EARSeL Manchester 2024
Abstract
Corresponding Author: luke.mcquade@plus.ac.at

An Approach for the Semantic Enrichment of Sentinel-1 Imagery Suitable for Large-scale Analysis

Luke McQuade[1], Martin Sudmanns[1], Dirk Tiede[1]

[1] UNIVERSITY OF SALZBURG, DEPARTMENT OF GEOINFORMATICS - Z_GIS, AUSTRIA

Keywords (5): Earth Observation, Synthetic Aperture Radar (SAR), Big Earth Data, Land Use, Copernicus

Challenge

Synthetic Aperture Radar (SAR) Earth observation (EO) satellites have several advantages over their optical counterparts, such as being able to observe the Earth's surface at night, and through a wide variety of weather conditions. However, due to the nature of their sensors and mechanisms of capture, the resultant imagery is often difficult to interpret and use in downstream analyses. Several approaches exist for the semantic enrichment of optical data, such as the Satellite Imager Automatic Mapper (SIAM[*]), which, coupled with their use in EO data cubes, can greatly improve accessibility and use of the original data. A system offering similar benefits for SAR EO data could be highly beneficial, especially considering the potential to complement optical data. Designing such a system to permit analyses across differing geographic areas globally presents an additional challenge which we have also attempted to address in this work.

Methodology

We devised an approach for the semantic enrichment of dual-polar Sentinel-1 radiometric-terrain-corrected (RTC) backscatter imagery (VV and VH polarizations). We refer to this as *polarimetric categorization*. It consists of binning the parameter space of VV and VH backscatter according to the scattering properties of known surface types; the result is that each pixel is assigned a category according to the scattering type(s) exhibited, e.g., surface scattering, volume scattering, double-bounce. The categorization/binning is performed with a decision tree algorithm, with set (constant) thresholds. These thresholds were determined in a part knowledge-based, part data-driven process. Figure 1 shows our preliminary categorization scheme.

Our categorization processor was implemented as a Python package, with working title, *dpolcat* (dual-polarimetric categorizer). It depends on standard EO packages such as xarray[†] and a STAC client. Just-in-time (JIT) compiler Numba[‡] was also utilized to improve computational performance.

Several trials were conducted in using categorized scenes generated by our algorithm in downstream analyses, such as flood mapping, burned-area delineation, and vegetation change mapping.

Expected results

[*] Baraldi, A., Humber, M. L., Tiede, D., & Lang, S. (2018). GEO-CEOS stage 4 validation of the Satellite Image Automatic Mapper lightweight computer program for ESA Earth observation Level 2 product generation – Part 1: Theory. Cogent Geoscience, 4, doi: 10.1080/23312041.2018.1467357

[†] https://xarray.dev/

[‡] https://numba.pydata.org/

Our preliminary version of *dpolcat* has facilitated several downstream trial analyses using models constructed in Python scripts. The first was an automatic flood mapping task. Observing the grid cells newly changed into strong surface scatterers (category 1 as in Figure 1), a model was built. Some simple spatial filtering and thresholding were also applied. The model was applied to map the flooding event of Duisburg, Germany, July 2021. Validated against the Copernicus Emergency Mapping Service reference, an F1 score of 0.80 was achieved. The source, intermediate, and resultant imagery are shown in Figure 2.

Other tasks include burned-area delineation and vegetation change mapping. Quantitative validation of these is ongoing. But, in all cases, the models using polarimetric categories were quick and simple to design and implement, compared to those in similar analyses using non-enriched images – models can be implemented in minutes rather than hours.

In terms of computational performance, it takes approximately 4 minutes to process an entire Sentinel-1 scene with *dpolcat*, using an Intel Xeon Platinum 8272CL CPU @ 2.60GHz.

Outlook for the future

As the category thresholds are globally constant, the algorithm can be applied to any Sentinel-1 scene, or across an area spanning multiple scenes, without a learning step - this is in contrast with similar techniques such as polarimetric decomposition and clustering. Also, compared to learning-based techniques, the algorithm is simple enough such that the computational cost for processing a scene is relatively low. This lends itself well to its use in semantic EO data cubes[§], where there is often a requirement to semantically enrich large numbers of scenes as a pre-requisite for further analyses.

Our preliminary results have identified some limitations of the approach. In time series analysis, especially with vegetated and mixed-use areas, the category assigned to a given location can vary unexpectedly, i.e., where there is seemingly no change to the underlying land-cover type. We refer to this as categorical instability. This could be due to, for instance, signal values close to category thresholds; and, with noise (which radar is prone to), a pixel can 'flip' between categories over time. Furthermore, there is scope to refine the implementation and integrate *dpolcat* processing into existing EO data cubes. We aim to address these in future work.

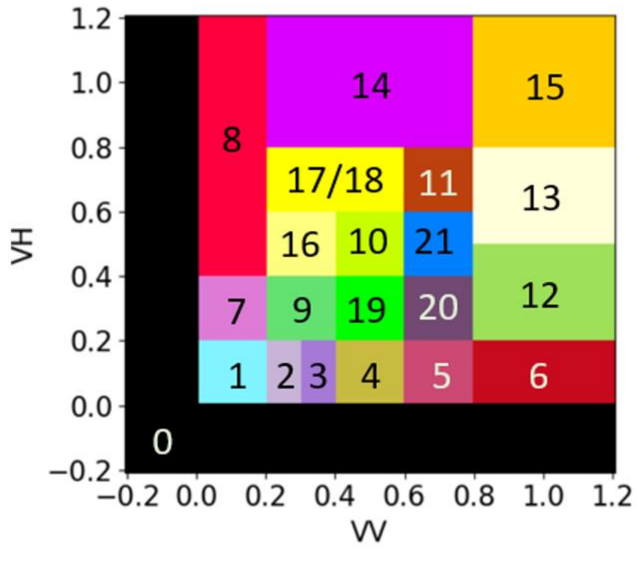

Figure 1 Division of the VV and VH parameter space into polarimetric categories: 1, 2, 3 and 4 represent mostly surface scattering; 9, 10, 16 and 19 represent mostly volume scattering; 8 is ill-defined (physically improbable); 0 is invalid or 'no data'; the remaining represent double-bounce and other phenomena.

§ Augustin, H., Sudmanns, M., Tiede, D., Lang, S., Baraldi, A., 2019. Semantic Earth Observation Data Cubes. Data 4, 102. https://doi.org/10.3390/data4030102

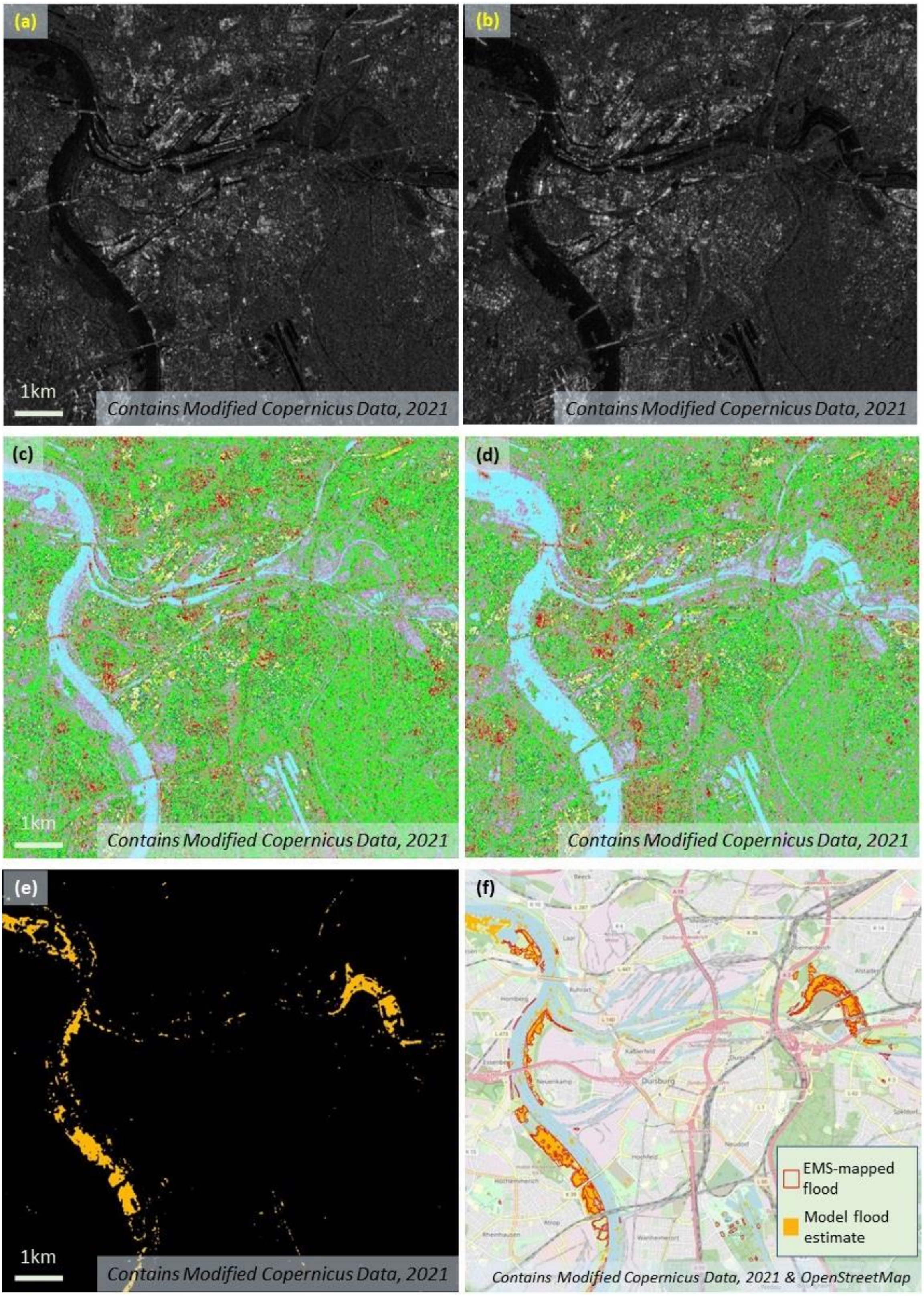

Figure 2 Example of using *dpolcat* in an automated flood mapping workflow – Duisburg, Germany, 13 July 2021 (pre-flood) to 16 July 2021 (in-flood). (a) Sentinel-1 VV backscatter, pre-flood. (b) Sentinel-1 VV backscatter, in-flood. (c) Pre-flood image categorized with *dpolcat*. (d) In-flood image categorized with *dpolcat*. (e) Flood map model output. (f) Flood map model output with Copernicus EMS reference**.

** Copernicus Emergency Mapping Service, product [EMSR517].