https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

HyBEAR: A Hyperspectral Benchmark for Bare Soil Detection

Agata M. Wijata[1,2*], Bogdan Ruszczak[3*], Adriana Niepala[3], Michal Gumiela[2], Krzysztof Smykala[3,4], Nicolas Longépé[5], and Jakub Nalepa[1,2]

[1]Silesian University of Technology, Akademicka 2A, 44-100 Gliwice, Poland
[2]KP Labs, Bojkowska 37J, 44-100 Gliwice, Poland
[3]Opole University of Technology, Prószkowska 76, 45-758 Opole, Poland
[4]QZ Solutions Sp. z o.o., Technologiczna 2, 45-839 Opole, Poland
[5] -Lab, European Space Agency, Largo Galileo Galilei 1, 00044 Frascati, Italy
[*]These authors contributed equally to this work.

Correspondence to: Agata M. Wijata (awijata@ieee.org), Bogdan Ruszczak (b.ruszczak@po.edu.pl), Adriana Niepala (a.niepala@po.edu.pl).

Abstract. Detecting bare soil areas is an important step in the analysis of Earth observation data in a variety of Precision Agriculture (PA) applications focused on quantifying soil properties and assessing soil quality. In this paper, we introduce the HyBEAR benchmark—a novel large-scale collection of high-resolution hyperspectral aerial images (with 2 m ground sampling distance) accompanied with manual bare soil annotations verified with domain experts. Usually, the bare soil detection problem is tackled at the pixel level, meaning that detection methods classify all pixels as either bare soil or background. In contrast to this approach, we provide pixel-level annotations for the entire agricultural parcels (if the parcel is labeled as bare soil, then all pixels within that parcel are labeled accordingly), and aim to support the development of methods that identify entire fields with no vegetation. Commonly, such fields undergo further analysis to determine specific soil parameters and characteristics that are important while planning various PA activities, such as fertilization. The HyBEAR 🐻 benchmark includes (*i*) the largest-to-date (108,064,591 pixels, corresponding to 43,225 hectares) and most heterogeneous dataset for bare soil detection, as well as (*ii*) the validation procedure (training-test splits and quality metrics) and a set of baseline results, obtained for a set of machine learning bare soil detection models. From the FULL collection of 1954 images in HyBEAR, which we divided into 5 spatially-disjoint folds, we additionally selected a random, stratified subset (MINI) of the images which may be useful for designing and verifying bare soil detection algorithms. Overall, HyBEAR is a step toward standardizing the way the community builds and confronts bare soil detection algorithms in a thorough, reproducible, and unbiased way.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

1 Introduction

Over the past decades, the agricultural sector has undergone significant transformations driven by human technological advancements aimed at meeting the escalating demands for food, fiber, and fuel from a rapidly expanding global population (Sishodia et al., 2020; Aijaz et al., 2025). With arable land becoming increasingly scarce, enhancing crop productivity has become crucial for ensuring food security (Pingali, 2012; Roy et al., 2024). Precision Agriculture (PA), emerging from technological advancements in fields such as Earth observation, data analytics, and in-situ field monitoring, plays a vital role in improving crop productivity. The global market for smart agriculture, which includes PA, was valued at approximately 15 billion U.S. dollars in 2022 and is projected to grow to 30 billion U.S. dollars by 2027 (Allied Market Research, 2023). This growth reflects the increasing adoption of PA, which is important for producing essential human necessities (Sishodia et al., 2020; Song et al., 2018), especially given the limited availability of arable land (Song et al., 2018). PA aims to optimize agricultural practices by monitoring various parameters spanning soil quality (Bünemann et al., 2018), soil composition (Chen et al., 2022; Nalepa et al., 2024; Seu et al., 2025), and moisture levels (Nowak, 2021; Boguszewska-Mańkowska et al., 2022; Ruszczak and Boguszewska-Mańkowska, 2022), and taking appropriate actions based on these insights, while also reducing the environmental impact of agriculture (Misara et al., 2022). Indeed, sustainable food production necessitates a consideration of its environmental impact, especially in the face of climate change and pollution. PA offers a framework for improving agricultural efficiency while accounting for environmental effects (Finger et al., 2019), emphasizing the assessment and monitoring of soil characteristics, temperature, and seasonal ecosystem dynamics (Ponnusamy and Natarajan, 2021; Sayão et al., 2020). Moreover, understanding the relationship between soil parameters and crop yield in specific regions can provide valuable insights into the effectiveness of implemented agricultural practices (Yue et al., 2021). Research has also highlighted the importance of soil class in influencing crop yields, further emphasizing the need for accurate soil assessment (Tunçay et al., 2021). The evolution and future trajectory of PA toward sustainable food systems are continuously being shaped by technological advancements (Xu et al., 2024), with Artificial Intelligence (AI) playing a significant role in enhancing crop productivity and resource management (Aijaz et al., 2025).

1.1 Limitations of traditional soil assessment methods

Traditional methods for estimating soil parameters often rely on field-point sampling, involving the localized collection of soil or crop samples for subsequent laboratory analysis (Rutter et al., 2022). However, the inherent limitation of these in-situ methods, which typically focus on a few selected locations, hinders the ability to map the spatial distribution of key indicators (Nalepa et al., 2022, 2024). Consequently, considerable research efforts have been recently directed toward establishing correlations between data acquired through field methods (serving as the ground-truth information), and data obtained from satellite imagery (Meng et al., 2020; Hong et al., 2020; Lu and He, 2019) or aerial platforms such as manned aircraft and drones (Zhang et al., 2021; Yue et al., 2021; Han et al., 2019; Ji

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

et al., 2018). Recent reviews highlight the limitations of traditional soil sampling in capturing spatial variability. Additionally, they are extremely difficult to scale to large regions, and thus developing the remote sensing techniques for high-resolution soil property mapping is the direct remedy to this challenge (Chen et al., 2022; Nalepa et al., 2024). Drone-borne hyperspectral imagery, for instance, enables high-resolution mapping of crucial soil nutrients (Yan et al., 2023). Similarly, airborne and satellite imagery can be effectively used to estimate various soil parameters at a (potentially global) scale, offering inherent spatial scalability (Nalepa et al., 2024).

1.2 Remote sensing for bare soil identification

Remote Sensing (RS) has emerged as a powerful tool for the identification of bare soil (Campos et al., 2022), which is a prerequisite for estimating soil parameters and crop monitoring, owing to its capacity to assess extensive areas repeatedly over time (Yue et al., 2021; Ponnusamy and Natarajan, 2021). In the agricultural context, the methods involving the acquisition of Multispectral (MSI) (Jin et al., 2020) and Hyperspectral (HSI) (Lu et al., 2020) Images are prevalent. The multispectral data facilitate the straightforward estimation of chlorophyll content using green vegetation indices (Jin et al., 2020; Lu and He, 2019). However, the broad bandwidth of multispectral imaging can limit the accuracy of early detection of subtle negative symptoms in PA, including nutrient deficiencies and plant diseases (Adão et al., 2017). In contrast, HSIs, with their high spectral resolution (narrow and contiguous bands), enable the acquisition of finer details in the spectral response of a given area. HSI-based techniques can potentially detect various anomalies, such as the aforementioned plant diseases or soil features, earlier than MSI due to the richer spectral information contained within its narrower bands (Lu et al., 2020).

1.3 Bare soil detection as a necessary step in the data processing chain

Following the removal of images obscured by clouds to ensure data quality (Grabowski et al., 2022, 2024), the accurate identification of bare soil areas becomes a crucial subsequent step in the RS data processing chain for agricultural applications (Campos et al., 2022). Therefore, isolating the spectral response directly from the soil surface is a critical step in the data processing chain for various agricultural and environmental applications. By accurately detecting and potentially masking out non-bare-soil pixels, researchers can enhance the reliability and accuracy of subsequent analyses aimed at estimating crucial soil properties, including the moisture content, nutrient levels, organic matter, and texture (Chen et al., 2022; Nalepa et al., 2024). Furthermore, the identification of bare soil is also essential for monitoring agricultural practices, such as tillage, tracking fallow land, and assessing soil erosion risks (Yue et al., 2021; Zhao et al., 2024). Additionally, pruning non-soil areas can play a key role in on-board processing, where MSIs/HSIs are analyzed on edge devices, e.g., satellites. In this scenario, removing the parts of the image that do not contain the objects of interest (here, bare soil areas) will substantially accelerate the entire analysis process and make it more memory-efficient (Wijata et al., 2023). Therefore, bare soil detection can be considered a “smart data compression” step, in which areas of interest are determined to guide further analysis in a data-driven manner (note that extracting soil parameters from non-soil pixels would obviously lead to noisy, inherently incorrect estimates).

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Finally, as already mentioned, the spatial information derived from RS-based bare soil maps provides a significant advantage over traditional point-based soil sampling methods, enabling a more comprehensive, spatially continuous understanding of soil conditions across large areas.

1.4 Methods for bare soil detection

Detecting bare soil areas is approached through various methodologies, categorized as (*i*) green vegetation index-based methods and (*ii*) pixel-level machine learning classification techniques. The first group of algorithms applies vegetation indices originally designed for broader MSI bands, with the Normalized Difference Vegetation Index (NDVI) being the most widely recognized (Zhang et al., 2021). Recent studies, such as the Intuition-1 satellite mission (KP Labs, Gliwice, Poland), explore in-orbit bare-soil detection using spectral vegetation indices derived from hyperspectral imagery (Wijata et al., 2024a). By applying a threshold to the NDVI values, it is possible to delineate areas with dense green vegetation, as well as regions with medium or no vegetation (Zhang et al., 2021; Wang et al., 2020). In this context, areas lacking vegetation include bare soil, water bodies, and infrastructure, e.g., roads or buildings. The filtering of these non-bare soil objects is often achieved using various spectral indices, such as the Visible, Green-Based Built-up Index (VGNIRBI) or Normalized Difference Built-up Index (NDBI) for buildings, and the Modified Normalized Difference Water Index (MNDWI) for water bodies (Kaur and Pandey, 2022). Other examples of green vegetation indices used to estimate crop volume are the Enhanced Vegetation Index (EVI) and the Optimized Soil Adjusted Vegetation Index (OSAVI) (Nejatian et al., 2022). Dedicated bare-soil indices, e.g., the Bare Soil Index (BSI) (Nguyen et al., 2021; Liu et al., 2022), have also been investigated in the literature.

The pixel-level classification techniques (i.e., those that are not based on thresholding selected spectral indices) frequently implement Machine Learning (ML) algorithms, with Random Forests (RFs) being a popular choice (Vlachopoulos et al., 2020; Zhu et al., 2022; Saha et al., 2020), along with its variations, such as the Guided Regularized RF (GRRF) (Izquierdo-Verdiguier and Zurita-Milla, 2020). Deep learning algorithms have been widely investigated and have indeed established the state of the art in a multitude of fields. Some studies have explored deep learning models of various architectures, such as detection models, Capsule Networks (CapsNets), and semantic segmentation models (Joshi et al., 2021), as well as the aggregation of results from different methods using ensemble techniques (Saha et al., 2020), for bare soil identification. It is worth emphasizing that some deep learning methods, e.g., convolutional neural networks, exploit contextual information while inferring a pixel-level prediction—this is in contrast to the pure pixel-level algorithms, operating on the pixel-level information (e.g., a selected set of spectral indices). Recent advancements in this area include the application of quantum-kernel support vector machines for the detection of bare soil in hyperspectral imagery (Wijata et al., 2024b; Miroszewski et al., 2026). To enhance the accuracy of bare-soil pixel detection and mitigate potential error sources, researchers have used specific spectral indices as features. For instance, a mask for distinguishing bare soil pixels in raster data has been effectively created using NDVI and the Cellulose Absorption Index (CAI) (Pechanec et al., 2021).

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Building on these established methodologies, recent research continues to advance bare-soil detection. There is a growing interest in leveraging advanced deep learning architectures to improve the accuracy and robustness of these techniques. For example, Zhao et al. (2024) introduced a novel Hybrid Attention Network (HA-Net) designed explicitly for bare soil extraction from optical RS images. This network incorporates attention mechanisms to effectively distinguish bare soil from complex backgrounds, including urban and agricultural landscapes, and to enhance the extraction of small bare soil areas. The significance of high-quality, diverse datasets remains a central theme, and the broader RS community recognizes the critical need for large-scale annotated datasets to train and evaluate the algorithms effectively in a standardized way, as highlighted in the context of RS object recognition (Liu et al., 2025). Furthermore, a review provides a comprehensive overview of the satellite RS techniques employed for identifying bare soil, discussing the latest advancements, limitations, and challenges associated with various methodologies (Delaney et al., 2025). We tackle the challenge of building a comprehensive and unbiased benchmark for confronting the known and emerging bare soil detection algorithms in a fully reproducible and fair way.

1.5 Conclusion from state-of-the-art

The literature review reveals a critical gap in the availability of standardized, publicly accessible benchmark datasets specifically tailored for bare-soil detection using hyperspectral imagery. While numerous studies explore the application of remote sensing for soil assessment and analysis, the absence of such datasets significantly hinders objective algorithm comparison and limits the reproducibility of research findings (Kapoor and Narayanan, 2023). Notably, to our knowledge, there are no datasets that would include the entire fields of bare soil, and would allow the community to develop and validate the algorithms for identifying such fields that are free of vegetation. Indeed, the existing methods and collections focus on pixel-level annotations (thus support building pixel-level classification algorithms), which can be misleading for certain precision agriculture processes, especially those that relate to the whole-field procedures (e.g., fertilization). In this article, we address this research and development gap.

1.6 Contribution

We introduce HyBEAR 🐻: a comprehensive benchmark for bare soil detection in hyperspectral imagery. It is composed of the following pivotal components:

- **Data**—We release a large-scale (1,954 hyperspectral patches, with approx. 108 million pixels), high-resolution, spatially-heterogeneous dataset of hyperspectral image patches accompanied with precise ground-truth delineations of bare soil areas. The HyBEAR dataset is available at the following link: https://doi.org/10.5281/zenodo.17607897.
- **Validation procedure**—We define the cross-validation procedures (applied to both the complete set of available hyperspectral patches and its reduced subset, referred to as the FULL and MINI versions, respectively),

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

accompanied by the quality metrics that shall always be calculated while confronting the emerging bare soil detection techniques.

- **Baseline results**—We establish a set of baseline results obtained using an array of classic machine learning models, strictly following the suggested validation procedures. These baseline results may become the point of departure for further research in developing bare soil detection techniques, as they are directly comparable with the emerging results obtained for the HyBEAR validation procedures.

- **Code**—To ensure full reproducibility of the bare soil research, we release our code and the baseline machine learning models. The code and models are available at the following link: https://doi.org/10.5281/zenodo.17607897.

Our contributions not only fill a crucial gap in existing resources but also provide the research community with an essential tool for rigorously validating and advancing novel techniques in this domain, ultimately fostering more reliable and impactful research in the application of hyperspectral remote sensing to precision agriculture.

1.7 Structure of the paper

The HyBEAR dataset, together with the procedures of (*i*) generating ground-truth delineations and (*ii*) extracting the image patches, as well as the cross-validation settings and quality metrics, are discussed in Section 2. In Section 3, we report and discuss the experimental results constituting the baseline results of the HyBEAR benchmark. Section 4 concludes the paper.

2 Dataset

2.1 Hyperspectral data collection

The HSIs were acquired by QZ Solutions, a company based in Poland and focused on making farming more productive and sustainable with new technologies[1], in the southern Poland on March 3, 2021. Data acquisition was carried out using the HySpex VS-725 hyperspectral imaging system (Norsk Elektro Optikk AS), which consists of two sensors: SWIR-384 (spectral range: 930-2500 [nm], number of bands: 288, spectral resolution: 5.45 [nm]) and VNIR-1800 (spectral range: 400-1000 [nm], number of bands: 186, spectral resolution: 3.26 [nm]). The system was placed on the Piper PA-31 Navajo aircraft (flight altitude 2550-2700 [m], cruising speed 61.8 [m/s], ground sampling distance (GSD) 2 [m], cloudless and windless weather). Finally, for each pixel, we capture 430 spectral bands in the range 414.1–2357.4 [nm] (with the spectral resolution of 3.26 [nm] for the Visible Near-Infrared [VNIR] range, and of 5.45 [nm] for the Short-Wave Infrared [SWIR] range).

[1]For more details concerning QZ Solutions, we refer to the official webpage of the company: https://qzsolutions.eu/.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

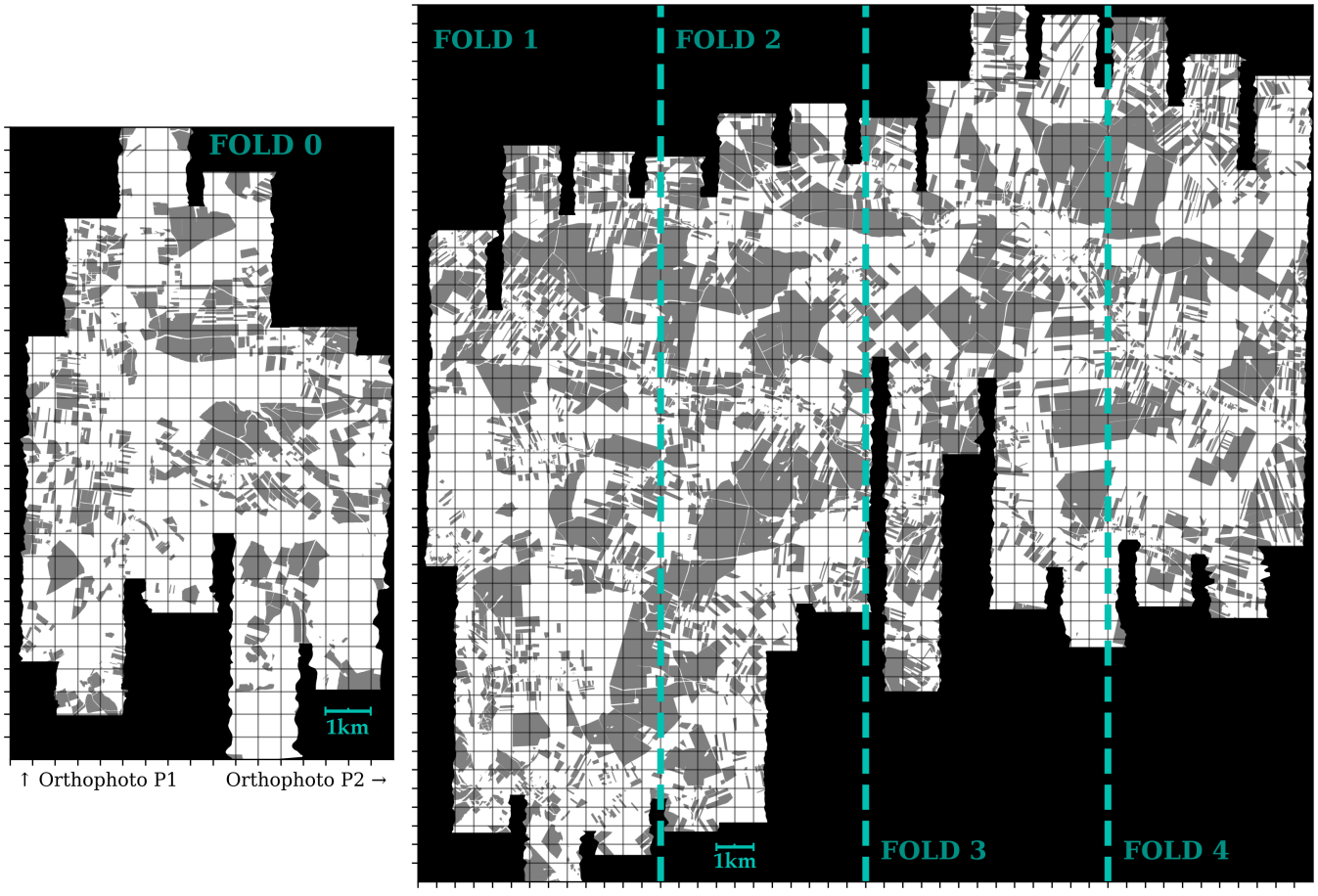

Figure 1. The HyBEAR 🐻 benchmark patches (grid) overlaid onto the two source orthophotomaps with indication of the divisions into the images and the folds (separated on the maps with the green dashed lines). We depict areas labeled as `SOIL` using the gray color, `NON-SOIL` with the white color, and for the regions where we have `NO DATA`, we use the black color.

Data were collected for two areas: the first one depicted with the (*i*) **P1** orthophoto map covering 7,637 [ha], that is 19,092,581 hyperspectral pixels, and the second one, referred to as (*ii*) **P2**, for which we provide imagery of 35,588 [ha] (88,972,010 pixels). These areas are presented in Figure 1. The maps match two different locations in Poland: Lower Silesian Voivodeship (P1), where the map covers the fields in the vicinity of the village of Przeworno, and Opolskie Voivodeship (P2), for which the map reveals hundreds of hilly fields in the area south of the town of Głubczyce. Both locations are more than 60 km apart, and the images were acquired within an hour of each other. Thus, due to the dynamic position of the sun and clouds, the lightning conditions differed, and the HyBEAR dataset is heterogeneous at the spatial and image acquisition levels, hence may be used to quantify the generalization abilities of ML algorithms. Figure 2 displays the significant difference between patches – compare images a) and b) with c) and d) to see how they differ in terms of reflectance levels. Finally, the collected data may be used

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

for two purposes: (*i*) for a regression task aimed at estimating soil component levels, which was the goal of the HYPERVIEW challenge (Nalepa et al., 2022, 2024), and (*ii*) for the bare soil detection, which conditions the accuracy of the estimated component levels. In this paper, we focus exclusively on the latter task.

2.2 Ground truth preparation

High-quality Ground Truth (GT) is paramount for the validity and utility of any benchmark dataset, particularly in remote sensing applications. Accurate GT labels are essential for training robust machine learning models and for objectively evaluating their performance. The preparation of GT for bare soil detection in HSIs poses unique challenges, including spectral overlap between bare soil and other non-vegetated surfaces (e.g., roads and buildings), as well as the potential for sparse or early-stage vegetation that may be difficult to discern.

In this benchmark, the GT was meticulously prepared through a combination of automated and manual interpretation techniques to ensure high accuracy and reliability. While preparing the manual bare-soil outlines, we considered information from vegetation indices, which are sensitive to chlorophyll and can effectively indicate the absence of mature vegetation. However, relying solely on vegetation indices has limitations. For instance, fresh or very sparse vegetation might not yield a strong enough signal in vegetation indices to be reliably excluded as bare soil. Conversely, non-vegetated areas such as roads, buildings, and artificial surfaces can exhibit low vegetation index values, potentially leading to misclassifications if only this information source is used. Additionally, RGB channels were insufficient for accurate bare soil mapping, primarily because the visual distinction between bare soil and other non-vegetated surfaces can be subtle and is dependent on flight altitude. Also, early or stressed vegetation might not be visually apparent in RGB imagery, leading to errors in GT.

In Figure 2, several examples of patches are presented. We prepared two different views of the hyperspectral patches: the RGB composition of the bands, and the CIR (Color Infrared using the near-infrared band) image. As one can see, bare soil fields are not easily distinguishable using only RGB images (examples (a) through (e)). Relying solely on a CIR image can also be misleading (examples (b) and (e)). Objects present in the field of view, such as dirt roads or tree shadows, can make identifying the precise borders of bare soil fields difficult, as seen in (c) and (d) of the attached figure.

To overcome these limitations, a strategy combining the analysis of both vegetation indices and visual information was applied. Initial outlines of potential bare-soil areas were informed by vegetation index analysis, helping identify regions where vegetation was likely absent. Subsequently, these outlines were carefully reviewed and refined by human experts (with 10, 4, and 1 years of hands-on experience), who visually inspected the HSI data and potentially high-resolution RGB imagery (if available). This manual refinement step was crucial for accurately delineating bare soil areas and excluding non-vegetated areas, such as roads, buildings, and shadows, that vegetation indices might have flagged. This iterative process of leveraging automated tools and expert knowledge ensures high accuracy and consistency in the GT labels.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

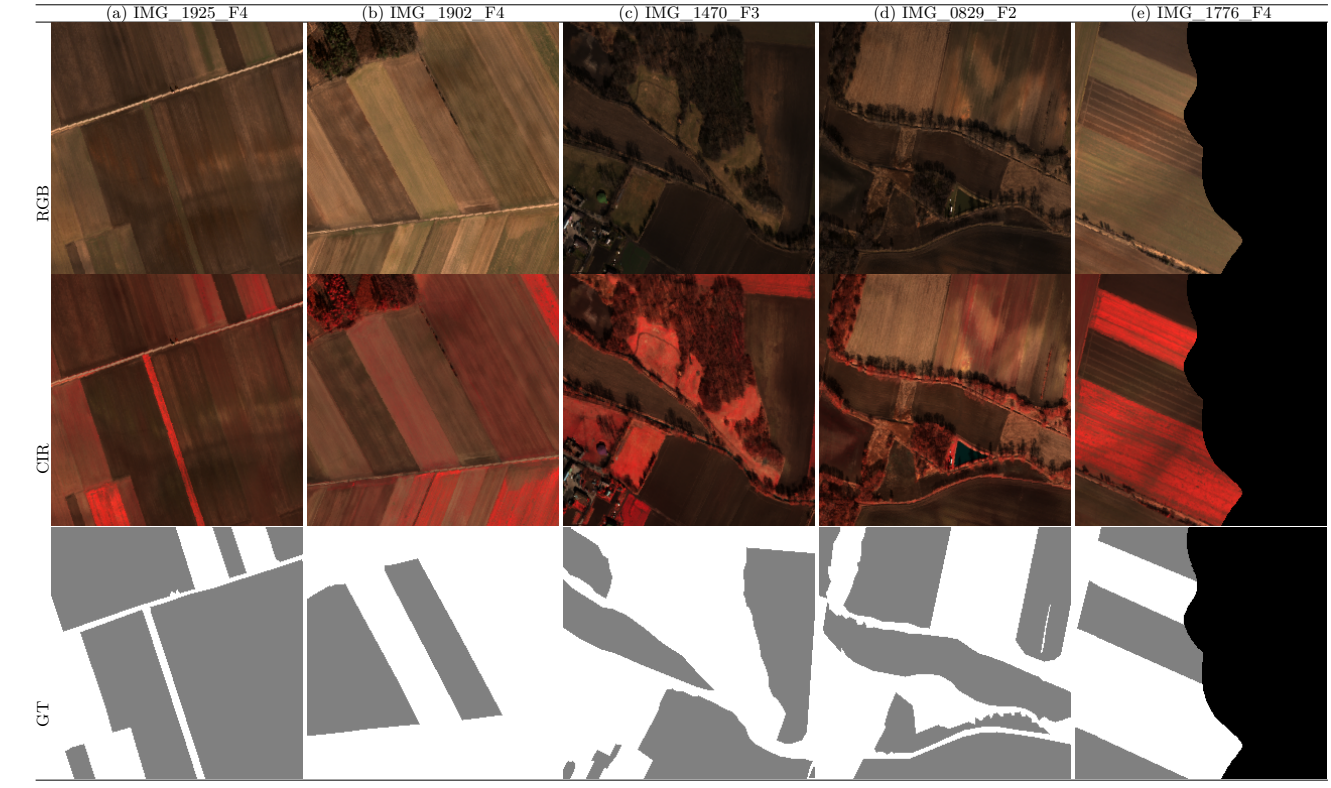

Figure 2. Examples of selected patches from the HyBEAR 🐻: (a) RGB images composed of bands 465.3, 532.4, and 628.3 [nm] of the source hyperspectral patches, (b) CIR images composed of bands 532.4, 628.3, 855.3 [nm], and (c) GT images. We depict areas labeled as `SOIL` using the gray color, `NON-SOIL` with the white color, and for the regions where we have `NO DATA`, we use the black color. The HyBEAR dataset provides original hyperspectral images, RGB, and CIR representations that are composed of 3 bands only and normalized prior to display to enhance their contrast.

The careful preparation of the GT is a critical aspect that contributes to the remote sensing community. By providing a reliable and accurate reference for bare soil, HyBEAR enables researchers to develop and evaluate algorithms for HSI data analysis, ultimately advancing the field of bare soil detection.

2.2.1 Details of the Manual Labeling Process

To facilitate the manual labeling process[2], we worked on RGB, CIR, and NDVI compositions of bands derived from the HSIs. The first image provided a natural view for the annotators, while the subsequent images emphasized vegetation, including that just beginning to grow. Consequently, the vegetation, which exhibits a high degree of reflectance in the infrared spectrum, was highlighted and rendered in vivid colors and pronounced contrast. The use

[2]The annotation process was performed manually in a Python application prepared for this activity, which was based on the LabelMe library (https://pypi.org/project/labelme).

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

of CIR proved particularly advantageous for analyzing fields where vegetation has already developed, enabling clear differentiation between young plants and other environmental components that could affect the manual delineation process.

The annotations initially included two classes: `SOIL` and `MAYBE-SOIL`. The `SOIL` class was assigned in cases where at least approx. 85% of the field surface constituted the visible soil, exhibiting no distinct indications of vegetation. The `MAYBE-SOIL` label was assigned to ambiguous areas where young plants partially covered the soil or where visual evidence indicated that the vegetation process had already begun or was likely to start soon. All instances of the `MAYBE-SOIL` class have been carefully analyzed and discussed with the domain experts, and subsequently converted into either the `SOIL` or `NON-SOIL` class. During the annotation process, efforts were made to avoid terrain obstacles, such as trees, bushes, or infrastructure. The factors influencing soil visibility were also considered, including terrain topography, shading, moisture levels, and the presence of plant residues or harvest remnants.

2.3 Extraction of image patches

To make the data analysis within HyBEAR more affordable and practical, the large source orthophotomaps were split into smaller image patches. The HyBEAR dataset is composed of 1,954 square patches of size `250  250`, that were extracted from two different source othophotops (`P1` and `P2`). We divided the images according to longitude into five *folds* (`F0`, `F1`, `F2`, `F3`, and `F4`) (Figure 1). Since the second map is roughly four times larger, we assigned one fold to `P1` and four folds to the images extracted from `P2` to evenly distribute the images among folds. For each of those patches, in the dataset folder `HyBEAR/images/`, we store two files: an HSI with 430 spectral bands (`IMG_****_F*.TIFF`) and the labels' file (`GT_****_F*.TIFF`), each with its metadata and geospatial information encoded. The name of the file is built using a consecutive file identifier (ID), and a fold number (e.g., `IMG_0555_F1.TIFF`). In addition to the full version of the HyBEAR dataset, we offer a lighter version of our collection, referred to as the MINI subset of HyBEAR. We randomly selected a subset of 50 images from each fold, stratifying by the proportion of soil pixels in each subset. The dataset statistics are summarized in Table 1.

It is worth noting that the primary reason we chose to use the `TIFF` file format for all images is that it allows us to include geolocation information for each patch. Also, it preserves the full image encoding precision and supports decent data compression. For HyBEAR, we selected a lossless deflate compression method that reduced the average patch size from approx. 215 [MB] per image to around 40-48 [MB].

The resulting image collection was tested using two different environments: `QGIS` (versions 3.28 and 3.40) and a Python-based application (employing `Rasterio`, `NumPy`, and `Scikit-Learn` libraries—the configuration file of the Python environment, as well as the code we use for testing, is delivered with this dataset to support reproducibility).

2.4 Cross-validation protocol

The HyBEAR dataset consists of two independent aerial hyperspectral scenes, `P1` and `P2`, which were carefully selected to enable a comprehensive and rigorous evaluation of algorithms dedicated to bare soil detection. One of

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

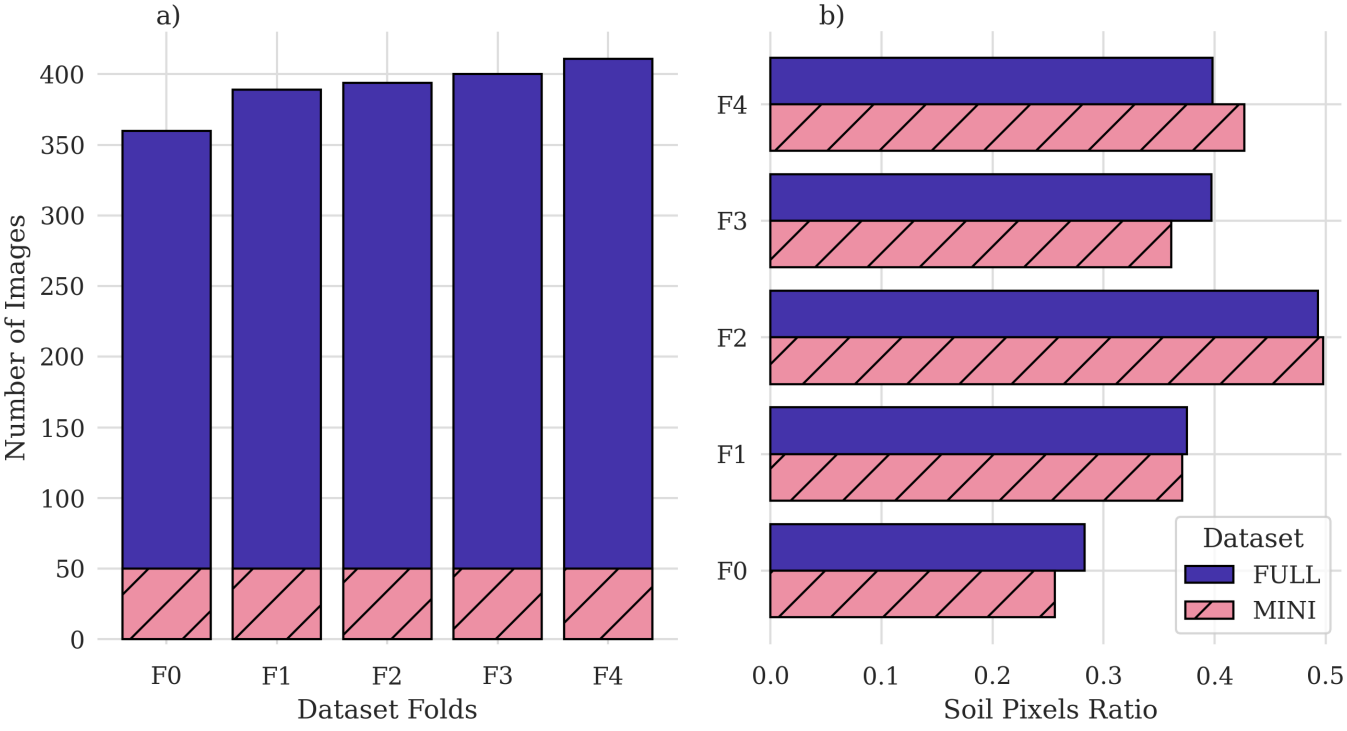

Figure 3. The HyBEAR dataset details summary. On the left: a) we depict the number of images for every fold and MINI subsets. On the right: b) we visualize the proportion of the soil and non-soil classes across all folds and subsets (for the FULL and MINI versions of the dataset).

the key aspects of HyBEAR is the evaluation protocol, based on the multi-fold cross-validation scenarios. To verify the generalization capability of the bare soil detection methods, as well as their robustness to the variability of the acquired data, we defined the experimental scenarios within a five-fold cross-validation framework:

- *Stage 1*: training an algorithm on the folds `F1` to `F4` ($\boldsymbol{T}_{F1:F4}$) → testing on the fold `F0` (Ψ_{F0}).
- *Stage 2*: training an algorithm on the folds `F=1` ($\boldsymbol{T}_{F0,F2:F4}$) → testing on the fold `F1` (Ψ_{F1}).
- *Stage 3*: training an algorithm on the folds `F=2` ($\boldsymbol{T}_{F0:F1,F3:F4}$) → testing on the fold `F2` (Ψ_{F2}).
- *Stage 4*: training an algorithm on the folds `F=3` ($\boldsymbol{T}_{F0:F2,F4}$) → testing on the fold `F3` (Ψ_{F3}).
- *Stage 5*: training an algorithm on the folds `F=4` ($\boldsymbol{T}_{F0:F3}$) → testing on the fold `F4` (Ψ_{F4}).

This evaluation protocol has been designed with several key objectives in mind, relevant to the evaluation of algorithms in the context of real-world applications of bare soil detection from aerial data:

- **Evaluation of generalization to spatially independent areas**—the hyperspectral scenes `P1` (`F0`) and `P2` (`F1:F4`) represent data acquired from different geographical locations. Testing on one area after training on the

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

other allows a direct assessment of the algorithm's ability to generalize knowledge to new, previously unseen areas. This is crucial, as in practice, models are rarely applied to exactly the same areas where they were trained.

- **Verification of robustness to the variability of acquisition conditions**—aerial data can exhibit significant spectral and spatial variability depending on atmospheric conditions, illumination, season, or sensor configuration. Our evaluation protocol, by testing on spatially-disjoint areas (`P1` and `P2`, see Figure 1), enables the assessment of the algorithms' robustness to such factors. An algorithm that performs well in both scenarios demonstrates greater reliability.

- **Preventing overfitting to the specifics of a single dataset**—training and testing on the same dataset can lead to inflated performance evaluations if the model learns specific characteristics of that dataset that are not universal. Our cross-validation protocol minimizes this risk by forcing the model to learn more general and representative features of bare soil.

- **Enabling comparability of results**—defining clear data splitting scenarios as part of the benchmark ensures that different bare soil detection methods can be evaluated in a consistent and comparable manner. This allows researchers to objectively quantify and assess the progress in this field.

To facilitate a fair comparison of different bare soil detection methods, the following commonly adopted classification and segmentation metrics will be used to quantitatively evaluate the models within the presented evaluation protocol—these metrics are presented in the next section.

2.5 Evaluation metrics

To evaluate the performance of the emerging ML models, a series of commonly adopted classification metrics is used (Powers, 2011). This choice is motivated by the need for a comprehensive analysis of the models' ability to correctly identify bare soil pixels.

- **Accuracy** (ACC) measures the overall percentage of correctly classified samples, and it is calculated as:

$$\mathrm{ACC} \quad \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{1}$$

  where TP, TN, FP, and FN represent true positives (correctly classified bare soil image pixels), true negatives (correctly classified non-bare soil image pixels), false positives (the non-bare soil image pixels incorrectly classified as those containing bare soil), and false negatives (bare soil image pixels incorrectly classified as the non-bare soil image pixels). Accuracy represents a fundamental measure of classification effectiveness.

- **Sensitivity** (SEN) determines the model's ability to identify all actual positive cases (bare soil image pixels):

$$\mathrm{SEN} \quad \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}}. \tag{2}$$

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

High sensitivity is crucial for minimizing bare soil omissions (false negatives).

– **Specificity** (SPE) measures the model's ability to correctly identify all actual negative cases (image pixels not containing bare soil):

$$\text{SPE} \quad \frac{\text{TN}}{\text{TN}+\text{FP}}. \tag{3}$$

High specificity ensures that the model rarely generates false positive detections.

– **F-score** (F1) represents the harmonic mean of precision ($\frac{\text{TP}}{\text{TP}\ \ \text{FP}}$) and sensitivity, providing a balanced measure of performance, particularly important for potentially imbalanced datasets:

$$\text{F1} \quad 2 \cdot \frac{\text{Precision} \cdot \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}}. \tag{4}$$

– **Intersection over Union** (IoU) measures the degree of overlap between the predicted and actual positive outcomes. It is particularly useful for evaluating the quality of segmentation or object detection. It is calculated as the ratio of the intersection of the sets of predicted positives and actual positives to their union:

$$\text{IoU} \quad \frac{\text{TP}}{\text{TP}+\text{FP}+\text{FN}}. \tag{5}$$

– **Matthews Correlation Coefficient** (MCC) is a measure of the correlation between observed and predicted binary classifications, more robust to imbalanced data than ACC or F1 (Chicco et al., 2021). It is calculated as:

$$\text{MCC} \quad \frac{\text{TP} \cdot \text{TN} \quad \text{FP} \cdot \text{FN}}{\sqrt{\text{TP}+\text{FP} \quad \text{TP}+\text{FN} \quad \text{TN}+\text{FP} \quad \text{TN}+\text{FN}}}. \tag{6}$$

– The **Receiver Operating Characteristic** (ROC) **curve** illustrates the relationship between SEN and SPE for various classification thresholds (Powers, 2011). The **Area Under this Curve** (AUC) quantifies the overall ability of the model to discriminate between classes, where higher AUC values indicate better performance.

Collectively, these metrics are crucial to provide a comprehensive, objective, and standardized way to assess and compare the performance of various methods in the bare soil detection task.

2.6 Dataset organization

The data is organized in the following directory structure:

– `HyBEAR/`—the main directory serving as the root for the entire dataset. It contains all the necessary files and subdirectories to organize the hyperspectral patches and their corresponding GTs for both `P1` and `P2`.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

- HyBEAR/images/—this directory is the core of the dataset, containing the extracted hyperspectral patches and their associated GTs. This organization allows for direct access to the data for training and testing purposes:
  - HyBEAR/images/F0_FULL/, …, HyBEAR/images/F4_FULL/, HyBEAR/images/F0_MINI/, …, HyBEAR/images/F4_MINI/—these directories contain the TIFF image patches and the GT patches for folds F0:F4, and for MINI or FULL subset.
- HyBEAR/code/—this directory stores several Python Jupyter Notebook files: for the initial dataset presentation, for displaying the hyperspectral patches and their associated GTs, for the reproduction of the reported benchmark results of the best-performing machine learning models, as well as the configuration file allowing the installation of the Python open-source libraries that are necessary to process the dataset.
- HyBEAR/models/—this directory contains the models trained using HyBEAR under the suggested 5-fold cross-validation regime. We deliver models for each validation phase and for two different machine learning algorithms (Logistic Regression and Support Vector Machines). Therefore, this folder contains 10 files. These models can be verified using the HyBEAR/code/model_evaluation.ipynb that was developed to support the classification of each hyperspectral image and pixel.

The hyperspectral image patches (files from IMG_0000_F0.tiff to IMG_1953_F4.tiff) contain information about the light reflectance values in different spectral ranges for each pixel. The GT patches (files from GT_0000_F0.tiff to GT_1953_F4.tiff) store binary information and indicate which pixels in the hyperspectral patch belong to:

- the SOIL class—encoded as (1),
- and which do **not** depict SOIL (hence, are NON-SOIL)—encoded as (0),
- additionally, for the pixels extracted from the edges of the main hyperspectral scene that do not contain data in the image, we use the code (-9999) and encode them in the GT images as background, thereby supporting their automated filtration.

The patches have fixed dimensions of 250 250 pixels and were generated using a grid-based method from the original hyperspectral scenes P1 and P2. The dataset (consisting of these patches) is available at https://doi.org/10.5281/zenodo.17607897. The entire dataset comprises 1,954 data tuples and has a total size of 96 [GB] (for the compressed files, with the average size of a single tuple being 48 [MB]).

3 Experiments and baseline results

3.1 Methods

In addition to the dataset, several machine learning models were evaluated under the introduced cross-validation regime. We trained models on the hyperspectral pixels extracted separately from each image, and we employed

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

Table 1. The HyBEAR 🐻 dataset structure and descriptive statistics.

Fold	Source Map Id	Number of Images	Number of Soil Pixels	Number of Non-Soil Pixels	Soil / Non-Soil Pixel Ratio [%]	Background Pixel Ratio [%]	Size [MB]
FULL dataset							
F0	**P1**	360	5,407,227	13,685,354	28.3	15.1	17785.35
F1	**P2**	389	8,079,009	13,469,713	37.5	11.4	19398.19
F2	**P2**	394	11,377,235	11,696,924	49.3	6.3	20196.45
F3	**P2**	400	8,557,984	13,012,410	39.7	13.7	18807.78
F4	**P2**	411	9,058,907	13,719,828	39.8	11.3	19837.53
	P1 ∪ P2	**1,954**	**42,480,362**	**65,584,229**	39.3	11.5	96025.30
MINI dataset							
F0	**P1**	50	700,992	2,040,112	25.6	12.3	2559.51
F1	**P2**	50	998,715	1,690,212	37.1	14.0	2429.20
F2	**P2**	50	1,397,507	1,410,897	49.8	10.1	2449.66
F3	**P2**	50	990,808	1,753,872	36.1	12.2	2394.34
F4	**P2**	50	1,250,892	1,680,835	42.7	6.2	2549.96
	P1 ∪ P2	**250**	**5,338,914**	**8,575,928**	38.4	10.9	12382.67

the standard machine learning model architectures: Logistic Regression (later denoted as LR), Linear models with L2 Regularization (L2), Adaptive Boosting (AB), Support Vector Machines with a linear kernel (SVM), Decision Trees (DT), and Random Forest (RF). All models operate on feature vectors of size 430, which are the values of all spectral bands captured within a particular pixel. The selected models enable the efficient processing of the extensive HyBEAR dataset, for which the number of pixels for an image could reach 62,500. Since the number of images in each fold varies, and as we exclude the background pixels that do not contain any data, the overall number of training samples varies accordingly. The ratio of "useful" pixels (excluding background) regarding every fold is reported in Table 1.

3.2 Baseline results

In the experiments, we exploit both versions of the HyBEAR benchmark, with all the images (FULL), and using only a subset of 50 images from each fold (MINI). Table 2 provides the aggregated metrics, together with their standard deviation. The detailed results of every fold are gathered in the appendix (Table A1). Additionally, the results are depicted in Figure 4. Here, in addition to the metrics computed for every fold (points on the dashed lines), we plotted the average metric levels (dotted lines). Following this visualization should allow for spotting the models that are more consistent across the folds, and therefore, would return a more robust and universal classifier.

The highest performance scores were obtained for the LR-based models for the majority of metrics, with the SVM-based models reaching nearly the same performance, both for the FULL and MINI versions of HyBEAR. In Figure 5, we present three example patches, with the GT labels, and the corresponding predictions elaborated using

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

the investigated machine learning models. Here, the red areas show the false positive bare soil detections, whereas the yellow areas correspond to the false negative detections. The average accuracy for some models reached a decent level of 0.926–0.927 (for SVM and LR, respectively), and the F-score was almost 0.9 (0.898 for the best-performing LR model). When previewing the predictions for these models over the example test patches (Figure 5), we can observe that there is still significant room for improvement for most of the models. For some less challenging patches (i.e., IMG_0717_F1), the results are satisfactory, but for some others (i.e., IMG_0022_F0), most of the models perform poorly. For this image, only L2, LR, and SVM work reasonably well. One could notice that, on average, the less complicated linear models allow for better solutions, less prone to noise, and more universal for the new data (Table A1). However, capturing intrinsic bare soil features, through e.g., automated representation learning in deep learning models, may help boost the capabilities of the detectors—this should be further researched.

Table 2. The HyBEAR 🐻 benchmark baseline results (the average and standard deviation, $\pm\sigma$) averaged after the 5-fold cross-validation procedure. The best results for each metric are **boldfaced**.

Model	ACC	SEN	SPE	F1	IoU	MCC	ROC
	σ	σ	σ	σ	σ	σ	σ
Evaluation using all images from each fold – FULL dataset							
AdaBoost	0.857 ± 0.025	0.825 ± 0.158	0.865 ± 0.044	0.804 ± 0.098	0.680 ± 0.125	0.691 ± 0.096	0.845 ± 0.062
Decision Trees	0.881 ± 0.033	0.776 ± 0.170	0.934 ± 0.017	0.817 ± 0.121	0.703 ± 0.153	0.733 ± 0.113	0.855 ± 0.077
Linear w $_2$ regularization	0.916 ± 0.013	**0.965** ± 0.028	0.882 ± 0.029	0.897 ± 0.023	0.814 ± 0.039	0.831 ± 0.024	**0.923** ± 0.008
Logistic Regression	**0.927** ± 0.016	0.910 ± 0.099	0.928 ± 0.025	**0.898** ± 0.051	**0.819** ± 0.080	**0.841** ± 0.052	0.919 ± 0.038
Random Forest	0.919 ± 0.038	0.840 ± 0.182	**0.953** ± 0.019	0.869 ± 0.117	0.782 ± 0.162	0.816 ± 0.112	0.897 ± 0.083
Support Vector Machines	0.926 ± 0.016	0.909 ± 0.108	0.927 ± 0.027	0.897 ± 0.054	0.817 ± 0.084	**0.841** ± 0.053	0.918 ± 0.041
Evaluation using 50 selected images from each fold – MINI dataset							
AdaBoost	0.823 ± 0.024	0.813 ± 0.133	0.822 ± 0.064	0.776 ± 0.074	0.638 ± 0.095	0.634 ± 0.070	0.818 ± 0.045
Decision Trees	0.866 ± 0.042	0.777 ± 0.173	0.913 ± 0.028	0.804 ± 0.118	0.684 ± 0.150	0.707 ± 0.116	0.845 ± 0.077
Linear w $_2$ regularization	0.890 ± 0.026	**0.970** ± 0.022	0.836 ± 0.042	0.872 ± 0.034	0.774 ± 0.055	0.789 ± 0.044	0.903 ± 0.018
Logistic Regression	**0.907** ± 0.019	0.923 ± 0.072	0.891 ± 0.047	**0.884** ± 0.034	**0.793** ± 0.055	**0.810** ± 0.039	**0.907** ± 0.022
Random Forest	0.906 ± 0.044	0.846 ± 0.183	**0.934** ± 0.032	0.860 ± 0.112	0.766 ± 0.155	0.797 ± 0.112	0.890 ± 0.080
Support Vector Machines	0.906 ± 0.018	0.922 ± 0.079	0.890 ± 0.041	0.882 ± 0.037	0.791 ± 0.060	0.808 ± 0.041	0.906 ± 0.025

4 Conclusions

Bare soil detection is an important task in precision agriculture, as it allows for determining the areas in Earth observation imagery that shall be further analyzed while estimating specific soil parameters and features. In this paper, we introduced HyBEAR 🐻—a comprehensive hyperspectral benchmark dataset, accompanied by the baseline results for automated bare soil detection from HSIs. The provided large-scale data collection delivers high-resolution hyperspectral imagery (with the 2 m GSD), together with the carefully prepared ground-truth delineations of bare soil areas. These annotations were meticulously verified in collaboration with domain experts who routinely analyze

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

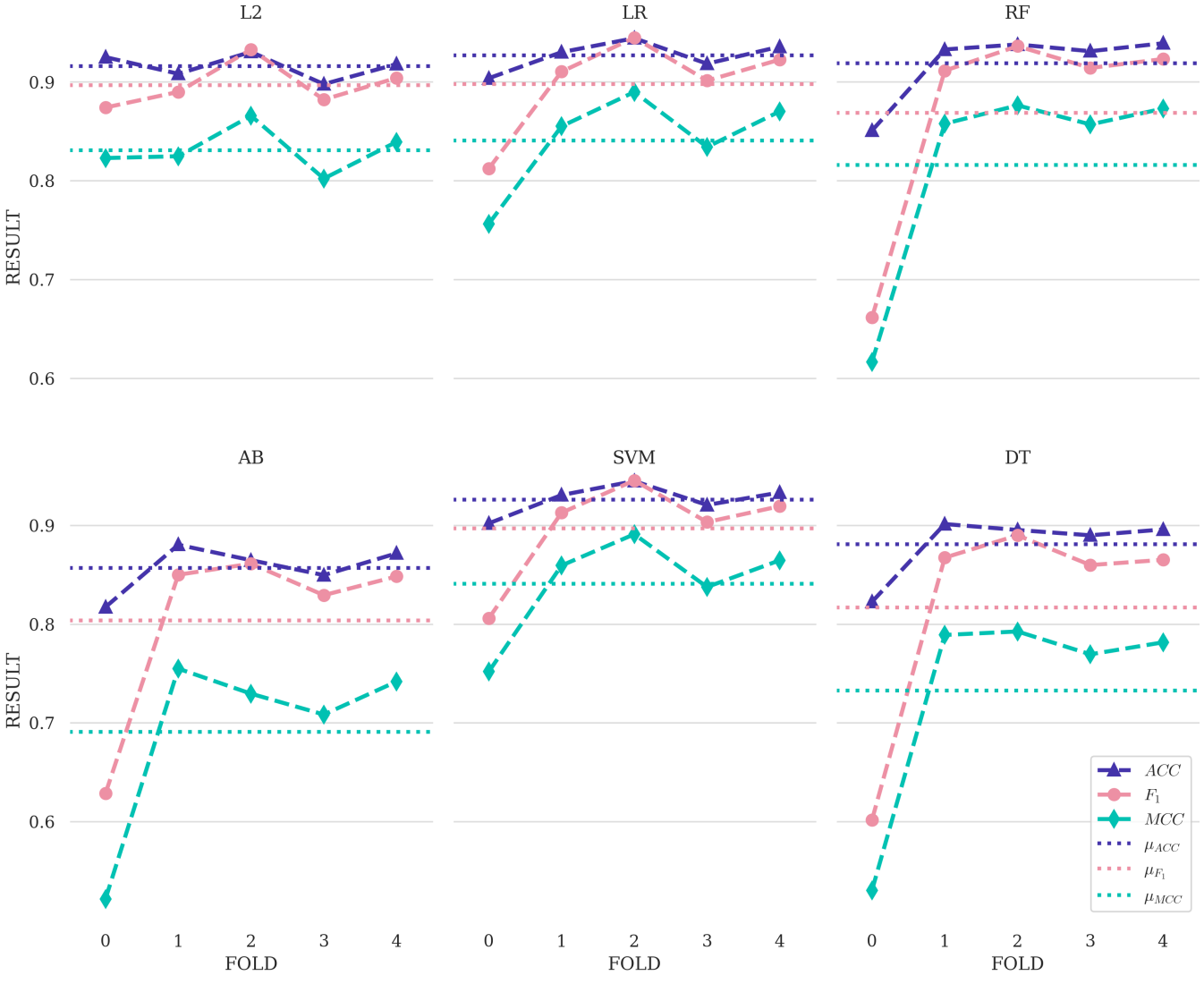

Figure 4. The benchmarking results for the investigated machine learning models. We present the results for each fold separately, along with the average result for each metric (dotted line).

soil samples. Finally, the provided airborne imagery is free of disturbances, with no cloud cover or dust in the camera's field of view. To our knowledge, the HyBEAR dataset is the first in the literature that offers bare soil pixels captured for specific agricultural parcels (i.e., it is not a set of spatially unaware multi/hyperspectral pixels). Therefore, it may be utilized in practical examples, where delineating the field boundaries is key. HyBEAR defines the five-fold cross-validation procedure and is accompanied by the baseline results. We strive to ensure full reproducibility of any research that will emerge based on the provided benchmark, and thus made our implementations and code examples publicly available.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

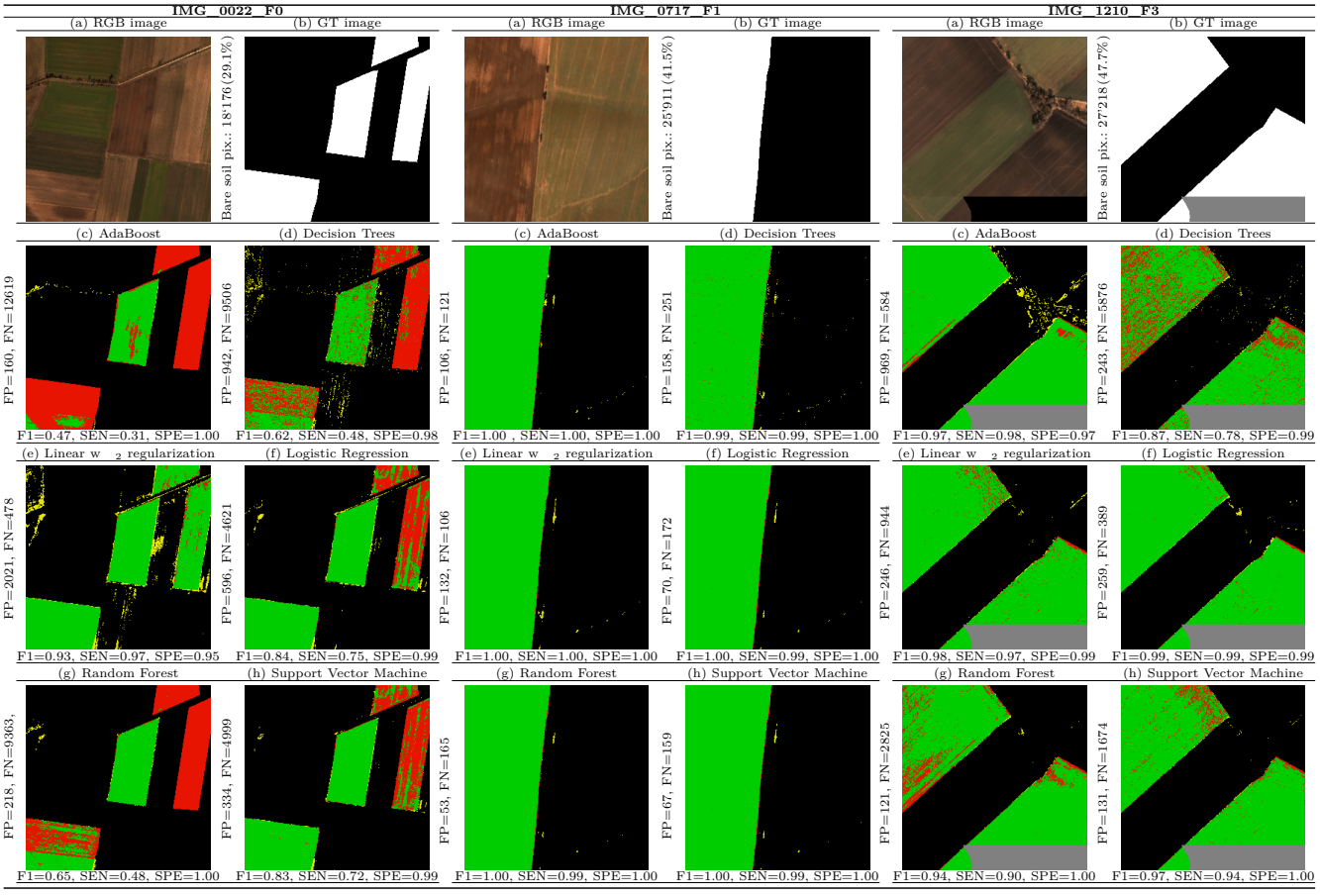

Figure 5. Example bare soil detection results with the quantitative metrics: (a) RGB images with (b) GT marked in white, (c-h) various prediction methods (TP—green, FP—yellow, FN—red, TN—black, missing data/background—gray).

The experimental results reported in this work will serve as the point of departure for future work, potentially in a multitude of use cases related to hyperspectral data analysis. The quantitative and qualitative results showed that there are areas where classic machine learning models (operating on spectral curves for each hyperspectral pixel) fail to appropriately identify bare soil. Thus, extracting more discriminative features and leveraging automated representation learning offered by deep learning (Guerri et al., 2024) would likely lead to higher-quality bare soil detection. It is important to emphasize that the HyBEAR dataset may be exploited for other tasks, such as feature extraction (Zhang et al., 2024), feature selection (Tan et al., 2025), unsupervised clustering/domain adaptation (Cai et al., 2024), and many more.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

5 Code and data availability

We published HyBEAR 🐻 – Wijata et al. (2025) on Zenodo: https://doi.org/10.5281/zenodo.17607897. In addition to the images and data, this package includes the Python code for evaluating the machine learning models with HyBEAR, as well as other code examples.

Appendix A: Detailed experimental results

Table A1: The detailed results for all evaluated algorithms, computed under a 5-fold cross-validation regime in two variants: using all available images (FULL) and the MINI subset.

Model	Fold	Performance Metrics						
		ACC	SEN	SPE	F1	IoU	MCC	AUC
Evaluation using all images from each fold – FULL dataset								
AdaBoost	0	0.817	0.546	0.925	0.629	0.458	0.522	0.735
	1	0.880	0.904	0.866	0.850	0.739	0.755	0.885
	2	0.865	0.852	0.878	0.861	0.756	0.730	0.865
	3	0.850	0.921	0.803	0.829	0.709	0.709	0.862
	4	0.872	0.902	0.852	0.848	0.737	0.742	0.877
Decision Trees	0	0.823	0.473	0.961	0.602	0.431	0.531	0.717
	1	0.901	0.860	0.926	0.867	0.766	0.789	0.893
	2	0.895	0.858	0.932	0.890	0.802	0.793	0.895
	3	0.890	0.851	0.916	0.860	0.754	0.770	0.883
	4	0.896	0.839	0.933	0.865	0.762	0.782	0.886
Linear w L_2 regularization	0	0.925	0.917	0.928	0.874	0.777	0.823	0.923
	1	0.908	0.989	0.860	0.890	0.802	0.825	0.925
	2	0.931	0.982	0.881	0.933	0.875	0.866	0.931
	3	0.898	0.964	0.854	0.882	0.789	0.802	0.909
	4	0.918	0.971	0.884	0.904	0.825	0.839	0.927
Logistic Regression	0	0.904	0.734	0.971	0.812	0.684	0.756	0.853
	1	0.930	0.946	0.921	0.911	0.836	0.855	0.933
	2	0.944	0.967	0.922	0.945	0.896	0.890	0.945
	3	0.919	0.939	0.905	0.902	0.821	0.834	0.922
	4	0.936	0.962	0.919	0.923	0.857	0.870	0.940
Random Forest	0	0.851	0.514	0.984	0.662	0.494	0.617	0.749
	1	0.933	0.920	0.941	0.912	0.838	0.858	0.931
	2	0.938	0.924	0.952	0.936	0.880	0.876	0.938
	3	0.931	0.924	0.936	0.914	0.842	0.857	0.930

Continued on next page...

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Table A1 – continued from previous page

Model	Fold	ACC	SEN	SPE	F1	IoU	MCC	ROC
	4	0.939	0.919	0.953	0.923	0.858	0.873	0.936
Support Vector Machines	0	0.902	0.718	0.975	0.806	0.675	0.752	0.846
	1	0.931	0.967	0.909	0.913	0.840	0.859	0.938
	2	0.945	0.971	0.920	0.946	0.897	0.891	0.945
	3	0.921	0.935	0.912	0.903	0.824	0.838	0.923
	4	0.933	0.956	0.918	0.920	0.851	0.865	0.937
Evaluation using 50 selected images from each fold – MINI dataset								
AdaBoost	0	0.801	0.577	0.908	0.652	0.484	0.524	0.742
	1	0.841	0.880	0.820	0.798	0.663	0.677	0.850
	2	0.854	0.851	0.857	0.853	0.743	0.708	0.854
	3	0.804	0.894	0.743	0.786	0.648	0.625	0.818
	4	0.816	0.866	0.782	0.790	0.653	0.636	0.824
Decision Trees	0	0.796	0.471	0.951	0.599	0.427	0.508	0.711
	1	0.886	0.870	0.894	0.844	0.730	0.754	0.882
	2	0.904	0.880	0.927	0.901	0.819	0.808	0.903
	3	0.864	0.843	0.878	0.834	0.715	0.719	0.861
	4	0.878	0.824	0.915	0.845	0.731	0.746	0.869
Linear w L_2 regularization	0	0.905	0.935	0.891	0.864	0.761	0.797	0.913
	1	0.873	0.992	0.808	0.847	0.735	0.766	0.900
	2	0.928	0.984	0.873	0.931	0.871	0.861	0.928
	3	0.866	0.968	0.797	0.853	0.744	0.750	0.882
	4	0.877	0.971	0.814	0.863	0.760	0.769	0.892
Logistic Regression	0	0.909	0.795	0.964	0.850	0.739	0.789	0.879
	1	0.904	0.942	0.883	0.875	0.777	0.803	0.913
	2	0.938	0.972	0.904	0.940	0.886	0.878	0.938
	3	0.888	0.948	0.847	0.872	0.773	0.781	0.897
	4	0.897	0.957	0.857	0.881	0.788	0.799	0.907
Random Forest	0	0.830	0.520	0.978	0.664	0.497	0.603	0.749
	1	0.916	0.924	0.912	0.887	0.797	0.822	0.918
	2	0.943	0.945	0.941	0.943	0.892	0.886	0.943
	3	0.911	0.932	0.896	0.894	0.808	0.819	0.914
	4	0.930	0.910	0.943	0.912	0.838	0.853	0.926
Support Vector Machines	0	0.902	0.785	0.958	0.838	0.721	0.772	0.871
	1	0.904	0.971	0.866	0.877	0.781	0.809	0.919
	2	0.938	0.975	0.902	0.940	0.886	0.878	0.938
	3	0.895	0.938	0.866	0.878	0.783	0.791	0.902
	4	0.894	0.944	0.860	0.877	0.781	0.790	0.902

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Open Access Earth System Science Data Discussions

Author contributions. Conceptualization – AMW, JN, BR, AN; Data curation – AMW, BR, AN, KS; Formal analysis – AMW; Investigation – AMW, BR; Methodology – AMW, BR, JN; Project administration – MG, NL; Resources – KS; Software – AN, BR; Supervision – AMW, BR, JN; Validation – AN, AMW, BR, JN; Visualization – AN, AMW, BR; Writing – original draft – AMW, BR, AN, JN; Writing – review & editing – AMW, BR, JN, AN, NL, KS.

Competing interests. The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgements. AMW and JN were supported by the Silesian University of Technology grant for maintaining and developing research potential. The Article Processing Charge was financed under the European Funds for Silesia 2021-2027 Program co-financed by the Just Transition Fund—project entitled "Development of the Silesian biomedical engineering potential in the face of the challenges of the digital and green economy (BioMeDiG)"; project number: FESL.10.25-IZ.01-07G5/23, and was co-financed by the European Regional Development Fund, through the project entitled "Development of a system enabling remote assessment of soil macronutrient content and soil pH using new machine learning algorithms based on hyperspectral imaging (SmartSoil)"; project number: POIR.01.01.01-00-0287/21. This work was financially supported by the Opole University of Technology as part of the DELTA project no. 314/25. JN was supported by the Silesian University of Technology Rector's grant: 02/080/RGJ25/0052.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

References

Adão, T., Hruška, J., Pádua, L., Bessa, J., Peres, E., Morais, R., and Sousa, J. J.: Hyperspectral Imaging: A Review on UAV-Based Sensors, Data Processing and Applications for Agriculture and Forestry, Remote Sensing, 9, https://doi.org/10.3390/rs9111110, 2017.

Aijaz, N., Lan, H., Raza, T., Yaqub, M., Iqbal, R., and Pathan, M. S.: Artificial Intelligence in Agriculture: Advancing Crop Productivity and Sustainability, Journal of Agriculture and Food Research, 20, 101 762, https://doi.org/10.1016/j.jafr.2025.101762, 2025.

Allied Market Research: Smart Agriculture Market by Type and Component: Opportunity Analysis and Industry Forecast, 2021–2027, Tech. rep., Allied Market Research, https://www.alliedmarketresearch.com/smart-agriculture-market, 2023.

Boguszewska-Mańkowska, D., Ruszczak, B., and Zarzyńska, K.: Classification of Potato Varieties Drought Stress Tolerance Using Supervised Learning, Applied Sciences, 12, https://doi.org/10.3390/app12041939, 2022.

Bünemann, E. K., Bongiorno, G., Bai, Z., Creamer, R. E., De Deyn, G., de Goede, R., Fleskens, L., Geissen, V., Kuyper, T. W., Mäder, P., Pulleman, M., Sukkel, W., van Groenigen, J. W., and Brussaard, L.: Soil Quality – A Critical Review, Soil Biology and Biochemistry, 120, 105–125, https://doi.org/10.1016/j.soilbio.2018.01.030, 2018.

Cai, M., Xi, B., Li, J., Feng, S., Li, Y., Li, Z., and Chanussot, J.: Mind the Gap: Multilevel Unsupervised Domain Adaptation for Cross-Scene Hyperspectral Image Classification, IEEE Transactions on Geoscience and Remote Sensing, 62, 1–14, https://doi.org/10.1109/TGRS.2024.3407952, 2024.

Campos, L. R., Demattê, J. A., Bellinaso, H., Poppiel, R. R., Greschuk, L. T., Rizzo, R., Rosin, N. A., and Rosas, J. T. F.: Detection of Bare Soils in Sugarcane Areas by Temporal Satellite Images: A Monitoring Technique for Soil Security, Soil Security, 7, 100 057, https://doi.org/10.1016/j.soisec.2022.100057, 2022.

Chen, S., Arrouays, D., Leatitia Mulder, V., Poggio, L., Minasny, B., Roudier, P., Libohova, Z., Lagacherie, P., Shi, Z., Hannam, J., Meersmans, J., de Forges, A. C. R., and Walter, C.: Digital Mapping of GlobalSoilMap Soil Properties at a Broad Scale: A Review, Geoderma, 409, 115 567, https://doi.org/10.1016/j.geoderma.2021.115567, 2022.

Chicco, D., Tötsch, N., and Jurman, G.: The Matthews Correlation Coefficient (MCC) is More Reliable Than Balanced Accuracy, Bookmaker Informedness, and Markedness in Two-Class Confusion Matrix Evaluation, BioData Mining, 14, 13, https://doi.org/10.1186/s13040-021-00244-z, 2021.

Delaney, B., Tansey, K., and Whelan, M.: Satellite Remote Sensing Techniques and Limitations for Identifying Bare Soil, Remote Sensing, 17, https://doi.org/10.3390/rs17040630, 2025.

Finger, R., Swinton, S. M., El Benni, N., and Walter, A.: Precision Farming at the Nexus of Agricultural Production and the Environment, Annual Review of Resource Economics, 11, 313–335, https://doi.org/10.1146/annurev-resource-100518-093929, 2019.

Grabowski, B., Ziaja, M., Kawulok, M., Cwiek, M., Lakota, T., Longépé, N., and Nalepa, J.: Are Cloud Detection U-Nets Robust Against in-Orbit Image Acquisition Conditions?, in: IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, pp. 239–242, https://doi.org/10.1109/IGARSS46834.2022.9884919, 2022.

Grabowski, B., Ziaja, M., Kawulok, M., Bosowski, P., Longépé, N., Le Saux, B., and Nalepa, J.: Squeezing Adaptive Deep Learning Methods with Knowledge Distillation for On-Board Cloud Detection, Engineering Applications of Artificial Intelligence, 132, https://doi.org/10.1016/j.engappai.2023.107835, 2024.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Guerri, M. F., Distante, C., Spagnolo, P., Bougourzi, F., and Taleb-Ahmed, A.: Deep Learning Techniques for Hyperspectral Image Analysis in Agriculture: A Review, ISPRS Open Journal of Photogrammetry and Remote Sensing, 12, 100 062, https://doi.org/https://doi.org/10.1016/j.ophoto.2024.100062, 2024.

Han, L., Yang, G., Dai, H., Xu, B., Yang, H., Feng, H., Li, Z., and Yang, X.: Modeling Maize Above-Ground Biomass Based on Machine Learning Approaches using UAV Remote-Sensing Data, Plant Methods, 15, 1746–4811, https://doi.org/10.1186/s13007-019-0394-z, 2019.

Hong, Y., Chen, S., Chen, Y., Linderman, M., Mouazen, A. M., Liu, Y., Guo, L., Yu, L., Liu, Y., Cheng, H., and Liu, Y.: Comparing Laboratory and Airborne Hyperspectral Data for the Estimation and Mapping of Topsoil Organic Carbon: Feature Selection Coupled with Random Forest, Soil and Tillage Research, 199, 104 589, https://doi.org/10.1016/j.still.2020.104589, 2020.

Izquierdo-Verdiguier, E. and Zurita-Milla, R.: An Evaluation of Guided Regularized Random Forest for Classification and Regression Tasks in Remote Sensing, International Journal of Applied Earth Observation and Geoinformation, 88, 102 051, https://doi.org/10.1016/j.jag.2020.102051, 2020.

Ji, S., Zhang, C., Xu, A., Shi, Y., and Duan, Y.: 3D Convolutional Neural Networks for Crop Classification with Multi-Temporal Remote Sensing Images, Remote Sens., 10, https://doi.org/10.3390/rs10010075, 2018.

Jin, X., Li, Z., Feng, H., Ren, Z., and Li, S.: Deep Neural Network Algorithm for Estimating Maize Biomass Based on Simulated Sentinel 2A Vegetation Indices and Leaf Area Index, Crop J., 8, 87–97, 2020.

Joshi, G. P., Alenezi, F., Thirumoorthy, G., Dutta, A. K., and You, J.: Ensemble of Deep Learning-Based Multimodal Remote Sensing Image Classification Model on Unmanned Aerial Vehicle Networks, Mathematics, 9, https://www.mdpi.com/2227-7390/9/22/2984, 2021.

Kapoor, S. and Narayanan, A.: Leakage and the Reproducibility Crisis in Machine-Learning-Based Science, Patterns, 4, 100 804, 2023.

Kaur, R. and Pandey, P.: A Review on Spectral Indices for Built-up Area Extraction using Remote Sensing Technology, Arabian Journal of Geosciences, 15, 391, https://doi.org/10.1007/s12517-022-09688-x, 2022.

Liu, Y., Meng, Q., Zhang, L., and Wu, C.: NDBSI: A Normalized Difference Bare Soil Index for Remote Sensing to Improve Bare Soil Mapping Accuracy in Urban and Rural Areas, CATENA, 214, 106 265, https://doi.org/10.1016/j.catena.2022.106265, 2022.

Liu, Y., Li, W., Liu, L., Zhou, J., Peng, B., Song, Y., Xiong, X., Yang, W., Liu, T., Liu, Z., and Li, X.: ATRNet-STAR: A Large Dataset and Benchmark Towards Remote Sensing Object Recognition in the Wild, https://arxiv.org/abs/2501.13354, 2025.

Lu, B. and He, Y.: Evaluating Empirical Regression, Machine Learning, and Radiative Transfer Modelling for Estimating Vegetation Chlorophyll Content Using Bi-Seasonal Hyperspectral Images, Remote Sens., 11, https://doi.org/10.3390/rs11171979, 2019.

Lu, B., Dao, P. D., Liu, J., He, Y., and Shang, J.: Recent Advances of Hyperspectral Imaging Technology and Applications in Agriculture, Remote Sens., 12, https://doi.org/10.3390/rs12162659, 2020.

Meng, X., Bao, Y., Liu, J., Liu, H., Zhang, X., Zhang, Y., Wang, P., Tang, H., and Kong, F.: Regional Soil Organic Carbon Prediction Model Based on a Discrete Wavelet Analysis of Hyperspectral Satellite Data, International Journal of Applied Earth Observation and Geoinformation, 89, 102 111, https://doi.org/10.1016/j.jag.2020.102111, 2020.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Miroszewski, A., Nalepa, J., Wijata, A. M., Mielczarek, J., Saux, B. L., and Sebastianelli, A.: Quo Vadis, Quantum Machine Learning?: Quantum kernel methods meet Earth observation, IEEE Geoscience and Remote Sensing Magazine, pp. 2–30, https://doi.org/10.1109/MGRS.2025.3596291, 2026.

Misara, R., Verma, D., Mishra, N., Rai, S. K., and Mishra, S.: Twenty-Two Years of Precision Agriculture: a Bibliometric Review, Precision Agriculture, 23, 2135–2158, https://doi.org/10.1007/s11119-022-09969-1, 2022.

Nalepa, J., Le Saux, B., Longépé, N., Tulczyjew, L., Myller, M., Kawulok, M., Smykala, K., and Gumiela, M.: The Hyperview Challenge: Estimating Soil Parameters from Hyperspectral Images, in: 2022 IEEE International Conference on Image Processing (ICIP), pp. 4268–4272, https://doi.org/10.1109/ICIP46576.2022.9897443, 2022.

Nalepa, J., Tulczyjew, L., Le Saux, B., Longépé, N., Ruszczak, B., Wijata, A. M., Smykala, K., Myller, M., Kawulok, M., Kuzu, R. S., Albrecht, F., Arnold, C., Alasawedah, M., Angeli, S., Nobileau, D., Ballabeni, A., Lotti, A., Locarini, A., Modenini, D., Tortora, P., and Gumiela, M.: Estimating Soil Parameters from Hyperspectral Images: A Benchmark Dataset and the Outcome of the HYPERVIEW Challenge, IEEE Geoscience and Remote Sensing Magazine, 12, 35–63, https://doi.org/10.1109/MGRS.2024.3394040, 2024.

Nejatian, A., Makian, M., Gheibi, M., and Fathollahi-Fard, A.: A Novel Viewpoint to the Green City Concept Based on Vegetation Area Changes and Contributions to Healthy Days: a Case Study of Mashhad, Iran, Environmental Science and Pollution Research, 29, 1–9, https://doi.org/10.1007/s11356-021-15552-4, 2022.

Nguyen, C. T., Chidthaisong, A., Kieu Diem, P., and Huo, L.-Z.: A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8, Land, 10, https://doi.org/10.3390/land10030231, 2021.

Nowak, B.: Precision Agriculture: Where do We Stand? A Review of the Adoption of Precision Agriculture Technologies on Field Crops Farms in Developed Countries, Agricultural Research, 10, 515–522, https://doi.org/10.1007/s40003-021-00539-x, 2021.

Pechanec, V., Mráz, A., Rozkošný, L., and Vyvlečka, P.: Usage of Airborne Hyperspectral Imaging Data for Identifying Spatial Variability of Soil Nitrogen Content, ISPRS International Journal of Geo-Information, 10, https://doi.org/10.3390/ijgi10060355, 2021.

Pingali, P. L.: Green Revolution: Impacts, Limits, and the Path Ahead, Proceedings of the National Academy of Sciences, 109, 12 302–12 308, https://doi.org/10.1073/pnas.0912953109, 2012.

Ponnusamy, V. and Natarajan, S.: Precision Agriculture Using Advanced Technology of IoT, Unmanned Aerial Vehicle, Augmented Reality, and Machine Learning, in: Smart Sensors for Industrial Internet of Things: Challenges, Solutions and Applications, edited by Gupta, D., Hugo C. de Albuquerque, V., Khanna, A., and Mehta, P. L., pp. 207–229, Springer International Publishing, Cham, ISBN 978-3-030-52624-5, https://doi.org/10.1007/978-3-030-52624-5_14, 2021.

Powers, D. M. W.: Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation, ArXiv, abs/2010.16061, https://api.semanticscholar.org/CorpusID:3770261, 2011.

Roy, A., Moradkhani, H., Mekonnen, M., Moftakhari, H., and Magliocca, N.: Towards Strategic Interventions for Global Food Security in 2050, Science of The Total Environment, 954, 176 811, https://doi.org/10.1016/j.scitotenv.2024.176811, 2024.

Ruszczak, B. and Boguszewska-Mańkowska, D.: Deep Potato – The Hyperspectral Imagery of Potato Cultivation with Reference Agronomic Measurements Dataset: Towards Potato Physiological Features Modeling, Data in Brief, 42, 108 087, https://doi.org/10.1016/j.dib.2022.108087, 2022.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Rutter, E. B., Ruiz Diaz, D., and Hargrave, L. M.: Evaluation of Mehlich-3 for Determination of Cation Exchange Capacity in Kansas Soils, Soil Science Society of America Journal, 86, 146–156, https://doi.org/10.1002/saj2.20354, 2022.

Saha, S., Saha, M., Mukherjee, K., Arabameri, A., Ngo, P. T. T., and Paul, G. C.: Predicting the Deforestation Probability Using the Binary Logistic Regression, Random Forest, Ensemble Rotational Forest, REPTree: A Case Study at the Gumani River Basin, India, Science of The Total Environment, 730, 139 197, https://doi.org/10.1016/j.scitotenv.2020.139197, 2020.

Sayão, V. M., dos Santos, N. V., de Sousa Mendes, W., Marques, K. P., Safanelli, J. L., Poppiel, R. R., and Demattê, J. A.: Land Use/Land Cover Changes and Bare Soil Surface Temperature Monitoring in Southeast Brazil, Geoderma Regional, 22, e00 313, https://doi.org/10.1016/j.geodrs.2020.e00313, 2020.

Seu, D., Longepe, N., Cioltea, G., Maidik, E., and Andrei, C.: Seeing Soil from Space: Towards Robust and Scalable Remote Soil Nutrient Analysis, https://arxiv.org/abs/2512.09576, 2025.

Sishodia, R. P., Ray, R. L., and Singh, S. K.: Applications of Remote Sensing in Precision Agriculture: A Review, Remote Sensing, 12, https://doi.org/10.3390/rs12193136, 2020.

Song, Y., Zhao, X., Su, H.-Y., Li, B., Hu, Y.-M., and Cui, X.-S.: Predicting Spatial Variations in Soil Nutrients with Hyperspectral Remote Sensing at Regional Scale, Sensors, 18, 3086, https://doi.org/10.3390/s18093086, 2018.

Tan, K., Zhu, L., and Wang, X.: A Hyperspectral Feature Selection Method for Soil Organic Matter Estimation Based on an Improved Weighted Marine Predators Algorithm, IEEE Transactions on Geoscience and Remote Sensing, 63, 1–11, https://doi.org/10.1109/TGRS.2024.3422475, 2025.

Tunçay, T., Şeref Kılıç, Dedeoğlu, M., Dengiz, O., Başkan, O., and İlhami Bayramin: Assessing Soil Fertility Index Based on Remote Sensing and GIS Techniques with Field Validation in a Semiarid Agricultural Ecosystem, Journal of Arid Environments, 190, 104 525, https://doi.org/10.1016/j.jaridenv.2021.104525, 2021.

Vlachopoulos, O., Leblon, B., Wang, J., Haddadi, A., LaRocque, A., and Patterson, G.: Delineation of Bare Soil Field Areas from Unmanned Aircraft System Imagery with the Mean Shift Unsupervised Clustering and the Random Forest Supervised Classification, Canadian Journal of Remote Sensing, 46, 489–500, https://doi.org/10.1080/07038992.2020.1763789, 2020.

Wang, X., Zhang, Y., Atkinson, P. M., and Yao, H.: Predicting Soil Organic Carbon Content in Spain by Combining Landsat TM and ALOS PALSAR Images, International Journal of Applied Earth Observation and Geoinformation, 92, 102 182, https://doi.org/10.1016/j.jag.2020.102182, 2020.

Wijata, A., Ruszczak, B., Niepala, A., Gumiela, M., Smykała, K., Longepe, N., and Nalepa, J.: HyBEAR, https://doi.org/10.5281/zenodo.17607898, 2025.

Wijata, A. M., Foulon, M.-F., Bobichon, Y., Vitulli, R., Celesti, M., Camarero, R., Di Cosimo, G., Gascon, F., Longépé, N., Nieke, J., Gumiela, M., and Nalepa, J.: Taking Artificial Intelligence Into Space Through Objective Selection of Hyperspectral Earth Observation Applications: To Bring the "Brain" Close to the "Eyes" of Satellite Missions, IEEE Geoscience and Remote Sensing Magazine, 11, 10–39, https://doi.org/10.1109/MGRS.2023.3269979, 2023.

Wijata, A. M., Lakota, T., Cwiek, M., Ruszczak, B., Gumiela, M., Tulczyjew, L., Bartoszek, A., Longépé, N., Smykala, K., and Nalepa, J.: Intuition-1: Toward In-Orbit Bare Soil Detection Using Spectral Vegetation Indices, in: IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pp. 1708–1712, https://doi.org/10.1109/IGARSS53475.2024.10640702, 2024a.

https://doi.org/10.5194/essd-2026-64
Preprint. Discussion started: 22 April 2026

Wijata, A. M., Miroszewski, A., Saux, B. L., Longépé, N., Ruszczak, B., and Nalepa, J.: Detection of Bare Soil in Hyperspectral Images Using Quantum-Kernel Support Vector Machines, in: IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, pp. 817–822, https://doi.org/10.1109/IGARSS53475.2024.10641442, 2024b.

Xu, J., Cui, Y., Zhang, S., and Zhang, M.: The Evolution of Precision Agriculture and Food Safety: a Bibliometric Study, Frontiers in Sustainable Food Systems, 8, 1475 602, https://doi.org/10.3389/fsufs.2024.1475602, 2024.

Yan, Y., Yang, J., Li, B., Qin, C., Ji, W., Xu, Y., and Huang, Y.: High-Resolution Mapping of Soil Organic Matter at the Field Scale Using UAV Hyperspectral Images with a Small Calibration Dataset, Remote Sensing, 15, https://doi.org/10.3390/rs15051433, 2023.

Yue, J., Zhou, C., Guo, W., Feng, H., and Xu, K.: Estimation of Winter-Wheat Above-Ground Biomass using the Wavelet Analysis of Unmanned Aerial Vehicle-Based Digital Images and Hyperspectral Crop Canopy Images, Int. J. Remote Sens., 42, 1602–1622, https://doi.org/10.1080/01431161.2020.1826057, 2021.

Zhang, Y., Xia, C., Zhang, X., Cheng, X., Feng, G., Wang, Y., and Gao, Q.: Estimating the Maize Biomass by Crop Height and Narrowband Vegetation Indices Derived from UAV-based Hyperspectral Images, Ecological Indicators, 129, 107 985, https://doi.org/10.1016/j.ecolind.2021.107985, 2021.

Zhang, Y., Liang, L., Li, J., Plaza, A., Kang, X., Mao, J., and Wang, Y.: Structural and Textural-Aware Feature Extraction for Hyperspectral Image Classification, IEEE Geoscience and Remote Sensing Letters, 21, 1–5, https://doi.org/10.1109/LGRS.2024.3357201, 2024.

Zhao, J., Du, D., Chen, L., Liang, X., Chen, H., and Jin, Y.: HA-Net for Bare Soil Extraction Using Optical Remote Sensing Images, Remote Sensing, 16, https://doi.org/10.3390/rs16163088, 2024.

Zhu, C., Ding, J., Zhang, Z., and Wang, Z.: Exploring the Potential of UAV Hyperspectral Image for Estimating Soil Salinity: Effects of Optimal Band Combination Algorithm and Random Forest, Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 279, 121 416, https://doi.org/10.1016/j.saa.2022.121416, 2022.