Statistical Office in Olsztyn	**Statistical Office in Olsztyn**	Created 2019-03-05

WPH ESSnet Big Data II meeting 1

to WPH Group ESSnet Big Data II

cc

from Marek Morze WPH leader

subject Final minutes WPH meeting 1 of ESSnet BD II by WebEX on 28-02-2019

Participants:

Marek Morze (PL WPH leader)	√	Clara Schartner (DE Case study 6)	√
Przemysław Slesiński (PL Case study 1)	√	Bertrand Ballet (FR case study 7)	
Magdalena Mleczko (PL Case study 1)	√	Stefano Mugnoli (IT Case study 8)	√
Maria Yli-Heikkila (FI Case study 2)	√	Ana M. Santos (PT Case study 9)	√
Marc Callens (BE Case study 3)		Francisco Caldeira (PT Case study 9)	√
Vianney Costemalle (FR Case study 4)	√		√
Lyana Curier (NL Case study 5)	√		
Tim de Jong (NL Case study 5)	√		
Natalie Rosenski (DE Case study 6)	√		

Replacements/Guests:

Marc Debusschere (BE Case study 3)	√	Herve Le Grand (FR case study 7)	√
Heikki Laurila (FI Case study 2)	√		

1. Opening, agenda and self-introduction of the meeting participants

Marek welcomed everybody to the first meeting of the WPH Group of the ESSnet Big Data II (CG). Then the representatives of each case study introduced themselves.

2. Case study 1 to 9: state of affairs

Case study 1: Przemysław Slesiński

Przemysław reported that Sentinel-1 and Sentinel-2 data for 2017/2018 season were downloaded and at this stage data are being processed. SNAP algorithms and processing parameters are tested to develop a workflow for calculation of sigma nought scattering coefficient.

Case study 2: Maria Yli-Heikkila and Heikki Laurila

Maria and Heikki have been preparing a tender for acquiring preprocessed Sentinel-1 data. They will wait until the Grant Agreement is signed. In the meantime administrative data (annual agricultural land use (LPIS) and agricultural payment entitlements) was received and preprocessed.

Case study 3: Marc Debusschere

Marc replaced the absent Marc Calens. He presented formal problems regarding the relationship between the federal statistical office and the regional statistical office, which is real partner of our work package. He said about main goal of case study that is use of the aerial photography with very high resolution to crop recognition in case of small size parcels.

Case study 4: Vianney Costemalle

Vianney outlined the main issues related to case study 4 and shortly described 3 tasks that would be investigated: implementing the UN-Habitat methodology for the whole France, bench marking this methodology with specifics data or concepts that are available in France or in Europe and promoting the results at the French, European and Global level.

Case study 5: Lyana Curier, Tim de Jong

Lyana and Tim Case informed that a characterization of urban sprawl across urban areas by means of traditional machine learning methods is currently being carried out. To this end, MODIS NDVI dataset running from January 2004 to June 2018 with a fortnightly temporal and 250 m spatial resolution was downloaded and meshed into the domain of interest which spans from 53.75 N to 50.70 N and 3.25E to 7.25E.

For each pixel statics parameters such as mean,std,extrema,Q25, median, Q75, and an histogram of the NDVI timeseries were computed and a land cover information retrieved from the CORINE database (Agricultural areas, Artificial surfaces, Forest and semi natural areas, Water bodies, Wetlands) was provided. First test have shown, that all tested classifiers (support vector machine, random forest Random, k-nearest-neighbors) reached an accuracy higher than 0.75. Confusion matrices of all classifiers shows that agricultural areas can be confidently classified with 97% of the agricultural areas pixels correctly classified. All classifiers misclassified at least 30% of the Artificial surfaces as Agricultural areas. which is a problem when trying to identify urban sprawl. In the meantime, processing tools to access and process Sentinel and other higher resolution satellite data are being investigated.

Case study 6: Clara Schartner and Natalie Rosenski

Clara presented the goal of Destatis and the German Federal Agency for Cartography and Geodesy (BKG) in this case study, which is the combination of official statistics and earth observation data to gain insight into the quality of urban life. At this stage Destatis and BKG are working on identifying a suitable aspect of urban quality of life and the corresponding data to measure it.

Case study 7: Herve Le Grand

Herve told about main objectives of the case study conducted by SSP France. The SSP plans to use the methodology for the fully automatic production of land cover maps at country scale developed by the CESBIO. The SSP aims at implementation of this technology to conduct the French statistical area-frame survey on land cover and land use - TERUTI.

Case study 8: Stefano Magnoli

Due to technical problems, there was no communication with Stefano.

Case study 9: Ana M. Santos

Ana and Francisco have prepared a slides for presenting the scope of work and people involved. At this time geometry of 2011 Census in order to identify Changed Areas and the use of Copernicus data for validation on Agricultural Census 2019 are being analyzed.

3. Presentation of WPH methodological framework

The main point of the meeting was presentation of WPH Methodological Framework (MF) by Magdalena Mleczko.

The methodological framework is divided into five stages: Pre-works and Stages from 1 to 4. The stages consist of tasks. Additionally the actions related to accuracy assessment, IT infrastructure, metadata are indicated. At the beginning, before stages from 1 to 4 description of the study background, which means literature, data sources, toolkit and software review is planned. This description of background will contribute to creation the State-of-Art. Within the pre-works the statistical products should be defined. Stage 1 is focused on specification of test area and collecting the EO and administrative data. Within Stage 1 the quality assessment needs to be done. For example assessment of completeness and usability of collected data. In case of satellite data it can be cloud cover, radiometric and geometric quality information. Additionally the metadata of collected data should be described. Stage 2 is the preparation of acquired and downloaded data for main processing and analysis. As in Stage 1 the quality assessment, IT infrastructure, metadata need to be done. Stage 3 is a development of the methods and procedures to be used for producing statistics. This stage includes data processing (e.g. image segmentation, classification, learning machine etc.) and complex analysis of the results. The used methods and procedures should be described. As in Stage 1 and 2 the quality assessment, IT infrastructure, metadata need to be done. The last stage is Stage 4 and includes pilot production, validation and final conclusions. All elements of methodological framework need to be detailed described. For this purpose the template of report is prepared. Template is organized stage by stage. The M12 is deadline for submitting the interim technical report. To remind interim technical report concerns the Reference Methodological Framework and the conditions for using the data, the methodology and the procedures to be

used for producing statistics. Summarizing, the points 1, 2 and description part of 3.1 and 4.1 should be filled for each case study before M12.

4. Adoption of WPH methodological framework

Each participant adopted the presented methodological framework concept, however, it required some clarification and refinement. Magdalena Mleczko explained and answered all the questions.

Natalie Rosenski commented that all activities together within pilot project is an experiment while the experiment statistics are indicated only for stage 3. Magdalena Mleczko answered that it is only suggestion to precise localize the experimental statistic, just for underline what is new and modern in our methodologies.

Vianney Costemalle had a question about data pre-processing, what does it mean exactly. Magdalena Mleczko said that pre-processing is preparation of acquired and downloaded data. It can mean: unzipping and importing original data to expected format, data reformatting, database re-shaping and necessary information extraction, radiometric and geometric correction of images, SAR pre-processing from single look complex/ground range data to calibrated sigma nought orthoimages etc. Usually pre-processing doesn't occur separately, but when radar data are used the preparation in a huge part of the whole processing as in case study no. 1.

5. Any remaining issues and closing.

At the end of the meeting, the approximate date of the next meeting between 15.04-15.05.2019 was set. Also Marek Morze told about literature review WPK template which will be sent to all WPH participants for providing information on literature which are relevant to each case study.