


Climate Change Initiative Extension (CCI+) Phase 1
New Essential Climate Variables (NEW ECVS)
High Resolution Land Cover ECV (HR_LandCover_cci)

End-to-End ECV Uncertainty Budget
(E3UB)

Prepared by:

Università degli Studi di Trento
Fondazione Bruno Kessler
Università degli Studi di Pavia
Università degli Studi di Genova
Université Catholique de Louvain
Politecnico di Milano
Université de Versailles Saint Quentin
CREAF
e-GEOS s.p.a.
Planetek Italia
GeoVille





	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	1	

Changelog

Issue	Changes	Date
1.0	First version.	02/07/2019
2.0	Updated version.	03/01/2020
3.0	Global update of the document.	20/11/2020
3.1	Updated version according to CCI_HRLC_Ph1_AR2_RID-ESA.xlsx	07/12/2020
4.0	Section 1.2. has been updated with LCC related information. Section 2.4 has been added with the description of the uncertainty in composite generation (input quality index). Section 3.2 has been updated according to the analysis of the geographical and temporal distribution of the SAR historic data. Section 6.4 has been updated with the addition of the input quality index to the uncertainty output. Section 6.5 has been added with description of the uncertainty in the multitemporal cascade model used for the historical production. Section 7 has been updated with the description of the uncertainty in abrupt change detection.	31/10/2022



Detailed Change Record

Issue	RID	Description of discrepancy	Sections	Change
3.1	FR-01	The bullets of paragraphs are wrong 1.1, 1.2, 1.3 instead of 3.1, 3.2, 3.3	Section 3	Fixed all wrong numbering of sections.

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	2	

Contents

1	Introduction.....	3
1.1	Executive summary.....	3
1.2	Purpose and scope	3
1.3	Applicable documents	3
1.4	Reference documents.....	3
1.5	Acronyms and abbreviations	3
2	Optical pre-processing	5
2.1	Radiometric correction.....	5
2.2	Cloud and cloud-shadow detection / restoration	5
2.3	Spectral filtering and harmonization	5
3	SAR pre-processing chain	6
3.1	SAR processing chain	6
3.2	SAR data.....	6
3.3	Geometric processing.....	6
3.4	Speckle noise	7
4	Multi-sensor geolocation	7
4.1	An example of RMSE computation.....	8
4.2	RMSE computation without ground truth.....	9
4.3	From registration error to uncertainty	9
4.3.1	Sub-Pixel Registration Error	9
4.3.2	Non-sub-pixel Registration Error.....	9
5	Classification.....	10
6	Decision fusion	12
6.1	Uncertainty in consensus theory	12
6.2	Uncertainty in Markov Random Fields	13
6.3	Uncertainty with Deep Learning.....	14
6.4	Uncertainty Output	15
7	Multitemporal change detection and trend analysis.....	16
8	References	16

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	3	

1 Introduction

1.1 Executive summary

This document provides an assessment of the end-to-end uncertainty budget the HRLC ECV products are associated with. HRLC products are based on a wide range of input data whose uncertainties propagate at different levels of dependency according to the data characteristics and the processing steps involved in the production. By taking into account the scarce availability of ground-measured reference information and the practical impossibility to collect physical measurements on wide areas as those selected for this project, the proposed uncertainty models will be, by necessity, theoretical.

Many steps of the processing chain (e.g., pre-processing, geolocation, classification, etc.) involve algorithms that come with uncertainty models associated to them. For instance, the classification task is able to output probabilistic posteriors that can be managed at the fusion level to infer uncertainty score pixel-wise. Both uncertainties of input data sets and processing model-related ones must be considered, including error propagation dynamics. The nature of the input data sets (discrete classes vs. continuous variables) and the associated error characteristics (random error/ bias, error distribution), including potential correlations between errors of different input variables should be evaluated. Finally, uncertainties related to the spatial scales of data sets, scaling issues related to the validation activity must be accounted for as well.

1.2 Purpose and scope

This document provides both a detailed overview of the main sources of uncertainty for each step of the full processing chain and representations of pixel-wise uncertainty for the final HRLC and LCC products.

This document deals with all known potential sources of error, uncertainty and known correlations in the data that are seen as potential contributors for the definition of an uncertainty product. The output for classification is given as a three layer data-structure that includes:

- Classification maps associated with both first- and second-best performers, in terms of posterior probability, as returned by the classification-fusion model.
- Actual values of posterior probabilities corresponding to the two above mentioned maps.
- Input quality index corresponding to the input optical-data quality, which is related to the number and temporal distribution of the images acquisitions used in the composite generation step.th

The uncertainty of LCC is affected by the propagation of the uncertainties coming from the previous steps. The probability of change contributes as a source of information to model the certainty associated to LCC products as well as the availability of images on a yearly bases. If images are not enough for one year the change information is provided over a longer time span and becomes more uncertain.

1.3 Applicable documents

Ref. Title, Issue/Rev, Date, ID

[AD1] CCI HR Technical Proposal, v1.1, 16/03/2018



[AD2] CCI_HRLC_Ph1-D2.2_ATBD, latest version

1.4 Reference documents



Ref. Title, Issue/Rev, Date, ID

1.5 Acronyms and abbreviations

AMI	Active Microwave Instrument
AOT	Aerosol optical thickness
ASAR	Advanced Synthetic Aperture Radar

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	4	

ATBD	Algorithm Theoretical Basis Document
BOA	Bottom of Atmosphere
CCI	Climate Change Initiative
DEM	Digital Elevation Model
ECV	Essential Climate Variables
ERS	European Remote Sensing
ETM	Enhanced Thematic Mapper
GCOS	Global Climate Observing System
GRD	Ground Range Detected
GT	Ground truth
GUM	Uncertainty in Measurement
HR	High Resolution
LaSRC	Landsat-8 surface reflectance code
LC	Land Cover
LCC	Land Cover Change
LEDAPS	Landsat Ecosystem Disturbance Adaptive Processing System
LOGP	Logarithmic opinion pool
MAP	Maximum a posteriori
MGRS	Military Grid Reference System
MODIS	Moderate Resolution Imaging Spectroradiometer
MPM	Marginal a posteriori modes
MR	Medium Resolution
MRF	Markov Random Field
MSI	Multispectral Instrument
MSS	Multispectral Scanner
NIR	Near infrared
OLI	Operational Land Imager
RMSE	Root mean square error
RST	Rotation scale translation
SAR	Synthetic Aperture Radar
SoW	Statement of Work
SWIR	Short-wave infrared
TIRS	Thermal Infrared Sensor
TM	Thematic Mapper
TOA	Top of Atmosphere
UTM	Universal Transverse of Mercator
VHR	Very High Resolution
WGS84	World Geodetic System 1984

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	5	

2 Optical pre-processing

Detailed analysis of pre-processing errors/accuracy related to harmonized Sentinel-2 / Landsat products is given in [1]. Although the final processing chain of the CCI HRLC project will not be identical to [1], this work provides the most complete reference for the prior modelling of the variables that contributes to pixel-level measurable uncertainty of the products coming from an integrated pre-processing stage.

2.1 Radiometric correction

LaSRC assumes a Lambertian, plane-parallel atmosphere, and uses the 6S radiative transfer model to invert directional surface spectral reflectance from observed top-of-atmosphere reflectance. Several atmospheric parameters are required for the inversion including surface pressure, column water vapor, ozone, and aerosol properties. LaSRC algorithm assumes two SR ratios, red to blue and red to ultra-blue, and uses the difference between these assumed ratios and observed TOA reflectance ratios to invert for AOT and Angstrom exponent. The two fixed SR ratios for the globe are derived from MODIS and MISR data, and expressed as a function of mid-infrared vegetation index.

Currently, uncertainty estimates for LaSRC are based on comparison with corrections based on in situ atmospheric parameters from the Aerosol Robotic Network [1]. These comparisons indicate improved performance compared to the LEDAPS algorithm or an alternative version of LEDAPS that used MODIS aerosol products as input. For Landsat 8 OLI, overall uncertainty varied from 0.11% absolute reflectance (SWIR1 band) to 0.85% absolute reflectance (blue band). For Sentinel-2/MSI, overall uncertainty varied from 0.3% absolute reflectance (SWIR1band) to 1.4% absolute reflectance (blue band).

2.2 Cloud and cloud-shadow detection / restoration



Cloud and cloud-shadow detection accuracy is intrinsically difficult because of the impossibility of directly measuring physical parameters related to clouds. Posterior evaluation of cloud detection accuracy can be figured by referencing to appropriate literature, but this does not provide a direct method to quantify uncertainty in a pixelwise manner. A possible strategy to mitigate this problem is to associate probabilities in the classification step related to cloud identification. This is similar to the idea implemented in processors like sen2cor.

2.3 Spectral filtering and harmonization

Given the differing solar and view angles associated with Landsat 8 and Sentinel-2, normalizing the BRDF effects is desirable. Retrieving the BRDF information directly from medium resolution optical remote sensing data is not feasible with the current temporal and angular distribution of the data. Instead, the BRDF information needs to be ingested a priori. It is currently on-going revision of the most appropriate technique to achieve this so that effects of different illumination conditions may be included in the model.

2.4 Composite Generation

When working at large scale, it is necessary to harmonize the times-series of images acquired over different tiles which are characterized by different lengths and are acquired at different times. This is mainly due to the irregular cloud coverage (which hampers the use of some images of the time-series) and the different orbit acquisitions (different temporal sampling). To solve this problem, in the pre-processing step we generate monthly, seasonal and annual composites. This condition allows us to mitigate cloud occlusions problem and minimize the processing resources. To this end, we consider a statistic-based approach that computes the median value for each pixel. This approach is able to generate consistent results at large scale in an automatic way by sharply reducing the spatial noise. However, in many cases the available data for the composite generation can be limited to few acquisitions (e.g., from none to as few as 2-3 acquisitions). In such cases, the uncertainty related to the median values increases significantly. To mitigate this problem, an input quality index is associated to the mapped pixels for each year. The input quality index is strictly related to the number of

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	6	

composites in a given year that are generated with at least three valid acquisitions (i.e., non saturated, cloud and shadow free pixels). In the case of annual composites, the input quality is proportional to the number of valid acquisitions of the only composite.

3 SAR pre-processing chain

The spaceborne synthetic aperture is a powerful key Earth observation technique for large-area monitoring, and in recent decades, its all-weather capability and fine ground resolution facilitated the development of a great variety of applications, such as the land mapping, for example. Due to the nature of the SAR range mapping and reflectance functions, the measurements of multi-channel SAR system can be biased by error originating from many deleterious factors in which significantly degrade the quality of SAR image. In this section we quantify sources of uncertainty on SAR results, i.e. all those aspects that leading doubts about the validity of the result of a measurement (or processing).

3.1 SAR processing chain

Undoubtedly a first uncertainty cause of uncertainty in SAR processing is to be researched in the application of techniques for reconstructing the observed scene, such as Level-1 data produced as Single Look Complex (SLC) and Ground Range Detected (GRD) from raw data (also so-called Level-0), i.e. the backscattered wave. This may be modeled for instance by the Uncertainty in Measurement (GUM) framework, introduced in [2] and used to evaluate the uncertainties of the amplitude values pixel by pixel, on the bases of a statistical analysis.



3.2 SAR data

The quality of land cover map is certainly changed by the number of the SAR images used to produce both static and historical maps. About the static production, in the PVASR document it is stated that at least five images per season are necessary to achieve reliable results. A smaller number of images degrades the quality of the results by reducing the discriminative power of SAR data sets among natural land cover classes. Concerning the historical production, the situation is even more complex because of the poor availability of ERS and ENVISAT products both in time and with respect to their geographical distribution. In addition, the processing of time series with data acquired by the same sensor (ERS or ENVISAT) but on different orbits, i.e., SAR sequences composed of both ascending and descending data, produces very noisy and distorted maps, very far by the best possible case.

The number and the type of the images used to produce the maps every five years has a strong impact on the quality of the final products. To reduce this impact the production of historical maps from SAR was restricted considering only to the tiles with images belonging to the same acquisition orbit (ascending or descending), and with at least a number of images available in the year of interest, where this number was selected according to the data availability for each historical area, aiming at the best possible quality of the results.

3.3 Geometric processing

A second source of uncertainty refers to next steps applied to SAR data sets, such as geometric processing. Specifically, radar image processing requires the geometrical overlaying of the remotely data sensed from different sensors and/or geometries, in order to mitigate very severe distortions over elevated and sloping terrain [3]. Before the change detection and surface classification, these distortions must be mitigated by terrain correction. SAR imaging requires precise determination of the relative position and velocity of the radar platform with respect to its target at all times. However, this information is not available at the time of imaging, since the platform must be moving for the azimuth resolution technique to work. The basic solution is to assume that the platform has a uniform velocity over a smooth geoid [4]. To remove the geometric distortions due to terrain relief, the *radiometric calibration* could be applied for compensating all spatial and time dependent variation, as well as the cross-track and image to image intensity inconsistencies due to signal attenuation by distance [5].

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	7	

The algorithm is based on a Digital Elevation Model (DEM) and creates a simulated SAR image based on an imaging radar model.

However, this process introduces a level of uncertainty because it simplifies the image creation considerably, introducing certain amount of distortions into the results. In fact, when there is significant relief in the area being imaged, for example, the DEM model that is being adopted will be based on a proper smooth geoid assumption and will lead to pixel placement, rendering it suitable (or not) for a quantitative analysis of terrain features.

Finally, to correct image orientations, SAR images are geocoded. This phase includes SAR image resampling to a spatial representation with known geometric properties. Standard map projections, like the Universal Transverse Mercator (UTM) mapping, are used. Processing involves the image rotation and scaling to properly transform it into the mapping coordinates chosen [6]. The geocoding represents another uncertainty source, since images are taken at varying pass angles, and each resulting image contains an approximately square rotated SAR image inside it, with unused image pixels set to black.

3.4 Speckle noise

In addition to geometrical features, another limitation of SAR data sets is the speckle noise. The speckle noise is an intrinsic feature of SAR data, and it is given by the consistent summation of signals from ground scatters randomly and loosely distributed within the scene. The existence of speckle noise in SAR images is an inherent and specific random characteristic. This noise has an impact on the interpretation of these images and introduces further limitations in applications exploiting SAR time series [7]. The speckle noise reduction has to strictly be carried out in order to preserve polarimetric properties, without introducing any image quality degradation and corrupting statistical characteristics. Consequently, not using a suitable speckle filtering involves a dramatic impact in terrain classification performance [8].

Instead, performing a multi-temporal analysis based on multiple images acquired over time, backscattered values in can be aggregated in both coherent and incoherent ways to reduce the effect of noise. For classification purposes, it has been proved [9] that the use of multitemporal sequences improves the accuracy of the final results, either thanks to the fusion at the decision level of the results for each image, or by combining multiple SAR images into a single input to the classification procedure.



4 Multi-sensor geolocation

Multi-sensor geolocation [10] aligns data collected from different sensors (a reference and still image and an input image to be transformed) in a common reference system in order to process them coherently further in the processing chain. Within the CCI+ HRLC processing chain, the two image data sources correspond to optical and SAR data. The result of the geolocation process may be more or less precise, yielding uncertainty associated with the output images. Such uncertainty generally also affects the subsequent blocks along the processing chain.

There exist different strategies for assessing the accuracy of the geolocation process: one of the possibilities is the computation of the root mean square error (RMSE) in pixel units [11]. In an experimental setup, where the correct transformation is known, the RMSE may be computed analytically. Otherwise, it is possible to estimate it by means of specific control points or landmarks.

The control points are identified in both images and the RMSE is computed based on the residual spatial distances. Ideally, in case the images are perfectly matched, the distances of the control points in the reference and registered images is equal to zero. In all the other cases, the control points may not be perfectly matched, and the distances are generally non-zero, although they can be smaller than one-pixel size, on average.

The following section describes the computation of the RMSE in case the transformation is known and is modelled as a “rotation-scale-translation” (RST) transformation. For lower-complexity transformations, like rigid or shift transformations, the same computation holds but it is necessary to fix unitary scale (rigid) and the rotation angle to zero degrees (shift).

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	8	

4.1 An example of RMSE computation

Let $Ref(x, y)$ and $In(x, y)$, $(x, y) \in \Omega \subset \mathbb{R}^2$, where Ω is a region of interest, be two images called reference and input, respectively. If they are both of size $A \times B$ pixels, then $\Omega = [0, A] \times [0, B]$. In the RST case, $T_p(x, y)$ is the geometric transformation described by the parameter vector $p = (t_x, t_y, \theta, k)$ and has the form:

$$T_p(x, y) = \begin{pmatrix} k \cos(\theta) & k \sin(\theta) & t_x \\ -k \sin(\theta) & k \cos(\theta) & t_y \end{pmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

where $\{t_x, t_y\}$ determine translations in the x and y directions, θ is the rotation angle, and k is the scaling factor. Thus, we can write $T_p(x, y) = Q_p \cdot [x, y, 1]^T$, where Q_p is the RST transformation matrix given above, and the superscript “ T ” indicates the transpose operator. There is a one-to-one correspondence between Q_p and p . To register $Ref(x, y)$ and $In(x, y)$ it is necessary to find the value of p such that $In(T_p(x, y))$, the input transformed by T_p , best matches the reference (see ATBD).

When accurate ground truth is available, such as when test images are created synthetically (a typical scenario when a geolocation method is developed and is being validated), a standard way of assessing registration accuracy is by using the RMSE $E(p_e)$ [11]. Suppose the ground truth (GT) transformation is given by $p_{GT} = (t_{x1}, t_{y1}, \theta_1, k_1)$ and the computed transformation is $p = (t_{x2}, t_{y2}, \theta_2, k_2)$, with the two RST matrices $Q_{p_{GT}}$ and Q_p respectively. It is possible to define the error transformation $p_e = (t_{xe}, t_{ye}, \theta_e, k_e)$, along with the corresponding RST matrix Q_{p_e} , and measure the discrepancy between p_{GT} and p .

According to the matrix formulation of the RST transformation, being Q_{p_e} the error transformation matrix, the following should hold [11]:

$$Q_{p_e} = Q_p \cdot Q_{p_{GT}}^{-1}$$

that yields:

$$\begin{cases} k_e = \frac{k_2}{k_1}, & \theta_e = \theta_2 - \theta_1 \\ t_{xe} = t_{x2} - k_e(t_{x1} \cos(\theta_e) + t_{y1} \sin(\theta_e)) \\ t_{ye} = t_{y2} - k_e(t_{y1} \cos(\theta_e) - t_{x1} \sin(\theta_e)) \end{cases}$$

Now, let $(x, y) \in \Omega$ and let $[x', y']^T = Q_{p_e} \cdot [x, y, 1]^T$. This can be equivalently written as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = k_e \begin{pmatrix} \cos(\theta_e) & \sin(\theta_e) \\ -\sin(\theta_e) & \cos(\theta_e) \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} t_{xe} \\ t_{ye} \end{bmatrix}$$



Then, the RMS error is defined as:

$$E(p_e) = \sqrt{\frac{1}{AB} \int_0^B \int_0^A (x' - x)^2 + (y' - y)^2 dx dy},$$

Substituting the formula for x' and y' and solving for $E^2(p_e)$ yields:

$$\begin{aligned} E^2(p_e) = & \frac{\alpha}{3} (k_e^2 - 2k_e \cos(\theta_e) + 1) + (t_{xe}^2 + t_{ye}^2) - (At_{xe}^2 + Bt_{ye}^2)(1 - k_e \cos(\theta_e)) \\ & - k_e (At_{ye} - Bt_{xe}) \sin(\theta_e) \end{aligned}$$

where $\alpha = A^2 + B^2$. This formula is used in this research to measure registration accuracy when the ground truth transformation is available. This scenario holds when a geolocation method is being developed and semi-

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	9	

simulated data sets are used for its tuning and validation. It obviously does not apply to the case in which two real image data sets are available because the reference GT transformation is not known.

4.2 RMSE computation without ground truth

The previous section describes how to compute analytically the RMSE in case ground truth data is available. In case no ground truth information is available, it is still possible to compute an estimate of the RMSE through control points (landmarks) [10]. It is possible to identify well-known control points in both the reference and the registered input image and estimate the accuracy of the registration process from them. Hence, a sample estimate of the registration RMSE can be computed by averaging the single RMSE computed with respect to each pair of control points, i.e., as a sample estimate of the RMSE functional. Of course, the GT collection step can be considerably time-consuming.

As an alternative, a common and effective approach to validate registration results when GT is unavailable is to use a visual (qualitative) assessment rather than a numerical (quantitative) calculation of RMSE figures. In the former case, the reference and registered images are combined in appropriate image composites that emphasize the spatial matching of the linear and curvilinear features in the image (e.g., colour composites with RGB channels drawn from both images or checkerboard spatial composites). This approach has been extensively used in the project to evaluate registration performance in the application to real image data pairs not endowed with any GT information.

4.3 From registration error to uncertainty

Once the registration error is estimated, one can indirectly derive information about the uncertainty generated by the geolocation process within the overall land-cover mapping process. In particular, it is convenient to distinguish two scenarios based on the achieved error.

4.3.1 Sub-Pixel Registration Error

There are cases where the registration error is less than a single pixel. Thus, the grid of the two images after registration is almost perfectly matched. Indeed, this is the goal of most image registration efforts.



In this case, every pixel is matched with the corresponding pixel in the other image. Obviously, there may still be a residual error. However, achieving sub-pixel accuracy implies that the Earth region associated with a pixel in the reference image is almost the same as the Earth region associated with the registered image, i.e., spatial mismatch between the optical and SAR sources is smaller than the pixel size after registration.

This is the best possible achievement in multi-sensor geolocation, and in this case, no uncertainty is deemed to be forwarded to the following processing blocks of the chain. Every pixel is correctly located, and no uncertainty needs to be propagated along the overall processing chain.

4.3.2 Non-sub-pixel Registration Error

If the registration RMSE is larger than one pixel, the pixels in the reference image are not correctly matched with the pixels in the registered image, on average, and the mismatch implies an actual misalignment of the data associated with the two pixel grids. Non-sub-pixel errors may cause artefacts in boundary regions, where pixels of different classes are superimposed due to the residual shift between the reference and registered images. Conversely, in a flat homogeneous image region, registration error may not cause problems, as the mismatch may not influence the resulting classification map.

In this case, the uncertainty is propagated to the data fusion block. Operatively, the probabilistic fusion that takes place in the decision fusion module may generally be affected by the residual non-sub-pixel misregistration. The goal of the CCI+ HRLC processing chain is to estimate land cover. In principle, the impact of non-sub-pixel registration error on land cover uncertainty may be assessed, on each pixel, by making use of the probabilistic per-pixel information available to the decision fusion module within a local window, whose size may be

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	10	

determined as a function of the registration RMSE. However, residual misregistration is expected to intrinsically translate *per se* into increased uncertainty in the pixelwise posterior probabilities obtained by the HRLC processing chain on each pixel. Moreover, the aforementioned local-moving-window process may add significantly to the overall computational burden. Therefore, consistently with the goal of assessing the uncertainty in the output HRLC product, the impact of residual misregistration on the overall uncertainty will be characterized through the pixelwise posterior distribution. The possible use of moving-window processes will be considered methodologically or experimentally in a tradeoff with computational burden.

5 Classification

Uncertainty is unavoidable in all classification domains: a certain amount of uncertainty is always involved in deciding the class a sample is assigned to. The unanimously recognized framework to represent uncertainty is probability. Specifically, the Bayesian concept of maximum posterior probability encloses the amount of uncertainty (measurable from data) that generates in the probabilistic decision of a classifier. Since no specific classifier has been selected yet (both for optical and SAR processing classification), the treatment is given in this section in a general way. In Figure 1, the general workflow associated with the classification part of the processing chain is recalled.

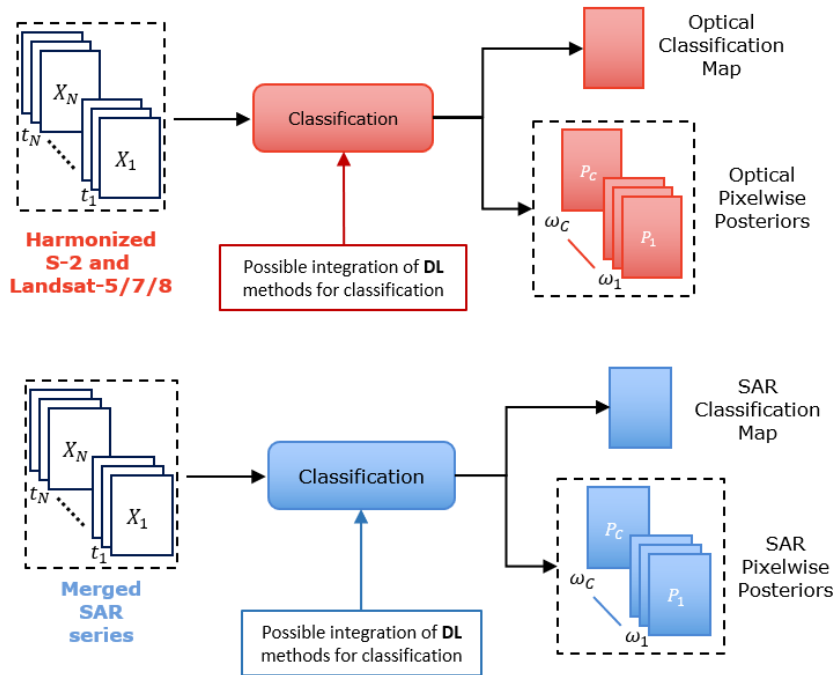


Figure 1. Workflow of the classification process for optical and SAR time series of images.

To model input of the Decision Fusion block dealing with integration of different sources of uncertainty, here we present a general framework to posterior probability definition that is algorithm independent [10]. We can model the posterior probabilities using m linear classifiers. Each linear classifier implements a hyperplane that separates its corresponding class from the other classes. The equation of each hyperplane is

$$f_j(\mathbf{x}) = \mathbf{w}_j \cdot \mathbf{x} + b_j$$

where \mathbf{w}_j is a weight vector and b_j is a bias term. Each source-specific posterior probability is typically computed using the softmax function:

$$P(\omega_j | \mathbf{x}_i, \theta_h) = \frac{\exp(f_h(\mathbf{x}_i))}{\sum_{k=1}^m \exp(f_k(\mathbf{x}_i))}$$

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	11	

Hence, the parameter set $\theta_h = (\mathbf{w}_h, b_h)$. For any probabilistic-based classifier, plugging the posterior probabilities into the cross-entropy function and solving the equation (gradient methods) is at the core of the well-known learning rules and or backpropagation algorithms.

As widely acknowledged, SAR images are an important source of information but the speckle noise gives SAR images a granular appearance that makes interpretation and analysis hard tasks. Furthermore, one of the major issues is the assessment of topographic information content in this kind of images that could be extrapolated by exploiting classification techniques. Classification accuracy values, which include user's accuracy (UA), producer's accuracy (PA), and overall accuracy (OA), are strongly influenced by the adopted input model, which could cause considerable errors in the model output.

An image contains an enormous amount of information, and the challenge is how to represent it in a more compact way, which is why features are originated. In other words, for a more compact and possibly more significant representation of the information embedded in image, it is usually decomposed into several features. Specifically, the extraction of spatial features from remotely sensed data and the use of this information as input to further processing steps has received considerable attention over the two last decades. Unfortunately, due to the complexity of the images and the existence of image noise and disturbances, the information derived from an image is always ambiguous. This source of uncertainty makes the following recognition/classification process more complex. The accuracy of spatial feature extraction can hardly be formulated due to intra-class variation and inter-class similarity [12]. This uncertainty of spatial features extraction is of course on top of the other source of inaccuracy, common to any image element, such as positional uncertainty, attribute uncertainty, topological uncertainty, computational inaccuracy, imprecision/inexactitude, inconsistency, incompleteness, repetition, vagueness, omission, misinterpretation, misclassification, abnormalities and knowledge uncertainty.


Land cover maps are generated by the classification of remote sensing data, and are frequently used as input to spatially explicit environmental models, and its quality is generally assessed at the global or class-specific. It is therefore clear that another uncertainty factor is closely correlated with the used classification algorithms. However, a clear assessment of classification accuracy is not easily provided, and several works have been proposed to evaluate the technical classification uncertainty [13], [14]. These studies demonstrate that uncertainty assessment provides valuable information on the performance of land cover classification models, both in space and time. Statistical inferences arise by supervised learning ensembles. Ensemble methods based on bootstrapping, such as random forests (e.g., used for the classification of SAR data), have improved the predictive accuracy of individual trees, but fail to provide a framework in which distributional results can be easily determined. Random forests are among the most popular machine learning techniques for prediction problems that are known for making predictions with low errors.

We grow N decision trees, and then our predictor \mathcal{Y} for an input data x is given by:

$$\mathcal{Y} = \sum_i^N \omega_i T_i(x)$$

where ω_i is some weight and $T_i(x)$ is the value predicted by the i -th tree of classifier.

When using random forests is very important to predict a quantitative response and to get a measure of uncertainty. No prediction is free from errors, as every model is a simplified representation of reality. The prediction error can be tracked down to uncertainty introduced in a model either as a result of input uncertainty or during incomplete construction of a model. Thus, the modelling process is very dependent on training data, not only because of its uncertainties but also because the data has to include representative samples of land covers [15]. Furthermore, wrong assumptions on the relations within the training data can also lead to an increase in prediction errors; the error related to bias. Due to their black-box nature, random forest models are difficult to interpret and the inherent modelling and input uncertainties are difficult to quantify. Within the last ten years statisticians discovered desirable properties of RF that make the models more transparent, especially with regards to the quantification of prediction uncertainties. In fact, the uncertainty of the RF predictions can be estimated using several approaches, one of them is the quantile regression forests method [16], which

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	12	

estimates the prediction intervals. Other methods include the U-statistics approach as in [17] and the Monte Carlo simulations approach [18]. Many of these works have addressed the important challenge of the determination of prediction intervals that will contain with a given probability the value of an unobserved response.

Finally, it must be noted that supervised classifiers are based on class assignment rules that derive from a set of multiclass training samples. A not negligible quantity of uncertainty stems from the choice of the training set and uncertainties related to its labelling. The training data are samples of individual classes and the class assignment rules are derived from the entire study area. A high quality training dataset is mandatory to train the classifier model. In our learning problem, we have a training set associated with labels taking values in the set of HRLC classes. Collecting training patterns whose actual class is known with certainty may be expensive, difficult, or even impossible. In particular, reference data for large-scale land cover map acquired by visual interpretation of remotely sensed data suffer from little but important factors influencing the quality of visually interpreted data. There are many effects tied to land cover class labelling of interpreters, whose activity strongly depends by multiple variables. Visual interpretation of high-resolution imagery, even when delivered by well-trained professionals, is subject to interpreters' variation. Due to their perception of different land cover types, interpreters may incur in errors. This may have a substantial impact on the later uses of this reference dataset, as mentioned in the work by McRoberts [19]. In summary, reference data in the form of visual interpretations of remotely sensed data, even by well trained professional interpreters, are subject to substantial interpreter disagreement and error. If the reference data are imperfect in the sense of being subject to error, then the stratified estimators may be biased, and sometimes substantially biased despite the errors may be small. The use of inputs from at least three experienced interpreters to mitigate this effect is recommended. In conclusion, in practice, the label (class name) in a training dataset may not be correct (when it is generated by a human interpreter, for instance, mistakes are going to happen). Therefore, if the training dataset is not of high quality, it may lead to lower classification performances. Thus, classification accuracy is inherently associated with uncertainty [20].

6 Decision fusion



Data fusion methodologies should consider source-specific uncertainties in order to estimate the overall uncertainty of the classification result. More in detail, decision fusion combines the posterior probabilities associated with the outputs of single classifiers when applied to the single data inputs, here namely optical and SAR data. Therefore, multiple decisions are combined into a final result by taking into account the level of uncertainty associated with each source, which is intrinsically expressed by the corresponding pixelwise posterior probability distribution.

As described in ATBD-v3, the whole class legend Ω is divided into: Ω_O , the set of classes that are distinguished only by using optical data ("optical-exclusive"); Ω_S , the set of classes that are distinguished only by using SAR data ("SAR-exclusive"); and Ω_C , the set of classes that are discriminated by the classifiers operating with both data modalities (common classes). The optical classifier works on the set of classes $\Omega_O \cup \Omega_C$, the SAR classifier outputs posterior probabilities for the set of classes $\Omega_S \cup \Omega_C$. The decision fusion stage first merges the optical and SAR outputs on the common classes Ω_C , then it takes into account the presence of the exclusive classes Ω_O and Ω_S through a class-specific combination rule.

The following subsections discuss uncertainty modelling issues with regard to the families of decision fusion methods that are developed, i.e., (i) weighted voting and consensus theory, and (ii) fusion based on Markovian modelling (both families are combined with the aforementioned class-specific combination rules).

6.1 Uncertainty in consensus theory

Consensus theory [21], [22] involves general procedures with the goal of combining multiple probability distributions to summarize their estimates. Since the use of consensus theory simply aims at fusing posterior

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	13	

probabilities coming from different classifiers, it is possible to obtain again a probability distribution. The source-specific uncertainties are therefore directly combined in the process, leading to an overall uncertainty.

In the case of the HRLC pipeline, the individual information sources correspond to the outputs of the optical and SAR processing chains. Since the two classifiers generally work on different sets of classes, this fusion is possible on the set of common classes Ω_C only. Considering this specific case, the two major consensus-theoretic approaches, i.e., linear and logarithmic opinion pool (LOP and LOGP) compute functionals $\mathcal{C}(\omega_j|\underline{x}, \Omega_C)$ and $\mathcal{L}(\omega_j|\underline{x}, \Omega_C)$ ($\omega_j \in \Omega$) that merge the pixelwise posteriors provided by the optical and SAR processing chains, conditioned on the common classes Ω_C and as a function of the multisensor feature vector \underline{x} (see ATBD-v2).

\mathcal{C} and \mathcal{L} do not determine proper probability distributions *per se*, except in special cases (e.g., when the weights sum to one in the case of \mathcal{C} or when they are uniform in the case of \mathcal{L}). However, both functionals can be normalized (linearly in the case of LOP and nonlinearly through a softmax operator in the case of LOGP) to derive a probability distribution $P_{\mathcal{F}}(\omega_j|\underline{x}, \Omega_C)$, that expresses a pixelwise measure of uncertainty on the set of common classes. As proven in ATBD-v2, this measure of pixelwise uncertainty is extended to the whole set of classes as:

$$P_{\mathcal{F}}(\omega_j|\underline{x}) = P_{\mathcal{F}}(\omega_j|\underline{x}, \Omega_C) [\lambda P(\Omega_C|\underline{O}, \Omega_O \cup \Omega_C) + (1 - \lambda) P(\Omega_C|\underline{S}, \Omega_S \cup \Omega_C)] \\ + \lambda P(\omega_j|\underline{O}, \Omega_O) P(\Omega_O|\underline{O}, \Omega_O \cup \Omega_C) + (1 - \lambda) P(\omega_j|\underline{S}, \Omega_S) P(\Omega_S|\underline{S}, \Omega_S \cup \Omega_C),$$

where purple terms result from consensus-theoretic fusion on the common classes, blue terms are derived from the output of the optical chain and regard the optical-exclusive classes, red terms are similarly associated with SAR-exclusive classes, and $\lambda \in [0,1]$ is a weight computed as a function of the prior probabilities. The resulting $P_{\mathcal{F}}(\omega_j|\underline{x})$ yields a probability distribution that expresses a pixelwise measure of uncertainty after the consensus processing stage.

6.2 Uncertainty in Markov Random Fields

Markov random fields (MRFs) are probabilistic graphical models able to include contextual information in the form of class interactions between neighbouring pixels. As discussed in ATBD-v3, an MRF is determined by an energy function, whose minimization with respect to the labels is equivalent to the application of the maximum *a-posteriori* (MAP) criterion [23]:

$$Y^{MAP} = \underset{Y}{\operatorname{argmax}} P(Y|X),$$

where Y and X indicate the random fields of all class labels and feature vectors, respectively, across the whole pixel grid I . In particular, the Hammersley-Clifford theorem specifies the relation between such energy $U(\cdot)$ and the posterior probability $P(Y|X)$:


$$P(Y|X) = \frac{1}{Z} \exp(-U(Y|X)), \quad Z = \sum_Y \exp(-U(Y|X)),$$

where Z is the normalization constant (named partition function). Considering the MRF model in which only up to pairwise clique potentials are non-zero, then the energy can be written as:

$$U(Y|X) = - \sum_{i \in I} \alpha \log P_{\mathcal{F}}(y_i|\underline{x}_i) - \gamma \sum_{\substack{i \in I \\ j \in \partial i}} \delta(y_i, y_j),$$

where $y_i \in \Omega$ is the class label of the i th pixel, \underline{x}_i is its feature vector, ∂i is its neighborhood, $P_{\mathcal{F}}(y_i|\underline{x}_i)$ is derived from pixelwise fusion (see above), α and γ are weight coefficients, $\delta(\cdot)$ is the Kronecker impulse, the first summation considers pixelwise contributions, and the second one represents the pairwise interactions.

The uncertainty associated with the class labels predicted according to such an MRF model can be computed as a function of the corresponding energy. Indeed, using the aforementioned Hammersley-Clifford theorem yields the global posterior probability $P(Y|X)$ that is not a pixel-wise measure of uncertainty and is generally hard to compute because the partition function Z is intractable except in special cases [23].

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	14	

However, the local contextual pixelwise probability $P(y_i | \underline{x}_i, \{y_j\}_{j \in \partial i})$, i.e., the distribution of the class label of each pixel, conditioned to its observations from all sources and to the labels of the neighbouring pixels, can be derived from the energy [23] and provides a spatial-contextual measure of uncertainty of the predicted land cover.

While LOGP directly conveys uncertainty in the form of $P_{\mathcal{F}}(\omega_j | x)$, the MRF formulation is based on an energy function. Therefore, the uncertainty in MRF should be estimated according to its energy. In order to do that, a softmax is applied:

$$\tilde{P}_{MRF}(y_i | \underline{x}_i, \{y_j\}_{j \in \partial i}) = \frac{\exp[-U(y_i | x_i, \{y_j\}_{j \in \partial i})]}{\sum_{\omega_k \in \Omega} \exp[-U(\omega_k | x_i, \{y_j\}_{j \in \partial i})]}$$

However, the MRF energy depends on the weights α and γ that tune the tradeoff among the various contributions to the energy function. This parameters also have an impact on the resulting uncertainty – an undesired behaviour given the different meaning of these parameters. In order to output an uncertainty distribution that both complies with the posteriors coming from LOGP and reflects the spatial structure of the MRF output, the uncertainty estimation is also parametrized with a further parameter μ so that the overall distribution remains similar, although with enhanced spatial regularity:

$$\tilde{P}_{MRF}(y_i | \underline{x}_i, \{y_j\}_{j \in \partial i}) = \frac{\exp[-\mu U(y_i | x_i, \{y_j\}_{j \in \partial i})]}{\sum_{\omega_k \in \Omega} \exp[-\mu U(\omega_k | x_i, \{y_j\}_{j \in \partial i})]}$$

The value for μ is determined by comparing the distributions (histograms) of the LOGP and MRF uncertainty.



In principle, a further possible measure of spatial-contextual pixel-wise uncertainty would be $P(y_i | X)$, i.e., the probability distribution of the label of each pixel, conditioned to all image observations used to compute prediction. However, the calculation, estimation, and optimization of $P(y_i | X)$ corresponds to the use of the marginal *a-posteriori* modes (MPM) criterion to MRF-based classification rather than to the MAP criterion. On one hand, MPM formulations for MRF-based classifiers are computationally convenient in the case of multiscale quadtree graphs. On the other hand, they are remarkably time-expensive in the case of planar graphs because of the need to iteratively run time-consuming stochastic samplers (Gibbs or Metropolis sampling) [24]. Accordingly, the use of this MPM-based uncertainty measure is deemed substantially disadvantageous in the HRLC pipeline.

6.3 Uncertainty with Deep Learning

Several formulations involving deep learning are discussed in the ATBD-v3 with regard to classification and fusion stages. Deep neural networks [25] allow the computation of pixelwise uncertainty measures. The output of the last layer may normally be interpreted probabilistically by using softmax activation functions. Let $\sigma(\mathbf{z})$ be the output quantities, where $\sigma(\cdot)$ is the softmax function and the vector $\mathbf{z} = [z_1, z_2, \dots, z_C]$ collects the inputs resulting from the last hidden layer, with C being the number of classes. The softmax output is computed as:

$$\sigma(\mathbf{z})_i = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}}, \quad i = 1, 2, \dots, C.$$

The softmax function takes as input a vector of real numbers and normalizes it into a new vector of numbers that can be interpreted as expressing a probability distribution associated with the predicted label. After applying the softmax, each component of the input vector will be in the interval $(0, 1)$, and the components will sum up to 1, so that they can be interpreted as probabilities. Accordingly, picking the class that yields the largest value of $\sigma(\mathbf{z})_i$ ($i = 1, 2, \dots, C$) is interpreted as a formulation of the MAP criterion.

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	15	

Such consideration allows the deep learning formulations discussed in the ATBD-v3 to be automatically associated with pixelwise uncertainty measures by inspecting the output values of the softmax activation function, before thresholding such values in order to determine the class labels.

6.4 Uncertainty Output

The complete posterior probabilities may contain redundant information with respect to the goal of representing uncertainty, since the least probable classes may be associated with probability values close to zero. The final formulation for uncertainty has taken into account this comment and especially climate modellers' indications and data size/budget constraints.

As the first indicator of the uncertainty, together with the output classification map (corresponding to the most probable class on each individual pixel), also the second most probable one is considered on a pixelwise basis. Moreover, the posterior probability values for these best and second best classes are also provided together with the corresponding classification maps. This allows to greatly reduce the data size for the uncertainty while keeping the information that has been deemed more relevant by the Climate Group. In fact, that formulation shows among which classes the major uncertainty is and the relevance of the possible doubt (higher when the posterior for these two classes are similar and lower as their difference increases). In addition, the optical input quality index is included in the uncertainty output, as the optical data is the main source of information and its quality directly affects the classifier performance. Therefore, the output for uncertainty includes:

- **Classification maps** corresponding to best and second-best thematic classes;
- **Posterior probabilities** corresponding to best and second-best thematic classes.
- **Input quality index** corresponding to the optical input quality, related to the number and temporal distribution of the optical acquisitions used in the composite generation step.

Their specific format has been defined in agreement with both the Climate Group and the Engineering Team to take into account both readability and storage constraints.

6.5 Uncertainty in the multitemporal cascade model

The adopted cascade approach is a rigorous probabilistic Bayesian strategy to incorporate temporal dependence information in the classification of a time series. In that sense, it combines the posterior probabilities coming from LOGP at time t_1 with the posterior probabilities at time t_0 . This process directly conveys uncertainty in the form of probabilities [26]:



$$P_F(y_i^1 | x_i^1, x_i^0) \propto \frac{P_F(y_i^1 | x_i^1)}{P(y_i^1)} \sum_{y_i^0} \frac{P_F(y_i^0 | x_i^0)}{P(y_i^0)} P(y_i^0, y_i^1),$$

where $P_F(y_i^0 | x_i^0)$ are the fused posteriors at t_0 (the ones to propagate backward in time; see ATBD-v3) and $P_F(y_i^1 | x_i^1)$ are the ones at t_1 . It is worth noting that $P_F(y_i^0 | x_i^0)$ coincides with the final posterior probabilities obtained at time t_0 , as described in Section 6.2. Then, $P(y_i^1)$ and $P(y_i^0)$ are the prior probabilities corresponding to t_1 and t_0 , respectively, which are often omitted considering that spatial MRF prior is already in the model. Then $P(y_i^0, y_i^1)$ is the joint probability matrix (JPM) representing the chances of having a temporal transition among certain classes.

The subsequent application of MRF at time $t = t_1$ follows the same criteria described in Section 6.2:

$$\tilde{P}_{MRF} \left(y_i^1 | x_i^1, x_i^0, \{y_j^1\}_{j \in \partial i} \right) = \frac{\exp \left[-\mu U \left(y_i^1 | x_i^1, x_i^0, \{y_j^1\}_{j \in \partial i} \right) \right]}{\sum_{\omega_k \in \Omega} \exp \left[-\mu U \left(\omega_k | x_i^1, x_i^0, \{y_j^1\}_{j \in \partial i} \right) \right]},$$

thus leading, in the historical case with the multitemporal cascade as well, to a pixelwise uncertainty measure that plays the same role as in the static case.

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	16	

7 Multitemporal change detection and trend analysis

As the last part of the CCI HRLC processing chain, the multitemporal change detection and trend analysis is highly influenced by the uncertainties coming from previous steps. In particular, uncertainty is co-related to: i) the decision fusion step, ii) the classification maps and iii) the multisensor geolocation part. In consequence, the analysis done in previous steps applies in the same way for this last step.

Since products are developed at pixel-level (see [AD2] for further details), uncertainty will be associated in the same way but with more specifically, to the products from abrupt change detection maps (30m). As described in [AD2], for abrupt/permanent change maps the Break For Additive Seasonal and Trend (BFAST) [27] allows the computation of pixelwise uncertainty measures and there is indeed a function to save the confidence/probability of the BFAST method. BFAST is a method that analyses a univariate time series and decomposes it into three components: trend (T_t), seasonal (S_t) and noise (e_t). The general model of BFAST is an additive decomposition of the tree components in the form of:

$$Y_t = T_t + S_t + e_t \quad (t = 1, 2, \dots, n)$$



In which Y_t is the observed data at time t . The model iteratively fits a piecewise linear trend and seasonal model, and the intercept and slope of consecutive linear models define breaks in the signal. The breaks detected in the pixels denote a potential change in the time series with a date associated. BFAST is available from the Comprehensive R Archive Network (CRAN) but is computationally heavy. As a result, it has implemented in python to increase the code efficiency and sophisticated functionalities to improve the performance of BFAST. BFAST provides several elements as the output. One of the outputs is bp.Vt (the output of the breakpoints function for the trend model) that represents the detected breaks in the trend component considering a segment size between potentially detected breaks. Also ci.Vt is the output of the breakpoints confint function for the trend model that defines lower and upper limits of the confint function in the date interval but always associated to a default 95% confidence level. Since a probability for the change date is required, the possible implementation is calling separately the confint function in order to retrieve the confidence level associated to a change happening within a day from the breakpoint. Starting from a given confidence value $lv = 0.99$, the model decreases the level of confidence (each time 0.01 percent) to reach a fixed "change interval" of length 4. Acquired confidence level refers to the percentage of probability, or certainty, that the confidence interval would contain the true date of change.

The slope of consecutive linear models in BFAST can be used to derive the magnitude and direction of abrupt changes. The magnitude of change tends to be relatively large for real abrupt changes. So, the magnitude of the breaks detected is key to distinguish real abrupt changes from other disturbances of the signal in the time series. Moreover, the magnitude of change can be applied to define the uncertainty associated with each pixel.



As described in [AD2], the feature magnitude is calculated between two adjacent years if enough data is available for each year (at least three acquisitions in consecutive months in a year). If enough data is not available for consecutive years, the algorithm keeps the first year (with enough data) and checks the number of acquisitions in the other years to find a year with sufficient acquisitions in the whole six years of data. If the data availability is verified, the year of the change and the probability of change are reported. Thus, considering this procedure the value of described as the reliability is introduced that shows the distance between the couple of years that the change information is provided and can be considered as a uncertainty value for the reported year of the change. The value is a number between 1 – 5 for the changed pixels and 0 for no change.

8 References

- [1] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, Eds., *Image Registration for Remote Sensing*. Cambridge: Cambridge University Press, 2011. doi: 10.1017/CBO9780511777684.

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	17	

- [2] S. Savastano and R. Guida, "Uncertainty Quantification in Synthetic Aperture Radar Remote Sensing Data Processing," in *IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium*, Jul. 2018, pp. 9193–9196. doi: 10.1109/IGARSS.2018.8518217.
- [3] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*. 1991. Accessed: Jul. 02, 2019. [Online]. Available: <https://www.wiley.com/en-ao/Synthetic+Aperture+Radar%3A+Systems+and+Signal+Processing-p-9780471857709>
- [4] T. L. Logan, "Terrain Correction of Synthetic Aperture Radar Imagery Using the Cray T3D," 1995.
- [5] C. E. Wivell, C. Olmsted, D. R. Steinwand, and C. S. Taylor, "An earth remote sensing satellite- 1 Synthetic Aperture Radar Mosaic of the Tanana River Basin in Alaska," 1993.
- [6] T. Bayer, R. Winter, and G. Schreier, "Terrain influences in SAR backscatter and attempts to their correction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 29, no. 3, pp. 451–462, May 1991, doi: 10.1109/36.79436.
- [7] T. Lillesand, R. W. Kiefer, and J. Chipman, "Remote Sensing and Image Interpretation, 7th Edition," *Wiley.com*, 2015. <https://www.wiley.com/en-it/Remote+Sensing+and+Image+Interpretation%2C+7th+Edition-p-9781118343289> (accessed Jul. 02, 2019).
- [8] Jong-Sen Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 5, pp. 2363–2373, Sep. 1999, doi: 10.1109/36.789635.
- [9] G. Lisini, A. Salentini, P. Du, and P. Gamba, "SAR-Based Urban Extents Extraction: From ENVISAT to Sentinel-1," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 8, pp. 2683–2691, Aug. 2018, doi: 10.1109/JSTARS.2017.2782180.
- [10] J. Le Moigne, N. S. Netanyahu, and R. D. Eastman, Eds., *Image Registration for Remote Sensing*. Cambridge: Cambridge University Press, 2011. doi: 10.1017/CBO9780511777684.
- [11] I. Zavorin and J. Le Moigne, "Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery," *IEEE Transactions on Image Processing*, vol. 14, no. 6, pp. 770–782, Jun. 2005, doi: 10.1109/TIP.2005.847287.
- [12] S. Ahmed and Md. M. Rahman, "Handling Uncertainty under Spatial Feature Extraction through Probabilistic Shape Model (PSM)," *International Journal of Science and Research (IJSR)*, vol. 9, no. 2, Sep. 2013, Accessed: Dec. 19, 2019. [Online]. Available: https://www.researchgate.net/publication/264974196_Handling_Uncertainty_under_Spatial_Feature_Extraction_through_Probabilistic_Shape_Model_PSM
- [13] L. Loosvelt *et al.*, "Random Forests as a tool for estimating uncertainty at pixel-level in SAR image classification," *International Journal of Applied Earth Observation and Geoinformation*, vol. 19, pp. 173–184, Oct. 2012, doi: 10.1016/j.jag.2012.05.011.
- [14] F. Löw, P. Knöfel, and C. Conrad, "Analysis of uncertainty in multi-temporal object-based classification," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 105, pp. 91–106, Jul. 2015, doi: 10.1016/j.isprsjprs.2015.03.004.
- [15] K. Baake, "Quantifying Uncertainty of Random Forest Predictions: A Digital Soil Mapping Case Study," p. 84.
- [16] N. Meinshausen, "Quantile Regression Forests," p. 17.
- [17] L. Mentch and G. Hooker, "Quantifying Uncertainty in Random Forests via Confidence Intervals and Hypothesis Tests," p. 41.
- [18] J. Coulston, C. Blinn, V. Thomas, and R. Wynne, "Approximating Prediction Uncertainty for Random Forest Regression Models," *Photogrammetric Engineering & Remote Sensing*, vol. 82, pp. 189–197, Mar. 2016, doi: 10.14358/PERS.82.3.189.
- [19] R. E. McRoberts, S. V. Stehman, G. C. Liknes, E. Næsset, C. Sannier, and B. F. Walters, "The effects of imperfect reference data on remote sensing-assisted estimators of land cover class proportions," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 142, pp. 292–300, Aug. 2018, doi: 10.1016/j.isprsjprs.2018.06.002.
- [20] L.-H. Hsiao and K.-S. Cheng, "Assessing Uncertainty in LULC Classification Accuracy by Using Bootstrap Resampling," *Remote Sensing*, vol. 8, no. 9, p. 705, Sep. 2016, doi: 10.3390/rs8090705.

	Ref	CCI_HRLC_Ph1-E3UB		
	Issue	Date	Page	
	4.rev.0	31/10/2022	18	

- [21] J. A. Benediktsson, "Hybrid consensus theoretic classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 35, no. 4, pp. 833–843, Jul. 1997, doi: 10.1109/36.602526.
- [22] J. A. Benediktsson and P. H. Swain, "Consensus theoretic classification methods," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 22, no. 4, pp. 688–704, 1992, doi: 10.1109/21.156582.
- [23] S. Z. Li, *Markov random field modeling in image analysis*. Springer, 2009.
- [24] I. Hedhli, G. Moser, J. Zerubia, and S. B. Serpico, "A New Cascade Model for the Hierarchical Joint Classification of Multitemporal and Multiresolution Remote Sensing Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 11, pp. 6333–6348, Nov. 2016, doi: 10.1109/TGRS.2016.2580321.
- [25] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*.
- [26] G. Storvik, R. Fjortoft, and A. H. S. Solberg, "A bayesian approach to classification of multiresolution remote sensing data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 3, pp. 539–547, Mar. 2005, doi: 10.1109/TGRS.2004.841395.
- [27] J. Verbesselt, R. Hyndman, G. Newnham, and D. Culvenor, "Detecting trend and seasonal changes in satellite image time series," *Remote Sensing of Environment*, vol. 114, pp. 106–115, 2010.