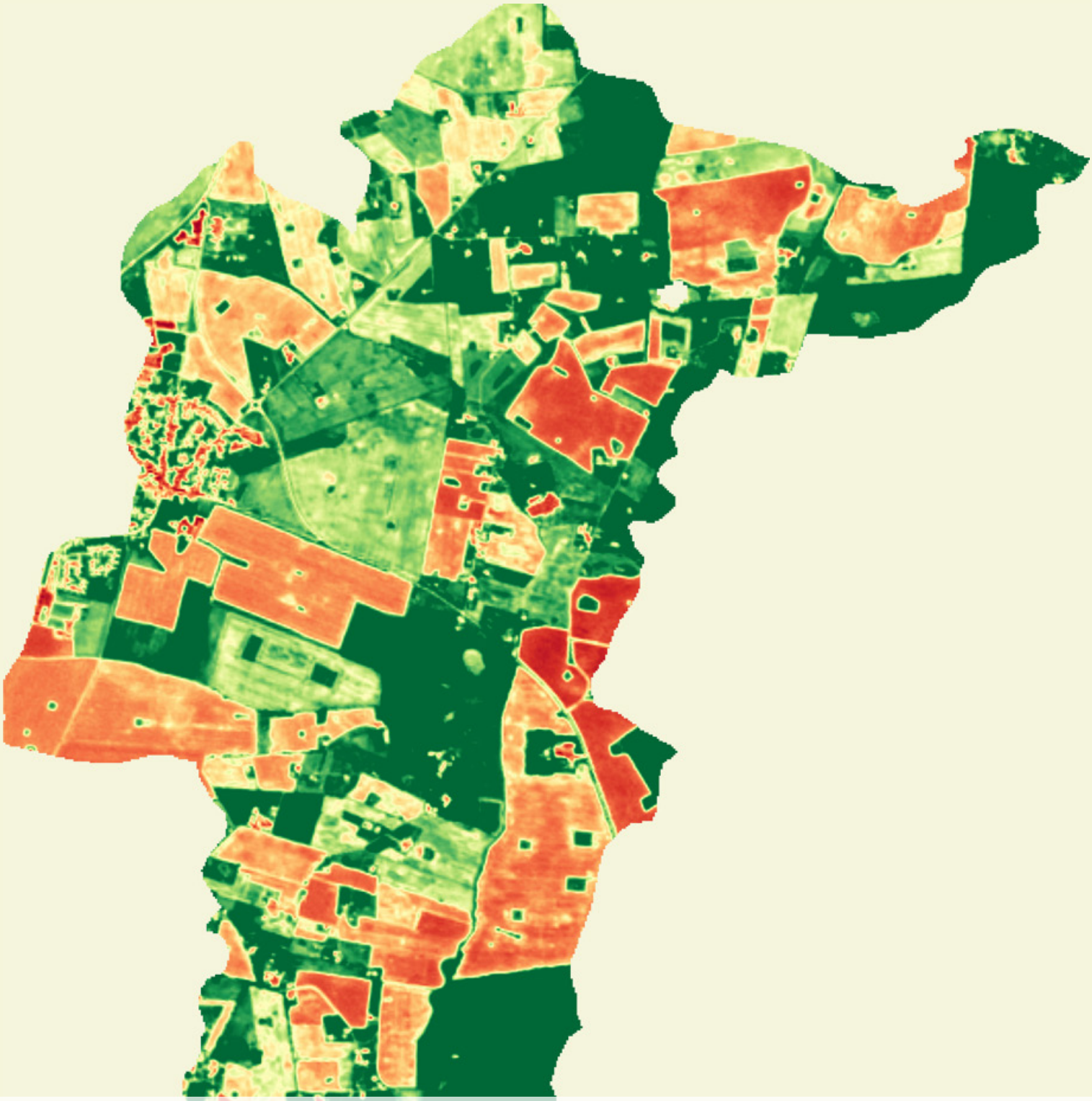


CLASSIFICATION OF DANISH AUTUMN VEGETATION USING SATELLITE DATA

Data from the Danish Agricultural Monitoring Programme

Technical Report from DCE – Danish Centre for Environment and Energy No. 370 2026

AARHUS
UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

CLASSIFICATION OF DANISH AUTUMN VEGETATION USING SATELLITE DATA

Data from the Danish Agricultural Monitoring Programme

Technical Report from DCE – Danish Centre for Environment and Energy No. 370 2026

Anastasia Katharina Kratschmer
Rasmus Rumph Frederiksen
Mette Thorsen
Gitte Blicher-Mathiesen

Aarhus University, Department of Ecoscience

AARHUS
UNIVERSITY
DCE – DANISH CENTRE FOR ENVIRONMENT AND ENERGY

Data sheet

Series title and no.:	Technical Report from DCE – Danish Centre for Environment and Energy No. 370
Category:	Scientific advisory report
Title:	Classification of Danish autumn vegetation using satellite data
Subtitle:	Data from the Danish Agricultural Monitoring Programme
Authors:	Anastasia Katharina Kratschmer, Rasmus Rumph Frederiksen, Mette Thorsen, Gitte Blicher-Mathiesen
Institution:	Aarhus University, Department of Ecoscience
Publisher:	Aarhus University, DCE – Danish Centre for Environment and Energy ©
URL:	https://dce.au.dk/en
Year of publication:	February 2026
Editing completed:	20 February 2026
Referee:	Robert Ladwig
Quality assurance, DCE:	Signe Jung-Madsen
Linguistic QA:	Anne-Mette Poulsen
External comments:	No comments received
Financial support:	Promilleafgiftsfonden
Please cite as:	Kratschmer, A. K., Frederiksen, R.R., Thorsen, M., Blicher-Mathiesen, M. 2026. Classification of Danish autumn vegetation using satellite data. Data from the Danish Agricultural Monitoring Programme. Aarhus University, DCE – Danish Centre for Environment and Energy, 34 pp. Technical Report No. 370

Abstract: This report presents a deep learning model to classify autumn vegetation in the six Danish LOOP-catchments over eight years (2017–2024) using optical and radar satellite data. The model also predicts main crop types to support vegetation classification. It was trained on 5,477 field–year combinations from 1,220 fields. Autumn vegetation was correctly classified into seven categories 67.8% of the time. Main crops could be classified into ten categories at an accuracy of 86.2%. The performance difference reflects challenges related to morphological overlaps among vegetation types in autumn, subtle autumn growth patterns, and limited optical data due to cloud cover.

Keywords:	Landovervågningsoplande, remote sensing, agriculture, catch crops, cover crops, vegetation, satellite data, classification, nitrate leaching, machine learning, deep learning
Drawings:	Anastasia Katharina Kratschmer,
Front page photo:	Sentinel Hub collection Sentinel-2 L2A
ISBN:	978-87-7648-034-9
ISSN (electronic):	2244-999X
Number of pages:	34

Contents

Preface

The following note is part of the project "Målrettede efterafgrøders effekt på kvælstofudledning" (English: "The effect of targeted catch crops on nitrogen leaching"), which aims to quantify whether there is a significant difference in nitrogen leaching to streams in catchments that are otherwise comparable but have different quantities of targeted catch crops. Targeted catch crops are catch crops that are required in addition to standard catch crop requirements in catchments that drain to coastal marine areas particularly vulnerable to high nitrogen loads. The project is funded by Promilleafgiftsfonden for Landbrug and is a collaboration between Ecoscience at Aarhus University and SEGES Innovation P/S and runs through 2025 and 2026.

This note outlines the work to develop a model that can classify the type of autumn vegetation present in the fields of catchment based on satellite data. The register data forming the basis of the model is derived from the six Landovervågningsoplande (English: Land Surveillance Catchments) in Denmark, which have detailed records of vegetation cover types. The model utilizes data from the two satellite constellations Sentinel 1 and Sentinel 2 from the years 2017-2024. The model will be used to qualify information in register data so that more precise information about the vegetation composition of Danish agricultural catchments in the autumn period can be obtained and included in the analysis.

Sammenfatning

Formålet med arbejdet, der beskrives i denne rapport, er at klassificere efterårsvegetation i seks oplande i en otteårs periode ved hjælp af remote-sensing-data og Deep Learning-teknikker.

Denne rapport beskriver udviklingen og testningen af en model til at klassificere forskellige afgrøders vækst om efteråret ud fra optiske og radarsatellitdata for danske marker. Modellen klassificerer desuden forskellige typer af hovedafgrøder for at understøtte efterårsmodellens konfiguration. Modellen er udviklet på data fra 5.477 mark/år-kombinationer fra 1.220 forskellige marker, der er lokaliseret i et af de seks danske LOOP-oplande (Landovervågningsoplande), og hvor data dækker en otteårig periode fra 2017 til 2024.

For LOOP-områder, der ikke var set af modellen under træningen, blev efterårsvegetationen klassificeret korrekt i 67,8 % af alle tilfælde på tværs af alle syv efterårsvegetationsgrupper. Den bedst identificerbare gruppe af efterårsbevoksning var vinterkorn, der blev identificeret korrekt 82 % af tiden. Gruppen, der var sværest at identificere, var enkimbladede efterafgrøder og blandinger, der blev klassificeret korrekt i 52 % af tilfældene. Hovedafgrøden blev klassificeret korrekt i 86,2% af alle tilfælde på tværs af alle ti hovedafgrødegrupper. Her var vinterraps den lettest identificerbare gruppe med en klassifikationssucces på 95 %. Den gruppe, der blev identificeret med den laveste succesrate, var bælgplanter med 26 %.

For marker fra områder, der allerede var blevet set af modellen under træningen, blev efterårsvegetation fra endnu usete marker klassificeret korrekt i 72,2% af tilfældene. Her var succesraten for korrekt klassifikation af efterårsbevoksningen højest for vinterkorn (88 %) og lavest for brak/tom (ingen registreret vegetation) (51 %). For hovedafgrøden var succesraten 90,6%.

Forskellen mellem præstation, afhængigt af om modellen havde set marker fra det pågældende LOOP-område under træningen eller ej, fremhæver udfordringen ved at generalisere i internt forskellige vegetationskategorier på tværs af geografisk adskilte områder. Forskellen i præstation mellem klassificering af efterårsvegetation og klassificering af hovedafgrøder kan skyldes, at der generelt er flere overlap i vegetationsmorfologi mellem efterårsvegetationsgrupperne, f.eks. vil brak med græs vegetationsmæssigt ligne en græsbaseret efterafgrøde i efterårsperioden. Mindre tydelige vækstmønstre opfanges af modellen i løbet af efterår og vinter, hvor tilgængeligheden af optiske satellitdata er mindre på grund af skydække. Disse begrænsninger påvirker succesraten, hvormed efterårets forskellige afgrøder og bevoksninger kan klassificeres ud fra satellitdata alene.

Summary

The purpose of the work described in this report is to classify autumn vegetation within six catchments for an eight-year period using remote sensing data and deep learning techniques.

This report describes the development and testing of a model for autumn vegetation category prediction using optical and radar satellite data for Danish agricultural fields. The model also predicts main crop types, which supports the autumn vegetation classification task.

The model was developed based on data from 5,477 field–year combinations from 1,220 different field-IDs in the Danish LOOP catchment areas (Landovervågningsoplande) from an eight-year period from 2017 to 2024.

For fields in LOOP catchments unseen during training, the autumn vegetation was classified correctly into the seven autumn vegetation groups 67,8% of the time. The most easily identifiable group of autumn vegetation was "Winter cereals", which was correctly identified 82% of the time. The group that was most difficult to identify was "Catch crop (monocot/mix/unknown)" with 52%. The main crop was classified correctly into the ten main crop groups 86,2% of the time. Here, "Winter oilseed rape" was the most easily identifiable group with a classification success of 95%. The group that was least successfully identified was "Legumes" with a success of 26%.

For areas already seen during training, autumn vegetation from unseen field tiles was classified correctly 72.2% of the time. Here, the success rate for correct classification of autumn growth was highest for "Winter cereals" (88%) and lowest for "Fallow or none" (no registered vegetation) (51%). The main crop was classified correctly 90.6% of the time.

The gap in performances depending on whether the model has already seen fields from the LOOP catchment in question highlights the challenge of generalizing across areas with internally diverse vegetation categories. The difference in performance between the autumn vegetation classification task and the main crop classification task may be caused by greater morphology overlaps among the autumn vegetation groups (for instance, a grass set-aside can be vegetationally highly similar to a grass-based catch crop in the autumn period), less distinct growth patterns during autumn and winter, and the much rarer availability of optical satellite data (due to the higher frequency of cloud cover in the autumn and winter periods). These constraints may create a natural ceiling for how well autumn vegetation can be classified from satellite data alone.

1 Introduction

1.1 Nitrate leaching and catch crops

Eutrophication reduces water quality and changes the ecological structures in aquatic ecosystems (Dodds and Welch 2000). Agricultural fertilization is a considerable source of nitrogen to streams, coastal, and marine environments. In Denmark, approximately 70% of the nitrogen load to coastal waters is estimated to originate from agriculture (Det Økonomiske Råd 2017).

Nitrate leaching from agriculture is strongly dependent on agricultural practices and management (Kumar et al. 2025; Li et al. 2023; Zhao et al. 2022). Cultivation of catch crops between main crop cycles is one such practice: Catch crops are non-harvested crops grown during autumn and winter, with the primary purpose of mitigating nitrate leaching (Kumar et al. 2025). Estimates of the reduction in nitrate leaching when catch crops are present compared to bare fields after the main crop has been harvested vary. A global review from 2022 estimated that catch crops can reduce nitrate leaching by as much as 68 % (Nouri et al. 2022). A review focusing on Nordic country conditions (Valkama et al. 2015) found highly variable effects on nitrate leaching of under-sown catch crops (catch crops sown with the main crop). The reported effects ranged from a 95% reduction to a 5% increase for non-legume catch crops such as ryegrass. The average reduction in nitrate leaching was 50% (Valkama et al. 2015). In practice, catch crops often do not simply replace bare fields, and their environmental impact, nutrient uptake, and decomposition dynamics depend on their type (Nouri et al. 2022), C:N-content ratio (Wanic et al. 2019), growth period (Kumar et al. 2025), and the vegetation they replace (Schwartzkopff et al. 2026).

To reduce nitrate leaching from the root zone and the subsequent nitrogen load to coastal areas, Denmark has specific requirements stipulating that catch crops in autumn must cover approximately 11% and 15% of the agricultural area in the autumn, depending on the farmer's application level of organic (biologically derived) fertilizer such as manure. Farmers applying less than 80 kg N ha^{-1} yr^{-1} from organic fertilizers must comply with the 11% requirement, whereas higher application rates are subject to the 15% threshold. Furthermore, some catchments that drain to nitrate-sensitive areas, such as Natura 2000 areas, have an additional requirement for establishing extra catch crops. In these areas, so-called "livestock catch crops" are intended to further compensate for the use of organic fertilizers (SGAV 2025). In areas especially vulnerable to high nitrogen loads, the planting of further so-called "targeted catch crops" (maximum 44,2%) is required (Landbrugsstyrelsen 2025).

An estimated 34-78 kg N ha^{-1} are each year leached from the root zone to streams and groundwater in five of the six Landovervågningsoplande (Land Surveillance Catchments) in Denmark, known as the LOOP catchments (Blicher-Mathiesen et al. 2024). These catchments are unique due to their long-term, high-resolution records of crop composition, agricultural management practices, and stream water chemistry. In addition, five of the six catchments include measurements of soil pore water and groundwater chemistry dating back to the 1990s. These datasets provide valuable decades-long information about Danish agricultural practices and their environmental effects, and they

are used for a wide range of environmental modeling tasks (Blicher-Mathiesen et al. 2024).

In Denmark, the level of detail in records of catch crop cultivation varies between areas: In the LOOP catchments, detailed field-level records of autumn vegetation are available for many years, but in most other areas before 2019, records were only available at the farm level. Even in the current registration system, only the presence or absence of a catch crop is recorded. It is, therefore, important to be able to obtain information on catch crop presence, type, and growth independently of any existing registration data.

1.2 Satellite data in vegetation monitoring and modelling

Purpose of using satellite data in vegetation monitoring

Satellite data provides an independent source of information on field-level vegetation. Sentinel-1 and Sentinel-2 offer widely accessible radar and optical data with 10–60 m resolution and are used for diverse vegetation-modeling tasks, often for main crops (Tufail et al. 2025; Campos-Taberner et al. 2019) or even trees (Immitzer et al. 2016). The interest in modelling catch crop growth has increased, both for biomass estimation (Goffart et al. 2021; Bendini et al. 2024) and type classification, especially in Northern European countries like Belgium and Germany (Vanpoucke et al. 2024; Jensch et al. 2025; Selvaraj et al. 2025).

Modeling autumn vegetation from satellite data

In a Belgian study from 2024 (Vanpoucke et al. 2024), agricultural autumn vegetation was classified into seven groups using a two-step hierarchical model based on Sentinel 2 (optical) satellite data. This was done using a 1D Convolutional Neural Network on field-level mean values derived from satellite data with thrice-a-month temporal resolution. The study reported F1-values (a measure of classification accuracy with values between 0 and 1, higher values being better) of 0.87-0.89 for the classification task. Spatial transferability was not tested.

In a German study from 2025 (Selvaraj et al. 2025), Sentinel 1 (radar) data was utilized as input to another two-step hierarchical model comprising an initial binary classification into "catch crop" or "other", followed by sub-category classification. The model type used was Random Forest (a decision tree-based model) with manually extracted temporal data variables. Very variable F1-values for the classification of catch crop sub-groups were obtained, ranging from around 0.5 for mustard to at least 0.95 for the mixture group in one of the test years. However, spatial transfer testing of the binary top-level classification yielded F1-values between 0.67 and 0.835. Spatial transferability for the sub-category classification was not tested.

In another German study from 2025 (Jensch et al. 2025), both Sentinel 1 (radar) and Sentinel 2 (optical) data were used as input into an ensemble of Random Forest models to categorize fields as "catch crop" or "other crop". At field level, the classification accuracy was 93%, with F1-values for catch crops of 0.95 and 0.89 for "other crops". Neither spatial transferability nor sub-class classification were explored. Integrating optical and radar data only slightly improved performance compared with using optical data alone.

1.3 Deep learning models

Deep learning models are sophisticated machine learning models that, given sufficient data, approximate complex functions present in the data (Abiodun et al. 2018). In deep learning models, several layers of linear and nonlinear transformations are applied to the data to yield representations of the data that are useful for fulfilling the task in question, such as regression or classification (Aggarwal 2023, ch. 4). Through repeated rounds of training, deep learning models learn how to parameterize these transformations (Abiodun et al., 2018). This allows the modeler to fit a model without needing in-depth hypotheses and knowledge about which data features (raw or derived) are relevant for the modeling task (Abiodun et al,. 2018). While deep learning-models are extremely powerful, in most implementations the modeler trades transparency for gains in performance, when choosing deep learning models over other, simpler, machine learning models (Assis et al. 2024).

Deep learning models are used for tasks in areas as diverse as medicine, agriculture, and business (Abiodun et al. 2018). They can be composed of a wide variety of building blocks. Relevant in this report are the three concepts: Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Gated Recurrent Units (GRU).

CNNs utilize filters that slide across the input, often an image. Using several iterations of many filters, the model builds a representation of the image's features (LeCun et al. 2015), which can be used for tasks such as classification or regression.

RNNs can process sequential input data, such as time series. They process the input from each sequence step sequentially but pass information between the sequence steps, so that elements of information from a previous step are available and can be considered at a later step (LeCun et al. 2015).

GRU networks are a type of RNN. They operate by passing information between sequence steps and include mechanisms that determine how much information from previous sequence steps should be retained and combined with information in the subsequent steps. Standard GRU processes the data sequence in a forward direction, whereas a bidirectional GRU processes the data both forwards and backwards (Perumal et al. 2024).

1.4 Purpose and aim

The purpose of this work is to classify autumn vegetation within six catchments for an eight-year period using remote sensing data and deep learning techniques. The six selected catchments are the Danish LOOP catchments with unique crop and management records. The utilized remote sensing data is radar data and optical data from Sentinel-1 and 2 satellites, and the deep learning techniques applied are CNNs and bidirectional GRUs.

2 Methods

2.1 Reference data

The six LOOP catchments (Landovervågningsoplande) are Danish agricultural catchments for which detailed crop, vegetation, and management records have been maintained since 1990 in five catchments and since 1998 in the sixth. For 31 fields across five of the catchments, chemical analysis of soil water has been performed weekly when sufficient soil water was available. For all six catchments, chemical analysis of stream water has been performed biweekly, providing long and detailed timelines of their chemical status (Blicher-Mathiesen et al. 2024). In our study, we used the detailed records of crops, catch crops, and other registered vegetation types in these six areas.

Danish agriculture features a wide range of crop, catch crop, and vegetation types, which we grouped into broader categories. Main crops were divided into 12 categories of which 10 were included in the project (tab. 2.1), and autumn vegetation types were separated into nine categories of which seven were included in the project (tab. 2.2).

Table 2.1. Main crop vegetation categories

Beet	Sugar beets and fodder beets, a very uniform category.	Spring cereals	Including several cereal species with varying cultivation practices: oats, spring wheat, spring barley etc. With and without under-sown grass. With and without co-cropping with legumes. Fairly homogeneous.
Fallow or none	Used for explicitly fallow areas, land taken out of production, or fields with no main crop registered in that year, very heterogeneous category.	Trees and bushes	Including forest, berry shrubs, orchards, and energy-crop willow, a very heterogeneous group.
Grass/forage	Including grass with or without clover in varying proportions, whether grazed or cut for hay and/or silage. A fairly heterogeneous category.	Winter cereals	Comprising different cereal species sown in autumn, such as winter wheat, winter barley and winter rye. Homogenous group.
Legumes	Encompassing crops such as peas, lupin, faba beans, and others. A very heterogenous category.	Winter oilseed rape	Autumn-sown oilseed rape. Very homogeneous group.
Maize	Covering different cultivation practices and end uses, but generally homogeneous.	Spring oilseed rape	Spring-sown oilseed rape. Very homogeneous group. Not included due to low group size.
Potatoes	Potatoes for different end uses. A homogeneous but small category.	Other	Mainly vegetables and flower production, but also wetlands and animal grazing. Very heterogeneous. Not included due to high heterogeneity and low sample size.

Table 2.2. Autumn vegetation categories

Catch crop (crucifer alone or mix)	Catch crops that are comprised fully or partly of crucifers such as oil radish and mustard.	Winter cereals	Comprising different cereal species sown in autumn, very homogenous group.
Catch crop (monocots, monocot mixes, and unknown species)	Catch crops that are not known to comprise crucifers, neither fully or partially. In practice, grasses, cereals (monocots), alone or in mixes with non-crucifers such as Phacelia. Also includes further unspecified catch crops.	Winter oilseed rape	Autumn-sown oilseed rape. Very homogeneous group.
Fallow or none	Used for explicitly fallow areas, land taken out of production, or fields with no main crop registered in that year, very heterogeneous category.	Other catch crops	Unspecified nitrogen-fixating catchcrop. Not included due to small category size.
Grass/forage	Including grass with or without clover in varying proportions, whether grazed or cut for hay and/or silage. A decently heterogeneous category.	Spring cereals	Spring cereals registered as autumn vegetation. Not included due to small category size.
Trees and bushes	Including forest, berry shrubs, orchards, and energy-crop willow, a very heterogeneous group.		

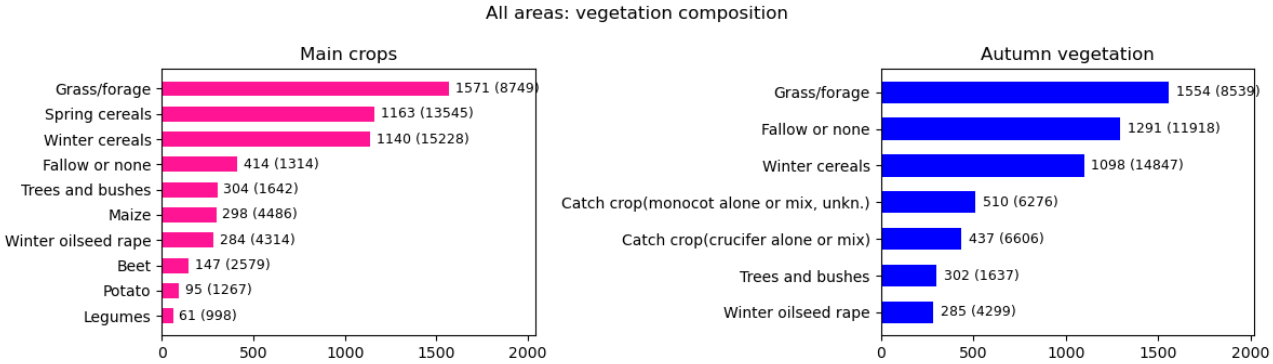


Figure 2.1. Number of distinct fields belonging to each main crop category and autumn vegetation category in the full dataset. The number in parenthesis is the number of distinct 80*80 meter field tiles assigned to each category.

In this project, fields were divided into non-overlapping square tiles with a side length of 80 m. For main crop, the most common type across all areas was ”Grass/forage” at field scale and ”Winter cereals” at tile scale, reflecting differences in field size for the main crop categories. Another large category was ”Spring cereals”. For autumn vegetation, the most common type across all areas was ”Grass/forage” at field scale and ”Winter cereals”) at tile scale, with “Fallow or none” also representing a substantial share (fig. 2.1). Area-wise vegetation compositions can be found in the appendix (fig. 7.1- fig. 7.6). A description of the procedure for dividing fields into 80 m x 80 m tiles can be found in section 2.3.

Figure 2.2. Overview of the LOOP catchments' locations. •: Annual interviews+ measurements in soil water, groundwater, drain water and streams Δ: Only annual interviews and measurements in streams. Adapted from Blicher-Mathiesen et al. (2024).

2.2 Satellite data

The satellite data was obtained by requesting data from the Sentinel Hub API for the six LOOP catchments for eight and a half years from January 1, 2017 to June 30, 2025, using the Python package "Sentinelhub" v. 3.11.3.

For Sentinel-2 (optical data) ('Sentinel-2 L2A' 2025, 2), 13 bands were downloaded, of which 12 were reflectance data from wavelengths 0.490 µm (blue) to 2.190 µm (shortwave infrared). The last band was a scene classification layer (SCL), which was later used in removing pixels affected by clouds, snow, or cloud shadows. For Sentinel-1 (radar data) (ascending mode) ('Sentinel-1 GRD' 2025, 1), two bands were downloaded: VV and VH polarizations. These downloaded bands (aside from the SCL) were integrated into a 14-channel data stack covering each LOOP catchment.

In addition, five vegetation indices were calculated pixel-wise based on the raw data and added to the data stack, resulting in a 19-channel stack spanning each area geographically.

From radar data, one index was calculated: RVI (Radar Vegetation Index) (Holtgrave et al. 2020).

From optical data, four indices were calculated: NDVI (Normalized Difference Vegetation Index); NDRE (Normalized Difference Red Edge); PSRI (Plant Senescence Reflectance Index); and NDMI (Normalized Difference Moisture Index) (IDB-NDVI 2025; IDB-NDRE 2025; IDB-PSRI 2025; IDB-NDMI 2025).

Table 2.3. Vegetation indices calculated pixelwise in addition to the raw bands

Index	Formula	Used bands	Good for
RVI (Radar Vegetation Index)	$\frac{4 * VH}{VH + VV}$	VV (radar) VH (radar)	Plant structure
NDVI (Normalized Difference Vegetation Index)	$\frac{NIR - Red}{NIR + Red}$	NIR: Band 08 (optical) Red: Band 04 (optical)	Plant health and density
NDRE (Normalized Difference Red Edge)	$\frac{NIR - Red\ edge_5}{NIR + Red\ edge_5}$	NIR: Band 08 (optical) $Red\ Edge_5$: Band 05 (optical)	Plant health and density, sensitive to chlorophyll
PSRI (Plant Senescence Reflectance Index)	$\frac{Red - Blue}{Red\ edge_6}$	Red: Band 04 (optical) Blue: Band 02 (optical) Red Edge6: Band 06 (optical)	Plant senescence
NDMI (Normalized Difference Moisture Index)	$\frac{NIR - SWIR_{11}}{NIR + SWIR_{11}}$	NIR: Band 08 (optical) $SWIR_{11}$: Band 11 (optical)	Vegetation water content and water stress

The full satellite data stacks were spatially clipped and assigned to individual agricultural fields based on yearly field polygon layers from Landbrugsstyrelsen **(LandbrugsGIS/Landbrugsstyrelsen 2025)**.

Each observation was made up of data from March 15th in the year associated with the field polygon to March 14th in the following year to reflect the period from one sowing of spring-sown main crops to the next.

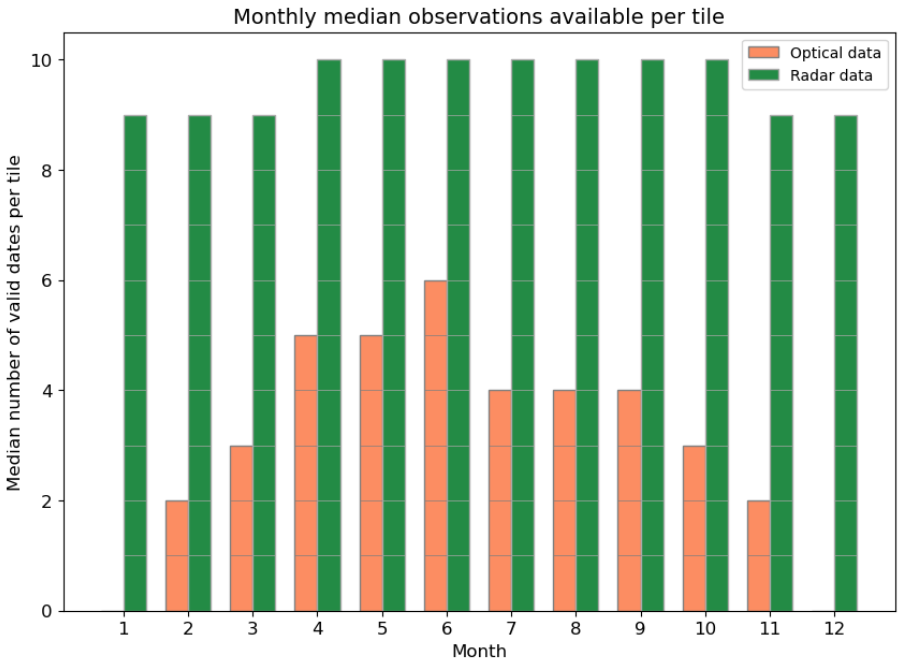


Figure 2.3. Median number of available data points per field tile after filtering out data points marked as affected by clouds. In some cases, however, not all remaining points were included in the final analysis, as a subset were identified as outliers.

The optical data acquisitions are significantly affected by clouds, leading to substantial reduction in data availability in winter compared to spring and summer. It is common for a field to have no valid optical data acquisitions in December and January, and only a few in November and February. In contrast, radar data acquisitions are much more evenly distributed throughout the year (fig. 2.5).

2.3 Preprocessing of the data

Field level data was split into 80 m x 80 m non-overlapping squares, which were treated as individual samples, assuming a sufficient degree of independence. This increased the amount of training data, captured a broader range of growth patterns within each crop type (as sample means become more variable), and provided sub-fields of equal size, as required by the model's fixed input size expectations.

After filtering out fields with main crops or autumn vegetation types of less than 500 tiles present in the dataset, 54,122 tiles were left, originating from 5,477 different field/year-combinations. The data was downsampled to be composed of the channel mean in a half-month time window. For each channel (time series), outliers were removed by filtering out observations more than two standard deviations in distance from the median away for the series, then linear interpolation to daily scale was performed, after which the half-monthly median of the daily values was calculated to represent the period. This ensured that each half-monthly period had exactly one value per band attributed to it, despite that the number of measurements obtained in the period ranged from zero to several.

The data was subsequently divided into two distinct representations of each tile (fig. 2.4):

- A difference-based $24 \cdot 24$ pixel 19-channel was constructed (one channel per raw satellite band and one per derived vegetation index). This representation consists of the signed (as opposed to absolute) differences between half-monthly means. The model treats this matrix as if it were a standard picture (just with 19 channels instead of the three in a classic RGB image) and attempts to extract meaning from the patterns in these different values.
- A half-monthly mean $1 \cdot 24$ pixel 19-channel representation of the absolute raw satellite data and vegetation index values.

Values in each representation were normalized by use of band-wise 5^{th} and 95^{th} percentiles computed from the absolute value training dataset according to the training-validation-testing-split of the training run in question. Values were subsequently scaled to the range 0 to 1 to support model convergence.

The difference matrix is an attempt to represent the temporal patterns of the data in a manner that the model can interpret as a textured surface - something that CNN models are effective at learning from (Geirhos et al. 2022). The absolute (but normalized) values were included as input to the model to provide a purer representation of the raw data.

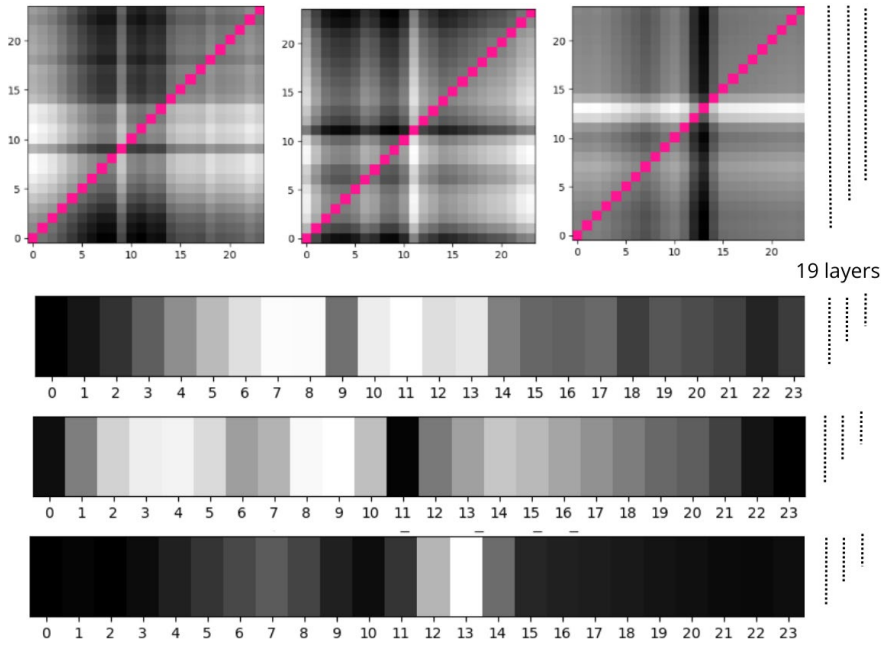


Figure 2.4. The structure of the data input to the model. In the upper part of the figure, three examples of the relative difference matrices are given. The lower part illustrates three examples of the absolute value matrices. Here, the NDVI band is shown in its unnormalized form. The dots to the side of the data signify that the data is layered in 19 layers, composed of raw satellite bands and derived vegetation indices. The pink line signifies flat zero values, as the difference of a half-month value to itself is equal to zero.

2.4 Model and training

The model combines two deep learning techniques (fig. 7.11).

The absolute values matrix, which contains the raw spectral and vegetation index time series, is processed via a bidirectional GRU with 128 hidden states (pieces of information) passed between time-steps in both forward and backward directions. The difference matrix, which encodes the differences between time steps in the raw spectral and vegetation index series, is processed using a CNN with four hidden layers with, respectively, 32 5*5-pixel filters, 64 3*3-pixel filters, 128 3*3-pixel filters, and 128 5*5-pixel filters. CNNs are highly effective at recognizing texture-like patterns (Geirhos et al. 2022), which the difference matrix emulates. The latent representations obtained from the GRU and CNN branches are combined and passed through fully connected (linear) layers to produce class-specific logits, which are then converted into class probabilities for the two classification tasks (main crop and autumn vegetation type classification). More information on the model architecture can be found in the appendix, including a sketch (fig. 7.11).

The project was implemented in Python 3.12.7. The data was stored in xarray-files (Xarray 2025.1.2), and the model itself was implemented with PyTorch (2.6.0+cpu).

The model's most important task is to classify the autumn vegetation into categories. A sub-task was to classify the main crop into main crop categories to:

- enable the model to learn more general representations of vegetation dynamics.
- obtain an indication of whether the model was working correctly for detecting vegetation types, as classifying main crops was assumed to be markedly easier for a multitude of reasons: the longer growth periods, the plants reaching higher developmental stages during the period, lower presence of clouds in the growth period, among others.

As there were substantial imbalances in vegetation category frequency, observations were randomly sampled during training using weights inversely proportional to their class frequencies in the training dataset for the training run. Since each observation belongs to both a main crop class and an autumn vegetation class, the weights per sample were calculated as the sum of the inverse training set frequencies of its two class labels.

The loss function, which quantifies how incorrect the model's predictions are, was cross-entropy loss. Cross-entropy describes how well the predictions align with the target classes. Cross-entropy loss for a sample can be calculated as:

$$l_n = -\log\left(\frac{e^{x_n,y_n}}{\sum_{c=1}^{C} e^{x_{n,c}}}\right)$$

where n is the sample, c is the class, and $x_n y_n$ is the raw score for the class produced by the model (PyTorch 29 2025). This formulation amounts to comparing the score assigned to the target category with the sum of scores assigned to all categories. The loss per sample is then averaged over all observations in the dataset. In our case, separate losses were calculated for the main crop and autumn vegetation classification tasks.

The model was allowed to train for 120 epochs, each consisting of 50 batches with 32 observations, totaling 1600 observations per epoch. Early stopping was employed – if no overall improvement in validation loss was observed for 10 successive epochs, training was stopped. Throughout the training, new best-performing models (evaluated based on the validation loss for both tasks) were saved. The initial learning rate was set to 0.005 and was reduced by a factor of 0.55 each time three successive epochs did not produce an overall improvement in the loss of the two tasks. This was done to let the model approach potential (local or global) minima in the loss landscape. Every 15 epochs, the training switched between updating (training) the GRU part of the model and updating (training) the CNN part, while keeping the other part frozen. The fully connected output layers were never frozen and could always be updated. This strategy allows each part of the model to be trained without being dominated by the other.

Testing spatial transferability

When training a model on data from specific geographical areas in order to apply it to other geographical areas, it is crucial to assess how well the model works on data from geographical areas not included in training. Because only six LOOP catchments are available, we adopted an approach in which the same areas were alternatingly used for both training and testing.

The model was initialized and trained independently six times. Each time, a different LOOP catchment was held out and not used in either training or validation but only as test data. The remaining field tiles were split into a training set, and a validation set such that no field tile originating from the same field–year combination was present in both datasets, ensuring a higher degree of independence. Accordingly, the proportion of test data was the proportion of observations originating from the held-out LOOP catchment. The proportion of validation data was $\left(1 - prop_{test\_area}\right) \cdot 0.10$, and, consequently, the proportion of training data was $(1 - prop_{test\ area}) \cdot 0.9$.

In this report, a test field tile is considered to originate from an *unseen area* if no field tiles from its geographical catchment were included in the model's training or validation data. Conversely, a test field tile is considered to originate from a *seen area* if other field tiles from the same catchment were included in the training and/or validation data.

All field tiles from fields from which a test tile originates were excluded from the model's training and validation data. This means that in the first case, test tiles come from *unseen fields in unseen catchment areas*. For the latter case, test tiles come from *seen catchment areas, but unseen fields*.

Testing performance in geographical areas used for training

It is also informative to evaluate how well the model performs when classifying vegetation in fields located in geographical areas that were included in the training data. This allows the potential performance gap between fields from seen geographical areas and those from new, unseen areas to be quantified, i.e. a quantification of the geographical generalizability of the model (or lack thereof).

The model was trained once on randomly selected field tiles from all LOOP catchments, ensuring that field tiles originating from the same fields were only present in data subsets. Test, validation, and training proportions were 0.12, 0.08, and 0.8, respectively.

Baseline model

As a baseline comparison model, a training-validation-test split was performed, and the autumn vegetation and main crop category frequencies in the training data were calculated and used as weights for a model that randomly guesses vegetation category, using training data frequencies as probabilities.

Performance metrics

Model performance was evaluated using two metrics: confusion matrices and F1-values. Confusion matrices (e.g. fig. 3.1) show how often categories in a classification task are predicted correctly or incorrectly, and which categories they are misclassified as. The confusion matrices in section 3 can be read as follows: Each row represents the true category, and each column represents the predicted category. Each cell in the matrix therefore corresponds to a specific combination of true and predicted categories.

The matrices in section 3 are row-normalized. This means that the values in each row sum to 1 (aside from minor rounding effects), and the diagonal values represent the proportions of observations belonging to the true category (indicated by the row) that were correctly classified. The off-diagonal values show the proportions of observations that were misclassified into other categories.

For classification tasks, a measure of classification success often used is "F1". These metric balances recall (the ratio of true positives to the sum of true positives and false negatives) and precision (the ratio of true positives to true positives and false positives). F1-values range between 0 and 1, where higher values indicate better performance. F1-values are closely related to the confusion matrix and can be viewed as taking into account both recall (as shown by the diagonal values in a row-normalized matrix) and precision (which would correspond to the diagonal values in a column-normalized matrix).

3 Results

3.1 Classification of unseen areas

Classification of autumn vegetation types solely from satellite data in LOOP catchments not seen by the model during training (fig 3.1, fig. 7.7) was partly possible. The two most difficult categories to identify correctly were "Fallow or none" and "Catch crop (monocot alone or mix, or unkn.)", with F1-values of 0.55 and 0.56, respectively. Notably, field tiles from all autumn vegetation categories were frequently misclassified as "Fallow or none", indicating potential overlaps between the very heterogeneous and diverse "Fallow or none" category and nearly all other autumn vegetation groups. The most easily identified groups were the homogeneous winter crops "Winter cereals" and "Winter oilseed rape", which reached F1-values of 0.83 and 0.76 (fig 3.1, fig. 7.7).

For main crops, the variation in classification success was more substantial. Some classes were very well classified – "Winter oilseed rape", "Winter cereals","Maize","Beet", and "Spring cereals" all reached F1-values above 0.9. Other classes proved more challenging – "Fallow or none" and "Legumes" performed poorly, with F1-values of only 0.32 and 0.29, respectively. "Potato" and "Trees and bushes" were also challenging for the model to identify (fig 3.2, fig 7.8**)** and were represented by relatively low numbers of observations in the dataset (fig. 3.3).

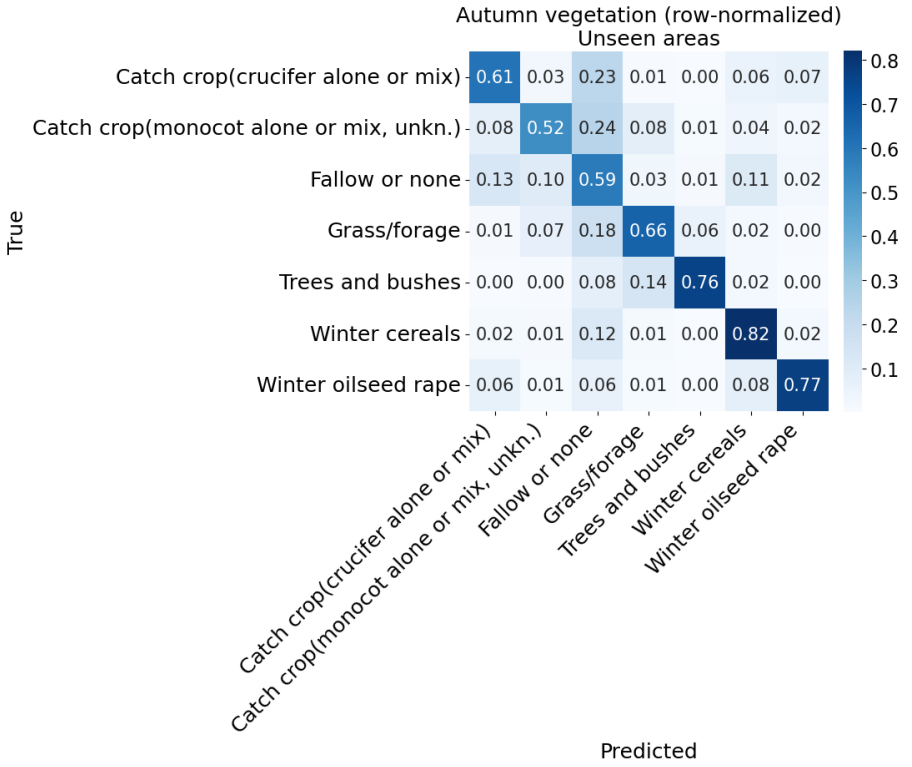


Figure 3.1. Confusion matrix for autumn vegetation classification for LOOP catchments not included in model training. The results are based on six different runs where one LOOP area was held out of training and used for testing in each of the six runs. In a row-normalized confusion matrix like this one, the diagonal values show the recall for the true category given by the row. Off-diagonal values describe how large a proportion of observations from the true category (given by the row) were misclassified as the categories given by the column.

Figure 3.2. Confusion matrix for main crop classification for LOOP catchments not included in the training of the model. The results are based on six different runs, each holding out one LOOP catchment for testing. In a row-normalized confusion matrix like this one, the diagonal values show the recall for the true category given by the row. Off-diagonal values describe how large a proportion of observations from the true category (given by the row) were misclassified as the categories given by the column.

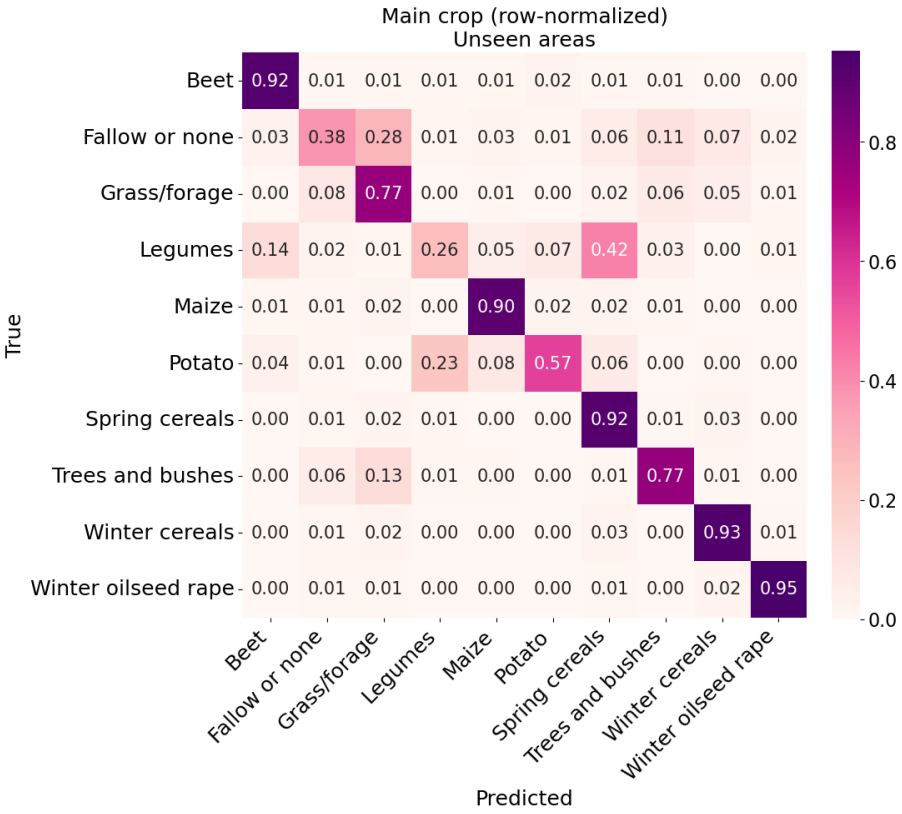
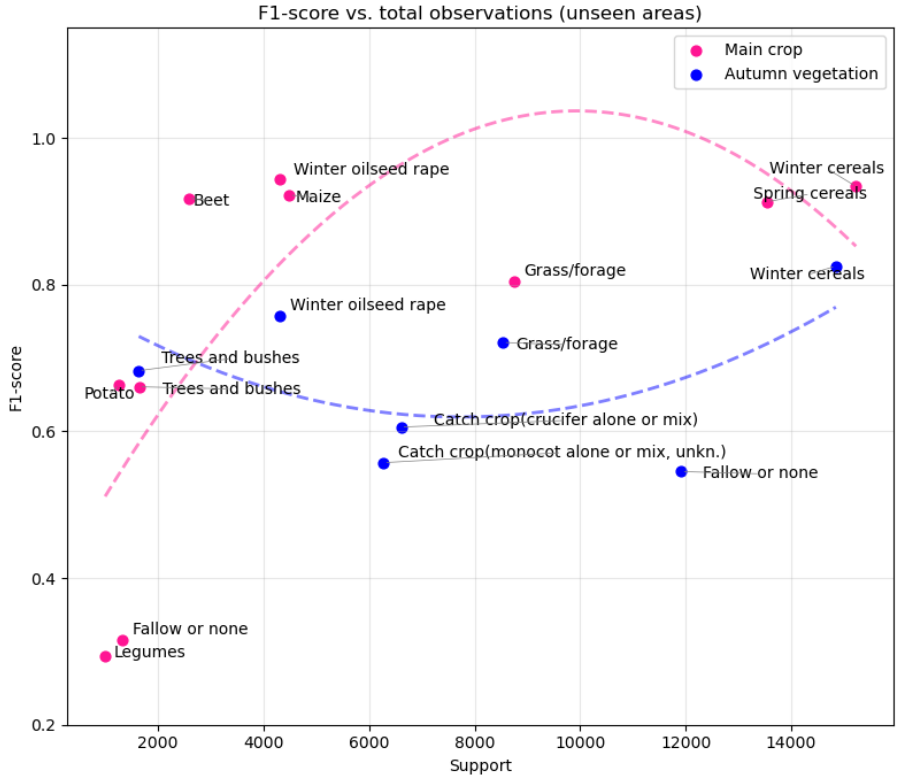


Figure 3.3. F1-values for autumn vegetation and main crop classification for LOOP catchments not included in the model training. The results are based on six different training runs, each holding out one LOOP catchment for testing. Support (number of observations in the dataset) on the y-axis reflects all the data and does not account for how large a proportion of each category was seen by the model during training. The lines are the trends in F1-value given by the size of the category in the data.



3.2 Classification of seen areas, unseen fields

When classifying the autumn vegetation of tiles originating from LOOP catchments that the model had already seen during training, the autumn-sown main crops "Winter oilseed rape" and "Winter cereals" were identified most accurately, with correct classification rates 79% and 88% of the time, respectively. The autumn vegetation category most challenging to identify was the highly heterogeneous "Fallow or none" category (identified correctly only 51% of the time) (fig. 3.4, fig. 7.9).

When classifying the main crops of tiles from LOOP catchments included in the training data, the model was able to correctly identify the majority of categories most of the time, with six of 10 achieving classification accuracies of 90% or higher. The most difficult category to identify was the highly heterogeneous "Fallow or none", with correct identification only 34% of the time (fig. 3.5, fig. 7.10).

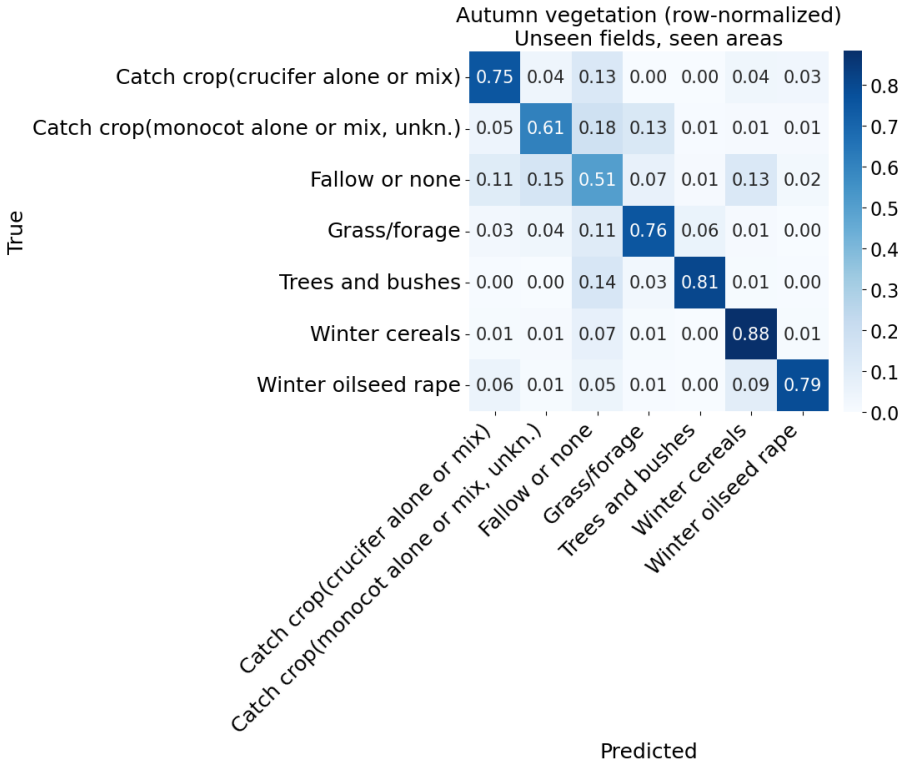


Figure 3.4. Confusion matrix for autumn vegetation classification. Here, test data was randomly selected, so no LOOP catchment was held out. It was, however, ensured that no field seen during training was used for testing. In a row-normalized confusion matrix like this one, the diagonal values show the recall for the true category given by the row. Off-diagonal values describe how large a proportion of observations from the true category (given by the row) were misclassified as the categories given by the column.

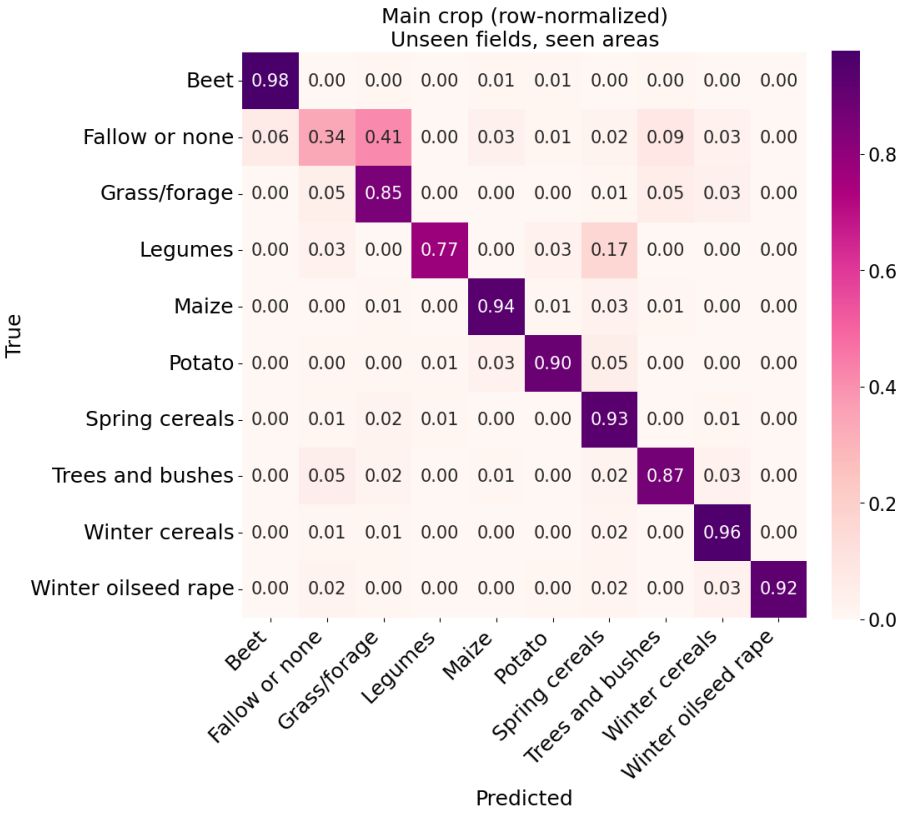


Figure 3.5. Confusion matrix for main crop classification. Here, test data was randomly selected, so no LOOP catchment was held out. It was, however, ensured that no field seen during training was used for testing. In a row-normalized confusion matrix like this one, the diagonal values show the recall for the true category given by the row. Off-diagonal values describe how large a proportion of observations from the true category (given by the row) were misclassified as the categories given by the column.

Correct classification was easier for fields from LOOP catchments that the model had already seen during training (fig. 3.4, 3.5), even though the model had not seen the exact fields before, than for fields from previously unseen geographical areas (fig. 3.1, fig. 3.2**)**. The difference was most noticeable for groups with few observations and high heterogeneity.

3.3 Comparison to baseline model

For all main crops and autumn vegetation categories, the trained model performed much better than the baseline model (a random guessing model that was parameterized by the vegetation category frequencies in the training dataset constructed from randomly chosen fields). The largest improvement in main crop task was observed for the "Beet" category, where the F1-value was 184 times higher than that of the baseline model (random guessing based on training dataset frequencies). The smallest improvement was found for "Spring cereals", a very well represented category that is often guessed correctly by chance, where the trained model achieved an F1-value 3.5 times higher than the baseline

In the autumn vegetation task, the largest improvement over the baseline model (ratio 22.67) was observed for the category "Trees and bushes" and the smallest improvement for by the very heterogeneous category "Fallow or none" (ratio 3.06) (tab. 3.1).

Table 3.1. Comparison of the baseline model's F1-value with the F1-values for LOOP catchment fields seen by the deep learning model during training and fields from LOOP catchments not seen by the model during training.

Main crop	Baseline model F1	F1 seen area	F1 unseen area	Ratio Seen/ Baseline	Ratio Unseen/ Baseline	Autumn vegetation	Baseline model F1	F1 seen area	F1 unseen area	Ratio Seen/ Baseline	Ratio Unseen/ Baseline
Winter oilseed rape	0.06	0.95	0.94	15.83	15.67	**Winter cereals**	0.25	0.85	0.83	3.40	3.32
Beet	0.005	0.96	0.92	192.00	184.00	**Winter oilseed rape**	0.05	0.82	0.76	16.40	15.20
Winter cereals	0.22	0.95	0.93	4.32	4.23	**Trees and bushes**	0.03	0.72	0.68	24.00	22.67
Maize	0.05	0.95	0.92	19.00	18.40	**Grass/ forage**	0.2	0.77	0.72	3.85	3.60
Spring cereals	0.26	0.94	0.91	3.62	3.50	**Catchcrop (crucifer etc.)**	0.1	0.76	0.61	7.60	6.10
Grass/ forage	0.2	0.86	0.8	4.30	4.00	**Catch crop (monocot etc.)**	0.07	0.61	0.56	8.71	8.00
Trees and bushes	0.04	0.74	0.66	18.50	16.50	**Fallow or none**	0.18	0.52	0.55	2.89	3.06
Legumes	0.02	0.75	0.29	37.50	14.50						
Potato	0.01	0.91	0.66	91.00	66.00						
Fallow or none	0.03	0.34	0.32	11.33	10.67						

4 Discussion

4.1 Potential intrinsic ceilings to classification performance

There are large performance differences between how well the model can classify the main crop and the autumn vegetation type. While the model performed reasonably well in identifying main crop type for the more frequently occurring categories, it was less convincing at classifying the autumn vegetation type for equally common categories. The model's successful ability to classify main crops indicates that the model setup was at least reasonably well suited for vegetation identification tasks. However, there may be a technical ceiling on how well models one of this kind can perform on tasks like this – one that lies below what would be needed for reliable and precise real-life applications. For main crops, the ceiling would primarily be driven by the trade-off between group sample size and intra-group homogeneity. For the autumn vegetation, the ceiling would be markedly lower due to the same trade-off, high cloud cover in autumn, less distinct growth patterns, and shorter growth periods of the autumn vegetation.

A further challenge that is difficult to overcome is that fields with very similar, or even identical, vegetation can be assigned to different categories in the agricultural register because they serve different agronomic purposes. For example, a field that is covered by grass may be recorded as a catch crop, a main crop, or a fallow field, depending on the agricultural and/or legal context.

Additionally, the 10–60 m spatial resolution of the satellite data reduces the ability to discern fine-scale plant morphological details, limiting the analysis to temporal dynamics of spectral profiles.

4.2 Potential for improving classifications

"Winter oilseed rape" and "Winter cereals" were more easily recognized as main crop than as autumn vegetation (fig. 3.1, fig. 3.2, fig. 3.4, fig. 3.5**)**. Including satellite data stretching further into the next calendar year than March 14th would improve accuracy in autumn vegetation classification due to additional information from late spring and early summer growth patterns of the winter-sown main crop.

Classification was more successful for field tiles originating from LOOP catchments that the model had already seen during training (fig. 3.1, fig. 3.2, fig. 3.4, fig. 3.5). This may reflect differences in the crops representing each vegetation group or differences in cultivation practices or growth patterns across LOOP catchments, or simply because the model did not receive enough and sufficiently diverse input to generalize to new areas.

This indicates that the model may not have reached its full generalization ceiling, and that its generalization ability could likely be improved by obtaining and including more training data from diverse geographical areas with diverse representatives of the main crop and autumn vegetation classes. The needed main crop data could easily be obtained for more geographical areas in Denmark, but autumn vegetation type data is not as readily available (which is why a model for autumn vegetation classification such as this one is relevant). However, since the main crop and autumn vegetation classification

tasks are strongly coupled, this inclusion may still support the model in learning good representations of vegetation data. This would certainly be true for autumn vegetation categories that are registered in the next agronomical year as main crops such as "Winter oilseed rape", "Winter cereals" and to similarly the "Trees and bushes" category, where the vegetation cover type is generally stable throughout the whole year.

4.3 Some vegetation groups may be naturally easier to discern

Some vegetation groups, for instance "Beet", "Potato", and "Maize", seem to be relatively easy for the model to discern. The model has been able to distinguish them from other main crops at much higher rates than training-data-frequency-based random guesses (the baseline model). For instance, the main crop "Beet" was correctly identified a notable 184 times more frequently than random guessing would have resulted in. "Potato" and "Maize" also seem to lend themselves well to the classification, being identified 66 and 18.4 times more often than training-data-frequency-based random guesses. These high ratios are partly due to the rarity of these crops: Since they occur infrequently in the data, a frequency-based random guess would rarely select them. While this inflates the ratio, a high ratio requires the model to correctly identify these crops with some regularity.

For autumn vegetation, "Trees and bushes" is identified correctly at a 22.67 higher rate than training-data-frequency-based random guesses. This may be caused by the category being stable throughout the year, allowing the model to base its prediction on data from all 24-time steps across the full year. "Winter oilseed rape" is identified at 15.20 times the rate than that of the baseline model, despite not sharing the clear advantage of "Trees and bushes", which suggests that it has an easily discernable spectral or temporal expression.

4.4 Other considerations and use cases

As very rare crop categories like "Spring oilseed rape" (main crop) were removed from the training, validation, and test datasets, the model cannot predict those categories, and requesting a classification of them would in all cases result in a misclassification. This could be solved by adding a non-informative "other" category to place everything not belonging to the 10 main crop groups and the 7 autumn vegetation groups into to the model. However, such a category would be difficult to predict because it would be highly heterogeneous internally.

The model in its current form will not outcompete sound register data, but it may provide an indication of the autumn vegetation category in cases where registration is missing, such as the limited time period from 2017-2018, for which satellite data are available but field-level registration of autumn vegetation is not. It may also be useful for identifying sub-types of catch crops that current registration does not provide.

The goal of this project was to develop a model that could support catchment-scale decomposition of autumn vegetation, which would in turn be useful for comparison with nitrogen levels in the catchment streams. However, the model may be more appropriate for supporting register data (where present) rather than substituting it. It could be used to detect discrepancies between register data and satellite data for certain catchment, year, - and autumn vegetation type combinations.

5 Conclusion

It was possible to train a model that classifies main crops and autumn vegetation from satellite data with moderate success, and considerably better than training-dataset-frequency-based random guesses. However, several serious constraints were encountered: limited availability of optical data during the autumn and winter periods, strong phenological similarity between vegetation belonging to different administrative vegetation classes, and a challenging trade-off between inter-group homogeneity and sample sizes.

The model performed worse on under-represented groups such as ”Potatoes” and “Legumes” in the main crop classification. For instance, the “Legume” group, represented by only 61 fields, achieved an F1 of 0.29, while the much larger “Spring cereals” group, with 1163 fields, achieved an F1 of 0.91. Autumn vegetation classification appeared to be affected by the vegetation types and spectral overlap associated with the heterogeneous “Fallow or none” group with several other groups: For LOOP catchments unseen during training, ”Grass/forage” was misclassified as “Fallow or none” 18% of the time, and the two groups “Catch crop (crucifer alone or mix)” and “Catch crop (monocot alone or mix, unkn.)” were misclassified as “Fallow or none” as much as 23 and 24% of the time, respectively. This hints at the fact that some register-based categories do not have as clear demarcation lines between them in the field as they do in registers.

Since the model performs better on unseen fields located in LOOP catchments represented during training than on fields from entirely new areas, this suggests that additional training data, covering a wider range of geographical areas and crops, could improve model performance. This is particularly relevant for improving the classification success for heterogeneous vegetation categories that may be represented by different crop types in different catchments. For instance, the “Trees and bushes” category is represented in LOOP catchment 3 mainly by Christmas tree production, while in LOOP catchment 4 it is represented by fruit and berries: apple, pear, and currant (red and black).

Autumn-sown crops like “Winter oilseed rape” and ”Winter cereals” were, as expected, classified more accurately as main crops than as autumn vegetation. This indicates that extending the satellite data time series into the following calendar year could provide additional information and potentially improve the accuracy of autumn vegetation classification.

However, in practice there is likely to be a ceiling to how much better the classification accuracy can become, the principal limitation being real-life vegetational overlaps of dissimilar register-based categories. Other important limitations are the relatively flatter and more heterogeneous growth curves of vegetation in the temperature-limited autumn-winter period and the limited accessibility of optical satellite data in autumn and winter due to frequent cloud cover.

Currently, information from satellite data can be used to predict autumn vegetation with accuracies of 77-82% for the following vegetation categories: “Winter oilseed rape”, “Trees and bushes”, and “Winter cereals”, while the lowest prediction accuracies of 52% and 59% are observed for “Catch crop monocot, mix, unknown” and “Fallow or none”, respectively.

While testing was only performed inside of the LOOP-catchments, the model is expected to generalize to other geographical areas, given that they are reasonably similar in terms of climate, soil and cropping systems to the LOOP-catchments.

6 References

Abiodun, O. I., A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed & Arshad, H. 2018. State-of-the-art in artificial neural network applications: A survey. Heliyon 4 (11): e00938. https://doi.org/10.1016/j.heliyon.2018.e00938.

Aggarwal, C. C. 2023. Neural networks and deep learning: A textbook. Springer International Publishing. https://doi.org/10.1007/978-3-031-29642-0.

Assis, A., J. Dantas & Andrade, E. 2024. The performance-interpretability trade-off: A comparative study of machine learning models. Journal of Reliable Intelligent Environments 11 (1): 1. https://doi.org/10.1007/s40860-024-00240-0.

Bendini, H. do N., R. Fieuzal, P. Carrere et al. 2024. Estimating winter cover crop biomass in France using optical Sentinel-2 dense image time series and machine learning. Remote Sensing 16 (5). https://doi.org/10.3390/rs16050834.

Blicher-Mathiesen, G., M. Thorsen, J. Wienke et al. 2024. Landovervågningsoplande 2023: NOVANA. No. 628. NOVANA. DCE – Nationalt Center for Miljø og Energi.

Børgesen, C. D., P. Sørensen, G. Blicher-Mathiesen, J. W. M. Pullens, J. Zhao & Olesen, J. E. 2020. An empirical model for predicting nitrate leaching from the root zone of agricultural land in Denmark. No. 163.

Campos-Taberner, M., F. J. García-Haro, B. Martínez et al. 2019. A Copernicus Sentinel-1 and Sentinel-2 classification framework for the 2020+ European Common Agricultural Policy: A case study in València (Spain). Agronomy 9 (9). https://doi.org/10.3390/agronomy9090556.

Det Økonomiske Råd. 2017. Økonomi og miljø 2017. Det Økonomiske Råd. https://dors.dk/files/media/rapporter/2017/M17/m17.pdf.

Dodds, W. K. & Welch, E. B. 2000. Establishing nutrient criteria in streams. Journal of the North American Benthological Society 19 (1): 186–196. https://doi.org/10.2307/1468291.

Geirhos, R., P. Rubisch, C. Michaelis, M. Bethge, F. A. Wichmann & Brendel, W. 2022. ImageNet-trained CNNs are biased towards texture; Increasing shape bias improves accuracy and robustness. arXiv:1811.12231. https://doi.org/10.48550/arXiv.1811.12231.

Goffart, D., Y. Curnel, V. Planchon, J.P. Goffart & Defourny, P. 2021. Field-scale assessment of Belgian winter cover crops biomass based on Sentinel-2 data. European Journal of Agronomy 126: 126278. https://doi.org/10.1016/j.eja.2021.126278.

Holtgrave, A.-K., N. Röder, A. Ackermann et al. 2020. Comparing Sentinel-1 and -2 data and indices for agricultural land use monitoring. Remote Sensing 12 (18). https://doi.org/10.3390/rs12182919.

IDB-NDMI. 2025. IDB – Information for sensor and index (NDMI). https://www.indexdatabase.de/db/si-single.php?sensor_id=96&rsindex_id=56.

IDB-NDRE. 2025. IDB – Information for sensor and index (NDRE). https://www.indexdatabase.de/db/si-single.php?sensor_id=96&rsindex_id=223.

IDB-NDVI. 2025. IDB – Information for sensor and index (NDVI). https://www.indexdatabase.de/db/si-single.php?sensor_id=96&rsindex_id=58.

IDB-PSRI. 2025. IDB – Index: Plant senescence reflectance index. https://www.indexdatabase.de/db/i-single.php?id=69.

Immitzer, M., F. Vuolo & Atzberger, C. 2016. First experience with Sentinel-2 data for crop and tree species classifications in Central Europe. Remote Sensing 8 (3). https://doi.org/10.3390/rs8030166.

Jensch, K., G. Ghazaryan, S. Ernst, P. Hostert & Nendel, C. 2025. Integrating Landsat, Sentinel-2 and Sentinel-1 time series for mapping intermediate crops. European Journal of Remote Sensing 58 (1): 2507738. https://doi.org/10.1080/22797254.2025.2507738.

Kumar, U., T. Engedal, V. Hansen et al. 2025. Late sowing of cover crops reduces potential for nitrate leaching reduction and carbon inputs. Agriculture, Ecosystems & Environment 393: 109858. https://doi.org/10.1016/j.agee.2025.109858.

LandbrugsGIS/Landbrugsstyrelsen. 2025. LandbrugsGIS – Download kortdata. https://landbrugsgeodata.fvm.dk/.

LeCun, Y., Y. Bengio & Hinton, G. 2015. Deep learning. Nature 521 (7553): 436–444. https://doi.org/10.1038/nature14539.

Li, J., W. Hu, H. W. Chau et al. 2023. Response of nitrate leaching to no-tillage is dependent on soil, climate, and management factors: A global meta-analysis. Global Change Biology 29 (8): 2172–2187. https://doi.org/10.1111/gcb.16618.

Nouri, A., S. Lukas, S. Singh, S. Singh & Machado, S. 2022. When do cover crops reduce nitrate leaching? A global meta-analysis. Global Change Biology 28 (15): 4736–4749. https://doi.org/10.1111/gcb.16269.

Perumal, T., N. Mustapha, R. Mohamed & Shiri, F. M. 2024. A comprehensive overview and comparative analysis on deep learning models. Journal on Artificial Intelligence 6 (1): 301–360. https://doi.org/10.32604/jai.2024.054314.

PyTorch. 2025. CrossEntropyLoss — PyTorch 2.9 documentation. https://docs.pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html.

Schwartzkopff, M. V., P. Abrahamsen, J. Eriksen & Jensen, L. S. 2026. The role of autumn crop-coverage in nitrate leaching: A Daisy model scenario analysis comparing catch-crops and winter cereals across multiple sites and climatic conditions in Denmark. Geoderma Regional 44: e01033. https://doi.org/10.1016/j.geodrs.2025.e01033.

Selvaraj, S., D. Bargiel, A. Htitiou & Gerighausen, H. 2025. A SAR-driven approach for winter catch crop classification in Germany. IEEE Access 13: 171642–171665. https://doi.org/10.1109/ACCESS.2025.3615739.

Sentinel-1 GRD. 2025. https://docs.sentinel-hub.com/api/latest/data/sentinel-1-grd/.

Sentinel-2 L2A. 2025. https://docs.sentinel-hub.com/api/latest/data/sentinel-2-l2a/.

Tufail, R., P. Tassinari & Torreggiani, D. 2025. Deep learning applications for crop mapping using multi-temporal Sentinel-2 data and red-edge vegetation indices: Integrating convolutional and recurrent neural networks. Remote Sensing 17 (18). https://doi.org/10.3390/rs17183207.

Valkama, E., R. Lemola, H. Känkänen & Turtola, E. 2015. Meta-analysis of the effects of undersown catch crops on nitrogen leaching loss and grain yields in the Nordic countries. Agriculture, Ecosystems & Environment 203: 93–101. https://doi.org/10.1016/j.agee.2015.01.023.

Vanpoucke, K., S. Heremans, E. Buls et al. 2024. Field-level classification of winter catch crops using Sentinel-2 time series: Model comparison and transferability. Remote Sensing 16 (24). https://doi.org/10.3390/rs16244620.

Vejledning om obligatoriske målrettede efterafgrøder 2024. 2025. Landbrugsstyrelsen. https://lbst.dk/Media/638640701100753573/Vejledning%20om%20obligatoriske%20målrettede%20efterafgrøder%202024_240724.pdf.

Vejledning om regler for pligtige efterafgrøder og husdyrefterafgrøder og dyrkningsrelaterede tiltag, planperioden 1. august 2025 til 31. juli 2026. 2025. Styrelsen for Grøn Arealomlægning og Vandmiljø. https://lbst.dk/Media/638907634605166192/Vejledning%20efterafgrøder%20og%20dyrkningsrelaterede%20tiltag%20for%202025_2026_version_august.pdf.

Wanic, M., K. Zuk-Golaszewska & Orzech, K. 2019. Catch crops and the soil environment – A review of the literature. Journal of Elementology 24 (1). https://doi.org/10.5601/jelem.2018.23.3.1638.

Zhao, J., J. W. M. Pullens, P. Sørensen, G. Blicher-Mathiesen, J. E. Olesen & Børgesen, C. D. 2022. Agronomic and environmental factors influencing the marginal increase in nitrate leaching by adding extra mineral nitrogen fertilizer. Agriculture, Ecosystems & Environment 327: 107808. https://doi.org/10.1016/j.agee.2021.107808.

7 Appendix

7.1 Vegetation composition all areas

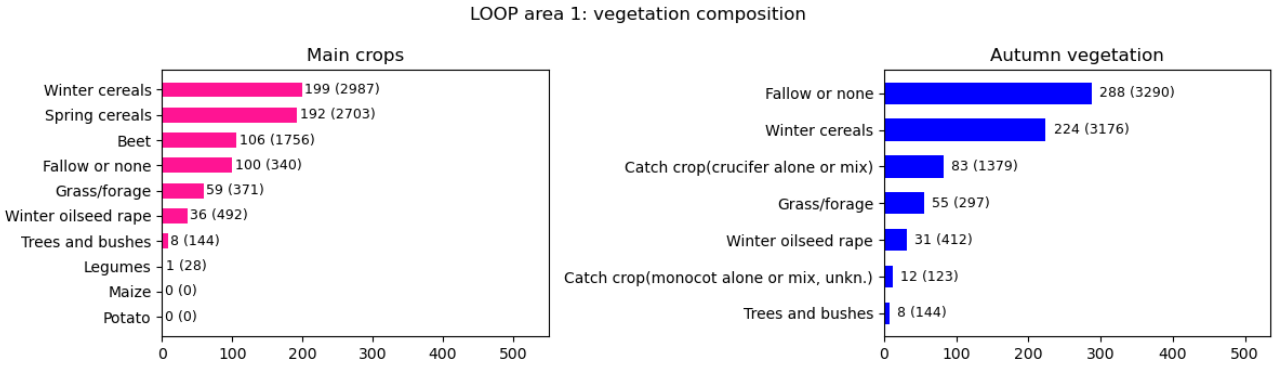


Figure 7.1. Vegetation composition in LOOP catchment 1. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 1. The number in parenthesis is the number of distinct 80*80 meter field tiles belonging to the category in the area.

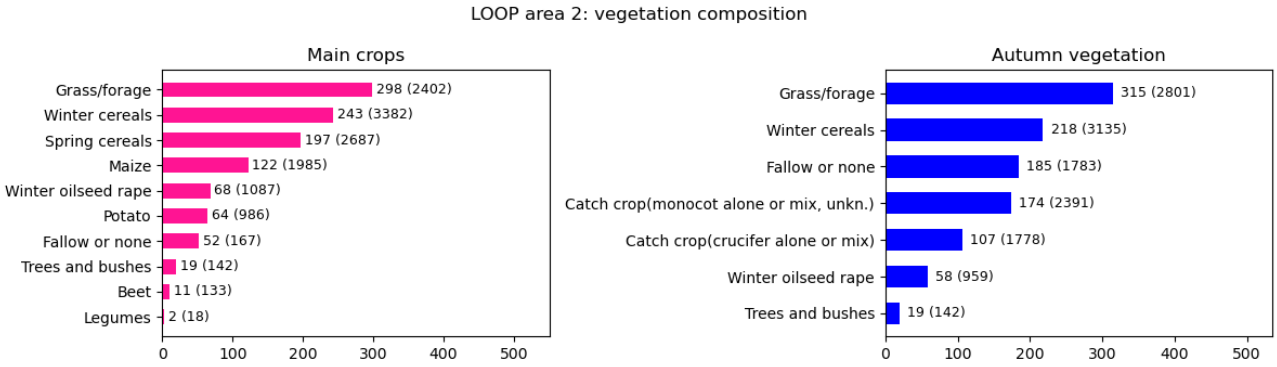


Figure 7.2. Vegetation composition in LOOP catchment 2. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 2. The number in parenthesis is the number of distinct 80*80 meter field tiles belonging to the category in the area.

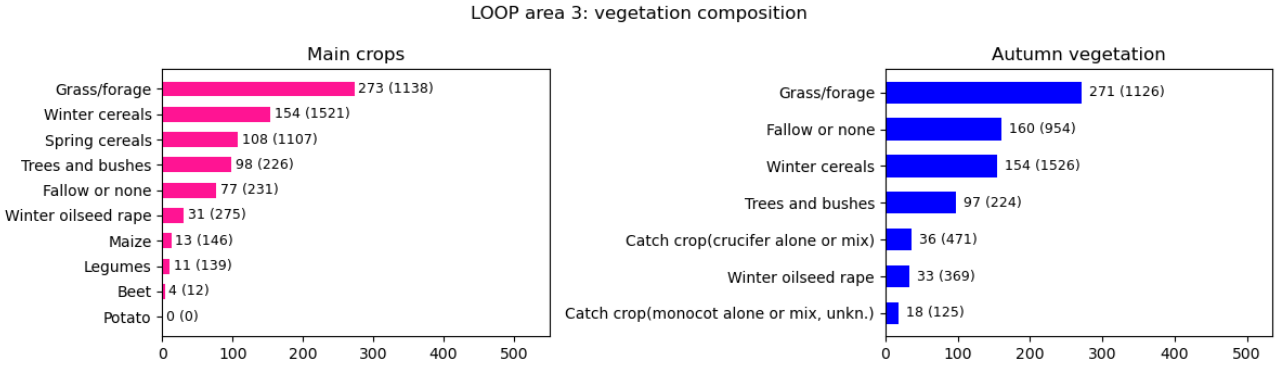


Figure 7.3. Vegetation composition in LOOP catchment 3. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 3. The number in parenthesis is the number of distinct 80*80 meter field tiles belonging to the category in the area.

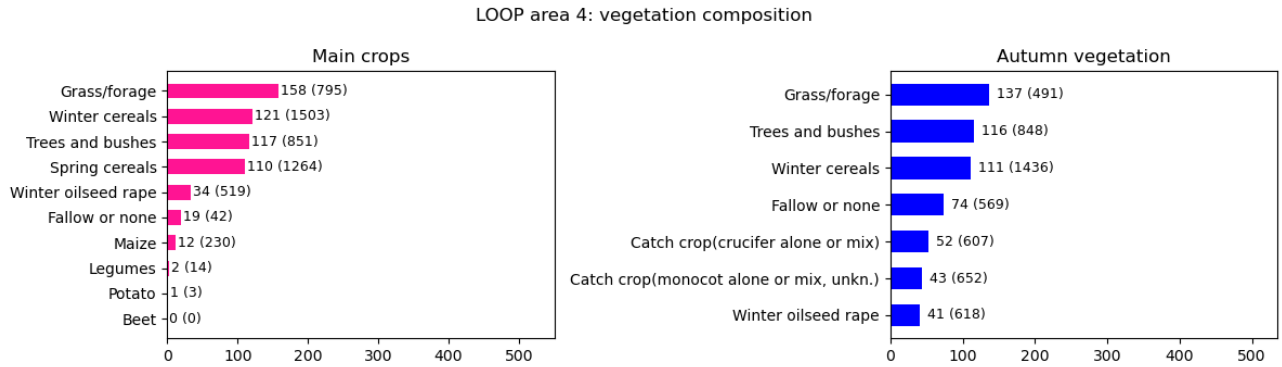


Figure 7.4. Vegetation composition in LOOP catchment 4. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 4. The number in parenthesis is the number of distinct 80*80 meter field tiles belonging to the category in the area.

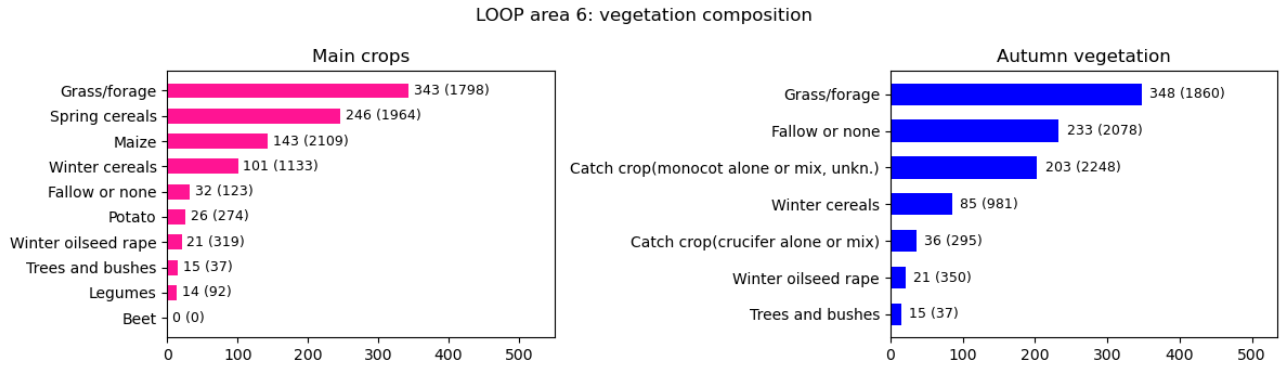


Figure 7.5. Vegetation composition in LOOP catchment 6. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 6. The number in parenthesis is the number of distinct 80*80-meter field tiles belonging to the category in the area.

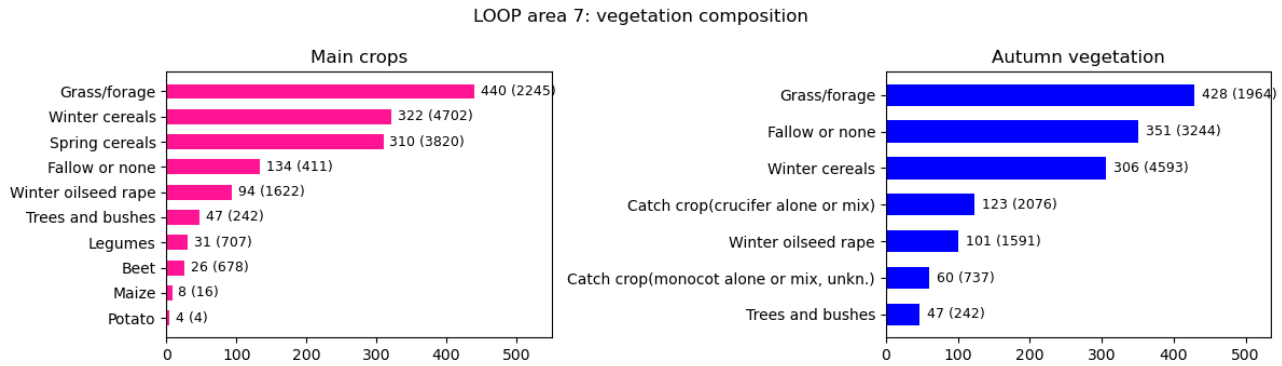


Figure 7.6. Vegetation composition in LOOP catchment 7. Shown are the number of distinct fields belonging to each main crop category and autumn vegetation category in LOOP catchment 7. The number in parenthesis is the number of distinct 80*80 meter field tiles belonging to the category in the area.

7.2 F1-value graphs

Unseen areas

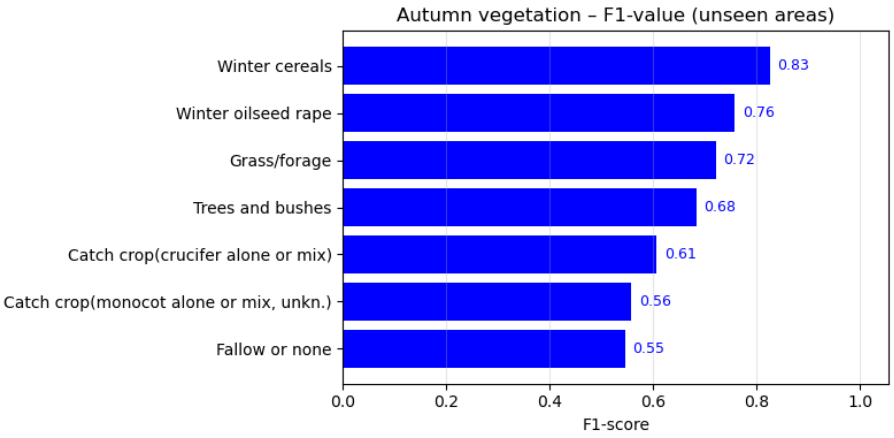


Figure 7.7. F1-values for autumn vegetation classification in LOOP- areas unseen during training. -The results are derived from six different runs where in each run one LOOP catchment was held out of training for use in testing.

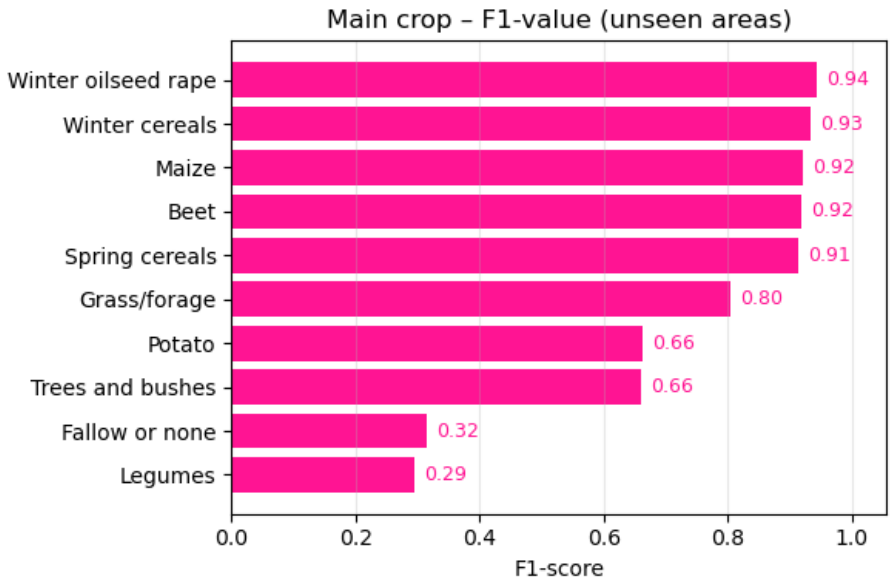


Figure 7.8. F1-values for main crop classification in LOOP catchments unseen during training. The results are derived from six different runs where in each run one LOOP catchment was held out of training for use in testing.

Seen areas

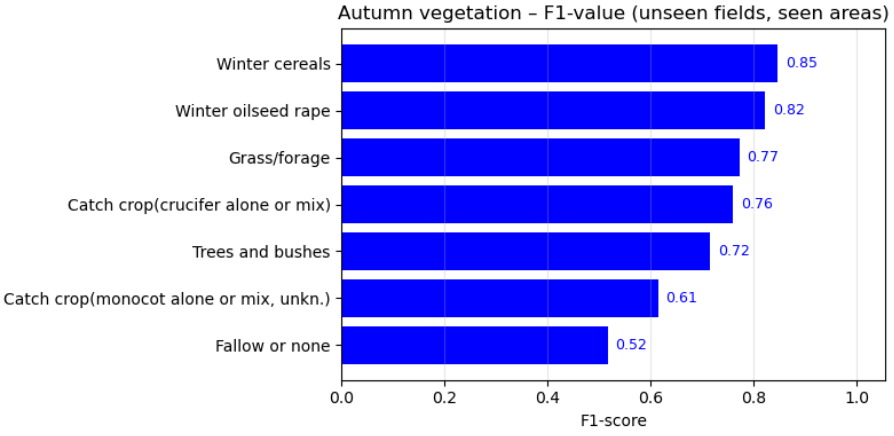


Figure 7.9. F1-values of autumn vegetation categories. Here, test data was randomly selected, so no LOOP catchment was held out. However, it was ensured that no field seen during training was used for testing.

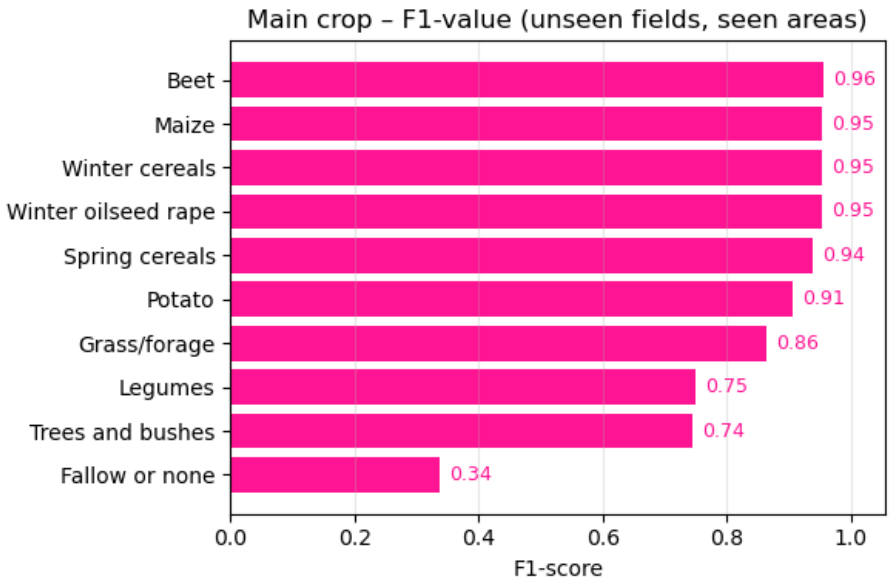


Figure 7.10. F1-values for main crop classification. Here, test data was randomly selected, so no LOOP catchment was held out. However, it was ensured that no field seen during training was used for testing.

7.3 Model diagram

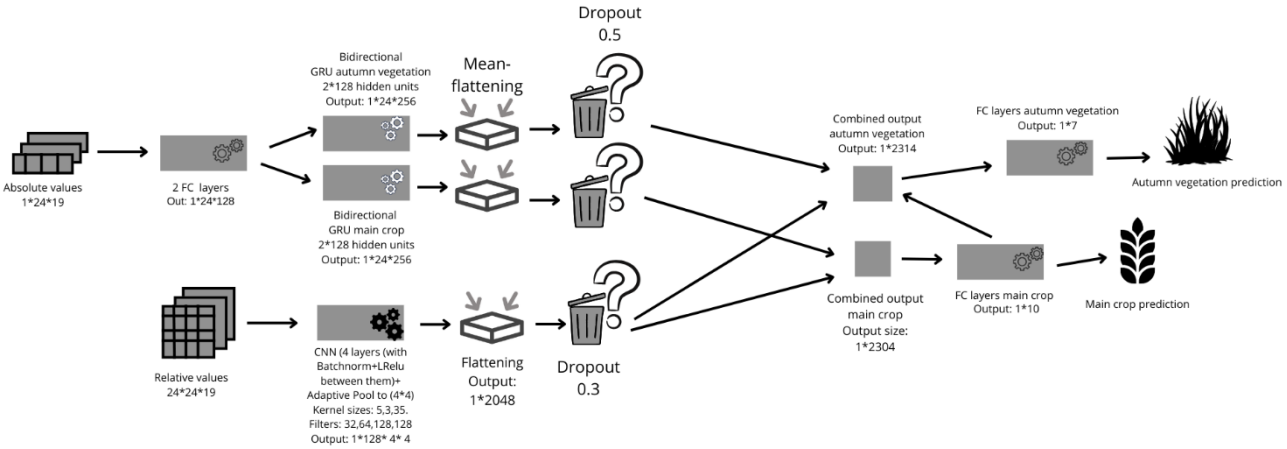


Figure 7.11. Overall sketch of the model structure (some details left out for brevity). The model includes two branches, a CNN branch and a GRU branch, and has two heads performing main crop and autumn vegetation predictions, respectively.

CLASSIFICATION OF DANISH AUTUMN VEGETATION USING SATELLITE DATA

Data from the Danish Agricultural Monitoring Programme

This report presents a deep learning model to classify autumn vegetation in the six Danish LOOP-catchments over eight years (2017–2024) using optical and radar satellite data. The model also predicts main crop types to support vegetation classification. It was trained on 5,477 field–year combinations from 1,220 fields. Autumn vegetation was correctly classified into seven categories 67.8% of the time. Main crops could be classified into ten categories at an accuracy of 86.2%. The performance difference reflects challenges related to morphological overlaps among vegetation types in autumn, subtle autumn growth patterns, and limited optical data due to cloud cover.

ISBN: 978-87-7648-034-9
ISSN: 2244-999X