EUROPEAN SOUTHERN OBSERVATORY

Organisation Européenne pour des Recherches Astronomiques dans l'Hémisphère Austral

Europäische Organisation für astronomische Forschung in der südlichen Hemisphäre

VERY LARGE TELESCOPE

User Manual for `molecfit`

VLT–MAN–ESO–19550-5772

Issue 3.11

Date 2018–10–31

Prepared:	Innsbruck ESO In-Kind Team	2018-10-31	
	Name	Date	Signature
Approved:	***TBD***		
	Name	Date	Signature
Released:	***TBD***		
	Name	Date	Signature

This page was intentionally left blank

Change record

Issue/Rev.	Date	Section/Parag. affected	Reason/Initiation/Documents/Remarks
3.11	31/10/2018	Sect. 4.9	Known issues
3.10	13/04/2018	Sect. 4.	Environment variables
3.9	16/01/2018	Sect. 4.9	Known issues
3.8	21/12/2017	All	Major revision of the code
3.7	01/04/2014	Sect. 3, 7	Updated installation instructions, documented expert mode
3.6	31/10/2013	Sect. 4.4, 4.5, 5.1, 6.1, 6.2, and 7.1.4	Profiles, FITS keywords, and test updates
3.5	26/09/2013	All	RIXes for final review considered
3.4	25/07/2013	All	Changes related to code-testing workshop, added appendix
3.3	27/05/2013	Sect. 4.5	Changes in parameter description
3.2	30/04/2013	Sect. 5.5.3	Change in input parameter restrictions
3.1	28/03/2013	All	Changes related to midterm review and more
3.0	02/12/2012	All	Major revision of the code
2.5	17/05/2012	Sect. 4.4, 6.3, and 6.4	Discussing RFM problems and modified output files
2.4	03/05/2012	Sect. 1, 2.2, 4.4.1, 6.5, and 7	Discussing telluric absorption correction, Figure 1 modified
2.3	21/03/2012	Sect. 4 and 5.5.3	Voigt profile approximation added
2.2	16/03/2012	Sect. 5	Updating information concerning GDAS profiles Comment about width zero components (convolution)
2.1	23/11/2011	Sect. 4 and 5.5.3	Adding comments of Alain Smette
2.0	06/04/2011	All	Second version
1.0	28/02/2011	All	First version

The authors of the Austrian ESO In-Kind Team Innsbruck:
Stefan Noll
Wolfgang Kausch
Marco Barden
Cezary Szyszka
Amy M. Jones
Stefan Kimeswenger
Institute for Astro- and Particle Physics, University of Innsbruck
Technikerstr. 25/8, A-6020 Innsbruck / Austria

The authors of the European Southern Observatory:
Alain Smette
Andrea Mehner
Sabine Möhler
Gurvan Bazin
Jose A. Escartín
European Southern Observatory, ESO Headquarters
Karl-Schwarzschild-Str. 2, 85748 Garching bei München / Germany

This page was intentionally left blank

Contents

1 Introduction

This document is intended to be a user documentation/manual of the code `molecfit` developed in the framework of the Austrian ESO In-Kind project SM-03 as described in the corresponding Detailed Specification Document [SM-03-DS].

`molecfit` aims at fitting tropospheric/stratospheric telluric features and deriving correction functions for the removal of these features from spectra of astronomical targets. The development of the Common Pipeline Library (CPL) workflows is based on an Interactive Data Language (IDL) prototype code by A. Smette (ESO). A preceding version of the code was developed in the framework of the Austrian ESO In-Kind DR06 project (see [DR06-UM]). `molecfit` is also available via a Reflex workflow with a graphical user interface. A separate document describes this option ([MF-GUI]).

This document is organised as follows: Section 2 gives a brief overview of the project and the incorporated algorithms. Section 3 provides information on the installation procedures. Section 4 contains a description on how to run the code, the required input parameter file, and the output files. In Section 5 the atmospheric model and its adaptation to the observed spectrum are described in detail. Finally, the code performance is evaluated in Section 6.

2 Overview

2.1 The project

Ground-based astronomical observations suffer from emission and absorption processes in the atmosphere, which deteriorate the quality of the obtained data. At wavelengths longer than $2 - 2.5\,\mu$m, where thermal radiation from molecules in the lower atmosphere dominates, the amount of this disturbing background radiation can determine whether scientific observations are feasible at all. For this wavelength regime, it is crucial to be able to estimate the intensity of the atmospheric emission in advance. Such a prediction requires good knowledge of the column densities of atmospheric constituents that significantly contribute to the greenhouse effect. The most important molecules are water (H_2O), carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), and ozone (O_3). In particular, a good knowledge of the water abundance, *i.e.* air humidity, is essential, since H_2O is the main contributor to the IR atmospheric spectra. Moreover, it is much more variable than the other important species. Due to the variability of molecular abundances, the removal of telluric absorption features from astronomical spectra depends on observations of telluric standard stars with relatively smooth continua at similar time and airmass as the scientific target. Even if those observations are available, the telluric absorption correction becomes tricky in the near- and mid-IR, where wavelength regions with negligible atmospheric absorptions are hardly present. In this case, a reliable determination of the shape of the unabsorbed continuum is very difficult if an interpolation approach is used. Therefore, a realistic model of the atmospheric absorption for given observing conditions would make the telluric absorption correction more reliable and could reduce the number of required observations of telluric standard stars.

For a periodic monitoring of the abundances of crucial atmospheric constituents as well as a high-quality correction of telluric absorption features in astronomical spectra a fast, user-friendly, and reliable software tool is needed. This project aims at providing such a tool. Thanks to the advances made in modelling the Earth's atmosphere, it is possible to compute realistic atmospheric emission and absorption spectra. Such model spectra can be fitted to observed atmospheric spectra with relative deviations of a few percent only. The code `molecfit` represents such a fitting tool for a reliable derivation of atmospheric transmission functions and molecular abundances. The incorporated model spectra are generated by the radiative transfer code Line-by-line Radiative Transfer Model (LBLRTM) [2]. The basic input for the radiative transfer code is an atmospheric profile, which is created from a standard atmosphere (containing information on height, pressure, temperature, and chemical composition for general tropical environments up to 120 km), modelled meteorological Global Data Assimilation System (GDAS) data (containing pressure, temperature, and relative humidity for the region of the observing site for elevations up to ≈ 25 km), and on-site meteorological measurements by the ESO Meteo Monitor (EMM). Using iterative techniques, the input atmospheric profile is varied to obtain a model spectrum that well fits the scientific input spectrum (see Section 2.2). In this process, it is also necessary to optimise the scaling, wavelength grid, and resolution of the model. For computing radiance spectra reliably, grey body radiation from the telescope itself has to be taken into account as well.

The resulting software product might be made accessible to the scientific community. Once adopted by astronomers for atmospheric monitoring purposes, the efficiency of astronomical instruments working in the thermal IR such as CRIRES or VISIR could be significantly improved. Since strong molecular absorption affects all observable wavelength regimes except for the visual, many ESO instruments could benefit from an improved correction of telluric features in spectra of astronomical targets. In this document, we will demonstrate the performance of the code by using CRIRES and X-Shooter spectra of bright sources and VISIR sky radiance spectra (see Section 6). The test spectra fulfil the minimum calibration criterion, *i.e.*, they are wavelength-calibrated

one-dimensional (1D) spectra provided as ASCII tables, FITS tables, or FITS images. See section 4.8 for more information about the format of the Molecfit input files.

The C code molecfit relies on third-party code (see Section 3.1):

- C version of the least-squares fitting library mpfit by C. Markwardt [4] based on the FORTRAN fitting routine MINPACK-1 by Moré et al. [1980]

- The radiative transfer code LBLRTM v12.8. This publicly available software is developed within the Radiative Transfer Working Group of the Atmospheric and Environmental Research Inc. (AER); see also Clough et al. [2005], [5], and [2]) for more details. It can handle all molecules incorporated in a line parameter database, *e.g.* HIgh-resolution TRANsmission molecular absorption database (HITRAN) [6], and offers a wide range of possibilities to adjust input parameters. An additional part of this software is the Line File (LNFL), which provides the required line information on the basis of the line parameter database.

- A line database of molecular parameters: Currently, the line parameter list aer_v_3.6 delivered with the LBLRTM v12.8 package is included in the molecfit package. This line parameter list is build from HITRAN 2008 [6] and contains some updates (see [2] for more details).

- An atmospheric profile valid for the time of observation: The profile is created from standard atmospheric profiles, on-site measurements by the EMM [7], and GDAS data (a product of the National Centers for Environmental Prediction (NCEP) model, created by the Air Resources Laboratory (ARL) of the National Oceanic and Atmospheric Administration (NOAA), see [5], [8], [9]). Such atmospheric profiles contain pressure, temperature, and humidity for a number of layers of the atmosphere, at any point on Earth for all recent dates (going back to December 2004), on a 3h-grid basis.

- Perl scripts get_inv.pl and get_grib.pl from the NCEP NOMAD server [10] for fast download of GDAS data in the GRIB2 format.

2.2   Algorithm

Schematic graphs of the basic functionality of the software are shown in Figures 1 and 2, which provide a general overview of the workflow and the structure of the molecfit routine in more detail, respectively.

First, the code reads the science spectrum from an ASCII table, FITS table, or FITS image. Furthermore, it reads an ASCII driver file with the user input. If available, ESO keywords including EMM data are directly taken from a FITS header. A single atmospheric profile is compiled from data from three sources: a standard atmospheric profile for a given climate zone, an appropriate GDAS model profile for the time of the observation and the telescope site, and the corresponding ground-based EMM measurements (see Section 5.1 for more details). Input for the radiative transfer model is the resulting merged atmospheric profile (with a possible pre-selection of relevant molecules) and the target airmass at the time of observation (see Section 5.2). To match the observed spectrum, the code adapts the atmospheric spectrum (either transmission or radiation; see Section 5.3) further by flux scaling, wavelength grid correction, and convolution with a suitable instrumental profile (see Section 5.5). For a radiance spectrum, the contribution of thermal emission from the telescope is also taken into account (see Section 5.4).

The central component of the algorithm is the comparison/fitting of the calculated and the input science spectrum by means of `mpfit` [4]. The χ^2 minimisation procedure of this routine is based on the Levenberg-Marquardt technique (see Moré et al. [1980]), an iterative search algorithm characterised by gradient-controlled jumps in parameter space. Since this technique is prone to finding local minima, reasonable starting values and constraints for the fit parameters are required. `mpfit` checks whether the desired fit quality is reached. If this is not the case, it changes fit parameters in an appropriate way to search for a better χ^2. Each function call causes a new calculation of the sky model. For a change of the molecular abundances, the profiles of molecules can be scaled by simple factors, which represent a subset of the fit parameters provided to `mpfit`. The other parameters are coefficients of polynomials for continuum scaling, coefficients of Chebshev polynomials for the correction of the wavelength solution, the FWHM of boxcar, Gaussian, and Lorentzian kernels that are used to build a realistic instrumental profile, and the telescope emissivity if a radiance spectrum is computed. When a satisfactory fit is reached after several iterations of the χ^2 minimisation procedure, the code writes the best-fit spectrum, atmospheric profile, and fit parameters to output files. The best-fit molecular profiles are integrated to get molecular columns in ppmv. `molecfit` also computes the water vapour content of the atmosphere in mm. The atmospheric abundances are part of a special output summary file (see Section 4.9).

The fit parameters are defined in the input driver file (see Section 4.6). Dynamically setting fit flags to include and exclude parameters from the fitting procedure allows the code to find a reasonable solution in a relatively fast and robust way. In detail, `molecfit` follows a plan consisting of six steps by default:

- Step 1: scaling of the continuum

- Step 2: wavelength and resolution fit

- Step 3: rescaling of the continuum

- Step 4: fitting of the molecules

- Step 5: joint continuum, wavelength, and resolution fit

- Step 6: fit of all components (molecules, continuum, wavelength, and resolution)

As described in Section 4.2, it is also possible to focus on the last step only. Moreover, individual parameters can be fixed for the entire fitting procedure.

For the telluric absorption correction, the executable `calctrans` has to be used with the same parameter file as for `molecfit`. `calctrans` calculates the atmospheric transmission for the full wavelength range of the input spectrum and corrects this spectrum using this function. These calculations are separated from the fitting procedure, since the run of a radiative transfer code for a wide wavelength range is very time consuming. This is particularly critical if the fit is optimised by several code runs with different input parameters. For this reason, the fitting procedure should be performed for several well-defined narrow wavelength ranges that can be provided by a special ASCII or FITS file (see Section 4.6) if spectra like those from X-Shooter are processed. Too wide fit ranges are also not recommended due to the probable failure of the polynomial continuum fit. The narrow wavelength range of CRIRES allows one to fit the full spectrum at once. `calctrans` writes a FITS table with the model transmission function, the telluric absorption corrected spectrum, and a quality flag. Moreover, these results are provided in the same file format as the input frame, *i.e.* ASCII table, FITS table (with multiple extensions for multiple chips), or 1D FITS image (with possible extensions for error and quality). For

the latter case, two FITS images are written: one for the correction function and one for the corrected spectrum (see Section 4.9).

Alternatively, two seperate executables `calctrans_lblrtm` and `calctrans_convolution` can be used to calculate the atmospheric transmission and convolve the spectrum respectively.

If other spectra as the fitted one shall be corrected by the same correction function, an ASCII list of file names can be provided by the input parameter file (see Section 4.6). The executable `corrfilelist` then corrects the listed spectra for telluric absorption and saves these spectra in the same file format as the input files. In addition to the allowed file formats for `molecfit`, 2D FITS images can be corrected in this way.
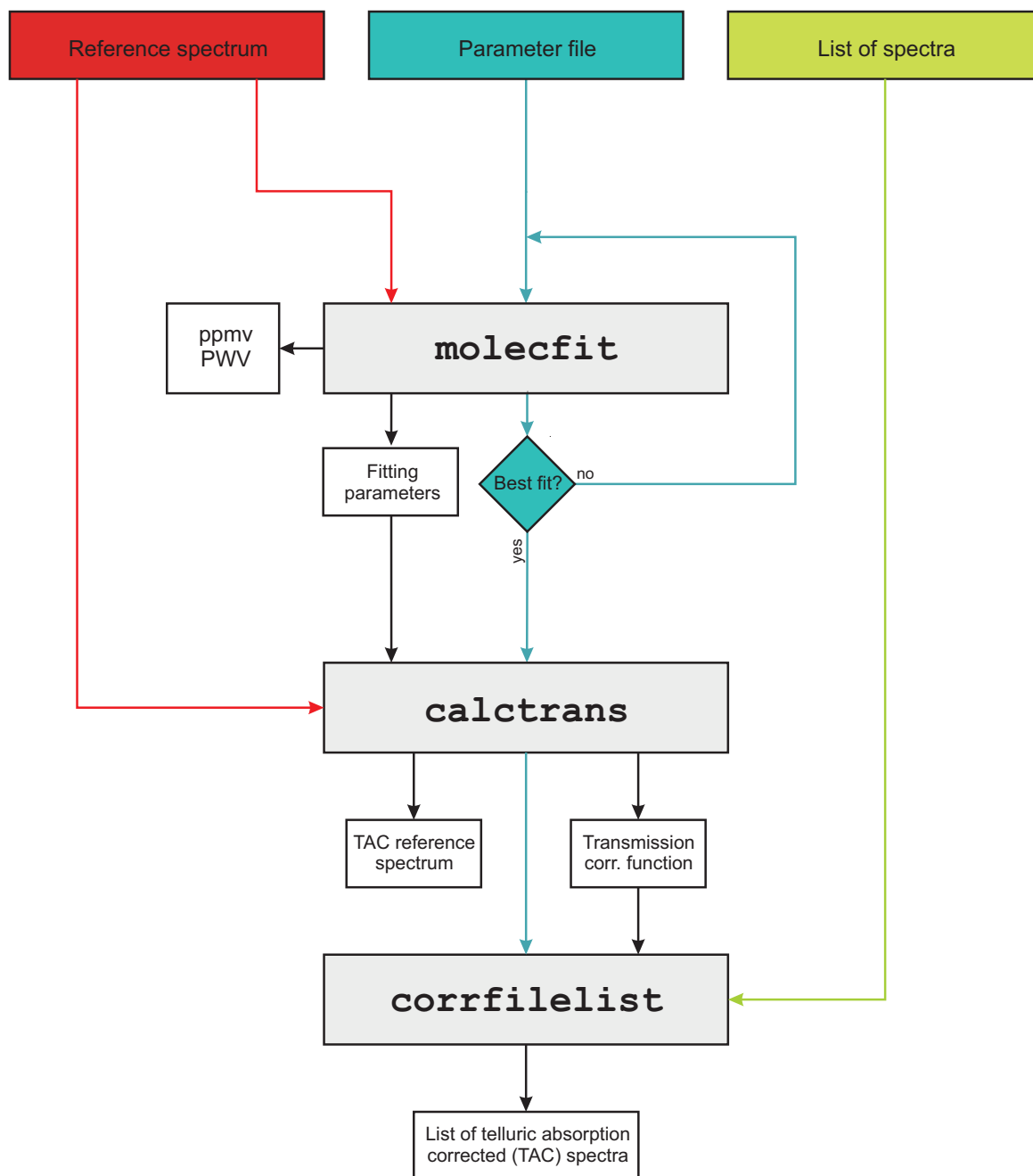
Figure 1: Overview of the software workflow. It shows the input and output for the three executables `molecfit`, `calctrans`, and `corrfilelist` and the connection between these routines.
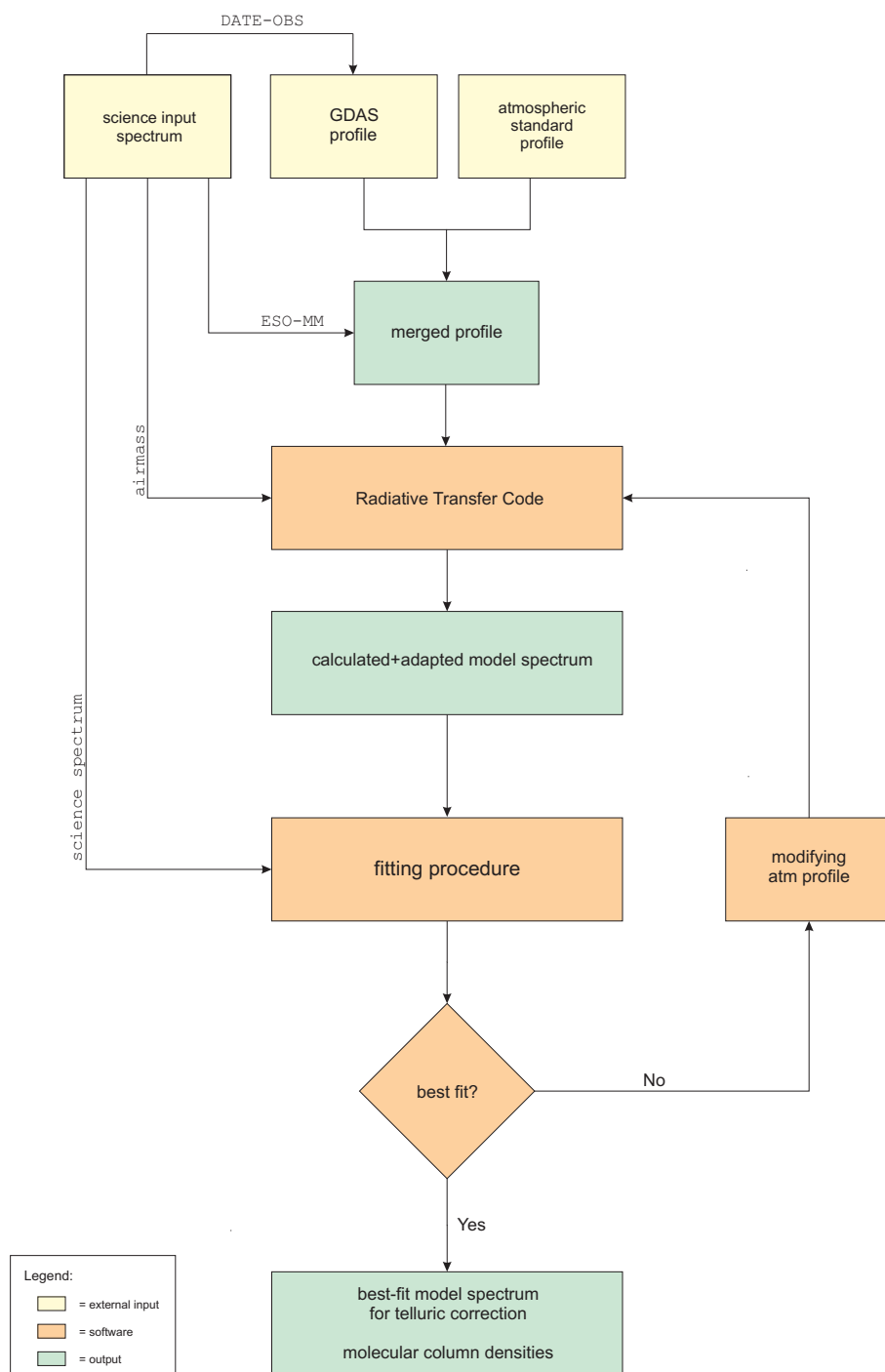
Figure 2: Workflow of the `molecfit` routine.

3 Installation procedure

3.1 Requirements

The installation of the basic `molecfit` binary package requires:

- C99 compatible compiler (e.g. gcc or clang)

- glibc 2.11 or newer on Linux or OS X 10.7 or newer

- common unix utilities (bash, tar, sed, grep, . . .)

The optional GUI to `molecfit` requires:

- Python v2.6 or v2.7 (but not Python v3.x)

- wxPython v2.8 or newer

- Python matplotlib v1.0 or newer

- PyFITS v2.4 or newer

The command line client also has optional display features which require:

- gnuplot v4.2 patchlevel 3 or newer

3.2 Binary installation

First the downloaded installer needs to be made executable. To do this change into the directory the installer was downloaded to and run following command (replacing molecfit_installer.run with the actual downloaded filename):

```
chmod u+x ./molecfit_installer.run
```

Now the installer can be executed from the same folder with:

```
./molecfit_installer.run
```

It will ask for an installation directory where it will extract its contents to. It is recommended to choose an empty directory to avoid overwriting existing files.

After the installer has successfully finished, the `molecfit` executables are installed into the `bin` subdirectory of the chosen installation folder. They can be executed by specifying the full or relative path. Also installed are a set of example parameter files for several instruments in the `examples/config` directory. To run a CRIRES example type:

```
<INST_DIR>/bin/molecfit <INST_DIR>/examples/config/molecfit_crires.par
```

For more details see Section 4.

For the GUI use the `molecfit_gui` executable instead of `molecfit`. For more details on the GUI see [MF-GUI].

3.3 GUI dependencies

The GUI requires some additional dependencies to be installed on the system. To check if the python installation is able to run the GUI, following commands can be run:

```
python -c 'import wx'

python -c 'import matplotlib; import matplotlib.backends.backend_wxagg'

python -c 'import pyfits'
```

If these commands fail please see following site for instructions on how to install these packages:

```
http://www.eso.org/pipelines/reflex_workflows/
```

3.4 Package contents

The installation package is a self extracting tarball containing the `molecfit` source code and pre-built versions of its third party dependencies:

- Common Pipeline Library v6.4.2 and its dependencies cfitsio v3.350, wcslib v4.16 and fftw3 v3.3.3 [16]

- Gridded Binary (GRIB) v1.9.9 [17]

- radiative transfer code LNFL v3.1 and LBLRTM v12.8 [2]

- AER line database v3.6 [5]

3.5 Source Installation

Advanced users may want to install everything from source, the basic instructions for this are outlined here. The installation from source additionally requires the gfortran compiler.

3.5.1 CPL compilation

The CPL sources can be obtained from [16]

For `molecfit` CPL only requires cfitsio. It can be installed as follows:

```
./configure --prefix=/install-location

make

make shared

make install
```

Then CPL can be install with:

```
./configure --prefix=/install-location --with-cfitsio=/install-location

make

make install
```

See the respective packages documentation for details on the installation procedure.

3.5.2 Non-ESO sources compilation

The rest of the required sources can be downloaded from the same place the `molecfit` binaries installers are available from. The third party source tarball contains a convenience Makefile to build and install all packages at once. It can be run with:

```
make -f BuildThirdParty.mk install prefix=/install-location
```

For installation on Mac OS please read the comments in this Makefile.

3.5.3 Molecfit compilation

After all dependencies have been installed `molecfit` can be compiled from source into the same location.

This is the only step required if one wants to update `molecfit` from source after previously installing the third party dependencies with the binary installer.

```
./configure --prefix=/install-location --with-cpl=/install-location
```

```
make

make install
```

In order to use `molecfit` from this location the environment variable `LD_LIBRARY_PATH` (or `DYLD_LIBRARY_PATH` on Mac OS) need to be set. With the bash shell this is done with following command:

```
export LD_LIBRARY_PATH=/install-location/lib
```

Now `molecfit` is ready to be used from `/install-location/bin`.

4 Running procedure

4.1 Environment variables `molecfit`

Before the execution you can to define some environment variables in order to change the behavior of molecfit.

- TMPDIR: Define the folder where create the temporary directory. If it is not define, molecfit use the default 'tmp' directory in the system

- MOLECFITDIR: Define the installation directory. If it is not define, molecfit use the path of the installation directory provide in compilation time.

- MOLECFITDIR_DATA: Define the data directory. If it is not define, molecfit use the path of the data directory create in the installation.

Example of use:

```
export TMPDIR=$PWD
export MOLECFITDIR=$PWD/dest
export MOLECFITDIR_DATA=$MOLECFITDIR/share/molecfit/data
```

4.2 Calling `molecfit`

After the compilation, the executable `molecfit` for the fitting of molecular transmission and emission is located in the `bin/` directory. It can be invoked via

```
cd <arbitrary_dir>
<fullpath>/bin/molecfit <parameter file> <optional mode>
```

where `<parameter file>` represents a user-defined parameter file (including paths). The `<optional mode>` parameter enables the user to run different modes of the fitting procedure:

- 'm' - multiple: Choosing this option, the fitting procedure follows all five steps described in Section 2.2.

- 's' - single run: Only step #5 is carried out (see Section 2.2).

Note that in both modes fitting steps that were designated to be skipped (by setting fit flag = 0 in the parameter file) are excluded (see Section 4.6).

The `<INST_DIR>/examples` directory contains several examples, including input spectra (in subfolder `input`) and configuration parameter files (in subfolder `config`) for the CRIRES, VISIR and X-Shooter instruments. The CRIRES spectrum represents the best data set for analysing molecular abundances and is optimally fitted using the parameter file listed in Section 4.6:

```
cd <arbitrary_dir>
<fullpath>/bin/molecfit examples/config/molecfit_crires.par
```

The other examples can be started synonymously.

The conversion of the input data file into a FITS table that is suitable for the fitting procedure can also be performed by means of the executable `preptable` plus `<parameter file>` as input parameter. The procedure is the same as for `molecfit`, but it stops before the fitting is started.

4.3 Calling `calctrans`

After the compilation, the executable `calctrans` for the derivation of the telluric absorption spectrum over the whole spectrum is also located in the `bin/` directory. It can be run by

```
cd <arbitrary_dir>
<fullpath>/bin/calctrans <parameter file>
```

where `<parameter file>` represents the same user-defined parameter file (including paths) as used for `molecfit`.

The telluric absorption correction can be calculated for a new spectrum with a different aimass (cf. Section 4.3) by changing the argument of FILENAME in the `<parameter file>` (see Section 4.7). If it is an ASCII file, the argument of TELALT must be also changed.

Note that the line kernel and the wavelength correction are not modified by this approach. This requires a direct fit with `molecfit`.

4.4 Calling `corrfilelist`

The `bin/` directory contains the executable `corrfilelist` for the telluric absorption correction for each of the spectra listed in an ASCII file given as argument to the LISTNAME keyword in the parameter file. Since the correction function is not recalculated, airmass differences are not considered (cf. Section 4.3). In the same way as `calctrans`, `corrfilelist` can be invoked by

```
cd <arbitrary_dir>
<fullpath>/bin/corrfilelist <parameter file>
```

where `<parameter file>` represents the same user-defined parameter file (including paths) as used for `molecfit`.

4.5 Calling `calctrans_lblrtm` and `calctrans_convolution`

Alternatively, correction of spectra affected by different (but known) kernels or different airmasses can be dealt with in the following way by calling two other executables:

```
cd <arbitrary_dir>
```

- `<fullpath>/bin/calctrans_lblrtm  <parameter file>`

  calculates the transmission spectrum from LBLRTM using the best fit parameter over the whole wavelength range, with a sampling 5 times better than the input spectrum;

- `<fullpath>/bin/calctrans_convolution  <parameter file>`

  performs the convolution of the spectrum produced by `calctrans_lblrtm` with the user-provided kernel.

A typical use is the following:

1. the user knows well the kernel for the spectra provided by each fiber of a multi-fiber spectrograph. `Molecfit` is called to fit the molecular content on an input spectrum (argument of FILENAME) corresponding to a given fiber; the parameter file includes the filename with the corresponding kernel (argument of KERNEL_FILE);

2. once a satisfactory fit is obtained, `calctrans_lblrtm` is run with the parameter file unchanged;

3. then `calctrans_convolution` is called with the same parameter file in which the name of the input file (argument of FILENAME) has been changed for the spectrum of another fiber, and where the kernel filename (argument of KERNEL_FILE) has been changed to provide the corresponding kernel. This step can be repeated for each spectra.

On the other hand, if the airmass for another fiber differs significantly from the first, `calctrans_lblrtm` should be called with the same parameter file where the name of the input file (argument of FILENAME) has been changed to the filename of the corresponding spectrum. The telescope altitude (argument of TELALT) must be changed also if the input file is an ASCII file.

4.6   The parameter file

All parameters needed for the fit are read from an ASCII parameter file, which contains parameter names, descriptions, and initial values. Example files can be found in the folder `examples/config/`.

The file is divided into sections, where parameters belonging together are grouped. These sections incorporate the directory structure (labelled DIRECTORY STRUCTURE), the required input (INPUT DATA), output options (RESULTS), the precision of the fit (FIT PRECISION), molecular information (MOLECULAR COLUMNS), the fit of the background and the continuum (BACKGROUND AND CONTINUUM), the wavelength fit (WAVELENGTH SOLUTION), the resolution fit (RESOLUTION), environmental parameters (AMBIENT PARAMETERS), and the atmospheric profiles (ATMOSPHERIC PROFILES). The individual parameters are given one per line. The parameter name plus a trailing colon and space is followed by one or more parameter values, which also have to be separated by spaces. Lines beginning with a hash key are assumed to be comments and are skipped. Thus, next to removing parameter lines, there are several options to "deactivate" a parameter in the parameter file:

Command "active":

`ftol: 1e-2`

Command "deactivated" (commented out):

```
#ftol: 1e-2
```

Command "deactivated" (no trailing colon):

```
ftol 1e-2
```

Command "deactivated" (no parameter value)

```
ftol:
```

or

```
ftol
```

In the following, `examples/config/molecfit_crires.par` is shown as an example of a parameter file. Individual parameters are explained in Section 4.7:

```
### Driver for MOLECFIT

## INPUT DATA

# Data file name (path relative to the current directory or absolute path)
filename: examples/input/crires_spec_jitter_extracted_0000.fits

# ASCII list of files to be corrected for telluric absorption using the
# transmission curve derived from the input reference file (path of list and
# listed files relative to the current directory or absolute path; default: "none")
listname: none

# Type of input spectrum -- 1 = transmission (default); 0 = emission
trans: 1

# Names of the file columns (table) or extensions (image) containing:
# Wavelength  Flux  Flux_Err  Mask
# - Flux_Err and/or Mask can be avoided by writing 'NULL'
# - 'NULL' is required for Wavelength if it is given by header keywords
# - parameter list: col_lam, col_flux, col_dflux, and col_mask
columns: Wavelength Extracted_OPT Error_OPT NULL

# Default error relative to mean for the case that the error column is missing
default_error: 0.01

# Multiplicative factor to convert wavelength to micron
# (e.g. nm -> wlgtomicron = 1e-3)
wlgtomicron: 1e-3
```

```
# Wavelengths in vacuum (= vac) or air (= air)
vac_air: vac

# ASCII or FITS table for wavelength ranges in micron to be fitted
# (path relative to the current directory or absolute path; default: "none")
wrange_include: none

# ASCII or FITS table for wavelength ranges in micron to be excluded from the
# fit (path relative to the current directory or absolute path; default: "none")
wrange_exclude: none

# ASCII or FITS table for pixel ranges to be excluded from the fit
# (path relative to the current directory or absolute path; default: "none")
prange_exclude: examples/config/exclude_crires.dat

## RESULTS

# Directory for output files (path relative to the current directory or absolute path)
output_dir: output

# Name for output files
# (supplemented by "_fit" or "_tac" as well as ".asc", ".atm", ".fits",
# ".par, ".ps", and ".res")
output_name: molecfit_crires

# Plot creation: gnuplot is used to create control plots
# W - screen output only (incorporating wxt terminal in gnuplot)
# X - screen output only (incorporating x11 terminal in gnuplot)
# P - postscript file labelled '<output_name>.ps', stored in <output_dir>
# combinations possible, i.e. WP, WX, XP, WXP (however, keep the order!)
# all other input: no plot creation is performed
plot_creation: XP

# Create plots for individual fit ranges? -- 1 = yes; 0 = no
plot_range: 0

## FIT PRECISION

# Relative chi2 convergence criterion
ftol: 1e-2

# Relative parameter convergence criterion
xtol: 1e-2

## MOLECULAR COLUMNS

# List of molecules to be included in the model
# (default: 'H2O', N_val: nmolec)
list_molec: H2O CH4 O3
```

```
# Fit flags for molecules -- 1 = yes; 0 = no (N_val: nmolec)
fit_molec: 1 1 1

# Values of molecular columns, expressed relatively to the input ATM profile
# columns (N_val: nmolec)
relcol: 1. 1. 1.

## BACKGROUND AND CONTINUUM

# Conversion of fluxes from phot/(s*m2*mum*as2) (emission spectrum only) to
# flux unit of observed spectrum:
# 0: phot/(s*m^2*mum*as^2) [no conversion]
# 1: W/(m^2*mum*as^2)
# 2: erg/(s*cm^2*A*as^2)
# 3: mJy/as^2
# For other units, the conversion factor has to be considered as constant term
# of the continuum fit.
flux_unit: 0

# Fit of telescope background -- 1 = yes; 0 = no (emission spectrum only)
fit_back: 0

# Initial value for telescope background fit (range: [0,1])
telback: 0.1

# Polynomial fit of continuum --> degree: cont_n
fit_cont: 1

# Degree of coefficients for continuum fit
cont_n: 3

# Initial constant term for continuum fit (valid for all fit ranges)
# (emission spectrum: about 1 for correct flux_unit)
cont_const: 1.

## WAVELENGTH SOLUTION

# Refinement of wavelength solution using a polynomial of degree wlc_n
fit_wlc: 1

# Polynomial degree of the refined wavelength solution
wlc_n: 3

# Initial constant term for wavelength correction (shift relative to half
# wavelength range)
wlc_const: 0.

## RESOLUTION

# Fit resolution by boxcar -- 1 = yes; 0 = no
fit_res_box: 0
```

```
# Initial value for FWHM of boxcar relative to slit width (>= 0. and <= 2.)
relres_box: 0.

# Voigt profile approximation instead of independent Gaussian and Lorentzian
# kernels? -- 1 = yes; 0 = no
kernmode: 0

# Fit resolution by Gaussian -- 1 = yes; 0 = no
fit_res_gauss: 1

# Initial value for FWHM of Gaussian in pixels
res_gauss: 1.

# Fit resolution by Lorentzian -- 1 = yes; 0 = no
fit_res_lorentz: 0

# Initial value for FWHM of Lorentzian in pixels
res_lorentz: 0.5

# Size of Gaussian/Lorentzian/Voigtian kernel in FWHM
kernfac: 300.

# Variable kernel (linear increase with wavelength)? -- 1 = yes; 0 = no
varkern: 0

# ASCII file for kernel elements (one per line; normalisation not required)
# instead of synthetic kernel consisting of boxcar, Gaussian, and Lorentzian
# components (path relative to the current directory or absolute path; default: "none")
kernel_file: none

## AMBIENT PARAMETERS

# If the input data file contains a suitable FITS header, the keyword names of
# the following parameters will be read, but the corresponding values will not
# be used. The reading of parameter values from this file can be forced by
# setting keywords to NONE.

# Observing date in years or MJD in days
obsdate
obsdate_key: MJD-OBS

# UTC in s
utc
utc_key: UTC

# Telescope altitude angle in deg
telalt
telalt_key: ESO TEL ALT

# Humidity in %
```

```
rhum
rhum_key: ESO TEL AMBI RHUM

# Pressure in hPa
pres
pres_key: ESO TEL AMBI PRES START

# Ambient temperature in deg C
temp
temp_key: ESO TEL AMBI TEMP

# Mirror temperature in deg C
m1temp
m1temp_key: ESO TEL TH M1 TEMP

# Elevation above sea level in m (default is Paranal: 2635m)
geoelev
geoelev_key: ESO TEL GEOELEV

# Longitude (default is Paranal: -70.4051)
longitude
longitude_key: ESO TEL GEOLON

# Latitude (default is Paranal: -24.6276)
latitude
latitude_key: ESO TEL GEOLAT

## INSTRUMENTAL PARAMETERS

# Slit width in arcsec (taken from FITS header if present)
slitw: 0.4
slitw_key: ESO INS SLIT1 WID

# Pixel scale in arcsec (taken from this file only)
pixsc: 0.086
pixsc_key: NONE

## ATMOSPHERIC PROFILES

# Reference atmospheric profile
ref_atm: equ.atm

# Specific GDAS-like input profile (P[hPa] HGT[m] T[K] RELHUM[%]) (path
# relative to the installation directory or absolute path). In the case of "none", no GDA
# profiles will be considered. The default "auto" performs an automatic
# retrieval.
gdas_prof: auto

# Grid of layer heights for merging ref_atm and GDAS profile. Fixed grid = 1
# (default) and natural grid = 0.
layers: 1
```

```
# Upper mixing height in km (default: 5) for considering data of a local meteo
# station. If emix is below geoelev, rhum, pres, and temp are not used for
# modifying the corresponding profiles.
emix: 5


# PWV value in mm for the input water vapour profile. The merged profile
# composed of ref_atm, GDAS, and local meteo data will be scaled to this value
# if pwv > 0 (default: -1 -> no scaling).
pwv: -1


end
```

4.7 The parameters

In the following, the individual parameters are explained in more detail by following the order as they appear in the parameter file.

```
## INPUT DATA:
```

- FILENAME: Path and name of the input spectrum. The path has to be relative to the current directory or absolute. The accepted file formats for the input spectrum are ASCII and FITS. The spectral data of the latter has to be provided either in table form (multiple FITS extensions for different chips) or as 1D image (with optional error and quality data in additional FITS extensions). See section 4.8 for more information about the input file format.

- LISTNAME: Input ASCII list of files for `corrfilelist`, which corrects spectra for telluric absorption using the transmission curve derived by `calctrans`. The path to the list and the listed files has to be relative to the current directory or absolute. In contrast to FILENAME, the listed files can also be 2D FITS images. By default, no file list is expected ("none").

- TRANS: type of input spectrum (1 = transmission, 0 = emission; default = 1).

- COLUMNS: Column names of the input file containing information on wavelength, flux, flux_err, and mask. The latter two are optional and can be disabled by setting them to "NULL". For ASCII files, the column names are irrelevant, with the exception of "NULL" input. In the case of FITS images, the given labels are compared to the FITS extension names (keyword "EXTNAME"). Since it is expected that the flux data is in the first layer of the FITS file (0th extension), the existence of a corresponding FITS keyword is not required for this column. Internally, the code denotes the columns as COL_LAM, COL_FLUX, COL_DFLUX, and COL_MASK.

- DEFAULT_ERROR: Default error relative to the mean in case the error column is missing (column name = "NULL", see previous record).

- WLGTOMICRON: Multiplicative factor to convert input wavelength unit into μm. For example, for nm this parameter has to set to 10^{-3}.

- VAC_AIR: Wavelengths in vacuum (= "vac") or air (= "air")? This parameter depends on the instrument and the wavelength calibration approach.

- WRANGE_INCLUDE: ASCII or FITS table for wavelength ranges to be fitted. The path to the file has to be relative to the current directory or absolute. Except for empty lines and comment lines starting with #, each line of an ASCII file has to provide a lower and an upper wavelength limit in μm. FITS tables must contain exactly two columns. The column labels are not fixed. If a range table is not desired and the full spectrum shall be fitted, "none" can be given. This is also the default value.

- WRANGE_EXCLUDE: ASCII or FITS table for wavelength ranges to be excluded from the fit. The path to the file has to be relative to the current directory or absolute. Except for empty lines and comment lines starting with #, each line of an ASCII file has to provide a lower and an upper wavelength limit in μm. FITS tables must contain exactly two columns. The column labels are not fixed. If the exclusion of wavelength ranges is not desired, "none" can be given. This is also the default value.

- PRANGE_EXCLUDE: ASCII or FITS table for pixel ranges to be excluded from the fit. The path to the file has to be relative to the current directory or absolute. Except for empty lines and comment lines starting with #, each line of an ASCII file has to provide a lower and an upper pixel. FITS tables must contain exactly two columns. The column labels are not fixed. The pixel counting starts with 1. If there are several chips, the pixel counting continues at chip limits. If a range table is not desired and all pixels shall be fitted, "none" can be given. This is also the default value.

RESULTS:

- OUTPUT_DIR: Output directory as path relative to the current directory or absolute path (see also Section `DIRECTORY STRUCTURE`). The folder is created if it does not exist.

- OUTPUT_NAME: Unique name space for output files. The extensions are supplemented by "_fit" or "_tac" as well as ".asc", ".atm", ".fits", ".par", ".ps", and ".res". See Section 4.9 for more details.

- PLOT_CREATION: `molecfit` invokes `gnuplot` to create control plots for verifying the fit. For each fit range, an individual plot is created, which contains a comparison of the observed and the modelled spectrum and their residual. An overview plot for the entire wavelength range is also produced. Finally, `calctrans` creates a plot that shows the initial and the telluric absorption corrected spectrum. This parameter allows specifying the output terminal of gnuplot: W - screen output incorporating wxt terminal, X - screen output incorporating x11 terminal, P - postscript files labelled '<output_name>_fit.ps', '<output_name>_fit_<fitrange>.ps', and '<output_name>_tac.ps' (stored in <output_dir>). Also, arbitrary combinations are possible, *e.g.* WP for wxt terminal and postscript, or WX for wxt and x11 terminal output.

- PLOT_RANGE: Flag for creation of plots of individual fit ranges (1 = yes, 0 = no). If the flag is set to 0, the files labelled '<output_name>_fit_<fitrange>.ps' are not produced.

FIT PRECISION:

- FTOL: relative χ^2 convergence criterion.

- XTOL: relative parameter convergence criterion.

MOLECULAR COLUMNS:

- LIST_MOLEC: List of molecules separated by blanks, which are included in the model. Note that only those molecules are valid, which are also present in the standard atmospheric profile REF_ATM (see also Section 5.1.1).

- FIT_MOLEC: Fit flags (1 = yes, 0 = no) separated by blanks for list of molecules. # has to match number of molecules in LIST_MOLEC. Example: 1 0 1 implies that only the first and third molecule are fitted. The middle molecule is included statically in the fit, *i.e.* its abundance is not varied in the fit.

- RELCOL: Relative molecular column densities separated by blanks, normalised to the values in the input atmospheric profile (1 = 100%). # has to match number of molecules in LIST_MOLEC.

BACKGROUND AND CONTINUUM:

- FLUX_UNIT: Conversion of fluxes from phot/(s·m^2· μm·arcsec^2) (emission spectrum only) to flux unit of observed spectrum: 0 = phot/(s·m^2·μm·arcsec^2) (–> no conversion), 1 = W/(m^2·μm·arcsec^2), 2 = erg/(s·cm^2·Å·arcsec^2), 3 = mJy/arcsec^2. For other units differing from the offered units by a constant factor, the constant term of the continuum fit parameters CONT_CONST has to be multiplied by the conversion factor. For non-flux-calibrated spectra provided in analogue-to-digital units (ADU), the default 0 is the best choice.

- FIT_BACK: fit of telescope background (1 = yes, 0 = no); used for emission spectrum fit only.

- TELBACK: initial emissivity value for telescope background grey body fit; used for emission spectrum fit only.

- FIT_CONT: Flag for polynomial fit of the continuum (1 = yes, 0 = no). For each fit range/chip, the model spectrum is multiplied by the resulting polynomial.

- CONT_N: degree of polynomial for the continuum fit.

- CONT_CONST: Initial constant term of the polynomial for the continuum fit. The same value is used for all fit ranges/chips. By default, 1 is assumed. Since all higher terms are set to 0 at the beginning, the fitting procedure starts without a continuum correction if the default setting is used. For sky emission spectra with correct FLUX_UNIT, the true continuum correction factor should be very close to the default value 1.

WAVELENGTH SOLUTION:

- FIT_WLC: flag for refinement of wavelength solution (1 = yes, 0 = no).

- WLC_N: degree of Chebyshev polynomial for refined wavelength solution.

- WLC_CONST: Constant term of the Chebyshev polynomial for the wavelength solution, which is derived for each chip independently. The provided constant term is used for all chips. The given value represents a shift relative to half the wavelength range of the input spectrum. By default, 0 is assumed. Since the linear term and the higher terms are set to 1 and 0, respectively, at the beginning, the fitting procedure starts without a wavelength correction if the default setting is used.

RESOLUTION:

The resolution fit incorporates a combined, boxcar, Gaussian, and Lorentzian convolution. The latter two convolutions result in a Voigt profile.

- FIT_RES_BOX: flag for resolution fit using a boxcar filter (1 = yes, 0 = no).

- RELRES_BOX: Initial value for FWHM of boxcar relative to slit width (≥ 0 and ≤ 2). A value of 0 combined with FIT_RES_BOX = 0 switches off the convolution of a boxcar.

- KERNMODE: By default (= 0), spectra are convolved by independent Gaussian and Lorentzian kernels. It is also possible to perform only one convolution with a kernel derived from a Voigt profile approximation (= 1), which also uses the FWHM of Gaussian and Lorentzian as input.

- FIT_RES_GAUSS: flag for resolution fit using a Gaussian filter (1 = yes, 0 = no).

- RES_GAUSS: Initial value for FWHM of Gaussian (in pixels). A value of 0 combined with FIT_RES_GAUSS = 0 switches off the convolution of a Gaussian.

- FIT_RES_LORENTZ: flag for resolution fit using a Lorentzian filter (1 = yes, 0 = no).

- RES_LORENTZ: Initial value for FWHM of Lorentzian (in pixels). A value of 0 combined with FIT_RES_LORENTZ = 0 switches off the convolution of a Lorentzian.

- KERNFAC: Kernel size in units of FWHM. Depending on KERNMODE either the Gaussian and Lorentzian kernels or the combined Voigt profile kernel are relevant. By default, a value of 3 is assumed.

- VARKERN: Flag for selecting a constant (= 0) or a variable kernel (= 1). In the latter case, all FWHM values are related to the central wavelength of the full wavelength range. The variable kernel increases linearly with wavelength, *i.e.* the resolution is constant. X-Shooter Echelle spectra show this behaviour. For spectra where the object profile in the slit mainly determines the kernel, the default constant kernel option is recommended.

- KERNEL_FILE ASCII file for fixed kernel elements (pixels) instead of a synthetic kernel consisting of boxcar, Gaussian, and Lorentzian components. The latter is the default case and requires "none". The path to the file has to be relative to the current directory or absolute. In the file, each kernel element has to be given on a separate line. It is not required that the kernel has been normalised to 1. Note that a given KERNEL_FILE overrules all other parameters related to the line profile. The kernel will not be fitted and will not depend on the wavelength (irrespective of VARKERN).

AMBIENT PARAMETERS:

The parameters of this section are only required if the input spectrum is provided as ASCII file or if the ESO FITS header keywords are missing or differ from the default names. The use of a specific value from the parameter file can be forced by setting the keyword name to "NONE". If no keyword is provided by the input file (*e.g.* ASCII file), it is not required to set the keyword names to "NONE".

- OBSDATE: Observing date in years, *e.g.* 2008.566, or MJD in days. This parameter is required for the retrieval of GDAS data.

- OBSDATE_KEY: FITS keyword name for OBSDATE (default: "MJD-OBS").

- UTC: UTC in s, starting at 00:00. This parameter is required for the retrieval of GDAS data.

- UTC_KEY: FITS keyword name for UTC (default: "UTC").

- TELALT: altitude angle of telescope in deg.

- TELALT_KEY: FITS keyword name for TELALT (default: "ESO TEL ALT").

- RHUM: Relative humidity in % for GEOELEV. This parameter is only relevant if EMIX is larger than GEOELEV.

- RHUM_KEY: FITS keyword name for RHUM (default: "ESO TEL AMBI RHUM").

- PRES: Pressure in hPa for GEOELEV. This parameter is only relevant if EMIX is larger than GEOELEV.

- PRES_KEY: FITS keyword name for PRES (default: "ESO TEL AMBI PRES START").

- TEMP: Ambient temperature in °C for GEOELEV. This parameter is only relevant if EMIX is larger than GEOELEV.

- TEMP_KEY: FITS keyword name for TEMP (default: "ESO TEL AMBI TEMP").

- M1TEMP: Temperature of primary mirror M1 in °C. This parameter is only relevant for emission spectra (*i.e.* TRANS = 0), where the thermal emission of the telescope has to be considered.

- M1TEMP_KEY: FITS keyword name for M1TEMP (default: "ESO TEL TH M1 TEMP").

- GEOELEV: elevation above sea level in m (default is Paranal: 2635 m).

- GEOELEV_KEY: FITS keyword name for GEOELEV (default: "ESO TEL GEOELEV").

- LONGITUDE: Longitude in deg (default is Paranal: -70.4051). This parameter is required for the retrieval of GDAS data.

- LONGITUDE_KEY: FITS keyword name for LONGITUDE (default: "ESO TEL GEOLON").

- LATITUDE: Latitude in deg (default is Paranal: -24.6276). This parameter is required for the retrieval of GDAS data.

- LATITUDE_KEY: FITS keyword name for LATITUDE (default: "ESO TEL GEOLAT").

INSTRUMENTAL PARAMETERS:

- SLITW: Slit width in arcsec. The provided value is only taken into account if the ESO keyword "ESO INS SLIT1 WID" does not exist and the instrument is not X-Shooter (special keyword finding routine). For CRIRES and VISIR, the stated keyword is usually present.

- SLITW_KEY: FITS keyword name for SLITW (default: "ESO INS SLIT1 WID").

- PIXSC: Pixel scale in arcsec. This parameter has to be provided manually, since this information could not be found in the ESO file headers of the investigated instruments.

- PIXSC_KEY: FITS keyword name for PIXSC. The default is "NONE", *i.e.* the parameter is read from the parameter file.

ATMOSPHERIC PROFILES:

- REF_ATM: Reference atmospheric profile (standard profile). By default, it is set to "equ.atm", which is located in the dedicated folder "data/profiles/mipas/". See Section 5.1.1 for more information.

- GDAS_PROF: Specific GDAS-like input profile (format: P[hPa] HGT[m] T[K] RELHUM[%], see Section 5.1.2). The path to the file has to be relative to the installation directory or absolute. In the case of "none", no GDAS profiles will be considered. The default option "auto" causes an automatic retrieval either from a local library or a web-server (see Section 5.1.4).

- LAYERS: Flag for grid of layer heights required for merging REF_ATM and GDAS profiles. The default option is 1, which selects a fixed grid with 50 layers for Cerro Paranal (see Section 5.1.4). If the parameter is set to 0, a natural grid is used, *i.e.* all layer heights of REF_ATM and the GDAS profile are combined, which tends to significantly increase the number of layers compared to the default option. If local meteo data are considered (see EMIX), GEOELEV is also added to the grid.

- EMIX: Upper mixing height in km for considering data of a local meteorological station (see Section 5.1.3). Above this level, the influence of the meteo data is expected to be zero (see Section 5.1.4). For Cerro Paranal, the default value of 5 km is assumed. If EMIX is below GEOELEV, no local meteo data are considered, *i.e.* the content of RHUM, PRES, and TEMP is not used for modifying the merged profiles.

- PWV: Precipitable water vapour (PWV) value in mm for the input water vapour profile. The merged profile composed of REF_ATM, GDAS, and local meteo data (see Section 5.1) will be scaled to this value if PWV is positive. By default (PWV = -1), the manipulation of the input water vapour profile is switched off. The RELCOL value for water vapour refers to the given PWV.

The keyword END is optional. It marks the last line that is considered by the file reading routine. Any text beyond END is ignored.

4.8   Format of the input spectrum

4.8.1   Accepted file formats

Molecfit currently accepts the following formats:

1. ASCII.

2. FITS binary table,

3. 1D FITS image,

4. Multi-dimentional FITS image,

and recognizes it in the following way:

i If the file is not a FITS file, Molecfit assumes an ASCII format.

ii If the file is a FITS file, Molecfit then looks at the presence of extensions. Molecfit assumes a 1D image format. (Note that the FITS keyword EXTEND only indicates that an extension may be present, not that it is present!)

iii If there is at least one extension, Molecfit uses the keyword XTENSION in the extension(s) to identify if the file is a—possibly multi-extension—FITS binary table (XTENSION=BINTABLE) or an image (XTENSION=IMAGE). If the file is an image, Molecfit then uses the NAXIS keyword to distinguish between 1D, 2D and 3D images.

4.8.2 A note regarding the mask values

The default behavior of Molecfit is that all values are either 0 or 1. A value of 0 indicates that the value must not be used by the fit and a value of 1 indicates that the value should be used by the fit. However, Molecfit also checks that values other than 0 and 1 are used. If this is the case, 0 is assumed to be ok and all other values cause pixel rejection.

Possible NAN in the flux and error columns are substituted by zero flux and the corresponding pixel is rejected for the fit. The same is performed for negative errors.

4.8.3 Format of the ASCII files

Two columns are mandatory: the first column must provide the wavelength and the second column, the flux. Two additional columns are optional: the 1-σ error on the flux and the mask.

The 'columns' entry in the parameter file should be a list of 4 names, which are unimportant, except for NULL. However, it is a good practice—if only for readability—to use

```
columns: Wavelength Flux Flux_Err Mask
```

If the error on the flux and/or the mask are not provided, FLUX_ERR and/or MASK should be replaced by NULL.

4.8.4 Format of the FITS binary tables

Two columns are mandatory: the first column must provide the wavelength and the second column the flux. Two additional columns are optional: the 1-σ error on the flux and the mask.

The 'columns' entry in the parameter file should be:

```
columns: wavelength_label flux_label flux_err_label mask_label
```

where the field is made of the label (title) of the column in the FITS binary table. Note that the labels are case-sensitive! If the error on the flux and/or the mask are not provided, FLUX_ERR_LABEL and/or MASK_LABEL should be replaced by NULL.

Example: reduced CRIRES (pre-upgrade) file using the optimally extracted fluxes.

```
columns: Wavelength Extracted_OPT Error_OPT NULL
```

4.8.5 Format of the FITS 1D image data

In the case of FITS image data, the wavelength information is provided through the FITS keywords CRPIX1, CRVAL1 and CRPIX1 if the wavelength vector is not provided.

The 'columns' entry in the parameter file should be:

```
columns: NULL Flux NULL NULL
```

4.8.6 Format of the FITS multi-extension image data

In the case of FITS image data, the wavelength information is provided through the FITS keywords CRPIX1, CRVAL1 and CRPIX1 if the wavelength vector is not provided.

The name of the extension is retrieved from the EXTNAME keyword of the corresponding extension.

The 'columns' entry in the parameter file should be:

```
columns: EXTNAME_for_wave EXTNAME_for_flux EXTNAME_for_error EXTNAME_for_mask
```

If the error or the mask is not provided, the corresponding entry should be replaced by NULL.

Examples:

- X-shooter visible spectrum reduced by the pipeline (esorex, gasgano, reflex), with the FITS keyword PRO.CATG including any of the MERGE1D string.

  ```
  columns: NULL FLUX ERRS QUAL
  ```

- X-shooter Internal Data Product spectrum (e.g. retrieved from the Phase 3 archive):

  ```
  columns: WAVE FLUX ERR QUAL
  ```

4.9 The output files

4.9.1 Output overview

The output files produced by `molecfit`, `calctrans`, `calctrans_lblrtm`, `calctrans_convolution` and `corrfilelist` are stored in the directory specified by the OUT-PUT_DIR parameter (see Section 4.7). The following output files (named corresponding to the OUTPUT_NAME parameter) are created by `molecfit`:

- <OUTPUT_NAME>.fits: input file converted to FITS table and with additional mask colum (1 = selected, 0 = rejected). This file can also be produced by `preptable`.

- <OUTPUT_NAME>_fit.par: copy of input parameter file.

- <OUTPUT_NAME>_fit.atm: final atmospheric profile for the best fit (incorporates standard profile + GDAS + EMM + fit of the molecular abundances).

- <OUTPUT_NAME>_fit.fits: FITS table containing the observed input spectrum and the modelled spectrum. There are 10 to 11 columns: #1: chip, #2: wavelength grid of input spectrum converted into μm and vacuum, #3: observed input spectrum (radiance/ transmission), #4: weight of observed data, #5: number of fit range (0 if not fitted), #6: model wavelengths in μm and vacuum, #7: continuum scaling function, #8: modelled output spectrum (radiance/transmission), #9: weight of model pixels (0 = no valid calculation), #10: weighted deviation between modelled and observed spectrum, #11: transmission curve for fitted wavelength ranges (only for transmission case).

- <OUTPUT_NAME>_fit.asc: ASCII version of the FITS table described above.

- <OUTPUT_NAME>_fit.ps: (optional) postscript plot showing a comparison of the best-fit model and input observed spectrum and the difference of both spectra. See parameter PLOT_CREATION.

- <OUTPUT_NAME>_fit_<fitrange>.ps: (optional) postscript plots for the individual fit ranges. See the parameters PLOT_CREATION and PLOT_RANGE.

- <OUTPUT_NAME>.res: results file containing information on the fit quality and the best-fit parameters (see Section 4.9.2).

The executable `calctrans` creates the following files:

- <OUTPUT_NAME>_tac.fits: FITS table containing the results of the telluric absorption correction of the input spectrum. There are 9 columns: #1: chip, #2: wavelength grid of input spectrum converted into μm and vacuum, #3: observed input spectrum (radiance/transmission), #4: weight of observed data, #5: model wavelengths in μm and vacuum, #6: best-fit transmission curve (correction function for telluric absorption), #7: weight of model pixels (0 = no valid calculation), #8: input spectrum corrected for telluric absorption, #9: flag indicating the quality of the telluric absorption correction (0 = very low transmission → zero-point uncertainties are crucial → numerical problems expected, 1 = probably OK). A model transmission curve (column #6) is also calculated for input sky radiance spectra. However, in this case, no telluric absorption correction is carried out (column #8).

- <OUTPUT_NAME>_tac.asc: ASCII version of the FITS table described above.

- <OUTPUT_NAME>_tac.ps: (optional) postscript plot showing a comparison of the input spectrum with and without telluric absorption correction. See parameter PLOT_CREATION.

- <FILENAME>_TAC.<filetype>: telluric absorption corrected input spectrum in the same file format as the input file. For ASCII and FITS tables, a corresponding column plus the quality flag column are added. If an error column exists, a column for the corrected error is also created. For FITS images, the flux spectrum is substituted by the corrected spectrum and existing flux error and mask/quality FITS extensions are modified.

- <FILENAME>_TRA.<filetype>: 1D FITS image without extensions that holds the correction function for telluric absorption. This file is only created if the input file is a FITS image.

Alternatively, the executable `calctrans_lblrtm` creates the following files:

- <OUTPUT_NAME>_N_T|R.fits: FITS table containing the result of the telluric absorption features where:

  - N is a number corresponding to the fit range,
  - T or R is appended to the filename according to the user's request: transmission or emission spectrum (`trans` keyword in the input parameter file).

- <OUTPUT_NAME>_code_stat.flags: flags used internally between `calctrans_lblrtm` and `calctrans_convolution`.

The executable `calctrans_convolution` creates the following files:

- <OUTPUT_NAME>_tac.fits

- <OUTPUT_NAME>_tac.asc

- <OUTPUT_NAME>_tac.ps

- <FILENAME>_TAC.<filetype>

- <FILENAME>_TRA.<filetype>

(see the paragraph above about `calctrans` for a detailed description of those files)

Finally, `corrfilelist` produces a <filename>_TAC.<filetype> file (see above) for each file listed in LIST-NAME.

4.9.2 Example of a `.res` file

The <OUTPUT_NAME>_fit.res file contains detailed information on the fit results. In particular, information on the fit quality, i.e. χ^2 and RMS values, all coefficients of the best-fit model, uncertainties of these coefficients (only if a parameter was fitted), and the final water column are given. For the provided `mpfit` status message, see the documentation in `mpfit.h`. In general, positive numbers imply that the code found a solution.

In the following, the output file belonging to the parameter file listed in Section 4.6 is shown:

```
DATA FILE:
examples/input/crires_spec_jitter_extracted_0000.fits

MPFIT RESULTS:
Status:                 2
Fit parameters:         38
Data points:            4096
```

```
Weight > 0:                   3896
Frac. of valid model pix.:    1.00
Iterations:                   13
Function evaluations:         154
Fit run time in min:          1.15
Avg. LBLRTM run time in s:    3.76
LBLRTM calls:                 16
Initial chi2:                 1.452e+09
Best chi2:                    4.562e+05
Reduced chi2:                 1.171e+02
RMS rel. to error:            1.082e+01
RMS rel. to mean:             1.990e-02

BEST-FIT PARAMETERS:


SPECTRAL RESOLUTION:
Rel. FWHM of boxcar (slit width = 1): 0.000
FWHM of boxcar in pixels:          0.000
FWHM of Gaussian in pixels:        1.940 +- 0.003
FWHM of Lorentzian in pixels:      0.500


WAVELENGTH SOLUTION:
Chip 1, coef 0: -5.082e-03 +- 4.356e-06
Chip 1, coef 1:  1.001e+00 +- 7.231e-06
Chip 1, coef 2:  2.085e-03 +- 7.277e-06
Chip 1, coef 3:  5.023e-04 +- 6.830e-06
Chip 2, coef 0:  2.314e-03 +- 4.339e-06
Chip 2, coef 1:  1.004e+00 +- 9.810e-06
Chip 2, coef 2:  5.840e-03 +- 6.325e-06
Chip 2, coef 3:  4.218e-05 +- 6.381e-06
Chip 3, coef 0: -3.487e-04 +- 7.627e-06
Chip 3, coef 1:  9.995e-01 +- 1.836e-05
Chip 3, coef 2:  1.551e-03 +- 9.098e-06
Chip 3, coef 3: -7.590e-05 +- 8.770e-06
Chip 4, coef 0: -1.360e-03 +- 3.974e-06
Chip 4, coef 1:  1.001e+00 +- 7.622e-06
Chip 4, coef 2: -1.099e-03 +- 6.119e-06
Chip 4, coef 3:  1.735e-05 +- 6.554e-06


CONTINUUM CORRECTION:
Range 1, chip 1, coef 0:  4.395e+01 +- 3.220e-03
Range 1, chip 1, coef 1:  6.602e+00 +- 1.070e+00
Range 1, chip 1, coef 2: -8.195e+03 +- 9.750e+01
Range 1, chip 1, coef 3: -5.589e+05 +- 2.507e+04
Range 2, chip 2, coef 0:  4.375e+01 +- 3.903e-03
```

```
Range 2, chip 2, coef 1:  1.293e+02 +- 1.375e+00
Range 2, chip 2, coef 2: -3.498e+04 +- 1.412e+02
Range 2, chip 2, coef 3: -2.385e+06 +- 3.431e+04
Range 3, chip 3, coef 0:  3.624e+01 +- 7.714e-03
Range 3, chip 3, coef 1: -1.796e+02 +- 2.219e+00
Range 3, chip 3, coef 2: -5.348e+03 +- 6.280e+02
Range 3, chip 3, coef 3:  2.646e+06 +- 8.653e+04
Range 4, chip 4, coef 0:  3.245e+01 +- 3.088e-03
Range 4, chip 4, coef 1:  5.918e+01 +- 1.094e+00
Range 4, chip 4, coef 2: -3.674e+03 +- 1.265e+02
Range 4, chip 4, coef 3:  2.851e+05 +- 3.342e+04


RELATIVE MOLECULAR GAS COLUMNS:
H2O: 0.885 +- 0.000
CH4: 0.946 +- 0.000
 O3: 0.932 +- 0.001


MOLECULAR GAS COLUMNS IN PPMV:
H2O: 2.051e+02 +- 4.099e-02
CH4: 1.639e+00 +- 3.884e-04
 O3: 3.909e-01 +- 4.666e-04


H2O COLUMN IN MM: 0.991 +- 0.000
```

4.10 Known Issues

- If the data have pixels with unusually high values, Molecfit may fail since the underlying algorithm uses single precision floats. A solution is to apply a scale factor to the data before the fit.

- The old MacOSX file system (HDF+) have case insensitivity. The user must be sure to put a different name in the 'output_name' parameter that the name of the input file in Molecfit. If you don't take care of that the GUI will not show any error message in the screen in the old MacOSX file systems, but it will not storage the output correction file.

5 The model

In this section, the atmospheric model used for `molecfit` is described in more detail. First, the building and properties of the atmospheric profiles required for the calculation of emission and absorption spectra are discussed (Section 5.1). Then, we explain the properties of the radiative transfer code used (Section 5.2). The contribution of the different molecules to the resulting atmospheric spectra is discussed in Section 5.3. Moreover, we provide some information on the modelling of the telescope emission (Section 5.4). Finally, we describe how the resulting model is adapted to the input science spectrum (Section 5.5).

5.1 Atmospheric profiles and meteorological data

Information concerning the composition of the atmosphere is available at various levels. To the end of creating a uniform profile with the variables temperature, pressure, and density of various molecular species as a function of geoelevation, three sources of input are merged: standard profile (produced for Michelson Interferometer for Passive Atmospheric Sounding (MIPAS) onboard the ENVISAT satellite), GDAS profile, and EMM data.

The largest amount of molecular density information is contained in the atmospheric standard profiles. However, they are only available for specific geographical latitudes and do not contain any time information whatsoever (see Section 5.1.1). To compensate the lack of time information, one can rely on the EMM (see Section 5.1.3). It provides the most frequent updates and is specific to the selected observing site. Unfortunately, it cannot provide molecular species information apart from water vapour (relative humidity measurements) and is restricted to a local on-site measurement, *i.e.* a single geoelevation data point only. To bridge the gap between these two data sources, GDAS provides a global grid of profile measurements (with approximate grid spacing of 110 km) to an altitude of ~ 26 km with updates every three hours. GDAS does not contain molecular species apart from H_2O, though (see Section 5.1.2).

These three data sources and the required processing for use with `molecfit` are described in detail below (see also Noll et al. [2012]).

5.1.1 MIPAS profiles

The atmospheric standard profiles provide the basis for the model atmosphere used in `molecfit` (see parameter REF_ATM in Section 4.7) including information on pressure, temperature, and molecular abundance as function of height (121 levels in the range 0-120 km). Up to now, the Reference Forward Model (RFM) homepage [1] provides standard profiles for mid-latitude (Lat = 45°, both, day and night), polar winter/summer (Lat = 75°) and equatorial day-time conditions in such a configuration (J. Remedios 2001). So far, the following molecules are included in this standard profile: N_2, O_2, CO_2, O_3, H_2O, CH_4, N_2O, HNO_3, CO, NO_2, N_2O_5, ClO, $HOCl$, $ClONO_2$, NO, HNO_4, HCN, NH_3, F11, F12, F14, F22, CCl_4, COF_2, H_2O_2, C_2H_2, C_2H_6, OCS, SO_2, and SF (see Table 2). Additional molecule profiles for F13 (CClF3), F21 ($CHCl_2F$), F113 ($C_2Cl_3F_3$), F114 ($C_2Cl_2F_4$), F115 (C_2ClF_5), and CH_3Cl are available. Apart from these data, less resolved profiles (a tropical, sub-arctic summer/winter and a US standard profile) are available with 50 geoelevation layers including the molecules H_2O, CO_2, O_3, N_2O, CO, CH_4, and O_2 only.

Comparison between atmospheric standard profiles

In this section, an equatorial day-time `equ.atm` and a mid-latitude profile `ngt.atm`, corresponding to a latitude Lat = 45°, will be compared. The location of Paranal (Lat = 24.6°) is between these two profiles. In Figure 3 both profiles are shown. Although the distribution of several molecules (*e.g.* N_2, O_2, CO_2) does not vary, significant differences between the two profiles are visible. To investigate the impact of the input profile differences on the output spectra, LBLRTM was run with the same input parameters, but with varying standard profiles.

The resulting spectra are shown in Figures 4/5. These plots reveal output radiance spectra differing by at most ±10%. The same is true for the transmission spectra, although somewhat less obvious due to numerical instabilities. Calculating the broad-band ratios in the main filter ranges $UBVRI_cJHKLMN$ indicates deviations of less than 2% (see Table 1). Hence, one can conclude that the differences between the two standard atmospheric profiles are negligible at this stage. Anu Dudhia [Anu Dudhia 2009, priv.comm.], the author of the RFM code, recommends the equatorial profile `equ.atm` to be used for typical applications at Cerro Paranal.

Table 1: Broad-band comparison of the relative ratios between the `equ.atm` and the `ngt.atm` atmospheric standard profiles.

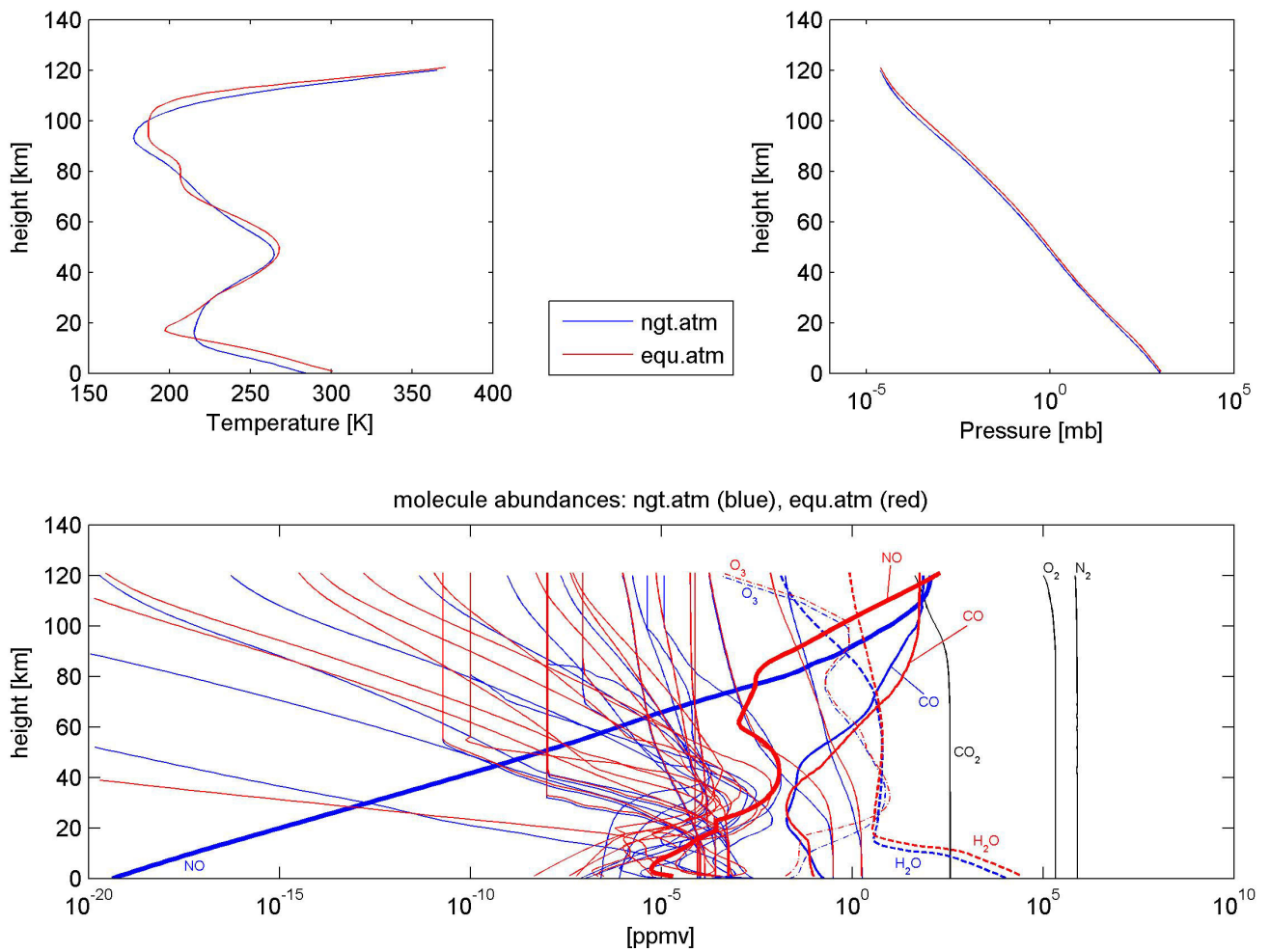| Filter	λ_{min} [µm]	λ_{max} [µm]	Radiance ratio [%]	Transmission ratio [%]
U	0.33	0.40	0.02	0.06
B	0.39	0.50	0.00	0.02
V	0.50	0.60	0.14	0.56
R	0.58	0.82	0.05	0.28
I_c	0.73	0.85	0.01	0.04
J	1.10	1.34	-0.00	-0.00
H	1.50	1.80	-0.00	-0.01
K	2.00	2.40	-0.05	-0.06
L	3.56	4.12	-0.00	-0.05
M	4.52	4.96	-0.69	1.37
N	7.40	13.60	-0.60	1.39

Figure 3: Comparison of the equatorial (`equ.atm`) and the mid-latitude night time atmospheric profile (`ngt.atm`). Red lines correspond to the equatorial, blue lines to the mid-latitude profile.
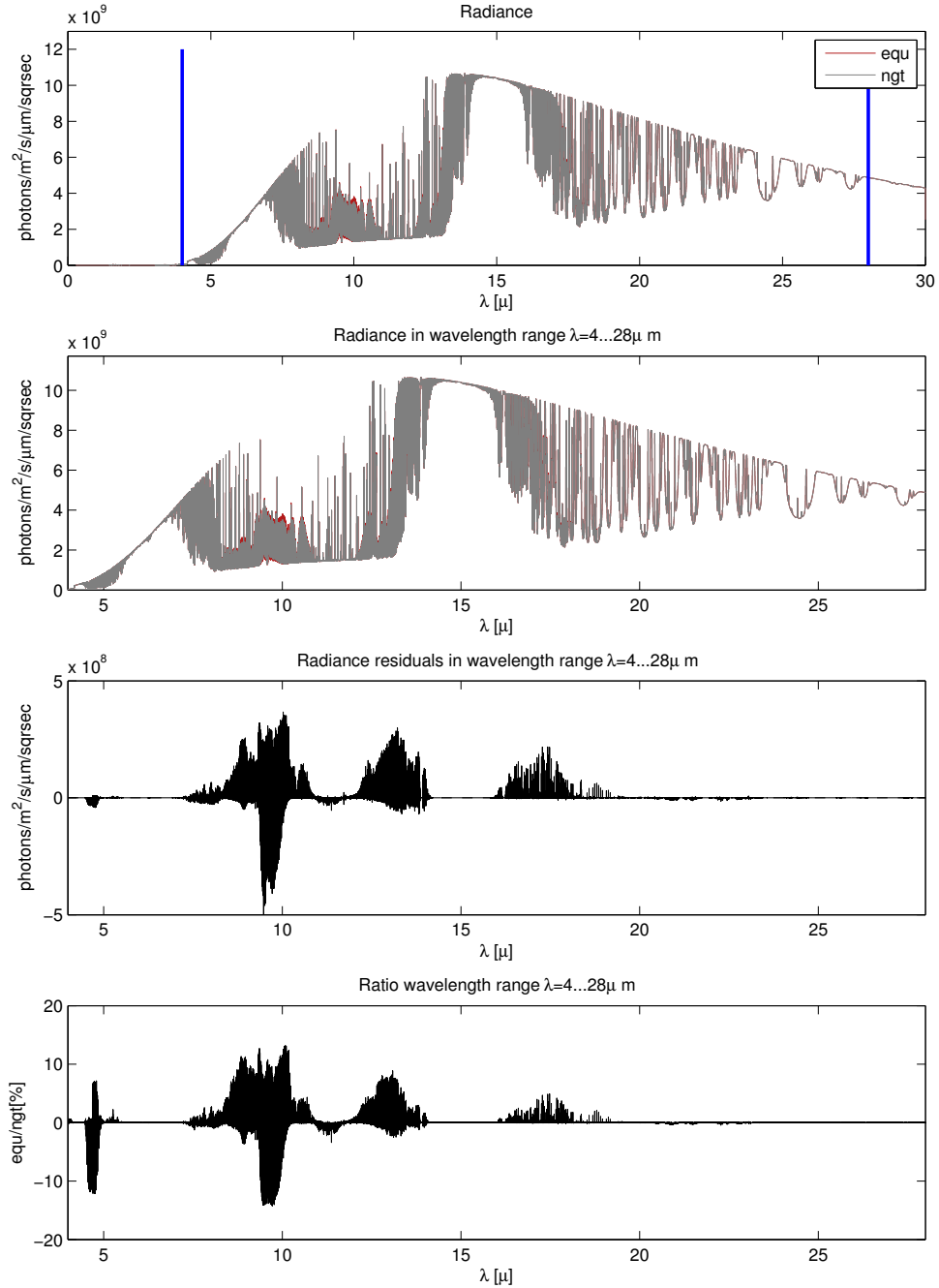
Figure 4: Direct comparison between the radiance spectra of equatorial day time (red line) and mid-latitude night time atmospheric standard profile (grey line) over the entire wavelength range $\lambda = 0.3 - 30\,\mu\text{m}$ (top panel). Blue lines mark the wavelength range ($\lambda = 4 - 28\,\mu\text{m}$) plotted in the three panels below. *Second panel*: Same as in top panel, but for the limited wavelength range. *Third and bottom panel*: Residuals equ-ngt and ratio equ/ngt of the radiance spectra, respectively.
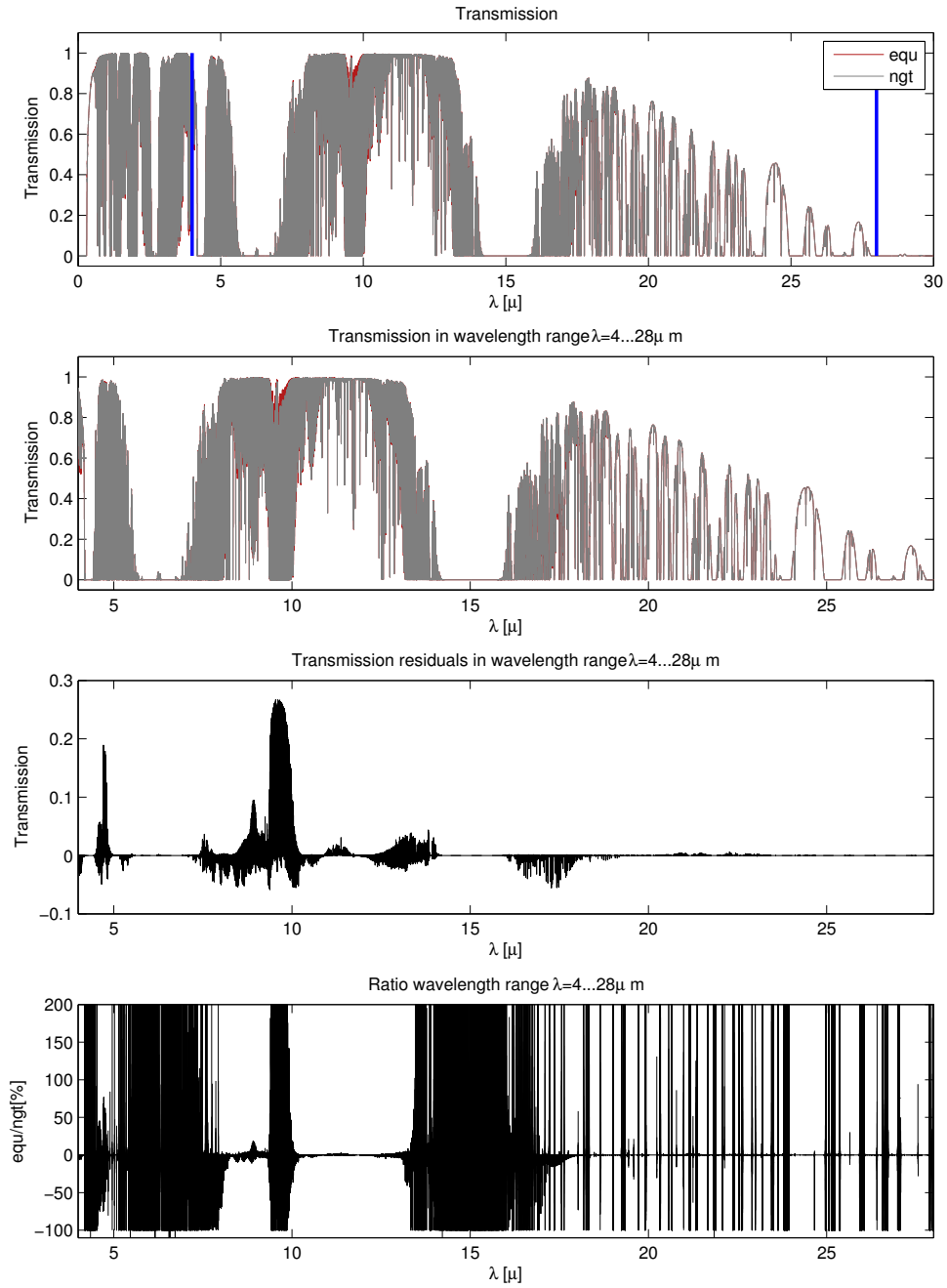
Figure 5: Same as Figure 4, but for the transmission.

5.1.2 GDAS profiles

The GDAS data provided by NOAA are a model-based set of meteorological data dedicated to weather forecast studies. The models are archived by the ARL, as a global, 1 degree latitude/longitude data set based on pressure surfaces (starting from Dec. 2004). Apart from various meteorological parameters for the surface, vertical profiles for 23 pressure levels ranging from 0 to about 26 km are provided for the geopotential height, temperature, relative humidity, and wind components (not used in `molecfit`) for three dimensions. An example is shown in Figure 6.

```
# P[hPa]  HGT[km]  T[K]    RELHUM[%]
    903    0.971   294.5    49.9
    900    0.976   295.8    35.5
    850    1.467   293.7    30.9
    800    1.985   291.0    29.8
    750    2.533   288.2    27.4
    700    3.112   284.5    26.0
    650    3.726   280.7    18.8
    600    4.379   276.2    11.4
    550    5.077   271.4     8.7
    500    5.827   266.0     7.5
    450    6.638   259.9     7.5
    400    7.522   252.7    10.5
    350    8.494   244.7    25.1
    300    9.578   236.0    53.4
    250   10.813   227.0    66.5
    200   12.267   218.8    37.7
    150   14.069   209.2    17.3
    100   16.489   200.3    32.4
     50   20.571   206.1     0.0
     20   26.324   221.4     0.0
```

Figure 6: *GDAS profile* with columns for pressure, geoelevation, temperature, and relative humidity.

The `molecfit` software package provides the entire GDAS data for the location of Cerro Paranal from Dec. 2004 to Sep. 2013 on a 3 h basis taken from the NOAA archive[1]. Later dates (on a 6 h basis) or data for a different site are automatically downloaded from an online archive[2].

5.1.3 ESO Meteo Monitor

The EMM provides information on the local meteorological conditions at the ESO sites La Silla and Paranal. The data at Paranal are taken by a local meteo station mounted on a 30 m high mast installed in October 1984 [7]. This meteo station provides the following meteorological information on a 20 min average basis:

```
s1,d1 = wind speed (m/s) and direction (0=North, 90=East..)
```

---

[1]ftp://arlftp.arlhq.noaa.gov/pub/archives/gdas1/
[2]http://nomad1.ncep.noaa.gov/pub/gdas/rotating/

```
        at 10m above ground (30m from 1998 onwards)
  rh = relative humidity (%) at 2m above ground
  t1 = air temperature (Celsius) at 2m above ground
  p  = pressure (mb) at 2m above ground
  td = dew point temperature (C) computed from rh and t1
```

Starting from January 1st 1985, currently ~ 400 000 data points are measured with the following accuracy [7]:

```
wind direction: ~5.63deg
wind speed:     ~2% over 10m/s
temperature:    ~0.1deg
humidity:       linearity about 1%
seeing:         better than 10% above 0.25 arcsec
```

The data can be retrieved online[3] on a daily basis, or as download provided by M. Sarazin[4] and are cumulatively shown in Figure 7. Thus, for any requested average time interval, like *e.g.* December and January, ample measurements are available. In compiling these data, care has to be taken to remove bad measurements before further processing.
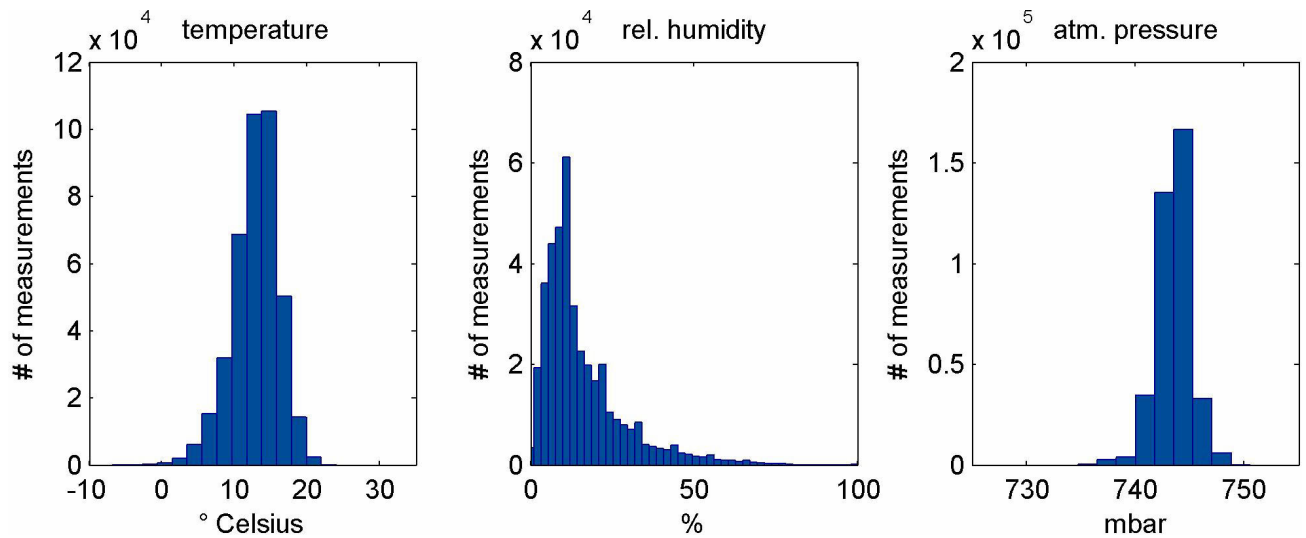

Figure 7: Histograms of EMM data (from Jan. 1985 to Jan. 2008). *Left panel*: temperature; *middle panel*: relative humidity; *right panel*: pressure.

---

[3]http://archive.eso.org/asm/ambient-server
[4]http://www.eso.org/gen-fac/pubs/astclim/paranal/database/

5.1.4   Processing of ESO Meteor Monitor data, GDAS, and MIPAS profiles

The main disadvantage of the GDAS profiles is that they do not represent the local atmospheric conditions of the geographical position and height of the observing site as accurately as provided by the EMM, and even more so for the MIPAS profiles. Therefore, one has to investigate how the three sources of information can be merged into a single profile.

GDAS profile processing

The GDAS profiles originate from a server at NOAA[5]. They are retrieved via a dedicated software package GRIdded Binary (GRIB) (see Section 3). GRIB downloads a large data set containing the specific GDAS information for the requested point in time. As this data set contains a model for the complete globe, subsequently, the data points for the specified geolocation are extracted. Moreover, as the GDAS data are taken on a 3 hr basis only, two profiles need to be retrieved surrounding the requested point in time. The parameters OBSDATE, UTC, LONGITUDE, and LATITUDE are required for this task (see Section 4.7). They are usually provided by standard and ESO FITS keywords. In the following, we will describe how the resulting two profiles are combined to best match the date of the observations.

If the profiles exist locally, no download from the web-server is required. In addition, the `molecfit` software distribution contains a compilation of all Cerro Paranal GDAS profiles for the dates from Dec. 01, 2004 to Sep. 30, 2013. Thus, before requesting the data from the web-server (in the case that they do not exist locally already), this database is checked for the existence of the appropriate profiles. For updates of this data set and the retrieval of data for other observing sites, see Section 7.2.4.

Unfortunately, the web-server does not provide GDAS profiles for all dates. Therefore, `molecfit` incorporates a fall-back alternative to ensure availability of GDAS data in all occasions. To that end, the monthly averaged profiles from the Cerro Paranal sky model are included as well (for a detailed description see [SM-01-UM] and Noll et al. [2012]). If after checking the local database or the web-server, a profile is still missing, the best-matching profile for the two-month bins Dec.-Jan., Feb.-Mar., ... are taken. This is also done if the site is not Cerro Paranal.

The described procedure for the GDAS profile retrieval is performed if the parameter GDAS_PROF (see Section 4.7) is set to "auto", which is the default. As an alternative, a specific GDAS-like profile (see Table 6) might be provided. Finally, it is possible to avoid the use of GDAS profiles by setting GDAS_PROF to "none".

Time averaged profiles

Typically, the requested observation date does not fall exactly onto a single GDAS time slot. Instead of simply retrieving the closest dataset the two neighbouring profiles are obtained. In order to combine the two, a time-weighted average is calculated, *i.e.* performing a linear interpolation.

Merging GDAS and MIPAS profiles

Next, the resulting GDAS profile is merged with the MIPAS standard profile. To that end, the MIPAS profile is regridded to a new irregular height grid with 50 levels (see Figure 8) spanning the whole geoelevation range

---

[5]http://140.90.198.158/pub/gdas/rotating/

from 2-120 km for Cerro Paranal. The GDAS profile is regridded to the same grid in the range 2-26 km and then used to substitute the respective columns in the MIPAS profile. In addition, the four height levels from 20-26 km are not only a simple substitute of the MIPAS data, but a weighted mix of GDAS and MIPAS profile, in order to provide a smooth transition from one dataset to the other. The influence of the GDAS profile decreases with increasing height: 80%, 60%, 40%, 20% at 20 km, 22 km, 24 km, 26 km, respectively. Beyond 26 km, no GDAS information is available.

The discussed fixed grid of layers is used if the parameter LAYERS (see Section 4.7) is set to 1, which is the default. A value of 0 will cause the building of a natural grid consisting of all layers of the MIPAS and the GDAS profile. If local meteo data are used, the observer altitude GEOELEV is also added. The transition from GDAS to MIPAS is performed by means of a decrease of the relative difference of pressure, temperature, and water vapour concentration of both profiles at the height of the uppermost valid GDAS layer up to an altitude which is 1.2 times higher. The resulting grid of layers is (slightly) more accurate than the fixed grid, but also consists of a significantly higher number of levels.
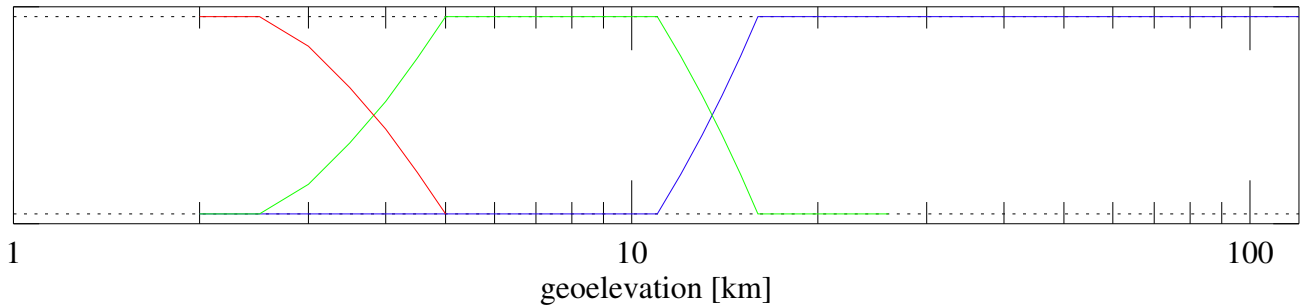

1	10	100

geoelevation [km]

Figure 8: *Composition of atmospheric profile:* relative importance of EMM data (red), GDAS data (green) and MIPAS data (blue) as function of geoelevation. Note that the interface boundary of GDAS and MIPAS data varies depending on the availability of the GDAS data.

Combining GDAS/MIPAS profiles with EMM data

Observed data from ESO telescopes provide on-site measurements at the ground layer for pressure, temperature, and humidity originating from the ESO meteo monitor at Cerro Paranal. A detailed study of the GDAS data (see Figure 9), which represent the local troposphere, including information concerning the dominant wind direction as a function of altitude reveals a gradual reversal (rotation of 180°) at a geoelevation of 5 km, the so-called mixing altitude h_{mix} (see [SM-01-UM]). Beyond this altitude, the wind direction remains constant independent of the observation date. Thus, it can safely be assumed that at this altitude the influence of the local environment (as determined from the EMM data) has diminished.

In order to smoothly integrate the EMM data, all GDAS values for pressure, temperature, and humidity below the altitude of the observatory are set to the EMM value. Values above the aforementioned mixing altitude are left untouched. Intermediate values are linearly interpolated resulting in a smooth transition. To this end, first, a logarithmically interpolated value of the GDAS data corresponding to the observatory's altitude h_{tel} is calculated for pressure, temperature, and humidity. These values describe the reference point for the linear decrease of the relative difference between EMM and GDAS data and in the interval $h_{\mathrm{tel}} - h_{\mathrm{mix}}$.

The resulting profile is a smooth combination of all input data, *i.e.* MIPAS, GDAS, and EMM.

The default mixing altitude of 5 km can be manipulated by changing the parameter EMIX (see Section 4.7). This could be interesting for other observing sites. Setting EMIX to a value lower than the observer altitude GEOELEV causes a profile building without local meteo data. In this case, the parameters PRES, TEMP, and RHUM are ignored.
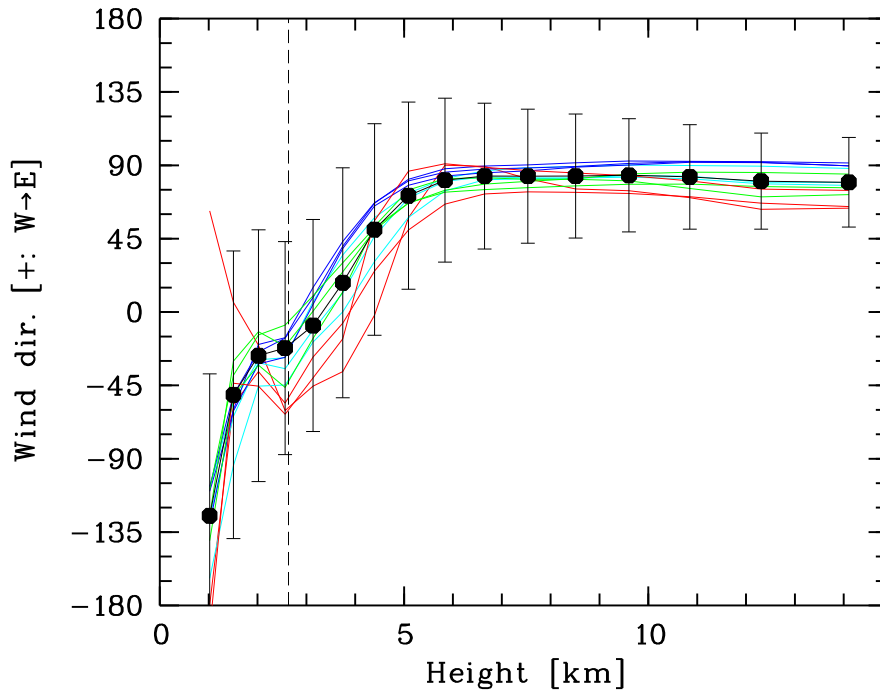

Figure 9: *GDAS wind direction as function of geoelevation:* red, cyan, blue, and green curves represent summer, autumn, winter, and spring, respectively. The black symbols show the all year average with the corresponding scatter. The vertical dashed line marks the geoelevation of Paranal. At ~ 5 km height, a constant plateau is reached.

Scaling of the merged water vapour profile to a given PWV

The height profiles of the molecular abundances are defined by the atmospheric standard profile. The only exception is water vapour, where the profile is a combination of modelled and observed data as described above. The GDAS and EMM data used cannot be controlled by the user. There might be cases where this is not satisfying. For example, the user might be interested to use water vapour columns derived from independent measurements for a better start profile. For this purpose, `molecfit` allows the user to enter a precipitable water vapour (PWV) value (parameter PWV, see Section 4.7), which is used to scale the merged water vapour profile. In this way, the input water vapour content of the atmosphere can be fixed. Only the shape of the profile is then ruled by the GDAS and EMM data. The profile scaling factor RELCOL, which is used in the context of the fitting procedure, refers to the modified profile. By default, the PWV option is switched off as indicated by a value of -1.

5.2 Radiative transfer code

`molecfit` uses the radiative transfer code Line-By-Line Radiative Transfer Model (LNFL v3.1 / LBLRTM v12.8), which is widely used in atmospheric and climate research studies.

5.2.1 Line File / Line-By-Line Radiative Transfer Model (LNFL/LBLRTM)

LBLRTM is developed within the Radiative Transfer Working Group of the AER (see also Clough et al. [2005], [5], and [2]). It is publicly available. LBLRTM can handle all molecules incorporated in the `aer_v_<version>` line parameter database [2] and offers a wide range of possibilities to adjust input parameters (see [12] for more details).

The AER code package used here consists of two programmes: (a) the "Line File" code LNFL, which extracts user selected spectral lines from the line parameter database, and provides these in appropriate form as input for (b) the radiative transfer code LBLRTM. Within `molecfit`, the most recent versions LNFL v3.1 and LBLRTM v12.8 are used.

Some LBLRTM key features are (taken from [2]):

- the Voigt line shape is used at all atmospheric levels with an algorithm based on a linear combination of approximating functions;

- it has been and continues to be extensively validated against atmospheric radiance spectra from the ultraviolet to the sub-millimeter;

- it incorporates the self- and foreign-broadened water vapour continuum model, MT_CKD as well as continua for carbon dioxide, and for the collision induced bands of oxygen at $1600\,\text{cm}^{-1}$ ($\lambda = 6.25\,\mu\text{m}$) and nitrogen at $2350\,\text{cm}^{-1}$ ($\lambda = 4.255\,\mu\text{m}$);

- all parameters of the line database are used including the pressure shift coefficient, the halfwidth temperature dependence, and the coefficient for the self-broadening of water vapour;

- a version of the Total Internal Partition Function (TIPS) programme is used for the temperature dependence of the line intensities;

- the effects of CO_2 line coupling are treated as first order with the coefficients for carbon dioxide generated from Niro et al. [2005];

- temperature dependent cross section data such as those available with the `aer_v_<version>` database may be used to treat the absorption due to heavy molecules, *e.g.* the halocarbons;

- an algorithm is implemented for the treatment of the variation of the Planck function within a vertically inhomogeneous layer as discussed in Clough et al. [1992];

- algorithmic accuracy of LBLRTM is approximately 0.5% and the errors associated with the computational procedures are of the order of five times less than those associated with the line parameters so that the limiting error is that attributable to the line parameters and the line shape;

- its computational efficiency mitigates the computational burden of the line-by-line flux and cooling rate calculation (Clough et al. [1992]), for example linear algebraic operations are used extensively in the computationally intensive parts of LBLRTM so that vectorisation is particularly effective with a typical vectorised acceleration of 20;

- FFT instrument function with a choice of 9 apodisation functions;

- includes a realistic spectral sea surface emissivity model in the infrared (Masuda et al. [1988]; Wu&Smith [1997]);

- input atmospheric profiles in either altitude or pressure coordinates;

- interfaces with other radiative transfer models (like RRTM), and also the forward model for inversion algorithms (*e.g.* Tropospheric Emission Spectrometer (TES) and Infrared Atmospheric Sounding Interferometer (IASI));

- these attributes provide spectral radiance calculations with accuracies consistent with the measurements against which they are validated and with computational times that greatly facilitate the application of the line-by-line approach to current radiative transfer applications.

In principle, the user can change the setup of the LBLRTM. A reduced but well documented list of input parameters and their default values is given by the file "config/lblrtm_setup". Some parameters are not fixed. They are updated by other input data. For example, the wavelength range is given by the input spectrum and possible cuts provided by the driver file (see Section 4.7). A change of the fixed LBLRTM input parameters is only recommendable for those users, which have a very good knowledge of the physics of atmospheric radiative transfer and/or LBLRTM.

5.2.2 `aer` **line database**

For calculating molecular spectra, the `aer_v_<version>` database [2] is used. It is built from HITRAN 2008 [6] and contains several updates. Currently, it covers the full spectral range from $0 - 25,232 \, \text{cm}^{-1}$ (*i.e.* up to $0.4 \, \mu\text{m}$) provides spectral information for 42 molecules. In total, more than 2,700,000 spectral lines are included. The majority is based on modelled data. However, only those 30 molecules are taken into account, which are present in the atmospheric standard profile. The remaining ones are minor trace gases and do not contribute significantly neither to radiance, nor transmission spectra (see Section 5.3). Table 5.2.2 provides an overview of all molecules based on the `aer` database (and known by LBLRTM) and those contained in the standard atmospheres (Column 4).

Table 2: List of molecules as provided by the `aer` line parameter database

# of molecule	molecule	alt. name	std. atmosphere	LBLRTM
1	H_2O	Water	X	X
2	CO_2	Carbon dioxide	X	X
3	O_3	Ozone	X	X
4	N_2O	Nitrous oxide	X	X
5	CO	Carbon monoxide	X	X
6	CH_4	Methane	X	X
7	O_2	Oxygen	X	X
8	NO	Nitric oxide	X	X
9	SO_2	Sulfur dioxide	X	X
10	NO_2	Nitrogen dioxide	X	X
11	NH_3	Ammonia	X	X
12	HNO_3	Nitric acid	X	X
13	OH	Hydroxyl		X
14	HF	Hydrogen fluoride		X
15	HCl	Hydrogen chloride		X
16	HBr	Hydrobromic acid		X
17	HI	Hydrogen iodide		X
18	ClO	Chlorine monoxide	X	X
19	OCS	Carbonyl sulfide	X	X
20	H_2CO	Formaldehyde		X
21	HOCl	Hypochlorous acid	X	X
22	N_2	Nitrogen	X	X
23	HCN	Hydrogen cyanide	X	X
24	CH_3Cl	Chloromethane		X
25	H_2O_2	Hydrogen peroxide	X	X
26	C_2H_2	Acetylene	X	X
27	C_2H_6	Ethane	X	X
28	PH_3	Phosphine		X
29	COF_2	Carbonyl fluoride	X	X
30	SF_6q	Sulfur hexafluoride	X	X
31	H_2S	Hydrogen sulfide		X
32	HCOOH	Formic acid		X
33	HO_2	Hydroperoxyl		X
34	O	Oxygen		X
35	$ClONO_2q$		X	X
36	NO+	Nitrosonium		X
37	HOBr			X
38	C_2H_4	Ethylene		X
39	CH_3OH	Methanol		X
40	BrO			X
41	C_3H_8	Propane		X
42	C_2N_2	Cyanogen		X

5.3 Molecular spectra

The process of fitting molecular spectra is a complex task, which requires optimised input parameters. One of the key inputs for the fitting is the number of molecules that are included in the fitting. Fewer molecules in the fitting process result in the code finding a solution in significantly shorter amounts of time. Also the fitting process is a lot more robust. However, if too few molecules are included in the fit, the results may not provide a satisfying residual for model and observation. Optimally, one should include exactly those molecules in the fit, which will significantly contribute over the wavelength range of interest.

To the end of providing some insight into what molecules are required for the fit at a specific wavelength interval, in the following, we provide guidelines with a specific focus on the test data provided with this software package for the instruments CRIRES, VISIR, and X-Shooter (see Section 6.1).

We have computed spectra with LBLRTM covering the complete wavelength range from $0.3 - 30\,\mu\mathrm{m}$ using all molecules (*i.e.* C_2H_2, C_2H_6, CH_4, CO_2, COF_2, CO, $ClONO_2$, ClO, $F14$, H_2O_2, H_2O, HCN, HNO_3, $HOCl$, N_2O, N_2, NH_3, NO_2, NO, O_2, O_3, OCS, SF_6, SO_2), individually, *i.e.* one at a time. At fixed wavelength, we have then calculated the normalised total radiance/transmission (by summing up the contributions of all molecules and subtracting the continuum) and the resulting relative importance of individual molecules.

The Figures 18-22 shown in the Appendix, give all molecules that, over the displayed wavelength range, exhibit at fixed wavelength λ a relative importance of at least 5%. The molecular data have been rebinned to 3000 data points for each individual wavelength range. This results in a varying resolution and in more molecules becoming important over smaller wavelength regimes. C_2H_6, *e.g.*, has a few significant lines in some of the wavelength ranges shown (*e.g.* Figure 19), but does not show up in the overview plot (Figure 18). Typically, transmission (blue) and radiance (red) plots do not differ significantly. Hence, we do not show them separately. Note that these plots do not allow calculation of absolute fluxes.

These plots can be used to identify the important molecules over any wavelength range. For the range shown in Figure 19, *e.g.*, the user ought to include H_2O, CH_4, and O_3. The fitting might mildly profit also from including C_2H_6.

5.4 Thermal emission by telescope

The telescope structure and the observing instrument cause unavoidable thermal emission in the IR. In particular, the telescope main mirror is a significant source of radiation. Hence, for radiance spectra, this background component has to be considered. A simple approximation is the calculation of a grey body spectrum, which equals a black body (BB) of temperature T times a wavelength-independent emissivity ϵ. Since the emitting source, *i.e.* the main mirror, also absorbs a fraction of $1-\epsilon$ of the incoming sky radiation, the relative contribution of the telescope emission to the observed radiation is increased. Therefore, the apparent grey body radiation in flux-calibrated spectra can be derived by

$$F_{\mathrm{tel}} = \frac{\epsilon}{(1-\epsilon)}\,\mathrm{BB}(T). \tag{1}$$

For T, `molecfit` uses the temperature of the primary mirror (input parameter M1TEMP, see Section 4.7). Since the mirror temperature is close to the ambient temperature of about 280-290 K (Paranal), the grey body emission is expected to be important at wavelengths longwards of the *H*-band. The emissivity is a free fit

parameter (see parameter `TELBACK`). An initial value can be provided in the parameter file. The default value is 0.1.

5.5 Adaptation of model to input spectrum

A good correspondence of the calculated model spectrum and the observed spectrum is usually prevented by the broadening of the spectral lines by the instrument, small errors in the wavelength calibration, uncertainties in the flux calibration in the case of emission spectra, or the non-flat standard star continuum in the case of transmission spectra. Hence, these unavoidable shortcomings of observed data have to be accounted for in the fitting procedure. For this reason, `molecfit` modifies the model spectrum using a polynomial fit of the continuum and the wavelength grid. In addition, the model gets convolved with a kernel mimicking the instrumental profile. In the following, we discuss the fit parameters related to this adaptation process in detail.

5.5.1 The continuum

The model spectrum is scaled by a polynomial of degree n_c (`CONT_N` in the parameter file; see Section 4.7)

$$F_{\text{out}}(\lambda) = F_{\text{in}}(\lambda) \sum_{i=0}^{n_c} a_i \lambda^i. \tag{2}$$

For deriving the $n_c + 1$ coefficients a_i, the zero point of the wavelength grid is shifted to the centre of the fit range. For $a_0 = 1$ and all other $a_i = 0$, the model spectrum remains unchanged. This is the default configuration for the initial coefficients. In the parameter file, the initial value of the constant term of the polynomial (parameter `CONT_CONST`) can be set manually. The continuum correction is carried out independently for each fit range listed in the `RANGE_INCLUDE` file if such a file is provided. A fit range (or the full spectrum) is further split if it is distributed over more than one chip.

Before correcting the continuum, optionally a flux conversion is carried out. Details on the options selected by the parameter `FLUX_UNIT` are given in Section 4.7. If the required data units are not included in `FLUX_UNIT`, this factor must be incorporated into the a_0 coefficient of the polynomial. As a general rule, it is advisable to choose a_0 close to the mean flux (emission) or maximum flux (transmission) of the input spectrum (after consideration of `FLUX_UNIT`) to optimise the performance of `molecfit`.

5.5.2 The wavelength solution

The wavelength grid of the model spectrum is adapted to that of the observed spectrum by applying a Chebyshev polynomial of degree n_w (`WLC_N` in the parameter file; see Section 4.7)

$$\lambda' = \sum_{i=0}^{n_w} b_i t_i, \tag{3}$$

where

$$t_i = \begin{cases} 1 & \text{for } i = 0 \\ \lambda & \text{for } i = 1 \\ 2\lambda t_{i-1} - t_{i-2} & \text{for } i \geq 2 \end{cases} \tag{4}$$

and λ ranging from -1 to 1. The temporary conversion of the wavelength grid to a fixed interval results in coefficients b_i independent of the wavelength range and step size of the input spectrum. For $b_1 = 1$ and all other $b_i = 0$, the model spectrum remains unchanged. This is the default configuration for the initial coefficients. In the parameter file, the initial value of the constant term of the polynomial (parameter WLC_CONST) can be set manually. This parameter corresponds to a wavelength shift relative to half the full wavelength range. For each chip or FITS extension, the wavelength fit is carried out independently.

For checks or improvements of the input wavelength grid, the model wavelengths rebinned to the input grid are provided in the results tables of `molecfit` and `calctrans` (column "mlambda", see Section 4.9.1). The wavelengths are always given in μm and vacuum. Note that the reliability of this absolute wavelength calibration depends on the quality of the fit. Outside the selected fitting ranges, the wavelengths have to be interpolated or extrapolated by the Chebyshev polynomial. In particular, at optical and near-IR wavelengths, where strong absorption bands suitable for fitting are rare, the provided wavelengths have to be taken with care. For this reason, `molecfit` does not provide an automatic wavelength solution correction.

5.5.3 The resolution

The model spectrum is convolved with up to three different profiles in order to get similar line shapes as in the observed spectrum. If a profile is not desired, it can be skipped by setting its width and fit flag in the parameter file to zero (see Section 4.7).

The first kernel is a simple boxcar

$$F_{\text{box}}(\lambda) = \begin{cases} 1 & \text{for } -w_{\text{box}}/2 \leq \lambda \leq w_{\text{box}}/2 \\ 0 & \text{for } \lambda < -w_{\text{box}}/2 \cap \lambda > w_{\text{box}}/2 \end{cases} \tag{5}$$

which is adapted to the pixel scale and normalised to an integral of 1. In the parameter file, the width w_{box} (parameter RELRES_BOX) has to be given as fraction of the slit width, which is determined by the parameters SLITW, the slit width in arcsec, and PIXSC, the pixel scale in arcsec (see Section 4.7). By default, RELRES_BOX is set to 1, *i.e.* the slit width. The fit parameter RELRES_BOX can only vary between 0 and 2.

The second convolution kernel is a Gaussian

$$F_{\text{gauss}}(\lambda) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{\lambda^2}{2\sigma^2}\right) \tag{6}$$

centred on 0, where

$$\sigma = \frac{w_{\text{gauss}}}{2\sqrt{2\ln 2}}. \tag{7}$$

The FWHM w_{gauss} is given by the driver file parameter RES_GAUSS in pixels (see Section 4.7). It is restricted to values below 100 pixels. The default value is 1 pixel. The number of pixels in the kernel amounts to KERNFAC (default: 3) times w_{gauss}.

Finally, the third kernel is a Lorentzian

$$F_{\text{lorentz}}(\lambda) = \frac{1}{\pi} \frac{\lambda}{\lambda^2 + (w_{\text{lorentz}}/2)^2} \tag{8}$$

centred on 0, where w_{lorentz} is the FWHM. The width is adjusted by the driver file parameter RES_LORENTZ, which has to be provided in pixels (see Section 4.7). It is restricted to values below 100 pixels. The default

value is 1 pixel. The number of pixels in the kernel amounts to KERNFAC (default: 3) times w_{lorentz}. Compared to a Gaussian, the Lorentzian approaches the 0-level flux significantly slower, at much larger distances from the maximum.

Note that width zero components can occur in typical conditions, *e.g.* the line profile is very close to a pure Lorentzian shape. If this is not intended, the user should reduce the number of degrees of freedom by fixing individual fit components. A zero here identifies a unity convolution (*i.e.* no change of the input spectrum).

The combination of a Gaussian and a Lorentzian is called a Voigt profile. The flag KERNMODE (see Section 4.7) allows the user to apply only a single Voigt profile kernel, which is calculated by an approximate formula that takes the FWHM of Gaussian and Lorentzian as input. In this case (KERNMODE = 1), KERNFAC gives the kernel size in FWHM of the derived Voigt profile and not the FWHM of Gaussian and Lorentzian, as it is done for the default mode of two independent convolutions (KERNMODE = 0). For the fit results, the KERNMODE selection should be less important than the relative contributions of boxcar, Gaussian, and Lorentzian to the fitted line profile. Significant changes in the line profile can cause deviations in the water vapour column of more than 10% (cf. Section 6).

The parameter VARKERN allows the user to fit a kernel that linearly increases with wavelength (see Section 4.7). If the flag is set to 1, this option is selected. It is suitable for dispersion-dominated kernels and constant wavelength bins. In this case, the initial FWHM parameters are given for the central wavelength of the full wavelength range (considering the data of all chips). The default VARKERN = 0 assumes a constant kernel for the entire wavelength range. This option is suitable for narrow wavelength ranges and slit/object profile-dominated kernels.

Finally, the user can provide an optimised kernel by an ASCII file which is provided by the parameter KERNEL_FILE (see Section 4.7). If this option is used (*i.e.* the default "none" is replaced by the corresponding file name), the fixed input kernel overrules the creation of a kernel based on boxcar, Gaussian, and Lorentzian components. Since there will be no fit of the kernel shape and width, the line profile has to be known well. Note that VARKERN = 1 will not have an effect, since the pixel-based input kernel is wavelength independent.

Table 3: Description of test sample

Label	Inst.	T/R[a]	Resolution	λ^{b} [μm]	Date	Time [UT]
C1	CRIRES	T	high	$3.274 - 3.352$	2008/07/26	02:50
V1	VISIR	R	low	$10.37 - 12.50$	2009/10/25	04:07
V2	VISIR	R	low	$11.17 - 13.30$	2009/10/25	04:26
X1	X-Shooter	T	medium	$0.53 - 1.02$	2010/11/16	06:49

[a] transmission or radiance.

[b] The full wavelength range is shown. This is also the fit range except for X1, which was fitted in the two ranges $0.762 - 0.770\,\mu$m (part of the O_2 A-band) and $0.940 - 0.951\,\mu$m (H_2O absorption).

6 Code performance

In the following, we illustrate the handling of `molecfit` and the quality of the resulting fits by means of several example spectra. In Section 6.1, the derivation of molecular abundances is discussed. In particular, the quality of the derived water vapour column densities is investigated. In Section 6.2, we briefly discuss the performance of `molecfit` in terms of the correction of telluric absorption features. Finally, Section 6.3 provides some general rules that should be considered for a successful application of `molecfit`.

6.1 Derivation of molecular abundances

We illustrate the use of `molecfit` for the derivation of molecular abundances by means of a small sample of spectra consisting of a CRIRES, an X-Shooter, and a pair of VISIR spectra (see Table 3). The high-resolution ($R \approx 50000$) CRIRES standard star spectrum is typical of spectra that are used for the nocturnal PWV measurements at the VLT [14]. Despite of the narrow wavelength range from 3.274 to $3.352\,\mu$m, the good resolution allows one to fit many individual telluric absorption lines, which usually results in convincing fits of the water vapour content and abundances of other significant molecules in the atmosphere. The medium-resolution ($R \approx 8800$) X-Shooter VIS-arm star spectrum demonstrates [14] measurements in a narrow wavelength range at relatively short wavelengths (0.940 to $0.951\,\mu$m). Finally, the VISIR LR-mode spectra ($R \sim$ a few hundred) are examples for sky radiance spectra in the mid-IR that are dominated by water vapour emission (see Section 5.3). Since the very low resolution is not optimal for deriving PWV values, these spectra can be used to test the limits of the method. The presence of spectra that were taken at similar time but for different wavelength regimes allows some interesting comparisons.

We evaluate `molecfit` discussing a series of test runs for the sample of spectra defined in Table 3. For all sample spectra, the code was started for a set of input parameters similar to those given in Section 4.6. By default, the spectra were fitted with LBLRTM and polynomials of degree 3 for the continuum and wavelength correction. Due to a challenging continuum shape, a continuum polynomial of degree 5 was chosen for spectrum V2. For X1, only a constant wavelength shift was allowed because of the very narrow fit ranges compared to the full spectrum (see Table 3). The width and shape of the instrumental function was partly constrained. While the width of the Gaussian was unconstrained, the Lorentzian was fixed and the width was set to 0.5 pixels to avoid a degenerated fit. Since this setting is somewhat arbitrary, the derived PWV values could be

Table 4: PWV results for the test sample in Table 3

Label	Fit molecules[a]	t_{fit} [min][b]	H$_2$O column [mm] ESO[c]	`molecfit`	PWV$_{fit}$/ PWV$_{init}$[d]	rel. RMS [%]
C1	H$_2$O, O$_3$, CH$_4$	1.2	0.89	0.99	0.885	2.0
V1	H$_2$O, CO$_2$, O$_3$	17.3	2.25	2.55	1.291	5.7
V2	H$_2$O, CO$_2$, O$_3$	16.8	2.25	1.87	0.982	9.0
X1	H$_2$O, O$_2$	0.6	0.96	0.92	0.605	4.1

[a] VISIR: CO$_2$ and O$_3$ calculated but not fitted; fixed RELCOL: CO$_2$: 1.06 (390 ppmv; global reference value for 2011 [15]), O$_3$: 1.08 (258 Dobson units; Patat et al. [2011]); X-Shooter: O$_2$ calculated but not fitted (RELCOL = 1)

[b] tested on a Core2Quad Q9550@2.83GHz, 8GB RAM, Fedora 16 (64 bit)

[c] CRIRES: taken from fit with the IDL `molecgui` code (A. Smette, ESO); VISIR: taken from the ESO PWV monitor (VISIR but different wavelength regime; closest measurement: $\Delta t \leq 41$ min); X-Shooter: taken from ESO PWV monitor (same spectrum, but measured in NIR regime by means of a curve-of-growth analysis).

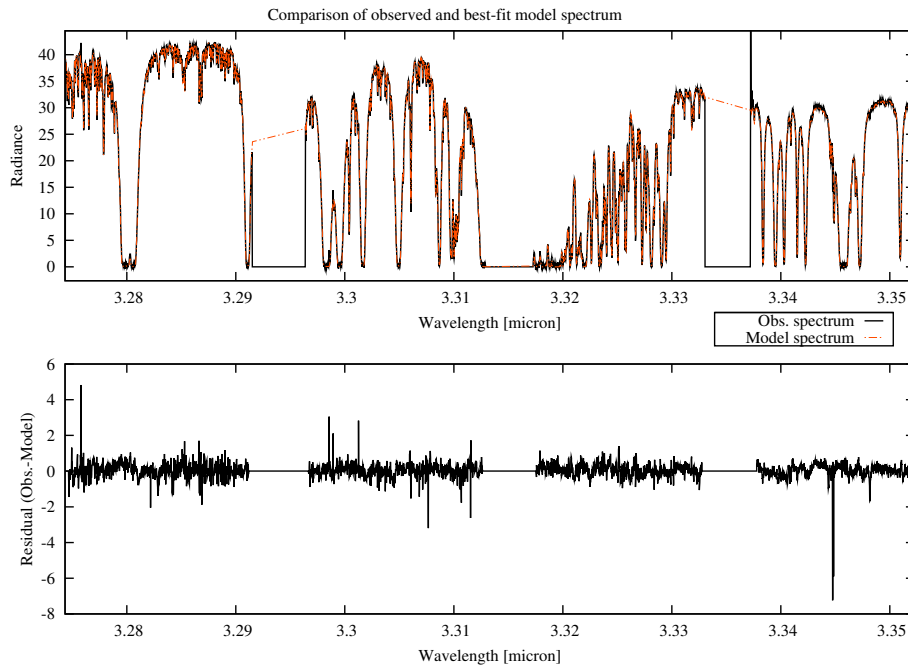[d] ratio of the PWV derived from best-fit and initial atmospheric model.

Figure 10: Comparison of the CRIRES spectrum C1 (black) and the best-fit model (red). Wavelengths not covered by the four CRIRES chips are indicated by zero radiance. The lower panel shows the difference of both spectra. Wavelengths excluded from the fit are indicated by a residual of zero.
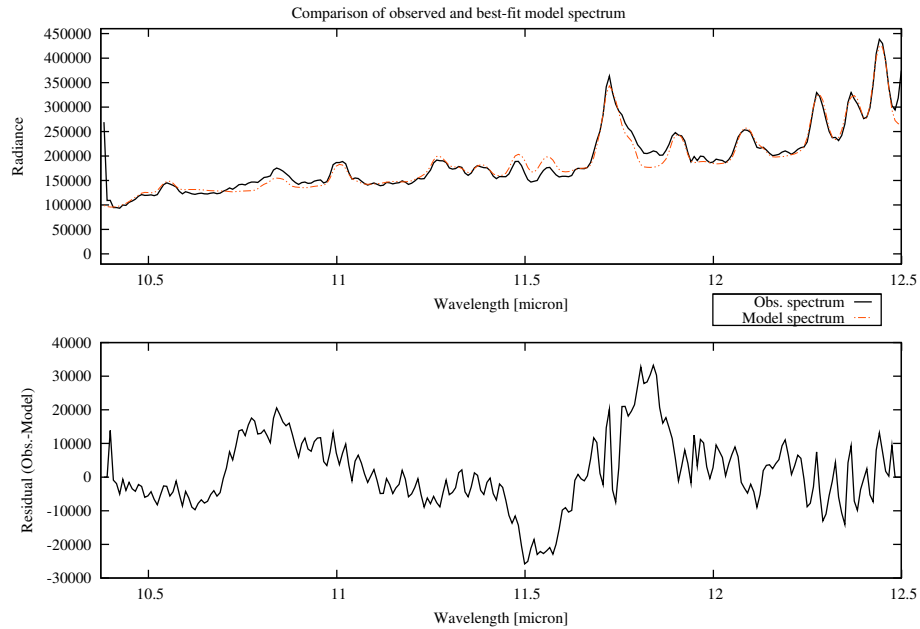
Figure 11: Comparison of VISIR spectrum V1 (black) and the best-fit model (red). The lower panel shows the difference of both spectra.
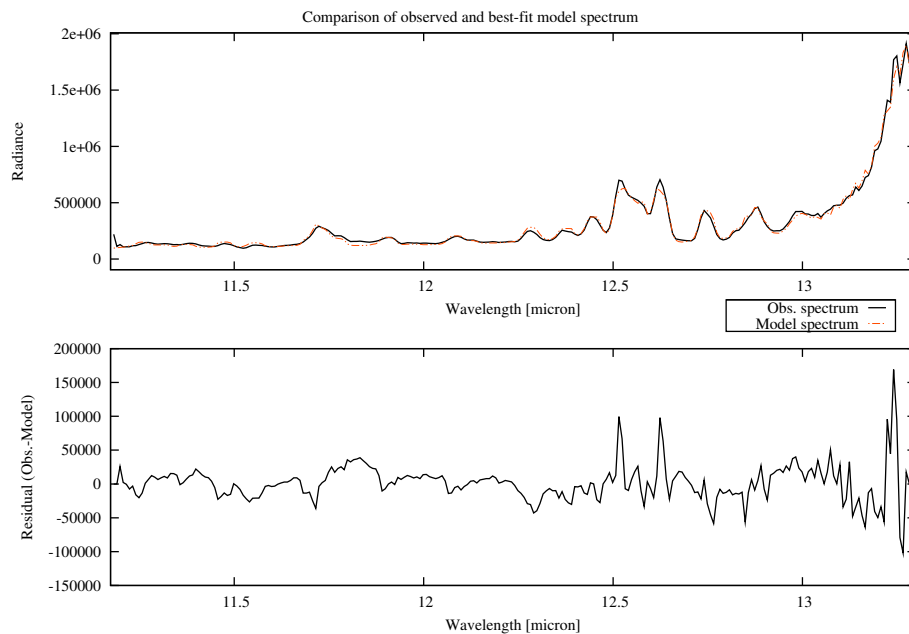

Figure 12: Comparison of VISIR spectrum V2 (black) and the best-fit model (red). The lower panel shows the difference of both spectra.
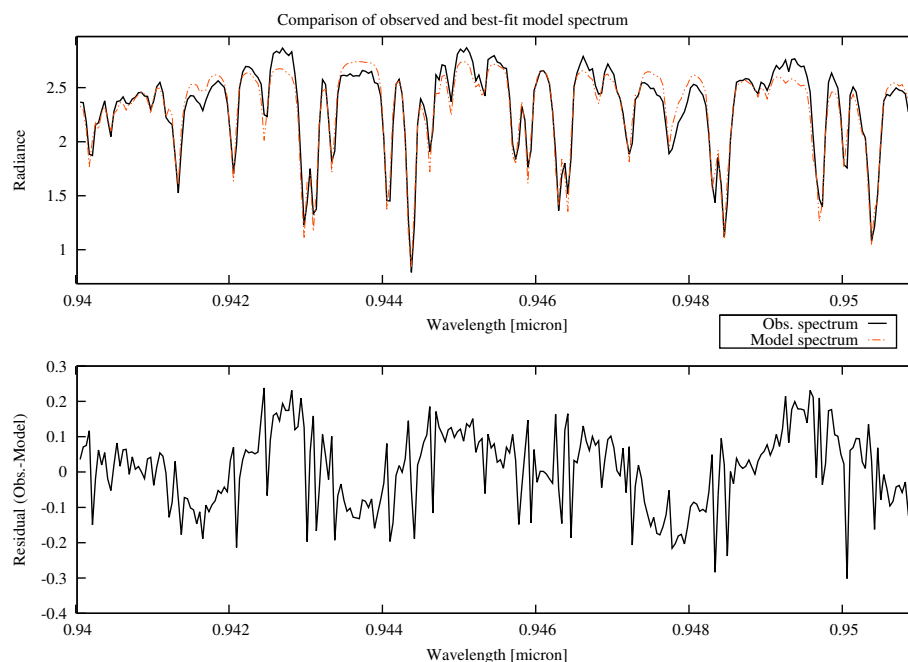
Figure 13: Comparison of X-Shooter spectrum X1 (black) and the best-fit model (red) for the fitted water vapour band. The lower panel shows the difference of both spectra.

offset compared to the real ones. In particular, this could be true for X-Shooter spectra, for which a significant Lorentzian component could not be detected, so far. Tests indicate uncertainties in the order of 0.1 mm. Due to the relatively small number of pixels in the fit range, the Gaussian/Lorentzian kernel radius of the X-Shooter and VISIR spectra was reduced from 300 FWHM for CRIRES to 30 FWHM. A boxcar was only fitted for the VISIR sky radiance spectra. Moreover, the VISIR fits were carried out with an initial emissivity of 0.1. For the fitting, the wavelength ranges provided in Table 3 were used, except for a few pixels at the edges (20/40 pixels for CRIRES, 3 pixels for VISIR). The selection of the fit molecules depended on the fit wavelength range. 1 to 3 molecules of H_2O, CO_2, O_3, and CH_4 were considered for a fit (see Section 5.3 for an explanation). Since the fit ranges of the VISIR spectra were optimised for water-related features (see Table 3), the abundances of the minor contributions CO_2 and O_3 were fixed and set to reasonable values. The selected molecules and the fit results are summarised in Table 4. Figures 10 to 13 show examples for comparisons of observed input spectra and best-fit models. The plots were created by `molecfit`. It should be noted that the setups were not tuned to provide the best possible fits, which might explain at least some of the fitting discrepancies.

As indicated by Table 4, the six test spectra are well fitted. The RMS relative to the mean flux ranges from 2% for the CRIRES spectrum to 9% for the redder VISIR spectrum V2. The strongest residual in Figure 10 is caused by a stellar line, which was not masked due to its negligible contribution to the global fit. In Figure 12, the strong emission at the red end of the spectrum is caused by CO_2. For this reason, spectrum V2 is less well suited for water vapour fitting than the spectrum V1 at shorter wavelengths. The fitting times are very different for the four setups. They depend on the width of the wavelength range and number of lines to be calculated by LBLRTM. For this reason, the fit of the X-Shooter spectrum just took 36 s. Note that the long code run times for the VISIR spectra could be significantly reduced if the basic (now outdated) HITRAN 2008 line list was used instead of the more complex `aer` line list (see Section 5.2.2 and Appendix 7.2.3). The derived PWV values

range from 0.92 to 2.55 mm. These values can be compared to those derived with different approaches and/or different spectra. Table 4 shows good agreement of the `molecfit` results and data taken from the ESO PWV monitor [14]. Typical deviations are in the order of 0.1 mm for the high resolution data, which agrees well with the reported uncertainties of the PWV monitor fits. Test data fits with different parameter sets suggest PWV errors in the same range for `molecfit`. For the VISIR data, the deviations are significantly higher. The direct comparison of the results for the two VISIR spectra confirms that these setups are not well suited for water vapour fits. On the other hand, the RMS values are convincing, which suggests that the fits of the low resolution VISIR spectra suffer from degeneracies. PWV measurements of distinctly higher quality can be expected for data like the CRIRES spectrum with many deep and narrow absorptions.

A reliable PWV measurement does not imply that all fit parameters are trustworthy. In particular, abundances of molecules without sufficiently strong spectral features and parameters for the adaptation of the model spectrum to the observed spectrum such as continuum, wavelength, and resolution parameters (see Section 5.5) are affected if the fitting problem is degenerated. An example is the convolution kernel, where the relative contributions of boxcar, Gaussian, and Lorentzian are difficult to determine. Another example is the telescope emissivity (see Section 5.4), which is a fit parameter for sky radiance spectra. Values of 0.079 (V1) and 0.039 (V2) for the VISIR spectra show a significantly higher uncertainty than the PWV measurements. This implies that the H_2O abundance is a relatively robust fit parameter and that there are parameters which require significantly higher quality data.

Table 4 also includes the ratio of the best-fit PWV and the PWV derived from the initial parameters. Since the initial H_2O profile is derived from an atmospheric standard profile, a suitable meteorological GDAS model profile, and ground-based meteorological data from the EMM (see Section 5.1), the derived PWV ratios can be used to estimate the quality of this profile-composition procedure, provided that the radiative transfer code produce realistic molecular spectra. A wide range of ratios between 0.605 and 1.291 was found, which indicates that the input profiles can lead to a significantly under- or overestimated atmospheric water content. This result shows the importance of the fitting procedure for obtaining realistic PWV values. The effect of the input water vapour profile on the fit quality is extensively discussed in the SM-03 Science Report [SM-03-SR].

6.2 Telluric absorption correction

Apart from deriving abundances of greenhouse gases, `molecfit` is designed to be used for the correction of telluric absorption features. For this purpose, the routine `calctrans` provides the best-fit model transmission (see Section 2.2). It also corrects the fitted spectrum.

`calctrans_lblrtm` and `calctrans_convolution` provide the best-fit model transmission and perform the convolution of the spectrum separately.

For correcting other spectra taken under similar atmospheric conditions and airmasses, the executable `corrfilelist` can be used. In the following, we briefly discuss telluric absorption correction for CRIRES (Section 6.2.1) and X-Shooter spectra (Section 6.2.2).

6.2.1 CRIRES

In order to evaluate the performance of the code in terms of telluric feature correction, we have studied two pairs of CRIRES observations in the wavelength range from 3.28 to 3.36 μm consisting of a telluric standard

Table 5: Description of data and results for the CRIRES telluric absorption correction test

Pair	Target[a]	Time[b] [UT]	Airmass	PWV [mm]	rel. RMS[c] [%]	Model RMS[d] [%]
1	T	2:49	1.569	0.83	3.6	0.9
1	S	3:08	1.546	0.90	2.0	
2	T	3:58	1.472	0.86	3.4	1.0
2	S	4:19	1.475	0.92	2.1	

[a] T = telluric standard star, S = science object

[b] All spectra were taken on July 30th, 2009.

[c] deviation of best-fit model from observed spectrum; relative to mean flux

[d] deviation of model transmission curve for telluric star from corresponding model for science spectrum; absolute value
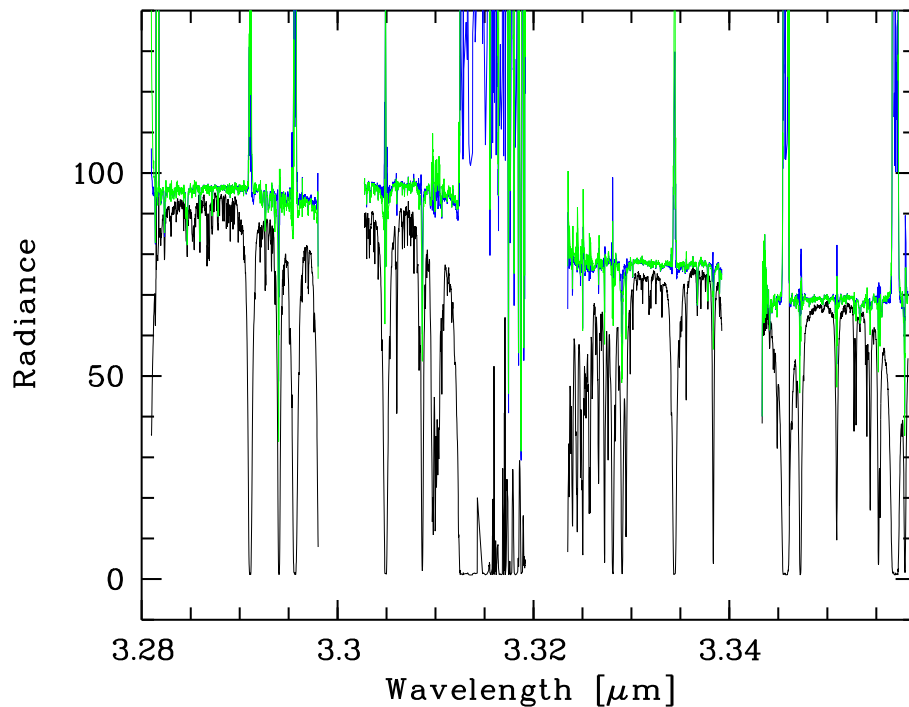

Figure 14: Correction of telluric absorption features in science spectrum 1 (black). The blue and the green curves display the corrected spectra obtained by transmission models fitted to the science and the telluric standard star spectra. No data is shown for wavelengths suffering from very high atmospheric opacity (radiance lower than 1 unit).
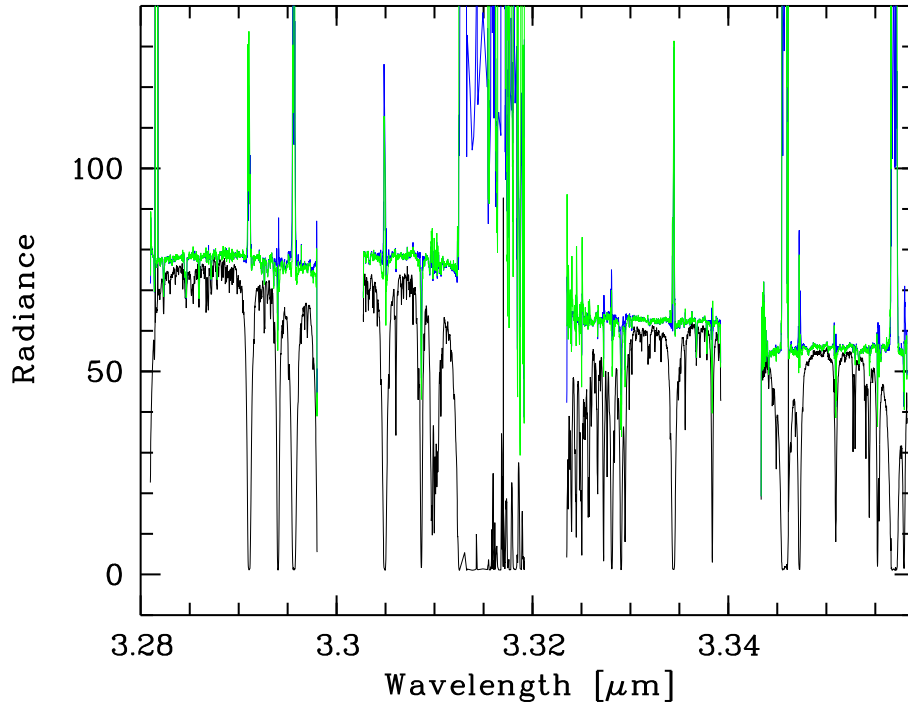
Figure 15: Correction of telluric absorption features in science spectrum 2. For details see Figure 14.

star spectrum and a spectrum of a science target. The spectra of a pair were taken at almost the same time and airmass (see Table 5). We run the code for all spectra to obtain model transmission curves. The input parameters were set as given in Section 4.6. The resulting PWV values provided by Table 5 confirm that the spectra were taken at similar atmospheric conditions. Moreover, these values suggest that the fits are reliable, which is also indicated by RMS values relative to the mean flux between 2.0 and 3.6% (cf. Table 4).

The best-fit transmission curves as obtained for the standard stars and the science targets were used to correct the spectra of the science targets. The resulting spectra are displayed in Figures 14 and 15. Except for wavelengths suffering from very low transmission and strong transmission gradients, the corrected spectra show only weak line residuals. In general, the correction appears to be satisfying for transmission fractions greater than 70 to 80%. The large errors for the centres of optically thick lines can be explained by the division by very small numbers and small errors by statistical noise and in the zero flux line of the observed spectra. Small errors in the wavelength grid and line profile fit cause residuals in the wings of the lines. Interestingly enough, the correction of a science spectrum by a fit of this spectrum itself is not worse than the result based on the standard star transmission curve. The similarity in the correction functions is also indicated by the very small RMS values of about 1% given in Table 5. Hence, the test data set could be corrected without the need of telluric standard star observations. However, it is not clear whether this is also true for spectra taken at a different time, other science targets, and other instrumental setups. Moreover, a systematic comparison with the results of classical methods of telluric absorption correction is missing. This is out of the scope of this document and will be discussed elsewhere.

Table 6: Fit ranges for the example X-Shooter NIR-arm spectrum

ID	Range [μm]	Molecule
1	1.12 – 1.13	H_2O
2	1.47 – 1.48	H_2O
3	1.80 – 1.81	H_2O
4	2.06 – 2.07	CO_2
5	2.35 – 2.36	CH_4

6.2.2 X-Shooter

For X-Shooter, an investigation of a large data set was carried out, which is summarised in the SM-03 Science Report [SM-03-SR]. In this document, we focus on the illustration of the telluric absorption correction for an example NIR-arm X-Shooter spectrum. The flux-calibrated 1D spectrum was obtained by the ESO X-Shooter standard pipeline (see Modigliani et al. [2010]). It was provided to `molecfit` as FITS image. See section 4.8 for more information about the Molecfit input spectrum format.

For the fitting, we used a similar parameter setup as described in Section 6.1 for the X-Shooter VIS-arm spectrum. However, there are some important differences. Since the telluric feature correction has to be carried out for the entire spectrum ranging from 0.99 to 2.48 μm, we have defined five narrow fit ranges in different parts of the spectrum, which are listed in Table 6. Each range of 10 nm in size comprises 100 pixels. The low coverage of the spectrum by fit ranges (about 3% of the full spectrum) is necessary to obtain `molecfit` results within a few minutes. Tests indicate that this limitation does not significantly deteriorate the correction compared with runs using more windows. The fit ranges cover lines from H_2O, CO_2, and CH_4 (see Table 6). However, we only fit the strongly varying water vapour lines. For the other molecules and O_2 and CO, which also have lines in the near-IR regime, we use the abundances from the atmospheric standard profile (see Section 5.1.1). In the case of CO_2, we used a RELCOL value of 1.06 instead of 1 (see Section 4.7) in order to take into account the global increase in CO_2 during the last decade. The use of fixed molecular column densities for all molecules but H_2O makes the fit more robust and faster. The low variability of the fixed species does not significantly affect the quality of the telluric absorption correction (see below). Another measure to stabilise the fit is the restriction to wavelength shifts, *i.e.* the neglection of higher order Chebyshev polynomials for the wavelength correction (see Section 5.5.2). Moreover, the continuum correction was limited to wavelength-independent factors for each fit range. Since the fit ranges are very narrow, it is a good assumption that the variation of the stellar continuum (strong stellar lines are not present) is significantly lower than the variation of the telluric absorption. As general start value 1×10^{-13} was chosen to avoid changes of the continuum parameter by many orders of magnitude in the course of the fitting procedure. Finally, X-Shooter spectra show a nearly linear increase of the line widths with increasing wavelength. We considered this correlation for the convolution kernel calculation by using the VARKERN = 1 option (see Section 4.7).

Figures 16 and 17 show the results of the fitting procedure and the telluric absorption correction. The figures were created by `molecfit` and `calctrans` respectively. Figure 16 illustrates the position of the fit ranges and shows the fit residuals. The relative RMS amounts to relatively low 4.2%. Further results of the fit are an intermediate PWV value of 2.4 mm, a moderate average wavelength shift of 0.35 pixels of the observed spectrum towards shorter wavelengths, and a best-fitting Gaussian kernel FWHM of 1.2 pixels for a fixed Lorentzian of 0.5 pixels. As Figure 17 demonstrates, the fit parameters derived from the five narrow windows are also suitable
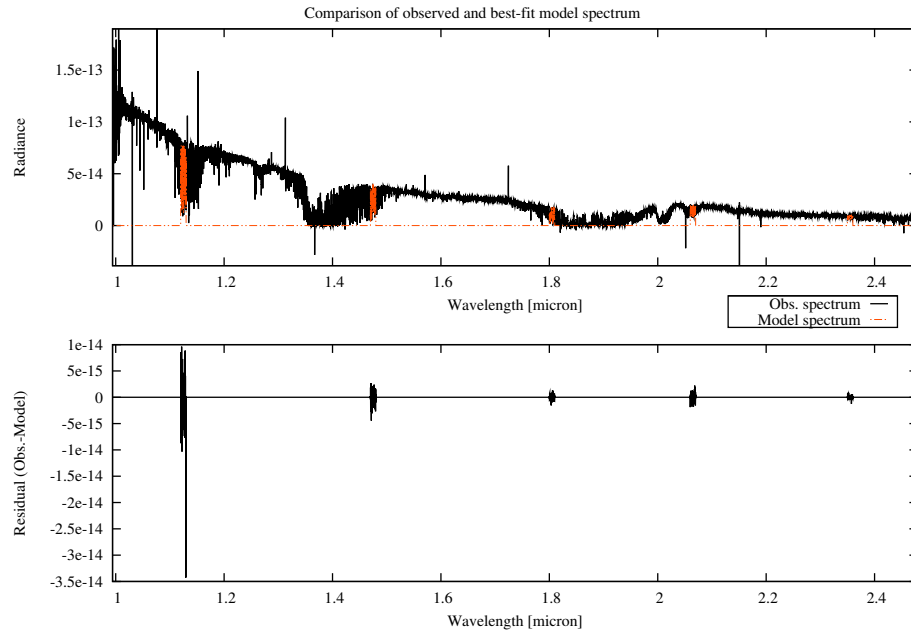
Figure 16: Comparison of a NIR-arm X-Shooter spectrum of a telluric standard star (black) and the best-fit model (red) for five fit ranges of 10 nm in size. The lower panel shows the difference of both spectra. Wavelengths excluded from the fit are indicated by zero transmission in the upper panel and a residual of zero in the lower panel.
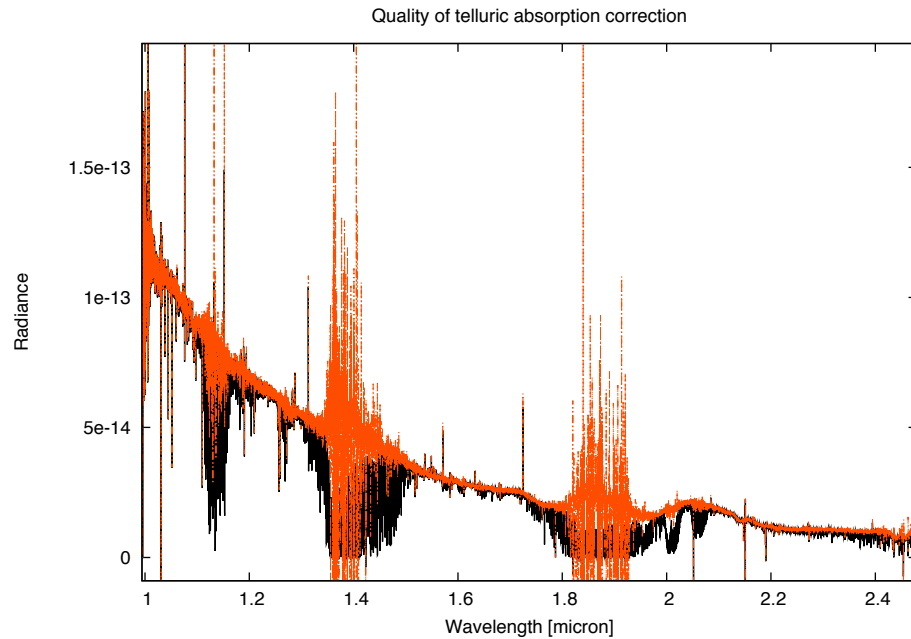

Figure 17: Comparison of a NIR-arm X-Shooter spectrum of a telluric standard star (black) and the same spectrum corrected for telluric absorption based on the fit shown in Figure 16 (red).

to correct the full spectrum for telluric absorption. Down to transmission values of less than 0.5, the residuals are relatively small in general. This is particularly true for the CO_2 bands at about $2\,\mu$m. Consequently, the assumed fixed CO_2 concentration has turned out to be correct. For very low transmissions close to 0, the residuals are significantly increasing. However, this effect is expected, since the correction in these ranges corresponds to the division of two very small numbers, which are affected by strong relative statistical and systematical errors. For this reason, the results for the cores of the H_2O bands between the *J*, *H*, and *K* bands have to be taken with care. It is expected that the near-IR spectrum of a hot telluric standard star is monotonic decreasing. However, the corrected spectrum shows bumps and troughs in regions of low atmospheric transmission. This is most probably caused by a poor flux calibration, since the removal of the bumps related to the H_2O bands would require an unrealistic water vapour reduction of about 40%. For this reason, a simple interpolation between regions of high transmission must fail in the discussed near-IR wavelength ranges. For improving the pipeline performance in future, the `molecfit` results are very promising, since they show that the wavelength regions with reliable continuum can be extended significantly.

6.3 Tips and tricks

In the following, we provide a summary of rules that should be taken into account for a successful application of `molecfit`:

- Pixels with possible defects that could affect the fit quality can be excluded from the fit in two ways. First, the critical pixels can be listed in an ASCII or FITS file that is provided by the parameter PRANGE_EXCLUDE (see Section 4.7). Second, pixels can be skipped by adding a mask column to an input ASCII or FITS table, or adding an mask extension to a FITS image. In both cases, the name of the column/extension has to be given by the fourth COLUMNS parameter.

- The resulting best-fit parameters of `molecfit` are written to a `.res` file (see Section 4.9.2). In the case of a complex fit, the more reliable fit parameters could be taken from this file and used as (fixed) input for another iteration of the fitting procedure.

- Changing the fit parameters FTOL and XTOL (see Section 4.7) can significantly affect the code run time and the quality of the fit. In the case of unsatisfying fit results, it may be an option to change the default values. However, the effect is often unpredictable, since more relaxed convergence criteria can lead to worse as well as better fit quality.

- To achieve an optimal performance of the code, one should fit only those molecules that significantly contribute to the wavelength range of the fit. We suggest to base the selection of relevant molecules on the information given in Section 5.3.

- In principle, PWV values can be measured by means of all kinds of spectra of bright standard stars which show telluric lines in absorption and atmospheric emission spectra in the thermal IR. However, a good fit requires significant H_2O features. This criterion cannot be fulfilled if the water lines are very weak as in the optical. Moreover, too low resolution can smooth out the crucial lines, which can make the fit very unstable or even impossible (see Section 6.1).

- For `molecfit` applications aiming at the derivation of the atmospheric water vapour content or the telluric absorption correction of astronomical spectra, it is often sufficient to set the abundances of other

molecules to a fixed value. For the more frequent molecules in the atmosphere (see Section 5.3), the column density from the input standard profile, *i.e.* RELCOL = 1, is usually relatively close to the true value. It has to be taken into account that the MIPAS standard profiles were created in 2001 (see Section 5.1.1), which causes deviations for greenhouse gases that indicate a significant long-term increase in atmospheric abundance. For example, the global CO_2 concentration increased by about 6% in one decade, which suggests RELCOL = 1.06 (see Section 6.1).

- The user can explicitly set the initial value of the constant term of the polynomials for the continuum correction and the wavelength solution. The higher-order coefficients are automatically set to reasonable start values if required (see Section 4.7). Setting WLC_CONST is only recommended if a wavelength shift towards a certain direction is expected. The CONT_CONST is more critical. If the continuum level of the input spectrum strongly deviates from 1, even after setting the scaling factor FLUX_UNIT (see Section 4.7), it is prudent to adapt this term.

- For a correct wavelength fit especially at high resolution, it is important to have a correct setting of the VAC_AIR parameter. The wavelength system depends on the instrument and the wavelength calibration approach by the data reduction pipeline. IR data (CRIRES, VISIR) tend to be provided in vacuum wavelengths, whereas data at shorter wavelengths tend to be provided in air wavelengths (X-Shooter). If the user does not know this necessary input, it can be easily derived by running `molecfit` with both VAC_AIR options.

- For the instrumental profile created by the convolution of boxcar, Gaussian, and Lorentzian kernels (see Section 5.5), a derivation from the fit can be difficult. If the χ^2 degeneration by the kernel parameters cause a bad fit, we recommend to modify the initial values of these parameters or to fix the width of a kernel element. Some knowledge on the true functional form of the instrumental profile can be quite helpful. For a first test, the width of the Lorentzian might be fixed. For CRIRES, a value of 0.75 plus a large kernel size could be reasonable (A. Smette 2012, priv. comm.). For our tests, we used 0.5 pixels for all instruments (see Sections 6.1 and 6.2), although X-Shooter spectra do not indicate a significant contribution of a Lorentzian. The boxcar kernel can probably be neglected if the slit width does not have a significant impact on the line FWHM. If the user has access to a suitable line profile kernel, it can be imported via the parameter KERNEL_FILE (see Section 5.5.3). The three-component profile fitting is switched off in this case.

- The run time of the radiative transfer code LBLRTM depends on the width of the fitted wavelength ranges. For a better performance, it is, therefore, recommended to use fit ranges as narrow as possible. It also makes the polynomial continuum fit more reliable. For *e.g.* X-Shooter, several representative ranges over the entire wavelength range covering lines of all critical molecules could be defined. A suitable bin size is 10 nm. CRIRES spectra are sufficiently narrow that the full wavelength range can be used for the fitting procedure.

- For spectra covering a wide wavelength range and narrow fit ranges, the degree of the polynomial for the wavelength solution WLC_N should be set to 0 to avoid unpredictable wavelength corrections outside the fit ranges if a telluric absorption correction is required. For X-Shooter, it is also strongly recommended to set VARKERN = 1 in order to consider the change of the instrumental profile with wavelength (see Section 4.7).

- For telluric absorption correction in the thermal IR, it could be advantageous to use the strong sky emission extracted as 1D spectrum as input for `molecfit`. `calctrans` will then calculate the transmission

function belonging to the best-fit sky radiance spectrum. Finally, the object spectrum can be corrected for telluric absorption via `corrfilelist`.

7 Appendix

7.1 Expert Fitting

The expert fitting mode allows one to fit individual chips and ranges and provides access to the coefficients of the polynomials for the continuum and wavelength correction. Moreover, parameter files with the best-fit values are written which can be used for another iteration of `molecfit`.

The example parameter file `molecfit_crires_expert.par` can be used to test the expert mode. The following rows differ from the standard parameter file:

```
# Fit flag for fitting ranges with respect to continuum correction
# (default: fit_cont)
fit_range1:
fit_range2:
fit_range3:
fit_range4:
```

The parameters `fit_range[range]` allow one to enter fit flags for the different fitting ranges with respect to the continuum correction. These parameters overrule the `fit_cont` parameter if set. Note that a change of the fitting ranges can change the numbering. User-defined ranges that extend over more than one chip will be split and numbered differently.

```
# Fit flag for chips with respect to wavelength correction
# (default: fit_wlc)
fit_chip1:
fit_chip2:
fit_chip3:
fit_chip4:
```

The parameters `fit_chip[chip]` are very similar to `fit_range[range]`. They allow for independent fits of the wavelength correction coefficients for each chip. They overrule `fit_wlc` if set.

```
# Range-specific coefficients for continuum correction (list separated by
# spaces; default: cont_const 0. 0. ...)
cont_range1:
cont_range2:
cont_range3:
cont_range4:
```

The parameters `cont_range[range]` contain the coefficients for the continuum correction for each fitting range. They have to be listed separated by spaces. The number of coefficients per parameter is not restricted. They can deviate for different ranges. However, note that the parameter `cont_n` sets a limit to the number of coefficients that are considered for the fit (`cont_n + 1`). The set coefficients overrule the standard values, which are `cont_const, 0., 0.,` and so on.

```
# Chip-specific coefficients for wavelength correction (list separated by
# spaces; default: wlc_const 1. 0. 0. ...)
wlc_chip1:
wlc_chip2:
wlc_chip3:
wlc_chip4:
```

The parameters `wlc_chip[range]` contain the coefficients for the wavelength correction for each chip. The maximum number of coefficients taken for the fit is `wlc_n + 1`. The set coefficients overrule the standard values, which are `wlc_const, 1., 0., 0.`, and so on.

Each provided fit flag or coefficient for the continuum and wavelength correction will be considered as start value for the fitting procedure. If no value is provided, `molecfit` runs in the standard mode, i.e. the default values are taken. There is a difference in the performance if all colons are removed from the parameter labels. In this case, the traditional `molecfit` mode runs as described in the user manual. However, a single colon will already activate the expert mode. As a consequence, an additional output file will be written to the output directory. It has the ending `_fit.rpar`. It equals the file for the `molecfit` input parameters labelled `_fit.par` except for the fit parameters, where the best-fit value is written. Apart from the `cont_range` and `wlc_chip` parameters, these are `relcol`, `telback`, `relres_box`, `res_gauss`, and `res_lorentz`. The advantage of this file is that it can directly be used as input for another run of `molecfit`. By modifying the fit flags (and parameter values) in this file before calling `molecfit`, the fit results can be improved by focusing on subsets of the fitting parameters, which makes the procedure more robust.

The expert mode offers full access to the `molecfit` fitting parameters and allows for more flexible iterative fitting. On the other hand, it increases the complexity of the fitting procedure and there is a chance of erroneous results if the mask files are changed and the range and chip-related parameters are not carefully adapted. For this reason, it has not been implemented as standard mode for `molecfit`.

7.2 Maintenance

7.2.1 Introduction

The package `molecfit` relies on some external codes and data, which might undergo a change due to development even after the SM-02 project ends. Although, it is not possible to foresee all possible modifications, some hints can be given to use `molecfit` even with updates from the external sources at a later stage. This particularly applies to the radiative transfer code LNFL/LBLRTM, and the GDAS data. In this appendix, we summarise the issues to be taken into account for a successful usage of `molecfit` with later versions of LNFL/LBLRTM, the line database, and the GDAS data.

7.2.2 Radiative transfer code LNFL/LBLRTM

The radiative transfer code package LNFL/LBLRTM is developed by AER[2] and can directly be obtained from there. `molecfit` is delivered with LNFL v3.1, LBLRTM v12.8, and the AER line database v3.6.

In principle, `molecfit` should also work with later versions as its principle usage has been unchanged for several years. This means that `molecfit` is expected to provide all functionality as long as it is retained. In

particular, the principle usage of the TAPE<xx> files [3], and the format of the LBLRTM input file `TAPE5` MUST remain unchanged. However, as LBLRTM is widely used in atmospheric research, we do not expect major changes in the near future. In the course of the former DR-, and later SM-projects, we used several versions without problems.

What might change is the installation procedure of the LNFL/LBLRTM codes. If a future release of `molecfit` is provided with a newer version of the radiative transfer code, the install script might need to be modified accordingly. For that purpose, the developers from AER provide a README file in the tree of the sources. In order to maximise flexibility, `molecfit` only searches for executables labelled `lnfl` and `lblrtm` in the `<INST_DIR>/bin/` directory. Therefore, a soft link to the actual binaries of LNFL and LBLRTM has to be provided by the installation script. To facilitate the modification in the script, the corresponding section is marked there.

We recommend the following procedure for the update of `molecfit` to later versions of LNFL/LBLRTM:

- Updates of the LNFL/LBLRTM are usually once or twice per year (see Section "What's new" at [2]). Therefore, a quarterly check for new versions is sufficient.

- Download the latest version of LNFL/LBLRTM from [2]

- Make sure that the principle usage has not changed. Check the documents in the `docs` directory in the tree of the LBLRTM sources. There, you should find a FAQ document, the release notes, and a detailed HTML description of the `TAPE5` input file for LBLRTM.

- Read the `README` files whether the installation procedure has changed. These files should be located in the `src/build` directory of LNFL and LBLRTM.

- Check whether there is also a line list update and follow the steps described in Section 7.2.3.

- Adapt the installation script of `molecfit`.

- Evaluate by means of the delivered examples.

7.2.3   AER Line parameter list

Usually the line parameter list delivered with the LNFL/LBLRTM package is updated at the same time as the code. We recommend to use the version which is delivered together with the corresponding code version.

To integrate the new line list, simply unpack the tarball and copy the new line parameter file (usually labelled `aer_v_<version>` to the directory `<INST_DIR>/data/hitran/` of your `molecfit` installation. This can be done also by modifying the install script accordingly. The corresponding section is marked there in a special way. Moreover, it is necessary to provide the new line list name to the `_LINE_DB` parameter in the `lblrtm_setup` file in the `<INST_DIR>/config/` directory.

Note that it has turned out that at long wavelengths relevant for VISIR, `molecfit` can be significantly slower if the recommended `aer` line list is used instead of the basic HITRAN 2008 database (see Section 6.1). Unfortunately, this line list is no more available at the HITRAN website [6], which now provides the 2012 version. The new version has not been tested in terms of performance in `molecfit`. Hence, it cannot be guaranteed that it works at all. If you want to test an original HITRAN line list, you have to put it in the `data/hitran/`

folder and modify _LINE_DB in the `lblrtm_setup` file. In addition, the parameter _LINE_DB_FMT has to be changed from 100 to 160 because of a different file format.

7.2.4 GDAS data

In contrast to the radiative transfer code, which is updated once/twice per year, the GDAS profiles are updated permanently. `molecfit` initially searches in a local tarball for the corresponding profiles (script `<INST_DIR>/bin/get_gdas_profiles.sh`), and tries to download it from the web in the case they are not found locally. The latter is done by the script `<INST_DIR>/bin/extract_grib.sh` with the help of the GRIB software.

However, the incorporated download server[6] does not seem to be very stable and organised in a predictable way. It might happen, that the profiles are not found. Therefore regularly updated data tarballs are available from http://www.eso.org/pipelines/skytools/molecfit. They can be downloaded and placed into the `<INST_DIR>/data/profiles/gdas` folder in case the automatic update mechanism fails.

Additionally `molecfit` contains the script used to produce the updated files for the website. It can also be used to update the local GDAS database. The script downloads the GDAS data archive from a different server[7]. The data there is organised on weekly basis, each stored in a single file labelled `gdas1.<mmm><yy>.w#`, being <mmm> the month (e.g. jan, feb, mar,...) and <yy> the year. The extension "w#" describes the number of the week of this month:

#=1 - days 1-7 of the month
#=2 - days 8-14
#=3 - days 15-21
#=4 - days 22-28
#=5 - days 29 - rest of the month

For example, the file `gdas1.apr07.w3` contains data of the third week of April 2007 (15th to 21st of April 2007). More information on the file structure is given here[8].

These files provide the GDAS profiles of the world wide grid. Therefore, it is necessary to extract the profiles in the same way as described in Section 5.1. This can be achieved with the shell script `update_gdas_db.sh`. It downloads the data on monthly basis, and extracts the profiles as required by `molecfit`, and adds them to the local database. However, this may take a while as the amount of data to be downloaded is large (up to 580 MB per week) and the server is fairly slow.

The shell script invokes a C-programme `extract_gdas_profiles`. Although, both tools are part of the molecfit package, they can be used independently of `molecfit` to provide the possibility to update the GDAS archive on different machines. This might be useful as some disc space is required temporarily. If this is intended, both tools have to be installed in the following way on the target machine:

- create subfolders `bin/` and `data/downloads/paranal/` in the target directory on the target machine.

---

[6]http://nomad3.ncep.noaa.gov/pub/gdas/rotating/
[7]ftp://ftp.arl.noaa.gov/archives/gdas1/
[8]http://www.ready.noaa.gov/gdas1.php

- Copy the shell script `update_gdas_db.sh` and the C source file `extract_gdas_profiles.c` to the `bin/` folder in target directory on the target machine.

- There compile the C programme by invoking

  ```
  gcc extract_gdas_profiles.c -lm -o extract_gdas_profiles
  ```

- Copy the tarball file `gdas_profiles_C-70.4-24.6.tar.gz` delivered with the `molecfit` package to the directory `<targetdir@targetmachine>/data/downloads/paranal/`. This tarball is being updated.

To update the local GDAS database please follow these steps:

- The final release of `molecfit` contains GDAS profiles up to March 2014. So, if this is your first update, invoke the update script with the parameter oct13:

  ```
  cd <INST_DIR>/
  ```

  OR

  ```
  cd <targetdir@targetmachine>
  ```

  (depending on your installation (see above)

- invoke the update script

  ```
  bin/update_gdas_database.sh oct13 P
  ```

  and proceed until the end of the last fully missing month, if necessary. Note: You only can download full months. The file `data/paranal/last_month.txt` shows a history of past updates.

- Copy the resulting tarball
  `data/downloads/paranal/gdas_profiles_C-70.4-24.6.tar.gz` to the folder
  `<INST_DIR>/data/profiles/gdas/` in your `molecfit` installation path.

- Optionally: delete the downloaded files in `data/downloads/gdas/`.

Some dates are missing. Please check the website[9] for more details. There are no GDAS data available before 1st of December 2004.

7.3 License issues

7.3.1 LNFL/LBLRTM

URL: http://rtweb.aer.com/lblrtm_frame.html

---

[9] `http://ready.arl.noaa.gov/archives.php`

LNFL Terms and Conditions

The LNFL software you are downloading has been copyrighted by Atmospheric and Environmental Research, Inc. (AER) and is proprietary to AER. AER grants USER the right to download, install, use and copy this software for scientific and research purposes only. By downloading this software, you are agreeing that:

- If you redistribute any portion of the software, whether in original or modified form, each source code file redistributed will include the AER copyright notice currently present in that file.

- The software or any modified version of the software may not be incorporated into proprietary software or commercial software offered for sale.

- Should you utilize the results of LNFL in your research, we request that you appropriately reference this usage in any associated publications.

- This software is provided as is without any express or implied warranties.

LBLRTM Terms and Conditions

The LBLRTM software you are downloading has been copyrighted by Atmospheric and Environmental Research, Inc. (AER) and is proprietary to AER. AER grants USER the right to download, install, use and copy this software for scientific and research purposes only. By downloading this software, you are agreeing that:

- If you redistribute any portion of the software, whether in original or modified form, each source code file redistributed will include the AER copyright notice currently present in that file.

- The software or any modified version of the software may not be incorporated into proprietary software or commercial software offered for sale.

- Should you utilize the results of LBLRTM in your research, we request that you appropriately reference this usage in any associated publications.

- This software is provided as is without any express or implied warranties.

Principal References

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, Atmospheric radiative transfer modeling: a summary of the AER codes, Short Communication, J. Quant. Spectrosc. Radiat. Transfer, 91, 233-244, 2005.

Clough, S.A., M.J. Iacono, and J.-L. Moncet, Line-by-line calculation of atmospheric fluxes and cooling rates: Application to water vapor. J. Geophys. Res., 97, 15761-15785, 1992.

7.3.2 cmpfit library

URL : http://www.physics.wisc.edu/ craigm/idl/cmpfit.html **URL :** http://cow.physics.wisc.edu/ craigm/idl/idl.html

DISCLAIMER

MPFIT: A MINPACK-1 Least Squares Fitting Library in C

Original public domain version by B. Garbow, K. Hillstrom, J. More' (Argonne National Laboratory, MINPACK project, March 1980) Copyright (1999) University of Chicago (see below)

Tranlation to C Language by S. Moshier (moshier.net) (no restrictions placed on distribution)

Enhancements and packaging by C. Markwardt (comparable to IDL fitting routine MPFIT

see http://cow.physics.wisc.edu/ craigm/idl/idl.html) Copyright (C) 2003, 2004, 2006, 2007 Craig B. Markwardt

This software is provided as is without any warranty whatsoever. Permission to use, copy, modify, and distribute modified or unmodified copies is granted, provided this copyright and disclaimer are included unchanged.

Source code derived from MINPACK must have the following disclaimer text provided.

==
Minpack Copyright Notice (1999) University of Chicago. All rights reserved

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

3. The end-user documentation included with the redistribution, if any, must include the following acknowledgment:

"This product includes software developed by the University of Chicago, as Operator of Argonne National Laboratory.

Alternately, this acknowledgment may appear in the software itself, if and wherever such third-party acknowledgments normally appear.

4. WARRANTY DISCLAIMER. THE SOFTWARE IS SUPPLIED "AS IS" WITHOUT WARRANTY OF ANY KIND. THE COPYRIGHT HOLDER, THE UNITED STATES, THE UNITED STATES DEPARTMENT OF ENERGY, AND THEIR EMPLOYEES: (1) DISCLAIM ANY WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE OR NON-INFRINGEMENT, (2) DO NOT ASSUME ANY LEGAL LIABILITY OR RESPONSIBILITY FOR THE ACCURACY, COMPLETENESS, OR USEFULNESS OF THE SOFTWARE, (3) DO NOT REPRESENT THAT USE OF THE SOFTWARE WOULD NOT INFRINGE PRIVATELY OWNED RIGHTS, (4) DO NOT WARRANT THAT THE SOFTWARE WILL FUNCTION UNINTERRUPTED, THAT IT IS ERROR-FREE OR THAT ANY ERRORS WILL BE CORRECTED.

5. LIMITATION OF LIABILITY. IN NO EVENT WILL THE COPYRIGHT HOLDER, THE UNITED STATES, THE UNITED STATES DEPARTMENT OF ENERGY, OR THEIR EMPLOYEES: BE LIABLE FOR ANY INDIRECT, INCIDENTAL, CONSEQUENTIAL, SPECIAL OR PUNITIVE DAMAGES OF ANY KIND OR NATURE, INCLUDING BUT NOT LIMITED TO LOSS OF PROFITS OR LOSS OF DATA, FOR

ANY REASON WHATSOEVER, WHETHER SUCH LIABILITY IS ASSERTED ON THE BASIS OF CONTRACT, TORT (INCLUDING NEGLIGENCE OR STRICT LIABILITY), OR OTHERWISE, EVEN IF ANY OF SAID PARTIES HAS BEEN WARNED OF THE POSSIBILITY OF SUCH LOSS OR DAMAGES.

7.3.3 HITRAN

URL: http://www.cfa.harvard.edu/hitran/
URL: http://www.cfa.harvard.edu/hitran/Updated/ref-table.pdf

Principal references

Rothman et al., "The HITRAN 2008 molecular spectroscopic database", Journal of Quantitative Spectroscopy and Radiative Transfer, vol. 110, pp. 533-572 (2009)

7.3.4 GDAS

URL: http://ready.arl.noaa.gov/gdas1.php

Air Resources Laboratory - Disclaimer

The information on government servers is in the public domain, unless specifically annotated otherwise, and may be used freely by the public. Before using information obtained from this server special attention should be given to the date and time of the data and products being displayed. This information shall not be modified in content and then presented as official government material. The user shall credit NOAA Air Resources Laboratory in any publications, presentations, or other derivative works that result from the use of this material. NOAA Air Resources Laboratory is not obligated to provide any support, consulting, training, or assistance of any kind with regard to the use of this material.

This material is provided "AS IS" and any express or implied warranties including, but not limited to, the warranties of merchantability and fitness for a particular purpose are disclaimed. In no event shall NOAA Air Resources Laboratory be liable to you or to any third party for any direct, indirect, incidental, consequential, special or exemplary damages or lost profit resulting from any use or misuse of this data.

This server is available 24 hours a day, seven days a week, however this server is not operational and is not maintained 24 hours a day, seven days a week. Timely delivery of data and products from this server through the Internet is not guaranteed.

Any references or links from the ARL web server to any non-Federal Government Web sites do not imply endorsement of any particular product, service, organization, company, information provider, or content. We are not responsible for the contents of any "off-site" web pages referenced from ARL servers.

Acronyms

AER Atmospheric and Environmental Research Inc.

AOPP Atmospheric, Oceanic and Planetary Physics, Oxford University, UK

ARL Air Resources Laboratory

ASCII American Standard Code for Information Interchange

CPL Common Pipeline Library

CRIRES CRyogenic high-resolution InfraRed Echelle Spectrograph

EMM ESO Meteo Monitor

Envisat Environmental Satellite

ESO European Southern Observatory

FITS Flexible Image Transport System

FWHM full width at half-maximum

GDAS Global Data Assimilation System

GRIB GRIdded Binary

GUI Graphical User Interface

HITRAN HIgh-resolution TRANsmission molecular absorption database

HTML Hypertext Markup Language

IDL Interactive Data Language

LBLRTM Line-by-line Radiative Transfer Model

LNFL Line File

MIPAS Michelson Interferometer for Passive Atmospheric Sounding

MIR mid-infrared

NCEP National Centers for Environmental Prediction

NIR near-infrared

NOAA National Oceanic and Atmospheric Administration

PWV precipitable water vapour

RFM Reference Forward Model

RMS root mean square

RPG Radiometer Physics GmbH

VISIR VLT Imager and Spectrometer for mid InfraRed

A   Molecular spectra

Figure 18: *Influence of individual molecules as a function of wavelength:* this figure shows the relative importance of all molecules at a given wavelength, which exceed more than 5% of the total radiance (in red) or transmission (in blue). See text for more detail.
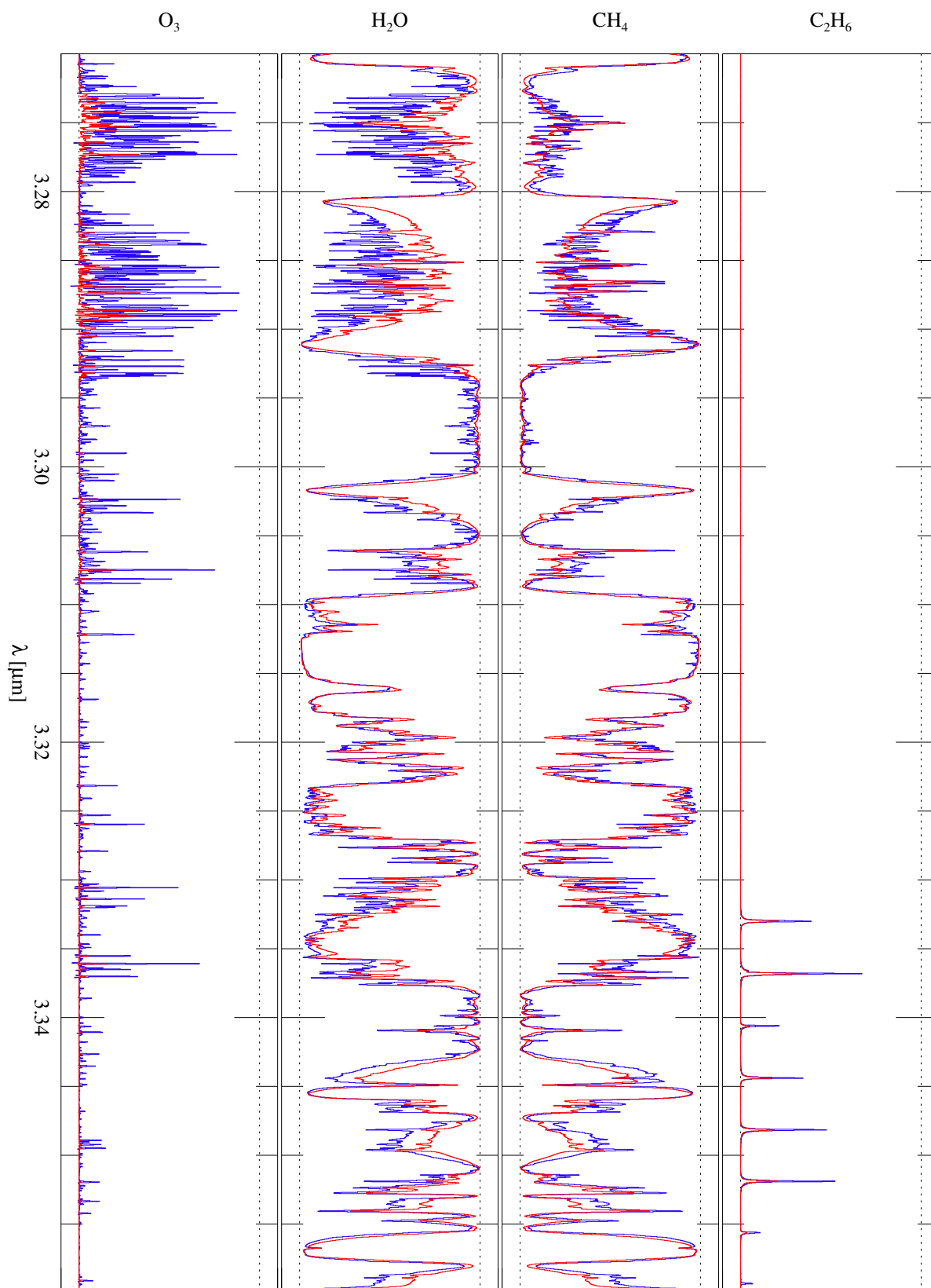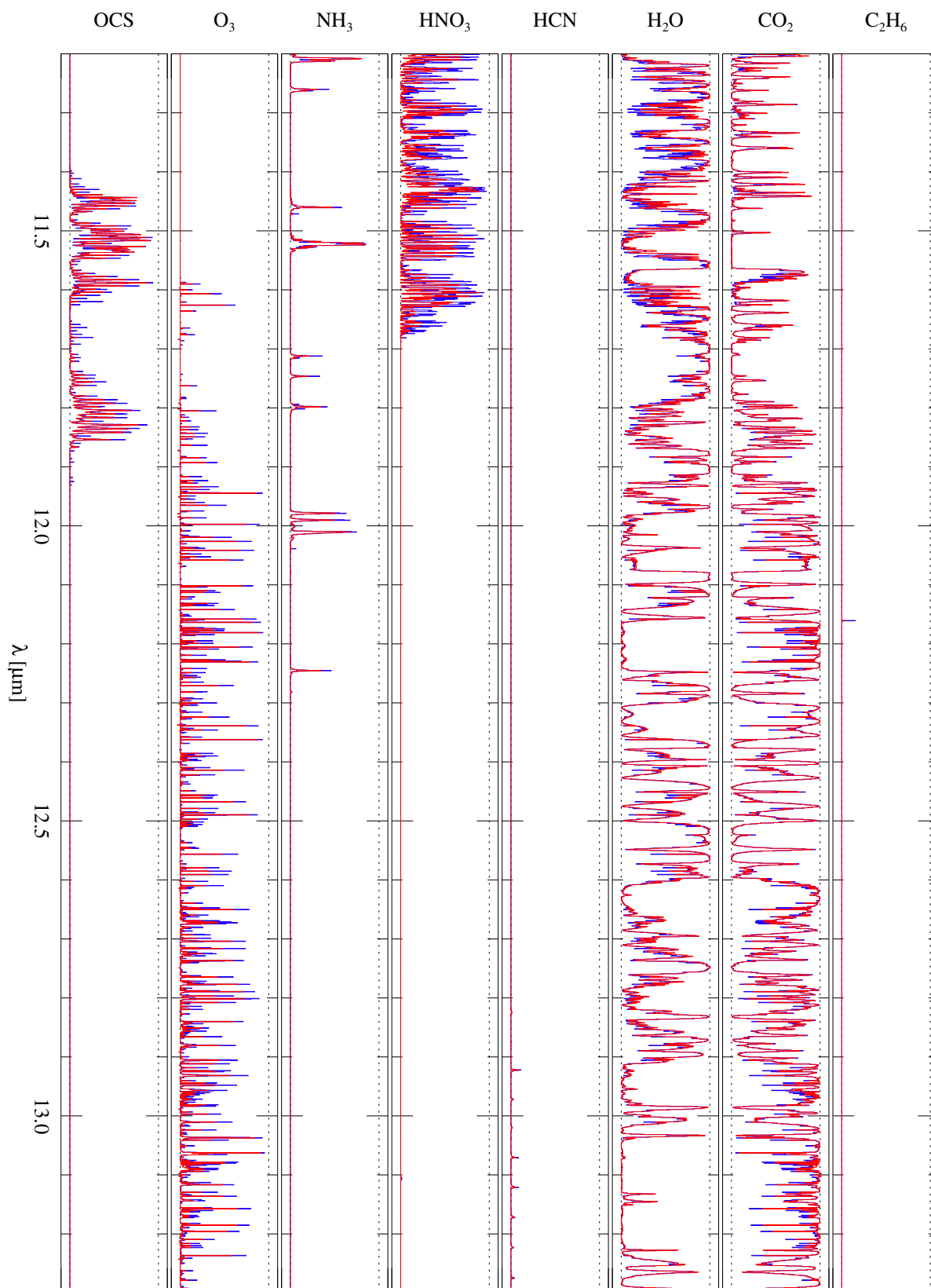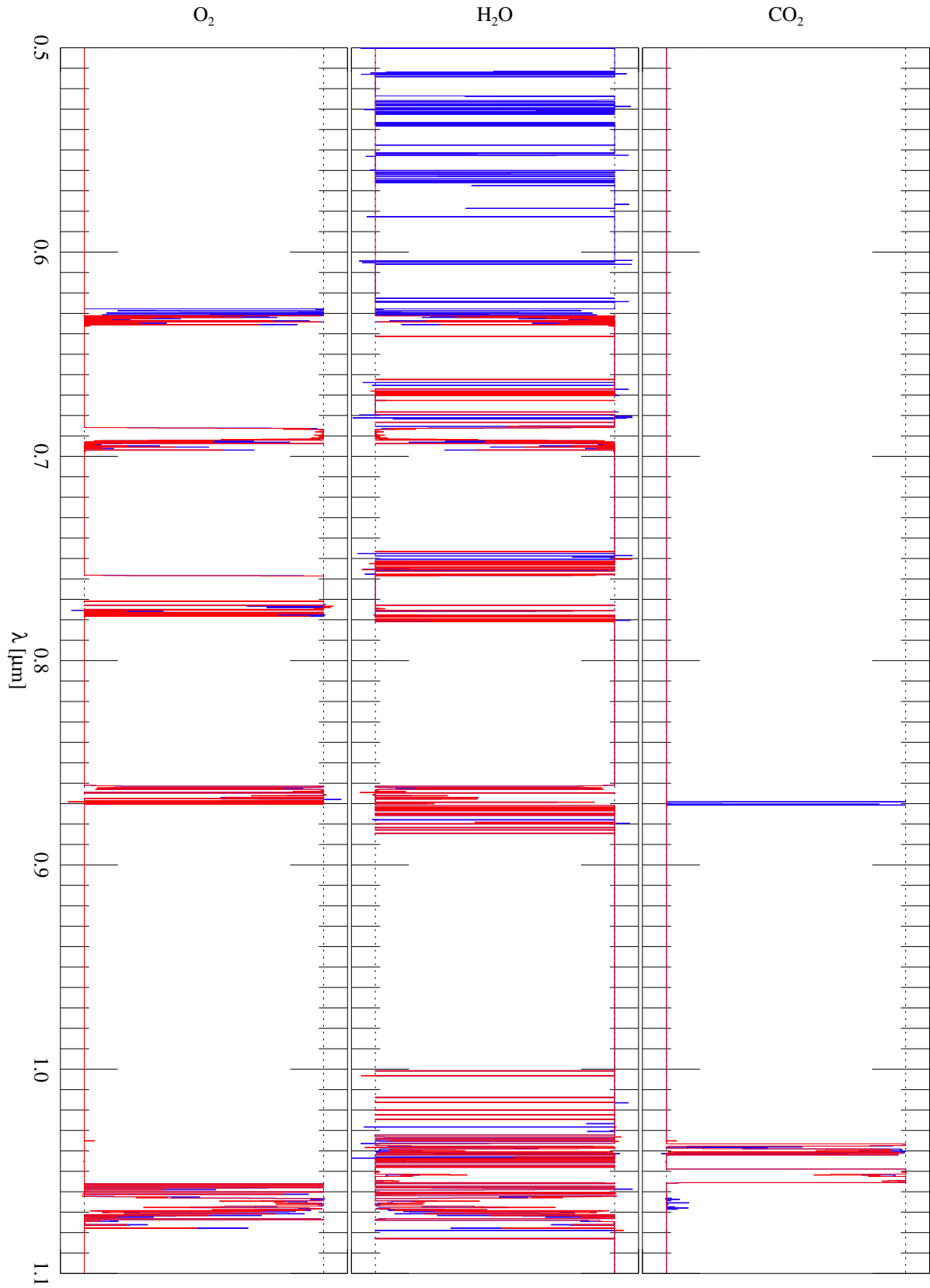
Figure 19: *Influence of individual molecules as a function of wavelength:* same as Figure 18, but for the CRIRES test data wavelength regime. See text for more detail.

Figure 20: *Influence of individual molecules as a function of wavelength:* same as Figure 18, but for the VISIR test data wavelength regime. See text for more detail.

Figure 21: *Influence of individual molecules as a function of wavelength:* same as Figure 18, but for the VISIR test data wavelength regime. See text for more detail.

Figure 22: *Influence of individual molecules as a function of wavelength:* same as Figure 18, but for the X-Shooter VIS-arm test data wavelength regime. See text for more detail.

References

[Anu Dudhia 2009] Anu Dudhia, private communication 2009

[1992] Clough, S.A., Iacono, M.J., & Moncet, J.-L. 1992, J. Geophys. Res., 97, 15761

[2005] Clough, S.A, Shephard, M.W., Mlawer. E.J., et al. 2005, J. Quant. Spectrosc. Radiat. Transfer, 91, 233

[1988] Masuda, K., Takashima, T., & Takayama, Y. 1988, Remote Sens. Environ., 24, 313

[2010] Modigliani. A., Goldoni, P., Royer, F., et al. 2010, Proc. SPIE, 7737, 773728

[1980] Moré, J.J., Garbow, B.S., & Hillstrom, K.E. 1980, User Guide for MINPACK-1, Argonne National Laboratory Report ANL-80-74, Argonne, Ill.

[2005] Niro, F., Jucks, K., & Hartmann, J.-M. 2005, J. Quant. Spectrosc. Radiat. Transfer, 95, 469

[2012] Noll, S., Kausch, W., Barden, M., et al. 2012, A&A, 543, A92

[2011] Patat, F., Moehler, S., O'Brien, K., at al. 2011, A&A, 527, A91

[1997] Wu, X. & Smith, L. 1997, Appl. Opt., 36, 2609

[DR06-UM] DR06 User Manual, VLT-MAN-ESO-19550-5286

[SM-01-UM] SM-01 User Manual, VLT-MAN-ESO-19550-5770

[MF-GUI] MOLECFIT: Graphical User Interface and Tutorial, VLT-MAN-ESO-19550-5928

[SM-03-DS] SM-03 Detailed Specification Document, VLT-SPE-ESO-19550-5769

[SM-03-SR] SM-03 Science Report, VLT-TRE-ESO-19550-5774

Links

[1] `http://www.atm.ox.ac.uk/RFM/`

[2] `http://rtweb.aer.com/lblrtm_frame.html`

[3] `http://rtweb.aer.com/docs/FAQ_LBLRTM.pdf`

[4] `http://www.physics.wisc.edu/ craigm/idl/cmpfit.html`

[5] `http://www.aer.com/`

[6] `http://www.cfa.harvard.edu/HITRAN/`

[7] VLT Astronomical Site Monitor, ASM Data, User Manual, VLT-MAN-ESO-17440-1773

[8] `http://www.arl.noaa.gov/READYamet.php`

[9] `http://ready.arl.noaa.gov/gdas1.php`

[10] http://nomads.ncep.noaa.gov/txt_descriptions/fast_downloading_grib.shtml

[11] ftp://arlftp.arlhq.noaa.gov/pub/archives/gdas1/

[12] http://irina.eas.gatech.edu/Lab_5560/lblrtm/lblrtm_inst.html

[13] https://www.wiki.ed.ac.uk/display/arcwiki/WP+1.1.1+Task+2

[14] http://www.eso.org/observing/dfo/quality/GENERAL/PWV/HEALTH/

[15] http://cdiac.ornl.gov/

[16] http://www.eso.org/sci/software/cpl

[17] http://www.cpc.ncep.noaa.gov/products/wesley/wgrib2

— End of document —