E-COLLABORATION FOR EARTH OBSERVATION: EXAMPLE CHALLENGE FOCUSED ON THE ATMOSPHERIC CORRECTION OF OCEAN COLOUR DATA

Samantha Lavender[(1)], Enguerran Boissier[(2)], Fabrice Brito[(2)], Rita Ribeiro[(3)], Christina Aas[(4)], Francesco Casu[(5)], Michele Iapaolo[(6)]

[(1)] Pixalytics Ltd, 1 Davy Road, Tamar Science Park, Derriford, Plymouth, Devon, PL6 8BX, UK. E-mail: slavender@pixalytics.com
[(2)] Terradue UK Ltd, 11 Old Jewry, London, EC2R 8DU, UK. E-mail: enguerran.boissier@terradue.com and fabrice.brito@terradue.com
[(3)] UNINOVA, Campus FCT/UNL, Monte da Caparica, 2829-516, Portugal. Email: rar@ca3-uninova.org
[(4)] Science and Technology AS, PO Box 72, Torshov, 0412 Oslo, Norway. Email: aas@stcorp.no
[(5)] IREA institute (CNR), Via Diocleziano 328, 80124 Napoli, Italy. Email: casu.f@irea.cnr.it
[(6)] ESA/ESRIN, ESA/ESRIN, Frascati, Italy. Email: Michele.Iapaolo@esa.int

ABSTRACT

The E-Collaboration for Earth Observation (E-CEO) project aimed to deliver a collaborative platform that, through data challenges, would improve the adoption and outreach of new applications and methods to process Earth Observation (EO) data. To test the E-CEO platform, a contest based on the Atmospheric Correction (AC) of ocean color data was proposed. Existing processors were tested, and the evaluation results analysed.

Overall, the challenge showed that the E-CEO platform can be used to simplify the process of comparing different processors. Once the different participants' software is uploaded and connected to the data packages, the processing runs automatically, and so the processing is quick to re-run and adjust. However, additional work has shown it's important to continue to have human involvement in the evaluation step as otherwise statistics may be incorrectly interpreted due to outliers.

1. INTRODUCTION

Data challenges have become a method of promoting innovation within data-intensive applications; building or evolving user communities, and potentially developing sustainable commercial services. The participants' can utilise the vast amounts of information (both in scope and volume) that's available online, and take advantage of reduced processing costs when virtual machines are provided for.

Round robin activities are part of the Ocean Colour community activities with recent examples having been carried out within the context of the European Space Agency (ESA) Ocean Colour Climate Change Initiative (OC-CCI) and CoastColour projects [1]. The OC-CCI project carried out a round robin for AC, which was focused on Case 1 (phytoplankton dominated) waters; Phase 2 started in 2014, and is being extended to Case 2 (coloured dissolved organic matter and / or suspended sediment dominated) dominated waters with the version 3.0 dataset due for release at the end of June 2016. However, there are limited opportunities for new approaches to be adopted as any approach needs to be ready to run on multiple satellite missions. There was also a theoretical / simulated data comparison carried out within the context of the International Ocean Colour Coordinating Group (IOCCG) [2]. This has since been extended to an 'Intercomparison of Atmospheric Correction Algorithms over Optically-Complex Waters' [3], which started in January 2014, but has yet to publish its results.

Within the frame of the E-CEO project, funded by ESA, the activity is called a challenge; a nominal challenge winner will be identified (i.e. the Application that scores the highest ranking overall), but it's accepted this will be under a specific set of criteria and what's of scientific interest for this Challenge is the underlying knowledge gained. For the participants, running their own code on a provided cloud platform might be a new experience and so it's expected that at least some may require support to have their in-house breadboards ported to this environment.

2. METHODOLOGY

2.1. E-CEO Platform

The backbone of the E-CEO platform is a common environment where the applications can be developed, deployed and executed. Then, the results are published via a common visualization platform for their effective validation and evaluation.

The E-CEO Web Portal and Data Repository has been developed from the ESA Grid Processing on Demand (G-POD) platform with the Cloud Interoperability Operational Pilot (CIOP) Sandbox and Runtime Environment (RTE) as the E-CEO Private Development Environment and Common Deployment / Evaluation environments, respectively. By utilising G-POD, the E-

CEO platform was provided with the capability to manage a large number of computing nodes and host EO processors that would need to access several tens of Terabytes of on-line EO data. Then, by also linking into CIOP, the G-POD system could be extended to allow participants' to develop and test their new applications within a virtualized environment prior to their deployment and exploitation.

In addition, the operational Single Sign-On system for ESA Web-based Applications (UM-SSO) is used for managing E-CEO users' access to the platform.

2.2. Example Challenge

The example challenge was based around the AC of ocean colour: processing Level 1 (L1) top of atmosphere calibrated radiances / reflectances to Level 2 (L2) bottom of atmosphere calibrated radiances / reflectances and derived products. With this being the first challenge that was run on the new E-CEO platform, it was decided to restrict this Challenge to a maximum of 10 participants for the 2014 activity.

The input data were the Medium Resolution Imaging Spectrometer (MERIS) L1 full resolution full swath (FRS) files and matchups were extracted for sites held within the National Aeronautics and Space Administration (NASA) bio-Optical Marine Algorithm Data set (NOMAD) [4], with processors tested including the:

● ESA BEAM Toolbox [5] SMAC processor - The Simplified Method for Atmospheric Correction [6] is primarily designed as an above-land rather than above-water processor and so only a continental or desert air mass is available.

● NASA SeaWiFS Data Analysis System (SeaDAS) [7] v7.1 standard AC processor: the processor that's used for the systematic processing of ocean colour missions by the NASA Ocean Biology Processing Group, which uses the near infra-red to estimate the aerosol contribution whilst accounting for a water signal in turbid waters [8].

● Optical Data processor of the European Space Agency (ODESA) MEGS® processor [9]: the prototype MERIS processor that's used to perform MERIS reprocessing activities.

All these applications were made available via the Data Challenge GitHub [10] organization repositories; used so that participants could potentially fork applications and link to their own GitHub repositories. In addition, participants would also have the possibility to integrate their own processing tools as long as they supported the extraction of a 3 by 3 kernel using ESA BEAM Toolbox PixEx operator or any other tool with a similar output format.

2.3. Participants Scenarios

As this was the first test of the E-CEO platform, a series of application scenarios were setup where different participants were created by varying the AC processor or how it was run i.e. varying the auxiliary data (meteorological information that includes mean sea level atmospheric pressure, surface wind speed, relative Humidity and total column ozone) and the provided Aerosol Optical Depth (AOD) for the SMAC processor.

The six participants were:

- A - BEAM (SMAC) with European Centre for Medium-Range Weather Forecasts (ECMWF) auxiliary data, held within the MERIS L1 file, CONTINENTAL aerosol and AOD of 0.1
- B - BEAM (SMAC) with ECMWF auxiliary data, CONTINENTAL aerosol and AOD of 0.2
- C - BEAM (SMAC) with ECMWF auxiliary data, CONTINENTAL aerosol and AOD of 0.3
- D - SeaDAS with climatological auxiliary data
- E - SeaDAS with National Centers for Environmental Prediction (NCEP) auxiliary data
- F - MEGS with ECMWF auxiliary data

2.4. Evaluation Data Packages

Three sites from the NASA NOMAD dataset [4], which are used for the vicarious calibration and validation of ocean colour data, were selected:

- "Aqua Alta" Oceanographic Tower (AAOT, 45°N 12°E) in the northern Adriatic Sea.
- BOUSSOLE (BOUSS, 43°N 8°E) buoy in the Ligurian Sea, one of the sub-basins of the Western Mediterranean sea.
- MOBY (Marine Optical BuoY, 21°N 157°W) moored off of the island of Lanai in Hawaii.

Then, 371 Envisat MERIS FRS scenes were selected to derive the statistics for the evaluation metrics and subsequent ranking. Plus a sub-set of twenty were included in a Data Package for computing software performance metrics.

2.5. Evaluation Process

This evaluation was based on a decision tree of evaluation criteria (Fig. 1), with three main branches that encapsulated the main criteria for the evaluation. The approach novelty included parallelization of the code, scalability & algorithm complexity, programming language and classes of algorithm. It was decided not to include this branch for this contest as pre-existing algorithms were being tested.

Figure 1. Top level evaluation criteria; taken from [11].

2.6. Evaluation Criteria

Therefore, the evaluation was undertaken in two steps, with the first step targeting the retrieval of the computing metrics. This was evaluated on the platform itself, as the participants' software was run within the E-CEO Evaluation environment. The criteria included the central processing unit (CPU) load, processing time, Random-access memory (RAM) usage and disk usage as specified within the E-CEO Evaluation Method technical note [11].

All participants' applications were ran using the twenty MERIS FRS Data Package to produce the MERIS L2 products. The applications were executed in the L2 production mode, with the computing metrics automatically published in the E-CEO Web Portal [12] for normalization and ranking.

The second step targeted the extraction of a 3 by 3 kernel of the mean reflectances (for 412, 443, 490, 510 and 560 nm) for the evaluation sites. To address this extraction, the participants' applications were ran against the 371 MERIS FRS products in an "evaluation" mode.

The Evaluator used an IPython notebook [13], now called Jupyter and with support for multiple languages, to derive the final R correlation coefficient (r_{final}) following the process below, with Fig. 2 showing what the IPython output looks like within GitHub:

1. For each waveband (412, 443, 490, 510 and 560 nm) and evaluation site (AAOT, BOUSS and MOBY) calculate the Pearson correlation factors.

2. For each waveband aggregate the sites' Pearson correlation factors using Eq. 1.

$$r_{cx} = \frac{\left(r_{AAOT\_x} + r_{BOUSS\_x} + r_{MOBY\_x}\right)}{3} \qquad (1)$$

Where x is the wavelength of each waveband

3. Derive the combined spectral Pearson correlation factor using Eq. 2 where the weightings have been calculated from the NOMAD dataset i.e. how good a correlation can be expected for each waveband.

$$r_{final} = \frac{\left(\left(r_{c412} \cdot 0.56\right) + \left(r_{c443} \cdot 0.73\right) + \left(r_{c490} \cdot 0.71\right) + \left(r_{c510} \cdot 0.36\right) + \left(r_{c560} \cdot 0.01\right)\right)}{\left(0.56 + 0.73 + 0.71 + 0.36 + 0.01\right)}$$

(2)

Figure 2. Example of the IPython output within GitHub.

The combined spectral r_{final} value was then manually inserted in the E-CEO Normalization and Ranking tool.

2.7. Criteria Weights

The philosophy used for E-CEO weighting was to elicit weights for criteria through five semantic terms [14], and then in the background they are transformed into numerical classifications that included a penalization or reward depending on the satisfaction of the respective criterion. For example, if a criterion is very important (lower to upper assigned weighting range is 0.8 to 1.0), but the normalised score for a participant is low (0.1) then the final score for that participant would be 0.82:

$$Final\ score = weighting_{lower} + \left(weighting_{upper} - weighting_{lower}\right) * score$$

(3)

If the normalised score for a different participant is high (0.9) then their final score will be 0.98. If, however, this criteria was only classed as of average importance (weighting range is 0.4 to 0.6) then a high score of 0.9 would result in a final (weighted) score of 0.58.

The weights were decided in advance of the evaluation being run, with a strong focus on the model / algorithm performance against the in-situ data for this Challenge.

3. RESULTS

Once the evaluation has been completed for each participant, the resulting scores are shown within the Evaluation results section of the E-CEO Web Portal [15] as the component scores for the individual participant (Fig. 3 Top) and overall set of scores for all participants (Fig. 3 Bottom).

3.1. Validation of the Pearson Correlation Component

To check whether the results produced by the E-CEO platform were correct, the output MERIS and in-situ text files have been read into Interactive Data Language (IDL) code and a Pearson correlation performed. For the IDL code, the filtering removes all the matchups where there is a zero value (missing data) for either the MERIS or in-situ data.

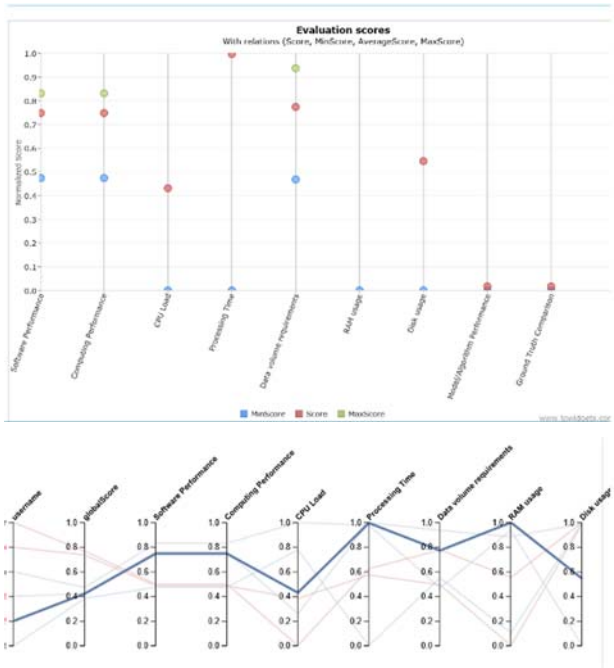

Figure 3. Individual scoring for participant a (top) and final ranking of the participants (bottom), with participant a being the highlighted one.

The results show reasonable agreement to those from the E-CEO platform, see Tab. 1. The differences from the platform were caused by the filtering approach used, but were not significant enough to influence the conclusion - that the platform was implemented correctly and that there was a difficulty in using automatically calculated statistics without intelligent filtering of the data also being applied.

Table 1: Correlation statistics calculated using the IDL implementation for participant a

site / waveband	412 nm	443 nm	490 nm	510 nm	560 nm
AAOT	0.008	0.164	0.371	-nan	0.468
BOUSSOLE	-0.145	-0.124	-0.085	0.120	0.212
MOBY	0.112	0.091	0.088	0.081	0.088

A value of at least 0.7 is normally needed for a correlation to be considered meaningful, and a positive correlation would be expected between the MERIS and in-situ data. The values for the different wavebands shown in Tab. 1 (results are for participant a) include both positive and negative correlations, with the highest value being 0.468 for the IDL code; for the platform it was 0.493, and for both it was the AAOT correlation at 560 nm that's the highest.

As shown in Fig. 4 (Left column), there were a number of outliers for especially the BOUSSOLE and MOBY match-ups that result in very poor correlation results across all wavebands. Therefore, an advanced method of filtering was applied: calculating the mean and standard deviation (stdev) of the remaining MERIS data, and then removing any points lower than the mean - stdev or higher than the mean + stdev; resulting in the plots shown in Fig. 4 (Right column).

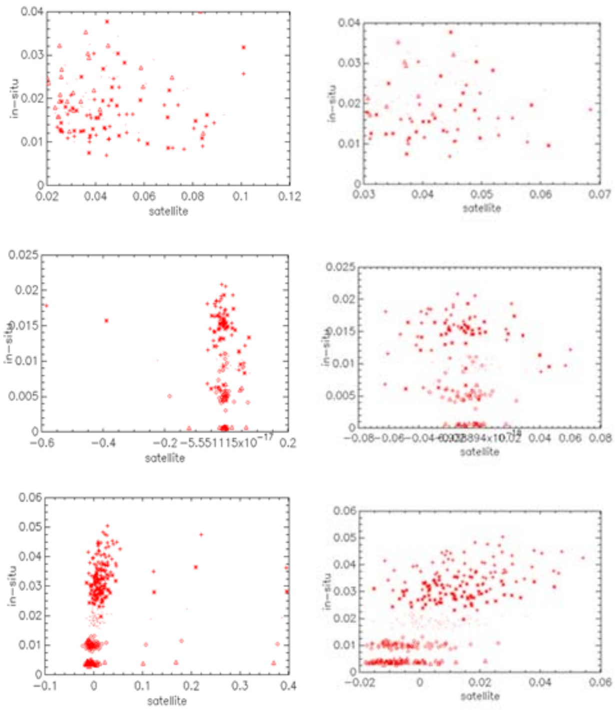

Figure 4. Multi-waveband correlation plots (the different symbols are the different wavebands) as the original results (Left) and after the removal of outliers (Right) for the AAOT (Top row), BOUSSOLE (Middle row) and MOBY (Bottom row) match-ups for participant a.

For participant a, the advanced filtering has improved the AAOT and overall correlation statistics, but some of the waveband datasets have been degraded (might be correctly so). The approach was also tested for participant f (the MERIS processor) with the results, and also correlations, improving much more consistently. This was to be expected as the MERIS processor was designed for above-water application, and has been developed and validated using these in-situ datasets, while the SMAC processor (with its continental aerosol) is designed for above-land application.

The r_{final} correlation coefficient for participant a was 0.135 from the platform and 0.065 for the IDL code before filtering; and 0.078 after filtering. For participant f it was 0.478 for the IDL code with filtering, and 0.267 from the platform.

4. CONCLUSIONS

The results showed that, overall, a collaborative platform can be used to simplify the process of comparing different processors. Once the different participant's software is uploaded and connected to the data packages, the processing runs automatically and collects information about the computing performance. Analysing the satellite versus in-situ matchups is quick to run and adjust.

The results from the IDL versus platform correlation demonstrate that there can be issues in using automatically calculated statistics without intelligent filtering of the data to remove outliers. Applying the advanced filtering resulted in the MERIS processor being the best participant for this component of the performance assessment, which was not the case when the unfiltered data were used. Therefore, it's important to continue to have human involvement in this step.

5. REFERENCES

1. Ruddick, K., Brockmann, C., Doerffer, R., Lee, Z., Brotas, V., Fomferra, N., Groom, S., Krasemann, H., Martinez-Vicente, V., Sa, C., Santer, R., Sathyendranath, S., Stelzer, K. & Pinnock, S. (2010). The CoastColour project regional algorithm round robin exercise. In. SPIE Vol. 7858 *'Asia-Pacific Remote Sensing conference'* (Ed. R. Frouin), Songdo Convensia Incheon, Republic of Korea, pp785807-1 - 785807-13.

2. IOCCG (2010). *Atmospheric Correction for Remotely-Sensed Ocean-Colour Products*. In reports of the International Ocean-Colour Coordinating Group, No. 10, (Ed. M. Wang), IOCCG, Dartmouth, Canada.

3. Jamet, C. (2016). Intercomparison of Atmospheric Correction Algorithms over Optically-Complex Waters at *http://www.ioccg.org/groups/atm_corr.html* (as of 3 May 2016).

4. Werdell, P.J. & Bailey, S.W. (2005). An improved bio-optical data set for ocean color algorithm development and satellite data product validation. *Remote Sensing of Environment*, **98**(1), 122-140.

5. Brockmann Consult (2016). ESA BEAM Toolbox at *http://www.brockmann-consult.de/cms/web/beam/* (as of 3 May 2016).

6. Rahman, H. & Dedieu, G. (1994). SMAC: a simplified method for the atmospheric correction of satellite measurements in the solar spectrum, *Int. J. Remote Sensing*, **15**(1), 123-143.

7. Fu, G., Baith, K.S. & McClain, C.R. (1998). The SeaWiFS Data Analysis System, *Proceedings of the 4th Pacific Ocean Remote Sensing Conference* (Qingdao, China, July 1998), pp73-79.

8. Bailey, S.W., Franz, B.A. & Werdell, P.J. (2010). Estimation of near-infrared water-leaving reflectance for satellite ocean color data processing. *Optics Express*, **18**, 7521–7527.

9. ACRI-ST. (2012). ODESA software distribution. *Quick start guide* at (as of 3 May 2016).

10. GitHub (2016). GitHub website. Online at *https://github.com/* (as of 3 May 2016).

11. Falcao, A.J., Ribeiro, R.A., Jassbi, J., Lavender, S., Boissier, E. & Brito, F. (Accepted for publication 2016). Evaluation support system for open challenges on Earth Observation topics. In *Real-World Decision Support Systems Case Studies*, Springer.

12. E-CEO (2016). E-CEO Web Portal. Online at *https://challenges.esa.int/* (as of 3 May 2016).

13. IPython (2016). IPython website. Online at *https://ipython.org/* (as of 3 May 2016).

14. Ribeiro, R.A. & Pereira, R.A.M. (2003). Generalized mixture operators using weighting functions: A comparative study with WA an OWA. *European Journal of operational research*, **145**, 329-342.

6. ACKNOWLEDGEMENTS

We acknowledge ESA and ACRI-ST for developing ODESA (http://earth.eo.esa.int/odesa), which includes MEGS, and ESA for collecting, processing and making available the MERIS dataset. NASA have led the development of SeaDAS and the NOMAD database, which is available thanks to NASA and the PIs for collecting the in-situ data.

ESA funded the E-CEO project under the Research Service Support activities (SOW TRPE-GSEG-EOPG-SW-12-0001), with Terradue UK Ltd as the prime.