Summary:

I believe it is important to publish satellite retrieval product papers to document the algorithms and to properly reference and cite the product. This paper documents the NOAA STAR GOES ABI TOA SW flux product algorithm. The paper highlights the challenges that need to be overcome to develop the GOES-ABI flux product, especially the spectral information needed to compute the broadband TOA albedo from ABI spectral channels. The paper is based on the product ATBD. The paper describes the current algorithm in depth, but the extent of the validation was lacking given that there are 3-years of validation opportunities. Once the product has been fully developed it will be a great asset to the remote sensing community in providing ABI 2-km pixel resolution TOA SW fluxes over the CONUS region. There is much work left to do for a viable product and documenting the progress is worthy of publishing in this journal. I will consider the paper for publication after the following concerns are addressed.

General comments.

The authors should be using the official CERES SSF L2 product not the CERES FLASHFlux L2 product for validation. The FLASHFlux product was designed for real-time processing, where many SSF inputs were replaced with real-time datasets, for example the GEOS 5.4.1 reanalysis rdataset was replace by the realtime FPIT dataset. The more realtime dataset algorithm datasets are often revised due to changing input quality. The CERES input datasets were designed for consistency across the record by limiting algorithm changes to avoid discontinuities in the parameter values. Also, the FLASHFlux fluxes do not employ the most up to date CERES instrument calibration coefficients.

I am assuming that the NOAA STAR GOES ABI TOA SW flux product is not available to the public.

I am having a hard time understanding why there is very little validation being performed. Table 7 simply is not sufficient, not even a full year of data is analyzed. It seems that exact time matching is necessary for validation. The ABI scans are every 15/10 minutes providing closely matched ABI and CERES fluxes.

There is no high-resolution TOA SW flux dataset ground truth dataset, agreed, that is the motivation for this product. The CERES dataset provides observed instantaneous SW fluxes at the 20-km nominal resolution. Linear interpolating the CERES footprint center fluxes across the ABI pixels does not represent a valid 2-km flux field. Cloud edges are not distinct.
This implies that the ABI high-resolution TOA SW fluxes should be mapped into the CERES footprint for validation. Or into lower resolution latitude and longitude grid such as performed by Akkermans, T.; Clerbaux, N. Narrowband-to-Broadband Conversions for Top-of-Atmosphere Reflectance from the Advanced Very High Resolution Radiometer (AVHRR). Remote Sens. 2020, 12, 305. https://doi.org/10.3390/rs12020305
Note they also stratify the validation results by IGBP type. The important part of the validation is determining whether the algorithm is not adding an overall bias to the TOA SW fluxes, while trying to reduce the RMS error. The instantaneous RMS error is a function of spatial scale. They also validate a several year's worth of TOA fluxes

The bin/channel regression rely on RTM results that have varying PW and ozone concentration. The PW water above the cloud or clear-sky surface is necessary to predict the NIR water vapor absorption, since none of the ABI band used are located inside absorption bands. It is not clear to me how current algorithm accounts for NIR water vapor absorption? This was unclear in section 2.

I believe the greatest uncertainty in the NTB algorithm is accounting for spectral information. Could the MODIS 2-week surface band BRDF be used in MODTRAN to update the predefined MODTRAN BRDFs? The MODIS BRDF product could be used to account for regional and seasonal variability.

Specific Comments:

Line 38 this paragraph seems out of place. Unless this study was used to for ABI channel selection it does not seem relevant.

Line 45. It would be beneficial for the reader to briefly outline the whole algorithm. To discuss both the indirect path and I am assuming a direct path. Perhaps to provide how this algorithm was developed and if it is used in any historical products.

Line 64 Does ground refer to truth dataset or to actual ground observations, since in the summary mentions "ground truth"

Line 95 Are the Kato and Loeb snow ADMs used as part of the CERES Ed2 ADMs?
Kato, S., and N. G. Loeb (2005), Top-of-atmosphere shortwave broadband observed radiance and estimated irradiance over polar regions from Clouds and the Earth's Radiant Energy System (CERES) instruments on Terra, J. Geophys. Res., 110, D07202, doi:10.1029/2004JD005308

Line 97 The Niu and Pinker are theoretical simulations, how do they translate to observation numbers in Eq. 3?

Line 131. I do not see how the Fig. 4 comparison adds value to the paper. The profiles were selected to get a sampling of the diversity of atmospheric profiles found on Earth.

Line 237 Is the Matlab stepwise fit used in the algorithm? If not this should sentence should be left out because it adds confusion.

Fig. 8 Could the spectral boundaries or band edges for each ABI band also be shown in Fig. 8. This way the reader can see the spectral range radiance that is predicted based on a single ABI band.

Line 260 could the band edges be given in μm in the text also.

Section 2.6 Which channel takes into account the bulk of the NIR water vapor absorption?

Line 264 Figure 9 is spelled out, whereas Fig. 8 is not on line 261

Line 266. I would agree that along the cloud edges there would be large differences between ABI and CERES TOA fluxes. These large differences would occur even if there were a perfect algorithm. However, over large spatial domains the ABI and CERES fluxes should be similar.

Table 6 and 7 are not referenced in the text.

Line 267 It would be nice to have statistics for Figure 9 similar to what is in Table 7. I do not see 2017/11/25, 17:57Z Fig 9 statistics with the 2019 statistics in Table 7.

Line 276. This is where Table 6 should be referenced to identify the CODC product

Line 283. The authors should use the CERES SSF Level 2 data, rather than CERES FlashFLUX footprint fluxes. As mentioned in the text, that FlashFLUX does not use the most up to date CERES instrument calibration coefficients. The CERES SSF product is available within 3-months of real-time.

What is limiting the number of validation match ups? Is the issue that your computing resources have limited computer storage that downloading all of the required datasets for ABI pixel level fluxes and comparisons with CERES is not possible after real-time when these products are no longer available at CLASS?

Line 304. The CERES footprint data has a resolution of 20-km at nadir, while the ABI pixel has 2-km resolution. By linear interpolating spatially the CERES fluxes across the ABI pixel does not properly distribute spatially the CERES flux observation (by not preserving cloud edges) and I would not consider that a truth dataset, since it does not represent the observed 2-km fluxes, It would be better to map the ABI pixels into the CERES footprint to validate the NTB algorithm. A CERES footprint at 60° view angle (near the scan edge) has a 40-km extent encompassing over 400 ABI pixels at nadir. Even better would be to evaluate the ABI product regionally, say for 1° regions, so that monthly regional comparisons can be made.

Fig 10 caption missing (e)

Line 326 Based on Table 6, the ABI radiances, aerosols, cloud mask, phase and optical depth are used as inputs. For clear-sky the surface spectral reflectance is based on 12 IGBP types, and 4 types for cloudy types. How is the pixel level above surface or cloud top amount to account for NIR atmospheric water vapor absorption. A lot of effort was used to define atmospheric profiles, I would assume this would be based on the ABI channel radiances. My other concern is that the 0.86 vegetation reflection is a function of season and region, in winter the leaves have fallen off the trees, where as in summer the trees have leaves. By simply relying on IGBP type does not account for the seasonal vegetation reflection.

Line 346 and line 33. Given that the ABI sampling is less than 15 minutes. The 7.5 minute difference is very small. Once the SW fluxes are compared at the footprint or regional scales the

time difference will not make much of a difference in the bias. All Terra and Aqua overpasses should be matched for well sampled validation results. The following paper Fig. 2 shows that the time difference does not dramatically increase the matching noise
B. A. Wielicki, D. R. Doelling, D. F. Young, N. G. Loeb, D. P. Garber and D. G. MacDonnell, "Climate Quality Broadband and Narrowband Solar Reflected Radiance Calibration Between Sensors in Orbit," IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium, 2008, pp. I-257-I-260, doi: 10.1109/IGARSS.2008.4778842

Line 348 I agree that seasonal/regional variation of the NIR vegetation reflection must be taken into account.

Line 358 The CERES edition 4 ADMs also rely on NDVI, which accounts for changes in the vegetation NIR reflectance. The CERES edition 2 relies on surface types only.

Su, W., Corbett, J., Eitzen, Z., and Liang, L.: Next-generation angular distribution models for top-of-atmosphere radiative flux calculation from CERES instruments: methodology, Atmos. Meas. Tech., 8, 611–632, https://doi.org/10.5194/amt-8-611-2015, 2015

Line 38- what is the source of the open shrub, desert, woody savanna and grassland spectral albedos? Are these TOA albedos?

Line 397. I agree there is no truth dataset for 2km resolution BB fluxes. That is the reason why this dataset is being produced. In order to perform a fair comparison, the high resolution ABI pixels fluxes must be mapped into the CERES footprint, or both reduced to a 100-km region in order to track the ABI and CERES over the record.

Line 405. “transformation of narrowband quantities into broadband ones” This sentence is ambiguous.

Line 414. What is this sentence trying to say? “The process of preparing for the usefulness of a new satellite sensor needs to be done in advance, the final configuration of the instrument becomes known at a much later stage.” This was not addressed in the paper

Line 416 What is this sentence trying to say? “As such, the evaluation of the new algorithms is in a fluid stage for a long time.” Usually there is an initial release and as the algorithms improve incrementally while the version number is updated over time. For example MODIS L1b C6.1 dataset is currently available and C7 is being developed and tested.

Line 417 This sentence is confusing. “Agreement or disagreement with know “ground truth” is not fully informative on the performance of the new algorithms to estimate desired geophysical parameters.” Are you talking about compensating errors?

Line 420 reliable cloud screening and cloud properties. What about non-retrieved cloud properties from cloud mask identified pixels? What about optically very thin clouds where the surface contributes to the TOA reflectances.

Line 421 The CERES SSF L2 Edition 4 product SW fluxes has been available prior to ABI and have not gone through any major revisions. The SSF1deg fluxes have been used to monitor global and regional SW flux variability over time. On the other hand the FlashFLUX has undergone revisions.

What is the application of high-resolution ABI TOA SW fluxes, that the low resolution CERES fluxes cannot fulfill? For the application, what is the required SW TOA accuracy?

Line 429 If the ABI aerosol algorithm does not ever reach stability in the future, will the TOA SW product ever be released?