

ATLAS (AuTomedated Lidar Analysis Software) manual

Author(s): Nikolaos Siomos

Last Update: 11/23/22

Version: 0.1

Table of Contents

1	Installation.....	3
2	Dependencies.....	3
3	Execution.....	4
4	The configuration file.....	8
4.1	Section: Lidar.....	9
4.2	Section: Channels.....	10
4.3	Channel Nomenclature.....	16
5	The settings file.....	17
5.1	Section: general.....	17
5.2	Section: converter.....	19
5.3	Section: preprocessor.....	22
5.4	Section: quicklooks.....	24
5.5	Section: rayleigh_fit.....	27
5.6	Section: telecover.....	29
5.7	Section: polarization_calibration.....	31
6	The radiosonde files.....	34

1 Installation

ATLAS can be cloned through git. In order to do so:

- Git must be installed in the computer
- ATLAS can be cloned from the stable repository: <https://github.com/nikolaos-siomos/ATLAS>
- Open a terminal and cd to the folder you want to place ATLAS to
- run: `git clone https://github.com/nikolaos-siomos/ATLAS.git`

Alternatively you can just download ATLAS from the link above.

2 Dependencies

The current version of ATLAS has been developed and tested in python 3.9. The following python packages should be installed:

- numpy – version 1.24.0
- pandas – version 1.5.2
- xarray – version 2022.12.0
- netcdf4 – version 1.6.2
- scipy – version 1.9.3
- matplotlib – version 3.6.2
- h5netcdf – version 1.1.0

Note that other package versions (or higher versions) might also coincidentally work.

The recommended way to install the dependencies and run ATLAS is currently through an Anaconda environment (tested with Miniconda). The dependencies for the current version were collected with conda – version 4.12.0

An installation example follows for Linux systems (it should also work for Windows and MAC systems – if problems arise the users are encouraged to provide feedback to the EARLINET forum):

- Download and install the latest [Miniconda](#) for python 3.9
- Activate the **base** environment (usually it is auto-activated by default)
- Create a new environment to install the dependencies of ATLAS (let's call it **atlas_box** here):
 - `conda install -n atlas_box python=3.9`
- It is highly recommended to install **mamba** at this stage that greatly speed ups conda.
 - `conda install -n base mamba`

- If mamba is not installed, please replace the word “mamba” with the word “conda in all commands below. Test were run with mamba – version 0.15.3
- Install all dependencies using mamba in the **atlas_box** environment
 - `mamba install -n atlas_box numpy=1.24 pandas=1.5.2 xarray=2022.12.0 netcdf4=1.6.2 scipy=1.9.3 matplotlib=3.6.2 h5netcdf=1.1.0`
 - If one or more of the packages are not available in the specific version then try installing the version the package manager recommends e.g.:
 - `mamba install -n atlas_box numpy pandas xarray netcdf4 scipy matplotlib h5netcdf`
- Before running ATLAS make sure that the **atlas_box** environment is activated. It is best to close the terminal after installation of the dependencies and open a new one to execute ATLAS

3 Execution

Execution of ATLAS is quite straightforward given that the necessary input files (settings file, configuration file) are properly set.

In general all necessary input and output paths can be provided as command-line arguments (see the argument list below). The default folder structure is used in the example bellow to make the definition of paths a bit more clear. Let’s assume the following two folders:

/mydrive/ATLAS: contains the ATLAS software downloaded from Git. The `__master__.py` is the main executable and it is located at */mydrive/ATLAS/program/__master__.py*

/anotherdrive/lidar_data: the input lidar data and the configuration and settings file are placed here

/anotherdrive/radiosondes: the input radiosonde ascii data

The default folder structure that ATLAS expects is given below. Directories are marked with bold and files with bold italic:

- **/mydrive**
 - **/ATLAS**
 - **/program**
 - ***__master__.py*** – main executable
- **/anotherdrive**
 - **/radiosondes** – must exist, see section 6 for the filename format
 - **/lidar_data**
 - ***/settings_file.ini*** – must exist, see section 5
 - ***/config_file.ini*** – must exist, see section 4

- **/nrm** – normal measurement, optional (must include a pol. Calibration for PollyXT systems)
- **/tlc** – telecover measurement with 4 sectors, optional
 - **/north** – north sector files
 - **/east** – east sector files
 - **/south** – south sector files
 - **/west** – west sector files
- **/tlc_rin** – telecover measurement with 2 rings, optional
 - **/inner** – inner ring files
 - **/outer** – outer ring files
- **/pcb** – polarization calibration measurement, optional (ignored for PollyXT systems – nrm folder is used instead)
 - **/+45** – files recorded with the calibrator at +45° (or equivalent, e.g. HWP at 22.5°)
 - **/-45** – files recorded with the calibrator at -45° (or equivalent, e.g. HWP at 22.5°)
- **/drk** – common dark measurement for all QA tests – optional

Basic execution example assuming anaconda was used for the installation of the python environment:

- Activate the **atlas_box** environment using:
 - `conda activate atlas_box`
- Get the absolute path to the `__master__.py` file, eg. **/mydrive/ATLAS/__master__.py**
- We will need at least the following argument:
 - `--parent_folder /mydrive/lidar_data/ --radiosonde_folder /mydrive/radiosondes --file_format licel`
- Run the following command:
 - `python /mydrive/ATLAS/__master__.py --parent_folder /mydrive/lidar_data/ --radiosonde_folder /mydrive/radiosondes --file_format licel`
- For a PollyXT system substitute **licel** with **polly_xt**

It is possible to run ATLAS also from the Spyder IDE. For execution using Spyder:

- Press **Ctrl + F6** OR go to **Run --> Configuration per File**
- Activate **Run file with custom configuration**
- Tick the **Command line options**
- Add all necessary arguments inside the box:
 - `--parent_folder /mydrive/lidar_data/ --radiosonde_folder /mydrive/radiosondes --file_format licel`
 - For a PollyXT system substitute **licel** with **polly_xt**
- Process using **F5** OR **Run File** like a normal python script

Sometimes it is more convenient to use different paths and/or filenames for the configuration file and the settings file. If for example they were all placed directly under the */mydrive* folder and *config_file.ini* and *settings_file.ini* were renamed to *config_file_polis.ini* and *settings_file_polis.ini*:

- We would need additionally the following arguments:
 - `--config_file /mydrive/config_file_polis.ini`
 - `--settings_file /mydrive/settings_file_polis.ini`
- And the execution command becomes:
 - `python /mydrive/ATLAS/__master__.py --parent_folder /mydrive/lidar_data/ --radiosonde_folder /mydrive/radiosondes --file_format licel --config_file /mydrive/config_file.ini --settings_file /mydrive/processing_options.ini`
 - For a PollyXT system substitute **licel** with **polly_xt**

The list of ATLAS arguments.

According to the default folder structure all paths are set with respect to a user provided parent folder path. Providing a path explicitly always overrides its default value. If the parent folder is not provided then all other paths must be explicitly defined. The file format is a mandatory argument, it must be provided.

-f or --parent_folder

The path to the parent folder that is the basis of the default folder structured. It is mandatory unless all the other paths are explicitly provided.

--dark_folder

The path to the dark folder. Defaults to: a drk folder inside the parent folder

--rayleigh_folder

The path to the rayleigh fit measurement folder. Defaults to: a nrm folder inside the parent folder

--telecover_sectors_folder

The path to the telecover folder that contains the sector files. Defaults to: a tlc folder inside the parent folder

--telecover_rings_folder

The path to the telecover folder that contains the ring (inner/outer) files – if any. Defaults to: a tlc_rin folder inside the parent folder)

--pcb_cal_p45_folder

The path to the polarization calibration +45 folder. Defaults to: a pcb/+45 folder inside the parent folder

--pcb_cal_m45_folder

The path to the polarization calibration -45 folder. Defaults to: a pcb/-45 folder inside the parent folder

--pcb_cal_stc_folder

The path to the polarization calibration folder for a calibration with a single calibrator position. Defaults to: a pcb/stc folder inside the parent folder

--radiosonde_folder

The path to the radiosonde folder. The radiosonde file that is closest to the measurement within 12h will be selected if more than 1 files are provided inside

--converter_out

The path to the folder where the converted netcdf files will be placed. This optional argument can be used if the folder must be placed out of the parent_folder. Defaults to:
parent_folder/netcdf

--preprocessor_out

The path to the folder where the preprocessed netcdf files will be placed. This optional argument can be used if the folder must be placed out of the parent_folder. Defaults to:
parent_folder/netcdf

--visualizer_out

The path to the folders where the plots and ascii folders will be created. This optional argument can be used if these two folders must be placed out of the parent_folder (default)

-c or --config_file

The path to the configuration file that contains the necessary metadata. This optional argument can be used if the configuration file must be placed out of the parent_folder (default)

-s or --settings_file

The path to the settings file that contains the processing options to run ATLAS. This optional argument can be used if the configuration file must be placed out of the parent_folder (default)

--file_format

Raw lidar file format. Currently only licel and polly_xt are supported. Choose one of:

- licel: for systems that use licel recorders
- polly_xt: for systems that use polly_xt recorders

4 The configuration file

The *configuration_file.ini* sample file is provided inside the *./templates* folder. The file consists of two main sections, Lidar and Channels. The file is structured according to the following format:

```
[Lidar]
variable_1 = <single_element_1>
variable_2 = <single_element_2>
.....

[Channels]
variable_A = <element_A1>, <element_A2>, ....., <element_An>
variable_B = <element_B1>, <element_B2>, ....., <element_Bn>
.....
```

According to the example, the first section consist of variables that take a single value as input while the later consist of variables that take coma separated values as input. For the latter, the number of values per variable is always equal to the number of signal channels that the user wants to process.

Some fields are optional. The raw file parser attempts to fetch all available metadata from the raw files. Default values are provided when the user either leaves them empty or does not include them in the configuration file at all. For example:

```
variable_1 = <single_element_1>
variable_2 =
variable_3 = <single_element_3>
```

Is equivalent to:

```
variable_1 = <single_element_1>
variable_3 = <single_element_3>
```

The variable_2 will not be taken at all into account and default values will be used instead. Please not that some of the fields are mandatory and there are others that are partly mandatory in a scene that using the default values can be quite error prone. More details on the variables themselves are provided the following sections.

4.1 Section: Lidar

Mandatory Variables

lidar_name: The lidar name. This field will be reported in the plots.

Example: lidar_name = POLIS

lidar_id: The lidar id according to the EARLINET DB.

Example: lidar_id = mun

Optional Variables

file_format: Define the parser for the raw files (Defaults to: licel). Currently licel and polly_xt are supported.

Example: file_format = polly_xt

altitude: The altitude of the station above sea level (in meters).

Example: altitude = 535

location: The station location. This field will be displayed in the plots.

Example: location = Munich

latitude: The station latitude (in degrees).

Example: latitude = 48.148

longitude: The station longitude (in degrees).

Example: longitude = 11.573

zenith_angle: Zenith angle of the lidar (in degrees, 0 at zenith, 90 at horizon).

Example: zenith_angle = 41

azimuth_angle: Azimuth angle of the lidar (in degrees, North = 0, E = 90).

Example: azimuth_angle = 0

4.2 Section: Channels

All variables provided for this section should include exactly as many channels as the `channel_id` variable. If a variable is not relevant for a specific channel (e.g. `dead_time` for analog channels) fill with “_”. Note that the number of channels provided has to be equal or less than the total number of channel in the raw input files. A subset of the raw channels can be provided in case the user doesn’t want to process all of them. For Licel systems, the Licel channel ID in combination with the laser identifier in the header (see the Licel manual) are used in order to link with the raw binary files. For PollyXT systems there is no specific identifier per channel. The user can use the channel index (first channel --> 1, second channel --> 2, etc) to link with the raw netcdf files. If nothing is provided, all PollyXT channels available in the raw files will be processed. For example, `channel_id = 1, 2, 4` means that the 3rd channel ordered as in the raw netcdf file will not be processed. It is therefore important that all input files **have exactly the same channels** when processing a measurement with ATLAS (e.g. same channels for all QA tests). This is also a good practice for Licel systems.

A detailed explanation of all the variables is provided below with examples. For all examples we have assumed the same system of 12 channels (6 physical) including analog and photon detection at 355 parallel and cross, 387 Raman, 532 parallel and cross, and 607 Raman. Please note that all example values presented here correspond to a fictional system and **should do not be blindly taken and applied to an actual system**.

Mandatory Variables

telescope_type: The telescope configuration. Choose among:

- n: near range
- f: far range
- x: single telescope

Example: `telescope_type = x, x, x, x, x, x, x, x, x, x, x, x`

channel_type: The channel type. Choose one among:

- p: co-polar linear analyzer
- c: cross-polar linear analyzer
- t: total (no depolarization)
- o: co-polar circular analyzer
- x: cross-polar circular analyzer
- v: vibrational Raman
- r: rotational Raman
- a: Cabannes
- f: fluorescence

Example: channel_type = p, p, c, c, v, v, p, p, c, c, v, v

channel_subtype: The channel subtype. Choose among:

- r: Signal Reflected from a PBS
- t: Signal Transmitted through a PBS
- n: N2 Ramal line
- o: O2 Ramal line
- w: H2O Ramal line
- c: CH4 Ramal line
- h: High Rotational Raman
- l: Low Rotational Raman
- a: Mie (aerosol) HSRL signal
- m: Molecular HSRL signal
- b: Broadband Fluorescence
- s: Spectral Fluorescence
- x: No specific subtype

Example: channel_subtype = r, r, t, t, n, n, t, t, r, r, n, n

dead_time: The dead time of the photon counting channels (in ns). (for analog channels set _).

Example: dead_time = _, 3.571, _, 4.545, _, 3.704, _, 3.167, _, 3.846, _, 3.846

trigger_delay_bins: The trigger delay per channel in bins. Provide negative values if the recording starts before the Q-switch (pre-triggering) and positive values if the recording starts after the Q-switch.

Example: trigger_delay_bins = -2020, -2011, -2020, -2011, -2020, -2011, -2025, -2014, -2025, -2014, -2025, -2014

Mandatory Variables for Licel

channel_id: IDs of each channel according to the raw file header. For licel provide the licel id per channel that is going to be include. Currently only BT and BC channels are supported. For PollyXT systems all channels are included by default. Provide only if you want to process specific channels

Example: channel_id = BT0, BC0, BT1, BC1, BT2, BC2, BT3, BC3, BT4, BC4, BT5, BC5

laser: The ascending laser number according to the licel manual. Use to link with the licel file. The channel_id is not a unique identifier for licel channels. A single channel can synchronize with more than one lasers. This variable is ignored when processing PollyXT files

Example: laser = 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

Mandatory Variables in the Future

scc_id: IDs of each channel according to the SCC configuration. In the future, linking with the HOI will be done through the scc_id (currently optional).

Example: scc_id = 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630

Partly Optional Variables

These variables take default values. It is highly recommended though to fill them in because might not be valid for the system.

dead_time_correction_type: The type of dead time correction. Choose among:

- 0: Non Paralyzable
- 1: Paralyzable

Defaults to: 0 for photon channels.

Example: dead_time_correction_type = _, 0, _, 0, _, 0, _, 0, _, 0, _, 0

emitted_wavelength: The wavelength of the originally emitted laser beam per channel. A very rough guess process is applied by default using that detected channel information that it is valid **ONLY for Nd:Yag lasers with elastic and vibrational Raman channels**. Please provide the values explicitly if your system does not fall in this category. Defaults to:

- 355 for channels with detected wavelength <520nm
- 532 for channels with detected wavelength <1000nm
- 1064 for channels with detected wavelength >=1000nm

If accurate molecular depolarization ratio calculations are required this variable should be filled in with the actual emitted wavelength with sub-nanometer accuracy.

Example: emitted_wavelength = 354.6, 354.6, 354.6, 354.6, 354.6, 354.6, 532.2, 532.2, 532.2, 532.2, 532.2

detected_wavelength: The wavelength of the detected signal according to the applied Interference Filter of each channel (in nm). This variable is provided by default from the raw file metadata for both Licel and PollyXT. However, if accuracy is required for molecular calculations the exact central wavelength should be provided.

Example: detected_wavelength = 354.6, 354.6, 354.6, 354.6, 386.5, 386.5, 532.2, 532.2, 532.2, 532.2, 607.5, 607.5

channel_bandwidth: Interference filter Bandwidth (in nm, Defaults to: 1nm). The default value is a very rough guess and it can easily lead to high uncertainty in the molecular depolarization value. Exact values provided from the manufacturers should be used instead.

Example: `channel_bandwidth = 0.5, 0.5, 0.5, 0.5, 1.0, 1.0, 0.5, 0.5, 0.5, 0.5, 1.0, 1.0`

background_low_bin: Starting bin of the background correction averaging range. Defaults to:

- The 600th bin before the end of the profile if the `trigger_delay_bins < 400`
- The 100th bin if the `trigger_delay_bins >= 400`

Using default values is currently risky. Please make sure that no spikes are appearing in the pretrigger range that would affect the background correction or manually provide this variable.

Example: `background_low_bin = 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100, 100`

background_high_bin: ending bin of the background correction averaging range. Defaults to:

- The 100th bin before the end of the profile if the `trigger_delay_bins < 400`
- The 300th bin if the `trigger_delay_bins >= 400`

Using default values is currently risky. Please make sure that no spikes are appearing in the pretrigger range that would affect the background correction or manually provide this variable.

Example: `background_high_bin = 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500, 500`

Optional Variables (include only to override the raw file metadata)

acquisition_mode: The mode of the recorded signals per channel. Choose among:

- 0: analog
- 1: photon counting

Example: `acquisition_mode = 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1`

bins: The total bins of the recorded signals per channel.

Example: `bins = 8192, 8192, 8192, 8192, 8192, 8192, 4096, 4096, 4096, 4096, 4096, 4096`

laser_polarization: The laser polarization of the recorded signals per channel. Choose among:

- 1: linear
- 3: circular

Example: laser_polarization = 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

laser_shots: The number of acquired laser shots per channel. Not recommended to provide it manually. If fetched from the metadata, this variable is different per file which is more accurate than providing a constant value here.

Example: laser_shots = 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200, 200

data_acquisition_range: The Data Acquisition Range of each analog channel [mV]. Use _ for photon channels

Example: data_acquisition_range = 100, _, 100, _, 20, _, 100, _, 100, _, 20, _

analog_to_digital_resolution: The analog to digital resolution (in bits) of each analog channel. Use _ for photon channels.

Example: analog_to_digital_resolution = 16, _, 16, _, 12, _, 16, _, 16, _, 12, _

range_resolution: The range resolution of each channel [m]. It will be used to calculate the Sampling Rate.

Example: range_resolution = 3.75, 3.75, 3.75, 3.75, 3.75, 3.75, 7.5, 7.5, 7.5, 7.5, 7.5, 7.5

pmt_high_voltage: The high voltage [V] provided to the detection unit per channel.

Example: pmt_high_voltage = 850, 850, 870, 870, 900, 900, 840, 840, 860, 860, 920, 920

laser_repetition_rate: The Laser Repetition Rate [Hz]

Example: laser_repetition_rate = 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 20, 20

4.3 Channel Nomenclature

Based on the input provided in the configuration file and the header of the raw files an ID is created per channel in ATLAS and accompanies it during processing and exporting. The ID consists of the following 8 characters:

- first 4 characters correspond to the **detected_wavelength** variable rounded up to the nearest integer and filled with leading zeros wherever applicable (e.g. 0355, 0607, 1064 etc)
- the 5th character corresponds to the **telescope_type** variable (refer to the respective table)
- the 6th character corresponds to the **channel_type** variable (refer to the respective table)
- the 7th character corresponds to the acquisition mode and can be “a” for analog or “p” for photon. The respective information is provided by the **acquisition_mode** variable
- the 8th character corresponds to the **channel_subtype** variable (refer to the respective table)

For example, channel **0355ftpr** would correspond to a far field photon counting co-polar channel at 355nm detected in the reflected path with respect to the PBS

In another example, channel **0387nvan** would correspond to a near field analog vibrational Raman channel at 387nm the corresponds to scattering from N₂ molecules.

Following the 12 channels defined in the examples of section 4.2, we would get the following channels in the ATLAS plots and netcdf files:

0355xpar, 0355xppr, 0355xcat, 0355xcpt, 0387xvan, 0387xvpn, 0532xpat, 0532xppt, 0532xcar, 0532xcpr, 0607xvan, 0607xvpn

5 The settings file

The *settings_file.ini* sample file is provided inside the *./templates* folder. Similarly to the *config_file.ini* it is also structured in sections and shares a similar syntax. The main difference is that the variables of the settings file do not have a specific number of values per section.

The file consists of seven main sections: general, converter, preprocessor, quicklooks, rayleigh_fit, telecover, and polarization_calibration. Note that non of these variables is truly mandatory. Unless provided they take default values that are explained below. However

5.1 Section: general

isday: Defaults to: False. If set to True all the vibrational Raman and the fluorescence signals will be excluded in the prepossessing stage

Example: isday = True

newdata: Defaults to: True. If set to False the converter and the preprocessing modules will not be called if the algorithm detects output files already produced by them for a specific measurement. This mainly saves time during execution.

Example: newdata = False

visualize: Defaults to: True. If set to False the visualizer modules that create the plots for the QA tests will not be called. Use when only processing of the data is required.

Example: visualize = False

process: The user can choose specific QA test(s) to process. It is only taken into account if **visualize** is set to True. Use any of:

- ray: Rayleigh Fit
- tlc: Telecover Test
- pcb: Polarization Calibration.

Defaults to: ray, tlc, pcb

Example: process = ray, tlc

This would create plots for the rayleigh fit and the telecover test (polarization calibration would be skipped)

process_qck: Choose which quicklooks to process. It is only taken into account if **visualize** is set to True. It must be a subset of **process**. Note that each visualizer module (ray, tlc, pcb) can create their own quicklooks. Choose among

- ray: Rayleigh Fit
- tlc: Telecover Test
- pcb: Polarization Calibration

Defaults to: ray

Example: process_qck = ray, pcb

This would create quicklooks only for both the Rayleigh fit test and the polarization calibration

5.2 Section: converter

debug: If set to True, debugging files will be generated in the `<path_to_the_converter_folder>/debug` folder. Note that the path to the converter folder Defaults to: `<parent_folder>/netcdf` but it can be different if manually provided by the user. These included the metadata gathered from the configuration file, the licel header, and the combination of the two. Defaults to: False

Example: debug = True

trim_overflows: This options determines how overflow values in the raw input files will be treated if detected. Choose among:

- 0: the algorithm will stop and provide a diagnostic error (default)
- 1: the files containing at least one overflow value will be screened out
- 2: overflows will be interpolated (use with care and only only a few bins per profile have overflow values)
- 3: overflows will not be excluded, use this only for debugging purposes

Example: trim_overflows = 2

files_per_sector: The number of telecover files per sector (integer). If provided, the telecover files must be placed in a single folder (this should be `<path_to_parent_folder>/tlc` according to the default folder structure). An automated assignment of the telecover files in different sectors will be attempted serially assuming the following temporal sequence of sectors: north – east – south – west. Note that the telecover test can have more than 1 rounds.

Example: files_per_sector = 5

files_per_ring: The number of telecover files per ring (integer). If provided, the telecover files must be placed in a single folder (this should be `<path_to_parent_folder>/tlc_rin` according to the default folder structure). An automated assignment of the telecover files in different sectors will be attempted serially assuming the following temporal sequence of sectors: inner - outer. Note that the telecover test can have more than 1 rounds.

Example: files_per_ring = 5

rsonde_skip_header: Radiosonde parser option. Number of lines to skip at the beginning of the radiosonde ascii file. Defaults to: 1 (1 line reserved for header info)

Example: rsonde_skip_header = 6 (skip the first 6 lines)

rsonde_skip_footer: Radiosonde parser option. Number of lines to skip at the end of the radiosonde file. Defaults to: 0 (no footer assumed)

Example: rsonde_skip_footer = 10 (skip the last 10 lines)

rsonde_delimiter: Radiosonde parser option. The delimiter that separates columns in the radiosonde file choose one of:

- S: space
- C: comma

Defaults to: S.

Example: rsonde_delimiter = C

rsonde_column_index: Radiosonde parser option. The index (integer) of the column that contains the Height, Pressure, Temperature, and Relative Humidity (optional) information. For example, rsonde_columns = 1, 3, 2, 6 means:

- Height: 1st column in the radiosonde file
- Pressure: 3rd column
- Temperature: 2nd column
- Relative Humidity: 6th column

The relative humidity column is OPTIONAL and can be omitted. Defaults to: 2 1 3

Example: rsonde_column_index = 2, 1, 3, 5

rsonde_column_units: Radiosonde parser option. The units of Height, Pressure, Temperature, and Relative Humidity (optional) columns in the radiosonde file. The number of values must be the same as in **rsonde_column_index**. Supported units for:

- height: m_asl (default), km_asl, m_agl, km_agl
- pressure: hPa (default), Pa, atm
- temperature: K (default), C, Cx10

- relative humidity: percent (default), fraction

If the Height units are in agl (above ground level) then the station altitude must be provided in the **rsonde_geodata**

Example: rsonde_column_units = m_asl, hPa, C, percent

rsonde_geodata: The radiosonde station latitude, longitude, and altitude. For example: rsonde_geodata = 35.33, 25.18, 39.0 means:

- Latitude: 35.33°
- Longitude: 25.18°
- Altitude: 39.0 m

Defaults to None for all values.

Example: rsonde_geodata = 48.25, 11.55, 492.0

rsonde_station_name: The name of the station where the radiosonde measurement was performed. Defaults to: None.

Example: rsonde_station_name = Muenchen-Oberschleissheim

rsonde_wmo_number: The WMO number of the radiosonde station. Defaults to: None

Example: rsonde_wmo_number = 10868

rsonde_wban_number: The WBAN number of the radiosonde station. Defaults to: None

Example: rsonde_wban_number =

5.3 Section: preprocessor

vertical_trimming: If set to True then bins above a certain altitude (20km by default) will be removed. Can speed up computations and save RAM. Defaults to: True

Example: vertical_trimming = False

vertical_limit: The maximum altitude in km above which no calculations will be performed. Solar background calculations are performed prior to vertical signal trimming to enable background calculations up to the maximum signal range. Defaults to: 20km

Example: `vertical_limit = 25.0`

channels: Provided it to process only selected channels. By default, no channel is excluded if this variable is not provided. The channel name must follow the nomenclature of section 4.3.

Example: `channels = 0355xcar, 0355xcpr`

In this case only the these 2 channels will be processed

exclude_field_type: Provide it to entirely exclude all channels of a specific **telescope_type**. By default, no channel is excluded if this variable is not provided.

Example: `exclude_field_type = x, n`

This will exclude all channels with x and n values provided in their **telescope_type**, that is all near field and unspecified channels

exclude_channel_type: Provide it to entirely exclude all channels of a specific **channel_type**. By default:

- if `isday = False` --> no channel is excluded if this variable is not provided
- if `isday = True` --> `exclude_scattering_type = v, f` (vibrational Raman and fluorescence channels excluded)

Example: `exclude_scattering_type = v, r, o, x`

This will exclude all channels with v, r, o, and x values provided in their **channel_type**, that is all vibrational and rotational Raman, co- and cross- circular polarized channels

exclude_detection_mode: Provide it to entirely exclude all channels of a specific **acquisition_mode**. By default, no channel is excluded if this variable is not provided.

Example: `exclude_detection_mode = a`

This will exclude all channels with 0 values provided in their **acquisition_mode**, that is all analog channels.

exclude_channel_subtype: Provide it to entirely exclude all channels of a specific **channel_subtype**. By default, no channel is excluded if this variable is not provided.

Example: `exclude_channel_subtype = w`

This will exclude all channels with w values provided in their **channel_subtype**, that is all water vapor channels.

5.4 Section: quicklooks

The following variables:

- **channels**
- **exclude_field_type**
- **exclude_channel_type**
- **exclude_detection_mode**
- **exclude_channel_subtype**

where defined in section 5.3 and are also applicable here. The usage is the same with the only exception that exclusions take place in the visualization phase and only for the quicklooks

use_distance: If set to False, the y axis units of the quicklook will correspond to the altitude in meters above sea level. If set to True, the y axis units corresponds to the distance from the system. This variable determines the units of the following variables:

- **y_lims**
- **z_max_zone**
- **smoothing_range**

Defaults to: True

Example: use_distance = False

x_lims: The x axis limits (lower and upper). Use two integers corresponding to the first and last temporal samples (not date!) that will be plotted in the quicklooks. Use 1 to start from the first sample. If values below 1 or above the total number of samples are used, they will be ignored

Example: x_lims = 100, 300

Assuming a Rayleigh Fit measurement with 1000 samples, the quicklook with start cover the zone between the 100th and the 300th.

x_tick: The x axis finest major tick in number of samples. Defaults to: automatic selection.

Example: x_tick = 2

t_lims: The x axis limits in time units. Use the following format hh:mm for both limits (not meant to be used for day-long quicklooks). Defaults to: automatic selection.

Example: *t_lims = 17:38, 21:56*

t_tick: The major tick for the time axis (same as x axis but with different units). Defaults to: automatic selection

Example: *t_tick = 5*

y_lims: The y axis limits (altitude/distance) in km (lower and upper). If values below 0 or above the maximum altitude/distance are used, they will be ignored. Defaults to: 0., 14.

Example: *y_lims = 0., 10.*

y_tick: The y axis finest major tick in km. Defaults to: 1km

Example: *y_tick = 0.5*

z_lims: The colorscale limits of the normalized range-corrected signal normalized with the mean value in the **z_max_zone** region. This setting is useful if for example the normalization takes place in a very strong aerosol layer making all other layers difficult to discern. Defaults to:

- *z_lims = 0., 1.* when **use_log_scale** = False and
- *z_lims = 1E-5, 1.* when **use_log_scale** = True

Example: *z_lims = 0., 0.9* (assuming **use_log_scale** = False)

z_max_zone: Provide a zone (min and max altitude/distance) in km. The maximum range-corrected signal value encountered in the zone will be used for the normalization of the range-corrected signal for the quicklook. Particularly useful in order to avoid scaling the colors with a cloud. Defaults to: 0.1, 2.

Example: *z_max_zone = 0.5, 1.5*

smooth: If set to True, a sliding average smoothing filter will be applied on the signals across y axis for better visualization. Defaults to: True

Example: *smooth = False*

smoothing_exponential: This variable is ignored if `smooth = False` and also if the upper and lower values of the **half_window** variable are the same. Choose one of:

- `True`: a smoothing window that exponentially increases with altitude/distance will be applied.
- `False`: a smoothing window that exponentially increases with altitude/distance will be applied

Defaults to: `False`.

Example: `smoothing_exponential = True`

smoothing_range: Set the first and last altitude/distance boundaries in km where smoothing should be applied. If they exceed the actual signal boundaries the actual boundaries will be used instead. Defaults to: 1., 14.

Example: `smoothing_range = 0.25, 14`.

half_window: Half smoothing window in the first and last bin of the smoothing region, in m. The window progressively changes between from the first to the last value. Use the same value twice to apply a constant window. Defaults to: `half_window = 100., 100`.

Example: `half_window = 5., 500`.

dpi: The dots per inch (dpi) resolution of the exported figures. Defaults to: 300 dpi

Example: `dpi = 100`

5.5 Section: rayleigh_fit

The following variables:

- **channels**
- **exclude_field_type**
- **exclude_channel_type**
- **exclude_detection_mode**
- **exclude_channel_subtype** (b, s, a, w, c)

where defined in section 5.3 and are also applicable here. The usage is the same with the only exception that exclusions take place in the visualization phase and only for the Rayleigh fit test. Respective default values are provided in parenthesis.

use_distance: If set to False, the x axis units of the Rayleigh fit will correspond to the altitude in meters above sea level. If set to True, the x axis units corresponds to the distance from the system. This variable determines the units of the following variables:

- **reference_height**
- **x_lims**
- **smoothing_range**

Defaults to: True

Example: use_distance = False

The following variables:

- **smooth** (True)
- **smoothing_exponential** (False)
- **smoothing_range** (1., 14.)
- **half_window** (100., 100.)
- **dpi** (300)

where defined in section 5.4 and are also directly applicable here. Respective default values are provided in parenthesis.

use_lin_scale: If set to True, a linear scale will be used for the y axis (signal axis). If set to False a logarithmic scale will be applied instead. Defaults to False:

Example: `use_lin_scale = True`

reference_height: The reference height/distance in km where the signals will be normalized for the Rayleigh fit. Defaults to: 9. In the future the reference height detection will be automated.

Example: `reference_height = 6`.

half_reference_window: The half window in meters used for the normalization. Defaults to: 100

Example: `half_reference_window = 1000`.

x_lims: Set the x axis (altitude/distance) limits in km (lower and upper). If values below 0 or above the maximum altitude/distance are used, they will be ignored. Defaults to: 0., 14.

Example: `x_lims = 0., 24`.

x_lims: The x axis finest major tick in km. Defaults to: 1km

Example: `x_tick = 2`.

y_lims: The y axis (signal) limits (lower and upper) of the normalized range-corrected signal in $\text{m}^{-1} \text{sr}^{-1}$ (pseudo-units). It is recommended to skip this variable and use the automatic selection because channels in different wavelengths have different optimal limits. Defaults to: automatic selection

Example: `y_lims = 5E-8, 5E-5`

5.6 Section: telecover

The following variables:

- **channels**
- **exclude_field_type**
- **exclude_channel_type**
- **exclude_detection_mode**
- **exclude_channel_subtype**

where defined in section 5.3 and are also applicable here. The usage is the same with the only exception that exclusions take place in the visualization phase and only for the Rayleigh fit test

The following variables:

- **use_distance** (True)
- **x_lims** (0., 2.4)
- **x_tick** (0.2)
- **y_lims** (automatic selection)
- **smooth** (True)
- **smoothing_exponential** (False)
- **smoothing_range** (0.2, 2.4)
- **half_window** (5., 200.)
- **dpi** (300)

where defined in section 5.5 and are also applicable here. The default values are not always the same and are provided in parenthesis. The **use_distance**, **use_lin_scale**, **x_lims** and **x_tick** variables correspond to the x axis that is the same for all the 3 subplots of the telecover test. The **y_lims** variable corresponds to the units of the middle telecover subplot (normalized signals). The left subplot has different units that are adapted accordingly depending on the provided **y_lims** values.

use_non_rangecor: If set to True, the non range corrected signals will be used for the left subplot of the telecover test. If set to False the range corrected signals will be used instead. Defaults to: False

Example: use_non_rangecor = True

use_last: If True additional black line with the difference between the first and the last sector is added in the right telecover subplot (deviations) e.g. $N2 - N1$. Defaults to False.

Example: use_last = True

normalization_height: The reference altitude/distance in km where the signals will be normalized for the Telecover test (relevant to middle and right subplots). It should be well above the suspected full overlap height. Defaults to: `normalization_height = 2`.

Example: normalization_height = 1.5

half_normalization_window: The half window in meters used for the normalization. Defaults to: 200.

Example: half_normalization_window = 500.

5.7 Section: polarization_calibration

The following variables:

- **use_distance** (True)
- **smooth** (True)
- **smoothing_exponential** (False)
- **smoothing_range** (1., 14.)
- **half_window** (100., 100.)
- **dpi** (300)

where defined in section 5.5 and are also applicable here. The default values are not always the same and are provided in parenthesis.

ch_r: Provide here channel names from the available ones with the reflected **channel_subtype** (e.g. 0355xpar). The number of reflected channels must be the same as the number of the respective transmitted channels provided by **ch_t** and by the GHK parameters. By default, all available reflected channels will be matched to all available transmitted channels that share the same telescope type, detection mode, and wavelength. WARNING! The field is mandatory if non-default GHK values are applied.

Example: ch_r = 0355xppr, 0532xcpr

ch_t: Provide here channel names from the available ones with the transmitted **channel_subtype** (e.g. 0355xcat). The number of reflected channels must be the same as the number of the respective transmitted channels provided by **ch_r** and by the GHK parameters. By default, all available reflected channels will be matched to all available transmitted channels that share the same telescope type, detection mode, and wavelength. WARNING! The field is mandatory if non-default GHK values are applied.

Example: ch_t = 0355xcpt, 0532xppt,

calibration_height: The calibration altitude/distance in km where the signals will be normalized for the pol. calibration. Defaults to: calibration_height = 3.

Example: calibration_height = 3.

half_calibration_window: The half window in meters used for the calibration. Defaults to: 500.

Example: half_calibration_window = 1000.

rayleigh_height: The calibration altitude/distance in km where the signals will be normalized for the comparison with the Rayleigh atmosphere. Defaults to: 9.

Example: rayleigh_height = 6

half_rayleigh_window: The half window in meters used for the comparison with the Rayleigh atmosphere. Defaults to: half_calibration_window = 500.

Example: half_rayleigh_window = 1000.

x_lims_calibration: The x axis (altitude/distance) limits in km (lower and upper) for the pol. calibration subplot. If values below 0 or above the maximum altitude/distance are used, they will be ignored. Defaults to: 0, 5

Example: x_lims_calibration = 0.1, 6.

x_lims_rayleigh: The x axis (altitude/distance) limits in km (lower and upper) for the Rayleigh comparison subplot. If values below 0 or above the maximum altitude/distance are used, they will be ignored. Defaults to: 0, 10

Example: x_lims_rayleigh = 0.1, 8.

x_tick_calibration: The x axis finest major tick in km for the pol. calibration subplot. Defaults to: 0.5

Example: x_tick_calibration = 1.

x_tick_rayleigh: The x axis finest major tick in Km for the Rayleigh plot. Defaults to: 1.

Example: x_tick_rayleigh = 2.

y_lims_calibration: The y axis limits (lower and upper) of the gain ratios at +-45. Used for the pol. calibration subplot. Defaults to: automatic selection

Example: y_lims_calibration = 0, 0.1

y_lims_rayleigh: The y axis limits (lower and upper) of the volume depolarization ratio. Used for the Rayleigh subplot. Defaults to: automatic selection

$y_lims_rayleigh = 0, 0.1$

K: The K value for each channel pair. Defaults to: 1 for all channels

Example: $K = 1.05, 1.$

G_R: The G value for the reflected channel of the pair. Defaults to: 1 for all channels (assuming no correction for the emission and the receiver)

Example: $G_R = 1., 1.$

G_T: The G value for the transmitted channel of the pair. Defaults to: 1 for all channels (assuming no correction for the emission and the receiver)

Example: $G_T = 1., 1.$

H_R: The H value for the reflected channel of the pair. Defaults to:

- 1 for all co-polar (p) reflected channels
- -1 for all cross-polar (c) reflected channels

(assuming no correction for the emission and the receiver)

Example: $H_R = 0.94, -0.98$

H_T: The H value for the transmitted channel of the pair. Defaults to: 1 or -1 for all co-polar (p) and cross-polar (c) transmitted channels, respectively (no receiver optics + emitted pcb. state correction)

- 1 for all co-polar (p) transmitted channels
- -1 for all cross-polar (c) transmitted channels

(assuming no correction for the emission and the receiver)

Example: $H_T = -0.94, 0.98$

R_to_T_transmission_ratio: The transmission ratio of the R to T channels per pair (T_R/T_T). $T_R = 1$ and/or $T_T = 1$ if no filter was applied during calibration. Defaults to: 1 for all pairs

Example: R_to_T_transmission_ratio = 8., 0.15

6 The radiosonde files

This section includes some more details about the radiosonde files and how ATLAS is handling them. Please refer also to section 5.2 and specifically to the following variables:

- **rsonde_skip_header**
- **rsonde_skip_footer**
- **rsonde_delimiter**
- **rsonde_column_index**
- **rsonde_column_units**
- **rsonde_geodata**
- **rsonde_station_name**
- **rsonde_wmo_number**
- **rsonde_wban_number**

The file format is general ascii. The parser is flexible and can handle headers and footers, two different delimiters (comma and space), custom column order, and custom units.

The Radiosonde folder

This folder should contain the radiosonde files. At this stage only ascii files can be parsed. Note that more than one files can be provided here. ATLAS scans all files and selects the one that it is closest in time with the Rayleigh measurement. Depending on the application, it can be more convenient to use a specific radiosonde folder where all radiosonde data is included and provided it as a command line argument when running ATLAS. Note that the time difference between the Radiosonde and the lidar measurement **cannot be larger than 18 hours**.

Filename Format

The name of the radiosonde file must follow a specific format so that it is recognized by ATLAS. It must start with the date and end with .txt Whatever comes after the date and before the extension is not important for parsing but can be helpful for sorting purposes for the user:

yyyymmdd_hhmm<add_whatever_you_want_here>.**txt**

Example: 20230219_1200_Munich.txt

This radiosonde file corresponds to 19.02.2023 at 12:00 UTC.