MesosphEO WP 4.3: MIPAS-IMKIAA Time Series ReadMe

S. Lossow, KIT

FMI-TN-MesosphEO-WP4-003
Version 2.1.1
25.5. 2017

1. Introduction

This document explains the netCDF format used for MIPAS-IMKIAA trace gas climatologies (CH_4, CO, CO_2, H_2O, N_2O, NO, NO_2, temperature) in the ESA's MesosphEO project. The files containing the time series are typically named as follows:

MESOSPHEO_parameter_L3_processor-name_retrieval-version_time-of-day_data-type_file-version.file-extension.

If the data are not separated by day or night the "time-of-tag" is omitted.

Examples:

MESOSPHEO_H2O_L3_MIPAS-IMKIAA_V220-221_time-series_fv0001.nc
MESOSPHEO_H2O_L3_MIPAS-IMKIAA_V220-221_daytime_time-series_fv0001.nc

The only exception currently exists for the NO V622 data. In this case two data sets exist. One data sets where only NO is retrieved and another data set where NO is jointly retrieved with temperature. For the latter data set the file name looks as follows:

MESOSPHEO_NO_L3_MIPAS-IMKIAA_V622_JOINT-T_time-series_fv0001.nc
MESOSPHEO_NO_L3_MIPAS-IMKIAA_V622_JOINT-T_daytime_time-series_fv0001.nc

2. Source data

Source data are the standard collection files that combine all retrieved profiles of a given parameter for the individual months.

3. Level2 screening

The level2 screening differs for the different parameters, depending upon if they are based on a linear-space or a log-space retrieval.

For the CH_4, N_2O and temperature data, which are retrieved in linear-space, the screening looks as follows:

(1) data points with a visibility flag of 0 are discarded
(2) data points with an average kernel diagonal element (aka AKD criterion) of less than 0.03 are discarded
(3) data above the uppermost tangent height are not considered any further.

For the CO, CO_2, H_2O, NO and NO_2 data, which are retrieved in log-space, only screening (1) and (3) are applied.

4. Data gridding, screening and statistical estimates

The data in the standard collection files are already on a fixed altitude grid. For convenience the data are interpolated on a 1 km grid extending from 50 km to 100 km. The data are binned monthly and for latitude bands of 10° (i.e. 90°S-80°S, 80°S-70°S, …, 70°N-80°N, 80°N-90°N).

Before the data in a given bin are averaged we apply an additional screening using the median and the median absolute difference. Any data outside the interval [median(x) +/- 7.5*median absolute difference(x)] are discarded, where x denotes the data that falls into a specific bin. This is a relatively weak filter, aiming to remove the most obvious outliers.

For the trace gases retrieved in log-space (i.e. CO, CO_2, H_2O, NO and NO_2) in addition to the binned data the mean of the averaging kernel diagonal is calculated. If this mean is less than 0.03 then the binned data are discarded. In general, binned data based on less than 20 observations are not considered any further, as are binned data that are a smaller than its associated standard error in absolute terms.

The time series data are separated for the observations with full (spectral) resolution (FR, 2002 - 2004) and reduced resolution (RR, 2005 - 2012). Version numbers for the FR data use two digits, for the RR data three digits are used. Likewise the data are separated for the different measurement modes, i.e nominal mode (NOM, version numbers starting with 2), middle atmospheric mode (MA, version numbers starting with 5), upper atmosphere mode (UA, version numbers starting with 6) and the NLC mode (NLC, version numbers starting with 7, just in January and July). An overview when observations in the different measurement modes were performed can be found on the following webpage: "http://eodg.atm.ox.ac.uk/MIPAS/L1B/". Beyond that time series data are made available for daytime (SZA 0°-97°), nighttime (SZA 97°-180°) and all times of day (SZA 0°-180°).

The temperatures are provided in K. The trace gases are given in volume mixing ratio, which is the basic unit. Missing values are represented by NaNs. Besides the averaged data the data files contain several additional estimators as the median, the standard deviation (STD) or the standard of the mean (SEM). The coverage of the data within a given bin is characterised by the mean in latitude, time, day of year and local time. The principal number of measurements going into a specific bin is also provided as function of the day of a month. A detailed description of all data fields is given in Sect. 6.

5. User guidance

The mesospheric coverage of the different MIPAS data sets is approximately as follows:

CH_4: NOM 50 km - 70 km, MA 50 km - 85 km, UA 50 km - 85 km, NLC 50 km - 85 km

CO: NOM 50 km - 70 km, MA 50 km - 100 km, UA 50 km - 100 km, NLC 50 km - 100 km

CO_2: NOM not retrieved, MA not retrieved, UA 75 km - 100 km (only daytime data exist), NLC not retrieved

H_2O: NOM 50 km - 70 km, MA 50 km - 90 km, UA 50 km - 90 km, NLC 50 km - 90 km

NO: NOM 50 km - 65 km, MA 50 km - 100 km (with some gaps where there is little NO), UA 50 km - 100 km (with some gaps where there is little NO), UA (with joint temperature retrieval) only at 100 km, NLC 50 km - 100 km (with some gaps where there is little NO)

N_2O: NOM 50 km - 60 km, MA 50 km - 60 km, UA 50 km - 60 km, NLC 50 km - 60 km

NO_2: NOM 50 km - 60 km (during nighttime, only very little coverage otherwise), MA 50 km - 60 km (during nighttime, only very little coverage otherwise), UA 50 km - 60 km (during nighttime, only very little coverage otherwise), NLC 50 km - 60 km (during nighttime, only very little coverage otherwise),

Temperature: NOM 50 km - 70 km, MA 50 km - 100 km, UA 50 km - 100 km, NLC 50 km - 100 km

6. NetCDF4 format for MIPAS time series

No	Variable	Unit	Dim	Description
1	time	days since 1900-01-01	21 for FR 88 for RR	time given at the month centre
2	time_bands	days since 1900-01-01	21 x 2 for FR 88 x 2 for RR	time bands
3	latitude	degrees north	18	centre of latitude bands
4	latitude_bands	degrees north	18 x 2	latitude bands
5	altitude	km	51	altitude grid 50:1:100 km
6	data_mean	VMR/K	time x 51 x 18	mean of binned data
7	data_median	VMR/K	time x 51 x 18	median of binned data
8	data_sem	VMR/K	time x 51 x 18	standard error of the binned data
9	data_std	VMR/K	time x 51 x 18	standard deviation of the binned data
10	data_obs	number	time x 51 x 18	number of observations binned together
11	avg_time	days since 1900-01-01	time x 18	average time of the binned data
12	avg_doy	days	time x 18	average day of year of the binned data
13	avg_latitude	degrees north	time x 18	average latitude of the binned data

14	avg_lt	hours	time x 18	average local time of the binned data
15	coverage	number	time x 18 x 31	number of observations binned together as function of day of month for every month 31 days are considered for simplicity

In addition the NetCDF files contain a number of global attributes. The most important ones are listed below:

No	**Attribute**	**Description**
1	date_created	file creation time as yyyymmddTHHMMSSZ
2	level_1_data_version	calibration version
3	level_2_data_version	retrieval version
4	value_for_nodata	fill value, NaNs are used
5	minimum_averaging_kernel_diagonal	averaging kernel diagonal screening applied, i.e. 0.03/-Inf for linear-space/log-space retrieved parameters (part of L2 screening)
6	visibility	visibility flag taken into account, i.e. yes (part of L2 screening)
7	data_above_the_highest_tangent_altitude	usage of data above the highest tangent altitude, i.e. no (part of L2 screening)
8	minimum_mean_averaging_kernel_ diagonal	minimum mean average kernel diagonal for the binned data, i.e. -Inf/0.03 for linear-space/log-space retrieved parameters (part of L3 screening)
9	outliers_removed	indicates if outliers were removed before binning the data, i.e. yes (part of L3 screening)

10	removal_method	the method used to remove the outliers, i.e the median and the median absolute difference (MAD, part of L3 screening)
11	factor	factor to determine the interval around the median outside which data are denoted as outliers, i.e. 7.5 x MAD (part of L3 screening)
12	iterations	number of iterations for the outlier determination, i.e. 1 (part of L3 screening)
13	minimum_number_of_observations	minimum number of observations required for the binned data, i.e. 20 (part of L3 screening)
14	time_of_day	time of day considered, i.e. either daytime, nighttime or all
15	solar_zenith_angle_min	minimum solar zenith angle considered
16	solar_zenith_angle_max	maximum solar zenith angle considered
17	file_version	the file version
18	file_version_description	description of file version
19	tracking_id	unique file identifier

7. Version history

version 1.0 / 14 March 2017:
- the initial version

version 1.1 / 16 March 2017:
- adding NO and temperature data

version 2.0 / 27 March 2017:
- adapt binning for trace gases retrieved in log-space, i.e. CO, H_2O and NO
- add information on dates and the corresponding number of observations going into a bin (see field “coverage”)
- add time bounds

- change dimension sequence so that time comes always first (according to the “ncdump” CLI)

version 2.1 / 11 April 2017:
- add CO_2 and NO_2 data
- add UA and NLC mode data for CH_4, CO, H_2O, N_2O and NO
- update of CO MA mode data set from V520 to V521
- update of NO MA mode data set from V520 to V521
- cosmetic changes to the figures accompanying the data

version 2.1.1 / 25 May 2017:
- minor bug-fixes of the document, data are unchanged

8. References

ESA MesosphEO project plan, version 1.5, 2016