

Development of an Earth Observation Cloud Platform in Support to Water Resources Monitoring

Andreea Bucur, Wolfgang Wagner, Stefano Elefante, Vahid Naeimi, and Christian Briese

Earth observation (EO) satellites collect verifiable observations that allow tracing natural and anthropogenic changes from local to global scale over several decades. Multi-decadal data sets are already available from various types of EO sensors, but their effective exploitation is hindered by the lack of data centres which offer dedicated EO processing chains and high-performance processing (HPC) capabilities. Recognizing this need, TU Wien founded the EODC Earth Observation Data Centre for Water Resources Monitoring together with other Austrian partners in May 2014 as a public–private partnership. The EODC aims at providing an independent science-driven platform that is transparent for its users and offering a high diversity and flexibility in terms of data sets and algorithms used. In this contribution, we describe the collaborative approach followed by EODC to build up its infrastructure and services and briefly introduce three pilot services.

A. Bucur (✉) • S. Elefante • V. Naeimi
Department of Geodesy and Geoinformation, Technische Universität Wien, Gußhausstraße
27-29/E120, 1040, Vienna, Austria
e-mail: andreea.bucur@geo.tuwien.ac.at; stefano.elefante@geo.tuwien.ac.at;
vahid.naeimi@geo.tuwien.ac.at

W. Wagner
Department of Geodesy and Geoinformation, Technische Universität Wien, Gußhausstraße
27-29/E120, 1040, Vienna, Austria

EODC Earth Observation Data Centre for Water Resources Monitoring GmbH, Gusshausstrasse
27-29/CA 02 06, 1040, Vienna, Austria
e-mail: wolfgang.wagner@geo.tuwien.ac.at

C. Briese
EODC Earth Observation Data Centre for Water Resources Monitoring GmbH, Gusshausstrasse
27-29/CA 02 06, 1040, Vienna, Austria
e-mail: christian.briese@eodc.eu

Introduction

Humans have changed the natural environment since their early existence. However, the scale of human impacts has become dramatic only in the past 60 years (Steffen et al. 2015). During this so-called “Great Acceleration” period (McNeill and Engelke 2016) scientific and technological progress lead to extensive production and offer of goods and services, and an overall improvement in the standard of living for billions of people (Bhaduri et al. 2014). This period also witnessed a sharp increase in the world population, going up from three billion in 1959 to seven billion in 2012 (US Census Bureau 2016). All this had a dramatic impact on the consumption of natural resources. One of the resources, which are increasingly under pressure, is water. Water is pivotal for the well-being of humans and natural ecosystems: agricultural and industrial production, biodiversity, human health etc. In the “Global Risks 2015” report of the World Economic Forum, the “water crisis” is rated as the risk with the highest societal impact (World Economic Forum 2015). Therefore, it is crucial to understand natural and anthropogenic influences on the water cycle and the factors that might determine changes over time (e.g. Oki et al. 2004; Tang and Oki 2016). Major attention must be given to the rise in global temperature – e.g. year 2016 (January–October) was reported as the warmest in historical records (NOAA 2016)—and consequently to a warmer climate which is generally acknowledged to prompt an increased occurrence of extreme events such as floods and droughts (e.g. IPCC 2013; Trenberth and Asrar 2014).

In this context continuous monitoring of water resources is essential. In order to improve water management practices reliable information about anthropogenic and natural impacts, and their interactions must be readily available. Ground-measurements are fundamental for this purpose. However, they have many shortcomings: sparse information over small areas, lack of representativeness at larger scale, high costs of maintenance, out of date or failed equipment and lack of funds to replace them etc. Complementing in situ networks, monitoring tasks are increasingly fulfilled by earth observation (EO) satellites which have been acquiring measurements of the land, atmosphere and oceans since the beginning of the 1970s. The new generation of sensors is able to collect an unprecedented amount and variety of observational data at high spatial resolution and short repeat intervals. A vast and diverse amount of EO data is, therefore, readily available to be mined for new insightful information; but this task is not short of challenges. As we will detail further in section “EODC: The Earth Observation Data Centre for Water Resources Management”, dedicated data centres that stimulate collaboration are needed for the effective exploitation of satellite images.

Here we present the EODC Earth Observation Data Centre for Water Resources Management which was founded as a public–private partnership with the aim to assist in water management by making use of earth observation data and big data cloud computing infrastructures. In the next section, the organisational and technical aspects of EODC are presented. In section “Pilot Services”, initial pilot services are briefly described.

EODC: The Earth Observation Data Centre for Water Resources Management

The EODC Earth Observation Data Centre for Water Resources Monitoring (www.eodc.eu) is a public-private partnership founded in May 2014, in Austria, by the Technische Universität Wien (TU Wien), the Austrian Meteorological and Geodynamics Institute (ZAMG), two private companies and individuals. The early idea of EODC was born already in 2011, and was prompted by the need to cope with exponentially growing data volumes and their scientific exploitation with increasingly complex algorithms (Wagner et al. 2014). EODC was set up as an international cooperation network which brings together scientific institutions, public organizations and several private partners from countries within and outside Europe.

Working with EO data on cloud platforms is not short of scientific, technical and organizational challenges as described in Wagner et al. (2014). The science is driven by the need to gain an integrated view of all processes driving the water cycle (Wagner et al. 2009). This requires analyses of many different geophysical parameters and their coupled feedbacks (e.g. soil moisture, temperature, precipitation, vegetation indices) based on data from multiple sensors (e.g. active and passive radar, optical imaging satellites) and their integration into earth system models. Thus, the information contained in satellite images becomes meaningful only after several specific processing steps.

Traditionally, the ground segments of EO missions have delivered raw images to remote sensing experts who, after high-level data processing (geo-referencing, normalization, radiometric correction etc. of data), have handed it out to application oriented users (hydrology, forestry, urban planning etc.). The later have extracted added-value information which can be further used for specific purposes (mapping for forest management etc.). This long-established system has assured that all parties had full control over the ownership of data and software. But this approach is inefficient because the data and resources (such as storage and processing capabilities or specific expertise for EO data processing) are basically duplicated for each user. Today, this traditional approach is reaching its limits. This is because, firstly, the latest generation of sensors generate huge amounts of data. To give one example, European Space Agency's Sentinel-1 satellites acquire in one year more data than their predecessor ENVISAT Advanced Synthetic Aperture Radar (ASAR) has done so in 10 years of operation (25 Terabytes in the first year of S1, 23.5 Terabytes in 10 years of ASAR). With a data capture rate of about 1.8 Terabytes per day, Sentinel-1 will acquire over its 7-year nominal mission lifetime over 1 Petabyte of raw data (Wagner 2015). Secondly, the algorithms used to transform the

EO data into useful information become increasingly more complex. Last but not least, a model may be run with more than just one data set or several complementary methods are combined into an ensemble in order to obtain the most reliable results and to estimate the uncertainty range of the predictions.

Considering the above, the way how EO data are stored, processed and distributed needs to be changed fundamentally. This has already been recognized by a number of private and public entities that have started to offer big data infrastructures for processing EO data (Wagner 2015). Some examples include private companies such as Google, and Amazon, and the public initiatives THEIA Land Data Centre in France or the Climate, Environment and Monitoring from Space (CEMS) initiative in the UK. Their solutions typically combine cloud technologies and high-performance computing (HPC) to allow users to explore large amounts of data via an internet connection. In other words, “the software moves to the data” rather than data being moved to the software on local working stations.

Similar to the above-mentioned entities also the EODC offers such a novel framework for working with EO data; its users have the possibility to access EO data via a cloud platform, process them with their own algorithms and extract the results. In order to be attractive for both scientific and operational users, the EODC infrastructure combines elements for operational data reception and processing, cloud platforms and storage system for scientific analysis collocated with advanced HPC capabilities. The EODC infrastructure has not been built from scratch but exploits as good as possible existing data centre capabilities by federating and integrating them. The main EODC data centre capabilities are currently located at the TU Wien Science Centre Arsenal collocated with the Vienna Scientific Cluster 3 (VSC-3). Figure 1 gives an overview of the status of this infrastructure at the end of 2016. It is planned to successively extend these capabilities in order to allow storing and processing the complete global Sentinel data archives.

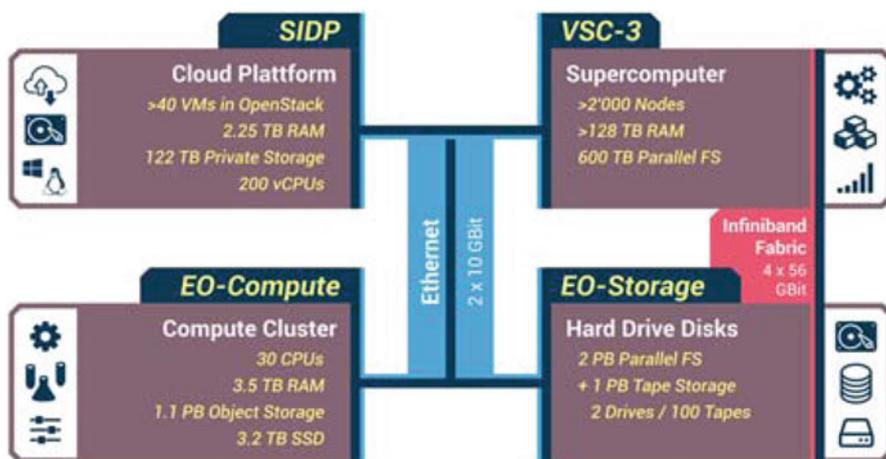


Fig. 1 The EODC infrastructure at the end of 2016

Several experiments carried out with Sentinel-1 SAR and ENVISAR ASAR data sets have already demonstrated the scalability of the EODC supercomputing environment. For example, a batch of 31, 978 Sentinel-1 images over Europe, with a total size of around 30 Terabytes (TB), was processed with TU Wien's SAR Geophysical parameters Retrieval Toolbox (SGRT). This Python package incorporates the ESA's Sentinel-1 Toolbox (S1TBX) and consists of modules for EO data pre-processing, model parameters extraction, and data production (Naeimi et al. 2016). Processing the 31, 978 Sentinel-1 images on the VSC-3 with around 300 nodes took roughly 10 days compared to more than 1 year that would be needed when processing the same data set with the same software with only 1 node (Elefante et al. 2016).

In terms of data availability, EODC hosts at the TU Wien Science Centre Arsenal a nearly complete and up-to-date data archive from its main sensors of interest (Sentinel-1, Sentinel-2, Sentinel-3). Additional data are available through the other EODC data centres operated by EODC cooperation partners (ZAMG, VITO NV, EURAC research). In this way the EODC decentralised IT infrastructure provides its users access to an extended and diverse number of data sets, trying to minimise the duplication of data as much as possible.

An ultimate goal of EODC is to encourage its partners and users to engage in collaborative science activities. The organisational structure was design to facilitate this by offering more than just access to performant processing resources. Thus, as described in Wagner et al. (2014), partners come together in so called communities which are formed around particular research topics (e.g. soil moisture), applications/services (e.g. drought monitoring, flood mapping) or tasks (e.g. software development, shared infrastructure resources). The participation in the EODC cooperation network is flexible according to one's interests and contribution, and can take one of the three forms of partnership: Principal Cooperation Partners, Associated Cooperation Partners or Developers. Facilitated by this bundling of interests, several EO data services are currently being developed jointly by several EODC partners.

Pilot Services

Several joint EODC services are already under development. These services typically rely on individual sensors, but ultimately the goal will be to benefit from the collocation of many diverse data sets by building multi-sensor data services. An example for single-sensor service is the Sentinel-2 data service platform developed by researches of the University of Natural Resources and Life Sciences (BOKU), Vienna, and run on the EODC infrastructure. As described by Vuolo et al. (2016), users of this service platform can submit processing requests and access the results via a user-friendly web page or using a dedicated application programming interface (API). Data products that can be produced in this way are atmospherically corrected

Sentinel-2 images and value-added products with a particular focus on agricultural vegetation monitoring, such as leaf area index (LAI) and broadband hemispherical-directional reflectance factor (HDRF).

An example for a multi-sensor service is the ESA CCI soil moisture data service as described by Dorigo et al. (2017). Soil moisture is an important component of the water cycle and the satellite-based products derived from active and passive microwave are increasingly being used for a wide range of applications (Dorigo and de Jeu 2016). For example, satellite-based soil moisture may be used for estimation of near-future vegetation health (Qiu et al. 2014), improved calculation of crop water requirement (McNelly et al. 2015) and operational drought warnings (Enenkel et al. 2016). EODC currently leads the second phase of the ESA Climate Change Initiative Soil Moisture project, providing the operational framework for merging more than a dozen of satellite data sets into consistent long-term soil moisture data records (Liu et al. 2012, 2011; Wagner et al. 2012).

As a last example, we note that several Sentinel-1 data services are currently being developed by the Remote Sensing Research Group of TU Wien in collaboration with other EODC partners. They range from a simple Sentinel-1 image-compositing service to processing services for the monitoring of soil moisture, water bodies, wetlands, and forests from regional to global scales. Figure 2 illustrates a false-colour composite of the Sentinel-1 data acquisitions before and after a flooding event on December 2015 (BBC News 2015) in Carlisle, UK.

Closing Remarks

In this subchapter we introduced the EODC Earth Observation Data Centre for Water Resources Management, which is a private–public entity founded for enabling the collaboration of scientific, public and private organizations for processing EO data in the cloud. As its name suggests, one of founding idea of EODC was to focus on the thematic area of water resources management, but thanks to the rapid growth of the EODC cooperation network, the number of application domains has been growing accordingly. In particular, agricultural monitoring and land use mapping applications have become important topics of collaboration between EODC partners. The experiences made over the short period since the foundation of EODC in 2014 show that EODC offers a framework for collaboration that can assist the development of long and complex data processing lines going from the raw EO data to the final model predictions (runoff forecast, crop yield etc.). Taking advantage of the big data technologies, latest scientific algorithms can be scaled up to process high-resolution EO data from regional to global scales. This is a crucial step towards operational applications, which are ultimately needed to enhance the social benefits of EO technology.

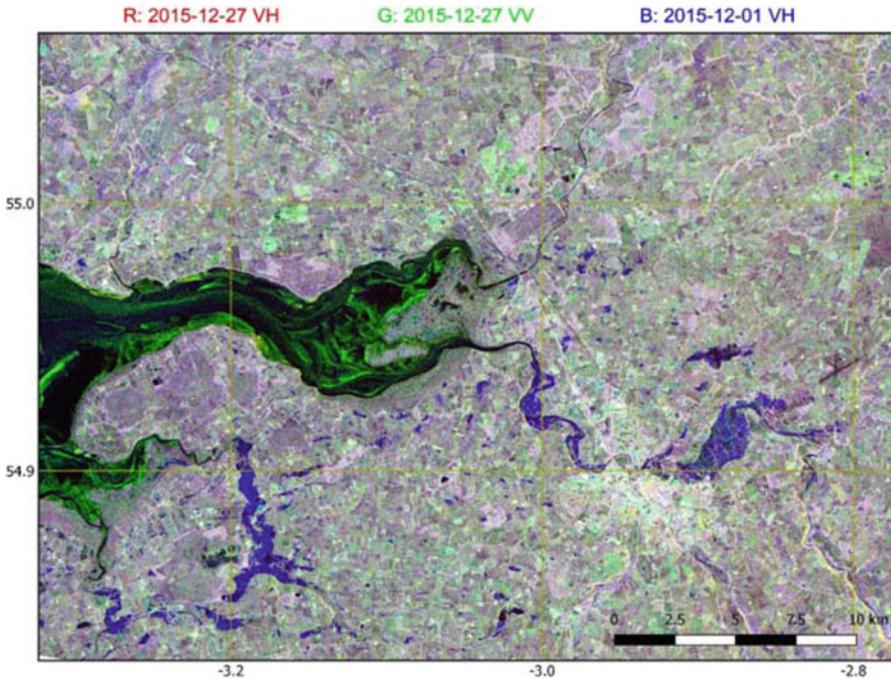


Fig. 2 False-colour composite generated from Sentinel-1 data acquisitions in Interferometric Wide swath (IW) mode, Ground Range Detected (GRD) high resolution product (20 m), VV and VH polarization from 1st and 27th of December 2015

References

- BBC News (2015) December storms' trail of destruction, 29 Dec 2015. Available from: <http://www.bbc.com/news/uk/35193682>. Accessed on 10 Dec 2016
- Bhaduri A, Bogardi J, Leentvaar J, Marx S (2014) The global water system in the anthropocene. Challenges for science and governance. Springer International Publishing, New York, NY. 437 pages
- Dorigo W, Wagner W, Albergel C, Albrecht F, Balsamo G, Brocca L, Chung D, Ertl M, Forkel M, Gruber A, Haas E, Hamer P, Hirschi M, Ikonen J, de Jeu R, Kidd R, Lahoz W, Liu YY, Miralles D, Mistelbauer T, Nicolai-Shaw N, Parinussa R, Pratola C, Reimer C, van der Schalie R, Seneviratne SI, Smo-lander T, Lecomte P (2017) ESA CCI soil moisture for improved Earth system understanding: state-of-the art and future directions. *Remote Sens Environ.* in press. <https://doi.org/10.1016/j.rse.2017.07.001>
- Dorigo W, de Jeu R (2016) Satellite soil moisture for advancing our understanding of earth system processes and climate change. *Int J Appl Earth Observ Geoinform* 48:1–4
- Elefante S, Wagner W, Briese C, Cao S, Naeimi V (2016) High-performance computing for soil moisture estimation. In Proceedings of the 2016 conference on Big Data from Space (BiDS'16), Santa Cruz de Tenerife, Spain, pp 95–98. <https://doi.org/10.2788/854791>

- Enenkel M, Steiner C, Mistelbauer T, Dorigo W, Wagner W, See L, Atzberger C, Schneider S, Rogenhofer E (2016) A combined satellite-derived drought indicator to support humanitarian aid organizations. *Remote Sens* 8(4):340
- IPCC (2013) Climate change 2013: the physical science basis. In: Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM (eds) Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on climate change. Cambridge University Press, Cambridge., 1535 pp. <https://doi.org/10.1017/CBO9781107415324>
- Liu YY, Dorigo WA, Parinussa RM, De Jeu RAM, Wagner W, McCabe MF, Evans JP, Van Dijk AIJM (2012) Trend-preserving blending of passive and active microwave soil moisture retrievals. *Remote Sens Environ* 123:280–297
- Liu YY, Parinussa RM, Dorigo WA, De Jeu RAM, Wagner W, Van Dijk AIJM, McCabe MF, Evans JP (2011) Developing an improved soil moisture dataset by blending passive and active microwave satellite-based retrievals. *Hydrol Earth Syst Sci* 15:425–436
- McNeill JR, Engelke P (2016) The great acceleration: an environmental history of the anthropocene since 1945. Harvard University Press, Cambridge, MA
- McNelly A, Husak GJ, Brown M, Carroll M, Funk C, Yatheendradas S, Arsenault K, Peters-Lidard C, Verdin JP (2015) Calculating crop water requirement satisfaction in the West Africa Sahel with remotely sensed soil moisture. *J Hydrometeorol* 16(1):295
- Naeimi V, Elefante S, Cao S, Wagner W, Dostalova A, Bauer-Marschallinger B (2016) Geophysical parameters retrieval from sentinel-1 SAR data: a case study for high performance computing at EODC. In Proceedings of the 24th High Performance Computing Symposium (HPC '16). Society for Computer Simulation International, San Diego, CA., Article 10, 8 pages. [10.22360/SpringSim.2016.HPC.026](https://doi.org/10.22360/SpringSim.2016.HPC.026)
- NOAA (2016) National Centers for Environmental Information, State of the Climate: Global Analysis for October 2016 (published online November 2016). Available from: <http://www.ncdc.noaa.gov/sotc/global/201610>. Accessed on 10 Dec 2016
- Oki T, Entekhabi D, Harrold TI (2004) The global water cycle. In: The state of the planet: Frontiers and challenges in Geophysics. American Geophysical Union, Washington, DC, pp 225–237
- Qiu J, Crow WT, Nearing GS, Mo X, Liu S (2014) The impact of vertical measurement depth on the information content of soil moisture times series data. *Geophys Res Lett.*, 2014GL060017. <https://doi.org/10.1002/2014GL060017>
- Steffen W, Broadgate W, Deutsch L, Gaffney O, Ludwig C (2015) The trajectory of the Anthropocene: the great acceleration. *Anthrop Rev* 2(1):81–98
- Tang Q, Oki T (eds) (2016) Terrestrial water cycle and climate change: natural and human-induced impacts, vol 221. John Wiley & Sons., New York, NY
- Trenberth KE, Asrar GR (2014) Challenges and opportunities in water cycle research: WCRP contributions. *Surv Geophys* 35(3):515–532
- Vuolo F, Žóltak M, Pipitone C, Zappa L, Wennig H, Immitzer M, Weiss M, Baret F, Atzberger C (2016) Data service platform for Sentinel-2 surface reflectance and value-added products: System use and examples. *Remote Sens* 8:938
- US Census Bureau (2016) International Data Base, Updated August 2016. Available from: www.census.gov. Accessed on 10 Dec 2016
- Wagner W (2015) Big data infrastructures for processing sentinel data. In: Fritsch D (ed) Photogrammetric week, pp 93–104
- Wagner W, Fröhlich J, Wotawa G, Stowasser R, Staudinger M, Hoffmann C, Walli A, Federspiel C, Aspetsberger M, Atzberger C, Briese C, Notarnicola C, Zebisch M, Boresch A, Enenkel M, Kidd R, von Beringe A, Hasenauer S, Naeimi V, Mücke W (2014) Addressing grand challenges in earth observation science: the Earth Observation Data Centre for water resources monitoring. *ISPRS Annals Photogram Remote Sens Spatial Inform Sci* 2(7):81

- Wagner W, Dorigo W, de Jeu R, Fernandez-Prieto D, Benveniste J, Haas E, Ertl M (2012) Fusion of active and passive microwave observations to create an Essential Climate Variable data record on soil moisture. In: XXII ISPRS Congress, Melbourne, Australia
- Wagner W, Verhoest N, Ludwig R, Tedesco M (2009) Remote sensing in hydrological sciences. *Hydrol Earth Syst Sci* 13(6):813–817
- World Economic Forum (2015) *Global risks 2015*, 10th edn. World Economic Forum, Geneva

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

