

PyCAMA - Support #28421

Prepare for version 2 quality control monitoring.

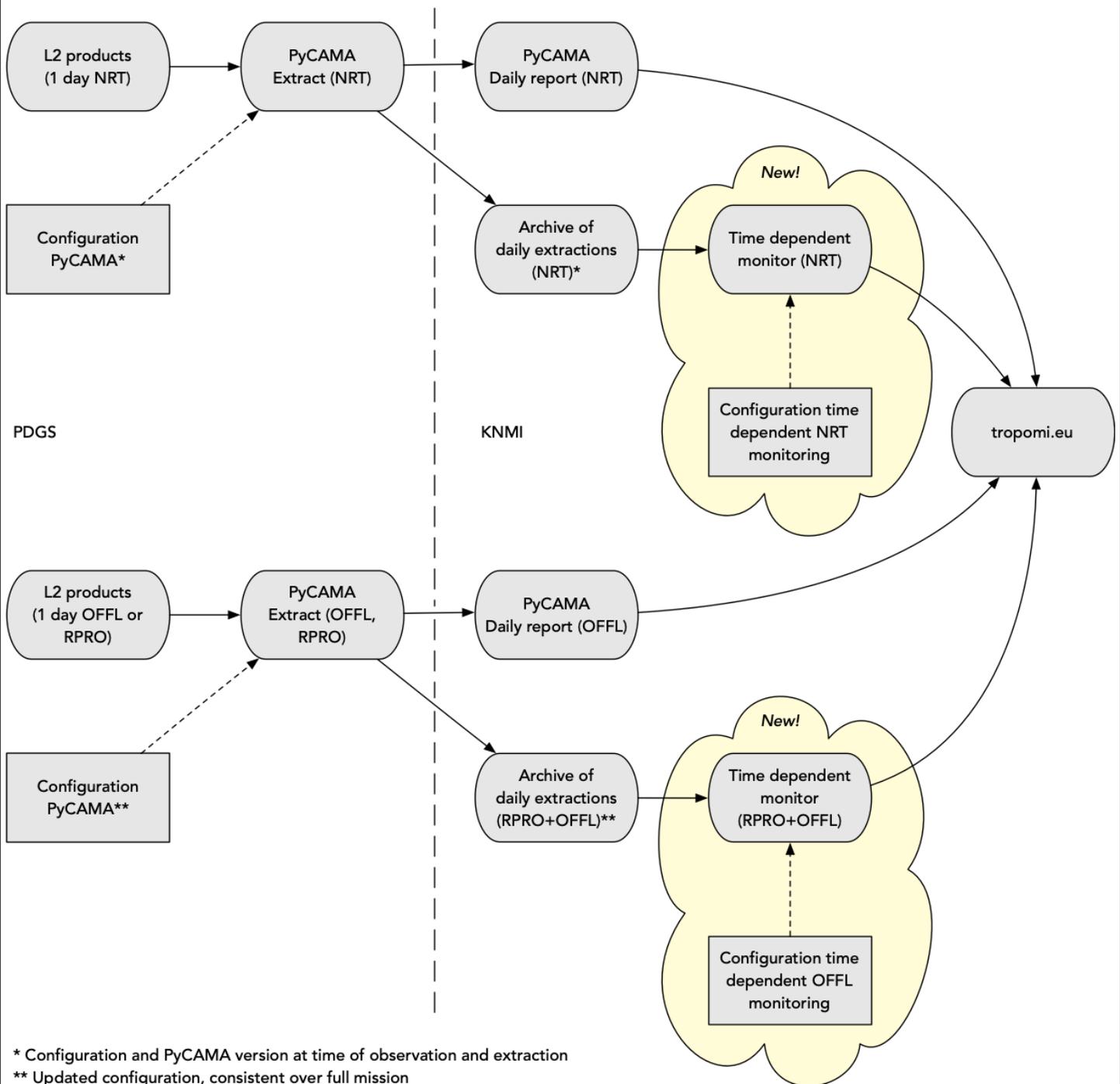
10/15/2020 10:09 AM - Maarten Sneepe

Status:	New	Start date:	10/15/2020
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 some background we refer to the flow chart below. The current level 2 quality control is done on a daily basis only. This provides a lot of detail, but it is also rather time-consuming to look into each report, and monitoring trends is tricky, to say the least.

Sentinel 5 precursor quality control monitoring flow



We (Jacques Claas, Maarten Sneep, Mark ter Linden) have been working to improve this situation, by adding time dependent monitoring facilities. These plots are based on the daily extractions. What isn't extracted can't be monitored. For the offline monitoring we will combine the reprocessed data and the ongoing forward production to have a fully consistent dataset. This also means that adjusting the settings for the daily extractions must be done this year. Any change made later will introduce a break in the record, not ideal for time-dependent monitoring.

For NRTI monitoring we have to rely on whatever was extracted at the time of observation, what wasn't extracted at the time can't be monitored. As the L2 data is no longer available, we can't run the daily extractions again. This means that either not all plots are available, or that plots aren't available for the full phase E2 of the mission. We will likely limit the plots in time, as in many cases the NRT monitoring is likely to be dominated by processor updates, especially where the NRT and OFFL algorithms differ and certainly by the upcoming L1B update. For monitoring the stability of the instrument, the offline dataset is more useful.

For simplicity these plots are generated on a daily basis using a "fixed" configuration. This configuration is different from the configuration of PyCAMA, but local and specific to the time dependent monitoring. It is not possible to request different plots yourself through the website, but for us it is fairly easy to update the configuration to have different plots the next day. However, "fairly easy" doesn't mean we're jumping for joy when a request comes in, we'd like to get this right, especially to make sure that the daily

extractions are configured correctly.

One of the important issues has to do with data filtering. At the moment all data is used, provided that all variables contain a value that is not a fill-value. So they are synchronised and not filtered on the qa-value. Either or both can be changed. The synchronisation is per product, the filtering is either per product (when synchronised) or even per variable. The possibilities of PyCAMA are pretty wild, as the same variable can even be read twice: once with and once without QA-value selection. Please get in touch to discuss the options for your product, if you need guidance.

Here are the sample figures we generated for your product, based on 7 days of DDS2B data. This represents the future situation, after reprocessing. Fine-tuning can be done on both sides here: the configuration of the daily extractions and the configuration of the time-dependent plots. If your product isn't there, then we have no idea at the moment what to monitor.

We will upload later is what we can generate from our NRTI archive. Here we have limited room to update the time-dependent monitoring plots, some extractions or plot types may not be available at all.

The issues linked below are to communicate request for changes in specific areas.

Related issues:

Related to PyCAMA - Support #26241: [PyCAMA] Supply list of up to 12 paramete...	In Progress	06/16/2020	12/11/2020
Related to PyCAMA - Support #27841: [PyCAMA] update filters for HCHO (possibl...	New	09/09/2020	12/11/2020
Related to PyCAMA - Support #26231: [PyCAMA] Verify configuration of PyCAMA f...	New	06/16/2020	12/11/2020
Related to PyCAMA - Support #28431: [PyCAMA] How long should the daily report...	Feedback	10/15/2020	12/11/2020

History

#1 - 10/15/2020 10:10 AM - Maarten Sneep

- Related to Support #26241: [PyCAMA] Supply list of up to 12 parameters you want to follow over time added

#2 - 10/15/2020 10:10 AM - Maarten Sneep

- Related to Support #27841: [PyCAMA] update filters for HCHO (possibly others) to exclude data with QA value < 0.5 added

#3 - 10/15/2020 10:10 AM - Maarten Sneep

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

#4 - 10/15/2020 10:14 AM - Maarten Sneep

- Due date set to 12/11/2020

- Assignee set to Jacques Claas

#5 - 10/15/2020 10:27 AM - Maarten Sneep

- Related to Support #28431: [PyCAMA] How long should the daily reports be available once the time-dependent monitoring is in place? added

#6 - 10/15/2020 12:20 PM - Maarten Sneep

Mark already prepared a test page with some sample images <http://mpc-l2.tropomi.eu/index-time.html#> (click 'timeseries' in the top of the window).

#7 - 11/10/2020 04:27 PM - isabelle de smedt

Dear Maarten and Jacques

Thank you for the HCHO plots.

I find the wish list from issue 26241.

Only the variable "formaldehyde_slant_column_corrected" is not clear to me, because it is called slant column correction in the plot you sent. Did you apply a formula or is it simply the formaldehyde_slant_column_corrected field?

On the line plots; there are probably too many lines, so I would suggest to remove the North and South lines, and keep only the land and sea lines that make more sense for HCHO.

And as we discussed in issue 27841, I would like to apply the qa filtering, but only on the following variables: formaldehyde_tropospheric_vertical_column, formaldehyde_tropospheric_vertical_column_precision, formaldehyde_tropospheric_vertical_column_correction and formaldehyde_tropospheric_air_mass_factor.

This is for the current plots. I will have a look at what is done for the other molecules, to get some ideas.

Kind regards

Isabelle

#8 - 11/11/2020 10:27 AM - Klaus-Peter Heue

The ozone plot looks quite nice.

However, also some line plots as for the other trace gases would be nice to have.

From both scientific and quality control point of view an interesting plot would be lines for:

*Antartica: mean and min ozone columns with latitude south of 60°S

*Artic: comparable

*Norther and southern mid latitude mean ozone columns between 60 to 20

*Tropical band (20 south to 20 north) mean ozone column

Especially for the minimum in the polar regions the qa-filter would be essential.

A separation between land and sea is not needed for ozone

Is this feasible?

Also some plots on the QA value and the RMS as for SO₂ or HCHO (figure0) are interesting time series

The surface albedo (effective_albedo for OFFL) as a function for latitude and time, comparable to the O₃ columns

The viewing angle dependency (HCHO figure10) for the total ozone columns (ideally only for the tropical band)

#9 - 11/11/2020 12:02 PM - Maarten Sneep

Klaus-Peter Heue wrote:

The ozone plot looks quite nice.

However, also some line plots as for the other trace gases would be nice to have.

From both scientific and quality control point of view an interesting plot would be lines for:

*Antartica: mean and min ozone columns with latitude south of 60°S

*Artic: comparable

*Norther and southern mid latitude mean ozone columns between 60 to 20

*Tropical band (20 south to 20 north) mean ozone column

These parameters are not available as a daily extraction, not even as a module to handle this. I'm afraid that this is not possible. We only have NH/SH as a split, plus the zonal mean.

Especially for the minimum in the polar regions the qa-filter would be essential.

We can add a qa_filter for individual parameters, please be specific to which parameters these should be added (I assume the total vertical column only).

A separation between land and sea is not needed for ozone

Is this feasible?

Some features are one size fits all, so you'll get land/sea in the daily extractions regardless. Of course, for the time-dependent plot we can select what to use.

Also some plots on the QA value and the RMS as for SO₂ or HCHO (figure0) are interesting time series

The surface albedo (effective_albedo for OFFL) as a function for latitude and time, comparable to the O₃ columns

A zonal mean can be arranged.

The viewing angle dependency (HCHO figure10) for the total ozone columns (ideally only for the tropical band)

Latitude bands are limited, as indicated above.

I have to ask you to be specific. This is a two step process, for a time-series we need to extract the parameter in the daily processor, otherwise there is nothing to plot (issue [#26231](#)). Then we have the parameters for the time-dependent monitoring (issue [#26241](#)). From your response I understand that the answer for issue [#27841](#) is 'yes, please'. May I ask you to put the requested information into the specific issues, as several parts will be handled by different people.

#10 - 11/11/2020 12:05 PM - Maarten Sneep

isabelle de smedt wrote:

Dear Maarten and Jacques

Thank you for the HCHO plots.

I find the wish list from issue 26241.

Only the variable "formaldehyde_slant_column_corrected" is not clear to me, because it is called slant column correction in the plot you sent. Did you apply a formula or is it simply the formaldehyde_slant_column_corrected field?

I have to double check. Will respond in [#26241](#).

On the line plots; there are probably too many lines, so I would suggest to remove the North and South lines, and keep only the land and sea lines that make more sense for HCHO.

This can be done.

And as we discussed in issue 27841, I would like to apply the qa filtering, but only on the following variables:
formaldehyde_tropospheric_vertical_column, formaldehyde_tropospheric_vertical_column_precision,
formaldehyde_tropospheric_vertical_column_correction and formaldehyde_tropospheric_air_mass_factor.

Yes, I'm working on that. Still encountering some issues there.

This is for the current plots. I will have a look at what is done for the other molecules, to get some ideas.

Very good, that is part of the idea for these plots here.

#11 - 11/12/2020 01:26 PM - Christophe Lerot

For ozone OFFL, after discussion with Klaus-Peter, we would like to have time series plots for the following variables:

- ozone_total_vertical_column
- ozone_total_vertical_column_precision
- fitted_root_mean_square
- effective_albedo
- ozone_effective_temperature
- ozone_ghost_column

qa_values filtering can be applied to all of them.

It would be nice to have both color zonal/time plots (as the example in the 03.zip) and time lines for all of them. For the time lines, both the land/sea and hemisphere separations can provide complementary information.

Thank you

#12 - 11/17/2020 02:40 PM - Nicolas Theys

SO2:

- Some graphs have too many lines. I would skip the "mean" curves and keep only "medians"
- For the plots where two variables are shown (eg corrected SO2 SCDs and SO2 SCD window1), would it be possible to keep the same colors (eg. North and North, South and South) and use solid and dashed lines instead. It is easier to read and compare.
- QA > 0.5 filtering should be applied to plots for variables VCD and VCD precision, for the rest it is fine.
- Figure SO2__timeseries-11.png looks odd. Here the filled values were not filtered out?
- Figure SO2__timeseries-12.png could be improved. I would use the color scale from 1e-2 to 1e1.
- Figure SO2__timeseries-13.png and Figure SO2__timeseries-14.png are empty.

#13 - 12/01/2020 05:39 PM - Ronny Lutz

Thanks for providing the plot examples for L2_CLOUD in the CLOUD_.zip.

Some remarks for each plot below:

CLOUD_-timeseries-0.png:

(this shows the variables cloud_fraction_apriori from OCRA, cloud_fraction from ROCINN_CAL and cloud_fraction_crb from ROCINN_CRB)
Would it be possible to omit the sea/land lines? This way the number of lines would decrease from 18 to 6 and the plot would look much less crowded?

If this is not possible, could maybe different linestyles be used, e.g solid for cloud_fraction_apriori, dashed for cloud_fraction and dotted for cloud_fraction_crb?

CLOUD_-timeseries-1.png:

(this shows the variables cloud_top_height and cloud_height_crb)

As above, could maybe different linestyles be used for the two variables, e.g. solid for cloud_top_height and dashed for cloud_height_crb?

CLOUD_-timeseries-2.png:

(this shows the variable cloud_optical_thickness)

Plot looks fine

CLOUD_-timeseries-3.png:

(this shows the variable cloud_albedo_crb)

Plot looks fine

CLOUD_-timeseries-4.png:

I assume the OCRA "blue" and "green" are taken from reflectances_ocra (index 0 for blue, index 1 for green)?

Is ROCINN "red" taken from continuum_reflectance_oxygen_Aband? If not, please take it from there. In [#26231](#) I have added this to the list of parameters for the daily extraction.

To make the plots more intuitive, could the linestyle and linecolors be adapted to e.g.

- solid lines for mean, dashed lines for median,

- red lines for ROCINN "red", green lines for OCRA "green" and blue lines for OCRA "blue"?

CLOUD_-timeseries-5.png:

(this shows the variables fitted_root_mean_square and fitted_rootmean_square_crb)

Plot looks fine

CLOUD_-timeseries-6.png:

(same as 5, but showing max, q99 and q95)

Plot looks fine

CLOUD_-timeseries-7.png:

This shows the cloud_fraction from ROCINN_CAL I assume? Will these plots be generated also for cloud_fraction_apriori and cloud_fraction_crb or do we have to choose one of the three cloud fractions?

Otherwise, the plot looks fine

CLOUD_-timeseries-8.png:

This shows the cloud_top_height from ROCINN_CAL. Will these plots be generated also for cloud_height_crb or do we have to choose one of the two cloud heights?

Otherwise, the plot looks fine

CLOUD_-timeseries-9.png:

This row plot shows the cloud_fraction from ROCINN_CAL I assume? Same comment as for plot 7, will these plots be generated also for cloud_fraction_apriori and cloud_fraction_crb or do we have to choose one of the three cloud fractions?

Since the mean cloud fraction will be somewhere around 0.5, the colorbar ranges could maybe be reduced from [0,1] to e.g. [0.2,0.8] to highlight more the variations.

CLOUD_-timeseries-10.png:

This row plot shows the cloud_top_height from ROCINN_CAL. Same comment as for plot 8, will these plots be generated also for cloud_height_crb or do we have to choose one of the two cloud heights?

Since the mean cloud height will be at lower levels, maybe the colorbar range could be reduced from [0,15000] to [0,10000] to highlight more the variations.

CLOUD_-timeseries-11.png:

I'm not sure where these numbers come from, but can the number of warnings (green) be higher than the total number of processed pixels (black)?

Or can a single pixel have several warnings? Maybe good to check with the UPAS people?

CLOUD_-timeseries-12.png:

Zonal plots for cloud_fraction look fine. 5 degree resolution seems ok. Same comment as for plots 7 and 9: Will these plots also be generated for cloud_fraction_apriori and cloud_fraction_crb?

CLOUD_-timeseries-13.png:

Zonal plots for cloud_top_height look fine. 5 degree resolution seems ok. Same comment as for plots 8 and 10: Will these plots also be generated for cloud_height_crb?

CLOUD_-timeseries-14.png:

Plot looks fine

Thanks,
Ronny

Files

S5P-L2-Flow-L2QC.png	188 KB	10/15/2020	Maarten Sneep
AER_AI.zip	969 KB	10/15/2020	Maarten Sneep
AER_LH.zip	158 KB	10/15/2020	Maarten Sneep
CH4____.zip	1.13 MB	10/15/2020	Maarten Sneep
CLOUD_.zip	2.25 MB	10/15/2020	Maarten Sneep
CO____.zip	1.22 MB	10/15/2020	Maarten Sneep
FRESCO.zip	1.54 MB	10/15/2020	Maarten Sneep
O3____.zip	98.7 KB	10/15/2020	Maarten Sneep
HCHO_.zip	1.76 MB	10/15/2020	Maarten Sneep
SO2____.zip	2.11 MB	10/15/2020	Maarten Sneep