

# CREATION OF HIGH RESOLUTION SUSPENDED PARTICULATE MATTER DATA IN THE NORTH SEA FROM SENTINEL-2 AND SENTINEL-3 DATA.

*A. Alvera-Azcárate, A. Barth, C. Troupin, J.-M. Beckers*

*D. Van der Zande*

GHER  
University of Liège  
Liège  
Belgium

RBINS,  
Direction Natural Environment  
Belgium

## ABSTRACT

A combined Sentinel-2 and Sentinel-3 Suspended Particulate Matter reconstruction is performed using an Empirical Orthogonal Function technique, called DINEOF (Data Interpolating Empirical Orthogonal Functions). The combination of these two datasets allows us to retain both the high spatial resolution of the Sentinel-2 data while increasing the temporal resolution thanks to the addition of Sentinel-3 data on days when no Sentinel-2 data are available. Results show an increased variability on the reconstruction of Sentinel-3 data, and a low error of the overall reconstruction.

**Index Terms**— Ocean Suspended Particulate Matter, Multisensor synergy, Missing data reconstruction, DINEOF

## 1. INTRODUCTION

High resolution ocean data, like those produced by the Sentinel-2 (S2) satellites, provide important information about ocean dynamics and water quality in the coastal zones. However, the narrow swaths of these satellites result in a long repeat cycle (5 days for the tandem S2 A and B). Given the transient nature of small scale processes typical of coastal regions, a combined use with other satellite data of lower spatial resolution but shorter temporal resolution becomes necessary in order to resolve these temporal changes, at the cost of losing the high-resolution information. Information from medium resolution ocean colour sensors like OLCI (Ocean and Land Colour Instrument) onboard Sentinel-3 (S3) can be used in days with no high spatial resolution data. Both S2 and S3 data are affected by clouds, so even in a combined dataset clouds are a major source of missing data.

In this work we propose an approach to perform a combined reconstruction of S2 and S3 data using DINEOF (Data Interpolating Empirical Orthogonal Functions, [1]). The aim is to obtain both a high spatial and temporal resolution data set, with no missing data. The approach is tested in the Belgian Coastal Zone.

## 2. DATA USED

S3 OLCI L1 full resolution data, with a spatial resolution of 300 m, were obtained from the CREODIAS infrastructure. First the IdePix software (v2.2.10), available on the Sentinel Application Platform (SNAP, v7.0), is used for detection of pixels containing land, clouds, cloud shadows, floating vegetation and other elements which can affect accurate retrieval of remote sensing reflectance products. Next, C2RCC [2, 3] is applied to remove the atmospheric signal from the top of atmosphere signal to retain the primary ocean colour variable: the remote sensing reflectance ( $R_{rs}$ ) for the water pixels. From this variable it is possible to derive bio-geophysical products such as SPM and chlorophyll-a concentrations. The SPM product is generated following [4] and the CHL product using [5].

L1C data from the Multispectral Instrument (MSI) onboard S2 have been processed similarly to the S3 processing chain to ensure product comparability. C2RCC (SNAP v7.0) was used to generate L2 remote sensing reflectance products. Suspended Particulate matter (SPM,  $\text{gm}^{-3}$ ) at 10 m resolution were calculated using the Nechad algorithm [4]. Cloud shadows in S2 SPM data have been detected and removed following [6].

In order to maximise the similarity between both datasets, identical atmospheric correction procedures and ocean colour algorithms were applied to both the S2 and S3 products to ensure maximal comparability. Unavoidable differences between both products are however still possible due to the difference in spectral band characteristics and acquisition time.

## 3. METHOD

To construct the data matrix a combination of S2 and S3 scenes is done, with S2 data used as the default. On days

where a S2 pass is not available, data from S3 are interpolated to the S2 spatial grid. If S3 data are too cloudy, with less than 2% of valid pixels, then that day is not included in the data matrix. The initial dataset is therefore daily except for those days where no S2 nor S3 data were available, and with the spatial resolution of the S2 dataset. The use of the two data streams in an alternate way allowed for maximising the use of the high resolution S2 data while not having temporal gaps due to their lower revisit time. In the Belgian coastal zone, revisit time is of about 2-3 days and therefore the gap between two S2 images is not very large. As mentioned, S3 data interpolated onto the S2 grid were used for those days, but of course these interpolated data do not resolve features at the S2 resolution.

A DINEOF reconstruction is then performed with the S2-S3 combined SPM data matrix. DINEOF uses a truncated EOF basis calculated iteratively to infer the missing data [1, 7]. The optimal number of EOFs is determined by cross-validation, by marking as missing about 2% of initially valid data. At each EOF mode calculation, the error between the DINEOF reconstruction and the cross-validation data is calculated, and the number of modes that minimises this error is retained for the final reconstruction. DINEOF is therefore a data-driven approach, in which the information necessary for retrieving the missing information is obtained from the data through the EOF basis. Temporal coherence in the data is reinforced by using a Laplacian filter developed for DINEOF [8].

#### 4. RESULTS AND CROSS-VALIDATION

Several DINEOF tests have been made with different filter lengths. Varying this filter has an impact in the number of EOFs retained [8], and therefore in the variability of the final reconstruction. The results from the best run are discussed here. For the present work, data from 5 February to 28 May 2018 on the Belgian Coastal zone are used, for a matrix size of  $3501 \times 4261$  (in space) and 55 temporal steps. The best reconstruction in terms of the cross validation comes from a 5-EOF basis, obtained with a temporal filter of 1.1 days. A cross-validation is performed following [9] to determine the RMS error of the reconstruction, by hiding initially available data in the form of clouds. This approach provides a better estimate of the reconstruction error than taking randomly-distributed cross-validation points. The cross-validation error of the final reconstruction was of 1.3 mg/l.

An example of the gap-free reconstruction data is shown in Figure 1 along with the initial data. The reconstruction is showing a good amount of spatial variability and realistic SPM distribution, based on the original data displayed. The initial data in this example are interpolated S3, and as can be

seen in the north-south transect shown in Figure 1, the smooth initial S3 data have been reconstructed with a higher variability, effectively increasing the spatial resolution of the data. This result is obtained because the EOFs used to reconstruct the missing data have retained the small-scale variability in the S2 frames, and this has been therefore used to reconstruct the S3 frames. Similar results were obtained at different time frames with initially S3 data.

#### 5. CONCLUSIONS

A joint Sentinel-2 (S2) and Sentinel-3 (S3) Suspended Particulate Matter (SPM) DINEOF analysis has been realised in order to remove missing data due to clouds (among other causes). The use of both data streams allowed us to maximise the use of the high spatial resolution S2 data while also minimising the temporal gaps due to the lower revisit time of S2. The results show that not only the reconstruction realistically represents the main patterns of SPM variability in the Belgian Coastal Zone and the small-scale patterns, but also that the initially smooth S3 data are reconstructed with an increased spatial variability, therefore generating a final product with high spatial and temporal resolution.

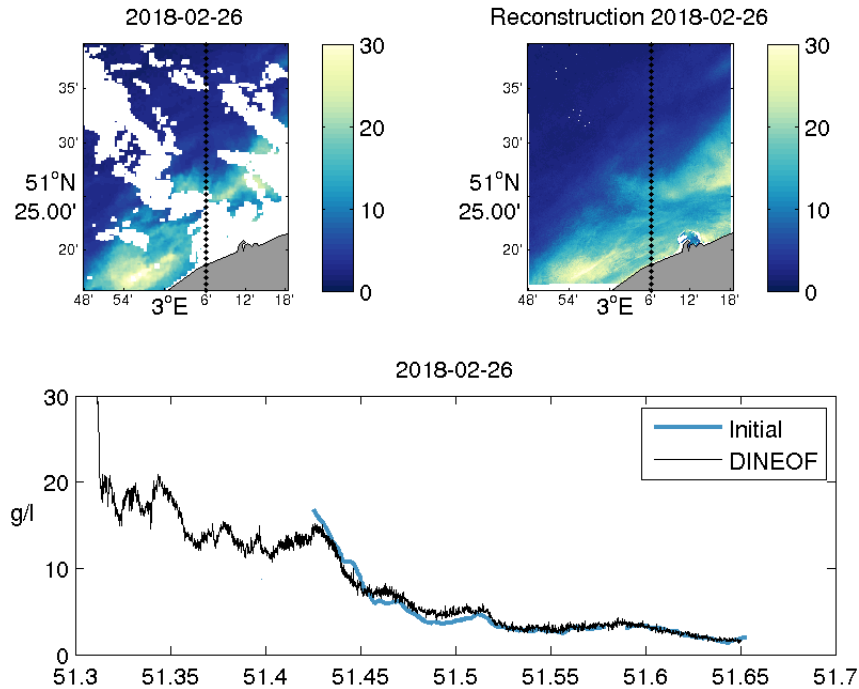
Future work needs to assess how long the S2 temporal gaps can be (up to 5 days in the Equator) so as to still obtain an increase in the final spatial resolution of the dataset. A thorough validation of the results needs to be done, by assessing the accuracy of the results, and also the effective spatial resolution achieved by the DINEOF interpolation. The proposed approach can be also extended to other variables like chlorophyll-a concentration.

#### 6. ACKNOWLEDGEMENTS

This research was performed with funding from the Belgian Science Policy Office (BELSPO) STEREO programme in the framework of the MULTI-SYNC project (contract SR/00/359). Computational resources have been provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifique de Belgique (F.R.S.-FNRS) under Grant No. 2.5020.11 and by the Walloon Region.

#### 7. REFERENCES

- [1] A. Alvera-Azcárate, A. Barth, M. Rixen, and J.-M. Beckers, "Reconstruction of incomplete oceanographic data sets using Empirical Orthogonal Functions. Application to the Adriatic Sea surface tempera-



**Fig. 1.** Top left: initial SPM data from S3 interpolated to the S2 grid. Top right: DINEOF reconstruction. Bottom: transect of SPM from the two datasets in the position marked by the dotted line in the top panels.

ture.” *Ocean Modelling*, vol. 9, pp. 325–346, 2005, doi:10.1016/j.ocemod.2004.08.001.

- [2] C. Brockmann, R. Doerffer, M. Peters, K. Stelzer, S. Embacher, and A. Ruescas, “Evolution of the C2RCC neural network for Sentinel 2 and 3 for the retrieval of ocean colour products in normal and extreme optically complex waters,” pp. Living Planet Symposium ESA, SP-470.
- [3] R. Doerffer and H. Schiller, “The MERIS Case 2 water algorithm,” *International Journal of Remote Sensing*, vol. 28, no. 3-4, pp. 517–535, 2007.
- [4] B. Nechad, K. Ruddick, and Y. Park, “Calibration and validation of a generic multisensor algorithm for mapping of total suspended matter in turbid waters,” *Remote Sensing of Environment*, vol. 114, no. 4, pp. 854 – 866, 2010.
- [5] H. Lavigne, D. Van der Zande, K. Ruddick, J. Cardoso Dos Santos, F. Gohin, V. Brotas, and S. Kratzer, “Quality-control tests for OC4, OC5 and NIR-red satellite chlorophyll-a algorithms applied to coastal waters,” *Remote Sensing of Environment*, 2021, In press.
- [6] A. Alvera-Azcárate, D. Van der Zande, A. Barth, J. F. Cardoso dos Santos, C. Troupin, and J.-M. Beckers, “Detection of shadows in high spatial resolution ocean satellite data using DINEOF,” *Remote Sensing of Environment*, vol. 253, pp. 112229, 2021.
- [7] A. Alvera-Azcárate, Q. Vanhellemont, K. Ruddick, A. Barth, and J.-M. Beckers, “Analysis of high frequency geostationary ocean colour data using DINEOF,” *Estuarine, Coastal and Shelf Science*, vol. 159, pp. 28–36, 2015.
- [8] A. Alvera-Azcárate, A. Barth, D. Sirjacobs, and J.-M. Beckers, “Enhancing temporal correlations in EOF expansions for the reconstruction of missing data using DINEOF,” *Ocean Science*, vol. 5, pp. 475–485, 2009.
- [9] J.-M. Beckers, A. Barth, and A. Alvera-Azcárate, “DINEOF reconstruction of clouded images including error maps. Application to the Sea Surface Temperature around Corsican Island,” *Ocean Science*, vol. 2, no. 2, pp. 183–199, 2006.