



PROGRAMME OF  
THE EUROPEAN UNION



# Validation Methodologies and Procedures

SERVICES SUPPORTING THE EUROPEAN ENVIRONMENT AGENCY'S (EEA) IMPLEMENTATION OF THE COPERNICUS EUROPEAN GROUND MOTION SERVICE – PRODUCT VALIDATION.



Date: **20/10/2023**

Doc. Version: **1.0**

Content ID: **SPECIFIC CONTRACT No 3506/R0-COPERNCA/EEA.58973**  
**Implementing Framework service contract No EEA/DIS/R0/21/009**

Kongens Nytorv 6  
1050 Copenhagen K  
Denmark  
Tel.: +45 3336 7100  
Fax: +45 3336 7199  
[eea.europa.eu](http://eea.europa.eu)



## Document Control Information

|                         |   |
|-------------------------|---|
| <b>Document Title</b>   | EGMS Validation Methodologies and Procedures  |
| <b>Project Title</b>    | Services supporting the European Environment Agency's (EEA) implementation of the Copernicus European Ground Motion Service – product validation. |
| <b>Document Authors</b> | Joan Sala Calero, Malte Vöge, Joana Esteves Martins, Daniel Raucoules, Marcello de Michelle, Amalia Vradi, Filippo Vecchiotti                     |
| <b>Project Owner</b>    | Lorenzo Solari  |
| <b>Project Manager</b>  | Joan Sala Calero  |
| <b>Document Code</b>    | D6.1_Validation_Methodologies_and_Procedures.docx   |
| <b>Document Version</b> | 1.0   |
| <b>Distribution</b>     | Public  |
| <b>Date</b>             | 20/10/2023  |

### Document Approver(s) and Reviewer(s):

| Name           | Role          | Action        | Date       |
|----------------|---------------|---------------|------------|
| Lorenzo Solari | Project owner | Approved v1.0 | 01/03/2023 |
| Lorenzo Solari | Project owner | Approved v1.1 | 20/10/2023 |

### Document history:

| Revision | Date       | Created by       | Short description of changes |
|----------|------------|------------------|------------------------------|
| Draft    | 30/09/2022 | Joan Sala Calero | Initial draft version        |
| v1.0     | 31/01/2023 | Joan Sala Calero | Final version                |
| v1.1     | 20/10/2023 | Joan Sala Calero | Sub-selection of sites       |

## Contents:

|   |    |
|---|----|
| 1. INTRODUCTION .....   | 4  |
| 2. INDEX OF AGREEMENT (IoA) .....                                       | 5  |
| 3. VALIDATION METHODOLOGY DESCRIPTION .....                             | 6  |
| 3.1. Point density check.....   | 6  |
| 3.1.1    Introduction.....  | 6  |
| 3.1.2    CLC groups and classes .....                                   | 6  |
| 3.1.3    Burst density approach.....                                    | 8  |
| 3.2. Comparison with other ground motion services.....                  | 10 |
| 3.2.1    Introduction.....  | 10 |
| 3.2.2    Pre-processing .....   | 10 |
| 3.2.3    Intercomparison .....  | 11 |
| 3.2.4    Computation of IoA .....                                       | 13 |
| 3.3. Comparison with inventories of phenomena/events.....               | 15 |
| 3.3.1    Introduction.....  | 15 |
| 3.3.2    Detection capability assessment .....                          | 15 |
| 3.3.3    Mapping capability assessment.....                             | 16 |
| 3.3.4    Monitoring capability assessment.....                          | 16 |
| 3.3.5    Estimation of the percentage of mapped phenomena covered ..... | 17 |
| 3.4. Consistency check with ancillary geo-information .....             | 18 |
| 3.4.1    Introduction.....  | 18 |
| 3.4.2    Validation Approach .....                                      | 18 |
| 3.4.3    Validation Procedures by Dataset Type and Phenomena .....      | 19 |
| 3.4.4    Result Interpretation and Determination of IoA .....           | 21 |
| 3.5. Comparison with GNSS data.....                                     | 23 |
| 3.5.1    Introduction.....  | 23 |
| 3.5.2    Methodology .....  | 24 |
| 3.5.3    Computation of the IoA.....                                    | 26 |
| 3.6. Comparison with in-situ monitoring data .....                      | 27 |
| 3.6.1    Introduction.....  | 27 |
| 3.6.2    Type of measurement available .....                            | 27 |
| 3.6.3    Workflow for in-situ/EGMS (velocity) .....                     | 28 |
| 3.6.4    Workflow for in-situ/EGMS (time series) .....                  | 28 |
| 3.7. Evaluation XYZ and displacements with Corner Reflectors .....      | 31 |
| 3.7.1    Introduction.....  | 31 |
| 3.7.2    The estimated height of the MPs around the CR location .....   | 31 |
| 3.7.3    Geopositioning accuracy by XY offset estimation .....          | 31 |
| 3.7.4    Quality of the EGMS time-series displacements .....            | 32 |
| Appendix 1: ADA common approach.....                                    | 33 |



# 1. INTRODUCTION

This document is part the documentation connected to the European Ground Motion Service (EGMS) products validation as specified in the framework contract EEA/DIS/R0/21/009. The validation of the Interferometric Synthetic Aperture Radar open datasets at continental scale has the following goals and considerations:

- It verifies the **usability** of the data for different applications according to initial user requirements and with respect to the fields of application foreseen by the Validation of the EGMS Product Portfolio<sup>1</sup> and the EGMS End User Requirement<sup>2</sup> documents.
- It determines if the **quality** of the products is consistent with the technical specifications for different areas and applications. It is used to confirm the conclusions of the EGMS Quality Assurance and Control Reports.
- It addresses the **completeness** and **consistency** of the data products together with their **accuracy**.
- It is performed **independently** from the EGMS production. The validation datasets used, and the chosen procedures/criteria are documented in deliverables D3.1, D5 and D6.1.
- It is based on the comparison of data of different nature. Therefore, a complete agreement is most likely impossible, and differences may not be related to a quality issue.

In the following sections, the data workflows and methodologies developed to carry out validation activities are described. Methodologies and approaches have been adequately referenced for the sake of reproducibility.

---

<sup>1</sup> <https://land.copernicus.eu/user-corner/technical-library/egms-end-user-requirements-document>

<sup>2</sup> <https://land.copernicus.eu/user-corner/technical-library/validation-approach-of-the-egms-product-portfolio>

## 2. INDEX OF AGREEMENT (IoA)

As stated in the introduction, the objectives of the validation exercise are strongly focused on **usability, completeness, and consistency**.

For the sake of **reproducibility**, validation data and EGMS data exports or subsets from different product levels are collected and stored in the EGMS validation environment together with the necessary scripts/notebooks to evaluate the aforementioned criteria.

This evaluation follows the high-level methodology described in section 3 of this document and establishes **agreement** between two different data sources (validation and EGMS). In no case agreement should be taken as comparison with absolute ground truth. For example, in the comparison of different ground motion services, the IoA should be regarded as a data science exercise to discover disagreements between two datasets complemented by experts' knowledge and available literature.

In order to guarantee **reproducibility** and yield the different validation methodologies **comparable**, the Index of Agreement (IoA) will be referenced across the validation documentation. This normalised index (0..1) is an indication of how validation and EGMS data agree based on expert analysis as it may contain a qualitative/empiric component:

Table 1: Index of Agreement (IoA) summary

| IoA                      | Description  |
|--------------------------|--|
| 1.00 to 0.75<br>(high)   | Validation data and EGMS subsets have a high level of agreement. Both measures are highly correlated and capture the same phenomena.   |
| 0.75 to 0.25<br>(medium) | The EGMS subsets and the validation data have a medium level of agreement. Noise and contributions to the disagreement should be detailed.   |
| < 0.25<br>(low)          | The validation data and EGMS subsets do not seem to agree. In these exceptional cases most probably one of the two sources is not capturing the observed phenomena and should be not taken into account. |



## 3. VALIDATION METHODOLOGY DESCRIPTION

In this section the workflows and algorithms used to validate EGMS data are detailed for each Validation Activity (VA). They take as inputs the validation datasets and sites described in deliverables D3.1 and D5.

### 3.1. Point density check

#### 3.1.1 *Introduction*

The objective of this validation activity is it to be a complement of the existing point density quality controls deployed in the EGMS production to ensure consistency over the EU territories. Sites have been selected to represent the four EGMS production data providers and with a special focus on mountainous areas.

Taking the coordinates of the EGMS timeseries enables the estimation of densities of each of the CLC Urban Atlas classes comparing the different algorithms. Moreover, other indicators (timeseries standard deviation, coherence) are studied to ensure they are consistent across the different EGMS coverage. As a validated land cover input, Urban Atlas (version 2018) verified datasets were selected.

The following table summarizes the different usability criteria to be evaluated:

Table 2: Usability criteria

| Criteria                   | Description  |
|----------------------------|--|
| Completeness               | Data gaps and density measurements should be consistent with land cover classes prone to landscape variation.  |
| Consistency                | Point density should be consistent across the same land cover class for different regions (e.g. urban classes shall have higher density than farming grounds). On top of that, consistency is also expected between ascending and descending products.                 |
| Pointwise quality measures | The output statistical and quality parameters of the EGMS results shall be key in assessing their quality. For instance, in general terms it is expected that the temporal coherence to be higher in urban classes whereas the root-mean-square error should be lower. |

#### 3.1.2 *CLC groups and classes*

To simplify the analysis, the IoA results have been aggregated and presented for each of the main CLC groups:

Table 3: CLC groups



| CLC group                    | Expected properties  |
|------------------------------|--|
| Artificial Surfaces          | The ideal reflectors are found in these classes. Point density should be <b>high</b> (above 1000 MP/km <sup>2</sup> ).   |
| Forest and seminatural areas | These areas might be subject to constant surface/roughness changes and their reflectivity properties are not ideal. Point density should be <b>medium</b> for these areas. |
| Agricultural areas           | These areas are changing yearly due to farming activities. Point density should be <b>low</b> for these areas.   |
| Wetlands                     | Point density should be <b>very low</b> for these areas since they are affected by inundation and vegetation changes.  |
| Water bodies                 | Classes listed in this section may be used to spot processing artifacts or land cover mapping errors. Point density should be close to zero.                               |

Some of the land cover density and quality parameters can be extracted from existing literature coming from the EGMS advisory board<sup>1</sup> and production team<sup>2</sup>. The remaining CLC class expected density values have been empirically determined by comparing the different values across the EGMS coverage and data producers (marked then as Empirical and column filled with estimates). The fixed thresholds were defined with the raster version of CLC which has a lower resolution than the Urban Atlas layers:

Table 4: Density values and thresholds available (defined with Corine Land Cover raster)

| CLC group           | CLC class  | MP/km2        | Fixed | Empirical |
|---------------------|--|---------------|-------|-----------|
| Artificial surfaces | <a href="#">1.1.1 Continuous urban fabric</a>                    | 5000 to 10000 | X [1] |           |
| Artificial surfaces | <a href="#">1.1.2 Discontinuous urban fabric</a>                 | 1000 to 5000  | X [1] |           |
| Artificial surfaces | <a href="#">1.2.1 Industrial or commercial units</a>             | 1000 to 5000  | X [1] |           |
| Artificial surfaces | <a href="#">1.2.2 Road and rail networks and associated land</a> | 1000 to 5000  |       | X         |
| Artificial surfaces | <a href="#">1.2.3 Port areas</a>                                 | 1000 to 5000  |       | X         |
| Artificial surfaces | <a href="#">1.2.4 Airports</a>                                   | 1000 to 5000  |       | X         |
| Artificial surfaces | <a href="#">1.3.1 Mineral extraction sites</a>                   | < 1000        |       | X         |
| Artificial surfaces | <a href="#">1.3.2 Dump sites</a>                                 | < 1000        |       | X         |
| Artificial surfaces | <a href="#">1.3.3 Construction sites</a>                         | < 1000        |       | X         |
| Artificial surfaces | <a href="#">1.4.1 Green urban areas</a>                          | < 1000        |       | X         |
| Artificial surfaces | <a href="#">1.4.2 Sport and leisure facilities</a>               | < 1000        |       | X         |
| Agricultural areas  | <a href="#">2.1.1 Non-irrigated arable land</a>                  | < 25          |       | X         |
| Agricultural areas  | <a href="#">2.1.2 Permanently irrigated land</a>                 | < 25          |       | X         |

<sup>1</sup> <https://land.copernicus.eu/user-corner/technical-library/validation-approach-of-the-egms-product-portfolio><sup>2</sup> <https://land.copernicus.eu/user-corner/technical-library//quality-assurance-and-control-report-2013-harmonisation-test>

|                              |  |       |   |   |
|------------------------------|--|-------|---|---|
| Agricultural areas           | <a href="#">2.1.3 Rice fields</a>  | < 25  |   | x |
| Agricultural areas           | <a href="#">2.2.1 Vineyards</a>  | < 25  |   | x |
| Agricultural areas           | <a href="#">2.2.2 Fruit trees and berry plantations</a>  | < 25  |   | x |
| Agricultural areas           | <a href="#">2.2.3 Olive groves</a>   | < 25  |   | x |
| Agricultural areas           | <a href="#">2.3.1 Pastures</a>   | < 25  |   | x |
| Agricultural areas           | <a href="#">2.4.1 Annual crops associated with permanent crops</a>   | < 25  |   | x |
| Agricultural areas           | <a href="#">2.4.2 Complex cultivation patterns</a>   | < 25  |   | x |
| Agricultural areas           | <a href="#">2.4.3 Land principally occupied by agriculture, with significant areas of natural vegetation</a> | < 25  |   | x |
| Agricultural areas           | <a href="#">2.4.4 Agro-forestry areas</a>  | < 25  |   | x |
| Forest and seminatural areas | <a href="#">3.1 Forest</a>   | < 25  |   | x |
| Forest and seminatural areas | <a href="#">3.2 Shrub and/or herbaceous vegetation associations</a>  | < 25  |   | x |
| Forest and seminatural areas | <a href="#">3.3 Open spaces with little or no vegetation</a>   | < 100 | x |   |
| Wetlands                     | <a href="#">4.1 Inland wetlands</a>  | < 5   |   | x |
| Wetlands                     | <a href="#">4.2 Coastal wetlands</a>   | < 5   |   | x |
| Water bodies                 | <a href="#">5.1 Inland waters</a>  | < 5   |   | x |
| Water bodies                 | <a href="#">5.2 Marine waters</a>  | < 5   |   | x |

### 3.1.3 Burst density approach

To establish a fair comparison between sites with different Sentinel-1 imagery coverage, a geospatial algorithm has been designed to handle overlapping bursts to generate fair density statistics. The procedure is run independently for ascending and descending orbits and identifies the common areas and outputs a collection of non-overlapping polygons. With these unique and non-overlapping polygons (Figure 2) all the collected bursts can be cropped covering each one of the sites (Figure 1).

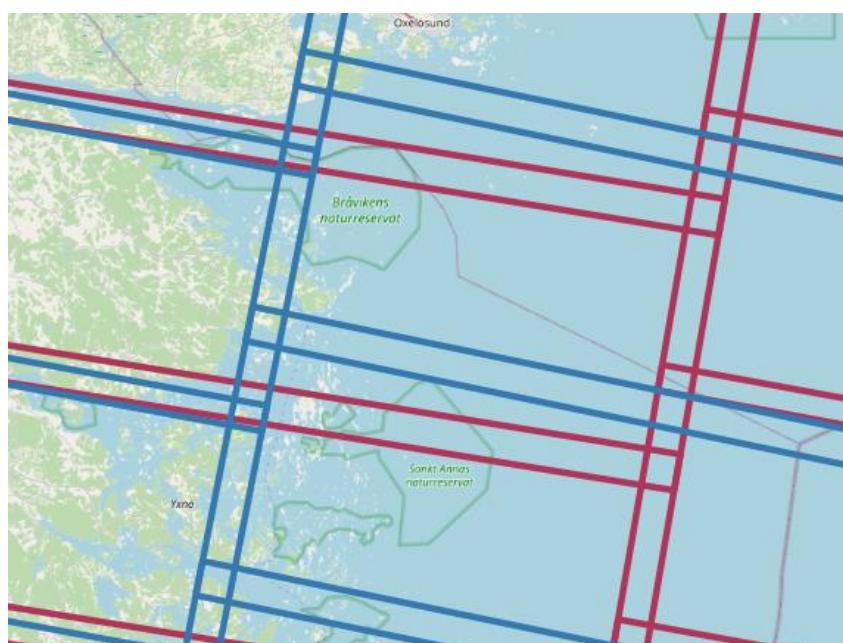


Figure 1: Overlapping descending bursts (IW\_022/IW\_095 in Sweden)

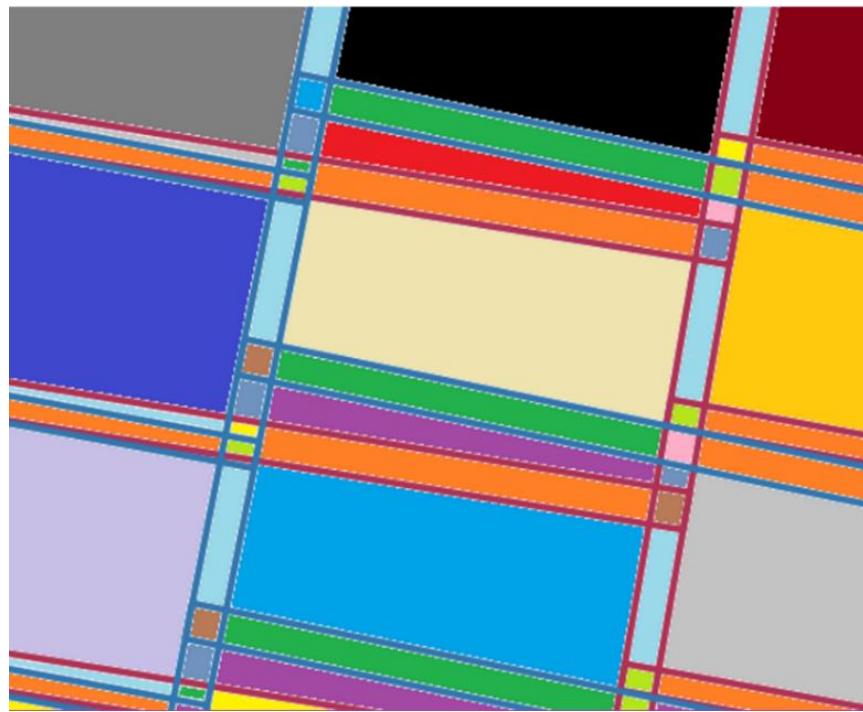


Figure 2: Unique polygons represented with random colours (IW\_022/IW\_095 in Sweden)

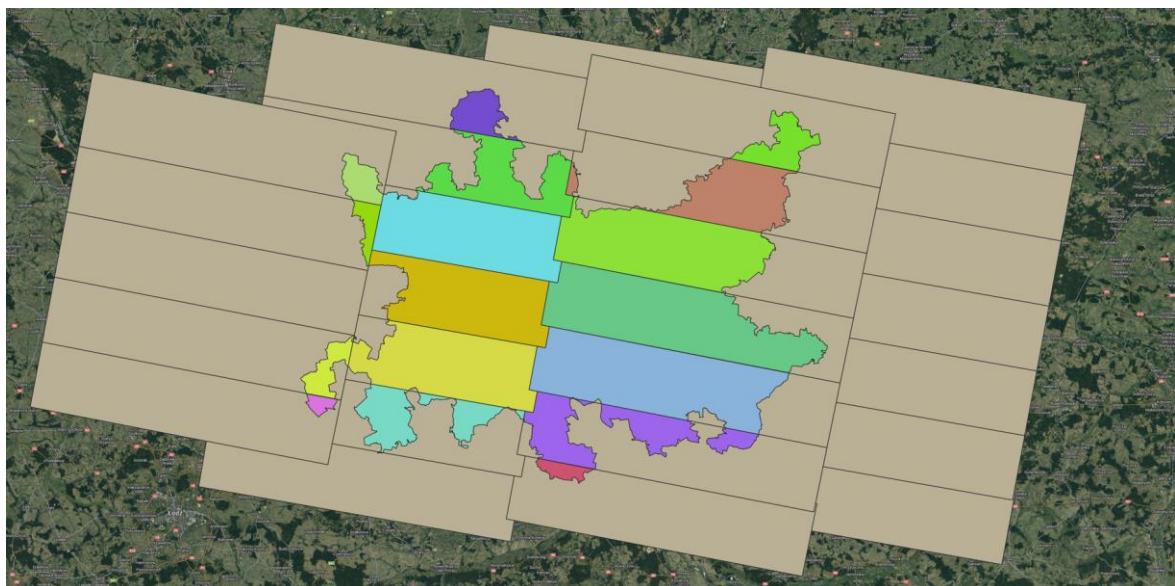


Figure 3: Unique bursts selected for MP density (Warsaw)



## 3.2. Comparison with other ground motion services

### 3.2.1 Introduction

The EGMS results have been validated against quality-controlled and validated datasets from existing operational national/regional Ground Motion Services (GMSs). The validation has been carried out separately for each product level and acquisition geometry. As the availability of product levels varies for the respective national/regional GMSs, the validation has been limited to the available datasets and performed only when levels are comparable. In addition, two more local datasets (Etna, Italy and Thyborøn, Denmark), both independently quality controlled and validated, are added to the validation system. This data is made available for this validation activity by data providers of the validation consortium.

Statistical measures have been calculated for Active Deformation Areas (ADAs), which have been identified in an automated procedure described in Appendix 1. The result of the ADA detection procedure is provided as polygons outlining the detected deformation areas.

The products compared will be dependent on product availability from the respective national GMS. For line-of-sight products, only datasets with matching track numbers, i.e., with identical acquisition geometry, will be compared. The EGMS datasets are supplied for each Sentinel-1 burst separately, while for most national GMSs all measurement points for a particular track number are merged to a single dataset. Therefore, the EGMS datasets have also been merged accordingly in the pre-processing stage (see below).

The non-GMS products used in this validation activity are based on well-established processing approaches and use processing chains that have been quality checked by comparison against GNSS data (IREA products for Etna) or by comparison against a GMS service referred to an earlier period (DTU products for Thyborøn). For further information on the quality checks for non-GMS data the reader can be referred to the respective section in the document "D3.1 Validation Data Collection".

For the intercomparison of InSAR datasets a focus has been made on statistical measures, following the approach proposed by Sadeghi et al. (2021), and have been composed of:

- Co-location of ADAs.
- Velocity intercomparison.
- Time-series intercomparison.
- Point density and coverage intercomparison.

### 3.2.2 |Pre-processing

In the pre-processing stage, all datasets are transformed, so that they are directly comparable with the EGMS datasets. This includes adjusting the temporal periods the datasets cover, including the actual temporal sampling, as well as adjusting the spatial sampling.

The pre-processing steps below (adapted from Sadeghi et al., 2021) describe the pre-processing steps for a pair of a national GMS datasets and the corresponding EGMS dataset:

- **Transformation of datasets to common coordinate reference system:** Both datasets have been transformed to "ETRS89-extended / LAEA Europe (EPSG:3035)" coordinates, as projected coordinates are more suitable for resampling data points to a uniform grid.

- **Clip of the time-series data to area of interest (AOI):** AOI defined by shapefile polygon in "ETRS89-extended / LAEA Europe (EPSG:3035)" reference system.
- **Clip of time-series data to maximum common overlap:** Determine common time-period covered by both datasets and (for L2 products) common dates. Clip time-series to common time-period. Re-reference time-series to zero deformation at first sample of common time-period.
- **Merge of all datasets with same track number covering AOI:** Re-estimate velocity for common time-window:
  - Mean velocity has been recalculated for the common time-period by linear regression.
  - Temporal coherence has been left unchanged, as it cannot be recalculated.
  - For L3 products: Interpolation of reference timeseries to sampling dates of EGMS dataset.
- **Spatial resampling to common regular grid with resolution of 30 x 30 m:** For each grid cell the properties of all original measurement points are averaged:
  - Mean velocity.
  - Mean coherence.
  - Mean time series.
 In addition, the number of original measurement points is counted and stored for later comparison of point densities.

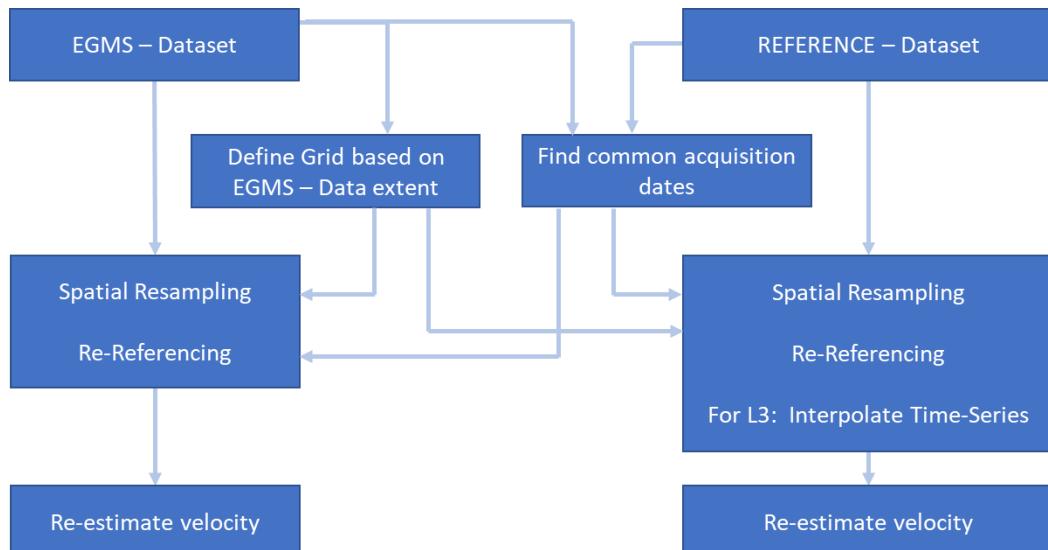


Figure 4: Flow-chart for the pre-processing stage.

### 3.2.3 Intercomparison

For each of the identified ADA within the validation areas, the velocities, deformation time series and point densities have been compared:

- **Spatial Overlap of ADAs:** To investigate what the general agreement of the datasets is with respect to the identification of the displacement phenomena; the identified ADAs have been compared. For each ADA in the EGMS dataset, all ADAs of the reference

dataset has been checked for an overlap and the overall coverage (in percent) of the ADA with ADAs from the validation dataset has been determined.

To define a robust measure of the overlap, that works well for comparing sites with a large number of rather small ADAs and sites with few large ADAs, the spatial overlap is calculated as the maximum of the following two measures:

- a) the percentage of ADAs that overlap by more than 30% of the combined area.
- b) the percentage of the total overlap area for all ADAs relative to the total area of all merged (EGMS and validation dataset) ADAs

To make the measure of overlap focus on differences in velocities and less dependent on differences in point coverage, only ADAs are considered, where both datasets have a certain minimum number of points (i.e., the minim number required for ADA detection) in the combined ADA area.

- **Velocity intercomparison:** Relative differences between the deformation velocities has been calculated for the common grid pixels. The relative velocity is measured in percent of the average of both dataset's velocity and at least 3 mm/year. The average is used instead of just the validation dataset's value to make the measure more stable in case the validation dataset's velocity value for a grid point is close to zero. Using a minimum of 3 mm/year ensures that the measure does return too high values, where both dataset's velocities are close to zero (e.g., at the edges of ADAs where velocity values are rather small). Mean and standard deviation of the relative velocity differences have been extracted for each ADA, as well as the correlation coefficient for estimated velocities of all common grid pixels.  
The spatial correlation of the velocity values is measured, i.e., the correlation of the velocity of individual grid points. The correlation coefficient is calculated using the Pearson correlation, which is a measure of linear correlation.  
In addition, noise levels have been calculated based on the measurement points outside of all ADAs, which have also been compared.
- **Time series comparison:** Differences between the time series for each common grid pixel have been calculated:
  - a) Mean and standard deviation for these time series differences have been extracted for each common grid pixel.
  - b) Correlation coefficients for the deformation time series have been estimated for each common grid pixel. The correlation coefficient is calculated using the Pearson correlation, which is a measure of linear correlation.
  - b) For each ADA the mean and standard deviation of the respective pixel-wise mean and standard deviation values have been extracted.
  - c) For each ADA the mean and standard deviation of the pixel-wise time series correlation coefficients have been extracted.
- **Point density and coverage intercomparison:** For comparing the point density and coverage, the measurements have been resampled to a coarser grid, e.g., 100 m x 100 m. For each of these pixels, the number of measurement points is counted in both datasets. This gives a direct comparison of the average point density within an

AOI/ADA. By counting the raster pixels that contain at least one InSAR measurement, a measure of the coverage has been retrieved, which can also be directly compared.

As the EGMS products and the various validation data products have been processed with different processing techniques and the choice of reference points, GNSS calibration and velocity reference frame differ for all datasets, an exact agreement between datasets has not been possible. This is an important aspect of the validation and should always be kept in mind when reviewing the results of this validation activity. To ensure a "fair" assessment of the validation measures produced in this VA, the translation of the above-described validation measures to IoA (see below) has been implemented accordingly.

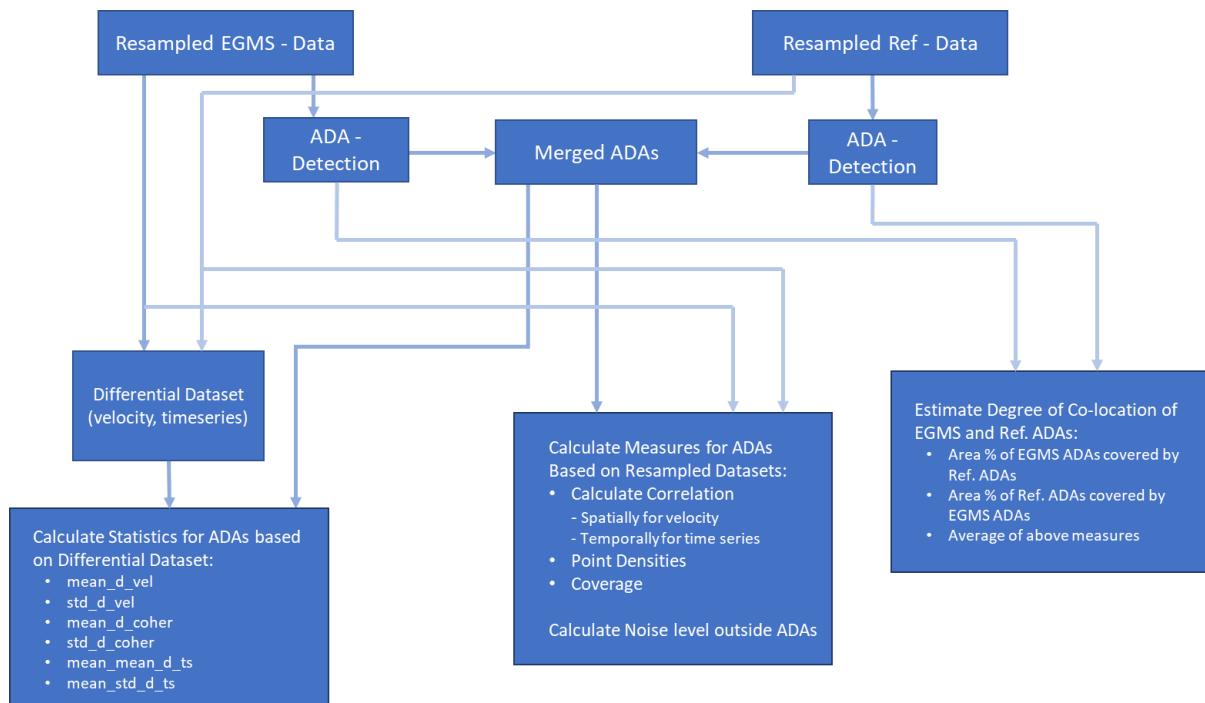


Figure 5: Flowchart for calculation of intercomparison measures

### 3.2.4 Computation of IoA

As mentioned above, these measures have been calculated for each ADA and for each product level and acquisition geometry. For each of the validation measures thresholds have been defined for  $\text{IoA} = 0$  and  $\text{IoA} = 1$ . Based on these thresholds,  $\text{IoA}$  have been interpolated linearly. Values resulting in  $\text{IoA}$  below 0 has been set to 0. Accordingly, values resulting in  $\text{IoA}$  above 1 have been set to 1.

It must be noted that not all above-mentioned validation measures have been used to calculate  $\text{IoA}$ , as they are not equally important for a meaningful validation of the consistency of the EGMS data with the validation datasets. The focus of the calculation of overall  $\text{IoA}$  has been on measures that describe if EGMS and validation dataset show the same displacement patterns with a good agreement in amplitude and overall temporal evolution of the displacement.

Table 5 shows preliminary thresholds for validation measure to  $\text{IoA}$  translation for several displacement phenomena and selected validation measures:

Table 5: Thresholds for validation measure to  $\text{IoA}$ .



| Validation Measure  |   | IoA = 0.0 | IoA = 1.0 |
|---------------------|---|-----------|-----------|
| Spatial_Overlap [%] | Maximum of:<br>-> $(\text{Number ADAs with } > 30\% \text{ overlap}) * 100 / \text{Number ADAs}$<br>-> $(\text{Total overlap area all ADAs} * 100) / (\text{Total area joined ADAs})$ | 30        | 80        |
| dV_rel_mean [%]     | $(\text{abs(dVel}) * 100) / \max(\text{abs(Vel}_{\text{egms}} + \text{Vel}_{\text{valid}})/2, 3.0)$   | 300       | 30        |
| Vel_Corr            | Spatial correlation of velocity values on common grid using Pearson Correlation Index (linear correlation)  | 0,3       | 0,8       |
| Disp_Corr           | Mean of Temporal correlation of time series using Pearson Correlation Index (linear correlation)  | 0,3       | 0,8       |

The spatial overlap calculation considers both phenomena with many small ADAs (e.g. landslides) as well as phenomena with few large ADAs (volcanism, mining), and the relative velocity difference measure normalizes the difference to the averaged displacement velocity, and thus, both measures are rather independent of the observed phenomenon. The same is true for the correlation indices for spatial and temporal correlation. Therefore, the same threshold for translation to Agreement Indices have been used for all validation sites (see Table 5). The thresholds have been chosen to reflect that a perfect agreement is not possible, and that the Agreement Indices should answer the question, if there is a general good agreement between the datasets, as described above.

IoA for these four measures has been calculated for each available acquisition geometry (track number) and each available product level, from which the corresponding four mean IoA have been calculated. The final IoA for a validation site has been calculated as average from the four averaged IoA. While IoA provide an objective measure of agreement between EGMS and reference datasets, they shall not be able to consider special circumstances and context. Therefore, it must be stressed that the oversight of an expert, by visual inspection of all outputs of the validation activity (maps, plots etc.), and the careful interpretation of the results is very important.

### References:

- Ester, M., Kröger, P., Sander, J., Xu, X. (1996). A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. KDD. 96. 226-231.
- Sadeghi, Z., Wright, T.J., Hooper, A.J., Jordan, C., Novellino, A., Bateson, L., Biggs, J. (2021). Benchmarking and Inter-Comparison of Sentinel -1 InSAR velocities and time series. Remote Sensing of Environment. 256. 112306. 10.1016/j.rse.2021.112306.

## 3.3. Comparison with inventories of phenomena/events

### 3.3.1 Introduction

The main activity carried out in VA3 aims at comparing the EGMS data with the information provided by inventories (points locating phenomena, polygons representing the geometry of the phenomena, expected velocity or qualitative characteristics of the motion, dates of events or damages) to test the capabilities on EGMS data. A particular and specific interest for geohazards management is the possibility of using EGMS data to complement (or even to build) inventories of phenomena as the existing inventories are not exhaustive (and sometimes do not exist on specific areas). In this perspective the capability of detecting previously known phenomena has been the main capability related to inventories to be tested. When available, the similarity between polygons derived from inventories and from EGMS is a second main validation indicator for the capability for mapping geohazards.

The characteristics of the addressed phenomena must be considered because of their impact on the detection from EGMS data but also on the way the events/phenomena are included in the inventory. For example, for landslides EGMS shall be very performant for large sized (e.g. several hm) very slow motions ( $\sim$ cm/year or less) whereas inventories can include in priority faster (and observable from ground-based observation) motions with potential consequences on the infrastructures. In addition, the threshold (on motion, e.g. on the velocity values) for detecting from EGMS data a large sized (several km) motion such as post-mining motion including a large number of measurement points inside will be different than a more localized phenomenon such as landslides that shall be characterized from few points. For these reasons, adaptation of the detection thresholds has been made to the geohazard type.

### 3.3.2 Detection capability assessment

This task is based on the extraction of the EGMS mean velocity (on level 2b and level 3) inside a polygon obtained from the inventory. In this perspective, a test is made to determine if the phenomena are positively identified.

#### Polygon data inventories

As in certain cases Multi-temporal Interferometry/Persistent Scatterers Interferometry results can underestimate the motion respect to inventory. That can occur when the actual displacement rates are too high. That can also occur when the Measurement Points are in sectors of the motion where the velocity is slower than provided by the inventory (that provides generally average velocity on a polygon or a maximum value). And finally, the inventory velocities can frequently be missing or estimated during a different period than EGMS. Two levels of positive detection are proposed to avoid rejecting potential positives whose velocity differs from the inventory data.

**i) Partial positive.** The detected motion measured inside the polygon is significantly different from its neighbourhood. A criterion of mean velocity difference detection between inside the polygon and outside is proposed (i.e. in a small square surface containing the polygon, excluding the MP inside the polygon) with values higher than 2 mm/yr. This threshold value is significantly higher than specified for EGMS data precision. During the project this threshold can be adjusted in accordance to the precision values obtained by the other VAs. In addition, and as mentioned previously, the detection threshold should be dependent on the kind of phenomena (and their

kinematics/geometry characteristics). During the project a table of empirical thresholds for the different phenomena included in the project will be maintained (jointly with VA4) to improve this activity by using more adapted values.

This approach can be also applied when velocity information is lacking in the inventory.

**ii) Full positive.** In addition to i), the motion must be consistent with the expected (or measured with ground instruments) inventory-derived motion. The information of velocity obtained from the inventory has been considered when available. The detection shall be identified as full positive if the difference with the average of the MPs' velocities is less than 2mm/year (value that can be modified as explained in i) ) or less than 30% of the inventory-derived velocity. This second threshold is applied mainly for faster motions for which a 2mm/year is not adapted (for example with a dm/year motion, a cm/year precision should be sufficient for a good detection).

#### **Point data inventories (case of excerpts of national databases)**

For each area of interest, EGMS data clustering has been performed using Classification of Active Deformation Areas (ADA) described in Appendix 1 to generate deformation polygons. For tests areas with few inventory polygons the EGMS polygons can alternatively be produced manually. The detection has been considered as partially positive if a point related to a phenomenon is inside an ADA derived polygon and negative if it is located outside the produced polygons. Metadata associated with the point information is important for interpretation of negative detection (e.g. when the phenomenon occurred; Size of the phenomenon; Slow motion or sudden and fast displacement), distance to the closest EGMS point must also be considered.

#### ***3.3.3 Mapping capability assessment***

In this task, the objective is to provide a simple indicator of the similarity of polygons derived for the inventories with deformation signatures visible in the EGMS timeseries. ADA polygons have been compared to the ones provided by inventories.

On polygons detected as positive (or at least partial positive) in task a) a computation of surface areas of the inventory polygon and of the EGMS derived polygons has been performed. A simple indicator for comparing both kind of polygons is proposed: the surface area of the intersection of the corresponding polygons (inventory and EGMS) divided by their average. The EGMS polygon has been considered as positive for mapping if this value is higher than 0.5.

#### ***3.3.4 Monitoring capability assessment***

This task aims at identifying events occurred at specific sub-periods of the period covered by the EGMS data. The focus has been on selected polygons identified as at least partial positive for detection (in sub-task a) and having information on events that can be temporally located (for instance, based on the inventories of damages and events). The EGMS Times Series on these polygons, have been visually examined to detect changes in the motion regime consistent with the known events. In this case the polygon has been considered as positive for Monitoring. Contrary to a) and b), this task requires interpretation of the data that cannot be automated.

The output of this task is the percentage of positive and false negative for each category of positives for the two Lines of Sight (level 2b data) and Horizontal-Vertical decomposition (level 3). To have a more detailed view of this validation, the results have been analysed taking into



account the theematics and expected velocity: the indicators have been defined (for level 2b and level 3) for each thematic and distributed in classes of velocities.

### **3.3.5 *Estimation of the percentage of mapped phenomena covered***

The objective of this additional task is to evaluate if the distribution of EGMS measurement points is sufficient to detect the on-going phenomenon.

Based on Solari et al. (2018) 5 points in the polygon are sufficient to potentially detect the phenomenon. However, our focus has been on the theematics resulting in motions of very different scales (from very local scales for engineering issues to phenomena covering several kilometres for mining activity). A simple threshold on the number of points inside the polygon without considering the size of the polygon could result as an insufficient characterization of the suitability of the points distribution. For this reason, it is proposed to compute the points' density (points/km<sup>2</sup>) inside the polygon as a secondary indicator.

The output of this task has been the percentage of polygons from the inventories having more than 5 EGMS and the average densities of measurement points in the polygons. The computation has been carried out for the two LOS data (level 2b) and the Horizontal-Vertical decomposition (Level, 3). The indicators have been analysed considering the theematics, the local relief (classes of slopes derived from the DEM), and land-cover classes. Those characteristics are required to understand the suitability of the EGMS results for each context/thematics.

## 3.4. Consistency check with ancillary geo-information

### 3.4.1 Introduction

This validation activity has assessed the consistency of EGMS results with geological, geomorphological, and geotechnical data based on the concept of "radar-interpretation" described in Farina et al. (2007). The approach consists of an integration of InSAR measurements along with other ancillary data (satellite images, aerial photos, topographic maps, etc.) to obtain an accurate analysis of the studied phenomenon. The ancillary data shall be used as support for the interpretation, e.g., in case of landslides, active faults or mining activity. There are not derived quantitative statistics or other metrics that are directly translated to IoA values.

The following general classes of ancillary geo-information have been used to validate EGMS products in this activity. For a detailed overview of the datasets used for the different validation sites, defined in the "Validation Data Collection" document:

- Geological, lithological, hydrogeological, geomorphological, and geotechnical maps.
- Corine Land Cover maps.
- Cadastral maps.
- Digital elevation models and derived slope data.
- Inventories of anthropic activities that can trigger surface motion, such as mining, quarrying, geothermal fluids extraction, oil & gas extraction, and water pumping.
- Web mapping services (Google Earth, ESRI World Imagery) for visual inspection and as map backgrounds.

### 3.4.2 Validation Approach

Depending on the validation site's characteristics and the ancillary datasets available, a selection of the following validation measures has been applied:

- Spatial overlap with spatial features in the ancillary geoinformation, e.g., geological units, topographic features, or spatial features in bedrock depth.
- Consistency of the amplitude of the ground motion signal with the geological asset, e.g., type of overburden or depth to bedrock.
- Consistency of the temporal evolution of the ground motion, e.g., compared to mining activity, oil/gas production or CO<sub>2</sub> injection.

To prevent too much thematic overlap with validation activities VA3 (Comparison with inventories of phenomena/events), comparison with inventoried data, is limited to cases, where also other auxiliary data is available. In this validation activity, inventory data was used to check for general agreement of EGMS data with the inventory data in general by means of visual inspection.

The consistency check relies on statistical values calculated for certain areas/units depending on the ancillary geoinformation, as well as visual inspection by an expert. It has to be emphasized that this validation activity shall provide a measure of plausibility of the EGMS data given the available ancillary geoinformation, and therefore the interpretation of the results by an expert is most important. Subsequently, IoA values are not directly calculated from statistical measures. Instead, the statistical measures are intended to help the expert in his interpretation of the data.

In addition to the statistical measures, the co-location/overlap of geographical / geomorphological (or similar) units with identified Active Deformation Areas (ADAs) has been checked. As for VA2 and VA3, the identification of ADAs has been carried out by an automated

procedure described in Appendix 1. The result of the ADA detection procedure is provided as polygons outlining the detected deformation areas.

### 3.4.3 Validation Procedures by Dataset Type and Phenomena

The validation approach is dependent on the specific validation area, i.e. the phenomenon that is observed within the area and the type of ancillary data that is available for the area. In the following the validation procedures for each type of phenomenon and dataset that are relevant for this validation activity are described.

#### ***Comparison with lithological/geological and geomorphological structures***

- *Objective:*
  - Validate consistency of ADAs found in EGMS data with geological/lithological/geomorphological situation in ancillary datasets.
- *Input:*
  - A range of geological maps, which come in various formats (raster/vector)
- *Approach:*
  - Create polygons (shapefiles) with delineated lithological (or geological) structures (manual work, no need for repetition in validation updates and not necessary if data is already in vectorial format)
  - Identify ADAs in validation area (automated procedure, see above).
  - Calculate mean velocity, std. dev and size for ADAs.
  - Calculate statistics for mean velocity within ADAs and overlap with ADAs per lithological /geological unit.
  - Check consistency of selected ADAs with geological structures by visual inspection.
- *Result:*
  - Measure of consistency of mean velocity with lithological region.
  - Measure of consistence of displacement patterns with lithological features, e.g. overburden thickness.

#### ***Comparison with numerical data, e.g., topography, slope, soil thickness***

- *Objective:*
  - Validate consistency of ADAs found in EGMS data with features in the numerical datasets and correlation with variations in EGMS velocities.
- *Input:*
  - Numerical dataset in raster or vector (points) format.
- *Approach:*
  - Identify ADAs in validation area (automated procedure, see above)
  - Qualitative comparison of variations in numerical dataset to localization of ADAs founds in EGMS data by visual inspection
  - Qualitative comparison of variations in numerical dataset to velocity variations in EGMS data by visual inspection.

- *Result:*
  - Overlay-maps with EGMS velocity data, detected ADAs and ancillary geoinformation

### **Tailings Dams / Waste Disposal sites**

- *Objective:*
  - Validate consistency of ADAs found in EGMS data with expected ground motion (subsidence or slope movements) connected to tailings dams and waste disposal sites.
- *Input:*
  - Shapefiles with delineated tailings dams and/or waste disposal site.
- *Approach:*
  - Create manual delineation of tailings dams / waste disposal site, where input data is not provided as shapes.
  - Identify ADAs in validation area (automated procedure, see above).
  - Qualitative comparison of localization of tailings dams and/or waste disposal site to ADAs founds in EGMS data by visual inspection.
- *Result:*
  - Overlay-maps with EGMS velocity data, detected ADAs and ancillary geoinformation.

### **Fault Lines**

- *Objective:*
  - Validate consistency of ADAs found in EGMS data with the presence of fault lines.
- *Input:*
  - Shapefiles with fault lines.
  - Information about surface motion regime from geological studies.
- *Approach:*
  - Identify ADAs in validation area (automated procedure, see above).
  - Qualitative comparison of localization of fault lines to ADAs founds in EGMS data by visual inspection. Focus is on correlation of ADAs boundaries with sections of fault lines.
- *Result:*
  - Overlay-maps with EGMS velocity data, detected ADAs and ancillary geoinformation.

### **Landslides / slope instability**

- *Objective:*
  - validate consistency of ADAs found in EGMS data with location of landslides / slope instabilities in ancillary datasets.

- *Input:*
  - Shapefiles with delineated landslide / slope instability areas.
- *Approach:*
  - Qualitative comparison of localization of landslides / slope instabilities to ADAs founds in EGMS data by visual inspection.
- *Result:*
  - Overlay-maps with EGMS velocity data, detected ADAs and landslide / slope instability areas.

### ***Oil and Gas extraction and piezometric data timeseries***

- *Objective:*
  - Validate consistency of EGMS time-series with documented oil/gas extraction activities and piezometric variations
- *Input:*
  - Time-series data.
- *Approach:*
  - Visual comparison of time-series for overlap period.
- *Result:*
  - Overlay plots for visual inspection.

### ***Comparison with Land cover maps / cadastral maps / Web mapping services (Google Maps, Bing Maps)***

- *Objective:*
  - Support interpretation of consistency with phenomena.
- *Input:*
  - Corine Land Cover maps, optical imagery.
- *Approach:*
  - Qualitative comparison of localization of map classes / map features to ADAs found in EGMS data by visual inspection
- *Result:*
  - Overlay-maps with EGMS velocity data, detected ADAs and ancillary geoinformation

#### ***3.4.4 Result Interpretation and Determination of IoA***

In the above section it is described how data is processed and what outputs have been produced to enable the expert to validate the consistency of the EGMS data with the ancillary geoinformation. Based on these, the expert provides a validation in terms of certain criteria called Indices of Agreement (IoA) for certain aspects of the validation.

To help the expert to come to a comprehensible and reproduceable decision, these criteria or IoAs are formulated as questions focusing on certain aspects of the validation, usually on the



comparison with a particular ancillary geoinformation dataset. The following list provides examples for questions formulating certain IoA criteria.

- Do the detected ADAs correlate as expected mostly with anthropogenic areas in the Corine Land Cover Map?
- Do ADAs significantly overlap with mine waste categories?
- Do the detected ADAs show a correlation with areas of higher depth to bedrock?

The selection of questions for the expert to answer is dependent on the dataset and the displacement phenomenon in question and is therefore unique for each validation site. The IoA is answered by the expert in textual form, describing the observations and drawing first conclusions. A final conclusion for a validation site is then drawn based on the observations made for the individual IoAs.

## References

- Farina, P., Casagli, N., Ferretti, A., 2008. *Radar-interpretation of InSAR measurements for landslide investigations in civil protection practices*. Proceedings of the 1st North American Landslide Conference. 272-283.
- Ester, M., Kröger, P., Sander, J., Xu, X., 1996. *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*. KDD. 96. 226-231.

## 3.5. Comparison with GNSS data

### 3.5.1 Introduction

The goal of this activity is to compare InSAR time series and velocities to corresponding values of GNSS stations. For the comparison between time series and velocities against GNSS data test statistics have been applied, to judge whether the differences are significant. Figure 5 sketches the expected geometry where InSAR observations are depicted as house and buildings. GNSS time series are sample at different time than InSAR and their stations are usually not located at the same place where InSAR observations are detected. This requires performing Spatio-temporal interpolation to make sure that comparisons are done in same (Spatio-temporal) sampling. In general terms, the workflow is the same for all data products L2a, L2b, L3. However, there are some differences. To compare L2a products to GNSS, double differences are calculated in space and time. This is not needed in L2b and L3 product because they are both spatially relative to the same reference system, i.e., ETRF 2000. In addition to that, GNSS time series have not been projected to LOS when compared to L3 product, because the latter already provide vertical and horizontal components. Furthermore, only those GNSS stations that are considered by the provider to be reliable are selected.

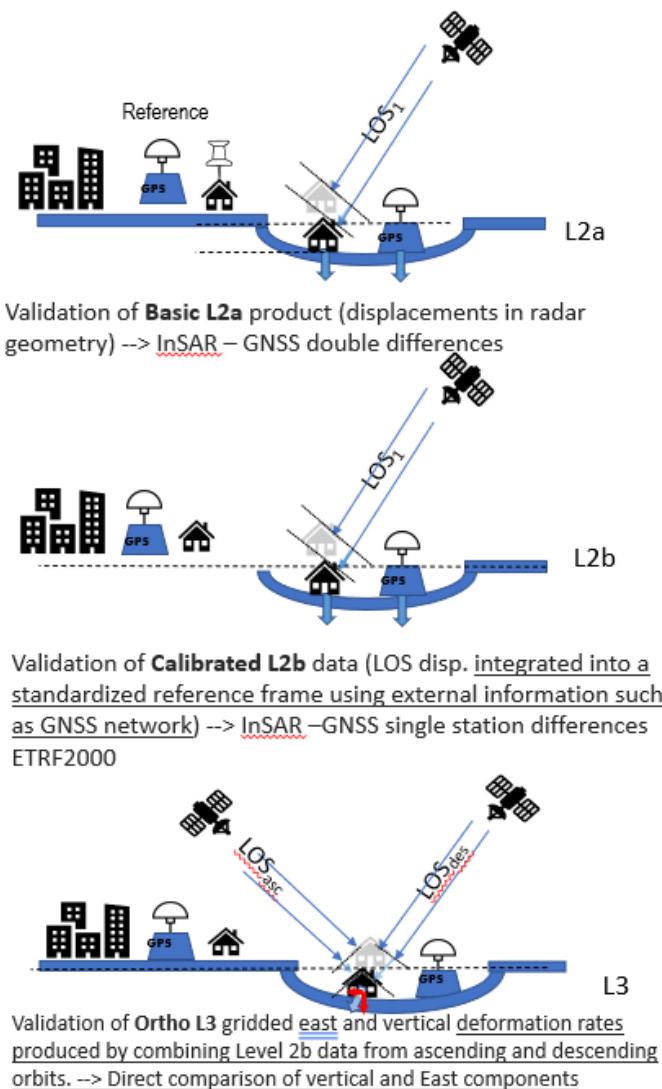


Figure 6: Validation of different product levels

### 3.5.2 Methodology

The workflow for the comparison between EGMS and GNSS data is summarized in the workflow shown in Figure 7 and described in the next sections:

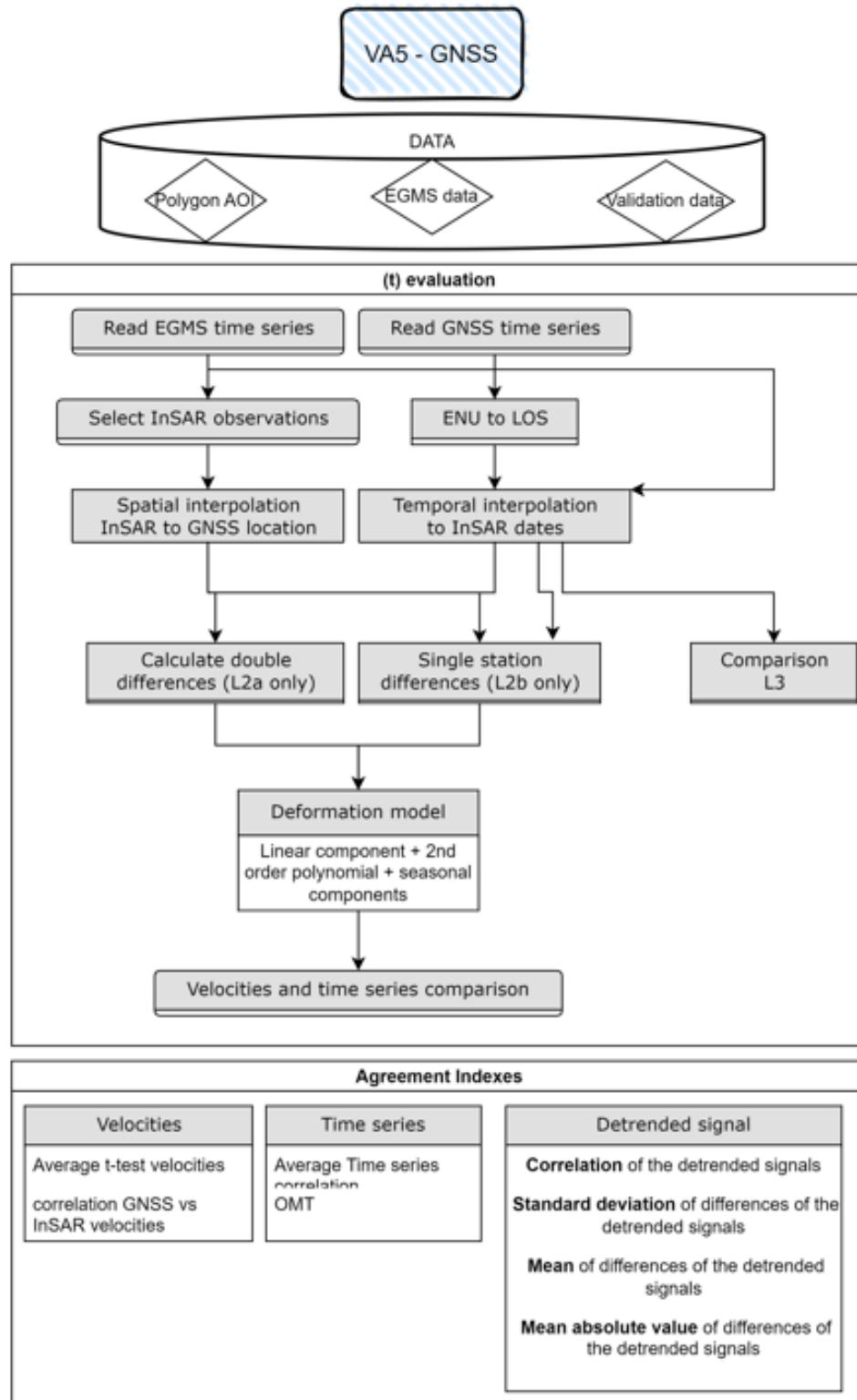


Figure 7: EGMS comparison with GNSS

The following steps have been followed:

1. **Temporal interpolation:** The comparison requires the interpolation of GNSS time series to InSAR acquisition dates, to make sure that the same time interval is covered. The interpolation is carried out using a 12-day window with weights the inverse of time difference.
2. **Time reference:** In both time series, GNSS and InSAR, the same reference date (also known as master date) has been used.
3. **Projection of GNSS time series to radar line-of-sight (LOS):** Given the observation geometry GNSS displacement to radar LOS for level 2a and 2b data is transformed. For level 3 data this conversion is not required.
4. **GNSS spatial referencing:** For level 2a, data has been selected on a GNSS station as reference station per thematic area. Then, the difference between the rest of the stations and the reference is calculated. For level 2b and 3 products, the velocity differences have been calculated between the GNSS reference frame and the reference frame used in the 2b and 3 products and subtracted from GNSS time series.
5. **InSAR MP selection:** The InSAR MP is selected to be compared with GNSS based on two criteria.
  - a) The first criterium is distance. MP must be at a distance less than selected limit from a GNSS, which assumes that the deformation signals do not change much over this distance. This maximum distance is expected to be in the order of few hundred meters, the exact value has been investigated, see Puggaard et al., 2022.
  - b) The second criterium is height w.r.t. ground. With it, an attempt has been made to select only those MP that reflect objects with similar foundations to the corresponding GNSS station.
6. **Spatial interpolation:** Using only the selected InSAR MP, their time series is interpolated spatially to the GNSS location. The interpolation error has also been estimated, considering the provided MP accuracy.
7. **Double differences:** For L2a products only. This is not needed in L2b and L3 product because they are both spatially relative to the same reference system, i.e., ETRF 2000.
8. **GNSS-InSAR comparison:** After the previous tasks are performed, the two data sets are compared in two ways:
  - a) Time series comparison. The correlation between GNSS and InSAR time series and standard deviation of the differences of time series is calculated.



- b) Comparison of deformation models. Using only the part of the time series where InSAR and GNSS overlap, a deformation model is estimated using BLUE (Best Linear Unbiased Estimation), which also provides the variance of the estimated parameters. Then, the same model for both GNSS and InSAR is estimated, and the estimated parameters are compared.

### 3.5.3 Computation of the IoA

The IoA are defined based on statistical analysis. Both classical statistical metrics such as standard deviation or RMS, and test statistics, are used. With test statistics and given a level of significance, differences between InSAR and GNSS are evaluated to determine if they are significant.

More specifically, the following IoA are considered:

1. An overall model test on the differences between InSAR and GNSS,
2. RMS of differences
3. Standard deviation of differences
4. Correlation between time series
5. Percent outliers of the differences
6. t-Test on velocities

The final IoA values need to be refined, and normalized withing the interval [0,1]. With this normalization, 0 means GNSS and InSAR are completely different and 1 very similar.

## 3.6. Comparison with in-situ monitoring data

### 3.6.1 Introduction

To evaluate agreement between insitu data and EGMS subsets, two types of measurements based on time have been considered:

1. Multiple consecutive campaigns (multiple dates regularly/irregularly taken).
2. Continuous monitoring systems (daily measurements).

On the other hand, the other two types are kept:

- A. Natural /ground monitoring systems.
- B. Infrastructure / buildings monitoring systems.

Criteria to discard validation sites:

1. Due to low measurement point density.
2. Due to the too quick deformation velocity on the in-situ data (which do not correspond to the precision / accuracy of the EGMS product).
3. Due to the too sporadic in-situ measurements (less than a yearly repetition measurement).
4. Due to a mismatch between the in-situ time series with the ones for the EGMS (too old in time).

### 3.6.2 Type of measurement available

The type of in-situ data available for the VA6 are mostly falling into the goal category A such as Levelling measurement (with campaigns taken every year), ground water monitoring (with piezometer campaigns taken every month) and GNSS or automatic total station (ATTS) systems (daily measurements). The first two types of multiple consecutive campaigns (levelling and piezometers) have been evaluated against the EGMS by converting the L2 of the interferometric product (ascending and descending) in up-down components.

Conversely, the measurements in X, Y and Z taken from the continuous monitoring systems (GNSS and ATTS) have been evaluated against the L2 of the EGMS once the 3-dimensional components have been converted in LoS (ascending and descending).

A type of measurement which fall into the goal category B is the automatic crack-meter daily measurements installed on two houses on an active landslide. In this case, since the measurement evaluate crack propagation and expansion (in the horizontal plane) the validation has been carried out by converting the L2 of the interferometric product (ascending and descending) in east-west components. For all the cases the EGMS L3 product (EW and UD) remain a viable alternative to be used for the VA6 activities in case the PS point distribution and location is close enough to the in-situ stations.

The workflow proposed to carry out the validation of the in-situ data against the EGMS service is divided in two main blocks:

- 1) Validation of the *velocity* of the phenomena (*Vel*).
- 2) Validation of the evolution in *time series* of the phenomena (*TS*).

### 3.6.3 Workflow for in-situ/EGMS (velocity)

A workflow is proposed for in-situ/EGMS aimed at evaluating their velocity. The series of scripts take the velocity and error quality flags from the EGMS and compares them with homologues flags of the in-situ measurements (validation 1 to 1). The Workflow has been programmed primarily in R scripting language and utilises R libraries, hosted on the GITLAB, prepared by Terrasigna, and executed programmatically via Jupiter Notebook instances. In figure 7 it is shown the different building blocks which compound the “VA6 velocity workflow”.

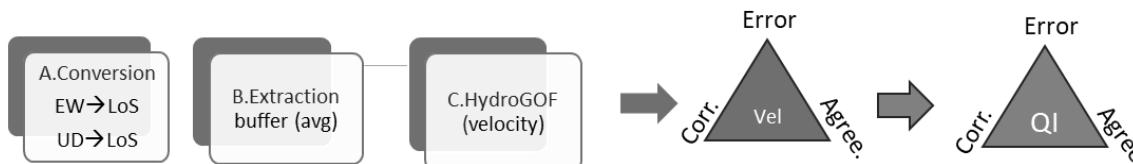


Figure 8: Workflow for in-situ/EGMS (velocity) – Corr. (Correlation), Agree. (Agreement)

Every single building block in detail as follows:

- Conversion of the ancillary data velocity in the same LoS of the EGMS (EW, UD, Ascending, Descending).

More in detail, the cosine directors already available in the EGMS data are used to carry out the X, Y, and Z conversion into the LoS once converted the coordinate systems of the in-situ data.

- Multiple buffer extraction of the information relative to the surrounding PS present near-by the in-situ stations.

The local coordinate system of the in-situ measurements have matched the EGMS and a multiple choice of buffer radius (20, 50, 100m) has been proposed based on the proximity of the in-situ data. The way the velocity field for the group of MPs land inside the buffer is aggregated is very flexible (average, median, max, min, standard deviation).

- Quantitative evaluation by plotting the velocity values of in-situ and EGMS data by means of the R Good of Fitness package (Hydro GOF).

This is a very special R library because of the ease offered to calculate up to 17 statistics of good of fitness of the observed (in-situ) against the simulated (EGMS) measurements. On the HydroGoF both statistical and graphical goodness-of-fit measures between observed and simulated values are implemented. The final output has allowed to compile “the triangle of the truth” which has three vertex each one dedicated to quantitative index (error, correlation and agreement).

- Translation of quantitative results in Quality Index (IoA)

In order to be consistent with the other VA tasks the quantitative index has been translated in IoA in accordance to the 5 hypothetical classes spanning between very strong and very weak performance.

### 3.6.4 Workflow for in-situ/EGMS (time series)

A workflow is proposed for in-situ/EGMS aimed at evaluating their time series. The series of scripts take the EGMS data falling inside the proximity buffer of each in-situ station and compares the time series associated to each PS with the in-situ time series (validation many to 1). The Workflow is implemented primarily in R scripting language and have utilised R libraries, hosted on the GITLAB, prepared by Terrasigna and executed programmatically via Jupiter

Notebook instances. In figure 8 it is shown the different building blocks which compound the “VA6 time series workflow”.

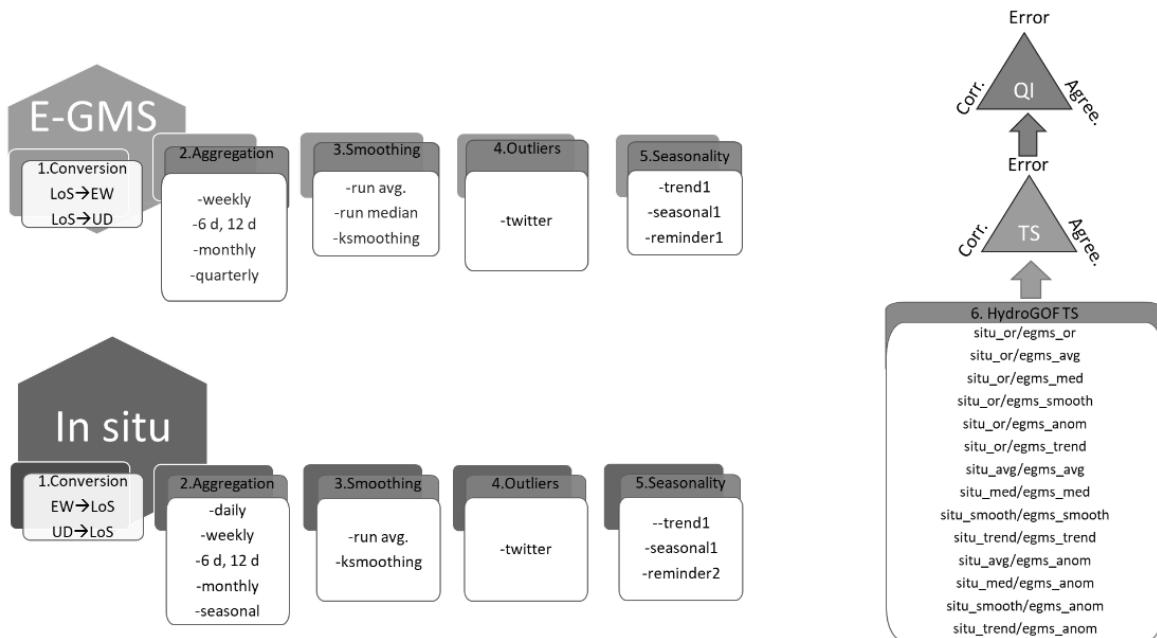


Figure 9: Workflow for in-situ/EGMS (time series)

The process to compare in situ and satellite time series is the following:

1. Conversion of the ancillary data TS in the same LoS of the EGMS (EW, UD, Ascending, Descending). In case of dealing with piezometers or levelling data, the opposite conversion is made: transformation of the LoS of EGMS in the same direction of the in-situ data (EW and UD).

Here only the InSAR timeseries that land inside the chosen buffer around the in-situ measurements are considered and the conversion based on the type of data is applied.

2. Temporal aggregation: 6 days, 12 days, monthly, quarterly, yearly aggregation applied to both TS (in situ and EGMS).

In this step, the temporal aggregation have been performed by considering the type of data; levelling consists mostly in yearly measurements, piezometers measurements are taken monthly whereas GNSS and ATTS are daily. For the latter a broader choice of aggregation is foreseen original (6 days), 12 days, monthly, quarterly, and yearly.

3. Smoothing: running mean, running average and k-smooth applied to the TS to smooth both time series

In this step, a series of smoothing algorithms have been applied in case the TS appear to be noisy.

4. Outliers' detection: filtering of outliers from both TS by means of the *anomalyze* method.

Again, in this step, in case some PS or the in-situ present some outliers on the TS the anomaly detection algorithm has identified data point in the TS that deviate from the dataset's usual behaviour. In case of our case studies those anomalies can indicate a change in the



measurement system or in case of the EGMS impact of weather on radar measurement or incorrect phase unwrapping.

5. Seasonality extraction: use of the Seasonal decomposition of Time series by Loess (STL) to extract the seasonality, the trend and to seasonal adjust applied to both TS.

The Seasonal decomposition of Time series by Loess (STL) is capable to separate the trend and seasonal component from the original TS. For this step only one variable is necessary to be changed based on the type of temporal aggregation of the data in analysis.

6. Quantitatively evaluation of (in-situ vs. EGMS) TS by means of the R Good of Fitness package (HydroGOF).

The same step described in step d) has been performed, however here several combinations of data correlation have been tested. As shown in figure 8 starting with the evaluation of the GOF for the original data against each other up to the trend seasonality for both type (in-situ vs. EGMS) has been tested. The idea behind this concept is that the simplification of the time series operated in both sense (for in-situ and EGMS) in several way could lead to a better final validation. More in detail, as follow the overview of the quality indicators subdivided in three groups, which has been calculated in step d) for the velocity and step 6) for the TS, is shown:

1. Error estimation indicators:

- i. MAE: Mean absolute error
- ii. RMSE: Root Mean Square Error

2. Correlation indicators

- i. R2: Coefficient of determination

3. Agreement indicators

- i. d: Index of Agreement

5. Translation of quantitative results to IoA.

Here the same concept explained for step e) has been applied to the TS in-situ validation results.

## 3.7. Evaluation XYZ and displacements with Corner Reflectors

### 3.7.1 Introduction

The goal of this validation activity is to validate the EGMS using the Level2a products concerning the **three validation requirements** (height, location, and displacements). Prior to validation, all the validation datasets have been converted whenever necessary to the same reference system as the EGMS. The proposed methodologies for each validation requirement are described as follows.

### 3.7.2 The estimated height of the MPs around the CR location

For this task, the CR with known heights are used or the ones derived by the levelling campaigns or GNSS if levelling is not performed. This has been our ‘ground truth’. Then the differences between the ‘ground truth’ heights (CR) and the MP’s estimated heights at the location of the CR are estimated. It is assumed that the differences between orthometric and geometric heights are negligible, given the small distances between CR (Marinkovic et al., 2008).

The procedure to achieve this goal has required the following steps:

1. Selection of the MPs heights at and around the CR locations using 100m buffer around the location of the CR. Ideally, only a single point at the CR location has been used if the CR is visible. The remaining MPs have then been used to perform statistics.
2. Calculate the relative height differences between the ‘ground truth’ locations (CR) and the corresponding MPs at and around the ‘ground truth’ locations (previously selected in step i). Then, the relative height differences between places with more than one CR is estimated. For locations with only one CR, this one is used as one of the ‘ground truth’ locations and other locations of known heights to estimate the relative height differences between ‘ground truth’ measurements. This is the case for one of our test sites at Calern multi-technical geodetic observatory in France (details in Table 5) where there is only one CR but multiple other precise geodetic measurements (DORIS, GNSS, precise levelling), which, given their infrastructure, it is expected to have MPs.
3. Estimate the corresponding accuracies after step (2).
4. Comparison between the estimated relative heights of the CR and the estimated relative heights of the MPs. Given the accuracies estimated in iii statistical testing is used (Fisher, 1925) to decide if the differences can be judged to be significant considering their corresponding standard deviations.

### 3.7.3 Geopositioning accuracy by XY offset estimation

For this requirement, the GNSS local observations have been selected given the high accuracy of GNSS in the horizontal component. The methodological procedure have included the following tasks:

1. The accurate position of the CR’s is known. From the selection of MPs performed in the task (1i), the distance (offset) is computed between the CR and the closest MP.



2. If CRs are not visible at the beginning of the validation activities, scatterers are used with known positions by the other measurements performed on the test sites.
3. Use of statistical tests to evaluate if the offset is significant given the required accuracy.

### **3.7.4 Quality of the EGMS time-series displacements**

To evaluate the quality of the EGMS time-series displacements, GNSS stations are used at the chosen sites. Therefore, part of the methodology for this validation requirement is described in VA5. The methodological procedure for this task has included the following steps:

1. Selection of the EGMS MPs located around each CR with a buffer of 100m assuming that the deformation within this buffer has been similar. We have also evaluated point statistics if more than three MPS's are selected. CRs MPs have been selected in different track modes oriented to the corresponding geometry.
2. Conversion of the GNSS to LOS (see VA5 task 3) is also performed in VA5.
3. Select one point where measurements have been performed as a reference.
4. Estimate the double differences between the points used for validation and the selected reference point. Perform the same for the MPs selected in (I)
5. Comparison between both time series following the same procedure described in VA5 task 9.

## Appendix 1: ADA common approach

A central aspect of validation activities is the identification of Active Deformation Areas (ADAs) for which each VA carried out their specific comparison to different kinds of validation data. Two different methods are applied, depending on the nature of the observed ground motion and its spatial distribution, as well as on the point density of the areas in question.

**Approach 1:** The first procedure for identification of ADAs is based on the DBSCAN algorithm by Ester et al. (1996), which is a well-known approach for identifying clusters in point data. Depending on the observed deformation phenomena, points below a certain threshold are first filtered out depending on whether detection of uplift or subsidence is selected.

In this filtered dataset, the DBSCAN algorithm finds clusters of points, while isolated points are discarded as they are being considered as noise. The two key parameters to the DBSCAN algorithm are epsilon and min-points (or minp). Epsilon (given in meters as eps\_m, see below) is the radius of the circle to be created around each data point to check the point density and minp is the minimum number of data points required inside that circle for this data point to be classified as a "core point". The parameters are used to divide points into "core points", "border points" and "noise points". If minp points are within a radius of eps\_m, the respective point is considered a core point. If the number is less, it is considered a "border point", and if there is none, it is considered a noise point. Like this, eps\_m and minp together control whether points are connected ("density-connected") and, thus, whether they can be considered to be part of the same cluster (see Ester et al., (1996) for more details).

After all point clusters have been identified, a concave hull (or alpha shape) is being calculated, which represents the envelope of the respective cluster. The "alpha" parameter can be used to "loosen" or "tighten" the fit to the points. An alpha parameter of 0 corresponds to the convex hull, which envelopes all points with the least possible level of detail. The alpha shape procedure applied for the validation has been adapted from Dwyer (2014; <https://gist.github.com/dwyerk/10561690>). Finally, each resulting cluster polygon is extended with a buffer zone (in meters) defined by the "buffer" parameter.

The resulting clusters are then further filtered. The "min\_cluster\_size" parameter defines the minimum allowed number of points within a cluster. Clusters below this threshold have been discarded. Similar, the "min\_cluster\_vel" parameter defines the threshold for the overall cluster-velocity, which is defined as the velocity for which a certain percentage of the points, called cluster\_vel\_quantile (e.g., 90%), are below this velocity and (100-cluster\_vel\_quantile)% of the points have a higher velocity. Compared to the simple mean velocity of the cluster, this provides a more stable measure of the "dominating" velocity of the cluster and reduces the effect of a large number of points just above the point velocity threshold, especially in the added buffer zone, and also prevents unwanted effects by outlier points with extremely high velocities. Clusters with a cluster-velocity below "min\_cluster\_vel" are also discarded.

Figure A1 shows example for the ADA detection with intermediate results for all processing steps:

- a) Original InSAR points with velocities.
- b) InSAR points filtered by velocity.
- c) Result of the DBSCAN algorithm on the filtered points.
- d) Final cluster envelopes determined by the alpha shape algorithm and subsequently extended by a buffer zone.

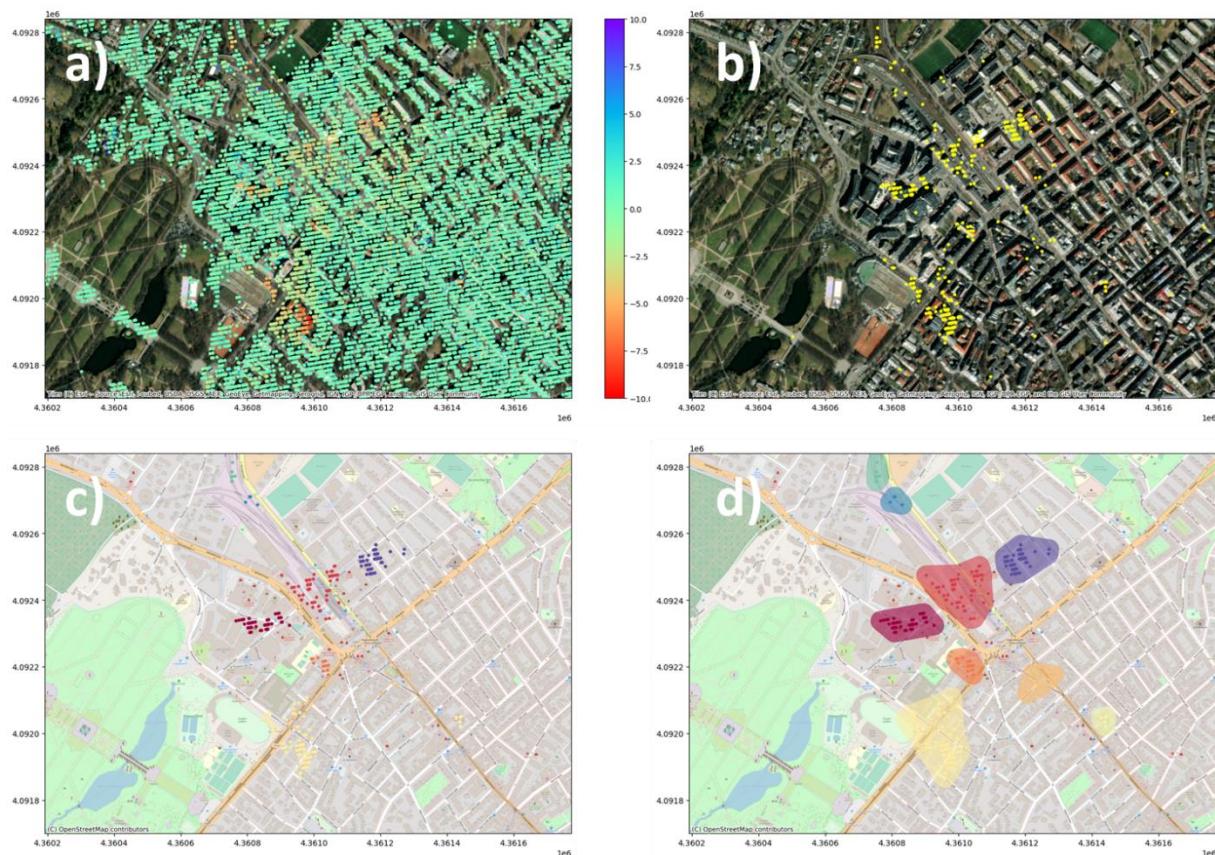


Figure A1: Example for the ADA detection and its processing steps.

A weakness of the above approach is that it ignores all points that are regarded as stable by eliminating them before applying the cluster detection. With this, no distinction is possible whether two groups of moving points are separated by an area of stable points or just by an area of low coherence, i.e., an area with no points at all. Figure shows an example of a large active deformation area in Germany.

The DBSCAN method produces a large number of clusters, where in fact one large deformation phenomenon is covering a large area with only patches of coherent measurement points.

**Note:** Specifically for the inventories of phenomena comparison, other tools like *ADAFinder* from *ADATools* package (Montserrat et al. 2022) were considered and tested during the development stages of the validation activity with similar approaches as the ones described above.

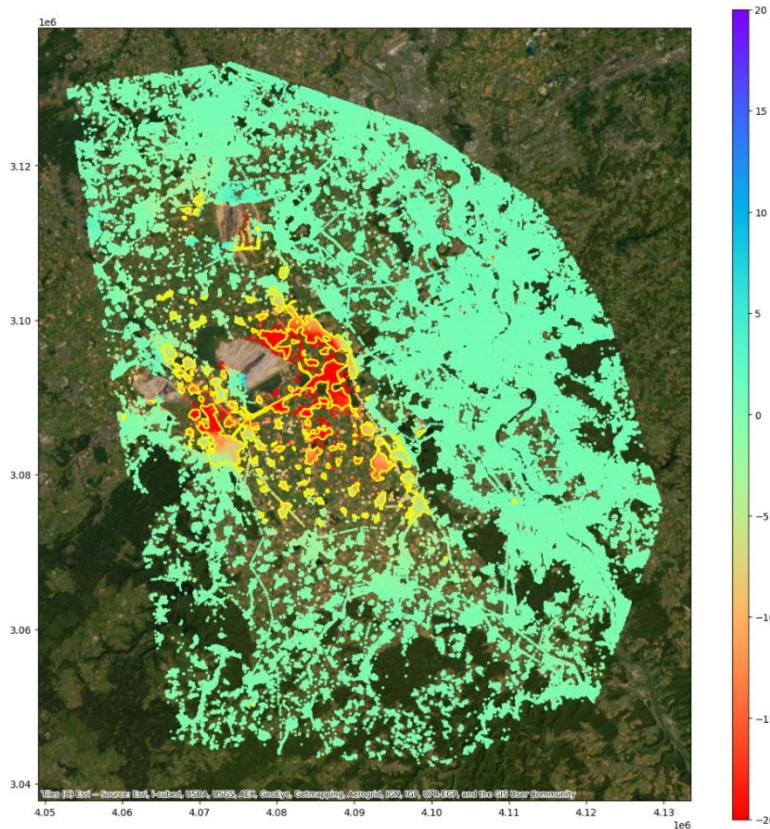


Figure A2: ADA detection for a mining area in Germany, using the DBSCAN method.

**Approach 2:** In order to account for the fact that these patches are separated by areas with no points instead of stable areas, a region-growing method is applied. For this, a Delaunay triangulation is first applied of the original dataset, including all stable points. This allows us to efficiently utilise information about point-neighbours. Then, all triangles that exclusively have stable corner points are removed, which means that those triangles represent a stable area. After these areas have been removed, region growing is applied by selecting a random (seed-) triangle and add its points to the current cluster. In the next iteration the same procedure is done for all neighbours of the original triangle. This is continued iteratively, each iteration adding the neighbours of the triangles added in the previous iteration, until no more neighbour-triangles are left, and the cluster is complete. The cluster triangles are removed from the original set of triangles and a new seed-triangle is selected to find the next cluster in the same way. The procedure is repeated until the original set of triangles is empty. Because the clusters' outlines can directly be determined from the sets of triangles, there is no need for applying the alpha shape method.

As for the DBSCAN method, each resulting cluster polygon is extended with a buffer zone and clusters below a certain minimum allowed number of points and a minimum allowed mean velocity are discarded, as described above. Also like for the DBSCAN method, the whole procedure is applied separately to detect downward and upward movements.

Figure A3 shows the same dataset as in Figure A2 but with ADAs detected by the region-growing approach. Instead of many small individual clusters, the main deformation area is now covered by one single cluster, which reflects the situation in the area much better. This, however, comes at slightly increased computational costs and, as Figure A3 shows, can result in slightly irregular or spiky envelopes. Therefore, for smaller ADAs, e.g., for landslides or urban subsidence, the DBSCAN method is preferred, where the level of detail of the envelopes can be adjusted using the "alpha" parameter.

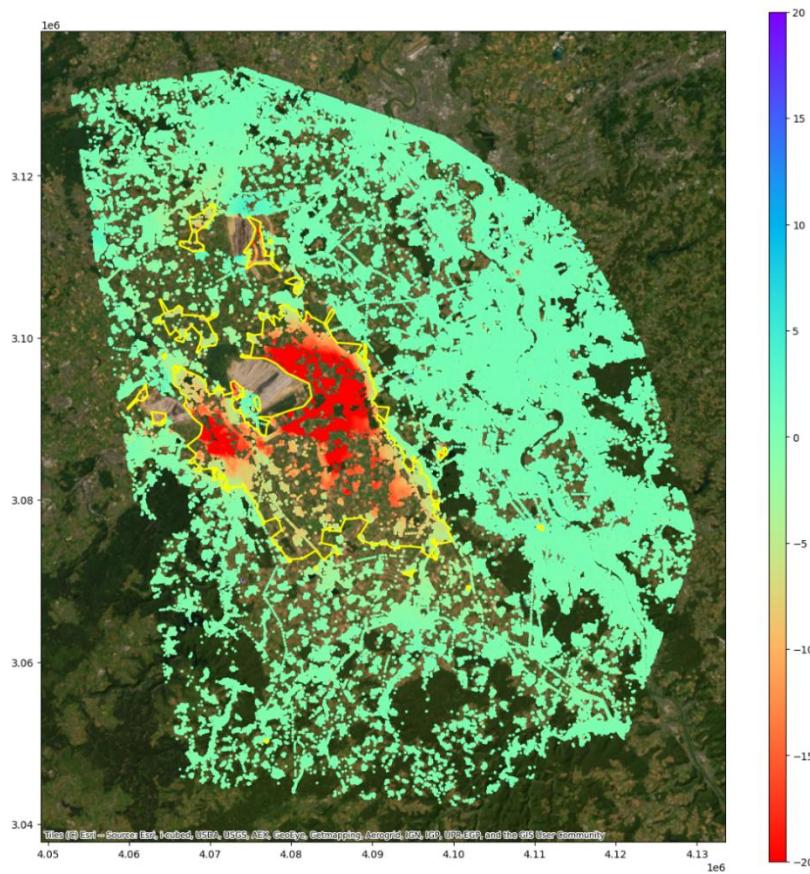


Figure A3: ADA detection for a mining area in Germany, using the region-growing method.

Below, all parameters for the ADA detection are listed, with short explanations of their meaning/role. Parameters starting with "dbSCAN" or "regiongrow" are only relevant for the respective method. Table shows preliminary parameters for different types of deformation phenomena. List of Parameters:

- **v\_min**: Velocity limit for cluster detection in either upward or downward direction.
- **dbSCAN\_eps\_m**: The maximum distance between two samples to be considered as in the neighbourhood of the other.
- **dbSCAN\_minp**: The number of samples in a neighbourhood for a point to be considered as a core point (see Ester et al., 1996).
- **dbSCAN\_alpha**: Controls the "level-of-detail" or roundness of the concave-hull drawn around each point-cluster.
- **regiongrow\_min\_cluster\_dist**: Minimum distance for cluster in region growing algorithm. Point groups with a distance larger than this, even if only separated by an area without points, are considered different clusters
- **min\_cluster\_size**: Minimum points allowed to be in a cluster.
- **min\_cluster\_vel**: Minimum allowed cluster-velocity (see above).
- **cluster\_vel\_quantile**: fraction of points with velocities below the cluster-velocity.
- **buffer**: Size of the buffer in meter, which is added to the resulting ADA polygon.



Table A1: Deformation phenomena and their default ADA detection parameters. Actual parameters used for particular sites may differ from this table to take into account the specific site's character and displacement phenomenon.

| Parameter                        | Urban subsidence | Landslides | Mining / Water or Gas Extraction |
|----------------------------------|------------------|------------|----------------------------------|
| v_min [mm/year]                  | 4                | 4          | 4                                |
| dbscan_eps_m [m]                 | 150              | 150        | 150                              |
| dbscan_minp [#]                  | 5                | 5          | 5                                |
| dbscan_alpha [no unit]           | 0,005            | 0,005      | 0,005                            |
| regiongrow_min_cluster_dist [m]: | 500              | 500        | 2000                             |
| min_cluster_size [#]             | 20               | 50         | 100                              |
| min_cluster_vel [mm/year]        | 5                | 5          | 5                                |
| cluster_vel_quantile [fraction]  | 0.95             | 0.95       | 0.95                             |
| buffer [m]                       | 30               | 30         | 30                               |

## References

Ester, M., Kröger, P., Sander, J., Xu, X., 1996. *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*. KDD. 96. 226-231.

Barra, A., Solari, L., Béjar-Pizarro, M., Monserrat, O., Bianchini, S., Herrera, G., ... & Moretti, S., 2017. A methodology to detect and update active deformation areas based on sentinel-1 SAR images. *Remote sensing*, 9(10), 1002.

Navarro, José & Cuevas, María & Tomás, Roberto & Barra, Anna & Crosetto, M.. (2019). A toolset to detect and classify Active Deformation Areas using interferometric SAR data.

Monserrat, O. & Barra, Anna & Reyes-Carmona, Cristina & Cuevas, María & Bejar, Marta & Navarro, José & Tomás, Roberto & Galve, Jorge & Solari, Lorenzo & Sarro, Roberto & Mateos, Rosa & Azañón, J. & Herrera, Gerardo & Crippa, B.. (2022). ADATools: a set of tools for the analysis of terrain movement maps obtained with SAR Interferometry. 10.4995/JISDM2022.2022.13865.