, reinforcing the idea that manure's spectral signature can be influenced by geographical factors. These findings highlight the importance of region-specific model training to enhance detection accuracy, as spectral signatures may differ across areas due to local environmental variations.

#### 3.2.3. Italy, Additional Dataset

In remote sensing of agricultural operations, a common issue is the scarcity of well-documented and detailed ground truth information. Detailed samples, especially when they include multiple observation along the agricultural season, are usually the result of specific inspections, which tend to be limited in geographical scope for obvious matters of resource and research personnel availability. However, a broader geographical context could be covered using a statistic analysis rather than an individual-result analysis. In our case, the DUSAF dataset [25] offered a chance to qualitatively evaluate the model's performance in a broader context within Italy. The DUSAF dataset reports tens of thousands of farmland polygons with agricultural land-use data on the Italian Region of Lombardy; albeit it lacks precise, field-specific records of manure application events, it offers a chance to leverage statistics. The model can indeed be assessed on a larger and more varied set of samples, giving insight into its behavior over diverse field types and conditions across Italy. Although the analysis relied on statistical observations on detection patterns rather than exact validation of manure application events, still it provided some interesting insights.

Results obtained on DUSAF polygons revealed patterns in manure detection that aligned with seasonal expectations in several cases. Manure detections were indeed more

frequent in spring, a typical period for spreading manure to enrich the soil before the growing season. The model also detected an unexpectedly high number of manure application cases during summer months, which could be due to specific types of crops requiring late manuring (e.g., *solanaceae* and *cucurbitaceae* types, both grown in northern Italy). Figure 8 illustrates these seasonal detection patterns, highlighting both the expected spring peak and the summer surge in the detected manure events. The summer surge, slightly greater than expected, could be due to environmental factors or spectral characteristics in summer that the model may misinterpret as signs of manure application.

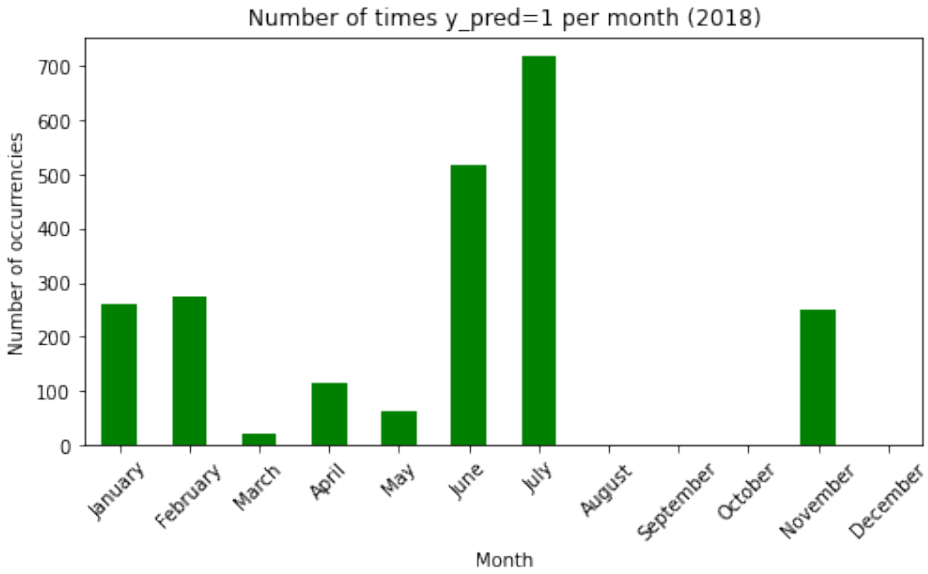



**Figure 8.** Number of occurrences detected, per month, on the DUSAF dataset using Sentinel-2 indexes only.

#### 3.2.4. Introduction of Thermal Data

Considering that manure may affect thermal properties of the observed fields, besides resulting into exothermal reactions itself, it was hypothesized that a thermal band could also contribute to identification of manuring operations. S-2 does not feature any thermal band, hence thermal indexes from Landsat-8 were combined with optical ones from Sentinel-2. The temporal sampling was naturally different; to create temporally consistent fused datasets, each thermal acquisition from Landsat-8 was associated with the nearest Sentinel-2 multispectral acquisition, assuming the time displacement negligible with respect to the temporal scale of the manuring effects. Indeed, notwithstanding the temporal misalignment between thermal and multispectral acquisitions, this approach enhanced the model performance in both the Spanish and the Italian context, as reported in Table 5, where accuracy figures for the Random Forest model are compared in the case without and with the integration of thermal data.

**Table 5.** Increase in accuracy of RF models from thermal data integration.

| Context | w/o Thermal | w/ Thermal | Improvement |
|---|---|---|---|
| Spain | 90% | 92% | +2% |
| Italy | 70% | 82% | +12% |

The benefits of jointly using multispectral and thermal data emerged also in the assessment over the DUSAF dataset, with the disappearance of November detections (Figure 9), which were assumed to be false positives based on local cultivation practices.

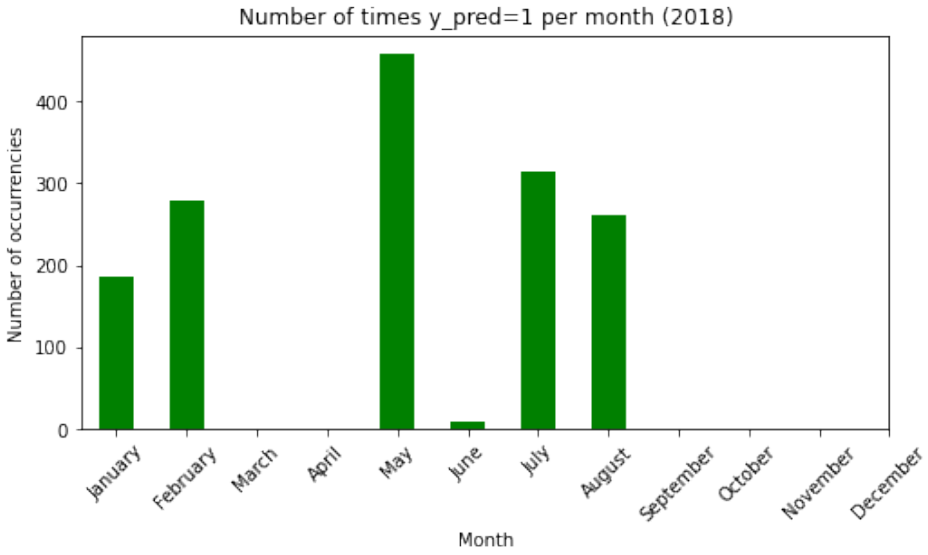



**Figure 9.** Number of occurrences detected, per month, on the DUSAF dataset (model combining Sentinel-2 and Landsat-8 indexes).

In summary, combining Landsat-8 and Sentinel-2 indexes, despite differences in resolution and band alignment, outperformed using Sentinel-2 alone, highlighting the value of thermal data in this context.

## 4. Conclusions

Space-based mapping of manuring events in predominantly agricultural areas can provide important information for large-scale environmental management. An effective method to monitor manuring would represent an important means for environmental monitoring and regulatory compliance, particularly with respect to the EU's *Nitrates Directive*. This paper presents a method for detecting manure application in agricultural fields using multi-source spaceborne Earth observation data. The proposed model achieved promising accuracy levels using Support Vector Classification on a suitable selection of optical-based indexes. Best accuracy levels (around 92%) were achieved when training and testing were carried out in the same geographical context, whereas training on one area and testing on a different geographical area led to a significant decrease in accuracy, suggesting that local practices' and environmental features' impact on spectral signatures associated to manuring events. The introduction of thermal bands contributed to increasing accuracy (by 2% to 12% in our experiments). Limitations of the method include reliance on optical data, which may lead to missed detections in case of extended cloud coverage, as well as testing on a limited dataset due to difficulty in collecting reliable ground truth information at scale.

## 5. Future Developments

Regarding future developments, plans include expanding the study to involve manured fields from diverse geographical regions, which could validate and refine the model's robustness in varied environmental conditions. Also, additional spectral indices may further enhance detection accuracy, particularly considering those sensitive to organic matter. The incorporation of cost-sensitive learning methods will address class imbalances and reduce data loss, improving model performance and applicability across larger datasets. Finally, expanding the dataset itself is also essential for training more generalizable models aimed at agricultural monitoring on a larger scale.

**Author Contributions:** Conceptualization, D.M. and F.D.; methodology, F.D.; software, D.M.; validation, D.M. and F.D.; formal analysis, D.M. and F.D.; investigation, D.M. and F.D.; resources, D.M. and F.D.; data curation, D.M.; writing—original draft preparation, D.M.; writing—review and editing, D.M. and F.D.; visualization, D.M. and F.D.; supervision, F.D.; project administration, F.D.; funding acquisition, F.D. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was partially supported by the "Nord Ovest Digitale e Sostenibile (NODES)" project, which has been granted funding through the MUR – M4C2 1.5 of PNRR, under the European Union's NextGenerationEU initiative (Grant agreement No. ECS00000036).

**Data Availability Statement:** To promote open and replicable research, all code, notebooks, datasets, and materials used in this study are publicly available in the following GitHub repository, allowing for replication and further research: https://github.com/Amatofrancesco99/organic-fertilizers-detection (accessed on 10 March 2025) .

**Acknowledgments:** The authors wish to thank Francesco Amato for his coding and data processing/management work. This research was partly funded by the European Union—NextGenerationEU, Mission 4 Component 1.5—ECS00000036—CUP F17G22000190007.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Food and Agriculture Organization (FAO) of the United Nations. Chapter 1. Economic Dimensions of Agriculture. In *Statistical Yearbook—2022*; FAO Statistics: Rome, Italy , 2022. [CrossRef]
2. Cleland, J. World Population Growth: Past, Present and Future. *Environ. Resour. Econ.* **2013**, *55*, 543–554. [CrossRef]
3. Chakravarty, S.; Ghosh, S.; Suresh, C.; Dey, A.; Shukla, G. *Deforestation: Causes, Effects and Control Strategies*; Intech: London, UK, 2012. [CrossRef]
4. Hignett, T.P. History of chemical fertilizers. In *Fertilizer Manual*; Springer: Berlin/Heidelberg, Germany, 1985; pp. 3–10.
5. Scherer, H.W.; Mengel, K.; Kluge, G.; Severin, K. Fertilizers Chapter 1: General. In *Ullmann's Encyclopedia of Industrial Chemistry*; John Wiley and Sons, Ltd.: Hoboken, NJ. USA, 2009. [CrossRef]
6. Kleinman, P.J.; Spiegal, S.; Liu, J.; Holly, M.; Church, C.; Ramirez-Avila, J. Managing Animal Manure to Minimize Phosphorus Losses from Land to Water. In *Animal Manure*; John Wiley and Sons, Ltd.: Hoboken, NJ, USA, 2020; pp. 201–228. [CrossRef]
7. Kim, D.K.; Kaluskar, S.; Mugalingam, S.; Arhonditsis, G.B. Evaluating the relationships between watershed physiography, land use patterns, and phosphorus loading in the bay of Quinte basin, Ontario, Canada. *J. Great Lakes Res.* **2016**, *42*, 972–984. [CrossRef]
8. Fahmy, F. Pollution Erodes fish Stocks and Livelihoods in Egyptian Lake. 2022. Available online: https://www.reuters.com/world/africa/pollution-erodes-fish-stocks-livelihoods-egyptian-lake-2022-09-01/ (accessed on 10 March 2025) .
9. EU Council Directive. 91/676/EEC of 12 December 1991 Concerning the Protection of Waters against Pollution Caused by Nitrates from Agricultural Sources. *Off. J. L 375 EU* **1991**, *375*, 12. Available online: https://faolex.fao.org/docs/pdf/eur19065.pdf (accessed on 10 March 2025).
10. Tzilivakis, J.; Warner, D.; Green, A.; Lewis, K. A broad-scale spatial analysis of the environmental benefits of fertiliser closed periods implemented under the Nitrates Directive in Europe. *J. Environ. Manag.* **2021**, *299*, 113674. [CrossRef] [PubMed]
11. Svoboda, J. Inspections in the agricultural practice from the Nitrate directive point of view. *Acta Agric. Slov.* **2006**, *87*, 155–158. [CrossRef]
12. The European Commission Implementing Decision (EU) 2022/696. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022D0696 (accessed on 10 March 2025).
13. Mateo-Garcia, G.; Veitch-Michaelis, J.; Smith, L.; Oprea, S.; Schumann, G.; Gal, Y.; Baydin, A.; Backes, D. Towards global flood mapping onboard low cost satellites with machine learning. *Sci. Rep.* **2021**, *11*, 7249. [CrossRef]
14. Jean, N.; Burke, M.; Xie, M.; Davis, W.M.; Lobell, D.B.; Ermon, S. Combining satellite imagery and machine learning to predict poverty. *Science* **2016**, *353*, 790–794. [CrossRef]
15. Verma, P.; Patil, S. A Machine Learning Approach and Methodology for Solar Radiation Assessment Using Multispectral Satellite Images. *Ann. Data Sci.* **2021**, *10*, 907–932. [CrossRef]
16. Girona, T.; Realmuto, V.; Lundgren, P. Large-scale thermal unrest of volcanoes for years prior to eruption. *Nat. Geosci.* **2021**, *14*, 396. [CrossRef]
17. Cunha, M.; Marçal, A.; Silva, L. Very early prediction of wine yield based on satellite data from vegetation. *Int. J. Remote Sens.* **2013**, *31*, 3125–3142. [CrossRef]

18. Fraternali, P.; Morandini, L.; González, S.L.H. Solid waste detection, monitoring and mapping in remote sensing images: A survey. *Waste Manag.* **2024**, *189*, 88–102. [CrossRef] [PubMed]
19. Dodin, M.; Smith, H.D.; Levavasseur, F.; Hadjar, D.; Houot, S.; Vaudour, E. Potential of Sentinel-2 Satellite Images for Monitoring Green Waste Compost and Manure Amendments in Temperate Cropland. *Remote Sens.* **2021**, *13*, 1616. [CrossRef]
20. Pedrayes, O.D.; Usamentiaga, R.; Trichakis, Y.; Bouraoui, F. Remote sensing for detecting freshly manured fields. *Ecol. Inform.* **2023**, *75*, 102006. [CrossRef]
21. Dubbini, M.; Belluzzo, M.; Zanni Bertelli, V.; Pirola, A.; Tornato, A.; Alessandrini, C. The Identification of Manure Spreading on Bare Soil through the Development of Multispectral Indices from Sentinel-2 Data: The Emilia-Romagna Region (Italy) Case Study. *Sensors* **2024**, *24*, 4687. [CrossRef] [PubMed]
22. Shea, K.; Schaffer-Smith, D.; Muenich, R.L. Using remote sensing to identify liquid manure applications in eastern North Carolina. *J. Environ. Manag.* **2022**, *317*, 115334. [CrossRef] [PubMed]
23. Grandini, M.; Bagli, E.; Visani, G. Metrics for Multi-Class Classification: An Overview. *arxViv* **2020**. [CrossRef]
24. Diaz Pedrayes, O.; Usamentiaga, R. *Satellite Imagery Dataset of Manure Application on Pasture Fields*; Elvesier: Amsterdam, The Netherlands, 2022. [CrossRef]
25. Territory and Civil Protection Department—Region of Lombardy. Destinazione d'Uso dei Suoli Agricoli e Forestali (DUSAF Database). Available online: https://www.dati.lombardia.it/Territorio/Dusaf-6-0-Uso-del-suolo-2018/7rae-fng6 (accessed on 10 March 2025).
26. Marzi, D.; Dell'Acqua, F.; Trichakis, I. Assessing Compliance to EU Nitrate Pollution Regulations by Detecting Manure Applications in Time Series of Sentinel-2 Acquisitions. In Proceedings of the IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium, Pasadena, CA, USA, 16–21 July 2023 ; pp. 3450–3453.
27. Amato, F. ee-Satellites Python Library. Available online: https://pypi.org/project/ee-satellites/ (accessed on 10 March 2025).
28. Amato, F. Revolutionize Your Satellite Data Acquisition with ee-Satellites. Available online: https://dev.to/amatofrancesco99/exciting-news-for-data-scientists-and-remote-sensing-enthusiasts-5g2g (accessed on 10 March 2025).
29. Google. Sentinel-1 SAR GRD: C-Band Synthetic Aperture Radar Ground Range Detected, Log Scaling. Available online: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S1_GRD (accessed on 10 March 2025).
30. Google. Sentinel-2 MSI: MultiSpectral Instrument, Level-2A. Available online: https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR (accessed on 10 March 2025).
31. Google. USGS Landsat 8 Collection 2 Tier 1 TOA Reflectance. Available online: https://developers.google.com/earth-engine/datasets/catalog/LANDSAT_LC08_C02_T1_TOA#description (accessed on 10 March 2025).
32. Mohammed, R.; Rawashdeh, J.; Abdullah, M. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. In Proceedings of the 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 7–9 April 2020; pp. 243–248. [CrossRef]
33. Venkatesh, B.; Anuradha, J. A Review of Feature Selection and Its Methods. *Cybern. Inf. Technol.* **2019**, *19*, 3. [CrossRef]
34. Zou, X.; Hu, Y.; Tian, Z.; Shen, K. Logistic Regression Model Optimization and Case Analysis. In Proceedings of the 2019 IEEE 7th International Conference on Computer Science and Network Technology (ICCSNT), Dalian, China, 19–20 October 2019; pp. 135–139. [CrossRef]
35. Tharwat, A.; Gaber, T.; Ibrahim, A.; Hassanien, A.E. Linear Discriminant Analysis: A Detailed Tutorial. *AI Commun.* **2017**, *30*, 169–190. [CrossRef]
36. Awad, M.; Khanna, R.; Awad, M.; Khanna, R. Support vector machines for classification. In *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*; Springer: Berlin/Heidelberg, Germany , 2015; pp. 39–66.
37. Guo, G.; Wang, H.; Bell, D.; Bi, Y. *KNN Model-Based Approach in Classification*; Springer: Berlin/Heidelberg, Germany, 2004.
38. Ali, J.; Khan, R.; Ahmad, N.; Maqsood, I. Random Forests and Decision Trees. *Int. J. Comput. Sci. Issues (IJCSI)* **2012**, *9*, 272–278.
39. Yu, T.; Zhu, H. Hyper-Parameter Optimization: A Review of Algorithms and Applications. *arXiv* **2020**. [CrossRef]
40. Turing Portal. How and Where to Apply Feature Scaling in Python? Available online: https://www.turing.com/kb/how-and-where-to-apply-feature-scaling-in-python (accessed on 10 March 2025).
41. Prusty, S.; Patnaik, S.; Dash, S. SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. *Front. Nanotechnol.* **2022**, *4*, 972421. [CrossRef]


**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.