Review #1

General comments:
1. Whereas the overall research is of great importance as the various EOS algorithms are moving into the JPSS era, this paper is plain and simple incomplete. The authors give a pretty good discussion of the footprint size difference estimation experiments, in the title and the abstract they speak of retrieval algorithm differences as well, yet in the text the treatment of retrieval algorithm barely gets a sentence or two. I feel like the entire discussion of the retrieval algorithm is missing. In research it is important for published result to be reproducible to at least some degree. This work is definitely not reproducible. The footprint difference evaluation needs more detail and the algorithm difference evaluation plain and simple needs to be written.
Thank you for your time to review this paper. We added a paragraph on page 9 to provide more details on the cloud retrieval differences between MODIS and Aqua. We also revised the method section extensively with the new diagram (Figure 6) to illustrate the simulation process.

2. Throughout, please consider the audience may not know what your different acronyms are. Please expand in text or provide an acronym table.
We defined all the acronyms in the paper.

Specific comments:

Line 7: CERES, first mention of the instrument, please expand
We defined CERES in the title, which should give reader a reference.

Line 12: same thing for VIIRS and MODIS, consider the audience may not know what the acronyms are.
Defined.

Line 26: "cloud property differences", elaborate what exact cloud properties you mean
Defined the cloud property as "cloud fraction and optical depth".

Line 48: "flied" is not a word
Thank you for catching the typo, corrected.

Line 49: "were launch" is grammatically incorrect
Corrected in the revised version.

Line 56-57: spacecrafts is not a word
We replaced "spacecrafts" with "satellites".

Lines 115-117: why have you not considered the scattering angle?
The scattering angle is intrinsically considered here, as cos(scattering angle) = cos((SZA)cos(VZA) + sin(SZA)sin(VZA)cos(RAZ), where SZA is solar zenith angle, VZA is viewing zenith angle, and RAZ is relative azimuth angle.

Line 128: "the nighttime LW radiance is simply from the total channel measurements" doesn't make any sense.

We don't quite understand the reviewer's comment. As we stated in the paper, the CERES instrument measures radiances in shortwave, window, and total channels. For daytime measurements, the longwave radiance is derived as the difference between total channel and shortwave channel. For nighttime measurements, the shortwave radiance is zero thus the longwave radiance is the same as the total channel. We changed this sentence to "the nighttime LW radiance is directly derived from the total channel measurements" to be consistent with the previous sentence.

Line 157: snow /no snow conditions, how are you getting this information?
Snow information was obtained from the national snow ice data center and is included in the CERES SSF data. Detailed description is added on page 5.

Line 159: "to derived" is grammatically incorrect
Corrected, thank you!

Line 163-166: how are you getting snow/no snow information?
Snow information was obtained from the national snow ice data center and is included in the CERES SSF data. Detailed description is added on page 5.

Line 178: replace 'assessed' with 'ascertained'
Changed.

You talk about the simulated footprint generation, but where is the discussion of the algorithm differences? "We tweaked it" is not good enough.
The cloud algorithm difference between MODIS and VIIRS cloud retrieval are discussed on page 9. However, please note that addressing the cloud retrieval difference is not the focus of this paper. The main focus of the paper is to address the effect of footprint size and cloud retrieval differences on flux inversion. We don't understand where "We tweaked it" is coming from, as we didn't use any wordings like that in the paper and therefore unable to address the reviewer's comment.

Line 192: "that consist with" is not English, "result in less footprints" is not either.
We modified the sentence to "as the narrowband-to-broadband regressions are only applied to footprints that consist of the same surface types which result in fewer footprints with valid fluxes for CERES-NPP than for CERES-Aqua."

Line 204-210: Why? Please explain
Explanations of the cloud fraction and cloud optical depth differences were given on lines 197-205.

Line 224-227: This does not make any sense to anyone who might theoretically even consider checking your results.

Readers who might want to check our results can download the CERES-Aqua and CERES-NPP SSF data and calculate the daily ratio for each 1 degree by 1 degree grid box using the cloud fraction and cloud optical depth of the CERES footprints.

Line 235: Why did your flux increase? Explain

CERES flux is derived as: π*radiance/anisotropic factor (see Equation 1 in the revised manuscript). As we explained in the paper, for oblique viewing angles, the anisotropy factors for thick clouds are smaller than those for thin clouds. As anisotropy factors decrease when retrieval cloud optical depths increase, the inverted fluxes increase (as the radiances didn't change). We understand this part might be confusing, to help reader better understand how we derive flux from CERES radiance measurements, we added some descriptions in the introduction section.

Line 273: replace 'is' with 'are'
Modified.

Line 275: replace 'these' with 'that'
We rewrote this sentence.

Line 282: replace 'thus' with 'to'
Changed.

Comments on figures:

Figure 2: so these CERES footprints are misaligned like that in real life or just in illustration. Please clarify.
We rewrote the paper, and Fig. 2 is replaced with the flow chart (Fig. 6). The original Fig. 2 is just for illustration, not how CERES footprints look like in real life.

Figures 3-6: can't be evaluated because I don't know what I'm looking at. There is no explanation in the text as to why the images look they way they do. Could be a thousand different reasons, so figures do not make sense.
We modified the paper and the captions of these figures to make it clear to readers.

Review #2

Major comments

This manuscript attempts to quantify the flux uncertainties in the NPP-CERES fluxes because NPP does not have a rotating CERES instrument to generate NPP specific ADM as the TERRA and AQUA satellites had and has to borrow the ADM generated from AQUA satellite. The study uses AQUA MODIS radiances to create simulated fluxes for both AQUA and NPP footprints - going through the same narrowband to broadband conversion and using the same ADM for radiance to flux conversion to isolate the differences due to footprint size alone. The differences due to both footprint size and cloud properties are examined with VIIRS-adjusted cloud properties and subsequent ADM. This study helps to understand and quantify the uncertainty in the current NPP CERES fluxes due to missing its own ADM. However, the description of the methodology could be much improved by adding proper flow chart and naming convention.

Thank you for your time reviewing this paper. We added a flow diagram (Figure 6) to illustrate the process of convoluting the high-spatial-resolution MODIS pixel-level radiances and cloud/aerosol retrievals with CERES-Aqua and CERES-NPP point spread functions to provide narrowband radiances and cloud/aerosol properties for the respective simulated CERES-Aqua and CERES-NPP footprints, and the simulated CERES-NPP footprints with VIIRS-like cloud properties. These narrowband radiances are then used to derive broadband radiances and they are converted to fluxes using the Aqua ADMs.

The other missing piece of information is how the uncertainties estimated with simulated fluxes starting from MODIS pixel radiances are measured against actual retrieved NPP CERES flux?

As illustrated in Fig. 3 in the revised manuscript, the overpass times for NPP are different from those of Aqua. The broadband radiances generated using narrowband-to-broadband coefficients also differ from the measured broadband radiances. Direct comparison between fluxes in the simulated CERES-NPP footprints and the actual CERES-NPP footprints will mostly show the difference caused by the afore mentioned differences (the simulated CERES-NPP used the Aqua Sun-viewing geometry). The focus of the paper is to assess the effect of footprint size difference and scene identification difference on flux inversion without having to account for the overpass time difference and calibration difference. This is the reason that we designed this study and used narrowband-to-broadband regression to derive the broadband radiances.

Similar comparisons between the matched NPP and AQUA radiances (Figure 1) could be applied to match fluxes. Likewise, gridded monthly flux differences between the operational CERES-AQUA and CERES-NPP retrievals (similar as Figures 3, 5, 6) would provide an all-inclusive uncertainty that the current simulated uncertainty estimates could put into perspective.

We compared the fluxes using matched NPP and Aqua footprints as suggested by the reviewer. Results are provided in Fig. 2 and Table 1, and descriptions are on page 8. However, as mentioned above and demonstrated in Fig. 3, the overpass time differences between the Aqua and NPP orbits dominate the monthly gridded flux differences, as shown below. The left plot

shows that monthly gridded SW flux for Aqua and the right plot shows the SW flux differences between NPP and Aqua. Most of the features in the difference plot resemble those in Fig. 3.
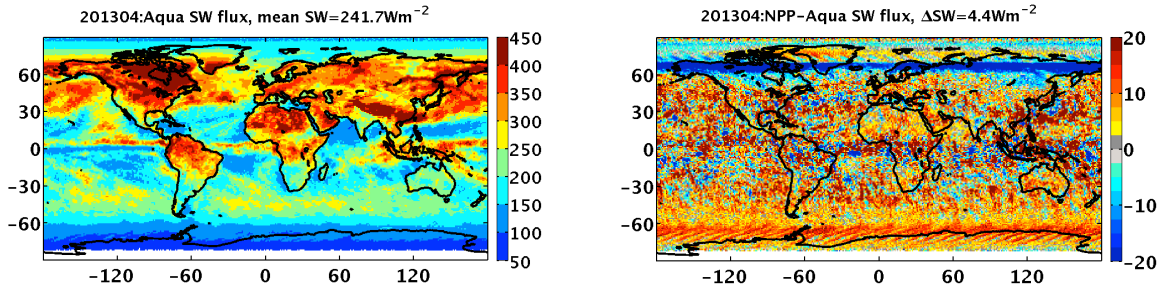
We can compare the albedo between NPP and Aqua to bypass the insolation difference, as shown below (figure on the left is the monthly mean albedo and figure on the right is albedo difference between NPP and Aqua). The global monthly mean albedo from CERES-NPP is about 0.003 (1.02%) greater than the global monthly mean albedo from CERES-Aqua (0.292). We believe these differences stem from the combination of calibration difference, ADM and scene type differences. It is impossible to attribute how much each of them contributed to the albedo difference just by comparing the flux/albedo from Aqua and NPP. This is the reason that we come up with this study to isolate the contribution of ADM and scene type differences to flux difference.

We added the albedo difference plot (Fig. 4) in the revised manuscript and some discussions on page 8.
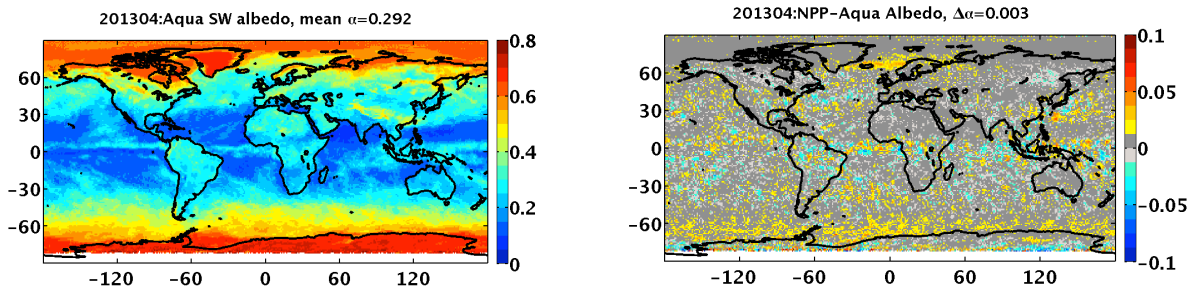
Detailed comments:

1. The methodology description is very confusing. Please explain clearly how each of the three products are produced, use an acronym for each of them and the intermediate products, i.e.,

SNR- Aqua - simulated narrowband radiance with Aqua footprint size

SBR -Aqua - simulated broadband radiance with Aqua footprint size

SBF-Aqua --- simulated broadband flux with Aqua footprint size

We rewrote the methodology section and defined that $I_a^s(\lambda)$ is the narrowband radiance for the simulated CERES-Aqua footprints, I_a^s is the broadband radiance for the simulated CERES-Aqua footprints, and F_a^s is the broadband flux for the simulated CERES-Aqua footprints. Their counterparts for the simulated CERES-NPP footprints are $I_n^s(\lambda)$, I_n^s, and F_a^s, and they are $I'^s_n(\lambda)$, I'^s_n, and F'^s_a for the simulated CERES-NPP footprints with VIIRS-like cloud fraction and

cloud optical depth. The CERES-Aqua measured broadband radiance is I_a^m, and the CERES-NPP measured radiance is I_n^m. The corresponding fluxes are F_a^m and F_n^m.

2. Need a flow chart to illustrate the methodology. Something like:
MODIS-Aqua -> SNN-Aqua -> (narrowband to broadband conversion) SBR-Aqua
->(MODIS ADM) SBF-Aqua
MODIS-Aqua -> SNN-NPP -> SBR-NPP -> (MODIS ADM) SBF-NPP
MODIS-Aqua -> SNN-NPP -> SBR-NPP -> (likeNIiRS ADM) SBF-NPP2
We added a flow diagram (Figure 6) to illustrate the process of convoluting the high-spatial-resolution MODIS pixel-level radiances and cloud/aerosol retrievals with CERES-Aqua and CERSE-NPP point spread functions to provide narrowband radiances and cloud/aerosol properties for the respective simulated CERES-Aqua, and simulated CERES-NPP footprints with MODIS cloud properties and with VIIRS-like cloud properties. These narrowband radiances are then used to derive broadband radiances and they are converted to fluxes using Aqua ADMs.

3. Once product acronym is defined, specify which two products are compared in Figures 3, 5, 6.
Modified the captions using the acronym (now Figures 7, 8, 9).

4. line 71-73 From this sentence, it is inferred that one of the CERES instrument on the TERRA/AQUA satellites used RAP scanning mode to create ADM at the beginning of the mission. How long was the RAP mode in operation? line 82-83 says ADM was created from multi-year CERES measurements from both RAP and crosstrack mode. Does that mean both CERES instruments were used for ADM? Please clarify.
We specified the total lengths of RAP data collect for Terra and Aqua on lines 83-84: "There are about 60 months of RAP data collected on Terra and about 32 months of RAP data collected on Aqua.". We clarified that the "Terra ADMs and Aqua ADMs were developed separately using multi-year CERES Terra and Aqua measurements in RAP mode and in cross-track mode using the scene identification information from MODIS" at lines 98-101.

5. line 119: why do you use different view zenith angle requirement for LW radiance?
We modified the SW radiance matching criterion of viewing zenith angle to be the same as LW radiance (i.e. 2 degrees). We updated the Table 1 and Figure 1 for the SW comparison.

6. Figure 1 and Table 1: could you add plots for total radiance comparison between CERES-NPP and CERES-AQUA? Also, the difference in radiances between CERES-NPP and CERES-NPP could be partly due to different footprint size besides the calibration?
CERES instruments measure radiances in total, SW, and window channels. These measured radiances are filtered radiances, which then are unfiltered by accounting for the spectral response function for each channel. Unfiltered radiances are provided for LW, SW, and window channels, but not for total channel, thus we are unable to compare the total radiances. The reviewer is correct, that footprint size difference also contributes to the radiance difference. However, the effect of footprint size on radiance should be random. The radiance bias is mostly from the calibration difference between CERES-Aqua and CERES-NPP, and the footprint size difference can increase the RMS error. We added this in the revised version.

7. Figure 2 Caption: add "CERES" before footprints
This figure is replaced with Figure 6 in the revised version.

8. line 170-173: The narrow-band to broad-band conversion coefficients are derived
from simulated MODIS AQUA radiances and CERES-AQUA broadband radiance. Could
these coefficients also be footprint size dependent? The same coefficients are now
applied to convert both simulated CERES-AQUA and CERES-NPP footprint.
*The narrowband-to-broadband (NB2BB) regression coefficients should not be dependent on the
footprint size, but the broadband and narrowband radiances and fluxes could change as the
footprint size changes which sometimes results in more/fewer clouds in the footprints.*

9. line 177-178: Actually it would be interesting to compare the simulated CERES-AQUA
and CERES-NPP fluxes at footprint level even with different footprint size. How would
this comparison differ from the comparison between matched CERES-AQUA and CERESNPP
observations in Figure 1?
*We compared fluxes using matched CERES-Aqua and CERES-NPP footprints in Figure 2 and
also included the mean fluxes and root-mean-square errors in Table 1. Some discussions are
also added to the paper (page 8). We also compared the fluxes from simulate CERES-Aqua and
simulated CERES-NPP at the footprint level using data of April 1, 2013. Flux differences and
RMS errors are provided on page 12, which are much smaller than the flux biases and RMS
errors listed in Table 1.*

10. Figure 4 (cloud property difference between MODIS and VIIRS) and the generation of
SBF-NPP2 are better introduced before comparison the three products (Figure 3.5, 6) to
ease the logic flow. Also be specific about difference (MODIS- VIIRS) or the other way?
*We revised the paper by adding a section dedicated to the comparison between NPP and Aqua
(section 2) and the cloud property differences are discussed in this section in the revised
manuscript (Figure 5).*

11. Paragraph: 238-250 how is uncertainty defined?
*We revised this paragraph to be more specific. The uncertainty refers to the differences between
simulate CERES-Aqua and CERES-NPP.*

12. Could you provide a map of difference in ADM?
*As shown in Equation (1), for a given scene type, ADM (R) is defined as a function of solar
zenith angle, viewing zenith angle, and relative azimuth angle. Additionally, ADM is also
different for different scene types. Averaging anisotropy values into monthly grid box means
could be misleading.*

13. Finally, please have native English-speaking colleagues check the grammar usage.
There are many awkward sentences.
We did a thorough editing of the paper.

Review #3

This paper studies the effects of different footprint sizes and cloud algorithms on the radiative flux calculation from CERES on Suomi-NPP. Quantifying uncertainties from different footprint sizes and cloud algorithms in CERES-NPP is an important issue in calibration.

However, major revision of the text is needed. The sources of uncertainties in these two CERES (CERES-NPP and CERES-Aqua) observations are not fully explained in the text. The collocated radiances themselves already have 1.5%, 0.5%, and 0.1% of intrinsic differences for each channel, respectively. The sampling errors from two different footprint sizes are also embedded in the radiance comparison.

The actual CERES-Aqua and CERES-NPP measured radiances differ, as mentioned above. These comparisons are based upon matched CERES footprints from Aqua and NPP, thus the sample number used here are the same for Aqua and NPP. The footprint size differences can contribute to the radiance differences, but these differences should be random. We believe that the footprint size differences can increase the RMS errors, but the mean radiance differences are mostly resulted from calibration differences between CERES-Aqua and CERES-NPP (page 7). The goal of the simulation designed in this paper is to isolate the effect of footprint size difference on flux inversion without having to unravel the contribution of calibration difference to the flux difference.

This study assumes that CERES-NPP and CERES-Aqua are identical instruments with compatible performance. But one may ask if there were improvements in CERES-NPP instrumentation/electronics or calibration. Possible degradation of CERES-Aqua instrument can be mentioned, if any. Before discussing uncertainties from footprint sizes and cloud algorithms, comprehensive uncertainty analysis of two CERES instruments is necessary to quantify errors from spatial sampling and cloud algorithm differences. I understand that this work is the first step towards identification of the two different CERES observations, but further uncertainty analysis and detailed descriptions are required. The methods to identify the effects of different footprints and cloud algorithms need to be described in detail.

Instrument degradation is indeed an issue if it is not accounted in the calibration. CERES instruments are designed with onboard calibration and instrument degradation is accounted for when the monthly gain factors are derived. We provided some discussion of the instrument gains on page 7. The calibration differences between CERES-Aqua and CERES-NPP are not fully understood yet. The team is working on to understand the differences and eventually place both of them on the same radiometric scale. The focus of the paper is not to address the calibration difference. The simulated CERES-NPP footprints are designed to quantify the flux uncertainty due to footprint size difference and cloud fraction and cloud optical depth differences without having to unravel them from the calibration difference.

1. Introduction: Line 71, Could you provide detailed explanation how two instruments help to construct ADMs that readers can visualize what you described?
One of the CERES instrument is placed in RAP mode, which is designed to maximize the angular coverage for constructing ADMs, while the other one is placed in cross-track mode to provide

daily global coverage. The objective here is to use one instrument to maximize the angular coverage without having to sacrifice the spatial coverage. We rewrote this part in the revised manuscript to make this clear.

2. Introduction: Line 88, "These pixel-level cloud properties are spatially and temporally matched with the CERES footprints ~" ,
What is the spatial and temporal window for scene type matching from MODIS?
The spatial matching is to include all MODIS pixels (1 km) in the CERES footprint (20 km at nadir). The temporal matching is within 20 seconds when CERES is in cross-track mode, and is within 6 minutes when CERES is in RAP mode. We specified these criteria used on page 5.

3. Line 126-128,
"These differences do not show any view zenith angle dependence".
Figure 1 compares the radiances of SW, daytime LW, and nighttime LW between two CERES measurements, but does not mention view zenith angle difference. Please explain why these differences do not show any view zenith angle dependence. Daytime LW is derived as the difference between total and SW channels. Nighttime LW is the same quantity with the total radiance because SW radiance is zero during nighttime. How daytime LW difference can be lower than SW difference? Are they offset with bias?
The comparisons shown in Figure 1 include data taken from nadir viewing angles to oblique angles with viewing zenith angle greater than 60 degrees. When we compare the radiances taken at different viewing zenith angle ranges (e.g. <20 degrees, >20 degrees, >40 degrees), the RMS errors of the radiances remain almost the same. This indicates that these differences do not show any viewing zenith angle dependence. We added these discussions in the revised manuscript (page 7).

The reason that the daytime LW difference is smaller than that of SW is because the difference of the total channel is slightly smaller than the difference of the SW channel. When the SW radiance is subtracted from the total radiance, the difference in LW is not as large as that of SW.

4. Table 1.
The difference between two cases is largest in SW radiance, but it is less in day time LW radiance. Authors can explain the reason why the difference in daytime LW radiance, which is derived as the difference between total and SW channels.
As mentioned above, the reason that the daytime LW difference is smaller than that of SW is because the difference of the total channel is slightly smaller than the difference of the SW channel. When the SW radiance is subtracted from the total radiance, the difference in LW is not as large as that of SW. We added some explanation on Page 7.

Line 212-14,
"Polar region cloud fraction differences are mainly because (that) VIIRS lacks the water vapor and CO2 channels which affect the polar cloud mask algorithm."
Please explain how lack of those channels cause polar region cloud fraction difference. While VIIRS shows less cloud fraction over northern high latitude

snow regions, it shows more cloud fraction over arctic. If all positive and negative difference is caused by large uncertainty of VIIRS, then more explanations on VIIRS cloud retrieval limitations including usage of different parameterization over snow/ice surfaces and uncertainty caused by this difference should be mentioned.

The CERES cloud working group developed sophisticated cloud detection algorithms using visible and infrared channels of MODIS separately for polar and non-polar regions and for daytime, twilight, and nighttime. However, these detection algorithms have to be modified to apply to the VIIRS observations, as some of the MODIS channels utilized for cloud detection are not available on VIIRS. These modifications include replacing the 2.1 um MODIS channel with the 1.6 um VIIRS channel, and replacing detection tests using MODIS 6.7 um and 13.3 um channels with VIIRS 3.7 um and 11 um channels, and supplement with tests utilizing VIIRS 1.6 um channel and the brightness temperature differences between 11um and 12um. These changes mainly affect the cloud detections over the polar regions, and as different tests and thresholds are used for daytime and twilight cloud detections, these changes will also affect different latitude zones differently.

Line 158 are used to derived → are used to derive
Corrected, thank you!

5. Conclusions
299 How these simulations helped to calibrate CERES-NPP observations?
These simulations are not designed to calibrate the CERES-NPP radiance observations. The focus of the paper is to assess the effect of footprint size difference and scene identification difference on flux inversion without having to account for the overpass time difference and calibration difference.

306 How much is the uncertainties of CERES ADMS as appeared in Su et al (2015)?
The biases and RMS errors provided in Su et al. (2015) is for diurnally averaged SW and LW flux. The bias for regional monthly mean TOA SW flux is less than 0.2 Wm-2 and the RMSE is less than 1.1 Wm-2. For TOA LW flux, the regional monthly mean bias is less than 0.5 Wm-2 and the RMSE is less than 0.8 Wm-2.

314-316 VIIRS and MODIS cloud retrieval algorithms are not consistent. Then it is important to quantify the differences and reflect them to improve the CERESNPP algorithm. Last sentence, "To maintain the consistency of the CERES climate data record, it is thus important to maintain the consistency of cloud retrieval algorithms." does not help much to solve the problem.
Figure 5 quantify the cloud fraction and cloud optical depth difference between Aqua and NPP. The CERES team developed cloud retrieval algorithm using MODIS observations for over ten years. Because VIIRS lacks some of the channels that were used for MODIS cloud retrieval, to make the cloud retrieval consistent between VIIRS and MODIS, the MODIS and VIIRS retrieval algorithms have to be reevaluated using only channels that are available for both MODIS and VIIRS. This is a grand challenge and will take time to develop the new algorithms and to process

data of more than 30 instrument years. We rewrote this sentence to reflect the need to develop algorithm considering the capabilities of both MODIS and VIIRS.