

# Recommendation for UV soft calibration from a study of the Slit function retrieval of Sciamachy

Sweta Shah, KNMI, sweta.shah@knmi.nl

December 12, 2016

## 1 Introduction

Parameters useful to modify the slit function of the instrument are: Offset, Gain, Shift/Displacement<sup>1</sup> and Squeeze as a function of wavelength for the range 260-330 nm. Our **goal** was to find the combination of the above parameters that give us the smallest residuals. Residual is defined as the difference between a convolved solar reference model subtracted from a Sciamachy solar measurement. In all our discussions and figures we show the *relative* residuals which is in percentage since this quantity is more meaningful in order to compare results between Channels 1 and 2 and etc. We have used two reference models: Kitt Peak (KP) and Dobber (DB). We found that the residuals obtained using the Dobber solar reference spectra are smaller compared to the Kitt Peak, thus we recommend to use DB reference for slit function retrievals in the UV.

## 2 Method

I will outline the method of slit function retrieval for one solar spectrum. This will be then applied to the rest of the solar spectra of each day for the length of the entire mission. We start with a SCIAMACHY solar spectra (Measurement), and a solar reference spectra (Reference). We want to find the optimum characterization of the SCIAMACHY instrument's slit function parameters for which a best match between Measurement and Reference spectra can be retrieved. The instrument slit functions for the case of SCIAMACHY are characterized by the following functional forms: Gaussian, Simple Hyperbolic and Voigt depending on the wavelength. The wavelengths of the instrument ranging from UV to infrared are further divided into Channels 1 through 8. Since we are interested in ozone retrieval and therefore the UV wavelength range, we concern ourselves with Channels 1 and 2. The end of Channel 1 and beginning of Channel 2 overlap giving us choice to choose the observations from the two channels for which the data is more reliable. The slit functions ( $S$ ) of Channels 1 and 2 are characterised by the Simple Hyperbolic functional form expressed as (taken from ATBD document):

$$S(\lambda) = \frac{a_1^2}{a_2^2 + (\lambda - a_3)^2} + a_4. \quad (1)$$

In the equation above,  $a_i$  are coefficients representing the parameters given in the level 1 data file: Normalisation factor ( $a_1 = 1$ ), full width half maximum (FWHM) (related to  $a_2$ ), central point ( $a_3$ ) and offset ( $a_4 = 0$ ). The coefficient  $a_2$  can be solved in terms of FWHM by using the fact that at the central point ( $\lambda_0$ ):  $S(\lambda_0) = 1/a_2^2$ . At each wavelength, this functional shape can be manipulated with the SF parameters (Gain, Offset, Shift and Squeeze) to modify the resulting (reference) intensity to best match the data.

---

<sup>1</sup>these terms are used interchangeably in the document; they mean the same

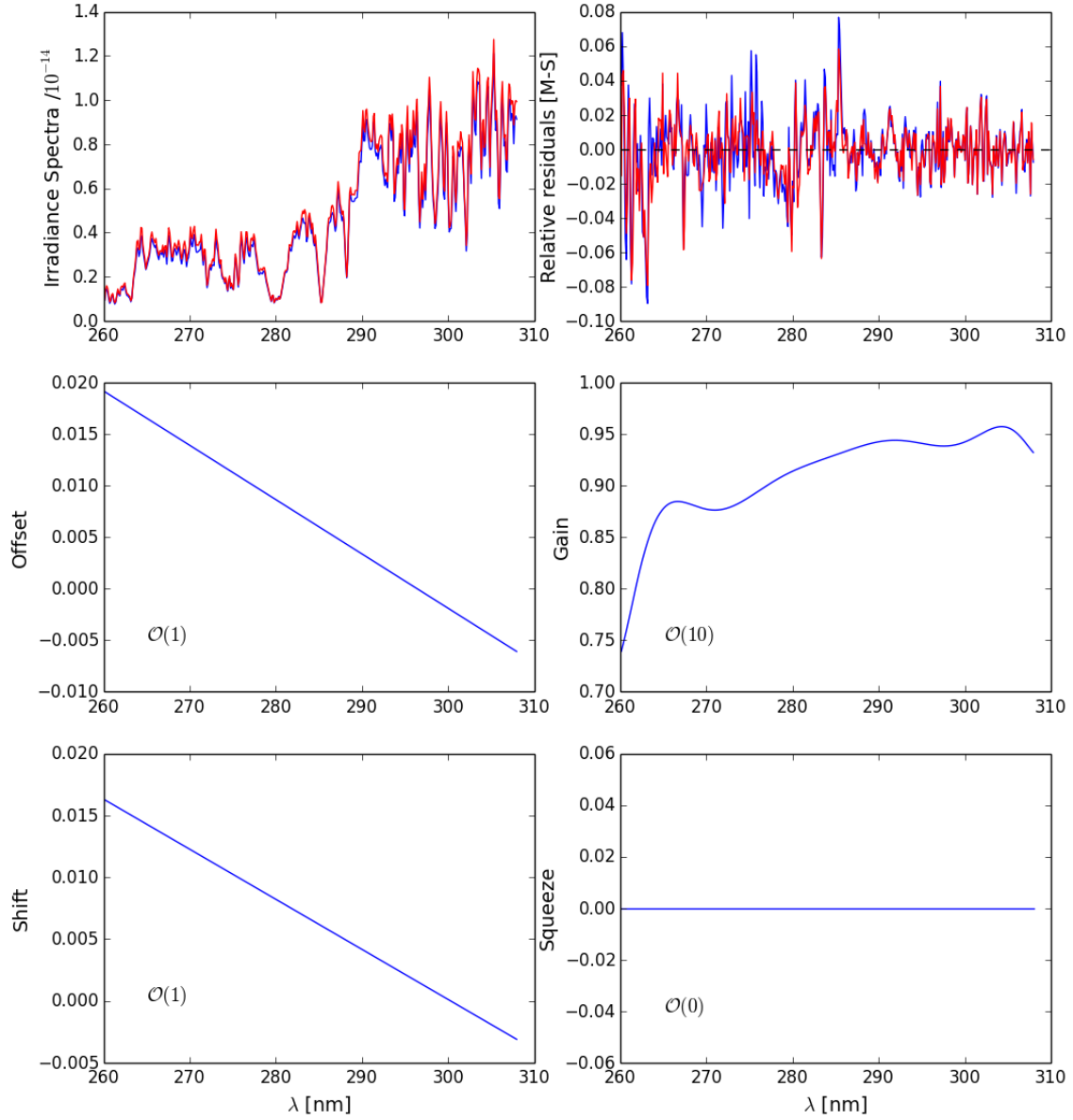


Figure 1: Relative Residual [in %] of SCIAMACHY Channel 1 of a single solar spectrum corresponding to a single day. The reference spectra used for convolution is from Dobber. The SF parameters: offset gain, shift and squeeze applied to the reference spectra are shown in the bottom panels. In each panel the polynomial order ( $\sum_i a_i \times x^i$ ) for each SF parameter are specified, such that  $\mathcal{O}(n)$  means a polynomial of power  $n$ .

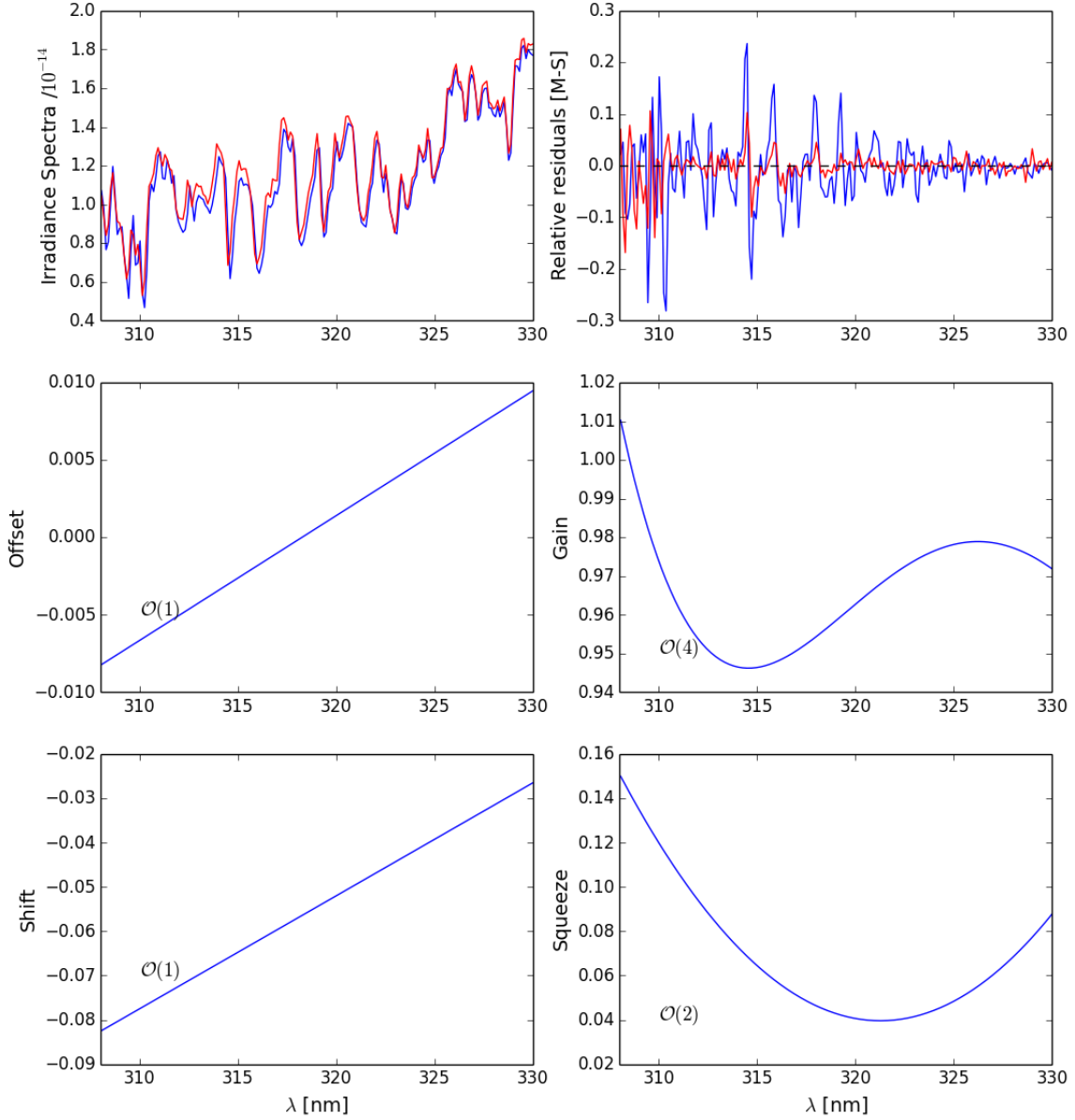


Figure 2: Same as in Figure 1 but for Channel 2.

The parameters FWHM and  $\lambda_0$  at each wavelength can be read off from the level 1 data which are identical for all solar spectra. These values are given for only certain wavelengths in the Channels 1 and 2 which can be interpolated for the wavelengths inbetween. The level 1 data provides the parameter values that were measured before the satellite's launch. We want to find parameter values after the launch (using the solar measurements) which more accurately characterise the instrument slit functions. As mentioned above, the parameter values for each slit give us the functional form which are a function of wavelength. This is used in the retrieval algorithm fed through the slit function model that has four degrees of freedom that each spectra peak can be transformed with: additive constant (Offset), multiplication factor (Gain), displacement of the peak along wavelength (Shift) and expansion and contraction of the spectral peak (Squeeze). Each of these degrees of freedom can be modelled as a polynomial of  $n$  order for each slit function of the instrument at the

desired channels. The optimal estimation is run using this model from which the best values of the Offset, Gain, Shift, Squeeze as a function of wavelength of Channels 1 and 2 are retrieved. Example for one solar spectrum is shown in Figures 1 and 2.

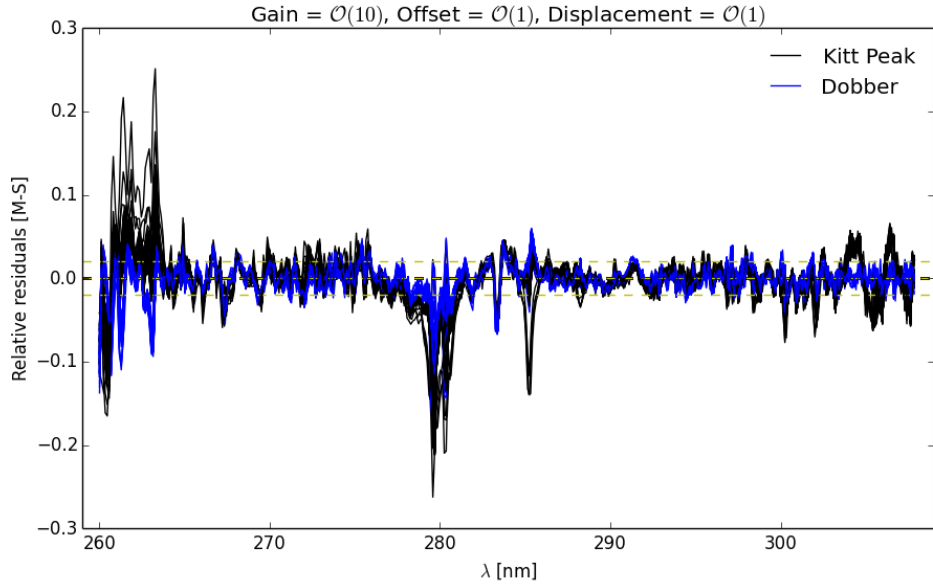


Figure 3: Comparison of the relative Residuals [in %] of SCIAMACHY Channel 2 solar spectra for each 21st day of every month of the entire mission computed using Dobber’s reference spectrum versus Kitt Peak reference spectrum. Note that the SF parameter Displacement is Shift.

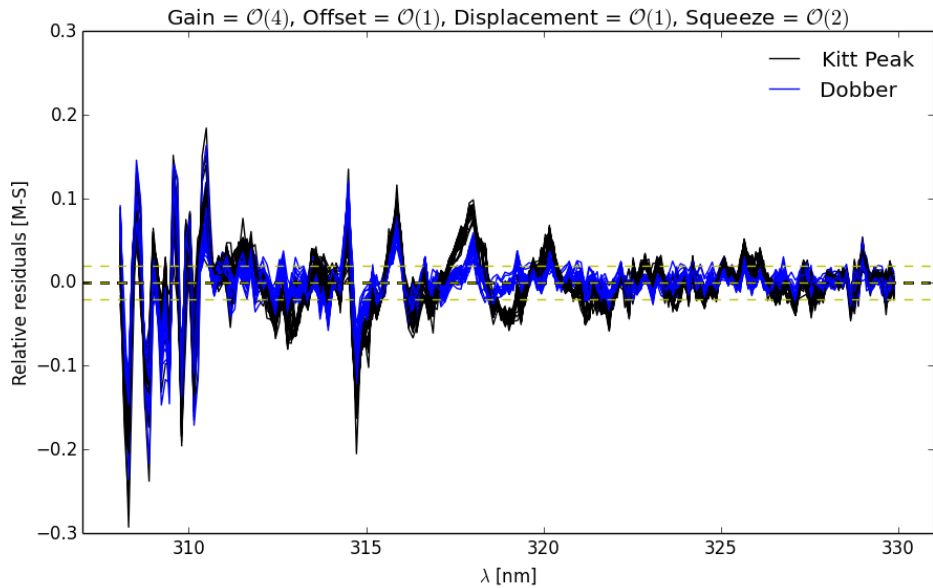


Figure 4: Same as in Figure 3 but for Channel 2.

### 3 Results

We find that in Channel 1, the wavelength range 265-308 nm is the optimal range for obtaining smallest residuals, however because the ozone retrieval algorithm requires data from 260 nm, we do the optimal estimation from 260-308 nm. The higher wavelength values around 314 nm are known to have had calibration problems and therefore we use until 308nm where the slit function retrievals are well behaved. The irradiance at 308-314 nm in Channel 1 is very hard to match with the reference. We ran the optimal estimation on all solar measurements of Sciamachy data for each day which amounted to 3463 spectra. We convolved all spectra with Offset, Gain, Shift, Squeeze of polynomial orders of = 4, 15, 1, None and found that the cost function for majority reaches less than 1 (as expected) within 10 loops of retrieval in the Optimal Estimation routine.

The residuals in Channel 2 are larger throughout. We performed exercises where we divided the relevant range of 308-330 nm into smaller ranges to get smaller residuals. We found that this was not effective and found the polynomial orders of the set of Offset, Gain, Shift, Squeeze to be 1, 4, 1, 2. We also find that using Shift in range 265-308 nm reduces the residuals significantly. Look at Figure 1. Similarly, in Channel 2, we find that using Shift, Squeeze in the range 308-330 nm reduces the relative residuals significantly. Please look at Figure 1. Similarly, in Channel 2, we find that using Shift, Squeeze in the range 308-330 nm reduces the relative residuals significantly.

The relative errors of the observation in the wavelength range we will use for ozone retrieval 265-330 nm is in the order of  $10^{-5}$ . The relative residuals are in the order of a few percent. So we expect an error of a few percent to propagate into the ozone retrieval despite more accurate solar irradiances.

There are anomalies in the residuals at around 280nm and 301nm. These are due to the strong Magnesium II lines from the solar spectra and will not be used for ozone retrieval.

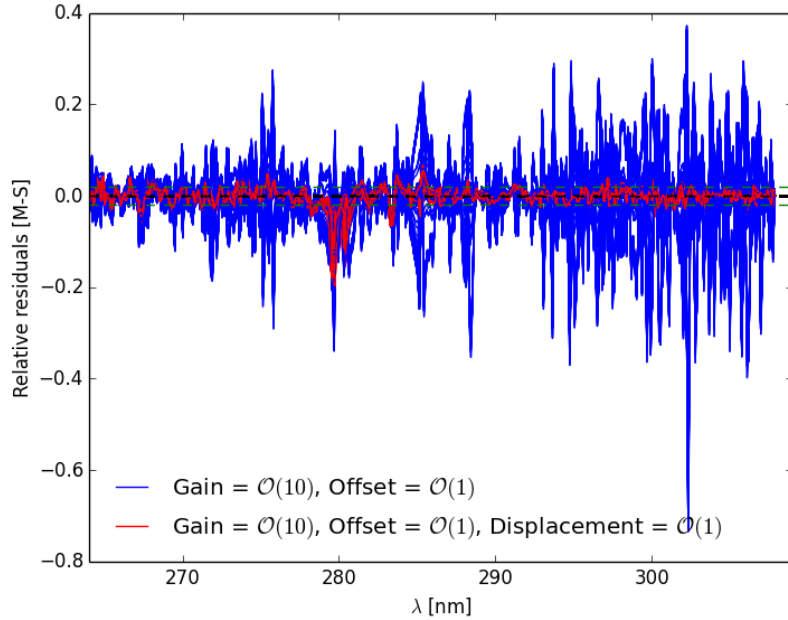


Figure 5: Relative Residuals [in %] of SCIAMACHY Channel 1 solar spectra for each day of entire mission. The reference spectra used for convolution is from Dobber.

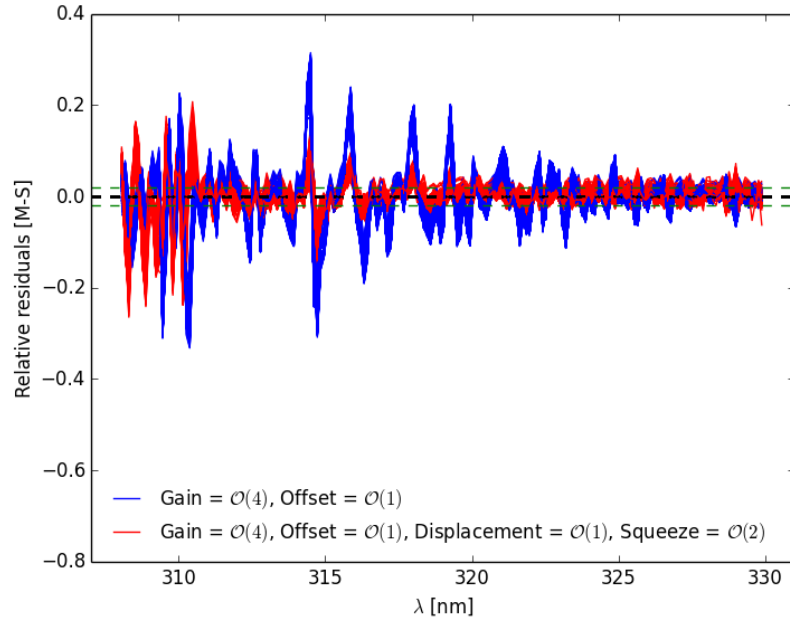


Figure 6: Relative Residuals [in %] of SCIAMACHY Channel 2 solar spectra for each day of entire mission. The reference spectra used for convolution is from Dobber.

Below are figures of the time dependence (day-to-day) of the fitted Gain, Offset, Shift and Squeeze for each solar spectrum as a function of wavelength. They are shown for both Channels 1 and 2.

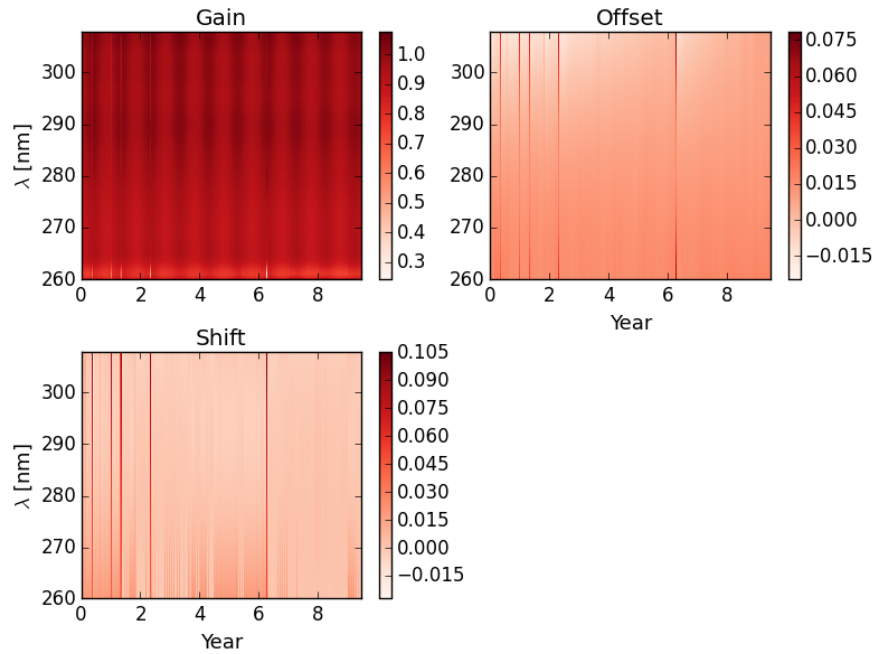


Figure 7: Time (yearly) dependence of Gain, Offset and Shift as a function of wavelength for Channel 1.

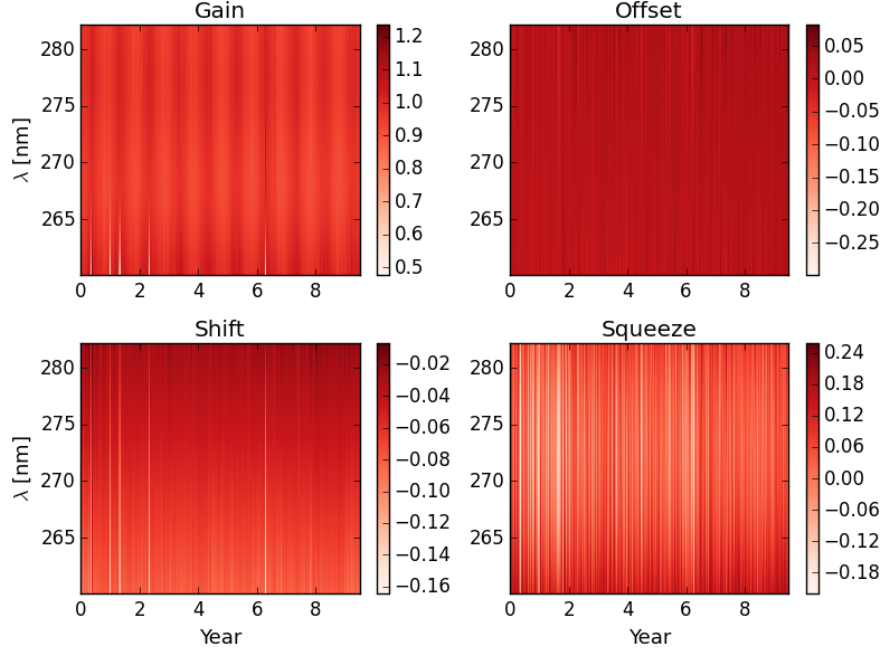


Figure 8: Time dependence of Gain, Offset, Shift and Squeeze as a function of wavelength for Channel 2.

## 4 Preparing Level 1 data for Ozone profile retrieval

The quantities passed for Forward model and the optimal estimation for the retrieval of the state-vectors (for e.g. O3 profiles, temperature profiles, etc.) are reflectance spectra defined as:

$$\mathcal{R} = \frac{R}{I}. \quad (2)$$

where,  $R, I$  are Earth radiance and Solar Irradiance respectively. We have two reflectances where one is based on measured quantity and the other on simulated quantity described as:

$$\mathcal{R}\Big|_M = \frac{R|_M}{I|_M}; \quad \mathcal{R}\Big|_S = \frac{R|_S}{I|_S} \quad (3)$$

where,  $M, S$  are measured and simulated quantities respectively.  $I|_S$  is simulated solar spectrum convolved with the spectral line slit function (shift and squeeze) for which we use Dobber's high resolution spectrum. Thus, the simulated solar wavelengths,  $\lambda_S$  are transformed according to:

$$\lambda'_S = \lambda_S * \hat{s}(\lambda) + \hat{d}(\lambda). \quad (4)$$

where the  $\hat{s}, \hat{d}$  are squeeze and displacement values at each solar wavelength. In Eq. 3,  $R|_S$  is simulated earth radiance which is derived using the forward model  $\mathcal{F}(\mathbf{x})$  give by:

$$R|_S = I|'_S \times \mathcal{F}(\mathbf{x}) \quad (5)$$

We use slit function Gain and Offset purely for comparison between the level of simulated spectrum (Dobber) and the measured solar spectrum. The correction that needs to be applied for ozone retrieval before the level 1 data is passed on to the retrieval algorithm should have spectral line slit function correction using only shift and squeeze in the manner shown in Eq. 4.

Note: The difference in the wavelength grids of  $R|_M$  and  $I|_M$  due to the Doppler shift of the Sun must be taken into account. Thus the measured earth and solar spectra are on same wavelength values. The simulated solar spectra is first convolved with the spectral slit function corrections and then interpolated to the (lower-resolution) measurement wavelength grid.

## 5 Slit function files

The shift and squeeze values retrieved using the KNMI optimal estimation code given in hdf5 files and the polynomial order and coefficient for each parameter (Gain, Offset, Shift, Squeeze) are listed in the two SF retrieval output files. For example, in the hdf5 file, each column inside the data field ‘Retrieval/SqueezeOrder’ lists the polynomial order fit for each solar spectrum and for the entire Sciamachy mission, there are 3465 solar spectra: one for each day. The corresponding wavelength values of squeeze then are listed in the data field ‘Retrieval/Squeeze.on\_IndepVar’. Thus, this lists the values of  $\hat{s}$  to be applied at each ( $\lambda_S$  as shown in Eq. 4). The values of  $\hat{s}, \hat{d}$  are produced for Channels 1 and 2 separately consistent with the Figures 7 and 8 above.

References: Dobber et al. 2008, Solar Phys, 24