Manuscript ID: amt-2022-37

Spectral replacement using machine learning methods for continuous mapping of Geostationary Environment Monitoring Spectrometer (GEMS)

Yeeun Lee, Myoung-Hwan Ahn*, Mina Kang, Mijin Eo

 # General response to referees' comments – major revision

It is highly appreciated for the detailed comments of Glen Jaross to greatly improve the revised manuscript. We tried to reflect the referee's comments as much as possible and the manuscript has been reorganized to take better into account the referee's comments. More specifically, the revised version includes the impact of ANN-filled radiance to the Level 2 products such as cloud and ozone. One thing to be noted is that the data used in this paper have been updated with the operational data (Level 1C) for the analysis of retrieval results. In this reply, the referee's comments are repeated in blue, our responses to the specific comment are given in red, and the revised manuscript is presented in *italic*.

Specific comments from Dr. Glen Jaross

#1. The authors have partially addressed my concerns expressed after my initial review. In their Version 2 they offer a quantitative evaluation of ANN radiance errors, which addresses one of my concerns.

⇨ Thanks a lot.

I repeat my earlier criticism that **the authors have not clearly stated the objective of this paper**. In the introduction the authors identify the fundamental problem they are trying to address: Level 2 products have spatial gaps due to bad pixels. Logically, the next step is to ask the question: what is the best approach to filling these gaps? The authors do not ask this question but instead proceed to discuss a radiance replacement approach in the Level 1 product. In my opinion there are multiple solutions that will better fill Level 2 gaps than radiance replacement in the Level 1 product. The authors should acknowledge there are alternative solutions to this problem and tell the readers this paper describes the investigation of one technique.

⇨ Okay. We acknowledged the objective of the study and its justification was not fully explained in the previous version of the manuscript. The starting point of this current work was to mitigate aesthetic annoyance that the released image would produce but we admitted that ultimately, filling the gaps in Level 2 products would be the final goal for the issue. It is true that the annoyance could be mitigated by other approaches such as simple interpolation of the Level 2 products (would be very efficient for the products having a smooth spatial variation). However, we wanted to touch upon a bit more the fundamental issue, "could the machine learning approach provide a new insight on the missing radiances caused by the bad pixels?" We didn't expect any algorithm would provide full content of information that only the actual measurements data could provide. However, some cases, the available nearby data (either in space or in spectral domain) could fill the gap introduced by the bad pixels with sufficient accuracy that could be used for further application. Furthermore, the carefully prepared machine learning algorithm could provide some information even on the rather complicated spectral range (as described below).

⇨ Reflecting our response, the revised part of manuscript is (or are);

Section 1 Introduction (Lines 50-64)

Especially, when a scene on the Earth dramatically changes, discontinuity caused by the interpolation becomes larger. This effect causes spatial discontinuity in Level 1B data and retrieved properties (Level 2) by affecting retrieval processes with contaminated spectral features.

As a way of filling in the spatial gaps, this study approaches the underlying problem by focusing on radiances with spectral replacement using machine learning methods. The spatial gaps found in Level 2 data can be filled in with various methods (e.g. variogram, empirical orthogonal functions or mathematical filters) and for each Level 2 product, there will be a more suitable method using multiple sources of information and distribution characteristics (Fang et al., 2008; Guo et al., 2015; Katzfuss and Cressie, 2011; Llamas et al., 2020; Yang et al., 2021). In this regard, this research places more emphasis on efficiency and further application of the approach because improving erroneous spectral features can be an efficient way to solve the issue for all products and also has the potential to be applied to various measurement issues of hyperspectral data. For that, further questions to be investigated here are whether non-linear relations could be accurately emulated with machine learning methods and input radiances have valid information for retrieval processes. For the investigation, cloud and ozone retrievals are performed with the reproduced spectra of GEMS to evaluate the effectiveness of the suggested approach and its limitations.

\

#2. The authors attempt a quantitative evaluation of the replacement radiances for the GEMS pixel gaps, but the reader is left wondering how important the residual errors are. A correlation coefficient of 0.82 in PC2 or PC3 doesn't sound good, but what does that really mean in terms of product performance? The paper contains some discussion regarding the Ring Effect, but otherwise little is said about how useful ANN-based radiances are. The very big question remaining is, is there any improvement in Level 2 performance and does that performance warrant further pursuing the ANN technique? The authors may argue that this is the subject of a future paper. If so, this should be made clear in both the stated objectives and in the conclusions.

⇨ The straightforward approach to answer the usefulness of ANN-based radiances is analyzing the impact of reproduced spectra on the Level 2 products. As we suggested in the previous reply and the referee pointed out, the task will incur a new research work which will certainly be carried out in the near future. However, to check the possibility and ensure future work, we analyzed the effects of spectral replacement for the Level 2 retrieval processes (in Sect. 3.3.1) as suggested, especially for ozone and cloud properties with the help of the GEMS Level 2 algorithm team. Considering that the PCA results still provide the information on how qualitatively successful the spectral replacement is for different spectral regions in terms of spectral relations, the PCA part is reorganized in Sect. 3.2.2 to support the retrieval results presented in the following section.

⇨ Reflecting our response, the revised parts of the manuscript are;

Section 3.2.2 PCA-based analysis (Lines: 294-305)

As presented in Table 4, comparing PC scores provides qualitative information on the effectiveness of the suggested method. The results show that the mean spectral pattern (the first PC) and some dominant patterns can be sufficiently reproduced with the suggested models, but other spectral features such as the third PC for Defect 1 or the second PC for Defect 2 have difficulty obtaining valid information from input radiances for accurate reproduction. The interesting finding is that only Defect 3 shows high correlation coefficients over 0.95 for the leading PCs having higher explained variance ratios. Each PC except the first one may contribute to a small extent to total radiances given the explained variance ratio. However, it also could be enough to determine subtle spectral patterns important for retrieval processes. The effectiveness of spectral replacement for each spectral region could be glimpsed in the results, which will be discussed further in the following section with retrieval processes.

Table 1 *Correlation coefficients of PC scores of GEMS and ML radiances and explained variance ratios of GEMS radiances for each target region in Fig. 8-10.*

Defects	Factor	PC 01	PC 02	PC 03	PC 04	PC 05	PC 06
1	Correlation coefficient	0.9999	0.9976	0.8172	0.9779	0.6846	0.6609
2		0.9999	0.8129	0.9876	0.4294	0.7035	0.5046
3		0.9999	0.9962	0.9787	0.6644	0.5399	0.2649
1	Explained variance ratio [%]	99.9905	0.0071	0.0007	0.0006	0.0001	0.0001
2		99.9524	0.0268	0.0141	0.0019	0.0012	0.0005
3		99.9954	0.0038	0.0003	0.0001	0.0001	0.0001

Section 3.3.1 Cloud and ozone retrieval (Lines 307-330)

In the previous section, it was found that the overall prediction error is about 5% except for the ozone absorption lines and dominant spectral patterns can be successfully reproduced with the suggested method. The next question to be discussed is whether the reproduced spectral features are applicable to the retrieval process. Even if the trained models accurately reproduce an absolute value at each wavelength, the Level 2 retrieval could be unsuccessful if non-linear relations are *too elusive to be properly emulated with the model. The radiances at O2-O2 absorption lines related to Defect 3 has the smallest prediction error of 0.5% and we checked that cloud information with the fitting window in 460.2-490.0 nm can be successfully retrieved with the reproduced spectra in Fig. 8. The difference of cloud centroid pressure retrieved with ML and GEMS spectra is about 1% on average for normal measurements but the cloud properties retrieved with ML spectra have weak stripped features. The spectral range of Defect 3 is very narrow and thus the input radiances provide enough information for successful spectral replacement and the retrieval process.*

The replaced radiances at ozone absorption lines showed high prediction error in the previous section. For the qualitative investigation of the effect, the reproduced spectra presented in Fig. 10 are applied to the ozone retrieval of GEMS. Figure 13 shows total ozone column density with unlagged bad pixel area for the comparison of spatial discontinuity. As previously mentioned in Sect. 2.1.2, the ozone properties retrieved with measured GEMS spectra show distinct spatial discontinuity over the bad pixel area as shown in Fig. 13a and the discontinuity is somewhat reduced in Fig. 13b with ML spectra. However, the retrieved properties show different spatial distribution patterns when comparing the surrounding areas which have true measurements. It seems the ozone properties are underestimated especially for higher radiances in Fig. 13b and the stripping features found in Fig. 10 may affect the retrieval process considering the features are also found in Fig. 13b. It is also clear that the angle conditions provide important information for the retrieval because without the conditions, the retrieval results show unrealistic features with higher variance. The results indicate that the spatial distribution can be approximately matched for the ozone properties retrieved with measured (true) and reproduced spectra, but it seems limitation still remains to get an exact retrieval value.

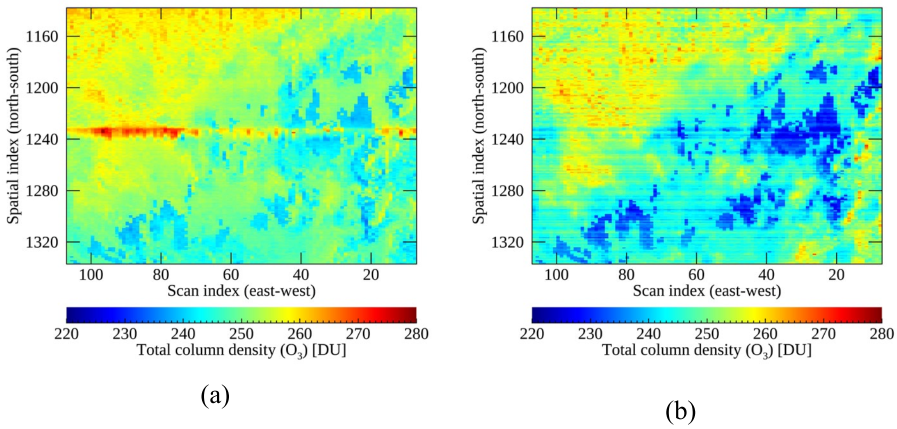


Figure 1 *Spatial distribution of total ozone column density retrieved with (a) GEMS and (b) ML radiances presented in Fig.11. The GEMS spectra are measured on 10 March 2021 (06 UTC).*

#3. My general impression of the paper is it spends too much time considering the specific GEMS data cube gaps, especially in the paper section concerning evaluation. As the authors point out in their conclusions, the radiance replacement approach may be useful for other instruments, and those instruments will almost certainly not share the same gaps as GEMS. The readers would be better served if the authors investigate a variety of data gaps and demonstrate the efficacy of the ANN replacement method in each. The GEMS gaps can be a subset of those tested. Such an approach ties in better with a modified objective of the paper, where the emphasis is the strengths and weaknesses of ANN radiances under a variety of circumstances.

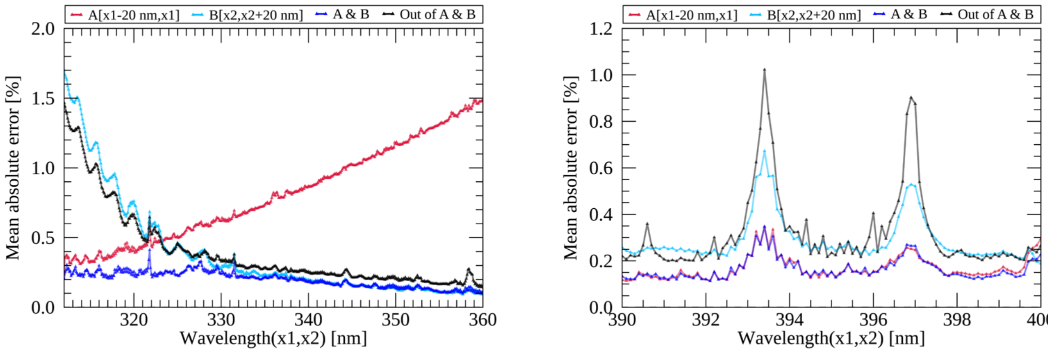
⇨ It is quite an interesting suggestion. By doing so, we may prepare for the bad pixel issues GEMS would face in the coming years. On the other hand, the data gaps identified in the current GEMS data could give representative examples of bad pixels that a hyperspectral instrument would have. One is the complex ozone absorption bands, another one is the rather smooth spectral bands in 400-500 nm, while the other is the rather flat and highly correlated (with neighboring wavelengths) bands. Thus, the characteristics revealed with the three gaps could represent general data gaps we would expect from a hyperspectral instrument in the UV/VIS spectral range. However, we fully agreed that the input conditions were not fully investigated because of the retrained conditions, and thus we have inserted the spectral analysis following the retrieval results for the cause analysis of retrieval results and further application of the method.

⇨ The revised parts of the manuscript are;

3.3.2 Cause analysis for further application (Lines 331-378)

The retrieval uncertainty found in Fig. 13b is attributed to the lack of information in the input data or insufficient model optimization. For Defect 2, the input spectral range (400-500 nm) may have deficient information for ozone properties and it could cause the unsuccessful replacement. To clarify this and investigate further for future applicability, we choose two output cases targeting ozone absorption lines in 312-360 nm and Fraunhofer lines in 390-400 nm to apply the suggested method with different input cases. In the Fraunhofer lines, the Ring effect caused by rotational Raman scattering can be found over two radiance peaks which is generally known to be very small and largely affected by the existence of clouds (Joiner et al., 1995). Together with ozone absorption lines, the analysis results could give a clear evidence on whether the small scattering features could be reproduced with machine learning depending on different input wavelengths. The PCA-ANN model is trained for each input case respectively with high-quality measurements in March 2021 (around 80,000 spectra after bad pixel masking and the elimination of saturation pixels).

Figure 14 presents mean absolute errors for ozone absorption and Fraunhofer lines with different input conditions including or not the closer sides of the output spectral bands (within the range of 20 nm). As assumed, prediction errors increase at the spectral peaks and overall error patterns along the output wavelengths largely differ for each input condition. It is clearly shown that the errors tend to be higher when reproduced with farther input spectral bands for both output spectral lines. The similar input condition (360-500 nm) with Defect 2 is plotted in black lines in Fig. 14a and the results clarify that the insufficient information from the input data may cause large errors for radiances at shorter wavelengths and subsequently the ozone retrieval process. Figure 14 verifies that each input case has a different amount of information which determines the effectiveness of the model to reproduce certain spectral features. It also can be deduced that the method could be quite useful even for strong absorption lines when the input and output spectral ranges are sufficiently close.

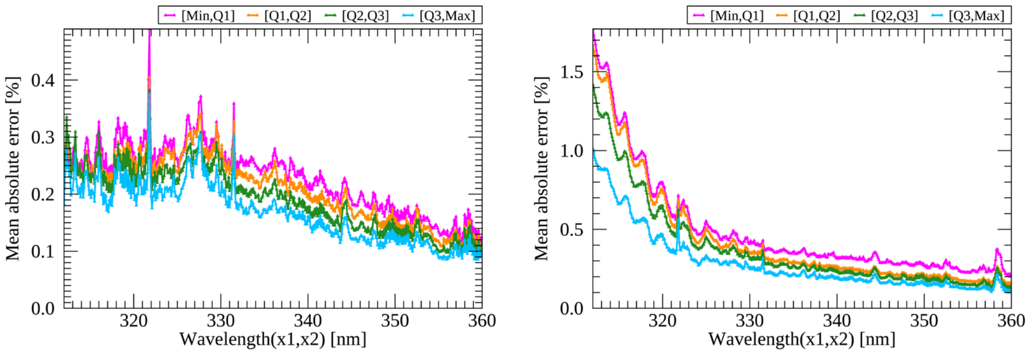


(a)

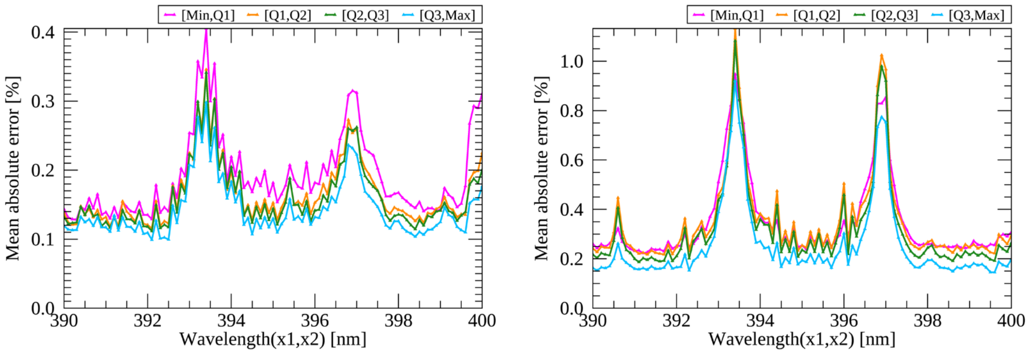
(b)

Figure 2 *Mean absolute errors for the reproduced and measured radiances at (a) ozone absorption and (b) Fraunhofer lines and the x1 and x2 indicates wavelengths at the edges of output spectral bands. The absolute error is calculated between the ML and GEMS radiances divided by the latter in percent.*

Figure 15 presents a closer inspection by dividing spectra into four groups depending on the scene brightness. Different scenes could have different error levels which could be ignored in the averaged values in Fig. 14. The analysis is performed with the spectra reproduced with the input conditions showing the smallest (blue lines) and the largest (black lines) errors in Fig. 14. The PCA-ANN model reproduces dominant spectral features with an error of 0.4% for all scenes with the best input condition including near sides of output spectral bands as shown in Fig. 15, but it seems the difference increases with darker scenes (weak signals). This indicates low signals would be generally less predictable even with the information extracted from the very close wavelengths. The error spectra show more distinguishable spectral features with farther input spectral bands, which shows that the spectral information from the input condition would be insufficient to properly reproduce exact spectral features.



(a)



(b)

Figure 3 *Mean absolute errors for the reproduced and measured radiances at (a) ozone absorption and (b) Fraunhofer lines with different input spectral bands including (the first column) or excluding (the second column) both near sides of output wavelengths within the range of 20-nm. The Q1, Q2 and Q3 represent the first, second and third quartile and each color indicates the average in the range of each quartile. The x1 and x2 indicates wavelengths at the edges of output spectral bands and the absolute error is calculated between the ML and GEMS radiances divided by the latter in percent.*

In this section, different output spectral bands containing absorption or scattering lines are compared with different input conditions. It seems the suggested method (PCA-ANN) could be quite effective when the input spectral ranges are closer to the target wavelengths to be reproduced. However, it is not necessarily true the wider the input spectral range is, the more accurate the replacement becomes. If input spectral ranges have some calibration issues (e.g. stray light or saturation) or provide conflicting features with other input spectral bands as shown in Fig. 14a, the reproduced spectrum would have diverse and inconsistent features causing higher error. In conclusion, the suggested method accurately predicts the overall magnitude of a spectrum, but reproducing a certain spectral feature with high accuracy would need more information especially for low signals or strong absorption lines. At least, the input and output spectral regions should be close enough to reduce the spectral error up to 0.5%, the uncertainty of the reproduced spectra at O2-O2 absorption lines presenting successful cloud retrieval results.

#4. If the objective of the paper is to propose a solution for the GEMS missing pixels, I believe the discussion presented in Version 2 of this paper is incomplete. But if the objective is to describe a radiance replacement technique that is the basis for further investigation, the discussion presented in this current version is appropriate. In that case the authors should spend less time analyzing the specific GEMS data gaps, though they can be discussed as the impetus for the investigations.

⇨ Again thanks for the raising the point. The ultimate goal is to increase usefulness of GEMS data for a longer time period, at least for designed lifetime of 10 years under the new operational environment of geostationary. The current study shows that the gap filling (in level 1) over the longer wavelengths are quite feasible and reliable, while it still haves limitations for strong absorption bands which may provide the reasons why we need actual observation data over such spectral bands. However, accumulation of observation and auxiliary data along with an improved nonlinear algorithm, the shortage of gap filling over the absorption bands could be improved, we hope.

⇨ As introduced in the earlier part of our response, the introduction part as well as overall sections of the manuscript have been revised by relating the resulting problem of bad pixels (i.e. spatial gaps in Level 2) and the method we chose to deal with (i.e. spectral replacement) with the broadened perspective based on the suggestion by applying ozone and cloud retrievals with the reproduced spectra and analyzing spectral gaps with different input conditions.

⇨ Reflecting our response, the revised parts in the manuscript are;

Section 4: Conclusions (Lines: 380-415)

In terms of accuracy, the spatial gaps found in Level 2 products could be better improved when applying a fitted method based on spatial distribution characteristics of each product. In this regard, we more focus on improving the erroneous spectrum itself on the radiance level to check whether the issue could be more efficiently resolved for both radiances and retrieved properties with improved spectral features. For the approach, this study suggests machine learning methods (PCA-ANN and PCA-Linear) to fill in various spectral gaps denoted as Defects 1-3 by investigating how much information could be obtained to reproduce spectral features without any additional information. The basic assumption of this approach is that radiances of a spectrum have strong linear and non-linear relations, which could be emulated with the ANN and multivariate linear regression. The spectral range of output radiances is set to the wavelengths of bad pixels, while the input radiances correspond to the rest part of a spectrum for Defects 1-3, respectively.

In the results, PCA-Linear model presents smaller prediction errors for the defective region having strong linear relations between input and output radiances (Defect 1) or having a narrower spectral gap (Defect 3). When applying the reproduced spectra for Defect 3 to the cloud retrieval, the cloud centroid pressure is successfully retrieved with an error of 1%, on average. This is because the output spectral range of Defect 3 is comparably narrower and thus the input wavelengths provide enough information to reproduce exact spectral features which are valid for the subsequent retrieval process. The PCA-ANN model is better for the output radiances having strong non-linear relations (Defect 2). Dominant spectral patterns and the overall magnitude of spectra could be successfully reproduced mostly with an error of 5% except for ozone absorption lines, while the exact spectral patterns would be insufficiently reproduced. When applying the reproduced spectra to the ozone retrieval, the spatial distribution of total ozone column density can be approximated but with high uncertainty. This indicates that additional information would be needed except for the spectral relations of radiances for the successful retrieval process at strong absorption lines ultimately reducing spatial gaps in the Level 2 products.

Considering that the number of bad pixels would increase in operation as did in Ozone

Mapping and Profiler Suite (OMPS) (Seftor et al., 2014), an efficient way of replacing bad pixels would be necessary for the long-term operation of GEMS. It is also highly possible that an unexpected issue could occur such as the row-anomaly of Ozone Monitoring Instrument (OMI) (Schenkeveld et al., 2017). The ultimate goal of this research is to increase the usefulness of GEMS data for a longer time period, at least for designed lifetime of ten years. The current work verifies that the gap filling (in Level 1) with certain spectral conditions shows quite reliable results even with the limitations for the strong absorption bands, which is natural and provides the reasons why we need observation data over such spectral bands. However, it is also anticipated that accumulation of observation data along with auxiliary data and improved nonlinear algorithm, the limitation could be improved in future study. For that, this paper provides the basis for further applicability of the method by evaluating the efficiency of machine learning methods to reproduce hyperspectral data especially in the UV/VIS spectral range.

Reference

Fang, H., Liang, S., Townshend, J. R. and Dickinson, R. E.: Spatially and temporally continuous LAI data sets based on an integrated filtering method: Examples from North America, Remote Sens. Environ., 112(1), 75–93, doi:10.1016/J.RSE.2006.07.026, 2008.

Guo, L., Lei, L., Zeng, Z. C., Zou, P., Liu, D. and Zhang, B.: Evaluation of spatio-temporal variogram models for mapping Xco2 using satellite observations: A case study in China, IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., 8(1), 376–385, doi:10.1109/JSTARS.2014.2363019, 2015.

Joiner, J., Bhartia, P. K., Cebula, R. P., Hilsenrath, E., McPeters, R. D. and Park, H.: Rotational Raman scattering (Ring effect) in satellite backscatter ultraviolet measurements, Appl. Opt., 34(21), 4513, doi:10.1364/AO.34.004513, 1995.

Katzfuss, M. and Cressie, N.: Spatio-temporal smoothing and EM estimation for massive remote-sensing data sets, J. Time Ser. Anal., 32(4), 430–446, doi:10.1111/J.1467-9892.2011.00732.X, 2011.

Llamas, R. M., Guevara, M., Rorabaugh, D., Taufer, M. and Vargas, R.: Spatial Gap-Filling of ESA CCI Satellite-Derived Soil Moisture Based on Geostatistical Techniques and Multiple Regression, Remote Sens. 2020, Vol. 12, Page 665, 12(4), 665, doi:10.3390/RS12040665, 2020.

Schenkeveld, V. M. E., Jaross, G., Marchenko, S., Haffner, D., Kleipool, Q. L., Rozemeijer, N. C., Veefkind, J. P. and Levelt, P. F.: In-flight performance of the Ozone Monitoring Instrument, Atmos. Meas. Tech., 10(5), 1957–1986, doi:10.5194/amt-10-1957-2017, 2017.

Seftor, C. J., Jaross, G., Kowitt, M., Haken, M., Li, J. and Flynn, L. E.: Postlaunch performance of the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite (OMPS) nadir sensors, J. Geophys. Res. Atmos., 119(7), 4413–4428, doi:10.1002/2013JD020472, 2014.

Yang, M., Khan, F. A., Tian, H. and Liu, Q.: Analysis of the Monthly and Spring-Neap Tidal Variability of Satellite Chlorophyll-a and Total Suspended Matter in a Turbid Coastal Ocean Using the DINEOF Method, Remote Sens. 2021, Vol. 13, Page 632, 13(4), 632, doi:10.3390/RS13040632, 2021.