AO11: Measuring Atmospheric Composition from CubeMAP

Annabel Chantry
Supervisor: Dr Anu Dudhia

Abstract

This report presents an error analysis for retrievals of vertical profiles of methane concentration from CubeMAP. CubeMAP is a European Space Agency (ESA) mission planned for launch in 2024, which will measure transmission spectra of solar radiation through the Earth's atmosphere with high resolution. The transmission spectra will be used to retrieve vertical profiles of atmospheric composition. Since the instrument has not yet been built, the purpose of this error analysis is to quantify the impact of different error sources, so that the instrument design can focus on minimizing the largest errors.

This analysis used the Reference Forward Model (RFM) to produce expected spectra from CubeMAP. Then, error spectra that represent the impact of different errors on the measured transmission spectra were produced and optimal estimation was applied to propagate errors in transmission spectra into vertical profiles of error in methane concentration. In order to compare the size of each error source across all the different heights, the information content was found, which is a measure of how uncertainty in a value is reduced by a measurement. Since the error profile is essentially a vector, information content is introduced as a scalar quantity to allow different error sources to be compared. Finally, information loss caused by each error source was found.

Twenty different error sources were considered. These were split into instrument errors caused by problems with the instrument, and atmospheric errors associated with the uncertainty in the model atmosphere used. All were found to have an information loss below the threshold, which was chosen to be the information loss from doubling the noise. This report also contains a comparison of the size of the information loss from different error sources and the factors that determine this.

1 Introduction

ESA plan to launch several Scout missions as part of ESA's Earth Observation FutureEO Programme [1]. These will consist of small satellites that can be developed and launched rapidly to prototype and demonstrate novel observation techniques. These missions will provide supplementary observations to existing satellite data and could be scaled up to larger missions if successful.

CubeMAP is one of these Scout missions which aims to produce regular vertical profiles of the concentration of gases in the stratosphere and upper troposphere to a higher precision than existing measurements. CubeMAP will be a constellation of three satellites which will each carry a High Resolution InfraRed Occultation Spectrometer (HIROS) and a and a near-infrared Hyperspectral Solar Disc Imager (HSDI) [2]. This report will focus on the HIROS instrument which will measure transmission spectra of the atmosphere with high resolution in a series of narrow spectral windows. A transmission spectrum is a measure of the intensity of radiation coming from a source over a spectral window. The spectral unit used in this report is the spectroscopic wavenumber, which is the number of wavelengths per unit distance measured in cm^{-1}. Every molecule in the atmosphere has characteristic absorption features at particular wavenumbers where radiation is highly absorbed which can be seen in a transmission spectrum and used to retreive the composition. Measurements of the composition of the atmosphere are important for monitoring the impact of emissions from human activities, understanding atmospheric processes, and climate modeling.

This report will describe the error analysis for vertical profiles of methane concentration from CubeMAP. The aim is to estimate the size of the uncertainty in methane concentration from 20 error

sources and find the information loss to the overall vertical profile. Since the HIROS instrument has not yet been built, estimates of the impact of the different instrument errors will help to specify instrument design. The estimates will also help quantify the limitations of methane measurements from CubeMAP and the impact of making certain assumptions when retrieving methane from this data.

Past vertical profiles of methane have been retrieved by the Atmospheric Chemistry Experiment (ACE), which like CubeMAP uses sun occultation, and is currently regarded as the most accurate measurement of stratospheric methane [3]. A validation of the ACE methane profiles found them to be accurate to within 10% in the upper troposphere and lower stratosphere; CubeMAP aims to improve upon this [4]. Other instruments such as the MIPAS instrument on the Envisat mission have measured stratospheric methane using limb emission but to a lower precision [5]. Limb emission involves an instrument pointing horizontally and measuring emission from the molecules themselves.

Methane has been chosen as an example because its concentration is relatively stable compared to many other trace species, and therefore imposes a stricter definition on what is considered a useful measurement.

1.1 Viewing Geometry

CubeMAP will orbit around the Earth's equator approximately once every 100 minutes. It will observe a sunrise and sunset every orbit separated by approximately 180 degrees in longitude. The HIROS instrument will point at the sun and measure the transmission of solar radiation through the atmosphere. By repeatedly measuring this during a sunrise or sunset, spectra can be obtained at different tangent heights through the atmosphere as shown in Figure 1.

The tangent height refers to the height of the line of sight of the instrument above the earth when the measurement was taken. A retrieval height refers to the height of a measurement of methane concentration in the vertical profile. Since the path along a particular tangent passes through multiple layers of the atmosphere, the transmission spectrum at one tangent height depends on the methane concentration at all the retrieval heights along the tangent. This error analysis uses 60 retrieval heights spaced 1 km apart between 10 km and 60 km, and 34 tangent heights spaced 1.5 km apart.

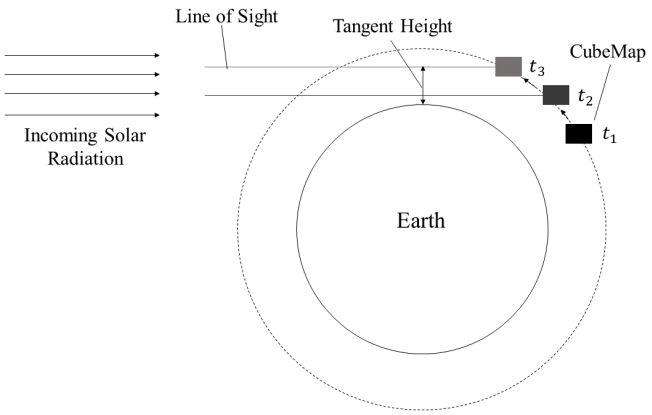

Figure 1: A diagram of the sun occultation viewing geometry used by CubeMAP

Many current missions, such as the Greenhouse gases Observing SATellite (GOSAT) and the Tropomi instrument on the Sentinel-5P satellite, use nadir viewing to measure methane concentrations[6],[7]. This involves a satellite pointing down at the Earth and measuring emission from the atmosphere. Another viewing geometry is limb viewing, where an instrument points sideways and measures emission from the atmosphere. Nadir viewing is useful for measuring the regional distribution of methane concentration, identifying point sources, and measuring composition of the troposphere between clouds, but it has low vertical resolution compared to limb viewing. Limb viewing has the advantage of higher vertical resolution and a longer optical path length through the atmosphere, making it easier to detect trace gases. However, cloud cover often prevents limb viewing in the troposphere, and the long path length results in low horizontal resolution. Solar occultation is a type of limb viewing with the advantage of a high signal to noise ratio due to the brightness of the sun. However, sun occultation produces vertical profiles twice per orbit, whilst limb emission can operate continually during the orbit. This is why CubeMAP will be used primarily for vertical profiles to a high accuracy in the stratosphere, where there is less horizontal variation in gas concentrations and lower cloud cover [8].

1.2 The Instrument

The HIROS instrument is a laser heterodyne spectroradiometer, the first of its kind used in an ESA satellite [2]. The instrument mixes incoming radiation with a laser at a known frequency. The laser scans through the narrow spectral window during a mea-

surement and the resulting beat frequency created by the interference of the two signals is a function of the intensity of the incoming radiation at the emission frequency of the laser [9]. This allows the instrument to produce high resolution spectra in a narrow spectral window of interest. HIROS will use spectral windows that are only 1cm^{-1} wide. Since the spectral window is so narrow it must be targeted towards the absorption features of the molecule or molecules being measured. The spectral window used for methane measurements is shown in Figure 2. The two strong methane lines make this window suitable for measuring methane, but the presence of other spectral lines is a potential source of error. The window has been chosen to include spectral lines of other molecules so that they can also be retrieved. The CO_2 line can also be used to establish the pressure of the atmosphere at the tangent point since CO_2 concentrations are well known.

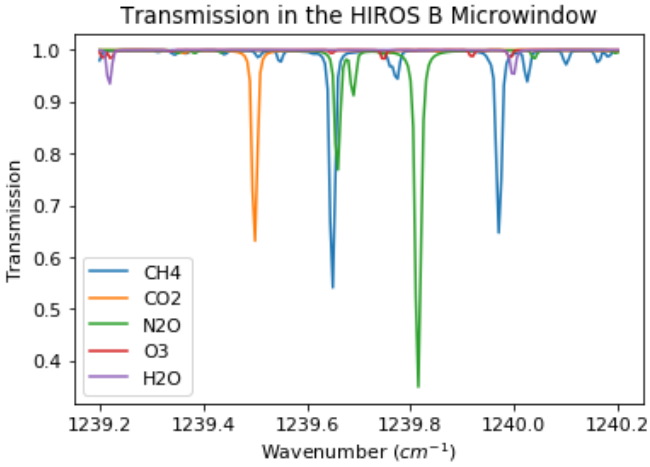

Figure 2: The HIROS B microwindow

2 Method

2.1 Optimal Estimation

This section will describe how methane concentrations are retrieved from transmission spectra using optimal estimation theory, and therefore how an error in the transmission spectra is propagated into an error in methane concentration.

The RFM describes how methane concentration at one retrieval height will change the expected transmission spectra at a particular tangent height. For the error analysis a linearized form of the forward model $\mathbf{K}$ is used:

$$\mathbf{y} = \mathbf{K}\mathbf{x}, \tag{1}$$

where $\mathbf{x}$ is a column vector of methane concentrations at all the chosen retrieval heights, $\mathbf{y}$ is a column vector of all the transmission spectra appended end to end, and $\mathbf{K}$ is a matrix that maps $\mathbf{x}$ onto $\mathbf{y}$.

Retrieving methane concentration from the transmission spectra requires finding the gain matrix $\mathbf{G}$ that describes the opposite relationship to $\mathbf{K}$. $\mathbf{G}$ maps transmission spectra onto the methane concentrations, the retrieval is then described by:

$$\mathbf{x} = \mathbf{G}\mathbf{y}. \tag{2}$$

A method needs to be chosen for constructing $\mathbf{G}$. A common approach to this is a least squares solution, where the difference between the measurments and the values corresponding to the solution are minimised [10]:

$$\chi^2 = (\mathbf{y} - \mathbf{K}\mathbf{x})^2. \tag{3}$$

This is the simplest form of χ^2 but it would be more suitable to use a least squares fit weighted by the variance, so that spectral points with a lower variance are weighted higher [10]:

$$\chi^2 = (\mathbf{y} - \mathbf{K}\mathbf{x})^T {\mathbf{S_y}}^{-1} (\mathbf{y} - \mathbf{K}\mathbf{x}), \tag{4}$$

where $\mathbf{S_y}$ is the covariance matrix for the transmission spectra.

The covariance matrix is a square matrix that contains the covariance between any two elements in a vector, where covariance is a measure of the joint variability of two random variables. In this case $\mathbf{S_y}$ describes the covariance between any two points in the transmission spectra, and the main diagonal contains the variance at each wavenumber.

Finally, it can be beneficial to include a priori information to constrain the fit. The a priori information $\mathbf{x_a}$ is an initial estimate for the methane concentration from a past measurement or climatology model. With optimal estimation, the retreival will weight the measured value of methane concentration with the a priori information by the size of the variance, and return a weighted combination. This is useful because the least squares fit by itself becomes unsuitable when the signal to noise ratio is small, such as high up in the stratosphere where methane concentration is low. The modified weighted χ^2 is given by [10]:

$$\chi^2 = (\mathbf{y}-\mathbf{K}\mathbf{x})^T {\mathbf{S_y}}^{-1} (\mathbf{y}-\mathbf{K}\mathbf{x}) + (\mathbf{x_a}-\mathbf{x})^T {\mathbf{S_a}}^{-1} (\mathbf{x_a}-\mathbf{x}), \tag{5}$$

where $\mathbf{S_a}$ is the a priori covariance matrix. $\mathbf{S_a}$ describes the covariance in the a priori information of

methane concentration at different retrieval heights. For this project it was assumed that methane concentration was initially known to 10%, reflecting the accuracy of the existing ACE measurements.

χ^2 is then minimized to find the form of G [10]:

$$\mathbf{G} = (\mathbf{K^T S_y K} + \mathbf{S_a^{-1}})^{-1}\mathbf{S_y^{-1}K^T}. \quad (6)$$

It was then assumed that covariance between points in the transmission spectra is caused only by noise from the instrument detector. The noise is assumed to be constant in size, and uncorrelated from one spectral point to another since these are measured consecutively in time. This means the covariance matrix, $\mathbf{S_y}$, can be written as a constant times the identity matrix, and $\mathbf{G}$ can be rewritten as:

$$\mathbf{G} = (\mathbf{K^T K} + \sigma_y^2 \mathbf{S_a}^{-1})^{-1}\mathbf{K^T}, \quad (7)$$

where σ_y is the constant noise in the transmission spectra.

The $\mathbf{K}$ matrix can be constructed using Jacobian spectra produced using the Reference Forward Model (RFM). One Jacobian spectra is the change in transmission at one tangent height caused by a small perturbation in methane concentration at a specific retrieval height, $d\mathbf{y_i}/dx_j$. For each tangent height, Jacobian spectra corresponding to perturbations at all the retrieval heights are produced and combined to construct the $\mathbf{K}$ matrix, and $\mathbf{G}$ is found. The error in methane concentration is then given by:

$$\delta\mathbf{x} = \mathbf{G}\delta\mathbf{y}, \quad (8)$$

where $\delta\mathbf{y}$ is a column vector of error spectra and $\delta\mathbf{x}$ is a column vector of error in methane concentration at the different tangent heights. For this project it was also useful to retrieve aerosol concentrations alongside methane in order to reduce the error in methane introduced by the presence of aerosols. The size of the $\mathbf{y}$ vector is therefore doubled to include aerosol concentrations at all the retrieval heights, as are the number of rows of the $\mathbf{G}$ matrix and the columns of the $\mathbf{K}$ matrix.

2.2 Covariance

The error in methane concentration due to noise in the transmission spectra is given by the diagonal of the noise covariance matrix $\mathbf{S_x}$. For scalar quantities, when two independent measurements are combined using an average weighted by the square of their standard deviations, the combined variance is given by [10]:

$$\sigma^2 = (1/\sigma_1^2 + 1/\sigma_2^2)^2. \quad (9)$$

This can be generalised to multiple measurements of vectors using the covariance matrices. The covariance of the retrieval error is given by a weighted combination of the covariance from the measured methane concentration and the a priori covariance [10]:

$$\mathbf{S_x} = (\sigma_y^2 \mathbf{K^T K} + \mathbf{S_a}^{-1})^{-1} \quad (10)$$

2.3 The Reference Forward Model

The forward model used for this project was the Reference Forward Model (RFM) [11]. The RFM predicts the expected transmission spectra given a particular atmospheric profile.

First the absorption coefficient at each wavenumber k_ν is calculated by summing the contributions from the local spectral lines. Then the path of the ray is approximated by splitting the atmosphere into homogeneous segments each with an equivalent temperature and pressure as shown in Figure 3. The optical thickness of each segment is then found using the Curtis-Godson approximation, which treats each segment as homogeneous with the same absorber amount throughout, using an absorber weighted average of p and T:

$$\chi_{seg} \approx k(\bar{p}, \bar{T})u, \quad (11)$$

where u is absorber amount.

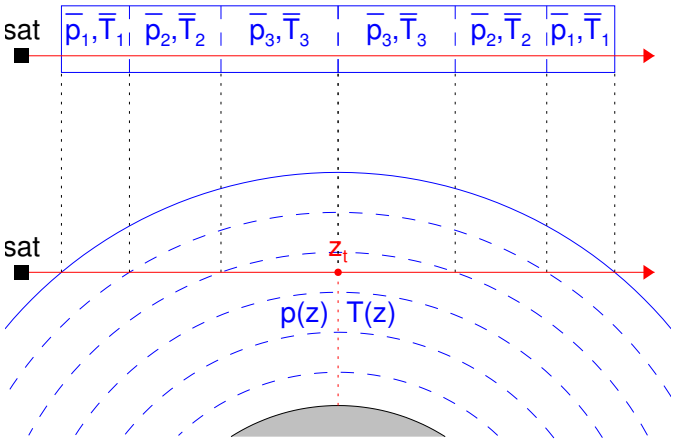

Figure 3: Construction of path segments for limb viewing geometry. The lower part of the diagram shows a ray path through a layered atmosphere. The upper part shows the equivalent set of homogeneous segments. Figure taken from [11]

The radiance at each wavenumber for solar occul-

tation is given by [11]:

$$L = B_0\tau, \tag{12}$$

where B_0 is the solar radiance and τ is the product of the transmittance of each path segment, which is equivalent to summing the optical thickness of each segment because of the relationship $\tau = \exp(-\chi)$.

To produce a transmittance spectra for a specific instrument, the radiance of the atmosphere must also be convoluted with an instrument lineshape function and a field of view function given by the user. The measured radiance given a lineshape function, ψ, and a field of view function, ϕ, is given by [11]:

$$R(\nu) = \int L(\nu')\psi(\nu' - \nu)d\nu', \tag{13}$$

$$R(z) = \int L(z')\phi(z' - z)dz'. \tag{14}$$

For this project the RFM has been used to produce expected transmission spectra and Jacobian spectra for HIROS given a set of inputs. Many of the error sources were modeled by changing the inputs in the RFM and producing new expected transmission spectra. The error spectrum is then the difference between the new spectrum and the nominal spectrum. The inputs changed were:

- The atmospheric profile: which includes the temperature, pressure and concentrations of atmospheric molecules
- The lineshape function: which describes the line shape of the instrument
- The field of view function: which describes the height and width of the line of sight of the instrument
- The radius of curvature of the Earth at the location of the measurement

2.4 Information Content

Multiplying these error spectra by the gain matrix produces a vertical profile of how error in concentration varies with height. In order to compare error sources it would be more useful to have a single number that quantifies how a particular error source impacts measurements of concentration, such as the information content H. H comes from information theory and is a measure of how much the uncertainty in a value has been decreased by a measurement [10]. Assuming the probability density function in methane concentration is a Gaussian distribution, H is given by [10]:

$$H = -\log_2(\mathbf{S_x}{\mathbf{S_a}}^{-1}), \tag{15}$$

where $\mathbf{S_x}$ is normally the covariance matrix for the total retreival. For this application $\mathbf{S_x}$ is being defined as the covariance matrix for the ideal instrument with only noise plus a covariance matrix for one additional error source constructed from the error profile:

$$\mathbf{S_x} = \mathbf{S_{noise}} + \delta\mathbf{x}\delta\mathbf{x}^T, \tag{16}$$

where $\mathbf{S_{noise}}$ is the noise covariance matrix described in Section 2.2 and $\delta\mathbf{x}$ is a column vector of the error profile for one error source. This allows information content to be calculated separately for each error source. The information content is measured in bits and can be interpreted as a measure of what knowledge has been gained from this measurement compared to what was already known. The information content with only noise is the maximum possible, and so it can be useful to consider the information loss for each error source given by:

$$H_{loss} = H_{error} - H_{noise}, \tag{17}$$

where a smaller information loss indicates lower error in the overall profile.

2.5 Degrees of Freedom

The number of degrees of freedom is another single number used to quantify the overall error in a profile due to a particular error source. It is a measure of the number of independent points one could retrieve methane at. It is given by [10]:

$$N = tr(1 - \mathbf{S_x}{\mathbf{S_a}}^{-1}) \tag{18}$$

3 Constructing the Error Spectra

To estimate the error caused by different sources, error spectra must be produced which are the expected transmission spectra with that source of error minus the spectra without it. An example of this is demonstrated in Figure 4. The following section will describe the physical origin of the error sources considered and how error spectra were calculated.

3.1 Offset

Measuring transmission involves a radiometric calibration procedure, where electronic signal from the

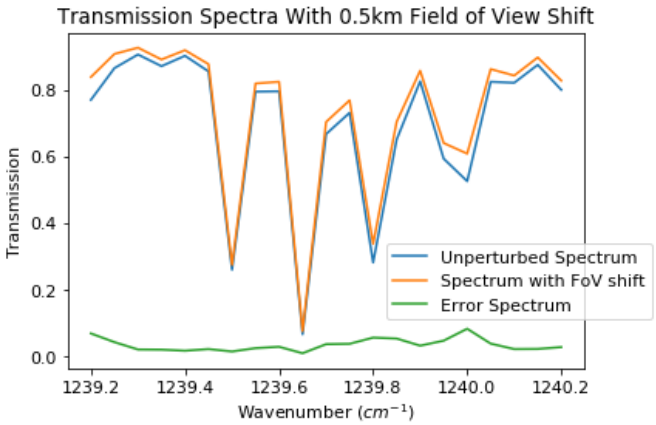

Figure 4: An example of an error spectra from field of view (FoV) shift

instrument is related to transmittance. The voltage measured by the instrument can be described as:

$$V = V_0 + g(\tau - 1), \tag{19}$$

where V_0 is the voltage when viewing the sun above the atmosphere (i.e. $\tau = 1$) and g is the radiometric gain which has to be determined for the instrument in order to find τ. Offset error is caused by the measured transmission being offset from the true value by a fixed amount for all wavenumbers. This would arise from an error in determining V_0. The resulting error spectra would be a constant line as seen in Figure 5.

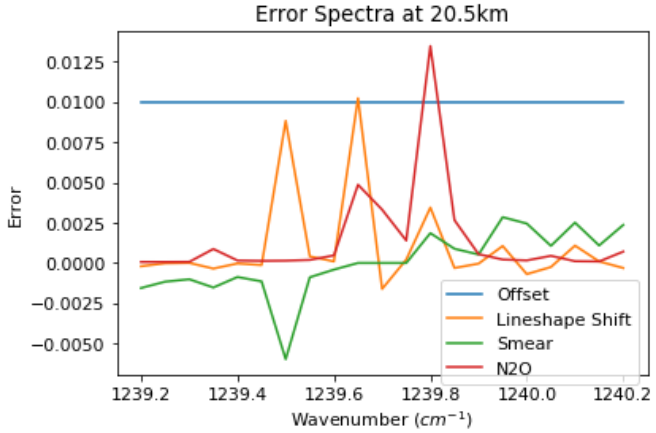

Figure 5: Some example error spectra for different error sources

3.2 Scaling

Scaling error also arises from the radiometric calibration procedure, specifically from an error in the radiometric gain g. A 1% error in g causes transmission to be scaled incorrectly according to:

$$y_s = y + (y - 1)0.01 \tag{20}$$

The scaling error spectra were calculated by scaling the expected spectra produced by the RFM and subtracting the unscaled spectra.

3.3 Lineshape Errors

Calibration errors occur when measurements at one wavenumber are incorrectly assigned to a different wavenumber. This would be caused by the instrument line shape being different than expected. The nominal lineshape function for the instrument is a box car function. Three perturbed lineshape functions were investigated: shift, spread and skew, as illustrated in Figure 6. These were selected with the goal of representing any lineshape error as a superposition of these three independent modes.

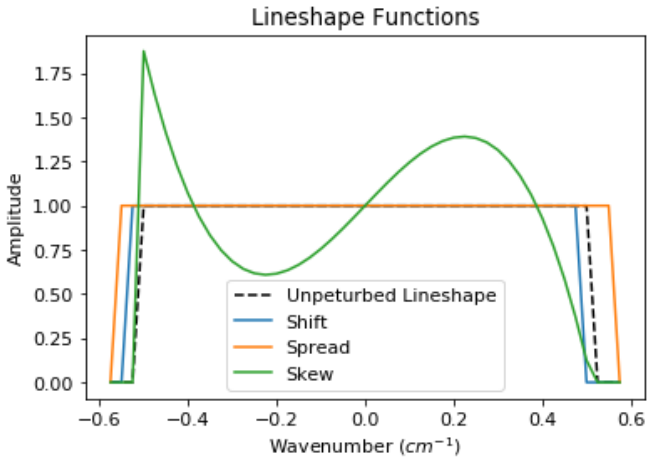

Figure 6: Graph showing the perturbed and unperturbed lineshape functions

3.4 Field of View Errors

Field of view errors are caused by uncertainty in the line of sight of the instrument, including its width and what tangent height it is looking along. The nominal field of view function is a trapezoidal function and three different perturbations were used: a shift of 0.1km, a 10% spread and a skew of 0.1.

3.5 Interfering Molecules

To predict the expected transmission spectra measured by HIROS, assumptions must be made about

the concentrations of other molecules in the atmosphere that also affect transmission. The uncertainty in these concentrations can be represented by increasing the concentrations in the profile input into the RFM and generating new expected spectra. Molecules that had lines in the spectral window that did not overlap with a methane line were still considered, because every point in the spectral window is weighted and included in the retreival and therefore could have an impact. CO_2 concentrations were increased by 1%, and H_2O, O_3 and, N_2O by 10%, reflecting the assumed knowledge of their concentrations. An example error spectrum from increasing N_2O is seen in Figure 5, where the largest difference is centered around the N_2O spectral line.

3.6 Aerosols

The Aerosol error is caused by the presence of aerosols in the atmosphere increasing the optical depth. The error spectra were produced by adding an aerosol profile to the assumed atmosphere in the RFM and generating new spectra. The effect of aerosols on transmission is spectrally smooth, and is therefore very similar to the offset error.

3.7 Temperature and Pressure

The absorption coefficient of the absorbers is dependent on the temperature and pressure of the atmosphere as described by the Curtis-Gordon approximation and equation 11. As a result, uncertainty in the temperature and pressure is a source of error in methane concentration. Similarly to the interfering molecules the error spectra for temperature and pressure were produce by increasing the temperature and pressure profiles input into the RFM by 10K and 2% respectively, which are reasonable estimates of the uncertainties with which these parameters are known.

3.8 Change in Radius of Curvature

Since the radius of curvature of the Earth is not uniform, the layers of the atmosphere that a tangent height will pass through could change depending on the location above the Earth. To find the error spectra the radius of curvature input into the RFM was changed by 10km and new transmission spectra were generated.

3.9 Horizontal Variation in Temperature

So far, all the errors have only considered vertical variation in temperature and pressure in an atmospheric column. In reality, the line of sight of the satellite spans a large horizontal distance along which there could be a change in surface temperature. It is possible for the RFM to take three vertical temperature and pressure profiles along with an angle between them, and generate transmission spectra with these as the near, middle, and far atmosphere. The original temperature profile was increased by 10K for the near profile, decreased by 10K for the far profile, and the corresponding pressure profiles were found by numerically integrating the hydrostatic balance equation, equation 21.

3.10 The Hydrostatic Assumption

The initial profile used in the RFM for the unperturbed spectra was assumed to be in hydrostatic balance meaning the relationship between pressure and temperature is given by:

$$\frac{\mathrm{d}p}{\mathrm{d}h} = -\frac{pg}{RT}, \tag{21}$$

where g is the gravitational acceleration and R is the specific gas constant for air. The hydrostatic balance assumption typically applies to a vertical column in the atmosphere. Since the satellite will move on the order of 100km horizontally during the occultation, it does not sample the atmosphere in a vertical line, meaning this relationship may not hold for the temperature and pressure profile sampled. To consider a non-hydrostatic atmosphere, a perturbation in the temperature profile was introduced which increased linearly from 0K to 10K with height and added to the unperturbed temperature profile.

3.11 Smear

Smear error is caused by the satellite continuing to rise whilst the laser scans through the spectral window, meaning that one end of a transmission spectrum is measured at a lower tangent height than the other. The effect of linear drift was simulated by interpolating between adjacent spectra. The original spectrum was then subtracted from the new spectrum to create an error spectrum such as the one shown in green on Figure 5. Since the transmission tends to increase with height, this results in an error spectrum that is negative for lower wavenumbers and positive for higher wavenumbers.

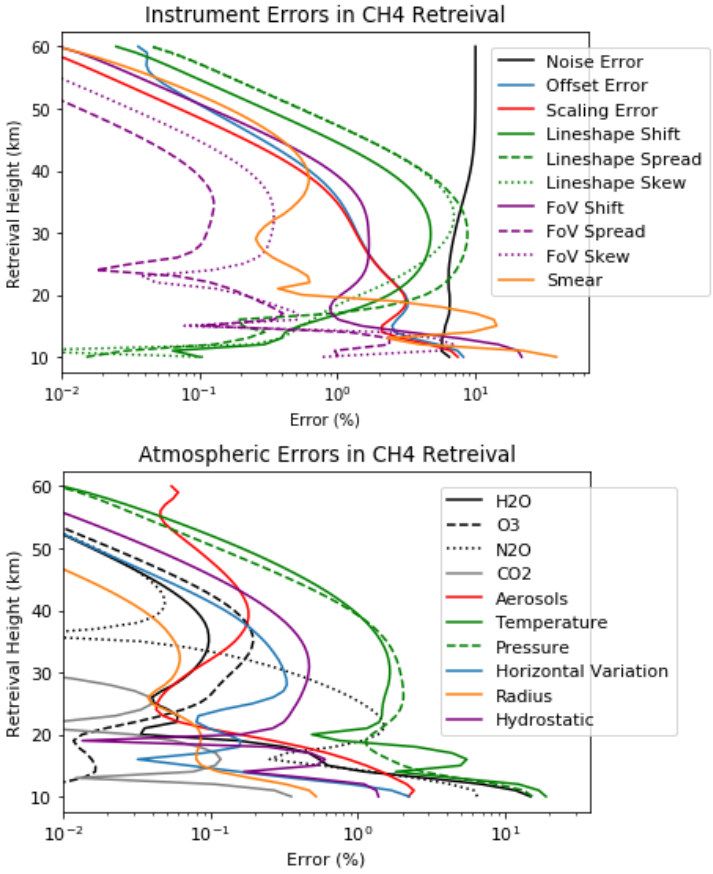

Figure 7: Error profiles for atmospheric and instrument errors

4 Results

4.1 Vertical Error Profiles

Figure 7 shows the vertical profiles of error in methane concentration found using equation 8.

Above 50km the error sources other than the noise are all close to zero, because the retreival is reverting to the a priori estimate for concentration. High up in the atmosphere there are not enough methane molecules to distinguish methane lines from the noise. As a result, the noise error becomes much larger than the a priori uncertainty in concentration. Because optimal estimation is being used the retreival then weights the a priori methane concentration higher as the height increases until it is returning only the a priori concentration. The systematic errors then become smaller because they are associated with the measured methane concentration which is being weighted much lower in the retreival than the a priori concentration. 50km is approximately where the mesosphere begins hence why CubeMAP is intended for measurements in the stratosphere and upper troposphere.

Below 15km some error sources such as smear, temperature and the field of view errors are large. In practice there may often be clouds in the way of the line of sight of the satellite below 15km preventing a measurement, but if a measurement is taken the error in methane concentration could be larger than the rest of the profile.

Overall, while the methane signal is the largest, the size of each error source varies with height making them difficult to compare, hence the need to find the information loss associated with each error.

4.2 Information Loss

Table 1 shows the information content, information loss, and degrees of freedom loss for all the errors considered. The information loss when the noise is doubled is included in row 1 as a threshold for when an error would become significant, and promisingly all the errors estimates lie below it. The error sources with the highest degrees of freedom also have the lowest information content and vice versa.

The largest information loss comes from the smear. This could be mitigated using a retrieval that considers the changing tangent height during a measurement, possibly by using Jacobian spectra where the tangent height changes over the spectral window. It is also possible that moving the spectral window so that the methane spectral lines are in the middle of the wavenumber range would reduce the size of the smear error, but this could also have a negative impact on the retrieval of other molecules using the same window.

The largest information loss from atmospheric errors was due to uncertainty in temperature and pressure. This is largely due to the size of the perturbation made to the temperature and pressure profiles which was 10K for temperature and 2% for pressure. In practice it may be that the uncertainty in temperature and pressure can be made smaller than these perturbations by measuring them simultaneously.

For the other atmospheric molecules, H_2O and N_2O had a much larger impact than O_3 and CO_2. This is likely because in the chosen spectral window both H_2O and N_2O had spectral lines that overlapped with the largest methane lines, unlike O_3 and CO_2. The size of the information loss from these molecules would be different depending on the choice of spectral window and might be reduced if methane was retrieved using all three of CubeMAP's spectral windows, increasing the number of spectral lines

Error Source	Error Size	Information Loss	DoF Loss
Double Noise	0.02	-9.69	3.90
Smear	See 3.11	-6.15	-18.91
FoV Shift	0.1km	-5.26	-5.76
Lineshape Spread	5%	-5.16	-2.92
Lineshape Skew	0.1	-3.95	-1.58
Lineshape Shift	0.01cm^{-1}	-3.51	-0.86
Offset	0.01	-3.42	-1.02
Scaling	1%	-3.19	-0.86
FoV Skew	0.1	-2.20	-0.66
FoV Spread	10%	-0.54	-0.13
Temperature	10K	-4.90	-4.73
Pressure	2%	-4.26	-2.77
H_2O	10%	-3.69	-2.37
N_2O	10%	-2.28	-0.55
Aerosol	See 3.6	-0.85	-0.09
Hydrostatic	See 3.10	-0.37	-0.04
Horizontal Variation	See 3.9	-0.36	-0.06
Radius	10km	-0.03	-0.01
O_3	10%	-0.01	-0.01
CO_2	1%	-0.01	-0.01

Table 1: Summary table of the information content, information loss and degrees of freedom for each error source, split into instrument and atmospheric errors and sorted by the information loss. With only noise the Information content was 18.52 bits and the degrees of freedom 6.59

used.

The information loss can also indicate the strength of certain assumptions, for example neglecting the variation in radius of curvature of the Earth has little impact on the information and so should be a strong assumption. Similarly neglecting horizontal variation in temperature and assuming the atmosphere is in hydrostatic balance seem to be reasonable assumptions to make.

Overall, it is important to consider that the relative size of the information loss across error sources is partially determined by the size of the instrument or atmospheric error initially assumed. Since the instrument has not yet been built this project used estimates of the size of the errors in the instrument that may be different in practice. For this reason it would be useful to find a tolerance for the instrument errors as described in the following section.

4.3 Tolerance

The tolerance is the size of error in the instrument that leads to a chosen threshold of information loss. The tolerances in Table 2 were found using the information loss from doubling the noise as a threshold.

Error Source	Estimated Size	Tolerance
Offset	0.01	0.09
Scaling	1%	10%
Lineshape Shift	0.025cm^{-1}	0.21cm^{-1}
Lineshape Spread	5%	24%
Lineshape Skew	0.1	0.73
FoV Shift	0.1km	0.46km
FoV Spread	10%	238%
FoV Skew	0.1	1.34

Table 2: Table of the tolerances for instrument errors

To find these tolerances it was assumed that there is a linear relationship between the size of the error profile and the size of the assumed error in the instrument. For the lineshape and field of view errors this relationship was found to be non-linear in practice. This may have caused the FoV and lineshape tolerance estimates to be unrealistic. Future work could include finding tolerances based on a non-linear relationship which is specific to each error source.

5 Conclusion

The aim of this projet was to complete an error analysis of methane retrievals from CubeMAP. Overall, estimates for the vertical error profile and information loss for twenty error sources have been found. These estimates will be used to specify instrument design and inform which assumptions to make about the atmosphere during the retrieval. All the information loss estimates were under the double noise threshold, although this also depends on the size of the error assumed. The largest contribution to information loss was from the smear error, followed by the uncertainty in the field of view or tangent height of the instrument. Tolerances were found for the instrument errors.

Future work could include an error analysis for the retrieval of the other molecules measured by CubeMAP which should be different due to the varying intensity and placement of spectral lines in the spectral window. It could also consider how the choice of spectral window impacts the error in the measurements and whether this could be optimized for a more accurate methane retrieval.

References

[1] European Space Agency. *Contract signed to build Scout CubeMAP*. URL: https://www.esa.int/Applications/Observing_the_Earth/FutureEO/Contract_signed_to_build_Scout_CubeMAP. (accessed: 06.11.22).

[2] European Space Agency. *ESP-MACCS / CubeMAP*. URL: https://www.eoportal.org/satellite-missions/esp-maccs. (accessed: 25.03.23).

[3] P. F. Bernathl. "The Atmospheric Chemistry Experiment (ACE)". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 186 (2017), pp. 3–16.

[4] M. De Mazière et al. "Validation of ACE-FTS v2.2 methane profiles from the upper troposphere to the lower mesosphere". In: *Atmospheric Chemistry and Physics* 8 (2008), pp. 2421–2435.

[5] European Space Agency. *MIPAS*. URL: https://earth.esa.int/eogateway/instruments/mipas. (accessed: 25.03.23).

[6] European Space Agency. *Tropomi*. URL: https://www.esa.int/Applications/Observing_the_Earth/Copernicus/Sentinel-5P/Tropomi. (accessed: 25.03.23).

[7] National Institute for Environmental Studies. *GOSAT Instruments and Observational Methods*. URL: https://www.gosat.nies.go.jp/en/about_2_observe.html. (accessed: 25.03.23).

[8] J. H. Seinfeld. *Atmospheric chemistry and physics : from air pollution to climate change*. John Wiley and Sons, Incorporated, 2016, p. 296.

[9] STFC Rutherford Appleton Laboratory. *Development of a Laser Heterodyne Radiometer (LHR)*. URL: https://ceoi.ac.uk/technologies/infra-red-radiometry/laser-heterodyne-radiometry/laser-heterodyne-radiometer/. (accessed: 01.11.22).

[10] C. D. Rodgers. *Inverse Methods For Atmospheric Sounding*. Singapore ; London : World Scientific, 2000.

[11] A. Dudhia. "The Reference Forward Model (RFM)". In: *Journal of Quantitative Spectroscopy and Radiative Transfer* 186 (2017), pp. 243–253.