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

Yeeun Lee[1], Myoung-Hwan Ahn[1], Mina Kang[1], Mijin Eo[1]

[1]Department of Climate and Energy Systems Engineering, Ewha Womans University, Seoul, 03760, Republic of Korea

Correspondence to: Myoung-Hwan Ahn (terryahn65@ewha.ac.kr)

Abstract. Earth radiances in the form of hyperspectral measurements~~data~~ contain useful information on atmospheric constituents and aerosol properties. The Geostationary Environment Monitoring Spectrometer (GEMS) is an environmental sensor measuring such hyperspectral data in the ultraviolet and visible (UV/VIS) spectral range over the Asia-Pacific region. After ~~successful~~ completion of the in orbit test (IOT) of GEMS in October 2020, bad pixels are found as one of~~a~~ remaining calibration issues because ~~to be updated with follow up treatment.~~ d~~D~~uring the IOT~~,~~ one-dimensional interpolation ~~is~~ performed ~~in operation~~ to replace the erroneous pixels ~~of GEMS, which~~ caused~~s~~ high interpolation error for a wide defective area on the detector array. The error results in obvious spatial gaps in the measured radiances and also retrieved properties. To resolve the fundamental cause of the issue, this study takes an approach reproducing erroneous spectral features of the defective spectra with machine learning methods with ~~To resolve the issue, this study suggests machine learning methods with~~ artificial neural network (ANN) and multivariate linear regression (Linear) ~~for filling in a spectral gap of defective spectra~~. The basic assumption of the methods is that radiances ~~of~~ in a spectrum have linear and non-linear relations and a finite range of radiances can be reproduced with the relations. The machine learning models are trained with ~~normal~~ defect-free measurements of GEMS after dimensionality reduction ~~for input parameters~~ with principal component analysis (PCA) for efficient model training ~~process~~. Results show that PCA-Linear has small prediction errors especially for a narrower spectral gap and less vulnerable to outliers ~~in the training data~~ with an ~~prediction~~ error of 0.5-5%. PCA-ANN shows better results emulating strong non-linear relations with an ~~prediction~~ error of ~~within~~ about 5% except for the shorter wavelengths around 300 nm. It is verified that ~~The~~ dominant spectral patterns can be successfully reproduced with the models ~~nearly~~ within the level of radiometric calibration accuracy of GEMS, but a limitation~~s~~ still remains when it comes to much finer spectral features~~particularly in the reproduction of the precise spectral features which needs additional information to be investigated further~~. When applying the reproduced spectra to retrieval processes of cloud and ozone, the cloud centroid pressure shows an error of around 1% while total ozone column density shows relatively higher variance. It seems to be clear that the effectiveness of the method can be improved further by optimally setting the input and output spectral bands for the spectral replacement. As ~~the~~ an initial ~~approach~~step in reproducing spectral patterns for erroneous spectra ~~reproducing missing radiances of GEMS~~, this study verifies ~~that spectral relations in the UV/VIS spectrum are successfully reproduced with a simple machine learning model, which~~machine learning

methods ~~–~~have~~s~~ high potential to be updated further for enhancing measurement quality of environmental satellite measurements.

1 Introduction

Earth radiances can provide useful information on the atmospheric chemical composition, especially when it is measured in the form of many contiguous spectral bands. This type of measurements is referred to as ‘hyperspectral’ (Bovensmann et al., 1999; Goetz et al., 1985) which is frequently sampled with high spectral resolution to accurately describe absorption lines of ~~a~~ targeted gaseous or particulate components (Boersma et al., 2004; Kang et al., 2020; Manolakis et al., 2019; Pan et al., 2017). The Geostationary Environment Monitoring Spectrometer (GEMS) on-board the Geostationary Korea Multi-Purpose Satellite-2B (GEO-KOMPSAT-2B) is an environmental sensor providing such a hyperspectral measurement in the ultraviolet and visible (UV/VIS) spectral region from 300 to 500 nm with a spectral resolution of finer than 0.6 nm (Kim et al., 2020). Following the launch of the satellite in February 2020, the in orbit test (IOT) of GEMS was successfully completed in October 2020 with some issues to be continuously monitored. The root cause of each issue is to be examined with collected long-term measurements on the radiance level (Level 1B), as it has been dealt with for other polar orbit satellite sensors having similar sensor characteristics (Ludewig et al., 2020; Pan et al., 2019, 2020; Schenkeveld et al., 2017).

One of the issues to be periodically monitored is ~~about~~ bad pixels, which refer to anomalous pixels having hot, cold, noisy or drifted readout values in raw data (Lo´pez-Alonso and Alda, 2002). The definition of bad pixels is not universal, and in this paper, it refers to all kinds of pixels having abnormal observation features. Bad pixel detection is based on the sensor characterization sorting out erroneous signals from a normal trend and ~~. A~~a few hot pixels were flagged as bad pixels during on-ground tests for GEMS. ~~and a~~Additionally, more pixels have been sorted out during the IOT because of the impacts from the launch of the satellite and different environment conditions in space. The number of bad pixels may increase as time goes by (Kieffer, 1996), which indicates a significant number of bad pixels could affect measurement quality during the operation period~~ of GEMS~~.

Subsequently, the detected pixels should be replaced~~Following the bad pixel detection, replacement of measurements on bad pixel positions needs to be performed. There are various ways to replace the measurements on bad pixels~~ (Boldrini et al., 2012; Burger, 2009; Rankin et al., 2018)~~,~~ and in the GEMS calibration system, it adopts ~~one-dimensional~~ spatial interpolation on the detector array along the spatial direction (Fischer et al., 2007; Schläpfer et al., 2007). However, the approach showed its limitation during the IOT, when an area consisting of bad pixels is quite large and the adjacent pixels valid for spatial interpolation are too far from the erroneous area. Especially, when a scene on the Earth dramatically changes, discontinuity caused by the interpolation becomes ~~larger~~more apparent. This effect causes ~~not only~~ spatial discontinuity in Level 1B data and retrieved properties (Level 2) ~~on two-dimensional measurements, but also to~~by affecting ~~a~~ retrieval processes ~~using the~~with contaminated spectral features~~ contaminated by bad pixels (Marchetti et al., 2019)~~.

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.

In this respect, ~~this study suggests machine learning methods to replace bad pixels on the radiance level using valid spectral features of normal measurements. As a way of replacement,~~ we ~~compare~~ suggest machine learning models with ~~approaches using~~ artificial neural network (ANN) and multivariate linear regression which is trained to emulate spectral relations. Theoretically, it has been verified that ANN can accurately emulate non-linear relations with a simple model structure when there are a large number of training data (Cybenko, 1989; Hornik et al., 1989). Machine learning methods also have a high chance to successfully process hyperspectral data because the abundant datasets make the training process more efficient after breaking the curse of dimensionality with a proper pre-processing step (Gewali et al., 2018). For that, p~~P~~rinciple component analysis (PCA) is applied in this study, as it has been known to be very~~is~~ useful to extract important information from hyperspectral measurements (Bajorski, 2011)~~.~~ and widely used to retrieve environmental and surface properties (Horler and Ahern, 1986; Joiner et al., 2016; Li et al., 2013, 2015).

For atmospheric remote sensing, the majority of researches ~~on hyperspectral measurements~~ has employed machine learning as a proxy of the radiative transfer model to retrieve geophysical states ~~with~~ from measured spectral radiances (Hedelt et al., 2019; Loyola et al., 2018; Zhu et al., 2018). There are fewer approaches applied to obtain radiation flux (Dorvlo et al., 2002; Zarzalejo et al., 2005) and even much fewer to obtain hyperspectral radiances for different purposes such as to accurately quantify radiative forcing in climate system (Taylor et al., 2016), increase spectral resolution (Le et al., 2020) and fill in a spectral gap for inter-calibration (Wu et al., 2018). A monochromatic radiance itself rarely contains any important meaning and thus seldom has it been a final target for machine learning. In this study, however, radiance at each wavelength of a targeted spectral region are~~becomes~~ an important output to be reproduced.

The following sections are organized as follows. Section 2 introduces sensor specification of GEMS ~~including an overview of bad pixel detection and replacement methods for GEMS. In the section,~~and a general description of machine learning models ~~suggested in this study is introduced as well as the~~with model structure and hyperparameter setting. Section 3 contains model optimization results and error analysis for wide defect regions. With the optimized model, the spatial and spectral inspection is performed ~~to actual measurements including bad pixel areas~~for reproduced radiances and retrieved properties. In Sect. 4, conclusions are presented with limitations as well as further application of the method~~s~~ in ~~the~~ future study.

2 Data and methods

2.1 Data description

2.1.1 GEMS

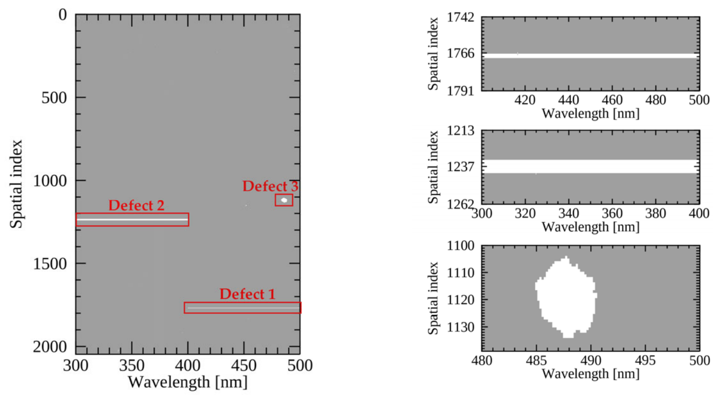
GEMS is a UV/VIS imaging spectrometer in the geostationary orbit observing the Asia-Pacific region (5° S-45° N, 75° E-145° E) with high spatial and spectral resolution to retrieve key atmospheric constituents such as ozone (O_3), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), formaldehyde (HCHO), glyoxal (CHOCHO) and aerosol properties ~~(Level 2)~~ (Kim et al., 2020). The observation targets of GEMS are the Sun (irradiance mode) and the Earth (radiance mode) and the description for each measurement mode is summarized in Table 1. In both measurement modes, incident light from a scene passing through a fore-optics and a spectrometer reaches to a two-dimensional detector array, the charge-coupled device (CCD) detector. The CCD of GEMS comprises 2,048 rows and 1,033 columns of photoactive pixels along the spatial direction from north to south and the spectral direction with a sampling interval of 0.2 nm, respectively. GEMS observes the Sun on the purpose of calibration once a day with a premise ~~of~~ for the measured solar irradiance being stable and nearly time independent. 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. The data used in this study are the operational data (Level 1C) which are used for the retrieval processes of Level 2 products.

Table 1 Top level measurement specifications of GEMS

Measurement mode	Solar irradiance	Earth radiance
Data dimension [spectral, spatial, scan]	[1033, 2048]	[1033, 2048, 695] (nominal scene)
Spectral range [nm]	300-500	
Spectral sampling [nm/pixel]	0.20	
Spectral resolution [nm]	< 0.60	
Spatial resolution [km^2]	-	3.5 × 8 (spatial × scan)
Measurement frequency	Once a day (13:00 UTC)	Every hour (00:45-07:45 UTC)

2.1.2 Bad pixel

Bad pixel detection is generally performed with dark-current measurements which are taken without exposure to light for a certain integration time (Howell, 2006), and for GEMS, the integration time corresponds to about 70 milliseconds. Figure 1 illustrates bad pixel positions (in white) on the GEMS CCD detector array identified during the IOT. A cluster and distinct line shapes of bad pixels shown in Fig. 1a are initially identified during on-ground calibration before the launch of the satellite. Following the suggestions made by the instrument developers, linear interpolation along the spatial direction (north-south) is applied to replace the unusable measurements on bad pixel positions and a single bad pixel could be properly substituted with such a simple procedure. However, it was found during the IOT that significant interpolation error could occur on the bad pixel positions denoted as Defects 1-3 (see Fig. 1b), especially when the invalid spatial width ~~of the invalid area~~ is too wide ~~as shown in~~such as Defects 2-3~~ and 3~~.



(a) (b)

Figure 1 ~~(a)~~ The t~~T~~wo-dimensional bad pixel map (a) on the GEMS CCD detector along the spectral (x-axis) and spatial direction (y-axis) and (b) zooming in the bad pixel positions from top to bottom rows for Defects 1-3. Bad pixels are marked in white.

The interpolation error ~~could~~ seriously affects 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. O~~The~~ zone ~~O_3~~ retrieval is also affected by Defect 2 (300-400 nm) as the spectral radiances within 300-380 nm are ~~provide~~ major ~~information for the~~ozone ~~O_3 retrieval of GEMS~~absorption lines in the UV/VIS spectral range (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 spatially interpolated radiances ~~at the certain wavelength~~ are homogeneous with its surroundings (see Fig. 2~~b~~), the spectral patterns are not properly reproduced with the existed method (spatial interpolation) which caus~~ing~~es ~~the~~ distinct horizontal lines to the retrieved products, ~~in Fig. 2c.~~ to be discussed in Sect. 3.2 2.

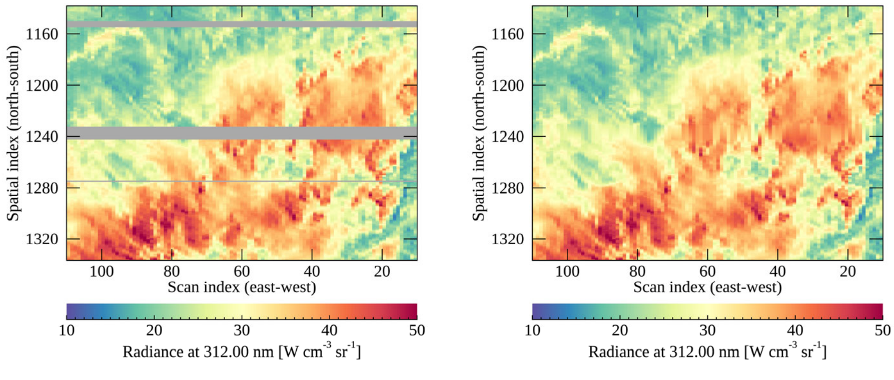
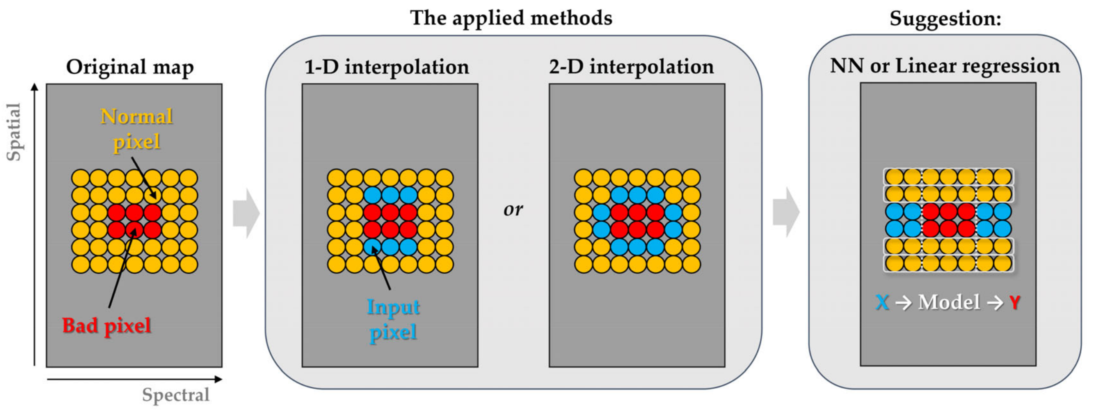


Figure 2 Spatial distribution of GEMS radiances at 312 nm with bad pixels (a) marked in dark gray and (b) reproduced with spatial interpolation~~ and (c) the total O_3 column of GEMS~~. The GEMS spectra are measured on 10 March 2021 (06 UTC)~~The measurements are on 10 March 2021 (06 UTC)~~.

~~To eliminate the bad pixel effects this study suggests machine learning methods to spectrally reproduce radiances on bad pixels instead of spatial interpolation as described in Fig. 3. The multivariate linear regression and ANN models are compared to evaluate model performance for reproducing Earth radiance corresponding to the bad pixel positions of Defects 1-3. Solar measurements have high spatial homogeneity resulting in small interpolation error even on the large bad pixel areas, and are not considered in this study.~~



~~**Figure 3** Schematic chart of the input (blue circle) and output pixels (red circle) on the GEMS CCD for the spatial interpolation and machine learning methods suggested in this study. Yellow circle indicates adjacent pixels to bad pixel position.~~

2.2 Replacement approach

2.2.1 General description

Upwelling radiances are determined by the interactions of light with trace gases, aerosols and clouds in the atmosphere and surface reflectivity. Especially, scene properties are the dominant factor resulting in strong linear relations among radiances in a spectrum. In other words, when a scene is dark (bright), the upwelling radiances of the scene over the whole spectral region tend to become generally low (high). S~~Reproduction of radiances on bad pixels~~spectral replacement is based on a fact that radiances at different wavelengths for a scene are highly correlated with each other and have certain spectral relations (Liu et al., 2006; Wu et al., 2018). If the relations are~~can be~~ accurately ~~established~~emulated, some missing values in a spectrum can be ~~properly~~ reproduced with radiances at the other wavelengths. The important question here is whether non-linear relations can be accurately reproduced with only the spectral information. ~~Further questions to be investigated are whether non-linear relations could be accurately emulated with machine learning methods and valid information for the reproduction may exist in the input radiances.~~ To investigate this~~emulate the relations of the input and output radiances in a spectrum~~, ~~the~~ randomly collected GEMS spectra measured on defect-free pixels ~~for various scenes~~ are used to establish the relations with ~~which are measured on normal pixels located closer to a bad pixel area on the detector array for Defects 1-3. The~~ the basic premise ~~of this approach is~~ that neighbor pixels on the detector array (set to within 100 spatial indices) would have similar measurement characteristics~~. After training a model with the normal spectra, a spectral gap (output radiances) could be reproduced through the model~~.

Because it is highly possible that input radiances have redundant information, PCA is applied for dimensionality reduction to compress the input radiances to low-dimensional principle components (PCs). The strong linear relations among radiances in a spectrum are compressed to the first PC, which has the largest variance. The non-linear properties caused by atmospheric scattering, absorption, different optical paths and sensor noise are projected onto the subsequent PC subspaces. The PCA process is given by the following Eq. (1):

$$\mathbf{Z}_{n\times p} = \mathbf{X}_{n\times\lambda}\mathbf{W}_{\lambda\times p} \qquad \mathbf{(1)}$$

where **Z, X** and **W** represent the PC scores, input and PC matrix, respectively. The PC scores matrix (**Z**) is obtained by projecting the input to the PC subspaces with **W**, which is obtained by applying eigenvalue decomposition to the **X**. The subscript n, λ and p ~~means~~ indicates the dimension of matrix corresponding to the number of datasets, input wavelengths and the number of PCs, respectively.

With the compressed data, multivariate linear regression (PCA-Linear) and ANN (PCA-ANN) models are trained to define the relations between input ($\mathbf{X_m}$) and output ($\mathbf{Y_n}$) radiances in a spectrum. The PCA-ANN model is constructed with a simple feed-forward model with a hidden layer as described in Fig. 34. In the model optimization process, the PCA-ANN model with a hidden layer showed faster and more effective convergence of loss function than the models having multi-hidden layers in this study. For PCA-Linear, it adopts a simple linear model structure consisting of parameters such as weight and bias having the minimum mean squared error (MSE) between the regressed and measured radiances. After model optimization, it can be used to replace bad pixels ($\mathbf{X'_m}$, $\mathbf{Y'_n}$) with reproduced radiances ~~to be~~ likely measured by the sensor.

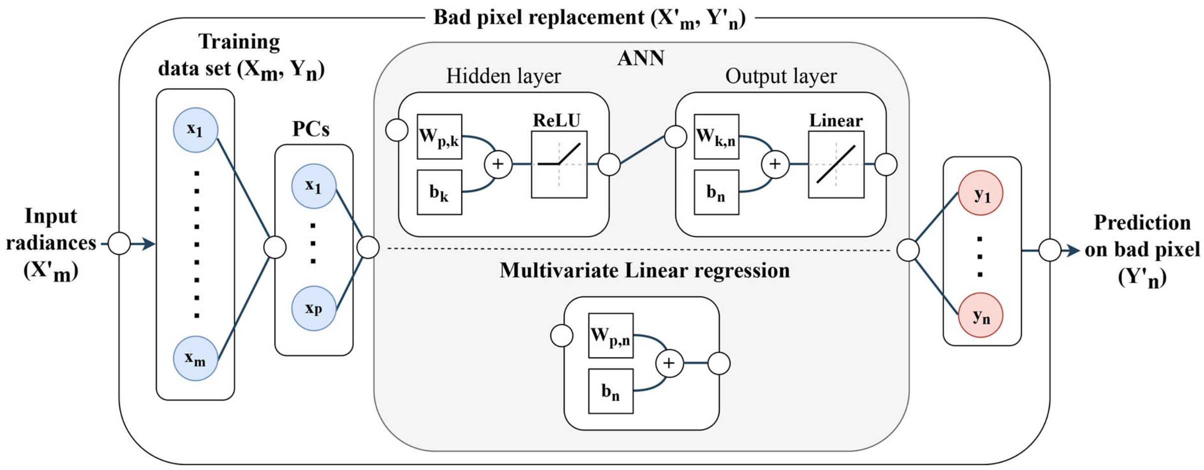


Figure 3 Schematic chart of the training and bad pixel replacement process. **W** and **b** represent weight and bias parameters in each layer. The subscript *m, n, p* and *k* is equal to the spectral dimension of input and output parameters, the number of PCs and hidden nodes of the ANN model, respectively.

2.2.2 Input/output and model optimization

For the ~~reproduction process~~model training, radiances of each spectrum are divided into input and output radiances based on the specified spectral ranges in Table 2. The spectral ranges of output radiances for Defects 1-3 are identical to each defective region and the rest part of a spectrum becomes input radiances. The constructed input and output datasets are split into training and test data to update model parameters and check for overfitting, respectively~~Training and test data are constructed with GEMS radiance data~~ which are randomly sampled out in March ~~April~~ 2021 ~~to update model parameter and check for overfitting, respectively~~. When training a machine learning model, it is important to sample the similar numbers of bright or dark scenes because the majority corresponds to dark scenes with random selection. Considering that the training process becomes unstable when the collected scenes are skewed to low radiances, oversampling is a necessary step before the scene selection by binning the datasets depending on the magnitude of spectra.

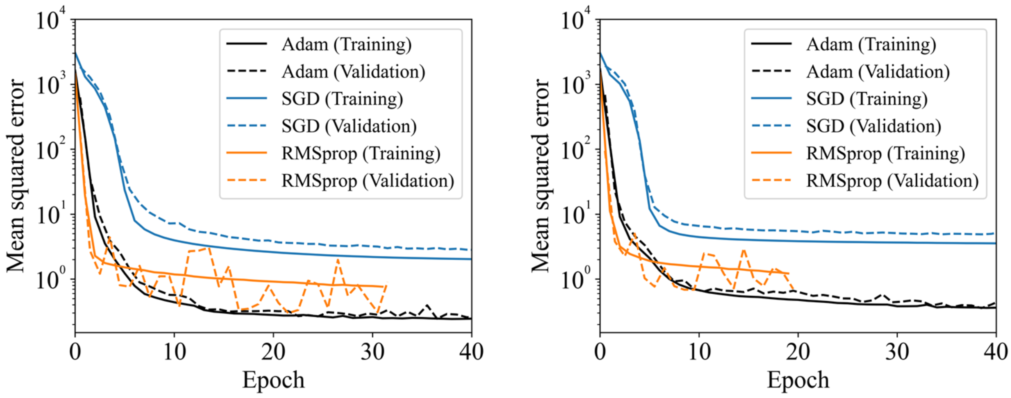
The datasets for the models should be sampled at identical spectral grids and for that, each spectrum is interpolated in a pre-processing step and after the reproduction, the spectra are reversely interpolated onto its original spectral grids. Considering that the intrinsic information a spectrum has 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. The solar zenith angle (SZA) and viewing zenith angle (VZA) are key variables determining optical paths of upwelling and downwelling radiances and thus are used as input variables together with radiances.

The neural network constructed with the hyperparameter setting presented in Table 2 is implemented with TensorFlow, a high-level Application Programming Interface (API) written in Python. As described in Fig. 4, the activation function is the Rectified Linear Unit (ReLU) in the hidden layer of the ANN model. The structure itself is not complicated but it has multiple nodes in the input and output layers, which makes ReLU more competitive (Nwankpa et al., 2018). The hyperbolic tangent (tanh) and sigmoid function show poor results especially when the output parameters have lower variance making the optimization stuck into the averaged value and preventing the model from being updated.

Table 2 Input and output (I/O) parameters for ANN training and hyperparameter for optimization of neural network.

Model	Parameter	Defect 1	Defect 2	Defect 3	Remark
I/O	Input (X_m)	SZA / VZA			Random selection (100,000 for training and test data)
		300-400 nm	400-500 nm	460-483.9 / 491.1-500 nm	
	Output (Y_n)	400.1-500 nm	300-399.9 nm	484-491 nm	
Hyper-parameter	Activation function	ReLU			
	Optimizer	Adam optimizer			
	Loss function	Mean squared error			
	Scaling	Standardization			

For the optimizer, Adaptive Moment Estimation (Adam) is used which shows stable results compared to Stochastic Gradient Descent (SGD) and Root Mean Square Propagation (RMSProp) (Kingma and Ba, 2015). It is empirically found that SGD without gradient clipping tends to cause exploding gradient and RMSProp has difficulty reaching the global minima compared to Adam. 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 addition~~s~~ of the angle conditions as input parameters speeds 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.



(a) (b)

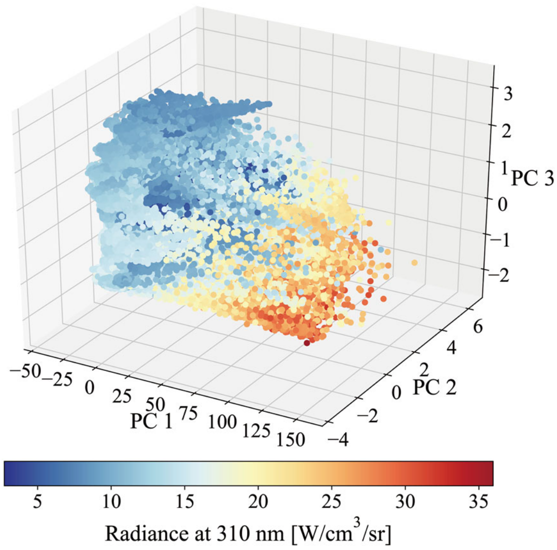
Figure 4~~5~~ Training and validation losses for Defect 2 (a) with and (b) without the angle conditions as input parameters. The results are obtained with different optimizers such as Adam (black), SGD with the gradient clipping value of 0.5 (blue) and RMSprop (orange).

3 Results and discussion

3.1 Model selection

3.1.1 Optimization results

~~Earth radiance is determined by the interactions of light with trace gases, aerosols and clouds in the atmosphere and reflected properties of a scene. The magnitude of a spectrum is dominantly determined by the scene properties which result in strong linear relations among radiances in a spectrum. In other words, when a scene is dark (bright), the upwelling radiances of the scene over the whole spectral region tend to become generally low (high). The PCA analysis performed for dimensionality reduction describes the characteristic with PC scores of input radiances in training data for Defect 2 (See Fig. 6). In the figure, it can be found that the first principal component (PC) is highly correlated with the magnitude of a spectrum represented by the radiance at 354 nm. This indicates that strong linear relations among radiances in a spectrum are compressed to the first PC, which has the largest variance. The non-linear properties caused by atmospheric scattering, absorption, different optical paths and sensor noise are projected onto the subsequent PC subspaces.~~



~~**Figure 6** PC scores of input training data from 400 to 500 nm for Defect 2 after dimensionality reduction. Colorbar represents the radiance at 310 nm of output training data.~~

Figure 5~~7~~ shows model optimization results depending on each model and the number of PCs as the input nodes. Because the spectral range of output radiances differs for each defect region (Defects 1-3), model optimization is also performed, respectively. The spectral ranges of output radiances for Defects 1 and 2 are wider than that of Defect 3 which results in higher MSE. PCA-ANN seems to be unstable for Defect 1 showing overfitting~~ed features~~ which might be caused by unfiltered outliers in output radiances of GEMS at the wavelengths longer than 480 nm. ~~It is empirically found that PCA-ANN is more vulnerable to outliers compared to PCA-Linear.~~ Defect 2 is at the wavelengths where the upwelling radiances are largely affected by ozone ~~O_3~~, which increases non-linearity between input and output radiances. Because of the strong non-linearity, PCA-ANN shows better performance than PCA-Linear for Defect 2. Defect 3 has the smallest number of output parameters in a narrow spectral gap which causes strong correlation between input and output radiances. The loss functions (MSE) in Fig. 5~~7~~c are small and converge quickly for both PCA-ANN and PCA-Linear models. With the results, the optimized number of PCs is set to 90 for all defect regions when loss functions for both training and test data efficiently converge, with PCA-Linear for Defects 1 and 3 and the PCA-ANN model for Defect 2.

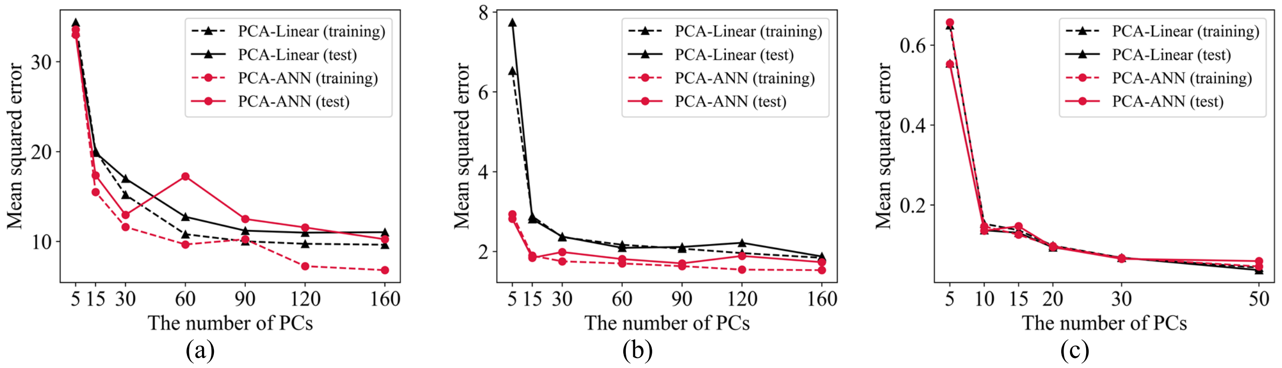
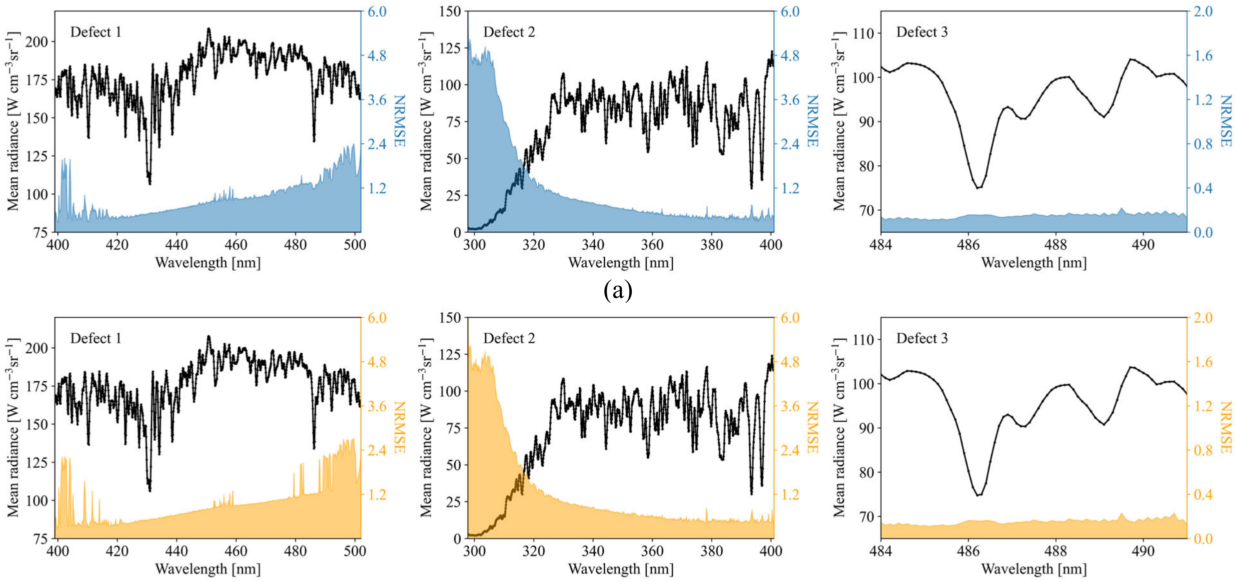


Figure 5 Loss function depending on the number of PCs with PCA-ANN (red) and PCA-Linear (black) models for spectral replacement with training and test datasets for ~~to predict the output radiances corresponding to the spectral range of~~ Defects 1-3 ((a): Defect 1, (b): Defect 2 and (c): Defect 3). ~~The dashed and solid line indicates training and test results, respectively.~~ The number of hidden nodes for ANN is double the number of PCs.

3.1.2 Statistical evaluation

The optimized model structures for Defects 1-3 are set as described in the previous section. Following that, in this section, model performance is statistically evaluated with training and test datasets specified in Table 2. Figure 6~~8~~ presents mean and normalized root mean squared error (NRMSE) of the ~~predicted~~ output radiances ~~with~~ of training and test data. The NRMSE is a statistical indicator normalized by the mean radiance at each wavelength and it can be found that radiances ~~affected by~~at strong absorption lines have relatively high uncertainty. Especially, information from the radiances in 400-500 nm is insufficient to properly represent ~~O_3~~ ozone absorption features at shorter wavelengths and it causes high uncertainty at the wavelengths shorter than 325 nm in Defect 2. Defect 1 also has higher errors around the edges of output spectral ranges where pixel saturation could be found. It is clear that prediction errors of Defects 1-2 increase at the output wavelengths far from the input spectral bands. Defect 3 shows~~has~~ the smallest NRMSE of around 0.2% because of strong linear relations between input and output radiances as previously mentioned in Sect. 3.1. ~~The NRMSE is less than 0.1% for both training and test data for Defect 3.~~ The results show that it is possible to successfully reproduce spectral features at a narrower spectral range with simple linear regression.

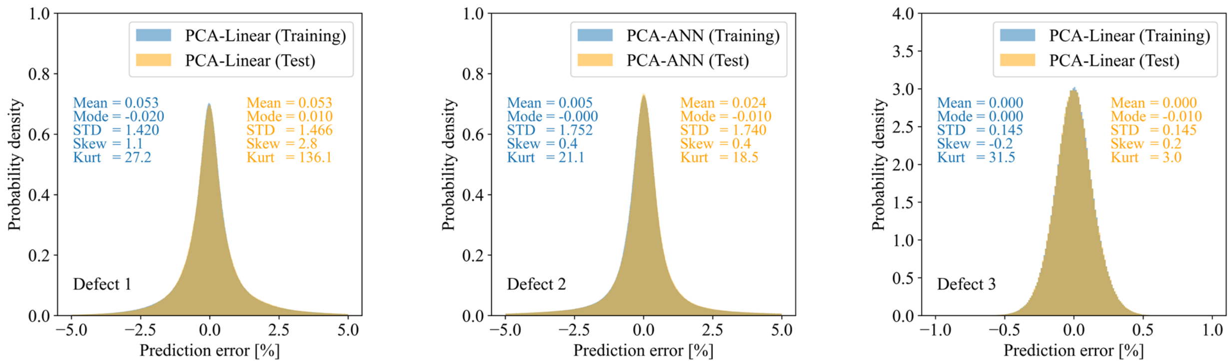


(a)

(b)

Figure 6~~8~~ O~~Mean (black) and NRMSE (blue) spectra of o~~utput radiances for Defect 1-3 with averaged spectra and NRMSE for~~with~~ (a) training~~-~~ and (b) test ~~-~~datasets measured in March~~-April~~ 2021. The unit of NRMSE is in percent.

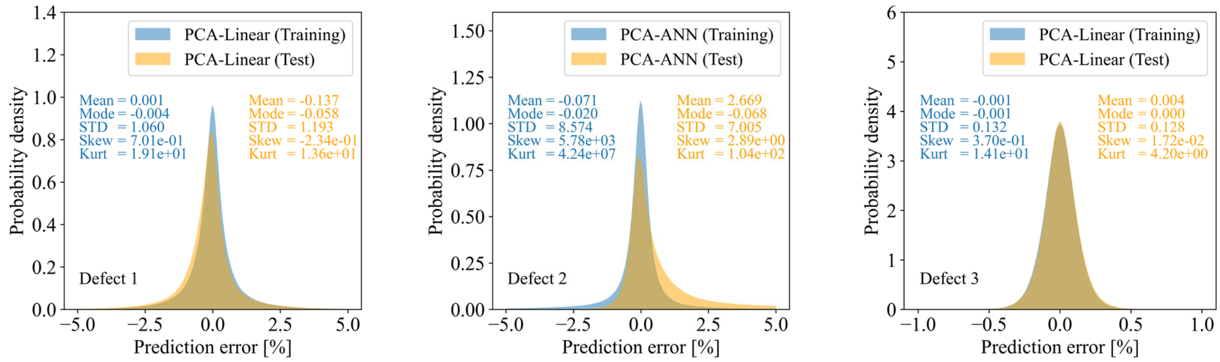
Figure 7~~9~~ ~~shows~~ presents ~~the~~ error histograms of each prediction model for Defects 1-3 with training and test data. The mode and mean of error histograms are on the order of 0.001-0.01 ~~for test and training data. The machine learning models are good enough to properly emulate spectral relations between input and output parameters,~~ and both training and test data show nearly identical distribution~~but it is somewhat over-fitted to the training data causing a few outliers for the prediction of test data~~. Defects 1-2 have~~s~~ ~~the~~ larger~~est~~ standard deviation, which is consistent with the higher NRMSE at the edges of output spectral range and shorter wavelengths around 300 nm in Fig. 6, respectively~~8~~. The largest kurtosis of Defect 1~~2~~ for ~~both training and~~ test data indicates tails of the distributions are heavy compared to normal distribution, mostly from the radiances of ~~at shorter wavelengths~~saturation pixels. ~~Considering that the overall prediction error is within 5% except for the O_3 absorption lines, the prediction models for Defects 1-3 are well constructed for further bad pixel replacement.~~



(a) (b) (c)

Figure 7~~9~~ Error histograms ~~Prediction error~~ of randomly collected training (blue) and test (yellow) datasets measured in March~~-April~~ 2021 with the optimized models for Defects 1-3 (PCA-ANN for Defect 2 and PCA-Linear for Defects 1 and 3). Prediction error and the statistics are calculated with the difference between the predicted and measured radiances divided by the latter.

~~The training and test datasets presented in Figs. 8-9 are randomly collected spectra of GEMS measurements in March-April 2021, which guarantees a basic assumption in machine learning that the underlying population should be identical for training and test data (Zhen and Li, 2008). However, in operation, the prediction model is obliged to be trained in advance with sufficient datasets for timely reproducing erroneous pixels of satellite measurements on a daily basis. This indicates the assumption might be violated if measurement characteristics of training and test data significantly change. To investigate further the effect, the prediction model is trained with training and test data measured in March and April 2021, respectively (see Fig. 10). The results show that histograms of test data for Defects 1-2 are more skewed than those of training data, when the measurement periods of training and test datasets differ. On the other hand, prediction results for Defect 3 are independent to the measurement periods of training and test data. This results indicate that the spectral features of GEMS spectra change profoundly as time goes by.~~



~~(a)~~ ~~(b)~~ ~~(c)~~

~~Figure **Same as Fig. 9 with training and test datasets measured in March and April 2021, respectively.**~~

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 area (Defects 1-3) and its surroundings where actual measurements regarded as ‘true’ exist ~~could be obtained~~are investigated. 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~~ to the sub-nadir longitude of GK-2B. Table 3 presents spectral ranges of Defects 1-3 and the target wavelengths for the analysis. ~~Specifying~~ Targeting the wavelengths for the analysis helps to specifically analyze the spectral patterns of absorption lines of trace gases and cloud properties.

Table 3 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. 8-10~~11-13~~, each column shows GEMS, ML radiances and the difference while the first and second rows show the radiances at the representative wavelengths showing the smallest and the largest differences, respectively. Figure 8~~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. 8~~11~~ is within the range of ~~-~~± 0.5% because the spectral gap of Defect 3 is narrower than the counterparts of Defects 1-2. For Defect 3, there is no distinct scene dependence over the output wavelengths and the difference shows noise-like features ~~and the spatial dependence~~ originated from instrument artifacts. One thing to be noted is that the results presented here is calculated at the finer spectral grids of 0.1 nm before interpolating to the original spectral grids. After the interpolation, the difference especially at strong peaks in a spectrum could increases by 0.5% for Fig. 8b.

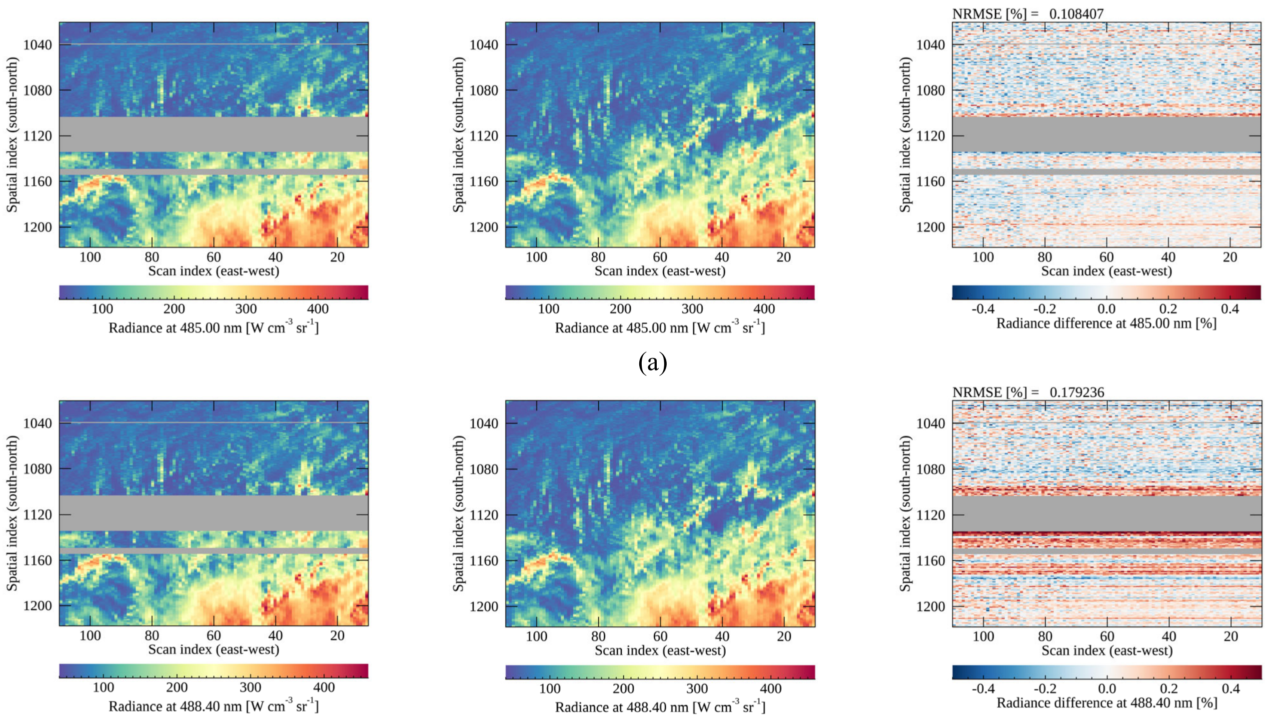


Figure 8~~11~~ Spatial distribution of ~~The~~ GEMS, ML radiances and the difference (from the first to the third column~~left to right~~) at the wavelengths presenting (a) the smallest and (b) the largest differences for the Defect 3 area. The difference is calculated between the ML and GEMS radiances divided by the latter in percent. Bad pixels are marked in dark gray and the color bar range is ± 0.5%. The unit of NRMSE is in percent divided by ~~the ~~mean radiance.

Figure 9~~12~~ shows the Defect 1 area where ~~the ~~differences between GEMS and ML radiances are within ~~within ~~about 5%~~ of the GEMS radiances~~. It ~~also ~~shows that dark targets (clear sky with ~~small ~~low radiance) show a positive difference while bright targets (mostly cloudy sky with ~~large ~~high radiance~~s~~) show ~~an ~~the opposite~~ tendency~~. The tendency~~ies~~ is~~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%.

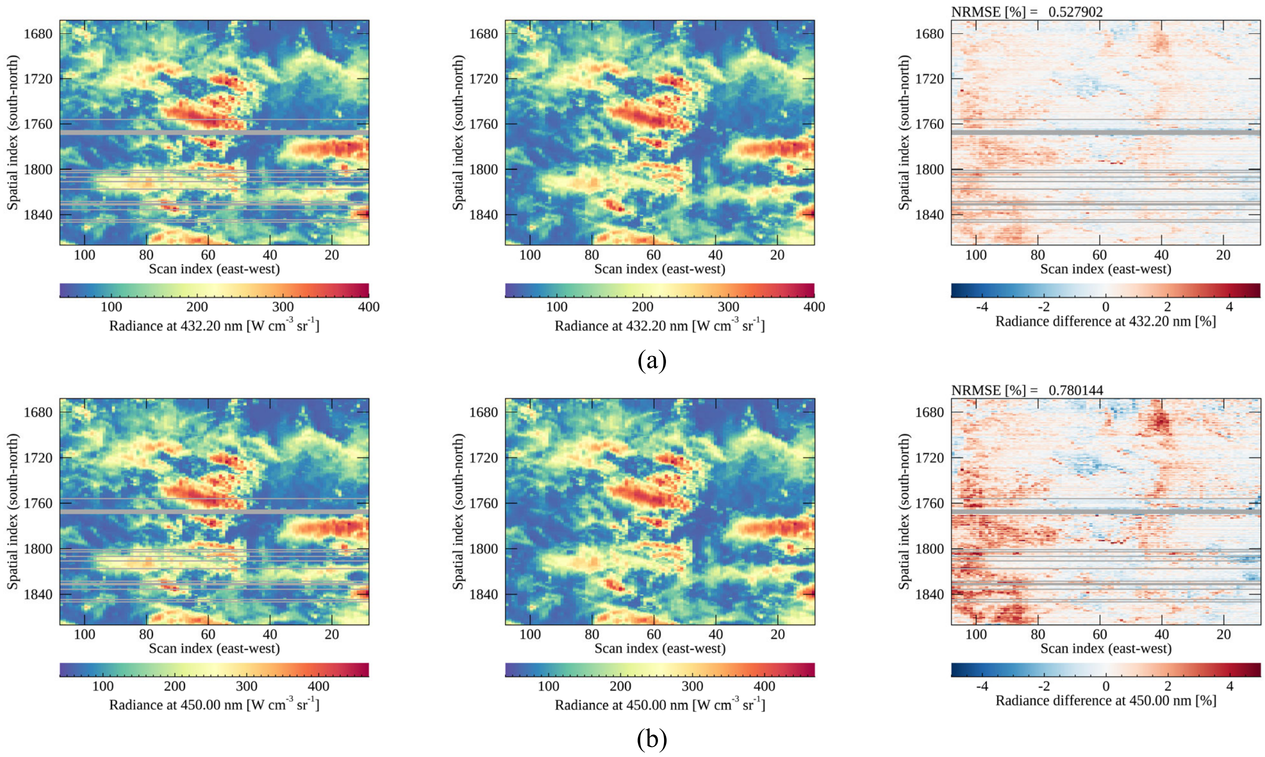


Figure 9 Same as Fig. 8 for the Defect 1 area with the color bar range of 5%.

~~**Figure 12** 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. 6-7~~8-9~~). Both Defects 2-3 have the output spectral ranges of ~~Defects 2-3 are ~~about ~~around ~~100 nm but it seems the output radiances near 300 nm for Defect 2 need more information to be successfully reproduced. The stripping features found in Fig. 10~~3~~b become~~s~~ significant at 312 nm for the ML radiances on the contrary to the radiances at 356.8~~7.2~~ nm in Fig.10~~3~~a. The stripping features seem~~s~~ to be added during the reproducing process especially for shorter wavelengths, and the reason is still unclear. We suspected that unpredictable noises from the instrument would cause the features and it seems more distinguishable in low signals. The scene dependence found in Fig. 9 is also dominant in Fig. 10 at shorter wavelengths, but with the opposite tendency. It is also found that some areas are undetected as bad pixels causing big differences over the areas close to the center in Fig. 10. ~~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.~~

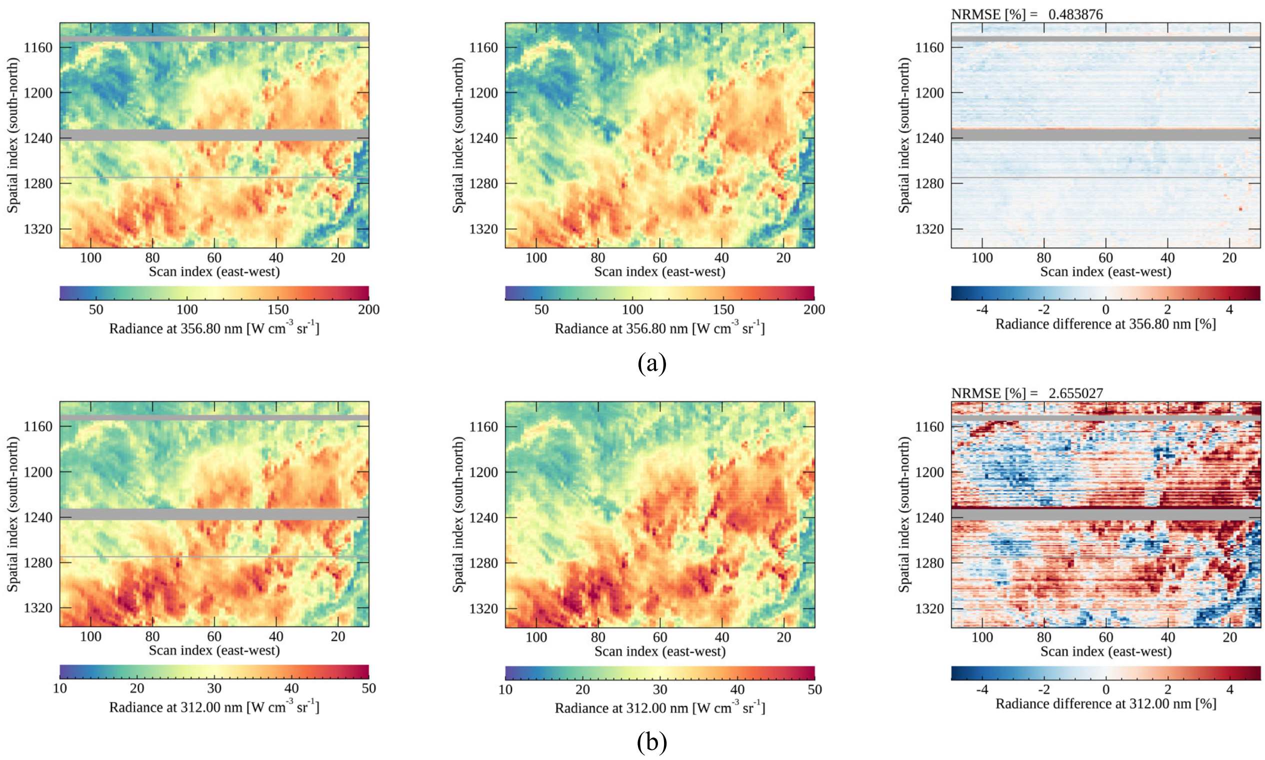
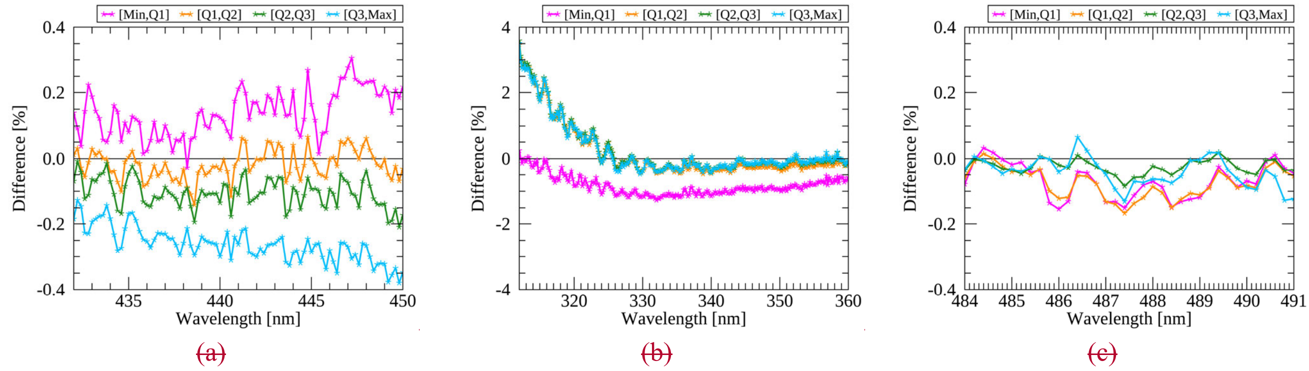


Figure 10 Same as Fig. 8 for the Defect 2 area with the color bar range of 5%.

~~**Figure 13** Same as Fig. 11 for the Defect 2 area with the color bar range of 10%.~~

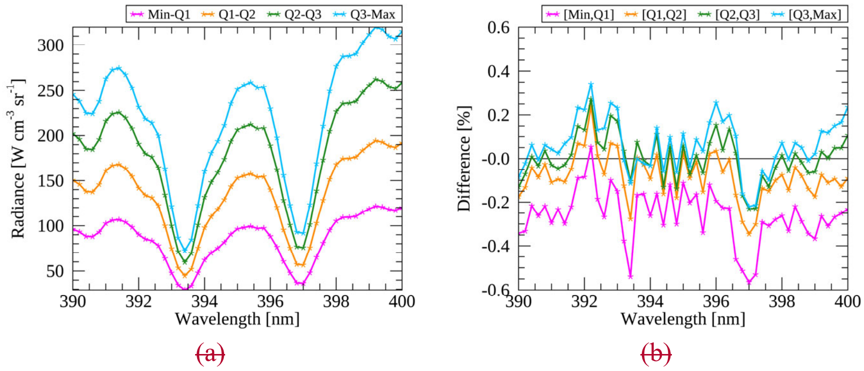
~~A closer inspection is performed to analyze the general spectral features over target wavelengths. Within each defect area in Figs. 11-13, the collected spectra are divided into four groups considering that ML radiances could have different systematic biases depending on the scene brightness as shown in Fig. 11. Figure 14a shows that the ML radiances for the Defect 1 area 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%~~



~~**Figure 14** 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.~~

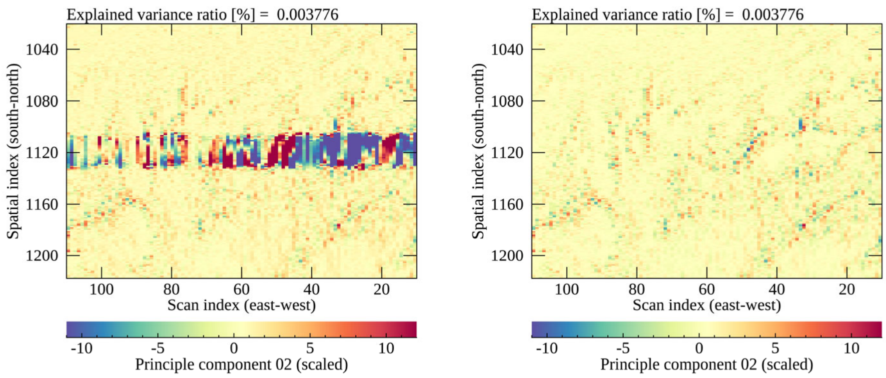


~~**Figure 15** (a) Mean ML radiances (b) and the difference with GEMS radiances at Fraunhofer lines for the Defect 2 area. Each color indicates the average for each quartile and Q1, Q2 and Q3 represent the first, second and third quartile, respectively. The difference is calculated between the ML and 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 reproduced spectral patterns before the retrieval process, we apply PCA to GEMS radiances collected within each area in Fig. 8-10~~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~~ defect-free spectra. Figure 11~~5~~ 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. The Defect 3 area having the widest defective width is ~~targeted~~ presented 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, t~~The radiances reproduced with spatial interpolation on the bad pixel area are projected together as shown in Fig.11~~5~~a, and it seems ~~. As assumed, the PC scores from reproduced spectra with spatial interpolation show~~ disparate values are found over~~-~~ the bad pixel area~~because the spectral patterns of the interpolated spectra are inconsistent with the patterns of normal spectra~~. The ML radiances in Fig. 11~~5~~b show spatially homogenous PC scores on the contrary which indicates that the machine learning methods could ~~properly~~ reproduce ~~the~~ dominant spectral patterns, in this case for the second PCs~~, in the case of the second PC~~.



(a) (b)

Figure 11 The second PC scores of (a) GEMS radiances and (b) ML radiances on the target area for Defect 3. The PC is scaled for ~~clear~~ clarity of presentation.

The dominant spectral patterns for each PC are presented in Fig. 12~~6~~ with the eigenvector matrix constructed from GEMS radiances for the ~~target~~ specified target wavelengths in Table 3~~wavelengths of Defects 1-3~~. Each color indicates the eigenvector~~s~~ for the first-sixth PCs ~~which determine the contribution~~contributing to total radiances ~~of radiances~~ at each wavelength~~s for each PC subspaces~~. Li et al. (2015) verified~~s~~ that the leading PCs from the UV/VIS backscattered radiation (shorter than 360 nm) ~~are highly correlated with~~mainly represent dominant absorption features and surface properties, ~~and~~ while the trailing PCs might be associated with instrument artifacts and other unresolved spectral features with PCA as similarly shown in Fig. 12. ~~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.~~

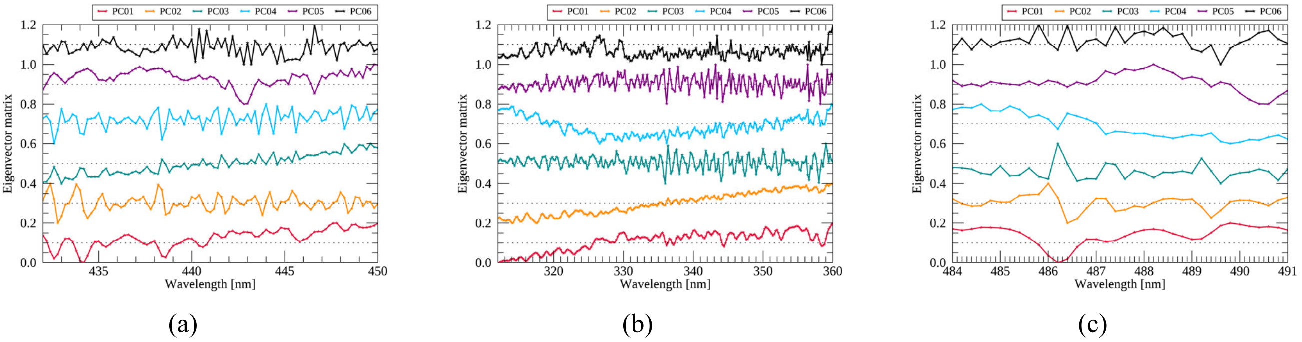


Figure 12~~16~~ Eigenvector of the first-sixth PCs applied to GEMS radiances for the target wavelengths of (a) Defects 1, (b) Defect 2 and (c) Defect 3. All eigenvectors are scaled (min-max scaling) and shifted for clarity of presentation.

As presented in Table 4, ~~the comparison of~~comparing PC scores ~~could indirectly~~ provides qualitative ~~the~~ information on the ~~similarity~~ effectiveness of the suggested method~~the dominant patterns between ML and GEMS radiances with the correlation coefficient~~. The results show that the ~~the~~ mean spectral pattern (the first PC) and some dominant patterns ~~could~~ can be sufficiently ~~well~~ 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. Given the explained variance ratio, each PC except the first one may contribute to a small extent to total radiances but it could be enough to determine subtle spectral patterns important for retrieval processes. The interesting finding is that only for Defect 3, the leading PCs having relatively higher explained variance ratio show high correlation coefficients over 0.95. 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 process. ~~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 4.** 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~~

Table 4 Correlation coefficients (Corr.) of PC scores of GEMS and ML radiances and explained variance ratios (EVR) of GEMS radiances for each target region in Fig. 8-10 excepting bad pixel area.

PC	**Defect 1**		**Defect 2**		**Defect 3**	
	Corr.	**EVR**	**Corr.**	**EVR**	**Corr.**	**EVR**

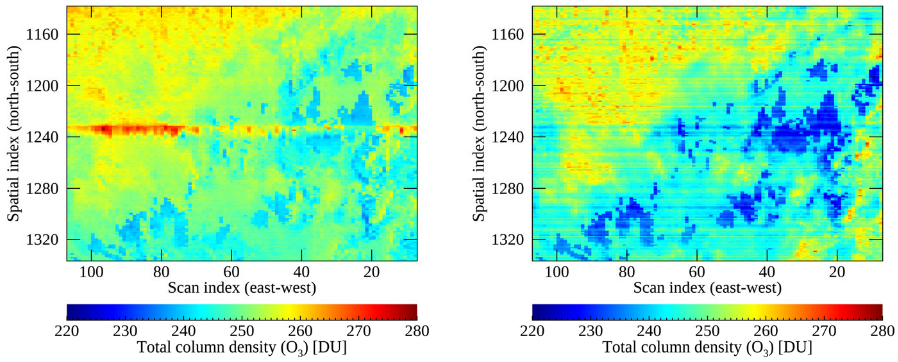
1	0.9999	99.9906	0.9998	99.9504	1.0000	99.9953
2	0.9983	0.0070	0.8672	0.0294	0.9976	0.0038
3	0.8511	0.0007	0.9857	0.0135	0.9863	0.0003
4	0.9731	0.0006	0.5469	0.0019	0.8147	0.0001
5	0.6646	0.0001	0.8454	0.0012	0.6079	0.0001
6	0.7999	0.0001	0.7197	0.0005	0.7815	0.0001

3.3 Level 2 retrieval results

3.3.1 Cloud and ozone retrieval

In the previous section, it was found that the overall prediction error is about 5% except for 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 O_2-O_2 absorption lines related to Defect 3 has the smallest 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 presented in Fig. 8. The difference of cloud centroid pressure retrieved with ML and GEMS spectra is about 1% on average while the cloud properties retrieved with ML spectra have weak stripping 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.

For qualitative investigation of the replaced radiances at ozone absorption lines having high uncertainties, the reproduced spectra presented in Fig. 10 are applied to the ozone retrieval algorithm of GEMS. Figure 13 shows total ozone column density with un-flagged 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 even for the surrounding areas. 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 much higher variance. The results indicate that the spatial distribution could be approximated with reproduced spectra, but the accuracy could not be guaranteed especially for obtaining an exact retrieval value.



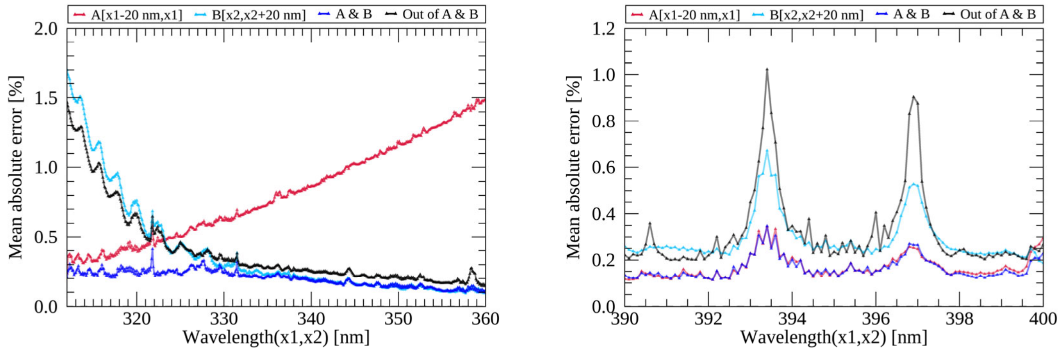
(a) (b)

Figure 13 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).

3.3.2 Cause analysis for further application

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 defect-free measurements in March 2021 (around 80,000 spectra after bad pixel masking and the elimination of saturation pixels).

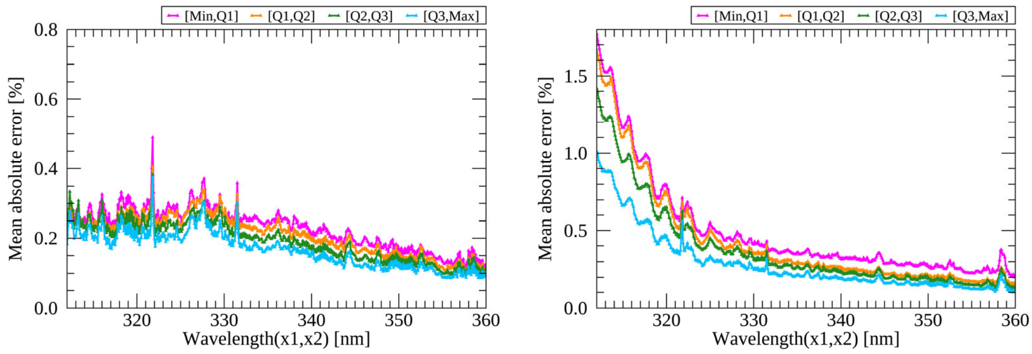
Figure 14 presents mean absolute errors for ozone absorption and Fraunhofer lines with different input conditions including or not the near sides of the output spectral bands (within the range of 20 nm). As assumed, prediction errors increase at the spectral peaks and overall error patterns along the output wavelengths largely differ depending on input conditions. It is clearly shown that the errors are higher when reproduced with farther input spectral bands from output spectral lines. In Fig. 14a, the similar input condition (360-500 nm) with Defect 2 is plotted in black lines 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 determining the accuracy 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.



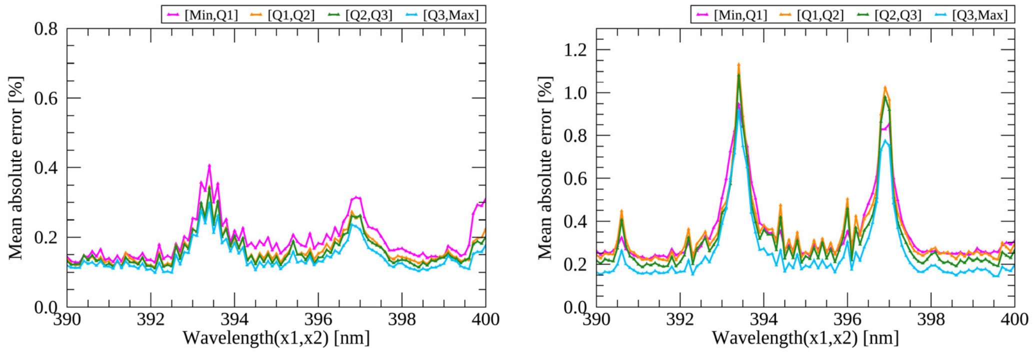
(a) (b)

Figure 14 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 15 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) 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 O_2-O_2 absorption lines presenting successful cloud retrieval results.

4 Conclusions

GEMS is an environmental sensor measuring hyperspectral radiances from 300 to 500 nm in the Asia-Pacific region for timely atmospheric monitoring. During the IOT of GEMS, one of calibration issues was found that erroneous values of bad pixels on the detector array are not properly replaced with spatial interpolation, the current operational method of GEMS. It is clear that when the bad pixel area is too large, the spatial interpolation tends to cause high interpolation error especially for a scene having large spatial inhomogeneity (i.e. cloud edges). The high interpolation error of bad pixels could affect the retrieval process, which causes horizontal discontinuity at a certain latitude for the retrieval of Level 2 products.

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. ~~To resolve the issue, this study suggests machine learning methods using PCA-ANN and PCA-Linear to fill in the spectral gaps caused by bad pixels, denoted as Defects 1-3 in this study.~~ 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 ~~corresponds~~ 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.

~~Considering that input radiances have strong linear relations, dimensionality reduction with PCA is applied in the pre-processing step to reduce linear relations of input radiances and to increase computational efficiency of training process.~~ 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 ~~with input radiances~~ (Defect 2). ~~. The narrower the spectral range of output radiances is, the smaller the prediction error is because the prediction error of Defect 3 is around 0.5%, while it is around 5% for Defects 1 and 2 except for the shorter wavelengths around 300 nm.~~ 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. ~~The comparison results between measured and reproduced radiances show that the dominant spectral patterns could be successfully reproduced for the spectral gap-filling mostly within 5%, while the spectral patterns determined by very small signals such as the Ring effect would be insufficiently reproduced with the suggested methods. The extracted spectral patterns using PCA also present the similar results showing highly correlated PC scores for the first PC for Defect 1-3 regardless of the models, while the PCs determining subtle spectral features are relatively less correlated especially for Defect 1-2.~~

~~To apply the methods in operation, however, it needs to be updated further to solve the following issues. The machine learning model, especially the PCA-ANN model, becomes highly unstable when measurement characteristics of training and test data significantly change.~~ This indicates that~~If measurements have high seasonal dependence, then the time lag between training and test data should be as shorter as possible to guarantee that both data are sampled from an identical population considering the basic premise of ANN. It is empirically found that the time lag between training and test data should not be over two weeks for GEMS which could be technically demanding in operation. Secondly, the radiance at shorter wavelengths (around 325 nm) of Defect 2 has high prediction error of over 5%, which is higher than the level of radiometric calibration accuracy of GEMS (4%). To increase prediction accuracy at strong absorption lines and describe precise spectral features,~~

additional information ~~would~~ would be~~be~~ needed ~~besides~~ except for the spectral relations of radiances ~~in a spectrum~~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).~~, machine learning methods suggested in this study can be a useful tool for filling in a spectral gap and increasing the number of data reserving measurement characteristics of the sensor.~~ 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.

Author contribution

M.-H.A. conceptualized and supervised the study; Y.L. conducted the research, performed the experiments and prepared the manuscript; M.K. contributed to the editing of the manuscript and developing methodology. M.E. contributed to the pre-processing of raw data.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgements

We wish to express our gratitude to Dr. Glen Jaross and the anonymous reviewer for their valuable comments to greatly improve the quality of this research. We also thank Dr. Kang-Hyeon Baek (Pusan National University) and Gyuyeon Kim (Ewha Womans University) for their assistance in retrieving Level 2 data for this study. ~~We~~ The authors acknowledge the contribution of ~~would like to thank~~ the Environment Satellite Center (ESC) of National Institute of Environmental Research (NIER) for providing GEMS Level 0-1C~~0-2~~ data.

Financial support

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2018R1A6A1A08025520).

References

Bajorski, P.: Statistical inference in PCA for hyperspectral images, IEEE J. Sel. Top. Signal Process., 5(3), 438–445, doi:10.1109/JSTSP.2011.2105244, 2011.

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.

Boersma, K. F., Eskes, H. J. and Brinksma, E. J.: Error analysis for tropospheric NO2 retrieval from space, J. Geophys. Res. D Atmos., 109(4) [online] Available from: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2003jd003962 (Accessed 20 September 2019), 2004.

Boldrini, B., Kessler, W., Rebner, K. and Kessler, R.: Hyperspectral imaging: a review of best practice, performance and pitfalls for inline and online applications, J. Near Infrared Spectrosc., 20(5), 438, doi:10.1255/jnirs.1003, 2012.

Bovensmann, H., Burrows, J. P., Buchwitz, M., Frerick, J., Noël, S., Rozanov, V. V., Chance, K. V. and Goede, A. P. H.: SCIAMACHY: Mission objectives and measurement modes, J. Atmos. Sci., 56(2), 127–150, doi:10.1175/1520-0469(1999)056<0127:SMOAMM>2.0.CO;2, 1999.

Burger, J.: Replacement of Hyperspectral Image Bad Pixels, NIR news, 20(7), 19–21, doi:10.1255/nirn.1151, 2009.

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.

Cybenko, G.: Approximation by superpositions of a sigmoidal function, Math. Control. Signals, Syst., 2(4), 303–314, doi:10.1007/BF02551274, 1989.

Dorvlo, A. S. S., Jervase, J. A. and Al-Lawati, A.: Solar radiation estimation using aritificial neural networks, Appl. Energy, 71(4), 307–319, doi:10.1016/S0306-2619(02)00016-8, 2002.

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.

Fischer, A. D., Downes, T. V. and Leathers, R.: Median spectral-spatial bad pixel identification and replacement for hyperspectral SWIR sensors, in Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIII, vol. 6565, edited by S. S. Shen and P. E. Lewis, p. 65651E, SPIE., 2007.

Gewali, U. B., Monteiro, S. T. and Saber, E.: Machine learning based hyperspectral image analysis: A survey, arXiv, 2018.

Goetz, A. F. H., Vane, G., Solomon, J. E. and Rock, B. N.: Imaging spectrometry for earth remote sensing, Science (80-.)., 228(4704), 1147–1153, doi:10.1126/science.228.4704.1147, 1985.

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.

Hedelt, P., Efremenko, D. S., Loyola, D. G., Spurr, R. and Clarisse, L.: Sulfur dioxide layer height retrieval from Sentinel-5 Precursor/ TROPOMI using FP_ILM, Atmos. Meas. Tech, 12, 5503–5517, doi:10.5194/amt-12-5503-2019, 2019.

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.

Hornik, K., Stinchcombe, M. and White, H.: Multilayer feedforward networks are universal approximators, Neural Networks, 2(5), 359–366, doi:10.1016/0893-6080(89)90020-8, 1989.

Howell, S. B.: CCD imaging, in Handbook of CCD Astronomy, pp. 66–101, Cambridge University Press, Cambridge., 2006.

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.

Kang, M., Ahn, M. H., Liu, X., Jeong, U. and Kim, J.: Spectral calibration algorithm for the geostationary environment monitoring spectrometer (Gems), Remote Sens., 12(17), 1–17, doi:10.3390/rs12172846, 2020.

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.

Kieffer, H. H.: Detection and correction of bad pixels in hyperspectral sensors, in Hyperspectral Remote Sensing and Applications, vol. 2821, edited by S. S. Shen, pp. 93–108, SPIE., 1996.

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.

Kim, J., Jeong, U., Ahn, M. H., Kim, J. H., Park, R. J., Lee, H., Song, C. H., Choi, Y. S., Lee, K. H., Yoo, J. M., Jeong, M. J., Park, S. K., Lee, K. M., Song, C. K., Kim, S. W., Kim, Y. J., Kim, S. W., Kim, M., Go, S., Liu, X., Chance, K., Miller, C. C., Al-Saadi, J., Veihelmann, B., Bhartia, P. K., Torres, O., Abad, G. G., Haffner, D. P., Ko, D. H., Lee, S. H., Woo, J. H., Chong, H., Park, S. S., Nicks, D., Choi, W. J., Moon, K. J., Cho, A., Yoon, J., Kim, S. kyun, Hong, H., Lee, K., Lee, H., Lee, S., Choi, M., Veefkind, P., Levelt, P. F., Edwards, D. P., Kang, M., Eo, M., Bak, J., Baek, K., Kwon, H. A., Yang, J., Park, J., Han, K. M., Kim, B. R., Shin, H. W., Choi, H., Lee, E., Chong, J., Cha, Y., Koo, J. H., Irie, H., Hayashida, S., Kasai, Y., Kanaya, Y., Liu, C., Lin, J., Crawford, J. H., Carmichael, G. R., Newchurch, M. J., Lefer, B. L., Herman, J. R., Swap, R. J., Lau, A. K. H., Kurosu, T. P., Jaross, G., Ahlers, B., Dobber, M., McElroy, C. T. and Choi, Y.: New era of air quality monitoring from space:

Geostationary environment monitoring spectrometer (GEMS), Bull. Am. Meteorol. Soc., 101(1), E1–E22, doi:10.1175/BAMS-D-18-0013.1, 2020.

Kingma, D. P. and Ba, J. L.: Adam: A method for stochastic optimization, in 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, International Conference on Learning Representations, ICLR. [online] Available from: https://arxiv.org/abs/1412.6980v9 (Accessed 7 May 2021), 2015.

Le, T., Liu, C., Yao, B., Natraj, V. and Yung, Y. L.: Application of machine learning to hyperspectral radiative transfer simulations, J. Quant. Spectrosc. Radiat. Transf., 246, 106928, doi:10.1016/j.jqsrt.2020.106928, 2020.

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.

Liu, X., Smith, W. L., Zhou, D. K. and Larar, A.: Principal component-based radiative transfer model for hyperspectral sensors: Theoretical concept, Appl. Opt., 45(1), 201–209, doi:10.1364/AO.45.000201, 2006.

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.

Lo´pez-Alonso, J. M. and Alda, J.: Bad pixel identification by means of principal components analysis, Opt. Eng., 41(9), 2152, doi:10.1117/1.1497397, 2002.

Loyola, D. G., Gimeno García, S., Lutz, R., Argyrouli, A., Romahn, F., Spurr, R. J. D., Pedergnana, M., Doicu, A., Molina García, V. and Schüssler, O.: The operational cloud retrieval algorithms from TROPOMI on board Sentinel-5 Precursor, Atmos. Meas. Tech, 11, 409–427, doi:10.5194/amt-11-409-2018, 2018.

Ludewig, A., Kleipool, Q., Bartstra, R., Landzaat, R., Leloux, J., Loots, E., Meijering, P., Van Der Plas, E., Rozemeijer, N., Vonk, F. and Veefkind, P.: In-flight calibration results of the TROPOMI payload on board the Sentinel-5 Precursor satellite, Atmos. Meas. Tech., 13(7), 3561–3580, doi:10.5194/amt-13-3561-2020, 2020.

Manolakis, D., Pieper, M., Truslow, E., Lockwood, R., Weisner, A., Jacobson, J. and Cooley, T.: Longwave infrared hyperspectral imaging: Principles, progress, and challenges, IEEE Geosci. Remote Sens. Mag., 7(2), 72–100, doi:10.1109/MGRS.2018.2889610, 2019.

Nwankpa, C., Ijomah, W., Gachagan, A. and Marshall, S.: Activation Functions: Comparison of trends in Practice and Research for Deep Learning, arXiv [online] Available from: http://arxiv.org/abs/1811.03378 (Accessed 8 May 2021), 2018.

Pan, C., Weng, F. and Flynn, L.: Spectral Performance and Calibration of the Suomi NPP OMPS Nadir Profiler Sensor, Earth Sp. Sci., 4(12), 737–745, doi:10.1002/2017EA000336, 2017.

Pan, C., Zhou, L., Cao, C., Flynn, L. and Beach, E.: Suomi-NPP OMPS Nadir mapper's operational SDR performance, IEEE Trans. Geosci. Remote Sens., 57(2), 1015–1024, doi:10.1109/TGRS.2018.2864125, 2019.

Pan, C., Yan, B., Cao, C., Flynn, L., Xiong, X., Beach, E. and Zhou, L.: Performance of OMPS Nadir Profilers' Sensor Data Records, IEEE Trans. Geosci. Remote Sens., 1–9, doi:10.1109/tgrs.2020.3026586, 2020.

Rankin, B. M., Broadwater, J. B. and Smith, M.: Anomalous pixel replacement and spectral quality algorithm for longwave infrared hyperspectral imagery, Int. Geosci. Remote Sens. Symp., 2018-July, 4316–4319, doi:10.1109/IGARSS.2018.8517461, 2018.

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.

Schläpfer, D., Nieke, J. and Itten, K. I.: Spatial PSF nonuniformity effects in airborne pushbroom imaging spectrometry data, IEEE Trans. Geosci. Remote Sens., 45(2), 458–468, doi:10.1109/TGRS.2006.886182, 2007.

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.

Taylor, M., Kosmopoulos, P. G., Kazadzis, S., Keramitsoglou, I. and Kiranoudis, C. T.: Neural network radiative transfer solvers for the generation of high resolution solar irradiance spectra parameterized by cloud and aerosol parameters, J. Quant. Spectrosc. Radiat. Transf., 168, 176–192, doi:10.1016/j.jqsrt.2015.08.018, 2016.

Wu, W., Liu, X., Xiong, X., Li, Y., Yang, Q., Wu, A., Kizer, S. and Cao, C.: An Accurate Method for Correcting Spectral Convolution Errors in Intercalibration of Broadband and Hyperspectral Sensors, J. Geophys. Res. Atmos., 123(17), 9238–9255, doi:10.1029/2018JD028585, 2018.

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.

Zarzalejo, L. F., Ramirez, L. and Polo, J.: Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index, in Energy, vol. 30, pp. 1685–1697, Elsevier Ltd., 2005.

Zhu, S., Lei, B. and Wu, Y.: Retrieval of hyperspectral surface reflectance based on machine learning, Remote Sens., 10(2), 1–15, doi:10.3390/rs10020323, 2018.