Spectral Calibration Algorithm for the Geostationary Environment Monitoring Spectrometer (GEMS)

Mina Kang¹, Myoung-Hwan Ahn²*, Xiong Liu³, Ukkyo Jeong⁴,⁵, and Jhoon Kim⁶

¹ Department of Atmospheric Science and Engineering, Ewha Womans University, 52 Ewhayeodae-Gil, Seoul, 03760, Korea; mina@ewhain.net
² Department of Climate and Energy Systems Engineering, Ewha Womans University, 52 Ewhayeodae-Gil, Seoul, 03760, Korea; terryahn65@ewha.ac.kr
³ Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts, 02138, USA; xliu@cfa.harvard.edu
⁴ Earth System Science Interdisciplinary Center, University of Maryland, 5825 University Research Ct, College Park, Maryland, 20740, USA; ukkyo.jeong@nasa.gov
⁵ Goddard Space Flight Center, NASA, 8800 Greenbelt Road, Maryland, 20771, USA
⁶ Department of Atmospheric Science, Yonsei University, Yonsei-Ro 50, Seoul, 03722, Korea; jkim2@yonsei.ac.kr

*Correspondence: terryahn65@ewha.ac.kr; Tel.: +82-2-3277-4462

Abstract: The Geostationary Environment Monitoring Spectrometer (GEMS) onboard the Geostationary Korean Multi-Purpose Satellite 2B was successfully launched in February 2020. GEMS is a hyperspectral spectrometer measuring solar irradiance and Earth radiance in the range of 300 to 500 nm. This paper introduces the spectral calibration algorithm for GEMS, which uses a nonlinear least-squares approach. To assess the performance of the algorithm, sensitivity tests for a series of spectral parameters such as shift, spectral range for fitting, signal-to-noise ratio, spectral response function (SRF), and reference spectrum have been conducted. To improve the assessment, a synthetic GEMS spectrum using the prelaunch GEMS SRF is adopted here. The test results show that the required accuracy (0.002 nm) is achievable for the expected uncertainties of the parameters except for the SRF and the choice of high-resolution reference spectrum, which degrade the algorithm performance by an order magnitude. To mitigate the sensitivity to SRF, retrieval of in-orbit SRF using an analytic function is suggested. Finally, a few candidates for the high-resolution solar reference spectrum are prepared for testing by the instrument during in-orbit tests.

Keywords: GEMS; spectral calibration; hyperspectral instrument;

1. Introduction

The Geostationary Environment Monitoring Spectrometer (GEMS) is a hyperspectral instrument developed for Korea’s next-generation geostationary multi-purpose satellite program, the GK-2 program, which consists of two satellites, GK-2A and GK-2B. They are collocated at 128.2° E over the equator, where the current GK-1 satellite is stationed. While the GK-2A satellite is dedicated to weather monitoring with a high-performing imaging instrument (16-channel imagers in the
visible and infrared bands [1]), the GK-2B satellite is for ocean and environmental observation [2-3]. The ocean mission is a follow-on mission of the current Geostationary Ocean Color Imager (GOCI) of the GK-1 with an improved GOCI-II spectrometer, while the environmental mission is a new in the geostationary. The GEMS is the key instrument for the environmental mission, as it provides information on trace gases such as ozone, nitrogen dioxide, sulfur dioxide, and formaldehyde, as well as aerosols [2-3]. Geostationary orbit allows the GEMS instrument to provide unprecedented spatio-temporal resolution of air quality information, at least eight times per day with about 3.5 km × 8 km spatial resolution from space. The field-of-regard (FOR) of GEMS, from 5° S to 45° N latitude and from 75° E to 145° E longitude, covers the areas of major pollution sources across East and Southeast Asia. Its spectral range includes ultraviolet and visible wavelengths, from 300 to 500 nm, and thus covers the absorption lines of the listed trace gases.

The primary products of GEMS are derived from geo-located and calibrated radiance and irradiance data—level 1b (L1b) data—using algorithms including differential optical absorption spectroscopy (DOAS), optimal estimation, principle component analysis, and look up tables [4-6]. The accuracy of the retrieval algorithms depends on the quality of the L1b data, many of which are highly sensitive to wavelength uncertainty [7-9]. Furthermore, the measurement signals of several trace gases such as sulfur dioxide and formaldehyde are quite low, and wavelength uncertainty could significantly degrade algorithm performance. For example, systematic uncertainties of the slant columns of formaldehyde for spectral calibration have been shown to be 0.92 ×10^{15} molecules/cm^2 [5]. Therefore, accurate and reliable spectral calibration of the L1b data is essential for derived information to be meaningful and reliable [10-12]. Currently, the operational L1b process for GEMS is prepared to provide a spectral accuracy of 0.02 nm (one tenth of the spectral sampling resolution of GEMS), which is expected to be sufficient for many applications. However, the requirements for trace gas and aerosol retrieval are much higher (0.002 nm, an order of magnitude higher); therefore, it is imperative to prepare an improved spectral calibration algorithm, hereafter called SPECAL.

SPECAL is based on nonlinear least-squares fitting, similar to that implemented for previous hyperspectral programs such as Global Ozone Monitoring Experiment (GOME-1, 2), Scanning Imaging Absorption spectroMeter for Atmospheric CHartography (SCIAMACHY), Ozone Monitoring Instrument (OMI), and Ozone Monitoring Profiler Suite (OMPS) [13-18]. The method basically finds the best fitting between the measured spectrum and a reference spectrum, such as a high-resolution solar reference, by adjusting spectral parameters such as shift, squeeze, and full width at half maximum (FWHM) of the spectral response function (SRF) [15, 17-18]. While the basic approach is the same, the specific fitting parameters and algorithm configuration including spectral fitting ranges are program specific. For example, the GOME program uses the shift and squeeze parameters for whole spectral range, while the OMI program uses multiple spectral fitting windows. On the other hand, the OMPS program adopts two different fitting parameter bases: spectral shift including the FWHM of SRF.
SPECAL is designed to account for the space environment and the operational concept of GEMS, which is different from those of polar orbiting programs. For example, scene inhomogeneity, an important source of wavelength uncertainty in a polar orbiter [16] is not relevant considering the scanning mechanism (step and dwell) and the fixed, geostationary platform. As this type of iterative fitting method requires significant computational resources, SPECAL must be optimized and instrument-specific to meet its requirements while incurring reasonable numerical costs. Our preliminary study [19] examined the significance of several spectral fitting parameters for SPECAL. Here, we describe further developments and improvements of SPECAL, comprising improved sensitivity and error analysis using actual measurement data in addition to the theoretical SRF functions derived from the measurement data. The new version is also shown to be numerically more efficient by reducing the required number of iterations. Section 2 presents a detailed description of the GEMS SPECAL method and data used, including processes used for the sensitivity test. Section 3 reports the performance of SPECAL along with the impacts on accuracy of uncertainties in the spectral parameters. Concluding remarks follow in Section 4.

2. Data and Method

2.1. GEMS Instrument

GEMS is a space-borne, nadir-viewing, imaging spectrometer, developed by the Korea Aerospace Research Institute and Ball Aerospace Technology Company to measure radiation reflected from the earth and direct solar irradiance in ultraviolet and visible wavelengths with 0.2 nm spectral sampling and about 0.6 nm FWHM resolution. While solar irradiance is to be measured once per day at night, Earth radiance is to be observed every hour during the day, from 9:00 to 16:00 Korea Standard Time. Note that the actual observation duration is only 30 min, even though the observation frequency is hourly, mainly due to GOCI-II’s stringent requirement of pointing stability. Therefore, the two payloads share the duty cycle, alternating observation and rest every half hour.

GEMS is designed to have the same optical paths for both Earth radiance and solar irradiance: Light passes through a calibration assembly, scan mirror, telescope, spectrometer, and detectors. The calibration assembly houses two diffusers consisting of a multi-layer structure including two panes of UV-grade ground fused silica to reduce the sensitivity on solar incident angle by flattening the bidirectional transmission distribution function. For daily solar irradiance measurements, the working diffuser is deployed, while the other diffuser is used semiannually as a reference to monitor changes in the working diffuser over the mission lifetime. The trends in the solar measurements from the two diffusers would suggest changes in the sensor throughout, given the assumption that the reference diffuser is stable. When the diffusers are not in use, they are placed inside the calibration housing to reduce exposure and contamination. As the diffuser is positioned in front of the scan
mirror, the solar irradiance measurements pass through the same optical light path used in Earth radiance measurements. Therefore, for retrieval algorithms using Earth reflectance, which is the ratio between the observed solar irradiance and Earth radiance, many of the calibration uncertainties are expected to cancel out.

The two-axis scan mirror with position-sensing heads ensures GEMS scans the entire FOR and controls line-of-sight motion. It scans from east to west with a fixed north–south field of view of approximately 7.78°. During the 30 min observation period, the scanner moves and stops 704 times and, once stopped, observations are performed for 2.53 s to cover one slit, resulting in the full FOR being observed from the 701 slits. The reflected radiation by the scan mirror is passed into a three-mirror Schmidt telescope, which projects light onto the spectrometer’s entrance slit. The Offner spectrometer with a high-precision grating system disperses the radiation, and illuminates onto a single two-dimensional charge coupled device (CCD), which has an array of 1033 spectral pixels by 2048 spatial pixels. The measurements imaged onto the CCD are converted to digital counts using a 14-bit analog-to-digital converter. After amplification, these digitized counts are transmitted through the host spacecraft’s instrument control electronics. Scientific data are then passed to a Consultative Committee for Space Data System encoder (protocol 133.0-B-1) and delivered to the GK2B spacecraft over the broadband SpaceWire data line. Choi et al. 2019 [3] provide more details on the GEMS instruments.

2.2. GEMS SRFs

SRFs play important roles, both in SPECAL and in the retrieval algorithms [17]. Essentially, they determine the spectral response of an instrument to the input signal as well as the center wavelength of a detector. The GEMS SRFs are characterized using the responses of GEMS to a series of monochromatic radiation generated by a tunable laser (with accuracy known to be better than 0.006 nm, and less than 5 cm⁻¹ of the FWHM) at multiple viewing angles with a discrete number of wavelengths [2]. The seven nominal wavelengths (301.8, 330, 365, 390, 435, 470, and 498.2 nm) are selected considering the retrieval windows used for the retrieval algorithms. The measured SRFs at each specific location and wavelengths are interpolated to derive a complete set of SRFs for each CCD pixel. Smooth variations of the SRF characteristics along both the spatial and spectral directions introduce negligible uncertainties in the interpolated SRFs. However, it is anticipated that the in-orbit SRF characteristics could vary owing to factors arising during the launch and/or operation in space [20-24]. For example, variations of the optics’ temperatures, the high and low frequencies all together, and instrument degradation could alter the effective shape and width of the SRFs [25]. As discrepancies between the assumed and actual SRF introduce significant fitting uncertainties in SPECAL, a measure to characterize in-orbit SRF is essential. This is described in Section 3.3.

2.3. SPECAL Algorithm
The main purpose of SPECAL is to assign accurate wavelength information to each CCD pixel. Initial spectral assignment is obtained during on-ground laboratory calibration by illuminating the instrument field of view with signals from the tunable laser. Similar to a spectral line source such as a PtCrNe hollow cathode lamp, a tunable laser generates monochromatic light to feed GEMS. Furthermore, it is insensitive to Doppler broadening, and provides very sharp, strong emission lines within a certain wavelength range. To determine the spectral band center wavelength for a CCD pixel, each line profile measurement of the tunable laser is fitted using a Gaussian function. The centroids of the line profiles with their wavelength are then fitted to a fourth-order polynomial as a function of the detector pixel. Using the polynomials, spectral registration at all CCD pixels are determined at the ground.

In the space environment, the wavelengths at CCD could be altered due to the changes in the temperature gradient of the spectrometer, the effective SRF caused by an inhomogeneous scene, for example, and the Doppler effect. The Doppler effect is predictable from the relative velocity between the spacecraft and the Sun [14]. Also, as mentioned earlier, spectral changes due to heterogeneous scenes are negligible in GEMS. The variations of optical bench temperature due to internal or environmental heat sources could induce changes in both the alignment of the optics and their optical dispersion [26]. Therefore, the spectral registration of GEMS is primarily perturbed by thermal changes of the instrument.

To account and correct for the spectral deviation, SPECAL is designed to run in two steps during the GEMS mission. The first step of the algorithm derives spectral information from prelaunch and early orbit calibration data, which are measured at various temperatures covering the full range of spectrometer temperatures applicable to the instrument [16]. The wavelength for each CCD pixel is modified as a function of the spectrometer’s temperature during the measurement: It is obtained from the difference in the optical bench temperature between the actual observation conditions and on-ground laboratory measurement ($T_{spect}$ and $T_{spec_ref}$ shown in Figure 1 represent the temperature during the measurement and during the on-ground test, respectively). The first correction is straightforward and fast, but previous studies have found that its calibration accuracy is insufficient for level 2 products such as ozone profiles [14-17]. Therefore, uncertainty induced in SPECAL during the first step can propagate to become unacceptable errors in DOAS-type retrieved products. Therefore, additional SPECAL process and a second step are required for level 1b data quality. Figure 1 shows the flowchart for GEMS SPECAL.
The second step of SPECAL uses cross-correlation [14], which fits the spectral features of the measured solar irradiance or radiance to those of an accurately known high-resolution reference spectrum. It iteratively finds spectral scale parameters (e.g., spectral shift and squeeze) based on nonlinear least-square fitting, optionally with iterative damping constraints (i.e., the Levenberg–Marquardt method) [27]. Using the on-ground calibration data as a first estimate, the algorithm calculates Jacobians with respect to these parameters and the cost function. The wavelength scale is optimally estimated by updating the spectral fitting parameters until the cost function is minimized and the parameters converge. As a reference spectrum for comparison with GEMS measurements, the high-resolution solar reference irradiance spectrum is convolved with the analytical SRF based on the on-ground calibration result. Several high-resolution reference spectra are currently available for the GEMS spectral range. Section 3.4 gives a detailed sensitivity analysis of the sources of reference spectra for SPECAL.

Equation (1) shows how the algorithm simulates measured irradiance \( I_M(\lambda) \) using the reference solar irradiance convolved by fitting parameters.

\[
I_M(\lambda) = I_R(\lambda + \Delta \lambda) \left[ S_0 + S_1 dG + S_2 (dG)^2 + S_3 (dG)^3 \right]
\]  

(1)

Here, \( I_R(\lambda) \) is the reference spectrum modulated and convolved from a high-resolution solar reference spectrum, \( I_M(\lambda) \) with an analytic SRF, as shown in Equation (2). It is compared directly with GEMS measurements \( I_M \) until it converges. The weighting functions for each fitting parameter are calculated at each iteration either analytically or using finite differences, and then used for the inversion.
In Equation 1, \( dG \) means the initial wavelength scale (\( \lambda \)) minus its averaged value (\( \lambda_{\text{avg}} \)), and \( S_i \) (\( i = 0 \)–3) are the coefficients of a third-order scaling polynomial for \( I_\delta \). The Earth radiance is affected by atmospheric absorption (especially by ozone) and scattering (Ring effect, and Rayleigh and Mie scattering). Additional fitting parameters are included for the SPECAL of radiances [15] (not presented in this paper).

To accurately characterize the in-orbit SRFs, the fitting algorithm also includes additional SRF parameters (i.e., shape, width, and asymmetry). For the GEMS SRF mode, we use empirical numerical functions such as asymmetric and broadened Gaussian [22] or Super-Gaussian [28], which are widely used and proven in prior studies [24, 29-30].

To examine the sensitivities of SPECAL to the shape parameters, we use the asymmetric Super-Gaussian (ASG) function [28] as shown in Equation (3).

\[
f(\lambda) = \frac{k}{2w\Gamma(1/k)} \begin{cases} 
\exp\left(\frac{-\Delta\lambda}{w-a_w}\right)^{1-a_k} & \text{for } \lambda \leq 0 \\
\exp\left(\frac{-\Delta\lambda}{w+a_w}\right)^{1+a_k} & \text{for } \lambda > 0
\end{cases}
\]

where, \( k \) and \( w \) are independent parameters that determine the shape and width of \( f(\lambda) \), respectively, \( \Gamma \) is the gamma function, and \( a_w \) and \( a_k \) are additional asymmetry parameters [28]. When \( k \) is 2, the ASG function becomes a Gaussian function. For \( k < 2 \), the function has a more pointed shape with longer tails.

Shift (\( C_h \)) and squeeze (\( C_q \)) are generally used to represent the actual wavelength shift relative to the initial wavelength scale. Shift describes linear changes in the wavelength scale, while squeeze denotes non-linear changes. Thus, the wavelength change for each detector pixel is different in the presence of a squeeze. The wavelength change due to shift and squeeze is modeled as follows:

\[
\Delta\lambda = C_h + dG(C_q - 1).
\]

Based on Equation (4), the actual wavelength grid is the same as the initial grid when the shift is zero and the squeeze is 1 (i.e., no squeeze).

On the other hand, there are frequent disadvantages to utilizing the squeeze parameter for the quantification of actual wavelength changes. When an algorithm calculates a value for squeeze in certain fitting windows, understanding the degree
of change in the whole wavelength range is not straightforward. Additionally, it might not be realistic to separate wavelength changes into a single shift and squeeze, considering the actual wavelength assignment is non-linear and often parameterized as a fifth- to tenth-degree polynomial [31]. Therefore, for simplicity in understanding wavelength changes and for more straightforward utilization in retrieval processes, it is often the case that only the shift parameter is fitted. To describe the changes over the entire observational wavelength range with shift only, we introduce high-order shift coefficients (here, fifth-order polynomial) as shown in Equation (5), to sufficiently account for the non-linear variation of the wavelength change:

$$\Delta \lambda = C_{50} + C_{51} dG + L + C_{52} dG^2 + C_{53} dG^3.$$  \hspace{1cm} (5)

### 2.4. Sensitivity Tests

A series of sensitivity tests is conducted to estimate the performance of the algorithm with respect to uncertainties in the fitting parameters. As the GEMS instrument is in the commissioning phase, observed spectra are currently not available; therefore we use a synthetic spectrum for the GEMS measurement. It is constructed by convolving a high-resolution solar reference spectrum with the measured GEMS SRF, with artificially perturbed wavelength parameters over the spectral range (300 to 500 nm). The synthetic spectrum also includes expected measurement noise, calculated based on the noise model of GEMS. The measured SRF at a nominal wavelength of 365 nm during laboratory calibration and the resulting synthetic GEMS solar irradiance used in this study are shown in Figure 2. For more realistic and precise error estimation of SPECAL, detailed spectral parameters are considered, including signal-to-noise ratio (SNR), wavelength scale parameter, fitting window, SRF, and high-resolution solar reference spectrum. The test results for each spectral parameter are described in the following subsections. (The algorithm is not meaningfully sensitive to the SNR; therefore, the results are not presented here).
Figure 2. (a) Prelaunch GEMS SRF measured at 365.0 nm during laboratory characterization. (b) Synthetic GEMS solar irradiance obtained by convolving a high-resolution solar reference spectrum (Chance and Kurucz, 2010) and the prelaunch SRF.

The test involves running the SPECAL algorithm with given wavelength perturbations (shift, $C_h = 0.001$ nm; squeeze, $C_q = 0.005$; and resolution of the SRF, FWHM, $\sigma = 0.6$ nm). Table 1 shows that the retrieved values of the algorithm are consistent with the given perturbations within the expected uncertainty of the spectral parameters. The performance of the algorithm is also evaluated using mean bias ($B_\lambda$) and root-mean-squared-deviation (RMSD, $S_\lambda$) between the known (true value before perturbation) and retrieved wavelength registration. The calculated $B_\lambda$ and $S_\lambda$ are less than the required spectral accuracy of GEMS (i.e., 0.002 nm), as listed in Table 1. Within expected wavelength perturbations the accuracy of the algorithm is reliable, and SPECAL produces optimal values for the given simulated spectrum and measurement. However, several additional uncertainties are possible with the calibration algorithm. Therefore, we next investigate and quantify the calibration performance resulting from each spectral fitting parameter.

Table 1. Input spectral perturbations and output results using simple spectral scale parameters (i.e., $C_h$ [nm], $C_q$ [unitless], and spectral resolution in $\sigma$ [nm]). The mean bias ($B_\lambda$), root-mean-squared-deviation ($S_\lambda$), and cost function ($\chi^2$) of spectral fitting results are also shown.

<table>
<thead>
<tr>
<th>Given perturbation</th>
<th>Retrieved perturbation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_h$</td>
<td>$C_q$</td>
</tr>
<tr>
<td>0.010</td>
<td>0.005</td>
</tr>
</tbody>
</table>

3. Results and Discussion

3.1. Wavelength Scale Parameter

The wavelength change compared with the initial wavelength scale is represented using wavelength scale parameters such as shift, squeeze, and shift polynomials. To examine the calibration’s sensitivity to the wavelength scale parameters, we apply the parameters differently to the synthetic data and the SPECAL algorithm. The algorithm’s performance corresponding to the shift and squeeze parameters or the shift polynomial is compared for the whole GEMS observational wavelength range.

The obtained cost function and fitted parameter values indicate that the outputs are consistent with the simulated inputs when the algorithm uses a shift polynomial rather than shift and squeeze (Table 2). The algorithm fitting a shift polynomial successfully follows the simulated wavelength change that consists of shift and squeeze, whereas the opposite case does not. As the squeeze parameter is related to the first-order term of the shift polynomial (equations (4) and (5)), fitting a shift and squeeze cannot perfectly capture the actual wavelength variation due to the high-order shifts. Furthermore, fitting a polynomial to account for high-order shifts might
converge more quickly with fewer iterations than fitting using shift and squeeze. When considerable nonlinear wavelength changes are expected, such as in this case where a sufficient number of fitting radiances are required, it is recommended that fitting is applied using shift polynomial parameters.

Table 2. Comparison of inputs and corresponding outputs to show the algorithm’s sensitivity to wavelength scale parameters (using shift/squeeze or shift polynomials).

<table>
<thead>
<tr>
<th>Simulated Input</th>
<th>Algorithm Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_h$</td>
<td>$C_{hq}$</td>
</tr>
<tr>
<td>0.010</td>
<td>0.005</td>
</tr>
<tr>
<td>$Ch_0$</td>
<td>$Ch_{11} \times 10^3$</td>
</tr>
<tr>
<td>0.010</td>
<td>1.00</td>
</tr>
</tbody>
</table>

3.2. Fitting Window

Sensitivity testing shows that the residual value using high-order shifts is smaller than fitting using shift and squeeze. However, a fitting window that is too wide for the wavelength range applicable to the calibration can introduce unacceptable fitting errors in the significant dynamic range of radiance measurements; the errors can be up to about three orders of magnitude in the GEMS spectral domain (i.e., 300–500 nm) [16]. Therefore, if the fitting can be performed sufficiently quickly, it is better to perform SPECAL over several narrow fitting windows across the spectral range rather than fitting a shift polynomial over the entire (single) range. Fitting with multiple sub-windows allows detailed monitoring of the detector response and provides L1b data with sufficient accuracy for specific trace gas retrieval. For each selected sub-window (e.g., 10 or 15 nm), using a single shift parameter ($C_h$) is appropriate, as nonlinear spectral change is typically negligible in this case. The derived shift at each sub-window is then fitted to a polynomial as a function of wavelength to determine the amount of shift for every spectral pixel [14]. The number of sub-windows should be sufficient to obtain reliable results compared with the use of shift polynomials for the entire window. However, increasing the number of windows with a single fit parameter increases the numerical cost for the algorithm to converge. A further consideration is that the polynomial approach applied to derive the shift for each CCD pixel could affect SPECAL’s accuracy. Changing the fitting windows while keeping a constant wavelength range (for example, 400–450 nm to 410–460 nm) has an insignificant effect on the calculation time. Therefore, we focus on the effects of different numbers of fitting windows on the calculation time and algorithm performance.

The effects of two types of polynomial (power and Chebyshev) are compared. The calibration is repeatedly performed over 6 to 10 small sub-windows selected with consideration of the GEMS retrieval range. The following sub-window configurations are used. Window 1: 10 sub-windows of 10 nm width and 10 nm intervals. Window 2: six sub-windows of 30 nm width at 30 nm intervals at 305–335,
335–365, 365–395, 395–425, 425–455, and 455–485 nm. Window 3: also six sub-windows, but selected to include the beginning and end of the wavelength range, at 300–310, 330–340, 370–380, 400–410, 450–460, and 490–500 nm. Results are shown in Figure 3. The orange lines are the applied simulated wavelength shifts ($Ch_0$: 0.01, $Ch_1$: 0.0001, $Ch_2$: 0.00002), and asterisks are fitted shifts from each fitting window. Purple and blue lines represent the derived wavelength shifts based on Chebyshev and power polynomials, respectively. The figure’s right panel (b) shows that the shifts obtained using Chebyshev polynomials achieve better consistency with the simulated wavelength shift than the power polynomials. Table 3 lists the corresponding mean bias and RMSD for each fitting window. The results for fitting windows 1 and 2 show that a sufficient number of fitting windows is required to reflect the actual wavelength variations for the whole wavelength range. Nevertheless, using fewer fitting windows is numerically efficient, particularly when wavelength variability can be explained by low-order nonlinearity that satisfies Nyquist sampling frequency.

<table>
<thead>
<tr>
<th>Algorithm Output</th>
<th>Polynomial Method</th>
<th>Power</th>
<th>Chebyshev</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$BL \times 10^4$</td>
<td>$SL \times 10^4$</td>
<td>$BL \times 10^4$</td>
</tr>
<tr>
<td>Fitting Window 1</td>
<td>12.5</td>
<td>20.7</td>
<td>6.99</td>
</tr>
<tr>
<td>Fitting Window 2</td>
<td>31.5</td>
<td>115</td>
<td>5.05</td>
</tr>
<tr>
<td>Fitting Window 3</td>
<td>2.71</td>
<td>15.8</td>
<td>1.29</td>
</tr>
</tbody>
</table>

Comparing the results of fitting windows 2 and 3, with the same number of sub-windows, shows the importance of the locations of the sub-windows. It is necessary to optimize not only the number of sub-windows but also their ranges. The behavior of the actual wavelength changes observed during the mission will contribute to the selection and optimization of the fitting windows. In addition, the wavelength shift derived for each spectral pixel could depend on the chosen polynomial. While the algorithm’s performance meets the requirement for GEMS using either Chebyshev or power polynomials, the former gives better agreement regardless of the fitting window. (For fitting window 2, $BL$ and $SL$ of the power polynomial are 2.71 and 15.8, respectively, and 1.29 and 5.44, for the Chebyshev polynomial, respectively).
Figure 3. Comparison of wavelength shifts using two polynomials. (a) The derived wavelength shifts for each wavelength based on fitted shift parameters (asterisks) using three different sets of fitting windows and different polynomial fitting methods. The fitting windows for each plot are as follows. Top: 10 sub-windows with 10 nm width at 10 nm intervals. Middle: six sub-windows of 30 nm width at 30 nm intervals. Bottom: six sub-windows of 10 nm width spaced to include the beginning and end of the wavelength range. Orange lines are the simulated wavelength shifts applied in the tests (C<sub>0</sub>: 0.01; C<sub>1</sub>: 0.0001; C<sub>2</sub>: 0.00002), and asterisks are fitted shifts from each sub-window. Purple and blue lines represent the derived wavelength shifts based on Chebyshev and power polynomials, respectively. (b) The difference between the simulated and derived wavelength shifts at each wavelength was obtained using either a power (dotted line) or Chebyshev (dashed line) polynomial. The Chebyshev polynomial gives more consistent shifts with the simulated wavelength shifts.

3.3. SRF

Figure 4 shows results for the sensitivity tests of the SRF with five ASG functions. Each synthetic spectrum is then derived by convolution of a high-resolution solar
reference spectrum with these five SRFs. First, we apply the algorithm to synthetic irradiances convolved with different SRFs in Table 4, without including SRF fitting. The algorithm assumes a fixed symmetric Gaussian SRF. The cost function and calibration results for several synthetic spectra with different SRFs are summarized in Table 4. The algorithm’s sensitivity increases with the changes of shape (k) and width (w) of the SRF. The results indicate that the algorithm is significantly more sensitive to the shape (skewness) of the SRF than its width SRF. This is because the width of the SRF is related to the resolution of the spectrum, while its shape greatly determines the representative spectral characteristics. The algorithm fails to meet the requirements of GEMS, except when the simulated SRF is fully Super-Gaussian (k = 2). Therefore, the uncertainty of the shape of the SRF is a decisive factor determining the algorithm’s accuracy.

![Figure 4](image_url)

**Figure 4.** Variation of SRFs with changing ASG coefficients: w determines the width, and k is related to the shape of the function. When k = 2, the ASG function is identical to a Gaussian. The values of a_0 and a_w for the ASG function are zero.

<table>
<thead>
<tr>
<th>Simulated Input</th>
<th>Algorithm Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>w</td>
<td>k</td>
</tr>
<tr>
<td>0.433</td>
<td>1</td>
</tr>
<tr>
<td>0.360</td>
<td>2</td>
</tr>
<tr>
<td>0.339</td>
<td>3</td>
</tr>
<tr>
<td>0.329</td>
<td>4</td>
</tr>
<tr>
<td>0.323</td>
<td>5</td>
</tr>
</tbody>
</table>

Nevertheless, in-flight characterization of the SRF and its variation/correction can be well described using an empirical function such as ASG, as it can simultaneously fit the shape and width of the SRF with sufficient accuracy. [28-30]. Table 5 shows that the SPECAL algorithm, with fitted SRF parameters, successfully estimates the actual SRF except for the unrealistic change of shape and width of the SRF (when it changes from a fully Super-Gaussian to a Triangular model). The listed biases between the input and derived SRFs from the algorithm (denoted as $B_i$) are...
less than 0.1%. Spectral convolution of the high-resolution solar reference spectrum and interpolation of the SRF to a different wavelength grid are the main factors affecting the computation time. As GEMS observes the Sun during the night, there is enough time to calibrate the data by deriving the SRF. However, when variations in the SRF are not significant in-orbit, using a pre-computed reference spectrum convolved with a known nominal SRF (characterized during commissioning) can reduce the numerical cost.

Table 5. Sensitivity results of the updated algorithm (including SRF parameters) applied to solar irradiances convolved with different SRFs.

<table>
<thead>
<tr>
<th>Simulated Input</th>
<th>Algorithm Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( w )</td>
<td>( k )</td>
</tr>
<tr>
<td>0.433</td>
<td>1</td>
</tr>
<tr>
<td>0.360</td>
<td>2</td>
</tr>
<tr>
<td>0.339</td>
<td>3</td>
</tr>
<tr>
<td>0.329</td>
<td>4</td>
</tr>
<tr>
<td>0.323</td>
<td>5</td>
</tr>
</tbody>
</table>

We also perform a similar sensitivity study for the two asymmetric factors (\( a_w \) and \( a_k \)) when the shape and width parameters (\( w \) and \( k \)) are fixed as 0.360 and 2, respectively. The SRF’s asymmetry is known to closely correlate with shift (\( C_{w0} \)) [19], which is also confirmed in our algorithm results. Both asymmetry parameters are similarly correlated with the shift parameter, shown as Table 6. Their variation effectively leads to a linear wavelength change. Therefore, increasing asymmetry in the SRF further changes the wavelength shift (i.e., \( C_{w0} \) varies from 0.001 to 0.081).

Table 6. Sensitivity results of the algorithm with changes of asymmetries (\( a_w \) and \( a_k \)) in the simulated input SRF.

<table>
<thead>
<tr>
<th>Simulated Input</th>
<th>Algorithm Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_w )</td>
<td>( a_k )</td>
</tr>
<tr>
<td>0.010</td>
<td>0.010</td>
</tr>
<tr>
<td>0.030</td>
<td>0.010</td>
</tr>
<tr>
<td>0.050</td>
<td>0.010</td>
</tr>
<tr>
<td>0.010</td>
<td>0.030</td>
</tr>
<tr>
<td>0.010</td>
<td>0.050</td>
</tr>
</tbody>
</table>

3.4. High-Resolution Solar Reference Spectrum

The high-resolution solar reference spectrum is critical to spectral and radiometric calibrations for satellite-based instruments. Specific applications
include determination of instrument SRFs, accurate radiative transfer calculations, and Ring effect corrections [32-36]. There are various solar reference spectra with different radiometric accuracy, spectral resolution, wavelength range, and other spectral characteristics. To make an appropriate selection for GEMS, it is necessary to consider the radiometric and spectral uncertainty of each reference spectrum. GEMS has the following requirements: the absolute radiometric uncertainty must be less than the GEMS radiometric accuracy (4%); the resolution of the solar reference spectrum must be higher than the GEMS resolution (0.6 nm); and, finally, fully Nyquist sampled spectra are preferred.

Compared to the numerous low-resolution spectra, few suitable higher-resolution spectra are available due to the limited observations. Here, three candidates are considered for the GEMS solar reference. One is the Royal Netherlands Meteorological Institute (KNMI) spectrum, developed by Dobber et al. (2008) [34], which is utilized as the solar reference spectrum for the calibrating OMI instrument. Another is the SAO2010 spectrum provided by Chance and Kurucz [35], an updated version of the SAO96 [32] that had been widely used in heritage satellite mission programs. The final potential spectrum has recently been determined from ground-based measurements by the Quality Assurance of Spectral Ultraviolet Measurements in Europe spectroradiometer and a Fourier transform spectroradiometer (hereafter referred to as the QAS spectrum) [36].

To characterize the differences associated with these spectra, they are compared at the spectral resolution and range of GEMS, as shown in Figure 5. Each spectrum is known to have similar radiometric accuracy (uncertainty of less than 4% for the GEMS observational range), and the average difference among them is less than 1.4%. However it should be noted that there are significant differences of up to 8% at shorter wavelengths (see Figure 5). As the hyperspectral solar reference irradiances considered here are composite spectra from ground-based and/or balloon-based measurements, they might contain atmospheric absorption features [14]. The large radiative differences below 310 nm, for example, are likely related to ozone absorption. In addition, as a result of these radiometric uncertainties, SPECAL is significantly sensitive to the choice of high-resolution solar reference spectrum, as shown in Table 7.
**Figure 5.** Comparison of high-resolution solar reference spectra at the resolution and spectral range of GEMS. (a) Black line represents the relative percentage difference between the QAS and KNMI spectra, and orange line shows that between the SAO2010 and KNMI spectra. There is a remarkable difference between SAO2010 and KNMI near 300 nm (up to 8%). (b) Black line represents the relative percentage difference between the QAS and SAO2010 spectra, and orange line shows that between SAO2010 and KNMI solar irradiances.

The calibration accuracy using the QAS spectrum is lower than that achieved with the SAO2010 and KNMI spectra. The minimum chi-squared, mean bias, and RMSD values (2.45, 6.59, and 3.81, respectively) are larger than those for the other spectra. The shift values ($C_{\lambda_0}$ and $C_{\lambda_1}$) retrieved in the QAS test also indicate that the absolute wavelength calibration of the QAS itself is insufficient compared with other sources of solar irradiances. The expected value is 0.01 nm for $C_{\lambda_0}$ but the output of 0.009 nm indicates a residual of 0.001. Second, using different high-resolution solar reference spectra for the derivation of the synthetic data and simulated spectrum (reference spectrum) significantly degrades SPECAL’s performance. The retrieved shift parameter shows an additional wavelength shift of up to 0.017 nm with the cost function increased by over an order of magnitude when using KNMI and QAS. This unacceptable calibration result arises owing to radiometric uncertainties between the solar reference spectra. However, the fact that the SPECAL error is greater when using KNMI and QAS rather using SAO and KNMI indicates that the calibration result is sensitive not only to the radiative accuracy, but also to the spectral accuracy of the solar reference data. Therefore, to further improve SPECAL, proper choice of the solar reference spectrum is of critical importance. The optimal solar reference spectrum for GEMS is to be determined using on-orbit solar measurements during the instrument’s in-orbit tests.

<table>
<thead>
<tr>
<th>Simulated Input</th>
<th>Algorithm Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference Spectrum</td>
<td>Synthetic Spectrum</td>
</tr>
<tr>
<td>SAO2010</td>
<td>SAO2010</td>
</tr>
<tr>
<td>KNMI</td>
<td>KNMI</td>
</tr>
<tr>
<td>QAS</td>
<td>QAS</td>
</tr>
<tr>
<td>SAO2010</td>
<td>KNMI</td>
</tr>
<tr>
<td>KNMI</td>
<td>SAO2010</td>
</tr>
<tr>
<td>QAS</td>
<td>SAO2010</td>
</tr>
<tr>
<td>SAO2010</td>
<td>QAS</td>
</tr>
<tr>
<td>QAS</td>
<td>KNMI</td>
</tr>
<tr>
<td>KNMI</td>
<td>QAS</td>
</tr>
</tbody>
</table>

### 4. Conclusions
A spectral calibration algorithm for GEMS is presented here. To achieve sufficient accuracy for level 2 products, the calibration is performed using spectral fitting. The algorithm is designed to minimize the differences between the simulated reference spectrum (derived by convolving a high-resolution solar reference spectrum with the instrument SRF) and the spectrum measured by GEMS. To estimate the performance of the algorithm, the effects of several spectral fitting parameters on the calibration are examined using cost function and fitting residual. Within the expected wavelength variation (up to a few nm per pixel), the algorithm shows acceptable performance (accuracy better than 0.002 nm) to the expected uncertainties associated with most of spectral parameters, such as wavelength perturbation and signal-to-noise ratio. However, the sources of the high-resolution solar reference spectrum and inaccurate knowledge of in-flight SRF are shown to be critical to spectral calibration. The sensitivity to the SRF is expected to be mitigated by derivation of in-flight SRFs using analytical functions or modification of prelaunch SRFs. On the other hand, the radiometric disagreement between the high-resolution solar reference spectra is non-negligible in the GEMS observational range, especially below 310 nm; therefore, selecting the optimal solar reference spectrum is vital for better calibration performance. Furthermore, the current candidate reference spectra for GEMS might include some atmospheric absorption features, so further improvements are possible for spectral calibration after updating the current references or using a pure high-resolution solar spectrum measured from space.

As GEMS measures solar irradiance once per day during the night, the algorithm can use a sufficient number of fitting parameters, while the number of parameters for radiance measurements is limited due to the numerical cost. Hence, the operational spectral calibration algorithm for solar irradiance can include accurate SRF fitting parameters together with shift polynomials or shifts based on carefully selected narrow fitting windows. GEMS is known to have a very stable optical bench temperature and thus the wavelength change during the day is expected to be very small. Therefore, wavelengths of level 1b radiances are determined by modifying the wavelengths obtained from solar irradiances. Although spectral calibration for solar observations includes the parameters of SRF used to convolve the high-resolution solar reference spectrum, it incurs huge computational cost. Therefore, the application of a pre-derived reference spectrum convolved with a known nominal SRF is possible when the amount of change in the SRF during the mission is insignificant. The effect of applying wavelength shift can be linearized as a shift spectrum, and then included as a pseudo absorber spectrum; such linearization can also further expedite the fitting process. These approaches are therefore being investigated during in-orbit tests.

5. Patent

There is a Korea patent registration (10-2100545) resulting from this work.
Author Contributions: M.K. led manuscript writing, contributed to research design, and analyzed the results. M.-H.A. supervised this study, contributed to conceiving the idea and writing, and serves as the corresponding author. X.L. contributed to the research design and manuscript writing. U.J. contributed to manuscript writing. J.K. contributed to the discussion of the results.

Funding: This work is supported by the Korea Ministry of Environment (MOE) as “Public Technology Program based on Environmental Policy (2017000160002)”

Conflicts of Interest: The authors declare no conflict of interest.

References


