Use Case implementation - Crop Diversification

Quick user guide

Introduction

Crop diversification is one of the three "greening measures" of the Common Agricultural Policy, along with the creation of Ecological Focus Areas (EFA) and the maintenance of permanent grasslands at farm-level. The crop diversification requirement applies to farmers with over 10 ha of Arable Land (AL):

- up to 30 ha: farmers have to grow at least 2 crops and the main crop cannot cover more than 75% of the AL;
- over 30 ha: farmers have to grow at least 3 crops, the main crop cannot cover more than 75% of the AL and the 2 main crops not more than 95%.

Checking the compliancy of these rules has been referred to as "Crop Diversification" use case within the Sen4CAP project.

The methods presented here are not "set in stone" and are expected to evolve.

Sen4CAP Earth Observation product supporting the Crop Diversification use case

The specific product that has been developed to support the analysis of farmers' declaration compliancy is the "**L4A – crop type product**".

The aim of the product is to provide information about the crop type observed by remote sensing over each parcel of the Land Parcel Identification System (LPIS) or Geo-Spatial Aid Application (GSAA) dataset. The remote sensing information at the parcel-level is then combined at the holding level to answer the question whether the crop diversification regulation has been correctly applied by the farmer.

Crop Diversification regulation

Following the Technical guidance for the On-The-Spot checks of Crop Diversification[1], each holding belongs to a specific crop diversification category. From the nine categories defined in the document, two have been left out either because it is very specific case (holding land to the north of the 62nd parallel), or because it needs information from last year (which is not yet implemented in our activities). The remained seven categories considered in this use case are presented in Figure 1.

---

[1] JRC (2015), Technical guidance for the On-The-Spot checks of Crop Diversification, DS-CDP-2015-08, JRC96614.
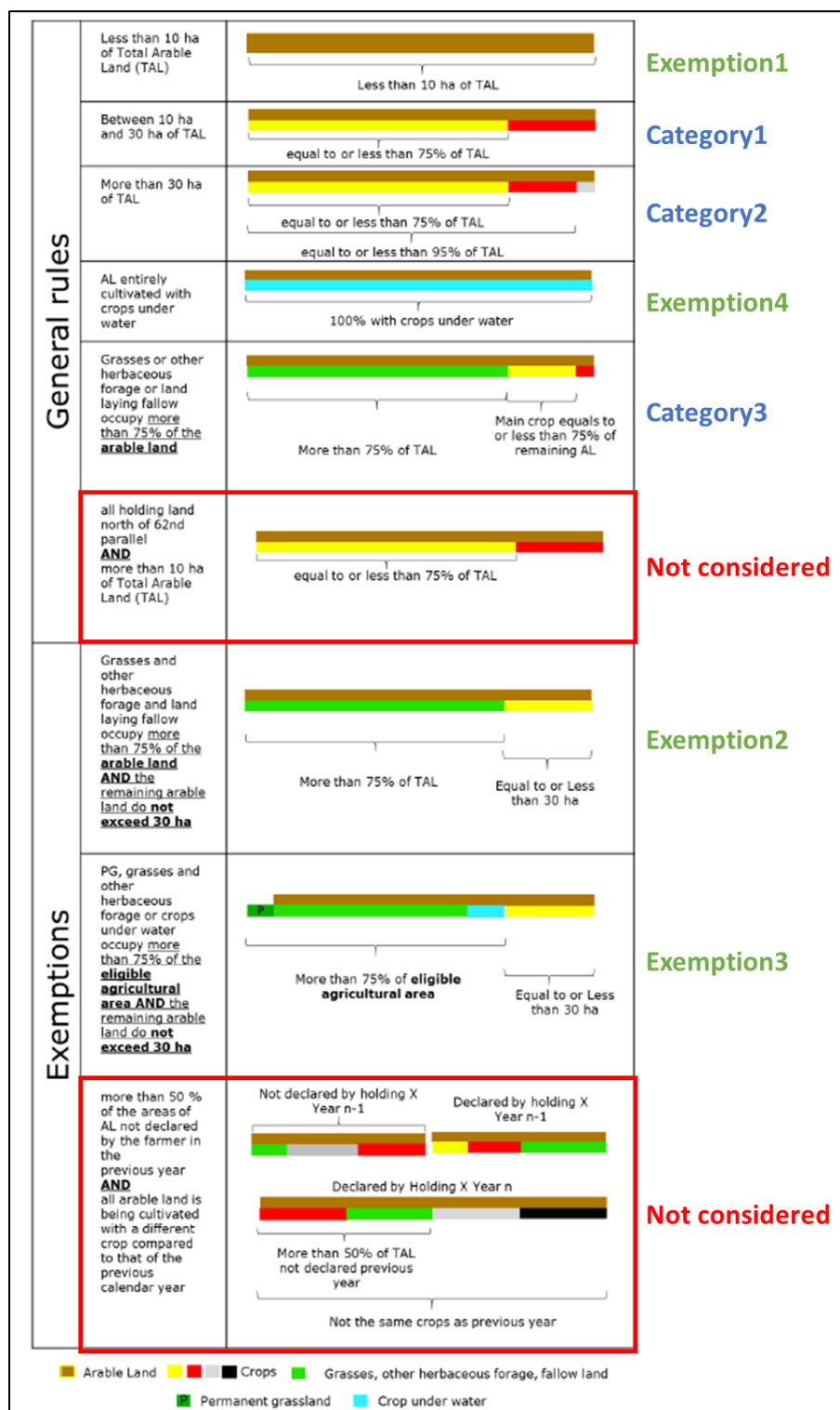
Figure 1. Crop diversification regulations from the Technical guidance for the On-The-Spot checks of Crop Diversification (JRC, 2015) considered in the Sen4CAP project and correspondence with the Sen4CAP crop diversification categories

The description of each category considered in the Sen4CAP crop diversification use case and the corresponding crop diversification rules are detailed in Table 1.

Table 1. Crop diversification categories considered in the Sen4CAP crop diversification use cases

Category	Description	Crop diversification rules
Category1	TAL between 10 and 30 ha	• At least 2 different crop types • Main crop ≤ 75% of TAL
Category2	TAL greater than 30 ha	• At least 3 different crop types • Main crop ≤ 75% of TAL • 2 main crops ≤ 95% of TAL
Category3	TGrass and Fallow greater than 75% of TAL	Main crop ≤ 75% of remaining AL
Exemption1	TAL less than 10 ha	No crop diversification required
Exemption2	TGrass and Fallow greater than 75% of TAL and remaining AL less than 30 ha	No crop diversification required
Exemption3	PGrass, TGrass and Cwater greater than 75% of EAA and remaining AL less than 30 ha	No crop diversification required
Exemption4	Cwater = TAL	No crop diversification required

TAL = Total Arable Land; AL = Arable Land; EAA = Eligible Agriculture Area; TGrass = Temporary Grassland; PGrass = Permanent Grassland; Fallow = Land Lying Fallow; Cwater = Crop Under Water

Satellite data signals

The classification methodology relies on a Random Forest algorithm applied on dense time series of both Sentinel-1 and Sentinel-2 images. The satellite data signal consists of the following set of raw values and composite indicators that are detailed in the Annex A.

These raw values and indicators are calculated first at the pixel-level and then, they are aggregated at the parcel-level. The per-parcel means and standard deviations constitute the input of the classification.

To improve the quality of the classification, some crop types from the original LPIS / GSAA dataset are merged in classes that make sense for both remote sensing and crop diversification regulation point of views. The decision of which crop type can be merged is taken in accordance with the Paying Agency. The aim is to facilitate the construction of the classification model and by doing so to improve the accuracy of the results.

The Random Forest algorithm is then applied to the data. This algorithm can be seen as a series of decision trees that identify the composite indicators derived from the Sentinel-1 and Sentinel-2 signal that best discriminate between the different crop types and that creates consecutive rules based on these composite indicators to define each crop type. The Sen4CAP **"L4A – Crop type product"** includes the two predictions with the highest degree of confidence. A confidence index is associated with each of these two predictions.

The classification algorithm is applied to all parcels which are (i) declared as Eligible Agricultural Area (EAA), (ii) monitorable by remote sensing (e.g. not greenhouses or nurseries) and (iii)

covered by at least 3 Sentinel-2 pixels (10-meter resolution) and 1 Sentinel-1 pixel (20-meter resolution). The way the numbers of Sentinel-1 and Sentinel-2 pixels are counted is explained in the Annex B.

Crop diversification monitoring approach

The Sen4CAP **"L4A – Crop type product"** is used to assess the compliancy of each holding regarding the crop diversification rules. The method presented here is based on (i) discussions made with partner pilot Paying Agencies and (ii) following advices from the Joint Research Center (JRC) given during the MARS conference of November 2018. More particularly, a "worst case scenario" approach has been implemented to handle the small parcels and the crop types that cannot be assessed by remote sensing[2].

The crop diversification monitoring approach relies on two consecutive assessments:

1) First, at the parcel-level, to verify that the crop type declared by the farmer is confirmed by the satellite signal;
2) Second, at the holding-level, to assess the compliancy with regard to the crop diversification rules.

Parcel-level assessment

A parcel is assessed as "conform" if the crop type declared by the farmer corresponds to one of the two outputs of the classification, i.e. with one of the two predictions associated with the two highest degrees of confidence. The parcel is assessed as "not conform" when the two outputs of the classification are different from the farmer declaration.

By default, the assessment at the holding-level will make use of the crop type declared by the farmer. Yet, in the case of a non conform parcel and if the confidence level associated with the first output of the classification is high (above a defined threshold), the user can decide to use this first crop type prediction for the assessment at the holding level. In this case, the information is also given in the assessment at the parcel level.

This parcel-level assessment is documented in the field "Classif_r" (for classification results) of the attribute table of the **"L4A – Crop type product"** (

Table 2).

Table 2. Classification results at the parcel level

Classif_r	Description
Classified_conform	Classified and conform
Classified_not_conform	Classified and not conform
Classified_not_conform_ prediction_used[*]	Classified and not conform, and the first prediction of the model is used for the crop diversification use case

---

[2] https://ec.europa.eu/jrc/sites/jrcsh/files/11-sifting_and_hhr.pdf

Not_classified_geometry	Not classified, problem in the geometry (no valid geometry, duplicate or overlapping with other parcels)
Not_classified_land_cover	Not classified, not monitored land cover class
Not_classified_minS2pix	Not classified, not covered at least by 3 S2 pixels
Not_classified_noS1pix	Not classified, not covered at least by 1 S1 pixel
Not_classified_undefined	Not classified, undefined reason (to investigate)

* this value is not included by default; it will only appear if the user decides to use the first crop prediction in the case of a non-conform parcel with a high degree of confidence in the first prediction (see the text before the table)

Holding-level assessment

This second assessment is implemented following a "worst-case scenario" approach. Only the parcels that are classified and conform (and classified and not conform with a high degree of confidence in the first prediction[3], if the user decided to rely on this prediction) are first used to assess the crop diversification compliancy of the holding. Then, the remaining parcels of the holding, which are assumed to be anything, are used to check if it has an impact on this assessment or not. Only the parcels declared as EAA by the farmer are used in the use case.

- Preparatory steps

Three preparatory steps are carried out prior to the assessment.

First, as for the classification, the crop codes of the Sen4CAP **"L4A – Crop type product"** are grouped in classes that make sense regarding the crop diversification rules. For example, two types of sugar beets are grouped together into a single crop diversification class. This grouping follows the indications provided by the Paying Agency. Indeed, as input for the use case, each pilot Paying Agency has provided the list of crop types / crop type groups eligible for crop diversification and their possible merging.

The second step defines if each crop diversification class belongs to Eligible Agricultural Area (EEA), Arable Land (AL), Permanent Grassland (PGrass), Temporary Grassland (TGrass), Land Lying Fallow (Fallow) and/or Crop Under Water (Cwater).

An example of these two steps is presented in Table 3.

---

[3] In this case, the crop diversification class corresponding to the first prediction of the classification is used in the crop diversification assessment.

Table 3. Example of L4A crop code LUT to define crop diversification classes

CTnumL4A	CTL4A	CTnumDIV	CTDIV	EAA	AL	PGrass	TGrass	Fallow	Cwater
180	ware potato	1	Potatoe	1	1	0	0	0	0
181	seed potato	1	Potatoe	1	1	0	0	0	0
185	industry potato	1	Potatoe	1	1	0	0	0	0
1	sugar beet	2	Sugar beet	1	1	0	0	0	0
2	beta vulgaris 'Mangelwurzel'	2	Sugar beet	1	1	0	0	0	0
166	beta vulgaris	2	Sugar beet	1	1	0	0	0	0
4	chicory	3	Chicory	1	1	0	0	0	0
132	cichorium intybus	3	Chicory	1	1	0	0	0	0
170	cichorium endivia	3	Chicory	1	1	0	0	0	0
260	cichorium intybus sugar loaf group	3	Chicory	1	1	0	0	0	0

CTnumL4A = crop type code used for the classification; CTL4A = crop name corresponding to CTnumL4A; CTnumDIV = crop type code used for the crop diversification use case; CTDIV = crop name corresponding to CTnumDIV; EEA = Eligible Agricultural Area; AL = Arable Land; PGrass = Permanent Grassland; TGrass = Temporary Grassland; Fallow = Land Lying Fallow and Cwater = Crop Under Water.

During the third preparatory step, a series of factors, detailed in Annex C, are summarized by holding, using the different categories defined above (EAA, AL, etc.). 11 factors, named "*_c", relate to the classified and conform parcels and 2 factors, named "*_nc", relate to the parcels not classified or classified but not conform. The "*_nc" factors are used in the "worst case scenario" approach as it can be anything. If the user has decided to work with the remote sensing prediction in the case of not conform parcels with a high degree of confidence in the first prediction, these parcels are also used to compute the "*_c" factors[3].

- Category identification

The first part of the holding-level assessment is to define the crop diversification category (out of the seven categories identified in Table 1) to which each holding belongs. The category is pre-defined using the "*_c" factors, i.e. accounting only for the classified and conform parcels[4]. The impact of the "*_nc" factors on the definition of the category is then checked, following the "worst case scenario" approach. If there is no impact, the category is confirmed. If there is an impact, all the possible categories to which the holding could belong are identified.

The output of this category assessment is stored in the field "CD_cat" (for crop diversification category) of the attribute table of the **"L4A – Crop type product"** (

Table 4).

Table 4. Crop diversification category assessment

CD_cat	Description
Exemption1	TAL less than 10 ha
Exemption2	TGrass and Fallow greater than 75% of TAL and remaining AL less than 30 ha

---

[4] and the classified and not conform parcels with a high degree of confidence in the first prediction, if the user has decided to work with the remote sensing prediction. In this case, the crop diversification class corresponding to the first prediction of the classification is used in the crop diversification assessment.

Exemption3	PGrass, TGrass and Cwater greater than 75% of EAA and remaining AL less than 30 ha
Exemption4	Cwater = TAL
Category1	TAL between 10 and 30 ha
Category2	TAL greater than 30 ha
Category3	TGrass and Fallow greater than 75% of TAL
Category1_or_2	Holding belongs to Category1 or Category2
Category1_or_3	Holding belongs to Category1 or Category3
Category2_or_3	Holding belongs to Category2 or Category3
Category1_2_or_3	Holding belongs to Category1, Category2 or Category3
Exemption_or_Category1	Holding belongs to at least one of the Exemption or Category1
Exemption_or_Category2	Holding belongs to at least one of the Exemption or Category2
Exemption_or_Category3	Holding belongs to at least one of the Exemption or Category3
Exemption_or_Category1_or_2	Holding belongs to at least one of the Exemption or Category1 or Category2
Exemption_or_Category1_or_3	Holding belongs to at least one of the Exemption or Category1 or Category3
Exemption_or_Category2_or_3	Holding belongs to at least one of the Exemption or Category2 or Category3
Exemption_or_Category1_2_or_3	Holding belongs to at least one of the Exemption or Category1 or Category2 or Category3

All details concerning the application of the "worst case scenario" approach can be found in the Sen4CAP ATBD for L4A crop type mapping.

- Crop diversification compliancy assessment

In the case of the exemption categories, no crop diversification is needed. For the other categories, different rules have to be respected to be compliant regarding crop diversification.

Like previously, the compliancy assessment is first used using the "*_c" factors, checking the rules specific to each category, and the holding is pre-assessed as "compliant" or "not compliant". The impact of the "*_nc" factors on the compliancy is then checked, following the "worst case scenario" approach. If there is no impact, the compliancy assessment obtained using the "_*c" factors is confirmed. If there is an impact, it means that there is not enough information to assess the holding compliancy regarding crop diversification and this situation is reported.

The output of the compliancy assessment is stored in the field "CD_diagn" (for crop diversification diagnostic) of the attribute table of the **"L4A – Crop type product"** (Table 5).

Table 5. Crop diversification compliancy assessment

CD_diagn	Description
Compliant	Holding compliant regarding crop diversification
Not_compliant	Holding not compliant regarding crop diversification
Not_required	Holding with no crop diversification required

Missing_info	Not enough information to assess the holding compliancy regarding crop diversification

When the category assessment associates different possible categories with a holding, the different corresponding rules are checked. The holding will be assessed as compliant (or not compliant) if it is compliant (or not compliant) in all possible categories. If it is not the case, the assessment concludes that not enough information is available to take a compliancy decision.

All details concerning the application of the "worst case scenario" approach are given in the Sen4CAP ATBD for L4A crop type mapping.

Result

The resulting Sen4CAP **"L4A – Crop type product"** consists of an ESRI shapefile containing the parcels and characterizing each of them with the following attributes:

- **Original attribute fields (list)**: Attribute fields from the original LPIS / GSAA dataset, including the original crop type code as provided by the PA
- **CT_decl**[5]: Sen4CAP L4A code of the crop type declared by the farmer (for the classified parcels)
- **CT_pred_1**[5]: Sen4CAP L4A code of the crop type predicted by the model with the highest degree of confidence
- **CT_conf_1**: Degree of confidence of CT_pred_1 (ranging from 0 to 1)
- **CT_pred_2**[5]: Sen4CAP L4A code of the crop type predicted by the model with the second highest degree of confidence
- **CT_conf_2**: Degree of confidence of CT_pred_2 (ranging from 0 to 1)
- **Classif_r**: Results of the conformity assessment at the parcel level
- **CD_cat**: Results of the crop diversification category assessment at the holding level (category)
- **CD_diagn**: Results of the crop diversification rules assessment at the holding level (compliancy)
- **LC**: General high-level land cover class defining if the parcel is monitorable or not:
    - 0: other natural areas
    - 1: annual crop
    - 2: permanent crop
    - 3: grassland
    - 4: fallow land
    - 5: greenhouse and nursery
- **S1pix**: Number of S1 pixels covered by the parcel
- **S2pix**: Number of S2 pixels covered by the parcel
- **Area_meter**: Area of the parcel (m²) measured using the polygon shape
- **GeomValid**: Identify parcels for which no polygon exists in the declaration dataset or with a not valid geometry (1 = valid geometry)
- **Duplic**: Identify parcels that have the exact same geometry as another (1 = duplicate parcel)
- **Overlap**: Identify parcels which overlaps with neighboring parcels (1 = overlapping parcel)
- **CTnumL4A**[5]: Sen4CAP L4A code of the crop type declared by the farmer (for all parcels)
- **CTnumDIV**[5]: Sen4CAP crop type code used for the crop diversification case (resulting from the possible merging of individual crop types belonging to a same group of crops in the specific context of the crop diversification rules)

---

[5] The correspondence between (i) the original crop type code provided by the PA, (ii) the Sen4CAP crop type code used for the classification and (iii) the Sen4CAP crop diversification code is given in a separate Look-Up Table named "Sen4CAP_L4A_NLD_2019_CropCode_LUT.csv".

- **CTnumDIV_p**[5]: Sen4CAP crop type code used for the crop diversification case (resulting from the possible merging of individual crop types belonging to a same group of crops in the specific context of the crop diversification rules), corresponding to CT_pred_1

Annex A: Satellite raw values and composite indicators derived from Sentinel-1 and Sentinel-2 data

From Sentinel-2

- Spectral values in green (B3), red (b4), NIR (B8), Red-edges (B5-6-7), SWIR 1 (B11) and SWIR 2 (B12)) interpolated every 10-day period of the year

- Normalized Difference Vegetation Index (NDWI), Normalized Difference Water Index (NDWI) and brightness interpolated every 10-day period of the year

From Sentinel-1

- Backscatter amplitude values in VV and VH polarizations and for ascending and descending orbits along with corresponding ration between VV and VH polarizations, averaged for each week of the year (AMPL_VV_asc, AMPL_VV_des, AMPL_VH_asc, AMPL_VH_des, RATIO_VVVH_asc, RATIO_VVVH_des)

- Coherence values in VV and VH polarizations and for ascending and descending orbits averaged for each week of the year (COHE_VV_asc, COHE_VV_des, COHE_VH_asc, COHE_VH_des)

- Mean of the backscatter amplitude values over iterative two-month periods: Jan-Feb / Mar-Apr / May-Jun / Jul-Aug / Sep-Oct / Nov-Dec

- Coefficient of Variation of the backscatter amplitude values over the same iterative two-month periods

- Mean of the coherence of each month of the period of interest

- Quantile 10 of the coherence values for each month of the period of interest (serving as a proxy for the minimun coherence value)

- Standard deviation of the coherence value along the whole period of interest (the length of this period evolving during the monitoring period)

Annex B: Counting Sentinel-1 and Sentinel-2 pixels covered by each LPIS/GSAA parcel

The number of Sentinel-1 and Sentinel-2 pixels covered by a parcel is calculated in five steps:

- Application of two inner buffers to the LPIS / GSAA datasets, of 5 and 10 meters respectively;
- Reprojection of these two new shapefiles in the UTM zone of the country (one or several);
- Rasterization of the two reprojected and buffered shapefiles;
- Counting of the number of Sentinel-2 pixels having their centroid inside the 5-meter buffer and of the Sentinel-1 pixels having their centroid inside the 10-meter buffer.

These steps are illustrated for Sentinel-2 in Figure 2.
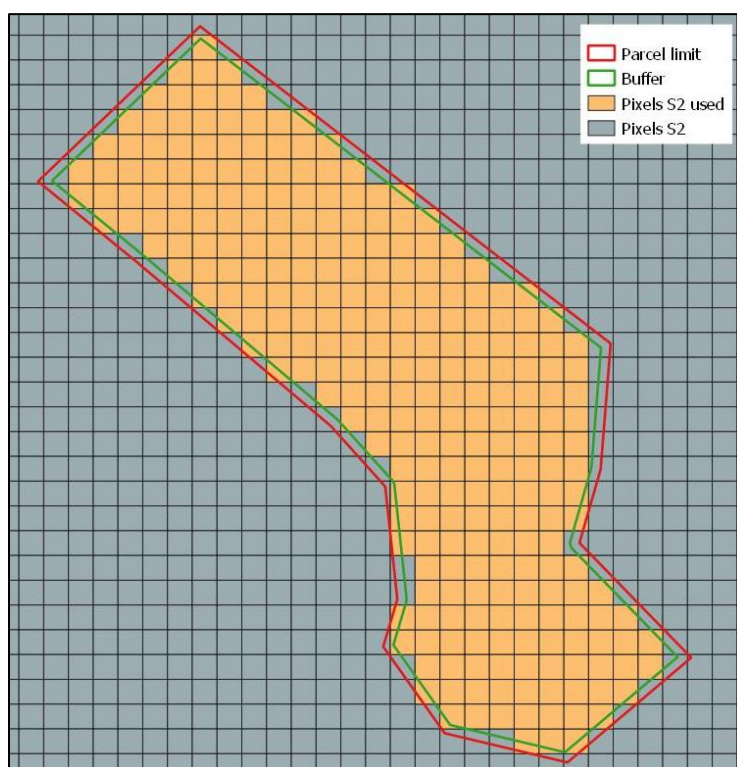
Figure 2. Identification of the Sentinel-2 pixels covered by the parcel and counted; this count is used as ancillary information for the crop type classification

Annex C: List of summarized factors by holding

Factor	Description
area_eaa_c	Classified and conform* Eligible Agriculture Area (EAA)
area_tal_c	Classified and conform* Total Arable Land (TAL) area
area_tempGrass_c	Classified and conform* Temporary Grassland (TGrass) area
area_permGrass_c	Classified and conform* Permanent Grassland (PGrass) area
area_llf_c	Classified and conform* Land Lying Fallow (Fallow) area
area_cwater_c	Classified and conform* Crop Under Water (Cwater) area
area_remAl_ex2_c	Classified and conform* remaining area of AL (in the case of exemption 2)
area_remAl_ex3_c	Classified and conform* remaining area of AL (in the case of exemption 3)
nb_types_c	Number of classified and conform* crop types in AL
area_mainCrop_c	Area of the main classified and conform* crop type in AL
area_2mainCrop_c	Area of the second main classified and conform* crop type in AL
nb_parcels_nc	Number of not classified or classified and not conform** parcels (all remaining parcels in EAA)
area_nc	Not classified parcels or classified or not conform** (all remaining parcels in EAA)

* and the prediction of the model in the case of the classified and not conform parcels with a high level of confidence in the first prediction, if it is activated by the user

** except the classified and not conform parcels with a high level of confidence in the first prediction, if it is activated by the user