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

We would like to thank Glen Jaross and the anonymous referee for the insightful comments and the detailed suggestions, which are significant to include some missed points not fully explained in the first draft. We have revised the manuscript considering the suggested comments and the referee's comments are presented in **bold**. In this document, we apply the figure and table numberings of the revised manuscript and the revised manuscript is presented in *italic*.

Specific comments from Dr. Glen Jaross

Comment – Part A-1:

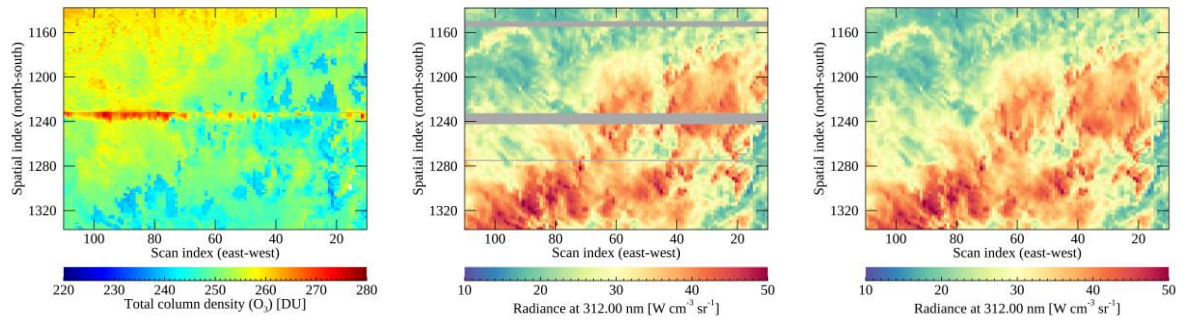
"I have difficulty understanding the goals of this paper. There seems to be a mismatch between what the authors are trying to achieve and the investigation they describe in the paper. The authors describe a problem on GEMS whereby "bad" pixels, i.e. missing pixel radiances, cause problems in the ensemble of measured radiances. … There is no mention of what criteria are used to assess the results."

Response A-1:

We agree with the comments that the goals of this paper are not fully explained. Bad pixels of GEMS could cause distinct spatial discrepancy in GEMS Level 1 data and it definitely introduces errors in Level 2 products (or make impossible to derive the Level 2 products). Thus the primary goal of the current work is to evaluate the applicability of machine learning methods for the spectral gap filling, which could reduce bad pixel effects to both Level 1 and 2 products. For the purpose, bad pixel effects are discussed first (as suggested by the referee) and thus we have added the following part to Section 2.1, describing the effects of bad pixels to radiances and retrieved properties (i.e., ozone):

The interpolation error could seriously affect the Level 2 products of which the spectral fitting windows are overlapped with bad pixel areas. For instance, cloud properties and aerosol effective height (AEH) of GEMS are retrieved from O_2-O_2 absorption bands around 477 nm (Choi et al., 2021; Kim et al., 2021) where the cluster of bad pixels is located (Defect 3). During the IOT, Defect 3 caused spatial discontinuity to the retrieved cloud and AEH distribution, which made the fitting window of the products modified to avoid bad pixel effects. The O_3 retrieval is also affected by Defect 2 (300-400 nm) as the spectral radiances within 300-380 nm provide major information for the O_3 retrieval of GEMS (Bak et al., 2019). The bad pixel effects in the Level 2 product are clearly shown in Fig. 2 which presents radiances at 312 nm and the retrieved total O_3 column of GEMS. Even though radiances at the certain wavelength are homogeneous with its surroundings (see Fig. 2b). The spectral patterns are not properly reproduced with the existed method (spatial interpolation) causing the distinct horizontal line in Fig.

2c.



(a) (b) (c)

Figure 1 *Spatial distribution of (a) the GEMS total O_3 column and radiances at 312 nm with bad pixels (b) marked in dark gray and (c) reproduced with spatial interpolation. The measurements are on 10 March 2021 (06 UTC).*

Comment – Part A-2:

"From an academic perspective the question of how well ML techniques can describe Earth backscattered radiances is an interesting one. If the authors approached this paper from that perspective they might provide a useful contribution to our ability to characterize the atmosphere through numerical techniques. But in doing so they must take a more rigorous approach to evaluating their radiance predictions.

The first thing the authors should do is to forget about the GEMS pixel defects. These are of no use in evaluating the efficacy of the technique since the true radiances remain unknown for these regions. Instead, choose regions of the detector where there are good measurements and treat them as missing for the purpose of deriving errors. There can be a variety of region shapes and sizes, including ones that look very similar to Defects 1, 2, and 3."

Response A-2:

The points are also raised by another referee and we appreciate the suggestions. Following the suggestions, we targeted certain areas including each defect (Defects 1-3) and its surroundings (100-indices toward both north and south directions) where actual measurements (regarded as 'true') could be obtained. With the analysis, it is possible to quantitatively evaluate how the ML radiances differ with the true measurements. In summary, the direct comparison between measured and reproduced radiances shows that the dominant spectral patterns could be successfully reproduced even though the spectral patterns determined by very small signals such as the Ring effect would be insufficiently reproduced with the suggested methods. The spectral patterns are also evaluated by applying PCA to the measured and reproduced spectra, and the extracted spectral patterns are highly correlated for the first PC for Defect 1-3 regardless of the models which are consistent with the direct comparison results.

The results are added to the 'Results and discussion' (Section 3) which has been re-organized with 'Model selection' (Section 3.1) and 'Evaluation' (Section 3.2). The following part has been inserted to Section 3.2:

3.2 Evaluation

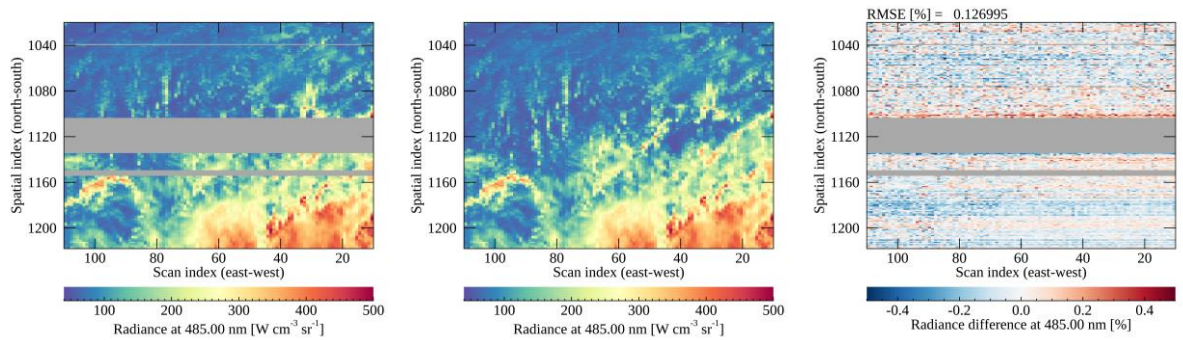
3.2.1 Spatial and spectral inspection

For the quantitative evaluation of the reproduced spectra, certain areas are targeted which include each defect (Defects 1-3) and its surroundings where actual measurements regarded as 'true' could be obtained. The evaluation is made with the data measured on 10 March 2021 (06 UTC), which are excluded for the model training. The center longitude of the areas is set to 128° E, which is identical with the sub-nadir longitude of GK-2B. Table 3 presents spectral ranges of Defects 1-3 and the target wavelengths for the analysis. Specifying the wavelengths for the analysis helps to specifically analyze the spectral patterns of absorption lines of trace gases and cloud properties.

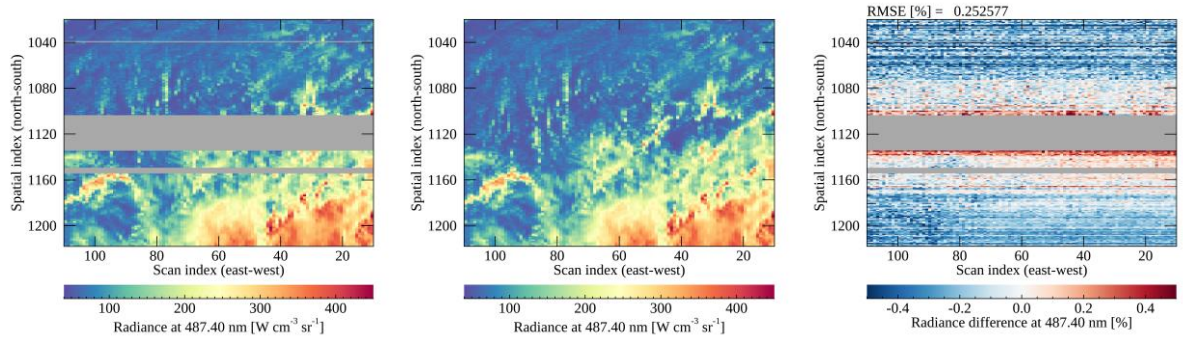
Table 1 *The spectral range of Defects 1-3 and target wavelengths for the analysis. The third column presents GEMS retrieval products of which each fitting window is overlapped with Defects 1-3.*

Defect	*Target wavelength*	*GEMS Level 2 product*	*Optimized model*
1 (400-500 nm)	*432-450 nm*	*CHOCHO, NO_2*	*PCA-Linear*
2 (300-400 nm)	*312-360 nm*	*O_3, HCHO, SO_2, NO_2, aerosol optical depth*	*PCA-ANN*
3 (484-491 nm)	*484-491 nm*	*Cloud, AEH*	*PCA-Linear*

The measured and the reproduced radiances with machine learning methods are directly compared, which are hereafter referred to as GEMS radiances and ML radiances. In Figs. 11-13, each column shows GEMS, ML radiances and the difference while the first and second rows show the representative wavelengths showing the smallest and the largest difference, respectively. Figure 11 shows the comparison results of the Defect 3 area, which presents the best performance compared to the Defect 1 and 2 areas. The difference in Fig. 11 is within $\pm$ 0.5% because the spectral gap of Defect 3 is narrower than the counterparts of Defects 1-2. The narrower the spectral range of the output radiances is, the more abundant information could be obtained from the input radiances. For Defect 3, there is no scene dependence over the output wavelengths and the difference shows noise-like features and the spatial dependence originated from instrument artifacts.



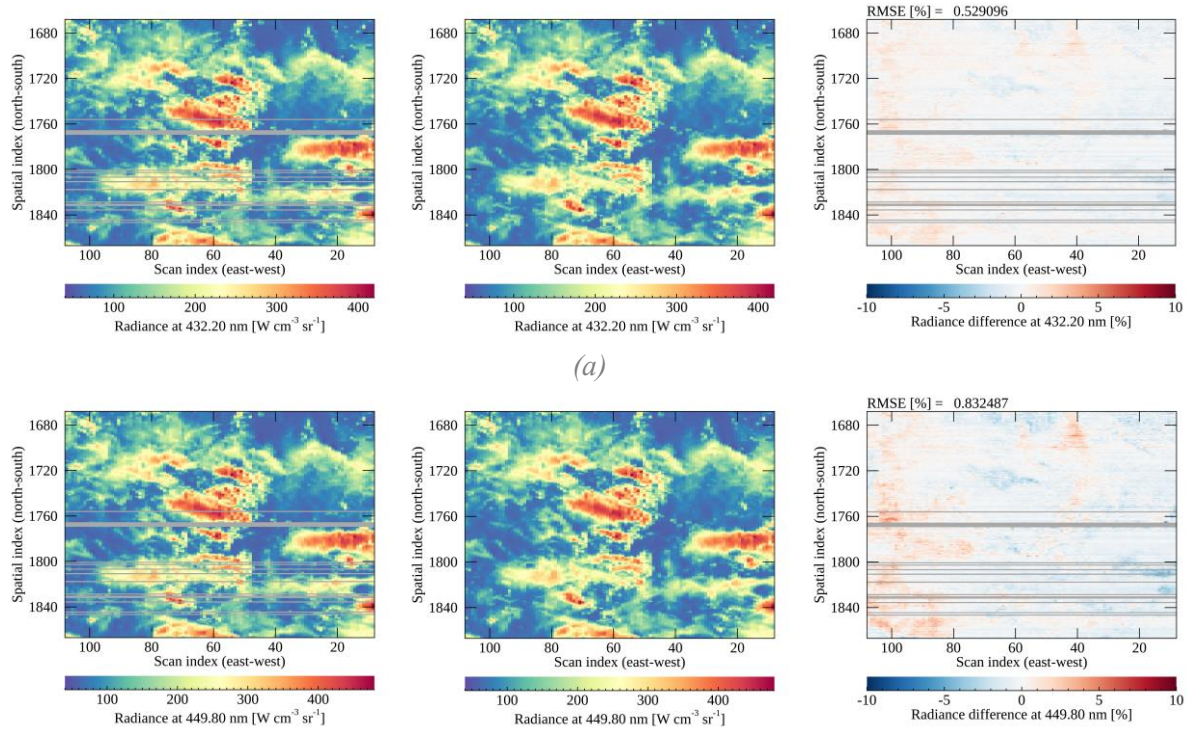
(a)



(b)

Figure 2 *The GEMS, ML radiances and the difference (from left to right) at the wavelengths presenting (a) the smallest and (b) the largest difference for the Defect 3 area. The difference is calculated between the ML and GEMS radiances divided by the latter in percent. Bad pixels are marked in dark gray and the color bar range is* $\pm$ *0.5%. The unit of RMSE is in percent divided by the mean radiance.*

Figure 12 shows the Defect 1 area where the ML radiances are within about 5% of the GEMS radiances. It also shows that dark targets (clear sky with small radiance) show a positive difference while bright targets (mostly cloudy sky with large radiances) show an opposite tendency. The tendencies are also found on the other dates for different angle conditions. It seems the applied machine learning model (PCA-Linear) might have its limitation in describing the non-linear relations of angle conditions, scene properties and radiances causing the difference of about 5%.



(a)

(b)

Figure 3 *Same as Fig. 11 for the Defect 1 area with the color bar range of 10%.*

For the Defect 2 area, it is clear that the information from radiances of wavelengths longer than 400 nm is insufficient to effectively reproduce the spectral features at shorter wavelengths (consistent results with Figs. 8-9). Both output spectral ranges of Defects 2-3 are around 100 nm but it seems the output radiances near 300 nm for Defect 2 need more information to be successfully

reproduced. The stripping feature found in Fig. 13b becomes significant at 312 nm for the ML radiances on the contrary to the radiances at 357.2 nm in Fig.13a. The stripping feature seems to be added during the reproducing process especially for shorter wavelengths, and the reason is still unclear. Another distinct feature found in Fig. 13 is that the difference in northern parts is very large with the difference of 10%. We suspect that the reason might be the VZA effect considering that VZA increases at the northern parts in the area. Without angle conditions as the input parameters for the model, the difference becomes doubled at 312 nm presenting similar patterns with the difference in Fig. 13b. This indicates the angle effect can be emulated in the model by applying VZA and SZA as the input parameters, but it is not fully resolved especially for the radiances at shorter wavelengths.

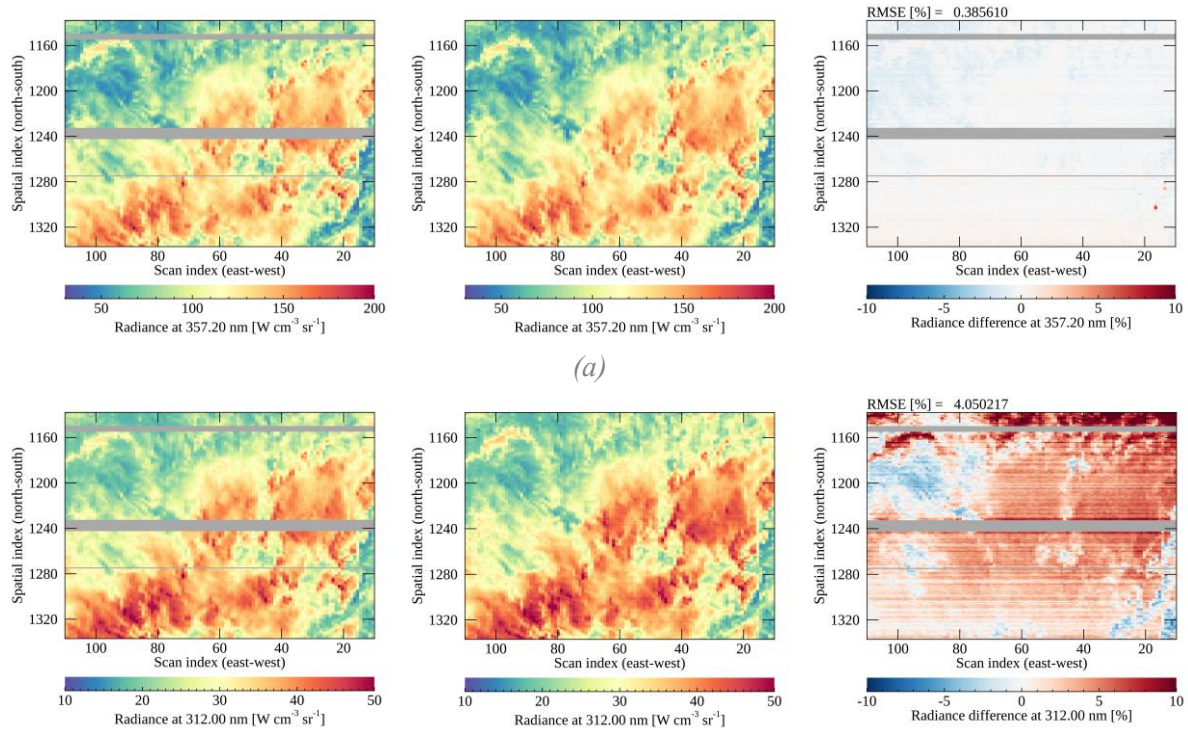
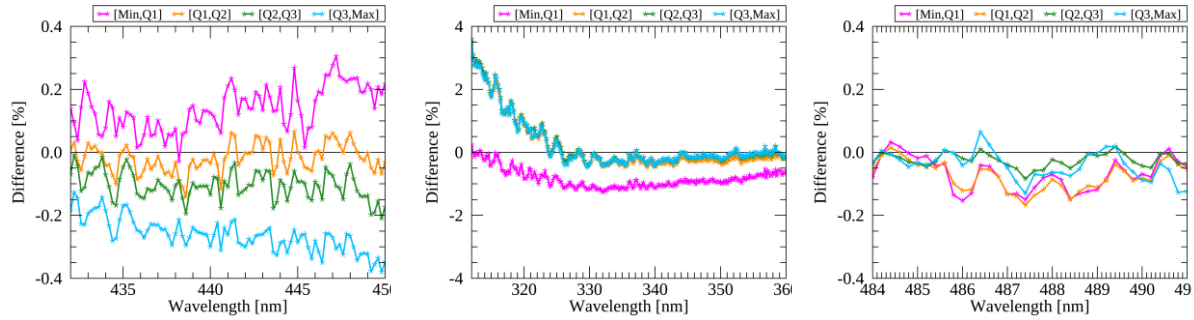


Figure 4 *Same as Fig. 11 for the Defect 2 area with the color bar range of 10%.*

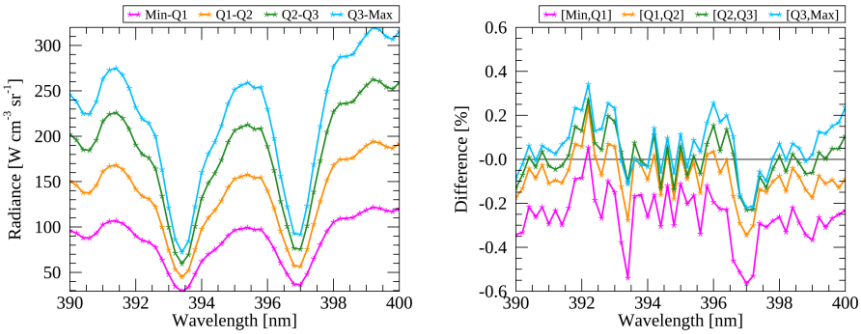
A closer inspection is performed to analyze the general spectral features over target wavelengths. Within each defect area in Figs. 11-13, the collected spectra are divided into four groups considering that ML radiances could have different systematic biases depending on the scene brightness as shown in Fig. 11. Figure 14a shows that the ML radiances over dark scenes have a positive bias while brighter scenes have a negative bias. It is interesting that the scene dependence is only significantly found for Defect 1. It should be noted that the y-axis range of Fig.14b is wider than the figures for Defects 1 and 3. Figure 14b indicates that the ML radiances are overestimated except for the very darker scenes especially at shorter wavelengths and it can be deduced that the complicated atmospheric effects involving clouds at the shorter wavelengths would affect the reproducing process. Figure 14c shows relatively large difference at the spectral peaks, but generally the difference is smaller than 0.2%



(a) *(b)* *(c)*

Figure 5 *Mean difference between ML and GEMS radiances within the target area for (a) Defect 1, (b) Defect 2 and (c) Defect 3. Each color indicates the average for each quartile and Q1, Q2 and Q3 represent the first, second and third quartile, respectively. The difference is calculated between the ML and GEMS radiances divided by the latter in percent.*

Besides the shorter wavelengths of Defect 2, the comparison between ML and GEMS radiances is presented by targeting Fraunhofer lines from 390 to 400 nm (see Fig. 14). The Ring effect caused by rotational Raman scattering can be found over the two peaks in Fig. 14a, which is generally known to be very small and largely affected by clouds (Joiner et al., 1995). Figure 14b shows that PCA-ANN reproduces the dominant features at the peaks very well on average within 0.6%, but it seems the difference increases with darker scenes where the Ring effect becomes stronger. This indicates that the ML radiances would need additional information to successfully reproduce the exact spectral features especially for the very small signals such as the Ring effect.



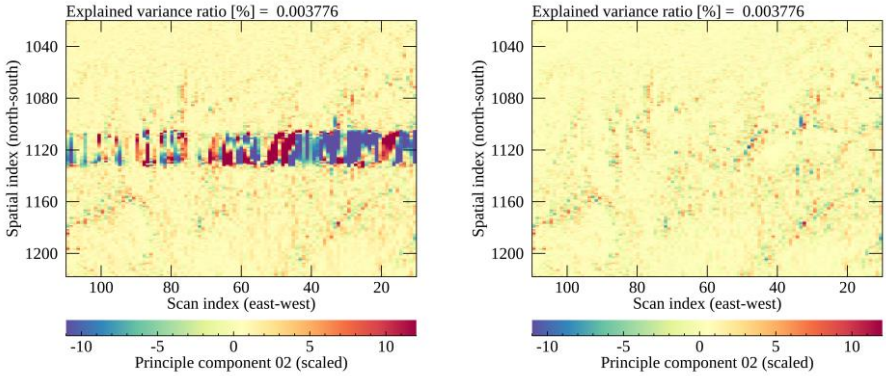
(a) *(b)*

Figure 6 *(a) Mean ML radiances (b) and the difference with GEMS radiances at Fraunhofer lines for the Defect 2 area. Each color indicates the average for each quartile and Q1, Q2 and Q3 represent the first, second and third quartile, respectively. The difference is calculated between the ML and GEMS radiances divided by the latter in percent.*

3.2.2 PCA-based spectral analysis

As applied in the pre-processing step for the present research, PCA is a very useful tool to capture the meaningful variances and it has been widely used to retrieve environmental and surface properties (Horler and Ahern, 1986; Joiner et al., 2016; Li et al., 2013, 2015). To investigate further the spectral patterns, we apply PCA to GEMS radiances collected within each area in Fig. 11-13 at the target wavelengths (see Table 3). With PCA, various spectral patterns are compressed to PC scores and this indicates that if a spectrum has disparate spectral patterns, the PC scores would also have distinct values when comparing with the PC scores of normal spectra. Figure 15 presents the PC scores of GEMS and ML radiances which are projected with the identical eigenvector matrix (corresponding to

X in Eq. 1) constructed from GEMS radiances with the exception of bad pixels. The Defect 3 area is targeted for the inspection which has the widest defective width along the north-south direction and the second PC scores are used for the analysis because the first PC scores represent mean radiances as discussed in Sect. 3.1.1. For the comparison, the radiances reproduced with spatial interpolation on the bad pixel area are projected together as shown in Fig.15a. As assumed, the PC scores from reproduced spectra with spatial interpolation show disparate values because the spectral patterns of the interpolated spectra are inconsistent with the patterns of normal spectra. The ML radiances in Fig. 15b show spatially homogenous PC scores which indicates that the machine learning methods could properly reproduce the dominant spectral patterns, in the case of the second PC.



(a) *(b)*

Figure 15. *The second PC scores of (a) GEMS radiances and (b) ML radiances on the target area for Defect 3. The PC is scaled for clarity of presentation.*

The dominant spectral patterns for each PC are presented in Fig. 16 with the eigenvector matrix constructed from GEMS radiances for the target wavelengths of Defects 1-3. Each color indicates the eigenvectors for the first-sixth PCs which determine the contribution of radiances at each wavelengths for each PC subspaces. Li et al. (2015) verifies that the leading PCs from the UV/VIS backscattered radiation (shorter than 360 nm) are highly correlated with dominant absorption features and surface properties and the trailing PCs might be associated with instrument artifacts and other unresolved spectral features with PCA. Similarly, Fig. 16 shows that the eigenvector for the first PC corresponds to the mean spectrum and the eigenvector for the second-sixth PCs show dominant spectral patterns originated from absorption features of trace gases, surface properties and unresolved features. Considering that the dominant patterns could be identically found in the eigenvectors constructed from GEMS reflectance (not shown), it can be deduced that the patterns are extracted from the spectral features caused by atmospheric interactions rather than instrument artifacts.

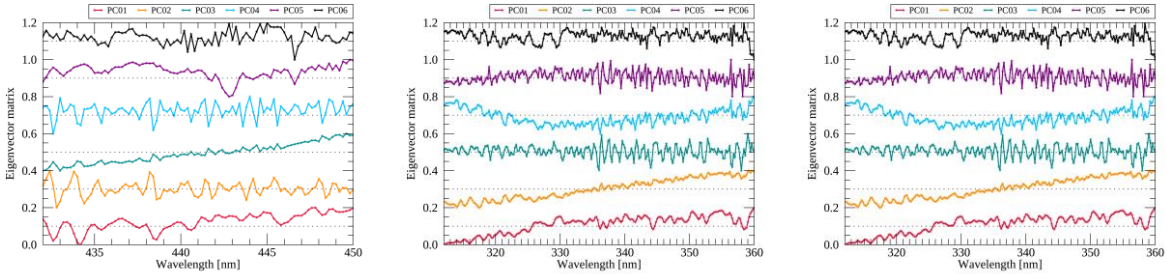


Figure 7 *Eigenvector of the first-sixth PCs applied to GEMS radiances for the target wavelengths of (a) Defects 1, (b) Defect 2 and (c) Defect 3. All eigenvectors are scaled (min-max scaling) and shifted for clarity of presentation.*

As presented in Table 4, the comparison of PC scores could indirectly provide the information on the similarity of the dominant patterns between ML and GEMS radiances with the correlation coefficient. The results show that the mean spectral pattern (the first PC) and some dominant patterns could be well reproduced with the suggested models, but other spectral features such as the second PC for Defect 2 have difficulty obtaining valid information from input radiances for accurate reproduction. The contribution to the original radiances from each PC might be very small except for the first PC because even the leading PCs have small explained variance ratio for hyperspectral data in UV/VIS spectrum. However, considering the results in Figs. 14-15, it would be enough to determine the exact spectral patterns significantly related to the important information for the retrieval process, which needs to be investigated further.

Table 2. *Correlation coefficient between PC scores of GEMS and ML radiances for the target areas of Defects 1-3 with the exception of bad pixels.*

Defects	*PC 01*	*PC 02*	*PC 03*	*PC 04*	*PC 05*	*PC 06*
Defect 1	*0.9999*	*0.9976*	*0.8172*	*0.9779*	*0.6846*	*0.6609*
Defect 2	*0.9999*	*0.8129*	*0.9876*	*0.4294*	*0.7035*	*0.5046*
Defect 3	*0.9999*	*0.9962*	*0.9787*	*0.6644*	*0.5399*	*0.2649*

Comment – Part A-3:

"But what is the real problem that the authors are trying to solve? In an instrument such as GEMS, designed to measure trace gas composition of the atmosphere through hyperspectral measurements, the goal is probably to produce trace gas products without spatial gaps caused by missing radiances. The authors do not discuss the issue of trace gas retrievals or other products derived from their predicted radiances. Is there any improvement at all in those products? …

The authors must also devise evaluation criteria that are more robust and quantitative than "these spectra look realistic." Since the goal for GEMS radiances is to derive atmospheric products such as trace gases, perhaps these trace gas retrievals can be used as the metric. Merely stating that predicted radiances agree on average with measured radiances to within X% ignores the subtle spectroscopic sensitivity of trace gases such as NO2, where the exact relationship between wavelengths is of utmost importance."

Response A-3:

As presented in the previous section, we have evaluated the reproduced spectra by comparing with actual measurements and applied PCA to analyze spectral features of reproduced spectra. With the analysis, it was found that the machine learning models properly reproduce dominant spectral patterns for Defects 1-3 with only radiances from the rest part of spectra and angle conditions. However, the exact spectral features (<1%) determined by small signal (the important information) may be accurately reproduced only if the spectral range of output radiances are closer to the input radiances enough to obtain sufficient information from the input radiances (such as Defect 3). Also, it seems additional information would be needed to reproduce exact spectral features when the output radiances are overlapped with strong absorption or scattering lines. Considering the ultimate goal of measuring hyperspectral data, we agree with the referee's suggestion to apply the retrieval algorithms for evaluating the reproduced spectra. However, as the initial approach reproducing missing radiance of GEMS, we hope to evaluate the applicability of machine learning methods for the GEMS measurements, which contain meaningful information as well as instrument artifacts. As the referee pointed out, the

effect of reproducing spectra for the retrieval process is a necessary step and based on the findings in this research, we hope to investigate further the step in a follow-up study.

Specific comments from the anonymous referee

Comment – Part B-1:

"I believe the demonstration of the method validity can be improved. In particular, the qualitative discussion about the performance of reproduction, described with Figs. 11-12, needs to be improved and more objective. Also, in addition to the prediction errors presented in Figs. 7–9, it is recommended to show how well the proposed methods reproduce known good spectra (i.e., actual measurements)."

Response B-1:

Firstly, thanks for the valuable comments and suggestions. As the referee pointed out, we acknowledged that the result parts referred in the comment definitely need improvements and thus the Sections 2-3 have been greatly revised. The applied analysis in the first draft could not quantitatively prove the validity of the suggested methods. Following the two referee's recommendations, we targeted certain areas including each defective region (Defects 1-3) and its surroundings (100-indices toward both north and south direction) where actual measurements (regarded as 'true') could be obtained.

Considering that this comment is also related to the updated section for evaluating the reproduction results of the suggested methods (mostly Section 3), the detailed responses and revised part in the manuscript would be identical to **Response A-2**.

Comment Part B-2:

"Besides, how can the spectral sampling of input/output (0.1 nm) be finer than the original GEMS data (0.2 nm)? More detailed descriptions about this are recommended. Overall, I suggest this manuscript be reconsidered after major revisions."

Response B-2:

The detailed description of the spectral interval of input and output has been added to Line 159 (w/o track changes):

The datasets for the models should be sampled at identical spectral grids and for that, each spectrum is interpolated in a pre-processing step and after the reproduction, the spectra are reversely interpolated onto its original spectral grids. Considering that the intrinsic information a spectrum has could be lost during the interpolation processes, the finer spectral grids (0.1 nm) are adopted for the model to minimize interpolation errors by preserving radiances at more frequent intervals than the original grids.

Comment Part B-3:

- **Line 78: Please give the full names of the gaseous species (i.e., O3, SO2, NO2, and HCHO).**

  Corrected.

- **Line 82: The authors refer to each of ~700 east-west pixels as a "scan," but probably this term is not accurate. Isn't the whole ~700 pixels considered to be in one scan? Also, can GEMS cover the entire field of regard by one scan? It seems that is what the authors are implying.**

The sentences have been revised as follows:

"For earth measurements, GEMS measures the backscattered radiation from east to west about 700 times by moving a scan mirror and for each scan, totally 2048 pixels are obtained along the north-south direction. All measurements at each scan position are combined together to cover the full field of regard (FOR) of GEMS."

- **Line 84: Do the CCD pixel numbers presented here represent those for only photoactive pixels?**

  The provided pixel numbers are designed to be photoactive pixels. However, signals from some pixels at the edges of the CCD are known to be invalid, which are flagged as low quality pixels. The point has been added to the revised manuscript.

- **Line 89: The general description of the bad pixel detection method is informative. But how about presenting how long the GEMS integration time is (by adding another sentence)?**

  The integration time of GEMS is 69.996409 milliseconds. The information has been updated to the manuscript.

- **Line 99: This sentence sounds as if the results of 1-D interpolation were presented earlier, which is not true. How about rephrasing this sentence, using a verb like "imply" instead of "indicate"?**

  We agreed to the point. It has been corrected.

- **Line 104: The subject affected by the defective pixels is the quality of ozone retrieval, not the ozone properties themselves.**

  Indeed. It has been corrected.

- **Line 148: How can the spectral interval of input and output (0.1 nm) be narrower than that of original GEMS measurements (0.2 nm)? How are the GEMS measurement spectra sampled onto the finer grids? Please give more details here.**

  As for the spectral intervals of GEMS spectra for the training process, the response for the comment has been addressed in the previous section.

- **Line 149: Did you investigate how much the results changed when trained without solar zenith angle (SZA) and viewing zenith angle (VZA)? Please describe the impacts of including these variables.**

  The impact of angle conditions as input has been analyzed and added to Line 177 (w/o track changes):

  Figure 5 presents the converging process of the PCA-ANN model for Defect 2 applying different optimizers with and without SZA and VZA conditions. The additions of the angle conditions as input parameters speed up the model convergence with smaller MSE because without the angle parameters, the information would be implicitly elicited during the optimization process. The model converges with angle conditions at 44, 98 and 33 epochs for Adam, SGD and RMSprop, respectively. Adam converges at the smallest MSE while the SGD converges with the highest MSE. RMSprop presents unstable loss for validation data and converges with higher MSE compared to Adam.

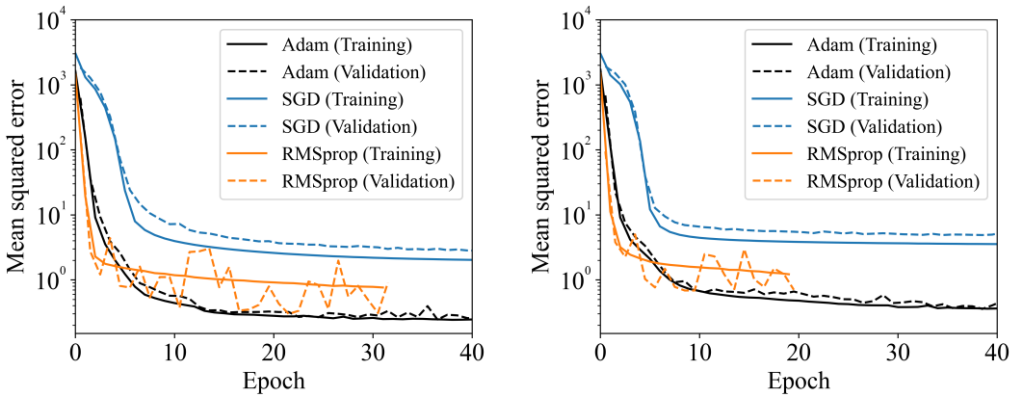


Figure 8 *Training and validation losses for Defect 2 (a) with and (b) without the angle conditions as input parameters. The results are obtained with different optimizers such as Adam (black), SGD with the gradient clipping value of 0.5 (blue) and RMSprop (orange).*

- **Figure 5: The caption and the color bar title do not correspond. Which wavelength was used between 310 and 354 nm?**

  Thanks for the correction. It is radiance at 310 nm and the caption has been corrected accordingly.

- **Line 264: How can we tell if spectra look "reasonable"? This statement is vague. Please consider changing Figs. 11-12 to include any reference (know, good, measured) spectra for the reconstructed parts.**

  The response for this comment is addressed in the previous section.

- **Line 269: I believe the term "noise" itself implies randomness, which would not necessarily be canceled in the normalized radiance. Please consider replacing the term with another, e.g., error, bias, artifact, etc.**

  Artifacts would be more proper expression, indeed. It has been updated.

- **Please consider re-writing the units in the figures as W cm–3 sr–1**

  Corrected.

- **Please consider minor English corrections below.**
  - **Lines 42, 49, 50, 100: affect to -> affect ?**
  - **Lines 109, 148, 184, 185, 199, 214, 221, 238, 241, 243, 250, 254, 258, 268: Defect -> Defects**
  - **Line 225: Fig. -> Figs.**
  - **Line 242: N-S -> North-South**
  - **Line 276: A period (.) missing between sentences**

  These comments have been addressed in the revised manuscript.

Reference

Bak, J., Baek, K. H., Kim, J. H., Liu, X., Kim, J. and Chance, K.: Cross-evaluation of GEMS tropospheric ozone retrieval performance using OMI data and the use of an ozonesonde dataset over East Asia for validation, Atmos. Meas. Tech., 12(9), 5201–5215, doi:10.5194/amt-12-5201-2019, 2019.

Choi, H., Liu, X., Gonzalez Abad, G., Seo, J., Lee, K.-M. and Kim, J.: A Fast Retrieval of Cloud Parameters Using a Triplet of Wavelengths of Oxygen Dimer Band around 477 nm, Remote Sens., 13(1), 152, doi:10.3390/rs13010152, 2021.

Horler, D. N. and Ahern, F. J.: Forestry information content of thematic mapper data, Int. J. Remote Sens., 7(3), 405–428, doi:10.1080/01431168608954695, 1986.

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.

Joiner, J., Yoshida, Y., Guanter, L. and Middleton, E. M.: New methods for the retrieval of chlorophyll red fluorescence from hyperspectral satellite instruments: Simulations and application to GOME-2 and SCIAMACHY, Atmos. Meas. Tech., 9(8), 3939–3967, doi:10.5194/amt-9-3939-2016, 2016.

Kim, G., Choi, Y. S., Park, S. S. and Kim, J.: Effect of solar zenith angle on satellite cloud retrievals based on O2–O2 absorption band, Int. J. Remote Sens., 42(11), 4224–4240, doi:10.1080/01431161.2021.1890267, 2021.

Li, C., Joiner, J., Krotkov, N. A. and Bhartia, P. K.: A fast and sensitive new satellite SO2 retrieval algorithm based on principal component analysis: Application to the ozone monitoring instrument, Geophys. Res. Lett., 40(23), 6314–6318, doi:10.1002/2013GL058134, 2013.

Li, C., Joiner, J., Krotkov, N. A. and Dunlap, L.: A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite, Geophys. Res. Lett., 42(7), 2515–2522, doi:10.1002/2015GL063204, 2015.