

PyCAMA - Support #26241

[PyCAMA] Supply list of up to 12 parameters you want to follow over time

06/08/2020 11:33 AM - Maarten Sneep

Status:	In Progress	Start date:	06/16/2020
Priority:	Normal	Due date:	12/11/2020
Assignee:	Jacques Claas	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	PyCAMA 1.0		

Description

For the time-dependent QC monitoring, we need a list of up to 12 parameters you want to follow over time. The PyCAMA user manual (<https://dev.knmi.nl/projects/pycama/dmsf>) has an overview of what is possible, but please don't hesitate to ask if you need examples. Because long term monitoring relies on stability, this list is only needed for the offline products.

The background to this issue will be shown in the progress meeting, on Tuesday 2020-06-16.

Related issues:

Related to PyCAMA - Support #26231: [PyCAMA] Verify configuration of PyCAMA f...	New	06/16/2020	12/11/2020
Related to PyCAMA - Support #28421: Prepare for version 2 quality control mon...	New	10/15/2020	12/11/2020

History

#1 - 06/08/2020 11:34 AM - Maarten Sneep

- Related to Support #26231: [PyCAMA] Verify configuration of PyCAMA for daily extractions added

#2 - 06/10/2020 09:46 AM - isabelle de smedt

Dear Maarten

I'm not sure to understand what you need compared to the list we had before.

Do you mean that we have to reduce the list to 12 parameters?

Do we start from scratch or from the excel sheet dating from Dec 2018?

Kind regards

Isabelle

#3 - 06/10/2020 09:50 AM - Maarten Sneep

isabelle de smedt wrote:

Dear Maarten

I'm not sure to understand what you need compared to the list we had before.

Do you mean that we have to reduce the list to 12 parameters?

Do we start from scratch or from the excel sheet dating from Dec 2018?

I will explain this issue in the progress meeting next Tuesday. This is only for the time-dependent display, and is a subset of the parameters available in the extractions. I hope that after the meeting this will be clear, if not ask questions then.

#4 - 06/10/2020 11:34 AM - Maarten Sneep

- File aerosol_index_340_380_mean_median_20180701_20200601.png added
- File aerosol_index_340_380_histogram_20180701_20200601.png added
- Description updated
- Start date changed from 06/08/2020 to 06/16/2020

Examples of the types of figures we want to generate. Note that for line plots multiple traces can be specified, but make sure that they cover a similar range of values, as otherwise you just see horizontal lines. The stats are available for land and sea separately, as well as for northern and southern hemisphere (so there are 5 variants in total).

aerosol_index_340_380_mean_median_20180701_20200601.png

aerosol_index_340_380_histogram_20180701_20200601.png

Once we receive your suggestion we will prepare samples for you to check and fine-tune.

#5 - 07/02/2020 02:01 PM - Maarten Sneep

- Assignee set to Jacques Claas

#6 - 07/09/2020 08:27 AM - Steven Compennolle

From Maarten's presentation at the PMR it is said that this is an AI "For algorithm leads and algorithm specialists". However, I think that for the validation contributions, this time-dependent monitoring can also be very useful (as proven for AAI at each VAL telecon). So validation contributors should be at least aware of the selection of parameters, and perhaps can add suggestions.

Perhaps it is a good idea to share an excel sheet with suggested parameters per product?

Obvious choices are the main retrieved parameters themselves, and perhaps also the qa_value

HCHO product: HCHO column number density, slant HCHO column number density, qa_value

NO2 product: stratospheric, tropospheric and total column number density (summed), slant NO2 column number density, qa_value

CLOUD product:

- CAL: cloud fraction, cloud top height, cloud optical thickness, qa_value

- CRB: cloud fraction, cloud height, cloud albedo, qa_value

- a priori OCRA cloud fraction

FRESCO: cloud_fraction, cloud_height, qa_value

#7 - 07/09/2020 09:01 AM - Maarten Sneep

- File 20200616T1130_Msneep_L2QC.pdf added

Hi Steven: as far as I'm concerned everyone is welcome to join in on the fun, but my primary contact are the people who perform the weekly checks. I will assume that they maintain contact with the validation team for further suggestions. Note that I need more than just the parameter, I also need the aspect to be looked at (median value over time, time dependent histogram, ...). For convenience I've attached the presentation from the PMR here.

#8 - 08/25/2020 03:24 PM - Athina Argyrouli

- Status changed from New to In Progress

The suggested 12 parameters (variable names in L2 CLOUD product) that we would like to follow over time are listed below:

1. Cloud fraction apriori (cloud_fraction_apriori) – mean value
2. Cloud fraction CRB (cloud_fraction_crb) – mean value
3. Cloud fraction CAL (cloud_fraction) – mean value
4. Cloud top height CAL (cloud_top_height) – mean value
5. Cloud optical thickness CAL (cloud_optical_thickness) – mean value
6. Cloud height CRB (cloud_height_crb) – mean value
7. Cloud albedo CRB (cloud_albedo_crb) – mean value
8. Continuum Reflectance Oxygen Aband (continuum_reflectance_oxygen_Aband) – mean value
9. Reflectances ocra (reflectances_ocra) 0 (blue color) – mean value
10. Reflectances ocra (reflectances_ocra) 1 (green color) – mean value
11. Fitted RMS CRB (fitted_root_mean_square_crb) – mean value (and if possible maximum value)
12. Fitted RMS CAL (fitted_root_mean_square) – mean value (and if possible maximum value)

#9 - 08/26/2020 02:23 PM - Maarten Sneep

- File cloud_fractions.png added

Hi Athina,

I've modified PyCAMA to allow for multiple variables in a single plot. This will allow us to combine the first three in a single plot, as shown below (this is a limited time, still downloading). The a priori cloud fraction isn't available in the current extractions, we will need the reprocessing for those.

cloud_fractions.png

Similarly both cloud heights can be combined as well, and the reflectances, and the RMS values.

The `continuum_reflectance_oxygen_Aband` parameter isn't extracted in the daily extractions, and can therefore not be plotted, if you want those you'll have to update [#26231](#).

I would encourage you to also include a histogram plot, to monitor the full distribution of values. There is more to clouds than just a mean value.

#10 - 08/26/2020 03:37 PM - Athina Argyrouli

Hi Maarten,

thank you for the update! The combined cloud fractions plot looks fine.
I thought that a histogram plot is not a possibility. Then, of course including all the moments of the distribution is welcome.

#11 - 08/26/2020 04:13 PM - Maarten Sneep

Athina Argyrouli wrote:

Hi Maarten,

thank you for the update! The combined cloud fractions plot looks fine.
I thought that a histogram plot is not a possibility. Then, of course including all the moments of the distribution is welcome.

The histogram plot is similar to the plot below for AAI. We don't really extract all moments, we have a few percentiles (1, 5, 10, 16, 25, median, 75, 84, 90, 95 and 99%), mean, min, max, standard deviation, and interquartile range as possibly useful for cloud. These are collected globally, and for north, south, land and sea separately. We also have these for each row separately (for those not separated into north, south, land and sea), these can be plotted in a similar way as the histograms themselves. And we can plot the zonal mean values as a function of time, again on a colour scale.

aerosol_index_340_380_histogram_20180701_20200601.png

#12 - 09/03/2020 09:31 AM - Alba Lorente Delgado

For CH₄, we thought about the following time-dependent line plots, mean values:

- *methane bias corrected
- *methane precision
- *chi squared NIR
- *chi squared SWIR
- *qa value

For the qa value, do you think is better the histogram plot rather than just line plot?

Would it be possible to add as well histogram plot on methane zonal mean values? Is it possible to select the latitude ranges or should just be NH and SH?

Thanks in advance,

SRON L2 Team.

#13 - 09/03/2020 10:06 AM - Tobias Borsdorff

For CO we would like to be as comparable as possible with the choices for CH4.
Hence the following line plots with time would be needed:

- *mean CO (global)
- *mean CO (zonal means NH,SH)
- *CO precision
- *chi squared SWIR
- *qa value

Additionally it would be good to have a plot showing the intensity of the stripes in the CO data with time.
(Maybe the variability of the stripes with swath direction?)

#14 - 09/03/2020 11:44 AM - Maarten Sneep

Thank you.

Alba Lorente Delgado wrote:

For CH4, we thought about the following time-dependent line plots, mean values:

- *methane bias corrected
- *methane precision
- *chi squared NIR
- *chi squared SWIR
- *qa value

For the qa value, do you think is better the histogram plot rather than just line plot?

Because we only assign 1, 0.4, or 0 from the methane processor (plus a few system values that occur only rarely), I'm not sure either is useful. A plot of the fraction of pixels processed successfully, and the fraction without warnings may be more useful. We may need to tweak the plots to get daily figures rather than by orbit, but this can be done.

Would it be possible to add as well histogram plot on methane zonal mean values? Is it possible to select the latitude ranges or should just be NH and SH?

No, we don't collect that information (histogram/stats for arbitrary latitude bands), just the mean is available. the subset we have is just SH and NH. We can plot a time-dependent zonal mean. We can adjust the resolution of the zonal mean if desired.

#15 - 09/03/2020 11:48 AM - Maarten Sneep

Tobias Borsdorff wrote:

For CO we would like to be as comparable as possible with the choices for CH4.
Hence the following line plots with time would be needed:

- *mean CO (global)
- *mean CO (zonal means NH,SH)
- *CO precision
- *chi squared SWIR
- *qa value

The same remark applies here, let me know what you think.

Additionally it would be good to have a plot showing the intensity of the stripes in the CO data with time.
(Maybe the variability of the stripes with swath direction?)

Plotting the stripe-correction as a function of time is available. I suggest that we start with that, and evaluate if changes need to be made later.

#16 - 09/03/2020 11:59 AM - Alba Lorente Delgado

Thank you for the feedback.

Maarten Sneep wrote:

Thank you.

Alba Lorente Delgado wrote:

For CH4, we thought about the following time-dependent line plots, mean values:

- *methane bias corrected
- *methane precision
- *chi squared NIR
- *chi squared SWIR
- *qa value

For the qa value, do you think is better the histogram plot rather than just line plot?

Because we only assign 1, 0.4, or 0 from the methane processor (plus a few system values that occur only rarely), I'm not sure either is useful. A plot of the fraction of pixels processed successfully, and the fraction without warnings may be more useful. We may need to tweak the plots to get daily figures rather than by orbit, but this can be done.

I agree; if the fraction of successes and w/o warnings is possible, that sounds fine.

Would it be possible to add as well histogram plot on methane zonal mean values? Is it possible to select the latitude ranges or should just be NH and SH?

No, we don't collect that information (histogram/stats for arbitrary latitude bands), just the mean is available. the subset we have is just SH and NH. We can plot a time-dependent zonal mean. We can adjust the resolution of the zonal mean if desired.

Then I think SH and NH histogram plot should be fine.

#17 - 09/09/2020 09:40 AM - isabelle de smedt

Hi

Here is the wish list for HCHO, from the most important to the less, depending on the possibilities.

- qa_value: histogram (if possible, one for lands, one for oceans)
- formaldehyde_tropospheric_vertical_column: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_slant_column_corrected: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_vertical_column_precision: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_slant_column_density_window1: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_vertical_column_correction: mean, 4 horizontal lines for NH, SH, lands and oceans
- fitted_root_mean_square_win1: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_air_mass_factor: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_clear_air_mass_factor: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- integrated_formaldehyde_profile_apriori: mean, 4 horizontal lines for NH, SH, lands and oceans

in PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION, there are 2 variables with dimension 450 that I would like to monitor over time.

These fields change only once a day, and they are already averaged, but if possible, we need to keep the row dimension in a colorplot (x: time, y: row, color values).

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets: can be called formaldehyde_slant_column_correction
- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets_scd0: can be called formaldehyde_model_background_slant_column_correction

Kind regards

Isabelle

#18 - 09/09/2020 02:27 PM - Nicolas Theys

Hi Maarten,

Inspired by HCHO, here is the list for SO₂:

- qa_value: histogram
- sulfurdioxide_tropospheric_vertical_column: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_slant_column_corrected: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_tropospheric_vertical_column_precision: mean+std, 6 horizontal lines for 30° zonal bands
- fitted_slant_columns_wind1 (1st variable): mean+std, 6 horizontal lines for 30° zonal bands
- fitted_root_mean_square_win1: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_total_air_mass_factor_polluted: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_clear_air_mass_factor_polluted: mean+std, 6 horizontal lines for 30° zonal bands
-

If possible, equatorial Pacific (10°S-10°N, 160°W-240°W) timeseries colorplots (x: time, y: row, color values) for:

- fitted_root_mean_square: mean
- sulfurdioxide_slant_column_corrected : mean+std

in PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION, there are 2 variables with dimension 450 that I would like to monitor over time.

These fields change only once a day, and should be averaged over the o3_grid dimension, and displayed as colorplot (x: time, y: row, color values).

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/window1_north can be called sulfurdioxide_slant_column_correction_north
- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/window1_south can be called sulfurdioxide_slant_column_correction_south

Best regards,
Nicolas

#19 - 09/09/2020 02:48 PM - Maarten Sneepe

isabelle de smedt wrote:

Hi

Here is the wish list for HCHO, from the most important to the less, depending on the possibilities.

- qa_value: histogram (if possible, one for lands, one for oceans)
- formaldehyde_tropospheric_vertical_column: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_slant_column_corrected: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_vertical_column_precision: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_slant_column_density_window1: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_vertical_column_correction: mean, 4 horizontal lines for NH, SH, lands and oceans
- fitted_root_mean_square_win1: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_tropospheric_air_mass_factor: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- formaldehyde_clear_air_mass_factor: mean+std, 4 horizontal lines for NH, SH, lands and oceans
- integrated_formaldehyde_profile_apriori: mean, 4 horizontal lines for NH, SH, lands and oceans

I'm going to see how we combine these, as values with similar order of magnitude will give a more readable result.

in PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION, there are 2 variables with dimension 450 that I would like to monitor over time.
These fields change only once a day, and they are already averaged, but if possible, we need to keep the row dimension in a colorplot (x: time, y: row, color values).

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets: can be called formaldehyde_slant_column_correction
- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets_scd0: can be called formaldehyde_model_background_slant_column_correction

I have to check that we extract these variables, but this is otherwise a standard feature.

#20 - 09/09/2020 02:57 PM - Maarten Sneep

Nicolas Theys wrote:

Hi Maarten,

Inspired by HCHO, here is the list for SO₂:

- qa_value: histogram
- sulfurdioxide_tropospheric_vertical_column: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_slant_column_corrected: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_tropospheric_vertical_column_precision: mean+std, 6 horizontal lines for 30° zonal bands
- fitted_slant_columns_wind1 (1st variable): mean+std, 6 horizontal lines for 30° zonal bands
- fitted_root_mean_square_win1: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_total_air_mass_factor_polluted: mean+std, 6 horizontal lines for 30° zonal bands
- sulfurdioxide_clear_air_mass_factor_polluted: mean+std, 6 horizontal lines for 30° zonal bands

30 degree zonal bands aren't extracted, so we can't do that. We have SH/NH only.

If possible, equatorial Pacific (10°S-10°N, 160°W-240°W) timeseries colorplots (x: time, y: row, color values) for:

- fitted_root_mean_square: mean
- sulfurdioxide_slant_column_corrected : mean+std

This isn't extracted, nor do we have facilities to extract this at the moment. Do you have this data available via the background correction in some way?

in PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION, there are 2 variables with dimension 450 that I would like to monitor over time.
These fields change only once a day, and should be averaged over the o3_grid dimension, and displayed as colorplot (x: time, y: row, color values).

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/window1_north can be called

sulfurdioxide_slant_column_correction_north

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/window1_south can be called sulfurdioxide_slant_column_correction_south

At last, this is possible. I have to verify that we read these variables, but otherwise this shouldn't be an issue.

#21 - 09/09/2020 03:20 PM - Nicolas Theys

Hi Maarten,

30° bands: sorry I forgot it was not possible. Hemispherical (N and S) timeseries are fine.
Equatorial Pacific: no problem, I can live with it.

Nicolas

#22 - 09/09/2020 09:07 PM - Maarten Sneep

Nicolas Theys wrote:

30° bands: sorry I forgot it was not possible. Hemispherical (N and S) timeseries are fine.
Equatorial Pacific: no problem, I can live with it.

I think I can add a module to extract the mean value (row-independent) for a specified region, this might be useful for other processors as well as a background monitor. Doing the rows correctly for a region can be done, but requires quite some care (and might need to move with the seasons, as is done with offline analysis for NO₂).

If more products are going to use this, then I can add such a module.

#23 - 10/04/2020 11:56 AM - Maarten Sneep

isabelle de smedt wrote:

in PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION, there are 2 variables with dimension 450 that I would like to monitor over time.
These fields change only once a day, and they are already averaged, but if possible, we need to keep the row dimension in a colorplot (x: time, y: row, color values).

- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets: can be called formaldehyde_slant_column_correction
- PRODUCT/SUPPORT_DATA/INPUT_DATA/BACKGROUND_CORRECTION/offsets_scd0: can be called formaldehyde_model_background_slant_column_correction

We currently don't extract these, but adding them isn't hard. While adding them I looked into the values. The claimed unit is 'mol m⁻²', but the values seem to match 'molecules cm⁻²', this holds for both variables. Can you follow up on that?

#24 - 10/05/2020 08:02 AM - isabelle de smedt

Ok, thank you Marteen.
Indeed, the unit should be molec/cm². This has to be corrected.
I will transfer to the operational processor.
Kind regards
Isabelle

#25 - 10/15/2020 10:10 AM - Maarten Sneep

- Related to Support #28421: Prepare for version 2 quality control monitoring. added

#26 - 10/15/2020 10:29 AM - Maarten Sneep

- Due date changed from 09/04/2020 to 12/11/2020

#27 - 11/11/2020 02:06 PM - Klaus-Peter Heue

*Total ozone as suggested in <https://dev.knmi.nl/issues/28421> O3____.zip

*same plot for effective Albedo

*QA histogram

*global mean per groundpixel

#28 - 12/01/2020 05:34 PM - Ronny Lutz

For L2_CLOUD, 12 suggested parameters for time dependent monitoring were already indicated in this thread above (post 8).
Would it be possible to add one more variable to this list (/INPUT_DATA/surface_albedo_nir)? This variable was also added to the wish list in issue [#26231](#), but is only available since UPAS 2.1.3.

The full list would then be (with the variable names in the L2 CLOUD product in parentheses):

1. Cloud fraction apriori (cloud_fraction_apriori)
2. Cloud fraction CRB (cloud_fraction_crb)
3. Cloud fraction CAL (cloud_fraction)
4. Cloud top height CAL (cloud_top_height)
5. Cloud optical thickness CAL (cloud_optical_thickness)
6. Cloud height CRB (cloud_height_crb)
7. Cloud albedo CRB (cloud_albedo_crb)
8. Continuum Reflectance Oxygen Aband (continuum_reflectance_oxygen_Aband), has been added also to the list in issue [#26231](#)
9. OCRA color B (reflectances_ocra, index 0 is the blue color)
10. OCRA color G (reflectances_ocra, index 1 is the green color)
11. Fitted RMS CRB (fitted_root_mean_square_crb)
12. Fitted RMS CAL (fitted_root_mean_square)
13. Input surface albedo from the G3_LER retrieval (surface_albedo_nir, only available since UPAS 2.1.3), has been added also to the list in issue [#26231](#)

In the line plots, the following could be shown together in one plot: (1,2,3), (4,6), (8,9,10), (11,12).

Thanks,
Ronny

Files

aerosol_index_340_380_mean_median_20180701_20200601.png	66.3 KB	06/10/2020	Maarten Sneep
aerosol_index_340_380_histogram_20180701_20200601.png	63.9 KB	06/10/2020	Maarten Sneep
20200616T1130_Msneep_L2QC.pdf	4.06 MB	07/09/2020	Maarten Sneep
cloud_fractions.png	83.3 KB	08/26/2020	Maarten Sneep