

## VALIDATING THE EUROPEAN GROUND MOTION SERVICE: AN ASSESSMENT OF MEASUREMENT POINT DENSITY

A. Vradi<sup>1\*</sup>, J. Sala<sup>1</sup>, L. Solari<sup>2</sup>, J. Balasis-Levinsen<sup>2</sup>

<sup>1</sup> Sixense Iberia, Barcelona, Spain - (amalia.vradi, joan.sala)@sixense-group.com

<sup>2</sup> EEA (European Environment Agency), Copenhagen, Denmark - (Lorenzo.Solari, Joanna.Balasis)@eea.europa.eu

**KEY WORDS:** EGMS, SAR interferometry, InSAR, Geospatial Data Validation, Point density

### ABSTRACT:

The European Ground Motion Service (EGMS) is an operational data service that uses interferometric analysis of Sentinel-1 radar images to provide high-resolution monitoring of ground deformation. The service is the first continental ground motion public dataset that is open and available for various applications and studies. This paper presents the validation workflow of the EGMS product in terms of spatial coverage and density of measurement points. The study employs open land cover data to evaluate the density per class and proposes statistical parameters associated with the data processing and timeseries estimation to ensure consistency across different selected sites.

The validation process involves twelve selected sites across Europe, representing the four processing entities, covering both rural and urban areas. The paper highlights the importance of ensuring completeness and consistency of the EGMS product for its effective use, along with pointwise quality measures critical in assessing the quality of the EGMS point timeseries results. Additionally, the paper discusses the challenges faced during the pre-processing of the data from EGMS and presents a custom algorithm designed to eliminate biases in the data. As an open and freely available dataset, the EGMS will provide valuable resources for further research and applications, enabling better understanding and management of geohazards, environmental monitoring, and urban planning.

### 1. INTRODUCTION

The European Ground Motion Service (EGMS) constitutes the first application of high-resolution monitoring of ground deformation for the Copernicus Participating States (Costantini et al., 2021). It provides valuable information on geohazards and human-induced deformation thanks to the interferometric analysis of Sentinel-1 radar images. This operational service constitutes the first European ground motion public dataset, open and available for various applications and studies.

The work presented here forms part of the activities taking place to validate the output of the EGMS service. The completeness and consistency of the products represent the usability of the service. Seven reproducible validation activities (VA) have been developed to evaluate the fitness of the EGMS ground motion data service. These activities collect validation data from various sources across 12 European countries (as shown in Figure 1)

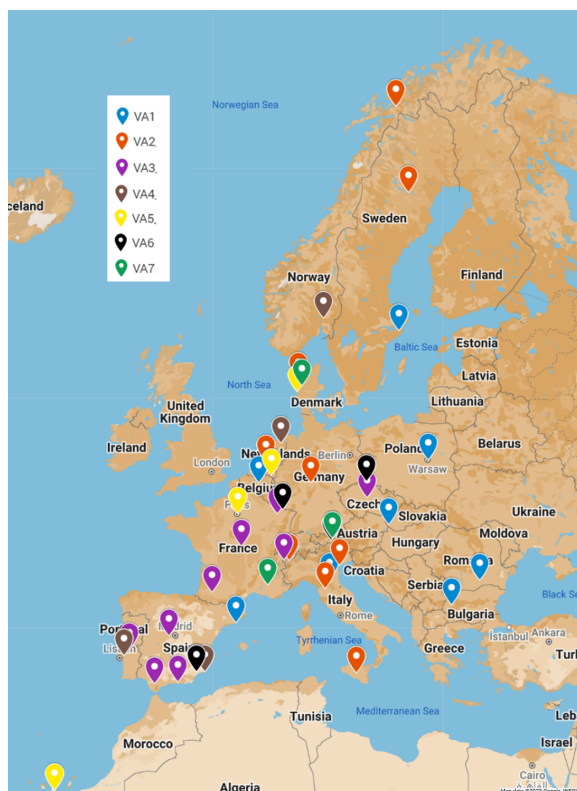
- **VA1 – Point density check performed by Sixense.** This activity evaluates the point density consistency across the different land cover classes defined in CLC Urban Atlas 2018 (high resolution land cover layer).
- **VA2 – Comparison with other ground motion services carried out by the Norwegian geotechnical institute (NGI).** This activity checks the performance of the continental ground motion service against the quality controlled and validated regional initiatives.
- **VA3 – Comparison with inventories of phenomena/events performed by the French geological survey**

(BRGM). This activity compares 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).

- **VA4 – Consistency check with ancillary geo-information carried out by the Norwegian geotechnical Institute (NGI).** This task makes use of national inventories of geomorphological, geotechnical and geological data together with expert judgement and automated procedures to discover active deformation areas on the EGMS timeseries datasets.
- **VA5 – Comparison with GNSS data performed by the Netherlands Organisation for Applied Scientific Research (TNO).** The goal of this activity is to perform a validation of the geocoding of the EGMS products together with ground motion timeseries comparison of GNSS measurements.
- **VA6 – Comparison with insitu monitoring data performed by the geological survey of Austria (GBA).** The objective of this task is to evaluate the insitu measurements coming from GPS campaigns, levelling data, extensometers, piezometers, inclinometers, geodetic monitoring, and tilt meters against the EGMS ground motion data.
- **VA7 – Evaluation XYZ and displacements with Corner Reflectors performed by the Dutch geological survey (TNO).** This activity aims to evaluate the precision of the EGMS timeseries (location, height and observed motion).

The objective of this paper is to present a workflow to validate the spatial coverage and density of measurement points (MP)

\* Corresponding author



**Figure 1.** Validation Activities (VA) and their validation sites (approximative) (from Google Maps)

in the EGMS product portfolio. A total of twelve sites have been selected for this activity, covering various areas of Europe. Additionally, these sites have been selecting so as to equally represent the four different processing entities of EGMS, as we explain in more detail in Section 2.2.

To measure the quality of the point density we employ open land cover data from the Copernicus Land Monitoring Service (CLMS). This allows us to evaluate the density of points per land cover class. For each measurement point there are associated quality parameters (Temporal Coherence, RMSE and Amplitude Dispersion). We perform statistical analysis to ensure they are consistent across the twelve different selected sites.

The EGMS timeseries are structured following the Sentinel-1 Interferometric Wide (IW) swath mode of the SLC source products. This creates a situation where there is a variable burst overlap depending on the latitude. To address this issue, a custom algorithm was designed to identify and extract the unique, non-overlapping polygon for each burst. This iterative algorithm was specifically designed to ensure a fair comparison among different areas and to guarantee a fair comparison between the 4 EGMS processing entities.

The availability of a high-resolution ground motion public dataset, such as EGMS, provides valuable resources for further research and applications, enabling better understanding and management of geohazards, environmental monitoring, and urban planning among other activities. The validation of this dataset also provides insights and a reproducible workflow that can be used for future validation activities of the EGMS product, as well as for other similar ground motion datasets.

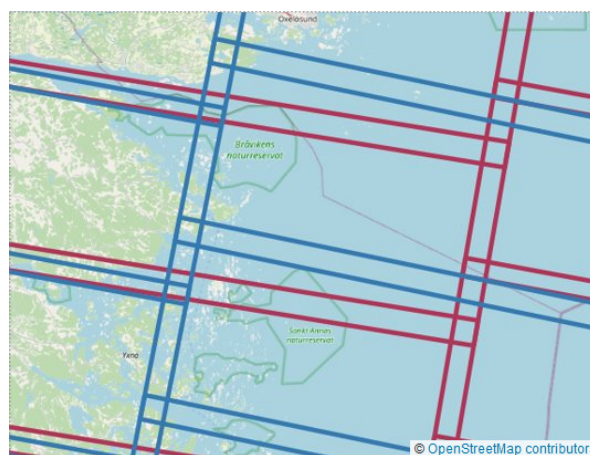
## 2. METHODS

### 2.1 EGMS Data Preprocessing

The EGMS data were produced by four different processing entities and their product specifications can be found here (European Environment Agency, 2022). The resulting products include three levels that are explained below in detail. In this work, all three levels of EGMS products were included in the validation (Solari et al., 2023):

- **Basic (L2a):** Line of sight velocity maps in ascending and descending orbits with annotated geolocalisation and quality measures per measurement point. Basic products are referred to a local reference point.
- **Calibrated (L2b):** Line of sight velocity maps in ascending and descending orbits referenced to a model derived from global navigation satellite systems time-series data. Calibrated products are absolute, being no longer relative to a local reference point.
- **Ortho (L3):** Components of motion (horizontal and vertical) anchored to the reference geodetic model. Ortho products are resampled to a 100 m grid.

One of the major challenges encountered during the preprocessing of the EGMS data was the presence of overlapping bursts from different Sentinel-1 satellite tracks for the full resolution products L2a and L2b. This issue was particularly prevalent in areas with high track overlaps, leading to a higher point density and potential biases in the analysis (for example, in Figure 3, the urban area of Stockholm is covered by multiple tracks).

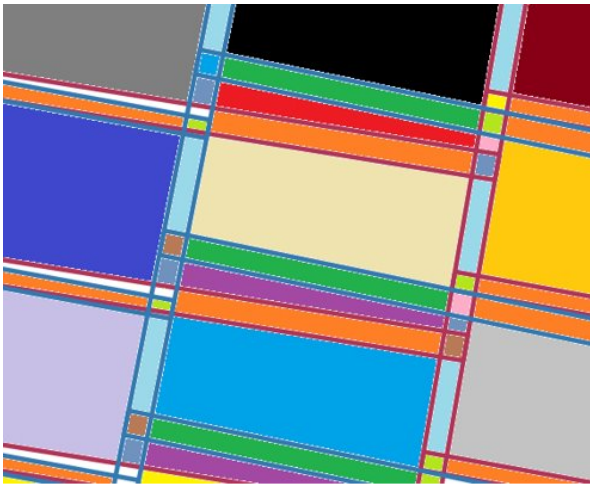


**Figure 2.** A collection of overlapping bursts/polygons covering the area of Stockholm (based map used from OpenStreetMap)

Intersection of polygons was used to determine the overlapping polygons. After that, a collection of unique non-overlapping polygons was created. All operations were performed with GeoPandas/Python (see Figure 3).

### 2.2 Data Used

A compilation of EGMS timeseries over 12 European sites was created for the point density validation activity. The sites were chosen to equally represent the four processing entities (see



**Figure 3.** Unique, non-overlapping polygon for each burst. Each polygon is depicted in different color to highlight that no overlap occurs.

section 2.1), with each entity processing three of the sites in the dataset. The validation activity aims to ensure the consistency across EU territories by comparing the point density at three sites for each algorithm. One site is located in a rural mountainous area, while the other two are urban. The dataset used in this study was obtained directly from the open data service Copernicus Land – Urban Atlas 2018 (Atlas, 2018), which contains validated Urban Atlas data with different land cover classes polygons, along with metadata and quality information.

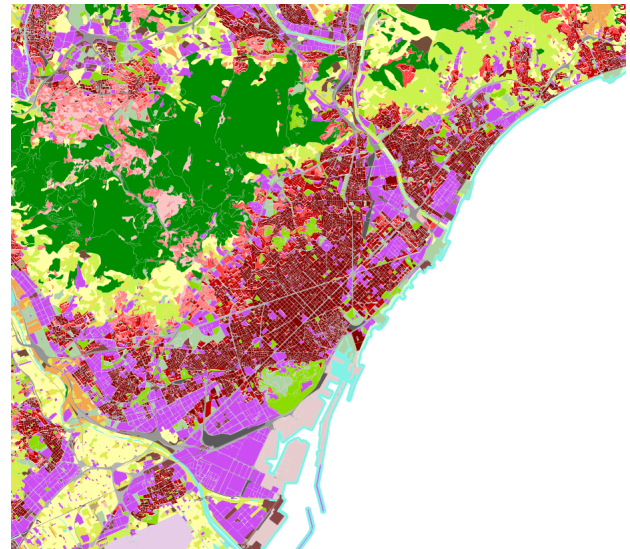
We have extensive Urban Atlas verified datasets on the cities of Barcelona, Bucharest, Bologna, Sofia, Stockholm, Warsaw, Brussels, Bratislava. In parallel we select four different rural and mountainous areas to analyse more challenging scenarios as well for the four processing chains of the providers. A summary of all sites used for the validation of point density can be found in Table 1.

There are 27 different land cover classes defined in Urban Atlas. To facilitate the analysis and the interpretation of the results, we aggregated and presented our findings for each of the main CLC groups: Artificial Surfaces, Forest and seminatural areas, Agricultural areas, and Water bodies. The classes names, and their corresponding code, of each category can be found in the Appendix, in Tables 2 and 3.

### 2.3 Usability Criteria

The usability criteria to be evaluated concern the completeness of the product, its consistency, and the associated quality parameters (temporal coherence, RMSE and amplitude dispersion).

- Ensuring the completeness and consistency of the EGMS product is essential to its effective use. To achieve completeness, it is important to ensure that the data gaps and density measurements are consistent with the land cover classes.
- Consistency is also vital for point density across the same land cover class for different regions. For instance, urban classes will have higher density than farming grounds, and this density should be consistent between the ascending and descending products.



**Figure 4.** The city of Barcelona segmented by the different land cover classes. A detailed legend of the class name and color can be found in the Appendix, Figure 10.

- Associated quality parameters are critical in assessing the quality of the EGMS PSI results. For example, the temporal coherence is expected to be higher in urban classes, and the root-mean-square error should be lower.

### 2.4 Performance Indicators

For the validation measures, key performance indices (KPI) are calculated, with values between 0 and 1. We normalise the estimated density values for each service provider with respect to the highest value for the grouped classes of Artificial surfaces, Agricultural areas and Forest and seminatural areas. Users expect consistent and good densities in these classes, specifically in the Artificial surfaces. On the other hand, for the group of Water classes we expect to see the lowest value for point density. Therefore, we employ the values of the group of Water as a metric for outlier detection, since the applied algorithms should barely produce any measurement points on these surfaces.

Summary of Point Density Validation Sites		
Area Name	Country	Type
Bologna	Italy	urban
Bucharest	Romania	urban
Stockholm	Sweden	urban
Barcelona	Spain	urban
Sofia	Bulgaria	urban
Warsaw	Poland	urban
Bratislava	Slovakia	urban
Brussels	Belgium	urban
Bolzano	Italy	rural/mointaneous
Las Palmas	Spain	rural/mointaneous
Zilina	Slovakia	rural/mointaneous
Tromso	Norway	rural/mointaneous

**Table 1.** Summary of Validation sites and their corresponding type (urban or rural/mountaneous)

## 2.5 Software Tools and Libraries Used

The development environment for this study was based on JupyterHub, which provides a platform for developing and running Jupyter notebooks (Kluyver et al., 2016). This allowed us to use popular scientific tool development environments such as GeoPandas, a library for working with geospatial data in Python, as well as Pandas and NumPy for data analysis.

These tools enabled us to perform data analysis and processing, including spatial data management and manipulation, as well as interfacing with the Sentinel-1 data. The use of open-source tools and platforms such as JupyterHub and GeoPandas (Jordahl et al., 2020), has allowed us to create a reproducible and transparent workflow for the validation of the European Ground Motion Service ground motion timeseries.

The validation environment was designed with system architecture considerations in mind. To facilitate reproducibility, we developed and deployed it on top of the Kubernetes container orchestration engine, leveraging state-of-the-art cloud-native technologies such as MinIO and Keycloak. This approach enables the system components to be easily transferred between cloud providers, making it compatible with EU cloud initiatives.

## 3. VALIDATION WORKFLOW

The point density validation workflow consists of several steps aimed at collecting, analyzing, and comparing the EGMS and Urban Atlas data. The first step of the workflow is to collect the necessary data for the analysis. A graphical overview of the steps is depicted in Figure 5. This involves obtaining the Urban Atlas<sup>1</sup> data for the selected areas of interest and downloading the corresponding data from EGMS<sup>2</sup>. The Urban Atlas data provides the ground truth against which the EGMS data will be compared.

In continuation, we perform preprocessing on both the EGMS and Urban Atlas data. On the EGMS side, we apply the custom algorithm for identifying and extracting the unique, non-overlapping polygons for each burst (described in Section 2.1) followed by clipping to the area of interest. On the Urban Atlas side, we calculate initial statistics and summaries, such as the distribution of land cover classes, for each area of interest. An example of the figures of class distribution can be found in Figure 9. These preprocessing steps are necessary to ensure that the data are properly prepared for subsequent analysis.

As a next step, we proceed to calculate the density of points for each land cover class in the areas of interest. Using aggregated groups of classes, we generate figures and graphs that can assist us in drawing conclusions. As a means of identifying outliers, we focus on the "Water" group of classes, where a low density of points is expected. In addition to the density of points, we also calculate statistics such as probability distributions for temporal coherence (as for example, in Figure 8). These steps allow us to evaluate the consistency of point density across different land cover classes and time periods, and to identify any significant deviations from the expected behavior.

With the point density and other statistics calculated, we move onto the final step of the workflow, where we produce Key

<sup>1</sup> <https://land.copernicus.eu/local/urban-atlas/urban-atlas-2018>

<sup>2</sup> <https://egms.land.copernicus.eu/>

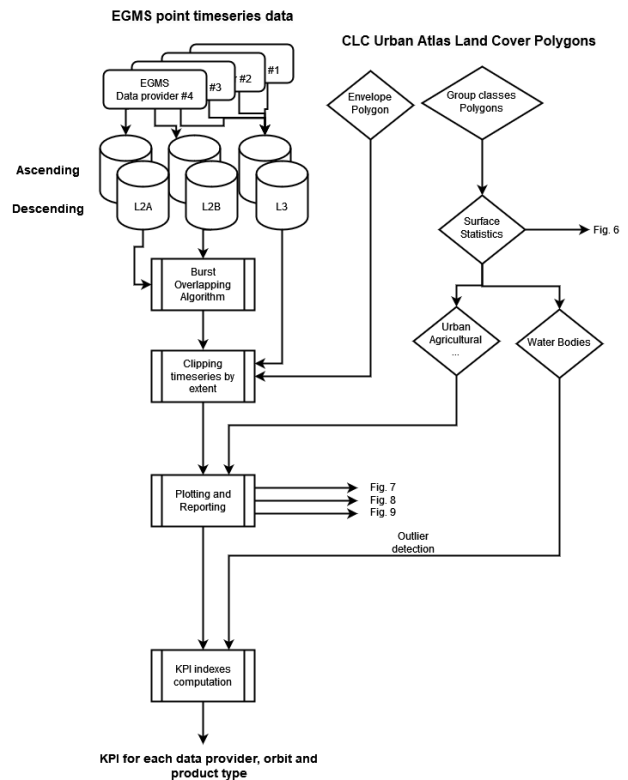


Figure 5. Point Density Validation Workflow

Performance Indicators (KPIs) for each land cover class, processing entity, and orbit (ascending or descending) for all three EGMS products. These KPIs provide a standardized way to compare the performance of each processing entity across different land cover classes and orbits, and can be used to identify potential issues or areas for improvement. We also generate summary reports to visualize and present the results of the validation activity in a clear and concise manner. Overall, this workflow allows for a systematic and reproducible approach to validating the point density of the EGMS service. The following figures summarize the results obtained for one of the validation sites (The class name that correspond to each color of the legend can be found in the Appendix, Figure 10.):

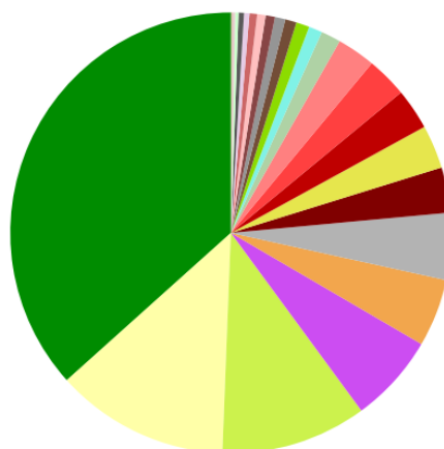
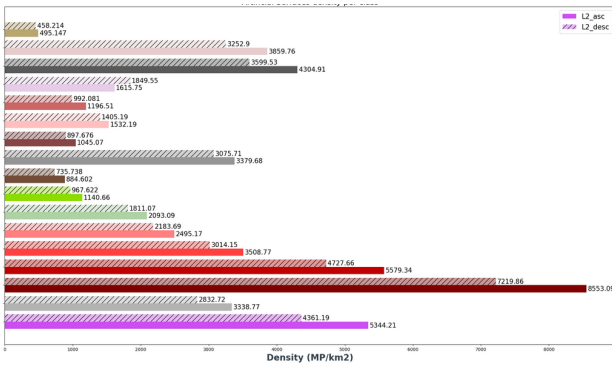
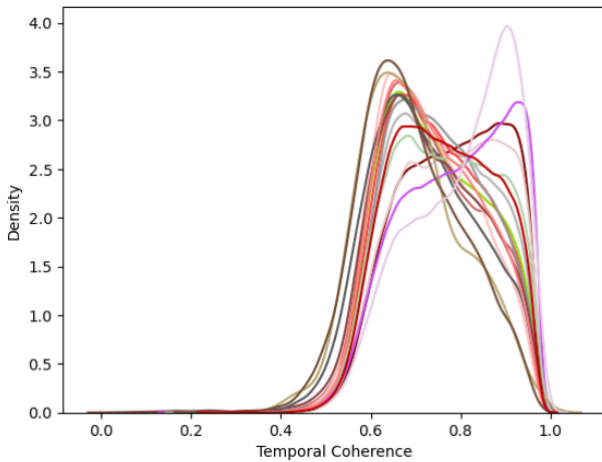


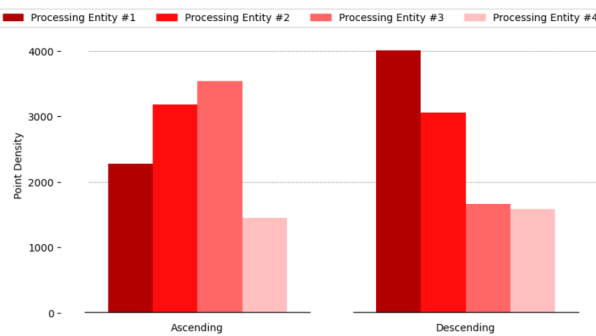
Figure 6. Distribution of CLC Urban Atlas classes (Barcelona, Spain). The legend can be found in the Appendix, Figure 10.



**Figure 7.** Density (MP/km<sup>2</sup>) for the group/class of Artificial Surfaces (Barcelona, Spain), also comparing asc/desc. The legend can be found in the Appendix, Figure 10.



**Figure 8.** Probability distributions for Temporal Coherence for the group of Artificial Surfaces, and Ascending orbit (Barcelona, Spain). The legend can be found in the Appendix, Figure 10.



**Figure 9.** Example of comparative analysis between the four processing entities, for the class of *Discontinuous very low urban fabric* (S.L.: < 10%). The legend can be found in the Appendix, Figure 10.

#### 4. CONCLUSIONS

The work presented in this paper describes a workflow for validating the output of the European Ground Motion Service (EGMS) in terms of spatial coverage and density of measurement points. We have developed a custom algorithm to address the challenge of overlapping bursts from different Sentinel-1 satellite tracks, ensuring a fair comparison among different

areas and eliminating any biases that could impact the results of the analysis.

The importance of validating the EGMS service cannot be overstated. As it provides crucial information on ground motion to a wide range of stakeholders, from scientists to policymakers, it is crucial that the service’s accuracy is verified and validated. Our workflow aims to contribute to this effort by providing a transparent, reproducible, and open approach to validate the point density of EGMS.

To achieve this goal, we employed open source data services of the European Environment Agency (EEA), specifically the Urban Atlas 2018 dataset (Atlas, 2018), to create a dataset of 12 selected sites across Europe, representing different urban and rural/mountainous areas. We also used open source software for geospatial analysis, ensuring that our methods can be easily replicated and built upon by others.

In conclusion, the workflow presented in this paper provides a valuable contribution to the validation of the EGMS service validation and demonstrates the importance of open science principles in ensuring the transparency and reproducibility of scientific research. We hope that our work will inspire further efforts to validate and improve the accuracy of the EGMS service, ultimately contributing to a better understanding of ground motion and its impact on our environment and society.

#### REFERENCES

Atlas, U., 2018. Copernicus Land Monitoring Service. *European Environment Agency: Copenhagen, Denmark*.

Costantini, M., Minati, F., Trillo, F., Ferretti, A., Novali, F., Passera, E., Dehls, J., Larsen, Y., Marinkovic, P., Eineder, M. et al., 2021. European ground motion service (egms). *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, IEEE*, 3293–3296.

European Environment Agency, 2022. Product description and format specification. [land.copernicus.eu/user-corner/technical-library/egms-product-description-document](http://land.copernicus.eu/user-corner/technical-library/egms-product-description-document).

Jordahl, K., den Bossche, J. V., Fleischmann, M., Wasserman, J., McBride, J., Gerard, J., Tratner, J., Perry, M., Badaracco, A. G., Farmer, C., Hjelle, G. A., Snow, A. D., Cochran, M., Gillies, S., Culbertson, L., Bartos, M., Eubank, N., maxalbert, Bilogur, A., Rey, S., Ren, C., Arribas-Bel, D., Wasser, L., Wolf, L. J., Journois, M., Wilson, J., Greenhall, A., Holdgraf, C., Filipe, Leblanc, F., 2020. *geopandas/geopandas: v0.8.1*.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B. E., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J. B., Grout, J., Corlay, S. et al., 2016. *Jupyter Notebooks-a publishing format for reproducible computational workflows*. 2016.

Solari, L., Balasis-Levinsen, J., Andersen, H. S., 2023. European ground motion service: production status and validation. Technical report, Copernicus Meetings.

## APPENDIX

In this appendix, we provide additional information that complements the main body of the article. Specifically, we include a color legend for the map used in the section 2.2 (Figure 4), in the figures in section 3, and tables that provide the names and codes of Urban Atlas classes used, split by groups (Tables 2 and 3)

Artificial Surfaces	
Code	Class Name
11100	Continuous Urban fabric
11210	Discontinuous Dense Urban Fabric
11220	Discontinuous Medium Density Urban Fabric
11230	Discontinuous Low Density Urban Fabric
11240	Discontinuous very low density urban fabric
11300	Isolated Structures
12100	Industrial, commercial, public, military and private units
12210	Fast transit roads and associated land
12220	Other roads and associated land
12230	Railways and associated land
12300	Port areas
12400	Airports
13100	Mineral extraction and dump sites
13300	Construction sites
13400	Land without current use
14100	Green urban areas
14200	Sports and leisure facilities

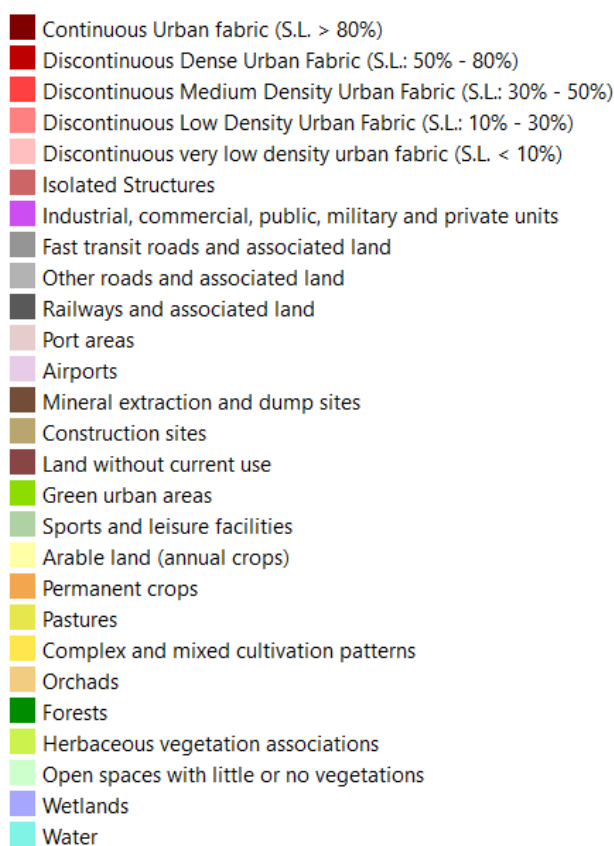
**Table 2.** Classes names and codes that have been grouped under Artificial Surfaces

Agricultural areas	
Code	Class Name
21000	Arable land (annual crops)
22000	Permanent crops
23000	Pastures
24000	Complex and mixed cultivation patterns
25000	Orchads

Forest and seminatural areas	
Code	Class Name
31000	Forests
32000	Herbaceous vegetation associations
33000	Open spaces with little or no vegetations

Water	
Code	Class Name
40000	Wetlands
50000	Water

**Table 3.** Classes names and codes for the groups of Agricultural areas, Forest and seminatural areas, and Water



**Figure 10.** Urban Atlas classes legend