European
Commission

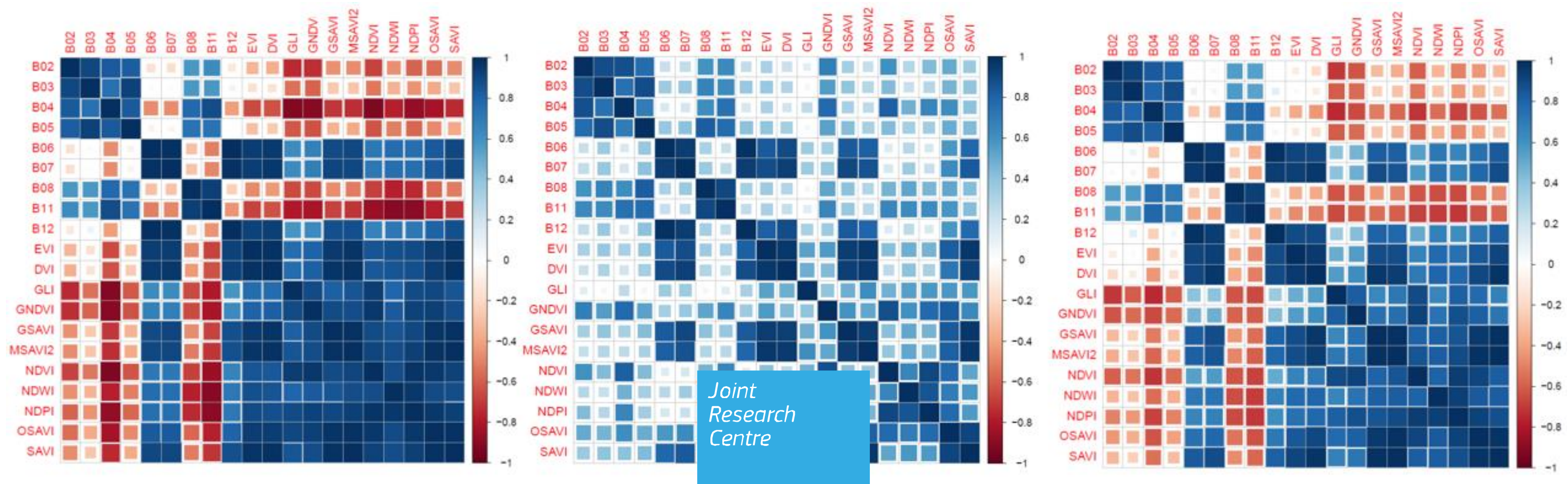
JRC TECHNICAL REPORT

Proposed workflow for optimization of land monitoring systems

Copernicus Sentinel time series analysis within the Checks by Monitoring framework

Zieliński R., Skøien J.O., Acquafresca L., Devos W.

2022



Joint
Research
Centre

Contact information
Name: Rafal Zielinski
Address: Via E. Fermi 2749, I-21027 Ispra/Italy
Email: rafal.zielinski@ec.europa.eu
Tel.: +39 0332-78 6327

EU Science Hub
https://joint-research-centre.ec.europa.eu

JRC130659

Ispra: European Commission, 2022

How to cite this report: Zieliński R., Skøien J. O., Acquafresca L., Devos W., Proposed workflow for optimization of land monitoring systems - Copernicus Sentinel time series analysis within the Checks by Monitoring framework, European Commission, Ispra, 2022, JRC130659.

Contents

Abstract

Member States in the European Union are responsible for implementing the payment schemes to farmers. Checks of farmer activity have traditionally been performed with on-the-spot-checks, but with the free access to wall-to-wall satellite imagery from the European Copernicus programme, Checks by Monitoring was introduced to reduce of the controls burden and facilitate the management of early warnings in combination with an option to correct the aid application that could prevent non-compliances. The automation of this process helped to move it from a sample-based approach to covering the full population.

Much work has already been done regarding cloud-based solutions and large-scale detection of activities, but less attention was given to understanding the relationship between the activity and the corresponding signals, and in general the signal behaviour. This report describes a framework that could be followed to analyse the signal behaviour from a large number of bands and indices from remote sensing sensors, and to select one or a few of them as potential candidates for marker development. In the analysed data sets, indices from Sentinel-2 appears to have better discriminative power than those from Sentinel-1, but also the latter can give good results for countries where clouds reduce the availability of Sentinel-2 images. Several different statistical descriptors of the remote sensing pixels appeared to give equally good results (P25, P50, P75 and mean), but because that the median is less affected by outliers, it would be the recommended choice.

The result of signal behaviour analyses highly depends on quality and level of details of input data and are based on ground truth data, the signal time series are computed from image pixels located inside the boundaries of a feature of interest (agricultural parcel) and dates of activities.

The signal selection approach is generic and can be applied to any land monitoring system where there is an interest in changes on a preselected field. The method is adjustable, exploring (by providing a sequence of statistical methods) links between local knowledge and practices on the one side, and remote sending images and the derived signals on the other side.

The report also demonstrates that the framework can be used to address common challenges during the CbM development, and gives an incentive to improve the ground data collection for CbM application.

An R package (programming language for statistical computing) that can handle most of the analyses in this report will be available at the start of 2023.

Acknowledgements

The Authors would like to thank the Members States' administration who made this work feasible by sharing the samples of national datasets and CbM related information. The data used in this work were shared in the frame of CbM Outreach 2021 initiative.

We would also like to thank the independent reviewers of this report, Aleksandra Sima, Pavel Milenov and the JRC editorial board for their feedback.

Authors

Rafał Zieliński, Jon Olav Skøien, Laura Acquafresca, Wim Devos

1 Introduction

Checks by Monitoring (CbM) is an automated and continuous process relying mostly on Copernicus Sentinel satellite data along with other data sources to determine the agricultural activity on all the agricultural parcels declared in defined payment schemes. Thanks to the European Copernicus programme's free, full and open data access policy (Breger, 2017; European Commission, 2013) the financial constraints of purchasing image data are no longer an obstacle for checks on the Common Agricultural Policy (CAP) payments and in particular on Checks by Monitoring applications.

In 2018 the European Commission introduced changes to the Implementing Regulation (EU) No 809/2014 that allowed Member States to choose the CbM as a substitute of the current on-the-spot-checks (OTSC). The main reasons for introducing the monitoring are reduction of the controls burden and facilitate the management to of early warnings in combination with an option to correct the aid application that could prevent non-compliances. The conceptual basis of check by monitoring were elaborated by Joint Research Centre (JRC) to guide the development and summarised in two discussion documents (Devos et al., 2017; Wim Devos et al., 2018a) and a technical guidance (W Devos et al., 2018b), and recently revised conceptual CbM document (Devos et al., 2021b).

The identification of an appropriate scenario, and corresponding Feature Of Interest (FOI) are the initial conditions of the CbM. The scenario brings the local context into the process of integrating image-based analyses as it manages available knowledge on agricultural practices. Understanding which activities are likely to happen on the ground is fundamental to make correct processing design choices and is critical for real case implementations. That knowledge of local practices allows to link a specific signal behaviour to a particular bio-physical manifestation in the FOI and anticipates reliable and correctly timed markers detection.

Even as the results of a farmer's activities and crop development stages are often distinctly manifested on the ground, not all of these manifestations are captured by the satellite imagery, especially when reduced to time series extracts for a single FOI (Voormansik et al., 2020). Therefore, dedicated signal analyses are necessary during the system setup and maintenance.

To enable the automated process that drives each CbM implementation, managed either regionally or nationally, the knowledge about the local practices including crop calendars should be considered as a pre-condition in the system design. Before developing operational monitoring methods, it is important to understand the temporal variations of the remote sensing signal of different crop types, land cover manifestations in a given region (Veloso et al., 2017). A CbM warning system relies on timely observations (recorded through markers) of satellite based signals. Marker development is based on the established relationship between a physical phenomenon and the observable spatio-temporal variation in remote sensing data. To link the two, knowledge of the physics behind the sensor and of the bio-physical aspects of local practice are needed. The former is common remote sensing knowledge but the latter will vary from one MS to another and even within a MS. Markers offer factual evidence in the monitoring system; thus, it is equally important to select the proper signal allowing for the best discrimination rate to detect the land cover manifestation resulting from required farmer's activities. The analyses that translate the local practice into corresponding signal behaviour (marker) are central in the design of the system.

Ground truth data are necessary to introduce the local knowledge into the design process. Such "in-situ" observations are mainly collected in situ, by inspecting parcels/FOIs in regular time intervals to capture the desired phenomenon or traces of farmers' activities. Farmers can also systematically provide ground information, recorded according to a dedicated data collection protocol. The calendar date of an activity, process, or event on the ground truth parcel is the key information in the CbM calibration and tuning. Such information provides an accurate time stamp of a phenomenon occurring on the ground, enabling correct interpretation of signal behaviour.

This work aims to provide a functional methodology for signal analyses needed at different stages of the CbM development process, and especially to facilitate a selection of the signal behaviours that can drive marker development, optimised to capture the activities happening on the ground. It can also contribute to the understanding of the potentials and limitations of remotely sensed imagery in capturing specific phenomena and to assess the detection capabilities of different markers being considered in the CbM development process.

This report uses the CbM concepts described in Devos et al (2021b) and further develop aspects related to signal analyses applied to the CbM, which might be applied during implementation or for an improvement of existing solutions. For the documentation of the CbM processes it refers to Zielinski et al (2022). It discusses the FOI related aspect previously described in Milenov et al (2021). Links to image stack processing and time series retrieval are provided in the JRC-CbM repository documentation by Lemoine and Anastasakis (2022). An overview of signal analyses are also provided in Zielinski et al (2022b).

In 2020, the European Court of Auditors (ECA) recommended that European Commission promotes Checks by Monitoring as the control system of choice for the future CAP (European Court of Auditors, 2020). A part of those recommendations was to set up and maintain a catalogue of best practices in CbM which this report partly addresses.

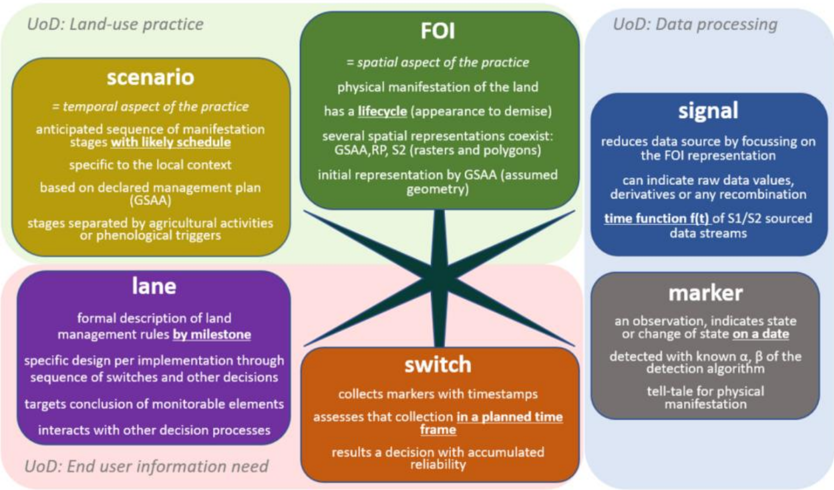
The research reported in this document would not have been possible without the data shared by the MSs in the frame of the CbM outreach 2021 initiative. The objective of this project was to support Paying Agencies in two directions: to provide better understanding of the overall potential of CbM for their individual landscape and CAP processes, and to lower the technology threshold by offering JRC's publicly available toolkit built on standardised access to data and customizable services.

The first chapter of this report present an approach to CbM signal analyses based on ground truth information. Chapter 2 provides a background to CbM with a focus on understanding of the scenario and agricultural practices and their links to signal. Chapter 3 introduces signals providing a brief description of Sentinel- 1 and 2 data and statistical data descriptors. Chapter 4 provides information about input data used to provide the examples to visualise the proposed approach. Chapter 5 starts with the workflow of the proposed approach and further details each processing step with examples. Chapter 6 provides three examples of further analyses using the proposed approach. Chapter 7 summarises the importance and characteristic of ground observations for the CbM application. Chapter 8 provides discussion on selected elements including considerations related to signals, ground truth data, proposed analyses and finally chapter 9 lists the conclusions of the research.

2 Checks by Monitoring background

In the following chapter the basic terms from the revised CbM concept presented in Devos et al. (2021) are presented to support further understanding of the subsequent chapters. Figure 1 shows six key CbM concepts and their short characteristics.

Figure 1. The six key CbM concepts



Source: (Devos et al., 2021b)

Below *an extract* with concise explanation of those concepts after Devos et al. (2021):

— **Scenario**: ''describes a sequence of stages that can be expected from the farmer's choice to use his land over a given timespan. For many practices, the scenario covers a single growing season (for most of arable crops), for some other practices, it can extend to multiple consecutive years (e.g. a full crop rotation cycle)''.

— **FOI**: ''is the surface of the earth where the specified practice will be performed. In a farmer's view, this surface would for example correspond to his/her particular plot, meadow or orchard (or part thereof, when appropriate)''.

— ***Signal:*** ''is a physical quantity that varies with time, space or any independent variable or variables (Proakis and Manolakis, 2006). The Copernicus program delivers a vast amount of imagery which are considered as a main source of data. The CbM-signal is designed to reduce this 'deluge' to something that is easily, reliably and accountably managed. A signal is capture for a specified FOI''.

— **Marker**: ''is an observation of a particular physical state or change of state. A marker records what has been observed on the signal. It holds both the nature of the observation and the time it was manifested in given FOI.''

— **Switch**: ''is a test mechanism to steer processing paths. A switch is an operational method to apply markers in a meaningful way, in line with the end user information needs. The markers originate from the signal data processing and relate to observations over the FOI in question. The relevance of that operation, in end user context, will be embedded in the lane.''

— **Lane**: ''is a processing path, executed by a given resource, which leads to a required conclusion. A lane integrates the formal land management rules necessary and sufficient to confirm the scenario for the FOI. The lane combines switches to collect the necessary external evidence (markers) for a conclusion on the compliance or non-compliance of all requirements relevant for a specific practice''.

The terms of lane and switch are not discussed in this report, but mentioned above for completeness of CbM concept information. This report discusses in more details the relationships between: scenario and agriculture practice; agricultural practice and land cover manifestation, land cover manifestation and signal, FOI and signal, signal and marker.

In other words, the CbM signal carries information that are detectable by a marker. A marker is an observation that confirm (or deny) a specific behaviour which is associated to a land cover manifestation or an activity. The selection and importance of which markers to follow in the process are provided in a scenario.

2.1 Building a scenario

Identification of an appropriate scenario is the first boundary condition of the CbM (Devos et al. 2021). The scenario relies on the aid application to specify the land management declared for an agricultural parcel and local context expressed by agricultural practices. A scenario consists of a sequence of stages (that could be understood as milestones in the farming process) reflecting the anticipated farmer's activities, the region-specific crop phenology or any other potential events (including natural events, such as disasters). The stages are ordered in a sequence as expected from a certain practice, e.g. mowing, growth of vegetation, second mowing etc. Their timeframe (calendar year, growing season or consecutive years/seasons) and approximate timing is known and serves as guide in the process with no pre-fixed dates nor deadlines.

If an activity, process or event does not induce an observable change in the physical manifestation on the field, it's considered not 'monitorable'. This excludes obvious cases of non-monitorable elements at the design stage (i.e. spraying to control insects), but we also have to consider the observability of the change in manifestation on an available signal. Not all physical manifestations are captured by a sensor due to specific sensor characteristics, revisit time or FOI characteristics. Therefore, the selection of monitorable stages considers the user information needs required to confirm/deny the requirement fulfilment. The selection needs to be optimised only for those stages critical for the relevant decision. Examples of scenarios can be found in (Devos et al., 2021b) as well as provided in the form of guidance throughout the structured template for documenting the land monitoring systems (Zielinski et al., 2022a).

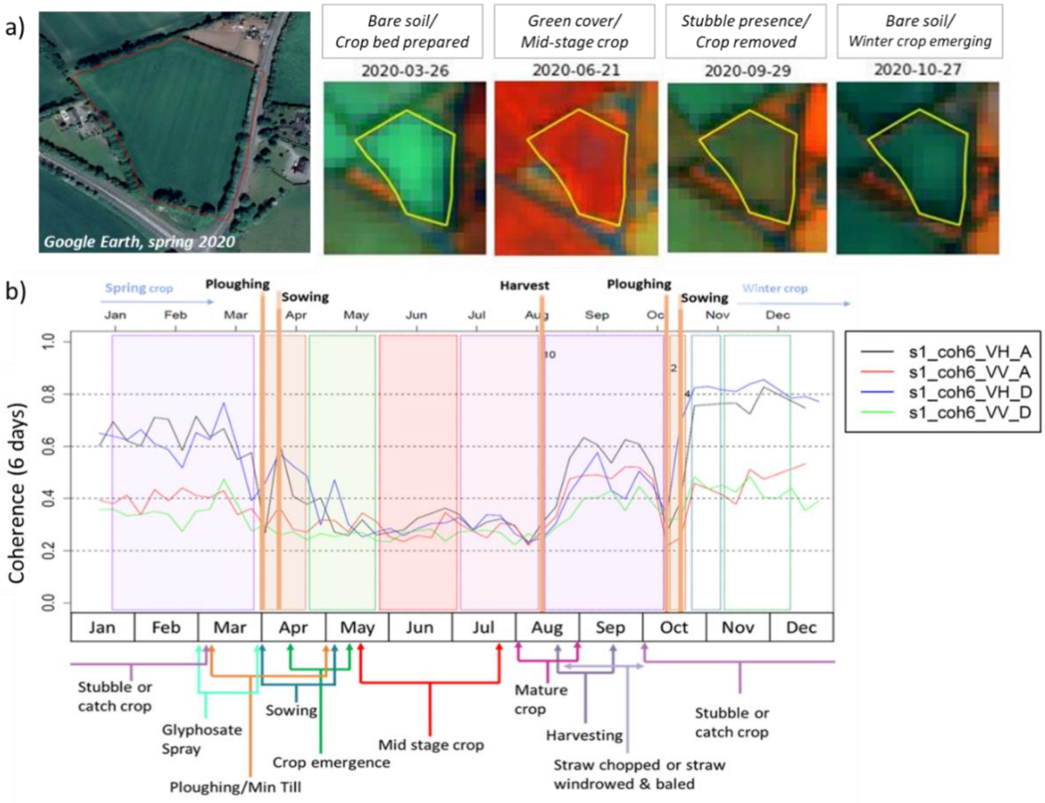
In CbM, the scenario brings the local context into the process of integrating image-based analyses with available knowledge on agricultural practices. Understanding the order of activities are likely to happen on the ground is crucial to make appropriate system design choices. The knowledge of local practice gives clues for reliable and correctly timed marker detection and enables link to assign a specific signal behaviour to a particular bio-physical manifestation in the FOI. The analyses that translate the local practice into signal and to select the corresponding signal behaviour (marker) are central in the design of the system.

The local context is brought in by incorporating agricultural practice in the scenario. Then that information is confronted with user information needs and included to support the link between a physical phenomenon and the observable variation in remote sensing data. This fully guides the monitoring development and support the selection of elements (i.e. markers, switches) to implement.

2.2 The role of the agricultural practice in scenario definition

The definition of the agricultural practice is in the realm of the farmer. The monitoring targets on detection of the actions of a farmer on a declared land with a management choice so there is a clear need of bringing this into the system at the implementation level. A description of an agricultural practice may consist of a crop calendar information with farmers activities, showing the schedule of agricultural operations (i.e. land preparation, sowing/planting, cultural operation, harvesting and post-harvest operation) needed in crop production with respect to time/growing season. This tool contains a collection of principles to apply for efficient and sustainable farm management, and allows farmers also to choose optimal time for the operations on the ground that are yearly adjusted to local and weather conditions. The calendar keeps a crop specific information that might include also vegetation stages of the crop itself. The information about agricultural practices are inherent parts of the CbM scenarios and provide a sequence of operations likely to happen on the ground with their indicative timing. Knowledge of the agricultural practice serves as guide in the process.

Figure 2. Documentation of agricultural practice depicted by Sentinel satellites: spring barley (IE). a) Subset of Sentinel-2 images documenting land cover manifestation associated to farmer's management activities, b) Sentinel-1 6-days coherence time series signals with five activities observed on the ground (marked with vertical orange lines). Crop calendar (spring barley) and potential periods for farmer's activities (bottom),



Source: GTCAP

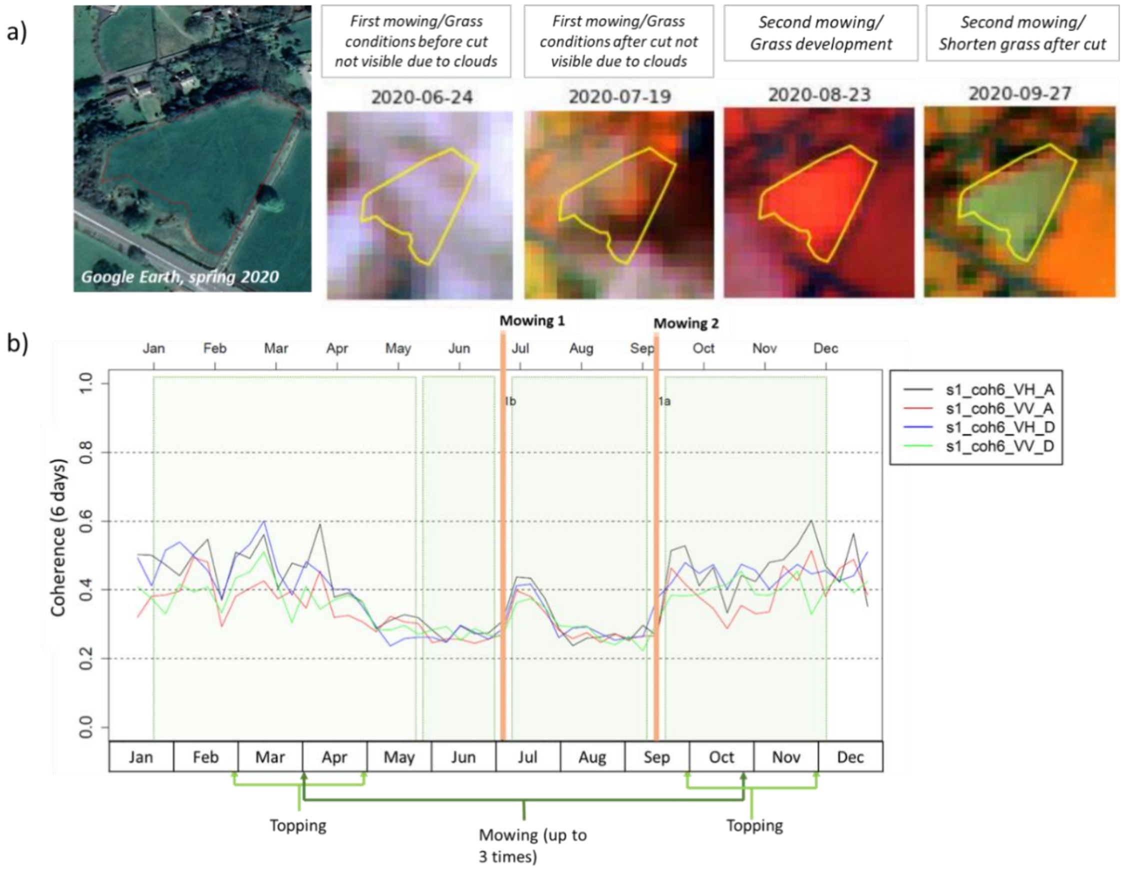
To visualise the role of the information about local practice on the declared management choices, Figure 2 and Figure 3 show examples of activities recorded in 2020 for two parcels, one on arable land and the other on permanent grassland, both located in Ireland. For these parcels a ground truth information is available with a specific date (or period) related to a specific farmer's activity. The recorded information facilitates the proper understanding of land cover manifestations, or their change (caused by farmers activities or natural events) and the link with the corresponding signal behaviour.

The first example shows spring barley followed by a winter crop in the second part of the season. In this case, it is possible to go through the seasonal cycle of a spring crop and name activities with corresponding land cover manifestations: at the beginning of the season bare soil is present due to ploughing activity, then sowing and the growth of the main crop until the harvest resulting in bare soil with stubble presence, followed by ploughing and sowing of a winter crop. From the ground perspective, the ploughing means that a soil is turned before sowing or planting to bring fresh nutrients to the soil surface while burying weeds and crop remains to decay. The land cover manifestation before the ploughing may vary from a fresh green vegetation (i.e. aftercrop, catch crop), through stubble shortly after main crop harvest, to any other state that develops on stubble left untreated before spring cultivation. The manifestation observable after this activity is uniform, bare soil with a rough surface with often-visible linear pattern (along the ploughing direction). Ploughing and cultivating soil treat the content of the upper 12 to 25 centimetres layer of soil (Schneider et al., 2017), where most plant-feeder roots grow. The uppermost soil turned by ploughing is drying out for couple of days. In case of a main crop, between ploughing and sowing, there might be other activities to flatten the soil surface for the crop bed preparation (i.e. harrowing).

The difference between a crop bed prepared and crop sown is rather difficult to spot as only the soil roughness is changed. Sowing is a process of planting seeds. When supported by machinery, it results in a pattern (rows) spaced regularly and those are well visible at the plant tillering stage. These elongated patterns unify to a homogenous green cover as the plant advances during the growing season until the ripening stage where a plant is fully-grown and ripens from green to golden colour.

Once a ripe plant is ready for harvest a collection of grain from the field is performed. During the harvest, a matured plant is cut and its residuals are both chopped and evenly distributed in the field, or the straw is windrowed, baled and removed from the field. The final stage observed after the harvested crop is stubble, where the initial sowing pattern is revealed with dried matter covering visible soil. The amount of the visible soil increases in the fields where straw residuals removed. If the stubble is left unattended for longer time, it becomes overgrown by vegetation and in this case, the appearance might change again, towards a green vegetation cover at the extreme case.

Figure 3. Documentation of agricultural practice depicted by Sentinel satellites: permanent grassland (IE). a) subset of Sentinel-2 images for documenting land cover manifestation associated with mowing activities, b) Sentinel-1 6-days coherence time series signals with two activities marked (with vertical orange lines). Crop calendar and potential periods for farmer's activities (bottom),

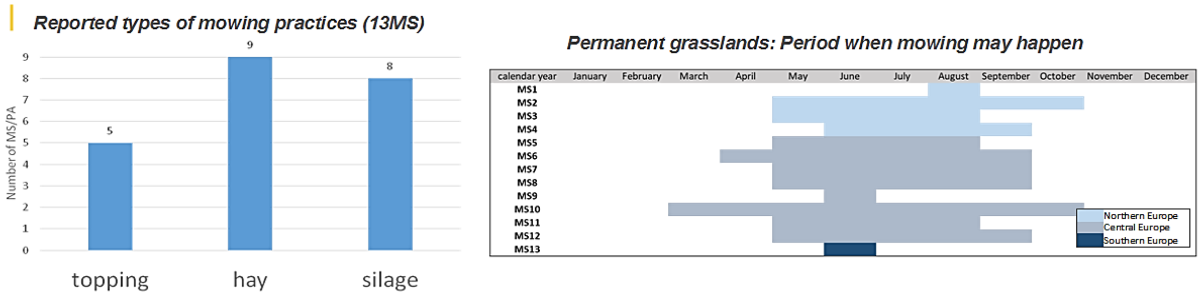


Source: GTCAP

From the agricultural practice point of view (see example in Figure 3), apart from grazing which is not discussed in this example, there are three main management activities linked with grass mowing: topping, mowing for silage and mowing for hay. The topping is the process of getting rid of the top of the plants, to induce growth and nutritional quality. The cut grass residuals are left on the ground. It can be performed at any time on pastures and often at the beginning /end of the season for grasslands. The mowing (either for silage and hay) is a process of harvesting the biomass by cutting the grass short, and collecting the residuals immediately or once they are dry. The haymaking process often includes other activities before the residuals collection like: tedding, raking, baling, which result in cut grass left on the field for several days. The distinction between these mowing types, might be helpful in understanding the land cover manifestation after the activity. Multiple activities can be performed on a single parcel in a season.

In the case of permanent grasslands, the presence of grass on a given field (i.e. at least five consecutive years) is a condition that might also be verified in the context of CbM, as continuous cover is expected. The seasonal deviation in appearance is caused by grass growth that occurs in the early vegetation season and after each mowing/grazing activity, so practically the grass length is a changing factor of the land cover manifestation during the session. The development stages vary from a primary growth stage, germination, vegetative, elongation, reproductive, and seed ripening (Moore et al., 1991). The first phases (vegetative, elongation) are characterised by a rapid plant growth and presence of a fresh and green biomass. This period is optimal for grazing and mowing for silage. Next, growth slows down at the maximum plant height, in the reproductive phase, when heading is developed. This period is considered as good time for making hay. After the seed ripening the plant dies, nutrient value declines sharply and leaves dry. Thus, in the later part of the season, dry grass is predominant (if not mowed/grazed). If the first mowing takes place at the end of the season, the yield is high, but the quality of hay for forage is low, in comparison to earlier stages. In the grasslands management a trade-off between yield quality and quantity is made. Therefore, the decision about when mowing/grazing practice should be performed, depends on individual farmer needs and local conditions, thus it is difficult to indicate/predict. An example of recorded management practices for permanent grassland is presented in the figure below (Figure 4). In ten Member States (out of 13 examined), the expected period of mowing mostly ranges from May to August, covering most of the vegetative period. The remaining three Member States indicated a single month, either June or August as the most probable period when mowing may happen.

Figure 4. Management practice for permanent grasslands: mowing types reported for 13 MS (left) and periods when mowing may take place (right).

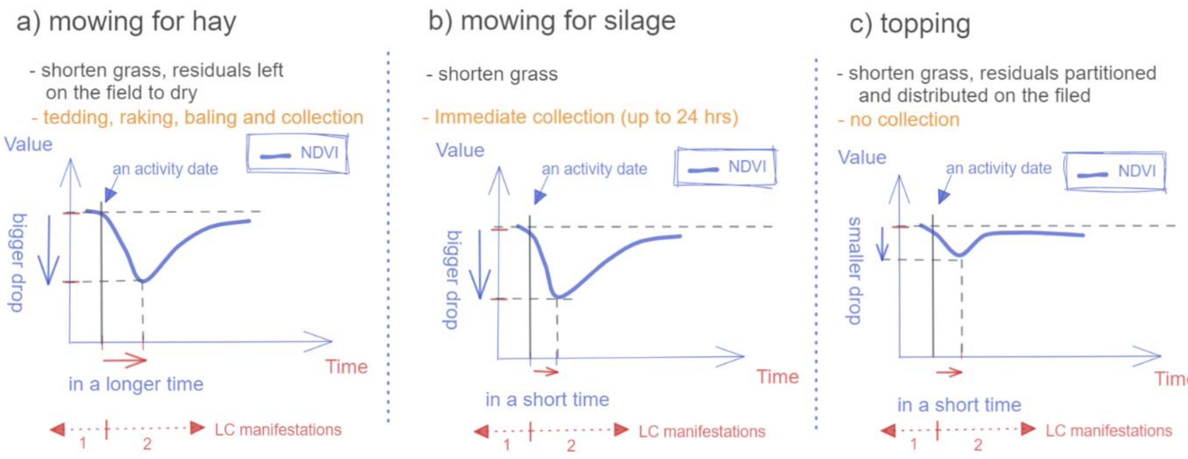


Source: (Zielinski, 2021)

In case of maintenance activities on permanent grasslands, a part of topping, mowing, grazing, there are other like muck spreading, grass residing and rolling. Some of them are manifested on the ground for a short period thus not captured by the satellite imagery extracts for a single FOI reduced to time series. An association of activity and corresponding timing together with a biomass loss provided a set of physical parameters that distinguish types of activities (Figure 5). However, there are in reality only a few types of activity (topping, mowing and grazing) that trigger a distinct change of biomass and land cover, making them monitorable with Sentinel data.

Based on the local practices, different signal behaviour is expected between three mowing types of activities. For topping the amount of the bio-mass reduced is the smallest as only the upper part of the plant is treated. Thus, only a small decrease of the signal (in a biomass monitoring signal such as NDVI) is expected, and only for a short time period, as the regrowth is stimulated by this activity. The mowing for silage reduces the grass length significantly and therefore results in a greater decrease of the signal values than topping. The plant recovery is quicker, as residuals are immediately collected after the cut thus, not preventing the development of fresh vegetation. The mowing for hay results in similar signal decrease as for mowing for silage but differs in time of plant regrowth as the grass residuals are left in the field to dry out. Thus, the signal is decreasing over a longer time period before it reaches its minimum value (Figure 5a). The information about activities and the differences between them might help in more effective marker development and parametrisation process. A lack of these considerations may result in greater omission rates in detection results (see chapter 6.2).

Figure 5. Signal behaviour of selected grassland maintenance activities (mowing for hay, silage and topping). There are two land cover manifestations marked for these activities; 1- grass cut/shorten, 2 – grass regrown.



Source: GTCAP

The mowing activity causes Synthetic Aperture Radar (SAR) coherence to increase, for all polarisations, when compared to values before the event. Movement and growth of vegetation cause temporal decorrelation that accumulates the changes captured between two acquisitions. The coherence may also decrease due to variations in the moisture content of soil and vegetation, which affects scattering centre location, as seen by SAR (Voormansik et al., 2020). This may explain the variability of observed changes between multiple observations. As also observable on Figure 2 and Figure 3, SAR coherence remains low throughout the developed phase. During the growth phase, physical and moisture changes are gradual. Similarly to mowing, after ploughing coherence increases, while NDVI records low values at the same time.

2.3 Analyses of practices – examples

Once the local information about the agricultural practice is collected together with the corresponding activities and the resulting land cover manifestations, the next would be to find optimal signal that holds enough information and variability to record the state or its change. In the Remote Sensing literature there are numerous records discussing relation between sensor-based observations and corresponding land cover manifestation, but most of the work is related to the traditional land cover or land use mapping initiative relying on nomenclatures, defined for domain-specific purposes (Devos and Milenov, 2015). However, in the CbM focus is given to agricultural practices that are either specifically requested or are banned by the payment schemes (a Member State specific ruleset) that farmer was applying for. Getting a confirmation whether a given practice is being implemented on the ground might be particularly challenging.

In the literature, there is already some significant work done in this subject. For example (Tamm et al., 2016; Voormansik et al., 2020) discuss the relation of interferometric coherence and NDVI time series to mowing and ploughing activities for CAP related applications. The ploughing and mowing activity causes coherence to increase when compared to values before an activity. In the case of mowing, as well, typical behaviour is observed, the coherence increases, while Sentinel-2 NDVI at the same time decreases.

Understanding of temporal behaviour of crops for agricultural application is discussed in (Veloso et al., 2017). In this study, a large number of temporal Sentinel-1 together with Sentinel-2 data were used to assess the potential of the Sentinel's for winter and summer crops monitoring (wheat, rapeseed, maize, soybean and sunflower). The authors point out the interest of SAR data and particularly the VH/VV ratio. A crop identification with Sentinel-1 data are presented by Beriaux et al., (2021).

Khabbazan et al., (2019) discuss the potential of Sentinel-1 SAR backscatter and interferometric coherence data for crop monitoring and the detection of key dates for important agricultural crops. Time series analyses showed that, for each of the crops considered, structural and biomass changes associated with crop development influenced the backscatter throughout the season. The VH/VV ratio proved useful as it reduces the influence of soil moisture. It is particularly sensitive to the increase in fresh biomass during the vegetative stages, and decreases during senescence as the vegetation water content decreases.

Mapping grassland management intensity using Sentinel-2 data for monitoring of agri-environmental measures is reported in (Bekkema and Eleveld, 2018). The method was developed using C5.0 decision tree classification on Sentinel-2 data. The Sentinel-2 Red-Edge Position vegetation index was found to be a good indicator of fertilization. The spectral responses of grassland types on peat and clay soils differ significantly.

Examples of work on detection of CAP related activities are already documented, including analytics and performance assessments and behavior of the signals used as markers (i.e. Beriaux et al., 2021; Lange et al., 2022; Vroey et al., 2022). Others focus on intensity analyses of land use (i.e. Bekkema and Eleveld, 2018; Griffiths et al., 2020; Lange et al., 2022; Ma et al., 2019).

Despite all development, as of today, there is no generic method that is capable to support translating the local practice, as described in CbM scenarios, into signal and select corresponding signal behaviour for marker design. Most of the solutions are based on common signals (NDVI, COH) with no attempts of discovering alternatives. Understanding the potentials and limitations of remotely sensed imagery in capturing specific phenomena in structural manner is a key element of the development. Exploration of new signals and in depth signal analyses are impossible without systematically collected ground truth data.

The ground truth-based signal analyses in the CbM may have several applications starting from:

— checking, if a land cover manifestation or its change is observable in the time series data given by a sensor/signal,

— finding the relation between a land cover manifestation and corresponding signal,

— deciding, whether a specific land cover manifestation is sufficiently reflected by a signal, or whether a given signal has provides evidence with high enough probability to become a valid marker,

— establishing the values of marker parameters, that could be as simple as change of signal value in time or expressed as more complex measures e.g. based on derivatives, or statistical trends, and

— quality assessment of marker detection performance.

The above list of applications is not exhaustive and there might be overlap between the points but it covers main areas where such analyses play importance and theirs results have a decisive role. For more information about the quality assurance aspects in checks by monitoring, where ground data play an important role, please refer to (Devos et al., 2021a)

3 Signals in the CbM

As discussed in chapter 2, a signal is derived for a specified FOI by reduction of data (pixels values) selected inside of the feature of interest outline, using a statistical data descriptor (i.e. median). The signal represents raw data values, derivatives or any combination of these. Such computation is performed for every valid acquisition in the time series. The example of signals and time series are presented in the following chapter.

In the context of CbM, Sentinel-1 and -2 dataset are considered as the main source of remotely sensed image data. Other data sources are possible, which might have similar or better characteristic in terms of temporal, spatial, and spectral resolutions.

This report does not cover the subject of large-scale automation in a signal production based on Sentinel imagery. For further information about methods and best practice solutions, please refer to the CbM toolbox source code and its relevant documentation provided by the EC JRC (Lemoine and Anastasakis, 2022).

3.1 Sentinel-1

The Sentinel-1 mission is based on a constellation of two satellites (A and B units) to achieve a short revisit time (Torres et al., 2012). Both satellites carry a C-band Synthetic Aperture Radar (SAR) and are operated in monostatic mode and are flying in a sun-synchronous orbit at an altitude of about 700 km. The front end of the instrument is based on an active phased array antenna driven by 280 transmit/receiver modules for each polarization and enabling electronic beam steering over a wide range of swath positions (up to 400 km ground range). There are four exclusive imaging modes of operations: Interferometric Wide Swath, Extra Wide Swath, Strip Map, with 6 possible incidence angles and Wave Mode (ESA, 2021). The first three modes can be operated in 4 different schemes of polarisation (2 in single and 2 in double): horizontal HH, vertical VV, HH+HV or VV+VH. The Wave mode can operate only in single polarisation, either in HH or VV. Overall this represents 34 possible sub-modes of operations resulting in different products and resolutions[1]. For example, the single polarization VV means that the product is obtained with a vertical transmitting and vertical receiving schemes. The product is obtained also for ascending and descending orbits.

Over land, the same SAR polarisation scheme is systematically used over a given area, to guarantee time series of data in the same conditions for routine operational services and to allow frequent interferometric SAR. Depending on the area, the selection is either vertical or horizontal, the choice being made according to the main application behind. As a general principle, the polarisation scheme uses the following logic:

- HH-HV or HH polarization for the monitoring of polar environments, sea-ice zones
- VV-VH or VV polarization for all other observation zones (with an exception for the Baltic Sea observed partially in HH-HV with Sentinel-1B during northern winter).

At the end of 2021 Sentinel -1B encountered power problems[2], thus the satellite is not operational since Dec. 23, when ESA noticed an issue with the system and switched off. With Sentinel -1B offline the repeat cycle of the radar system decreased from six to 12 days. As the 1B mission officially ended (03/08/2022)[3], the 6-day coherence product is not available for any applications, including CBM, until a new satellite (1C) is operational. The launch is currently scheduled for first half of 2023[4].

For specific details about the high level operations concept and strategy of the Sentinel missions please see the Sentinel High Level Operations Plan (ESA, 2021).

Synthetic Aperture Radar (SAR) has the advantage of operating at wavelengths not impeded by cloud cover or a lack of illumination and can acquire data over a site during day or night under all weather conditions. SAR observations are key for operational applications over oceans, seas, polar areas, as

1 Copernicus https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/instrument-payload/resolution-swath

2 ESA News https://scihub.copernicus.eu/news/News00980?s=09 (accessed 03/11/2022)

3 ESA News https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1/Mission_ends_for_Copernicus_Sentinel-1B_satellite (accessed 03/11/2022)

4 ESA News https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-1/Ride_into_orbit_secured_for_Sentinel-1C (accessed 03/11/2022)

well as, are also used for land applications (i.e. agriculture) and provide data for emergency response and security, in particular under adverse weather conditions (ESA, 2021).

The CbM signal may be constructed in many ways, either as a simple product, two polarizations difference, two polarization ratios, and normalised difference of two products or with more complex mathematical construction. The Table 1 shows a list of selected SAR-based products that are potentially useful in a vegetation and soil context.

Table 1. Examples of Sentinel-1 derived values and derivatives.

Index	**Full name**	**Equation**	**Reference**
BS *	Backscatter	See the following: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-1/data-products (accessed 03/11/2022)	(i.e. Bauer-Marschallinger et al., 2021)
COH *	Coherence	See the following: (Nasirzadehdizaji et al., 2021)	(Ferretti et al., 2007)
RVI	Radar vegetation index	See the following implementation: https://custom-scripts.sentinel-hub.com/custom-scripts/sentinel-1/radar_vegetation_index/ (accessed 03/11/2022)	(Szigarski et al., 2018)
NRPB	Normalised ratio Procedure between Bands	(S1_bs_VH – S1_bs_VV) / (S1_bs_VH + S1_bs_VV)	(Filgueiras et al., 2019)
CPR	Cross polarization ratio	S1_bs_VV / S1_bs_VH	(Vreugdenhil et al., 2020)
CPRI	Inverse cross polarization ratio	S1_bs_VH / S1_bs_VV	(Vreugdenhil et al., 2020)
S1-AD-ratios**	Ratios between A and D orbits for S1 backscatter and coherence	i.e. S1_coh6_VV_A/ S1_coh6_VV_D	
S1-AD-N-ratios**	Normalised ratios between A and D orbits for S1 backscatter and coherence	i.e. (S1_bs_VH_A - S1_bs_VH_D)/(S1_bs_VH_A + 1_bs_VH_D)	

* For coherence and backscatter, we are looking at six signals. Four of them are combinations of polarization (VH and VV) and orbit (ascending (A) and descending (D)), two are averages for each polarization (VH and VV) and an average for all combinations.

** For the ratios and normalised ratios, there are three signals for each of backscatter and coherence. One for each polarization (VH and VV) and one for the average of the two polarizations.

Source: GTCAP

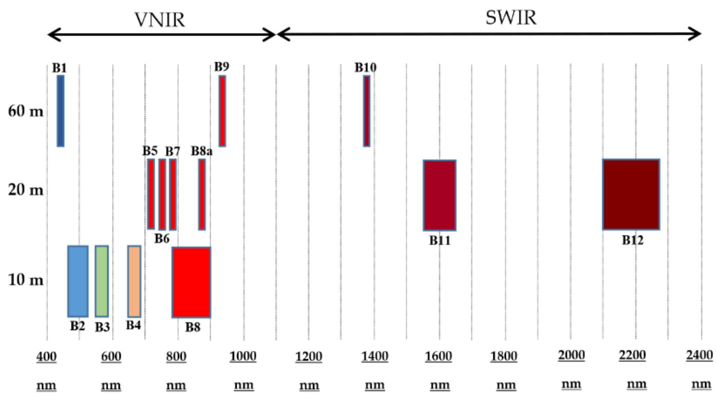
3.2 Sentinel-2

The Sentinel-2 mission provides enhanced continuity to services relying on optical multi-spectral high spatial resolution observations over global land and coastal regions. The mission is based on a constellation of two satellites (A and B units). The combination of the large swath (290km), spectral range (13 coupled with the global and continuous acquisition requirement with high-revisit frequency, leads to the daily generation of 1.3 TB of orthorectified top-of-atmosphere reflectance products (L1C) per satellite unit. This corresponds to an average continuously sustained raw-data supply rate of 160Mbps (ESA, 2021).

The Sentinel-2 Multi-Spectral Instrument (MSI) is a filter-based push-broom imager. It collects data in 13 spectral bands spread over the visible to the short-wave infrared range of the electromagnetic spectrum, with spatial resolutions ranging from 10 to 60 m (Figure 6). For more information about the Sentinel-2 products and validation protocols please refer to (Gascon et al., 2017).

The Sentinel-2 based CbM signal can be constructed in many ways, ether as a simple band reflectance, two bands difference, two bands ratios, normalised differences of two or more bands, or with more complex mathematical constructions. Whereas the band reflectance is inherited from the sensor and cannot be changed or improved, the derivatives can be developed for a specific application or phenomenon. There are many band derivatives or indices available, and extensive selection can be found in an online database of indices (Henrich et al., 2009) (IDB: www.indexdatabase.de accessed 03/11/2022) including the choice of wavelengths and coefficients depending on the selected sensors or applications. The IDB is extensive resource, but the literature may also provide with others not included. The database for Sentinel-2A sensor provides about 250 different indices.

Figure 6. Sentinel-2, MSI spectral bands vs. spatial resolution.



Source: (Gascon et al., 2017)

Vegetation Indices (VIs) obtained from remote sensing imagery are quite simple and effective algorithms for quantitative and qualitative evaluations of vegetation cover, growth dynamics, among other applications. The study of (Xue and Su, 2017) introduces the spectral characteristics of vegetation and summarises the development of VIs and their advantages and disadvantages. This paper reviews more than 100VIs, discussing their specific applicability and representativeness according to the vegetation of interest, environment, and implementation precision. For further reading, see also (Sishodia et al., 2020). Table 2 shows a list of selected spectral indices used in a vegetation and soil analysis context.

Table 2. Example of spectral indices derived fromSentinel-2 data.

Index	Name	Formula	Reference
NDVI	Normalised difference vegetation index	NDVI=(B08-B04)/(B08+B04)	(Rouse et al., 1973)
DVI	Difference Vegetation index	DVI = B08 – B04	(Tucker, 1979)
GNDVI	Green normalised Difference vegetation index	GNDVI = (B08 – B03) / (B08 + B03)	(Gitelson et al., 2005)
GSAVI	Green soil adjusted vegetation index	GSAVI = 1.5*(B08-B03)/B08+B03+0.5)	(Sripada, 2005)
GARI	Green atmospherically resistant vegetation index	GARI = (B08-(B03-1.7*(B02-B04)))/ (B08+(B03-1.7*(B02-B04)))	(Gitelson et al., 1996)
GARI2	Green atmospherically resistant vegetation index - alternative	GARI2 = (B08 - (B03 - (B02 - B04))) / (B08 - (B03 + (B02 - B04)))	(Sonobe et al., 2018)
NDWI	Normalised difference water index	NDWI = (B08-B11)/(B08+B11)	(Gao, 1996)
NDPI	Normalised Difference Phenology Index	NDPI=(B08-(0.74*B04+0.26*B11)/ (B08+(0.74*B04+0.26*B11)	(Wang et al., 2017)
GLI	Green leaf index	GLI = ((B03 – B04) + (B03 – B02))/ (2*B03 + B04 + B02)	(Louhaichi et al., 2008)
EVI	Enhanced vegetation index	EVI = 2.5*(B08 – B04) / ((B08 + 6*B04 -7.5*B02) + 1	(Huete et al., 2002)
BSI	Bare soil index	BSI = ((B11+B04)-(B08+B02))/ ((B11+B04)+(B08+B02))	(Diek et al., 2017) (Nguyen et al., 2021)
SAVI	Soil adjusted vegetation index	OSAVI = (B08-B04)/B08+B04+0.5)	(Huete, 1988)
OSAVI	Optimised soil adjusted vegetation index	OSAVI =(B08-B04)/B08+B04+0.16)	(Rondeaux et al., 1996)
MSAVI2	Modified soil adjusted vegetation index 2	MSAVI2 = (2*B08+1 – SQRT((2*B08+1)^2 -8*(B08-B04))) / 2	(Qi et al., 1994)

Source: GTCAP

The presence of cloud cover on one or more dates can be is a major limiting factor in exploiting time-series data acquired by optical space- and air-borne sensors.

The Sentinel’s data quality and availability are not spatially and temporally homogeneous due to effects related to cloudiness, varying acquisition frequencies at different geographic locations, overlapping acquisitions and the acquisition plan (i.e. Serco, 2021). The spatio-temporal inhomogeneity of the underlying data may therefore affect any analysis and is important to consider before designing the algorithm and interpreting the results (Sudmanns et al., 2020). Potential application of the same method on data located in northern, central or southern part of Europe (with different average cloud cover or acquisition frequency) might give different results. Therefore an analysis of the availability of scenes, average cloud cover and cloudiness might be crucial (i.e. (Sudmanns et al., 2020)).

An automated cloud and cloud shadow detection is a key component of the processing needed to prepare optical satellite imagery for scientific analysis. There are several algorithms available for Sentinel-2 cloud detection which were a subject of performance comparison to understand which algorithms best identify clouds and their shadows in images (Skakun et al., 2022; Tarrio et al., 2020). The Tarrio et al.(2020) examines relative performance of five different cloud-masking algorithms (Sen2Cor, MAJA, LaSRC, Fmask and Tmask) and provides an accuracy assessment that show a trade-off between omission and commission errors in cloud detection for individual algorithms. No single algorithm outperforms the others for both clouds and shadows, some algorithms are better at detecting either clouds or cloud shadows. The authors suggested an ensemble approach, aggregating the results from multiple algorithms to provide fewer undetected clouds and higher overall accuracy than any single algorithm, which may be the most useful for processing of Sentinel-2 data. Skakun et al., (2022) evaluates ten different algorithms that vary in their approach and concepts adopted. They were based on various spectral properties, spatial and temporal features, as well as machine learning methods. As part of the overall performance analyses (across algorithm, overall accuracy varied from 80.0 - 89.4% for Sentinel-2), the strengths and weaknesses of existing algorithms and potential areas of improvements are identified.

The performance result might help in selection of the most suitable algorithm, but consideration of the strengths and weaknesses helps in understanding and practical usage in CbM application.

This report deals mainly with the Sentinel-based signals, but there are also other sources of data that might be considered or already used by the Member States in their development work or national CbM implementations. An example is the use of geotagged photos in the CbM application. Further details are given by Sima et al. (2020).

3.3 Signal descriptor

In the CbM, the signal is calculated by applying different statistical data descriptors on the values within a polygon or set of pixels that are representative for the FOI. The selection of values to be used in the calculation is designed to capture any spatial aspect of the phenomena within the FOI. The feature of interest (FOI) is the main spatial element in the context of CbM, and it refers to a land surface where a specified practice is performed (Devos et al. 2021).

As a consequence, the CbM signal should not be considered as a derived DN value associated with an individual pixel tracked in time, but as result of the reduction of the image data stack targeting a single surface (the FOI) and represented as a function of time. This means that any spatial variability/pattern reflected by the individual image pixels at given time, is "reduced" to a single value at the level of the FOI. The ability of this value to reflect the spatial variability/pattern of the signal within the FOI, depends on the type of descriptor used – standard deviation, inter quartile range or more complex one, such as moments or fractal dimension. Obviously, the quality of the signal depends on the correctness of the FOI.

We can start by assuming that there is a parcel homogeneously used by farmer under a single practice. This parcel is for the CbM needs defined by the FOI boundary and a set of statistics is calculated based on the corresponding satellite data pixels for each valid acquisition. The term “valid acquisition” refers to the result of pre-processing of satellite data, at the FOI level, and application of defined protocols, for instance for the Sentinel-2, exclusion of acquisitions where presence of clouds or cloud shadows is detected. This step is particularly important because the presence of clouds (very high values) or cloud shadows (very low values) on the image and thus potentially inside the parcel strongly influences the FOI statistics. This influence grows with the proportion of unwanted (cloudy) to wanted (cloud-free) pixels within the parcel boundary for a single image acquisition. Presence of such artefacts leads to outliers in the time series, and may result in doubtful and erroneous interpretation of it. For example, while looking for mowing activities on grassland, inclusion of cloud-shadowed pixels in the statistics calculation results in a drop of the Sentinel-2 signal value in the time series, which might be wrongly interpreted as a mowing event.

To facilitate a time series analyses, the values of full image pixels for a single band located within the FOI boundary are selected. These are reduced to a single value with the help of the descriptive statistics, which summarise the enclosed data set. From a statistical view point, the data set can be either the entire population or a sample of a population (Mann, 2018). The descriptive statistics can be grouped in measures of central tendency and measures of variability. The measures of central tendency include the mean, median, and mode, while measures of variability include standard deviation, variance, minimum and maximum variables, kurtosis, skewness and interquartile range (IQR). In this report, we are looking at the 18 descriptors, listed in Table 3.

Table 3. Overview of statistical data descriptors used in the analyses in this report. The descriptor is applied on the set of full image pixels, located inside FOI, to calculate the resulting signal.

Short name	Description	Short name	Description
Min, P25, P50, P75, Max	a value for which 0, 25%, 50%, 75% and 100% of the pixels have a lower value	Skewness	Refers to asymmetry of pixel values distribution
Mean	The average of all pixel values	Skew.2SE	Standard deviation of skewness
SE.mean	The standard deviation of the mean estimate	Kurtosis	Descriptor describing the flatness/sharpness of a histogram
CI.mean	The 95% confidence interval of the mean estimate	Kurt.2SE	Standard deviation of skewness
Var, std.dev	Variance and standard deviation	Normtest.W Normtest.P	The test statistic W and the P-value P of Shapiro-Wilk test of normality of the data
Coef.var	Coefficient of variation, being the standard deviation divided by the mean	Entropy	Shannon's entropy is a measure of randomness in a system

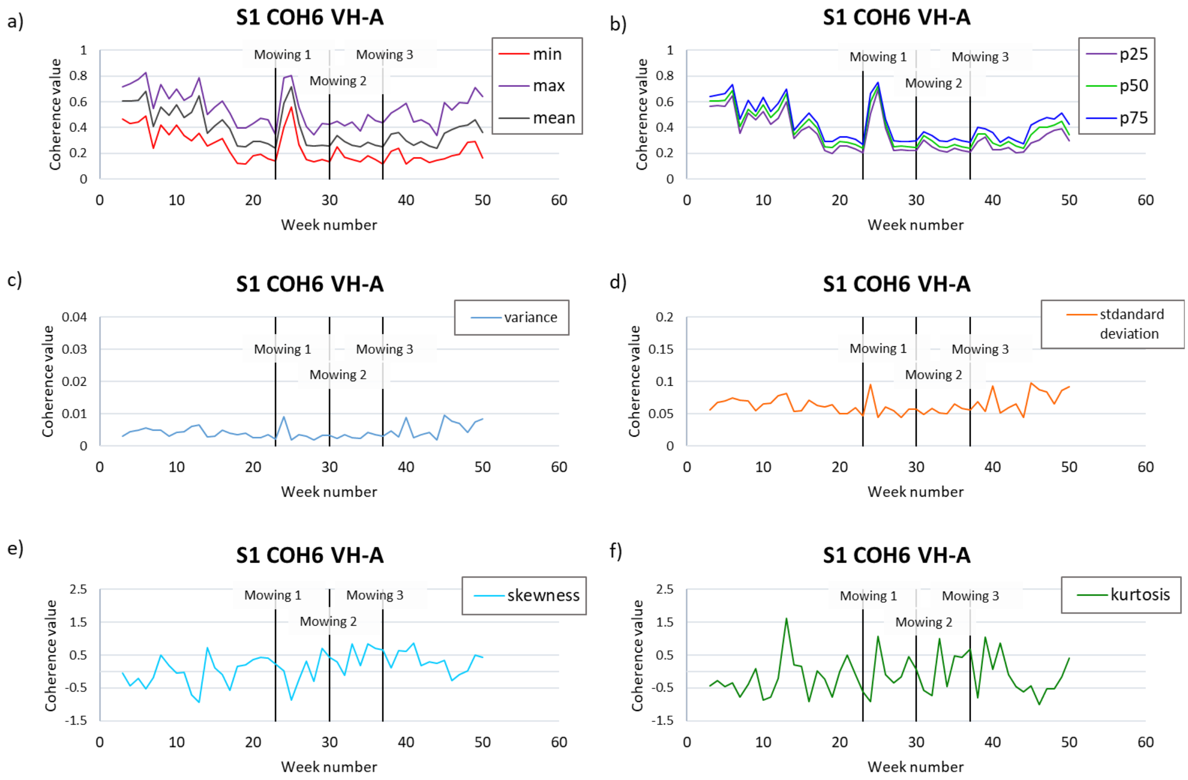
Source: GTCAP

The number of image pixels that are within the boundaries of a FOI will depend on a combination of FOI size and shape together with image resolution (i.e. Sentinel-2, ground sample distance of 10m). More complex shapes of parcels (especially elongated) will result in fewer pixels selected in comparison to a square/rectangle shaped parcel. The shapes and sizes of parcels to be monitored are influenced by local landscape characteristic and historical agriculture patterns.

The number of pixels available per FOI influences the calculation and reliability of the statistical descriptors in the CbM reductive approach. Formally, most of the descriptors can be calculated based on only two pixels, some of them even from one pixel. However, there is a risk that random errors will dominate when the number is low. The minimum number of pixels is a balance between data validity and having enough FOIs for analyses. If a region has an abundance of FOIs with less than 3 pixels available for the monitoring, these FOIs should be separated from the rest, and the estimation of typical signal behaviour in the system should be done separately, as random errors are more likely to affect their behaviour.

Figure 7 shows several time series of different statistical data descriptors applied on one Sentinel -1 product (6-days coherence with VH polarisation acquired from ascending orbit) for the same grassland parcel/FOI, together with lines representing ground truth observations of mowing activities (see chapter 4 for further information). The changes of signal values of mean, minimum and maximum as well as the quantiles p25, p50 and p75 of (Figure 7a, b) confirm the behaviour expected in a presence of mowing during the season. Before the event, the values decrease, to increase significantly after the event. The magnitude of changes is different among the following activities. An interesting pattern is observable for the first event (mowing for hay) with respect to the standard deviation and variance where its value increases during the mowing, followed by windrowing and collection. The changes in skewness and kurtosis values around the mowing are not sufficient to support the events detection because the observed patterns do not allow to link them directly to the farmer's activities in the field.

Figure 7. Examples of signal's time series based on Sentinel-1 6-days coherence with VH polarization (ascending orbit) calculated for a single grassland FOI with 10 descriptors, a) minimum, maximum and mean, b) quantiles: P25, P50 and P75, c) entropy, d) standard deviation e) skewness f) kurtosis. Three mowing activities observed on the ground are marked as black horizontal lines.



Source: GTCAP

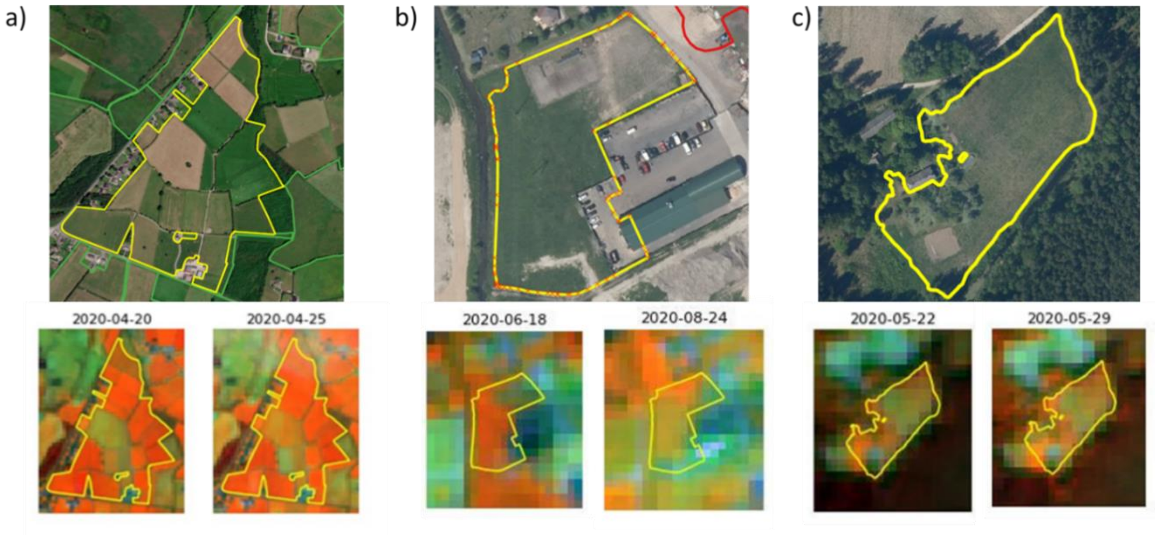
This report does not discuss the potential approaches for optimal data selection inside the feature of interest for the CbM application. There are, however, some pending questions concerning the optimal pixel selection approach, as well as the validation procedure of the FOI. For the status and progress report on the CbM FOI validation subject, please refer to (Milenov P., 2021).

4 Data used and data preparation

The data used to document the analysis and explain the proposed approach are subsets of real datasets available in the Member States and shared in the frame of the CbM outreach 2021 initiative.

The data consists of a set of parcels/FOIs and corresponding information about the ground truth observations from three Member States together with information about local practices. As the FOI geometry is based on Geo-spatial Aid Application (GSAA) and is not a subject of any verification, the selected FOIs were visually checked for correctness, and, if needed, the boundaries were improved to reflect the area under a single practice. Potential inaccuracies in boundary placement (Figure 8), and exclusion of ineligible features was performed to make sure that the analysed signals were free of such influences. Further details are given in chapter 6.3.

Figure 8. Examples of three features of interest potentially requiring boundary improvements. The agricultural parcels (considered as FOIs) marked in yellow are overlaid with Google Earth imagery and Sentinel-2 data acquired in the same season: (a) case of several practices within one FOI, (b) change of land use, non-agricultural area inside, (c) different management practices within the same FOI.



Source: GTCAP

Because the FOI defines the signal input of all analyses, not even the ‘smartest’, sophisticated analytical approaches will provide proper results if the FOIs are not correct. Therefore, validation and analyses of the FOI cardinality are fundamental boundary conditions of the CbM. It is a key to understand the relationships between the land cover manifestation and the signal as well as to the performance of the systems. For more information on this subject please refer to (Milenov et al., 2021).

The Sentinel-2 image data were checked for the presence of clouds/clouds shadows. We have discarded the observations for all days where at least one pixel in the FOI has been classified as cloud or shadow, as it is likely that also more pixels might be impacted. We have also limited the analyses to FOIs where at least 3 pixels per FOI were available for all Sentinel products, to balance the number of valid FOIs and data validity, as discussed in chapter 3.3. The ground truth data enables the establishment of a link between the farmer’s activity date and the corresponding changes in the signals. The existing observations available at the Paying Agencies (a collection of on-the-spot-checks, rapid field visits, and geo-tagged photos captured during various controls) shared within the Outreach project, are not always fit-to-purpose for signal analysis. The data vary with respect to the completeness of the information, about the activity performed as well as in relation to the typology of the dates recorded.

Among the three datasets presented in this report, one is based on the most complete dedicated survey, and the other two are based on existing data available at the Paying Agencies. For each data set, 60 FOIs with corresponding observations were selected.

The first dataset is from Ireland (for details about the survey see [5]), collected in a dedicated survey covering entire season with a weekly revisit interval. Such approach might be considered as an example of best practice in collecting ground data for the CbM purposes. The other two datasets are located in Czechia and Latvia, and contain observations about the activity type (only mowing) with observation frequencies corresponding to the ones of a typical OTSC setup. In these cases the date of activity was not always available directly, and data needed an enhancement based on auxiliary sources. Such input data enhancement can be done using the time series and known behaviour (i.e. from literature, like the usage of coherence for mowing Tamm et al., 2016) or other data sources like satellite/aerial orthoimagery, oblique images acquired by unmanned aerial systems or geo-tagged photos. For more discussion about ground observation, see chapter 7.

To facilitate systematic documentation and information exchange, a CbM documentation template was recently proposed (Zielinski et al., 2022a). This document cover all the design components and aim for a uniform description of information held by Paying Agencies' (PAs) regarding the system implementation including practices and relevant requirements. It was developed to allow for transparent, systematic, and structured documentation of the key elements of the monitoring systems relying on earth observation data to facilitate information exchange between different stakeholders. Understanding the farmer's activities on the ground is fundamental to make proper design choices for monitoring system.

However, it is important to remember that the two enhanced datasets have different degree of uncertainty in comparison to the data resulting from a dedicated ground survey, thus direct comparison of the result should be done with caution.

The absolute values presented in the analyses are calculated based on a given sample set, not from the full FOI population in the country, and should not be understood as representative for an entire territory of a Member States (it was not an objective of the project).

It is also important to underline here that this development work was possible thanks to Member States collaboration and their willingness to share the data.

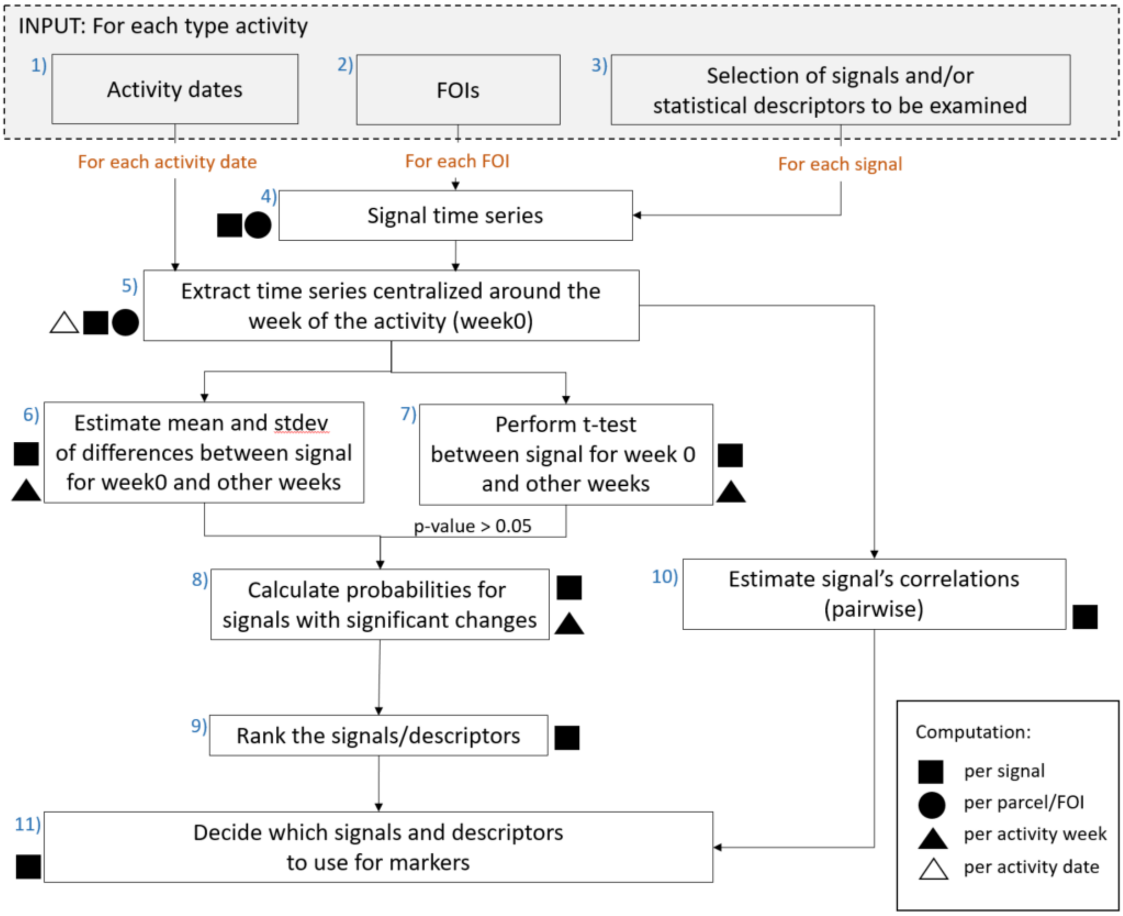
[5] Reported by Eoin Dooley, DAFM, Ireland https://marswiki.jrc.ec.europa.eu/wikicap/images/1/1b/IE_Monitoring_%26_AG_Practices_05_04_22.pdf (accessed on 03/11/2022)

5 Signal analyses for land monitoring systems

As discussed in chapter 2, it is assumed that for each scenario a list of activities resulting in potentially observable land cover manifestations is given. Next, the signal behaviour is analysed to understand and describe the relationship between signals and land cover manifestations on the ground. As a result, for each activity a list of candidate marker (mere observation of the signal value or its evolution in time) is created.

To accommodate versatility of this approach we start with no prior setup of the pre-, mid- and post-conditions on bio-physical aspects on the ground as discussed by Zielinski et al (2022a). These conditions might be introduced at the later stage of analyses. The proposed signal analyses workflow is shown in Figure 9. This method is not limited to any specific activity or land cover manifestation and can be applied at any data location. A single activity type or land cover manifestation can be analysed when the date of ground events are known. The number of signals that can be analysed for a single activity in question is also unlimited.

Figure 9. Proposed workflow of signal analyses to derive a ranked list of signals optimised to detect land cover manifestations and their corresponding activities.



Source: GTCAP

To analyse a single activity type (or land cover manifestation) the methodology may be summarised in the following steps:

Input data

1. **Activity dates**: the date of when an activity took place is a main information around which the analyses are centred. This information makes it possible to group multiple observations in a common workflow.

2. **The feature of interest (FOI)** provides the geographical boundary for which corresponding image pixels are selected, and those pixel values are used for time series computation using statistical descriptors. The content of a FOI should correspond to a single practice. This condition needs to be verified before the analyses.
3. **Signal(s)**: this is a list of one or more signals and statistical descriptors to be computed for an FOI and analysed. Depending on the scope of the analyses, this selection might include hundreds of signals, which can be analysed for a limited number of parcels. The main selection should be based on remote sensing knowledge with signals that reflects the activity, land cover manifestation needs or characteristic, to optimise the time needed for processing and analysing the results.

Analyses

4. **Extract signal time series**: extraction of signal time series is the first part of data preparation process needed for analyses. In this case, the time series are time ordered signal values extracted for each combination of FOI, signal and signal descriptor for every satellite acquisitions in a selected period (i.e. calendar year). The signal is calculated by applying different statistical data descriptors on the pixel values within a polygon or set of pixels that are representative for the FOI. The selection of pixel values to be used in the calculation is designed to capture any spatial aspect of the phenomena within the FOI. For more information about this process and best practices please consult the JRC CbM tools and corresponding documentation (Lemoine and Anastasakis, 2022). This is covered in chapter 5.1.
5. **Extract activity-centred time series:** by centring subsets of the time series around the activity dates, observations on multiple FOIs can be used in a single analyses flow. In this step, time series data are ordered in a weekly pattern, centred around the date of activity (indicated as a day of activity). The weekly pattern might increase the number of comparable valid Sentinel observations (i.e. only one valid acquisition per week necessary, reducing gaps due to clouds). The week 0 is the week when the activity took place in the field, week -1 is a week preceding the activity and week +1 is the week following the activity. This is covered in chapter 5.1.
6. **Estimating mean and standard deviation**: in order to understand how the signal behaves in relation to an activity, all the activity centred observations for the particular type of activity are analysed to visualise the potential change in the signal recorded before and after the activity took place. This approach shows the signal behaviour expressed as a mean value of the signal in the following weeks after the observed activity and also provides information about variability of observations. This is covered in chapter 5.1.
7. **Perform t-test**: using the activity centred signals, a statistical test (T-test) is used to verify whether the change between the week of the activity and the weeks before/after is actually significant. The test detects significant differences of the mean between two samples, to test if the average change of signal value after the activity is significant. Only signals with significant change are considered ($p > 0.05$) in the next step. This is covered in chapter 5.2.
8. **Probability of signals**: once the significant change of a signal for an activity type has been identified, the next step is to analyse whether this change is detectable, and which of the signals could be the most suitable to depict it. The type of activity observed in various parcels/FOIs, even from the same region, exhibit a strong variability. Thus, an important aspect of the signal behaviour is its repeatability, i.e. that the observed post-activity signal behaviour is visible also for other FOIs. For each signal, we have estimated a mean and a standard deviation for the changes for each week (7). Then we can calculate the probability of the difference values in a given week being greater/smaller than zero (greater: for increase of signal, smaller: for a decrease in signal after the activity). When a high probability is observed in consecutive weeks, such a signal is a good candidate to provide observation (marker) in the CbM context. This is covered in chapter 5.3.
9. **Rank signals**: after analysing several signals for a given land cover manifestation or its change caused by an activity, it is then necessary to summarise these results and ideally, to rank the signals from potentially useful to the ones that should be omitted in further considerations. To conclude, the probabilities and significance measures are summarised

for all tested indices and descriptors in chapter 5.4 and with some additional analyses in 6.1.

10. **Correlation analyses:** it is assumed that a marker derives from one signal, but a signal could source many markers. But we can also have several markers based on several different signals. A phenomenon or land cover manifestation observed by multiple markers reinforces the performance reliability only if the extracted information is not strongly correlated. In a case of strong correlation between the signals used to observe the same land cover manifestation state/change, the reliability reinforcement is minimal. This step can be considered as optional, and helpful if there is an in interest of defining several markers to detect the same activity (and thus, strengthening the outcomes of a switch in a lane). In result, a pairwise correlation measure is provided for each pair of the signals and visualised graphically, covered in chapter 5.5.
11. **Select signals for marker:** in the last step, the results from point 9 and potentially 10 are analysed to select one more than signals with a behaviour that was optimal for given activity, and which at the same time have limited correlation with each other between each pair. If it is not necessary to select several markers to detect the same activity, then only result from point 9 are used for the selection.

5.1 Analysis of signal behaviour

The analyses should be arranged according to the temporal resolution of available ground observations. Assuming that the dates of activities are available (ground observations) for the same type of activity observed on multiple FOIs, these observations can be used in a single analyses flow using a specific data arrangement that relies on the exact day of the activity. In this case, the value for the week before the activity will be calculated based on the acquisitions up to seven days before the activity date. The value for the first week after the activity will be based on the acquisitions up to seven days after the activity took place. This is similar to what was proposed by Voormansik et al (2020).

For the ground data observed in a weekly pattern assumes availability of every week field observation. In this case the signal values are arranged for each week, following the International Organization for Standardization (ISO) week approach[6]. The values for the activity week will be based on acquisitions for this week, and similarly for the weeks before and after the activity week. These observations can be used in a single analyses flow using a specific data arrangement presented in Figure 10.

The data are managed in a weekly pattern based on the optimal ground truth data frequency (see chapter 7). The weekly layout has a number of advantages:

— the averaging tends to smooth the signals and reduce effects of artefacts;

— the analyses limits the effect of irregular or missing data, as only one valid acquisition is necessary per week, and makes it easier to group and compare values for the same period;

— analyses based on the acquisition dates would require some kind of interpolation between acquisitions, as the acquisition dates after an activity would vary between the activities.

The weekly approach might be reconsidered:

— for cases where temporal resolution cannot be reduced, although for most of the datasets related to agriculture, this is not an important issue. Most land cover manifestations (or changes of them) of interest have a timespan longer than a week, limiting the need for sub-weekly data. In other words, a weekly signal sampling rate is sufficient to capture the activity through the behaviour of the affected land phenomenon. Otherwise this should be adjusted for specific cases;

— when the exact day of the activity is unknown, averaging observations into a single value may result in a combination of pre- and post-activity acquisitions for the activity week, as the dates of the acquisitions are not synchronised with the activity date. As a part of the change is already incorporated in the signal, the pre-activity signal value will be further away from the date of the activity.

For both cases, to find the changes in signal value as a result of the activity, it is necessary to compare the values after the activity with the value from the week before the activity took place.

[6] https://en.wikipedia.org/wiki/ISO_week_date

Following the figure below the *week 0* is the week when the activity took place in the field, *week -1* is the week preceding the activity and *week +1* is the week following the activity. The signals are arranged according to this pattern. For each activity a date and type of activity is prepared. However, in the analysis we are looking at changes between *week –1* and the weeks after the activity. The reason for this is that the signal from the acquisitions of *week 0* might be based on a combination of acquisitions before and after the activity. Hence, we chose to analyse the signal of the week before the activity, and refer to this week as week 0 further down in this report. In the rest of the report, for simplicity, we will refer to the week before the activity as *week 0*, and the week values in the figures will show zero for this week and 1 for the week of the activity.

Figure 10. Weekly data arrangement for ground truth analyses.

Source: GTCAP

We can distinguish between three principal post-activity signal responses:

— increase of the signal observed after the activity date (i.e. case of mowing observed in Coh6 data- see example chapter below),

— decrease of signal observed after the activity date (i.e. case of mowing observed in the NDVI data – see example chapter below)

— no change, which might lead to a conclusion that a signal does not capture the given land cover manifestation (i.e. case of mowing observed on S1 - backscatter – see example chapter below)

For some activities where the entire land cover or phenomenon is changed rapidly (e.g. ploughing) the change of visual/physical state is visible without delays. In a case of mowing (topping, silage, hay) the observable change between before/after the activity performed might be visible after a few days (especially in optical data), when the grass residuals start to dry out. In the case of mowing, the visual effect grows proportionally with the amount of biomass cut. For example, light topping activities might not be even visible, whereas the first mowing in the season, with a maximised hay harvest, clearly stands out due to significant biomass reduction and increased proportion of soil contribution in the response (after the activity). There are some activities (e.g. spraying) that are not easily observable due to no sufficient change in bio-physical manifestation and limitation in the sensor spectral/spatial resolution or satellite revisit time.

The state of a land cover manifestation affects the signal shape and magnitude; these values are signal dependent and are spatially variable as a result of geographic differences and different practices in different MS. Those properties can be estimated for a MS by analyses of the signals together with ground truth data.

In order to understand how the signal behaves in relation to an activity, a single (i.e. Figure 2 - Figure 3) or multiple observations are analysed. In comparison to a single FOI analysis, overlaying multiple observations of the same activity type (i.e. mowing) supports the visualization of the potential change in the signal recorded before and after the activity took place. This approach does not only show the signal behaviour expressed as a mean value of the signal in the weeks following the observed activity but also provides information about variability of observations. A spread of the values in the corresponding weeks may be analysed. However, also individual time series should be visualised together with activity dates, as a first indication of the correctness of the data and their representability for defining a typical behaviour.

Here are some initial considerations when analysing multiple observations (weekly pattern) of the same activity recorded in different parcels/FOIs:

— The data collection frequency: ground truth collected once a week, where a person is present on the spot to record the status of a crop and information about the farmer's activities in the field since the last inspection. Effectively, the longest time between a potential activity that happened on the ground and its record by the inspector may be 7 days.
— Sentinel data acquisition: data from Sentinel (1 and 2) have a certain temporal resolution, depending on the revisit time (ESA, 2021). In addition, the gap between two consecutive acquisitions is dependent on the geographical location, as some of the orbits may be partly overlapping (ESA, 2021). The optical sensor acquisitions are subject to presence of clouds and cloud shadows, and thus the input data to the process are even more irregularly spaced in time. In Europe, there are locations (e.g. Ireland) where, during the May-August period, only a few valid acquisitions are left after the cloud filtering. Therefore, such a signal may become unreliable in the context of the time series analysis.
— Time stamp of signal: another important element is to understand a signal in terms of content and time relevance. For example, Sentinel-1 based coherence (COH6) signals are calculated from a pair of consecutive acquisitions with 6-days lag time (Ferretti et al., 2007). Therefore, the observed signal behaviour should be interpreted accordingly, and not as a single date acquisition-based signal.

Signal behaviour: example

For a set of parcels/FOIs located in Latvia with known date of mowing activities, a number of median signals were extracted from Sentinel-1 and -2 data (Figure 11), including: a) S1- backscatter, b) S1 6-days coherence, c) S2 – Red band (B04), d) S2 – Near infrared band (B08), e) S2 – Short wave infra-red band (B11), f) S2- Enhanced Vegetation Index, g) S2- Bare Soil Index, and h) S2- Normalised difference vegetation index. Note that this is just a subset of the possible signals. In the analyses of signal behaviour, the same activities observed on different parcels are analysed together. In Figure 11, eight different signals are shown with the median response of multiple observations and the corresponding uncertainty bands. The signatures are presented for a limited time range (15 weeks centred around the activity) to focus the analyses on a single activity and exclude potential variability due to presence of other activity/land cover phenomena. It is clearly visible that a change of the signal can be observed after the activity for most of the signals. The exception is the Sentinel-1 backscatter where no significant change is observable for mowing activity. Thus, this signal has low potential for detecting this activity. The others have potential for further analyses as an increase of the signal is evident for four signals (S1-Coherence, S2- bands 4, 8 and 11 and BSI), whereas a significant decrease is visible for the indices: EVI and NDVI. The change is observable for 3-4 weeks after an activity.

From the behaviour point of view, the 6-day coherence starts to increase 2 weeks after week 0, and then it stays above the pre-activity level for a long period. The variability of the changes is relatively high, as also a decrease of the signal is within the uncertainty bands. The Sentinel-2 bands: B04, B08 and B11, increase after the activity. In absolute numbers, B08 increases more than the other two.

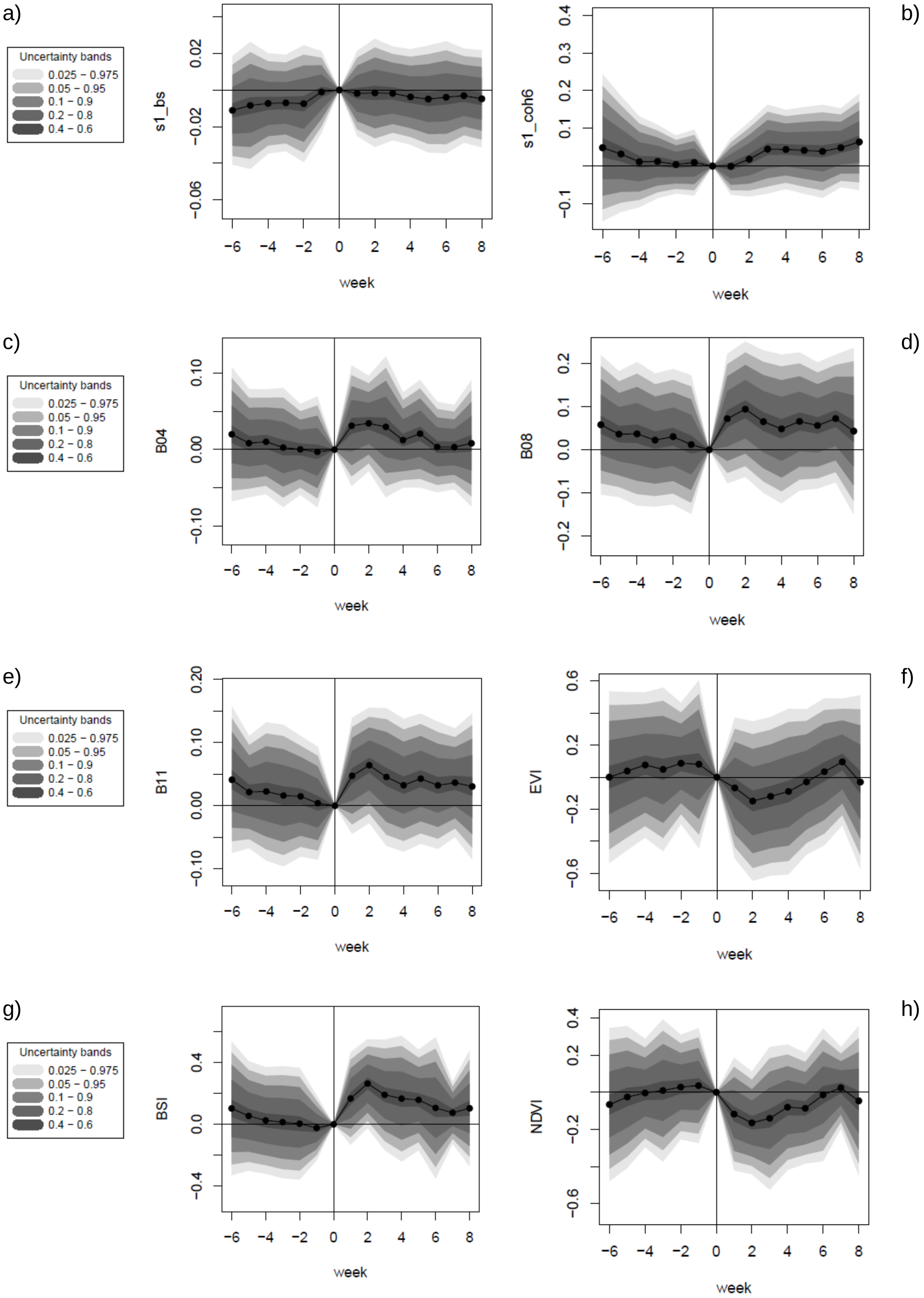
However, their shape and uncertainty bands are quite similar. The uncertainty seems to increase a little bit with time, but not much. The strongest increase appears to be two weeks after week 0. The EVI in Figure 11f decreases after the activity, a decrease which actually starts before the activity. The decrease reaches a minimum two weeks after week 0, and then the EVI increases again. The BSI quickly increases after the activity, before decreasing again. The uncertainty bands are relatively narrow after two weeks, meaning there are fewer cases where the index decreases after an activity. The NDVI quickly decreases after the activity. Also this index has relatively narrow uncertainty bands, meaning that there are few cases where the index increases after the activity.

All of the signals that are used for showing the changes in Figure 11 are normally in the range 0-1 or -1 – 1 (S2-indices). However, we can see that the range of the changes themselves are in a much smaller range for all of them. Backscatter, which is almost constant, has the smallest change, both for the mean and the uncertainty bands. B08 is the Sentinel-2 band with the largest change (around 0.1), whereas the change in coherence and the other bands is around half of this. The largest changes in absolute value can be seen for BSI (up to 0.3), whereas also EVI and NDVI have larger changes than the bands themselves (around 0.2). Whereas the absolute change is not the only thing to look for in signal that could possibly be used to develop a good marker, it could be a first indication of a signal that might be considered.

The standard deviations of the changes is of similar importance. The standard deviation is zero for week 0 (our data preparation method implies that all signals equal zero for this week), but for the remaining weeks, the standard deviation reflect to which degree different parcels observe the same change in signal value after the same period of time. The uncertainty bands are illustrating the standard deviation in Figure 11. Assuming a normal distribution of the changes, the bands in the legend on the right indicate within which range we expect 95%, 90%, 80%, 60% and 20% of the FOIs, respectively. Those bands vary between signals, giving us an indication of the real uncertainties of these signals, captured before/after activity took place. The uncertainties are results of a range of factors, including observation errors, random errors from the selected pixels of an FOI, boundary effects, ground truth errors and others.

The uncertainty is important when we compare it with the absolute values of the changes, described above. If the standard deviation is low, compared with the changes (giving narrow uncertainty bands in Figure 11), we have a signal that is a good candidate for marker development. This is particularly the case for BSI in the figure below and to some degree for NDVI.

Figure 11. Examples of signal behaviour, centred for mowing activity, based on a set of grassland FOIs located in Latvia. Sentinel signals (median) presented: a) S1- backscatter, b) S1 - 6-days coherence, c) S2 – Red band (b4), d) S2 – Near infrared band (b8), e) S2 – Short wave infra-red band (b11), f) S2- Enhanced Vegetation Index, g) S2- Bare Soil Index, and h) S2- Normalised difference vegetation index.



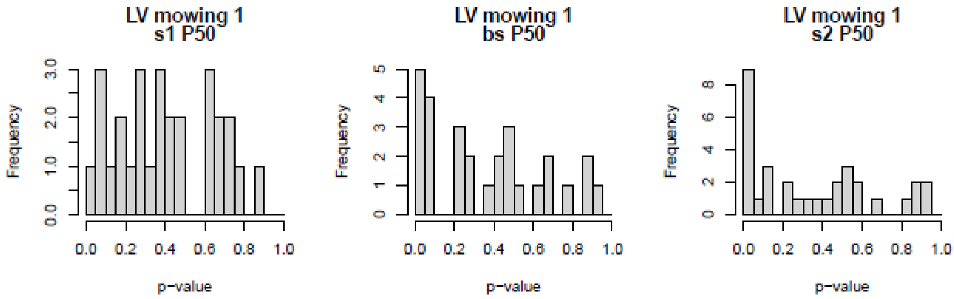
Source: GTCAP

5.2 Statistical significance of the difference

For signals arranged in a weekly pattern a statistical test can be used to verify whether the change between the week of the activity and the weeks before/after is actually significant. The T-test[7] is used to detect significant differences of the mean between two samples. In this case, we test if the average change of signal value after the activity is significant. The null hypothesis is that the true difference is zero (any difference between the weeks is just by chance), whereas the alternative hypothesis is that there is a true difference that is different from zero. We are looking at the values of the same parcels before and after an activity.

This is a paired t-test, also often referred to as dependent t-test. A necessary assumption is that the difference between the weeks (the dependent variable) should be normally distributed. To test this assumption, it is for example possible to run a Shapiro Wilks[8] test of normality. The test rejects the hypothesis of normality when the p-value is less than 0.05. In our case we have a large number t-tests, with each their p-value. In this case we would ideally like to see that the frequency of non-significant p-values is not much higher than for other p-values. This is the case for S1-signals in Figure 12, and to some degree for backscatter, whereas the frequency of low p-values slightly violates this assumption for S2-signals. However, it will add considerable complexity to the analyses if we consider alternative distributions. At the same time, the t-test is regarded as a rather robust test (Mann, 2018), particularly if the number of samples is above 25. We will therefore continue to use the t-test here, as we have more than 25 samples in most cases, and the deviation from normality is limited.

Figure 12. P-values of the Shapiro-Wilks test for median (P50) values for coherence (S1), backscatter and S2 indices for a large number of parcels located in Latvia.



Source: GTCAP

It should also be mentioned here that the number of samples (more parcels or activities) will increase the significance for a similar change. Hence, it is easier to identify a significant change when a high number of samples is present. But whereas a significant change is a necessary requirement for our ability to detect changes, it does not tell how easy it is to detect activities in individual parcels/FOIs.

Statistical test of significance: example

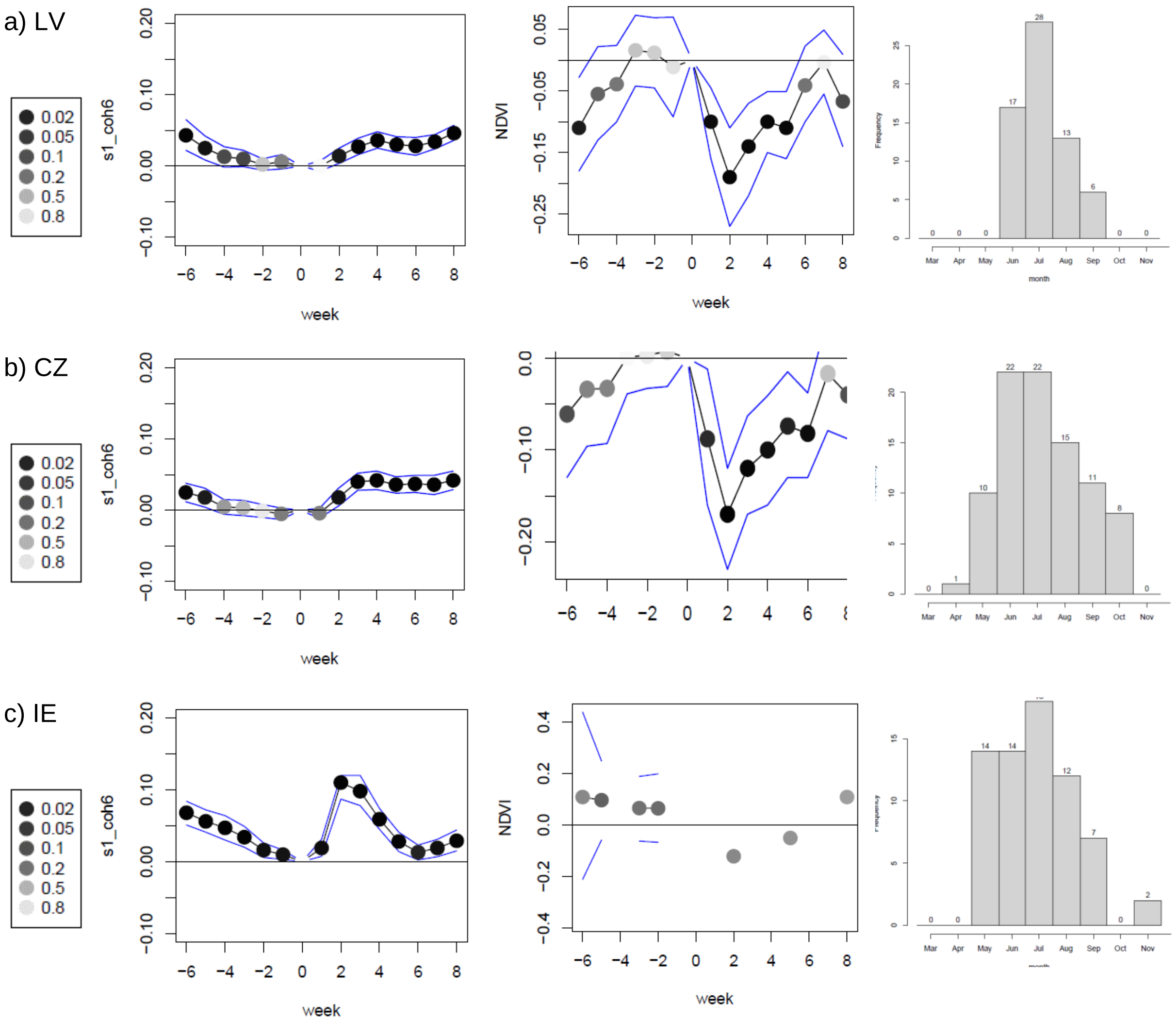
In this example, a set of grassland parcels located in Latvia, Czechia and Ireland is selected. We have extracted the ground truth information (mowing activity dates) together with two median signals (S1 – 6-day coherence and S2 - NDVI) for these parcels. Figure 13 shows the estimated mean difference in signal values between week 0 and the weeks before and after the activity. The blue line shows a confidence interval for the difference (the difference is within the limits of the interval with a certain probability, 95% is commonly used). The grayscale of the dots reflects the significance of the estimated difference being different from week 0; the darker the colour the higher significance reported. The p-value gives the probability that the difference occurred by chance. It is common to see p-values less than 0.05 as significant (less than 5% probability to observe the change by chance),

[7] https://en.wikipedia.org/wiki/Student%27s_t-test (accessed 03/11/2022)

[8] https://en.wikipedia.org/wiki/Shapiro%E2%80%93Wilk_test (accessed 03/11/2022)

whereas larger p-values are not regarded as significant. The last part of the figure shows the distribution of activity dates to indicating the periods with most frequent event in the analysed dataset.

Figure 13. Statistical test of significance of two median signals for sets of grassland parcels located in a) Latvia, b) Czechia, and c) Ireland. In case of the third dataset (shown in c), due to cloud/cloud shadow presence in Sentinel-2 data there is not enough NDVI observations to conclude the analyses. The dots refer to the significance levels (see legend for values, lower than 0.05 is usually seen as significant) of the change being different from zero. The histogram shows distribution of reported activity dates for the three datasets.



Source: GTCAP

In the first dataset, (Latvia), most of the activities were performed between June and September, in the second dataset (Czech) from May to September with a few in October and in the third dataset (Ireland) from May to September.

During mowing activity, a significant biomass reduction takes place. Therefore, we can expect a reduction in the values after the activity date for the signals emphasising the vegetation, such as the NDVI. Such a drop in the values from the weeks before the activity to the weeks after for the NDVI index is notable in two datasets. In the case of Ireland, due to cloud/cloud shadow presence in Sentinel-2 data there is not enough NDVI observations available to conclude the analyses. The LV and CZ dataset give a decrease of the NDVI value when comparing the week before to weeks after the activity, with the maximum reached in the second week (0.18). The significance level of the differences is significant (0.02 for both countries). Two weeks after mowing, the vegetation starts to recover and the NDVI values increase. The 95% confidence interval band includes a zero difference;

this means that it is not unlikely that there is no difference between the signal for week 0 and for the following weeks.

In case of 6-days coherence signal (Sentinal-1 based) a similar behaviour across the datasets is observed. After the activity, the coherence values increase for up to 3-4 weeks. Ideally, a visible decrease of the value before the activity weeks is observable, like in a case of IE data. After the activity, there is a significant increase for all three datasets, particularly for the acquisitions two weeks after week 0. The increase is largest for IE (0.10) and around 0.05 for the two other datasets. Analyses of the same type of signals acquired at different locations result in a similar behaviour, which confirms its stability and repeatability. However, the differences in the values or duration depend on the local variability captured in the FOIs. The significant difference between the week of activity and one or more of the surrounding weeks may be good indication of a signal that supports an activity detection.

5.3 Repeatability of the signal behaviour

Once the significant change for a signal for an activity type has been identified, the next step is to analyse whether this change is detectable, and which of the signals could be most suitable for detecting a change. The multiple activities observed on various parcels/FOIs, even from the same region, exhibit a strong variability. Thus, an important aspect of the signal behaviour is its repeatability, i.e. that the post-activity signal behaviour is visible also for other FOIs. This can be derived from further analyses of the signals.

For each signal, we have estimated a mean and a standard deviation for the changes for each week. Assuming a normal distribution, we can then calculate the probability of the difference values in a given week being greater/smaller than a threshold value, in this case greater than zero (greater: for increase of signal, smaller: for a decrease in signal after the activity). This shows the percentage of parcels that have a change in the same direction as the average change.

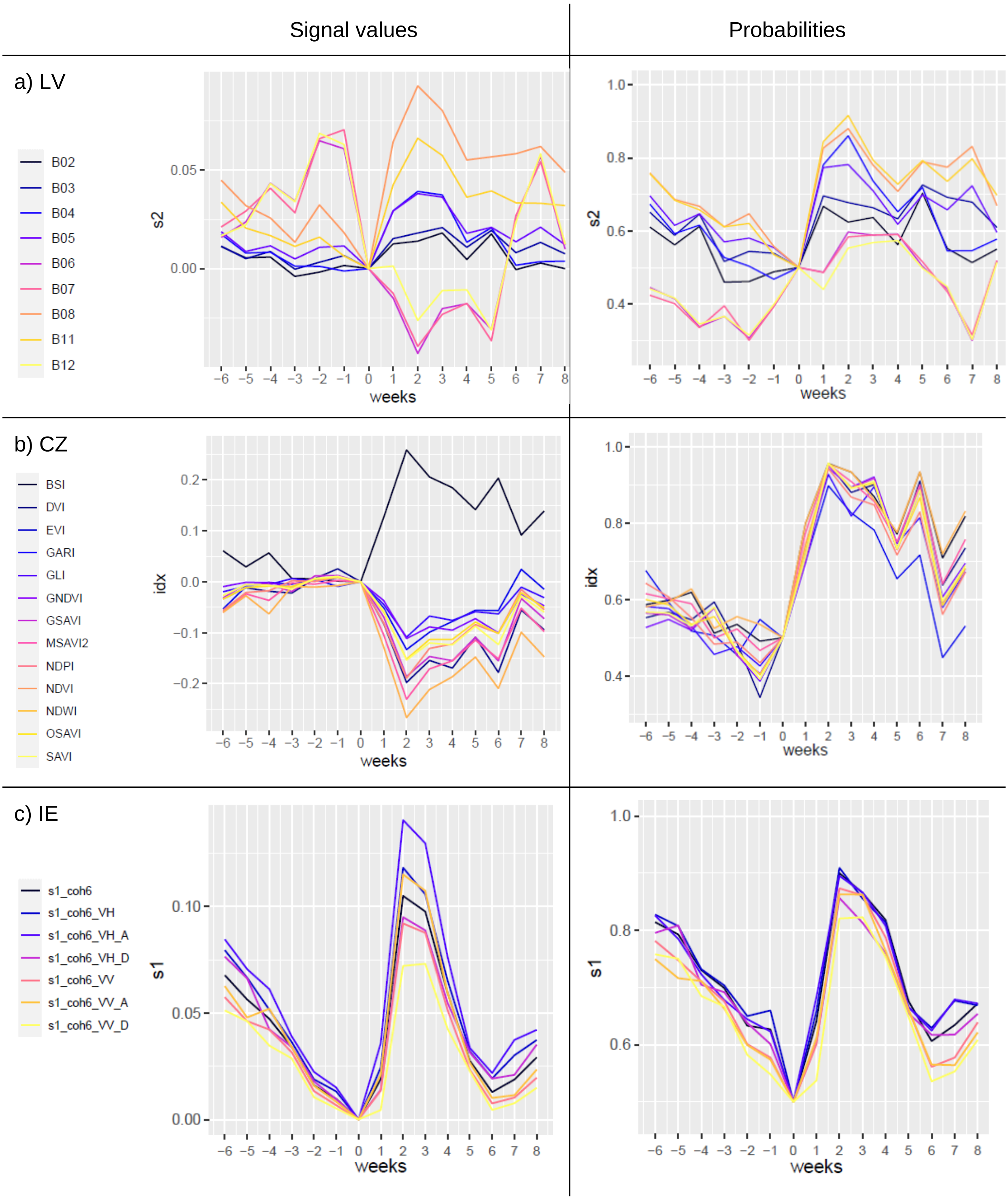
The selected threshold can be adjusted in a more advanced stage when predicted marker detection is concerned (i.e. % of single change, or specific threshold value). Then this would show the percentage of parcels that have a change that is larger than what could be a typical marker, which is then used for detecting events. A higher probability of signal increase shows stronger agreement with expected signal behaviour. Similar to previous analyses this one is also done on a weekly basis. When a high probability is observed for consecutive weeks, such a signal is a good candidate to provide base for a marker in the CbM context.

Statistical probability analyses: example

Data from Latvia, Czechia and Ireland illustrate the probability analyses. Three independent datasets consist of parcels with permanent grassland with mowing activities assigned (Figure 14). For the first two datasets Sentinel-2 spectral bands and spectral indices are presented, whereas the third shows Sentinel-1 6-days coherence signals. For each subplot (a, b, c), the left panel shows the difference between week of activity and the previous/following weeks for each S1/S2-signals (median value). The right panel shows the probability of the signal value being above the threshold, which is zero in this case. This value should ideally be as close to one as possible, and reflects the part of the uncertainty bands above the threshold.

Figure 14a shows results for the different Sentinel-2 bands for the Latvian dataset. In the left panel, we can see the changes according to the weeks. B06, B07 and B12 decrease after the activity, whereas the other bands increase. B08 and B11 show the strongest increase after the activity. The right panel shows the probabilities of observing an increase for signals that typically increase, and a decrease for signals that typically decrease. We are looking here for the highest probabilities, hence B11 would be a good candidate for developing a marker. Also B08 and B04 have relatively high probabilities. On the other hand, B06, B07 and B12 have the lowest probabilities of a FOI showing a change in the direction of what we would expect. The probabilities of these are close to 0.5, which means that it is almost random whether the signal will increase or increase after an activity.

Figure 14. Signal probability analyses for grassland parcels located in a) Latvia – S2 spectral bands, b) Czechia – S2 – spectral indices and c) Ireland - S1 – coherence 6-days. The left panels show signal difference between week 0 and the previous/following weeks, the right panels the corresponding probabilities of the signal change being above the threshold (zero in this case).



Source: GTCAP

The results for the spectral indices derived from Sentinel-2 for the Czech dataset are presented in the second part of the figure (Figure 14b). The median of all these signals shows a decrease after the activity, except for BSI, which strongly increases. The probabilities for a FOI showing a change in the same direction as what we expect are high for all indices for week 2. However, there are some small differences between the presented signals with the spread of 5% between most of them. The probabilities decrease after week 2, although they are also high for week 6. This could be considered an artefact or related to another land cover manifestation. Figure 14c shows results for the 6-day coherence derived from Sentinel-1 for the Irish dataset. The median of these signals shows an increase after the activity, reaching a maximum for week 2 or 3 after week 0, depending on the polarization and orbit. In general, the signals from the ascending orbits have larger increases than the signals from the descending orbits, and the signals from the VH-polarizations increase more than the signals from the VV-polarizations. These changes in absolute values are here partly reflected in the probabilities of a signal for a FOI showing the right behaviour. These probabilities are lowest for the signals from the descending orbits, and better for the ascending orbits.

Based on the different signal behaviour examples shown in this chapter and others that are not shown here, we can summarise some observations. First of all, it is clear that signals originated from various sensors and bands derivatives react differently to an activity presence. Some increase, some decrease, and some do not have a characteristic behaviour after a specific activity. There is also a difference in the strength of a signal increase or decrease. We will not discuss the exact behaviour for each signal, as this can depend on the country and typical geographic/meteorological conditions, and also difference in practices. The change in signal value is a first indicator of which signals would be good candidates for developing a marker. However, we can see that the magnitude of the signal changes on the left are only partly reflected in the probability panels on the right. This is caused by the standard deviation of the changes. If the signal values after an activity is highly variable between different FOIs, this can reduce the potential usage of a signal even when it exhibits a significant difference between the signal values before and after the activity. It should also be mentioned that we have analysed here only the probability of a signal going in the expected direction. As a second step, we should also analyse the probabilities of being above or below a particular value of a hypothetic marker. For more details, see the last part of chapter 5.4.

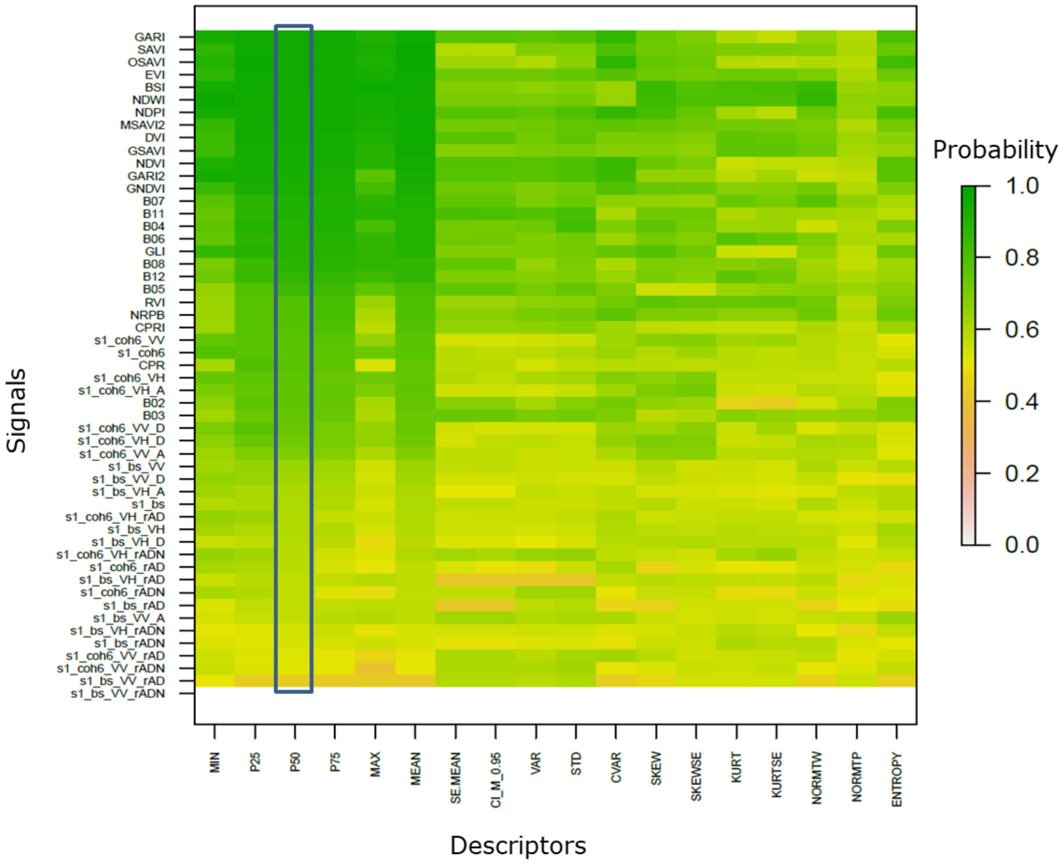
5.4 Summary of signal analyses

Assuming that one is interested in analysis of several signals for a given land cover manifestation or its change caused by an activity as shown above. It is then necessary to summarise these results and ideally, to rank the signals from potentially useful to the ones that should be omitted in further considerations.

To conclude the analyses of multiple signals, we summarise the probabilities in Figure 14 for all tested indices and descriptors in Figure 14. As discussed, a signal's behaviour varies for the same activity, depending on the input sensor data and statistical descriptor used (some signal values decrease, while others increase after activity). To make comparison of probabilities possible, the direction of the decreasing signals was changed. Hence, for a typical signal like the NDVI, we rather look at –NDVI. When we refer to the probability of increasing below, this will also include the probability that a decreasing signal (as NDVI) is decreasing.

In Figure 14, the highest probabilities for an increasing signal can occur after a different number of weeks. The differences are a combination of signal responses, local context and random variability. Instead of looking at each week separately (as discussed in previous chapter, see also Figure 11), Figure 15 summarises the maximum probabilities of the direction of the signal value for the first 4 weeks after a mowing activity for the Czech dataset, for 53 source signals and 18 descriptors (in total nearly 1000 signals). Hence, this gives the probability for when the signal is the strongest, independent of which week this happens. The probabilities have been sorted according to the P50 descriptor, making it easy to visualise differences between descriptors. When none of signal's descriptors has a significant increase or decrease detected, or if not enough acquisitions are available (due to clouds presence) then no probabilities are shown, and such case is marked with white colour.

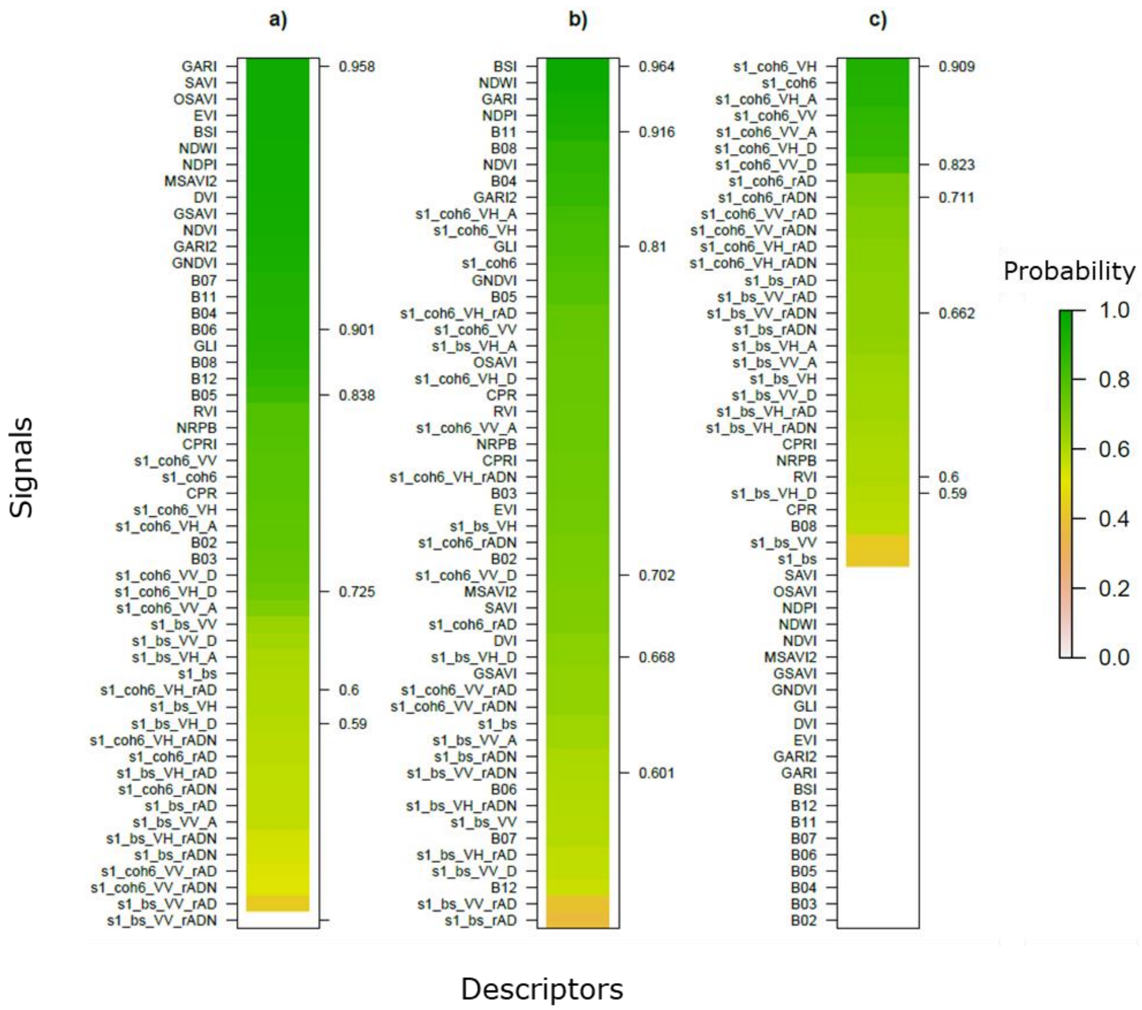
Figure 15. The highest probabilities of the change for different signals/descriptors being of the same sign as the average for selected parcels in Czechia, during any of the four weeks after mowing took place, sorted according to the P50 descriptor (marked with the box). Signals for which none of the descriptors give changes (from the t-test) and signals without sufficient data are shown as completely blank (only one signal at the bottom in this case).



Source: GTCAP

First, we can notice that the highest probability is 0.968 (the MIN descriptor of NDWI), meaning that we can on average expect to observe an increase of the signal when an increase is expected for 29 out of 30 FOIs. The second is that the probabilities are highest and rather similar for P25, P50, P75 and mean. The highest value overall is from MIN, but in general this descriptor gives lower probabilities than the other ones. The other descriptors also give high probabilities for some signals (including single cases), but in general they appear to be less suited for developing a marker than the mean or median descriptors. With the values sorted, we can also easily rank signals for this particular activity type. Some of these numbers are easier visualised in Figure 16, where the Czech data set is shown leftmost. Numerically, GARI ranked at the first place, but we can see that there are relatively small differences between GARI and almost all the other S2-based indices, maybe except for GLI. The difference between the first ten signals is less than 1.5% (0.943-0.958). This is likely within random variability.

Figure 16. The highest probabilities of the change for different P50-signals being of the same sign as the average for selected parcels in Czechia (a), Latvia (b) or Ireland (d), during any of the 4 weeks after mowing took place. Signals for which none of the descriptors give changes (from the t-test) and signals without sufficient data are shown as completely blank.



Source: GTCAP

Figure 16 also shows sorted maximum probabilities for the first 4 weeks after a mowing activity for the P50 descriptor, but for three countries: Czechia (a), Latvia (b) and Ireland (c). For better readability of the colour marked values, additionally some numbers on the right give the probability values for selected indices (i.e. NDVI and S1_coh6) and the first index with a probability above 0.6, 0.7, 0.8 or 0.9. The probabilities for the Latvian dataset are somewhat similar to the Czech dataset, but fewer indices give a high probability. The BSI and NDWI give the highest probabilities (0.964), whereas GARI is slightly lower (0.945). The highest probability is slightly higher than for the P50 descriptor of the Czech dataset. The difference between the Sentinel-2 indices and the others are not as clear as for the Czech dataset. Some of the S2-bands (B11, B08 and B04) are equally good or better than some of the derived indices, such as GLI, the different DVIs and the different SAVIs, which are in the order of 0.7-0.8. Also some of the 6-day coherence signals are better than some of the S2-bands and indices. See the figure for further details.

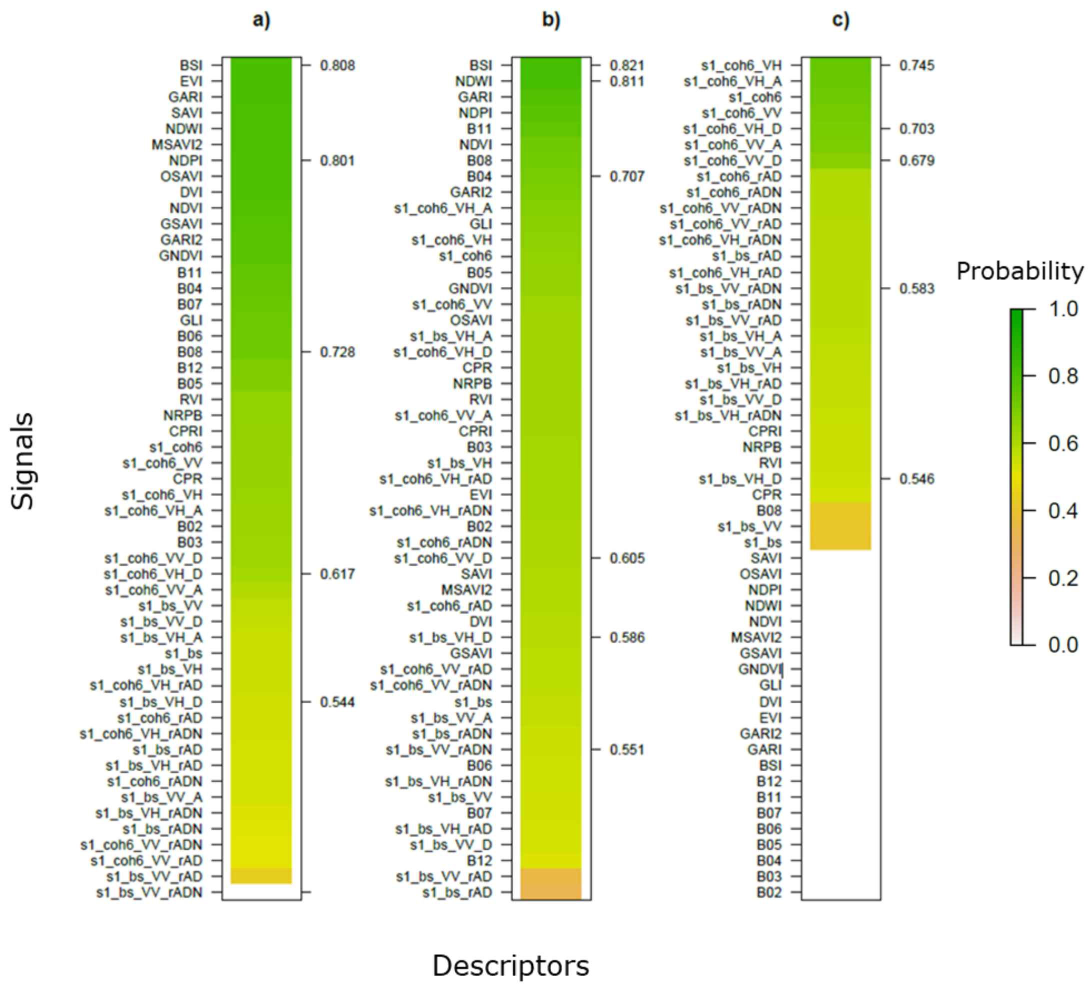
The results for the Irish dataset (c) are substantially different. First, much fewer signals can be used, due to the lack of valid acquisitions of S2-images. There are a couple of S2-bands that showed significant changes, but for most of them (and the S2-indices), there are too few observations, giving too high uncertainty of the estimates. Among the remaining indices, different COH6 signals perform best. The highest probability is 0.909, noticeably lower than what was achievable for the S2-indicesin Czechia and Latvia. This is on the other hand higher than the values for S1_coh6 in Czechia and Latvia. The highest probability for any of the coherence signals is 0.825 for Latvia and 0.776 for

Czechia. Different orbit ratios (from both coherence and backscatter) perform better than the backscatter itself, but all with probabilities below 0.7.

It is worth noting, that while interpreting results from three datasets, some difference in the ranking of the signals are expected due to local soil types, landform and climatic/weather conditions as well as the farmer's practices. Another element is differences in ground truth data uncertainty level among the datasets, where the Irish dataset are most reliable due to its fit-to-purpose and weekly field visits. This can likely explain the higher probabilities for the comparable signals (COH6-signals in particular).The probabilities of increasing or decreasing signals provide an interesting insight into the signal usability, but we would largely overestimate the number of detected events by using a threshold of zero. In the following figure (Figure 17), we test the detection possibility for a marker defined as half the largest change during the four weeks after the event. The figure shows the highest probabilities of a change being above this marker for Czechia (a), Latvia (b) and Ireland (c).

First, we can notice that the signals that performed well in Figure 16 are mostly among the high ranked ones also in Figure 17, even if the probabilities are lower. In Czechia, the S2-derived indices ranked top, most of them quite similar and in the range from 0.828-0.861 (GNDVI). In Latvia, NDWI, GARI and NDPI are the placed first (0.832-0.868), with NDVI a little bit behind (0.792). The top ranked signals are quite similar for Latvia and Czechia, but there are fewer good signals for Latvia. However, in this figure we can see more clearly that most of the S2-derived indices perform better than the S2-bands themselves. The probabilities for the Irish data are considerably lower, as there were not enough S2-acquisitions available. The average coherence is ranked first, but with a probability of 0.745, meaning that 1 of 4 FOIs are likely to have lower increases (or decreases) than this test marker.

Figure 17. The highest probabilities of the change for different P50-signals being more than half of the maximum signal change for selected parcels in Czechia (a), Latvia (b) or Ireland (d), during any of the 4 weeks after mowing took place. Signals for which none of the descriptors give changes (from the t-test) and signals without sufficient data are shown as completely blank.



Source: GTCAP

The probability values are used as a way to rank the signals. This is a method for looking for true positives. We have not analysed the risk of false positives or false negatives in this subsection, what would also be related to the typical changes of a signal when there is no activity.

5.5 Signals correlation

As previously mentioned, in the marker development process it is assumed that a marker derives from one signal, but a signal could source many markers. Different markers could also be derived from different signals. A phenomenon or land cover manifestation observed by multiple markers reinforces the performance reliability only if the extracted information is not strongly correlated. In a case of strong correlation between the signals used to observe the same land cover manifestation state/change, the reliability reinforcement is minimal. In case of optical sensors, utilisation of information from different part of the spectrum might be already helpful. However, the use of signals from independent sensors is advised (Sentinel -1 and -2) to take an advantage of different sensors characteristics and properties. The image data captured by different sensors might secure the automated process from weather dependencies and limitations of passive sensors and result in higher repeatability of observations (markers).

For every single sensor, there are many possibilities of signals and their further derivatives to explore in the course of a signal analyses and marker development. Once signals with repeatable post-activity behaviour are identified, they can be further analysed to check their correlation (pairwise). The lower the correlation between them the higher the potential of adding new information in the detection process.

Signal correlation analyses example:

For a set of grasslands parcels located in Czechia, the 54 Sentinel-1 and 2 based median signals are derived. The signal selection includes single image bands and their derivatives (image indices). The correlation coefficient[9] calculation between two signals is performed at the FOI level using the temporal subset (15 weeks around the mowing activity date available for each parcel), then averaged to a single value and arranged graphically (Figure 18). The values of correlation coefficients are between −1 and 1. Correlations equal to +1 or −1 correspond to signals behaving exactly alike or exactly the opposite, and imply that a linear equation describes the relationship between signal 1 and signal 2 perfectly, with all data points lying on a line. In such case, these two signals have the same information content. A value of zero implies that there is no dependency between the signals, and therefore using both signals is advantageous in terms of information content.

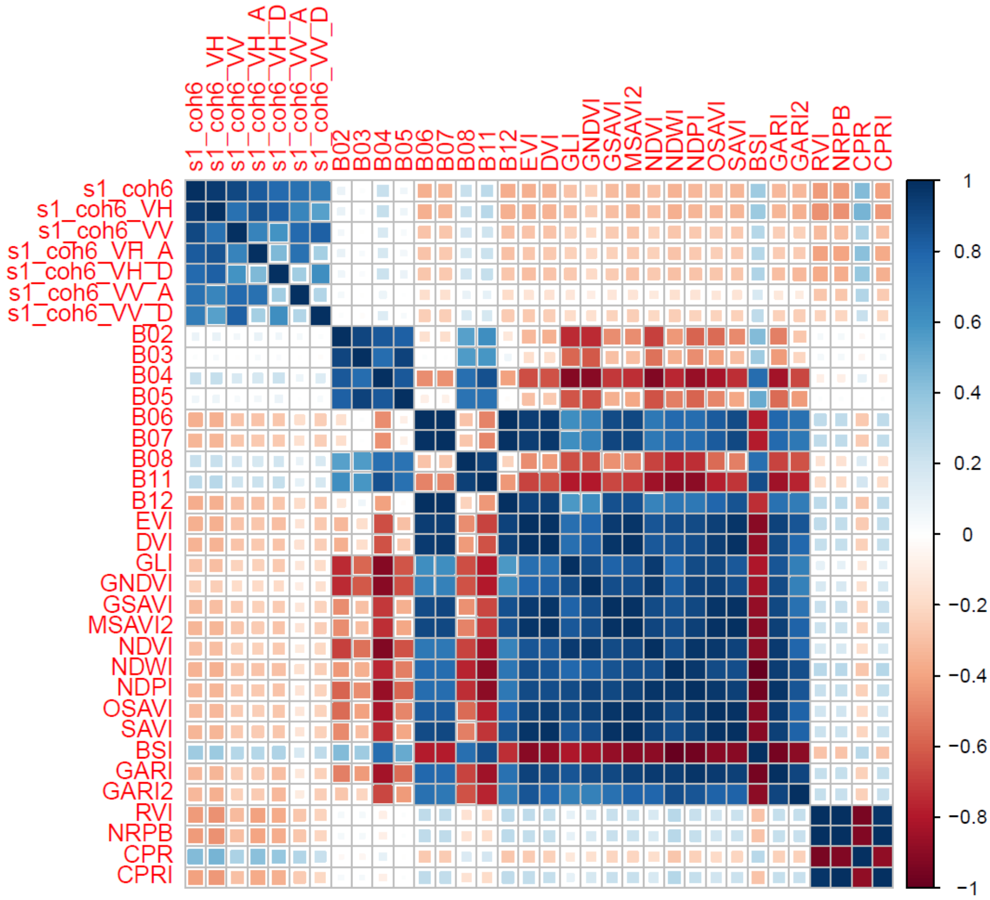
In this case, there are noticeable strong positive correlations between image bands and the image indices marked in blue colour. Negative correlations are marked in red, and are, for example, found between several S2-derived indices and the S2-bands themselves. The coherence signals are mostly well correlated with each other, maybe with the exception of ascending versus descending orbits. The same is the case for most of the S2 bands (except for B06, B07 and B12), for the S2-derived indices and for the backscatter derived indices. There is a high (positive or negative) correlation between the S2-derived indices and most of the bands, with the exception of B02, B03, B05, and to some degree B08. Still, the correlations between the bands and the derived indices are not as strong as between the derived indices. On the other hand, we see that there is weak correlation between the coherence and all other signals. In addition, the backscatter-derived signals are weakly correlated with all other signals.

This way, the correlation analyses can be used to select potential signals that introduce added value to the development of markers to detect an activity. With the correlations shown in Figure 18, it might not make sense to choose several S2-based indices for marker development, even if they are all among the optimal ones. If choosing more than one, it would make more sense to use one of the coherence signals, and then maybe one of the backscatter-based indices. This approach is applicable

9 https://en.wikipedia.org/wiki/Pearson_correlation_coefficient (accessed 03/11/2022)

to a greater number of signals (i.e. derived from Sentinel -1 and -2) and performed for any activities/land cover manifestation of interest.

Figure 18. Sentinel-1 and -2 based signals pairwise correlations for data from Czechia. High correlation values (dark blue for positive correlation, dark red for negative correlation) indicate similar content. Value of close to zero (marked in white or pale blue and pale red) implies that there is no dependency between the signals.

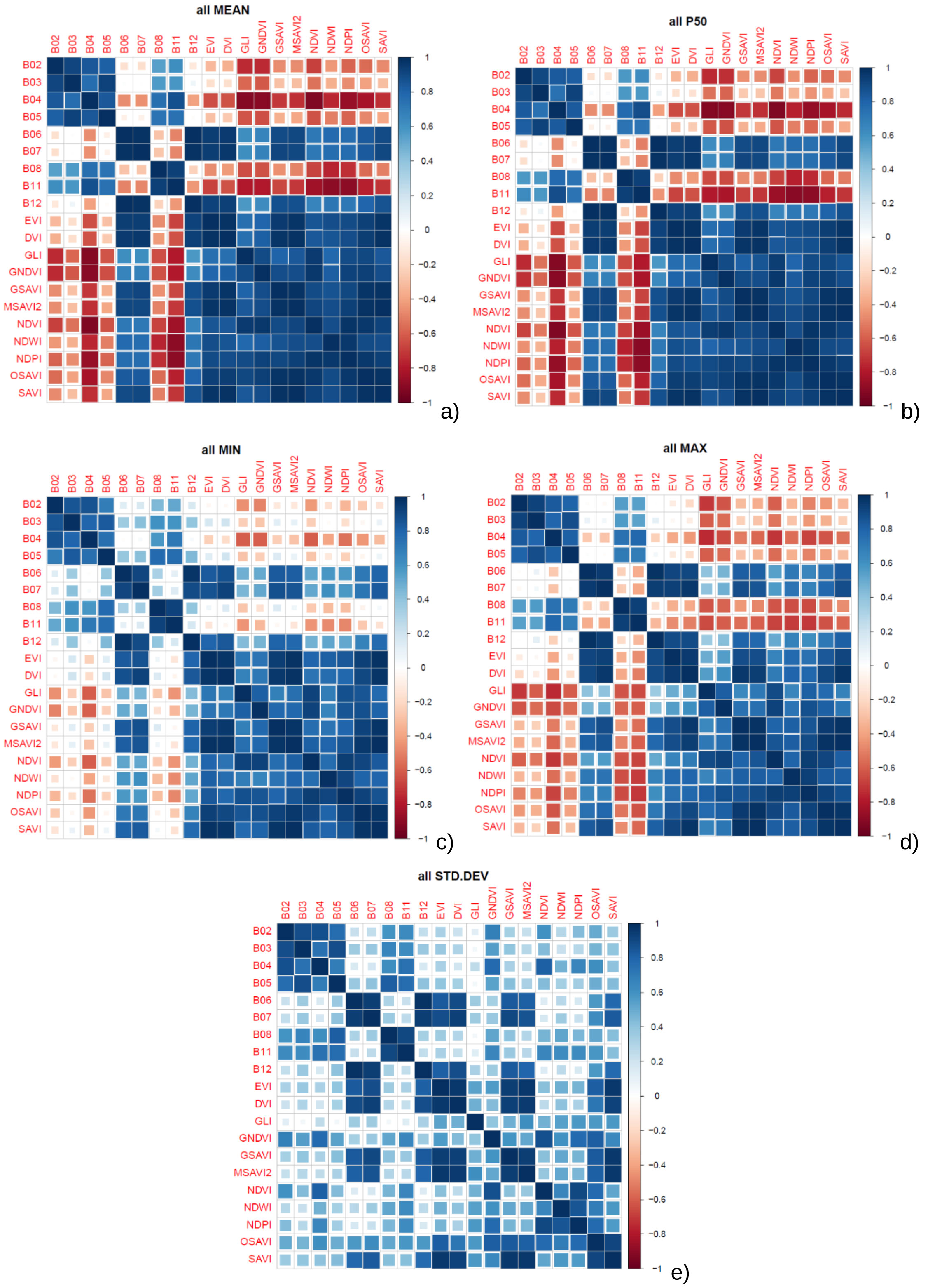


Source: GTCAP

Figure 19 shows the correlation between S2-bands and some of the S2-derived indices for different statistical descriptors calculated at the level of the FOI, for the Czech data set. It is clear that the correlations depend on the descriptor. In general the correlations are higher (positive or negative) for the median and the mean, Figure 19 a) and b). The difference between these two are marginal though. The minimum and maximum descriptors in Figure 19 c) and d) show somewhat similar patterns, but some correlations are higher for the maximum descriptor in this case, particularly the negative correlations between some of the S2-bands and their derived indices. The correlations of standard deviations are quite different from the others though. First, there are no negative correlations. Then there are considerably fewer highly correlated signals than for the other descriptors. This could be a result of this descriptor not being able to capture effect of the activities in the signal. Nevertheless, if the standard deviation proves to be a 'good' descriptor for another data set, the results here for example indicate that it might be possible to choose more than one S2-based index in this case.

We have not analysed or shown cross correlations between different signals and descriptors here. There are good reasons to expect that the correlations between different descriptors and different signals would be lower than between the same descriptors for these signals. An opportunity could then also be to use different descriptors for different signals.

Figure 19. Example of signals correlation (mowing activity in Czechia, 15 weeks period), with 5 of the main statistical data descriptors (mean, P50, min, max, stdev). High correlation values (dark blue for positive correlation, dark red for negative correlation) indicate similar content. Value of close to zero (marked in white or pale blue and pale red) implies that there is no dependency between the signals.



Source: GTCAP

6 Examples of further analyses

In this chapter, examples of further analyses using the modules described above are presented. These are examples showing that the methodology can be used for analysing other features than the activity behaviour itself. The first deals with data descriptor analyses to see which can be optimal for the CbM signals. The second explores the richness of detailed and complete ground truth dataset from Ireland and analyses different mowing types. The last presents preliminary results of a comparison between two sets of FOIs with verified and non-verified boundaries to assess and emphasise the impact of FOIs not following the requirements.

6.1 Signal data descriptor analyses

In these example several descriptors are analysed to find which of them have the greatest potential in the CbM signal behaviour analyses. The results presented in Figure 15 can also be analysed in a different way, leading to finding the optimal descriptor(s) for each signal. The optimal descriptors were summed up for all signals and for the second and third week after the activity. Figures 20 - 21 show a count of the most and least optimal descriptors, respectively. The first figure identifies the data descriptor with strongest evidence (best among others) for mowing activity for the different signals. Here, the mean, P25 and P75 appear to be the most optimal ones. Somewhat surprisingly also the minimum gives a good result. On the other hand, the second figure shows that minimum and maximum are also among the descriptors that frequently gives the poorest results, together with some of the other types of descriptors, such as kurtosis and entropy.

Figure 20. Quantitative comparison of the 18 data descriptors, the optimal (highest probability) for a mowing activity, based on Sentinel data acquisitions in the period 2-3 weeks after the activity takes place.

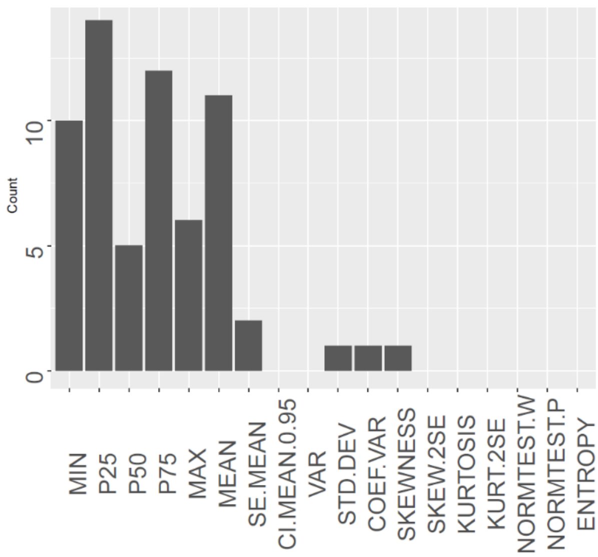
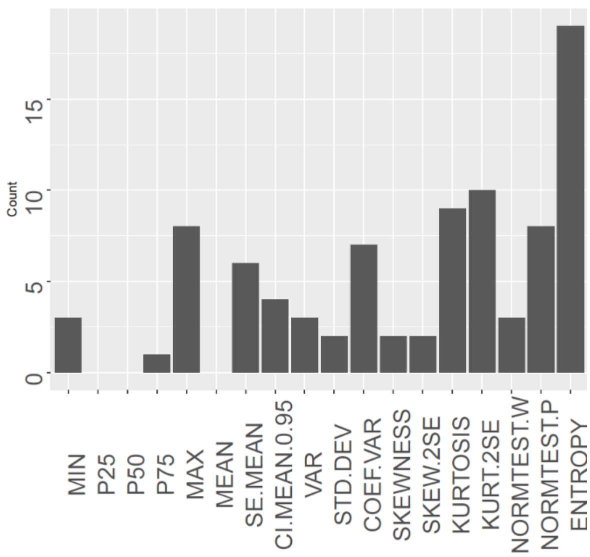


Figure 21. Quantitative comparison of the 18 data descriptors, the least optimal (lowest probability) for a mowing activity, based on Sentinel data acquisitions in the period 2-3 weeks after the activity takes place.

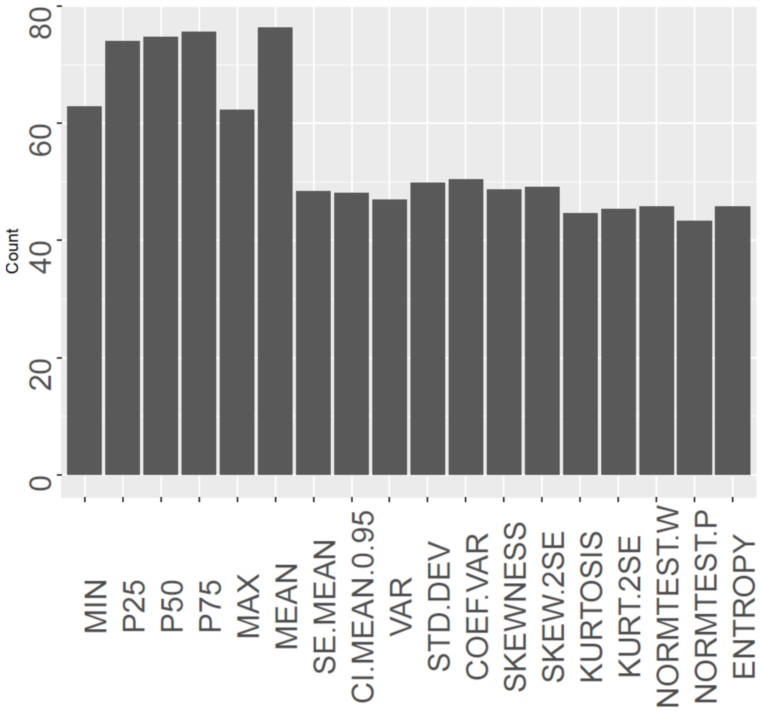


Source: GTCAP

The results in Figures 20 - 21 only consider the most optimal descriptor, but can be somewhat random in cases where several descriptors in reality are equally good. To overcome this, Figure 22 is a summary of how good the different descriptors are, relative to the one that is ranked first. For each signal (and week), a descriptor gets a score which is the probability of the descriptor divided by the probability of the most optimal signal descriptor. The perfect score is one, and lower if the descriptor gives a lower probability than the best descriptor. These scores are summed up for all signals for the second and third week after the activity took place. The results are almost similar for P25, P50, P75 and mean. Minimum and maximum give relatively good results, whereas the results are considerably poorer for the other descriptors.

The results are not surprising, as the mean is commonly used as the prime descriptor for analyses of signals. However, there is always a risk that the mean is affected by outliers, which typically can be found on the border of the FOI. Therefore, in this report we used median (P50 – 50th percentile) as the main descriptor. It might also be beneficial to consider more descriptors during the data discovery and analyses, especially for more complex cases of land cover manifestation.

Figure 22. Quantitative comparison of the 19 data descriptors. Descriptors classified according to their relative goodness (compared with highest probability for a signal) for a mowing activity, based on Sentinel data acquisitions in the period 2-3 weeks after the activity takes place.



Source: GTCAP

The analyses above show that several descriptors can be considered as a signal that can be used to detect ground activities. There were relatively small differences between the three quantiles and the mean, whereas the extremes and the more complicated descriptors were less suitable. However, it should be noted that this is a general analyses, and the results could be different for some of the signals. We can for example assume that signals like min, max and mean are more likely to be subject to boundary issues and random effects for FOIs with a low number of pixels than the quantiles (P25, P50 and P75). They are also more likely to be affected by deviation from normality of the signal values, particularly for bands and indices where the extreme values can strongly deviate from the "typical" values. These includes most of the ones that are not bounded by 0 - 1 or -1 – 1.

6.2 Analyses of types of mowing activities

This example is based on a detailed and complete ground truth dataset from Ireland. As mentioned before, this dedicated survey was designed to serve the purpose of further understanding land cover manifestations and activities represented locally. Three types of mowing (topping, mowing for hay and silage) activities were recorded separately, thus it gave a possibility of exploring the differences between activity types as discussed in chapter 2.2 for the Sentinel -1 signals. The Sentinel -2 data acquisitions, for this region, have persistent cloud cover recorded thus there are insufficient observations with cloud free data.

Figure 23 shows a comparison of different S1-coherence signals for three different mowing types: a) silage, b) hay and c) topping. The panels in the figure are similar to the ones presented in Figure 13 and discussed in the chapter 5.2. The signals behaviour depends on the type of the activity.

For all activity types, an increase of the signals is observed after a recorded activity. All signals have a higher increase after the activity for silage and hay than for topping. Also for the topping, the p-values from the t-test are considerably lower for many of the weeks. Topping is an activity where only a top part of the grass is cut, and the difference in land cover manifestation before/after the activity is smaller compared to the other two cases. This results in a less apparent change of the signal.

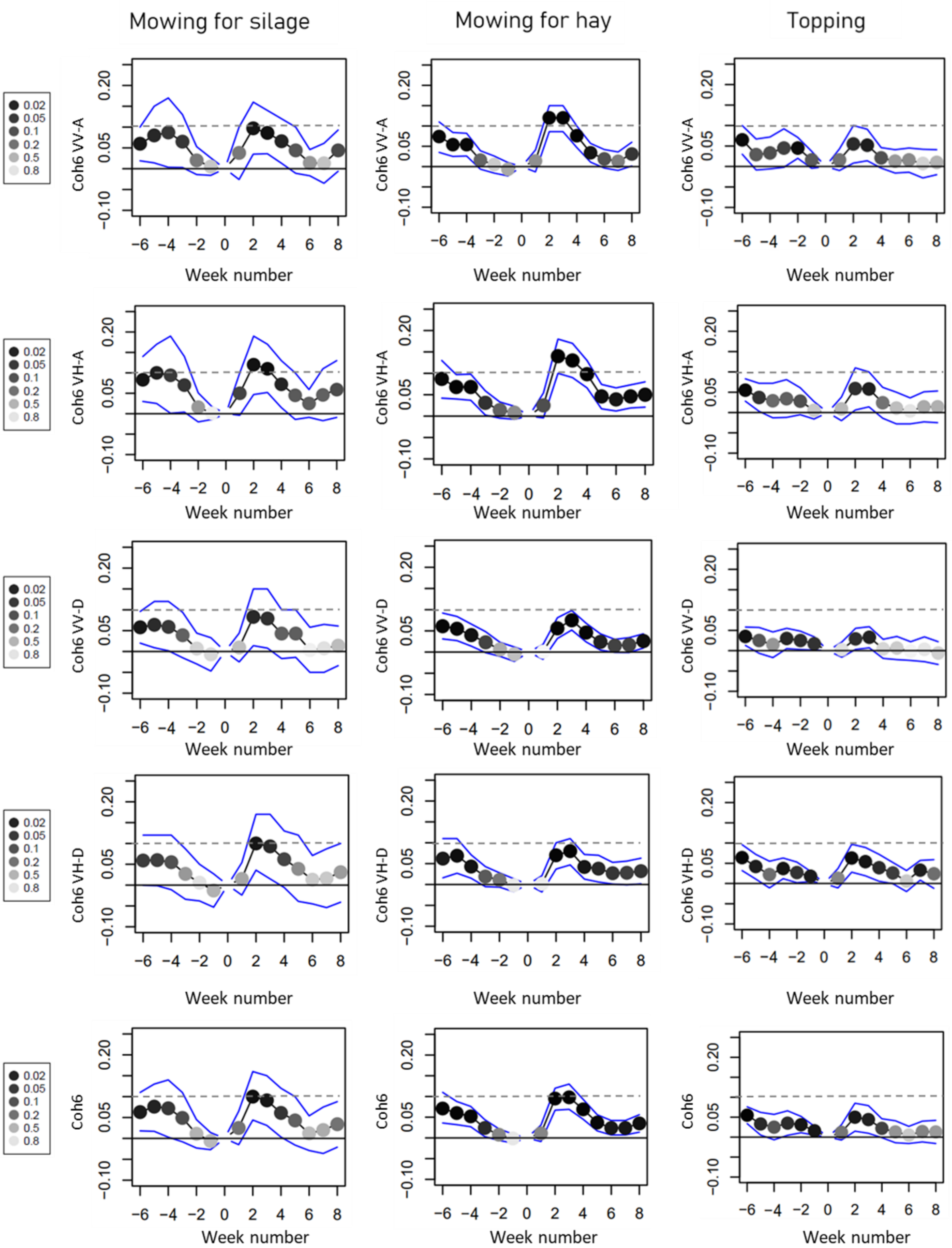
There are relatively small differences between the mowing for hay and silage cases but the latter gives higher values from the ascending orbits for both polarisations. On the other hand, it appears as if the descending orbits have a higher peak for hay. That also means that there is a very small difference between ascending and descending orbits for silage, whereas the ascending signals for hay are considerably higher than the descending signals.

We can further analyse the differences with ANalysis Of VAriance (ANOVA)[10]. This is based on a test which examines the differences in means between more than two groups, and can be seen as a generalisation of the previously used t-test. In this case, a two-way ANOVA is used to see if the mowing types provides different responses in respect to the polarization/orbit variables of the coherence signals. We found a statistically significant difference in the increase of the signal value by both coherence type and mowing type (both with $p < 0.001$). A Tukey post-hoc test (Mann, 2018) could then be used to examine which of the differences between groups that can be seen as significant. First of all, there is a significant difference between topping and the two other types of mowing, with low p-values ($p<0.001$) for both.

However, the difference between silage and hay is not significant overall in 6-days coherence signals (coh6). A reason for this could be that the relationship between the two varies for different polarizations/orbits. Differences between orbits (ascending and descending) are always significant (0.01 or lower), except for S1_coh6_VV_A vs S1_coh6_VH_D. The differences between polarizations for the same orbits have higher p-values (around 0.5), which means that the 0-hypothesis (they are equal) cannot be excluded. We can also see if there is a significant difference between silage and hay for any of the individual signals. This can be done with a t-test between the values for silage and hay. The lowest p-value for these tests is 0.22, which is not significant.

[10] https://en.wikipedia.org/wiki/Analysis_of_variance (accessed 03/11/2022)

Figure 23. Comparison of different mowing activities and different S1-signals (median of FOIs): a) mowing for silage; b) mowing for hay and c) topping. The dots refer to the significance levels (see legend for values, lower than 0.05 is usually seen as significant) of the change being different from zero.

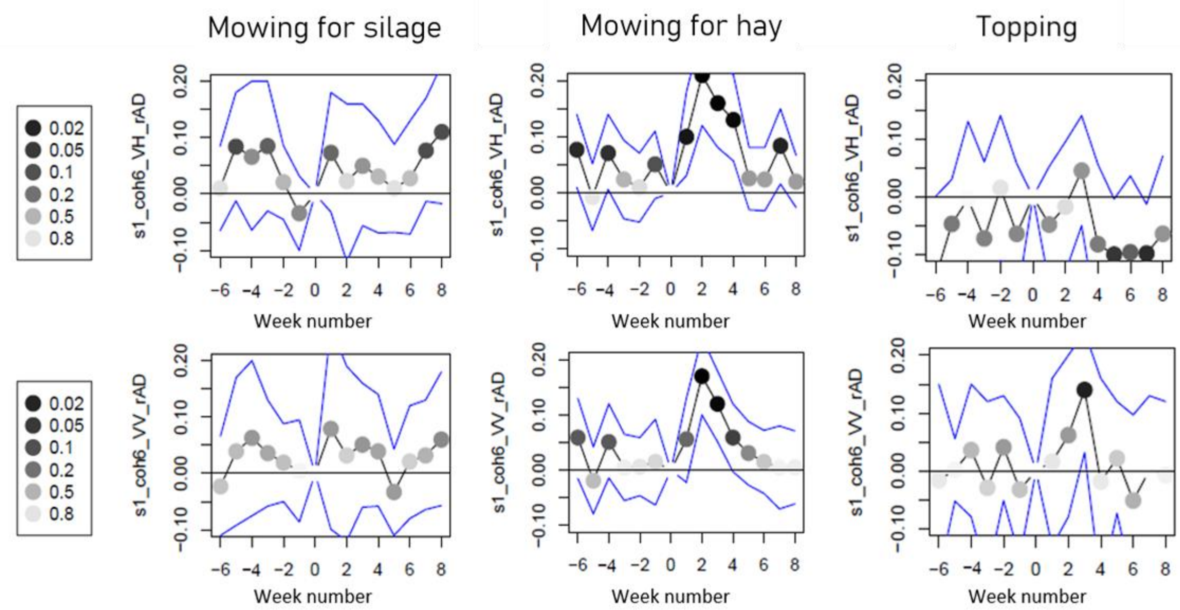


Source: GTCAP

The differences in values for different orbits suggests that the ratios between the two could also be used as an indicator. Figure 24 shows the ratios between the ascending and descending orbits for the different ratios and for the different mowing types. We can see that the signal does not have a significant behaviour for silage and topping, but it increases significantly for hay. Whereas the ascending/descending ratio itself was not among the best signals for detecting the mowing activities (Figure 16), it might be a good supporting signal for identifying the type of mowing.

The analyses of variance here still reveals a significant difference between the changes for the mowing type, but not between the different polarizations. If we run the Tukey post-hoc test after analyses of variance on the ratios between ascending and descending orbits, we find that there is a significant difference in the ratios between silage and hay, with a p-value of 0.001, and also between hay and topping, with an even lower p-value.

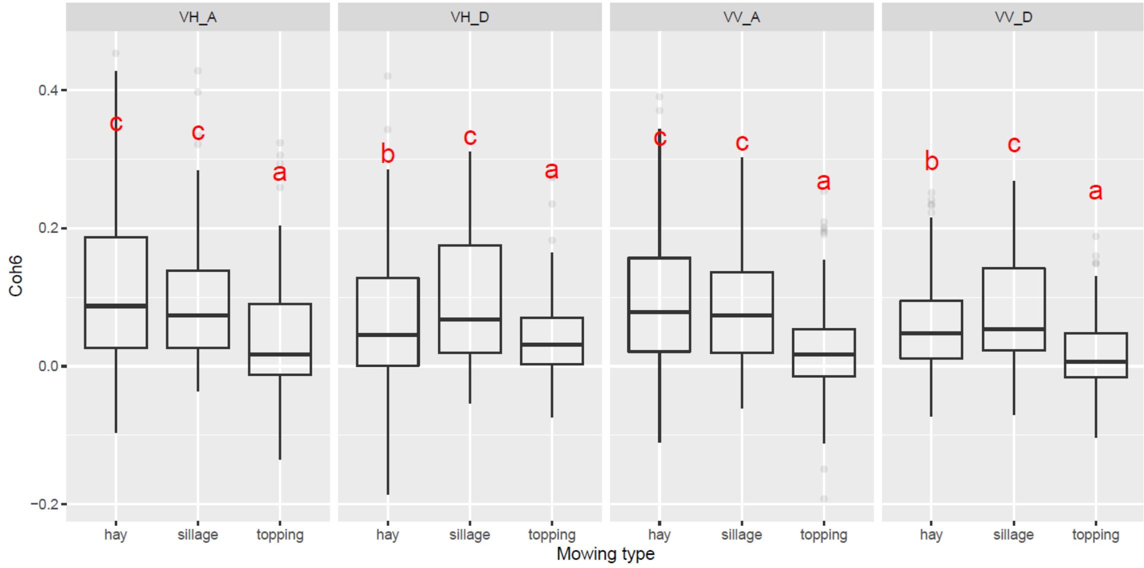
Figure 24. Comparison of different mowing activities for different ratios of S1 ascending/descending coherence signals (median of FOIs): a) mowing for silage; b) mowing for hay and c) topping. The dots refer to the significance levels (see legend for values, lower than 0.05 is usually seen as significant) of the change being different from zero.



Source: GTCAP

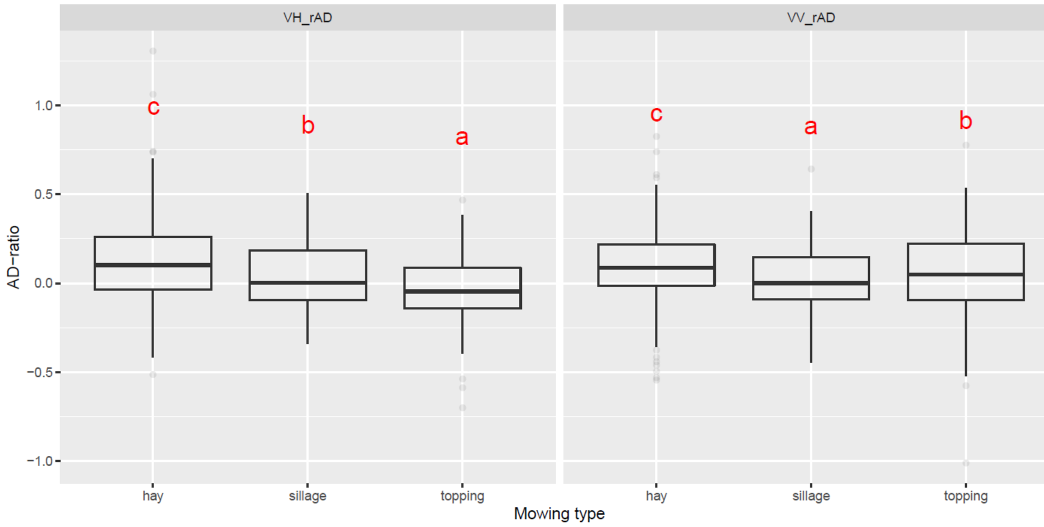
Figure 25 shows a box plot of the signal values from the first four weeks after mowing. Here we can see clearer some of the differences observed in the previous figure. The signal change for hay is considerably higher for the ascending orbits than for the descending orbits for both polarizations, whereas there is barely any difference for silage. The letters represent a grouping depending on the mean value, where the range of the values are split in three. The means belonging to one group may be assumed to be similar. The different behaviour of ascending and descending orbits in Figure 24 and Figure 25 suggests that we should also analyse the ratio between the ascending and descending orbits. Figure 26 is similar to Figure 25 for the orbit ratios of two polarisation (VH and VV), and shows how hay is generally increasing more than mowing for both ratios.

Figure 25. Boxplots of the average changes in signal values for the first four weeks after three mowing types in Ireland (silage, hay and topping) presented for S1 coherence signals. The median of the group is shown with a black line, the boxes represent the first and third quartiles of the values. The characters in red refer to the group of mean level.



Source: GTCAP

Figure 26. Boxplot of changes in the ascending/descending ratios for different S1 coherence polarisations for different types of mowing in Ireland (silage, hay, topping) for the first four weeks after the activity. The median of the group is shown with a black line, the boxes represent first and third quantiles of the values. Characters in red refer to the group of mean level.



Source: GTCAP

The result for the ratios are interesting, as this signal did not appear as one of the most optimal signal for detecting activities in chapter 5.4. However, it seems it can be useful for identifying the type of activity after some kind of mowing activity has been detected through other signals.

The difference in time when a field is free of grass residuals (discussed in chapter 2.2), between mowing for silage and hay is most visible in the change of the signal value in second week for all signals. The mowing for silage has higher change than mowing for hay. So far these three mowing types have been considered in the CbM without distinction. Such approach leads to setting the mowing detection parameters less tailored to cover the variety and in consequence potentially results in more omission/commission detection errors.

Furthermore, upon analyses of the user information needs, most of the MS's consider the occurrence of any of these mowing interventions (topping, mowing for hay and silage) sufficient to conclude on the farmers' fulfilment of the minimal requirements to get the subsidies.

As demonstrated in this example, the higher level of details in ground data enables specific analyses that would not be possible otherwise. Thinking about the CbM system development, and inclusion of more sophisticated measures or schemes, as already observed, the use of high quality and detailed ground truth data might give an advantage in finding solution to complex problems.

6.3 The influence on FOI design issues on the signal

The signal in the CbM is derived from a selection of pixels that are contained in a single FOI boundary. As presented in the Figure 8 in chapter 4, there are some FOIs with perimeters that do not follow the optimal design, including presence of ineligible features, inclusion of areas with non-agricultural use or with more practices performed within a single FOI (potentially separate units of management). The signal computed based on a geometry that does not follow an optimal design may vary from the expected behaviour for an activity signal, and it will therefore be much more difficult to identify this activity in the monitoring system. The FOIs with potential design issues should be identified during the FOI verification procedure, performed at least at the beginning of the annual CbM process.

In order to understand the influence of such cases, an independent test was performed on two permanent grassland datasets. The first consists of 30 original FOIs with perimeter issues, similar to the ones indicated above. The second one consists of the same FOIs but with geometries modified manually according to available image data from the same year. In the modification process the areas of the original FOIs were reduced by 60% on average, ranging from 1-96%. Typically, the largest FOIs were separated into smaller entities that each reflect a single practice. Each dataset has about 40 mowing observations assigned. Table 4 gives an overview of the sizes of the FOIs before and after correction.

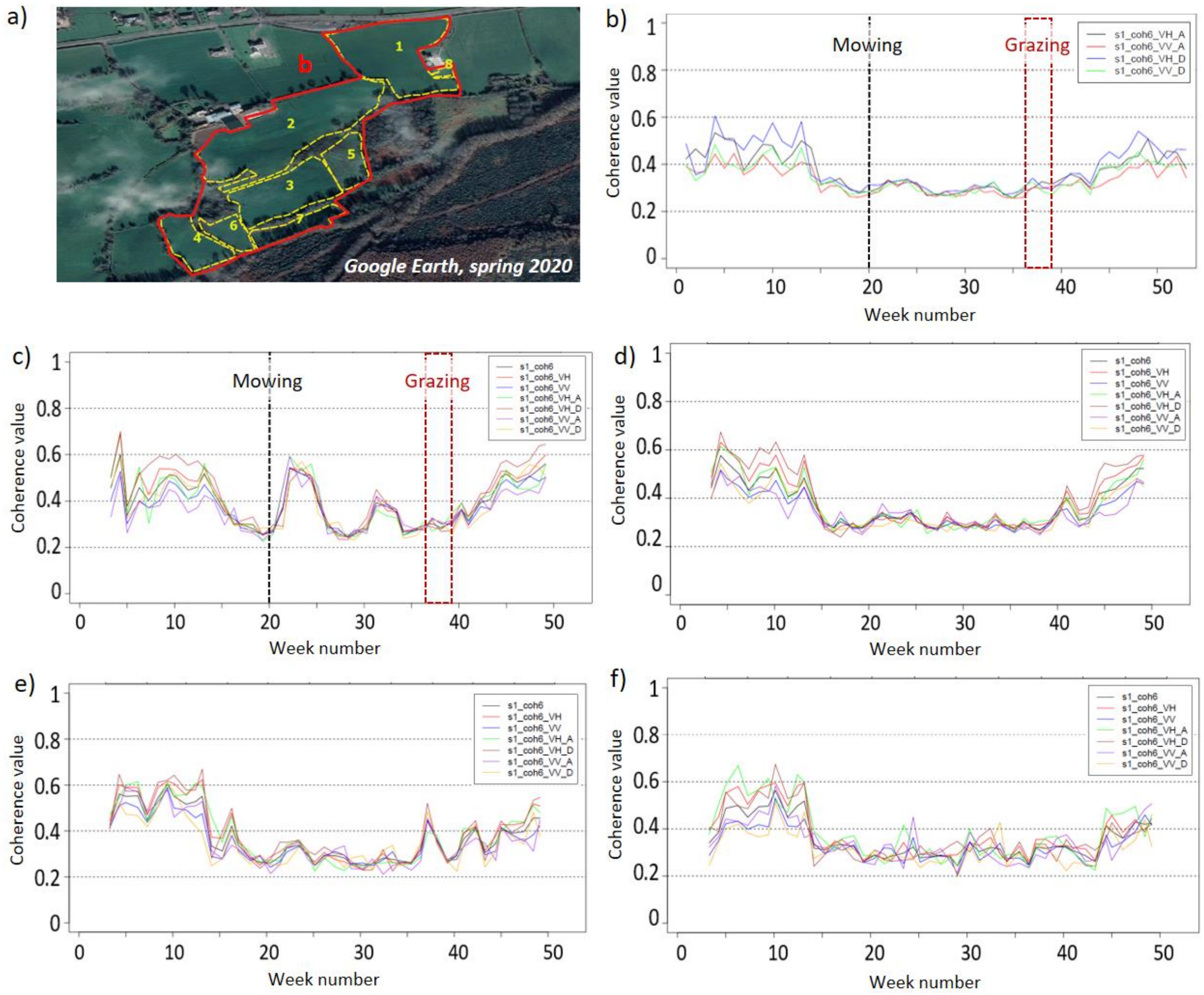
Table 4. Overview of the areas' changing between the dataset with the original geometries and the dataset with modified boundaries

FOIs	N. of FOIs	Total area (ha)	Min area (ha)	Max area (ha)	Mean area (ha)
Original geometries	30	266.1	1.2	28.7	8.9
Modified geometries	30	71.3	0.6	8.2	2.4

Source: GTCAP

Figure 27a shows an example of an original parcel in red (area 9.2ha) composed of grassland units used for mowing and grazing, and what could be identified as separate FOIs in yellow. Panel b) shows the time series of 6 days coherence signals and the recorded activities for the parcel. There is barely any increase in the signal after the activities. The ground truth information refers to the geometry of FOI 1 in a), and the signal shown in panel c) in the figure. Extracting the signal only for this FOI, we see, as expected, a much stronger signal increase for the mowing activity, whereas the increases for grazing are still limited. There is also another increase after week 30, which is not explained by the recorded activities. For the remaining FOIs, we can notice that there is an increase in e), partly corresponding to the recorded grazing activities.

Figure 27. An example of an FOI boundary and the relationship to the signal retrieved; a) FOIs overlaid on a satellite orthophoto, in red original grassland parcel (total area 9.2ha), in yellow the FOIs created after the split, Sentinel-1 coh6 time series signals for b) original grassland parcel, c-f) selected parcels created after splitting the original FOIs numbered from 1 to 4.

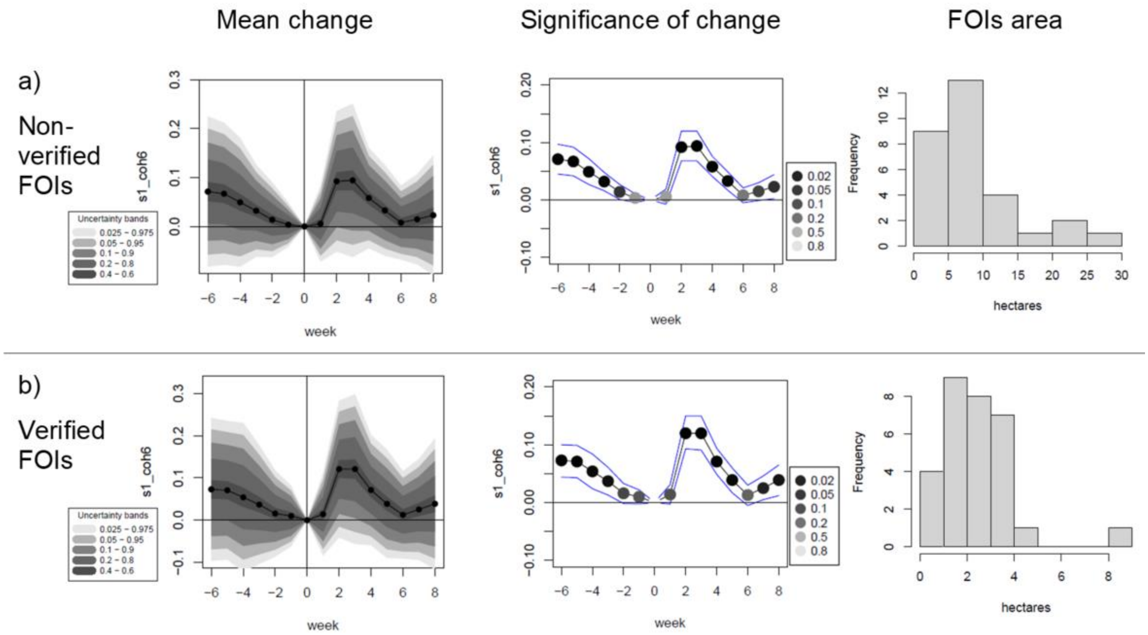


Source: GTCAP

This setup gives a good opportunity for comparing the results before and after boundary verification, visualised in Figure 28 for mowing activities (silage and hay). The panels on the top (a) are based on the non-verified FOIs, whereas the panels at the bottom (b) are based on the FOIs after boundary verification. Three panels are shown for each case: mean change with uncertainty bounds, the results of the t-test and a histogram of the FOI areas. At first, we can see that the signals generated for the verified FOIs are stronger (larger amplitude) in the response comparing to the ones which were not verified. The 'unwanted' elements that are included in the FOIs made the response weaker with values smaller than expected. This smoothing or contamination effect is stronger when the proportion between areas of the FOI and 'unwanted' elements included inside increases, and when activities in the "unwanted" lead to different manifestations or take place at different times. With a high proportion of those elements within a FOI, a random response will be added to all the signals, leading to a smoothed average response.

The largest change increases from 0.094 to 0.121 after the FOI boundary modification. The probability of being above zero increases from 0.90 to 0.93. This corresponds to a reduction in the risk for false negatives from 10 to 7%, a reduction of 30%. The changes expressed in averaged values for the datasets might not appear substantial for the entire population, which is composed of FOIs with various reductions of their areas. Therefore, to see the influence of 'unwanted' elements in the FOIs, the results are also considered according to the sizes of the reduced areas in Figure 29.

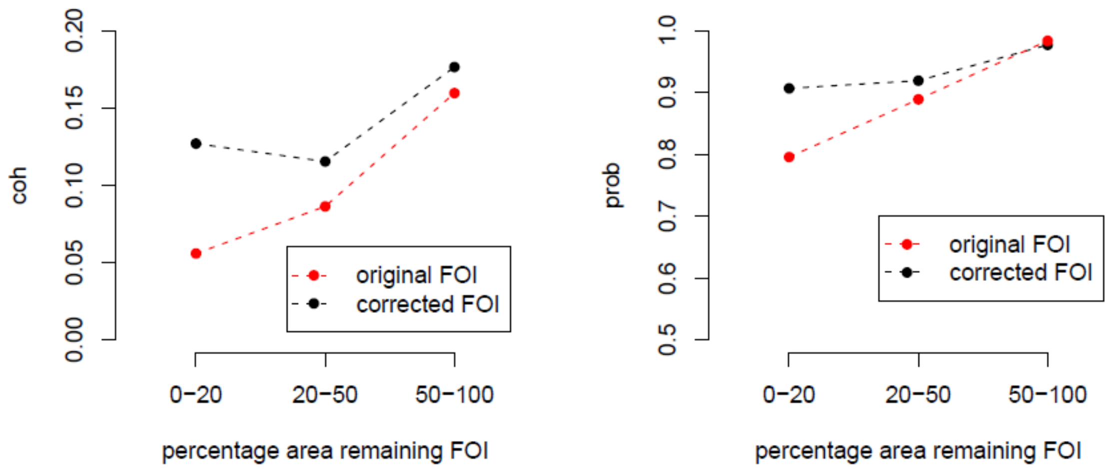
Figure 28. Comparison of results from two datasets, a) using the FOIs with boundary issues b) and corrected FOI boundaries. This figure provides for each dataset the following: mean and confidence bands for S1_coh for mowing, mean with p-value of t-test for the two cases, and histogram distribution of the FOIs areas.



Source: GTCAP

The average changes of signal responses observed 4 weeks after activity have been separated into three classes, depending on the percentage of the original FOI that is actually left after modifying the boundaries (Figure 29). The left panel shows the change of signal values and the right panel shows the changes in probabilities, both as a function of the remaining percentage of the FOI between two datasets (before and after changes).

Figure 29. Changes in signal and probabilities as a function of the size of the changed FOI for mowing activities (hay and silage), red: based on the original geometries of the FOIs dataset, black: based on the corrected FOIs geometries. The x-axis shows the percentage of remaining area of the FOI



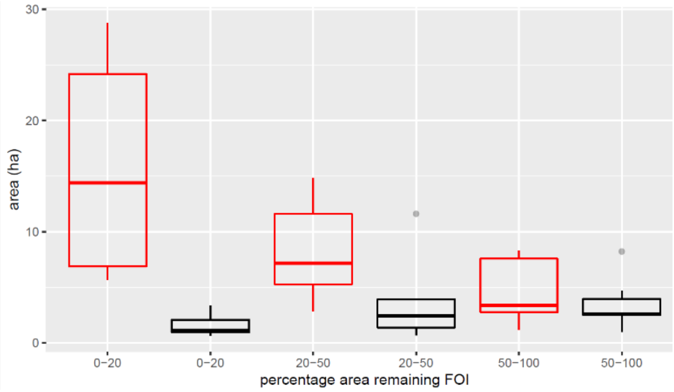
Source: GTCAP

First of all, the differences between the results from the original FOIs and the modified boundaries are not surprisingly a function of the area that has been removed in the correction process. In other words, the signal response decreases when the proportion of ‘unwanted’ elements in the FOI increases. The difference in signal change after an activity between the two datasets for the same events observed on the ground, relative to the modified ones are 54%, 27% and 18%, respectively, for the three area classes, with the largest difference for the largest area reduction.

There is a relatively small difference in signal change (only permanent grassland also in the removed parts), and almost no difference in probabilities when the corrected FOI contains more than 50% of the original FOI. This proportion may change for crops enclosed in one geometry with different land cover manifestation (i.e. bare soil vs green vegetation), but this is not discussed further in this chapter. The differences increase when the corrected FOI is less than half of the original FOI.

The second noticeable feature of the figure is that both the signal changes and the probabilities are lower for the FOIs that have been through larger modifications. This dependency might be related to strength and repeatability of the signal pattern which is suppressed (i.e. compare time series a) and c) of Figure 27) by a larger proportion of 'unwanted' elements inside the FOI. The proportion of the areas can be seen in Figure 30, which shows boxplots of the areas before and after modification of the FOI geometries. As smaller FOIs are more likely to be affected by random effects, it is therefore more likely that the dependency is caused by the FOI sizes rather than the percentage changes.

Figure 30. Boxplots of areas before (red) and after (black) correction of FOI geometries



Source: GTCAP

The influence of FOIs' design issues on the signals will depend on several things. We have not analysed all of this in detail, but from the analyses and an understanding of how the signals are created and affected, we can assume the following:

— We are likely to detect larger changes in the signals with increasing size of the removed part. Additionally, the magnitude of the change depends on the type of land cover and activity on the part that was removed.

— This effect is less if similar activities take place at similar times, larger if the no activities take place in the removed parts.

— The effect is not likely to affect all signals in the same way and it depends on the sensor characteristics and sensibility.

This chapter concludes the first analyses results of signal quality in the context of CbM, further investigation on larger sample set with a better quantification of area under 'unwanted' elements are needed to estimate the limits for which the FOI can show the desired behaviour.

7 Ground truth data sources and collection protocols

To make the CbM a reliable and performant tool, the marker need to be optimised based on ground truth data. The observations are collected by visiting and inspecting the parcels/FOIs in regular time intervals to capture the phenomenon of interest or traces of farmers' activities. Ideally, farmers can also systematically provide ground information, captured according to a dedicated data collection protocol (see examples in chapter 4). The date and type of an activity is the key information in the CbM development and implementation. This provides an accurate time stamp of an event observed on the ground, enabling correct interpretation of signal behaviour. Ground observations are also crucial in understanding the potentials and limitations of remotely sensed imagery in capturing specific phenomena and to assess the detection capabilities of different markers that might be considered in the CbM development process.

Dedicated survey:

Several data sources can provide information supporting signal analyses, as discussed before, but first and most valuable is a dedicated survey where fit-to-purpose data are collected. The survey needs to be planned with caution, and the design process should include at least the following elements:

— scope of survey, including a **list of activities**/events/land cover manifestations to observe,

— frequency of observations, **field revisit time** (i.e. weekly, monthly, or other, but optimise for the type of activity observed),

— type of observations to collect during field inspection, (i.e. crop type, crop phenological status, **activity type and day of observation**) and/or additional documentation, i.e. geo-tagged photos. A good practice would be to create a list of the codes with description that the field inspectors will annotate,

— survey extent, **spatial coverage of data collection** (i.e. the specific region or crop), often the parcels might be clustered to optimise the cost and time factor (limit long travels between parcels),

— sample size (**number of parcels to visit**) depends on the population size (total number of parcels), the number of different activities to detect and the budget,

— duration of survey (i.e. vegetative season, calendar year, crop specific stage), **adjusted to type of phenomena or activity to be observed.**

Careful planning is crucial because the complete survey results will only be available at the end the season, if systematic and long term ground observations are required. Missing or too scarcely collected data cannot be easily substituted or complemented. It is also important to include verification of the parcel/FOI boundaries in the survey protocol, so that field data can indisputably be associated to the proper FOI geometry.

Other potential data sources:

Most regions and PAs will have existing field surveys, designed for another purpose, such as the OTSC and RFV.

— The advantages of using already existing control data are obvious, not only from economical view point, but also that the quality of the available information is usually based on established protocols.

— The challenge lays in matching information content, time frame and the frequency of observations with the needs of the signal analysis. It is also worth noticing that the existing survey data are characterised with a sample size and spatial distribution designed for another purpose, i.e. for control of farmer's applications and thus the local practice variability might be not captured.

Some regions/MSs will have other image based sources providing single observations, either acquired on the ground (geo-tagged photos), aerial oblique surveys or typical wall-2-wall aerial campaigns providing systematic orthophoto updates.

— Those data sources do not deliver ready-to-use observation, but require interpretation and further data collection to derive the date of the activity and/or land cover status/change observations.

However, due to superior image resolution, such datasets are particularly useful i.e. to distinguish between mowed/grazed grassland. The images will show observable presence of animals/hay residuals, not available in Sentinel data.

— Another advantage of aerial acquisitions is their coverage, which may vary between a small region and part of the country. However, the data are collected in a short time span. The usage of an image-based data source to collect the ground truth observation might be complex and often it may not be possible to derive precisely the date of the activity, as the image acquisition times are not synchronised with farmer's activity calendar. Thus, deriving information about the land cover manifestation at the time of image capture is more feasible than indicting an exact date of an activity.

The geotagged photos, similarly to other image-based sources, provide information snapshots taken from a known position with limited field of view and a single direction (pointing).

— An advantage of this approach is that it takes little effort to get more photos from different positions and/or directions to provide more context and information.

— The images need to be further analysed to retrieve required observations (i.e. about practice, or a specific land cover presence) and thanks to ground orientation are intuitive to interpret. For more information about geotagged photos (Sima et al. 2020).

Ground truth can be observed by inspectors with a range of different frequencies, if they are not based on trustworthy reporting of the exact days of activities. The frequencies will affect the type of activity/land cover manifestation that can actually be reported.

— Daily (or multiple weekly) observations are necessary for detection of rapid changes, but very labour intensive and difficult to set up for larger sample collection covering all seasons. For an average arable crop scenario and the corresponding number of farmer's activities of interest. Usually not necessary for CbM.

— Weekly observations will be considerably less labour intensive than daily, but still give data that are sufficient for most CbM applications. Such data collection may cover the entire season/year, or just a selected period of time (e.g. a specific vegetation growth stage, a farmer activity - crop harvest or post-harvest condition/treatment of the field). Optimal for most of CbM applications.

— Monthly observations can be used for analyses of persistent phenomena, and in some cases used to confirm or rebut an activity reported by the farmer. The date of an activity that happened between the inspections is mostly not retrievable. However the interpretation of the conditions on the ground might indicate whether a given activity happened in the period (e.g. sowing, mowing or harvest). More applicable in quality assessment context than in signal analyses.

— Annual observation, the single visit approach gives good and reliable data on persistent of a phenomena, but has limitations in spotting wanted activity and does not provide accurate information about the date, unless the activity was shortly before or during the day of inspection. The usability of such observation highly depends on the date of inspection. Applicable in some aspects in quality assessment context.

The existing data are good starting point in the process, but fit-for-purpose ground observations, with detailed enough field observations, provide much more possibilities for data analyses and potentially will return the investment faster and give more reliable results when the process is automated.

8 Discussion and practical considerations

The Copernicus satellite data offer a great opportunity for automation in the checks by monitoring programs being implemented by the Paying Agencies. Their implementation work was preceded by research and development efforts, and some of those were published. There are examples of development work that supports cloud based solution for CbM (i.e. Lemoine and Anastasakis, 2022), large scale detection of selected activities in (i.e.; Lange et al., 2022; Vroey et al., 2022) or crop classification studies (i.e. Beriaux et al., 2021), which does not provide direct judgment on activity trough a marker. Instead they provide a crop class along with traditional control with a remote sensing approach (part of OTSC). Less attention was given to understanding the relationship between the activity and corresponding signal (i.e. Tamm et al., 2016; Voormansik et al., 2020). There is not much published documentation that provides methodological insight of a generic marker creation process by linking the practice and the signal. Neither has there been much work on how to compare suitability of a large group of indices and descriptors, and select one or more with the highest discrimination rates. This research aimed to bridge this gap.

This report elaborates an analysis workflow that starts from local information guided by scenario and agricultural practices, through data extraction and reaches an optimised selection of signals for marker development. An important aspect is that such a tiered and reductive workflow should be performed step-by-step, and after each step consider whether the results support further analyses of the activity for the data available. The result of such analyses highly depends on quality and level of details of input data, especially the in-situ and signal data, any incomplete or imprecise data may reduce the quality of the output and the conclusions that can be drawn from the analyses.

The implementation of a CbM system needs local adaptations. These should respond to local variabilities starting with user information needs where different schemes and interventions are considered. The practices and landscapes vary not only locally, but often across a single region or a country. Other elements are the weather and climatic conditions which continually change across the season and more importantly differ between the years. And finally, from a system point of view, there is a series of variabilities caused by the FOIs characteristic, signals and marker selection, the detection parameters and the corresponding timing.

The proposed approach for a workflow supports those local adaptations with usage of ground truth data and analyses driven by FOI location, practice or a specific activity.

8.1 Signal

The analyses are based on signals derived from Sentinel-1 and Sentinel-2 data streams which are considered the primary image source for the CbM. However, the proposed approach can theoretically be applied on data acquired by other sensors, because it operates directly on a signal time series. However, there are currently no other constellations that can offer the same combination of resolution and revisit time free of charge.

Still, there are some potential complementary data sources, for example;

- Planet Labs[11] can offer higher revisit times and high resolution images in the visual spectre plus infra-red, but as commercial product. A selected use case of this data applicable to the CbM workflow were discussed during the 26th Mars conference 2022[12].
- The Landsat programme also offers free data, but with a slightly lower revisit time for the visual bands (8 vs 5 days for Sentinel-2 after the launch of Landsat 9[13]) and does not have SAR (Li and Roy, 2017).
- The Pleiades Neo constellation will offer high resolution images with a possibility of high revisit times at targeted locations (Chouteau et al., 2022). It will be a commercial product.

The number and the capabilities of land observing satellites is rapidly improving (Belward and Skøien, 2015). It should also be noted that whereas the raw data of Landsat and Sentinels are free, some

[11] Planet webpage https://www.planet.com/ (accessed 03/11/2022)

[12] 26th Mars conference https://marswiki.jrc.ec.europa.eu/wikicap/index.php/Barcelona,_12-14_September (accessed 03/11/2022)

[13] NASA webpage https://landsat.gsfc.nasa.gov/satellites/landsat-9/ (accessed 03/11/2022)

tools to make it easier to access these and fit-for-purpose products for e.g. agriculture might not be free. It is therefore important to promote open-source solutions (i.e. Lemoine and Anastasakis, 2022).

The quality of the signal depends on a series of factors and have influence on signal understanding, marker development, detection result and the overall monitoring system performance. From a source of image point of view, data are acquired with a certain acquisition plan that results in a coverage that vary across the territories. The data have different availability due to revisits time, system health (see chapter 3.1) or, i.e. for optical sensors like Sentinel-2, the cloud cover that create gaps in the data and cause irregularly spaced time series data (see chapter 3.2). Pre-processing of the remote sensing dataset in the context of CbM is necessary to correctly identify the acquisitions with cloud cover, shading in the images for each FOI.

As a signal is derived for an FOI, a feature geometry guides the pixel selection, thus any inclusion of pixels that do not belong to the specific agricultural practice are reflected in the value derived for each calculation. The influence of such elements is proportional to their areas, bigger areas will give more deviations from the expected signal (see chapter 6.3).

The shape and area of the FOI determine the number of resulting pixels (i.e. Sentinel-2, ground sample distance of 10m and 20m, depending on band) available for the monitoring. More complex shapes of parcels (especially elongated) will result in a lower proportion of pixels selected in comparison to a square/rectangular shaped parcel of the same size. Due to this effect, some FOIs may become not monitorable using the given sensor data. The resulting number of pixels available per FOI drives the calculation and reliability of the statistical descriptors, thus the signal itself. For the FOIs with less than 3 pixels available for the monitoring (see chapter 3.3) a separate handling in the CbM system should be undertaken to capture the resulting signal behaviour (i.e. for a single pixel representation a value of the signal instead of mean/median). Other studies, such as Vajsová et al. (2020) suggested for the Sentinel-2 a minimum of 8 full pixels, and that the application of a negative buffer should not decrease the number of pixels with more than 60%.

In a single analysis, a large number of potential signals (image sources and descriptors) can be considered for a land cover manifestation or an activity detection. The fact that we only found the mean and the quantiles to be good descriptors for the cases analysed in this report (see chapter 6.1), does not exclude that the other descriptors can be of good use for other cases or in other regions. Whereas it could be possible to use all of these in a machine learning procedure, too many explanatory variables (signals) relative to the number of activities can cause overfitting, i.e., that we find a too complex relationship, which fits very well with the training data, but is likely to perform poorly on a verification data set (Hawkins, 2004). The methods presented here can therefore also be of good use to reduce the amount of possible input to a machine learning algorithm.

8.2 Ground truth data

Ground truth data, despite being important for both understanding the signal behaviour and quality control of the outcome, is often a fairly neglected element in the CbM development. In the course of the CbM outreach 2021 activities, only one out of 17 active participants already had a dedicated ground observation set to support CbM analyses. A few participants were in the process of collecting it in the course of 2021, and the majority considered it for their future task list. The existing datasets available to all Member States, e.g. OTSC/RFV, control with remote sensing or geo-tagged photos captured during various controls, are not always fit-for-purpose. In most cases, they needed enhancement because they do not capture directly the farmers' activities relevant for the need of signal analysis or provide information in sub-optimal timing and limited frequency, such as once a year.

When available, the ground truth data with a specific calendar date (or a period) for a given farmer's activity facilitates correct understanding of land cover manifestations (or their change) and allows for linking them with the corresponding image signal behaviour. Such analyses allow for selection of the optimal signal for each activity type. There are already experiences from a well-designed and complete dataset[14] with weekly ground observations collected to support CbM development. In the design of a survey for CbM development or quality assessment, the following elements require careful planning: future functionalities (type of activity/event that needs to be confirmed), required observation

[14] Reported by Eoin Dooley, DAFM, Ireland https://marswiki.jrc.ec.europa.eu/wikicap/images/1/1b/IE_Monitoring_%26_AG_Practices_05_04_22.pdf (accessed on 03/11/2022)

frequency, type of information collected, spatial coverage, and sample size. Careful planning is crucial because the complete survey result will only be available at the end the season. Where multiple ground observations are required, missing or too scarcely collected data cannot be easily substituted or complemented. The fit-for-purpose ground data with detailed field observations provides more possibilities for data analyses, and potentially will return the investment faster. Potential data collection with the help of farmers should be considered.

8.3 Signal analyses method

The activity centred approach made it easy to join a group of activities and analyse the similarities between them. The analyses were done on a weekly basis, which poses both advantages and disadvantages compared with using the exact acquisition dates (see chapter 5.1). The signals were centred around a known activity week. With exact knowledge of all activity dates, it would have been possible to centre the signals around this single date (i.e. Voormansik et al., 2020). The choice of the approach depends on the characteristics of available data.

The significance analyses (t-test) test whether it could be possible to find a typical signal for an activity type, or whether an activity type could be classified as non-detectable. Where the sample size is too small (typically less than 10), a non-significant result here does not necessarily mean that the activity type does not produce a particular signal behaviour. Similarly, a significant t-test does not mean that it is easy to identify an activity, or a lack of an activity. For a marker to be developed, it is also necessary to select a land cover manifestation that is detectable with earth observation data. This depends on the variability of the signal after an event, as described in chapter 5.3. The variability of the signal is partly depending on the quality of the FOIs in the sample. Normally, FOIs with verified boundaries are expected to give lower variability, although there was not much difference in the example in chapter 6.3. It would be worth to repeat these analyses for another dataset, to see if this effect can be confirmed, or if it a result of random behaviour.

Following the workflow, signals and descriptors can be ranked according to the probabilities of exceeding a set threshold. As an example we have presented results for two thresholds – zero and one which is relative to the maximum change of the averaged signal. The latter is a more realistic estimate of the ability to detect a possible event when using a marker for detection. Using this approach it would also be possible to further test settings of marker parameters. The marker parametrisation is not discussed here, as it considered a separate research topic.

Small differences in the probabilities between signals are likely caused by influence of local characteristics of the test sites rather than one signal being significantly better than others. However, the larger differences between signals are likely the result of different abilities to detect a certain activity.

The pairwise correlation analyses of the signals is applicable when simultaneously developing several markers for the detection of a single activity. The idea is that more than one signal will strengthen the detection rate of the markers. Highly correlated signals (positive or negative) used to observe the same land cover manifestation state/change will not reinforce the reliability of the detection (see chapter 5.5). However, signals with low correction may add non-redundant information and thus increase the detection rate. Using signals from independent sensors will often be advantageous, as they have various sensors characteristics and properties. Detailed analyses on quantifying the information gain in the detection process have not been conducted.

8.4 Lessons learned and suggestions for implementation

The lessons learned during the method development and the analyses of signals for different activities cover several aspects.

Preparation:

— Understand the local practice starting from a single activity or land cover manifestation (or its change) and describe it from the field phenomena perspective. Identify the ones that are potentially observable in satellite data and may provide evidence that the farmer completed the required activities.

— Use the CbM template (Zielinski et al., 2022a) to document the processes and components (scenario, user information needs, rules, signal, and more).

- Collect ground truth data to optimise the monitoring results performance and reliability. Incorrectly dated events or poorly delineated FOIs will make it harder to find the typical signal behaviour after an activity.
- Plan and collect the ground truth data, fit-to-purpose, that are relevant for the analyses (or question to be answered). According to the budget plan, consider defining: scope and type of information needed to be collected, the frequency of observations, sample size, spatial coverage of data collection and duration of the survey to assure data fit-to-purpose.
- Detailed and high quality ground truth data gives a wider spectrum of analyses that can be performed. As the data require less adaptation or manual enrichment, this should give higher reliability of the judgment.

Data analyses:

- Plot and analyse individual time series of signals together with the reported activity dates, and evaluate whether they reflect the expected behaviour.
- Adapt the analysis method to the ground truth data available. For each question to investigate select appropriate analysis and statistical method.
- Ensure that FOIs contain at least 3 pixels for the descriptors to be valid.
- Calculate the signal statistics (based on image bands/products and their derivatives) for each FOI from these selected pixels only. The proposed workflow gives an indication of an order of analyses to perform signal ranking for an activity. Different research questions, like the ones in chapter 6, might require some adaptations.
- Make sure that the analyses are performed with an adequate level of statistical rigor, including sample size. Uncertainty is likely to hamper the analyses if less than 10 samples of each activity type is available, whereas a recommendation is 20-30 samples for each activity type.
- Test the local variability in the FOI data. If the sample size of ground truth allows, consider dividing the analyses into meaningful segments driven by a location, a specific practice or an activity type.
- Test the results of the analyses against the logic of expected characteristic from the practice and remote sensing knowledge.
- Evaluate data quality and output of analysis at every step: remember that nonsense input produces nonsense output.
- Document relevant components of the practice and analyses.

Analytical environment implementation:

- Be aware of data formats while changing the environments (i.e. as simple as format of a date)
- Ensure consistent naming of bands, indices etc.
- An R[15] package that can handle most of the analyses in this report will be available at the start of 2023.

Effort of the analyses; lessons learnt.

- The main effort of the analyses is in-situ data gathering and preparation. If ground truth observations are collected via a planned ground survey, often a full season is needed to complete this part.
- The signal input data has to be extracted from remote sensing images. The process can be automatized, but organizing data will demand some effort.
- Most of the functionalities can be found in an R package. However, building the analytical environment will require some effort and specific expertise.

[15] The R Project for Statistical Computing, https://www.r-project.org/ (accessed 03/11/2022)

— Once an analytic environment is prepared, the processing time will depend on the number of samples and signals per activity to be processed, but this will be automatized.

— Analyses of the results can require a substantial part of the effort. Careful planning and knowledge driven signal selection will reduce time necessary for calculation and analyses of the outputs.

8.5 Possible outlook

The workflow in this report describes the steps for using remote sensing data for marker development in the context of the CbM. By following the described methodology, a ranking of signals for an activity or a land cover manifestation is created, based on probabilities and statistical significance of the difference observations. A next step could be to include also the pairwise correlations of signals in the automated framework, which is currently a separate step. The logical extension of this work would be analyses related to marker parametrization and quality assessment of detection based on various datasets.

Where it becomes apparent that some activities are not applied to the entire FOI area at the same time a set of analyses that deals with intra-parcel/FOI diversity could solve the detection problem. Development of spatial-temporal variability markers could be another research topic. From the temporal perspective, this research has looked at the phenomena with duration from a few days to several weeks, thus an extension to annual and multi-annual scenario verification might also be considered another development challenge.

This report is about the approach and its description, a further attention should be given to application of the presented approach and focus on marker development for other than agricultural land cover manifestations and corresponding activities, but this would require availability of additional data and further collaboration with the Member States.

9 Conclusion

This report aims to bridge a gap in the methodological background for the CbM marker development, by a workflow starting from the data and its quality (scenario definition and local practices, sensor availability, observed activities and correctness of FOIs), going through different visual and statistical analyses of individual and multiple time series, resulting in a rank and evaluation of the potential of different signals and descriptors. Identification of the optimal signals and descriptors is a requirement for development of markers, which are later used for activity/manifestation detection purposes. The report shows how the same framework can be used also for analysing the quality of FOI delineation and for analyses of differences between different types of activities. This framework also helps identifying the monitorable and non-monitorable cases of land cover manifestation or corresponding farmer's activities for a set of signals or a sensor.

The purpose of this report is more to describe a workflow and the relationships between the single steps of the analyses rather than provide final numerical results for a set of popular markers. These will depend on the type of manifestations, the availability of data and local factors, ranging from soil quality to weather conditions. An unlimited number of signals and descriptors can be analysed through this workflow, standardizing the way how the optimal signals for marker development are selected.

The proposed signal analysis workflow was developed as a method to optimise the use of remote sensing data in CbM development. However, it is not limited to any specific activity or land cover manifestation in the agricultural sector. It can be applied at any data location for any other applications. The only assumptions are that the entire area of interest can be assumed to undergo the same processes and the dates of land cover manifestations or ground events are known. This signal selection approach is equally applicable at the initial design phase as well as at the stage of system upgrade or improvement.

References

Bauer-Marschallinger, B., Cao, S., Navacchi, C., Freeman, V., Reuß, F., Geudtner, D., Rommen, B., Vega, F.C., Snoeij, P., Attema, E., Reimer, C., Wagner, W., 2021. The normalised Sentinel-1 Global Backscatter Model, mapping Earth's land surface with C-band microwaves. Sci. Data 8, 277. https://doi.org/10.1038/s41597-021-01059-7

Bekkema, M., Eleveld, M., 2018. Mapping Grassland Management Intensity Using Sentinel-2 Satellite Data. GI_Forum 1, 194–213. https://doi.org/10.1553/giscience2018_01_s194

Belward, A.S., Skøien, J.O., 2015. Who launched what, when and why; trends in global land-cover observation capacity from civilian earth observation satellites. Glob. Land Cover Mapp. Monit. 103, 115–128. https://doi.org/10.1016/j.isprsjprs.2014.03.009

Beriaux, E., Jago, A., Lucau-Danila, C., Planchon, V., Defourny, P., 2021. Sentinel-1 Time Series for Crop Identification in the Framework of the Future CAP Monitoring. Remote Sens. 13. https://doi.org/10.3390/rs13142785

Breger, P., 2017. The Copernicus full, free and open data policy.

Chouteau, F., Gabet, L., Fraisse, R., Bonfort, T., Harnoufi, B., Greiner, V., Le Goff, M., Ortner, M., Paveau, V., 2022. JOINT SUPER-RESOLUTION AND IMAGE RESTORATION FOR PLÉIADES NEO IMAGERY. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLIII-B1-2022, 9–15. https://doi.org/10.5194/isprs-archives-XLIII-B1-2022-9-2022

Devos, W., Fasbender, D., Lemoine, G., Loudjani, P., Milenov, P., Wirnhardt, C., 2017. Discussion document on the introduction of monitoring to substitute OTSC - Supporting non-paper DS/CDP/2017/03 revising R2017/809. https://doi.org/10.2760/258531

Devos, W, Lemoine, G., Milenov, P., Fasbender, D., 2018. Technical guidance on the decision to go for substitution of OTSC by monitoring. Publications Office. https://doi.org/10.2760/693101

Devos, Wim, Lemoine, G., Milenov, P., Fasbender, D., Sima, A., Griffiths, P., Loudjani, P., Wirnhardt, G., 2018. Second discussion document on the introduction of monitoring to substitute OTSC : rules for processing applications in 2018-2019. Publications Office. https://doi.org/10.2760/344612

Devos, W., Milenov, P., 2015. Applying tegon, the elementary physical land cover feature, for data interoperability. CRC Press, Columbia, Ohio (United States). p. 243-270, JRC92804

Devos, W., Milenov, P., Zielinski, R., Vajsová, B., 2021a. ANNEX IV Ground data, version 1.2 of Checks by Monitoring Quality Assurance Framework. https://marswiki.jrc.ec.europa.eu/wikicap/images/c/cb/1_2_Annex_IV_Groud_data.pdf

Devos, W., Sima, A., Milenov, P., 2021b. Conceptual basis of checks by monitoring (JRC Technical Report No. JRC127678). European Union.

Diek, S., Fornallaz, F., Schaepman, M.E., De Jong, R., 2017. Barest Pixel Composite for Agricultural Areas Using Landsat Time Series. Remote Sens. 9. https://doi.org/10.3390/rs9121245

ESA, 2021. Sentinel High Level Operations Plan (HLOP) (No. COPE-S1OP-EOPG-PL-15-0020).

European Commission, 2013. Commission Delegated Regulation (EU) No 1159/2013 of 12 July 2013 Supplementing Regulation (EU) No 911/2010 of the European Parliament and of the Council on the European Earth Monitoring Programme (GMES) by Establishing Registration And Licensing Conditions for GMES Users and Defining Criteria for Restricting Access to GMES Dedicated Data and GMES Service Information. Off. J. Eur. Union L 309/1.

European Court of Auditors, 2020. Using new imaging technologies to monitor the Common Agricultural Policy: steady progress overall, but slower for climate and environment monitoring, Special Report 04/2020, doi:10.2865/46869.

Ferretti, A., Monti-Guarnieri, A., Prati, C., Rocca, F., Massonet, D., 2007. InSAR Principles - Guidelines for SAR Interferometry Processing and Interpretation. ESA Train. Man. 19.

Filgueiras, R., Mantovani, E.C., Althoff, D., Inácio, E., Filho, F., França Da Cunha, F., 2019. Remote sensing Crop NDVI Monitoring Based on Sentinel 1. https://doi.org/10.3390/rs11121441

Gao, B.C., 1996. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sens. Environ. 58, 257–266. https://doi.org/10.1016/S0034-4257(96)00067-3

Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languille, F., Alhammoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., Fernandez, V., 2017. Copernicus Sentinel-2A Calibration and Products Validation Status. Remote Sens. 9. https://doi.org/10.3390/rs9060584

Gitelson, A.A., Kaufman, Y.J., Merzlyak, M.N., 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens. Environ. 58, 289–298. https://doi.org/10.1016/S0034-4257(96)00072-7

Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. Geophys. Res. Lett. 32, 1–4. https://doi.org/10.1029/2005GL022688

Griffiths, P., Nendel, C., Pickert, J., Hostert, P., 2020. Towards national-scale characterization of grassland use intensity from integrated Sentinel-2 and Landsat time series. Remote Sens. Environ. 238, 111124. https://doi.org/10.1016/j.rse.2019.03.017

Hawkins, D.M., 2004. The Problem of Overfitting. J. Chem. Inf. Comput. Sci. 44, 1–12. https://doi.org/10.1021/ci0342472

Henrich, V., Götze, C., Jung, A., Sandow, C., Thürkow, D., Cornelia, G., 2009. Development of an online indices database: Motivation, concept and implementation, in: 6th EARSeL Imaging Spectroscopy SIG Workshop Innovative Tool for Scientific and Commercial Environment Applications. EARSeL, Tel Aviv, Israel.

Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sens. Environ. 83, 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

Huete, A.R., 1988. A soil-adjusted vegetation index (SAVI). Remote Sens. Environ. 25, 295–309. https://doi.org/10.1016/0034-4257(88)90106-X

Khabbazan, S., Vermunt, P., Steele-Dunne, S., Ratering Arntz, L., Marinetti, C., van der Valk, D., Iannini, L., Molijn, R., Westerdijk, K., van der Sande, C., 2019. Crop Monitoring Using Sentinel-1 Data: A Case Study from The Netherlands. Remote Sens. 11. https://doi.org/10.3390/rs11161887

Lange, M., Feilhauer, H., Kühn, I., Doktor, D., 2022. Mapping land-use intensity of grasslands in Germany with machine learning and Sentinel-2 time series. Remote Sens. Environ. 277, 112888. https://doi.org/10.1016/j.rse.2022.112888

Lemoine, G., Anastasakis, K., 2022. Repository of the JRC-CbM for the CAP. https://github.com/ec-jrc/cbm

Li, J., Roy, D.P., 2017. A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. Remote Sens. 9. https://doi.org/10.3390/rs9090902

Louhaichi, M., Borman, M.M., Johnson, D.E., 2008. Spatially Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat. http://dx.doi.org/10.1080/10106040108542184 16, 65–70. https://doi.org/10.1080/10106040108542184

Ma, Q., Chai, L., Hou, F., Chang, S., Ma, Y., Tsunekawa, A., Yunxiang, C., 2019. Quantifying Grazing Intensity Using Remote Sensing in Alpine Meadows on Qinghai-Tibetan Plateau. Sustainability 11, 417. https://doi.org/10.3390/su11020417

Mann, P.S., 2018. Introductory Statistics, 9th Edition. ed. Wiley.

Milenov, P., Vajsová, B., Voican, G.-M., Acquafresca, L., Anastaskis, K., Lemoine, G., 2021. Assessing the Validity of the Feature of Interest in the context of the CAP Checks by Monitoring (JRC Technical Report No. JRC123711). European Union.

Moore, K.J., Moser, L.E., Vogel, K.P., Waller, S.S., Johnson, B.E., Pedersen, J.F., 1991. Describing and Quantifying Growth Stages of Perennial Forage Grasses. Agron. J. 83, 1073–1077. https://doi.org/10.2134/agronj1991.00021962008300060027x

Nasirzadehdizaji, R., Cakir, Z., Balik Sanli, F., Abdikan, S., Pepe, A., Calò, F., 2021. Sentinel-1 interferometric coherence and backscattering analysis for crop monitoring. Comput. Electron. Agric. 185, 106118. https://doi.org/10.1016/j.compag.2021.106118

Nguyen, C.T., Chidthaisong, A., Kieu Diem, P., Huo, L.-Z., 2021. A Modified Bare Soil Index to Identify Bare Land Features during Agricultural Fallow-Period in Southeast Asia Using Landsat 8. Land 10, 231. https://doi.org/10.3390/land10030231

Proakis, J.G., Manolakis, D.K., 2006. Digital Signal Processing (4th Edition). Prentice-Hall, Inc., USA.

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., Sorooshian, S., 1994. A modified soil adjusted vegetation index. Remote Sens. Environ. 48, 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. Remote Sens. Environ. 55, 95–107. https://doi.org/10.1016/0034-4257(95)00186-7

Rouse, J.W., Haas, R.H., Scheel, J.A., Deering, D.W., 1973. Monitoring Vegetation Systems in the Great Plains with ERTS, in: In Proceedings of the 3rd Earth Resource Technology Satellite (ERTS) Symposium. Washington, DC, USA, pp. 48–62.

Schneider, F., Don, A., Hennings, I., Schmittmann, O., Seidel, S.J., 2017. The effect of deep tillage on crop yield – What do we really know? Soil Tillage Res. 174, 193–204. https://doi.org/10.1016/j.still.2017.07.005

Serco, 2021. Copernicus Sentinel data access annual report 2020.

Sima, A., Loudjani, P., Devos, W., 2020. Use of geotagged photographs in the frame of Common Agriculture Policy checks (JRC Technical Report No. JRC120223). European Union.

Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batič, M., Frantz, D., Gascon, F., Gómez-Chova, L., Hagolle, O., López-Puigdollers, D., Louis, J., Lubej, M., Mateo-García, G., Osman, J., Peressutti, D., Pflug, B., Puc, J., Richter, R., Roger, J.-C., Scaramuzza, P., Vermote, E., Vesel, N., Zupanc, A., Žust, L., 2022. Cloud Mask Intercomparison eXercise (CMIX): An evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. Remote Sens. Environ. 274, 112990. https://doi.org/10.1016/j.rse.2022.112990

Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., Mochizuki, K., 2018. Crop classification from Sentinel-2-derived vegetation indices using ensemble learning. J. Appl. Remote Sens. 12, 026019.

Sripada, R., 2005. Determining in-season nitrogen requirements for corn using aerial color-infrared photography (Ph.D. dissertation). North Carolina State University, USA.

Sudmanns, M., Tiede, D., Augustin, H., Lang, S., 2020. Assessing global Sentinel-2 coverage dynamics and data availability for operational Earth observation (EO) applications using the EO-Compass. Int. J. Digit. Earth 13, 768–784. https://doi.org/10.1080/17538947.2019.1572799

Szigarski, C., Jagdhuber, T., Baur, M., Thiel, C., Parrens, M., Wigneron, J.P., Piles, M., Entekhabi, D., 2018. Analysis of the Radar Vegetation Index and Potential Improvements. Remote Sens. 2018 Vol 10 Page 1776 10, 1776. https://doi.org/10.3390/RS10111776

Tamm, T., Zalite, K., Voormansik, K., Talgre, L., 2016. Relating Sentinel-1 Interferometric Coherence to Mowing Events on Grasslands. Remote Sens. 8. https://doi.org/10.3390/rs8100802

Tarrio, K., Tang, X., Masek, J.G., Claverie, M., Ju, J., Qiu, S., Zhu, Z., Woodcock, C.E., 2020. Comparison of cloud detection algorithms for Sentinel-2 imagery. Sci. Remote Sens. 2, 100010. https://doi.org/10.1016/j.srs.2020.100010

Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E., Potin, P., Rommen, B., Floury, N., Brown, M., Traver, I.N., Deghaye, P., Duesmann, B., Rosich, B., Miranda, N., Bruno, C., L'Abbate, M., Croci, R., Pietropaolo, A., Huchler, M., Rostan, F., 2012. GMES Sentinel-1 mission. Sentin. Missions - New Oppor. Sci. 120, 9–24. https://doi.org/10.1016/j.rse.2011.05.028

Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sens. Environ. 8, 127–150. https://doi.org/10.1016/0034-4257(79)90013-0

Vajsová, B., Fasbender, D., Wirnhardt, C., Lemajic, S., Devos, W., 2020. Assessing Spatial Limits of Sentinel-2 Data on Arable Crops in the Context of Checks by Monitoring. Remote Sens. 12, 2195. https://doi.org/10.3390/rs12142195

Veloso, A., Mermoz, S., Bouvet, A., Toan, T.L., Planells, M., Dejoux, J.-F., Ceschia, E., 2017. Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications. Remote Sens. Environ. 199, 415–426. https://doi.org/10.1016/j.rse.2017.07.015

Voormansik, K., Zalite, K., Sünter, I., Tamm, T., Koppel, K., Verro, T., Brauns, A., Jakovels, D., Praks, J., 2020. Separability of Mowing and Ploughing Events on Short Temporal Baseline Sentinel-1 Coherence Time Series. Remote Sens. 12, 3784. https://doi.org/10.3390/rs12223784

Vreugdenhil, M., Navacchi, C., Bauer-Marschallinger, B., Hahn, S., Steele-Dunne, S., Pfeil, I., Dorigo, W., Wagner, W., 2020. Sentinel-1 Cross Ratio and Vegetation Optical Depth: A Comparison over Europe. Remote Sens. 2020 Vol 12 Page 3404 12, 3404. https://doi.org/10.3390/RS12203404

Vroey, M.D., Vendictis, L. de, Zavagli, M., Bontemps, S., Heymans, D., Radoux, J., Koetz, B., Defourny, P., 2022. Mowing detection using Sentinel-1 and Sentinel-2 time series for large scale grassland monitoring. Remote Sens. Environ. 280, 113145. https://doi.org/10.1016/j.rse.2022.113145

Wang, C., Chen, J., Wu, J., Tang, Y., Shi, P., Black, T.A., Zhu, K., 2017. A snow-free vegetation index for improved monitoring of vegetation spring green-up date in deciduous ecosystems. Remote Sens. Environ. 196, 1–12. https://doi.org/10.1016/J.RSE.2017.04.031

Xue, J., Su, B., 2017. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. J. Sens. 2017, 1353691. https://doi.org/10.1155/2017/1353691

Zielinski, R., 2021. Outreach: translating practice to signal behaviour. CbM Outreach 2021, 19/11/2021, https://marswiki.jrc.ec.europa.eu/wikicap/images/4/41/Outreach_Part1_Practice2signal_19Nov21.pdf

Zielinski, R., Milenov, P., Sima, A., Devos, W., 2022a. Structured template for documenting agricultural land monitoring systems (No. JRC130662). Ispra, Italy.

Zielinski, R., Sedano, F., Devos, W., Borio, D., 2022b. Collection of Methods for Data Analysis: from Signal Selection to Marker Detection (JRC Technical Report No. JRC130244). Ispra, Italy. https://marswiki.jrc.ec.europa.eu/wikicap/images/5/52/JRC130244_collection_of_methods_for_data_analysis_19sep2022.pdf

List of abbreviations and definitions

CbM	Checks by Monitoring
FOI	Feature of Interest
LC	Land Cover
DIAS	Data and Information Access Services
VI	Vegetation Indices
FOI	Feature Of Interest
IQR	Inter Quartile Range RFV/OTSC
RFV	Rapid Field Visit
OTSC	On The Spot Checks
UAV	Unmanned Aerial Vehicle
MSI	Multi-Spectral Instrument
CAP	Common Agricultural Policy
SAR	Synthetic Aperture Radar
BS	BackScatter
COH	COHerence
COH6	COHerence 6-day
VV/VH	Vertical – Vertical vs Vertical – Horizontal of Sentinel-1 polarizations
A/D	Ascending / Descending orbits of Sentinel-1
RVI	Radar Vegetation Index
NRPB	Normalised Ratio Procedure between Bands
CPR	Cross Polarization Ratio
CPRI	Cross Polarization Ratio Inverse
NDVI	Normalised Difference Vegetation Index
DVI	Difference Vegetation Index
GNDVI	Green Normalised Difference Vegetation Index
GSAVI	Green Soil Adjusted Vegetation Index
GARI	Green Atmospherically Resistant Vegetation Index
NDWI	Normalised Difference Water Index
NDPI	Normalised Difference Phenology Index
GLI	Green Leaf Index
EVI	Enhanced Vegetation Index
BSI	Bare Soil Index
SAVI	Soil Adjusted Vegetation Index
OSAVI	Optimised Soil Adjusted Vegetation Index
MSAVI2	Modified Soil Adjusted Vegetation Index 2
MS	Member States
GSAA	Geo-spatial Aid Application
ECA	European Court of Auditors
ANOVA	ANalysis Of Variance

GTCAP	Geodata and technologies for the common agricultural policy
ESA	European Space Agency
JRC	Joint Research Centre of European Commission
ISO	Organization for Standardization

List of figures

List of tables

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

— by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

— at the following standard number: +32 22999696,

— via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.