Retrieval of trace gases using IASI

R. Anthony Vincent
Oriel College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2016

This thesis is dedicated to
my daughters,
Isabella and Katarina

Acknowledgements

I would sincerely like to thank Anu Dudhia for his support and timely feedback throughout the entirety of my thesis research. Anu went out of his way to welcome my daughters and I to the United Kingdom and was consistently available to provide help no matter how busy his schedule became. Always the educator, Anu never shied away from a teaching moment and persisted until the lesson was learned.

I also thank Don Grainger for his interest, comments, and enthusiasm in my work and studies. Within the Earth Observation Data Group, I thank Andy Smith, Lucy Ventress, and Elisa Carboni for much help sorting out the algorithmic details during the early part of my studies. Additionally, I much appreciate the comments to my first and second year reports provided by Simon Calcutt, Leigh Fletcher, and Neil Bowles. Sarah Harrington was very helpful during the admissions process, for which I am grateful.

Externally, I am indebted to Richard Pope and Martyn Chipperfield of the University of Leeds for providing timely OCS model fields from TOMCAT. I thank David Moore from the University of Leicester for further testing the look-up tables we generated in their retrieval scheme of carbon monoxide. Also, I thank Steve Montzka of NOAA for giving me permission to use their OCS flask data.

Additionally, I thank my examiners, Pat Irwin and Brian Kerridge, for agreeing to examine my thesis and provide their expert opinions.

Finally, I thank my parents, brothers and sisters, and my daughters for supporting and believing in me all these years. It was a long road diverging in a yellow wood.

Abstract

Trace gases comprise less than 1% of Earth's atmosphere, yet they dominate the radiation budget and drive atmospheric chemistry. The Infrared Atmospheric Sounding Interferometer (IASI) is a set of two nadir viewing Fourier transform spectrometers aboard the MetOp-A and B satellites in polar orbit that measure spectral radiances in the thermal infrared and, therefore, observe the spectroscopic signatures of numerous trace gases. This thesis focuses on the practicalities of fully utilizing the vast amounts of data recorded via satellite with limited computational resources and prior knowledge of the target gas.

First, a method for selecting a vertical retrieval grid is presented that minimizes the amount of prior knowledge appearing in the estimate due to applied constraints. This selection method uses the information content of the retrieval to optimally determine the number and spacings of vertical levels, while accounting for correlations between the various layers. Failing to select an appropriate vertical grid for water vapour and temperature retrievals, such as using levels equally spaced in pressure, was found to reduce the information content by as much as 30% on average.

Overwhelmingly, the slowest part of an iterative retrieval is evaluating the forward model. As a result, the computation may fail to keep pace with data acquisition. Procedures to parametrise the Reference Forward Model (RFM) by creating a set of pre-tabulated look-up tables (LUTs) of absorption cross-sections for individual gases were developed. Since such LUTs can be unruly in size, a simple compression scheme based on linear interpolation was analysed that reduced the total LUT size to just 4.4% of the original. Additionally, the fine spectral grid for radiative transfer was also reduced using similar methods. Model comparisons to a global atmospheric ensemble showed negligible increases in error compared to IASI instrument noise and a factor of 130 increase in computational speed. These LUTs are openly available for use by the scientific community, whether using the RFM or to be incorporated into any forward model.

Finally, a linear retrieval scheme was developed to estimate total column amounts of carbonyl sulphide (OCS) at a rate roughly 10^4 times faster than a typical iterative retrieval. This scheme incorporates two concepts not utilised in previously published linear estimates. First, all physical parameters affecting the signal are included in the state vector and accounted for jointly, rather than treated as effective noise.

Second, the initialisation point is determined from an ensemble of atmospheres based on comparing the model spectra to the observations, thus improving the linearity of the problem. The entirety of IASI data from 2014 was analysed and showed spatial features of OCS total columns including depletions over tropical rainforests, seasonal enhancements over the oceans, and distinct OCS features over land. Error due to assuming linearity was found to be on the order of 11% globally for OCS. Comparisons to surface VMR *in situ* samples taken by NOAA show seasonal correlations greater than 0.7 for five out of seven sites across the globe. Furthermore, this linear scheme was applied to OCS, but may also be used as a rapid estimator of any detectable trace gas using IASI or similar nadir-viewing instruments.

Contents

List of Figures

List of Abbreviations

ACC	Antarctic Circumpolar Current
ACE	Atmospheric Chemistry Experiment
AKM	Averaging kernel matrix
AVHRR	Advanced Very High Resolution Radiometer
BBT	Blackbody brightness temperature
BRDF	Bidirectional reflectance distribution function
CA	Carbonic anhydrase
CG	Referring to the Curtis-Godson approximation
CDOM	Chromophoric dissolved organic matter
c-k	Referring to the correlated-k method
CNES	Centre National d'Etudes Spatiales
CTM	Chemistry transport model
DFN	Degrees of freedom for noise
DFS	Degrees of freedom for signal
DMS	Dimethyl sulphide
DRM	Data resolution matrix
EUMETSAT	European Organization for the Exploitation of Meteorological Satellites
FTS	Fourier transform spectrometer
IASI	Infrared Atmospheric Sounding Interferometer
IFOV	Instantaneous field of view
ILS	Instrument line shape
IMG	Interferometric Monitor for Greenhouse Gases
IR	Infrared, referring to wavelengths of light longer than visible, but shorter than microwave and radio

L1B	Level 1B, referring to calibrated and unapodized spectral radiances
L1C	Level 1C, referring to calibrated and apodized spectral radiances
L2	Level 2, referring to quantities retrieved from L1 spectral observations
LBL	Line-by-line
LUT	Look-up table
MIPAS	Michelson Interferometer for Passive Atmospheric Sounding
NEΔT	Noise equivalent differential temperature
NH	Northern Hemisphere
NLSF	Non-linear least squares fit
NWP	Numerical weather prediction
OCS	Carbonyl sulfide
OPD	Optical path difference
RFM	Reference Forward Model
RTM	Radiative transfer model
RTTOV	Radiative Transfer model for the Television infrared observation satellite Operational Vertical sounder
SH	Southern Hemisphere
SNR	Signal to noise ratio
SVD	Singular value decomposition
TES	Tropospheric Emissions Sounder
TOMCAT	Tropospheric chemistry transport model developed by the University of Leeds
TPTR	Total path transmittance regression
UV	Ultra-violet radiation
VMR	Volume mixing ratio

Excellence is an art won by training and habituation. We do not act rightly because we have virtue or excellence, but we rather have those because we have acted rightly. We are what we repeatedly do. Excellence, then, is not an act but a habit.

— Aristotle

1

Introduction

Contents

1.1 Motivation

Earth's modestly thick atmosphere is surprisingly transparent when compared to Venus and the gas giants. There are numerous regions in the electromagnetic spectrum where the majority of photons pass through the entirety of Earth's atmosphere without being absorbed. This connotes the fact that over 99% of the atmosphere comprises molecular nitrogen, molecular oxygen, and atomic argon. Molecular nitrogen and oxygen are homopolar and thus have no electric dipole moment. In other words, when these molecules vibrate they only weakly interact with light through higher order electric moments and the even weaker magnetic moments [*Bransden and Joachin*, 2003]. Argon is even more transmissive, because an atomic gas only absorbs and emits light according to electronic orbital transitions in the visible and ultraviolet. Amazingly, the overwhelming majority of thermal absorption that creates the greenhouse effect and dominates climate is controlled by the composition of less than 1% of Earth's atmosphere. These remnants are referred to as trace gases.

Furthermore, molecular nitrogen is extremely inert, while molecular oxygen generally requires conditions much hotter than terrestrial temperatures or the presence of catalysts to combust. Therefore, atmospheric chemistry is also driven by trace gas composition. Identifying the sources and sinks of trace gases provides knowledge of fundamental physical process while tracking abundances is crucial towards forecasting air quality assessments. However, given the shear size of the Earth, atmospheric sampling becomes a daunting task.

Remote sensing from satellite provides unmatched spatial and temporal coverage of the Earth. However, physical quantities must be inferred from the observed spectral radiance as opposed to *in situ* sampling (e.g., radiosondes), while the estimation process (colloquially known as a retrieval) tends to be ill-posed and computationally intensive. Prior constraints are normally required to stabilize the retrieval, which ultimately biases the resulting estimate with prior assumptions. Additionally, instruments aboard polar orbiting satellites, such as the infrared atmospheric sounding interferometer (IASI), may record over 1 million observations and push upwards of 30 Gbytes worth of data per day. For decades the mathematical theory and method of inverting a physical model based on observations for optimal results has been established. However, optimality takes time and resources; sometimes more than is available. Therefore, the problem facing trace gas retrievals (as well as many current fields of study) is how to make signal driven products in a timely manner while sifting through "big data." Given the fact that computer transistors are approaching quantum limits, it cannot be assumed that satellite observations will be fully utilized in the future by simply waiting long enough for processing speed to increase.

1.2 Thesis outline

The first part of this thesis research investigates the impact constraints have upon retrieving temperature and trace gas vertical profiles independent from prior assumptions. Atmospheric retrievals are sometimes referred to as a "dark art" if only because one can artificially improve the condition of the problem by tightening constraints or reducing the number of estimates. As the constraints tighten, less of the true state appears in the estimate. If the constraints are too weak, then the retrieval may attempt to fit signal noise and provide spurious results. As implied, there is indeed an optimal balance of constraint versus attempted number of estimates where explicit constraints are traded for reduced representations such that little to no prior knowledge appears in the estimate. This balance is found not through art, but through methodical evaluation.

A new and robust method for selecting vertical levels is proposed, evaluated, and compared against two other simpler methods [*Vincent et al.*, 2015]. Aside from quantifying and recommending which selection method to use, such studies are important for two reasons. First, artefacts of prior assumptions can be removed post estimation with minimal loss of content. This is crucial when assimilating satellite data into forecasting models, especially when forecast data itself was used as the prior knowledge, so such models do not create circular evaluations. Second, increasing the number of estimated parameters increases computation time. Vertical level selection performed before initiating a retrieval directly saves numerical resources.

The second portion of research looks at parametrising the slowest part of an iterative retrieval, the forward model, for increased performance with a negligible increase in error. The forward model generally represents the physics of the problem at hand and, in this case, takes the atmospheric state and simulates what is observed by the instrument (i.e., spectral radiance). Retrievals seek to invert the forward model. Currently, evaluating the forward model tends to be several orders of magnitude slower than the other sections of the retrieval combined, depending somewhat on the forward model method. By creating sets of look-up tables (LUTs) for individual gases, the forward model can be sped up by a factor of 100 while LUT interpolation error remained well below instrument noise levels [*Vincent and Dudhia*, 2016]. Furthermore, these LUTs are independent of the forward model that created them and could be used as a separate record of data for the radiative transfer community.

Finally, a new linear trace gas retrieval method was developed and applied towards estimating carbonyl sulphide (OCS). While the equations of radiative transfer are non-linear, if the initial atmospheric state is sufficiently close to the true state, then the retrieval is performed linearly in one step. Such methods are clearly not as accurate as iteratively accounting for non-linearities and do not presume to be. However, if the forward model is precomputed without having to simulate spectra several times per observation, then computations are $10^4 - 10^6$ times faster. The main benefit of exploring linear retrievals is to fully utilize the spatial and temporal sampling of IASI to identify interesting areas and trends across the globe, so that the more expensive iterative retrievals are efficiently targeted.

Furthermore, OCS is a poorly understood trace gas with a variety of physical implications. The majority of atmospheric sulphur is found in OCS, which makes it crucial towards understanding the global sulphur cycle. Additionally, since plants absorb OCS without respiring it during the night, observing land sinks of OCS is an indirect way to remotely deduce photosynthetic activity of vegetation. Furthermore, the majority of OCS is produced in the oceans from microscopic organics, living and deceased, and therefore connects oceanic to atmospheric processes. However, routine detection of OCS is challenging and only possible with the latest generation of hyperspectral satellites. As such, the presented linear detection method shows interesting distributions of OCS estimates that may prove useful towards understanding the fundamental mechanisms driving the creation and destruction of OCS. The limitations of estimating OCS using this linear method are explored and quantified based on the current understanding of OCS abundances.

While retrieval methods, such as optimal estimation, approach theoretical limits of detectability, there is still a vast gulf between connecting the abundance of observations to unknown physics processes. This thesis explores the space between optimally accurate results and pragmatic achievability.

Flying is learning how to throw yourself at the ground and miss.

— Douglas Adams

2

Overview of IASI

Contents

2.1 Design and orbital specifics

The Infrared Atmospheric Sounding Interferometer (IASI) is a Fourier transform spectrometer (FTS) designed for remote sounding of the atmosphere and surface. There are currently two identical versions of IASI on board the MetOp-A and MetOp-B satellites, with a third planned for launch by 2016 to 2017 aboard MetOp-C. The MetOp network is a series of polar low-earth orbiting satellites carrying multi-sensor payloads, including IASI, to conduct operational weather forecasting. Data collected by the satellite payloads are ingested incrementally throughout the day into weather models and used to create continuously updated forecasts [*Hilton et al.*, 2009].

The IASI sensors were developed by the Centre National d'Etudes Spatiales (CNES) in association with the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT). The primary operational mission of IASI is to measure vertical temperature and water vapour profiles across the globe on a daily basis for meteorological forecasting. IASI was designed to measure temperature to within 1 K accuracy and water vapour within 10 % at a vertical resolution of 1 km [*Hilton et al.*, 2011], although in practice this is only achievable in conjuction with *a priori* data from forecast models. In other words, the desired vertical resolution of IASI retrievals for temperature and water vapour is not realisable from the observation itself and must be interpolated by referencing prior weather fields.

The details of IASI's optical design are complicated, but fundamentally it is a Michelson interferometer. A generic schematic of a FTS is shown in Fig. 2.1. Light from the source enters the optical system and is sized and collimated before passing through a 50/50 reflector where 50% of the light is reflected while the remaining 50% is transmitted. One of the optical paths is then reflected off a stationary mirror, while the other is reflected off a movable mirror. The two are then recombined to form an interference pattern. Adjusting the position of the movable mirror changes the optical path of one beam relative to the other, thus introducing a variable phase shift. By scanning the movable mirror, the phase shift changes with time and creates an interferogram (i.e., an interference pattern on the detector that also changes with time). Taking the Fourier transform of the interferogram then yields the observed spectrum as a function of frequency. A detailed diagram of IASI's optical design is shown in Fig. 2.2a for reference.

IASI uses three separate detectors to record a large spectral range of the thermal infrared (IR) from 645 to 2760 $\mathrm{cm^{-1}}$ (15.5 µm to 3.6 µm), so range is not limited by the spectral sensitivity of a single detector. The first and second detectors are made of HgCdTe while the third is composed of InSb. Data are collected with a spectral sampling resolution of 0.25 $\mathrm{cm^{-1}}$ (determined by the maximum displacement of the movable mirror) and apodized to 0.5 $\mathrm{cm^{-1}}$ as part of level 1 (L1) processing. Since IASI records an exceptional amount of data for a satellite based sensor, it is necessary to pre-process the raw interferograms (level 0) prior to transmission. Data are radiometrically calibrated using both an on-board blackbody (Fig. 2.2b) and a view of deep space background that provides a reference of both hot and

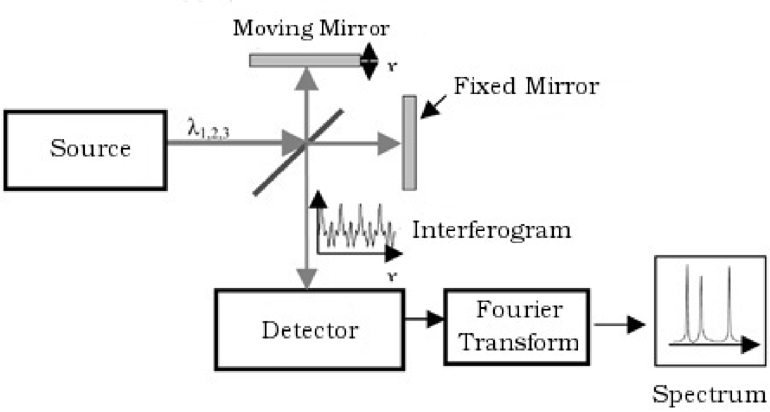
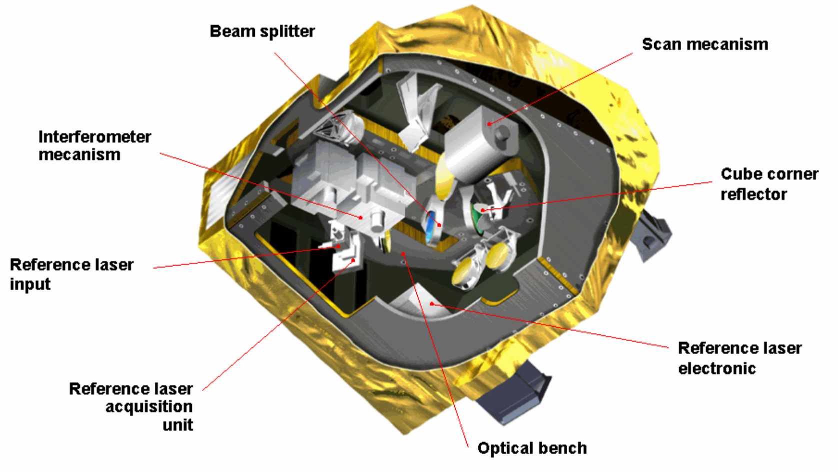


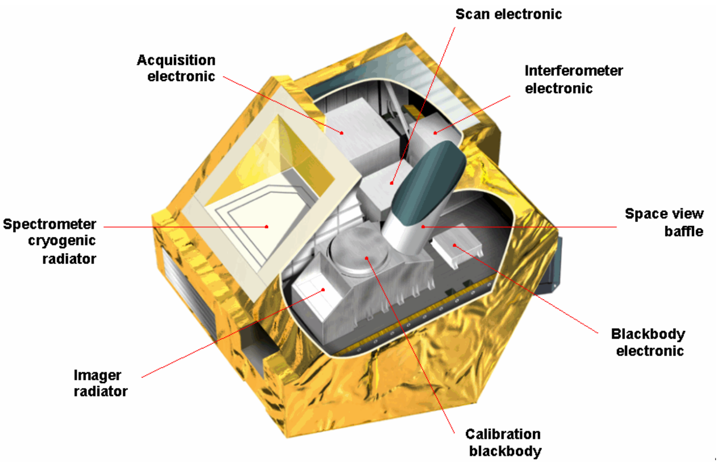
Figure 2.1: A schematic representation of generating spectra from the interferograms produced by the Michelson interferometer. This modified figure was taken from *Saptari* [2004].

cold sources. Next, Fourier transforms are performed by internal processors to generate raw spectra. The frequency axis is then calibrated to the reference laser tracking the position of the pivoting corner cube reflector and the three spectra are combined to produce one continuous spectrum. Finally, calibrated spectra and quality data are transferred to the ground station where the sampled spectrum is apodized to localize spectral mixing and make the observed radiance easier to model. Processing spectra on board before transmission reduces the data rate demand from 45 Mbit/s to 1.5 Mbit/s.

IASI is a “whisk broom” imaging spectrometer; the spatial dimensions are acquired by scanning along the direction perpendicular to the velocity of the orbiting satellite, as depicted in Fig. 2.3a. One horizontal scan takes 8 s to complete and covers a swath of ±48.3° with respect to nadir (approximately 2200 km) before moving on to the next line. During one scan 30 Earth frames are recorded followed by two calibration frames. One IASI frame is observed by four pixels in a 2×2 array. Each pixel has an instantaneous field of view (IFOV) of 14.65 mrad, corresponding to a circular observation 12 km in diameter at nadir. As the scan proceeds from nadir towards the edge, the IFOV becomes elongated and elliptical in shape. At the far edge of the scan the IFOV has a minor axis of 20 km and major axis



(a) Top down view.

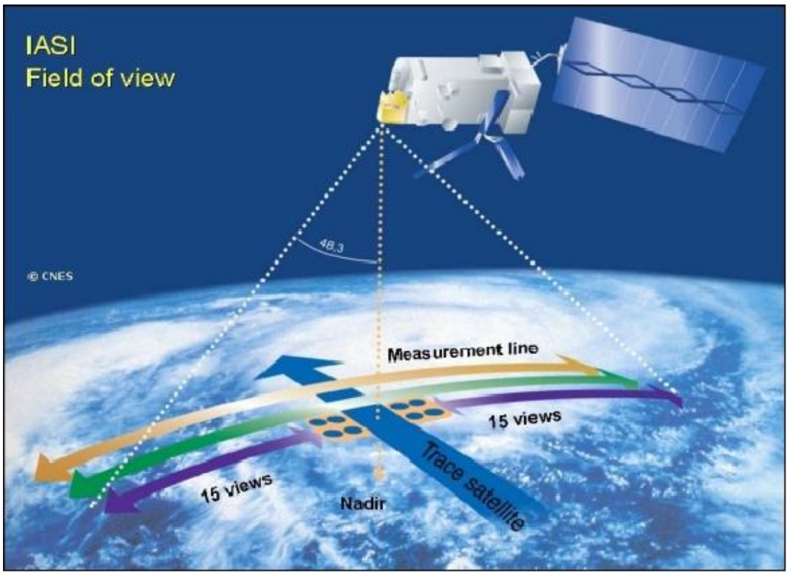


(b) Bottom up view.

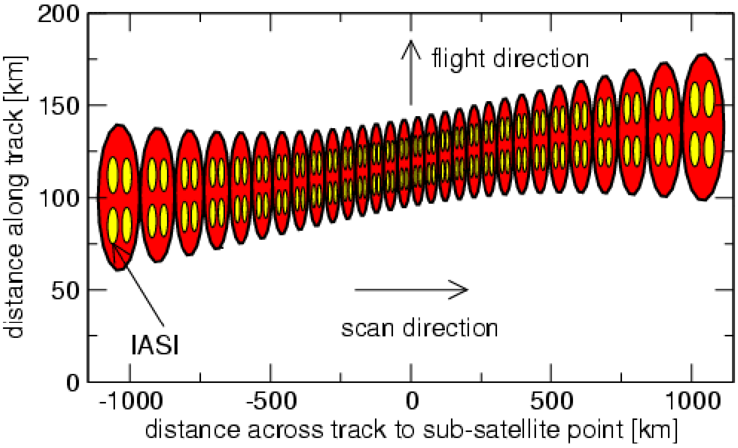
Figure 2.2: Internal drawings of IASI's optical design. (Pictures credit: CNES, 2009)

of 39 km. Figure 2.3b shows the change in IFOVs and combined FOV as IASI flies while scanning perpendicularly.

Near complete Earth coverage is achieved twice daily, providing 12 hour temporal resolution of the atmospheric state with a descending node local equator crossing time of 09:30 for IASI-A and 10:30 for IASI-B. A list of IASI sensor and orbital properties are displayed in Table 2.1 for reference. Additionally, IASI observations are collocated spatially (although with different IFOV sizes) and temporally with other instruments aboard the MetOp satellite; such as the Advanced Very High Resolution Radiometer (AVHRR), the Advanced Microwave Sounding Unit (AMSU),



(a) The IASI field of view depicted as the sensor scans along the measurement line.



(b) A graphical representation of how the field of view changes during one scan while the satellite flies forward in orbit.

Figure 2.3: IASI's field of view during orbit while scanning the transverse direction. (Figures credit: CNES, 2009)

the Microwave Humidity Sounder (MHS), and the Global Ozone Monitoring Experiment-2 (GOME-2).

Table 2.1: Characteristics of the IASI sensor and orbit (taken from CNES and EUMETSAT).

Parameter	Characteristic	Parameter	Characteristic
Spectral Range	$645 - 2760\,\mathrm{cm}^{-1}$	Polar Orbit	Sun-synch.
Spectral Resolution	$0.5\,\mathrm{cm}^{-1}$ apodized	Altitude	$\sim 817\,\mathrm{km}$
NEΔT at 280 K	$0.2 - 0.35\,\mathrm{K}$	Eccentricity	0.0012
Line Scan Rate	8 s	Inclination	98.7° to equator
Stare Interval	151 ms	Nadir IFOV shape	Circular
Earth scans/line	30	Nadir IFOV	12 km diameter
Calibrations/line	2	Pixel Number	4 Pixels
Orbital Period	101 min	Data Rate	1.5 Mbits/s
Reliability	$> 80\%$ over 5 yrs	Availability	$> 97\%$ over 5 yrs

Combining observations from both IASI-A and B allow for contiguous sounding of the atmosphere without spatial gaps along the equator. Figure 2.4 shows a typical daily Earth observation in a thermal window channel using both A and B data. Complete morning daytime and evening night-time sampling across the globe occurs within a *local* time span of approximately 2.5 hours. In other words, all data is collected between approximately 8:40 - 11:20 AM or PM locally. While no region of the Earth is left unobserved, IASI never observes the atmosphere during the afternoon as the orbits are sun-synchronous. The afternoon orbit is served by NOAA satellites, currently *NOAA-19*.

Spectral radiances that have been calibrated and apodized are referred to as level-1C (L1C) data. Figure 2.5 shows a typical observed radiance spectrum for a tropical atmosphere where the surface is approximately 295 K. Notice that the IASI spectral range samples the higher energy side of peak terrestrial emission such that spectral radiance decreases with wavenumber. In this example, spectral radiance at $2000\,\mathrm{cm}^{-1}$ is just 10% the radiance at $1000\,\mathrm{cm}^{-1}$.

Blackbody radiation represents an ideal situation of quantum harmonic oscillators radiating with perfect efficiency at a uniform temperature. While Fig. 2.5 shows that the atmosphere is clearly not a perfect blackbody, the concept is still useful. The Planck function at a given wavenumber (ν) for blackbody spectral

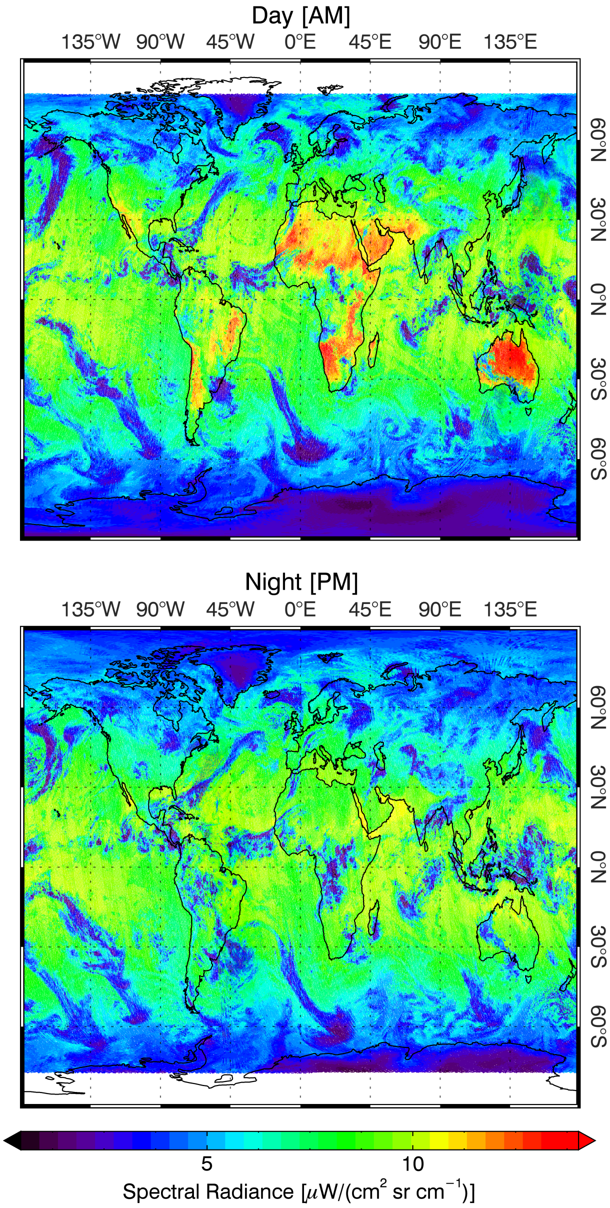


Figure 2.4: Typical daily observations combining both IASI-A and B instruments looking through a thermal window channel at 964 cm^{-1}. The top picture shows sun illuminated observations along the descending node during the day for surface solar zenith angles less than 90° and the bottom picture shows night observations along the ascending node for surface solar zenith angles greater than 90°.

radiance is given by

$$B_\nu (T) = \frac{2hc^2\nu^3}{\exp\left(hc\nu/k_B T\right) - 1}, \tag{2.1}$$

where h is the Planck constant, c is the speed of light, and k_B is the Boltzmann

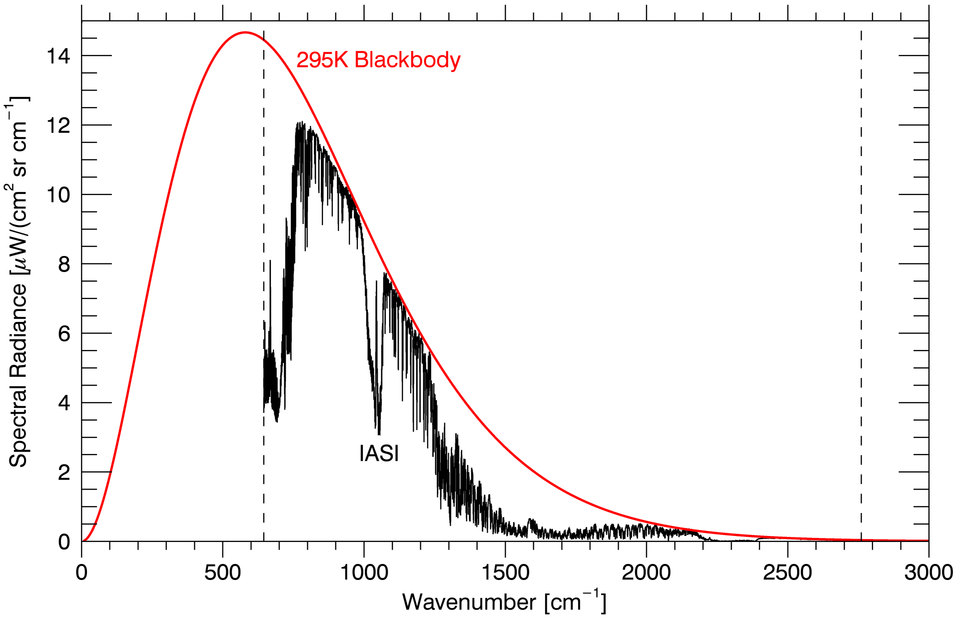


Figure 2.5: A typical spectral radiance observed by IASI of a tropical atmosphere is shown in black. The dashed lines represent the boundaries of the recorded spectral range, while the red line illustrates the curve an ideal blackbody at 295 K would radiate.

constant. For a given wavenumber, B is a function only of T. Therefore, Eq. (2.1) can be inverted to provide what is commonly referred to as blackbody brightness temperature (BBT);

$$T_b\left(L_\nu\right) = \left[\frac{k_B}{hc\nu}\ln\left(1+\frac{2hc^2\nu^3}{L_\nu}\right)\right]^{-1}, \tag{2.2}$$

where L_ν is the observed spectral radiance. Thus, BBT represents the temperature a blackbody would have in order to produce the observed radiance values at each specific wavenumber. Often it is desirable to work in units of BBT rather than radiance, because there is a linear correspondence between the observed BBT and the temperature of the state, while radiance is a non-linear function of temperature. When estimating an atmospheric quantity from satellite, transforming to BBT thus improves the linearity of the retrieval as discussed in *Rodgers* [2000, Ch. 5.1].

Figure 2.6 shows an example of transforming the spectral radiance in Fig. 2.5 to BBT. Notice that the curvature of the Planck function is removed and the higher baseline closely represents surface temperature for nadir viewing situations. One consequence of flattening the sampled spectrum is that it couples the scene

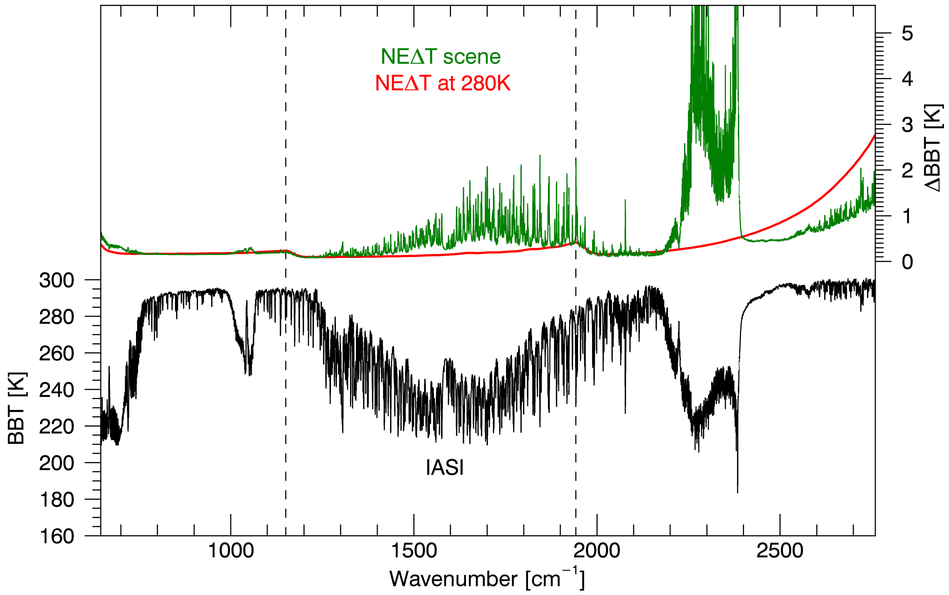


Figure 2.6: Brightness temperature is shown in black by converting the spectral radiance in Fig. 2.5 using Eq. (2.2). The NEΔT at 280 K for IASI is shown in red, while the scene specific NEΔT is shown in green. Additionally, the dashed lines represent the breaks between the three different detectors.

independent radiometric noise to the observed radiances. For example, the noise equivalent differential temperature (NEΔT) represents the spectral noise of the instrument in terms of variations about the BBT when observing a blackbody at a constant temperature (normally 280 K). However, the true NEΔT of the scene is a function of the observed radiance and is related to the independent NEΔT at 280 K by applying the product rule to Eqs. (2.1) and (2.2):

$$\mathrm{NE\Delta T}_s\,(L) = \left.\frac{\partial B}{\partial T}\right|_{280\,\mathrm{K}} \times \frac{\partial T_b}{\partial L} \times \mathrm{NE\Delta T}_{280\,\mathrm{K}}. \tag{2.3}$$

It is common for the $\mathrm{NE\Delta T}_{280\,\mathrm{K}}$ to be reported as a metric of instrument noise [*Hilton et al.*, 2011]. However, care must be taken when developing retrieval schemes using units of BBT. Figure 2.6 shows that the $\mathrm{NE\Delta T}_{280\,\mathrm{K}}$ is only approximately accurate compared to the scene specific NEΔT at wavenumbers below roughly 1300 cm^{-1}. In general, the NEΔT of the scene should be used to accurately represent the noise of the observation in BBT space.

2.2 Cloud detection

Photographs of Earth taken from space provide obvious evidence that at any given time, approximately 50% of the Earth is covered by clouds. Of course this varies hourly, but the fact is that much satellite observational data of the surface is obscured by cloud cover. While the tropopause acts as an effective lid preventing much water vapour from ascending into the stratosphere, the presence of cloud in the scene affects both tropospheric and stratospheric sounding. In particular, identifying sources and sinks of gases from surface processes is crucial for understanding annual carbon budgets and pollution cycles, so even low level clouds disrupt the scene. As a result, it is common practice to simply flag cloud-contaminated pixels, remove them from the dataset, and move on to the next pixel. For pixels with thin cloud or partial cloud coverage it may be possible to radiatively model the cloud optical properties along the observed path and retain the ability to estimate gas quantities. However, if the pixel is completely covered by dense cloud, then the area below the cloud is totally inaccessible via thermal IR.

Cloud fraction is a key parameter for characterising cloud cover in an observed scene and is the percent (or ratio) of the pixel that is filled by cloud. By itself, cloud fraction only estimates the apparent amount of cloud that could reside at any vertical tropospheric level. When paired with additional multispectral information further cloud properties may be deduced, such as the vertical location of the cloud top (cloud top pressure), cloud top temperature, cloud phase (liquid versus ice cloud), optical depth, and multiple cloud layering. EUMETSAT currently uses both the collocated AVHRR derived product and a numerical weather prediction (NWP) threshold detection method in conjunction to conservatively determine if a scene is clear to a high degree of confidence [*August et al.*, 2012]. As a consequence, cloud coverage false positives surely occur, but the degree they do is largely uncertain without global cloud truth data for comparison. A third detection method based on training an artificial neural network (ANN) [*Li et al.*, 2001] is planned for operational use in the near future.

AVHRR is a high spatial resolution radiometer that observes in six spectral channels spanning the visible out to 12.5 µm and is used to remotely determine cloud cover and surface temperature. In fact, one IASI pixel overlays roughly 100 AVHRR pixels. Towards the end of 2010 the AVHRR derived cloud fraction was smoothed to the spatial resolution of IASI and packaged in the IASI L1C data structure for ease of analysis. While the AVHRR processor chooses from several possible methods, it commonly uses a spatial coherence test with the 11 µm channel to find pixels deviating beyond an empirical threshold and tag them with a binary cloud flag [*Saunders*, 1986]. Therefore, cloud fraction for an IASI pixel is determined by taking the ratio of cloudy AVHRR pixels to the total number of AVHRR pixels in the IASI field of view.

Additionally, the NWP test compares forecast surface temperature estimates with observed brightness temperatures from the IASI window channels at 832.5 and 900.25 cm^{-1}. If the observed temperature is less than 1 K below the predicted skin temperature, the pixel is flagged cloudy. Such a strict threshold of 1 K is more appropriate over water where emissivity is close to unity with little variation. However, variations and errors in surface emissivity over land frequently lead to false positives. As a result, approximately 75 % of the pixels tested are declared cloudy with this method [*August et al.*, 2012]. Threshold difference values are not set in stone and depend upon the tolerance to cloud in the retrieval. For example, *Hadji-Lazaro et al.* [2001] used threshold values of 8 K over sea and 15.3 K over land when retrieving total column amounts of CO, which tend to flag the densest and highest clouds while tolerating thinner and lower cloud coverage.

Ideally, all three cloud detection methods should yield the same result, but radiometric and spectroscopic techniques sense cloud indirectly. Without direct measurements of cloud properties, such as estimating the index of refraction via polarimetry or effective extinction with LIDAR, indirect measurements will generally agree, but with systematically noticeable differences. Figure 2.7 shows a comparison of the current cloud detection techniques used by EUMETSAT's L2 processor. Not surprisingly, AVHRR cloud detection tends to flag cloud clear scenes around

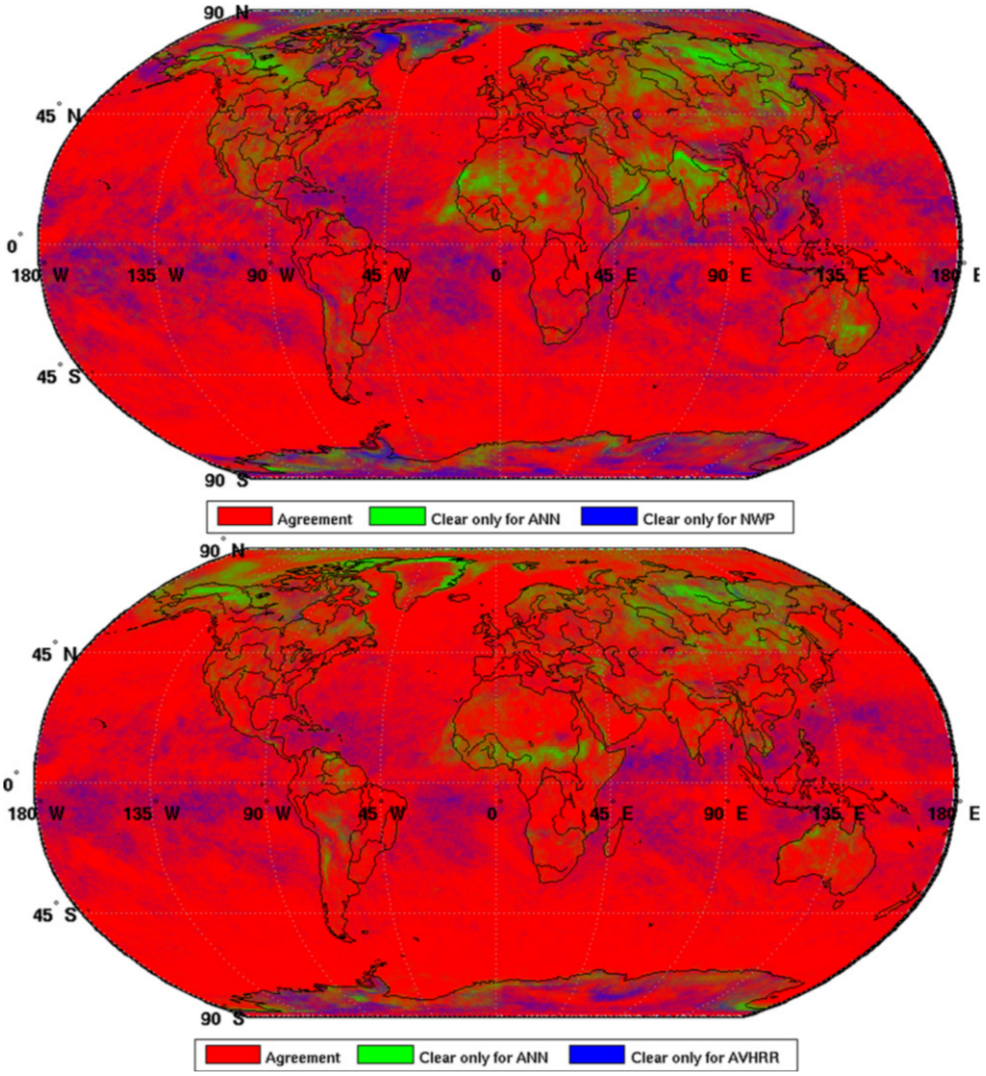


Figure 2.7: A comparison of three cloud detection routines used in the IASI L2 processor provided by EUMETSAT. This figure was taken from *August et al.* [2012].

coastlines and forest transition areas due to smaller scale changes in geography. Furthermore, the ANN method appears to miss cloud in high latitude areas over land. Perhaps the ANN training set lacked such a cloud scenario. Criticality aside, all three methods tend to agree the vast majority of the time.

One particular advantage of using AVHRR cloud fraction is that, since embedding this product within the IASI L1C data in 2010, there is no requirement to download and access a separate dataset, such as the EUMETSAT L2 results. Additionally the AVHRR cloud fractions are available at full spatial resolution for every IASI

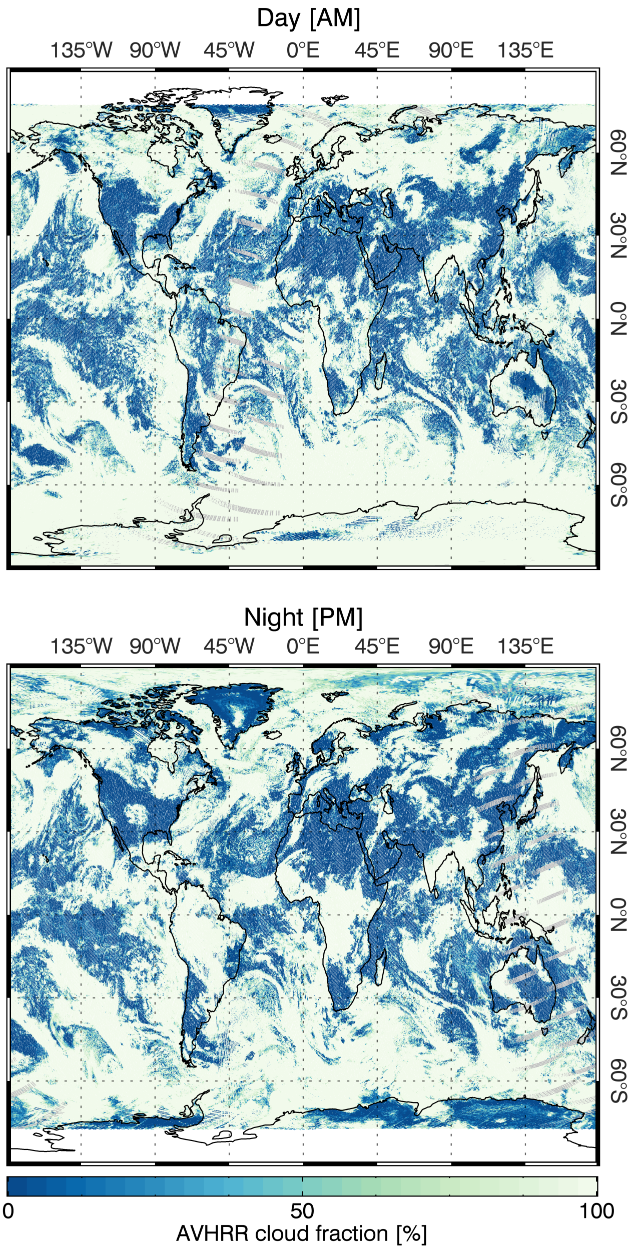


Figure 2.8: IASI cloud fractions derived from the co-located AVHRR instrument are shown for same scene represented in Fig. 2.4.

pixel, while the EUMETSAT L2 cloud fraction presents only the least cloudy pixel of the four per interval. Figure 2.8 shows an example of the AVHRR cloud fraction calculated for each IASI pixel previously displayed in Fig. 2.4. Comparing the two figures shows that cloud fraction alone contains no information on whether IASI is observing a low-warm cloud or a high-cold cloud.

2.3 Solar reflection angle

When observing Earth in the IR, solar reflections contribute as much irradiance as terrestrial emission between approximately 3 and 4 µm ($3300 - 2500\,\mathrm{cm}^{-1}$) for sun-illuminated Lambertian surfaces. Wavenumbers lower than $2500\,\mathrm{cm}^{-1}$ are generally considered to be dominated by thermal emission where solar influences are negligible. However, the degree to which this is true depends upon the surface temperature, surface type, and satellite viewing angles relative to the sun. Surely solar influences are zero during the night, but may be important for trace gas retrievals down to around $2000\,\mathrm{cm}^{-1}$ during the day [*Amato et al.*, 2002].

For surfaces that are perfectly Lambertian (i.e., reflections distributed evenly over the 2π steradian hemisphere) it makes no difference what direction IASI looks relative to the sun. However, most realistic surfaces have a significant specular lobe that distributes light preferentially about the specular path where the angle of reflection is equal to the angle of solar incidence. This creates a cone about the specular line where ray paths are more probable within this cone than outside. Therefore, if IASI observes along this path it is possible to image a partial reflection of the sun. This is often referred to as sun glint, which is a particular concern over water surfaces.

Designers of retrieval schemes using wavenumbers above approximately $2000\,\mathrm{cm}^{-1}$ have two choices: 1) Model the reflected solar component using prior knowledge of the surface bidirectional reflectance distribution function (BRDF) [*Torrance and Sparrow*, 1967], or 2) Avoid analysing scenes affected by sun glint. Unfortunately, global BRDFs specific to Earth's surfaces (especially in the IR) are poorly sampled and modelled. Idealised assumptions about specular versus Lambertian can be made, but to varying degrees of accuracy. Often it is easiest to simply avoid viewing geometries where sun glint may be observed.

Figure 2.9 shows a diagram of the problem. IASI L1C data contains surface zenith (θ) and azimuth (φ) angles for both IASI and the sun at the centre location of the observed pixel. From the zenith and azimuth angles the angle from solar

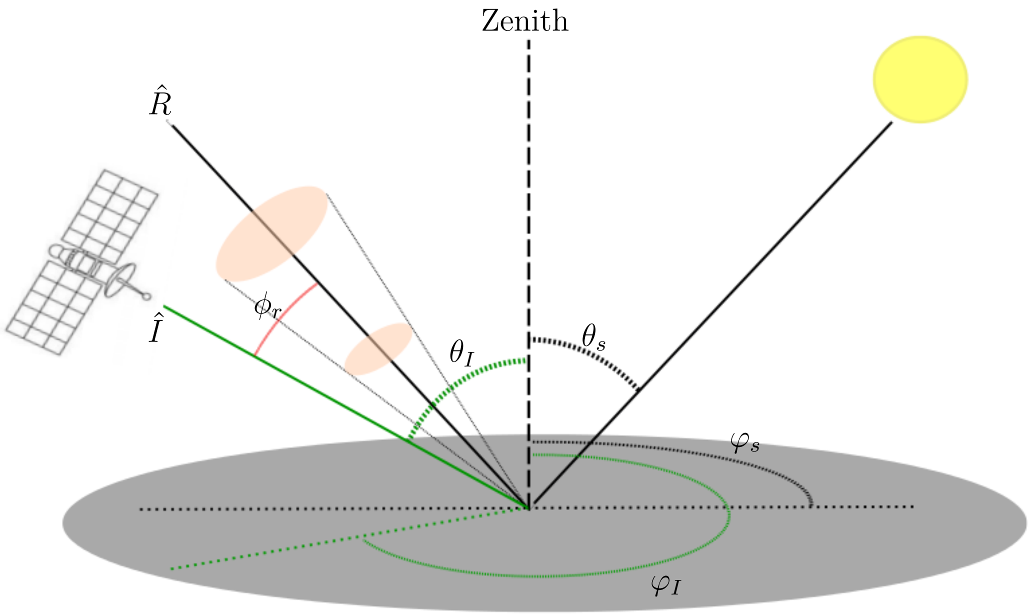


Figure 2.9: This diagram depicts the viewing geometry of IASI relative to the sun and specular solar reflection path at the centre surface location of an IASI pixel.

reflection is calculated according to

$$\begin{aligned}\cos(\phi_r) = \hat{R}\cdot\hat{I} = &-\sin(\varphi_s)\sin(\theta_s)\sin(\varphi_I)\sin(\theta_I) \\ &-\cos(\varphi_s)\sin(\theta_s)\cos(\varphi_I)\sin(\theta_I) + \cos(\theta_s)\cos(\theta_I)\,,\end{aligned} \tag{2.4}$$

where $\hat{R}$ is a unit vector pointing along the direction of specular reflection from a flat surface and $\hat{I}$ is a unit vector pointing from the observed surface to IASI. Values of ϕ_r close to zero represent IASI viewing directly along the solar specular reflection path where sun glint is maximum. Therefore, a cone of exclusion about $\hat{R}$ can be created where IASI observations close to the specular path are ignored.

An illustration of the solar reflection angle for IASI is shown in Fig. 2.10. While the true surface is rough, the large $12-39\,\mathrm{km}$ diameter of an IASI pixel averages over many sub-pixel surface facets such that even extended mountain ranges will appear relatively smooth. The latitude of minimum solar reflection angle varies with season and will be above the equator during Northern Hemisphere (NH) summer and below in NH winter.

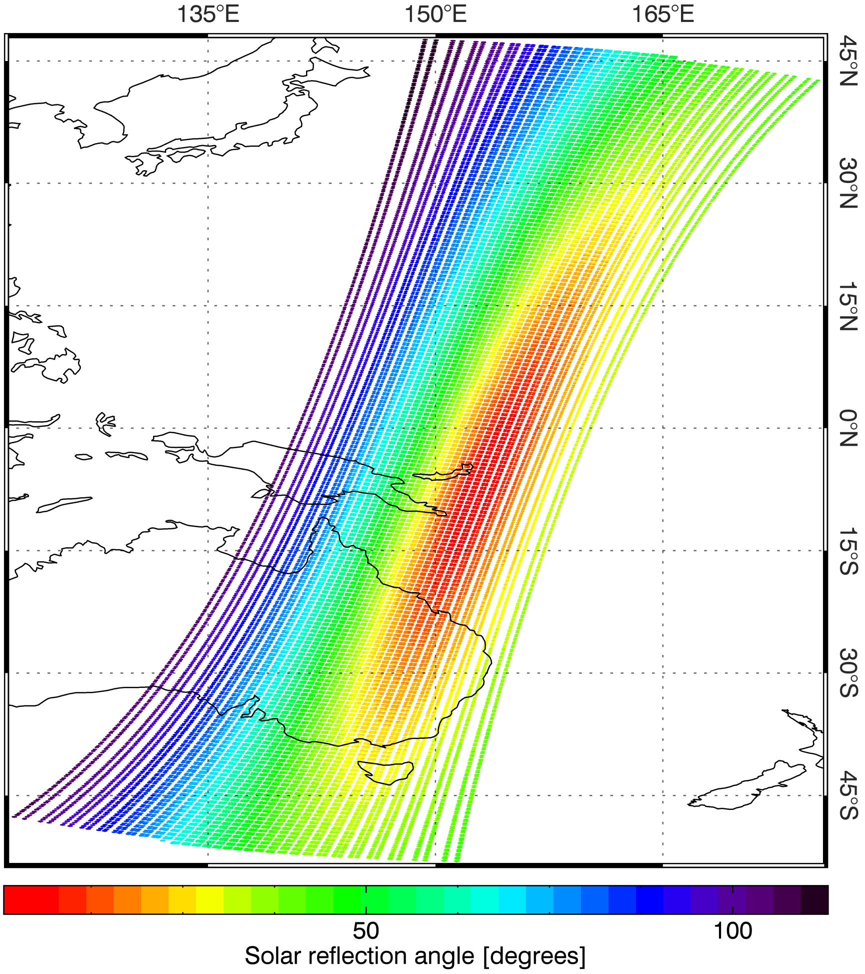


Figure 2.10: The IASI angle from solar reflection is shown for a portion of the daytime descending orbit. Notice that the minimum solar reflection angle is below the equator because this orbit occurred on the 26th of October, after the fall equinox.

2.4 Current IASI level 2 processors

Level 2 (L2) processors ingest calibrated and apodized L1C spectra and return products estimating the atmospheric state that generated the observed spectra. Those products could be temperature, gas concentrations, cloud properties, surface properties, or all mentioned products bundled together. Currently, there is an operational IASI L2 processor developed by EUMETSAT to support weather forecasting. Specified primarily to deliver estimates of surface temperature, atmospheric temperature, and water vapour profiles, the IASI L2 processor has also been adapted to retrieve ozone profiles along with column amounts of CO_2, CO, N_2O, and CH_4 [*August et al.*, 2012].

While the EUMETSAT product is the only one currently available in near-real time through the same pipeline as the L1C data, there are other L2 processors developed for IASI, as shown in Table 2.2. Aside from the specifics, there are two components of a L2 algorithm that determine the bulk trade-off between precision and computation. First, numerical radiative transfer calculations are normally the slowest retrieval component by far. Line-by-line (LBL) radiative transfer models (RTMs) reference a spectroscopic database, such as HITRAN or GEISA [*Rothman et al.*, 2013; *Jacquinet-Husson et al.*, 2008], and calculate monochromatic transmission along a vertically layered atmospheric path for each spectral frequency considered. As a result, LBL radiances are highly accurate, but computationally expensive. This expense can be mitigated by pre-calculating LUTs on a pressure and temperature grid while managing accuracy degradation. Alternatively, methods using spectral averaging for radiative transfer reduce computer time and, subsequently, accuracy. Historically, as processing power evolved, analytic band models gave way to the correlated-k method [*Goody et al.*, 1989] and now radiances parametrised as a function of profile dependent predictors, such as RTTOV, is a leading alternative method to LBL routines [*Matricardi et al.*, 2004]. While fast, parametrised RTMs do not achieve the same accuracy as LBL methods, particularly for atmospheric states not represented in the training set.

Secondly, inverse methods are not unique and different routines produce estimates at various rates and accuracies. Perhaps the most accurate and time consuming is the non-linear least squares fitting (NLSF) method where a cost function is iteratively minimised. When applied to atmospheric retrievals the RTM is rerun each iteration. Sometimes many iterations are necessary depending upon the applied constraints and starting points. Two popular surrogates are training artificial neural networks and using principle component regression, both statistical in nature [*Schlüssel et al.*, 2005]. While statistical methods prove to be substantially faster than NLSF, their fundamental shortcoming is their inability to retrieve anomalous states not found in the training ensemble. Thus, routine atmospheric states are estimated with high statistical certainty, but increasing the range of possible retrieved states fundamentally decreases the resulting certainty.

Table 2.2: Currently available L2 processors utilizing IASI data. This list is not comprehensive. Institutional acronyms not previously defined are: Rutherford-Appleton Laboratory (RAL), Université libre de Bruxelles (ULB), and Nello Carrara Institute of Applied Physics (IFAC).

L2 Processor	Developer	Inverse method	RTM type
Real-time IASI	EUMETSAT	Statistical/NLSF	Parametrised
Global CH_4	RAL	NLSF	Parametrised
FORLI (O_3, CO, HNO_3)	ULB	NLSF	LBL (LUTs)
Global NH_3	ULB	Linear	LBL (LUTs)
CO Retrieval	Univ. Leicester	NLSF	LBL
KLIMA-IASI (CO_2)	IFAC (Italy)	NLSF	LBL
KOPRAFIT	Karlsruhe Tech.	NLSF	LBL
φ-IASI	Univ. Basilicata	NLSF	LBL (LUTs)
Atmosphit	ULB	NLSF	LBL

You can have data without information, but you cannot have information without data.

— Daniel Keys Moran

3

Atmospheric retrieval theory

Contents

Atmospheric profile retrievals with a nadir-viewing satellite tend to be significantly ill-conditioned. In other words, the attempted number of estimated parameters (n) is greater than the degrees of freedom that come from the signal rather than the prior knowledge about the expected state and its covariance (also known as *a priori*). Therefore, constraints must be applied to stabilize the retrieval. Constraints in vector and matrix form can be chosen in a variety of ways [*Kulawik et al.*, 2006]. When constraints are applied to an ill-conditioned problem, the *a priori* information inevitably becomes an artefact of the resulting estimate [*Rodgers*, 2000]. At this point the designer of the retrieval has two choices, (1) tolerate artefacts from the *a priori* or (2) move to a representation that is better conditioned.

This chapter reviews inverse theory, applied to ill-conditioned atmospheric sounding. While there are at least two separate notations commonly used, we adopt the notation of *Rodgers* [2000], where numerous derivations supporting the following discussion can be found.

3.1 Optimal estimation

In general, a radiance spectrum is observed by an instrument, in this case IASI. The goal of a retrieval is to estimate the atmospheric state producing that spectrum by inverting the forward model describing the physics of radiative transfer through the atmosphere. When the radiative transfer function is sufficiently linear about a reference state vector ($\boldsymbol{x_0}$) of length n, the forward model (F) can be linearised according to

$$\boldsymbol{y} - F\left(\boldsymbol{x_0}\right) = \mathbf{K}\left(\boldsymbol{x} - \boldsymbol{x_0}\right) + \boldsymbol{\epsilon}, \tag{3.1}$$

where $\boldsymbol{y}$ is the measured spectrum of length m, $\boldsymbol{x}$ is the true state to be estimated, and $\boldsymbol{\epsilon}$ is the error in the measured signal relative to the forward model. Furthermore, $\mathbf{K} \in \mathbb{R}^{m \times n}$, referred to as the Jacobian matrix, is defined to be a matrix of partial derivatives such that $K_{ij} = \partial F_i\left(\boldsymbol{x}\right)/\partial x_j$.

Solutions to Eq. (3.1) can be estimated in the maximum *a posteriori* framework (also known as optimal estimation) by considering a linearisation about an *a priori* reference state ($\boldsymbol{x}_\mathrm{a}$). Estimates of an atmospheric state ($\hat{\boldsymbol{x}}$) are given by

$$\begin{aligned}\hat{\boldsymbol{x}} &= \boldsymbol{x}_\mathrm{a} + \left(\mathbf{K}^\mathrm{T}\mathbf{S}_\epsilon^{-1}\mathbf{K} + \mathbf{S}_\mathrm{a}^{-1}\right)^{-1}\mathbf{K}^\mathrm{T}\mathbf{S}_\epsilon^{-1}\left(\boldsymbol{y} - F\left(\boldsymbol{x}_\mathrm{a}\right)\right) \\ &= \boldsymbol{x}_\mathrm{a} + \mathbf{G}\left(\boldsymbol{y} - F\left(\boldsymbol{x}_\mathrm{a}\right)\right),\end{aligned} \tag{3.2}$$

where $\mathbf{G}$ is referred to as the gain matrix. The covariance matrix of the stochastic error in the measurements is denoted as $\mathbf{S}_\epsilon$. Since raw spectra from a Fourier transform spectrometer (FTS) such as IASI are generally uncorrelated, $\mathbf{S}_\epsilon$ has zeroes in the off-diagonal elements while the diagonal elements are the variances of the signal at that spectral position. However, because IASI spectra are apodized [*Amato et al.*, 1998], off-diagonal spectral correlations are thus introduced into $\mathbf{S}_\epsilon$ (discussed further in Sect. 3.6).

The term *a priori* is meant to include both a mean state, $\boldsymbol{x}_\mathrm{a}$, and its covariance, $\mathbf{S}_\mathrm{a}$. Inverting $\mathbf{S}_\mathrm{a}$ in Eq. (3.2) applies a “soft” constraint upon the solution, penalizing estimates that deviate greatly from the profile provided in the prior estimate. One method to determine $\mathbf{S}_\mathrm{a}$ for atmospheric temperature and trace gases is to download analysis data from a NWP source and calculate statistical covariances from a global ensemble or about the local region considered. However, statistical covariances calculated this way may not be invertible if the ensemble does not contain enough truly independent sample atmospheres. In this case, $\mathbf{S}_\mathrm{a}^{-1}$ might be replaced with an alternative method, such as Twomey–Tikhonov regularization, where smoothness constraints are imposed by considering first and second derivatives of the profile and treated as tuning parameters [*Kulawik et al.*, 2006]. While such methods are common, they include less prior knowledge in the sense that higher order physical correlations are intentionally ignored and suggest that the dimensionality of the retrieval should be reduced to improve the condition of the inverse problem.

In the Gaussian linear case, the posterior covariance (i.e., the estimated covariance of $\hat{\boldsymbol{x}}$) is found to be

$$\hat{\mathbf{S}}_x = \left(\mathbf{K}^\mathrm{T}\mathbf{S}_\epsilon^{-1}\mathbf{K} + \mathbf{S}_\mathrm{a}^{-1}\right)^{-1}. \tag{3.3}$$

This is a convenient result, because it means that the uncertainties and correlations between retrieved parameters are generated as a by-product of the retrieval process. Eq. (3.3) also highlights the fact that if $\hat{\mathbf{S}}_x = \mathbf{S}_\mathrm{a}$, then the retrieval has done nothing to improve upon the *a priori* and is completely insensitive to the estimated parameters.

Further diagnostic information about the retrieval is succinctly contained in a unitless $n \times n$ matrix known as the averaging kernel matrix (AKM), defined as

$$\mathbf{A} = \frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{x}} = \mathbf{GK}. \tag{3.4}$$

In this case, the rows of $\mathbf{A}$ correspond to the retrieved profile levels and can be thought of as smoothing functions ideally peaking at the referenced level, but with finite widths providing a measure of vertical resolution. Columns of $\mathbf{A}$ depict the response of the retrieval to δ-function perturbations in the true state profile levels. Furthermore, Eq. (3.2) can be rewritten in the more insightful but less practical form,

$$\hat{\boldsymbol{x}} = \left(\mathbf{I}_n - \mathbf{A}\right) \boldsymbol{x}_{\mathrm{a}} + \mathbf{A}\boldsymbol{x} + \mathbf{G}\boldsymbol{\epsilon}, \tag{3.5}$$

where $\mathbf{I}_n$ is the identity matrix with n diagonal elements. Written this way, it becomes clear that the estimate of state, $\hat{\boldsymbol{x}}$, is a weighted average of the true state and the prior state. Ideally, $\mathbf{A}$ approaches the identity matrix and no prior state appears in the estimate. However, this is seldom the case for nadir-viewing unless performing a maximum likelihood retrieval where there is by definition no *a priori* information.

3.2 Scalar figures of merit

Repeated analysis of $\mathbf{A}$ can be unwieldy when developing a retrieval algorithm. Therefore, a scalar "figure of merit" is often desirable that allows for multiple matrices of $\mathbf{A}$ to be compared in a straightforward manner. The degrees of freedom from the signal (DFS) is one such possible metric and is calculated by taking the trace of the averaging kernel matrix,

$$d_{\mathrm{s}} = \mathrm{Tr}\left(\mathbf{A}\right). \tag{3.6}$$

Perfectly conditioned inverse problems will have DFS values equal to the number of state parameters, n. Additionally, the "degrees of freedom for noise" (DFN) is a useful quantity to consider and tells how many degrees of freedom are allocated specifically to the noise. As expected, addition of the two yields,

$$d_s + d_N = \mathrm{Tr}\left(\mathbf{I}_m\right) = m, \tag{3.7}$$

the number of spectral measurements. Ultimately, an efficient retrieval does not simply maximize the DFS, but maximises the signal to noise ratio (SNR) of the retrieval while maintaining a target threshold of the DFS. Therefore, it is often necessary to consider removing any spectral channels that marginally increase the DFS and add mostly to the DFN.

The term ‘information’ is rather vague and depends greatly on the field of study. The DFS is one metric to quantify the information of a retrieval. Another possibility is the Shannon information content [*Shannon*, 2001], which fundamentally is a measure of the change in entropy of the problem before and after a measurement is made. To put it another way; the posterior uncertainty should be less than the prior uncertainty as a result of the measurement. Therefore, the Shannon information content is the ratio of the prior to posterior uncertainties,

$$H = \frac{1}{2}\log_2\left|\hat{\mathbf{S}}_x^{-1}\mathbf{S}_\mathrm{a}\right| = -\frac{1}{2}\log_2\left|\mathbf{I}_n - \mathbf{A}\right|, \tag{3.8}$$

where the base two logarithm is conventionally used in information theory to express H in dimensionless units of bits. The factor of $1/2$ results from defining one bit as reducing the standard deviation of the uncertainty by a factor of two, rather than the variance. The interpretation of H differs from the DFS in that 2^H represents the number of different atmospheric states that can be distinguished, whereas the DFS is the number of independent quantities that can be measured.

Finally, another common scalar metric to evaluate is the trace of the posterior covariance ($\hat{\mathbf{S}}_x$), i.e., the sum of the estimated uncertainties. While both the DFS and Shannon information content are only useful metrics in a constrained retrieval, it is desirable to minimize the posterior uncertainty for both constrained and unconstrained inverse problems. Alternatively, if many parameters are jointly estimated, but only one is of interest, then the uncertainty of that particular parameter can be evaluated at the expense of the other elements of $\boldsymbol{x}$.

3.3 Constraint mapping

Solutions to ill-conditioned inverse problems can be improved by simply reducing the number of estimated parameters. However, information content may be lost if the retrieved state is reduced too much. Therefore, as the original parameter space is reduced, the information content should be monitored in a consistent mathematical way. This is done by defining operators that map the retrieval between the original and reduced state space.

Consider two vertical grids for the problem of retrieving atmospheric profiles. First, a fine grid from the discretization of the full state vector, $\boldsymbol{x}$, chosen with enough vertical resolution to accurately calculate the equations of radiative transfer. Second, a coarser grid on which the retrieval is carried out, here on referred to as the retrieval grid ($\boldsymbol{z} \in \mathbb{R}^l$), where $l < n$. A coarser retrieval grid is necessary when the DFS are significantly less than n, in order to improve the condition of the retrieval. Mapping from the fine to the coarse grid imposes a "hard" constraint, where only solutions to the reduced representation are considered.

For convenience we apply a linear mapping from the retrieval to the fine grid so that

$$\boldsymbol{x} = \mathbf{W}\boldsymbol{z}. \tag{3.9}$$

Here $\mathbf{W} \in \mathbb{R}^{n \times l}$ is a mapping matrix that is commonly a piecewise linear interpolation operator. In fact, $\mathbf{W}$ could be any general linear transformation that maps the full state vector to a reduced retrieval vector, such as a truncated right singular vector matrix of $\mathbf{A}$ or the signal-to-noise matrix [*Ceccherini et al.*, 2009; *Rodgers*, 2000]. While singular value decomposition methods guarantee maximum retention of DFS, they transfer the full state vector into a reduced space that has no direct physical meaning. Here we apply linear interpolation to maintain a physical link between elements of the state vector and levels in the atmosphere.

To transform the *a priori* to the retrieval state space, the following averaging operation is required,

$$\boldsymbol{z} = \mathbf{W}^*\boldsymbol{x}, \tag{3.10}$$

where $\mathbf{W}^*$ is the pseudo inverse of $\mathbf{W}$. While there are infinite ways in which $\mathbf{W}^*\mathbf{W} = \mathbf{I}_l$, the most common is to define $\mathbf{W}^*$ in the least-squares sense,

$$\mathbf{W}^* = \left(\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)^{-1}\mathbf{W}^{\mathrm{T}}. \tag{3.11}$$

With these operators, any number of mathematical transformations are possible that interpolate or average parameters between the fine and coarse grids. However, care must be taken that such operations have sound physical reasoning. For example, the prior covariance of the retrieval can be expressed on the coarse retrieval grid

$$\mathbf{S}_{z\mathrm{a}} = \mathbf{W}^{*}\mathbf{S}_{\mathrm{a}}\mathbf{W}^{*\mathrm{T}}, \tag{3.12}$$

but interpolating $\mathbf{S}_{z\mathrm{a}}$ to a finer grid cannot be assumed valid if the prior covariance is only known on the coarse grid. This is because off-diagonal elements in $\mathbf{S}_{\mathrm{a}}$ represent physical correlations between various pressure levels and may not follow Gaussian statistics [*von Clarmann*, 2014].

More rigorous derivations and discussions of mapping between states can be found in *Worden et al.* [2006], *Bowman et al.* [2006], and *Rodgers* [2000, ch. 10]. Summarising the key mapping relationships relevant to this application, the following list is helpful:

$$\{m \times l\} \qquad \mathbf{K}_z = \mathbf{K}_x\mathbf{W}, \tag{3.13}$$

$$\{l \times m\} \qquad \mathbf{G}_z = \left(\mathbf{K}_z^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}\mathbf{K}_z + \mathbf{S}_{z\mathrm{a}}^{-1}\right)^{-1}\mathbf{K}_z^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}, \tag{3.14}$$

$$\{n \times n\} \qquad \mathbf{A}_x = \mathbf{W}\mathbf{G}_z\mathbf{K}_x = \mathbf{G}_x\mathbf{K}_x, \tag{3.15}$$

$$\{l \times l\} \qquad \mathbf{A}_z = \mathbf{G}_z\mathbf{K}_x\mathbf{W} = \mathbf{W}^{*}\mathbf{A}_x\mathbf{W}. \tag{3.16}$$

While the retrieval is performed on the coarse grid, the resulting atmospheric profile is interpolated to the fine grid providing the final estimate. Therefore, when calculating the DFS in Eq. (3.6) we must use the true averaging kernel, $\mathbf{A}_x$, and not $\mathbf{A}_z$. Alternatively, one could also retrieve coarse perturbations to a fine grid profile to ensure sharp features, such as the tropopause, were maintained. In which case, these mapping relationships still hold true.

3.4 Linear error characterization

Error from the posterior covariance matrix of the atmospheric state (e.g., a temperature profile) can be characterised by summing its individual error components,

$$\hat{\mathbf{S}}_x = \mathbf{S}_r + \mathbf{S}_f + \mathbf{S}_s, \tag{3.17}$$

where the right hand covariance terms are referred to as retrieval noise, forward model parameter error, and smoothing error respectively. Each error covariance is assumed to be uncorrelated with the others. Forward model error inherent to the modelling assumptions and spectroscopic parameters can also be included in the error budget. However, forward model error is difficult to quantify and requires comprehensive validation campaigns where the true state of the atmosphere is well sampled by external sources, such as radiosondes, that are coordinated with satellite overpass observations. Previous work has shown that error inherent to the RFM is well below IASI noise levels, with the exception of Q-branch transitions for CO_2 and CH_4 due to the mixing of spectral lines [*Ventress*, 2014].

Measurement noise specific to IASI is projected into the state space of the estimate, thus providing the retrieval noise covariance, given by

$$\mathbf{S}_r = \mathbf{G}_x \mathbf{S}_\epsilon \mathbf{G}_x^{\mathrm{T}} = \mathbf{W}\mathbf{G}_z \mathbf{S}_\epsilon \mathbf{G}_z^{\mathrm{T}} \mathbf{W}^{\mathrm{T}}. \tag{3.18}$$

Retrieval noise is often considered to represent the lower limits of achievability and is useful when performing sensitivity analyses prior to retrieval algorithm design. To account for apodization the measurement noise covariance ($\mathbf{S}_\epsilon$) is convolved with IASI's instrument line shape function. In the absence of apodization, measurement noise is spectrally uncorrelated for FTIR instruments so the measurement noise covariance is diagonal.

Error in the retrieval due to forward model parameter variability is referred to as forward model parameter error, or “systematic" error for short, and is

$$\mathbf{S}_f = \sum_b \mathbf{W}\mathbf{G}_z \mathbf{K}_b \mathbf{S}_b \left(\mathbf{W}\mathbf{G}_z \mathbf{K}_b\right)^{\mathrm{T}}. \tag{3.19}$$

The summation is taken over the elements in $\boldsymbol{b}$ (a vector of physical parameters which influence the measurement) with covariances $\mathbf{S}_b$ that are not included in the retrieval. Treating model parameter error in this way assumes that each contributing parameter is physically uncorrelated with all others. In the case of a temperature retrieval using spectral channels between 700 and 800 cm^{-1}, for example, one could consider contributions from the total columns of water vapour and ozone to be elements of $\boldsymbol{b}$ that contribute systematically to retrieval error where each $\mathbf{S}_b$ may be approximated from climatological databases or forecast models. Alternatively, water vapour and ozone could be jointly retrieved with temperature as elements of $\boldsymbol{x}$ where $\mathbf{S}_\mathrm{a}$ is expanded to include each $\mathbf{S}_b$.

Finally, the error introduced into the retrieval by representing fine spatial structure on a coarse vertical grid is called the smoothing error. Its covariance is

$$\mathbf{S}_s = \left(\mathbf{I}_n - \mathbf{A}_x\right) \mathbf{S}_t \left(\mathbf{I}_n - \mathbf{A}_x\right)^\mathrm{T}, \tag{3.20}$$

where $\mathbf{S}_t$ is the covariance of an ensemble of true states about the mean. However, if the true states were known, then there would be no point in the retrieval. Therefore, one must assume that the statistics of $\mathbf{S}_t$ are sufficiently well represented in observed or modelled databases, like ECMWF re-analysis data, such that $\mathbf{S}_t \approx \mathbf{S}_\mathrm{a}$. This is a safe assumption for well known physical parameters like temperature, water vapour, and ozone, but less so for the minor trace gases where a chemical transport model may be the only means to study the variability of abundances.

Notice that there appears to be a contradiction between Eqs. (3.17) and (3.3). *Rodgers* [2000, ch. 3.4] shows that Eq. (3.17) reduces to Eq. (3.3) when $\mathbf{S}_f$ is zero and $\mathbf{S}_\mathrm{a}$ is used in place of $\mathbf{S}_t$. If all parameters that noticeably contribute to the measurement are included jointly as elements in $\boldsymbol{x}$, then $\mathbf{S}_f$ is indeed zero. However, *von Clarmann et al.* [2001] showed that a joint retrieval is equivalent to incorporating an effective total measurement covariance,

$$\mathbf{S}_\epsilon^\mathrm{tot} = \mathbf{S}_\epsilon + \sum_b \mathbf{K}_b \mathbf{S}_b \mathbf{K}_b^\mathrm{T}, \tag{3.21}$$

where unwanted parameters are excluded from the retrieval and treated as sources of observed spectral error [*Walker et al.*, 2011]. The validity of this total measurement covariance method is contingent upon the linearity of the retrieval and accuracy of $\mathbf{S}_b$.

3.5 Nonlinear approach

Radiative transfer is a nonlinear process. Therefore, the validity of linear theory depends upon how close the initial guess and estimated solution are to the true solution. Many numerical methods are available for finding solutions to nonlinear equations. Of the possible methods, the Levenberg-Marquardt (LM) method is particularly well suited for solving atmospheric inverse problems [*Marquardt*, 1963]. In the linear case, Eq. (3.2) was found to estimate the solution by minimizing the chi-squared cost function analysed on the coarse grid,

$$\chi^2\left(\boldsymbol{z}\right) = \left[\boldsymbol{y} - F\left(\boldsymbol{x}\right)\right]^{\mathrm{T}} \mathbf{S}_{\epsilon}^{-1} \left[\boldsymbol{y} - F\left(\boldsymbol{x}\right)\right] + \left[\boldsymbol{z} - \boldsymbol{z}_{\mathrm{a}}\right]^{\mathrm{T}} \mathbf{S}_{z\mathrm{a}}^{-1} \left[\boldsymbol{z} - \boldsymbol{z}_{\mathrm{a}}\right]. \tag{3.22}$$

In the nonlinear case, the cost function in Eq. (3.22) is iteratively minimized with respect to the retrieval vector to estimate the atmospheric state:

$$\begin{aligned} \boldsymbol{z}_{i+1} = \boldsymbol{z}_i + & \left(\left[1+\gamma\right]\mathbf{S}_{z\mathrm{a}}^{-1} + \mathbf{K}_{zi}^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}\mathbf{K}_{zi}\right)^{-1} \times \\ & \left(\mathbf{K}_{zi}^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}\left[\boldsymbol{y} - F\left(\boldsymbol{x}_i\right)\right] - \mathbf{S}_{z\mathrm{a}}^{-1}\left[\boldsymbol{z}_i - \boldsymbol{z}_a\right]\right), \end{aligned} \tag{3.23}$$

where γ is a dimensionless parameter that effectively toggles between the Gauss-Newton method when $\gamma \Rightarrow 0$ and the method of steepest descent when $\gamma \Rightarrow \infty$. While the Gauss-Newton method converges quickly, it may overshoot a local minimum and then diverge or oscillate about the minimum without further convergence. Therefore, when the cost function in Eq. (3.22) increases, γ increases to shorten the step size and dampen such effects.

Convergence is determined by checking that changes in the state vector are less than its estimated error by some threshold factor. Given that the chi-squared value of the difference between the estimate and true state is expected to be the number of state elements on the coarse grid,

$$\chi^2\left(\hat{\boldsymbol{z}} - \boldsymbol{z}\right) = \left(\hat{\boldsymbol{z}} - \boldsymbol{z}\right)^{\mathrm{T}} \hat{\mathbf{S}}_z^{-1} \left(\hat{\boldsymbol{z}} - \boldsymbol{z}\right) \simeq l, \tag{3.24}$$

a conservative convergence criteria is established that terminates the iteration when the following condition is met, for example,

$$\left(\boldsymbol{z}_i - \boldsymbol{z}_{i+1}\right)^{\mathrm{T}} \hat{\mathbf{S}}_z^{-1} \left(\boldsymbol{z}_i - \boldsymbol{z}_{i+1}\right) \leq l/10. \tag{3.25}$$

Of the states calculated during the iterative process, the atmospheric state that minimizes Eq. (3.22) is selected as the retrieved estimate. Alternatively, convergence can also be expressed as a change in $F(\boldsymbol{x}_i)$.

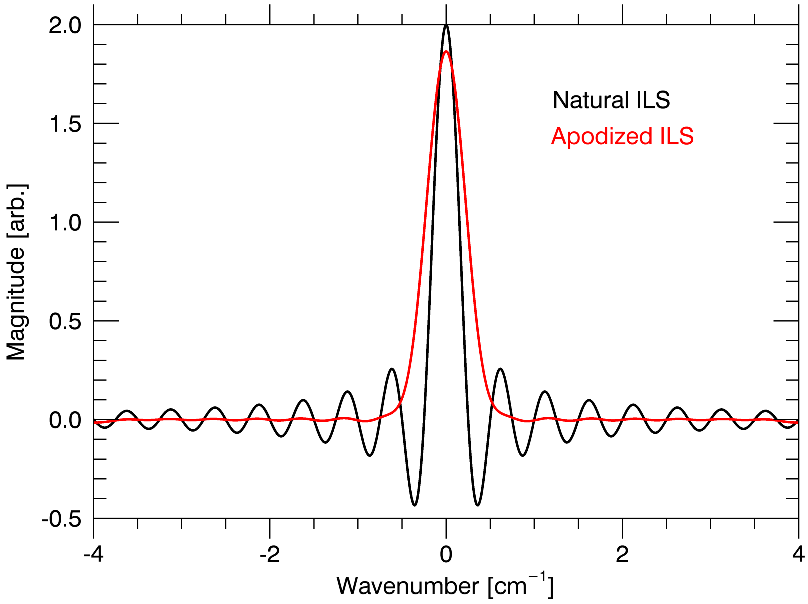


Figure 3.1: Instrument line shape functions specific to IASI are shown for the natural (sinc) ILS found in L1B data and the apodized ILS in L1C data.

3.6 Effects of apodization on retrieved quantities

Fourier transform spectrometers directly measure an interferogram by changing the optical path with a movable mirror. The spectral resolution is thus determined by the maximum optical path difference (OPD), i.e., $\delta\nu = (2L)^{-1}$, where L is 2 cm for IASI. Mechanically, the mirror can only translate a finite distance, which truncates the interferogram. This truncation can be thought of mathematically as multiplying an infinite OPD interferogram with a rectangular step function. In spectral space, the Fourier transform of a rectangle function is the cardinal sine (sinc) function, which convolves the infinite resolution true spectrum to produce what is observed by IASI. Therefore, the sinc function is the natural instrument line shape (ILS) that broadens measured spectral lines, as shown in Fig. 3.1.

Sinc functions are not numerically desirable ILS functions, because they oscillate with side lobes that slowly decay in magnitude. As a result, sharp spectral features will be followed by a 'ringing' pattern that is purely a result of the finite OPD, making it pragmatically difficult to model FTS spectra. The slow amplitude decay of the sinc means that numerical convolutions of model spectra require kernel widths up to $25\,\mathrm{cm}^{-1}$ or more in order to avoid substantial numerical rounding error, which may be computationally prohibitive. Furthermore, as the sinc includes both positive and negative values, it is possible to actually model negative atmospheric transmission values as a result of rounding error. On the plus side, FTS spectra are uncorrelated because the IASI sinc function sampled at $0.25\,\mathrm{cm}^{-1}$ is simply a delta function.

The usual solution to this modelling problem is to further convolve the spectra with an 'apodization' function that suppresses the sinc side lobes. Of the many possibilities, it is most common to use a Gaussian as the apodization function. Therefore, the apodized spectral radiance is given as

$$\begin{aligned} R_{ap}(\nu) &= \int_{-\infty}^{\infty} R_t(\nu_0)\,\mathrm{d}\nu_0 \int_{-\infty}^{\infty} w(\nu' - \nu_0)\, g(\nu - \nu')\,\mathrm{d}\nu' \\ &= \int_{-\infty}^{\infty} R_t(\nu_0)\, f(\nu - \nu_0)\,\mathrm{d}\nu_0, \end{aligned} \tag{3.26}$$

where R_t represents the true spectral radiance with infinite resolution, w is the sinc function, g is the Gaussian apodization function, and f is the resulting apodized ILS displayed in Fig. 3.1. IASI L1C data has been apodized using a Gaussian of full width at half height of $0.5\,\mathrm{cm}^{-1}$. Notice that the side lobes are almost entirely reduced at the expense of broadening the central peak. For IASI purposes the apodized ILS is adequately represented numerically with a full kernel width of only $2\,\mathrm{cm}^{-1}$. While the modelling problem is resolved, the consequence of this action is to introduce correlation between neighbouring spectral channels which must be properly accounted for in the retrieval.

Amato et al. [1998] showed that f can be written as a linear apodization operator ($\mathbf{O}$) such that

$$\boldsymbol{y}_{ap} = \mathbf{O}\boldsymbol{y} \tag{3.27}$$

$$F(\boldsymbol{x})_{ap} = \mathbf{O}F(\boldsymbol{x}) \tag{3.28}$$

$$\mathbf{K}_{ap} = \mathbf{O}\mathbf{K} \tag{3.29}$$

$$\mathbf{S}_{\epsilon}^{ap} = \mathbf{O}\mathbf{S}_{\epsilon}\mathbf{O}^{\mathrm{T}}. \tag{3.30}$$

Furthermore, they showed that when these relations are substituted into Eq. (3.2), $\mathbf{O}$ drops out of the equation if it is invertible. In other words, apodization has no effect on the retrieval assuming it is modelled properly and the apodizing function does not have singular eigenvalues, which may occur under 'heavy' apodization. If apodization is not modelled, then the problem can be avoided by selecting spectral channels with more than two positions of separation.

One common retrieval approximation is to apodize the Jacobians and model spectra in-line with the forward model, but to leave $\mathbf{S}_\epsilon$ unapodized as a diagonal matrix [*Bowman et al.*, 2006]. Clearly this introduces error into the retrieval, but to varying degrees of tolerance depending upon what parameter is estimated and if spectral channel selection is performed while avoiding adjacent channels. Modelling apodization in $\mathbf{S}_\epsilon$ informs the retrieval that spectral channels are correlated. Figure 3.2 shows an example of the spectral channel correlations appearing in $\mathbf{S}_\epsilon$ as a result of apodization by $\mathbf{O}$. Channels separated by more than three indices are correlated by less than 0.4%. All work presented in this thesis includes full treatment of apodization as presented in Eqs. (3.27) - (3.30).

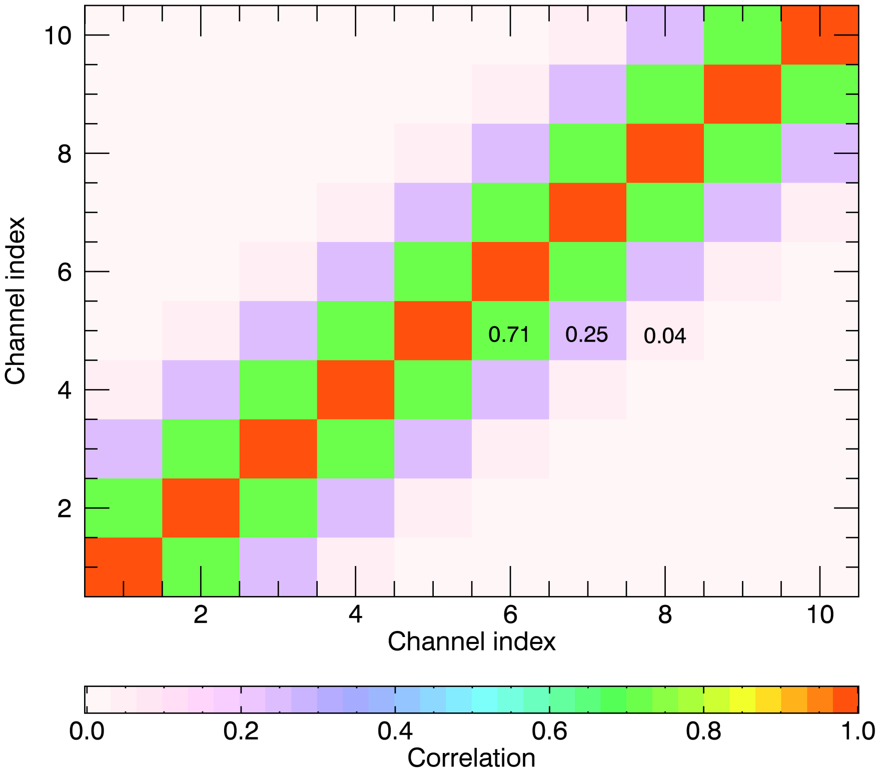


Figure 3.2: This shows a graphical representation of the spectral correlations in $\mathbf{S}_\epsilon$ due to Gaussian apodization.

When your values are clear to you, making decisions becomes easier.

— Roy E. Disney

4

Vertical level selection for temperature and trace gas retrievals

Contents

4.1 Problem overview

Retrieved profiles of temperature and composition from nadir-viewing instruments are often presented on a grid much finer than can be justified given the actual vertical resolution of the measurements. The DFS is a natural metric to use for determining the number of retrieval levels, but does not say where these levels should be located.

Analysing the cumulative trace of the AKM in whole number increments is one way to select the vertical level locations. However, this only uses the diagonal of the AKM and ignores all correlations between various levels. Therefore, a new method is proposed to determine the optimal subset from a fine vertical grid by selecting levels according to their contribution to the DFS while accounting for vertical level correlations. The entirety of this chapter was published in *Vincent et al.* [2015].

When designing a retrieval scheme it is useful first to determine the subset of coarse vertical levels that efficiently contribute to the estimate. By reducing the number of attempted estimates the retrieval relies less on formal prior knowledge and becomes more sensitive to the true state at the reduced representation. There are also computational benefits during the retrieval due to the improved conditioning of the problem, possibly faster convergence and greater tolerance of ad hoc assumptions in the *a priori*.

Consider, for example, the IASI L2 product; where temperature, water vapour, and ozone profiles are presented on a vertical grid of 100 pressure levels ranging from surface pressure up to 0.016 hPa, an altitude of approximately 80 km [*August et al.*, 2012]. While vertical resolution at this scale is highly desirable, retrievals can only be performed on such fine grid spacing at the expense of heavy dependence on the *a priori* estimate or other constraints.

Post-processing methods are developed to reduce the reliance upon *a priori* information. Possible *a priori* sources include Numerical Weather Prediction (NWP) data and chemical transport models such as the Goddard Earth Observing System Chemical transport model (GEOS-Chem) [*Bey et al.*, 2001]. While similar sources are generally of high quality, modelling artefacts do appear in atmospheric retrievals. For example, *Bowman et al.* [2006] recognized during the development of the Tropospheric Emission Spectrometer (TES) retrieval method that the full state vector grid (67-levels) used in calculating radiative transfer may be too fine for the purposes of a retrieval. They therefore decided upon a coarser retrieval grid of 14 pressure levels. However, since the DFS in their methane retrieval ranged from 0.5 to 2.0 [*Payne et al.*, 2009], the subsequent grid was still too fine for methane and discontinuities in volume mixing ratio (VMR) were observed originating from their modelled *a priori* fields [*Brasseur et al.*, 1998]. To improve the TES results, *Payne et al.* [2009] remapped the estimates to a single "representative tropospheric VMR" level that effectively removed *a priori* artefacts from the retrieval.

Previous work by *von Clarmann and Grabowski* [2007]; *Ceccherini et al.* [2009], and *Joiner and Da Silva* [1998] show that post-processing can transform a regularized retrieval to a maximum likelihood estimate of the atmospheric state. However, this is not only for use in post-processing. The idea of using a coarse retrieval grid to constrain the solution set and reduce *a priori* in the estimate can be implemented during the retrieval algorithm, while minimizing the loss of information.

The presented work is applied to profile retrievals of temperature, H_2O, O_3, CH_4, and CO using IASI. However, this methodology can be readily applied to other infrared atmospheric nadir sounding instruments, e.g. TES, Atmospheric InfraRed Sounder (AIRS) [*Aumann et al.*, 2003], and Cross-track Infrared Sounder (CrIS) [*Han et al.*, 2013] as well as different species. Comparisons between the new vertical selection method and two other common methods are presented.

4.2 Profile level selection

First, consider that there is a fine vertical grid used to discretely compute the equations of radiative transfer within a numerical framework. The spacings of this fine grid are determined by the ability to reproduce accurate radiance spectra, specifically the line shapes of individual absorbing gases. Thus, the fine grid of radiative transfer is decoupled from the vertical grid used in the retrieval, which may only be sensitive to much bulkier layers. Selecting the reduced vertical grid used in the retrieval from the fine radiative transfer grid is the topic under consideration.

From inspection of Eqs. (3.14) and (3.15), selecting an appropriate vertical grid depends upon three things; (1) the radiative response of the atmosphere, (2) the spectral resolution and noise characteristics of the observation instrument, in this case IASI, and (3) the chosen prior covariance matrix, $\mathbf{S}_\mathrm{a}$. Since IASI is already designed and collecting data, the only free parameter is the prior covariance.

Therefore, $\mathbf{S}_\mathrm{a}$ on the fine grid should be decided upon before attempting to optimize a coarse retrieval grid.

When determining a coarse retrieval grid the number and location of profile levels can be chosen in an ad hoc manner or decided based upon the distribution of information in the profile. The DFS (Sect. 3.2) is a natural scalar metric of information to use when constructing and comparing different vertical grids, because it can be directly compared to the number of attempted retrieval levels. When the DFS is approximately equal to the number of levels, then little prior knowledge appears in the estimate. Other possible scalar metrics of information include the Shannon information content and the trace of the posterior covariance [*Rodgers*, 2000, ch. 2] as discussed in Sect. 3.2. For this particular application the DFS is the most appropriate. With this in mind, three vertical level selection methods are described, ranging from the simple to the complex.

4.2.1 Equal spacing

The simplest possible selection method is to segment the atmosphere into layers of equal thickness. Levels of equal pressure may be used for better tropospheric sensitivity, or levels of equal height for stratospheric sensitivity. Nadir-viewing instruments such as IASI are typically more sensitive to the troposphere, so equal pressure spacing will be assumed for these comparisons, even though it is clearly inappropriate for species with stratospheric concentrations such as ozone.

4.2.2 Cumulative trace

Alternatively, a vertical selection method proposed by *von Clarmann and Grabowski* [2007], originally for removing hidden *a priori* from retrieved estimates, was used by *Payne et al.* [2009]. This method utilizes the vertical distribution of the DFS through the profile on the fine grid. From the averaging kernel matrix, $\mathbf{A}$, the cumulative trace is calculated as a function of the vertical axis, showing the contribution of each level to the total DFS of the retrieval. Next, the cumulative trace is segmented into equal spacings of number l and the selected coarse vertical levels interpolated from this curve. The resulting coarse pressure grid is thus irregularly spaced based on the vertical density of the DFS.

While using the diagonal of $\mathbf{A}$ is a clear improvement over equal pressure spacings, ignoring the off-diagonal sensitivities is a concern. This is because the original AKM expressed on the fine grid is likely to change morphology as levels are combined, hence the impact on the diagonal will not generally be the simple cumulation assumed. As a result, the vertical partitioning of DFS for a given atmospheric profile retrieved over, for example, 100 levels may differ from a 10 level retrieval due to correlations between levels introduced both spectrally and with $\mathbf{S}_\mathrm{a}$. Thus, a vertical selection method is desired that accounts for off-diagonal changes in $\mathbf{A}_x$ during the selection process.

4.2.3 Iterative selection

A proposed new vertical selection method is outlined as follows:

1. Calculate the DFS on the fine grid by making $\mathbf{W} = \mathbf{I}_n$. Doing so sets the coarse grid equal to the fine grid.

2. Next, a single level is removed by modifying $\mathbf{W}$ in accordance with the chosen interpolating method. In this case, piecewise linear is used for simplicity.

3. The resulting DFS from removing that level are determined from Eqs. (3.15) and (3.6). Each possible level is removed individually and its effect on the DFS noted. The removed level that decreased the DFS the least is then discarded. Removed levels are merged into neighbouring levels by redefining the $\mathbf{W}$ matrix.

4. This process is repeated to find the next least important level until all vertical levels have been ranked and discarded down to the final levels that contribute the most to the DFS.

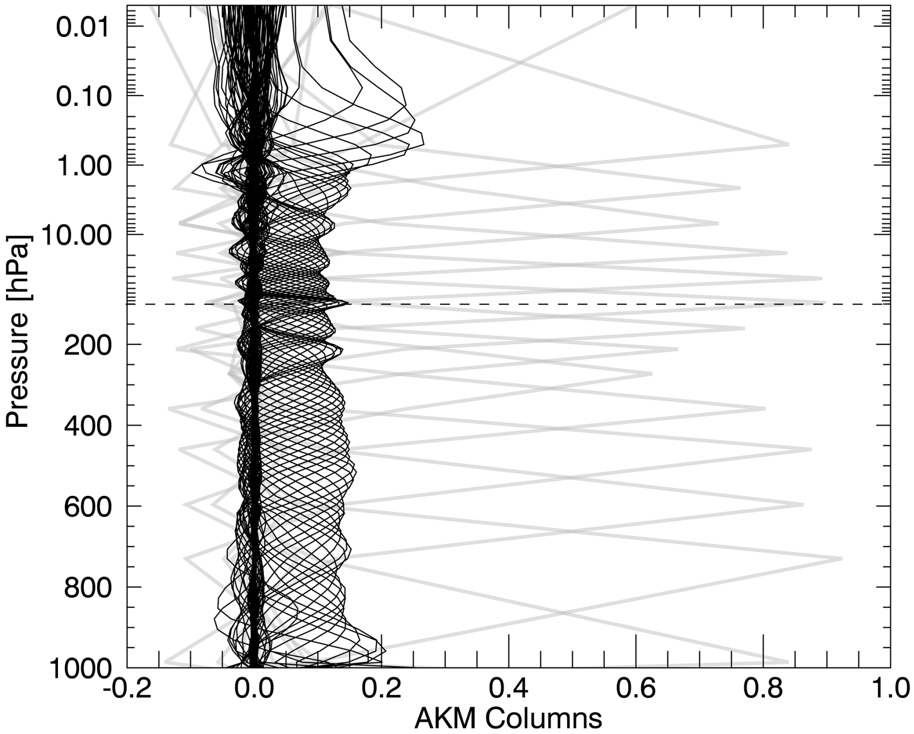


Figure 4.1: AKM columns are plotted for a sample temperature retrieval with IASI using a fine vertical grid (black) compared to a coarser grid (grey) selected using the iterative vertical selection method. Note that the pressure axis changes from linear to logarithmic above 100 hPa (dashed line).

Following this method results in a ranking of the vertical levels on the fine grid that can be truncated to produce an optimal representation on the retrieval grid for any number of levels.

To visualize the effect, Fig. 4.1 shows the columns of two AKMs for a temperature retrieval with IASI using part of the CO_2 spectral feature between 675 and 800 cm^{-1}. The smaller amplitude responses are for a retrieval on a fine 100 level pressure grid. While the larger more peaked responses are for a 15 level grid ($d_{\mathrm{s}} \approx 12$) chosen using this iterative selection method. Figure 4.1 highlights a fundamental trade-off inherent to constrained retrievals with fixed information content: More parameters can be retrieved with less sensitivity to the true state or fewer estimates attempted with greater sensitivity to the true state.

Clearly, the downside to this iterative selection method is the added complexity and computational cost of checking each available vertical level during the merging process, as compared to the method of interpolating from the cumulative trace of **A**. However, the iterative steps may be necessary to maximize DFS when there are significant off-diagonal sensitivities in **A**. In the absence of off-diagonal sensitivities the two methods converge on the same coarse vertical grid.

4.2.4 Deciding upon the number of levels

Choosing the number of retrieval levels ultimately depends upon the tolerance for *a priori* appearing in the estimate. For an effectively *a priori* free retrieval, the number of estimates should be at most equal to the DFS on the fine grid rounded down [*von Clarmann and Grabowski*, 2007], but this comes at the cost of a modest reduction in DFS. However, if it is acceptable for some *a priori* to appear in the estimate, then how many levels beyond the DFS should be included?

Figure 4.2 helps shed some light upon this question. The surface plot shown is also for a temperature retrieval with IASI using the CO_2 region previously mentioned. Pressure levels have been ranked by the iterative method just described and spectral channels have been ranked according to the channel selection method found in *Ventress and Dudhia* [2013]. The vertical and spectral selection methods bear much similarity. For this particular constrained retrieval, $d_\mathrm{s} = 7.6$ on a fine vertical grid.

Notice that the DFS increases nearly linearly with pressure level number initially, but quickly reaches a plateau once $l > d_\mathrm{s}$. The DFS axis begins at $d_\mathrm{s} = 1.99$, where retrieving temperature at two large bulk layers (troposphere and stratosphere) results in a well-constrained problem, using little *a priori*. Once more than approximately 10 pressure levels are included in the retrieval, then adding further levels negligibly increases the DFS and introduces only *a priori*.

Figure 4.2 can be extrapolated to any trace gas or temperature profile retrieval. Including many more vertical levels than DFS adds increasingly negligible information. Therefore, a straightforward way to decide upon the number of vertical grid levels is to add a few more levels than DFS on the fine grid and check via Eq. (3.15) that the loss in DFS is acceptably small.

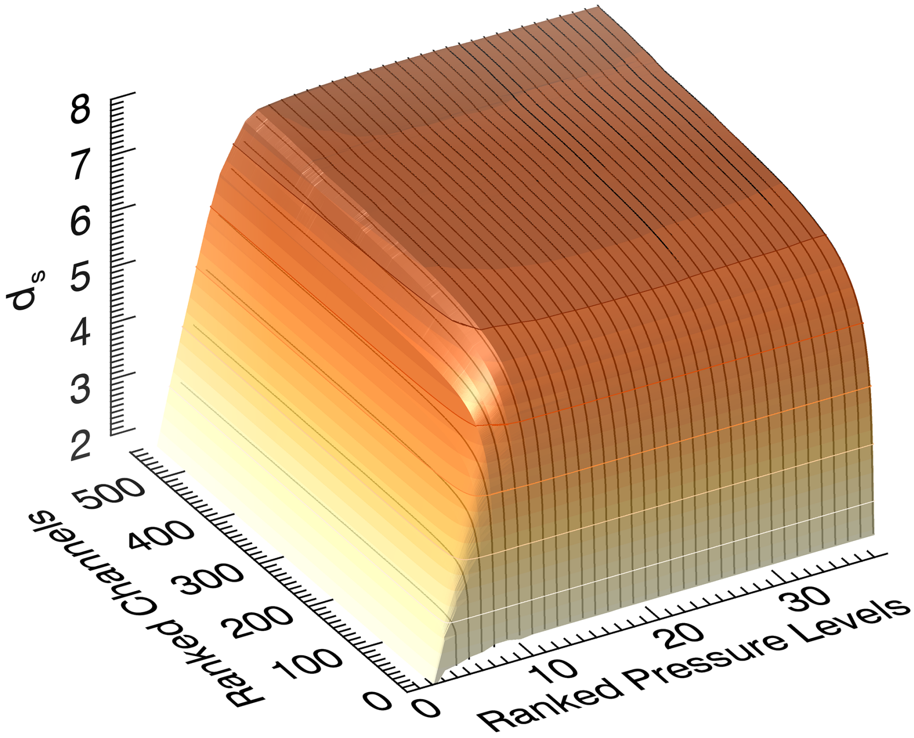


Figure 4.2: DFS for a temperature retrieval with IASI versus both ranked atmospheric pressure levels and ranked spectral channels from the CO_2 spectrum. The black lines show contours of number of pressure levels and the coloured highlight lines show contours of integer DFS.

4.3 Comparison of grid selection methods

In this section the three different selection methods are compared, using a variety of atmospheric conditions for simulated IASI retrievals of temperature, H_2O, O_3, CH_4, and CO. While there are other trace gases detectable with IASI, these four are the species with DFS values routinely greater than one. If DFS values are less than one, then conclusions about vertical distributions separate from the *a priori* cannot be made with confidence. Thus, minor trace gases are omitted from this study.

4.3.1 Computational details

Radiances measured by IASI are accurately reproduced by configuring a radiative transfer model to the specifics of IASI's orbital geometry and instrument response [*Ventress*, 2014]. Spectral modelling in this study was performed with the Reference Forward Model (RFM) using the HITRAN 2012 spectral database [*Rothman et al.*,

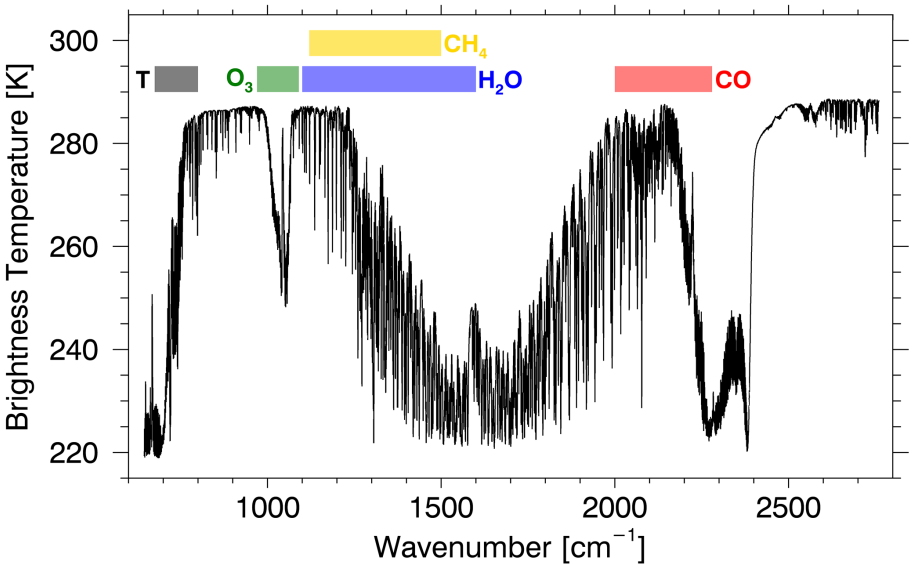


Figure 4.3: Simulated IASI spectrum showing the spectral ranges considered in this study, which are typical for temperature and trace gas profile retrievals with this instrument.

2013]. Jacobians, $\mathbf{K}$, for gas retrievals were created for estimates of the natural logarithm of VMR. Further information on the RFM can be found in the online manual [*Dudhia*, 2008].

Instrument noise was provided by the Centre National d'Etudes Spatiales (CNES) 2008 post-launch estimate [*Hilton et al.*, 2009] and used to construct the measurement covariance matrix, $\mathbf{S}_\epsilon$. Cross channel correlations in $\mathbf{S}_\epsilon$ due to apodization were accounted for as discussed in Sect. 3.6. Spectral ranges typical for IASI retrievals were selected for this study and are shown in Fig. 4.3. All spectral channels in the given ranges were included. Furthermore, for simplicity of analysis, joint retrievals of combined states were not considered.

4.3.2 Atmospheric ensemble description

Continuing work from the Thermodynamic Initial Guess Retrieval (TIGR) database [*Chevallier et al.*, 2000], an ensemble of 83 atmospheres was created by *Matricardi* [2008] specifically for the development of the RTTOV regression coefficients. Care was given to ensure that the atmospheric states represent the statistical variability of Earth's atmosphere, including extreme weather events. The 83 atmospheres were

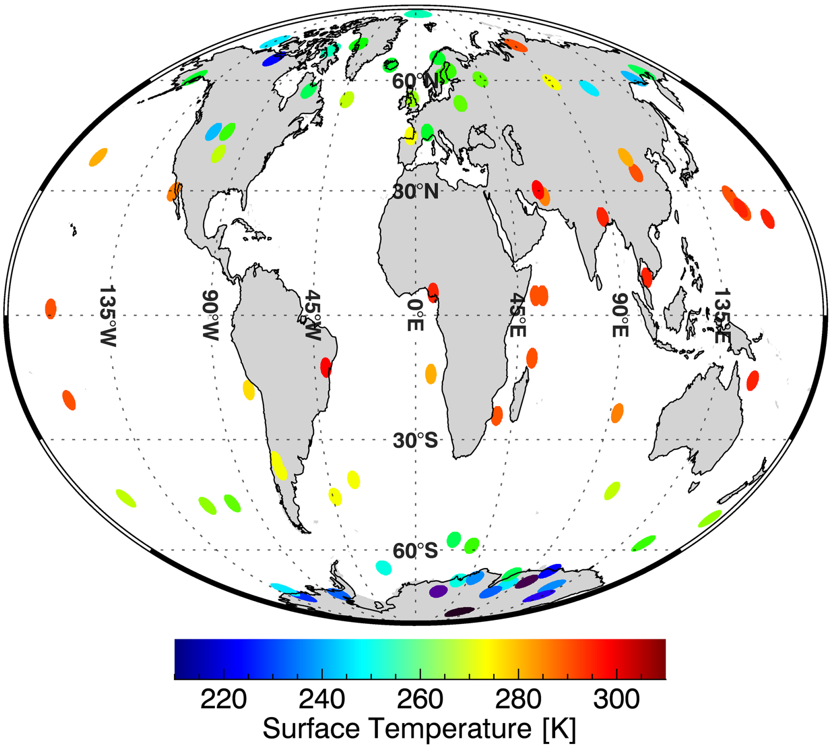


Figure 4.4: Sampling locations of the 80 atmospheres comprising the RTTOV training ensemble colour coded according to the atmospheric surface temperatures of the profiles.

originally sampled from over 120 000 000 profiles generated by the European Centre for Medium-Range Weather Forecasts (ECMWF) cycle30R2 forecasting system, creating sets for temperature, H_2O, and O_3. Other trace gas profiles, such as CH_4 and CO, were generated from the Global and regional Earth-system Monitoring using Satellite and *in situ* measurement (GEMS) database [*Hollingsworth A. et al.*, 2008].

Temperature and gas profiles are presented on a fixed fine grid of 101 pressure levels ranging from 1100 to 0.005 hPa. Of the total, three atmospheres represent minimum, maximum, and mean values of the 80 independent profiles. Therefore, those three were not used in this study. Figure 4.4 is a visual reference showing the sampling locations and atmospheric surface temperatures for the 80 atmospheres.

The original purpose of the TIGR database was to sample representative *a priori* information for accurate modelling of forward or inverse radiative transfer problems. As such, *a prior* covariance matrix was created for each constituent studied here by calculating the sample covariance from the 80 atmosphere ensemble. Therefore, the prior covariances used represent global statistics and include full off-diagonal correlations often neglected in retrieval schemes. This choice was made primarily to highlight the effect off-diagonal sensitivities in the AKMs have upon the level selection results. While it is unlikely these particular prior covariances will be used in an operational retrieval, they may provide the most interesting case study.

Furthermore, using a global prior covariance includes greater variation than a latitude specific atmosphere may experience. Therefore, the fine grid DFS represents an upper limit to what may be achieved in an operational retrieval. Figure 4.5 shows the spread of DFS for each atmospheric constituent considered as a function of its surface quantity and colour coded by surface temperature. Note that the DFS of the gases vary by about a factor of two. Additionally, there is an interesting correlation between DFS and surface temperature for H_2O, O_3, and CH_4 that is not apparent with CO. Both O_3 and CH_4 may be more detectable against hot surface air, however CO values of DFS show little relation to lower tropospheric temperature. H_2O concentration is physically correlated with temperature, i.e., warmer air holds more water, so the fact that both surface temperature and concentration coincide with the DFS is expected.

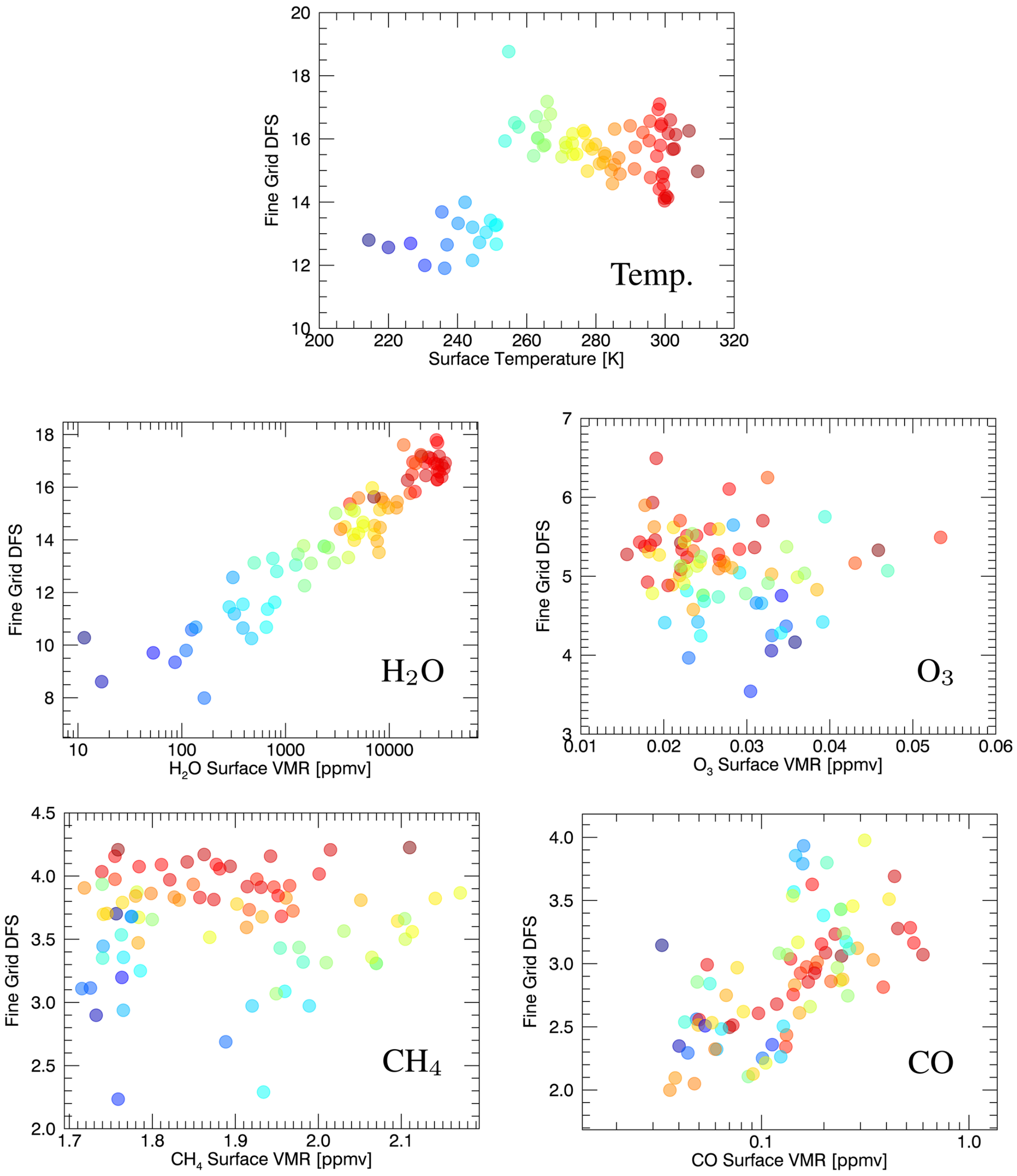


Figure 4.5: The DFS from the ensemble for temperature and trace gases as a function of their surface values. The points are colour coded by atmospheric surface temperature using the colour bar shown in Fig. 4.4.

4.3.3 Mid-latitude grid comparisons

A typical mid-latitude summer atmosphere was chosen from the ensemble to help visualize the differences of selecting coarse grids from the methods in Sect. 4.2. Contour plots of AKMs for the mid-latitude atmosphere are shown in Fig. 4.6. Consider the first two rows showing AKMs for temperature and H_2O. On the fine grid, the retrieval responses fall mostly along the diagonal in a largely symmetric pattern. Comparing the iterative to the cumulative trace coarse grid AKMs, there are differences, but they appear slight. For the following three rows of AKMs for O_3, CH_4, and CO, as the off-diagonal sensitivities increase the differences between coarse grids become more noticeable. In fact, these off-diagonal sensitivities are the motivating factor for using the iterative vertical selection method. Finally, AKMs for the equal spacing method are substantially different than the other two methods, because the location of information content, particularly in the stratosphere, is completely ignored.

These off-diagonal responses in the fine grid AKMs for O_3, CH_4, and CO are attributed to both spectral correlations from $\mathbf{K}^{\mathrm{T}}\mathbf{S}_{\epsilon}^{-1}\mathbf{K}$ and the off-diagonal terms in $\mathbf{S}_{\mathrm{a}}$ from Eqs. (3.2) and (3.4). If the full sample covariance were replaced with a simpler tridiagonal matrix for $\mathbf{S}_{\mathrm{a}}$ (i.e., smoothness constraint) or just a diagonal matrix, then the fine grid AKMs become more diagonal and symmetric. However, there are still spectral correlations to consider. In this case, the iterative selection method and the cumulative trace method will produce different grids.

Figure 4.7 shows the cumulative diagonal of the interpolated AKMs for each coarse grid method compared to the fine grid AKM for a temperature retrieval. Notice that creating a coarse grid from equal spacings of pressure reduces stratospheric DFS by approximately four. The cumulative trace coarse grid maintains greater DFS, but shows an immediate decrease in DFS between the surface and 800 hPa. This is because the morphology of the AKM changes as reducing vertical levels bins off-diagonal elements into the new diagonal components on the coarser grid. The iterative method grid keeps the greatest DFS along the profile compared to the other two methods, because it accounts for off-diagonal responses during

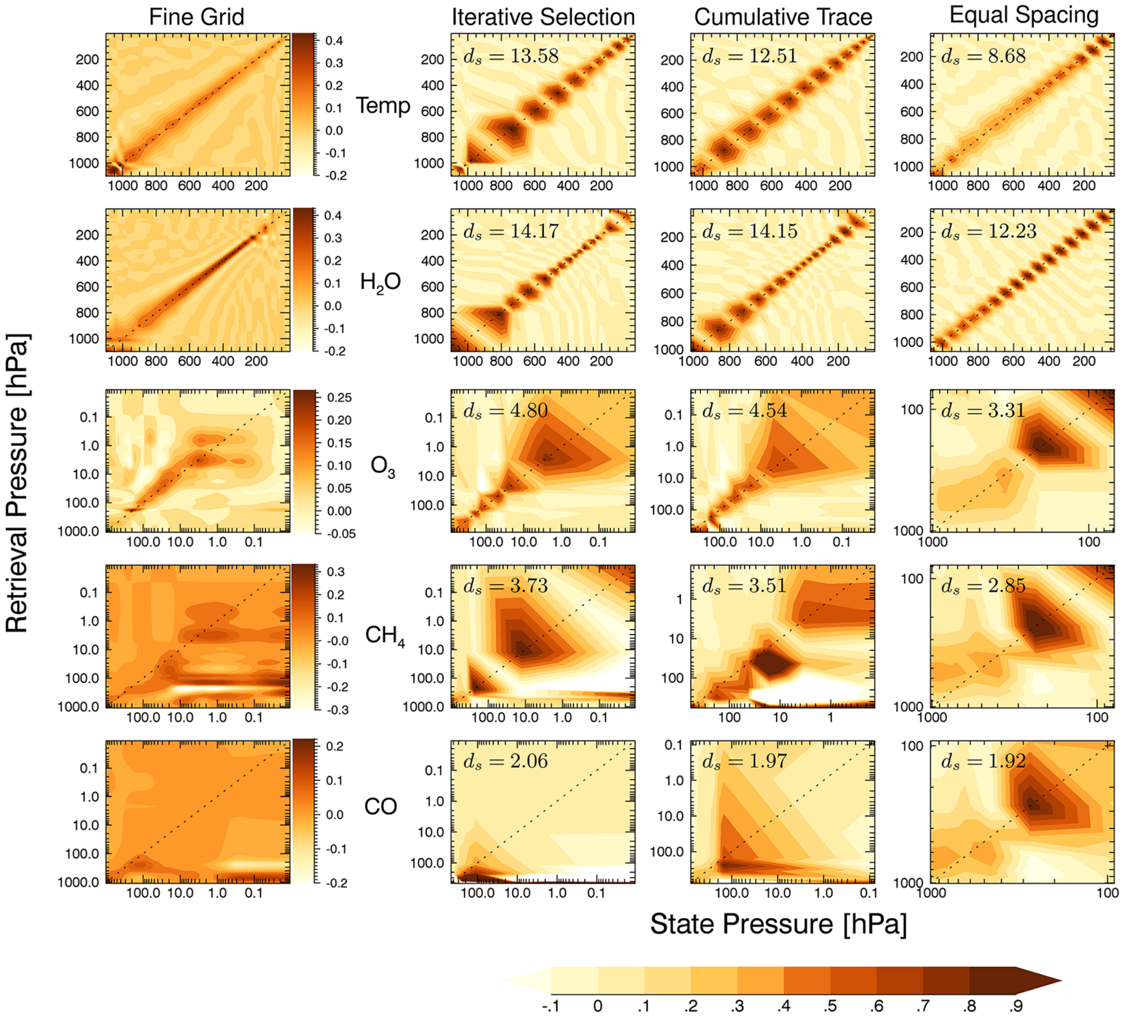


Figure 4.6: Contours of AKMs for a mid-latitude summer retrieval of temperature, H_2O, O_3, CH_4, and CO using IASI. Fine grid refers to AKMs evaluated on the 101 level vertical grid without interpolation from a coarser grid. The right three columns show AKMs on coarser vertical grids selected according to the methods discussed in Sects. 4.2.1–4.2.3. All three coarse grids have the same number of levels for each retrieval, chosen to be four more than the rounded DFS on the fine grid. DFS values specific to each coarse grid AKM are annotated in the top left corners. Notice that the pressure axes are specific to the AKM shown, especially the fourth column.

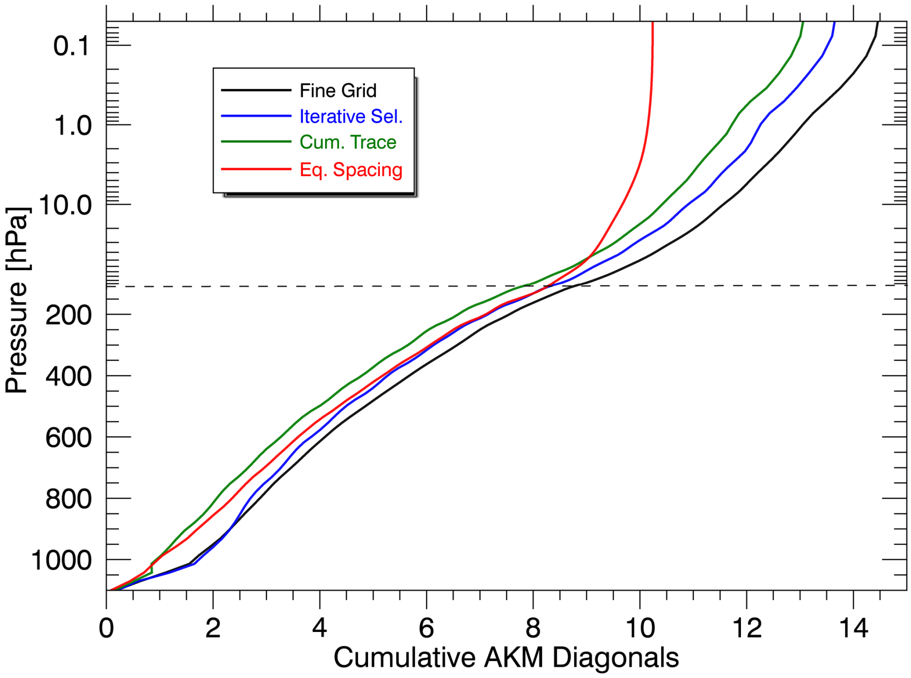


Figure 4.7: Cumulative diagonals of the AKM for a temperature retrieval on the 101 level fine grid as well as 18 level coarse grid AKMs (Sects. 4.2.1–4.2.3) interpolated to the fine grid.

the selection process. The next section discusses if the differences are significant to justify the more complex iterative selection method.

4.3.4 Ensemble comparisons

Rather than draw conclusions from a single atmospheric scenario, an ensemble of atmospheres were used to compare the different vertical grid selection methods. Comparisons between methods were made using the same number of vertical levels, but with different spacings, specific to each atmosphere to isolate differences solely due to the method. After analysis, the selection method that retained the greatest DFS on average for the 80 atmospheres was the iterative method from Sect. 4.2.3.

Figure 4.8 shows histograms of the loss in DFS by using an equally spaced vertical pressure grid as compared to the iterative method. For the majority of atmospheres the difference in DFS for CO and H_2O is less than 10 %. In fact, there are 10 atmospheres for a CO retrieval in which the equal spacing grid provided more DFS than the iterative grid, which highlights the fact that the iterative method

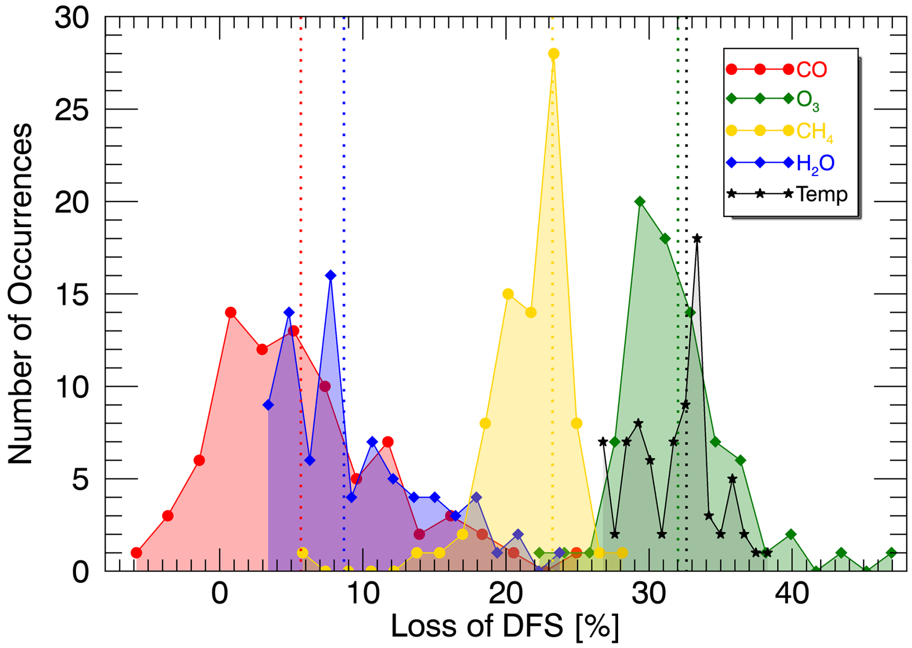


Figure 4.8: Histograms of 15 equally spaced bins showing the percent loss in DFS by using a coarse vertical grid with equal pressure spacing versus the coarse grid from the iterative method. The dotted vertical lines show the median values of the 80 atmosphere distribution.

does not guarantee finding the global maximum of DFS. However, the majority of atmospheres with equal spacing reduce the DFS by over 20 % for CH_4 and over 30 % for O_3 and temperature. This significant loss in DFS results from ignoring stratospheric information content. While using equal spacings in pressure seems like an obviously poor idea for retrieving O_3, there is also significant stratospheric information for CH_4 and temperature that should be considered.

The loss of DFS by using a vertical grid selected from the cumulative diagonal of **A** compared to the iterative method is shown in Fig. 4.9. For the majority of temperature retrievals the loss in DFS is less than 10 %, while the loss for trace gases is less than 5 %. This shows that the benefit from considering off-diagonal sensitivities in the selection process is slight, especially for the trace gases considered.

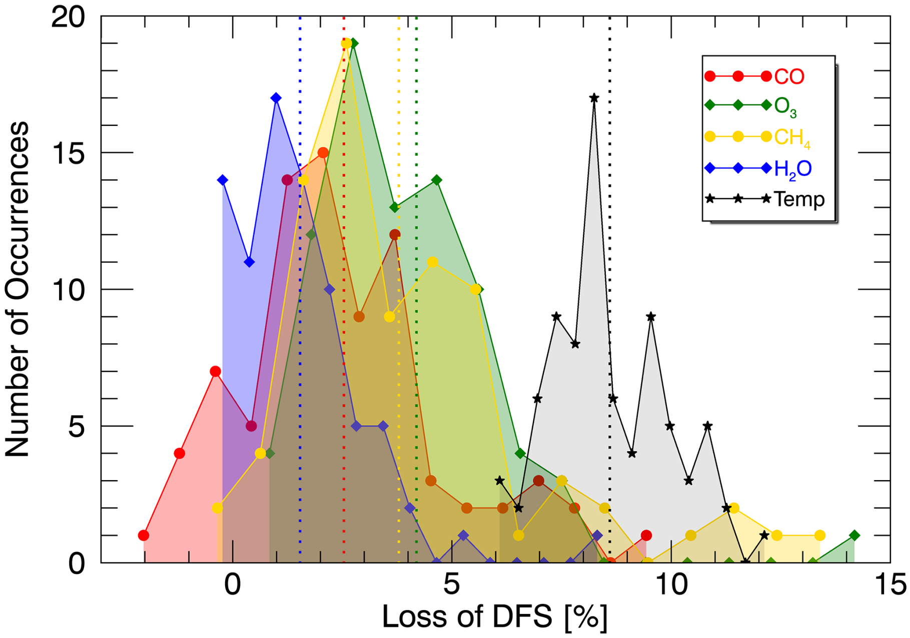


Figure 4.9: Histograms of 15 equally spaced bins showing the percent loss in DFS by creating a coarse vertical grid from the cumulative trace of **A** versus the coarse grid from the iterative method. The dotted vertical lines show the median values of the 80 atmosphere distribution.

4.4 Globally designed grid

Finally, when constructing a retrieval algorithm with IASI or another infrared sounding instrument, the developer may consider whether to use a constant vertical grid or one that optimally adapts to the observed atmosphere based on the prior state and converging profile. Figure 4.10 are histograms showing the loss of DFS when using a constant vertical grid optimized globally versus an atmosphere specific grid. The global vertical grid was constructed following the iterative method, but considering an average DFS from a tropical, mid-latitude, polar summer, and polar winter atmosphere as the metric of quality.

Notice that the median DFS losses are less than 3 % for O_3, CH_4, and CO. However, there are long tails extending past 10 % for the more extreme atmospheres. Additionally, temperature and H_2O show DFS loss values between 10 and 20 %, suggesting that the vertical location of information varies more significantly than the other gases considered. Therefore, an adaptive atmosphere specific vertical

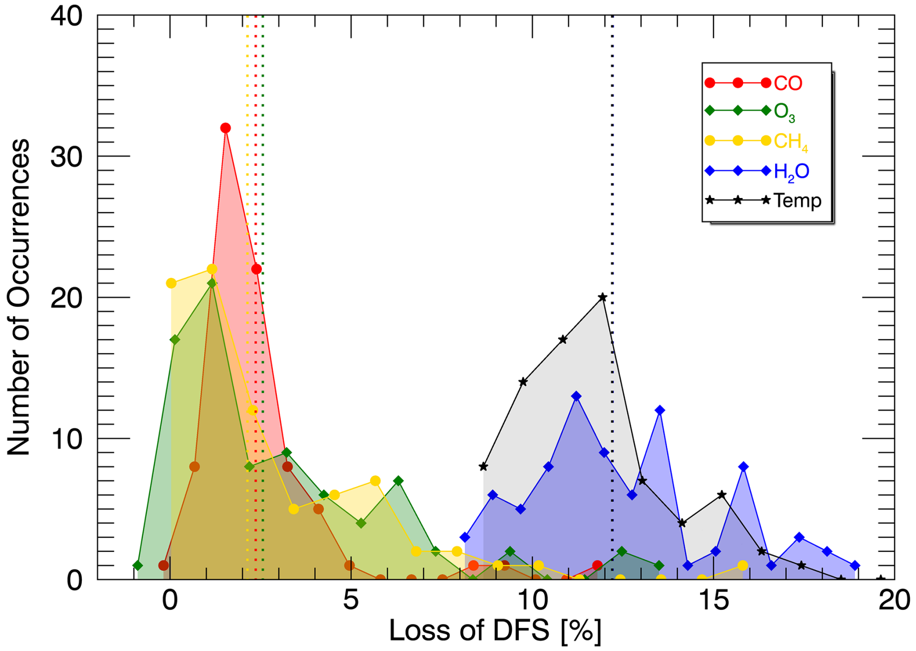


Figure 4.10: Histograms of 15 equally spaced bins showing the loss in DFS by using a constant globally optimized vertical grid versus an atmosphere specific grid. The dotted vertical lines show median values of the 80 atmosphere distribution, where the temperature and H_2O medians overlay each other.

grid may improve retrievals of temperature and water vapour from an information perspective. However, the practicalities of implementing an adaptive grid may make this increase in DFS an undesirable trade-off when attempting to produce time averaged or cross-platform analyses.

To summarize Figs. 4.8–4.10, the ensemble median values are shown in Table 4.1 along with the mean DFS values for the referenced iterative selection method.

Table 4.1: Summaries of the ensemble results from Figs. 4.8–4.10. The second column shows the ensemble mean DFS for each retrieval using the iterative method. The last three columns show the median percent loss of DFS for the histograms displayed.

Retrieval	$\overline{d_s}$	Eq PRE [$\%_{\mathrm{loss}}$]	Diag($\mathbf{A}$) [$\%_{\mathrm{loss}}$]	Global [$\%_{\mathrm{loss}}$]
Temperature	15.7	32.6	8.6	12.1
H_2O	13.8	8.7	1.5	12.1
O_3	5.1	32.0	4.2	2.5
CH_4	3.8	23.1	3.8	2.1
CO	2.9	5.8	2.5	2.3

4.5 Summary

When retrieving atmospheric profiles of temperature and trace gases from infrared spectral radiances, it is important to consider where in the vertical profile the estimates are made. A new iterative method for selecting a vertical grid was proposed and shown to outperform previously used selection methods by accounting for correlations and sensitivities between different vertical levels. Other compared methods of establishing a vertical grid coarser than the radiative transfer grid were using levels equally spaced in pressure and selecting levels by interpolating along the cumulative diagonal of the fine grid AKM.

The 80 atmosphere ensemble created to parametrise RTTOV was used to systematically compare the different vertical grid selection methods for temperature, H_2O, O_3, CH_4, and CO. For the majority of atmospheres, using a vertical grid with equal pressure spacings resulted in a 20–40 % loss of DFS for temperature, O_3, and CH_4. Median DFS losses for H_2O and CO were less than 10 %. In general, this shows that a significant proportion of DFS can be retained by considering the vertical location of the information content as opposed to choosing a vertical grid based on convenience.

Comparing to the cumulative diagonal of $\mathbf{A}$ method shows that greater DFS can be achieved with the iterative method, but median losses are less than 5 % for the trace gases and less than 10 % for temperature retrievals. This slight reduction in DFS is unlikely to affect the quality of retrievals in a noticeable way for the majority of atmospheric cases. Therefore, the simpler and less expensive method of interpolating along the cumulative diagonal of $\mathbf{A}$ will likely be sufficient, except for possibly extreme atmospheric scenarios with high inter-level correlations.

Finally, much effort is spent making retrieval schemes run faster. Naturally, one would prefer to design a coarse vertical grid just once and apply that to all scenarios, rather than optimize a grid for each retrieved atmosphere. This depends upon the variability of vertical information content. The median loss of DFS for O_3, CH_4, and CO when using a globally optimized grid versus an atmosphere specific grid was less than 3 %. For the majority of atmospheres, using a static grid results in a negligible retrieval difference for these three gases. However, the loss of DFS for temperature and H_2O is more appreciable, greater than $10\,\%$, implying the location of vertical information is more variable than the other trace gases considered. Whether or not to account for information variability with temperature and H_2O ultimately depends upon the motivation for computation speed and tolerance of the DFS reductions presented.

There is more to life than simply increasing its speed.

— Mahatma Gandhi

5

Fast monochromatic radiative transfer

Contents

5.1 Background

The retrieval of atmospheric temperature and composition from infrared measurements is usually an iterative process requiring repeated calculations of the radiative transfer equation to simulate the observations. In its simplest form the radiance L reaching the satellite at wavenumber ν from viewing geometry (e.g., tangent height, scan angle) z can be expressed as the sum of atmospheric and background terms:

$$L(\nu, z) = \int_{\tau_0}^{1} B(\tau)\ \mathrm{d}\tau + B_0\tau_0, \tag{5.1}$$

where τ is the transmittance along the line-of-sight from the satellite ($\tau = 1$) to the remote boundary of the atmosphere ($\tau = \tau_0$), B is the Planck function along this path, and B_0 is the Planck function of the background at the far side of or beyond the atmospheric path.

The measurement itself, R_{ij}, for nominal wavenumber ν_i and viewing geometry z_j, is modelled as a convolution of this radiance with the appropriate instrument functions

$$R_{ij} = \int\int L(\nu, z)\Phi(z - z_j)\Psi(\nu - \nu_i)\,\mathrm{d}\nu\mathrm{d}z. \tag{5.2}$$

Φ_j represents the instrument field-of-view (omitted for nadir-viewing with IASI) and Ψ_i represents the spectral response (the instrument line shape for Fourier transform spectrometers, or channel response for radiometers).

The three main geometries for infrared remote sensing can be distinguished by the relative importance of the background term in the radiative transfer equation; for solar occultation $B_0\tau_0$ is dominant, for limb-viewing it is negligible, and for nadir-viewing it is comparable to the atmospheric term. However, in all cases, since B is only a function of wavenumber and temperature, information on the composition and pressure (as well as temperature) is derived from the transmittance.

Calculating transmittance is usually the time-consuming part of the retrieval process. The most accurate method is to use a line-by-line (LBL) model. However, this is slow and only feasible in real-time data processing for solar occultation measurements, which typically have a maximum of 28 profile acquisitions per day.

Nadir-viewing instruments, such as IASI, provide measurements at typically 10^5–10^6 locations per day and this presents a greater challenge. With spectrally broad measurements from filter radiometers, the usual method is to apply the radiative transfer to pre-computed spectrally integrated quantities, effectively:

$$R_i = \int \bar{B}\,\mathrm{d}\bar{\tau} + \bar{B}_0\bar{\tau}_0 \tag{5.3}$$

where

$$\bar{\tau} = \int \tau(\nu)\Psi(\nu - \nu_i)\,\mathrm{d}\nu \tag{5.4}$$

and similarly for $\bar{B}$ [*Goody et al.*, 1989; *Lacis and Oinas*, 1991]. This differs fundamentally from the monochromatic approach in that Beer's Law does not hold for spectrally integrated transmittances. In other words, the net transmittance of a composite path is no longer simply the product of the transmittances of the component paths when dealing with averaged quantities. This can be readily appreciated if one considers the net transmittance through two consecutive identical cells, each opaque over the lower half of the spectral range and transparent in the upper half. The spectrally averaged transmittance of each cell is 0.5, but the transmittance of the two cells combined is, clearly, also 0.5, rather than the product $0.5 \times 0.5 = 0.25$. To overcome this it is necessary to parametrise spectrally averaged transmittances not only as a function of the 'natural' quantities such as pressure and temperature, but also a number of ad-hoc 'predictors' such as geometry and contaminants, relating to each channel and generally obtained by statistical regression from a large sample (with all the risks that entails when dealing with anomalous atmospheric cases) [*Moncet et al.*, 2008].

The recent trend has been towards replacing broad band filter radiometers, such as the High-resolution Infrared Radiation Sounder (HIRS; order of 10 channels of widths $\sim$ 10's of cm^{-1}), with Fourier transform or grating spectrometers (order of 1000 channels with resolution $\sim$1 cm^{-1}). Extending the spectrally averaged approach requires treating each spectral sample as an independent channel with its own set of tailored predictors. While this remains the 'fast' solution, it becomes unwieldy and, noting the steady increase in computing speeds, it is anticipated that a monochromatic approach will gradually be adopted even for the operational nadir-viewing instruments.

Significant time savings can be made by using the line-by-line model to precompute monochromatic transmittances $\tau(\nu)$ (or, more usually, a related quantity such as absorption cross section), over a variety of atmospheric conditions, and storing these in a 'look-up table' (LUT) for use in the retrieval forward model. Such an approach has, for example, been used for the Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) [*Fischer et al.*, 2008] limb-sounding instrument and IASI. The use of monochromatic LUTs of absorption cross sections is not new [*Scott and Chedin*, 1981]. However, in many cases the LUTs themselves are closely tied to their application and details of their generation, testing, and optimisation are not widely published.

The aim of this chapter is to review the criteria required to construct such LUTs and the various methods chosen for their implementation. Simulations of IASI radiances are used as a test case. Furthermore, a detailed description is given of the construction and testing of LUTs which can be implemented in the Reference Forward Model (RFM) for radiative transfer [*Dudhia*, 2008] as a direct replacement for the line-by-line calculation, which therefore allows comparisons of the speed/accuracy trade-off.

Finally, given that these LUTs require a relatively simple file format (even if the files are large) it is hoped that such LUTs may be regarded as independent databases, decoupled from their originating line-by-line models and any particular instrumental application. For the work presented in this thesis, these LUTs were crucial in the timely development and analysis of the OCS retrieval method presented in the next chapter.

5.2 Molecular spectroscopy primer

Atmospheric retrievals exploit the physical knowledge of how light propagates through the atmosphere in order to estimate the state that produced an observation. This section briefly touches upon the underlying physics of molecular spectroscopy for context. While discussions here are qualitatively summarized, further in-depth

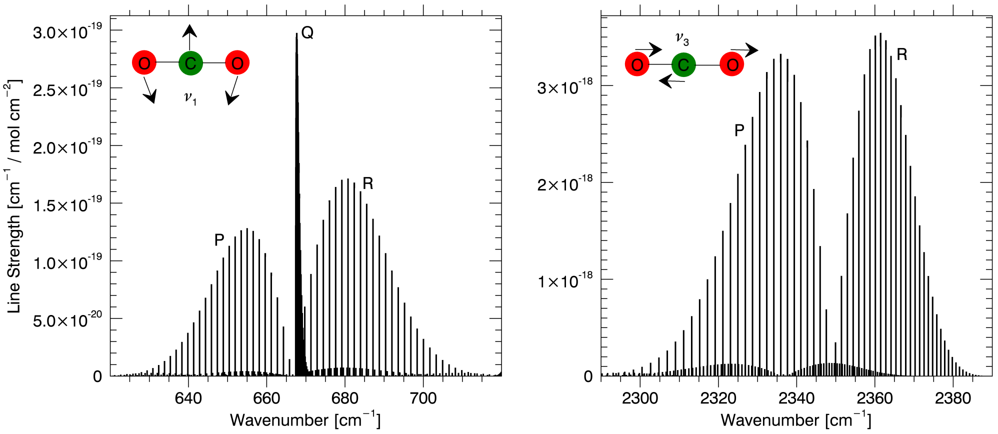


Figure 5.1: Rotational-vibrational spectral line strengths are shown for the fundamental vibrational modes of CO_2. These data were taken from the HITRAN 2012 database [*Rothman et al.*, 2013].

reading can be found in *Hollas* [2004] with full quantum mechanical treatment in *Bransden and Joachin* [2003].

In atmospheric remote sensing, "spectral lines" are fundamentally the observation of photons that are either absorbed or emitted from quantum transitions of atoms or molecules in the atmosphere. Schrödinger's wave equation revealed that, because of the wave-nature of all particles, energy levels of bound states and angular momentum are discretely quantised through constructive and destructive interference of the particles with themselves. Therefore, every molecule absorbs and emits radiation in a unique fashion specific to its configuration, where the spectra from each molecular type acts like an identifying fingerprint.

Consider CO_2 as an example. All gas particles translate through the air, but since this motion is unbound on small scales, kinetic energy is continuously distributed. However, molecules also rotate and vibrate in quantised states, which produce the spectral bands observed in the thermal infrared. Electronic transitions of orbital levels occur at energies and wavenumbers much higher than the range observed by IASI. It is the combination of rotational and vibrational (rotational-vibrational) transitions that produce the CO_2 spectral lines shown in Fig. 5.1.

Vibration between the atoms of a molecule tends to require greater energy than rotation. Therefore, pure rotational spectral features in absence of vibration normally interact with radiation in the microwave region. In the thermal infrared, the quantum vibrational transition determines the centre location of a spectral band. Figure 5.1 shows the spectral line strengths of the primary vibrational modes as recorded in HITRAN 2012 for the most abundant isotope of CO_2. The left panel shows the lowest energy primary vibrational mode (ν_1) where the atoms bend in and out of linear alignment. The word 'primary' here refers to transitions between the first excited state and the ground state. The right panel shows the highest primary vibrational mode (ν_3) for CO_2 where the bonds between carbon and oxygen asymmetrically stretch. The symmetric stretching mode (ν_2) does not interact with radiation by electric dipole moments (first order) and, therefore, has no associated spectral feature.

The individual lines shown in Fig. 5.1 for each vibrational band are separate quantum transitions between rotational levels. Spectral lines at wavenumbers left of centre make up what is referred to as the P-branch. These are transitions where the rotational quantum number of the excited vibrational state is less than the rotational quantum number of the vibrational ground state, while the R-branch is vice-versa for the greater energy difference. The Q-branch represents the vibrational transition where the rotational state remains the same. Notice that there is no Q-branch in the ν_3 vibrational mode, because these specific transitions are not allowed, i.e., have zero probability of occurring.

The magnitude of the lines in Fig. 5.1 are known as the line strengths, which are required knowledge in order to model atmospheric transmission. Atmospheric line strengths are complicated to represent, because they depend upon the propensity of a molecule to interact with radiation, the probability of a quantum energy transition, the redundancy of that transition (degeneracy), and the population of the varying energy states. In thermal equilibrium the relative line strength proportions follow a Boltzmann distribution, which is temperature dependent. Therefore, line strengths for each gaseous species are generally referenced from a pre-tabulated database, such as HITRAN, which were generated at typical terrestrial temperatures. As a result, temperatures much greater than around 300 K require referencing a separate line strength database.

Observations of spectral features clearly show that these lines are not delta-functions, but rather they show a broader distribution about line centre. This smoother shape is referred to as the line shape function. In general, there are three types of physical broadening affecting the line shape. First, emission or absorption of a photon can be interrupted by collisions between molecules that temporarily shift the quantum energy levels. This is known as pressure or collision broadening. The basic approximation of this effect assumes instantaneous elastic collisions and is given by the Lorentzian function:

$$f_L\left(\nu-\nu_0\right)=\frac{1}{\pi}\frac{\alpha_L}{\left(\left(\nu-\nu_0\right)^2+\alpha_L^2\right)}, \tag{5.5}$$

where ν_0 is the centre wavenumber and α_L is the half-width of the line shape and is linearly proportional to pressure and inversely proportional to the square root of temperature. In reality, molecules interact with each other during collisions such that no collision is truly elastic. However, corrections to Eq. (5.5) tend to be empirical in nature as analytic line shapes of non-ideal gases are largely unsolvable outside of numerical approximations.

Secondly, line shapes can be broadened according to Doppler shifts of the emitted radiation due to the random kinetic motion of particles. If the molecules are in thermodynamic equilibrium, then the kinetic energies obey Maxwell-Boltzmann statistics such that the line shape function is a Gaussian with regards to wavenumber:

$$f_D\left(\nu-\nu_0\right)=\frac{1}{\alpha_D\pi^{1/2}}\exp\left[-\left(\frac{\nu-\nu_0}{\alpha_D}\right)^2\right], \tag{5.6}$$

where α_D is the $1/e$ width and is proportional to the square root of temperature while independent of pressure.

Finally, spectral lines are naturally broadened by the fundamentals of the Heisenberg uncertainty principle. In other words, the finite lifetime of an excited energy level cannot be known exactly and this uncertainty translates into a spread of wavenumber values about line centre. Natural line broadening also results in a

Lorentzian line shape function. However, the width of this line shape is roughly 10^3 times smaller than pressure and Doppler broadening for Earth atmospheres and is generally ignored.

Pressure broadening tends to dominate in the lower troposphere while Doppler broadening dominates in the upper stratosphere. However, both collisional and Doppler effects contribute to the line shape functions throughout the majority of the atmosphere and must be accounted for jointly. The Voigt line shape is the convolution of collisional and Doppler broadening and must be computed for accurate modelling of atmospheric transmission. Of all the numerical steps in radiative transfer, calculating the Voigt profile for each spectral line is normally the dominating expense in LBL procedures.

5.3 Overview of nadir-viewing radiative transfer

Work presented in this chapter was performed using the RFM to numerically simulate atmospheric radiative transfer in the thermal infrared and to create the LUTs mentioned. The RFM is a line-by-line model originally written to simulate MIPAS observations. Based upon the GENLN2 model [*Edwards*, 1992], the RFM has since been adapted to numerous viewing geometries, including nadir, and is now a general purpose RTM intended to be flexible, accurate, and easy to use. One possible use of the RFM is to simulate IASI observed spectral radiances. Configuration steps necessary to simulate IASI with the RFM are documented in *Ventress* [2014, ch. 5].

5.3.1 The line-by-line approach

The aim of any RTM is to numerically solve the equation of radiative transfer for spectral radiance at each wavenumber increment. Equation (5.1) is expanded for IASI simulations to yield

$$\begin{aligned} L(\nu) = &\epsilon_s B(T_s)\, \tau_{s\to obs} + \int_s^{obs} B(T_z)\, \frac{\partial \tau_{z\to obs}}{\partial z} dz \\ &+ (\epsilon_s - 1)\, \tau_{s\to obs}^2 \int_s^{obs} B(T_z)\, \frac{\partial \tau_{z\to obs}^{-1}}{\partial z} dz, \end{aligned} \tag{5.7}$$

where we consider terrestrial thermal emission as observed through a clear non-scattering sky by nadir looking instruments. The first term accounts for emitted surface radiation reaching the observer, where ϵ_s is surface emissivity, $B\left(T_s\right)$ represents the Planck function for blackbody radiation at surface temperature T_s, and $\tau_{s \to obs}$ is the atmospheric transmittance from the surface to the observer. The dependence upon wavenumber for ϵ, B, and τ is implicitly assumed for brevity. The second term represents upwelling atmospheric radiation where z is a vertical height coordinate. Finally, the third term accounts for downwelling atmospheric radiation reflected upwards towards the observer assuming specular reflection. This is written in terms of upwelling transmittance for ease of integration by making use of the fact that $\tau_{s \to z} = \tau_{s \to obs} / \tau_{z \to obs}$ [*Amato et al.*, 2002, Eq. (7)]. Note that reflected downwelling radiation is second order in transmittance and generally only noticeable when observing in window channels over lower emittance geographies (e.g., the Sahara at $\approx 900\,\mathrm{cm}^{-1}$). Solar radiation reflected from the surface to the observer is only important at wavenumbers greater than $\approx 2000\,\mathrm{cm}^{-1}$ and currently not included in the RFM.

As discussed in *Scott and Chedin* [1981], most of the computation time spent solving Eq. (5.7) is consumed by calculating transmittance. Atmospheric transmittance for a vertically inhomogeneous atmosphere from the surface to a height z is expressed as

$$\begin{aligned} \tau_{s \to z} &= \exp\left(-\sum_g \int_s^z k_g\left(\nu, z'\right) \rho_g\left(z'\right) dz'\right) \\ &= \exp\left(-\chi\left(\nu\right)\right), \end{aligned} \tag{5.8}$$

where the summation is taken over all absorbing gases considered and ρ_g is the density of the g^{th} gas. The negative argument of the exponent (χ) is referred to as the total optical depth. Numerically, the integrals in Eqs. (5.7) and (5.8) are approximated by considering layers of finite vertical thickness, known as the "layer-by-layer" approximation. Figure 5.2 shows the contribution to optical depth for a water vapour spectral line in a tropical atmospheric scenario for each 100 hPa tropospheric layer. Notice that the optical depth shape changes depending upon

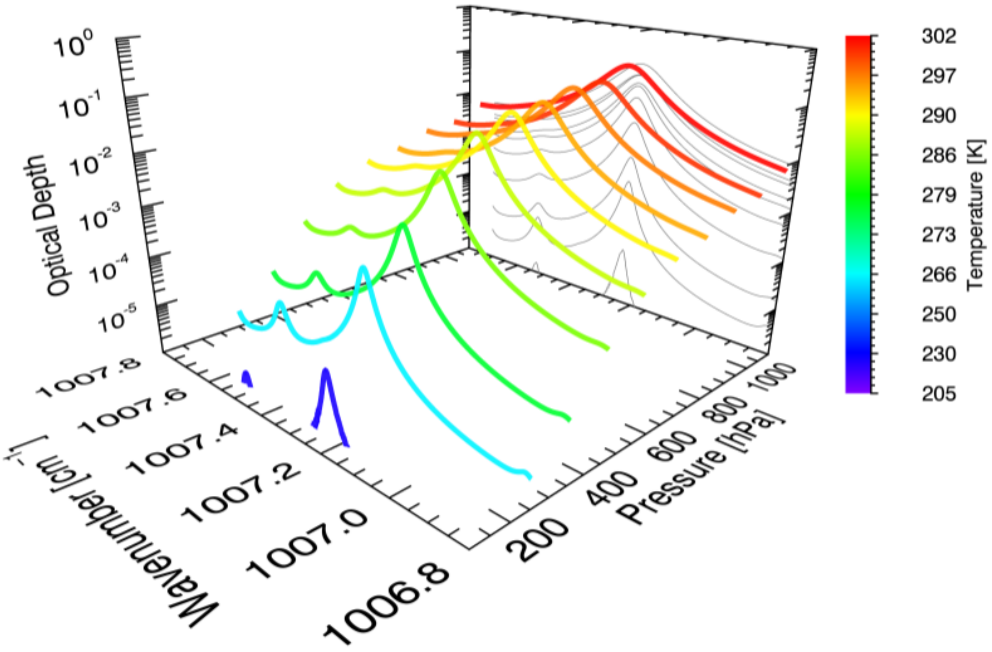


Figure 5.2: Optical depths of a water vapour spectral line are shown for tropospheric layers of 100 hPa for a typical tropical atmosphere. Each optical depth curve is colour coded to the approximate atmospheric temperature of that layer. These optical depths were calculated with the RFM using the 2012 HITRAN database. All spectroscopic fields listed in HITRAN were used, including empirically tested self-broadening effects.

the physical atmospheric conditions of that layer. The total optical depth is then the sum of the optical depths of each layer.

Within Eq. (5.8), determining k, the absorption cross section (units of $\mathrm{m}^2\,\mathrm{molec}^{-1}$), is the dominating expense. Referring to a model as 'line-by-line' means that the absorption coefficient includes contributions from each individual spectral line (i) within a given range;

$$k\left(\nu, p, T\right)_g = \sum_i S\left(T\right)_{ig} f\left(\nu - \nu_i, p, T\right)_g , \tag{5.9}$$

where S_{ig} is the strength of line i for gas g and $f\left(\nu - \nu_i\right)_g$ is the line shape function [*Galatry*, 1961]. Equation (5.9) explicitly shows the dependence upon pressure (p) and temperature (T), while there is also a noticeable dependence upon partial pressure (e) for H_2O through self-broadening effects [*Tobin et al.*, 1996; *Tipping and Ma*, 1995]. Line positions and strengths are referenced from a spectroscopic database, such as the HITRAN database [*Rothman et al.*, 2013]. The process of calculating optical depths is quite laborious because it involves analysing every single spectral line, for each atmospheric layer, and for all gases considered. However, since these calculations are so repetitive, much of the heavy burden can be pre-calculated in a LUT and referenced when analysing Eq. (5.8).

This chapter describes only the aspects of radiative transfer necessary to understand how LUTs are employed to speed up calculations of absorption or transmittance. A more complete discussion of numerically modelling radiative transfer is found in *Amato et al.* [2002] and *Hurtmans et al.* [2012] with encyclopedic theoretical details found in *Goody and Yung* [1989].

5.3.2 The Curtis-Godson approximation for finite vertical layers

The absorption cross section itself, as well as its spectral dependence, is also a function of pressure, temperature and, sometimes, absorber concentration if the absorbing molecule has a permanent electric dipole moment. Typically k would be evaluated for absorber weighted mean values of these quantities within the segment, i.e., approximating the optical depth of the inhomogeneous path segment as being the same as an equivalent homogeneous path characterised by these mean quantities. The Curtis-Godson (CG) approximation [*Curtis*, 1952; *Godson*, 1953; *Goody and Yung*, 1989] is commonly used to set an equivalent pressure, $\overline{p}_g$, for gas g as

$$\overline{p}_g = \frac{\int_z^{z+\Delta z} p\,(z)\,\rho_g\,(z)\,dz}{\int_z^{z+\Delta z} \rho_g\,(z)\,dz} \tag{5.10}$$

and an equivalent temperature as

$$\overline{T}_g = \frac{\int_z^{z+\Delta z} T\,(z)\,\rho_g\,(z)\,dz}{\int_z^{z+\Delta z} \rho_g\,(z)\,dz}. \tag{5.11}$$

Each absorbing gas has layer values distinct from those defined by the atmospheric levels. This potentially complicates the process of setting LUT grids, as each gas has its own CG pressure and temperature grid.

Figure 5.3 shows the RFM results of the CG approximation using the same water vapour line and tropical atmosphere from Fig. 5.2. The top plot shows that using just one homogeneous layer with the CG pressure and temperature yields a surprisingly accurate optical depth spectral shape compared to a 100 layer inhomogeneous atmosphere. The bottom plot residuals also show the effect that empirical H_2O pressure shifts and self-broadening [*Tobin et al.*, 1996; *Tipping and Ma*, 1995; *Koshelev et al.*, 2007] have upon the accuracy of the CG approximation. Interestingly, including self-broadening from the High Resolution Transmission (HITRAN) database [*Rothman et al.*, 2013] in the optical depth calculation leads to less residual error for the one layer example. Caution is needed, however, to avoid concluding that since the optical depth residual error is $\approx 3\%$ that only one or a few layers are necessary. The spectral line considered in Fig. 5.3 is largely Lorentzian. For lineshapes falling in the Voigt regime, the resulting radiance error is often significant compared to instrument noise for anything less than at least 30 layers [*Ventress*, 2014, ch. 5].

Finally, the difference between line-by-line and LUT calculations is in the evaluation of k: For a line-by-line model, k is evaluated as the sum of contributions of all local transitions appropriately adjusted for the CG mean values; while for a LUT model, k is interpolated from pre-tabulated axis values.

5.4 Previous approaches

In principle, precomputing LUTs for a particular molecule is straightforward: Use a line-by-line model to calculate $k(\nu, p, T, e)$ at a sufficient range and density of axis points that atmospheric radiance calculations with interpolated LUTs are of acceptable accuracy compared to the equivalent uninterpolated line-by-line calculation. However, there are two immediate practical problems.

First, in order to meet the ‘monochromatic’ criterion that atmospheric spectral features can be adequately resolved (set by the Doppler width of atmospheric lines), a wavenumber spacing of 0.001 cm^{-1} is generally required for nadir-viewing and 0.0005 cm^{-1} for limb-viewing. Assuming of the order of a 1000 points to represent the p, T, e domains, any LUT with a uniform spectral axis will require 10^6 real numbers just to represent the absorption cross section of a single absorber in a 1 cm^{-1} interval.

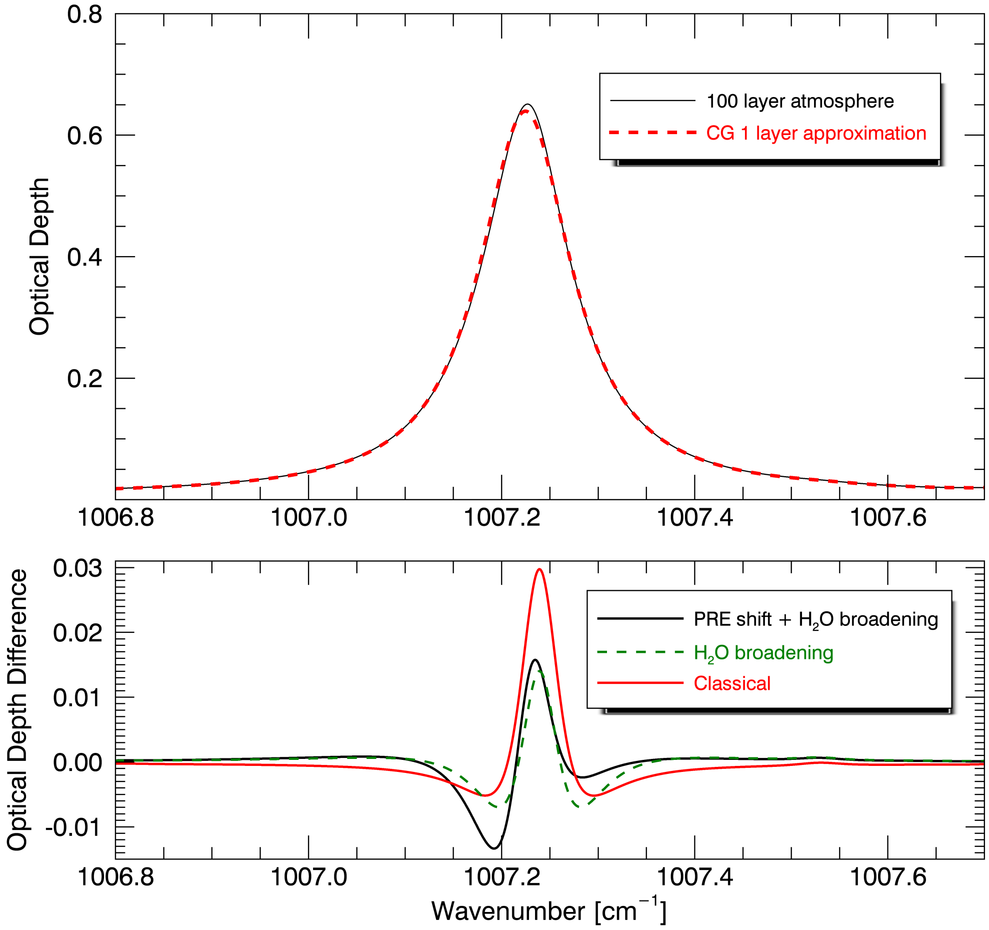


Figure 5.3: Top: Optical depth for a H_2O spectral line from a tropical atmosphere, the sum of layer contributions from Fig. 5.2. The solid curve shows the optical depth for a 100 layer inhomogeneous atmosphere and the dashed line shows optical depth for an equivalent one layer homogeneous atmosphere using the CG approximation. Bottom: The difference between the top curves (residual) is shown for three separately modelled conditions, which are classical ideal gas pressure broadening, H_2O self-broadening, and self-broadening including the H_2O collisional pressure shift. The reference spectra for the green and black curves are not plotted in the top diagram as they are indistinguishable by eye.

The second problem is defining the required accuracy. It is quite simple to define axis spacing such that k can be reconstructed with a given relative or absolute accuracy. The problem is that the monochromatic radiance reaching the satellite is not linearly related to k, and that is without even considering the effects of other absorbers or spectral convolution.

Scott and Chedin [1981] were among the first to bypass LBL processing time by creating a set of LUTs known as the Automatized Atmospheric Absorption Atlas for use with HIRS and the Infrared Interferometer Spectrometer (IRIS). Transmittance values were predetermined over a range of plausible pressure, temperature, and H_2O partial pressure scenarios and stored in LUTs at each spectral increment. By interpolating transmittance from the LUTs for the various atmospheric conditions, radiance calculation speeds were increased by a factor of 15 to 40. Parametrisation error was found to be negligible for the desired applications when compared to instrument noise.

The base steps to create a faster monochromatic RTM based on LUTs are summarised in the paper by *Amato et al.* [2002] describing the σ-IASI RTM designed specifically for atmospheric retrievals using IASI. In this model, individual optical depths for H_2O, CO_2, O_3, N_2O, CH_4, N_2, and O_2 were tabulated from the Line-By-Line Radiative Transfer Model (LBLRTM) [*Clough et al.*, 2005] and thinned in size using methods to be discussed in Sect. 5.6.4. Grid thinning reduced the final LUTs from 27.7 Gbytes to 726 Mbytes. However, self-broadening effects of H_2O line shapes due to variable partial pressures were not included, resulting in radiance error approaching IASI noise levels in spectral regions associated with H_2O absorption for the six analysed atmospheric scenarios.

A similar process was repeated when creating the forward model for the Fast Optimal Retrieval on Layers for IASI (FORLI) [*Hurtmans et al.*, 2012], which is based upon previously defined LUTs and includes a partial pressure dimension for H_2O. As of publication, there were three versions of this scheme specific to retrievals of O_3, HNO_3, and CO. Since the spectral ranges considered in FORLI were small ($< 150\,\mathrm{cm}^{-1}$), spectral compression of the LUTs was not needed. Absorption cross-sections at a given pressure and temperature interpolated from the LUTs were claimed to match the true values to less than 0.1 % error. However, validation in radiance space was not mentioned, in particular for comparing radiances over a host of atmospheric conditions to those produced by a full LBL model. Validation of radiances generated from LUTs is a crucial step, because forward model errors inevitably become artefacts of estimated quantities.

Optical depth LUTs were also created for retrievals using the Atmospheric Infrared Sounder (AIRS) in much the same fashion, but *Strow et al.* [1998] implemented a singular value decomposition (SVD) scheme to compress the ν dimension and make the LUT size more manageable. They showed that optical depth interpolation can occur in the compressed space, which sped up transmittance calculations by a couple orders of magnitude while keeping interpolation errors below a desired threshold in radiance. This compression method is further discussed in Sect. 5.6.4.

Finally, specific LUTs for MIPAS were generated for each microwindow used in the retrievals [*Dudhia et al.*, 2002]. The procedure was to start with a high-density tabulation of $k(\nu, p, T)$ for each absorbing species (the maximum microwindow width of 3 cm^{-1} makes this feasible). These LUTs were then progressively sub-sampled in the p, T domain, compressed using SVD and finally sub-sampled in the ν domain, at each stage comparing the simulated radiances with those of a line-by-line model to ensure the error budget remained within 10% of the instrument noise.

5.5 Alternative radiative transfer methods

As atmospheric spectral measurements from ground, aircraft, and satellite systems have dramatically increased over the years, so too has the demand for faster RTMs to assist in analysing the observed data. The current trend is that observations may continue to increase in number, at finer spectral resolution, and with decreasing instrument noise. Therefore, future RTMs will certainly need to be fast, highly accurate in comparison to LBL methods, and robust over a host of atmospheric conditions. Calculating radiative transfer with monochromatic LUTs maintains the desired robustness of LBL methods by carrying out the full integration in Eqs. (5.7) and (5.8) for each absorber individually. However, additional gains in speed can be made by statistically parametrizing portions of radiative transfer, but with reduced accuracy. With this in mind, three popular schemes that use statistical parametrisation are briefly discussed for context.

5.5.1 Total path transmittance regression

Spectroscopic line parameters can be bypassed all together in the method referred to as total path transmittance regression (TPTR). In this method an "effective" optical depth is statistically parametrised according to pressure, temperature, and gas concentration for each atmospheric layer (l) as given by

$$\log\left(\frac{\overline{\tau}_l}{\overline{\tau}_{l-1}}\right) - \log\left(\frac{\overline{\tau}'_l}{\overline{\tau}'_{l-1}}\right) \approx \sum_n a_{l,n} X_{l,n}, \tag{5.12}$$

where $\overline{\tau}$ is the mean transmittance over a spectral interval and the prime symbol denotes an exact value as determined from a more accurate model, such as LBL. The goal is to empirically choose a set of predictors (X_n) and their weights (a_n) as determined by least squares regression from an ensemble of atmospheric scenarios so that the difference between effective and true optical depth is minimized on average. One popular scheme over recent years is the RTTOV radiative transfer model, where detailed information and an extended list of references can be found in *Matricardi* [2008].

Forward model error is then a direct function of how variable these weights are. On average, the TPTR method can be highly accurate while extremely fast, as only a single exponent is evaluated for each layer per gas. One major limitation currently is that down-welling radiances reflected into the observation path are not modelled. However, this can potentially be parametrised as well in future versions. Perhaps its greatest shortcoming is the complete dependence on the training ensemble to capture the natural variability of Earth's atmosphere, including extreme events such as volcanic eruptions and forest fires. Even if a rare event is captured in the ensemble, the resulting forward model error should have an approximate inverse relationship to frequency of occurrence.

5.5.2 Correlated-k

Originally, satellite measurements in the infrared were from broad filter radiometers, so all spectral integrals had to be parametrised through computational necessity. The correlated-k (c-k) method [*Goody and Yung*, 1989] allows for substantial computational economy when calculating an optical property, such as transmittance, for a spectral band compared to LBL methods. The crux of the c-k method is based on the observation that most radiative parameters tend to be monotonic functions of absorption coefficient, and that these k values are well correlated along an atmospheric path for a single absorber. Therefore, the wavenumber, or frequency, variable (ν) is swapped with the cumulative distribution function (G) of k for each spectral interval ($\Delta\nu$), such that

$$\overline{R} = \int_0^1 R\left(k_\nu\right) d\left(\frac{\nu}{\Delta\nu}\right) = \int_0^1 R\left(k_G\right) dG, \tag{5.13}$$

where $\overline{R}$ is the mean spectral radiance. The cumulative distribution function is by definition a monotonic function of k,

$$G\left(k\right) = \int_0^k f\left(k'\right) dk', \tag{5.14}$$

where $f(k)$ is the distribution function of the absorption coefficient.

Numerically integrating Eq. (5.13) with G instead of ν generally requires substantially fewer summation terms as the cumulative distribution is so smoothly monotonic [*Lacis and Oinas*, 1991]. According to *Goody et al.* [1989], this economy scales as the number of spectral lines in the given interval, which may be greater than 10^4. However, they found the resulting error of assuming that k distributions are vertically correlated for a non-homogeneous atmosphere is on the order of 1 %. This is indeed small, but operational instrument noise is now on the order of 0.1 % or less. Furthermore, as spectral resolution continues to increase, the distribution function within an interval becomes less well populated, spoiling the smoothness of G, and requiring more summation terms asymptotically approaching those of the LBL method. Additional error may be accumulated when considering different absorbers with overlapping spectral regions. The c-k methodology is likely appropriate still for radiative flux calculations in climate models, however it fails to meet the accuracy needed to model current hyperspectral atmospheric sounders.

5.5.3 Optimal Spectral Sampling

Monochromatic RTMs allow for physically robust and flexible calculations of transmittance as they reduce Eq. (5.8) to Beer's Law when considering finite homogeneous layers. An alternative monochromatic method to computing radiances for every fine spectral grid point is to statistically parametrise a channel-averaged value based upon a limited subset. This is the method employed by Optimal Spectral Sampling (OSS) [*Moncet et al.*, 2008]. For each observed spectral channel, the OSS method seeks to approximate the channel averaged radiance as

$$\overline{R} \approx \sum_i w_i R\left(\nu_i\right), \tag{5.15}$$

where $R\left(\nu_i\right)$ are monochromatic radiances at a limited set (i) of selected "nodes" with weights, w_i. Nodes and weights are determined by statistical inference from an ensemble of LBL generated radiances representative of atmospheres displaying the natural variability of the detectable physical parameters. The OSS method is monochromatic and therefore maintains robust accuracy by physically modelling the spectroscopic parameters. Combining the OSS method with LUTs is likely slower than a TPTR method, but still significantly faster than LBL methods. However, the OSS is a statistical parametrisation method and shares the same limitations as those mentioned in the TPTR discussion, which are the dependence upon the accuracy and bounding of the training ensemble. If extreme or anomalous events are excluded from the training set, then the forward model is unlikely to accurately represent such events.

5.6 Selecting look-up table axes

Before creating LUTs, one must first decide upon the grid points along each dimension, i.e., temperature, pressure, and spectral wavenumber. For water vapour, there is also a partial pressure axis. Absorption cross sections, $k(\nu, p, T, e)$, must be tabulated at sufficient resolution in all four domains to avoid interpolation error. As seen in Eqs. (5.1) and (5.8), interpolation errors in k map non-linearly into errors in radiance space. Given a model atmosphere at defined pressure levels, it is somewhat trivial to create a LUT at the exact CG pressure and temperature values and reproduce highly consistent model radiances. The difficulty arises when attempting to create LUTs that encompass all possible atmospheric conditions and produce simulated radiances to a desired accuracy. As currently implemented, no extrapolation is performed for points lying outside the tabulated domain. Instead, the value at the closest boundary is used. As such, the LUTs were iteratively tested against an ensemble of atmospheres while constructing these grids by systematically comparing the changes in modelled spectral radiance to those produced in full LBL mode. Brightness temperature differences less than 0.02 K were desired, which is an accuracy criteria approximately an order of magnitude less than IASI instrument noise levels in much of the spectral range.

The 80 atmospheres created by *Matricardi* [2008] to parametrise the regression coefficients for RTTOV as discussed in Sect. 4.3.2 were used as the validation cases. Skin temperature was modelled to be 10 K greater than atmospheric surface temperature in order to enhance sensitivity of weak spectral lines and the pressure dependence of all lines in the lower atmosphere, which might otherwise be invisible in the absence of any temperature difference. Additionally a straight nadir viewing geometry was assumed. The objective is to meet the accuracy criteria over the entire 80 atmosphere ensemble while maintaining manageable file sizes with appreciable gains in computation speed.

5.6.1 Temperature

Creating robust LUTs require that the ranges of temperature, pressure, and partial pressure bound all possible atmospheric scenarios. Figure 5.4 shows the possible range of temperature for each pressure level represented in the RTTOV 80 atmosphere ensemble. The profile area within the solid black lines show the absolute maximum and minimum ensemble temperature values, which is the space that needs to be sampled and not the total area represented in grey shading. Therefore, a temperature grid was used relative to the midpoints between maximum

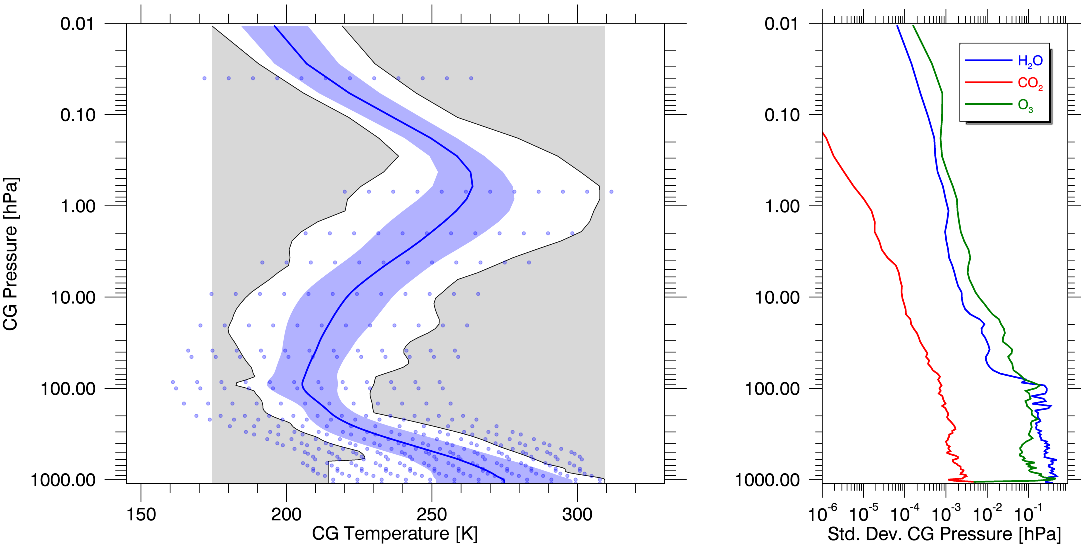


Figure 5.4: Left: The mean temperature profile from the 80 RTTOV atmosphere ensemble is shown as the solid blue line with lighter blue area representing the standard deviation of temperature profiles about the mean. The thin black lines depict the absolute minimum and maximum temperature values from the ensemble at those pressure levels. The blue dots help visualize the pressure and temperature grid points used in the H_2O LUT. Right: The standard deviations of the CG pressure levels specific to H_2O, CO_2, and O_3 for the 80 atmosphere ensemble.

and minimum profile temperatures as opposed to a regular grid populating the entire space between the absolute minimum and maximum temperatures. Relative temperature spacing uses approximately two thirds the memory as compared to sampling the entire grey area with the same sampling density. For reference, the dots in Fig. 5.4 show the temperature and pressure grid points for water vapour.

Absorption cross sections were calculated for all trace gases included in the ensemble atmospheres and are listed in Tab. 5.1. All relative temperature grids extend from ±50 K from the midpoint temperature profile. Temperature sampling increments are displayed in the second column of Tab. 5.1 for each trace gas. These were determined by first starting with an overly dense sampling and increasing the temperature spacing width until the difference in radiance compared to full LBL mode became appreciable relative to the IASI instrument noise level.

While temperature affects both the relative line strengths and line shapes of an absorber, absorption cross sections smoothly vary with temperature for a given

Table 5.1: Tabulated properties of the final and compressed gas specific LUTs. The first column shows the sample grid spacings of temperature where absorption cross sections are precalculated. The second column shows the absorption threshold from Eq. (5.17) used to thin the LUTs over the fine spectral grid. The third column displays the compressed size of the LUTs on hard disk and the fourth column shows the percent of the compressed size compared to the uncompressed LUTs.

	ΔT [K]	h	Size [Mbytes]	Compression [%]
H_2O	8.3	10^{-4}	2298	7.8
CO_2	4.5	10^{-5}	1738	9.1
O_3	5.5	10^{-5}	881	5.7
N_2O	11	10^{-5}	135	1.7
CH_4	11	10^{-5}	101	1.3
HNO_3	11	10^{-7}	359	4.6
CO	11	10^{-7}	127	1.6
NO_2	11	10^{-7}	147	1.9
NO	11	10^{-7}	15	.19
OCS	11	10^{-7}	57	.73
SO_2	11	10^{-7}	18	.23
Total			5623	4.4

layer. As an example, Fig. 5.5 shows water vapour contour plots of absorption cross section for two different wavenumbers. For any given pressure level notice that absorption cross section changes with low curvature as a function of temperature. Therefore, *Strow et al.* [1998] parametrized the temperature axis by fitting a 5th order polynomial through the optical depth values, while *Amato et al.* [2002] used a 2nd order polynomial, thus reducing the temperature grid by a factor of 2 to 3. For the LUTs created and analysed in this study, simple linear interpolation between temperature grid points was used. While linear interpolation requires a denser temperature sampling and larger file sizes, the LUTs are portable and rapidly decompressed to specified grids.

5.6.2 Pressure

Pressure primarily affects absorption cross sections by changing the spectral line shape functions due to collisional broadening. Line strengths are pressure dependent through collisional excitation, but this effect is minor in comparison and generally neglected in spectroscopic databases. When deciding upon the pressure axis for a particular gas, spacings can be incremented linearly with pressure, by the logarithm of pressure, or set with irregular spacings according to the resulting interpolation error. Figure 5.5 shows that the dependence of absorption cross sections upon pressure is not monotonic and is unlikely to be parametrised well by a low order polynomial or other simplistic functions. It may be that pressure is best interpolated in a piecewise manner. Since the H_2O LUT requires an extra e dimension, it is significantly larger in size than the other gas LUTs. Therefore, a tailored pressure grid was created for the H_2O LUT to minimize interpolation error while reducing the number of grid points to keep the file size relatively low. During the modelling of radiative transfer, absorption cross sections are linearly interpolated in the RFM along the pressure axis according to the logarithm of the given LUT pressure.

Reduced axis points for pressure were selected by starting with the 101 CG pressure levels and removing levels individually while checking the change in radiance over a small spectral range around $1500\,\mathrm{cm}^{-1}$. This pressure level removal process was repeated until the changes in brightness temperature approximately met the $0.02\,\mathrm{K}$ accuracy criteria. Finally, the 32 pressure levels shown in Fig. 5.4 were selected specific to H_2O.

The right plot in Fig. 5.4 shows the standard deviation of CG pressure from the 80 atmosphere ensemble. These standard deviations may be plotted as y-axis errorbars for the left figure, but are too small to be visualized if not displayed separately. Variability of CG pressure from one atmosphere to another is quite small when the vertical profile is approximated with many layers. In fact, the maximum standard deviation for H_2O is $\approx 0.3\,\%$ at $100\,\mathrm{hPa}$. Therefore, any atmosphere from the ensemble may safely be used to generate absorption cross sections without considering variations in CG pressure specific to the absorbing gas.

While an irregular pressure grid was used to create the H_2O LUT, the 101 level pressure grid native to the RTTOV ensemble modified by the CG approximation was used for the 10 remaining trace gases. Without a partial pressure axis increasing the size of the LUT, more pressure levels can be included while maintaining a

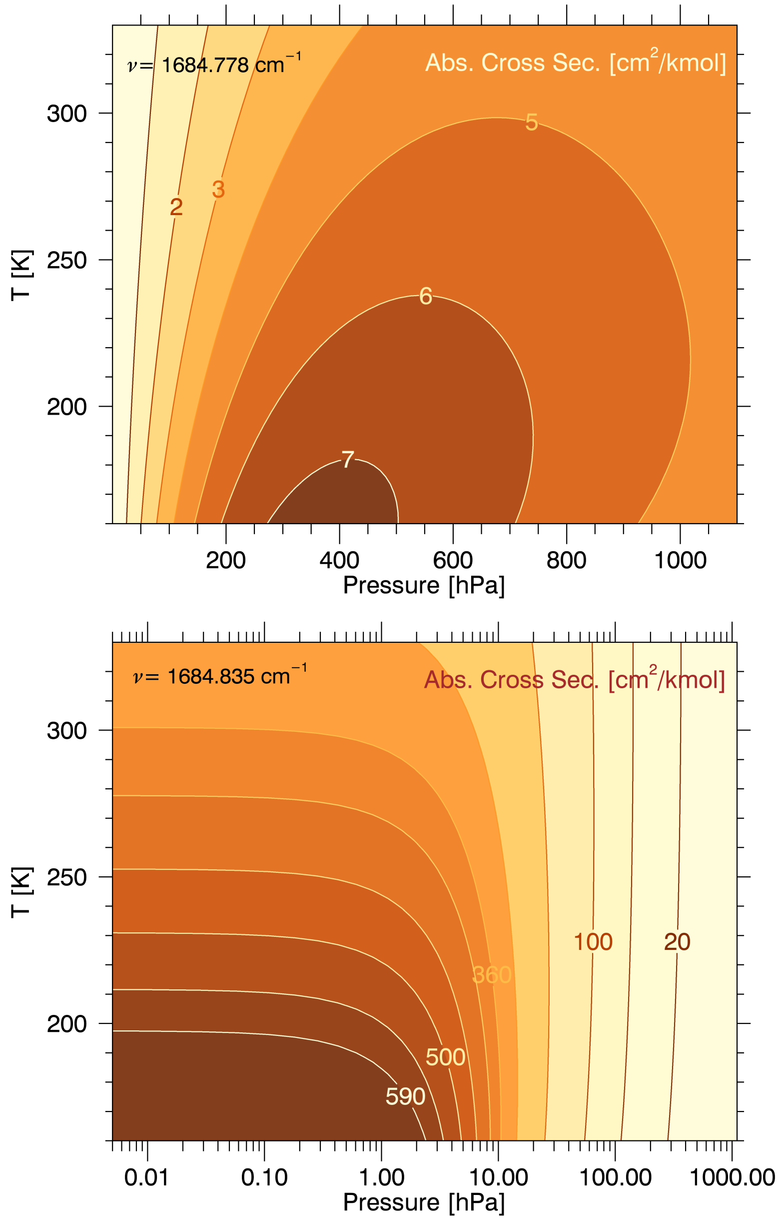


Figure 5.5: Contour plots of absorption cross section at line wing (top) and line centre (bottom) for water vapour using a mid-latitude VMR. Pressure is expressed linearly in the top plot and logarithmically in the bottom. Note the smooth temperature dependence at any pressure in contrast to the highly variable pressure dependence at fixed temperature.

manageable size on hard disk. A similar tailoring process was attempted for CO_2 and O_3, however it was found that far fewer pressure levels could be removed throughout the profile before interpolation error became noticeable.

5.6.3 Partial pressure for water vapour

Partial pressure axis values (units of hPa) for H_2O were selected in a manner similar to its pressure grid; where an overly dense e grid was reduced one point at a time, linearly interpolated between the remaining points, and verified by monitoring the change in radiance. The mean volume mixing ratios (VMRs) from the ensemble with extreme midpoint temperatures were used as the reference atmosphere when creating absorption cross sections. The axis points are scaling factors relative to the ensemble mean VMR and were selected to be 10, 35, 60, 80, 120, 200, 345, 600, and 1000%. Continuum effects are the extension and cumulation of self-broadening far from line centre and are not included in the water vapour LUT. These are best handled in the RFM calculations after referencing the saved absorption cross sections.

Partial pressure axes were not included for the remaining gases as spectral line shapes are only weakly dependent upon e for molecules without permanent dipole fields. Since the optical depth for a homogeneous layer is directly proportional to the density of the absorbing gas, these LUTs should provide accurate results regardless of e assuming the gas densities are properly adjusted when evaluating Eq. (5.8). To verify this posit, LUTs for CO_2 and O_3 at different partial pressures were created and found to yield negligibly different radiances as computed with the RFM.

5.6.4 Spectral Compression

Accurately reproducing the narrowest of spectral features detectable from nadir sounders, i.e., Doppler broadened lines in the upper stratosphere, require modelling the spectral grid at approximately 1,000 points per wavenumber. As atmospheric sounding instruments like IASI span thousands of wavenumbers, LUT size can push upwards of tens to hundreds of gigabytes on hard disk space. This can be unwieldy to share over network connections and the input/output operations alone may offset the potential gain in speed. Considering that each LUT originally contained over 2 million spectral points, the majority of LUT compression is achieved along the spectral axis. The fundamental question is should compression be determined by considering absolute or relative errors? A few examples of previous spectral compression techniques may help to guide the decision.

Substantial reductions can be achieved simply by taking a threshold of optical depth or absorption [*Scott and Chedin*, 1981; *Amato et al.*, 2002]. For each gas and layer, every spectral point that is optically thin beyond some threshold is removed from storage and reconstructed with either a small predetermined number or zero. *Scott and Chedin* [1981] used an absorption threshold value of 0.005. Alternatively, *Amato et al.* [2002] used 10^{-4}, which removed 88.5 % of all spectral points from storage, compressing the LUTs to 11.5 % of the original.

Sparks [1997] developed a technique to analytically select a reduced spectral grid by considering the change in line shape from removing spectral points and replacing the region with three-point Lagrangian interpolation. One benefit to this method is that the difference in absorption cross sections due to interpolating from a coarse to fine spectral grid need not be explicitly evaluated during grid selection, making it highly efficient. However, if the reduced spectral grid is fixed, then the cost to systematically raster through each fine grid point while checking interpolation error is only paid once. Furthermore, *Kuntz and Höpfner* [1999] pointed out that this spectral sampling is proportional to the line strengths and number density of the absorbing gas. In other words, the densest spectral sampling occurs near the centre of the strongest line for the most abundant gas, so errors are analysed in the absolute sense. One potential drawback to quantifying interpolation error in absolute terms is that a spectral grid constructed for a highly varying trace gas may produce unexpectedly large errors if that gas is significantly enhanced compared to the referenced amount.

Showing the limits of compressibility, *Strow et al.* [1998] developed a truncated SVD method to compress their LUTs to 1.4 % of their original size. However, LUTs compressed in this manner need to be expanded to their original dimensionality with the left and right singular vectors in order to extract the required absorption information. Thus, the truncated SVD method itself may not increase speed and likely add further processing time for the benefit of smaller LUTs. Interestingly, this problem was overcome by interpolating within the compressed space and regaining the speed increase. Additionally, *Strow et al.* [1998] found that the dynamic ranges for the optical depths of CO_2 and H_2O were so great that large optical depth values in the lower atmosphere dominated the reconstruction. Therefore, they reduced the range to compress by taking the 4th root of the optical depths and then applied SVD truncation, which required far fewer singular vectors to represent. The 4th root was decided upon by trial and error, however this SVD method seems to suggest that a relative compression metric performs better by treating all atmospheric layers with equal importance.

The key point to the method employed in this work is to use each pressure level in the LUT to define an atmospheric layer and then set the threshold criteria as the error in absorption for each layer independently. The spectral range computed extends continuously from 644 – 2761 cm^{-1} to cover the entire IASI range plus an extra wavenumber on either end to allow for apodization [*Amato et al.*, 1998]. A study conducted by *Ventress* [2014] showed that a convolution kernel width of 1 cm^{-1} was adequate to represent a 0.5 cm^{-1} full width at half maximum Gaussian apodization function. Spectral increments were set to spacings of $\Delta\nu = 0.001\,\mathrm{cm}^{-1}$ to capture all fine structure necessary for nadir-viewing scenarios. The total size of the 11 combined LUTs at this spacing was 124 Gbytes, which is unwieldy for storage and sharing.

This method to remove spectral grid points is numerical in nature. For an individual gas, at the l^{th} CG level the difference in absorption by removing a spectral point, k_i, and replacing that with a linear interpolation from neighbouring points,

$$k_i' = \left(\frac{k_{i+1} - k_{i-1}}{\nu_{i+1} - \nu_{i-1}}\right)\nu_i + \frac{\nu_{i+1}k_{i-1} - \nu_{i-1}k_{i+1}}{\nu_{i+1} - \nu_{i-1}}, \tag{5.16}$$

is evaluated and compared to a threshold value:

$$\left|\exp\left(-k_{igl}'u_{gl}\right) - \exp\left(-k_{igl}u_{gl}\right)\right| < h_g, \tag{5.17}$$

where h is the set threshold. If Eq. (5.17) is true for *all* CG levels, then the spectral point is removed from the LUT and the thinning process proceeds to the next point with the updated grid. Absorption threshold values used for each gas are displayed in the second row of Tab. 5.1.

An example of the thinned spectral grid (blue stars) is shown in Fig. 5.6 for the same H_2O line from Figs. 5.2 and 5.3. Using a threshold value of $h = 10^{-4}$ reduces the number of spectral grid points to just 3.7 % that of the original, while the residual interpolation error in total absorption is less than 0.2 %. However, there is an opportunity to remove even more spectral points if one considers that instruments such as IASI have much broader instrument line shape (ILS) functions than the irregular grid spacings shown. The apodized ILS for IASI is a convolution of a sinc and Gaussian function with a full width at half maximum (FWHM) of $0.5\,\mathrm{cm}^{-1}$ [*Amato et al.*, 1998], as highlighted in Fig. 5.6 for reference. Therefore, positive and negative error contributions separated by less than $\approx 0.5\,\mathrm{cm}^{-1}$ may be offsetting in the final apodized spectrum. To account for this, the thinning method was rerun with Eq. (5.17) evaluated after convolving the absorption cross sections over a local spectral cell with a simple triangular apodization function having a full base width of $0.5\,\mathrm{cm}^{-1}$. This absorption residual is also shown in Fig. 5.6 where irregular grid points are further reduced to 1.6 % of the original. Notice that the increase in positive and negative interpolation error on the fine grid will roughly cancel out once convolved with the IASI ILS and sampled on a grid of $0.25\,\mathrm{cm}^{-1}$.

Compressed LUT file sizes are shown in the third column of Tab. 5.1 with the percent relative to their uncompressed size shown in the fourth row. The total size on hard disk for all 11 trace gas LUTs is 5.62 Gbytes, which is just 4.4 % the size of the uncompressed LUTs. This is a significant compression of the originals, and can now be readily shared with the scientific community over the internet as opposed to transferring physical hard drives. It is important to note that interpolation error for this method is evaluated relative to each layer and treats all atmospheric levels with equal importance. This is surely a conservative approach and there are alternative methods available to further compress the spectral axis. However, if the lower limit using SVD truncation is approximately 1.4 %, then the conservative

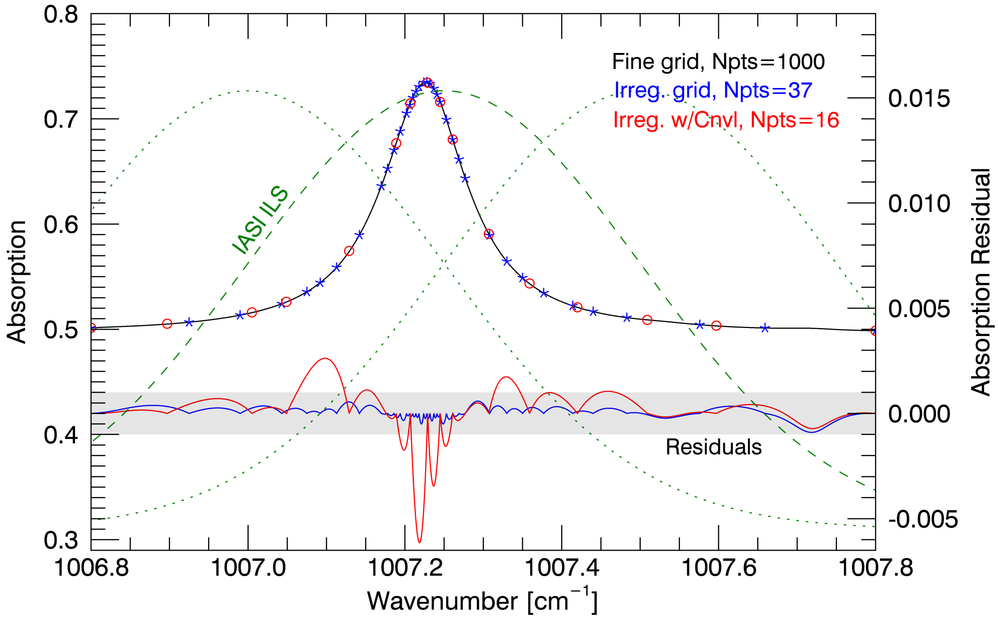


Figure 5.6: This plot shows the absorption spectrum for an isolated H_2O line in a tropical scenario with self-broadening and continuum effects included. The irregular grid produced by the spectral thinning method described in Sect. 5.6.4 ($h = 10^{-4}$) is annotated by the blue stars. The IASI ILS is overlain in green at each quarter wavenumber for reference. The irregular grid considering convolution with an ILS is shown as the red circles. Absorption residual errors (right axis) from linearly interpolating the irregular grids to the fine grid are overlain as the solid blue and red lines.

approach may be worth the extra size if the interpolation errors are small over a wide variety of atmospheric conditions.

Finally, care was taken to ensure that the minor trace gas LUTs were not overly compressed. Some of the weaker spectral lines for the less abundant gases may produce optical depth shapes that are essentially flat and entirely removed from the LUT. This will not be an issue for naturally occurring levels of that gas as represented in the GEMS database, but may lead to significant modelling errors for extreme and anomalous events such as forest fires and volcanic eruptions. Therefore, a far stricter threshold ($h = 10^{-7}$) was used for the last six minor trace gases listed in Tab. 5.1 to force retention of additional spectral grid points in the event of anomalously high mixing ratios. This was verified for SO_2 levels from a volcanic scenario following the validation method described in Sect. 5.8.

5.7 Reducing the radiative transfer spectral grid

Spectral compression reduces each LUT onto an irregularly thinned spectral grid. In order to calculate Eq. (5.1), the LUT spectral axis is interpolated to the regularly spaced fine grid, in this case $\Delta\nu = 0.001\,\mathrm{cm}^{-1}$. However, there is another opportunity to speed up the model by reducing the fine spectral grid used to calculate radiative transfer. The simplest option is to take the spectral points from the individual compressed LUTs and combine them together to form the union. Unfortunately, the resulting interpolation error from the union of LUT grid points for IASI purposes is unacceptably high. This is likely due to the effect apodization has on locally mixing spectral features from separate gases. Therefore, another methodology must be implemented.

Rather than thinning the fine spectral grid in absorption cross section or absorption space, as in Sect. 5.6.4, the common grid of radiative transfer is thinned by monitoring interpolation errors according to the blackbody brightness temperature (BBT) spectrum. The RTTOV atmosphere ensemble was used to generate 80 BBT spectra using all the listed gases in the atmospheric profiles. Then the error from removing a spectral point and replacing that with a linear interpolation from neighbouring points is calculated similar to Eq. (5.16), but using BBT instead of k. Threshold criterion for this method is set to

$$\left|\mathrm{BBT}_i - \mathrm{BBT}_i'\right| < 0.01\,\mathrm{K}. \tag{5.18}$$

If Eq. (5.18) holds true for all 80 spectra, then the i^{th} spectral point is removed. The resulting irregular grid for radiative transfer was reduced to 12.5% of its regularly spaced original size. Note that spectral radiance could be used in Eq. (5.18) instead. However, BBT is chosen because it is a common representation of the observation and it conveniently allows accuracy criteria to be expressed as a single value applicable across the whole IASI range, unlike radiance which would require a spectral dependence.

Evaluating Eq. (5.1) on the fine spectral grid is the top level calculation. Therefore, reductions in this grid translate proportionately to reductions in computation time. If the radiative transfer grid is reduced by a factor of ten, then one could expect computation speed to increase by approximately that same factor. However, this reduction is not free. The implicit cost of a reduced grid is further dependence upon the atmospheric ensemble and the explicit cost is a slight increase in interpolation error.

5.8 Look-up table performance

The desired outcome of using LUTs of absorption cross sections is to produce simulated radiances significantly faster than LBL methods with a negligible increase in error. Indeed, the computational time required to simulate an entire IASI spectrum using the 11 trace gases listed in Tab. 5.1 decreased by an average factor of 22 compared to LBL results. Including the reduced spectral grid from Sect. 5.7 further decreased computation time by an additional factor of six. Total computation time is 130 times faster than LBL mode on average. The RFM now computes full-range IASI radiances (8461 apodized channels) for the 11 combined gases in approximately 10 seconds on a standard desktop computer. Undoubtedly, there are numerous faster RTMs available to simulate IASI. However, the objective is to maintain uncompromising accuracy while becoming more operationally plausible.

Figure 5.7 shows the individual gas differences in BBT (ΔBBT) of the LUT results compared to the LBL results for the first five trace gases. Also, the noise equivalent differential temperature (NEΔT) of IASI is displayed for reference [*Hilton et al.*, 2011]. There are actually 80 separate lines of ΔBBT shown in these plots, one for each atmosphere from the RTTOV ensemble. Notice that all atmospheres studied produced ΔBBT values well below the IASI noise level. The remaining six trace gases are not shown because all of their NEΔT values were below 0.002 K. Upon close inspection of the H_2O plot, there appear to be a few spikes where ΔBBT approaches the IASI noise level. However, after analysis these spikes are purely artificial and result from selecting different methods of interpolating (polynomial or inverse polynomial) continuum effects when using the LUTs versus LBL mode, where one can argue equally well for either interpolation method.

Finally, all 11 gases were combined and calculated with the RFM to simulate IASI radiances over the 80 atmospheres. The resulting ΔBBT values from using LUTs versus full LBL are shown in Fig. 5.8. Over 99 % of the spectral channels have ΔBBT errors due to LUT interpolation less than 0.02 K for all atmospheric scenarios. As evident, many of these values are an order of magnitude or less than IASI noise. Additionally, if the LUTs were not appropriately gridded for any of the 80 atmospheres, then this would be apparent in Fig. 5.8. Similarly, Fig. 5.9 shows the resulting errors when including both LUTs and the reduced fine spectral grid from Sect. 5.7. As a result, over 99 % of all spectral channels have ΔBBT errors less than 0.04 K. Since the interpolation error from using these LUTs is several factors less than the IASI NEΔT, it is unlikely that the decrease in accuracy will be noticeable for any part of the spectrum over the vast majority of possible Earth atmospheres.

5.9 Summary

Look-up tables of absorption cross sections were created for the 11 most abundant trace gases along with an irregularly reduced spectral grid that allow the RFM to simulate IASI radiances approximately 130 times faster on average than its traditional LBL mode. These LUTs were spectrally compressed according to a simple linear interpolation scheme that accounts for offsetting positive and negative errors during apodization with the IASI ILS function. Thus, reducing total LUT file size from 120 Gbytes to 5.6 Gbytes, a compression to 4.4 % of the original. Furthermore, the resulting LUT error, as measured in brightness temperature, was rigorously validated for the 80 different atmospheres found in the RTTOV ensemble. More than 99 % of all IASI simulated spectral channels had a LUT interpolation error less than 0.02 K and a combined error including the reduced spectral grid less than 0.04 K, several factors below the IASI noise level.

As mentioned, several RTMs already incorporate LUTs or other “fast” methods in their numerical schemes. Perhaps the novelty in this work is the attention paid to error in the resulting radiances over a wide variety of atmospheric conditions with the flexibility afforded by creating a LUT separately for each individual trace gas. By carefully selecting the range and spacings of the LUT grids, it should be quite difficult to devise an atmospheric Earth scenario that results in large radiance errors due to LUT interpolation.

In fact, Dr. David Moore of the University of Leicester has already incorporated these LUTs and the irregular spectral grid into their IASI iterative retrieval scheme for CO that uses the RFM as the forward model. Even though their CO retrieval is distributed over many computing nodes, it is too slow for operational use and thus limited to scene specific analysis. By utilizing these LUTs he sped up the retrieval process by a factor of 10 and observed CO retrieval error due to LUT interpolation to be less than 1 %, far below the estimated uncertainty of CO. They did not observe the factor of 130 increase because the CO retrieval uses only a $40\,\mathrm{cm}^{-1}$ range where the computation of transmittance is fast enough that initialization overhead from creating arrays becomes noticeable. In short, the University of Leicester CO retrieval can now observe 10 times more IASI pixels in a given time frame with negligible added error.

These LUTs of absorption cross sections were created with the intent of using the RFM as a forward model for IASI in a more operational context. However, these LUTs are not specific to IASI and may be used with the RFM in any nadir-viewing scenario from $645\,\mathrm{cm}^{-1}$ to $2760\,\mathrm{cm}^{-1}$ where the instrument noise and spectral spacings are roughly greater than or equal to that of IASI. In fact, these LUTs could potentially be incorporated into any existing or future RTM as the absorption cross section points were designed for simple linear interpolation within the LUT grids.

Finally, aside from presenting a generalized method of creating, compressing, and validating LUTs, these are also freely available to the scientific community. The 11 gases listed in Tab. 5.1 are just a starting point. Other trace gas species including ammonia, CFCs, and several organic compounds are also detectable by IASI. Instructions on how to use the RFM to generate LUTs of absorption cross sections are posted on the RFM website along with the source code for spectral compression and several other helper functions for manipulating the LUTs. Each gas-specific LUT takes approximately two days to generate per computation node from creation to compression if covering the entire IASI range. This process is much faster for smaller spectral intervals. Further information on how to download the latest version of the RFM and the presented LUTs can be found by browsing to `www.atm.ox.ac.uk/RFM`.

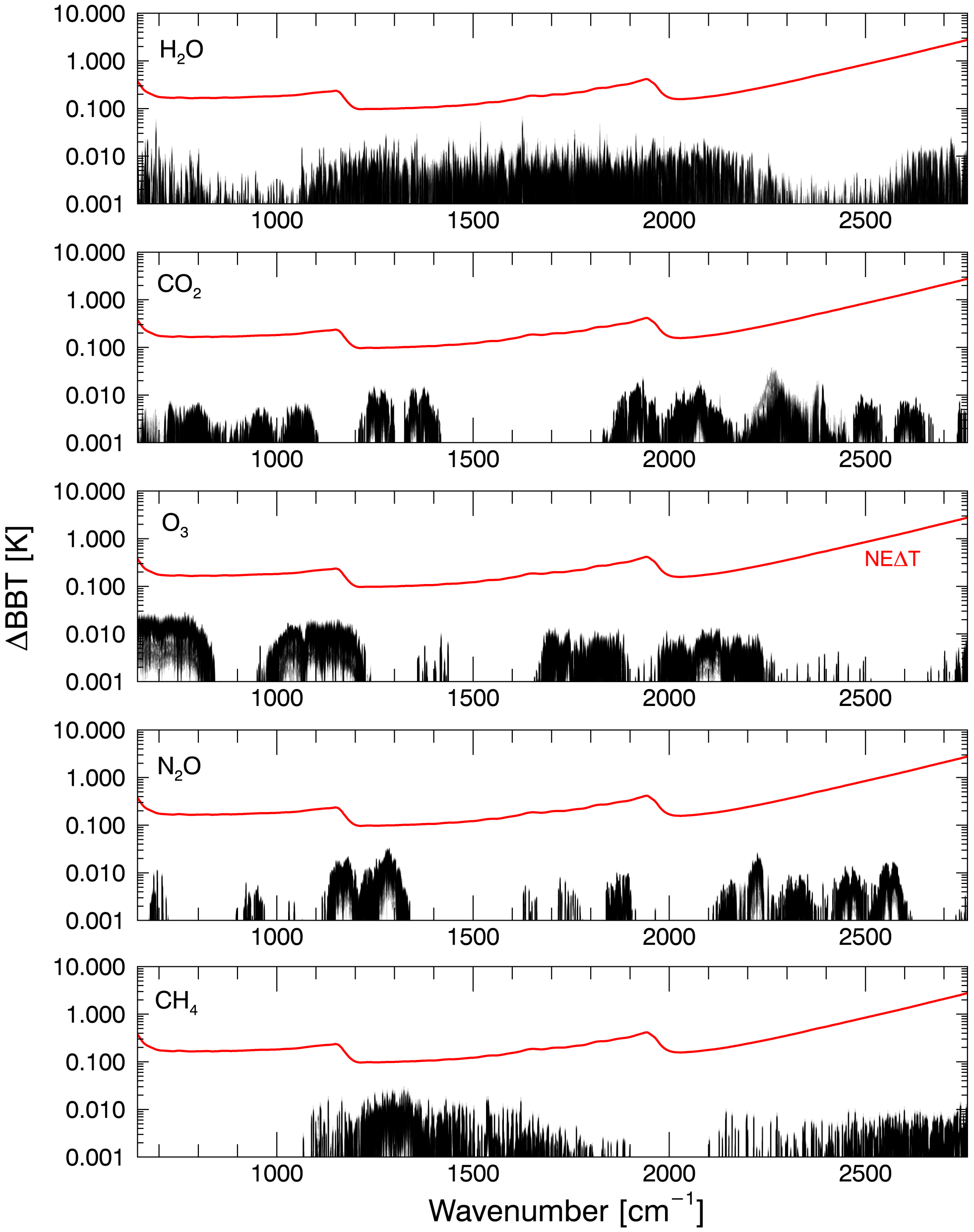


Figure 5.7: The magnitude of differences in brightness temperature from using LUTs to simulate IASI spectra versus full LBL mode with the RFM (black lines). The plots are specific to the individual brightness temperatures of H_2O, CO_2, O_3, N_2O, and CH_4. Each plot shows 80 ΔBBT lines over-plotted, one for each atmosphere from the RTTOV ensemble. The IASI NEΔT noise level at 280 K is shown in red for reference.

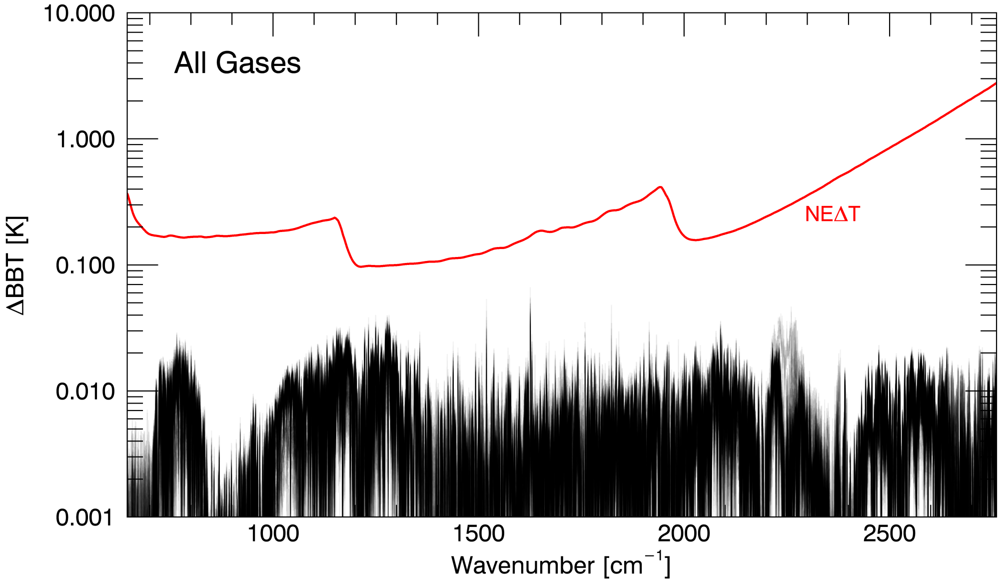


Figure 5.8: The magnitude of differences in brightness temperature from using LUTs to simulate IASI spectra versus full LBL mode with the RFM (black lines). All 11 gases were included in the radiative transfer calculations. In this figure, 80 ΔBBT lines are over-plotted, one for each atmosphere from the RTTOV ensemble. The IASI NEΔT noise level at 280 K is shown in red for reference.

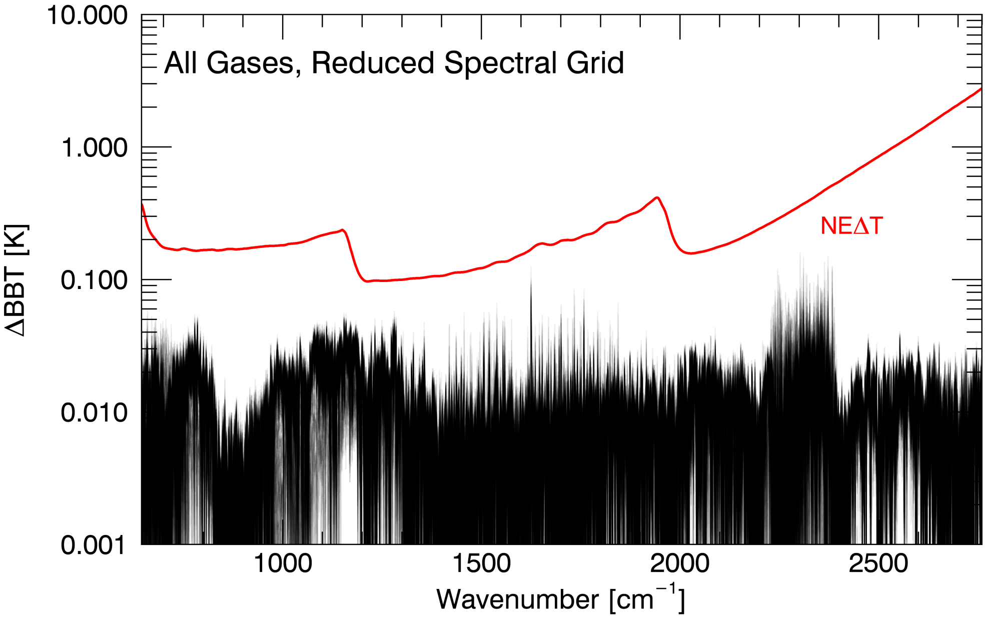


Figure 5.9: Same as in Fig. 5.8, but also utilizing the reduced fine spectral grid from Sect. 5.7.

The day you stop doing the small things is the day you think you're above everyone else.

— Kevin Hart

6

Atmospheric carbonyl sulphide (OCS): Background

Contents

6.1 Motivation

Carbonyl sulphide (OCS) is a long lived trace gas and, as such, is the most abundant sulphur containing gas in the atmosphere [*Montzka et al.*, 2007]. Therefore, estimates of OCS are crucial towards modelling the global sulphur cycle. While tropospheric OCS is a greenhouse gas of minor importance, stratospheric OCS converts to sulphate aerosols which play a significant role in Earth's radiative budget by scattering incoming short-wave radiation. In fact, OCS is one of the major contributors to stratospheric sulphates during volcanically inactive periods. Work by *Notholt et al.* [2003] suggested that elevated OCS concentrations in the upper tropical troposphere increase the stratospheric aerosol surface in the tropical lower stratosphere by a factor of two.

Aside from radiative effects, OCS is also linked to the carbon cycle of the biosphere. Uptake by vegetation is the major sink of OCS, therefore it may be possible to use OCS as a tracer for CO_2 absorption and photosynthetic productivity of ecosystems [*Berry et al.*, 2013]. Additionally, micro organisms in soil may use OCS in a similar manner to vegetation, thus providing additional insight on subsurface biological activity [*Wingate et al.*, 2008; *Seibt et al.*, 2006].

Finally, the oceans are the dominant source of global OCS, from both dissolved organic material and phytoplankton activity. Current models estimate the bulk of OCS originates from dissolved organics [*Launois et al.*, 2015a], but the proportion originating from this versus living phytoplankton is still uncertain. Given that phytoplankton provide the primary mechanism of long-term atmospheric carbon sequestration, monitoring the level, distribution, and breakdown of organic carbon in the ocean would be a major breakthrough in our understanding of the biosphere.

6.2 OCS sources

Carbonyl sulphide is emitted from a wide range of sources from processes occurring in land, ocean, and atmosphere. For nearly two decades it was thought that one half of atmospheric OCS originated from the oceans either directly or indirectly [*Chin and Davis*, 1993]. Now current models estimate that nearly 70% of OCS comes from the oceans [*Berry et al.*, 2013]. Another quarter of OCS arises from anthropogenic activity and biomass burning, but the proportion of these sources is largely unknown. The remaining contributions to OCS production come from anoxic soils and a trace amount from volcanic activity. Figure 6.1 shows a pie

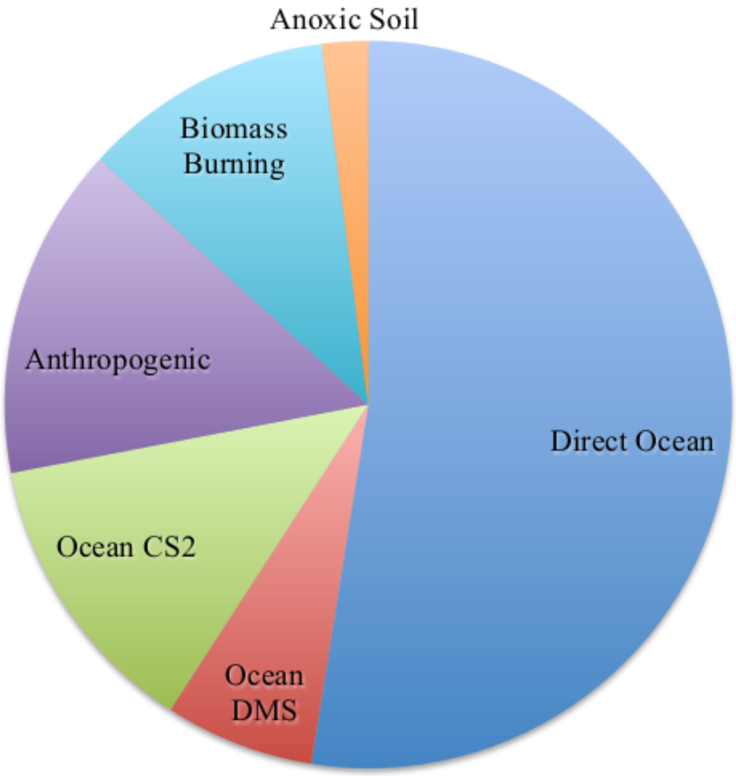


Figure 6.1: The proportion of OCS sources is shown according to current understanding. The total source flux of sulphur per year is estimated to be approximately 1.2 Tg [*Berry et al.*, 2013].

chart of the relative proportions of OCS sources as presently known. The total annual flux of OCS is currently estimated to be in the range of 0.9 – 4.5 Tg of sulphur [*Launois et al.*, 2015a] with best estimates matched to surface samples of 1.2 Tg [*Berry et al.*, 2013] and further validated by *Kuai et al.* [2015]. The physical mechanism behind each source is discussed in turn.

6.2.1 Direct ocean emissions

Direct emissions of OCS from the ocean are now estimated to be approximately 800 Gg of sulphur per year. Current modelling efforts [*Launois et al.*, 2015a] show that the majority of oceanic OCS likely originates from photochemical breakdowns of dissolved organic matter. While several organic sulphur precursors have been suggested [*Xu et al.*, 2001], knowledge of these precursors is tenuous and there is currently no mechanism to measure them directly. However, OCS production can be parametrized and quantified according to physically measurable quantities, i.e., chromophoric dissolved organic matter (CDOM) and ultra-violet (UV) radiation impinging the ocean surface. In short, direct OCS ocean emissions are proportionate to the amount of UV absorption by surface organics. These quantities can be estimated from satellite using ocean colour products, such as those found at `http://oceancolour.org` and `http://oceancolor.gsfc.nasa.gov`.

While surface incident radiation is a relatively easy quantity to estimate from solar position and retrievals of cloud cover, CDOM is a higher level ocean colour product not directly measured. In lieu of CDOM, *Launois et al.* [2015a] chose to parametrise their emissions model based upon the estimated oceanic absorption coefficient at 350 nm (a_{350}) [*Para et al.*, 2010]. However, a_{350} is not routinely retrieved as satellite sensors typically used for ocean colour do not observe at wavelengths below 400 nm, e.g., MODIS and SeaWiFS. To overcome this fact, *Launois et al.* [2015a] developed three parametrisation schemes to either extrapolate a_{350} from a_{440} or infer from chlorophyll concentrations. Therefore, a_{350} is used as an indicator of CDOM concentrations in the top levels of marine surfaces.

Additionally, measurements of marine OCS concentrations in low light scenarios imply that there is a 'dark' production pathway that also depends upon the amount of organic matter [*Von Hobe et al.*, 2001]. Therefore, *Launois et al.* [2015a] also parametrized this pathway to estimates of a_{350}. Surely dark production suggests an independence from light sources, but a_{350} ultimately provides an indirect measure as to the richness of near surface organic matter. The mechanism describing how OCS is directly produced in the ocean in the absence of light is poorly known. Figure 6.2 shows the current model of OCS direct ocean emissions by *Launois et al.* [2015b] averaged for the months of January and July. This includes both the photochemical and the dark pathways for OCS production as well as hydrolysis, which removes oceanic OCS in deep waters through reactions with H_2O and OH^-.

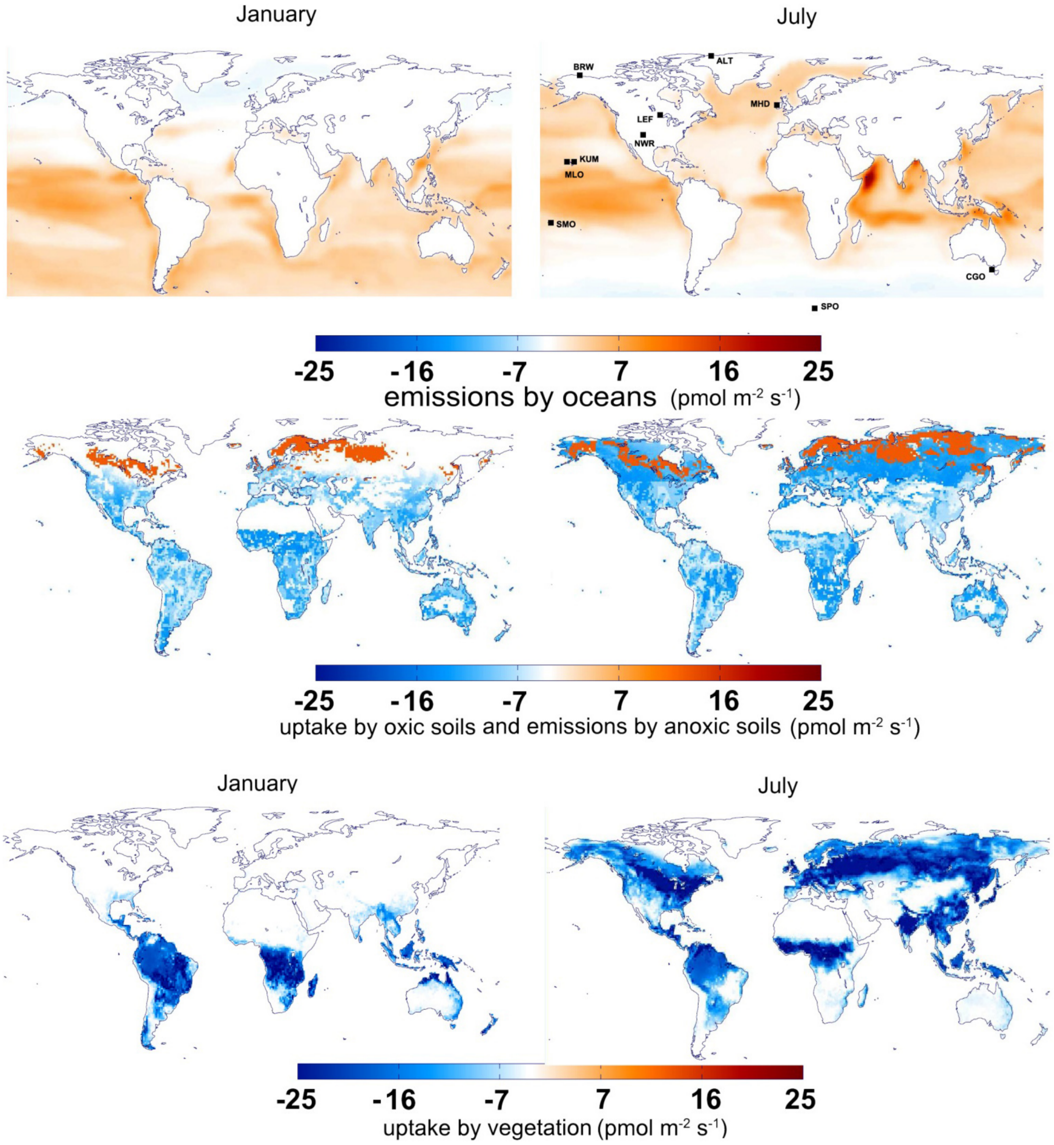


Figure 6.2: The top row shows modelled [*Launois et al.*, 2015a] monthly mean direct ocean emissions of OCS from January (left) and July (right), with uptake and emission by soils in the middle row, and vegetation uptake in the bottom row. This figure is reproduced from *Launois et al.* [2015b].

6.2.2 Conversion from marine CS_2

CS_2 is a short lived trace gas with an atmospheric lifetime of only days. *Chin and Davis* [1993] reported that the entirety of atmospheric CS_2 is rapidly lost to oxidation with the OH radical. While less is known about the creation of CS_2 than OCS, it is thought to largely follow the same mechanism as OCS oceanic production (photolysis of dissolved organics). In fact, models by *Kettle et al.* [2002] showed that oceanic CS_2 fluxes follow the same yearly pattern as OCS direct emissions. Experiments suggest that as much as 87% of CS_2 is ultimately converted into OCS [*Barnes et al.*, 1994, references therein]. However, oceanic fluxes of CS_2 will not be entirely correlated with OCS direct fluxes as the majority of ocean CS_2 is unaffected by temperature dependent hydrolysis [*Kettle et al.*, 2002].

6.2.3 Conversion from marine DMS

Dimethyl Sulphide (DMS) is another short lived trace gas emitted from the oceans that quickly reacts with the OH radical to form reduced sulphur compounds. Unlike CS_2, only 0.7% of DMS converts into OCS [*Barnes et al.*, 1994]. However, much more DMS is emitted from the oceans such that approximately the same amount of OCS originates from DMS as oceanic CS_2. Interestingly, nearly all DMS in its precursor form (dimethylsulphoniopropionate, DMSP for short) originates from phytoplankton that may utilize the reactivity of DMS and its subsequent sulphur products as a cellular pressure regulator and antioxidant to neutralize free radicals [*Sunda W. et al.*, 2002]. The intracellular concentration of DMSP is highest in the dinoflagellates and haptophytes, such as the coccolithophores. Therefore, OCS originating from DMS is associated with phytoplankton blooms, which tend to be maximum in during the local hemisphere spring [*Kettle et al.*, 2002].

6.2.4 Biomass burning

The amount of OCS released from biomass burning is still largely uncertain. *Kettle et al.* [2002] estimated biomass burning to yield between $11 - 64$ GgS per year while [*Berry et al.*, 2013] found that 136 GgS was a better match to observed surface measurements. Latitudinal distributions of OCS from biomass burning are still approximated in models as a constant flux term throughout the year.

Approximately 90% of biomass burning is a result of human activity, e.g., crop field management, firewood and peat consumption, deforestation, and the burning of agricultural waste. While CS_2 has not been directly observed from biomass burning, this does not mean all OCS is directly emitted. The lifetime of CS_2 may be quite short in the vicinity of combusting biomass. In the past, OCS from biomass was inferred from observed ratios of OCS to CO_2 [*Chin and Davis*, 1993]. The bulk of CO_2 is thought to be released during the firing stage, but the majority of OCS is emitted from biomass during the smoldering stage. How these two stages are physically linked to separate production of CO_2 and OCS is not known. Of course, the amount of OCS emitted depends upon the sulphur content of the specific biomass. Accurate models of this process would require a global atlas of sulphur content by forest and field type, which is not yet available.

6.2.5 Anthropogenic emissions

There are numerous proposed sources of anthropogenic OCS emissions with each source representing a highly unique man-made physical process. *Kettle et al.* [2002] suggested that twice as much atmospheric OCS originates from anthropogenic emissions of CS_2 than direct emissions. *Chin and Davis* [1993] discuss a myriad of sources, but conclude that the majority of industrial emissions result from the use of CS_2 as a solvent carrier in the regeneration of cellulose rayon and cellophane [*Smith and Timmerman*, 2003]. These two account for roughly 75% of industrial sources of CS_2 with the remaining portion coming from CCl_4 production, rubber chemicals, pesticides, and chemical reagents.

Combustion of coal in power plants is a detected source of OCS, however no appreciable amounts of CS_2 have been measured from smoke stack plumes. The amount of OCS emitted from coal burning has likely changed significantly over the past several decades due to increased regulation on atmospheric pollutants. The majority of inorganic sulphur (ash) is removed by crushing and washing coal before combustion. Additional sulphur may be removed by chemically treating the coal to liberate and capture sulphur containing gases and also by scrubbing the flue-gas before leaving the smokestack, which requires significant maintenance and cleaning [*Srivastava and Jozewicz*, 2001]. Most countries have policies implemented to enforce, at least, the washing of coal prior to combustion, but the additional sulphur removal steps require a significant increase in cost and engineering skill. Furthermore, while these processes markedly reduce SO_2 emissions to the lower troposphere, it is not entirely clear whether CS_2 or OCS are emitted as a by-product of the coal cleaning process.

Oil refineries and natural gas processing facilities are another source of anthropogenic OCS as a result sulphur recovery efforts [*Peyton et al.*, 1978]. As mandated by either national or local environmental regulations, sulphur compounds (mostly H_2S) are removed from oil and gas and converted back to elemental sulphur. This greatly reduces SO_2 production during combustion. However, OCS and CS_2 are produced during sulphur recovery and small amounts leak into the atmosphere. The amount of total OCS emitted from oil and gas processing is proportional to the leakage and production capability of the processing center.

Finally, automotive engines are emitters of both OCS and CS_2 [*Peyton et al.*, 1978]. The physical process is specific to the engine design, fuel type, and catalytic exhaust systems of a particular engine brand. However, *Fried et al.* [1992] found high correlation between CO and OCS emissions. Just like oil and gas refineries, recent atmospheric regulations seek to decrease the amount of sulphur-containing gases emitted from automotive exhaust. Diesel tends to contain higher levels of sulphur than petrol gasoline because of the added processing that is required to refine petrol from crude oil. However, most advanced countries enacted policies to reduce sulphur in diesel from historical mandates of 5000 ppm (prior to 1993 in the United States) to as low as 15 ppm of sulphur or below. Developing countries are also following suit in order to reduce local SO_2 emissions.

It is worth noting that countries like the United States provide tax-free or reduced cost diesel to farmers for the operation of tractors and other machinery which may not have the low sulphur regulation. Sulphur in diesel reacts with nickel in the metal alloys to produce lubricants that protect fuel injection systems. As sulphur is reduced in transportation diesel, other lubricants must be added or fuel injection systems re-engineered to withstand greater friction. This would increase the cost of farm diesel and thus food production. Therefore, farm diesel may continue to contain higher levels of sulphur over the near future until either regulations or diesel supply forces the full switch to “ultra-low sulphur” diesel. Interestingly, bio-diesel is naturally high in lubricity and very low in sulphur content.

6.2.6 Anoxic soils

Wetland and estuarine soils have been identified as sources of OCS [*Watts*, 2000]. Salt marshes, rice paddies, and peatlands are particularly high emitters of OCS. However, the presence of dense vegetation may overcome the emissions process and turn the area into a net sink. Subtropical salt marshes are thought to be the only continuous net soil emission area of OCS, but occupy small geographic areas [*Whelan et al.*, 2013]. In general, emissions of OCS from soil tend to be correlated with methane emissions suggesting that the process of decomposing organics in lower oxygen environments out-gasses several products.

The middle row in Fig. 6.2 shows the areas of OCS flux in red from anoxic soils based upon current soil inventories [*Launois et al.*, 2015b]. Note that the areas of greatest OCS emission are located in far north latitudes where the soil is, at least historically, in permafrost for the majority of the year. Much effort is currently spent quantifying methane emissions from melting permafrost. In fact, the model generating Fig. 6.2 uses estimated methane emissions to map anoxic soil activity. Therefore, as permafrost declines and methane emissions rise, so will the modelled emissions of OCS from these areas.

6.2.7 Volcanic emissions

Since volcanic activity accounts for the majority of atmospheric SO_2 emissions, one might think that volcanic plumes are also rich in OCS and CS_2. However, volcanoes actually contain very little of either of these products in their gaseous plumes [*Chin and Davis*, 1993]. Reactions to SO_2 and sulphate aerosols from OCS are energetically favourable. Therefore, it is unlikely to observe the reverse reaction of OCS production from SO_2. While volcanoes do emit detectable amounts of OCS, (approximately the same amount as anoxic soils per year) their injection is sporadic and generally ignored in chemical flux models [*Berry et al.*, 2013].

6.3 OCS sinks

Sinks of OCS are broadly categorized into three categories; uptake by vegetation, uptake by oxic soils, and loss due to atmospheric processes. Figure 6.3 shows current best estimates of the relative proportion of OCS sinks [*Berry et al.*, 2013; *Launois et al.*, 2015a]. Data collected by *Montzka et al.* [2007] since 2001 shows no upward or downward trend in OCS VMRs across numerous sampling sites. Therefore, the sinks of OCS are thought to balance the sources (≈ 1.2 TgS per year). However, recent observations by *Kremser et al.* [2015] show a positive upward trend in tropospheric OCS in three locations across Australia from 2001 to 2015. It is possible these results are local to Australia such that the global sinks balance OCS sources, but further estimates with enhanced spatial and temporal resolution, like those from satellite observations, are required to resolve the discrepancy.

Assuming global flux balance, *Montzka et al.* [2007] estimated that the implied atmospheric lifetime of OCS is between 1.5–3 years. However, this is a global average and the actual lifetime of an OCS molecule from creation to destruction depends upon the strength of and proximity to a sink of OCS. The tropospheric lifetime of OCS emitted from the oceans is likely close to the 1.5–3 year range, but the lifetime of OCS emitted near a large forest during peak growing season will be substantially shorter. Therefore, with regards to satellite estimates, it would be unrealistic to attribute an observed diurnal signal to areas that are weak sinks of OCS given its long atmospheric lifetime. However, if a diurnal signal of OCS were detectable, then the most likely area to show physical diurnal variation would be over dense forests during local summer where photosynthetic processes are directly connected to solar radiation.

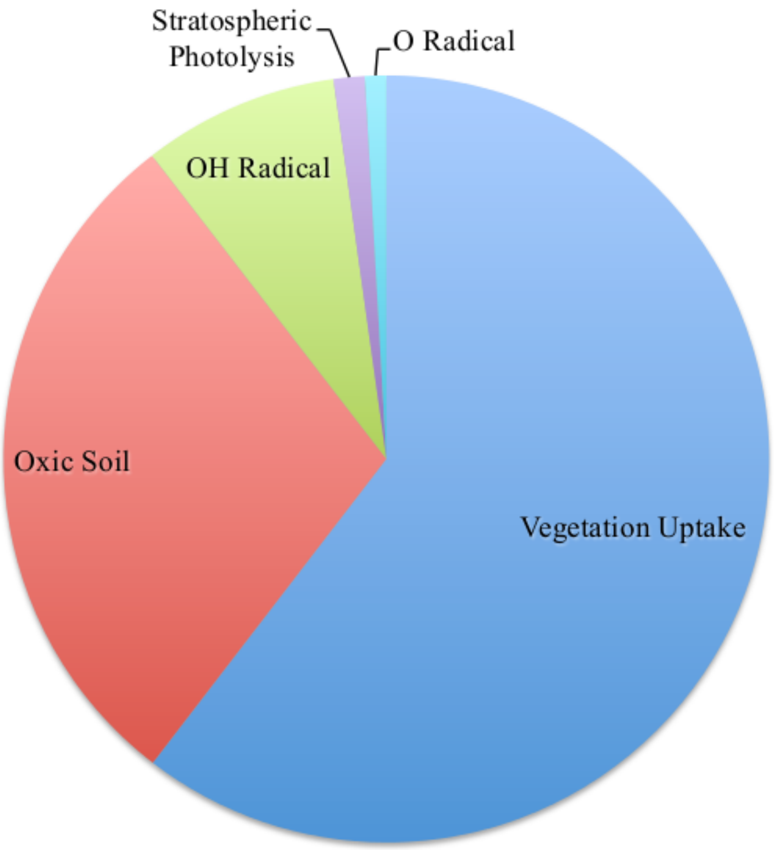


Figure 6.3: The proportions of relative OCS sinks are shown according to current understanding. The estimated proportion of vegetation uptake was greatly increased after the findings presented in *Suntharalingam et al.* [2008].

6.3.1 Vegetation uptake

As shown in Fig. 6.3, the majority of OCS is removed from the atmosphere by vegetation uptake. OCS takes the same diffusive pathway as CO_2 through plant stomata to the reaction sites in the chloroplasts. OCS is then consumed by the enzyme carbonic anhydrase (CA), which splits OCS into CO_2 and H_2S [*Protoschill-Krebs and Kesselmeier*, 1992]. Ingestion of OCS via leaf uptake is a one way ticket, meaning, plants do not respire OCS as they do unused CO_2. Therefore, as this process is physically linked to photosynthesis, OCS uptake by vegetation will be maximum during peak day-time activity and essentially zero after sunset. Current methods to model OCS leaf uptake are summarised in *Berry et al.* [2013].

The latest model of OCS uptake by vegetation is shown in the bottom row of Fig. 6.2. While the Amazon and Congo rainforest areas are continual sinks of OCS, the majority of vegetation uptake clearly occurs in the Northern Hemisphere summer due to the greater extent of plant covered land mass. The leaf uptake of OCS is species specific, therefore [*Launois et al.*, 2015b] included maps of plant functional types in climate specific zones to model the results shown in Fig. 6.2.

6.3.2 Oxic soils

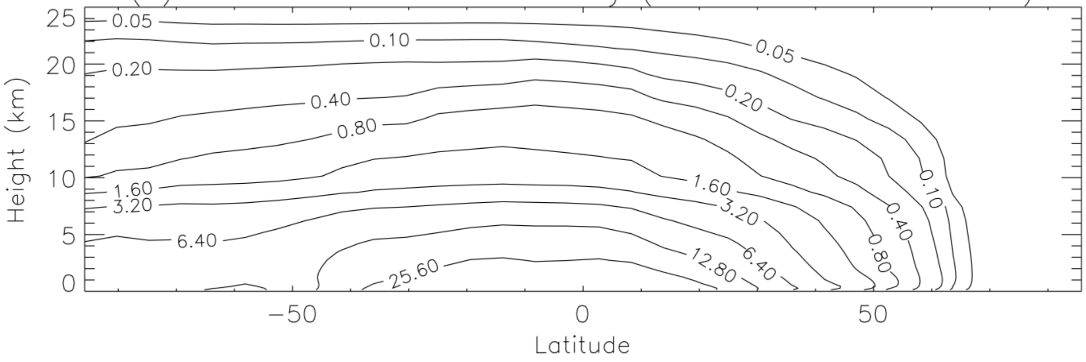
Soil uptake is the second most important sink of OCS, as shown in Fig. 6.3, and accounts for nearly 30 % of OCS removal from the atmosphere. Certain soil bacteria also contain the CA enzyme and those associated soils will act as a sink if OCS can diffuse into the top layer. The rate of absorption is a function of soil porosity, temperature, CA activity, and water content [*Van Diest and Kesselmeier*, 2008]. In a study conducted by *Kato et al.* [2008], the genera *Mycobacterium* was found to strongly absorb OCS at climatological concentrations. In a 30 hr time span, *Mycobacterium* in soil reduced air concentrations of OCS from 530 ppt to 330 ppt for the tests conducted. Additionally, *Yi et al.* [2007] observed that soil uptake of OCS was proportional to CO_2 respiration by soil production.

Modelled uptake by oxic soils is shown in the middle row of Fig. 6.2. *Launois et al.* [2015b] modelled these results not by proportionality to CO_2 production, but rather to soil uptake of H_2. This is because the *Streptomyces* bacteria have a high-H_2 affinity and share striking similarities with the development and morphological life cycles of *Mycobacterium* [*Scherr and Nguyen*, 2009]. Therefore, as H_2 uptake by soils is well-modelled, it acts as a proxy for estimating OCS uptake.

6.3.3 Oxidation with radicals

Loss of OCS due to oxidization with radicals is the dominant atmospheric sink. The primary radical species affecting OCS are OH and O in both the troposphere and stratosphere. Approximately 10 times more OCS is reduced because of OH interactions than O. The destruction of OCS due to OH is greatest near the surface

(a) COS oxidation rate: January (units: molecule cm^{-3} s^{-1})



(b) COS oxidation rate: July (units: molecule cm^{-3} s^{-1})

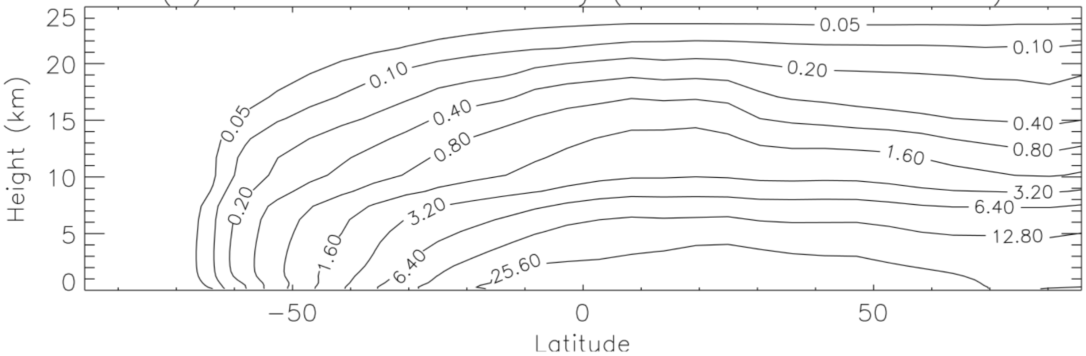


Figure 6.4: The oxidation rates of OCS by the OH radical are shown according to height and latitude for the months of January and July. This figure was taken from *Kettle et al.* [2002].

along latitudes of maximum solar zenith angle and is a function of OCS concentration, OH concentration, and the temperature dependent reaction rate constant.

The latitudinal and yearly dependence of the OCS oxidation rate due to OH was calculated by *Kettle et al.* [2002] and is shown in Fig. 6.4. Many generalizations were made to generate this plot, specifically the OCS and OH profiles and their latitudinal trends. However, this figure shows that the global sink of OCS due to OH is maximum near the tropical surface and minimum in the upper stratosphere and in high latitude winter. Figure 6.4 implies that high latitude injections of OCS from land (perhaps biomass burning or anthropogenic activity) blown over ocean during winter may persist until the season changes and OH oxidation increases.

6.3.4 Stratospheric photolysis

Finally, OCS can be photochemically reduced in the atmosphere by UV radiation with wavelengths below 388 nm [*Chin and Davis*, 1993]. Stratospheric loss due to photolysis is 15 times greater than in the troposphere. However, this effect accounts for only a small portion of the OCS flux budget and is generally lumped in with oxidation and accounted for as total atmospheric loss [*Berry et al.*, 2013]. For altitudes above the troposphere, the intensity of UV radiation accounts for the decrease in OCS VMR with altitude from both direct and indirect processes, such as the production of O in the upper stratosphere.

6.4 Previous estimates of OCS from satellite

There is a growing network of ground samples [*Montzka et al.*, 2007], zenith-viewing FTS observations [*Kremser et al.*, 2015], and aircraft campaigns [*Wofsy*, 2011] that estimate OCS concentrations or total column amounts. For example, there are 14 NOAA research sites (10 are annotated in the top right of Fig. 6.2) collecting flask samples to measure trace gases on a near weekly basis. While point samples of OCS are useful to constrain chemical transport models and identify potential global-scale mechanisms of OCS production and termination, they are sparse in spatial and temporal sensitivity. Observations by satellite are the only practical means by which to fill in the gaps and map OCS distributions across all latitudes and seasons. However, OCS is a challenging gas to estimate from satellite, as will be discussed further in the next chapter.

OCS was first observed from satellite by the Interferometric Monitor for Greenhouse Gases (IMG) [*Clarisse et al.*, 2011]. Since IMG collected data for less than two years (1996-1997), OCS was not pursued further until the launch of the Atmospheric Chemistry Experiment (ACE) instrument that began operation in 2003 [*Barkley et al.*, 2008]. ACE is a solar occultation instrument that views the sun through the limb of Earth's atmosphere. Therefore, ACE is well designed for stratospheric sensitivity, but cannot reliably sound the troposphere below an altitude of approximately 8.5 km. There were two major results from this work. First, they showed that OCS vertical profiles above the tropopause decrease steadily with altitude, thus confirming that there is no appreciable source of OCS due to stratospheric chemistry. Secondly, stratospheric OCS tends to be greater towards the equator and less at the poles. This general trend was also confirmed for tropospheric OCS based on a compilation of zenith-viewing ground observations and balloon campaigns [*Krysztofiak et al.*, 2015].

In retrospective analysis, *Glatthor et al.* [2015] used the limb sounding MIPAS instrument to retrieve OCS concentrations at the lowest-most detectable level, 250 hPa, using standard optimal estimation techniques. The compiled results from 2002 – 2012 in 5° by 15° latitude-longitude bins showed clear evidence of elevated ocean sources and tropical rainforest sinks that vary with season. However, limb sounding instruments are not ideal for tropospheric sounding and the 250 hPa level fails to probe the troposphere at high latitudes as the tropopause decreases in altitude from the equator.

Most recently, *Kuai et al.* [2014] developed an optimal estimation retrieval scheme as discussed in Ch. 3 to retrieve OCS amounts using the Tropospheric Emissions Sounder (TES). TES is a nadir-viewing FTS instrument aboard NASA's Aura satellite that was launched into polar orbit in 2004. TES observes in the thermal IR with an unapodized spectral resolution of $0.06\,\mathrm{cm}^{-1}$. Since TES conducts science missions instead of operational missions, it views global swathes one half of the time and operates in a special "stop and stare" mode the other half for events of particular scientific interest. Therefore, IASI collects similar global observations at a temporal rate four times denser than TES when considering both IASI-A and B instruments, but with coarser spectral and spacial resolution.

This retrieval first estimates a vertical profile of OCS on many vertical levels and then averages the levels between 900 and 200 hPa, because the DFS of the profile is less than one with a prior constraint of 20 % OCS variability. Therefore, only one bulk level of OCS is ever distinguishable and even then it is a weighted combination of the true estimate and the *a priori*, which was taken to be spatially flat across all locations. The OCS retrieval is carried out after the routine retrieval of temperature, H_2O, O_3, CO, CO_2, CH_4, surface temperature, emissivity, cloud optical depth, and cloud pressure. Only scenes with a cloud optical depth less than 0.5 are considered in the OCS retrieval as cloudy scenes further reduce the OCS information content. The OCS retrieval itself then jointly includes CO_2, H_2O, surface temperature, cloud optical depth, and cloud pressure in the state vector and uses the posterior covariances from the preprocessed retrieval as the constraints for these extra parameters in the OCS retrieval.

TES retrievals of OCS were validated using the NOAA flask measurements taken at Mauna Loa and the NOAA led HIAPER Pole-to-Pole Observations (HIPPO) air campaigns. In these scenarios they found the average OCS uncertainty to be approximately 7 ppt (1.4 %) with a bias of no more than 15 ppt. However, operating a multi-stage iterative retrieval is an expensive process (ch. 5), in part, because the RTM must be re-evaluated for each iteration per pixel. Therefore, the TES retrieval is currently used forensically rather than routinely.

A monthly mean of TES OCS results from June 2006 was published in *Kuai et al.* [2015], which further validated the concept that direct ocean emissions of OCS are much greater than previously thought [*Berry et al.*, 2013]. The published data included retrievals over ocean between ±40°, because the DFS rapidly fell to values less than 0.5 outside of this range. This means that the majority of the estimates at higher latitudes were dominated by the flat prior OCS field rather than the true OCS concentrations. To put it another way, the uncertainties from an unconstrained retrieval outside of this latitude range would be greater than the prior constraint of 20 % variability. An alternative approach would be to lessen the prior OCS constraint to extend the detectable latitude range, but at the expense of greater uncertainty in the retrieved values. However, the increase in uncertainty can be mitigated by averaging over a greater number of pixels which reduces uncertainty by the square root of the sample size. On the other hand, if the retrieved estimates are mostly *a priori* from tight systemic constraints, then no amount of averaging changes this fact. Nonetheless, the TES product created by *Kuai et al.* [2014] is the current leading retrieval scheme of OCS.

However beautiful the strategy, you should occasionally look at the results.

— Winston Churchill

7

Linear retrieval of OCS

Contents

7.1 Overview

The standard approach to estimating OCS with IASI would be to develop an iterative retrieval similar to that of *Kuai et al.* [2014] (TES) where a forward model, such as the RFM, is rerun every iteration per pixel while testing for convergence around the nonlinear solution. However, even utilising the LUTs developed in Ch. 5, the process of synthesizing a radiance spectrum is computationally expensive by current hardware standards. For example, hundreds to thousands of computation nodes may be required to keep pace with each IASI instrument (80000 pixels per 100 minutes) in an iterative scheme depending upon the forward model. This is indeed possible with dedicated processing centres, but at high cost. Therefore, the work presented here investigates an alternative method of pre-computing radiance spectra and Jacobians for an ensemble of atmospheres and selecting an initial atmosphere that makes the problem most linear. Thus, accepting the increased retrieval error due to residual nonlinearity for an increase in computational speed by a factor of roughly 10^4.

Furthermore, constraints are often necessary in iterative retrievals to prevent a feedback of noise and errors that cause the solution to diverge. For linear (one-step) estimates, constraints upon the solution are not as critical. The OCS estimates presented here are essentially unconstrained and all contaminating parameters that noticeably contribute to the signal are jointly estimated and accounted for. Therefore, this linear OCS scheme is not intended to be the most accurate for each individual pixel, but rather to rapidly process months worth of data on a single machine for analysis of the mean or median values where retrieval noise and highly variable weather events are mitigated by the large sample size.

7.2 Method description

This section describes in detail the proposed method of estimating OCS using IASI observations. The retrieval is a direct application of Eq. (3.2), where the RFM is represented by F and used to model IASI radiances and create Jacobian spectra ($\mathbf{K}$). Rather than use a climatological static mean value for $\boldsymbol{x}_\mathrm{a}$ as the linearisation point, the RTTOV 80 atmosphere ensemble discussed in Ch. 4 was used to create a subsequent ensemble of initial states ($\boldsymbol{x}_\mathrm{a}$), model spectra ($F(\boldsymbol{x}_\mathrm{a})$), and gain matrices ($\mathbf{G}$). The model spectrum that most closely matches the observed IASI spectrum is used to select the initialisation point from the ensemble, which will be discussed further. The point of this process of selecting a model atmosphere from an ensemble is to make the retrieval as linear as possible without iterating the forward model.

Brightness temperature spectra were intentionally used instead of radiance spectra because removing curvature from the Planck function improves the linearity of the retrieval [*Rodgers*, 2000, Ch. 5.1]. The downside to this is that the measurement noise (NEΔT) becomes a function of the observation, as shown in Fig. 2.6. Therefore, the measurement covariances ($\mathbf{S}_\epsilon$) were adjusted specifically for each atmosphere based on the model spectra when computing the gain matrices. Apodization was modelled in the off-diagonal elements of $\mathbf{S}_\epsilon$ according to the discussion in Sect. 3.6.

Additionally, the LUTs discussed in Ch. 5 were used in the radiative transfer modelling by the RFM. While in theory these atmospheres need only be evaluated once and the results saved for future access, in practice the process was repeated numerous times as the selection of parameters was improved. Realistically, these absorption cross section LUTs created for the RFM saved weeks worth of processing time.

Previous work has shown that enhanced atmospheric OCS from extreme events are detectable by IASI [*Clarisse et al.*, 2011]. Further to this, *Shephard et al.* [2009] demonstrated using radiance closure that climatological amounts of OCS are, in fact, retrievable from IASI observations when jointly estimating the additional parameters that contribute to the signal, such as temperature and water vapour. Retrievals can either estimate a single target while using a "total" measurement covariance that shuffles weight away from lines heavily affected by contaminating species or retrieve the target and contaminants jointly where the measurement covariance reduces to the instrument noise. *von Clarmann et al.* [2001] showed that the two are identical in the idealised scenario that the retrieval is linear and the physical covariances of the contaminants are exactly known. Current retrievals of ammonia exploit isolated spectral lines and use the former [*Walker et al.*, 2011; *Van Damme et al.*, 2014], while this work investigates the latter.

7.2.1 Spectral range considered

Identifying OCS spectral features is a straightforward process. Figure 7.1 shows the spectral range targeted in this study ($2000 - 2300\,\mathrm{cm}^{-1}$), including the dominant ν_3 rotational-vibrational band of OCS in the thermal infrared centred at $2060\,\mathrm{cm}^{-1}$.

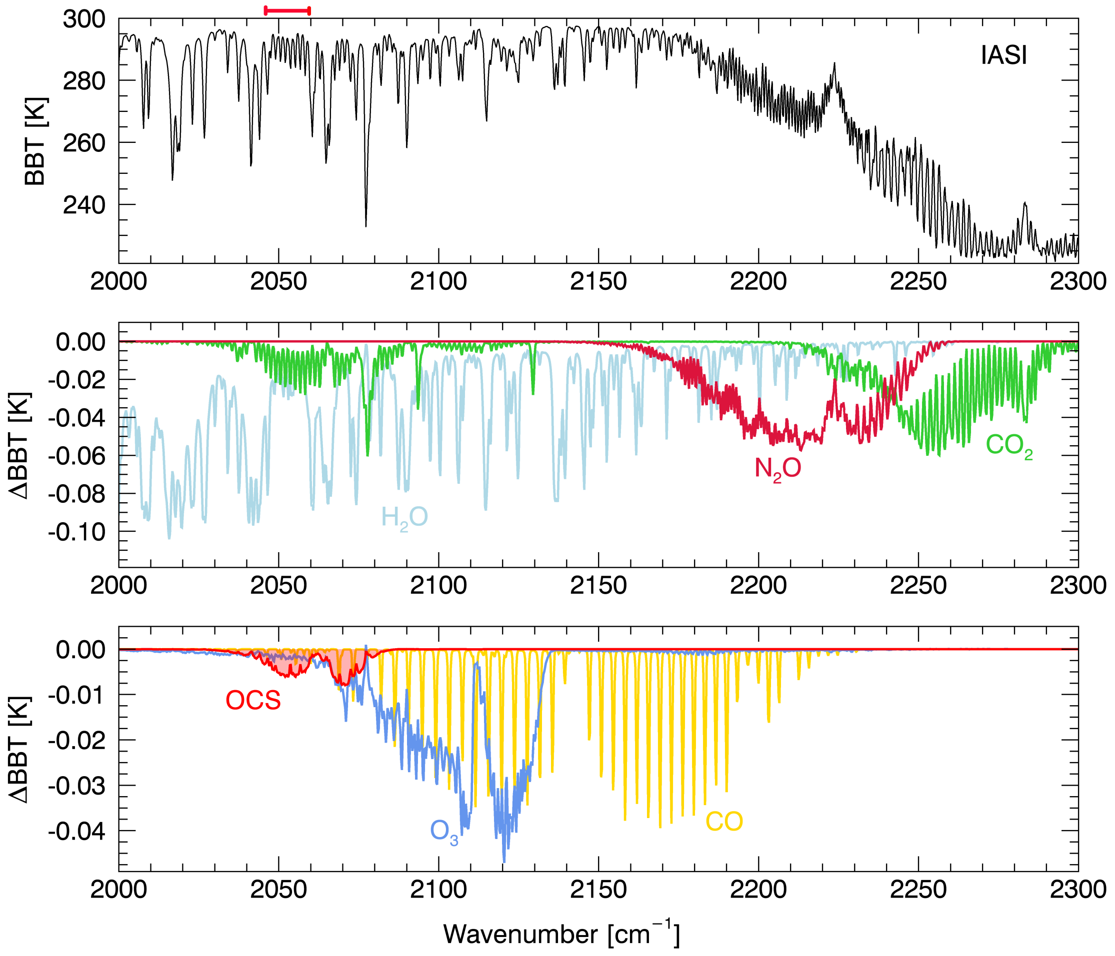
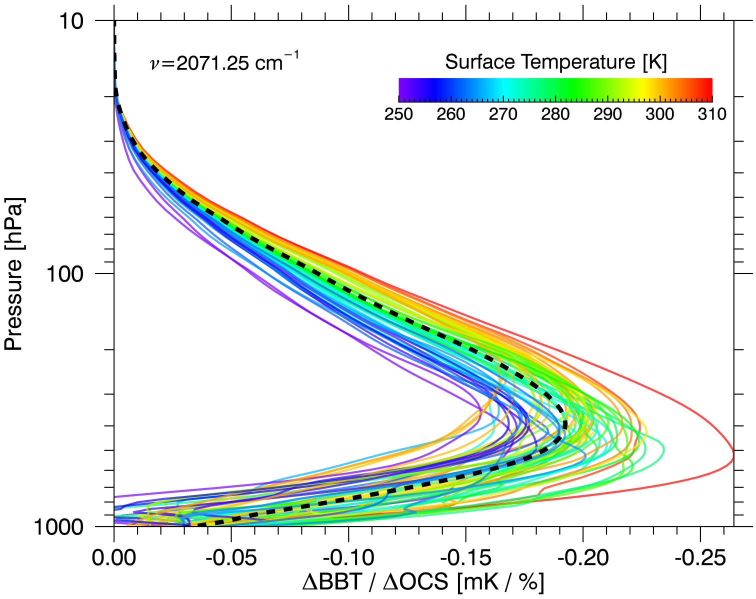


Figure 7.1: Top: A simulated IASI BBT spectrum from a desert (i.e. low humidity) atmosphere covering the spectral range used in the linear retrieval. Middle and Bottom: Jacobian spectra showing the change in BBT for a 1% increase in VMR for the gases listed. The CO_2 and N_2O spectra represent tropospheric perturbations while the remaining four are total column perturbations. The red bar denotes the area between two H_2O lines where a large portion of OCS information comes from.

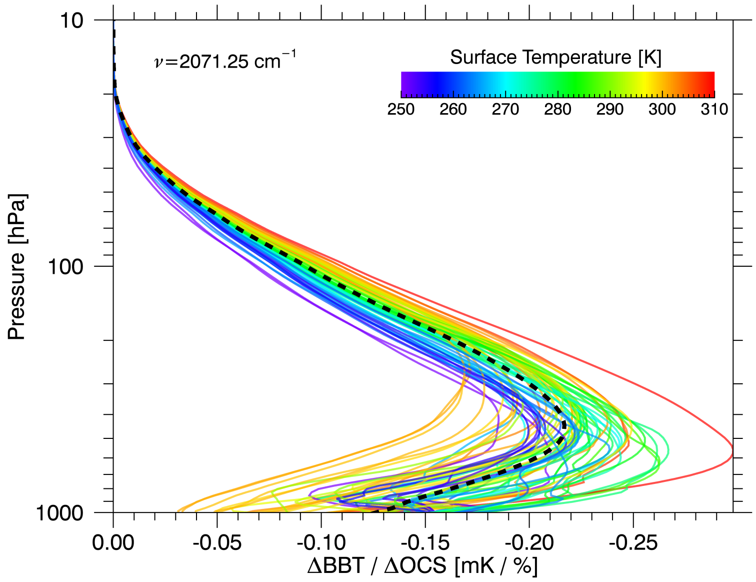
Notice that H_2O and CO_2 are the primary contaminants here with additional contributions from CO and O_3. This also shows there are no isolated OCS spectral lines and that the other detectable species must be accounted for explicitly during the retrieval.

The spectral range included in this retrieval is much larger than the OCS spectral band, which runs from $2040 - 2080\,\mathrm{cm}^{-1}$. This is to provide temperature and contaminating gas information from the spectrum as location specific *a priori* are not used. In particular, the CO_2 and N_2O spectral features are of various line strengths which saturate at different effective altitudes throughout the vertical profile. Since these two gases are well mixed with low natural variability, they provide robust information on atmospheric temperature. In an iterative retrieval, a much narrower spectral region would be used and the additional information would be supplanted by weather specific *a priori* to save time computing the forward model. Since the forward model is pre-calculated in this method, the added spectral range only increases the number of linear algebra operations.

The spectral characteristics of the observation and the applied constraints determine the vertical sensitivity of the retrieval. The weighting functions, i.e., the Jacobian values from perturbing each individual vertical layer, for OCS at the strongest spectral point ($2071.25\,\mathrm{cm}^{-1}$) are shown in Figs. 7.2a and 7.2b. Notice that peak OCS sensitivity with IASI is consistently around 500 hPa for all atmospheres and both surface temperature contrast scenarios. This is consistent with the OCS analyses published in *Shephard et al.* [2009] and *Kuai et al.* [2014]. However, when the surface ground temperature is significantly warmer than the surface air temperature (positive thermal contrast), then lower-most tropospheric OCS becomes up to three or four times more detectable. This is because thermal contrast between the surface and the atmospheric temperature accentuates spectral absorption or emission features, which makes them easier to distinguish.



(a) OCS weighting functions where the ground temperature is 3 K warmer than the surface air temperature.



(b) OCS weighting functions where the ground temperature is 15 K warmer than the surface air temperature.

Figure 7.2: Weighting functions of OCS are shown from vertical layers 1 km thick at 2071.25 cm^{-1} for an ensemble of atmospheres, which are colour coded by the individual surface air temperature. The dashed line represents the mean weighting function of the ensemble.

7.2.2 Defining the state vector and prior covariance

Even though OCS is the desired target, the intent of the joint retrieval is to simultaneously account for all physical parameters that affect the observed spectrum above the noise level. Mathematically this is handled via the extra rows in the gain matrix ($\mathbf{G}$) and the resulting cross-terms when multiplied by $\boldsymbol{y}$. If they are not accounted for, then the other physical parameters become biased into the target estimate. If this were to happen then one could not say with confidence whether an OCS enhancement or depletion was actually due to OCS or something else, such as water vapour or surface temperature.

With this in mind, the state vector is chosen to be

$$\boldsymbol{x} = [\,\mathrm{OCS},\, \mathrm{CO},\, \mathrm{O_3},\, \mathrm{CO_2/N_2O},\, \mathrm{H_2O^1},\, \mathrm{H_2O^2},\, T_s,\, T^1,\, T^2,\, T^3,\, T^4\,]^{\mathrm{T}}, \tag{7.1}$$

where the superscript indexes the vertical location of the retrieved atmospheric layer, as visualised in Fig. 7.3, and the absence of a superscript for a gas implies a total column amount. Specifically, the natural logarithm of the VMRs is retrieved to enforce positivity in all of the gases and dampen the effect water vapour variability may have upon the results. The term T_s represents ground surface temperature. Emissivity is not included in the state vector because the emissivity Jacobian is highly spectrally correlated (> 0.9) with the surface temperature Jacobian and indistinguishable from the other without strict use of *a priori*. Therefore, from analysis of Eq. (5.7) it is clear that the retrieved quantity is effectively $\epsilon_s T_s$ for spectrally grey emissivity. It is important to note that spectrally changing surface emissivity across the range $2040 - 2080\,\mathrm{cm^{-1}}$ is currently not accounted for and may influence the OCS results over land. However, spectral features of solids and liquids tend to be much broader than gases such that a grey approximation may be valid. Another source of error that may be more important than non-grey emissivity is the fact that all atmospheres in the ensemble were modelled with a surface emissivity of 0.99, which neglects downwelling radiation reflected back into the optical path. In both cases, observations over desert will be affected the most with minimal emissivity impact over water and dense vegetation.

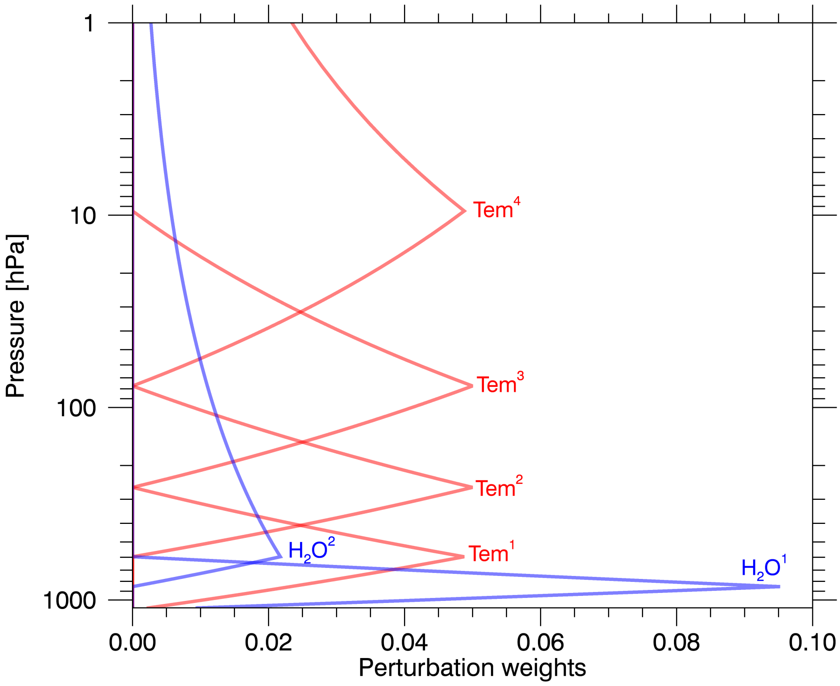


Figure 7.3: The vertical representations of the temperature and water vapour Jacobians are shown. These represent triangular pertubations as opposed to rectangular (even weighted) vertical perturbations.

The ratio of CO_2 to N_2O is included instead of the two separately to improve the conditioning of the inverse problem; which means that there is not enough independent information in the measurement to estimate both gases and atmospheric temperature without added constraints. Whilst N_2O is a low variability gas that does not overlap with the OCS spectral features, the point of including it in the ratio is to account for variations in CO_2 that may affect the OCS estimate. The downside to retrieving a ratio of two gases is that the knowledge of whether the numerator is enhanced versus a depletion of the denominator, and vice versa, is sacrificed for the improved independence of elements in the state vector. Any variability in N_2O due to high stratospheric effects, such as collisional excitation, were not observed in the retrieved estimates and, thus, assumed negligible for these purposes.

As shown in Fig. 7.3, four bulk layers of atmospheric temperature are retrieved ranging from the lower troposphere through the stratosphere. Additionally, two layers of water vapour are retrieved. The first layer is the lower-most troposphere that primarily accounts for water vapour continuum effects between absorption lines in the Jacobian as a result of self-broadening from $H_2O - H_2O$ collisions. The perturbation of the second layer peaks at 600 hPa, but includes contributions from the remaining upward levels of the atmosphere and accounts more for the absorption feature centres.

Since the statistical distributions of temperature and water vapour vertical profiles are well known, the resulting estimates can be constrained to scenarios found on Earth where clearly unphysical profiles are excluded with a negligible loss of sensitivity. Furthermore, since atmospheric temperature and water vapour are physical correlated, it is possible to represent this effect in the prior covariance. Thus, the 80 atmosphere ensemble was vertically binned down to the bulk layers of the retrieval and used to calculate the sample covariance matrix, which includes the cross-state physical correlation terms. The subsequent correlation matrix is shown in Fig. 7.4 with the standard deviations annotated along the diagonal elements. This sample covariance is then used as a sub-matrix (6×6) within the prior covariance (11×11) to constrain the water vapour and temperature portion of the retrieval to physical values within the global range. Further correlation terms between the remainder of the state elements are assigned to values of zero.

As a caveat, all elements of the state vector, including OCS, are technically constrained with finite values in the diagonal of the prior covariance. This is primarily for the purposes of developing a test-bed iterative retrieval that utilizes the Levenberg-Marquardt method, which will be discussed next. OCS variability in the prior covariance is assigned to be 200%. CO and O_3 variability is assigned 100%, the CO_2/N_2O ratio is set to 10%, and surface temperature 20 K. However, this is such a weak constraint that the DFS for the OCS total column is close to one for all atmospheres and, therefore, effectively unconstrained.

7.2.3 Parameter validation using an iterative retrieval

Validation of the retrieval framework, as previously defined, is crucial towards developing confidence in the resulting estimates. Without analysing external data,

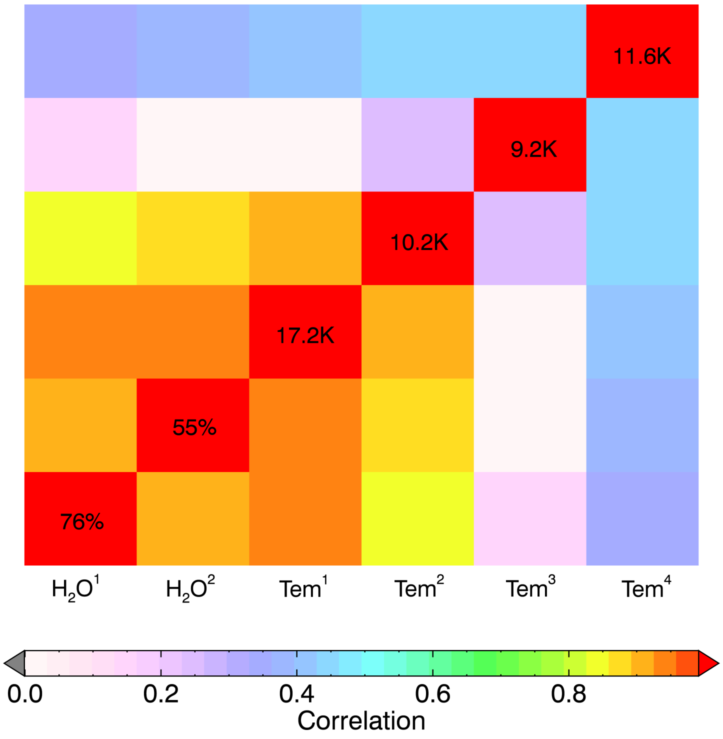


Figure 7.4: The correlation matrix is shown for the sample covariance of the H_2O and atmospheric temperature layers calculated from the 80 atmosphere RTTOV ensemble. The standard deviations are annotated along the diagonal elements for reference.

one can show using an iterative retrieval that:

1. The estimates converge during iteration.
2. The OCS spectral signature is noticeable in the residual spectrum of the converged result when excluded from the state vector and all other parameters are retrieved.
3. The variability of the converged residual spectrum over many pixels is similar to the expected instrument noise.

Each point is discussed in turn.

The iterative retrieval was written as a test-bed for the faster linear scheme; so the spectral range, state vector, and prior covariance are the same as previously defined. This nonlinear approach is based on the Levenberg-Marquardt method as shown in Eq. (3.23). The prior state is taken to be equal to the initial state, which is selected on a pixel-by-pixel basis from the ensemble of atmospheres. Each model atmospheric spectrum is compared against the measured spectrum and the j^{th} atmosphere that minimizes the spectral cost, i.e.,

$$\chi_j^2 = \left[\boldsymbol{y} - F\left(\boldsymbol{x}_j\right)\right]^{\mathrm{T}} \mathbf{S}_\epsilon^{-1} \left[\boldsymbol{y} - F\left(\boldsymbol{x}_j\right)\right], \tag{7.2}$$

is chosen as the starting point. For atmosphere selection, only the diagonal of $\mathbf{S}_\epsilon$ is used to save computation time when rastering through the 80 atmospheres. Scenes with an AVHRR cloud fraction greater than 20% are not included. Based on this methodology it was found that the majority of IASI pixels converged on a result that reduced the χ^2 cost function. In reality, this test for convergence was repeated numerous times as the state vector and prior covariance were modified until settling on the parameters defined in the previous section.

OCS signatures can be shown in the converged residual spectrum (IASI minus RFM) if all other contributing parameters are retrieved. This is done by removing OCS from the state vector while retrieving the other 10 in its absence. Figure 7.5 shows an example of this for an IASI pixel in the North Atlantic off the coast of Iceland where the retrieved surface temperature is 281 K. Notice that the OCS spectral signature is clearly above the IASI noise level for a particularly low surface VMR estimate of 404 ppt and matches well to the predicted OCS residual of the same VMR. It is important to keep in mind that Fig. 7.5 is for a single pixel without any spectral averaging to reduce instrument noise. Also apparent, is a substantial spike in the residual centred at 2077 cm^{-1}. This feature is presumably due to line mixing errors within the RFM for the CO_2 Q-branch located at this position. Therefore, these particular channels should be avoided as they are poorly modelled.

Once all physical parameters that contribute to the signal above the noise level are accounted for through the joint retrieval, then the standard deviation of the spectral difference between the observation and the model, i.e., the residual, should be equal to the instrument noise. If this is not the case, then any parameters

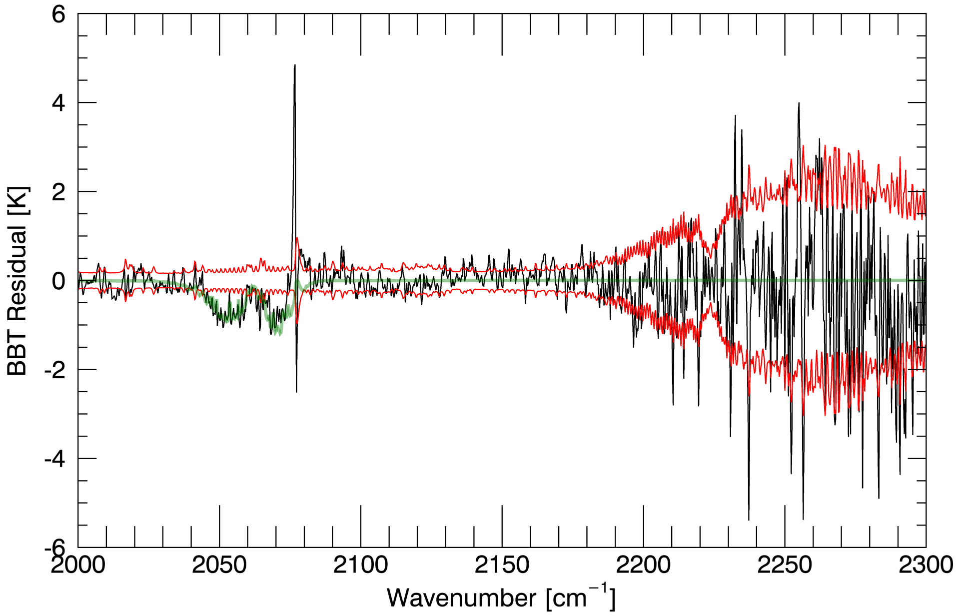


Figure 7.5: The residual spectrum between the IASI observation and the converged estimate from the RFM without modelling OCS is shown as the black line. The red line depicts the instrument noise level specific to the observed surface temperature of 281 K and atmospheric temperature profile. The green line represents the expected OCS signal in the residual for the retrieved VMR of 404 ppt.

that are not completely accounted for will show an associated spectral feature in the standard deviation of the residual spectra. To test this posit, the iterative retrieval was run over 600 pixels in a 10° × 10° latitude and longitude box in the Equatorial Pacific Ocean. The variability of the sample residuals is shown in Fig. 7.6 along with the average instrument noise profile in units of brightness temperature. Observe that the variability of the residuals matches closely to the average instrument noise with the exception of a few spectral features due to water vapour (Fig. 7.1). Therefore, the retrieval and associated state vector sufficiently account for the noticeable physical parameters aside from water vapour, which could be further resolved with more vertical levels along the profile.

There are three options to pursue with regards to unresolved, but influential, H_2O levels in the retrieval. Firstly, this effect could be tolerated as an unaccounted source of error in the OCS estimates. Secondly, additional layers of H_2O could be

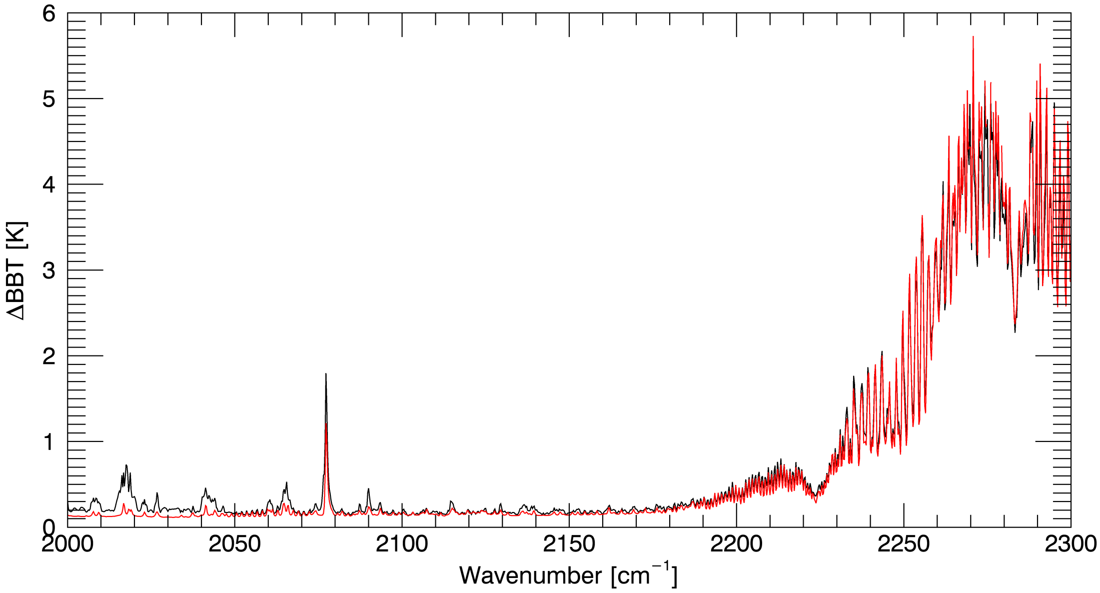


Figure 7.6: Black: The sample standard deviation of the residual spectra between the IASI measurements and the converged model spectra for an ensemble of 600 pixels from the tropical Southern Pacific Ocean. Red: The average instrument noise (NEΔT) for the IASI observations.

included in the state vector and jointly retrieved with an updated prior constraint. Thirdly, these associated H_2O spectral features could be treated as effective noise within the measurement covariance, thus decreasing the sensitivity of the retrieval to variations in the water vapour vertical profile.

The first option is undesirable because there is clearly evidence supporting further treatment of H_2O. At first attempt, three layers of H_2O were included in the state vector with a new prior covariance derived from the 80 atmosphere ensemble. However, it was found that this formulation did not converge unless a much stronger prior constraint was constructed. Therefore, these spectral variations for H_2O were instead treated as noise by creating a vector of scaling factors that increased the variances of the measurement covariance accordingly. This was accomplished by taking the ratio of the variance of the residual spectra over the square of the IASI instrument noise and setting any values less than one to unity. Thus, making the retrieval less sensitive to unretrieved layers of water vapour. All estimates of OCS from this point further include scaled variances in the measurement covariances due to water vapour variability.

7.2.4 Channel selection

Spectral channels in remote sensing tend to be highly correlated, not only by the gas specific rotational-vibrational energy transitions, but through other physical effects such as temperature and pressure. In other words, each channel does not normally add independent information and contains a certain amount of redundancy. In theory, adding more channels to the estimate always increases the total information content to varying degrees. In practice, there are spectral channels that contain more information than others such that adding channels of negligible importance does little to improve figures of merit (Sect. 3.2), but increases sensitivity to unaccounted physical parameter errors. One method to improve the robustness of a retrieval by reducing sensitivity to unaccounted parameters is to select a subset of spectral channels that contains the majority of information while excluding the remaining channels that negligibly contribute.

Channel selection was performed over the $2000 - 2300\,\mathrm{cm}^{-1}$ range in order to remove these spectral channels of little importance. One option is to remove channels while maximizing a figure of merit for the joint retrieval as a whole. Another is to maximize just the OCS portion of the retrieval at the expense of the other retrieved parameters. Since the other states are included just to improve the OCS estimates, the latter option is chosen here.

OCS is so weakly constrained that attempting to maximise the DFS is not appropriate in this case. In the unconstrained case, the DFS is not defined for maximum likelihood estimates. However, it is always desirable to minimize the posterior uncertainty, whether constrained or not. In this case, just the uncertainty component of OCS is considered:

$$\hat{\sigma}^2_{\mathrm{OCS}} = \hat{\mathbf{S}}^x_{1,1}, \tag{7.3}$$

where $\hat{\mathbf{S}}_x$ is defined in Eq. (3.3) and the subscript index denotes the first diagonal element corresponding to OCS.

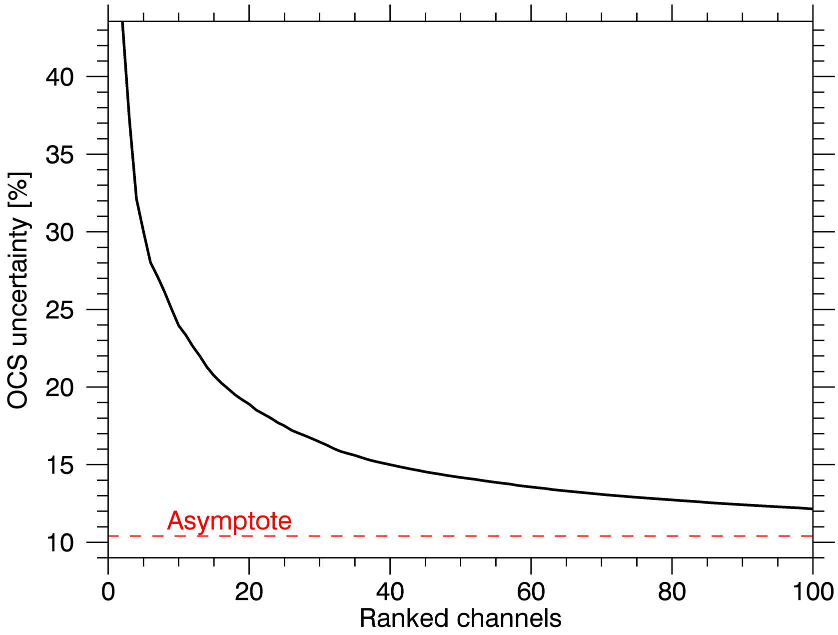


Figure 7.7: Ranked spectral channels are shown for a mid-latitude atmosphere based on their contribution towards minimizing the posterior uncertainty of OCS. The asymptote from including all 1201 channels is shown as the dotted red line.

The selection begins by first finding the best two spectral channels that minimise $\hat{\sigma}^2_{\mathrm{OCS}}$ after calculating all possible two channel combinations. Then a third channel is selected by adding all remaining channels individually and choosing the one which reduced $\hat{\sigma}^2_{\mathrm{OCS}}$ the most. This process is repeated until all spectral channels have been ranked according to their contribution towards minimizing the posterior uncertainty of OCS. The resulting channel ranking for a mid-latitude atmosphere is visualised in Fig. 7.7 where the best two channels estimate OCS uncertainty to be nearly 50% while including all 1201 channels reduce the uncertainty to just over 10%. Notice that the first 20 channels reduce uncertainty by a factor of two from the initial pair, but it takes the remaining channels to gain another factor of two reduction. For this retrieval the top 100 channels were retained, which yield an uncertainty of just 12% (versus 10%) for this particular atmosphere with 12 times fewer channels.

The resulting selected channels are shown in Fig. 7.8 for reference. Channels are selected from this method covering the entire spectral range, rather than just

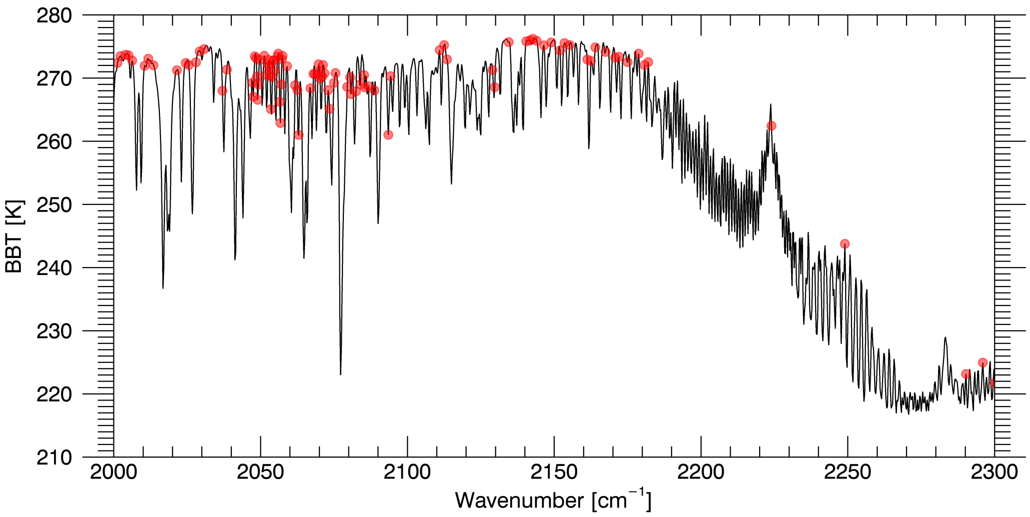


Figure 7.8: The top 100 spectral points (red circles) ranked in Fig. 7.7 from the channel selection are shown for reference as compared to all 1201 channels from the observed spectrum (black line) for a mid-latitude atmospheric scenario.

the 40 cm^{-1} OCS interval, because these outside channels contribute to the other 10 parameters jointly estimated that help improve the OCS retrieval. Channels are only selected in so far as they contribute to better OCS estimates. The CO_2 Q-branch at 2077 cm^{-1} was avoided by heavily penalizing these channels within the measurement covariance prior to running the selection. Notice that the selected channels largely avoid the majority of H_2O absorption features and frequently select the between band channels associated with water vapour continuum.

7.2.5 Selecting the initial atmosphere

The validity of the linear retrieval is contingent upon the choice of initial atmosphere. The initialisation point should be sufficiently close enough to the observed atmosphere that a single step places the estimate within the uncertainty level of the true state being observed. Failure to do so results in retrieval error due to the nonlinearity of the formulated problem. So how should an initial atmosphere be selected in order to minimize the nonlinearity error? Three possible techniques are analysed for determining the initial atmosphere that do not require rerunning the RTM.

1. Select the initial atmosphere whose model spectrum minimizes the spectral cost function in Eq. (7.2), as previously discussed. This method essentially picks the atmosphere whose mean spectrum is closest to the IASI observation for the selected spectral channels. The diagonal may be used to approximate $\mathbf{S}_\epsilon$ to speed up the process of running through the 80 atmospheres for each pixel since the selected channels contain few adjacent pairs.

2. Another method is to estimate what the model spectrum would be after the retrieval, within the linear framework of the problem, and then select the atmosphere which minimizes the projected spectral cost. The retrieved state can be linearly projected back into spectral space to estimate the posterior spectrum,
$$\hat{\boldsymbol{y}} = \mathbf{K}\left(\hat{\boldsymbol{x}} - \boldsymbol{x}_0\right) + \boldsymbol{y}_0. \tag{7.4}$$
If $\hat{\boldsymbol{y}}$ is substituted for $F(\boldsymbol{x})$ in Eq. (7.2) and $\hat{\boldsymbol{x}}$ is expanded using Eq. (3.2), then the resulting projected cost is given by
$$\chi^2_{pr} = \left(\boldsymbol{y} - \boldsymbol{y}_0\right)^{\mathrm{T}} \left(\mathbf{KG} - \mathbf{I}\right)^{\mathrm{T}} \mathbf{S}_\epsilon^{-1} \left(\mathbf{KG} - \mathbf{I}\right) \left(\boldsymbol{y} - \boldsymbol{y}_0\right). \tag{7.5}$$
In some texts **KG** is referred to as the Data Resolution Matrix (DRM) and, unlike **GK**, is generally not equal to the identity matrix in the unconstrained least-squares retrieval.

3. Finally, the third method considered is to train a vector operator to predict the non-linear error in OCS based upon the spectral difference between the initial model and measurement spectra. To do this, all possible permutations ($80 \times 79 = 6320$) of using one state from the 80 atmosphere ensemble as the initial point to retrieve another atmosphere from the ensemble are calculated to yield two matrices; an array of initial spectral differences ($\boldsymbol{\Delta}\mathbf{BBT}$ of size $6320 \times m$) and a vector of corresponding linearly retrieved OCS errors ($\delta\mathbf{OCS}$ of size 6320×1). The goal is to determine a prediction vector ($\boldsymbol{a}$ of size $m \times 1$) that approximates the following equation:
$$\delta\mathbf{OCS} = \boldsymbol{\Delta}\mathbf{BBT} \times \boldsymbol{a}. \tag{7.6}$$

> However, since there are only 80 independent atmospheres considered, the dimensionality of the problem must be reduced if Eq. (7.6) is to be successfully inverted to find $\boldsymbol{a}$. Therefore, $\Delta\mathbf{BBT}$ is decomposed into singular vector components, $\Delta\mathbf{BBT} = \mathbf{U}\Lambda\mathbf{V}^{\mathrm{T}}$, where $\mathbf{U}$ and $\mathbf{V}$ are the left and right singular vectors, respectively, and Λ is a diagonal matrix of its singular values. The inner dimensions of $\mathbf{U}$ and $\mathbf{V}^{\mathrm{T}}$ are then ranked in order of decreasing singular values and truncated at 79. Equation (7.6) is then recast as
>
> $$\mathbf{U}^{\mathrm{T}}\,\delta\mathbf{OCS} = \mathbf{U}^{\mathrm{T}}\,(\Delta\mathbf{BBT})\,\mathbf{V} \times \boldsymbol{a}', \tag{7.7}$$
>
> where the truncated least-squares solution to $\boldsymbol{a}'$ is calculated. Finally, the prediction vector is found to be $\boldsymbol{a} = \mathbf{V}\,\boldsymbol{a}'$.

These three initial atmosphere selection methods are compared using the RTTOV ensemble in the absence of instrument noise and contaminating parameters so that the error due solely to non-linearity is assessed. Each atmosphere of the 80 is used as a test case where the objective is to select an initial atmosphere from the remaining 79 which minimizes the error in the estimate while knowing the true OCS model value. In the ensemble all atmospheres contain the same OCS profile, because of the lack of information about its distribution and variability. The model OCS profile is 590 ppt at the surface, constant up to the tropopause, and steadily decreasing with altitude through the stratosphere. Therefore, the best initial atmosphere yields a retrieval step closest to zero, because OCS is a flat field throughout the model atmospheres.

Figure 7.9 shows histograms of the linear assumption error for the three selection methods discussed. Furthermore, a method of randomly selecting the initial atmosphere was also analysed to provide a baseline for comparison. Evidently, the method that most frequently yielded retrieval errors near zero was from selecting the atmosphere that minimized the projected cost. This was followed by matching the mean spectrum (minimizing the initial spectral cost). Predicting the retrieval error worked in the sense that it outperformed the baseline of random selection, but provides larger error than the other two methods.

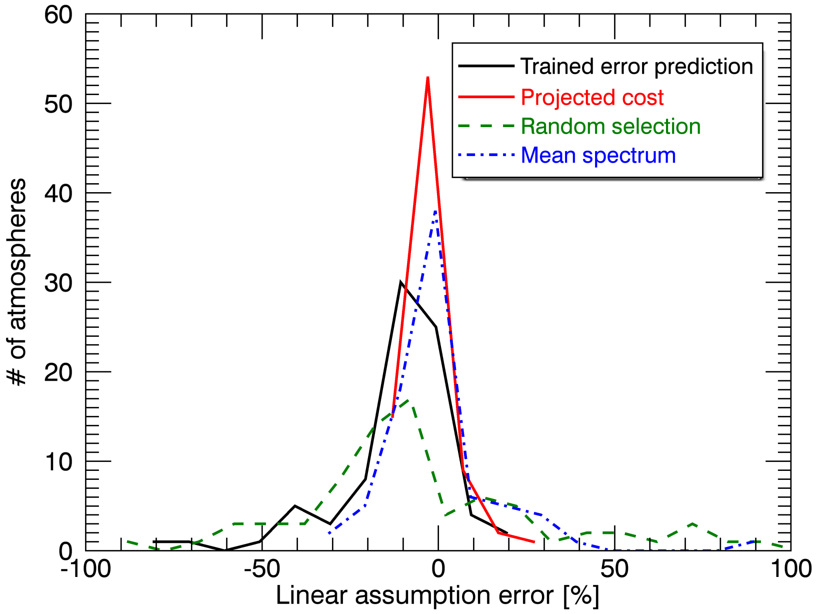


Figure 7.9: Histograms of the OCS retrieval error due to non-linearities are shown for four methods of selecting the initial atmosphere, $\boldsymbol{x}_0$. The first enumerated method in this section is labeled 'mean spectrum', the second 'projected cost', and the third 'trained error prediction'.

One may conclude that the initial atmosphere should be selected based upon minimizing the linearly projected cost. However, this was attempted with real data and in practice it became clear that a grossly non-linear starting point, such as using an initial polar atmosphere for an observation in the tropics, may occasionally be projected to outperform all other atmospheres. This is because linear analysis is only valid in the nearly linear to moderately linear regimes. Therefore, the method of selecting an initial atmosphere by minimizing the difference in the mean spectra, Eq. (7.2), was used in the work presented here because it avoids this particular problem.

The expected non-linearity error is given by the width of the distributions in Fig. 7.9 and the mean spectrum method was found to have an error of 11% on average. This analysis was also performed using all spectral channels, i.e., without channel selection, and found to yield an average error of 19% for this method.

Thus, channel selection is crucial towards improving the OCS retrieval, because it makes the problem almost twice as linear.

7.2.6 Geographical considerations

Lower-most tropospheric pressure is influential in the OCS retrieval not just through the direct effect of pressure broadening the spectral features near the surface, but also because of pressure dependent water vapour continuum effects in the lower troposphere that overlap with all OCS spectral lines. Surface pressure variations due to geographical altitude must, therefore, be accounted for in some way. If not, then the Jacobians and initial column amounts will misrepresent the observation, especially over mountain ranges and high plateaus.

To do this, separate atmospheric ensembles of model spectra, gain matrices, and initial values were created for surface pressure scenarios of 1030, 900, 800, and 700 hPa. Average surface pressure was tabulated from ECMWF reanalysis data and stored as a reference field. Prior to computing the linear estimates of OCS, the surface pressure for each IASI pixel based on its latitude and longitude is interpolated from the saved map. Then the appropriate ensemble is selected based upon whether the interpolated surface pressure falls outside or within the bounds of 950, 850, or 750 hPa. Applying this method noticeably removed any high terrain artefacts that systematically appeared in the OCS estimates.

Additionally, temperature contrast between the ground surface and lowest atmospheric layer affects the sensitivity of the OCS estimates, as shown in Figs. 7.2a and 7.2b. Thermal contrast is a particular problem over land, and especially deserts in the summer, where the surface is heated by solar absorption to values occasionally greater than 15 K above the atmospheric surface temperature. In the deep Antarctic, there can be a negative thermal contrast where the surface is actually colder than the atmosphere and absorption lines switch to emission features. This effect is far less important over the oceans, because the heat capacity of water is so great that thermal contrast tends to be slightly positive with less variability.

Therefore, the method employed in this work is to treat IASI observations over ocean as having a routine thermal contrast of +3 K, while allowing for greater variation over land. Instead of selecting from 80 atmospheres over land with one thermal contrast option, the ensemble is grown to include scenarios of -5, 3, 10, and 15 K of thermal contrast. So an observation over land has 320 possible atmospheric initialization points to select from. As previously mentioned, the model atmospheric spectrum that most closely matches the observed spectrum determines which atmosphere is selected as the initial point.

7.2.7 Quality filtering

In an iterative retrieval, high confidence in the estimate is obtained by verifying that the retrieval converged on a minimum χ^2 value. This may not be the correct minimum, but the fact that a minimum was found suggests that the framework of the problem is behaving in a consistent way. In a one-step linear retrieval the forward model is not recalculated for each individual pixel in order to save computation time. Incidentally, other metrics of quality must be evaluated in order to identify and exclude retrievals that have likely gone awry. The steps to filter the OCS estimates for quality are described in detail.

First, any IASI pixels with an AVHRR cloud fraction of 20% or greater are excluded from consideration prior to computing the retrieval. The presence of cloud introduces highly non-linear behaviour that must be modelled properly if the OCS estimates are to be trusted. This AVHRR cloud fraction product is not perfect and routinely flags sea ice as cloud. However, the vast majority of the time it provides a robust and accurate estimate of the amount of cloud filling the IASI pixel. Therefore, cloudy scenes are simply avoided in favour of clear sky observations.

Next, viewing angles noticeably affected by sun glint are excluded from the retrieval by calculating the solar reflection angle in Eq. (2.4) and removing pixels where this angle is less than 18°. Additionally, there is a slight overestimation of OCS when observing towards the limb. Rather than attempting to parametrise or mitigate this effect, observations with an air mass factor relative to nadir greater than

1.47 are avoided. This removes the very far edges of the IASI scan where the surface zenith angle is greater than 47°. For surface zenith angles less than this value, limb effects were not noticeable. Fortunately, the overlap of IASI-A with IASI-B is greater than this angular width, so no spatial gaps in coverage are introduced as a result.

Since the retrieval jointly estimates other physical parameters in conjunction with OCS, there is further opportunity for common sense filtering for quality. For example, if the retrieved surface temperature falls outside of the range between $230-340\,\mathrm{K}$, then that pixel is removed from consideration. Furthermore, if the lowest level of retrieved water vapour has a VMR greater than 4%, then the observation is clearly not represented properly and those OCS estimates are excluded.

Finally, the projected spectral cost from Eq. (7.5) can be used as a retrieval diagnostic given the fact that the atmosphere with the smallest initial spectral cost was selected as the initialization point. The expectation value of the projected cost should be approximately equal to the number of spectral channels ($m = 100$) if the retrieval were ideally linear. Since the problem is not linear, the average projected cost will certainly be greater than m. However, the magnitude of the projected cost provides a useful prediction as to how well the retrieval may perform. Thus the normalized criteria for accepting a retrieved pixel is given by

$$\frac{\chi^2_{pr}}{m} < 2. \tag{7.8}$$

Aside from filtering against cloudy scenes, this provides the strictest quality test of those mentioned and highlights geographical areas that are poorly represented by the modelled atmospheric ensemble.

7.3 OCS results from 2014

The entirety of IASI-A and B data from 2014 (19.4 Tbytes) were downloaded and processed in this study using the previously described linear retrieval technique. OCS total column median values are shown in Figs. 7.10–7.15 for two-month intervals in latitude-longitude bins of 0.5° × 0.5°. Median OCS values combined from all data in 2014 are shown in Fig. 7.16. The median was chosen instead of the mean, because the retrieval actually estimates the logarithm of the total column to enforce positivity and when raised to the exponential introduces positive skewness into the distribution of estimates. In other words, the spread of OCS estimates does not follow a Gaussian (normal) distribution and the heavy tail towards overestimation is mitigated by taking the median rather than the mean. The median also dampens the effect anomalous cases have upon the statistics of the distribution, whereas one bad pixel resulting in a wildly high or low total column amount could artificially dominate the mean.

The number of pixels per bin passing the quality and cloud free criteria is also shown for reference. Only bins containing three or more observations are shown and any areas with two or less observations are considered missing and coloured grey. Areas that are systematically low in number of observations are either routinely flagged as cloudy or routinely predicted via the projected cost to poorly model the observation. Notice that areas of sea ice towards the poles are consistently absent, likely due to AVHRR cloud flagging. However, persistent glaciers over land contain many more observations and do not experience this false-positive cloud flagging effect. Alternatively, desert areas during the day-time in local summer are frequently cloud free and marked as such, but routinely fail the quality check and contain few estimates. This signifies that the model atmospheres in the ensemble fail to closely match summer desert scenarios that are sun illuminated, perhaps because of lower surface emissivity that increases solar and downwelling reflections that are currently not modelled.

The sample standard deviation of OCS per spatial bin over the two month period is shown in the bottom row of Figs. 7.10–7.15. This gives an estimate of the width of the OCS distribution based upon the sampling of retrieved values. In an iterative retrieval, the posterior uncertainty from $\hat{\mathbf{S}}_x$ is normally used to represent the error of the retrieval. However, within the linear framework of this method the posterior uncertainty derived from the initial guess will systematically underestimate the true error of the OCS retrieval. Thus, the sample standard deviation provides a metric that is a combination in quadrature of retrieval noise, natural OCS variability, and errors due to unaccounted parameters. At a minimum, the sample standard deviation of OCS will be no less than the retrieval noise, assuming there is a sufficient number of samples. Areas that are clearly dominated by retrieval noise are the Antarctic plateaus, Greenland, and high latitude land in the Northern Hemisphere during winter.

7.3.1 Estimates over ocean

Beginning with the oceans, there is a clear correspondence of OCS estimates observed between day and night. Prior to filtering based on the solar reflection angle, it was apparent that sun glint was an issue for estimates over water, especially near the equator. However, by excluding observations along the specular path this issue was mitigated such that the day and night estimates resemble each other. This is the expected result because variations in thermal contrast from the day to night over water should be fairly small. Therefore, OCS should be equally detectable over water regardless of the time of day.

OCS estimates throughout the year show that there is a consistent feature of elevated OCS in the South Pacific off the coast of South America between 0 and −30° latitude that matches well to the direct OCS emissions modelled in Fig. 6.2. While there is some variation throughout the year, this particular feature remains relatively constant regardless of season. In contrast, further South there is a large OCS signal that appears to align with the Antarctic Circumpolar Current (ACC) in both day and night observations. This particular feature shows a large seasonal variation with maxima occurring during southern hemisphere summer and minima during winter when incident solar radiance is low. This is the seasonal cycle one would predict if the primary source of OCS were photochemically reduced CDOM. Before too much is concluded, it is important to acknowledge the fact that this OCS signal at −60° latitude may be a false positive resulting from a temperature artefact specific to the ACC. However, it is certainly worth further investigation.

Northern hemisphere ocean areas appear to have maximum OCS signal between March and June (local spring) with minimum values as the season approaches winter. Once again this is consistent with how the incident solar radiation varies with season for photochemical production. OCS features that particularly stand out in these areas are the tropical enhancement during May to June coming off the coast of Baja California and the high latitude structures south of Greenland and north-east of Iceland in this same time period. Additionally, there appears to be a consistent enhancement of OCS in the northern Indian Ocean by the Arabian Peninsula, which also resembles the model in Fig. 6.2.

Interestingly, there is an OCS feature over the Pacific Ocean between Japan and Alaska that is in phase, but one month delayed, with the high OCS signal over the east of China and the Tibetan Plateau. This ocean feature begins in January and February, reaches maximum in March and April, and then dissipates by August. Whereas the OCS land signal over China grows substantially in November and December and then is closer to background levels in May and June. One possibility is that the enhancement over the ocean between Japan and Alaska is an OCS plume originating from China transported by the easterly zonal winds that dissipates when OH concentrations increase during spring and summer. On the other hand, the two signals may be purely coincidental and indicative of two unrelated sources of OCS or other atmospheric characteristics that produce artificially high estimates over these regions.

7.3.2 Estimates over land

Satellite retrievals over land are subject to a greater number of surface type variations than over ocean. As a result, there are more variants contributing to the signal that may require a modelled response; such as emissivity, altitude, surface facets, BRDF, and snow cover. Therefore, one must analyse spatially sharp OCS gradients over land coinciding with geographical features and overly distinct land-sea boundaries with a certain amount of scepticism.

Recent work by *Glatthor et al.* [2015] and *Berry et al.* [2013] have shown that there should be a noticeable depletion of OCS over the Amazon and Congo rainforest areas due to strong vegetative uptake. This is indeed what is observed in these data, especially for the observations made during the day. The Amazon and Congo areas show OCS total columns approximately 10–20% less than what is estimated over nearby oceans at the same latitude. Therefore, these results are consistent with the idea that vegetative uptake is a significant sink of OCS. However, notice that the night-time estimates tend to be slightly greater than the day-time, which may be indicative of a physical OCS process with a diurnal signal. It is also possible that this effect is an artefact of the retrieval. One may quickly blame thermal contrast between day and night observations; except that it is the wrong way around from what is expected. For example, areas over desert like the Sahara and much of Australia show low OCS at night and higher OCS during this day. This is because solar heating increases thermal contrast, which makes trace gases more detectable. During the night, these low humidity areas quickly radiate away their heat and come closer to thermal equilibrium between the surface and the lower atmosphere, thus decreasing sensitivity to OCS. Therefore, if the higher night-time OCS signal over the rainforests is not physical, then it is unlikely to be solely due to thermal contrast.

Along this same vein, notice that the high latitude areas over land near the Arctic show substantially less OCS in the Winter months than at any other time of year. The high standard deviations of the estimates during this time show that these low OCS values may be due to a loss of detectability as the signal drops from cold temperatures over land. The same can be said for estimates over Greenland and most of Antarctica throughout the entire year, i.e., the SNR of OCS is too low to have much confidence in retrievals over these areas. However, sensitivity appears to return for estimates over northern Canada and Russia during spring, summer, and fall.

Additionally, there are several areas over land where there are particularly high OCS signals. Much of the continental United States show OCS estimates greater than ocean values at similar latitudes. The United States OCS signal appears to be maximum during March to April and minimum during July and August with a slow build up back to March. If these estimates are indicative of the true OCS levels, then the July to August minimum coincides with peak vegetative uptake for regions at this latitude. Sources of OCS in the United States, especially anthropogenic and biomass burning, are currently poorly understood.

Many regions in the Middle East and the north African Mediterranean coast also show very specific enhancements of OCS estimates. It is possible that there exists a surface emissivity feature in these regions that routinely yields spurious elevated OCS values. However, some of this effect is likely mitigated by the process of calculating the projected cost of the retrieval and removing pixels where the model initial atmospheres are predicted to poorly represent the scene. Therefore, it may also be possible that these signals are real and there are large sources of OCS creating local enhancements. If this signal represents physical OCS amounts, then the source is more likely to be anthropogenic in nature given that the detail closely follows geographical boundaries of human population.

Finally, the areas of high OCS signal over China and the former Soviet republics east of the Caspian Sea especially stand out in displayed estimates. These are also areas of known SO_2 emissions due to industrial processes and energy production that are routinely modelled in chemical transport models, such as TOMCAT [*Spracklen et al.*, 2005]. While it is energetically unfavourable for SO_2 to convert to OCS, the two may be positively correlated in many physical situations, especially in anthropogenic processes that do not have strict methods in place to reduce SO_2 emissions.

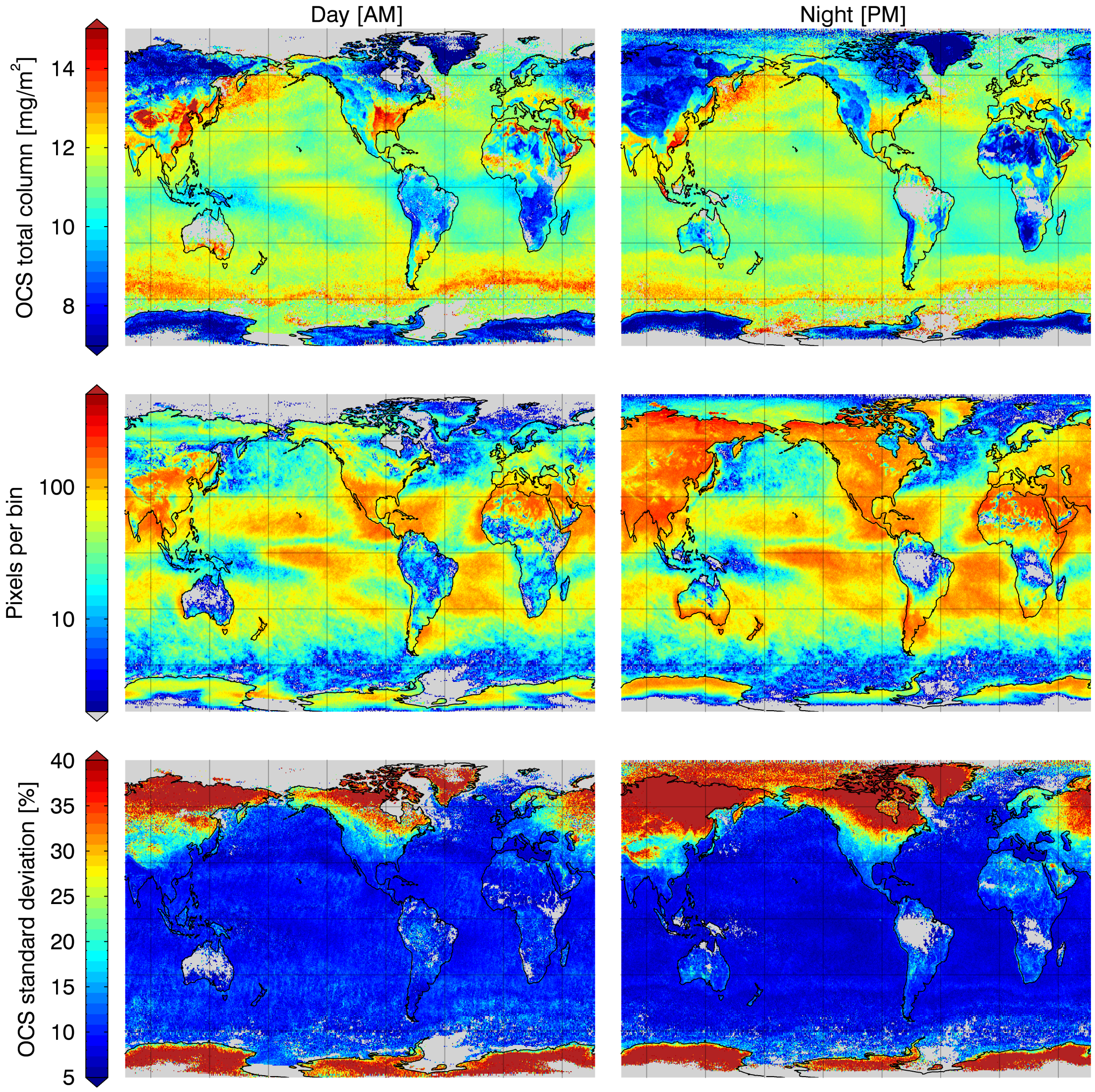


Figure 7.10: January–February 2014: Linear estimates of total column OCS median values are shown in the top row for sun illuminated morning (left column) and night-time evening (right column). The results are binned by latitude-longitude widths of 0.5° × 0.5°. The middle row shows the number of pixels per spatial bin that passed the quality control checks. The bottom row shows the sample standard deviation of OCS per bin for the two month interval. Spatial bins with missing data are coloured grey.

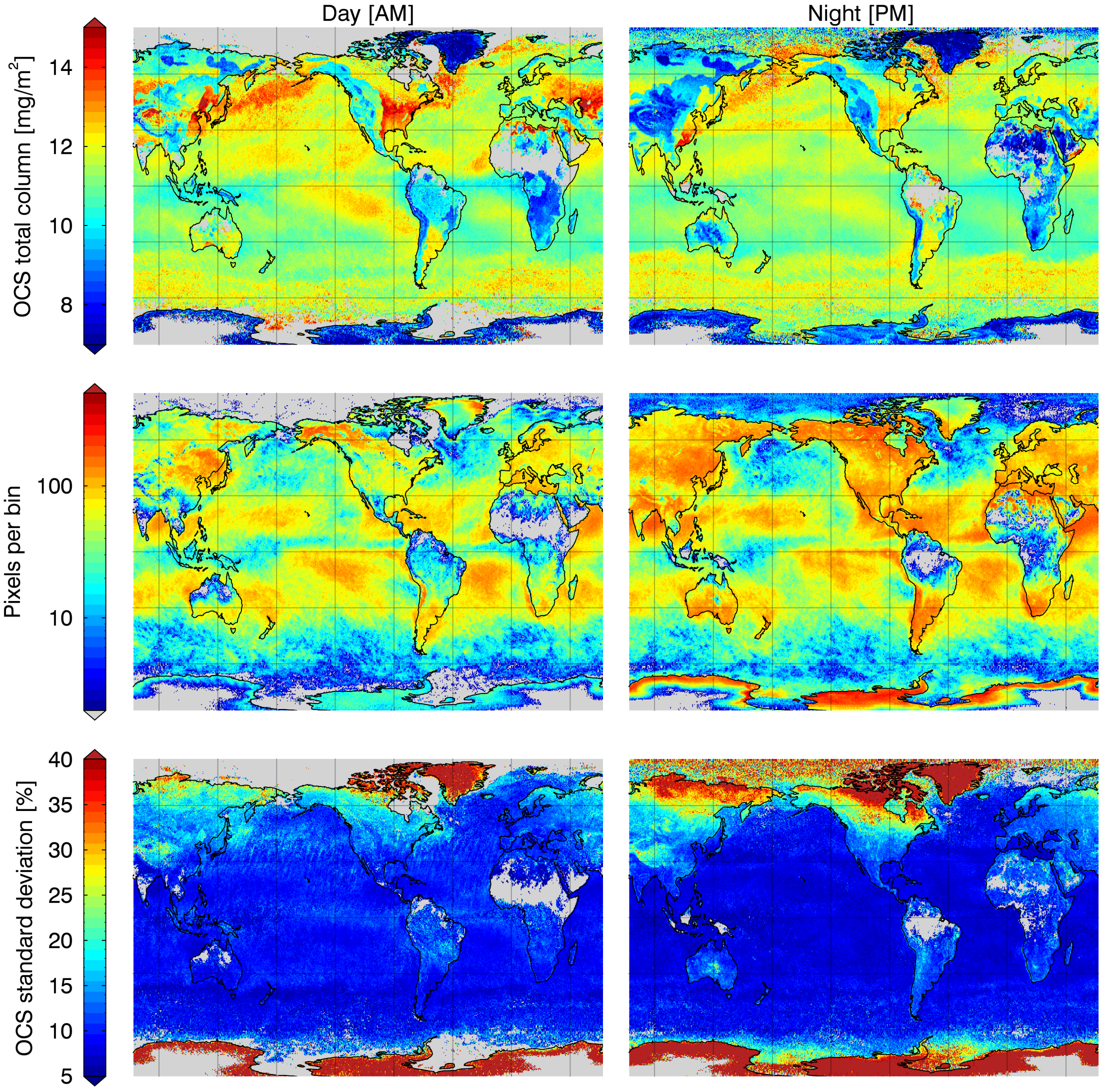


Figure 7.11: Same as in Fig. 7.10, but for March–April 2014.

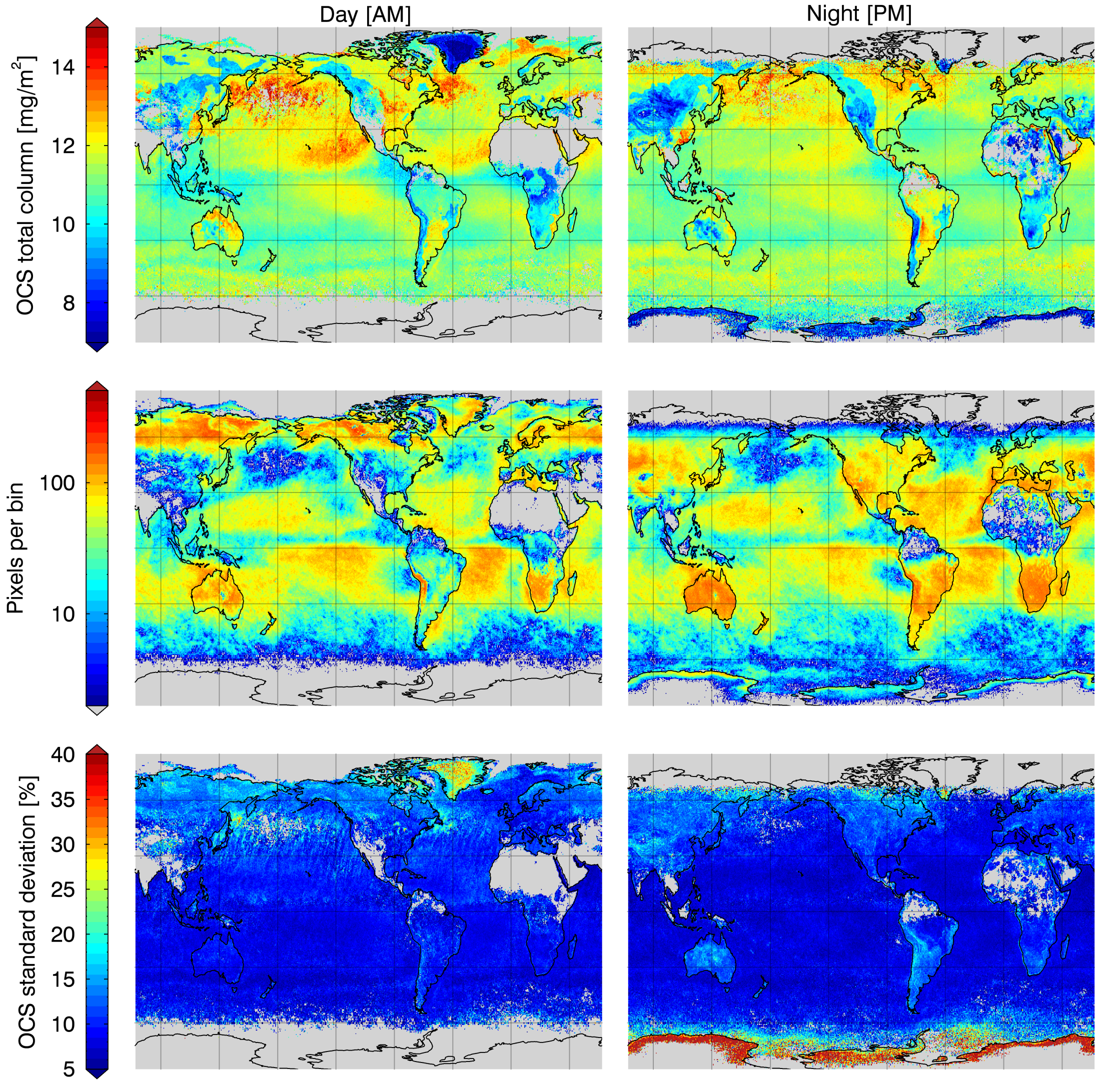


Figure 7.12: Same as in Fig. 7.10, but for May–June 2014.

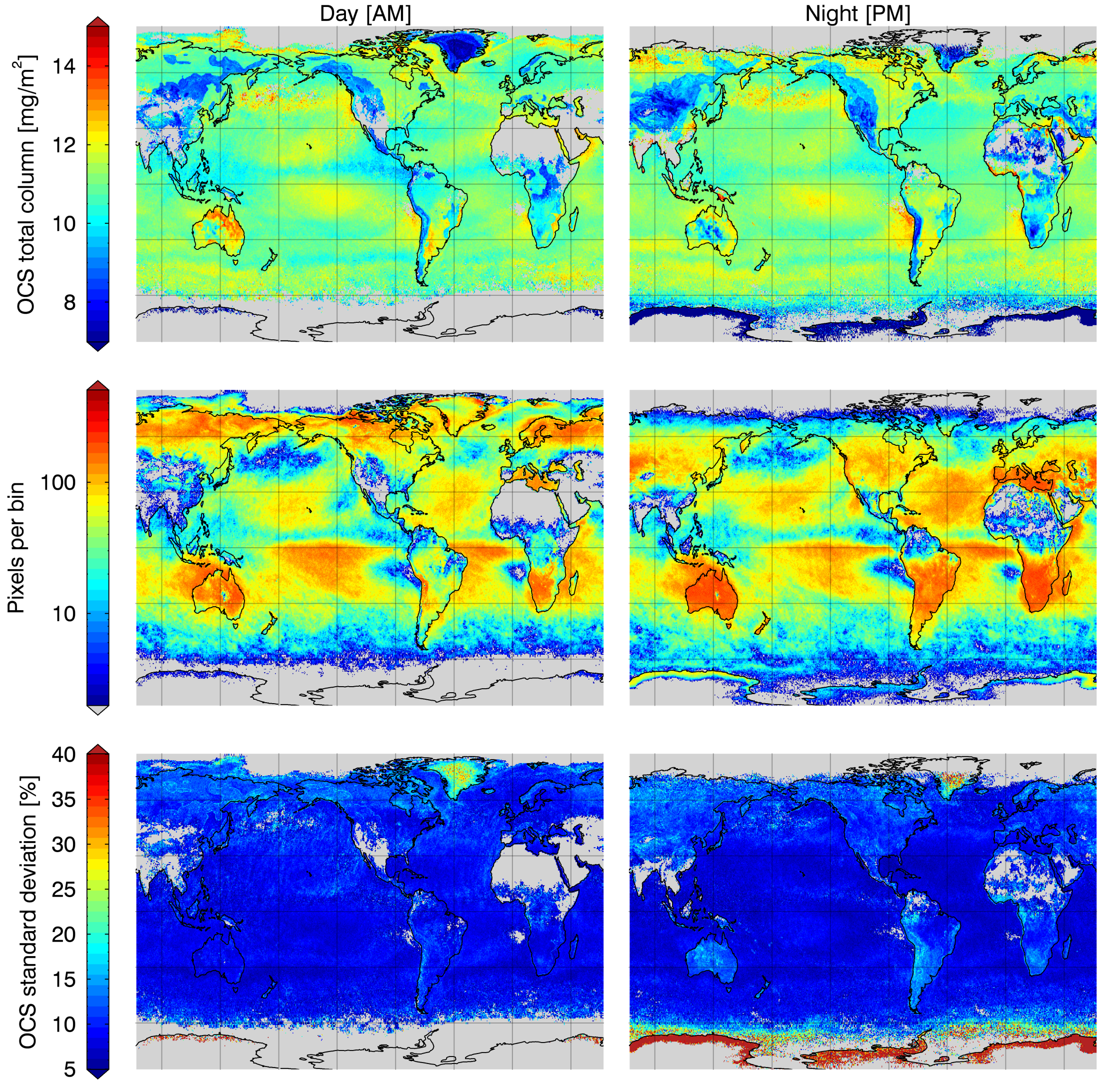


Figure 7.13: Same as in Fig. 7.10, but for July–August 2014.

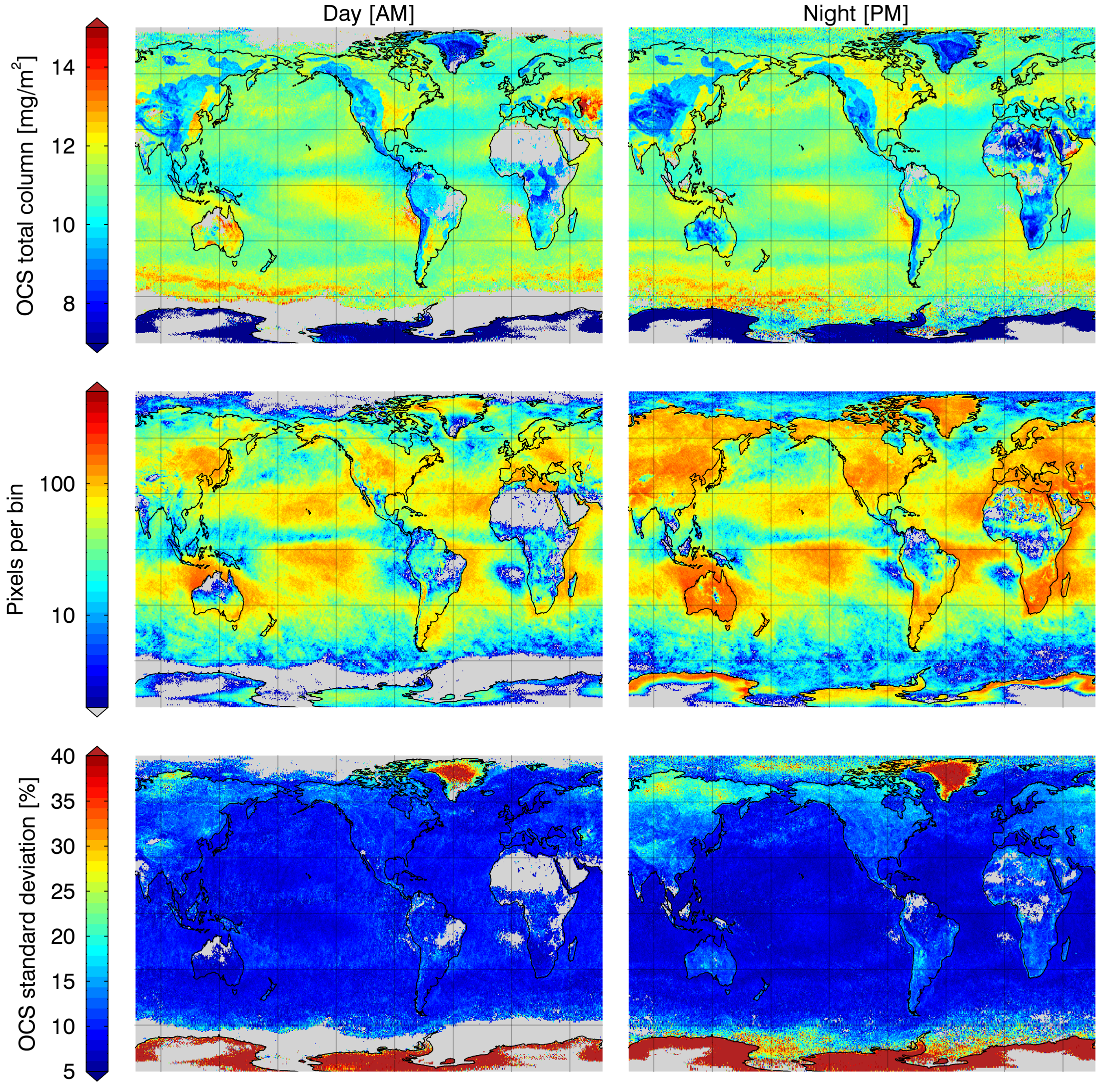


Figure 7.14: Same as in Fig. 7.10, but for September–October 2014.

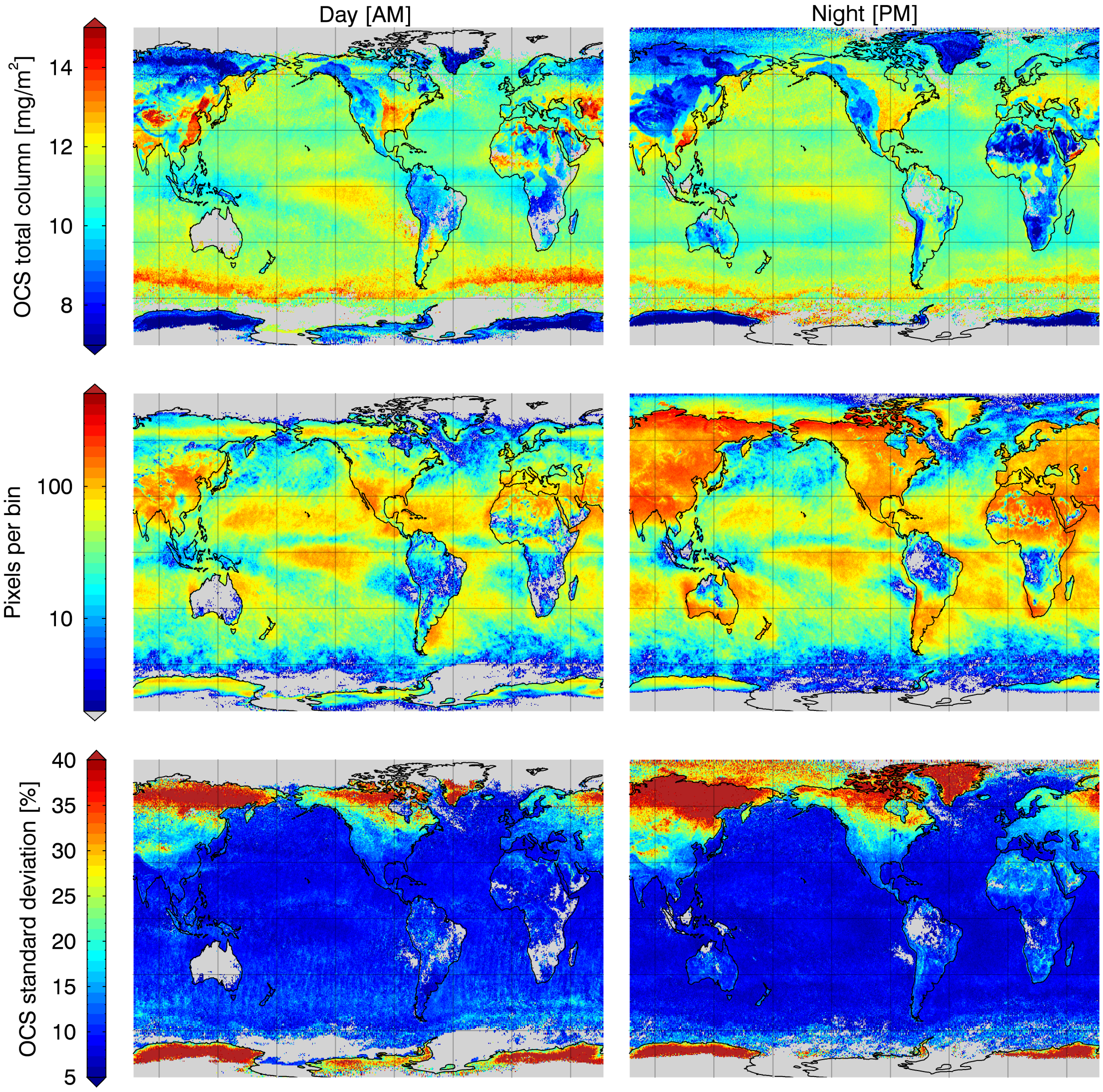


Figure 7.15: Same as in Fig. 7.10, but for November–December 2014.

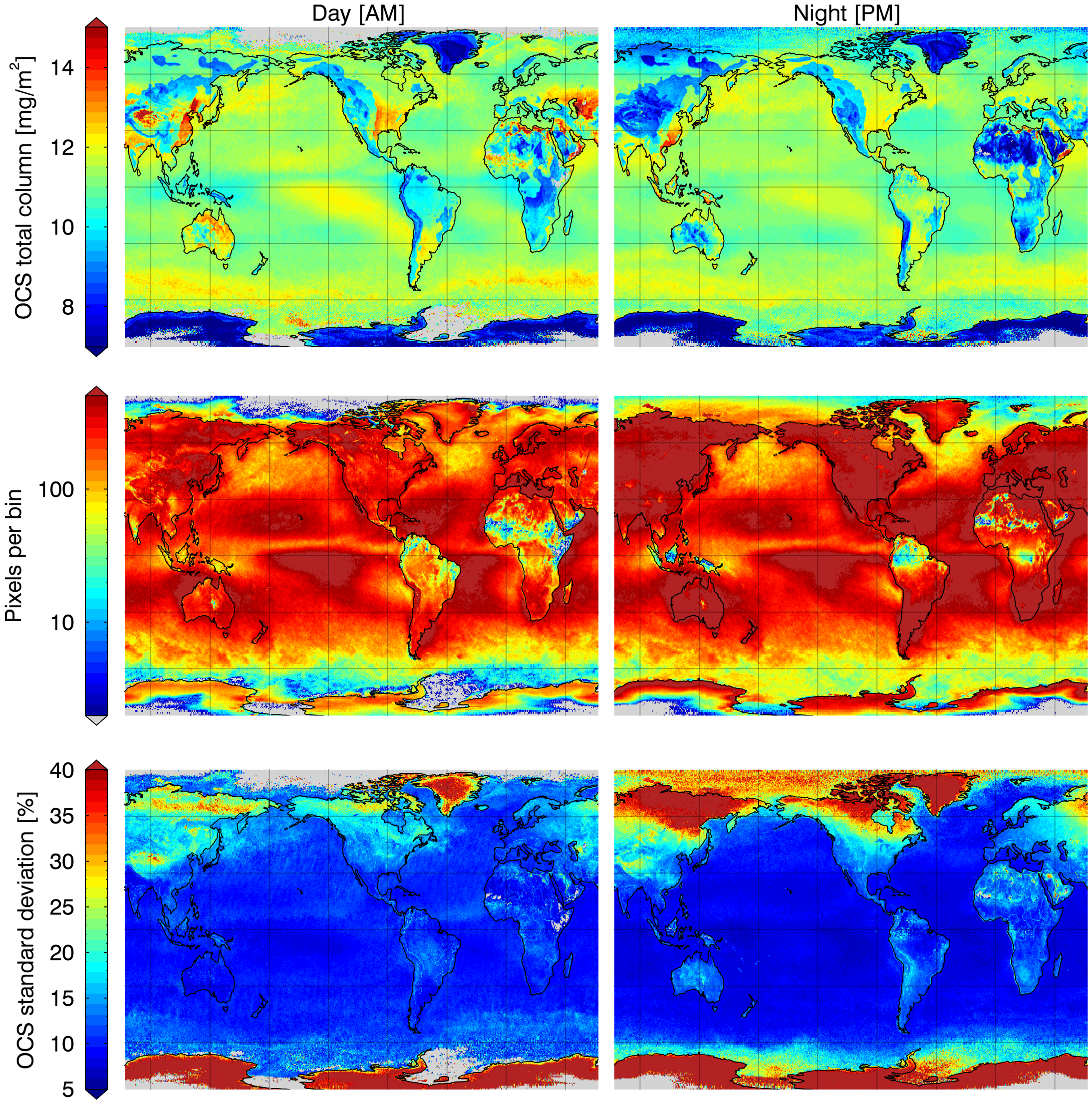


Figure 7.16: Same as in Fig. 7.10, but for all of 2014 combined.

7.4 Comparisons to NOAA flask samples

Total column estimates of OCS from the linear retrieval were also compared to VMR flask measurements of OCS collected by NOAA [*Montzka et al.*, 2007]. Although IASI total columns are different from localized point samples, the intent is to compare seasonal cycles to see if the two are temporally correlated. The Earth System Research Laboratory of NOAA collects surface air samples by flask from network sites located across the world to measure seasonal trends of numerous trace gases, including OCS. Flask measurements of OCS tend to have uncertainties within the range of $0.1 - 6$ ppt and are normally sampled on a weekly basis, but may occur less frequently depending upon location. Further information and the OCS flask data themselves are found on-line at `http://www.esrl.noaa.gov/gmd/hats/gases/OCS.html`.

Figure 7.17 shows the seasonal trend comparisons for IASI total columns against the NOAA flask measurements for seven sample sites; four in the Northern Hemisphere and three in the Southern Hemisphere. The results are displayed in monthly increments throughout 2014 where the IASI retrieved total columns are binned within a 2° radius about the location of the NOAA site. The site abbreviations along with their latitudes and longitudes are shown in the plot titles. Additionally, the monotonic (Spearman's) correlation coefficient (R) and its associated p-values (p_{val}) are also displayed. Note that p-values represent statistical significance on a scale of $0 - 1$ where values close to zero show significance in the sampled correlation, while higher values fail to reject the null hypothesis. As a rule of thumb, p-values less than 0.05 are generally regarded as statistically significant [*Hung et al.*, 1997].

Comparing pressure specific VMR to total column amount can be tenuous if the true shape of the vertical profile differs greatly from the referenced profile. Furthermore, the flask samples are not exactly coincident with the IASI observations in space and time, so this combines to introduce a certain level of natural error that is difficult to isolate and quantify. However, by analysing on a monthly basis,

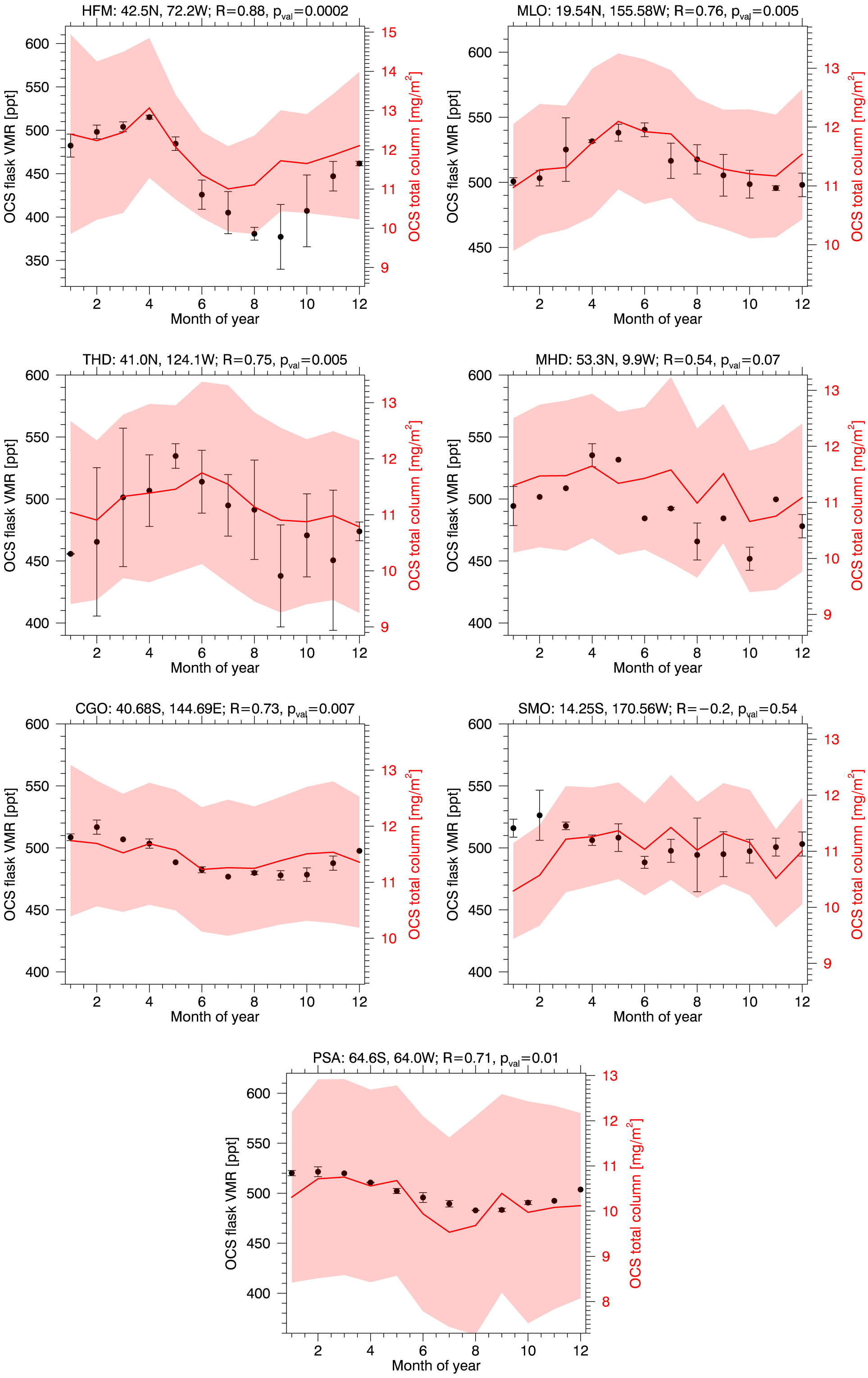


Figure 7.17: OCS total column median estimates (red) from the linear retrieval are compared to NOAA flask measurements of OCS surface VMR (black) binned by 12 month increments throughout 2014. Retrieval estimates are taken from a 2° radius about the NOAA site locations. The shaded red area represents the sample standard deviation of the total column estimates and the black error bars are the standard deviation of the flask samples within that month (*not* divided by the square root of the sample size).

these effects may be mitigated where the desired outcome is to show correlation and consistency between the seasonal signals of the two.

Of the seven, the Harvard Forest (HFM) site shows the greatest correlation at $R = 0.88$. It is important to point out that the flask samples here are taken immediately above the forest canopy at 30 m, while the IASI observations are most sensitive at mid-troposphere. Notice that the OCS flask VMR closely follows the total column trend during the winter months, but then drops proportionately much lower from June to September. Work discussed in the previous chapter suggests that forests are strong sinks of OCS and, therefore, most active during peak summer-time photosynthesis. Therefore, one would expect this sort of surface drop at canopy level compared to the total column of OCS.

Perhaps the most import comparison is the Mauna Loa Observatory (MLO), because the air is sampled closer to the peak sensitivity of IASI at an altitude of 3.5 km. Both flask VMRs and total columns show a clear seasonal cycle of OCS reaching maximum in late spring and minimum in early winter with a correlation coefficient of 0.76. A similar comparison was made for the OCS retrievals using TES [*Kuai et al.*, 2014] to NOAA flask measurements over Mauna Loa during 2011. They found a slightly higher correlation coefficient of 0.80 for their seasonal analysis, which is expected given that the TES retrieval accounts for non-linearities by iteratively minimizing the joint cost function.

Correlations similar to Mauna Loa are found at Trinidad Head (THD), Cape Grim Observatory (CGO), and Palmer Station in Antarctica (PSA). However, the site at Mace Head (MHD) shows a lower correlation of only 0.54 between the surface VMRs and the total columns. Inspection of both indicates that the OCS values at Mace Head are quite variable throughout the year with no clear seasonal behaviour. In this case, coincidence between flask samples and IASI observations becomes much more important due to the variable nature of OCS at this specific location on the west coast of Ireland.

Finally, the NOAA site located in American Samoa (SMO) actually shows a negative correlation between flask samples and IASI estimates. This is entirely due to the first two months of the year, January and February, while the remainder of the year shows a positive correlation. This early year depletion in the total column estimates can be visualized in Fig. 7.10. Notice that there is a spatial low in OCS total column that extends from the Indonesian islands well into the middle-south Pacific during this time of year. Since this is peak season for photosynthesis in the Southern Hemisphere, it is possible that American Samoa is downwind of Indonesia and northern Australia, strong OCS sinks for January and February, while the ocean surface near American Samoa is emitting OCS or its precursor gases. On the other hand, it is possible this feature is an artefact of some unsensed physical parameter or a weather effect yielding a nonlinear error biased consistently low.

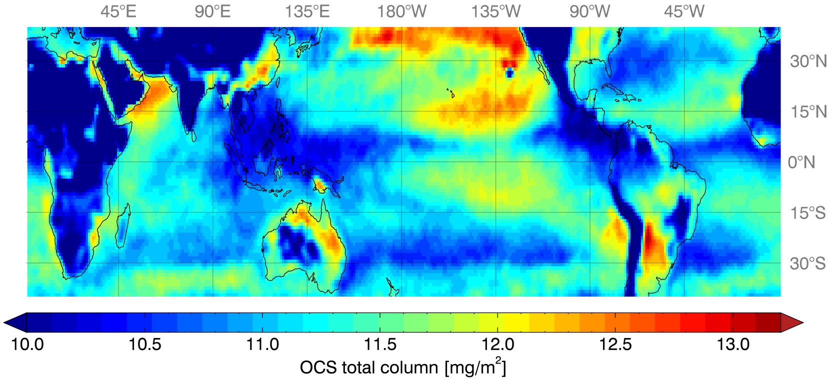
7.5 June comparisons to retrievals from TES

To date, the only nadir-viewing retrievals of tropospheric OCS to compare with are the TES results from observations in June 2006 published in *Kuai et al.* [2015]. Since IASI-A was not available in orbit until October 2006, a direct comparison for June 2006 cannot be made. However, the IASI results from June 2014 can be compared to the TES results of 2006 to identify any broad scale similarities or contradictions. Figure 7.18a shows the linear OCS median estimates from IASI for June 2014 over the same geographical range as the TES results from June 2006 displayed in the top row of Fig. 7.18b. Data from both day and night observations were used in both. The IASI results were binned in 1° latitude and longitude boxes and then smoothed to a spatial resolution of approximately 2×2° to more closely resemble the TES representation.

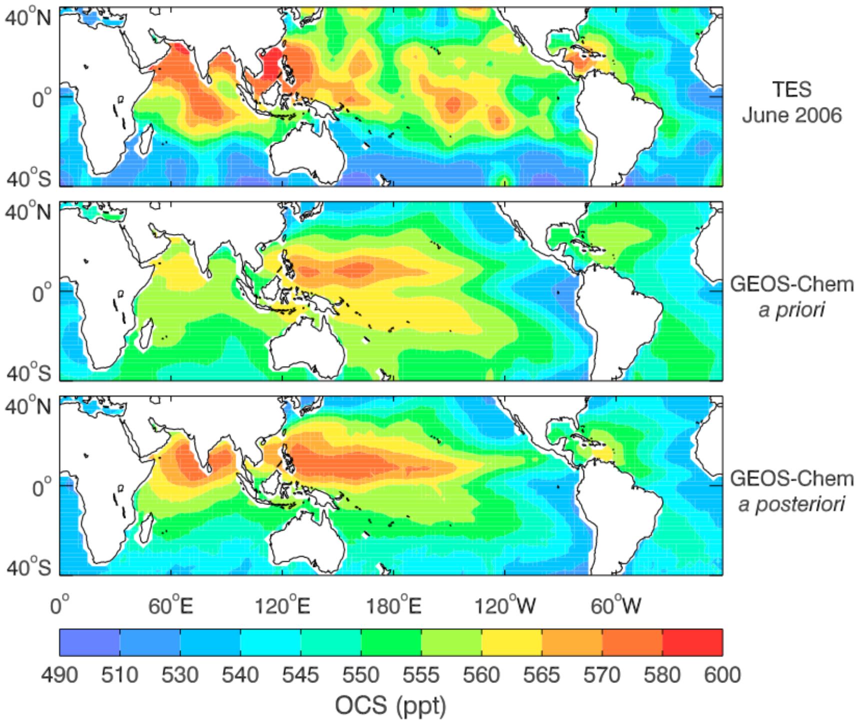
Care must be taken when comparing the two OCS products. Not only are the two maps showing different years from different instruments, but the IASI linear estimates of OCS are essentially unconstrained while the TES retrieval is constrained to a spatially flat prior OCS field. In other words, the TES results are smoothed to the prior field by a certain percentage that is not spatially uniform and dependent upon the DFS of that specific region. This is why *Kuai et al.* [2015] chose to only display retrieval results from ±40° in latitude, because the DFS in this area is closer to one. A spatial map of the mean DFS per bin would be helpful, but was not included in the publication. Future comparisons of retrieved quantities from differing methods and instruments (e.g., IASI to TES) should either be transformed into a maximum likelihood representation [*Rodgers*, 2000, Ch. 10.4.2], where the *a priori* is removed, or transform both products into a comparison specific representation determined by the *a priori* [*Rodgers and Connor*, 2003].

Nonetheless, there are OCS features that appear to be common between the IASI linear estimates and the fields from TES. In the eastern Pacific Ocean there are two features that appear to originate off the coast of northern Mexico and also Chile that extend towards the Equator along the westerly direction. While the two maps are certainly not an identical match, this general feature is seen in both. Additionally, both show OCS enhancements in the Indian Ocean off the coast of the Arabian Peninsula. However, the TES field here extends well past the southern tip of India, while the IASI results are more local to the coastline.

One area that is clearly contradictory between the two is over the Indonesian Islands and along the coast of Vietnam. The TES product shows high amounts of OCS here, while the IASI results show below average OCS total columns. It is unlikely this difference is physically accurate with only eight years between the two products. Furthermore, since the prior OCS field used in TES is 500 ppt, this enhanced feature approaching 600 ppt is driven by the signal and not the prior OCS field due to low information. Therefore, there is likely something physical other than OCS over this region that is accounted for in the TES retrieval and missed out in the IASI algorithm. One hypothesis is that this region over Indonesia may contain elevated aerosol abundances and high frequency partial cloud cover below 20% cloud fraction. Thus, extinction of the signal could be better accounted for in the linear IASI estimates, which is one area for future improvements.



(a) OCS results from IASI for June 2014, including both day and night observations.



(b) TES retrieval results of OCS are shown in the top panel from June 2006. The following two panels compare TES estimates to the GEOS-Chem CTM. This figure was taken from *Kuai et al.* [2015].

Figure 7.18: A comparison is shown between the IASI OCS result using the fast retrieval method developed in this chapter and the TES iterative retrieval for the month of June.

7.6 Seasonal comparisons to TOMCAT

The linear estimates of total column OCS from IASI are compared to the TOMCAT model developed at the University of Leeds. TOMCAT is a global off-line chemistry transport model (CTM) forced by ECMWF ERA-Interim temperature, winds, and humidity reanalyses [*Spracklen et al.*, 2005]. However, there is no direct emissions inventory of OCS currently implemented in TOMCAT, because only recently (past 15 years) have there been significant gains in understanding its sources and sinks. Therefore, sources of OCS are modelled as a constant molar emission ratio from anthropogenic SO_2 over land and oceanic OCS as an emission ratio of DMS as taken from the *Kettle et al.* [2002] inventory. Therefore, the current implementation of OCS in TOMCAT does not include the prevailing hypothesis that the major source of OCS is photochemically reduced CDOM in the oceans. Additionally, TOMCAT does not yet directly take into account sinks of OCS due to vegetative uptake.

Figure 7.19 shows the OCS total column median estimates from IASI next to the TOMCAT model of OCS by three month seasonal increments over the year 2014. These plots include both day and night observations compiled together. After inspecting the magnitude of the two, it appears that the TOMCAT model underestimates global OCS by approximately 50%. Surface VMRs of 590 ppt correspond to total column amounts of 13.1 $\mathrm{mg/m^2}$ for the reference profiles in the model atmospheres. Ground samples of OCS taken by NOAA show that surface VMRs tend to fluctuate in the range between $400 - 600$ ppt, so the total column estimates should fall somewhere between $10 - 13\,\mathrm{mg/m^2}$ on average, depending upon location and vertical profile.

The oceans are modelled with little seasonal variation in the TOMCAT model and show small changes in OCS total column with latitude in comparison to the linear retrieval. TOMCAT fixes OCS quantities to modelled DMS, which is difficult to resolve spatially and temporally with measurements due to the short lifetime of approximately one day for DMS. Therefore, there is no obvious correspondence between the estimates from IASI versus TOMCAT over ocean, with the exception of the high latitude area east of Greenland. This area does show a similarity between

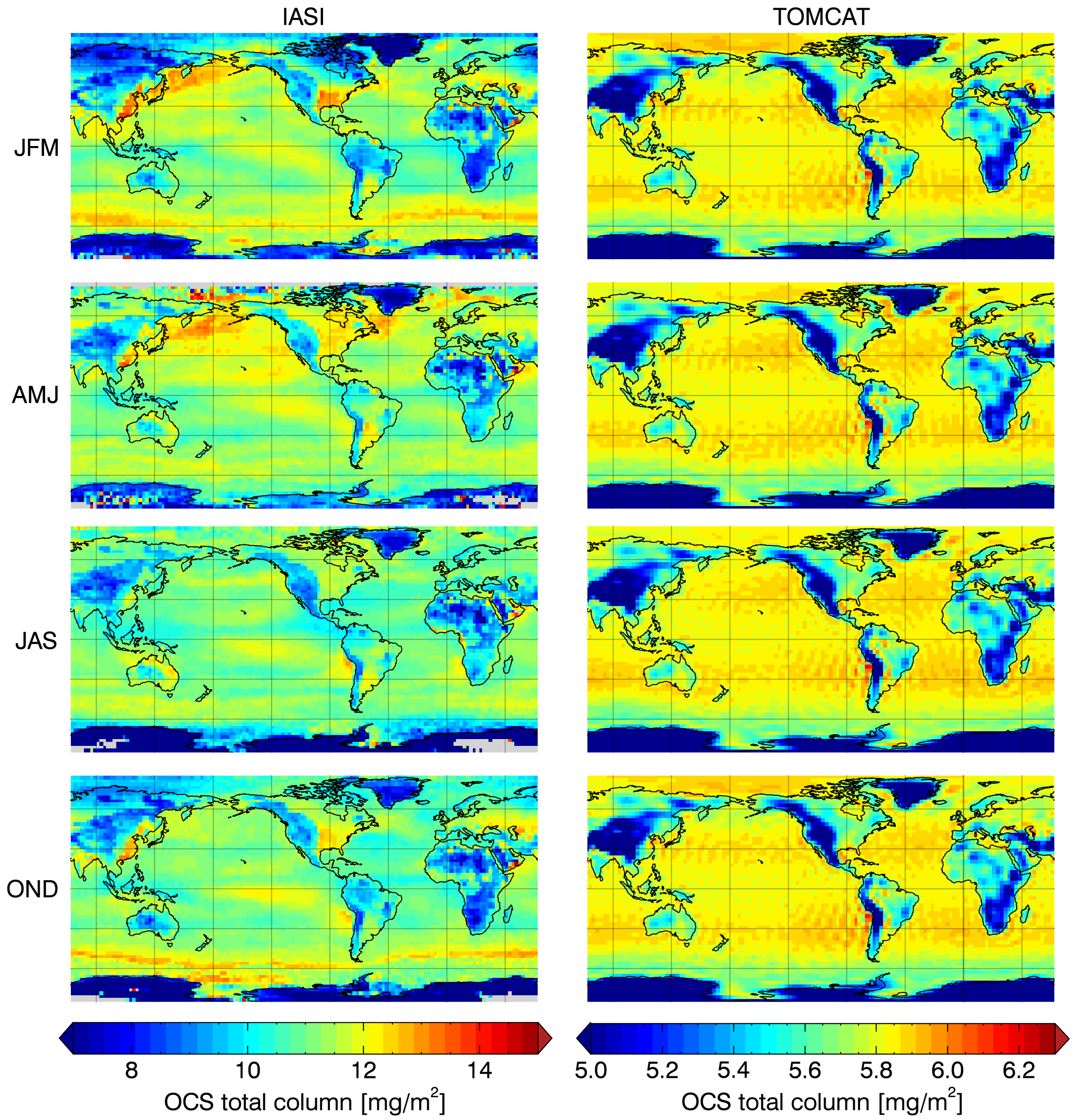


Figure 7.19: OCS total column medians from 2014 IASI observations (left column) are compared against the TOMCAT model of OCS (right column) for three month seasonal increments. The displayed results include both day and night observations. Spatial resolution is set by the TOMCAT model and each bin has a width of approximately 2.8° in latitude and longitude. Notice that the colour bars differ in scale.

the two for March through June. If the linear estimates from IASI are accurate here, then the comparison to TOMCAT suggests a DMS source for this specific feature.

Additionally, there do appear to be a few areas of elevated OCS over land that match between IASI and TOMCAT. In particular the spatial features of OCS by the Caspian Sea during fall and winter and the enhancement over northern Argentina during April to June. In TOMCAT, these areas are known for emitting substantial amounts of SO_2 and, therefore, may also emit OCS from the same physical process as detected from IASI. However, the large enhancement of detected OCS over eastern China is not represented well by parametrisation to SO_2 in TOMCAT. If this signal is in fact meaningful, then there exists a large source of OCS in China poorly correlated to SO_2 production. As mentioned in Sect. 6.2.5, industrial production of cellulose rayon and cellophane is found in numerous sites across eastern China and is one possible mechanism that could be incorporated in the TOMCAT model to see if spatial correlation of OCS between the two improves.

7.7 Summary

A novel linear retrieval method was developed and applied towards making timely estimates of OCS total columns for the entirety of IASI observations from 2014. There are two components that make this retrieval scheme unique in comparison to current linear methods. First, physical parameters that influence the spectral observations over the wavenumber range used for OCS are directly accounted for by jointly retrieving them along with OCS. This differs from previous methods in that they tend to use an effective measurement covariance that treats the physical parameters not directly retrieved as noise. Second, an initial linearisation point is selected from a global ensemble of atmospheres based on minimizing the spectral difference between the IASI and the modelled spectral radiances. This step is intended to make the retrieval more linear, thus reducing the need for iterative steps that rerun the forward model several times per pixel.

Additionally, an iterative retrieval for OCS was used as a test-bed to develop and validate the framework of the retrieval; i.e., the state vector, prior constraints, and initial atmosphere selection. Once this was accomplished, an ensemble of IASI observations over the Pacific Ocean was used to quantify the mean spectral residual for the converged estimates and showed that the majority of spectral channels match to within instrument noise, except the stronger water absorption features. Water vapour channels were then treated as noise by modifying the measurement covariance diagonals accordingly based on the mean spectral residual. Finally, channel selection was performed based on the OCS posterior uncertainty, reducing the number of channels from 1201 to 100, which ultimately made the OCS retrieval almost twice as linear.

The OCS estimates visualized in two month increments display many interesting features consistent with prior knowledge of its sources and sinks. For example, the day-time total columns show depletions in the OCS signal over tropical rainforests, which is consistent with the idea that vegetation is the strongest sink of OCS. The Pacific Ocean displays spatial features of elevated OCS that vary seasonally and appear to match the prediction made by *Berry et al.* [2013] that there is a large source in the Pacific Ocean, especially in the southern hemisphere. Interestingly, there is a clear band of high OCS estimates following the circumpolar current north of Antarctica, which is well known for consistent upwelling sustained by turbulent gyres. Additionally, regions of land showing high OCS estimates were found over China, the area east of the Caspian Sea, and northern coastal Africa leading to the Middle East. It is possible these land regions are emitting anthropogenic OCS or that there is some surface property unaccounted for that consistently leads to elevated estimates.

To validate the linear retrieval on a monthly basis, these OCS results were compared to surface VMR samples collected via flask by NOAA stations across the globe. It was found that five (three northern and two southern hemisphere) NOAA sites out of seven had seasonal cycle correlation coefficients greater than 0.7. Seasonal comparisons of model OCS fields to IASI total column estimates were also made from the CTM referred to as TOMCAT. However, it was difficult to draw too many conclusions from this analysis as numerous advances in understanding OCS sources and sinks have occurred in just the past three to eight years, which have not yet been incorporated into TOMCAT.

A great accomplishment shouldn't be the end of the road, just the starting point for the next leap forward.

— Harvey Mackay

8

Conclusion

The next generation of IASI sensors are specified to have twice the spectral and spatial resolution at one half the current instrument noise. Thus, opening up a further frontier of tropospheric trace gas estimation. In the future, with at least four times the data compared to current operations it becomes even more important to produce accurate estimates in a timely manner if the physical observations are to be fully utilized. Already there is a deficit between IASI data acquisition and estimation processing unless using a highly parametrised forward model. One cannot expect this problem to be solved solely by breakthroughs in computer hardware. Therefore, gains in retrieval efficiency, improved modelling, and novel estimation methods are all necessary towards making the most of the satellite remote sensing infrastructure.

Carbonyl sulphide is an example of a trace gas whose sources and sinks are not well known and currently on the edge of detectability. Ultimately, there needs to be an iterative optimal estimation scheme of OCS for IASI that is fast enough to keep pace with real-time acquisition. All steps of this thesis are necessary precursors towards this end goal. Since constrained OCS retrievals usually have DFS values no greater than one, it does not make sense to retrieve OCS on many vertical levels. Doing so guarantees prior knowledge artefacts appearing in the estimate. Since OCS is poorly known, efforts to minimize the reliance upon *a priori* are desirable. The procedure described in Ch. 4 can be followed to select an appropriate vertical retrieval grid that accomplishes this task.

Additionally, the LUTs of absorption cross-section and the irregular spectral grid from Ch. 5 may be incorporated into the RTM to increase the speed of generating model spectra by a factor of 100. This parametrisation of the forward model negligibly increased the spectroscopic error compared to the instrument noise of IASI for the entire spectral range, including the portion used for OCS. These LUTs encompass many different atmospheric scenarios and maintain the monochromatic nature of the radiative transfer approach. Therefore, this parametrisation is robust under a multitude of Earth-type physical variations.

In the absence of a large computational cluster, the forward model enhanced with LUTs may still be too time consuming to evaluate IASI data beyond individual and area specific events. In this case, one can reduce the accuracy of the retrieval by treating the problem within the linear framework presented in Ch. 7 while speeding up the computation process by a factor of roughly 10^4 (depending upon the specific retrieval). Analysis of model scenarios suggests that the error due to ignoring non-linearities is about 11% globally for OCS. Since these linear estimates can be generated so rapidly, it is possible to use a monthly median or mean of linear OCS fields as the initial point ($\boldsymbol{x}_0$) or even the *a priori* ($\boldsymbol{x}_\mathrm{a}$) to improve the efficiency and data quality of a constrained iterative retrieval.

Work presented in this thesis paid particular attention to OCS as an interesting test case. However, it is important to note that the linear retrieval method developed in Ch. 7, of using a multi-element state vector to jointly account for other physical parameters and selecting an initialisation point from an atmospheric ensemble, can be applied to any trace gas for any nadir viewing instrument similar to IASI. While the OCS results require further validation, the presented OCS spatial fields are intriguing and may lead to future understanding of its sources and sinks. Furthermore, this method can potentially provide additional insights for minor trace gases that are, as of yet, poorly quantified.

8.1 Future work for OCS retrievals

Within the framework of the rapid linear retrieval methodology discussed in Ch. 7, there is potential for future improvements. Perhaps the area with the greatest room for improvement is in the implementation of choosing an initial atmosphere from the given ensemble. The 80 atmosphere RTTOV ensemble was used because it contains a wide variety of Earth-like scenarios, but is small enough to keep the retrieval algorithm fast and efficient. However, greater accuracy of OCS estimates may be possible if an atmospheric ensemble was created specifically for this purpose. Open questions remain as to how many atmospheres should be included in the ensemble before the computational cost outweighs the improved OCS accuracy. Whether this number is 20, 80, or 2000 has yet to be decided. Furthermore, sectioning the selectable ensemble into latitudinal regions will certainly increase efficiency and may even improve the accuracy of the results.

Along this same vein, the linearity of the problem is highly sensitive to atmospheric thermal contrast between the ground surface temperature and the lower atmospheric temperature. In other words, the OCS Jacobians are dependent upon thermal contrast. This is linearly accounted for in the state vector by jointly retrieving surface and atmospheric temperature. However, there is likely a better way to estimate thermal contrast from the IASI spectrum prior to the OCS retrieval step. This information could be used to guide the initial atmospheric selection, rather than the method employed in this work of expanding the ensemble to include states of varying thermal contrast over land.

Furthermore, the AVHRR cloud fraction product was used to discriminate between cloudy and clear scenes. Cloudy scenes with greater than 20% cloud fraction were simply ignored and the remaining scenes were assumed to contain ideally clear skies. Therefore, any effects due to thin or partial cloud cover were neglected, along with extinction due to aerosol scattering. Since the RFM can produce a generalized extinction Jacobian, it may be possible to include this in the state vector and account for extinction effects jointly in the retrieval. Many of these effects appeared to show themselves in the CO_2 to N_2O ratio, which effectively acted as a mode to segregate unaccounted for extinction from the OCS estimates. However, retrieving extinction is a more direct method and may be worth future analysis.

Finally, once an iterative scheme for OCS using weather specific *a priori* is developed, then the fast retrieval method presented here can be further optimised in tandem with the more accurate estimates from the converged retrieval. *In situ* measurements are always desirable for external validation, but are generally sparse and suffer from coincidence sampling problems. A full iterative retrieval analysing a common set of IASI pixels is necessary to further improve upon faster retrieval methods which utilise the abundance of IASI measurements.

Bibliography

Amato, U., D. De Canditiis, and C. Serio, Effect of apodization on the retrieval of geophysical parameters from fourier-transform spectrometers, *Applied optics*, *37*(27), 6537–6543, 1998.

Amato, U., G. Masiello, C. Serio, and M. Viggiano, The σ-IASI code for the calculation of infrared atmospheric radiance and its derivatives , *Environmental Modelling & Software*, *17*(7), 651 – 667, 2002.

August, T., et al., IASI on Metop-A: Operational Level 2 retrievals after five years in orbit, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *113*(11), 1340–1371, 2012.

Aumann, H. H., et al., AIRS/AMSU/HSB on the Aqua mission: design, science objectives, data products, and processing systems, *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 253–264, doi:10.1109/TGRS.2002.808356, 2003.

Barkley, M. P., P. I. Palmer, C. D. Boone, P. F. Bernath, and P. Suntharalingam, Global distributions of carbonyl sulfide in the upper troposphere and stratosphere, *Geophysical Research Letters*, *35*(14), doi:10.1029/2008GL034270, l14810, 2008.

Barnes, I., K. H. Becker, and I. Patroescu, The tropospheric oxidation of dimethyl sulfide: A new source of carbonyl sulfide, *Geophysical Research Letters*, *21*(22), 2389–2392, doi:10.1029/94GL02499, 1994.

Berry, J., et al., A coupled model of the global cycles of carbonyl sulfide and CO2: A possible new window on the carbon cycle, *Journal of Geophysical Research: Biogeosciences*, *118*(2), 842–852, doi:10.1002/jgrg.20068, 2013.

Bey, I., et al., Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *106*(D19), 23,073–23,095, 2001.

Bowman, K. W., et al., Tropospheric emission spectrometer: Retrieval method and error analysis, *Geoscience and Remote Sensing, IEEE Transactions on*, *44*(5), 1297–1307, 2006.

Bransden, B. H., and C. J. Joachin, *Physics of atoms and molecules*, 2 ed., Prentice Hall, 2003.

Brasseur, G., D. Hauglustaine, S. Walters, P. Rasch, J.-F. Müller, C. Granier, and X. Tie, MOZART, a global chemical transport model for ozone and related chemical tracers: 1. Model description, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *103*(D21), 28,265–28,289, 1998.

Ceccherini, S., P. Raspollini, and B. Carli, Optimal use of the information provided by indirect measurements of atmospheric vertical profiles, *Optics express*, *17*(7), 4944–4958, 2009.

Chevallier, F., A. Chédin, F. Chéruy, and J.-J. Morcrette, TIGR-like atmospheric-profile databases for accurate radiative-flux computation, *Quarterly Journal of the Royal Meteorological Society*, *126*(563), 777–785, 2000.

Chin, M., and D. D. Davis, Global sources and sinks of OCS and CS2 and their distributions, *Global Biogeochemical Cycles*, *7*(2), 321–337, doi:10.1029/93GB00568, 1993.

Clarisse, L., Y. R'Honi, P.-F. Coheur, D. Hurtmans, and C. Clerbaux, Thermal infrared nadir observations of 24 atmospheric gases, *Geophysical Research Letters*, *38*(10), doi:10.1029/2011GL047271, l10802, 2011.

Clough, S., M. Shephard, E. Mlawer, J. Delamere, M. Iacono, K. Cady-Pereira, S. Boukabara, and P. Brown, Atmospheric radiative transfer modeling: a summary of the AER codes , *Journal of Quantitative Spectroscopy and Radiative Transfer*, *91*(2), 233 – 244, 2005.

Curtis, A. R., Discussion of "A statistical model for water-vapour absorption" by R.M. Goody, *Quarterly Journal of the Royal Meteorological Society*, *78*(338), 638–640, 1952.

Dudhia, A., Reference Forward Model (RFM), `http://www.atm.ox.ac.uk/RFM`, last access: June 2015, 2008.

Dudhia, A., P. E. Morris, and R. J. Wells, Fast monochromatic radiative transfer calculations for limb sounding, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *74*(6), 745–756, doi:10.1016/S0022-4073(01)00285-0, 2002.

Edwards, D. P., GENLN2: A general line-by-line atmospheric transmittance and radiance model. Version 3.0 description and users guide, *Tech. Rep. NCAR/TN-367+STR*, National Center for Atmospheric Research, Boulder, Colorado, USA, 1992.

Fischer, H., et al., MIPAS: an instrument for atmospheric and climate research, *Atmospheric Chemistry and Physics*, *8*(8), 2151–2188, 2008.

Fried, A., B. Henry, R. A. Ragazzi, M. Merrick, J. Stokes, T. Pyzdrowski, and R. Sams, Measurements of carbonyl sulfide in automotive emissions and an assessment of its importance to the global sulfur cycle, *Journal of Geophysical Research: Atmospheres*, *97*(D13), 14,621–14,634, 1992.

Galatry, L., Simultaneous effect of doppler and foreign gas broadening on spectral lines, *Phys. Rev.*, *122*, 1218–1223, 1961.

Glatthor, N., et al., Tropical sources and sinks of carbonyl sulfide observed from space, *Geophysical Research Letters*, doi:10.1002/2015GL066293, 2015GL066293, 2015.

Godson, W. L., The evaluation of infrared radiative fluxes due to atmospheric water vapour, *Quarterly Journal of the Royal Meteorological Society*, *79*(341), 367–379, 1953.

Goody, R., R. West, L. Chen, and D. Crisp, The correlated-k method for radiation calculations in nonhomogeneous atmospheres, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *42*(6), 539–550, 1989.

Goody, R. M., and Y. L. Yung, *Atmospheric radiation: theoretical basis*, 2 ed., Oxford University Press, 1989.

Hadji-Lazaro, J., C. Clerbaux, P. Couvert, P. Chazette, and C. Boonne, Cloud filter for CO retrieval from IMG infrared spectra using ECMWF temperatures and POLDER cloud data, *Geophysical research letters*, *28*(12), 2397–2400, 2001.

Han, Y., et al., Suomi NPP CrIS measurements, sensor data record algorithm, calibration and validation activities, and record data quality, *Journal of Geophysical Research: Atmospheres*, *118*(22), 12,734–12,748, doi:10.1002/2013JD020344, 2013JD020344, 2013.

Hilton, F., N. Atkinson, S. English, and J. Eyre, Assimilation of IASI at the Met Office and assessment of its impact through observing system experiments, *Quarterly Journal of the Royal Meteorological Society*, *135*(639), 495–505, 2009.

Hilton, F., et al., Hyperspectral Earth Observation from IASI: Five Years of Accomplishments, *Bulletin of the American Meteorological Society*, *93*(3), 347–370, doi:10.1175/BAMS-D-11-00027.1, 2011.

Hollas, J. M., *Modern spectroscopy*, John Wiley & Sons, 2004.

Hollingsworth A., et al., Toward a monitoring and forecasting system for atmospheric composition: The gems project, *Bulletin of the American Meteorological Society*, *89*(8), 1147–1164, 2008.

Hung, H. M. J., R. T. O'Neill, P. Bauer, and K. Kohne, The behavior of the p-value when the alternative hypothesis is true, *Biometrics*, *53*(1), 11–22, 1997.

Hurtmans, D., P.-F. Coheur, C. Wespes, L. Clarisse, O. Scharf, C. Clerbaux, J. Hadji-Lazaro, M. George, and S. Turquety, FORLI radiative transfer and retrieval code for IASI , *Journal of Quantitative Spectroscopy and Radiative Transfer*, *113*(11), 1391 – 1408, three Leaders in Spectroscopy, 2012.

Jacquinet-Husson, N., et al., The GEISA spectroscopic database: Current and future archive for Earth and planetary atmosphere studies, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *109*(6), 1043–1059, 2008.

Joiner, J., and A. Da Silva, Efficient methods to assimilate remotely sensed data based on information content, *Quarterly Journal of the Royal Meteorological Society*, *124*(549), 1669–1694, 1998.

Kato, H., M. Saito, Y. Nagahata, and Y. Katayama, Degradation of ambient carbonyl sulfide by Mycobacterium spp. in soil, *Microbiology*, *154*(1), 249–255, 2008.

Kettle, A. J., U. Kuhn, M. von Hobe, J. Kesselmeier, and M. O. Andreae, Global budget of atmospheric carbonyl sulfide: Temporal and spatial variations of the dominant sources and sinks, *Journal of Geophysical Research: Atmospheres*, *107*(D22), ACH 25–1–ACH 25–16, doi:10.1029/2002JD002187, 4658, 2002.

Koshelev, M., M. Tretyakov, G. Golubiatnikov, V. Parshin, V. Markov, and I. Koval, Broadening and shifting of the 321-, 325- and 380-GHz lines of water vapor by pressure of atmospheric gases, *Journal of Molecular Spectroscopy*, *241*(1), 101–108, doi:10.1016/j.jms.2006.11.005, 2007.

Kremser, S., N. B. Jones, M. Palm, B. Lejeune, Y. Wang, D. Smale, and N. M. Deutscher, Positive trends in southern hemisphere carbonyl sulfide, *Geophysical Research Letters*, *42*(21), 9473–9480, doi:10.1002/2015GL065879, 2015GL065879, 2015.

Krysztofiak, G., Y. V. Té, V. Catoire, G. Berthet, G. C. Toon, F. Jégou, P. Jeseck, and C. Robert, Carbonyl Sulphide (OCS) Variability with Latitude in the Atmosphere, *Atmosphere-Ocean*, *53*(1), 89–101, doi:10.1080/07055900.2013.876609, 2015.

Kuai, L., J. Worden, S. S. Kulawik, S. A. Montzka, and J. Liu, Characterization of Aura TES carbonyl sulfide retrievals over ocean, *Atmospheric Measurement Techniques*, *7*(1), 163–172, doi:10.5194/amt-7-163-2014, 2014.

Kuai, L., et al., Estimate of carbonyl sulfide tropical oceanic surface fluxes using Aura Tropospheric Emission Spectrometer observations, *Journal of Geophysical Research: Atmospheres*, *120*(20), 11,012–11,023, doi:10.1002/2015JD023493, 2015JD023493, 2015.

Kulawik, S. S., G. Osterman, D. B. Jones, and K. W. Bowman, Calculation of altitude-dependent Tikhonov constraints for TES nadir retrievals, *Geoscience and Remote Sensing, IEEE Transactions on*, *44*(5), 1334–1342, 2006.

Kuntz, M., and M. Höpfner, Efficient line-by-line calculation of absorption coefficients, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *63*(1), 97–114, 1999.

Lacis, A. A., and V. Oinas, A description of the correlated-k distribution method for modeling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres, *Journal of Geophysical Research*, *96*(15), 9027–9064, 1991.

Launois, T., S. Belviso, L. Bopp, C. G. Fichot, and P. Peylin, A new model for the global biogeochemical cycle of carbonyl sulfide – Part 1: Assessment of direct marine emissions with an oceanic general circulation and biogeochemistry model, *Atmospheric Chemistry and Physics*, *15*(5), 2295–2312, doi:10.5194/acp-15-2295-2015, 2015a.

Launois, T., P. Peylin, S. Belviso, and B. Poulter, A new model of the global biogeochemical cycle of carbonyl sulfide – Part 2: Use of carbonyl sulfide to constrain gross primary productivity in current vegetation models, *Atmospheric Chemistry and Physics*, *15*(16), 9285–9312, doi:10.5194/acp-15-9285-2015, 2015b.

Li, Z., A. Khananian, R. H. Fraser, and J. Cihlar, Automatic detection of fire smoke using artificial neural networks and threshold approaches applied to AVHRR imagery, *Geoscience and Remote Sensing, IEEE Transactions on*, *39*(9), 1859–1870, 2001.

Marquardt, D. W., An algorithm for least-squares estimation of nonlinear parameters, *Journal of the Society for Industrial & Applied Mathematics*, *11*(2), 431–441, 1963.

Matricardi, M., The generation of RTTOV regression coefficients for IASI and AIRS using a new profile training set and a new line-by-line database, *Tech. Rep. 564*, ECMWF Research Dept, Shinfield Park, Reading, Berkshire RG2 9AX, England, 2008.

Matricardi, M., F. Chevallier, G. Kelly, and J.-N. Thépaut, An improved general fast radiative transfer model for the assimilation of radiance observations, *Quarterly Journal of the Royal Meteorological Society*, *130*(596), 153–173, 2004.

Moncet, J.-L., G. Uymin, A. E. Lipton, and H. E. Snell, Infrared Radiance Modeling by Optimal Spectral Sampling, *Journal of the Atmospheric Sciences*, *65*(12), 3917–3934, doi: 10.1175/2008JAS2711.1, 2008.

Montzka, S., P. Calvert, B. Hall, J. Elkins, T. Conway, P. Tans, and C. Sweeney, On the global distribution, seasonality, and budget of atmospheric carbonyl sulfide (COS) and some similarities to CO2, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *112*(D9), 2007.

Notholt, J., et al., Composition of the upper tropical troposphere and its influence on the stratospheric aerosol formation, in *EGS - AGU - EUG Joint Assembly*, 2003.

Para, J., P. G. Coble, B. Charrière, M. Tedetti, C. Fontana, and R. Sempéré, Fluorescence and absorption properties of chromophoric dissolved organic matter (CDOM) in coastal surface waters of the northwestern Mediterranean Sea, influence of the Rhône River, *Biogeosciences*, *7*(12), 4083–4103, doi:10.5194/bg-7-4083-2010, 2010.

Payne, V. H., S. A. Clough, M. W. Shephard, R. Nassar, and J. A. Logan, Information-centered representation of retrievals with limited degrees of freedom for signal: Application to methane from the Tropospheric Emission Spectrometer, *Journal of Geophysical Research: Atmospheres (1984–2012)*, *114*(D10307), 2009.

Peyton, T. O., R. V. Steele, and W. R. Mabey, *Carbon disulfide, carbonyl sulfide: literature review and environmental assessment*, Environmental Protection Agency, Office of Health and Ecological Effects, 1978.

Protoschill-Krebs, G., and J. Kesselmeier, Enzymatic pathways for the consumption of carbonyl sulphide (COS) by higher plants, *Botanica Acta*, *105*(3), 206–212, doi:10.1111/j.1438-8677.1992.tb00288.x, 1992.

Rodgers, C. D., *Inverse methods for atmospheric sounding: Theory and practice*, vol. 2, 2 ed., World Scientific Singapore, 2000.

Rodgers, C. D., and B. J. Connor, Intercomparison of remote sounding instruments, *Journal of Geophysical Research: Atmospheres*, *108*(D3), doi:10.1029/2002JD002299, 4116, 2003.

Rothman, L., et al., The HITRAN2012 molecular spectroscopic database, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *130*, 4–50, 2013.

Saptari, V., *Fourier transform spectroscopy instrumentation engineering*, vol. 61, SPIE press, 2004.

Saunders, R., An automated scheme for the removal of cloud contamination from AVHRR radiances over western Europe, *International Journal of Remote Sensing*, *7*(7), 867–886, 1986.

Scherr, N., and L. Nguyen, Mycobacterium versus Streptomyces—we are different, we are the same, *Current Opinion in Microbiology*, *12*(6), 699–707, doi:10.1016/j.mib.2009.10.003, growth and development: eukaryotes/prokaryotes, 2009.

Schlüssel, P., T. H. Hultberg, P. L. Phillips, T. August, and X. Calbet, The operational IASI Level 2 processor, *Advances in space research*, *36*(5), 982–988, 2005.

Scott, N. A., and A. Chedin, A Fast Line-by-Line Method for Atmospheric Absorption Computations: The Automatized Atmospheric Absorption Atlas, *Journal of Applied Meteorology*, *20*(7), 802–812, 1981.

Seibt, U., L. Wingate, J. Lloyd, and J. A. Berry, Diurnally variable δ^{18}O signatures of soil CO_2 fluxes indicate carbonic anhydrase activity in a forest soil, *Journal of Geophysical Research: Biogeosciences*, *111*(G4), n/a–n/a, doi:10.1029/2006JG000177, g04005, 2006.

Shannon, C. E., A mathematical theory of communication, *ACM SIGMOBILE Mobile Computing and Communications Review*, *5*(1), 3–55, 2001.

Shephard, M. W., S. A. Clough, V. H. Payne, W. L. Smith, S. Kireev, and K. E. Cady-Pereira, Performance of the line-by-line radiative transfer model (LBLRTM) for temperature and species retrievals: IASI case studies from JAIVEx, *Atmospheric Chemistry and Physics*, *9*(19), 7397–7417, doi:10.5194/acp-9-7397-2009, 2009.

Smith, D. E., and R. W. Timmerman, Carbon disulfide, *Kirk-Othmer Encyclopedia of Chemical Technology*, 2003.

Sparks, L., Efficient line-by-line calculation of absorption coefficients to high numerical accuracy, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *57*(5), 631–650, 1997.

Spracklen, D. V., K. J. Pringle, K. S. Carslaw, M. P. Chipperfield, and G. W. Mann, A global off-line model of size-resolved aerosol microphysics: I. Model development and prediction of aerosol properties, *Atmospheric Chemistry and Physics*, *5*(8), 2227–2252, doi:10.5194/acp-5-2227-2005, 2005.

Srivastava, R. K., and W. Jozewicz, Flue gas desulfurization: the state of the art, *Journal of the Air & Waste Management Association*, *51*(12), 1676–1688, 2001.

Strow, L., H. E. Motteler, R. G. Benson, S. E. Hannon, and S. D. Souza-Machado, Fast computation of monochromatic infrared atmospheric transmittances using compressed look-up tables, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *59*(3–5), 481 – 493, Atmospheric Spectroscopy Applications 96, 1998.

Sunda W., Kieber D. J., Kiene R. P., and Huntsman S., An antioxidant function for DMSP and DMS in marine algae, *Nature*, *418*(6895), 317–320, doi:http://dx.doi.org/10.1038/nature00851, 10.1038/nature00851, 2002.

Suntharalingam, P., A. J. Kettle, S. M. Montzka, and D. J. Jacob, Global 3-D model analysis of the seasonal cycle of atmospheric carbonyl sulfide: Implications for terrestrial vegetation uptake, *Geophysical Research Letters*, *35*(19), doi:10.1029/2008GL034332, l19801, 2008.

Tipping, R., and Q. Ma, Theory of the water vapor continuum and validations, *Atmospheric Research*, *36*(1–2), 69 – 94, 1995.

Tobin, D. C., L. L. Strow, W. J. Lafferty, and W. B. Olson, Experimental investigation of the self- and N_2-broadened continuum within the ν_2 band of watervapor, *Applied Optics*, *35*(24), 4724–4734, 1996.

Torrance, K. E., and E. M. Sparrow, Theory for off-specular reflection from roughened surfaces, *J. Opt. Soc. Am.*, *57*(9), 1105–1114, doi:10.1364/JOSA.57.001105, 1967.

Van Damme, M., L. Clarisse, C. L. Heald, D. Hurtmans, Y. Ngadi, C. Clerbaux, A. J. Dolman, J. W. Erisman, and P. F. Coheur, Global distributions, time series and error characterization of atmospheric ammonia (NH_3) from IASI satellite observations, *Atmospheric Chemistry and Physics*, *14*(6), 2905–2922, doi:10.5194/acp-14-2905-2014, 2014.

Van Diest, H., and J. Kesselmeier, Soil atmosphere exchange of carbonyl sulfide (COS) regulated by diffusivity depending on water-filled pore space, *Biogeosciences*, *5*(2), 475–483, 2008.

Ventress, L., Atmospheric sounding using IASI, Ph.D. thesis, Oxford University, 2014.

Ventress, L., and A. Dudhia, Improving the selection of IASI channels for use in numerical weather prediction, *Quarterly Journal of the Royal Meteorological Society*, 2013.

Vincent, R. A., and A. Dudhia, Fast radiative transfer using monochromatic look-up tables, *Journal of Quantitative Spectroscopy and Radiative Transfer*, pp. –, doi:10.1016/j.jqsrt.2016.04.011, 2016.

Vincent, R. A., A. Dudhia, and L. J. Ventress, Vertical level selection for temperature and trace gas profile retrievals using IASI, *Atmospheric Measurement Techniques*, *8*(6), 2359–2369, doi:10.5194/amt-8-2359-2015, 2015.

von Clarmann, T., Smoothing error pitfalls, *Atmospheric Measurement Techniques*, *7*, 3023–3034, 2014.

von Clarmann, T., and U. Grabowski, Elimination of hidden a priori information from remotely sensed profile data, *Atmospheric Chemistry and Physics*, *7*(2), 397–408, 2007.

von Clarmann, T., U. Grabowski, and M. Kiefer, On the role of non-random errors in inverse problems in radiative transfer and other applications, *Journal of Quantitative Spectroscopy and Radiative Transfer*, *71*(1), 39–46, 2001.

Von Hobe, M., G. A. Cutter, A. J. Kettle, and M. O. Andreae, Dark production: A significant source of oceanic COS, *Journal of Geophysical Research: Oceans*, *106*(C12), 31,217–31,226, doi:10.1029/2000JC000567, 2001.

Walker, J., A. Dudhia, and E. Carboni, An effective method for the detection of trace species demonstrated using the MetOp Infrared Atmospheric Sounding Interferometer, *Atmos. Meas. Tech*, *4*(8), 1567–1580, 2011.

Watts, S. F., The mass budgets of carbonyl sulfide, dimethyl sulfide, carbon disulfide and hydrogen sulfide, *Atmospheric Environment*, *34*(5), 761–779, doi:10.1016/S1352-2310(99)00342-8, 2000.

Whelan, M. E., D.-H. Min, and R. C. Rhew, Salt marsh vegetation as a carbonyl sulfide (COS) source to the atmosphere, *Atmospheric Environment*, *73*, 131–137, doi:10.1016/j.atmosenv.2013.02.048, 2013.

Wingate, L., U. Seibt, K. Maseyk, J. Ogee, P. Almeida, D. Yakir, J. S. Pereira, and M. Mencuccini, Evaporation and carbonic anhydrase activity recorded in oxygen isotope signatures of net CO_2 fluxes from a Mediterranean soil, *Global Change Biology*, *14*(9), 2178–2193, doi:10.1111/j.1365-2486.2008.01635.x, 2008.

Wofsy, S. C., HIAPER Pole-to-Pole Observations (HIPPO): fine-grained, global-scale measurements of climatically important atmospheric gases and aerosols, *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *369*(1943), 2073–2086, doi:10.1098/rsta.2010.0313, 2011.

Worden, J., et al., Tropospheric Emission Spectrometer observations of the tropospheric HDO/H2O ratio: Estimation approach and characterization, *Journal of Geophysical Research: Atmospheres*, *111*(D16), 2006.

Xu, X., H. G. Bingemer, H.-W. Georgii, U. Schmidt, and U. Bartell, Measurements of Carbonyl Sulfide (COS) in Surface Seawater and Marine Air, and Estimates of the Air-Sea Flux from Observations During two Atlantic Cruises, *Journal of Geophysical Research*, *106*(D4), 3491–3502, 2001.

Yi, Z., X. Wang, G. Sheng, D. Zhang, G. Zhou, and J. Fu, Soil uptake of carbonyl sulfide in subtropical forests with different successional stages in south China, *Journal of Geophysical Research: Atmospheres*, *112*(D8), n/a–n/a, doi:10.1029/2006JD008048, d08302, 2007.