NEMESIS for Newbies

Leigh N. Fletcher

Abstract

An attempt to introduce new starters to the magical world of spectral forward models and optimal estimation retrievals.

Contents

1. Introduction

NEMESIS is the collective name for a suite of radiative transfer and spectral retrieval codes developed at the University of Oxford. The principle developer is Prof. Patrick Irwin, and NEMESIS is now used in a number of institutions across the UK (Leicester, Bristol, UCL, Open University) and USA (Goddard Space Flight Center and Jet Propulsion Laboratory).

NEMESIS is a tool for the analysis of observed imaging and spectroscopy of planetary atmospheres, both in our Solar System and, more recently, for exoplanets and Brown Dwarfs. The typical workflow of a NEMESIS team member has three stages:

1. Data Acquisition: Observations of planetary spectra from the ultraviolet to the microwave regime are obtained using Earth-based observatories, space-based telescopes, and interplanetary spaceprobes. These data are typically a measure of the spectral radiance as a function of location (latitude and longitude) on a planetary disc, but can also come in the form of disk-integrated (i.e., not spatially resolved) irradiance from a point source.

2. Spectral fitting: With a suitable choice of forward model (thermal emission, scattering, etc.), the assembled sources of opacity (spectral line data, continuum contributions, aerosol properties), and a known geometry (e.g., nadir or limb geometries, disc integration, transit spectra), the forward model can be iterated over multiple times, tweaking the atmospheric parameters using the calculated spectrum and measured spectrum match within the data uncertainty.

3. Interpretation: The final atmospheric profiles (temperature, cloud and composition) are the optimal estimates of the true environmental properties within the planet's atmosphere. These properties might be zonally averaged (i.e., a two-dimensional measurement of the derived quantity as a function of latitude and altitude), or spatially mapped across the disc. Beyond the NEMESIS inversions, these atmospheric properties can then be used to examine atmospheric origins, dynamics and meteorology, photochemistry and thermochemistry, and the physics of cloud formation.

These three categories represent the broadest of interpretations of the work of the NEMESIS group, and team members are continually updating the code, acquiring new data, and using the tool to solve new problems in planetary physics. This document is meant to supplement the Radtrans (forward model) and NEMESIS (inverse model) manuals, as well as the primary paper on this topic (Irwin et al., 2008)

2. Background

This section is derived from a number of excellent textbooks on the topic, and can be skipped by those wishing to get straight

to the code and its application. Here we introduce the basics of a planetary spectrum (using giant planets as an example), discuss forward modelling, molecular spectroscopy, the radiative transfer equation, sources of opacity (lines, collision induced absorption, aerosols) and methods of handling them (line by line and correlated-k). We then introduce spectral retrievals, looking at the definition of optimal estimation and potential sources of uncertainty.

2.1. A Basic Planetary Spectrum

The processes determining the observed radiance as a function of wavelength (or wavenumber cm^{-1}, $\nu = 1/\lambda$) vary as we move from the ultraviolet through to the microwave and radio. Aerosols, which both absorb and scatter light, will have a dominant effect in regions where the wavelength and the aerosol size are comparable - thus micron-sized particles won't have any influence on centimetre-wave measurements. Using the example of a giant planet spectrum, we identify the following spectral regions, broadly divided into reflected sunlight shortward of 4 μm and thermal emission longward of 4 μm.

- Ultraviolet: Shortward of 300 nm, dominated by Rayleigh and Raman scattering, with some absorption features from gases present in the upper troposphere and stratospheres.

- Visible: 400-900 nm Sunlight reflected from aerosols in haze and cloud layers, with absorption from gaseous methane becoming more and more prominent towards the redward (long-wavelength) end. The stronger the methane absorption, the higher in the atmosphere the reflection comes from.

- Near-infrared: 1-4 μm, continuing the reflected sunlight spectrum with strengthening absorption bands of methane. Additional species (ammonia, phosphine, hydrocarbons) can also absorb some of the reflected light. This range is typically used to study the 3D distributions of aerosols (i.e., clouds and hazes) Fluorescent effects from H$_3^+$, CO, CO$_2$ can also contribute, sensing the upper atmosphere.

- Mid-infrared: 5-30 μm, the region where thermal emission dominates over reflected sunlight, so the latter can be reflected. Spectra are formed from gaseous absorption and emission (vibration-rotation) features superimposed onto a collision-induced continuum of hydrogen and helium, which can be used to measure the temperature of the atmosphere and derive atmospheric composition.

- Far-infrared: 30-500 μm, continuing the dominance of the collision-induced continuum but with simpler rotational line features of atmospheric gases and a declining influence of aerosols (as they are now considerably smaller than the wavelength). These lines can be tropospheric absorptions (broad due to pressure broadening) or stratospheric emissions (narrow due to Doppler broadening). Merges into the sub-millimeter range (i.e., less than 1000 μm).

- Microwave: 1 mm to centimetre ranges and radio. The opacity of aerosols is negligible, with much of the opacity formed by gases like ammonia, allowing spacecraft to probe deep into the atmospheres, below the clouds that obscure shorter wavelengths. Synchrotron radiation from gas giant radiation belts, and the difficulties of achieving spatial resolution on planetary discs at these long wavelengths, mean that these wavelengths are less commonly used.

Given the design of instruments, it is typical for a NEMESIS project to work in just one of these spectral ranges in isolation (i.e., it's very rare to have UV and infrared observations of the same atmospheric feature at the same time). NEMESIS therefore adopts slightly different strategies in each spectral range, although they all employ the same optimal estimation procedure at the heart.

2.2. Forward Modelling

Measurements of atmospheric structure and composition from satellite instruments requires a knowledge of the way the atmosphere absorbs, emits and scatters electromagnetic radiation. Different molecular species have characteristic spectral signatures in the infrared whose position, shape and strength can be measured in planetary spectra. These can then be compared to laboratory measurements to determine the temperature, pressure and composition of the region of atmosphere in which the line was formed. The theory of *radiative transfer* describes how the radiation field interacts with the atmosphere and allows the calculation of the spectral characteristics of radiation leaving the atmosphere as a function of the distribution of the gases, aerosols and temperature, a process known as *forward modelling*.

2.2.1. Molecular Spectroscopy

As radiation propagates through an atmospheric medium it is modified by absorption, emission and scattering of radiant energy on a microscopic scale. A molecule's energy state can be separated into several quantized degrees of freedom: electronic (movement of electrons between atomic energy levels), vibrational (stretching and bending of molecular bonds), rotational (depending on a molecule's moment of inertia) and translational states. At IR wavelengths, photons do not have the energy to interact directly with atoms to promote electronic transitions. However, the energy can instead promote transitions between different rotational and vibrational states within molecules (e.g., Irwin, 2003), resulting in the manifold of absorption and emission features observed in planetary IR spectra.

Rotational absorption features dominate the far-IR, as transitions between closely spaced rotational energy states can be produced by long-wavelength, low-energy photons. Vibration-rotation transitions require a higher energy, and bands of PH$_3$, NH$_3$ and CH$_4$, and of numerous hydrocarbons, occur at shorter wavelengths in the mid-IR (e.g. the methane ν_4 vibration-rotation band near 7 μm). The band-shape depends upon the population of states and the symmetry of the molecule, which determines the quantum-mechanical selection rules for allowed

transitions between states. For spherical or linear diatomic molecules there are typically three branches for the band - the central Q branch at the fundamental vibrational frequency where the rotational quantum number, J, is unchanged in the transition between energy levels; the P branch at lower frequencies where $\Delta J = -1$ and the R branch at higher frequencies where $\Delta J = +1$.

Whether these bands appear in absorption or emission depends on the environmental conditions of the line-forming region. Both processes occur simultaneously, resulting in both gains and losses of spectral radiance, L_ν, at a particular wavenumber. The radiative transfer equation incorporates all of these physical processes to calculate the net change in radiance and hence the transmitted radiation at the top of the atmosphere that might be measured by an orbiting spacecraft.

2.2.2. The Radiative Transfer Equation

A brief outline of radiative transfer is described here, but for a complete overview of the subject the reader is referred to Goody and Yung (1989) and Hanel et al. (2003). Lambert's law states that the attenuation of spectral radiance due to absorption, dL_ν^{abs}, along a path at zenith angle θ is given by;

$$dL_\nu^{abs} = -\alpha L_\nu = -L_\nu k_\nu \rho_a(z) \sec \theta dz \tag{1}$$

where α is the absorption along the slant path ($\alpha = k_\nu \rho_a dz/\mu$), $\rho_a(z)$ is the mean density of the absorber at altitude z, k_ν is the absorption coefficient for the medium and $\mu = \cos \theta$.

Defining $L_{\nu 0}$ as the radiant intensity at the base of the atmospheric model, z_0, then the upward radiation at the top of the atmosphere (z_1, where the optical depth is zero), is calculated by integrating over the altitude, dz;

$$L_\nu^{abs} = L_{\nu 0} \exp\left(-\int_{z_0}^{z_1} \frac{k_\nu(z)\rho_a(z)}{\mu} dz\right) \tag{2}$$

$$L_\nu^{abs} = L_{\nu 0} \exp\left(\frac{-\chi_\nu(z_0, z_1)}{\mu}\right) \tag{3}$$

$$L_\nu^{abs} = L_{\nu 0} \tau_\nu(\mu, z_0, z_1) \tag{4}$$

where the second line recasts the integral in terms of the dimensionless optical thickness between the two limits;

$$\chi_\nu = -\int_{z_0}^{z_1} k_\nu(z)\rho_a(z) dz \tag{5}$$

and the third line is in terms of the transmission (the proportion of the radiance which will reach the top of the atmosphere along this slant path);

$$\tau_\nu = \exp(-\chi_\nu/\mu) \tag{6}$$

.

In the absence of thermal emission, the last equality gives the total radiance leaving the top of the atmosphere. However, Kirchoff's law states that the emissivity of the layer will equal the absorptivity, hence the radiance emitted along the slant path is;

$$dL_\nu^{em} = +k_\nu \rho_a(z) J(\nu) dz/\mu \tag{7}$$

where $J(\nu)$ is known as the *source function*. Assuming that the atmosphere is in local thermodynamic equilibrium (LTE), the temperature of the Boltzmann distribution describing the population density of the excited vibrational and rotational energy levels of the molecule is equal to the kinetic temperature, T. Neglecting scattering[1], the emitted radiation is given by the Planck function at the local kinetic temperature;

$$B(\nu, T) = \frac{2hc^2\nu^3}{\exp(hc\nu/k_B T) - 1} \tag{8}$$

where ν is the wavenumber, h is the Planck constant, c is the speed of light and k_B is the Boltzmann constant. The approximation is valid where kinetic collisions are frequent in the lower atmosphere, so that the mean time interval between collisions is shorter than the radiative time constant (Andrews, 2000). Lower collision rates and the increasing importance of solar irradiation in the stratosphere and mesosphere allow emissions to have excitation temperatures different from the local kinetic temperature.

Assuming LTE, the source function is equivalent to $B(\nu, T)$. The radiance of Eq. 7 is also attenuated by the overlying atmosphere so that the contribution reaching the top is given by;

$$dL_\nu^{em} = \frac{k_\nu \rho_a(z) B_\nu(z) dz}{\mu} \exp\left(-\int_z^{z_1} \frac{k_\nu(z)\rho_a(z)}{\mu} dz\right) \tag{9}$$

$$dL_\nu^{em} = B_\nu(z) d\tau_\nu(\mu, z, z_1) \tag{10}$$

To calculate the radiance at the top of the atmosphere, accounting for both gains (Eq. 10) and losses (Eq. 4), we use the *Schwarzschild equation*, $dL_\nu = dL_\nu^{em} + dL_\nu^{abs}$, and integrate over all atmospheric layers;

$$L_\nu = L_{\nu 0} \tau_\nu(\mu, z_0, z_1) + \int_{\tau_\nu(z_0, z_1)}^{1} B_\nu(z) d\tau_\nu(\mu, z, z_1) \tag{11}$$

$$L_\nu = L_{\nu 0} \tau_\nu(\mu, z_0, z_1) + \int_{z_0}^{z_1} B_\nu(z) \frac{d\tau_\nu}{dz} dz \tag{12}$$

The first term represents the radiance from the lowest layer of the atmosphere (either the surface, or the bottom of the atmospheric model in the case of the giant planets), attenuated exponentially due to extinction by the absorbing medium. The second term gives the cumulative transmission between some level and the top of the atmosphere due to the opacity of each atmospheric level. The derivative $d\tau_\nu(z)/dz$ is known as the *transmission weighting function*, which peaks roughly where the optical depth is unity (Irwin, 2003). The product $B_\nu(z) d\tau_\nu(z)/dz$ is commonly referred to as the *contribution function*. The total radiance can be viewed as a weighted average of the black body radiance from each atmospheric level (Houghton et al., 1984). These equations are only applicable where scattering from aerosols is negligible, and must be substantially modified when scattering particles are present.

---

[1]Scattering is important at wavelengths shorter than 5 μm where reflected sunlight forms a large component of the radiation field. Aerosol-scattering can still lead to errors in the thermal-IR if cloud particles have an equivalent size to the wavelength of the radiation.

2.2.3. Sources of Opacity

Gaseous Absorption and Emission. Observations of the rotational and vibrational lines of molecules with permanent electric dipoles[2] allows us to deduce the mole fractions of gases in a planet's atmosphere from measurements of the spectrum. The absorption coefficient (k_ν) of a single line, centred at wavenumber ν_0, is a function of the linestrength, S, and the shape function, g;

$$k_\nu(T, \rho) = S(T)g(\nu - \nu_0, T, \rho) \tag{13}$$

where ρ is the density of the absorber in question and T is the temperature of the line-forming region. The line strength is the integration of k_ν over all ν. The shape function, g, depends on the physical conditions of the atmosphere. Spectral lines are not delta-functions, but instead have finite widths determined by various broadening processes: natural broadening (due to the uncertainty principal), temperature-dependent Doppler broadening, or by collisions with other molecules (pressure or Lorentz broadening, which is pressure- and temperature-dependent). The first broadening mechanism is always negligible; Doppler broadening dominates at low pressures and Lorentz broadening dominates at high pressures. In practice, the lineshapes observed are a combination of the latter two effects, convolving the Doppler and Lorentz lineshapes to produce a Voigt lineshape, which is generally used for giant planet atmospheres, with some exceptions. In gas giant atmospheres, molecular lineshapes can undergo both self-broadening (collisions with identical molecules) and foreign broadening (due to collisions with the H_2-He mixture), though the latter dominates because of the high probability of collisions with H_2 and He in gas giant atmospheres.

The opacity due to spectral lines for each pressure level, χ_g, is calculated within a wavenumber bin, $\Delta\nu$, surrounding a central wavenumber, ν_c, by summing the absorption α along a slant path, ($\alpha = k_\nu \rho_a dz/\mu$), over the whole wavenumber range. The absorption coefficient for each spectral line is calculated explicitly, and this in turn depends on the line strength and intensity, and on the environment of the line forming region (temperature, pressure, etc.). Strong lines near to the wavenumber bin are also considered in the calculation, as their wings may contribute a significant opacity. The calculated opacity is then summed over the N layers to give the total optical depth due to absorption and emission.

A number of databases list the molecular absorption lines of important gases, as well as the intensities, half-widths and temperature-dependent broadening parameters. The `NEMESIS` group uses lines from HITRAN (Rothman et al., 2005, 2013), GEISA (Jacquinet-Husson et al., 2005), and other sources such as EXOMOL.

Collision Induced Absorption (CIA). Symmetric, diatomic, homonuclear molecules like H_2 have no permanent dipole moment and would not otherwise interact with photons (aside from weak features associated with the electric quadrupole moment) because the excitations are formally forbidden. However, collisions between two non-polar molecules can lead to a deformation of the molecule and the production of a transient electric dipole. The breaking of molecular symmetry allows the otherwise forbidden transitions to occur, providing an electromagnetic interaction that results in a weak collision-induced absorption (CIA). The abundance of H_2 in gas giants is so high that their far-IR spectra are dominated by CIA opacity (χ_{CIA}) from H_2-H_2 and H_2-He collision pairs. Molecular pairs may be free (collisional complexes) or bound by van der Waals forces (e.g., dimers), and undergo bound-bound, bound-free or free-free transitions between rotational or translational states. As a consequence of the Uncertainty Principle, the short lifetimes of the states creates a broad continuum feature in planetary spectra (Irwin, 2003).

A. Borysow and collaborators have developed numerical models for the quantum mechanical lineshapes for collision pairs contributing to giant planet spectra, allowing a pre-tabulation of χ_{CIA} at a given temperature, pressure and frequency range. The absorption is then interpolated onto the exact T, p and mole fractions used by the code. Collision-induced absorption of H_2-H_2, H_2-He, H_2-CH_4 and CH_4-CH_4 pairs are included using CIA coefficients pre-calculated from the tabulations of Borysow (1991, 1993); Orton et al. (2007); Borysow et al. (1988); Borysow and Frommhold (1986, 1987) and references therein.

Aerosol opacity. The radiative effects of clouds and aerosols are governed by their microphysical properties and both absorption and scattering can modify the radiation field detected by an orbiting instrument. The molecular extinction coefficient, σ_ν^{ext}, is the sum of the cross sections due to scattering and absorption;

$$\sigma_\nu^{ext} = \sigma_\nu^{sca} + \sigma_\nu^{abs} \tag{14}$$

The effects of scattering on the thermal-IR are assumed to be negligible (Nixon, 1998; Irwin et al., 2004).

Transparent regions of the spectrum permit radiation from deeper, hotter levels to penetrate through, leading to higher radiances. Cloudy regions blocks radiation from below and lowers the radiance, obscuring emission and absorption features. The opacity due to these particles in the absence of scattering, χ_p, is given at each layer in the atmosphere by:

$$\chi_p(z, \nu) = \sigma_\nu^{abs} N(z) \tag{15}$$

where σ_ν^{abs} is the absorption cross-section (cm^{-2}) pre-calculated from Mie theory by the `Makephase` code for a single particle, $N(z)$ is the number density of the aerosols (particles/cm^2) within each layer of the atmosphere along a particular path.

2.2.4. Line by Line Models

The most accurate method of calculating the monochromatic L_ν at a particular wavenumber is to take all the opacity sources (χ_g from all the spectral lines contributing to the radiance at this wavenumber, χ_{CIA} and χ_p) and to integrate them over a

---

[2]Electric dipoles are required for a molecule to interact with the radiation field. Molecules without a permanent dipole will show no spectral features unless temporary dipoles are generated (by collisions, for example).

wavenumber range appropriate to the resolution of the instrument. This is accomplished using a radiative transfer code, `Radtrans`, which was developed in Oxford, principally by S. Calcutt and P. Irwin. Once the atmospheric profiles have been defined, the atmosphere is split up into layers equally spaced in $\Delta \log(p)$ so that mean properties can be evaluated along different slant paths through the atmosphere. `Radtrans` calculates the opacity from each layer of the atmosphere and sums them along the path to obtain the transmission at the top of the atmosphere, but such a calculation can have a high computational burden. There are faster, alternative methods of simulating finite resolution spectra.

2.2.5. Correlated-k

One such alternative for calculating finite resolution spectra is to replace the rapidly-varying absorption-coefficient function, $k(\nu)$, with a smoother function of the absorption coefficients which is more easily integrated - the k-distribution - speeding up computations at the expense of diminished accuracy Goody and Yung (1989). The mean transmission ($\overline{T}(m)$) for a path of absorber amount m in a homogeneous atmosphere of uniform pressure and temperature, is given by;

$$\overline{T}(m) = \frac{1}{\Delta \nu} \int_{\nu}^{\nu + \Delta \nu} \exp(-k(\nu)m)d\nu \tag{16}$$

where the $k(\nu)$ contains the sum of all individual line contributions to the absorption coefficient at that particular wavenumber. As $k(\nu)$ is a rapidly varying function of wavenumber, a fine $d\nu$ is required to evaluate the transmission numerically. However, within a particular interval, the absorption coefficients may be replaced by a frequency distribution of the absorption coefficients, $f(k)$. This allows Eq. 16 to be rewritten as;

$$\overline{T}(m) = \int_{0}^{\infty} \exp(-km)f(k)dk \tag{17}$$

Here the transmission is independent of the ordering of the absorption coefficients, which can be reordered from lowest to highest absorption strengths without affecting the calculated transmission. In essence, it is sufficient to know what fraction of the wavenumber range is occupied by a particular absorption coefficient, without knowing exactly at what ν it occurs at (Irwin, 2003; Lacis and Oinas, 1991).

Specifying the cumulative frequency distribution, $g(k)$, as;

$$g(k) = \int_{0}^{\infty} f(k)dk \tag{18}$$

and noting that $k(g) = g^{-1}(k)$ (because $g(k)$ is a monotonic function of k), the transmission can be rewritten as;

$$\overline{T}(m) = \int_{0}^{1} \exp(-k(g)m)dg \tag{19}$$

where $k(g)$ is known as the k-distribution. The equation can be discretized by dividing the k-distribution into N intervals (Gaussian quadrature points) such that the shape of the distribution is sufficiently sampled to calculate the mean transmission as;

$$\overline{T} = \sum_{i=1}^{N} \exp(k_i m) \Delta g_i \tag{20}$$

where Δg_i are the quadrature weights for each of N quadrature points at which k_i is calculated within the spectral band. Given a sufficient number of Gaussian quadrature points for accurate sampling, integration of the k-distribution offers an sufficiently-accurate alternative to integrating Eq. 16 over all ν. k-distributions can be pre-calculated for a range of temperatures and pressures and stored in look-up tables for rapid interpolation and calculation of the transmission along a homogenous path. Absorption lines of different gases may be assumed to be uncorrelated (Irwin, 2003) so that the k-distributions can be combined.

However, real atmospheres are inhomogeneous, so that the pressure and temperature varies rapidly with position. To circumvent this problem in monochromatic calculations, the path would be split into sub-paths whose mean transmissions are multiplied together at the end. This becomes more complex when k-distributions are used, but fortunately we can assume that frequency-regions of high and low absorption in the k-distributions are correlated between layers at all locations along an inhomogeneous path (Lacis and Oinas, 1991). As transmissions are approximately correlated between atmospheric levels, we can multiply k-distributions together as if they were monochromatic, so that the combined mean transmission (splitting the inhomogeneous path into M subpaths) is;

$$\overline{T} = \sum_{i=1}^{N} \exp\left(-\sum_{j=1}^{M} k_{ij}m_j\right) \Delta g_i \tag{21}$$

This is the correlated-k approximation, which may be applied to both thermal emission and scattering atmospheres (Irwin, 2003).

Given the linedata, a reference atmosphere and a physical model for calculating the opacity at each wavenumber, synthetic spectra can be generated for comparison with observations. But to *retrieve* atmospheric parameters, we must utilise a retrieval algorithm to produce a good fit to the spectrum - the next section.

2.3. Inverse Modelling - Spectral Retrieval

2.3.1. The Retrieval Problem

The most direct approach in the analysis of planetary spectra is to calculate synthetic spectra for a particular atmospheric structure and composition, known as forward-modelling. Retrieval theory concerns the inverse method - calculating the *plausible* atmospheric structures which would result in the observed radiance and modifying the atmospheric structure to minimise the difference between synthetic spectra and observations. The problem is ill-conditioned, as relatively small changes in observed radiance can be amplified into finite changes in the retrieved profile, so even small measurement errors can have catastrophic effects. It is also ill-posed and under-constrained, as we are attempting to derive continuous profiles using discrete sets of measurements. In fact, there may be an infinite manifold of solutions which can satisfy the measured spectra within the error! Retrieval theory allows us to select the most likely solutions based on our prior knowledge of the physics and chemistry of the atmosphere.

The Forward Model. Consider a set of *n* atmospheric parameters to be determined (the state vector, $\mathbf{x}$, which may be a continuous atmospheric profile of temperature or composition, or a scaling factor of a known profile) from a set of *m* measurements of radiance over a range of wavenumbers and viewing geometries, $\mathbf{y}$. The measurements are linked to the atmospheric state via the forward model, $\mathbf{F}(\mathbf{x})$;

$$\mathbf{y} = \mathbf{F}(\mathbf{x}) + \epsilon \tag{22}$$

The forward model spectrum is calculated by considering the physics of the atmosphere, the line data and the effects of the instrument lineshape and field of view. ϵ is a vector containing the estimates of the noise on each of the measured radiances, with a measurement error covariance matrix, $\mathbf{S}_\epsilon$, defined as the expectation value of the error $E(\epsilon\epsilon^{\mathbf{T}})$. Our goal is to gain as much information as possible about the atmospheric state given the measurements.

Linearisation. We can linearise the solution around a reference state, $\mathbf{x_0}$, by Taylor expanding so that;

$$\mathbf{y} - \mathbf{F}(\mathbf{x_0}) = \frac{\partial \mathbf{F}(\mathbf{x_0})}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{x_0}) + \epsilon = \mathbf{K}(\mathbf{x} - \mathbf{x_0}) + \epsilon \tag{23}$$

where $\mathbf{K}$ is the $m \times n$ sensitivity kernel (also known as the Jacobian or functional derivative), containing the partial derivatives of each forward-model calculated spectrum with respect to all state vector elements, $K_{ij} = \frac{\partial F_i(\mathbf{x_0})}{\partial x_j}$.

Least-squares Solutions. With a complete knowledge of the Jacobian from the physics of the system, it is simple to extract the required state vector from Eq. 23. The simplest possible method of fitting the measurements is to use a least-squares fit, minimising a quadratic cost function;

$$\chi^2 = (\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}}))^T \mathbf{S}_\epsilon^{-1}(\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})) \tag{24}$$

where $\hat{\mathbf{x}}$ is an estimate of the true state vector, $\mathbf{x}$. The solution for the state vector (Rodgers, 2000) is then;

$$\hat{\mathbf{x}} = \left(\mathbf{K^T S}_\epsilon^{-1}\mathbf{K}\right)^{-1} \mathbf{K^T S}_\epsilon^{-1}\mathbf{y} \tag{25}$$

with covariance matrix;

$$\hat{\mathbf{S}} = \left(\mathbf{K^T S}_\epsilon^{-1}\mathbf{K}\right)^{-1} \tag{26}$$

The least squares method works best when the number of unknown parameters is less than the number of independent measurements and when the measurements contain sufficient information to retrieve each parameter independently. However, with more parameters to fit than measurements, this is an ill-posed problem. The retrieval problem does not have one unique solution, instead a physically-realistic solution must be selected from a range of possible answers. The least-squares method will be unstable, with physically unrealistic oscillations in retrieved profiles. To select physically-meaningful solutions from the family of non-unique solutions requires additional constraints, achieved by *optimal estimation*.

2.3.2. *Optimal Estimation and Tuning*

Linear Optimal Estimation. Additional constraints may be introduced by incorporating 'virtual measurements,' such as independent measurements of the atmospheric structure from different experiments, or estimates derived from previous studies, providing an *a priori* state vector, $\mathbf{x_a}$, with error covariance matrix, $\mathbf{S_a}$. The inclusion of these additional constraints is known as 'optimal estimation,' and creates a modified cost-function (Irwin et al., 2004);

$$\phi = (\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}}))^T \mathbf{S}_\epsilon^{-1}(\mathbf{y} - \mathbf{F}(\hat{\mathbf{x}})) + (\hat{\mathbf{x}} - \mathbf{x_a})^T \mathbf{S_a}^{-1}(\hat{\mathbf{x}} - \mathbf{x_a}) \tag{27}$$

The first term requires that the synthetic spectra must give a close fit to the measured spectrum, the second term requires the solution to lie close to an *a priori* profile. *Retrieval tuning* allows the user to specify how closely the measurements must be fitted, and how reliable the user deems the *a priori* estimates to be.

For linear models, Rodgers (2000) show the optimal-estimation solutions for the state vector becomes;

$$\widehat{\mathbf{x}} = \mathbf{x_a} + \left(\mathbf{K^T S}_\epsilon^{-1}\mathbf{K} + \mathbf{S_a^{-1}}\right)^{-1} \mathbf{K^T S}_\epsilon^{-1}(\mathbf{y} - \mathbf{Kx_a}) \tag{28}$$

$$\widehat{\mathbf{x}} = \mathbf{x_a} + \mathbf{G}(\mathbf{y} - \mathbf{Kx_a}) \tag{29}$$

where $\mathbf{G}$ is the 'gain matrix,'

$$\mathbf{G} = \left(\mathbf{K^T S}_\epsilon^{-1}\mathbf{K} + \mathbf{S_a^{-1}}\right)^{-1} \mathbf{K^T S}_\epsilon^{-1} \tag{30}$$

The gain matrix is the derivative of the estimated state with respect to the measurement ($\mathbf{G} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{y}}$), and gives the contribution of each measurement in $\mathbf{y}$ to the estimated state vector.

Non-Linear Optimal Estimation. The retrieval problem is generally non-linear, because small changes in the composition can strongly affect the atmospheric transmission and $\mathbf{K}$ (Irwin, 2003). So linearising the solution tends to lead to large errors, greater than the desired accuracy of the solution. Instead a solution must be obtained iteratively, and the Oxford `NEMESIS` (**N**on-Linear Optimal **E**stimator for **M**ultivariat**E** Spectral Analy**SIS**) retrieval algorithm uses Newtonian iteration to converge to a solution (Irwin et al., 2004). The optimum solution is;

$$\mathbf{x_{n+1}} = \mathbf{x_a} + \mathbf{S_a}\mathbf{K_n}^T \left(\mathbf{K_n S_a K_n}^T + \mathbf{S}_\epsilon\right)^{-1} ((\mathbf{y} - \mathbf{y_n}) - \mathbf{K_n}(\mathbf{x_a} - \mathbf{x_n})) \tag{31}$$

$$= \mathbf{x_a} + \mathbf{G_n}(\mathbf{y} - \mathbf{y_n}) + \mathbf{A_n}(\mathbf{x_a} - \mathbf{x_n}) \tag{32}$$

where the subscript *n* refers to the matrix or state vector calculated in the previous iteration, and $\mathbf{A_n} = \mathbf{G_n K_n}$ is the averaging kernel matrix.

Vertical Smoothing. To understand the role played by the averaging kernel, we consider a case where the measurements, $\mathbf{y}$, have zero error (ϵ), so that the Gain matrix $\mathbf{G}$ is zero. The optimal solution of Eq. 32 becomes;

$$\mathbf{x_{n+1}} = \mathbf{x_a} + \mathbf{A_n}(\mathbf{x_a} - \mathbf{x_n}) \tag{33}$$

i.e. a weighted mean of the *a priori* state and the retrieved estimate of the state. Hence the averaging kernel is a measure

of the sensitivity of the retrieval to the true atmospheric profile (the 'resolution' of the retrieval), which depends upon how the true profile is smoothed by the *a priori*, defined as $\mathbf{A} = \frac{\partial \hat{\mathbf{x}}}{\partial \mathbf{x}}$. Smoothing is provided in $\mathbf{S_a}$ by having adjacent pressure levels correlated with one another (Rodgers, 2000; Conrath et al., 1998), where off-diagonal elements (representing the correlation between different pressure levels) decrease away from the diagonal according to a correlation length, l_c, the length scale in units of $\ln(p)$ over which the inter-level correlation drops off by $1/e$ (p.55, Rodgers, 2000);

$$\mathbf{S_a(i,j)} = (\mathbf{S_a(i,i)S_a(j,j)})^{1/2} \exp\left(-\frac{\left|\ln(p_i/p_j)\right|}{l_c}\right) \quad (34)$$

For example, by testing a range of possible values for l_c on Saturn, a value of 1.5 was selected for the vertical smoothing, as it represents smoothing over approximately one atmospheric scale height, similar to the maximum vertical resolution for retrievals from nadir spectra. The final precision of the retrieval depends upon the extent of the smoothing, but the problem has been successfully reposed into a constrained form - the atmosphere has been represented by a finite set of parameters, smoothed according to the averaging kernel. The smooth constraint is a reasonable physical assumption, as atmospheric properties can be assumed to vary smoothly with altitude.

Braking Parameter. Non-linear retrievals could potentially become rapidly unstable, with the solution to each iteration rapidly diverging from the optimal solution so that the cost function, ϕ, increases without bound. A Levenberg-Marquardt type braking parameter, a_λ, is used to stabilise the model (Irwin et al., 2004);

$$\mathbf{x'_{n+1}} = \mathbf{x_n} + \frac{\mathbf{x_{n+1}} - \mathbf{x_n}}{1 + a_\lambda} \quad (35)$$

a_λ is initially set to one, and altered according to the change in the cost function ϕ at each stage of the iteration. If ϕ decreases then $\mathbf{x_n}$ is updated to $\mathbf{x'_{n+1}}$ and a_λ is reduced by an arbitrary factor of 0.3. If ϕ increases (the solution is attempting to diverge) then $\mathbf{x_n}$ is left unchanged and a_λ is multiplied by an arbitrary factor of ten. The estimate $\mathbf{x_n}$ is used in the forward model to generate a new (a) synthetic spectrum, (b) Jacobian $\mathbf{K}$, (c) gain matrix $\mathbf{G}$, and (d) averaging kernel, $\mathbf{A}$. The retrieval is deemed to have converged when ϕ alters by less than 10% (an arbitrary limit), a_λ tends to zero and the solution tends to the optimal estimate. With each iteration, the state vector is adjusted so that eventually the only remaining difference between the forward model and the measurements is the random noise on the measurements.

Tuning Retrievals. When the errors on the measured spectrum are large, the solution will be weighted towards lying close to the *a priori*. When the measurement errors are smaller than the constraints imposed by the *a priori*, the solution tends to the ill-conditioned 'least-squares' estimate, often with non-physical oscillations due to the lack of smoothing. The *a priori* profiles are often poorly constrained, and are used primarily to provide smoothing to prevent the occurrence of non-physical, high-frequency oscillations in the retrieved profiles. In this respect,

the NEMESIS retrieval algorithm is similar to the constrained linear inversion technique used by Conrath et al. (1998).

2.3.3. Retrieval Error Analysis

The formal error on retrieved temperature and composition from a planetary spectrum must take numerous sources of error into account. Random measurement noise, ϵ, which is the Noise Equivalent Spectral Radiance (NESR) of the measurement, is a source of measurement error (assumed to be uncorrelated between measurements). This is reduced by coaddition of the appropriate number of measurements.

However, other errors may arise from systematic errors in our modelling of the spectra, summarised below.

- Incomplete physical description of the radiative transfer: for example, the omission of scattering of radiation in our forward models, or the use of the Voigt lineshape to describe the broadening of the spectral features. The modelling of the collision-induced continuum may also be in error (Orton et al., 2007), and unaccounted-for processes like fluorescence may play a role.

- Inaccurate linedata: poor reproduction of spectral lineshapes might result from omission of weak spectral lines from the laboratory measurements, inaccurate intensities or broadening parameters, or poorly characterised partition functions (a cubic polynomial expression is fitted to the partition function data).

- Instrument Calibration: Data reduction requires the removal of effects related to electrical interferences, ripples, instrument lineshapes, spectral or radiometric calibrations is required before we can accurately reproduce the radiance of a planet's atmosphere.

- Correlated-k: The accuracy of this approximation depends on how well the distribution of absorption coefficients is sampled by the chosen number of Gaussian quadrature points. Too few points, and the k-distribution is not being accurately represented, but too many points will increase the computational burden. The correlated-k approximation has difficulties simulating very small features, for which line-by-line techniques should be employed to obtain high precision estimates of abundances.

- Instrument pointing errors: the orbital geometry of an observation is vital to its interpretation - incorrect latitudes, longitudes and viewing geometries will lead us to simulate the wrong radiance to fit the observed measurements.

- Line-of-sight effects: at high viewing geometries, the slant path through the atmosphere may be sampling regions of different atmospheric composition and temperature, leading to spurious results when a single atmospheric column is considered. However, the large ratio of a planet's scale height (indicating the vertical resolution) to its radius implies that the composition may not be changing significantly along these high-inclination slant paths. It is more likely to be a problem on terrestrial planets and Saturn's

moon, Titan, where a full two-dimensional retrieval algorithm (which takes both altitude and position into account) should be applied to such observations.

- Cumulative errors: the temperatures, para-hydrogen fraction and aerosol opacity of the line-forming region must be constrained before gaseous absorption and emission features of other gases can be fitted. Errors in $T(p)$ and the other retrieved quantities from the first stage of the retrieval are incorporated into the errors on mole fractions retrieved in the second stage.

These uncertainties are represented by adding a forward-modelling error ϵ_f to the measurement error ϵ_m, so that $S_\epsilon(i, i) = \epsilon_m(i)^2 + \epsilon_f(i)^2$. Random measurement errors on $\mathbf{y}$ are assumed not to be correlated with each other, so that S_ϵ can be taken as diagonal. The contribution of the measurement error term to the final error on the retrieved state vector, $\mathbf{x}$, is given by (Rodgers, 2000);

$$\mathbf{S}_m = \mathbf{G}\mathbf{S}_\epsilon\mathbf{G}^T \tag{36}$$

Similarly the smoothing could introduce bias towards *a priori* models, and cause unnecessary removal of real vertical structures, so the smoothing error term is given by;

$$\mathbf{S}_n = (\mathbf{A} - \mathbf{I})\mathbf{S}_a(\mathbf{A} - \mathbf{I})^T \tag{37}$$

The total retrieval error is then the sum of the two terms, $\mathbf{S}_t = \mathbf{S}_m + \mathbf{S}_n$. Substituting the expressions for $\mathbf{A}$ and $\mathbf{G}$ (Eq. 30 and $\mathbf{A_n} = \mathbf{G_n}\mathbf{K_n}$) into $\mathbf{S}_t$ gives the formal retrieval error as;

$$\widehat{\mathbf{S}} = \left(\mathbf{K^T}\mathbf{S_\epsilon^{-1}}\mathbf{K} + \mathbf{S_a^{-1}}\right)^{-1} \tag{38}$$

Unless otherwise stated, the errors quoted in many articles are derived directly from the diagonals of this formal retrieval error matrix. These errors represent the range of possible values that could fit the observed spectrum. However, some caution should be applied to these uncertainties - it is far better to adopt a full Monte Carlo approach, varying the *a priori* assumptions to explore the range of parameter space that can truly fit a spectrum. For example, if we retrieve a gas by simply scaling the vertical profile, we should remember that alternative vertical profiles might fit the data equally well but with a very different scaling factor! In that case, the small uncertainties on the retrieved scale factor do not adequately represent the uncertainty on the fit.

2.3.4. Functional Derivatives and Vertical Sensitivity

The partial derivative, $\frac{\partial L_\nu}{\partial x_i}$, of the spectral radiance with respect to perturbations of the some model parameter, x_i, at each pressure level is called the functional derivative. This determines the sensitivity of the observed spectra to variations of a given parameter at a particular altitude, which gives an idea of the information content of the spectra. A functional derivative of temperature at a given wavenumber is the same as a 'contribution function.' Two dimensional plots of the functional derivative as a function of wavenumber/wavelength reveal how the altitude sensitivity varies as we move in and out of absorption/emission features.

2.4. A Brief History of NEMESIS

The `Radtrans` model calculates the transmission, absorption and thermal emission spectra of gaseous paths using line-by-line, band model or correlated-k techniques. It can compute multiply-scattered spectra to study cloud reflectvity, and hemispherically-averaged spectra for use in Brown Dwarf and exoplanetary studies. The `Radtrans` software employed today started life in 1994, but was a development of earlier code called `GENLBLm` a general-purpose line-by-line transmission calculataion routing written by Dr. Simon Calcutt. The `Radtrans` input and output files are largely shared with `NEMESIS`, which is built upon the `Radtrans` libraries and subroutines.

`NEMESIS` itself was modified by P. Irwin from retrieval codes used for Galileo NIMS in the late 1990s, and overhauled for modelling of Cassini CIRS in the early 2000s (hence some of the forward model subdirectories still feature the CIRS name). Over the years it has been updated and modified to handle all of the forward model types represented by `Radtrans`, from the fast gradient-only method for non-scattering calculations to the slower scattering calculations and, most recently, to exoplanetary transit observations. The retrieval model's name comes from the rather excellent acronym: **N**on-Linear Optimal **E**stimator for **M**ultivariat**E** Spectral Analy**SIS**.

The `NEMESIS` retrieval algorithm is general-purpose, without being hardwired to any particular planet. It can invert spectra to estimate continuous profiles of temperatures, composition and clouds; use either nadir or limb viewing geometries; add thermal emission from solid surfaces if necessary; retrieve from multiple viewing geometries simultaneously; average over a finite field of view; perform scattering calculations; use filter functions to replicate photometry; use line-by-line calculations where correlated-k calculations are insufficient; and has been recently extended to model exoplanets in primary and secondary transits. Updates continue to be applied - most recently the incorporation of pre-tabulated monochromatic gas absorption and Markov-Chain Monte Carlo techniques.

Nemesis is traditionally known as the Goddess of Vengeance and Retribution. However, the word Nemesis originally meant the distributor of fortune, whether good or bad, in due proportion to each man according to his deserts. It is the ?goddess of fortune? view of Nemesis, which inspired the naming of this retrieval code in her honour. It is hoped that `NEMESIS` will bring good fortune and will considerably improve the retrieval of atmospheric properties from remotely-sensed infrared planetary spectra.

`NEMESIS` has been successfully applied to studying (among others):

- Jupiter's phosphine and ammonia distribution from Cassini CIRS (Irwin et al., 2004; Fletcher et al., 2009a) and VIMS (Giles et al., 2015).

- Titan's stratospheric nitriles (Teanby et al., 2006) and oxygen compounds (de Kok et al., 2007)

- Venus' CO distribution (Tsang, 2007) and Mars' dust and water distributions.

- Saturn's atmospheric temperatures (Fletcher et al., 2007, 2015), composition (Fletcher et al., 2009b) and dynamics (Fletcher et al., 2008, 2011b) from Cassini CIRS (Fletcher, 2007) and VIMS (Fletcher et al., 2011a).

- Ice giant atmospheric structure (Fletcher et al., 2014) and cloud structure (e.g., Irwin et al., 2017).

- Spectroscopy of exoplanet transits (Lee et al., 2012; Barstow et al., 2014).

3. Code Structure

The suite of spectral modelling and inversion software is written in FORTRAN, with additional pre- and post-processing tools often written in IDL. Examples of the former include the preparation of sources of opacity data; examples of the latter include the plotting and mapping of retrieved properties and spectral fits. In this section we provide a broad overview of the code. The main `radtrancode` directory is now maintained under version control with GITHUB. However, for new starters it is wise to simply make a copy of the working code and to run through the installation. The code works on 32- and 64-bit machines, and with both the Intel Fortran Compiler and the free GFORTRAN. A README file is available in the top-level directory that explains the process, but there are some pre-requisites:

- Ensure that the necessary paths are set for the creation of the libraries and binaries that result from the compilation. This should happen in your `.bashrc` or `.cshrc` file, depending on the shell you're using.

- The software uses numerical recipes fortran programs (e.g., those for calculation of matrix mathematics) that are standard, but need to be compiled first.

- Another pre-requisite, although it is unusual by the majority of `NEMESIS` users, is compiling the libraries within the `FOVgreg` directory, which is only really needed for Mars Climate Sounder, but the installation complains if it isn't present.

Once these pre-requisities are met, there are three major components of the software: raddata, radtran and nemesis, as outlined below.

3.1. Raddata

The most straightforward directory is `raddata`, which is simply a repository for the reference files used by the NEEMESIS and Radtrans subroutines. Here you'll find collision-induced absorption spectra (i.e., the continuum underlying infrared spectra), solar spectra for reflected sunlight calculations, gravitational parameters for each planet so that the temperature-pressure grid can be properly calculated from hydrostatic equilibrium; saturated vapour pressure information to understand the formation of clouds; and stellar spectra for some known exoplanet host stars.

Ultimately it might be wise to have spectroscopic line data and k-distributions also contained within this reference directory. These are located elsewhere in peoples file systems (and each institution uses slightly different libraries, depending on their focus).

3.2. Radtran

`Radtrans` is the forward model that underlies the `NEMESIS` inverse model, and so this directory and all its subdirectories are required before an inverse calculation can be performed. Here we briefly outline the content of each of the subdirectories, with the exception of `makefiles` (which contains the files needed to generate all the libraries and executables using FORTRAN), `includes` (which contains files of array size definitions and block definitions that are used throughout the code), `matrices` (which contains matrix manipulation routines) and `monteck` (unused code for calculating scattered spectra using a Monte Carlo method).

- `ciatable`: Contains the programs for generating collision-induced opacity tables. Some of these are the original subroutines of Borysow and colleagues, adapted to be run by `Makeciatable` (for generating CIA tables without distinction between ortho- and para-hydrogen), `Makeisotable` (same as previous, but adapted to longer wavelengths), `Makefptable` (for including para-hydrogen) and `Makefptable_allcia` (for combining para-hydrogen with all other sources of opacity. The resulting CIA files are then stored in `raddata` for use in the spectral simulations.

- `cirsrad`: Subroutines for scattering calculations that cannot use the implicit calculation of Jacobians (i.e., the gradient version of the code). The two executables are `CIRSdrv_wave` (a front-end to the correlated-k calculations) and `Lbldrv_wave` (for the line-by-line calculation), in case these need to be used as forward models outside of `NEMESIS` itself.

- `cirsradg`: Similar to the previous directory, but for those calculations that are non-scattering and can make use of the implicit calculation of the Jacobians (i.e., the gradient approach). The executable `CIRSdrvg_wave` provides a front-end for a correlated-k calculation.

- `path`: The distribution of temperature, aerosols and gaseous species along an atmospheric path determines the top-of-atmosphere radiance, so this subdirectory contains a series of subroutines for creating and manipulating those paths. They include `Profile` for setting up the temperature, pressure and composition grid (the `*.ref` and `*.prf` files); `Parah2_profile` for the para-hydrogen profile; `Dust_profile` for the specific aerosol concentration (i.e., the number of aerosol particles per gram of atmosphere); and a number of others that might be of use during retrieval testing. The code `Path` allows the user to read in a `*.pat` file, generate the layer temperatures and amounts, and output to a `*.drv` file.

- radtran: This is the core subdirectory for radiative transfer calculation routines, and a number of them are also shared with the cirsrad(g) and nemesis directories. Executables are available to calculate and manipulate *k*-distributions from a variety of different sources (Calc_fnktablec is one of the most current for non-exoplanet applications); as well as for reading and plotting the output of Radtrans itself. The input files (spectral databases and the driver file *.drv created by Path) for a Radtrans run are described in the manual - in many cases, most users simply run Nemesis with zero iterations in order to compute a forward model. The Radtrans executable is among the oldest, developed by Calcutt in the 1980s and expanded by Irwin in the 1990s to combine line-by-line, band and correlated-k models.

- rtm_util: A series of general-purpose utility routines used throughout the software are contained in this directory - no executables, but things like interpolation codes and line shape codes are all found here, including a file called datarchive.f which simply points to the users raddata directory.

- scatter: Directory containing subroutines used for scattering calculations throughout the code, including executables like Makephase (used to calculate the phase functions, cross sections and single scattering albedoes of spherical particles via Mie theory), Readphase (for reading in phase functions output earlier), and Normxsc (for normalising an aerosol cross-section to an optical depth at a specific wavenumber). Makephase was based on a pre-1994 code by Kamp and Collard, and also fits Henyey-Greenstein functions to the calculated phase functions to speed up calculations.

- spec_data: The final subdirectory contains subroutines and programs for manipulating line data, band data and correlated-k data. These are typically only used when assembling the databases in the first place, and key executables include Makedb (for making the line-by-line direct access data base from a variety of line lists, ahead of generating k-tables), Pl_lines (for plotting individual lines from a database) and Select (for copying a subset of line data from one database into a new one).

The sub-directories listed above represent the core of the forward model and all the tools necessary to manipulate the opacity sources and the atmospheric profiles, as well as computation of the spectral transmission. The NEMESIS inversion algorithm was built over the top of all these codes.

3.3. NEMESIS

The final directory is the nemesis directory, which simulates spectral radiance in either wavenumber (W cm-2 sr-1 (cm-1)-1) or wavelength (W cm-2 sr-1 micron-1) space. There are a few distinct flavours of Nemesis depending on the particular observation geometry: NemesisL for limb-observing geometries, NemesisMCS for Mars Climate Sounder (which required additional information on the field-of-view); Nemesisdisc for disc-integrated spectra or secondary transit observations (either in spectral irradiance W cm-2 um-1 or as a planet/star flux ratio if a solar reference file is included); NemesisPT for primary transit spectra (output units are 100*planet_area/stellar_area); and NemesisMC for the Monte Carlo scattering scheme. There's also a general purpose program for generating test spectra Generatespx, but many users simply use Nemesis as a forward-model and extract the spectrum from the output files.

The basic process of a NEMESIS calculation is as follows. The code reads in the input files to identify the type of calculation required, the sources of opacity (line by line, correlated-k, etc.), and any additional reference files (like stellar files). It then reads in the measured spectrum with readnextspavX.f and adds any forward modelling uncertainty as required to calculate the measurement covariance matrix $\mathbf{S}_\epsilon$. Next, it reads the inputs to set up the *a priori* and its covariance matrix $\mathbf{S}_a$ using readapriori.f. These are all passed to the subroutine coreret.f where the core retrieval is performed, calling the particular forward model (forwardXXX.f) relevant to the desired calculation, which calculates the spectrum using the forward models like cirsrad(g) in the radtrans directories. The 'g' stands for the gradient version of the forward model, which internally calculates the partial derivatives (the Jacobian matrix $\mathbf{K}$) of the synthetic spectra with respect to the model parameters, rather than calculating numerous perturbed spectra and taking the difference, which is much slower.

The output spectra and the $\mathbf{K}$ matrices and covariance matrices are passed to subroutines that calculate the gain matrix, cost function ϕ, and goodness of fit to the input spectrum. This loop is iterated until convergence is reached, using the Marquardt-Levenberg braking parameter, and then the outputs written into a series of files for plotting. But before we discuss the outputs, we'll now review the basic operation of NEMESIS.

4. Running NEMESIS

The detailed operation of NEMESIS is best learned via trial-and-error, manipulating the input files to see what effect they have on the output calculated spectrum. To that end, a series of template directories should be available for you to play with. Users generally have their own mechanisms for running NEMESIS, with some developing pipelines in IDL and Python to distribute large batch jobs to supercomputer clusters. But the basic operation remains the same - generating input files, running NEMESIS, then using code in IDL or Python to plot and interpret the output. Although some codes in IDL are available to read NEMESIS output, it is often best to develop these independently so that the user understands all the individual steps.

Note that the details below are not meant to replace the NEMESIS manual. All input files listed below usually start with a runname - this can be anything you choose, so I've represented it as a '*'.

4.1. Input Files

Spoiler alert - there are far too many input files to realistically keep track of, but almost all of them need to be there. Some are far more important than others, and some are only relevant to a very niche-case use of NEMESIS. They subdivide into four categories: those setting up the atmospheric structure for the planet of interest; those telling NEMESIS how to run; those specifying the sources of opacity; and those specifying the measurement. For many of these, you'll simply copy the input files across from a previous directly, as they don't all change from run to run (in fact, some you'll set up once and then never need to alter).

4.1.1. Atmospheric Inputs

NEMESIS is not hard-wired to a particular planet, so needs to be given the basic atmospheric structure as a starting point for fitting a spectrum. The most important file is the *.ref file (and its associated intermediate *.prf file), which contains the reference height, pressure, temperature and composition grid for NGAS gases on NPRO levels. This file also has a flag to say which planet we're looking at (Mars is 4, Neptune is 8, etc.), and a list of identifiers for the gases to be included in the reference model. Each gas has a unique identifier that can be found in the Radtrans manual, and a flag associated with it to say whether all isotopologues are combined (flag=0), or whether they are considered separately (i.e., methane might have flag=1 for normal CH_4, flag=2 for $^{13}CH_4$, and flag=3 for CH_3D). This *.ref file is the first place to look when seeking to modify temperatures and composition ahead of a spectral calculation.

Additional files include the parah2.ref file containing the vertical distribution of para-H_2 (only needed if you are modelling giant planets), the aerosol.ref file (the vertical distribution of aerosols for however many cloud profiles are being used), and the fcloud.ref file (fractional cloud cover profile specifying whether we have a uniform haze or thicker clouds with gaps, only needed if a scattering run is being performed). The *.set file contains information on the scattering angles to be used and how the atmosphere should be split into layers - these are specific to a particular planet, containing information on the distance to the planet, the surface temperature, the number of layers, etc. The *.vpf file contains a list of the gases that are to be limited by saturation profiles rather than be freely varying. For terrestrial planets, a *.sur file can be used to specify the surface emissivity as a function of wavelength. Together, these define the reference model atmosphere, and tells NEMESIS which planet it is looking at.

One further set of files that may be required are the *a priori* files if you are retrieving a continuous vertical profile. For example, when retrieving temperature you need a tempapr.dat file containing the reference temperature profile and an estimate of the *a priori* uncertainty as a function of altitude. This can be a copy of the $T(p)$ profile from the *.ref file, but is needed as the starting point for the inversion. It is not needed if you are simply retrieving a scaled or parameterised profile.

4.1.2. NEMESIS Flags

The second category of input files tells NEMESIS how the calculation should be run. Again, rather than being combined into a single input file, these are spread across multiple text files depending on the type, as described below:

- Input file *.inp: This file is a list of flags allowing you to specify whether wavenumber or wavelength space is to be used (ISPACE), whether to use multiple scattering (ISCAT) and therefore the non-gradient forward model; whether to use line-by-line instead of ktables (ILBL), whetrher to add any wavenumber offsets (WOFF); whether to add forward modelling error; the number of iterations (NITER, set to zero for a forward model, 15-39 otherwise); the percentage convergence limit (PHILIMIT, usually around 0.1); the number of retrievals to perform (NSPEC); and whether or not to read in the results of a previous retrieval (LIN). The latter is used if you already have the results of, say, a temperature inversion, and you wish to then use these to retrieve a gaseous abundance. Finally, the IFORM flag defines the units of the output spectrum, either as spectral radiance, planet/star flux ratio, spectral irradiance, etc.

- Prior file *.apr: When you run a retrieval, NEMESIS will calculate the spectrum based on the contents of the reference profile but modified with the profiles you're actually retrieving. In this file, you can specify exactly what you wish to retrieve, whether it be a continuous temperature or aerosol profile, or a parameterised or scaled gaseous abundance. The number of variable profiles NVAR is typically 1-5. Gases and clouds can be parameterised using parameters like deep abundances, transition pressures, and vertical gradients (fractional scale heights), with each of these parameters assigned a first guess to allow the calculation of the first spectrum in the iteration scheme. For continuous profiles like temperature, you'd also need a text file containing the *a priori* and an estimate of its uncertainty. To loosen the retrieval constraint and achieve a closer fit to the data, try increasing the size of the *a priori* uncertainty to weight away from the prior in favour of the data (or vice versa). There are a wide range of potential gas and cloud parameterisations now available and described in the manual.

- Flag file *.fla: Yet more flags for the run, defining whether para-H_2 should be equilibrium or normal (INORMAL); whether to use Rayleigh scattering opacity (IRAY); whether to use modified partition functions for exoplanets (IPTF); whether to use Henyey-Greenstein or Mie theory directly for scattering (IMIE); and whether to add further continuum opacity from certain gases.

- Zenith angle file *.zen: A single flag defining where in the atmosphere the zenith angle is defined, either at the bottom of the deepest level, at 0 km, or at the very top.

- Abort file *.abo: A file that just has to be present in case of an early termination.

- Previous retrieval file `*.pre`: The raw fitted parameters and their covariance matrices copied over from a previous retrieval, if desired and if LIN is non-zero above. These are identical to the output `*.raw` file, see below.

4.1.3. Spectral Inputs

The third category of inputs help to define the sources of opacity for the spectral inversion - i.e., which files to use for the spectral lines, the aerosol properties, and the continuum absorptions. For a run relying on k-tables, the `*.kls` file contains the paths to the tables (one for each gas) that need to be read in. These k-tables will have been pre-computed using the spectral sampling, resolving power and lineshape relevant to the instrument you are trying to model. Sometimes an additional `*.fil` file is used to perform channel integrations - i.e., integrating the calculated spectrum over a filter function. However, this will often have been accounted for in advance by calculating k-tables with the relevant filter file in place.

As an alternative to using k-tables, a line-by-line run would need a `*.lbl` file specifying the wavelength range and resolution, a `*.sha` file defining the instrument lineshape to convolve the spectra with (e.g., Hamming, Hanning, square, triangle, gaussian); and a `*.key` file pointing towards the database containing all the line data. A `*.pra` file can also be used to instruct the code to use different lineshapes for different species (e.g., NH_3 might use an exotic lineshape compared to the standard Voigt lineshape).

In addition to the gases with line data and k-tables, the `*.cia` file points the code to the source of collision-induced opacity, commonly from H_2, He and CH_4 for the giant planets. This continuum source is added to the opacity of each layer in the atmosphere. If the calculation involves reflectivity or transit spectroscopy, then a `*.sol` file contains the name of the solar or stellar spectrum that is located in the `raddata` directory.

Finally, we must specify how the aerosols in the `aerosol.ref` file interact with the radiative transfer. A `*.xsc` file contains the aerosol cross-sections and single scattering albedoes (calculated by `Makephase`) as a function of wavelength or wavenumber. This is required for both aerosol absorption and multiple-scattering cases. If multiple scattering is required, then a Henyey-Greenstein phase function is required for each particle type as a function of wavelength, contained within `hgphase(1-n).dat` files.

4.1.4. Measurement Inputs

Last, but not least, we need to give NEMESIS the spectrum to be fitted. This will have come from pre-processing of the raw data into spectra averaged onto a latitude grid, a lat-lon grid, or maybe just a spectral irradiance representing the disc average. The file structure is contained within Section 3.3 of the manual, but primarily consists of a table of the measured spectral radiance and its uncertainty (i.e., the diagonal elements of the measurement covariance matrix) on a wavenumber/wavelength grid. This will contain NCONV wavelengths specified for a particular observing geometry (latitude, longitude, emission angle, solar angle and azimuthal angle), which are specified on the first four lines of the file. The instrument lineshape (FWHM) us

also defined at the start of the file, and is used to convolve the input opacity sources to the right spectral resolution. Multiple latitudes and emission angles can be included in the same file, and NEMESIS can fit over multiple geometries at the same time to generate the output files.

With these four categories of input files, you now have everything you need for a NEMESIS forward model (zero iterations in the `*.inp` file) of a full spectral retrieval.

4.2. Intermediate Files

NEMESIS is built around the forward model types contained in the `radtran` directory, and so with each successive iteration, it produces a set of intermediate files that are used by those `Radtrans` subroutines. For example, NEMESIS will create a `*.pat` file (the Path file specifying how the atmosphere is to be split into layers and which forward model to run), which creates a `*.drv` file (the driver file containing the atmospheric properties on every layer). Others you'll see appearing are the `*.sca` and `*.str` files passing scattering and retrieved parameters between subroutines. Finally, each reference file will be transcribed into a new version that has been slightly tweaked by the code - this can be as simple as a recalculation of the height-pressure-temperature grid via the hydrostatic equation. This produces the output `*.prf`, `aerosol.prf`, `parah2.prf` and `fcloud.prf` files. At the very last step of the iteration, those PRF files contain the final results of the NEMESIS inversion.

4.3. Output Files

At the very last stage of the process (either a forward model or an inversion), NEMESIS produces a `*.mre` file containing the final result. This file effectively contains two matrices. The first is a matrix showing the fit of the spectrum to the data - it contains columns with the wavenumber/wavelength, the measured spectrum and it's uncertainty, and the fitted spectrum, and can be used to plot the results of the spectral fit and to estimate the goodness-of-fit (χ^2). The second matrix contains the retrieved properties (i.e., whatever you told it to fit in the `*.apr` file). There are four columns - the *a priori* state vector and its uncertainty (i.e., the diagonal elements of the *a priori* covariance matrix, S_a), the retrieved state vector and its uncertainty (i.e., the diagonal elements of the final covariance matrix). The state vector might contain a complete vertical profile on NPRO levels, but it might also contain parameterised gas/cloud distributions, or as little as a single scaling factor for a particular gas.

The entries in this second matrix can be extracted from the `*.mre` file for plotting in a variety of ways, and is probably the first thing you should learn how to do. With multiple retrievals, you might be able to plot how the retrieved parameters vary with latitude, or map them spatially across the planet. You might experiment with the size of the *a priori* uncertainty and then plot a range of different retrieved profiles to see how this affects things. Or you might remove, modify and adapt gaseous abundances to see how this affects the calculated spectrum. There's a wide variety of experiments to try, but learning how the plot the results in the `.mre` file is key.

There are some additional outputs that are of use:

- Raw file `*.raw`: The `*.mre` file contained only the diagonal elements of the covariance matrix, but for some cases we have to investigate the correlations between retrieved parameters. For this reason the `*.raw` file contains all the covariance matrices needed, as well as the fitted state vectors. Sometimes these are used as a foundation for a second retrieval (i.e., when multiple steps are used in the retrieval process), in which case they are renamed as the `*.pre` file and included as an input file.

- Iteration file `*.itr`: For every iteration of NEMESIS, this text file contains the state vector, fitted spectrum and Jacobian matrix **K** so that we can analyse the convergence with the IDL code `plotiternewX.pro`.

- Covariance file `*.cov`: Very similar to the `*.raw` file, but contains the Jacobian, the gain matrix **G**, the averaging kernels **A** and error covariance matrices for display using the IDL code `imagecovariance.pro`.

- Jacobian file `kk.out`: An unformatted binary file containing solely the output Jacobian **K** matrix for plotting functional derivatives with `plotkkimageX.pro`. However, this can also be accessed (along with a lot of additional information) in the covariance file.

Finally, NEMESIS can be run in a number of different ways. If you just type `Nemesis` at the command line, it will prompt you for the run name and then spit out a whole host of diagnostics as the model is running. We often like to save that output as a 'process' file to look back at, which you can do by redirecting the output to a simple text file:

```
>< runname.nam > run.prc &
```

The contents of `run.prc` can then be searched to understand the steps involved in the model, and contains an estimate of the goodness-of-fit to the data as the model converges, which should be around $\chi^2/N \approx 1$. It also contains a crude test to balance the weighting of the retrieval between the *a priori* and the measurements, by reporting the ratio of $\mathbf{K_n S_x K_n^T}$ to the diagonal elements of $\mathbf{S_\epsilon}$. This ratio should be of order unity, otherwise you run the risk of weighting too far to the prior or to the data.

5. Summary

This document attempts to bring together the background information, code structure, and operational considerations for a new user of NEMESIS. It is not meant to replace the NEMESIS and Radtrans manuals, not your wider reading of textbooks on this subject. But it can provide a crude starting point for your introduction to the NEMESIS project.

The best way to learn how to use NEMESIS is by playing with the files. To that end, we intend to set up a directory of common examples to get you started, with some useful tutorial-style ideas and exercises. *This is a work in progress... October 2017.*

References

Andrews, D.G., 2000. An Introduction to Atmospheric Physics. CUP, Cambridge UK.

Barstow, J.K., Aigrain, S., Irwin, P.G.J., Hackler, T., Fletcher, L.N., Lee, J.M., Gibson, N.P., 2014. Clouds on the Hot Jupiter HD189733b: Constraints from the Reflection Spectrum. ApJ 786, 154. doi:10.1088/0004-637X/786/2/154, arXiv:1403.6664.

Borysow, A., 1991. Modelling of collision-induced infrared-absorption spectra of H_2-H_2 pairs in the fundamental band at temperatures from 20K to 300K. Icarus 92, 273–279.

Borysow, A., 1993. Erratum: "Modeling of collision-induced infrared absorption spectra of H_2-H_2 pairs in the fundamental band at temperatures from 20 to 300K" [Icarus, Vol. 92, No. 2, p. 273 - 279 (Aug 1991)]. Icarus 106, 614.

Borysow, A., Frommhold, L., 1986. Theoretical collision-induced rototranslational absorption spectra for the outer planets: H_2-CH_4 pairs. Astrophys. J. 304, 849–865.

Borysow, A., Frommhold, L., 1987. Collision-induced rototranslational absorption spectra of CH_4-CH_4 pairs at temperatures from 50 to 300K. Astrophys. J. 318, 940–943.

Borysow, J., Frommhold, L., Birnbaum, G., 1988. Collison-induced rototranslational absorption spectra of H_2-He pairs at temperatures from 40 to 3000 K. Astrophys. J. 326, 509–515. doi:10.1086/166112.

Conrath, B.J., Gierasch, P.J., Ustinov, E.A., 1998. Thermal Structure and Para Hydrogen Fraction on the Outer Planets from Voyager IRIS Measurements. Icarus 135, 501–517. doi:10.1006/icar.1998.6000.

de Kok, R., Irwin, P.G.J., Teanby, N.A., Lellouch, E., Bézard, B., Vinatier, S., Nixon, C.A., Fletcher, L.N., Howett, C., Calcutt, S.B., Bowles, N.E., Flasar, F.M., Taylor, F.W., 2007. Oxygen compounds in Titan's stratosphere as observed by Cassini CIRS. Icarus 186, 354–363. doi:10.1016/j.icarus.2006.09.016.

Fletcher, L.N., 2007. Saturn's Atmosphere: Structure and Composition from Cassini/CIRS. Ph.D. thesis. University of Oxford.

Fletcher, L.N., Baines, K.H., Momary, T.W., Showman, A.P., Irwin, P.G.J., Orton, G.S., Roos-Serote, M., Merlet, C., 2011a. Saturn's tropospheric composition and clouds from Cassini/VIMS 4.6-5.1 μm nightside spectroscopy. Icarus 214, 510–533. doi:10.1016/j.icarus.2011.06.006.

Fletcher, L.N., de Pater, I., Orton, G.S., Hammel, H.B., Sitko, M.L., Irwin, P.G.J., 2014. Neptune at summer solstice: Zonal mean temperatures from ground-based observations, 2003-2007. Icarus 231, 146–167. doi:10.1016/j.icarus.2013.11.035, arXiv:1311.7570.

Fletcher, L.N., Greathouse, T.K., Moses, J.I., Guerlet, S., West, R.A., 2015. Saturn's Seasonally Changing Atmosphere: Thermal Structure, Composition and Aerosols. ArXiv e-prints arXiv:1510.05690.

Fletcher, L.N., Hesman, B.E., Irwin, P.G.J., Baines, K.H., Momary, T.W., Sánchez-Lavega, A., Flasar, F.M., Read, P.L., Orton, G.S., Simon-Miller, A., Hueso, R., Bjoraker, G.L., Mamoutkine, A., del Rio-Gaztelurrutia, T., Gomez, J.M., Buratti, B., Clark, R.N., Nicholson, P.D., Sotin, C., 2011b. Thermal Structure and Dynamics of Saturn's Northern Springtime Disturbance. Science 332, 1413–1417. doi:10.1126/science.1204774.

Fletcher, L.N., Irwin, P.G.J., Orton, G.S., Teanby, N.A., Achterberg, R.K., Bjoraker, G.L., Read, P.L., Simon-Miller, A.A., Howett, C., de Kok, R., Bowles, N., Calcutt, S.B., Hesman, B., Flasar, F.M., 2008. Temperature and Composition of Saturn's Polar Hot Spots and Hexagon. Science 319, 79–82. doi:10.1126/science.1149514.

Fletcher, L.N., Irwin, P.G.J., Teanby, N.A., Orton, G.S., Parrish, P.D., Calcutt, S.B., de Kok, R., Howett, C., Bowles, N., Taylor, F.W., 2007. Characterising Saturn's Vertical Temperature Structure from Cassini/CIRS. Icarus 189, 457–478.

Fletcher, L.N., Orton, G.S., Teanby, N.A., Irwin, P.G.J., 2009a. Phosphine on Jupiter and Saturn from Cassini/CIRS. Icarus 202, 543–564. doi:10.1016/j.icarus.2009.03.023.

Fletcher, L.N., Orton, G.S., Teanby, N.A., Irwin, P.G.J., Bjoraker, G.L., 2009b. Methane and its isotopologues on Saturn from Cassini/CIRS observations. Icarus 199, 351–367. doi:10.1016/j.icarus.2008.09.019.

Giles, R.S., Fletcher, L.N., Irwin, P.G.J., 2015. Cloud structure and composition of Jupiter's troposphere from 5-μm Cassini VIMS spectroscopy. Icarus 257, 457–470. doi:10.1016/j.icarus.2015.05.030, arXiv:1506.01608.

Goody, R.M., Yung, Y.L., 1989. Atmospheric radiation : theoretical basis. Atmospheric radiation : theoretical basis, 2nd ed., by Richard M. Goody and Y.L. Yung. New York, NY: Oxford University Press, 1989.

Hanel, R.A., Conrath, B.J., Jennings, D.E., Samuelson, R.E., 2003. Exploration of the Solar System by Infrared Remote Sensing: Second Edition. Exploration of the Solar System by Infrared Remote Sensing, by R. A. Hanel and B. J. Conrath and D. E. Jennings and R. E. Samuelson, ISBN 0521818974. Cambridge, UK: Cambridge University Press, April 2003.

Houghton, J.T., Taylor, F.W., Rodgers, C.D., 1984. Remote sounding of atmospheres. CUP, Cambridge UK.

Irwin, P., Teanby, N., de Kok, R., Fletcher, L., Howett, C., Tsang, C., Wilson, C., Calcutt, S., Nixon, C., Parrish, P., 2008. The NEMESIS planetary atmosphere radiative transfer and retrieval tool. Journal of Quantitative Spectroscopy and Radiative Transfer 109, 1136–1150.

Irwin, P.G.J., 2003. Giant planets of our solar system: atmospheres, composition, and structure. Praxis, Chichester UK.

Irwin, P.G.J., Parrish, P., Fouchet, T., Calcutt, S.B., Taylor, F.W., Simon-Miller, A.A., Nixon, C.A., 2004. Retrievals of jovian tropospheric phosphine from Cassini/CIRS. Icarus 172, 37–49. doi:10.1016/j.icarus.2003.09.027.

Irwin, P.G.J., Wong, M.H., Simon, A.A., Orton, G.S., Toledo, D., 2017. HST/WFC3 observations of Uranus' 2014 storm clouds and comparison with VLT/SINFONI and IRTF/Spex observations. Icarus 288, 99–119. doi:10.1016/j.icarus.2017.01.031, arXiv:1611.03257.

Jacquinet-Husson, N., Scott, N.A., Chedin, A., Garceran, K., Armante, R., Chursin, A.A., Barbe, A., Birk, M., Brown, L.R., Camy-Peyret, C., Claveau, C., Clerbaux, C., Coheur, P.F., Dana, V., Daumont, L., Debacker-Barilly, M.R., Flaud, J.M., Goldman, A., Hamdouni, A., Hess, M., Jacquemart, D., Kopke, P., Mandin, J.Y., Massie, S., Mikhailenko, S., Nemtchinov, V., Nikitin, A., Newnham, D., Perrin, A., Perevalov, V.I., Regalia-Jarlot, L., Rublev, A., Schreier, F., Schult, I., Smith, K.M., Tashkun, S.A., Teffo, J.L., Toth, R.A., Tyuterev, V.G., Vander Auwera, J., Varanasi, P., Wagner, G., 2005. The 2003 edition of the GEISA/IASI spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 95, 429–467.

Lacis, A.A., Oinas, V., 1991. A description of the correlated-k distribution method for modelling nongray gaseous absorption, thermal emission, and multiple scattering in vertically inhomogeneous atmospheres. J. Geophys. Res. 96, 9027–9064.

Lee, J.M., Fletcher, L.N., Irwin, P.G.J., 2012. Optimal estimation retrievals of the atmospheric structure and composition of HD 189733b from secondary eclipse spectroscopy. Monthly Notices of the Royal Astronomical Society 420, 170–182. doi:10.1111/j.1365-2966.2011.20013.x, arXiv:1110.2934.

Nixon, C.A., 1998. Remote sounding of the atmosphere of Titan. Ph.D. thesis. University of Oxford.

Orton, G.S., Gustafsson, M., Burgdorf, M., Meadows, V., 2007. Revised Ab Initio Models for H_2-H_2 Collision Induced Absorption at Low Temperatures. Icarus 189, 544–549.

Rodgers, C.D., 2000. Inverse Methods for Atmospheric Remote Sounding: Theory and Practice. World Scientific.

Rothman, L.S., Gordon, I.E., Babikov, Y., Barbe, A., Chris Benner, D., Bernath, P.F., Birk, M., Bizzocchi, L., Boudon, V., Brown, L.R., Campargue, A., Chance, K., Cohen, E.A., Coudert, L.H., Devi, V.M., Drouin, B.J., Fayt, A., Flaud, J.M., Gamache, R.R., Harrison, J.J., Hartmann, J.M., Hill, C., Hodges, J.T., Jacquemart, D., Jolly, A., Lamouroux, J., Le Roy, R.J., Li, G., Long, D.A., Lyulin, O.M., Mackie, C.J., Massie, S.T., Mikhailenko, S., Müller, H.S.P., Naumenko, O.V., Nikitin, A.V., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E.R., Richard, C., Smith, M.A.H., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G.C., Tyuterev, V.G., Wagner, G., 2013. The HITRAN2012 molecular spectroscopic database. J. Quant. Spectro. Rad. Trans. 130, 4–50. doi:10.1016/j.jqsrt.2013.07.002.

Rothman, L.S., Jacquemart, D., Barbe, A., Benner, D.C., Birk, M., Brown, L.R., Carleer, M.R., Chackerian, C., Chance, K., Coudert, L.H., Dana, V., Devi, V.M., Flaud, J.M., Gamache, R.R., Goldman, A., Hartmann, J.M., Jucks, K.W., Maki, A.G., Mandin, J.Y., Massie, S.T., Orphal, J., Perrin, A., Rinsland, C.P., Smith, M.A.H., Tennyson, J., Tolchenov, R.N., Toth, R.A., Vander Auwera, J., Varanasi, P., Wagner, G., 2005. The HITRAN 2004 molecular spectroscopic database. Journal of Quantitative Spectroscopy and Radiative Transfer 96, 139–204.

Teanby, N.A., Irwin, P.G.J., de Kok, R., Nixon, C.A., Coustenis, A., Bézard, B., Calcutt, S.B., Bowles, N.E., Flasar, F.M., Fletcher, L.N., Howett, C., Taylor, F.W., 2006. Latitudinal variations of HCN, HC_3N, and C_2N_2 in Titan's stratosphere derived from Cassini CIRS data. Icarus 181, 243–255.

doi:10.1016/j.icarus.2005.11.008.

Tsang, C., 2007. The Composition and Circulation of the Atmosphere of Venus using Venus Express/VIRTIS Data. Ph.D. thesis. University of Oxford.