Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC2
https://doi.org/10.5194/amt-2022-127-AC2, 2022

Reply on RC2

Simon Whitburn et al.

Author comment on "A CO_2-independent cloud mask from Infrared Atmospheric Sounding Interferometer (IASI) radiances for climate applications" by Simon Whitburn et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2022-127-AC2, 2022

We would like to thank the reviewer very much for its constructive feedback on the paper and its useful comments and suggestions, which all have been answered or addressed and significantly improved the paper. Point-by-point responses are provided below.

The reviewed manuscript presents a cloud detection algorithm for the hyperspectral spaceborne IASI instrument(s). The method uses a neural network approach and uses only IASI radiances as an input. The authors paid specific attention to avoid the channels affected by trace gases with the concentration variable over time. When applied to a series of IASI measurements, the method shows physical results and its robustness is demonstrated over the whole period of all MetOp/IASI instruments' observations up to now.

The research is topical, the methodology presented in the article is sound, and the paper is well structured and written (except for some minor issues listed in "Technical corrections" section). Still, there are several issues I'd like to clarify/fix before recommending the manuscript for publication.

I have chosen "major revision", but the changes I suggest are easy to implement. I believe, if the authors add the suggested information to the manuscript, it will become irreproachable from the methodological point of view and it will have a broader impact.

General comments:

- Even though the average cloud amount shown in Fig. 5 looks physical and reasonable, I see a general methodological issue in using IASI L2 cloud product for training the neural network. I do not question the quality of this product – as follows from lines 64–79 of the manuscript, the methodology is mature and the results are generally good, but there is one caveat. I would not hesitate if the training were based on some "ground truth" dataset coming from in situ measurements, ground- or space lidar, or some other instrument, but I see an inconsistency in using IASI itself as a reference, given that its time series (Fig. 6) shows certain artefacts in cloud cover. I do not ask the authors to redo the whole work, but it would be good to supplement it with some kind of validation of the training dataset using, e.g. CALIPSO clouds as a reference. To avoid the diurnal effects related to different overpass times, one can focus only on the clouds over the ocean. Perhaps, it would be sufficient to show several representative profiles of the training dataset and to compare them to overlapping CALIPSO cloud profiles (e.g. Chepfer et al., 2013). Or, better yet, show the 3-month average of the IASI L2 and CALIPSO (3 months are required to get a full spatial coverage from CALIPSO). One can also use GEWEX Cloud Assessment files (https://climserv.ipsl.polytechnique.fr/gewexca/index-2.html) for the comparison, but in this case IASI L2 cloud product should be processed in accordance with GEWEX CA rules (Stubenrauch et al., 2013). Either way, this cross-validation of the training dataset seems necessary to wrap up the methodological part.

Indeed, not a lot of details were given in the manuscript to justify the use of the version 6.5 of the IASI L2 cloud product as a reference, as also pointed out by the referee#1.

It is true that the IASI NN cloud product was not validated against CALIOP. However, the L2 cloud product is routinely monitored and validated against CALIOP data since 2019 and public monitoring reports are available at *https://www.eumetsat.int/iasi-level-2-geophysical-products-monitoring-reports.* The long-term statistics are updated each month, by appended new matchups into the evaluation. The current version of the L2 cloud product is the result of many years of improvement to reach a mature state today. As it can be seen on Fig. 2.3 of the last report (August 2022), the release of the version 6.5 has brought a significant improvement in the cloud detection against version 6.4. Hanssen Kuiper's skill score and the Percent Correct comparing the L2 cloud mask and CALIOP have increased of about 35% and 15%, respectively (to reach about 75% and 87%).

Here, the goal of the proposed algorithm is to produce an IASI cloud mask matching as best as possible the version 6.5 of the IASI L2 cloud product but consistent on the whole IASI time series (as the complete L2 cloud product has not been officially reprocessed yet with the latest version of the L2), and not affected by the changes in the atmospheric CO_2 (and the other long-lived species) concentrations. The good performance reached by the training and the other intercomparisons performed with the L2 cloud product show that the two products match very well and, therefore, the level of correspondence with CALIOP data should also be close to the one of the L2.

The manuscript has been updated to emphasize the good correspondence between CALIOP and the IASI L2 cloud product:

Lines 96-102: *The goal of the proposed algorithm is to produce a sensitive and consistent (unbiased) cloud mask over the entire IASI lifespan using as a reference dataset the version 6.5 (v6.5) of the operational IASI L2 cloud product (August et al., 2012). The latter shows a clear improvement over the previous version (v6.4) when compared to the cloud products from the Cloud-Aerosol LIdar with Orthogonal Polarization (CALIOP) onboard the Cloud-Aerosol LiDAR and Infrared Pathfinder Satellite Observation (CALIPSO, Winker et al., 2007), as reported in the CALIOP-CALIPSO IASI Level 2 geophysical products monitoring reports available at https://www.eumetsat.int/iasi-level-2-geophysical-products-monitoring-reports (accessed online on August 10, 2022). The retrieval method presented here is based on a supervised NN relying on the IASI radiance spectra only, […].*

- As far as I understand, the present requirements to articles published in EGU open access journals include the distribution of the codes and/or the data used in the article. To my knowledge, the trained neural network of a kind applied in the manuscript can be represented by a couple of pages of ASCII-text in pseudocode (variables of the first layer = linear combination of the variables of zero-layer, second layer =…, …, result= linear combination of the variables of N-1th layer). It can be added to the manuscript itself or be provided as a supplement, but it should be certainly doable and it will be useful for the community.

This is indeed a good suggestion. We have generated a MATLAB function (provided in ASCII-text format) allowing to easily run the network on MATLAB and containing all the training variables (see "supplement.zip" in attachment here). An example code with one input (45 IASI channels and the orography) is provided as well.

- The contribution functions for the channels centered at the same or close wavelengths for IASI and AIRS should be close to each other, see e.g. (Feofilov and Stubenrauch, 2017) mentioned in the manuscript. Correspondingly, I'm almost sure that the neural network trained for IASI will be applicable to AIRS and maybe even to other instruments like HIRS. It would be good to apply the NN to AIRS L1 data to show the potential and versatility of the method. Just a single map in the appendix would generalize the approach and the NN file explained in comment #2 will enable the other researchers to calculate their own cloud mask on the fly.

Indeed, this would have been a great addition to the paper. Unfortunately, not all the wavelengths used in the neural network to detect the presence of clouds are covered in the AIRS spectra (in particular, 13 channels around 2100 cm^{-1} fall in the gap present in the AIRS spectrum between 1650 cm^{-1} and 2175 cm^{-1}). It would of course be possible to develop an algorithm for the AIRS measurements using the same approach but it would then require to select a complete new set of AIRS channels and to train a new neural network. While this would be interesting, it is clearly out of the scope of this paper.

Specific comments:

- Lines 105−120: this text could be significantly simplified if Fig. 1 were supplemented with the vertical contribution functions (averaging kernels) for each channel. One can put these curves side by side, on top or at the bottom of Fig. 1. The actual scale is not that crucial, but the center and the halfwidth of the averaging kernel should be clearly visible. This is an important methodological point, which would explain the information content of the signals used in the approach. The NN itself has enough "black box" features, so anything that could be clarified should be clarified.

It is true that by taking spectral regions affected by H_2O, O_3 and CO, the selected channels cover a range of altitude of vertical sensitivity comprised between 0 and about 30 km in clear sky conditions. However, the focus is on the detection of cloud scenes and the vertical contribution functions will have completely different shapes in the presence of clouds. Given this, we therefore prefer not to include the vertical contribution functions in the Figure 1 as we think it can be misleading. Nevertheless, as many of the input channels are located in the atmospheric windows, we can confidently assume that the network will be sensitive to clouds at any altitudes since the atmosphere at these wavelengths is transparent in the absence of clouds. We have clarified this in the manuscript:

Lines 123-125: […] *Note that, as most of the selected channels are located in the atmospheric window regions, the network should be sensitive to clouds at any altitudes since the atmosphere at these wavelengths is transparent in the absence of clouds.*

- Lines 130−132: I'm not sure I've got the idea here. Normally, the thresholds of this kind should be selected basing on the minimization of an error (or maximizing of the correlation coefficient) for two datasets. That's what is written in lines 134−136 below, and I agree with this approach. I'd suggest to leave only this part in this paragraph since the beginning is misleading.

Indeed, the selected threshold is the one that minimize the difference between the IASI NN and the L2-derived cloud amount. The first sentence (Lines 130-132) was added only to emphasize that this threshold would be equal (or close) to 0.5 if the number of cloud-free and cloud scenes were equivalent on Earth, as the neural network was trained with half of the training set containing clouds and the other half without. We think this is interesting to be mentioned.

It would be also interesting to have a look at the difference curve mentioned in line 135 to estimate the uncertainty of the threshold, but the authors can just do it themselves and provide a ±value along with the threshold.

This is indeed a good suggestion. We have estimated an uncertainty corresponding to an increase of 1% in the difference between the L2 and the NN for land and sea observations separately. We have adapted the manuscript to include this:

Lines 146-150: *A separate threshold, of respectively 0.175 (± 0.020) and 0.275 (± 0.015), was defined for land and for sea measurements. Those are recommended when the NN cloud product is used for cloud removal preprocessing phase in satellite data retrieval schemes. However, as we demonstrate in Sect. 4.1, this threshold can be adjusted depending on the application. The uncertainty on the thresholds is estimated by evaluating the change in the threshold for a 1% increase in the difference between the L2 and the NN.*

- 2: In the left-hand side panel, I do not see the cloud structures of the right-hand side one. There are a lot of yellow circles in a cloud-free (?) area in the lower left section of the image. Perhaps, the light "haze" which one can see in the right-hand side corresponds to a real cloud, but it is not clear from the image. What is the correlation coefficient for these two panels and what is the r.m.s. of their difference?

The Figure 2 was given as an illustration more than for validation purposes. In fact, cloud coverage comparison with MODIS (Terra) is only possible for large cloud structures. In the region that you mentioned, the sky seems characterized that day by thin cirrus and sparse cumulus clouds. Because of the high sensitivity of the IASI NN cloud mask algorithm, an IASI field-of-view will be declared as cloudy even if only a small fraction of the pixel is occupied by a cloud or in the presence of a very thin cloud. In these specific cases, the comparison between MODIS and IASI is very difficult as the two instruments have different overpass times (10h30 vs 9h30) and clouds distributions can evolve rapidly (transport, precipitation, evaporation). We have adapted the manuscript to clarify this:

Lines 160-165: […] *The MODIS Terra corrected reflectance imagery for the same day is shown as well. An excellent correspondence is found for the large structures of high opaque clouds and the cloud-free regions (e.g. the North and South Africa, the Arabian Peninsula, …). For the regions characterized by sparse cumulus or thin cirrus clouds, the comparison is more difficult because of the different overpass time of the two instruments (10.30 AM/PM for MODIS Terra and 9.30 AM/PM for IASI) as the spatial distribution of clouds can evolve very rapidly (due to evaporation, precipitation and transport).*

- Lines 175−180 : I wonder if the NN training with CALIPSO would improve the agreement.

The agreement would potentially be better if CALIOP was used as the reference

dataset. However, it would not be possible to develop a global cloud mask product using CALIOP as collocations with IASI are only possible at high latitudes because of the different overpass time (about 4 hours in the tropics) of Metop and CALIPSO (CALIOP-CALIPSO IASI Level 2 geophysical products monitoring reports, available at *https://www.eumetsat.int/iasi-level-2-geophysical-products-monitoring-reports*).

- Lines 280−289 : as in general comment #1, I stumble here because there is a certain issue in the source dataset used for training, and at the same time we have a NN based on this dataset, which doesn't have this issue. I understand that this is possible, but one has to discuss this inconsistency because methodologically the neural network is not supposed to be different from the training dataset.

We think there might be a misunderstanding here. The discontinuities in the IASI L2 time series are the result of version changes in the retrieval algorithm (as the entire time series has not officially been reprocessed yet). As we mention in the manuscript (line 95), the neural network was trained only with output data from the version 6.5 of the IASI L2 cloud product (year 2020, which shows an excellent match with the NN in Fig. 6) and it was therefore expected that we obtain a consistent time series for the whole IASI period. To avoid a spurious trends over time, we also took care to avoid, for the input parameters of the NN, any channel that is affected by the absorption of changing greenhouse gases (lines 97-100). As the period considered for the training is short (one year), the existence of a trend in the L2 due to CO_2 concentration changes in the atmosphere is not affecting the training.

- 5: It is not clear whether the average cloud cover here was calculated considering "shrinking" of the lat/lon box when moving towards the poles. I made this exercise for Fig. 5d (see below) and I've got close values (67.0% and 66.9% for area-weighted and simple averages, respectively), but the differences might be larger for the instruments with better coverage of the polar areas. In any case, it is recommended to use the area-weighted values. Please, check.

The average cloud cover provided were indeed simple averages. We have replaced them by the corresponding area weighted values. Differences are not very large (about 1-2%), except for the AIRS L2 (from 87% to 94%) and the PATMOS (from 75% to 69%) cloud coverage.

- 5: this is more a comment rather than an issue and I do not require the authors to squeeze another panel to Fig. 5, which is already busy, but I think that this plot is worth providing here. In Fig. 1 below, I show the mean CIRS-LMD IASI and mean CIRS-LMD AIRS for 2015. As one can see, they are quite similar because the channels of (almost) the same wavelength were processed with the same methodology. The remaining difference is due to diurnal variation (Feofilov and Stuberauch, 2019). What is important, the similarity of these plots indirectly proves the point made in general comment #3.

Thank you for this very interesting remark. Unfortunately, as we explained in response to general comment 3, it is not possible to directly apply our network to the AIRS measurements because not all wavenumbers used in the NN are available in the AIRS spectrum.

- Line 305 : please, mark these areas on the maps in Fig. 7.

We have updated the Figure 7 to clearly show the IASI pixels that are flagged as

cloudy by the L2 and clear by the NN in presence of high dust loads (dust index > 10) and we have adapted the manuscript accordingly:

Lines 317-320: […] *For each of them, we plotted the cloud flag derived from the L2 (left) and from the NN (middle column). On top, we also displayed the contour plot of the dust index for two different levels (index of 10 and 20, respectively).* ***For the L2, the IASI pixels seen as cloudy by the L2 but as cloud free by the NN in presence of high dust loads (dust index higher than 10) are shown in pink.***

And Line 331: *In the area, the NN (middle panel) correctly flag the IASI measurements as cloud-free while the L2 reports the presence of clouds for most of the pixels (**pink circles on Fig. 7**).*

The figure is provided in attachment here (see "supplement.zip").

- Lines 345−346 and 125-129: please, provide the information on training time and number of samples.

The size of the data set is mentioned line 127 and the retrieval time for one day of IASI measurements is mentioned line 169-170. We have added a sentence about the training time in the manuscript, line 137-138:

[…] *In total, we performed 10 different trainings and we selected the least affected with dust (see Sect. 4.3).* ***The training of the network takes, depending on the run, about 100-150 iterations and is completed in about 30 minutes on a typical personal computer.*** *The performance of the selected training reaches 87.3% with an equivalent number of misdetections in the clear and the cloud group.*

Technical corrections:

Thank you for pointing out these technical mistakes. All have been corrected.

- Line 18: please, change "on the weather" to "for the weather"
- Line 21: either "in detection … and in derivation" or "to detect … and to derive"
- Line 66: "performance" would be better here
- Line 75: "retrieves"
- Line 88: comma is missing after "In the next section"
- Lines 247, 255, 270, and elsewhere: the terms "more conservative" and "less conservative" require too much thinking, especially if clear sky attribution is mixed up with cloud fraction in one sentence. I would just write "assigns clear sky flags more (or less) often than …" to avoid any ambiguity.
- Lines 341−342: please, reformulate to make a distinction between AVHRR products and CIRS-LMD IASI product. In the current version, the sentence reads as if they are in the same category. As far as I understand from the manuscript, this is not the case.

References used:

- Chepfer H., G. Cesana, D. Winker, B. Getzewich, and M. Vaughan, 2013: Comparison of two different cloud climatologies derived from CALIOP Level 1 observations: the CALIPSO-ST and the CALIPSO-GOCCP, J. Atmos. Ocean. Tech., doi.10.1175/JTECH-D-12-00057.1
- Feofilov, A. G. and Stubenrauch, C. J.: Diurnal variation of high-level clouds from the synergy of AIRS and IASI space-borne infrared sounders, Atmos. Chem. Phys., 19, 13957–13972,
- Stubenrauch, C., and 22 co-authors, Assessment of Global Cloud Datasets from Satellites: Project and Database Initiated by the GEWEX Radiation Panel, Bull. Amer.

Meteorol. Soc., 94(7), 1031-1049, doi:10.1175/BAMS-D-12-00117.1, 2013

Please also note the supplement to this comment:
https://amt.copernicus.org/preprints/amt-2022-127/amt-2022-127-AC2-supplement.zip