

Line 241:

I would begin the sentence with, “In the middle panel, at TOD = 0.25,:::”

This clarification has been made.

Line 252:

I’m not sure that “robust” is the word that should be used here. Perhaps “effective” would be a better choice. Wait at least until the validation results are discussed!

Agreed. Switched “robust” with “effective”.

Line 277:

What is “B7”? Please insert a few descriptive words.

B7 refers to the Build number of the OCO-2 operational algorithm. We decided that this information is not strictly necessary/informative for the current study so we removed reference to it.

Line 292:

It becomes clear later in the text, but please indicate here what it means to “pass” the cloud flag (is a given pixel is clear or cloudy if it passes?)

Modified the sentence to read;

An analysis of the full OCO-2 data set spanning 06-Sep-2014 to 01-Aug-2015 (orbit numbers 958 to 5762) showed that, on average, slightly more than one third (36%) of the nadir-land soundings pass the ABP operational cloud flag, i.e. are identified as clear, on a per-orbit basis.

Line 360:

Are the bins 2x 2 or 4x 4 as indicated in the figure caption?

This typo was corrected to 4x4.

Line 412:

Since the quantity Throughput is a percentage, shouldn’t the eqn. for it include a “/Ntotal”?

The omission of the divisor has been corrected.

Line 429:

“expensive” should be “expense”

Typo corrected.

Line 435:
“sounding” should be plural

Slight rewording of this sentence to put in singular case;

It is crucial that as many of the scenes as possible are correctly classified, while limiting the number of false negative cases (MODIS clear, OCO-2 cloudy), as once a sounding has been identified as cloudy by either ABP or IDP, it will not to be run in the operational L2 X_{CO_2} retrieval algorithm.

Line 483:
I don't agree that TNR>TPR means that the OCO-2 cloud mask is more sensitive to cloud than MYD35. It might be the case but I believe all it means for sure is that the OCO-2 cloud detection thresholds are set “tighter” than those of MYD35.

Agreed. The discussion was modified to;

This indicates that the OCO-2 cloud screening algorithms, as configured in this study, are more aggressive, i.e., use more stringent filtering thresholds, than does MODIS. This makes sense, as OCO-2 is sensitive to both clouds and aerosols, while the MODIS cloud mask product identifies only water and ice clouds.

Line 556:
Please include a few words explaining why the cloud screening thresholds are different for the CALIOP comparison.

Modified sentence to;

The cloud screening thresholds were set to similar, but not identical, values as those reported in Sect.~\ref{sec:threshold_optimization} in order to provide a throughput of $\geq 30\%$.

Figure 1:
Please label panel a and b.

Added labels to the figure as suggested.

In addition to the corrections given above, we have corrected what turned out to be a minor deficiency in the analysis related to the ABP filtering criteria. In the analysis code, the filtering on the chi-squared and albedo parameters had been inadvertently disabled. The overall effect to the results was minor, the primary one being a reduction in the number of reference clear scenes being predicted clear for glint-land scenes due to the surface albedo filters. This yields a loss in throughput but an increase in fractional agreement with MODIS. We have added some brief description of the strength of the individual filters in Sect. 2.3.

In addition, the suggested surface pressure threshold for glint-water has been revised from 75hPa to 20hPa to be in-line with the nadir-land and glint-land filtering criteria. The main effect is a slight decrease in the throughput.

The values in Table 1 and the statistics in Table 3 and Fig. 6 have been updated accordingly.