Abstract, page 1, line 21.
After reading the full paper, I think I understand what was implied by
the cryptic sentence. Suggest change to: "With tuning of algorithmic threshold
parameters that allows for processing of ~30% of all measurements,..."

Modified to read;

> With tuning of algorithmic threshold parameters that allows for processing of ~
> 20-25\% of all OCO-2 soundings,
> agreement between the OCO-2 and MODIS cloud screening methods
> is found to be ~ 85\%
> over four 16-day orbit repeat cycles in both the winter
> (December) and spring (April-May) for
> OCO-2 nadir-land, glint-land and glint-water observations.

Page 2, lines 45-51.
Mention that each footprint is associated with spectra for all three OCO2
bands, and reference the March 2015 ATBD, plus citing appropriate Figure(s).

Modified to read;

> Each footprint contains a single sounding,
> comprised of spectra for all three OCO-2 bands.
> Further details of the instrument and satellite viewing modes can be found in
> Sections 2.2 and 2.3 of
> \citep{OCO_L2FP_ATBD_B6_2015}.

Page 2, line 54.
Explain (perhaps in a separate sentence on line 56) that optical thickness refers to
gas+cloud+aerosol, and that "cloudy" refers to cloud plus aerosol in the paper.

Added the following discussion;

> In this work the definition of optical thickness
> includes the contribution from aerosols,
> as well as from both ice and water clouds,
> except where noted.
> Therefore, for OCO-2,
> labeling a scene as cloudy,
> indicates the detection of either cloud or aerosol, or both.

Page 3, line 81.
What paper does "this validation study" refer to?

Changed wording to;

However, the current validation study addresses only the global nadir and glint mode data.

Page 3, line 82.
Change to "reported in Taylor et al., (2012)"

See above.

Page 4, line 106.
Change to "Since OCO-2 collects almost"

Corrected typo.

Page 4, line 126.
Define subscript s and a.

Added sentence;

Here, the subscript s refers to the surface,
while a refers to a priori.

Page 5, line 147.
Change to "light to the OCO-2 radiance. Secondly, the 1.6 μm and 2.0 μm band strengths"

Modified discussion to;

This yields a variable fractional contribution of scattered light to the OCO-2 radiances.
Secondly, the 1.6\,μm and 2.06\,μm band strengths are highly variable, resulting in different sensitivities to atmospheric scattering.

Page 6, line 191.
Why are two thresholds displayed (instead of just one threshold)?

Added the following description;

The $\Delta p_{s,cld}$ test is two sided;
deviations from the ECMWF a priori, either high or low, will cause the scenes to be flagged as cloudy.

Page 7, line 206.
Change to "quicker fall of in the ILS wings"

Typo corrected.

Page 7, lines 212-215.
**This sentence is difficult to read. Rephrase (and then have someone else
also read the sentence to make sure it is clear).**

Rephrased the discussion as;

In the absence of scattering the respective CO_2 and H_2O ratios should converge
to unity as the light path distributions in the strong and weak bands will be identical,
irrespective of differences in surface albedos. For cases with larger TOD, however, the
light path distributions will differ between the bands, resulting in ratios that deviate from
1. We found that the ratios almost exclusively deviated in the positive direction, meaning
that the PPL in the weak band was larger than in the strong band. This is most likely a
consequence of generally lower surface albedos in the strong band as well as higher
aerosol sensitivity owing to nearly-saturated absorption lines.

Page 7, line 220.
**The phrase "significant albedo contrast" is not clear. Is "when there is
significant differences in the surface albedos of the two CO2 bands
" the intended meaning?**

Modified to;

In other words, when there are significant differences in the surface albedos of the
two CO$_2$ bands, the IDP has higher fidelity in identifying contamination by
cloud and aerosol.

Page 7, line 226.
**Use the information from Page 8, line 238, to define what is meant by low and high
clouds.**

Added the sentences;

Here, high (low) cloud is defined as cases where 95\% of the TOD resides in the
top 40\% (bottom 30\%) of the atmosphere. About 4\% and 18\% of the
soundings were classified
as high cloud and low cloud cases, respectively.

Page 7, line 226.
**In a separate paragraph, identify in words and meaning, the various screening
variables.**

Added the following discussion;

Details of the ABP $\Delta p_{s,cld}$ and χ^2 parameters
can be found in Sect. III.C. of
\cite{Taylor:TGRS:2012:GOSAT_cloud_screen_comp}.
In summary,
$\Delta p_{s,cld}$ detects changes in the retrieved vs. a priori
surface pressure brought about by scattering-induced PPL modification.
The multiplicative χ^2 scale factor allows the dynamically
calculated χ^2 threshold to be scaled.
Setting this parameter near unity indicates high confidence in
the instrument calibration and spectroscopy,
while very large values (say 20 or greater) effectively disables this test.
Moderate values, like those used in this study,
cause highly contaminated soundings to be screened,
but puts most of the burden on the surface pressure check.

The IDP R_{CO_2} and R_{H_2O} center and halfwidth values,
which are empirically determined,
simply define the acceptable range of R_{CO_2} and R_{H_2O}.
Soundings with calculated R_{CO_2} and R_{H_2O}
that fall outside the allowed range are flagged as cloudy.

Page 7, line 232.
If you know which type of features trigger the two algorithms, provide this information.

The use of the phrase "…are triggered by different features in the spectra…" was ambiguous and unsubstantiated. It has been replaced with;

This indicates that ABP and IDP are not flagging identical soundings
and are therefore complimentary.

Page 8, line 233.
Have you numerically determined this fact? If so, provide a little more information.

The sentence was replaced with;

The smooth curves indicate that all three cloud screening combinations
(ABP-only, IDP-only and ABP+IDP) exhibit a smooth decay
toward zero fraction passing with increasing TOD.

The data in the graph speaks for itself so we are not sure exactly what the reviewer's are asking. Hopefully the above sentence clarifies our meaning.

Page 8, line 235, and Figures 3 and 7.
The "hump" near forward model and/or CALIOP optical depths of 3 is very odd. If you have a good explanation for this, please provide it when Figure 3 is first discussed.

We do not currently have a good explanation for this feature that appears in both the simulation data set (which is created from real CALIOP profiles) and the CALIOP data set. We added the following discussion;

> As mentioned previously, the histogram (gray shading)
> indicates that there is a large number of scenes with TOD $\simeq$ 3.
> This feature also appears in the real CALIOP data to be presented in
> Sect.\ref{sec:comparison_to_calipso}.
> This odd feature in the data set is not currently understood.

Page9, line 279.
Renumber Figure 1 to Figure 3 (and renumber 2 and 3 to 1 and 2). It does not make sense to encounter Figure 1 in the text after first discussing Figures 2 and 3.

We agree. These corrections were made during the typesetting process for the on-line AMTD version.

Page 12, line 384.
The pc may be missing an – overscore mark.

This typo has been corrected.

Page 12, line 392.
The usage of "cloudy" to include cloud plus aerosol becomes problematic, since my mindset in reading the previous pages reverted to thinning of clouds as water and/or ice cloud particles, not micron sized aerosol particles. The use of the Angstrom coefficient implies that you are referring to the aerosol index (i.e. AOD x Angstrom coefficient).

Point well taken. The OCO-2 cloud screening algorithms detect both cloud and aerosol. Throughout the paper the term "cloudy" generally refers to contamination by either cloud or aerosol. Some text has been added throughout the paper in an attempt to make this point more clear.

For example, in the Abstract;

> These estimates can be biased by clouds and aerosols,
> i.e., contamination,
> within the instrument's field of view (FOV).
> Screening of the most contaminated soundings
> minimizes unnecessary calls to the computationally expensive

Level 2 (L2) \mbox{X_{CO_2}} retrieval algorithm.

For example, in the Introduction;

> It is therefore necessary to provide reliable
> cloud screening on all of the approximately
> one million OCO-2 measurements collected each day.
> In this work the definition of optical thickness
> includes the contribution from aerosols,
> as well as from both ice and water clouds,
> except where noted.
> Therefore, for OCO-2,
> labeling a scene as cloudy,
> indicates the detection of either cloud or aerosol, or both.

Lines 391-392
brought my reading cadence to a standstill, since too much is said in too little space.
Suggest some expansion of the sentence to clarify the sentence.

Reworded the sentence;

> It is possible that this could lead to disagreements
> in classifying contaminated soundings,
> especially for scenes containing small aerosol particles,
> i.e., large Angstrom coefficients,
> a condition in which the measurements from the two sensors
> need to be made at the same spectral points.

Page 13, line 412.
Is Throughput missing a divisor (e.g. Ntotal)?

The omission of the divisor has been corrected.

Page 14, line 435.
Change to "as once soundings have been .."

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 \mbox{X_{CO_2}}
> retrieval algorithm.

Page 14, line 442.
May make more sense to discuss the three metrics in terms of the post-launch settings (nstead of the pre-launch settings).

Reworded and added additional information to this discussion;

> For this particular data set,
> setting $\Delta p_{s,cld}$ to 25\,hPa and
> χ^{2} scale factor to 5
> allows a throughput $\simeq$ 42\%,
> with agreement $\simeq$ 77\% and $\mbox{PPV}$ $\simeq$ 52\%.
> The operational settings of the OCO-2 ABP since
> the on-orbit instrument checkout phase (September, 2014)
> have been 25\,hPa and χ^{2} scale factor $=$ 20.
> Studies showed that for nadir-land and glint-ocean viewing
> the $\Delta p_{s,cld}$ filter alone flags approximately 98\%
> of the soundings determined cloudy by ABP,
> while the surface albedo check provides significant filtering
> (up to 25\% of cloudy scenes) for glint-land viewing.

Page 15, line 487.
Suggest change to "It is critical to avoid latitudinal gradient biases in the ..."and/or other phrases. The phrase "spatial sampling biases" is too ambiguous. Clarify.

Replaced "spatial sampling biases" with "latitudinal sampling biases" as suggested.

Also, removed the following sentence as it does not seem to add much to the discussion;

> To this end,
> it is desirable for the OCO-2 cloud screening algorithms
> to pass as many of the clear-sky scenes as possible
> in persistently cloudy regions such as the convective tropics.
> There is only a modest penalty if the cloud screening thresholds are set too loosely,
> and some cloudy scenes are passed,
> as most will be identified and rejected by the OCO-2 L2 retrieval algorithm.

Finally, combined this paragraph with the one beginning with "To assess the spatial distributions…" for brevity.

Page 16, line 516.
Does the MODIS cloud mask utilize the Deep-Blue AODs? If not, this can be mentioned.

Modified and added to the discussion as follows;

> Specifically, the MODIS cloud mask correctly identifies these scenes as clear,
> but a single case study of the MODIS Deep-Blue derived AODs
> \citep{Hsu:JGR:2013:MODIS_deep_blue}
> revealed that sometimes these scenes are heavily aerosol laden.
> Implementation of the MODIS Deep-Blue AODs into the definition of
> cloudy/clear used in this work may provide slightly improved agreement
> between OCO-2 and MODIS cloud screening.
> But the collocated product was not available at the time this
> research was performed.
> As stated above,
> the OCO-2 screening algorithms do not discriminate between aerosol and cloud,
> and hence identify any scenes that are contaminated by cloud and/or aerosol.

Page 17, lines 574 –579.
I was not convinced by the reason presented in the text, that sub 5 km inhomogeneity in the cirrus cloud field accounts for the discrepancy stated on line 574. Upper tropospheric cirrus have horizontal length scales usually larger than 5 km.

The discussion here was in reference to the "all-clouds" case, not just high (cirrus) cloud cases. The CALIOP analysis was limited to data collocated with 5km of OCO-2. This was mentioned in Sect. 3.3. An initial analysis (not shown) of scenes collocated beyond 5km revealed very little differences in the statistical results.

Page 18, line 601.
Has the adeptness of ABP to detect very thin scattering layers high in the atmosphere been confirmed by recent calculations (after OCO-2 was positioned to more closely be in the CALIOP orbit)? If this is incorrect, revise and/or delete the claim.

As stated above, the CALIOP analysis was limited to those soundings with collocation distances < 5km. We therefore hypothesize that there will be no significant change in the global statistics when using a data set from the more recent period when OCO-2 and CALIOP have been flying in tighter formation. It is possible that looking at a different time period, i.e., season, would yield some differences in the results, but that work is currently beyond the scope of the current project.

Page 18, lines 600 and 604.
These two sentences are inconsistent ("unable" in line 600, "identifies most" in line 604). It may make more sense to move the sentence at the end of line 603 to the Conclusions.

Modified the discussion to;

These results suggest that the ABP,
which relies on PPL modification to detect cloud,
is unable to discern cloud near the surface,
even when the optical thickness is large.
Conversely, ABP is sensitive to
very thin scattering layers when they are located high in the atmosphere
due to the strong PPL modification.
Again, both of these behaviors were first identified in simulations
as seen in
Fig.~\ref{fig:cloud_hist_oco_land_june},
and have now been demonstrated with real data.

Figure 1.
There is a black shading along the orbit track over clear ground in the middle of Panel (b). The blue part of the track makes sense, and the black part of the clouds makes sense, but north of the clouds near 25.3 N, the black shading is not expected.

The following comment was added to the caption;

> The cloudy frames north of the visible cloud deck presumably
> contain sub-visible clouds or aerosols.

Figure 5.
The green shading is of little help. Would black and white panels (remove the color) make more sense?

Agreed. The contour plots were updated to black and white for clarity.

Figures 3 and 7.
The "hump" near forward model and/or CALIOP optical depths of 3 is very odd. If you have a good explanation for this, please provide it when Figure 3 is first discussed.

Unfortunately, we do not have a good explanation for this feature. Several comments to that effect have been added in the Simulations and Data discussion at appropriate points in the paper.

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.