European Commission

JRC TECHNICAL REPORTS

Checks by Monitoring quality inspection: EU requirements and methodology

Document set v.1.1
December 2020

Wim DEVOS, Natasa LUKETIC, Pavel MILENOV, Daniele BORIO

Joint Research Centre

This publication is a Technical report by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The scientific output expressed does not imply a policy position of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication.

EU Science Hub
https://ec.europa.eu/jrc

JRC 119733

How to cite this report: DEVOS Wim *et al.*, Checks by Monitoring quality inspection: EU requirements and methodology, Ispra, 2020, JRC 119733

Release notes

January 2020: pre-draft: chapter 5 (inspection methodology) distributed to CbM adopters

July 2020: first draft, distributed to CbM adopters, part of document set v1.0:

1. Focus on Sentinel inspection replacing, but still based on, independent ground truth
2. Introduction of the 3 step approach: detection, eligibility and financial impact
3. Separate quality element for pre-conditions, area management and eligibility detection
4. Extension to the AAR reporting on unmonitored and inconclusive parcels.

December 2020: second draft, GTCAP community distribution, document set v1.1:

1. A priori need for dedicated ground truth validation removed
2. Elaborated handling of the three steps in terms of sampling, relevance, expectations and error propagation
3. Additional considerations on spatial scale
4. Quality thresholds reformulated based on the 2019 experiences and insights
5. The term pre-condition removed to mirror changes in Regulation 2014R809

Contents

1. Executive Summary

1.1.1. This is a second, draft version of the document; it is neither final nor complete.

1.1.2. The monitoring approaches, checks by monitoring (CbM) and the future area monitoring system (AMS) approaches allow the CAP to be simplified and modernised. The simplification originates from the preventive aspects inherent to monitoring; the modernisation potential comes from a control mechanism that is as effective, but fairer and more efficient in managing agricultural activities than the so called 'On-The-Spot Checks' (OTSC) methods.

1.1.3. The expectations of monitoring approaches provide some continuity from the OTSC tradition, but they are formulated in the updated context, where Land Parcel Identification System (LPIS) and Geo-Spatial Aid Application (GSAA) take over the area and administrative checks dimension allowing the Sentinel data to focus on the eligibility checks and the prevention mechanism. The resulting prime elements can be formulated as.

- Continued assurance on the applicability of relevant CbM boundary conditions.
- Appropriate management of the spatial and area component
- A reliable assessment of eligibility within an application;
- An estimation of potential financial or any other system level impact (the CbM contribution) by processing the CbM QA findings

1.1.4. A methodology is proposed whereby majority of the automatic eligibility checks are verified by comparing process flow decisions against an independent (mostly visual) inspection on the freely available Sentinels time stacks. The interpretation keys for this inspection are based on readily available ground truth, without an a priori need for a dedicated validation mechanism of that ground truth.

1.1.5. Using Sentinel as primary inspection source for flow decisions allows for full random sampling without bias, but with small sample sizes and simple statistics.

1.1.6. The inspection results are then used to assess the alpha (false positives) and beta (false negatives) error rates of detections, compliance errors and the resulting financial or other impact. Some general elements regarding the analysis and reporting of this quality assessment results are introduced.

1.1.7. The technical documentation is structured in modules and encourages using the available cloud solutions to keep investments in software and hardware minimal. The method and set-up described here is co-developed and tested with the help from the l on-boarding Member States.

1.1.8. This particular draft relates to a version v1.1 of the CbM QA document set.

2. Introduction

2.1. Report objectives

2.1.1. This is a draft version of the document; it is neither final nor complete. Its release serves to sketch an overall framework and a skeleton wherein the draft proposals can develop further to maturity.

2.1.2. The document targets monitoring operators and quality control officers. As a result, it discusses the issues with some technical details. The two main topics are an elaboration of a generic checks by monitoringinspection methodology, and the launch of discussions on how this generic process can be appropriately customized for an individual implementation.

2.1.3. It is however clear that the core technology and methodology developed for the CbM QA are expected to be re-used, for its larger part, for the future area monitoring system that will become mandatory under the CAP2020+ proposal.

2.2. Rationale for quality assurance (QA)

2.2.1. CbM currently replace, on an optional basis, the on the spot checks (OTSC) for direct payments totalling over €40B for 2019. These checks target the correctness of the payments to the farmers and cooperate with two additional administration systems; the land parcel identification system and the geospatial aid application. The CbM emerged to prevent and resolve incorrect aid applications and hence reduce the number of non-compliant control results observed under the on the spot check system.

2.2.2. The CbM option was introduced as a substitute for the on the spot check for Member States to take advantage of the new technologies, in particular Sentinel 1 and 2 towards reducing costs (higher cost-efficiency), reducing burden (by automation and reducing field visits), increasing fairness (by avoiding sampling) and increasing awareness and compliance (by prevention through early warnings and a mechanism to modify the aid application by the farmer).

2.2.3. The Commission services acknowledge these potential advantages and desire to verify that the system actually delivers these. At the same time, the Commission is determined not to jeopardize the achievements of the OTSC efforts (i.e. maintain a low financial error rate). CbM is expected to lead to an increase of the overall compliance as it allows all farmers to remedy non-compliance of parcels and to withdraw or correct incompliant parcels from the application.

2.2.4. Given the wide range of anticipated advantages of CbM, implementation will inevitably have some second-order impact on the design of the payment procedures and the current procedures in the integrated administration and control system The vast majority of the relevant changes fall under the autonomous MS responsibility and are therefore not subject to this quality discussion.

2.3. Role of QA in the quality policy

2.3.1. The direct output (numbers and results) of this exercise is intended to enable the MS to report to the Commission about the state of its control and management system. The common basis of the methodology ensures an objective and comparable reporting from all Member States. For this common purpose, templates and procedures are provided.

2.3.2. However, the outcome of this exercise should also, not in the least, lead to an honest and more detailed self-assessment of the functioning of the system itself, in a process that may be separate from the reporting to the Commission under the proposed QA framework. This reflection could involve investigating weaknesses both of the quality assessment method and of the monitoring system itself, but also beyond. Without an open and honest self-assessment, it will be very difficult to optimize operational processes or identify appropriate remedial actions.

2.3.3. In the particular context of CbM self-assessment, Member States can use the underlying QA methodology to identify weaknesses in their CbM processes early in the season and potentially address them long before any reporting to the Commission.

3. Prime quality elements

A prime quality element expresses the quality expectation linked to a certain functional requirement. For CbM, the most urgent requirement relates to the ability to perform:

- (1) a check on the payment applications, through a correct assessment of the eligibility of the application and
- (2) an appropriate handling of the area values that ultimately lead to the financial impact of the errors.

3.1. Persistence of the boundary conditions for Checks-by-monitoring

3.1.1. The Technical Guidance (TG) on the decision to go for monitoring (DS-CDP-2018-17) identifies several pre-conditions that set the boundary that delimits the scope wherein CbM can be expected to function. Several systems have explicit validity rules that control the veracity of these conditions and so assure that the system operates as expected. However, in the absence of explicit validity tests in a CbM design, dedicated tests should confirm the appropriateness of the boundary. As CbM manage area based schemes during the full duration of a campaign year, it is obvious that both spatial and temporal aspects matter.

3.1.2. On the spatial aspect, the TG (chapter 1) expresses the need for the area declared to be truthful, matching agricultural land cover and having a graphical outline that is truthful. The combination can be assured if there is a close spatial cardinality and congruency between the GSAA's agricultural parcel and the physical object, management unit or field (called feature of interest - FOI). JRC has developed a dedicated test, called "P1" that checks such spatial cardinality. It involves a visual comparison between the GSAA and the FOI as portrayed itself over the season. Automatic Sentinel processing allows for the creation of an appropriate FOI portrayal. Problematic P1 test results should lead to the introduction of an explicit spatial validity test in the CbM design or/and an update or even upgrade of the LPIS or the GSAA system.

3.1.3. On the temporal aspect, the TG (chapter 2) assumes a desirable accuracy (discrimination ability for crop identification based on phenology) of 95%. Although crop identification is certainly not the main output of the CbM processes, such proxy target for conclusion remains valid and the purpose of estimating the discriminatory performance is easily recognised. Indeed, those parcels that are not targeted by the CbM flow will end up in the "pool of non-conclusive parcels" and will be subject to non-Sentinel follow-up including sampling. The size and nature of this pool is a topic of major concern for many CbM adopters and candidates and justify some basic assessment. A dedicated test, called "P2", checks whether inclusion into that pool was justifiable. The test is performed by either automated machine learning analysis or manual (mostly visual) inspection of the pool. The results primarily offer feedback into the tuning of the CbM algorithms. For practical reasons, this P2 test will not be required for the 2019 and 2020 campaigns.

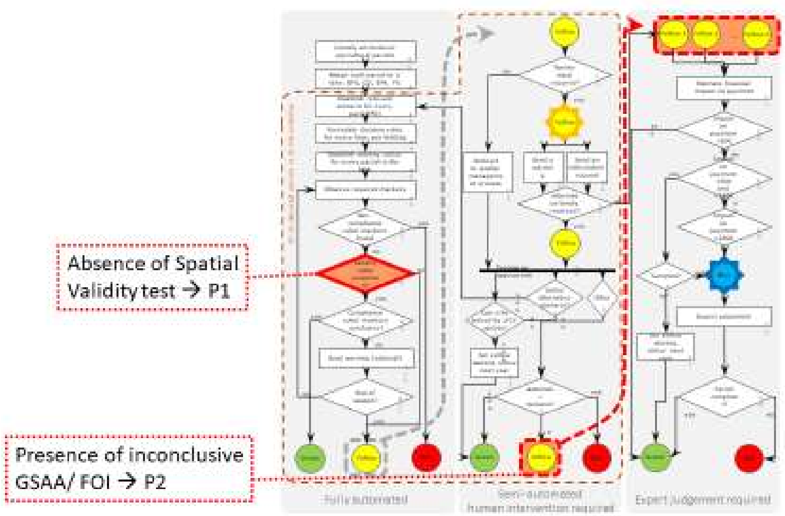

Figure 1. Scope and triggers of the boundary condition tests P1 (cardinality of spatial objects) and P2 (targeting performance)-

3.1.4. Figure 1 illustrates how the P1 and P2 address the CbM functioning boundary and its implied assumptions. The absence of a spatial validity check implies that all area values would be correct as they enter the process. A large pool of inconclusive parcels implies that none of the previous processes would have provided evidence to arrive at a conclusion.

3.1.5. The remainder of the quality elements can thus focus on the explicit decisions made by the CbM system, which is the stricter scope of the QA exercise. In the spatio-temporal domain, these automatic or semi-automatic decisions can relate to the spatial aspect (characteristics of the area) or the temporal aspect (detecting discriminating behaviour of the land) or a combination of both. Such explicit decision can either set a conclusion (compliance “green” or non-compliance “red” rules) or simply change the processing pathway (validity “yellow” rules).

3.2. Management of the area component

3.2.1. The relevance of area management in the context of area aid schemes is obvious. Good area management structures are essential for the decision to introduce CbM (see P1 above). However, given that landscapes and field boundaries are dynamic, it is much more straightforward and preferable to manage the spatial aspects explicitly during the CbM.

3.2.2. A first component of the spatial aspect is the spatial cardinality and congruency between GSAA and FOI representation. If this is managed in the system, e.g. as a validity rule, the decisions from the system can be analysed via a specific test, called G1. The G1 test is similar in nature to the P1 test, but differs in scope (it evaluates decisions not parcels) and processing outcome. Any FOI that has been subject to such explicit spatial validity rule on cardinality should not be included in the lot for the P1 test above.

3.2.3. A second spatial component deals with heterogeneity of the field (FOI). FOI heterogeneity is well described in the 2nd discussion document DS-CDP-2018-18 (chapter 4) and stable heterogeneity is an intrinsic property on some agricultural land (pro-rata grasslands, some permanent crops) whereas it is antithetical to some other land uses (such as several arable crops). Detected heterogeneity could be relevant to conclude on compliance, on validity and on non-compliance rules that are part of the process flow. Furthermore, any detection of distinct crops or management zones in larger FOIs could be the starting point for validating area shares for greening measures. A dedicated test, called G2 assesses the robustness of such decisions.

3.2.4. Heterogeneity can also be ephemeral, i.e. reflecting an expected or desirable form of land use. Examples are the partial mowing of grasslands or incremental harvests of vegetables. Such non-permanent heterogeneity, although expressed in spatial terms, is nevertheless considered a temporal aspect.

3.2.5. Indeed, in this CbM QA context, heterogeneity in the field mostly relates to the variability of the agricultural area from morphological or management causes. This has to be considered as separate from contamination of that area by distinct ineligible features, which do not represent agricultural area at all.

3.2.6. The CbM QA spatial testing procedures operate at a scale that is compatible with Sentinel 1 and 2 imagery.

3.3. The reliable assessment of eligibility within an application

3.3.1. An eligible application is an application where, within the requirements and tolerances set by the regulations and guidelines, parcels are compliant and all schemes' requirements are met. The monitoring approach observes the physical manifestation or behaviour of the parcels in that application. In contrast to the OTSC, which observes the state at one or only a few dates, CbM relies on observing the behaviour throughout the season. In addition, CbM relies on an existential feedback loop with the farmer that dissociates this detection process from its direct financial implications by allowing early warnings and remedial actions to ensure compliance. The temporal aspects of the CbM processing are therefore dominating the quality inspection and assessment methodology.

3.3.2. The variety of landscapes and schemes and their corresponding practices, imply that, with the currently available remote sensing technologies, no single "black box" algorithm will be able to provide a reliable and robust verdict for all scenarios. It is much more likely that several processes will co-exist in parallel e.g.:

- Continuous lookout for urbanisation and abandonment phenomena;
- Time relevant marker detection for anticipated activities on grassland or cropland, and for assessing the applicable deadlines of temporary EFA or arable crops ;
- End of season crop classification as ex-post evidence of activity on arable land or occurrence of a particular crop;
- Planned (by validity checks) and automatic (by expiry of a deadline) reception of ad-hoc and follow-up evidence by the farmer.

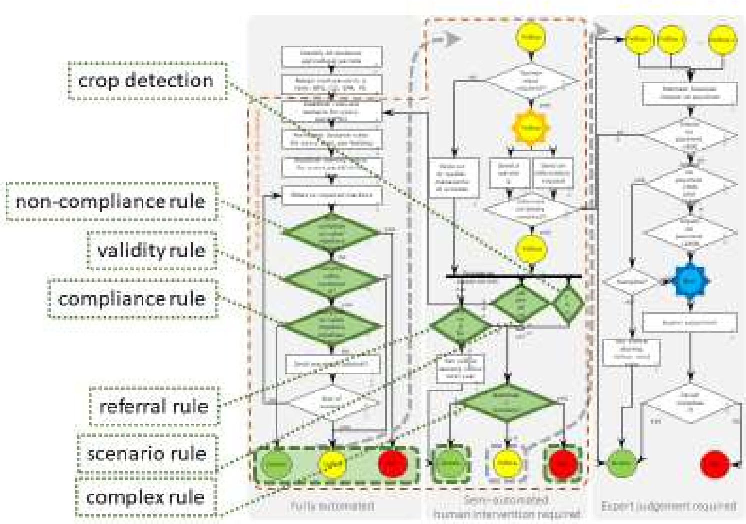

Figure 2. Identification of the kind and scope of rules whose robustness of detection and decision is tested in the CbM QA..

3.3.3. In principle, each decision flow process should be based on the detection of a single or a series of observable phenomena or patterns because that is the essence of monitoring with the Sentinels. As figure 2 shows, a wide variety of rules can be implemented; a particular test can rely on spatial or temporal properties of the FOI but what all rules have in common is that they rely on input detected on the Sentinel image stacks. However, the nature of the observed phenomena as detected from the Sentinel data, is not infinite; there is limited number of information extraction types. JRC currently identifies four types of temporal behaviour as well as one type of combined spatio/temporal behaviour. The typology, essentially a typology of marker core, is based on the timescale and context of the observable behaviour. Each type is assumed homogenous and can thus be evaluated by a single dedicated test, called T1 to T4 and C1:

- T1: occurrence of an abrupt land cover change;
- T2: evidence of a gradual land cover transition over the years;
- T3: observation of a tell-tale event;
- T4: identification of a crop through classification or other;
- C1: emergence and evolution of distinct spatial patterns in the FOI.

3.3.4. Each of these types of phenomenon or behaviour in a given season can hold several distinct properties, processes or events, and the detection of such phenomenon is embedded in a FOI scenario to steer it processing towards a conclusion. E.g.:

- abrupt land cover change: *urbanisation*, *forest clearing;*
- gradual land cover transition*: scrub regrowth, rewilding, afforestation;*
- tell-tale event: *roughness change, vegetation drop or vegetation variability that is associated with annual farmer activity;*
- crop class: *cereal* (possibly also *wheat, barley*), *hay land, pasture*;
- spatial pattern: *partial mowing, harvesting strips.*

3.3.5. What matters in this setup is the decision in the CbM process flow, rather than the individual detection of an event or activation of a marker. As a result, if several events need to be detected, or not be detected, before a decision is made (similar to the "flagging") only the final combined outcome that triggers the decision needs to be tested. The rationale for this is that concatenation does not fundamentally alter the process flow but merely delays the decision.

3.3.6. A distinction between the land cover and land use concepts is critical. Abrupt land use changes can trigger gradual land cover changes and vice versa; afforestation changes the land use overnight but the planted trees need years to close a forest canopy. A forest clearing is abrupt in land cover terms but does not necessarily involve a land use change.

3.4. A tiered approach

3.4.1. The quality elements cover the boundary conditions, the management of the area component as well as deriving eligibility conditions from a parcel's temporal behaviour. They operate on a number of inspection procedures ("tests") that identify the basic decision and lots as illustrated in table 1. The applicability of a test depends on the options within the individual design of a system.

Test (lot)	Rationale	Topic	Targeted decision	Test Output
P1	DS/CDP/2018/17 ch 1.2	GSAA/FOI (spatial) cardinality	Implied validity	Binary
P2	DS/CDP/2018/17 ch 2.1	Discriminatory power	Implied validity	Binary
G1	DS/CDP/2018/18 ch 2.4	Spatial congruency	Validity rule	Paired binary
G2	DS/CDP/2018/18 ch 2.4	Spatial Pattern variability and/or composition	Validity rule/ Dataflow rule	Paired binary
T1	DS/CDP/2018/18 ch 3	Land cover conversion	Dataflow rule	Paired binary
T2	DS/CDP/2018/18 ch 3	Land cover transition	Dataflow rule	Paired binary
T3	DS/CDP/2018/18 ch 3	Activity detection	Dataflow rule	Paired binary
T4	DS/CDP/2018/18 ch 3	Crop identification	Dataflow rule	Paired binary
C1	DS/CDP/2018/18 ch 3	Evolution of spatial pattern	Dataflow rule	Paired binary

Table 1. Overview of the CbM QA tests, with rationale, scope, target and output. Binary output is "passed/failed, paired binary output combines "positive/negative" and "true/false".

3.4.2. However, the relation between detection (by Sentinels), compliance (as described in the scenarios and processing lanes) and financial impact (from the resulting payment) is not straightforward.

- A given detection (ploughing) could mean compliance for arable land (a manifestation scenario) and non-compliance for environmentally sensitive permanent grassland (an absence scenario);
- Detecting an incorrectly declared crop can have financial repercussion if Voluntary Coupled Support (VCS) or Crop Diversification (CD) is concerned but is often irrelevant if only Basic Payment Scheme (BPS) matters

3.4.3. As a result, the QA methodology needs to follow a tiered approach that first assesses the detection robustness of the CbM algorithms and procedures. This basic tier, called "step", is the most technical one of the processes. Through the selection of tests applicable for the system, it demonstrates the effectiveness of the field (FOI) detections made by the monitoring system to keep track of the behaviour of the agricultural land in its landscape. Although there are 9 theoretical inspection tests (P1-P2, G1-G2, T1-T4, C1), first experience shows that less than half that number are applicable in an operational system. The tests will show that the vast majority of detections will be correct, that only a minority proves erroneous.

3.4.4. A second tier or step must translate the detection result to a compliance decision. This translation considers the agricultural parcel, the lane for the scheme and its scenario. As demonstrated by the example on ploughing, detection or non-detection could work either way and has to be disambiguated for every decision into manifestation or absence scenario. This disambiguation does not require any inspection effort because all needed information is already recorded into the system. A main result of this tier is an identification of errors that are abatable (by potential action of the farmer if he would have received an early warning) and errors that are end-stage (expected to be final).

3.4.5. The third tier or step continues with that minority of errors that is end-stage and propagates them into impact. For the financial impact, this involves adding declared area and applicable payment rates to the agricultural parcel, and determine the amounts involved. The result of this tier is maximum estimation of potential financial impact attributable to the errors in the CbM process.

3.4.6. At the time of writing, focuses and details of such third steps are still under development.

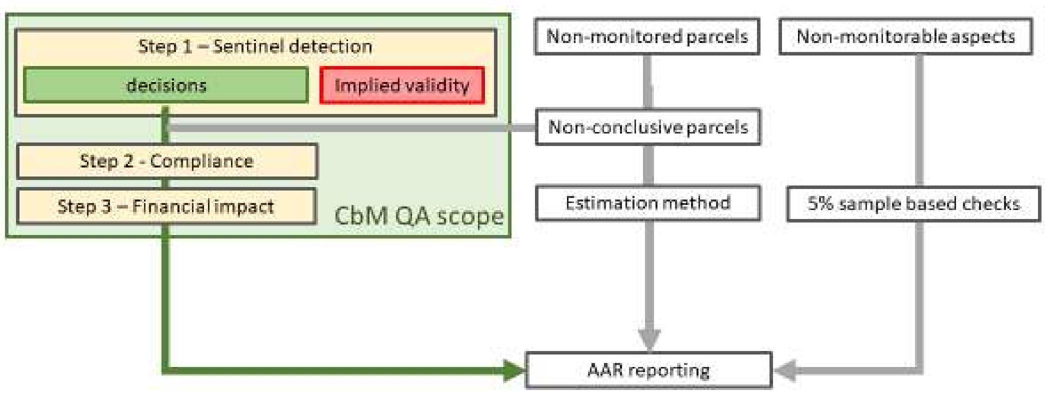

Figure 3. The scope of the QA methodology proposed in this document (in colour) and how it contributes into the financial reporting. Inconclusive parcels and non-monitorable aspects require separate processing paths not discussed here. AAR stands for annual activity report.

3.4.7. TG (DS-CDP-2018-17) and the discussion document (DS-CDP-2018-18), introduced false positive and false negative errors in terms of farms receiving undue payment or missing out on payments. This addressed a single universal expectation at the end of the CbM and subsequent payment process. In absence of real CbM data, the values were set at the industry default settings of 5% and 10%.

3.4.8. The 2019 CbM QA developments revealed that the proposed universal expression is not directly applicable at the level of each tier or step. This is because the error quantifications at the various steps are not proportionally related and also because the very concept of positive and negative depends on the formulation of the algorithm or rule. As a result, different thresholds were set at each tier and the respective proposals are represented in table 2. Limiting qualities (LQs) for false positives and false negatives (step 1) are now both set to 10% to reflect the symmetry and interchangeability between the two error types. This provides better protection against excessively strict tests in the case of small sample sizes (zero acceptant tests in the ISO 2859-2).

Step	Tier	Error type	LQ	Error type	LQ
1	Detection	Non-existing event (false positive)	10	Missed event (false negative)	10
2	AP Rules	Erroneous compliance conclusion	5	Erroneous non-compliance conclusion	5
3	Impact	End-stage amount	2%*	Abatable amount	-

Table 2. Conceptualization and acceptance threshold expressed in LQ (limiting quality index of ISO 2859-2) of false positives and false negatives in each CbM QA tier. LQ thresholds further developed from discussion document DS-CDP-2018-18. * : financial error in terms of total payment to the farmer.

3.4.9. Table 2 only elaborates the contribution to impact from within the CbM QA scope, it does not apply to those aspects addressed in the following chapter 3.5. The 2% value at step 3 reflects a materiality threshold, often used in financial reporting, and is introduced as a starting point for the ongoing development.

3.5. Non-monitorable, inconclusive and non-monitored aspects

3.5.1. Figure 3 scopes the CbM QA steps that also culminate into a quantification of the potential financial impact. Although not strictly part of the CbM QA scope, the other lines that also contribute to the overall annual activity report (AAR) need some clarification. Note that the AAR, where applicable, also involves the results of any remaining OTSC processes, which are not discussed here.

3.5.2. Non-monitorable aspects are defined in Regulation 2014R809 art. 40a.1(c) and are by default subject to checks of a 5% sample.

3.5.3. Inconclusive parcels are those parcels that are subject to some aspects of monitoring but where no final conclusion could be made on their eligibility compliance. The backstop method allows for some to be considered compliant (and hence payment) without further processing during the campaign.

3.5.4. Non-monitored parcels are those subject either to incomplete monitoring or to no monitoring at all. In fact, the system could have a priori decided to deal with them on a later moment via a referral rule. A referral rule is any rule that redirects the further processing of the FOI to a procedure or time beyond the annual analysis of the Sentinel data. Figure 4 shows where those categories can be identified.

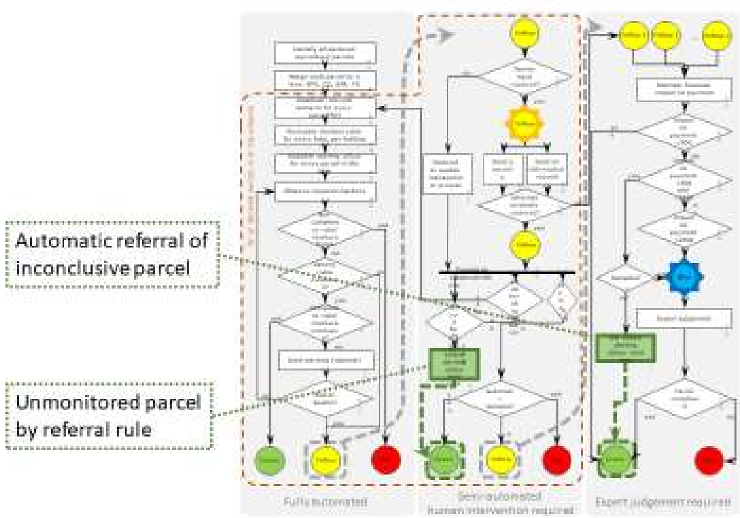

Figure 4. Decisions leading to inconclusive parcels and non-monitored parcels.

3.5.5. For the inconclusive parcels and unmonitored parcels, it is not possible to propose detailed guidelines on assessing the short and long term impact of potential errors from the referral decisions. Some basic reporting will be needed, but this is not within the scope of this document.

4. Integration in an IACS quality policy

The CbM control approach differs from the original OTSC that the monitoring elements are designed to interoperate and collaborate with existing prevention systems such as the LPIS or GSAA. The other systems relieve the CbM to focus on eligibility checks and activity detection, because former take care of the area component (area correctness) and link to the payment procedure respectively. Although it is not in the scope of this document to specify the LPIS and GSAA quality expectation, it is necessary to point out the obvious consequences.

4.1. LPIS QA impact

4.1.1. LPIS reference parcels are and remain the primary container for eligible area for the direct payments. Hence, the area measurement related elements of the LPIS QA are not affected by a choice to go for CbM.

4.1.2. However, quality elements that relate to the cardinality (QE3) and purity of the reference parcel (QE2b, QE2c) could possibly be revised to target a better fitness for purpose. Indeed, the monitoring approaches rely on an appropriate representation of the agricultural land cover (i.e. each individual "production block" in the correct sense, not an amalgamate of fields) that may be less effected by certain contaminations.

4.1.3. It therefore could make sense to develop a new LPIS QA conformity class profile for CbM, that will co-exist with the conformity class profile applicable for the continued OTSC implementation. This new profile should represent the requirements of an LPIS that is suitable for wall-to-wall monitoring, but it should not be mixed up or generate hybrids with the current OTSC-driven profile.

4.2. GSA QA development

4.2.1. The GSAA provides the farmer's access to the subsidies. It represents a formal, albeit temporary, link between agricultural area, beneficiary and scheme conditions and the resulting, unique combination leading to a payment.

4.2.2. As a result, it is conceivable to define GSAA quality expectations as:

- correctly identify and/or capture agricultural land, based on a LPIS that correctly quantifies area values.

- perform effective administrative crosschecks on the legitimacy of the beneficiary (e.g. with farm register, land tenure register, formalities….).

4.2.3. Most of the GSAA functionality relates to the functioning of internal mechanisms and could be assessed by internal inspection methods although there is no such assessment today. The future GSA is expected to inherit many of those functions and those aspects that would require external quality control (agricultural area values, final compliance with scheme conditions) included in a future GSA QA will interact with elements of the future LPIS QA and AMS QA respectively.

5. Agricultural parcel inspection methodology

5.1. Overall setup

5.1.1. The inspection methodology must provide an independent verification of the decision on detection, compliance evaluation and financial impact. It obviously starts with the detection aspect and works its way up the three steps of table 2.

5.1.2. Given that detectability of monitoring elements is the "alpha and omega" assumption for CbM, most underlying real world phenomena should be detectable (visually or otherwise) on appropriately selected and processed Sentinel data stacks. For the verification of the detection, the coverage and availability of the Sentinels allow for a completely random sampling, offering some key advantages

- It eliminates the risk of bias from spatial clustering or other sampling imperfections.
- It has no logistic consequences.
- Inspecting the full population has no impact on the inspection costs.

5.1.3. In the boundary condition tests, the aspect that is already verified is geospatial aid application quality through the P1 test; in the many CbM dataflow tests, it is the system decisions based on observable phenomena. In the testing nomenclature, these inspection units are called items. The outcome of a tested item is either "pass/fail" for a boundary condition test and true/false positive/negative for a dataflow decision test.

5.1.4. CbM accommodates local landscapes and farming conditions, so there is no single common inspection procedure. As a result, the robustness of a visual (or other) inspection protocol (what? and how? to look for when?) should be based on and validated by separate ground truth. Fortunately, relevant ground truth has already been collected in clusters and pilots spread over the territory during the preparation for the decision to adopt CbM. This often complements a vast amount of ground conditions and local photointerpretation knowhow that was collected during the past on the spot check campaigns, As a result, this existing ground truth and knowledge can be reused and a ground truth collection campaign does not have to be launched annually. Simple maintenance and tuning suffices.

5.1.5. The proposed CbM QA validity and dataflow tests were designed around the nature of the detectable phenomenon, so to facilitate the programming of an inspection environment and to increase ease of processing. The resulting inspection lot contains all decisions based on a detection characteristic and is assumed homogeneous, until analysis of inspection would indicate the contrary. If heterogeneity of results would be suspected, a lot should be disambiguated along the identified factors in a subsequent QA. Analysis of the CbM system will identify which tests are applicable for a particular implementation.

5.1.6. Grouping the QA items in such lots based on the nature of the detected information, means that the underlying populations become large and that the corresponding samples are very small. In fact, JRC has established that with the currently proposed 5% and 10% threshold, sample sizes per lot will never exceed 365 items. However, one should remain alert for the presence of subpopulations that could warrant the identification of a dedicated lot e.g. based on context, location in the dataflow, or any other risk or heterogeneity factor.

5.1.7. Acceptance testing is done at every step by comparing the "failed" or "false" numbers from the tests against the probability functions used by the ISO acceptance sampling schemes.

5.1.8. Any well targeted (specificity) and highly performant (sensitivity) CbM algorithm or decision will produce a very small number of negatives and obviously even less false negatives to test. Their abundance in the sample can be so low that statistical testing becomes unreliable. However, such formal acceptance test can be skipped because the low numbers imply that no remedial action is needed.

5.1.9. All "false" outputs of the detection tests are carried over into step 2, to establish the impact on the compliance decision. This involves a parcel based linking to a lane and corresponding manifestation or absence scenario. That linkage doesn't involve additional inspection activities but an appropriate and automatic connection to the outcome of the CbM path. The resulting reshuffling of errors allows estimating the false compliance and false non-compliance decisions.

5.1.10. The "false" outputs of the compliance assignment on each agricultural parcel are further divided into "abatable", where the farmer is expected to come forward and demand correction, and "end-stage" where the farmer has no incentive to contest. The "end-stage" compliance errors can be linked to area values and payment rates that are likely to represent paid amounts. It is at the overall system level that potential financial impact comes into sight.

There is no need to go into holding detail and accommodate thresholds and actual payments to an applicant. Step 3 elaborates those calculations.

5.1.11. The technical motivation and detail of these inspection steps are elaborated by JRC in a set of technical guidance documents listed in table 3.

Par.	**Topic**	**Document**
5.1.3	Item identification	Annex III
5.1.4.	Ground truth	Annex IV
5.1.5.	Testing element	Annex I
5.1.5.	Lots	Chapter 4 of TG
5.1.5.	Inspection	Chapter 5 of TG
5.1.6	Statistics	Annex V
5.3.2	Reporting	Annex VI
5.1.8	Acceptance test	TG tables
5.1.9	Compliance test	TG tables
5.1.10	Impact assessment	Chapter 6 of TG (tbd)
5.5.3	Cloud tools	JRC DIAS Github

Table 3. Summary of documents with technical details covering the CbM inspection methodology. Annex II (not in the table) holds the data flow overview.

5.2. Analysis

5.2.1. The above methodology produces acceptance verdicts at the level of each step

- An α and β for the detection (Sentinel detection step)
- A true and false error rate for the compliance decisions (Compliance step)
- An estimate of the maximum financial exposure from end-stage compliance decisions

5.2.2. The set of numbers provides an overall view of the system, but the applicable thresholds might hide relevant details. Indeed, the principles of fairness to all farmers, level playing field and good governance, all imply that any given number should not derive from merging a majority of perfect CbM results with a small minority of completely irrelevant CbM results. Phrased differently, no farmer or parcel should be exempted from the control because the system de facto ignores them and hides them "in the masses". When a particular CbM sub-component will be publicly known to be deficient, one can expect that some stakeholders will exploit that particular failure. MS should remedy such known conditions by adapting their CbM setup.

5.2.3. Furthermore, MS should analyse any test result where the individual errors exceed the requirement and propose remedial action in the CbM set up for the issue concerned. Such remedial action, can take the form of improving the monitoring algorithms, adapting the process flow, organizing external evidence collection and henceforth.

5.2.4. At overall system level, these initial α and β over the various steps are probably too generous to maintain in the long term, the good assurance level currently achieved by the OTSC. This will, when or if appropriate, be reassessed considering the outcome of the ongoing quality assessment trials with the CbM adopters.

5.2.5. The Commission believes that managing these various errors appropriately, provides the instrument to redesign the IACS and the control system in particular so it becomes an efficient instrument in the communication between administrations and farmers. The proposed, expandable, method allows the MS to tailor the inspection for its own self-assessment needs, before and beyond the mandatory reporting required by the Commission Services.

5.3. Reporting

5.3.1. As the Member States themselves perform the actual assessment, all relevant documentation of the inspection procedure should become available for information and later verification (audit). This includes the documentation of preparatory processes such as the ground truth collection and validation.

5.3.2. This reporting involves the Member States delivering the following data for which the Commission and Member States will jointly develop templates and schemas for this reporting:

- A short, schematic description of the CbM data flows, elaborating on the content of the formal CbM notification
- The numbers of items in the CbM data flow.
- A scoreboard with the error findings for each lot of step 1 and for the overall system on steps 2 and 3;
- The individual results of the inspected CbM items;

- If applicable or agreed, a remedial action plan or textual description of remedial actions following the analysis above;

5.4. Implementation and planning

5.4.1. The CbM quality assessment should be implemented once a year, timely so that remedial actions can be duly implemented. There is however no particular regulatory specification or deadline.

5.4.2. The Commission may consider, in justified cases, adapting the procedures for imagery provision under 1306/2013 art 21, to accommodate a reasonable amount of HHR imagery for tuning the ground truth collection database.

5.5. Costs

5.5.1. Given that the vast majority of ground truth data has already been collected and given that the Sentinel imagery and their access are free of charge, costs are limited to a once-off investment of the inspection environment (both technology and know-how) and recurrent inspection costs.

5.5.2. The once-off investment in expertise involves the analysis of the CbM set up (identifying the procedures, scenarios and corresponding items). It also requires the training of the inspectors. The analysis component should happen irrespective of the CbM exercise; it is mentioned merely pro-forma.

5.5.3. The technical investments in software will depend on the complexity or diversity of the system, which could in turn depend on the landscape and the administrative context. Still, in the cloud environment, which is the natural habitat of Sentinel imagery, many tools, algorithms and software components are already available, so the development cost should be reduced to customization only. JRC has developed a library of components that it uses for Sentinel based inspection and freely distributes these. Using the cloud means that there is no investment needed to acquire and maintain the physical infrastructure in-house.

5.5.4. All tests are either binary or paired binary, i.e. the inspector always reports a zero "O" or a one "1". In the proposed methodology, there is no need for mapping or elaborate data collection. Trials have shown that a properly equipped and trained inspector can easily process 50-200 items a day, depending on the type of decision checked, meaning that any lot can be inspected within ten working days. If there are no errors, the work stops there. If there are errors, the additional cost will depend on the internal organisation of the IACS system and the performance and transparency of the CbM design. The reporting costs should be estimated in person weeks rather than person months.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

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, or

- by electronic mail via: https://europa.eu/european-union/contact_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 at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub