The Global Radiative Energy Budget in MERRA and MERRA-2: Evaluation with Respect to CERES EBAF Data

LAURA M. HINKELMAN

Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, Seattle, Washington

(Manuscript received 11 July 2018, in final form 16 November 2018)

ABSTRACT

The representation of the long-term radiative energy budgets in NASA’s MERRA and MERRA-2 reanalyses has been evaluated, emphasizing changes associated with the reanalysis system update. Data from the CERES EBAF Edition 2.8 satellite product over 2001–15 were used as a reference. For both MERRA and MERRA-2, the climatological global means of most TOA radiative flux terms agree to within \( \pm 3 \) W m\(^{-2}\) of EBAF. However, MERRA-2’s all-sky reflected shortwave flux is \( \approx 7 \) W m\(^{-2}\) higher than either MERRA or EBAF’s, resulting in a net TOA flux imbalance of \( \approx 4 \) W m\(^{-2}\). At the surface, all-sky downward longwave fluxes are problematic for both reanalyses, while high clear-sky downward shortwave fluxes indicate that their atmospheres are too transmissive. Although MERRA-2’s individual all-sky flux terms agree better with EBAF, its net flux agreement is worse (\( \approx 8.3 \) vs \( \approx 3.3 \) W m\(^{-2}\) for MERRA) because MERRA benefits from cancellation of errors. Analysis by region and surface type gives mixed outcomes. The results consistently indicate that clouds are overrepresented over the tropical oceans in both reanalyses, particularly MERRA-2, and somewhat underrepresented in marine stratocumulus areas. MERRA-2 also exhibits signs of excess cloudiness in the Southern Ocean. Notable discrepancies occur in the polar regions, where the effects of snow and ice cover are important. In most cases, MERRA-2 better represents variability and trends in the global mean radiative fluxes over the period of analysis. Overall, the performance of MERRA-2 relative to MERRA is mixed; there is still room for improvement in the radiative fluxes in this family of reanalysis products.

1. Introduction

Reanalysis products are now accepted by the majority of atmospheric scientists as the best available long-term representations of the large-scale dynamical and thermodynamical states of the atmosphere. By incorporating observations such as pressure, wind, temperature, and humidity profiles from radiosondes as well as radiances and other quantities retrieved by satellites into an atmospheric model via data assimilation, model processes are closely constrained to the true atmospheric state. This improved fidelity makes reanalysis data attractive for use in a range of scientific studies.

Both the original and second versions of NASA’s premier reanalysis product, the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Reinecker et al. 2011; Gelaro et al. 2017), were developed at the Global Modeling and Assimilation Office (GMAO) with the goal of providing a superior depiction of the global water cycle (Reinecker et al. 2011). MERRA output has been used to investigate a wide range of topics including basic climatology, variability on diurnal to decadal scales, climate trends, and dynamical processes. For example, Urankar et al. (2012) used MERRA to study the feedbacks of aerosol and clouds on the energy budget over India. L’Ecuyer et al. (2015) relied on multiple variables from MERRA in their assessment of the state of the global energy budget in the early twenty-first century, while Trenberth et al. (2015) used it to examine top-of-atmosphere (TOA) radiative fluxes and surface temperature in the context of climate variability. Applications of MERRA data for polar studies include the work of La and Park (2016) on the effect of clouds on phytoplankton in Antarctica, as well as a paper by Letterly et al. (2016) concerning the impact of winter clouds on summer sea ice in the Arctic. Hundreds of other papers relying on MERRA data are in the literature. While use of MERRA...
continues, the number of studies based on MERRA-2 has grown continuously since its release in 2015. Examples include Shekhar and Boos (2017), examining the dependence of the Saharan meridional circulation on the West African monsoon; Cullather and Nowicki (2018), concerning the impact of atmospheric conditions on Greenland ice sheet melt; and the AMS State of the Climate report for 2016 (Blunden and Arndt 2017).

The reliability of all of these studies depends on the accuracy of the MERRA reanalyses’ representations of the atmosphere. In particular, all of the listed studies make use of radiative flux variables. However, in MERRA and MERRA-2, as in most reanalysis products, cloud properties and radiative energy budgets are unconstrained by data assimilation. Instead, cloud properties are derived from the assimilated dynamical fields, and radiative fluxes are computed from these combined with the assimilated thermodynamic profiles. Thus, the accuracy of radiative fluxes is mainly dependent on cloud parameterizations and radiative models rather than measured data, even though these fluxes are crucial drivers of the water cycle. For this reason, radiative fluxes from reanalyses must be thoroughly evaluated before they are used in climate or atmospheric process studies.

Certain aspects of the reanalyses’ radiative energy budgets have been evaluated previously. As the older of the two products, MERRA has received more attention in this regard. Bosilovich et al. (2011) examined MERRA’s TOA and surface radiative fluxes, looking not only at global values, but also at averages over the land and ocean. Other work by Wang and Dickinson (2013) concentrated on LW fluxes, while Dolinar et al. (2016) evaluated TOA cloud radiative effect, cloud amount, and precipitation using satellite, reanalysis, and surface observational data. Evaluations focused on specific regions have also been carried out. These include Cullather and Bosilovich (2012), looking at the polar energy budgets; Zib et al. (2012), comparing surface radiative fluxes and cloud fraction to surface measurements in the Arctic; and Wang and Zeng (2012), concentrating on the Tibetan Plateau. In addition, Posselt et al. (2012) evaluated clouds and radiative fluxes in MERRA under a range of ENSO conditions.

Significantly fewer studies have assessed the radiative fluxes in the MERRA-2 data product. A preliminary assessment of MERRA-2 that included radiative fluxes in addition to many other atmospheric variables is described in a NASA technical report (Bosilovich et al. 2015). Collof and Miller (2016) included comparisons between MERRA-2 and CERES-SYN1Deg (Rutan et al. 2015) radiative fluxes in their study of the radiation budget of the Amazon rain forest. Draper et al. (2018) evaluated all of the surface energy budget terms over land in MERRA-2, focusing on improvements to hydrology by emphasizing turbulent fluxes and their relation to temperature and precipitation. Of all of these reports, only Bosilovich et al. (2015) included comparisons between MERRA and MERRA-2 radiative terms.

A thorough evaluation of the radiative energy budget in MERRA-2 is thus absent from the literature, and it is the gap that we address here. This evaluation covers variability in both time and space and is made relative to the Clouds and the Earth’s Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Edition 2.8 satellite radiative flux product, which is designed specifically for climate model assessment and estimation of the global mean radiative energy budget. Comparisons between EBAF and the original version of MERRA are included in order to help identify where algorithm changes have improved or degraded the product. Simultaneous evaluation of MERRA and MERRA-2 may also aid researchers in deciding which product is best for their purposes.

2. Data and method

a. MERRA

NASA’s MERRA (Reinecker et al. 2008, 2011; no longer in production) was created to provide context for NASA satellite data, with a specific goal of better representing the hydrologic cycle than existing reanalyses. It combines version 5.2.0 of the Goddard Earth Observing System atmospheric general circulation model (GEOS-5) with a 3D variational data assimilation (3DVAR) algorithm to ingest a range of observational data. Notable features include an incremental analysis update (IAU) procedure, which applies corrective tendencies gradually over time, extensive use of radiance data from satellites, and an emphasis on accurate land surface modeling. MERRA spans the time period January 1979–February 2016 with a native spatial resolution of 0.5° × 0.66°.

The variables of greatest interest in this study will be those describing clouds and radiation. GEOS-5 incorporates a moist physics scheme that includes large-scale prognostic clouds (Bacmeister et al. 2006) and a modified version of the relaxed Arakawa–Schubert convective scheme (Moorthi and Suarez 1992). Layer clouds are grouped using three height ranges, and maximum overlap is assumed within the groups, while random overlap is assumed among the groups. Broadband radiative transfer is then computed using Chou and Suarez (1999) for the shortwave fluxes and Chou et al. (2001) for the longwave. Additional details about the MERRA system can be found in Rienecker et al. (2008, 2011). All two-dimensional fields, such as cloud-top pressure and TOA and surface radiative fluxes, are available hourly on the native 0.5° × 0.66° grid.
b. MERRA-2

The differences between the atmospheric global climate models underlying MERRA and MERRA-2 are described by Molod et al. (2012, 2015). The updates made to GEOS-5 were intended to improve the performance of the model in both the atmosphere-only and atmosphere–ocean configurations. Accordingly, changes were made to parameterizations of several physical processes. Most important to this study were modifications to the representation of turbulence and moisture-related processes. Cloud development is affected by the resulting impacts on precipitation at the grid scale and phase transitions of water in both clouds and precipitation, as well as the addition of a stochastic cumulus convection trigger. The operation of the turbulence parameterization under conditions of strong wind shear was also altered.

All of these changes were limited to adjustments to the parameterizations already used in GEOS-5. Only the surface layer parameterization was completely replaced, with the Helfand and Schubert (1995) scheme being incorporated in the place of Louis (1979). Another significant modification to GEOS-5 was the addition of observational aerosol optical depth data to the assimilation system (Randles et al. 2017), which clearly affects both cloud formation and the propagation of radiation.

The two-dimensional cloud and radiative variables provided by the MERRA-2 dataset are essentially the same as those from the original version. The major difference visible to the user is the change in base resolution to 0.5° × 0.625°.

Comparisons between MERRA and MERRA-2 performed at the GMAO have shown improvements in cloud amount and atmospheric moisture content that have been traced to greater evaporation of frozen precipitation and cloud ice water content (Molod et al. 2015). Incorporation of a probability distribution function for total water based on AIRS measurements also improved the accuracy of moisture and cloud variables. Changes to the critical relative humidity that controls micro- and macrophysical cloud processes reduced the cloud fraction in the 300–600-hPa range in the high latitudes, giving better agreement with the zonal mean cloud cover from AIRS.

c. CERES EBAF

The CERES (Wielicki et al. 1996) EBAF product (Loeb et al. 2009; Kato et al. 2013) was specifically created for use in climate model evaluation and energy budget estimation. It is based on the global 1° × 1° CERES SYN1deg flux product, with modifications that improve its similarity to numerically modeled datasets. In SYN1deg, TOA clear-sky irradiance is retrieved for a grid cell only when most of the observations falling within that cell indicate that the sky is actually clear. Thus, clear-sky irradiance values may be missing in areas that are consistently cloudy. MODIS observations from the clear portions of partly cloudy scenes are used to extend the coverage of clear-sky irradiance in EBAF. The MODIS narrowband radiances are converted to broadband irradiance by applying empirical narrowband-to-broadband conversion relationships and the CERES angular distribution models. Interpolated irradiance values are used when no clear-sky observations are available. In addition, an overall top-of-atmosphere imbalance of 6.5 W m⁻² was found in an earlier version of the SYN1deg product (known at that time as SRBAVG-GEO). Since an imbalance of this size cannot be explained by the observed storage of heat by the Earth system, adjustments constrained by known product uncertainties were made to the TOA radiative fluxes to align them as much as possible with the observed heat storage (Loeb et al. 2009). When first applied to the SYN1deg data in 2009, this resulted in a final imbalance of about −0.9 W m⁻². Elimination of most of the imbalance improves the product’s utility for evaluation of climate model output, since the net TOA imbalance is constrained to be zero in these models.

All-sky surface fluxes in SYN1deg are computed for each grid cell based on descriptions of the atmospheric column obtained from satellite retrievals (e.g., cloud height and optical depth), GEOS-5 reanalysis (atmospheric temperature and humidity profiles, skin temperature), and other datasets (surface albedo and emissivity). Clear-sky fluxes are then computed by removing the cloud parameters from each atmospheric column, as is generally done in climate models and reanalyses. The uncertainty of each of the input variables contributes to the overall uncertainty of the irradiance values computed from them. To minimize this uncertainty and increase the consistency with the optimized TOA values, the surface irradiance computations were constrained to the TOA fluxes (Kato et al. 2013). In this process, the atmospheric and surface properties used in the computations were adjusted according to their uncertainty, as determined from independent observations. A consequence of the adjustments is that the computed clear-sky surface fluxes take on some of the characteristics of the clear-sky-only sampled TOA fluxes. The resulting surface fluxes were evaluated by comparison to surface observations.

Uncertainty estimates for the EBAF-TOA and EBAF-Surface Edition 2.8 are given in Table 1 and will serve as references to assess significance in the comparisons that follow. Note that the uncertainties at the surface are 2–6 times higher than at the top of the atmosphere due to the many inputs used in the surface flux calculations.
d. Computations

MERRA (GMAO 2008), MERRA-2 (GMAO 2015), and EBAF (Loeb 2014; Kato 2014) irradiance data for the 15 years from 1 January 2001 to 31 December 2015 were obtained from the appropriate NASA data centers. A consistent set of radiative flux terms was computed from the variables provided in the three datasets. A conservative regridding operation was then used to interpolate the MERRA and MERRA-2 data to the CERES EBAF 1° x 1° grid, selected as the highest common spatial resolution. Time series, annual cycles, and climatologies of gridcell, zonal, and global mean fluxes were computed from the gridded data. Mean fluxes for the ocean and land were also computed, omitting grid cells containing both ocean and land. Antarctica was also omitted from the land average because its high, perpetually snow-covered interior regions are extremely different from other land areas. Finally, the MERRA datasets were evaluated by comparison to the EBAF data, as described below.

3. Results

a. Long-term mean global energy budget

We begin by examining the long-term global mean energy budget in MERRA, MERRA-2, and EBAF. Averages of each of the radiative flux components over the time period January 2001–December 2015 are listed in Table 2 (for the TOA) and Table 3 (at the surface). For acronyms referring to fluxes, see the appendix.

1) Top of the atmosphere

For both MERRA and MERRA-2, most TOA radiative flux terms agree to within about 3 W m⁻² of EBAF under all-sky conditions. MERRA-2 incoming shortwave (SW) and outgoing longwave (LW) are slightly improved over MERRA. However, MERRA-2 has a substantially higher (~8%) planetary albedo, leading to a reflected SW flux that is ~7 W m⁻² higher than in either MERRA or EBAF. As a result, MERRA-2 has a TOA net flux imbalance of ~4 W m⁻², which is nearly 5 W m⁻² lower than that from EBAF or 4 times the disagreement for MERRA.

Both the MERRA outgoing LW and MERRA-2 reflected SW flux discrepancies exceed the estimated EBAF uncertainties of 1.75 and 1.0, respectively.

Under clear-sky conditions, the albedo and all of the fluxes from MERRA and MERRA-2 agree better with EBAF. In particular, the albedo of MERRA-2 is much closer to that of EBAF, eliminating the large discrepancy in the reflected SW flux. This suggests that the differences between the TOA all-sky values from MERRA-2 and from the other datasets are due to differences in cloud variables such as amount or optical depth. This is borne out by the cloud radiative effect (CRE) values. MERRA and EBAF CRE values are very similar, but MERRA-2 clearly overrepresents cloud: MERRA-2 SWCRE is ~8 W m⁻² higher than that from either MERRA or EBAF. For its part, MERRA-2 LWCRE is somewhat low, offsetting part of the SWCRE to leave a difference of ~5 W m⁻² in total CRE between MERRA-2 and either MERRA or EBAF.

Note that the total extraterrestrial solar irradiance (TSI or TOA ASWDN) is smaller in CERES EBAF than in either of the reanalysis products. This is because EBAF uses time-varying TSI values from the Total Irradiance Monitor (TIM) on NASA’s Solar Radiation and Climate Experiment (SORCE) satellite (Kopp and Lawrence 2005), which has measured noticeably lower TSI than previous instruments (Kopp and Lean 2011). Over the period considered here, TIM gives a mean TSI of 1359.6 W m⁻², compared to the values of 1365.2 and 1362.0 W m⁻² used in MERRA and MERRA-2, respectively. These differences affect the overall energy budget of each dataset, as discussed in section 3a(3) below.

2) Surface

At first glance, the agreement between MERRA/MERRA-2 and EBAF is worse for the surface fluxes, as is generally expected. (The EBAF surface fluxes are inherently more uncertain because they are computed from multiple modeled and retrieved parameters, while the TOA fluxes are derived more directly from measurements.) Looking more closely, however, it is just the downward terms that deviate noticeably from EBAF. The discrepancies for surface ASWDN from MERRA

<table>
<thead>
<tr>
<th>Scale</th>
<th>SW up</th>
<th>LW up</th>
<th>SW down</th>
<th>SW up</th>
<th>LW down</th>
<th>LW up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly gridded</td>
<td>4</td>
<td>2.5</td>
<td>10</td>
<td>11</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Monthly global</td>
<td>—</td>
<td>—</td>
<td>6</td>
<td>6</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Annual global</td>
<td>1</td>
<td>1.75</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>3</td>
</tr>
</tbody>
</table>

Table 1. Uncertainty estimates for the EBAF irradiance products (W m⁻²). Monthly mean TOA values are taken from the EBAF TOA Ed2.8 data quality summary (CERES Science Team 2014), and the values for annual global means at TOA are derived by Loeb et al. (2009), while surface values are from Kato et al. (2012). Uncertainties are assumed to apply to both clear-sky and all-sky fluxes.
Table 2. Global, land, and ocean mean TOA radiation budget terms over the period Jan 2001–Dec 2015. All are defined as positive with the direction of propagation given in the variable name except for the net flux, which is positive when there is a net input into the Earth system (i.e., downward). TOA cloud radiative effect is defined as positive in the outgoing direction because this is more intuitive. Global mean differences that exceed the EBAF uncertainties of Table 1 and the corresponding flux values are indicated by bold type. Uncertainties are not applied to land and ocean fluxes.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>All-sky radiative fluxes</th>
<th>Clear-sky radiative fluxes</th>
<th>Cloud radiative effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SW down</td>
<td>SW up</td>
<td>LW up</td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERES EBAF</td>
<td>339.9</td>
<td>99.6</td>
<td>239.6</td>
</tr>
<tr>
<td>MERRA</td>
<td>341.3</td>
<td>98.9</td>
<td>242.9</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>340.5</td>
<td>106.3</td>
<td>238.2</td>
</tr>
<tr>
<td>MERRA–EBAF</td>
<td>1.4</td>
<td>−0.7</td>
<td>3.3</td>
</tr>
<tr>
<td>MERRA–EBAF</td>
<td>0.6</td>
<td>6.7</td>
<td>−1.4</td>
</tr>
<tr>
<td>Land excluding Antarctica</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERES EBAF</td>
<td>339.6</td>
<td>109.5</td>
<td>239.8</td>
</tr>
<tr>
<td>MERRA</td>
<td>340.8</td>
<td>104.2</td>
<td>241.0</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>339.9</td>
<td>107.1</td>
<td>239.6</td>
</tr>
<tr>
<td>Ocean</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERES EBAF</td>
<td>346.7</td>
<td>94.8</td>
<td>242.9</td>
</tr>
<tr>
<td>MERRA</td>
<td>347.9</td>
<td>95.9</td>
<td>247.0</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>347.0</td>
<td>105.1</td>
<td>240.8</td>
</tr>
</tbody>
</table>

and ALWDN from both reanalysis versions are greater than the EBAF uncertainties of 4 and 7 W m$^{-2}$, respectively. As we expect, the differences for ASWUP are small because the fluxes themselves are small, even though surface albedo is low in the reanalyses. The ALWUP differences should also be small, as observed, because the values of the primary controls of this variable (e.g., surface temperature and near-surface specific humidity) are taken from some version of the GEOS model for all three datasets. Thus, the question arises as to why the downward LW fluxes agree so poorly. The answer may again be differences in cloud parameters. Although most of the all-sky surface fluxes have improved going from MERRA to MERRA-2, the disagreement from EBAF in ATOTNET is much worse in MERRA-2. This is because the large biases in MERRA ASWDN and ALWDN are of opposite sign and thus offsetting, while the smaller biases in MERRA-2 reinforce each other.

Overall, the MERRA and MERRA-2 clear-sky surface radiative fluxes agree better with EBAF. The chief exception is CSWDN term in MERRA-2, for which the difference has increased from −0.9 to 5.0 W m$^{-2}$. Apparently, cloud effects offset the clear-sky deviation in this case. Since the signs of the clear-sky downward SW and LW flux differences are opposites, this large error has little effect on the CTOTNET, which has improved from the all-sky value by more than 5 W m$^{-2}$. The CSWDN is also too high in MERRA; however, this still represents an improvement from the all-sky flux deviation. As for MERRA-2, the clear-sky downward SW and LW flux errors cancel, yielding just a 1 W m$^{-2}$ difference in CTOTNET. The positive CSWDN biases exceed the EBAF uncertainties and reflect the fact that the reanalysis atmospheres are too transparent.

Unlike the values from the top of the atmosphere, the discrepancy in the surface net LWCRE from both MERRA and MERRA-2 is substantial, at about −7 W m$^{-2}$ out of the reference value of 28.9 W m$^{-2}$. This reflects the large differences from EBAF found for surface ALWDN. The magnitudes of the deviations in the net SWCRE are about half this value, leading to a disagreement in total net CRE of −4.4 W m$^{-2}$ for MERRA and −11.1 W m$^{-2}$ for MERRA-2. (The SW and LWCRE errors are of opposite sign for MERRA but of the same sign for MERRA-2.) The occurrence of such large discrepancies for global mean values indicates that there are definitely inaccuracies in the representation of clouds in both reanalyses. Clearly, there is more cloud in MERRA-2 than MERRA, although this does not affect the LW cloud effect greatly, suggesting that the additional clouds are at high altitude.

3) EFFECT OF DIFFERENCES IN SOLAR CONSTANT

In principle, the inconsistency among the TOA incoming solar flux values used by EBAF and the reanalyses could account for some of the biases in the other fluxes. The effect on the SW terms can be estimated by applying the atmospheric transmittance and reflectance values computed for the reanalysis products to EBAF’s global mean TOA ASWDN of 339.9 W m$^{-2}$. At the TOA, all of the changes to the upward and downward
Table 3. Global, land, and ocean mean surface radiation budget terms over the period Jan 2001–Dec 2015. All terms are defined as positive with the direction of propagation given in the variable name except the net flux and cloud radiative effect, which are positive when there is a net input into the Earth system (i.e., downward). Global mean differences that exceed the EBAF uncertainties of Table 1 and the corresponding flux values are indicated by bold type. Uncertainties are not applied to land and ocean fluxes. "Trans." = Transmittance.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>All-sky radiative fluxes</th>
<th>Clear-sky radiative fluxes</th>
<th>Cloud radiative effect</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SW down</td>
<td>SW up</td>
<td>LW down</td>
</tr>
<tr>
<td>Global</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CERES EBAF</td>
<td>186.5</td>
<td>24.1</td>
<td>345.2</td>
</tr>
<tr>
<td>MERRA</td>
<td>193.4</td>
<td>23.7</td>
<td>334.9</td>
</tr>
<tr>
<td>MERRA-2</td>
<td>185.6</td>
<td>23.2</td>
<td>336.7</td>
</tr>
<tr>
<td>MERRA-EBAF</td>
<td>6.9</td>
<td>−0.4</td>
<td>−10.3</td>
</tr>
<tr>
<td>MERRA-2-EBAF</td>
<td>−0.9</td>
<td>−0.9</td>
<td>−8.5</td>
</tr>
</tbody>
</table>

|                  | SW down  | SW up  | LW down | LW up  | Total net | Albedo | Trans. | SW | LW | Total |
| Land excluding Antarctica |          |        |         |        |           |        |        |    |    |       |
| CERES EBAF      | 192.5    | 41.2   | 328.4   | 393.7  | 86.0      | 0.214  | 0.567  | 240.4    | 49.5   | 304.2   | 395.8  | 99.4      | 0.206  | 0.708  | −39.6 | 26.3| −13.3 |
| MERRA           | 204.1    | 40.6   | 320.1   | 392.5  | 91.0      | 0.199  | 0.599  | 245.5    | 47.8   | 302.5   | 392.3  | 108.0     | 0.195  | 0.720  | −34.3 | 17.4| −16.9 |
| MERRA-2         | 202.1    | 40.7   | 318.9   | 391.4  | 88.9      | 0.201  | 0.595  | 240.0    | 49.1   | 301.9   | 391.3  | 105.9     | 0.198  | 0.730  | −37.5 | 16.9| −20.6 |

| Ocean           | SW down  | SW up  | LW down | LW up  | Total net | Albedo | Trans. | SW | LW | Total |
| CERES EBAF      | 186.6    | 14.3   | 359.4   | 409.0  | 122.7     | 0.077  | 0.538  | 249.5    | 18.8   | 329.2   | 408.2  | 151.7     | 0.075  | 0.720  | −58.5 | 29.5| −29.0 |
| MERRA           | 192.1    | 14.4   | 348.6   | 409.9  | 116.4     | 0.075  | 0.552  | 253.8    | 19.7   | 325.3   | 409.7  | 149.7     | 0.078  | 0.730  | −56.5 | 23.1| −33.4 |
| MERRA-2         | 181.9    | 13.5   | 351.3   | 409.5  | 110.3     | 0.074  | 0.524  | 253.7    | 19.7   | 327.0   | 409.3  | 151.7     | 0.078  | 0.731  | −65.5 | 24.1| −41.4 |
fluxes are less than 0.5 W m\(^{-2}\), but improvement is only seen in ASWUP for MERRA-2, where the lower incoming flux combined with MERRA-2’s high planetary albedo reduces the discrepancy from EBAF. The differences between EBAF and MERRA’s clear-sky and all-sky net fluxes grow to –2.2 and –2.3 W m\(^{-2}\), respectively, due to loss of cancellation from the downwelling term.

The only all-sky flux that improves at the surface is ASWUP in MERRA, where the smaller downward flux compensates for high transmittance. However, the 0.8 W m\(^{-2}\) improvement still leaves a significant difference of 6.1 W m\(^{-2}\) from EBAF. Thus, smaller ASWDN values at the TOA lead to worse surface ATOTNET for both reanalyses. CSWDN is notably improved for both MERRA and MERRA-2. As under all-sky conditions, the MERRA and MERRA-2 atmospheres transmit more radiation than EBAF’s (14% and 19%, respectively), and the lower incoming fluxes compensate for these discrepancies. The CSWDN error is reduced by 1.0 W m\(^{-2}\) in MERRA and no longer exceeds the EBAF uncertainty, while the error in MERRA-2 is reduced by 8%. CSWNET improves correspondingly in both cases. Thus, except for clear-sky fluxes at the surface, the reanalysis radiative energy budgets benefit from their high solar “constants.”

b. Global mean time series

Particularly for climate studies, it is important that the reanalyses capture both the short-term and long-term variabilities of the radiative energy budget. We thus examine time series of top-of-atmosphere and surface global mean monthly flux anomalies (i.e., with seasonal cycles removed) over the entire 15-yr study period. Time series for the TOA and surface fluxes are presented in Figs. 1 and 2, respectively.

1) TOP OF THE ATMOSPHERE

The top-of-atmosphere outgoing clear-sky SW and LW flux anomaly time series (Figs. 1a,d) from the reanalyses generally agree well with the EBAF values. However, the agreement is somewhat better for MERRA-2, for which overall variability is within 0.02 W m\(^{-2}\) of EBAF and correlations are at least 0.7. Agreement between the reanalyses and EBAF is worse for the all-sky radiative components. This is caused not only by markedly larger monthly variations in both of the MERRA datasets, but also by multiyear trends in MERRA that are not seen in the EBAF data. These include a large downward trend in the reflected SW flux (Fig. 1b) between 2008 and 2016 and upward trends in the outgoing LW flux (Fig. 1e) over the entire time series, except for the period between 2003 and 2008. Interestingly, the SW and LW trends and excess short-term variability tend to cancel in both MERRA products, so the all-sky net flux variations (Fig. 1h) compare more favorably to EBAF. The SW and LW cloud radiative effects (Figs. 1c,f) are generally dominated by the all-sky fluxes because the clear-sky time series are less variable. As a result, the CRE variability is too large, particularly for MERRA. The multiyear trends in the all-sky fluxes also affect CRE; however, the combination of the clear-sky and all-sky LW fluxes leads to improved correlation for MERRA and a steep decline in correlation for MERRA-2, which lacks the slight upward trend found in the other two datasets. This is in contrast to the CTOTNET (Fig. 1g), where agreement is better than in the SW and worse than in the LW. Overall, MERRA-2 captures the temporal variability of the TOA fluxes represented by EBAF better than MERRA does. This includes the incoming SW flux (Fig. 1i) because both MERRA-2 and EBAF use measurement-based total solar irradiance, which includes the 11-yr sunspot cycle and other observed variations.

2) SURFACE

Because of the greater number of radiative flux components at the surface, only the most interesting global mean surface irradiance time series are shown in Fig. 2. As at the top of the atmosphere, the variability of the clear-sky SW downward flux (Fig. 2a) in MERRA is smaller than in either of the other two datasets. While propagation through the atmosphere increases the variability in CSWDN by 0.47 W m\(^{-2}\) in EBAF and 0.36 W m\(^{-2}\) in MERRA-2, the change is only 0.11 W m\(^{-2}\) in MERRA. These differences are likely due to the use of a static aerosol climatology in MERRA. On the other hand, ASWDN at the surface (Fig. 2b) is essentially the inverse of the TOA reflected flux for all three products. The variability is again too high for the reanalyses, and there are multiyear trends that differ from those in EBAF. In spite of this, both versions of MERRA show relatively good agreement in variability and correlations greater than 0.5 for ASWUP (Fig. 2c).

Correlations are much better for the all-sky LW flux components (Figs. 2d–f), except that there is essentially no correlation between MERRA-2 and EBAF ALWNET (Fig. 2f). This is likely due to a flat and then slightly increasing trend in MERRA-2, as opposed to slight decreases in MERRA and EBAF. The magnitude of the variations in ALWUP (Fig. 2e) is moderately smaller in the two reanalyses, which carries over into ALWNET. When all SW and LW flux terms are combined (ATOTNET; Fig. 2g), cancellation of errors brings MERRA-2 into reasonably good agreement with EBAF, but the large trends in MERRA’s ASWDN remain in the total net flux, now opposing a very slight upward trend in both EBAF and MERRA-2. This
yields a negative correlation between MERRA and EBAF, as well as 50% larger overall variations. The all-sky downward SW fluxes dominate the ATOTNET time series.

The net SWCRE (Fig. 2h) most closely resembles the ASWDN term; however, combination with the other SW components leads to reduced variability in the EBAF and MERRA-2 time series. The tiny anomalies in MERRA’s CSWDN term do not counteract the all-sky variations much, with the result that the variability in MERRA’s SWCRE is more than twice that found in EBAF, and the correlation with EBAF is essentially zero. In the LW, the shapes of the ALWNET and CLWNET (not shown) time series are more similar for all datasets. This leads to cancellation, so the magnitude of the LWCRE (Fig. 2i) variations is smaller than in the SWCRE, particularly for MERRA-2. The correlations between the satellite and the two reanalysis LWCRE time series are quite small because of differences in magnitude among the datasets, as well as differences in multiyear trends. Given that the ASWDN fluxes dominate the SWCRE time series, it appears that most SW errors can be traced to cloud differences. The same cannot be said for the LW fluxes.

c. Annual cycles of global mean fluxes

The annual cycles removed to create the global mean flux anomaly time series are depicted in Fig. 3. Again, only the most notable results are included. Note that TOA terms are shown in Figs. 3a–d, while surface terms are in Figs. 3e–i. As expected, the TOA reflected SW flux (Fig. 3a) is largest during the Northern Hemisphere
winter, when the sun shines on Antarctica and the clouds of the Southern Ocean, and smallest in the Northern Hemisphere late summer, when the sun is farthest from Earth and vegetation is at its greatest extent. (Lack of landmass in the midlatitudes minimizes the seasonal effects of vegetation in the Southern Hemisphere.) However, in MERRA, the late summer dip is minimized, and a springtime depression is exaggerated, resulting in two minima of similar magnitude. The steeper decrease in springtime is likely due to the use of a lower albedo for sea ice in the Northern Hemisphere in the earlier version of MERRA (Gelaro et al. 2017), while the reduced summer dip corresponds to lower cloud radiative effect, as discussed below. The annual cycle of the outgoing LW flux (Fig. 3b) is roughly opposite that of the reflected SW flux, with a lesser amplitude. In this case, differences from EBAF are evident for both MERRA and MERRA-2: MERRA shows a smaller magnitude all year, while MERRA-2 has a larger amplitude during Northern Hemisphere spring and a smaller amplitude the rest of the year. When these are combined with the annual cycle of the incoming SW flux (not shown), there are notable discrepancies in both the MERRA and MERRA-2 total net fluxes (Fig. 3c). The magnitude of the TOA ATOTNET seasonal cycle is consistently too small for MERRA-2, while the summer minimum is delayed by about 1 month in MERRA.

The annual cycle of surface ASWDN (Fig. 3e) closely follows that of the incoming SW flux at the TOA (not shown), with the overall magnitude determined primarily by cloud radiative effect on this term. Relative to EBAF’s, the magnitude of the cycle is slightly large in MERRA and notably small in MERRA-2, although the
phases match. Albedo acting on the ASWDN creates the annual cycle of the upward SW flux (Fig. 3f), giving it an overall shape similar to that of the TOA ASWUP, including the effect of lower springtime ice albedo in MERRA. Cancellation between the upward and downward terms shifts the timing of the summer minimum forward in EBAF’s and MERRA-2’s ASWNET (Fig. 3g) and flattens the peak that develops in the fall, producing an asymmetrical shape resembling that for the TOA ATOTNET. The deeper minimum in MERRA’s ASWDN is less affected by cancellation with ASWUP, so the annual cycle of global mean ASWNET in MERRA is only slightly shifted and is notably large from June through December. Since the surface net LW flux (Fig. 3h) has almost no annual variability, the annual cycle of the overall total net flux (Fig. 3i) largely duplicates that of the all-sky net SW flux.

The TOA (Fig. 3d) and surface (not shown) net SWCRE are nearly mirror images of each other, with MERRA’s magnitudes too small between July and December and MERRA-2’s magnitude too large in June and July but otherwise in good agreement with EBAF. As a result, neither reanalysis product includes the double peaks present in the EBAF annual cycle that are due to a low value in June. The surface net LWCRE annual cycle (not shown) deviates from zero even less that the ALWNET. Thus, clouds have very little effect on the seasonality of the global mean surface net LW flux.

d. Global distribution

Global maps of the long-term mean radiative fluxes from EBAF, MERRA, and MERRA-2 allow us to identify specific regions where the reanalysis fluxes differ from the satellite-based values. Examining multiple
radiative energy budget terms also provides a means of appraising the sources of the observed differences. Figures 4 and 5 show global maps of selected radiative fluxes and flux differences for EBAF, MERRA, and MERRA-2 and form the basis for this discussion.

1) SHORTWAVE FLUXES

Maps of four SW flux terms are given in Fig. 4. First is the reflected flux at the top of the atmosphere in Fig. 4a. From the EBAF map, we see that the natural pattern is for lower values over the tropical and subtropical oceans and higher values over the northwest Pacific, ITCZ, much of the tropical and subtropical land (southeastern Asia, the islands of the Maritime Continent, northwestern South America, and most of Africa), the Antarctic, and portions of the Southern Ocean. Since TOA reflected SW flux is determined mainly by clouds, surface albedo, and aerosols, the remaining three terms give a picture of how this pattern is created. The contribution of clouds can be deduced from patterns of...
SWCRE, as shown in Fig. 4b. Over much of the globe, the cloud radiative effect pattern matches that of the reflected SW, but with lower amplitudes. For some land areas, the story is different. Despite the high reflection shown for central and northern Africa, the eastern portion of the Arabian Peninsula, the Amazon, and Antarctica, there is low-to-moderate cloud effect in these regions. There is also a region of higher cloud effect off the east coast of North America that does not appear in the reflected SW flux.

Surface reflection, illustrated by the surface clear-sky upward SW flux in Fig. 4c, also contributes to the TOA ASWUP. Large surface albedos occur over the Saharan Desert and eastern Saudi Arabia, Greenland, and Antarctica. Somewhat higher-than-average albedos also occur over the Arctic, Antarctic marginal ice zone, and Mongolia. Summing these fluxes with the cloud reflection accounts for most of the observed TOA ASWUP pattern. The effects of aerosol and water vapor absorption and reflection can be deduced from the surface clear-sky
downward SW flux. In the EBAF CSWDN map of Fig. 4d, areas brighter than expected solely from consideration of solar geometry appear over high-elevation mountain ranges such as the Andes and Himalayas. Lower values occur where greater amounts of aerosol are found—biomass burning areas such as central Africa, the Amazon, and the islands of the Maritime Continent; dusty regions such as northern Africa; and polluted areas in southeastern China and northern India. Reflection from these aerosols also contributes to the TOA ASWUP, bearing in mind that reduced fluxes at the surface correspond to increased fluxes at the TOA.

We next examine the differences between the EBAF and reanalysis global flux patterns. In MERRA, the TOA ASWUP flux (Fig. 4a) tends to be high over the tropical oceans, Maritime Continent, and Amazon and low in mid-to-high latitudes, particularly over land. In MERRA-2, the flux deviations are generally more positive than in MERRA, with new areas of positive bias over the Arctic, Southern Ocean, and central Africa as well as a large area of extremely high values over the tropical warm pool, extending from the western Indian Ocean to the date line. The latter biases exceed 50 W m$^{-2}$ at some locations and are strongest during the boreal summer. Consistent with the results of Bosilovich et al. (2015), these new positive features and the overall increase combine to produce the large (6.7 W m$^{-2}$) bias in MERRA-2’s global mean TOA ASWUP.

Errors in the other SW terms should explain these biases. For example, based on the TOA SWCRE difference plots (Fig. 4b), it appears that differences in the representation of clouds account for most of the observed ASWUP biases, as their patterns are nearly identical. (Note that the difference plots are all on the same scale and that as defined in this paper, positive TOA SWCRE bias indicates the presence of excess “cloudiness,” either in terms of cloud fraction or optical depth, and vice versa.) However, there are a few locations where cloud differences cannot explain the MERRA biases relative to EBAF. Relative to the TOA ASWUP errors, the magnitude of MERRA SWCRE biases is smaller over North America, central Africa, and the oceans below 60°S but larger over the Amazon, the islands of the Maritime Continent, and precisely over the Mediterranean, Black, and Caspian Seas. In MERRA-2, cloud errors are smaller than the ASWUP errors over land in the northern midlatitudes and in the Arctic, and they are larger over the Amazon and the oceans below 60°S. As for the ASWUP, the SWCRE generally increases going from MERRA to MERRA-2, leading to the high global mean value discussed previously. We note that the magnitudes of the SWCRE biases are smaller in MERRA-2 than MERRA over Northern Hemisphere land, the eastern tropical Pacific, and the Mediterranean Sea, so the general increase does lead to improvement in some areas.

Differences between the reanalysis products and EBAF are much smaller for surface upward clear-sky SW flux (Fig. 4c), which serves as a proxy for surface albedo. (Albedo differences were also computed, but using CSWUP allows the presentation to be made consistently in W m$^{-2}$.) In MERRA, there is an overall tendency for the surface CSWUP to be high over the oceans and low over land. In particular, CSWUP is somewhat small in northeastern North America and the Arctic, as well as in Antarctica and the surrounding marginal ice zone, and high in the general vicinity of the Indian subcontinent. The differences over land, including Antarctica, are reduced in MERRA-2 such that MERRA-2’s average CSWUP over land agrees better with EBAF than MERRA’s does. The improvements over Antarctica and the Arctic suggest a better treatment of ice cover in MERRA-2, although it appears that identification of marginal sea ice around Antarctica is not consistent between the reanalyses and EBAF, as was also noted by Bosilovich et al. (2015). However, this inconsistency may also be due to the high degree of cloudiness in this area, which limits the clear-sky sampling in EBAF. Further material about issues related to sea ice is provided in the discussion section. The only notable discrepancy between the CSWUP and albedo patterns (not shown) is that the albedo over Antarctica is high for MERRA, while CSWUP is low, likely because CSWDN is also low.

The clear-sky downward SW fluxes at the surface (Fig. 4d) have a more interesting story to tell. In addition to latitude, surface CSWDN is controlled by altitude, atmospheric aerosols, and water vapor, in that order. The large fluxes caused by the high elevation of the Himalaya, Sierra Madre, southern portion of the North American Cordillera, African highlands, and Antarctica are obvious in the EBAF figure. The effect of aerosols and possibly water vapor are evident in tropical rain forest and desert areas as reductions in CSWDN relative to the latitudinal pattern. MERRA is unable to capture some of these features. In particular, aerosols appear to be a problem. This is not surprising, given that the aerosol input to MERRA is just a monthly average annual cycle. MERRA’s CSWDN is high in tropical rain forest areas, where particulate organic matter from biomass burning is common, and in India and southeastern China, where dust is mixed with anthropogenic aerosols. Colarco et al. (2010) noted differences between MODIS aerosols and those found in the dataset used as the basis for the MERRA aerosol climatology that correspond to these high biases. Low CSWDN
values are found over dusty regions, such as the Sahara Desert, the Tarim basin of China, and off the western coast of northern Africa. Corresponding excesses of aerosols in these areas were not found by Colarco et al. (2010). MERRA’s values are also low directly over Greenland and Antarctica. It seems unlikely that this is caused by inaccurate elevations, but it might be related to excess water vapor amounts. Finally, low fluxes are present precisely over the Mediterranean, Black, and Caspian Seas, a phenomenon that suggests the presence of excess aerosol. The sharp boundaries of this area are difficult to explain, but there is a corresponding area of excess TOA CSWUP (not shown), leading to the difference in TOA SWCRE noted over these same water features. Overall, high biases predominate, yielding a global mean more than 4 W m\(^{-2}\) greater than in EBAF.

In MERRA-2, no additional regional problems are apparent, and many of the deviations found in MERRA are reduced. MERRA-2’s CSWDN over dusty regions nearly matches EBAF and has decreased over India and Southeast Asia, where dust is part of the overall aerosol mix. These improvements were noted in Buchard et al.’s (2017) study of the impacts of assimilating aerosol into MERRA and attributed to better dust optical properties and the inclusion of MISR aerosol data over bright desert areas. Although CSWDN has decreased over the tropical forests of Africa, no improvement is seen over the Amazon or the tropical islands of Southeast Asia. At the same time, the deficits over Greenland and Antarctica are somewhat reduced. In spite of the improvements, the net result is that biases are positive over most of the globe, which results in a global mean difference of about 5 W m\(^{-2}\) between MERRA-2 and EBAF, again suggesting insufficient aerosol overall. A little mental addition allows one to see how biases in TOA SWCRE, surface CSWUP, and surface CSWDN combine to produce those in the all-sky TOA reflected SW flux.

2) LONGWAVE FLUXES

Comparisons between LW fluxes in EBAF and the reanalyses are shown in Fig. 5. Much like the latitudinal distribution of incident SW flux, EBAF shows a general pattern of high TOA outgoing LW flux (Fig. 5a) close to the equator, decreasing toward the poles. However, large sections of moderate ALWUP are evident in many areas along the equator. In addition, the ALWUP is much lower in the Antarctic than the Arctic. Relative to EBAF, MERRA exhibits higher ALWUP values in most parts of the globe, particularly over the African rain forest and the ITCZ. The only significant low values occur over the oceanic tropical, Caribbean, and African warm pools, extending into the Amazon rain forest. These do not completely offset the high values, which results in a positive bias of about 3 W m\(^{-2}\) in the global mean. ALWUP is universally lower in MERRA-2 than MERRA, leading to a neutral global mean, but worsening the low biases in warm oceanic areas.

As in the SW, differences between the other flux terms in the satellite and reanalysis data can be used to explain much of the error detected in the outgoing LW flux. We first examine the surface emitted clear-sky upward LW flux (Fig. 5c), which would translate directly to the TOA outward flux without the presence of clouds. Since LW fluxes are controlled largely by emitting temperature, there is a nearly geometric pattern as a function of latitude in the EBAF surface CLWUP, with the primary differences being low values over regions of high elevation. In nonpolar areas, the agreement between the reanalyses and EBAF is excellent over the oceans, while the signs of the errors over land are mixed. In particular, low values occur over the drier areas of the globe, while high values occur over tropical rain forests, particularly in central Africa. These discrepancies are likely due to differences between the skin temperature \(T_{\text{skin}}\) values used in the EBAF and reanalysis CLWUP computations. Although the original source of \(T_{\text{skin}}\) in EBAF is also GEOS, differences can be created as the EBAF inputs are adjusted to obtain better agreement with the TOA fluxes. In particular, Draper et al. (2015) have demonstrated low \(T_{\text{skin}}\) values over dry land areas and high values over central Africa in GEOS-5 during the boreal summer and speculated that these are corrected in the EBAF adjustments (Draper et al. 2018). There are noteworthy positive biases in CLWUP over the Arctic and Antarctic Oceans and negative biases over polar land areas, with a clear demarcation of these biases at $\pm 60^\circ$. Possible reasons for flux differences at high latitudes are addressed in the discussion section.

The TOA LW cloud radiative effect (Fig. 5b) shows how the fluxes from the surface are modified during propagation through the atmosphere. The effect of clouds on LW radiation is more complex than on SW because both cloud temperature (height) and amount come into play: high clouds decrease outgoing LW flux, while a lack of high clouds exposes the surface or low clouds. Since the temperature of low cloud tops may be similar to that of the surface, it can be difficult to draw conclusions about the presence of low clouds from TOA LWCRE. Nevertheless, we can see that the patterns of LWCRE and ALWUP in EBAF are well correlated. There is little CRE over the SE Pacific and southern Atlantic and Indian Oceans, the African and Middle Eastern deserts, and polar areas, indicating the dominance of either clear skies or boundary layer clouds. Larger (negative) cloud effect occurs over the tropical
warm pool and tropical rain forests. MERRA captures the majority of this pattern, except for lower CRE over the warm tropical oceans and higher over the northern branch of the eastern Pacific ITCZ. The LWCRE is more negative nearly everywhere in MERRA-2, most notably a much broader swath of strong negative CRE in the tropical warm pool that was also identified by Bosilovich et al. (2015). Strong negative CRE in both MERRA and MERRA-2 suggests excess high cloud in the areas of deep convection, with insufficient or lower clouds over much of the land area. This leads to an overall negative bias of about 3 W m\(^{-2}\) in MERRA-2. We note that the error patterns in the MERRA and MERRA-2 TOA LWCRE are very similar to their ALWUP errors, suggesting that cloud errors are more important than surface errors. However, in some locations, such as the Arctic and Antarctic, opposing errors in the surface CLWUP counter the CRE errors, eliminating some features of the CRE error maps.

The surface all-sky downward longwave flux (Fig. 5d) is an important component of the surface energy budget that is only weakly coupled to TOA fluxes. This is because only the lowest kilometer or so of the atmosphere contributes to surface ALWDN. We had previously found that the MERRA and MERRA-2 global mean surface ALWDN had large biases relative to EBAF: approximately \(-10\) and \(-9\) W m\(^{-2}\), respectively. The plots in Fig. 5 show that these strong biases are spread over most of the globe. The deficits are particularly large over the Antarctic continent, Himalayan Mountains, stratocumulus areas, and Southern Ocean. In going from MERRA to MERRA-2, the magnitudes of the biases over the aforementioned areas are reduced, which decreases, but does not eliminate, the global mean negative bias. This is in part due to an increase in negative biases over northern Africa, the Arabian Peninsula, and India, and the elimination of slight positive biases over some tropical areas, which are not offset by the new positive bias over the Arctic Ocean. It is notable that overall, the regional biases for surface ALWDN in the reanalyses are larger than for any of the other LW terms.

e. Long-term mean land and ocean energy budgets

It is clear from this discussion that the performance of MERRA and MERRA-2 differs over ocean and land. Indeed, given the differences between the properties of land and ocean and how they interact with the atmosphere, it would be surprising if the reanalyses represented the radiative energy budget over both surfaces equally accurately. We thus review the long-term aggregate radiative fluxes above the land and ocean in MERRA, MERRA-2, and EBAF over the same January 2001–December 2015 time period. Land and ocean averages of the TOA radiative flux components are listed below the global means in Table 2. Note that the land and ocean flux values presented cannot be combined to yield the exact global means because of the elimination of coastal and Antarctic grid cells. In addition, the global mean uncertainties provided for EBAF are unlikely to characterize these aggregated fluxes owing to the smaller spatial extents considered and possible consistencies in errors across them. Thus, significance is not ascribed to the computed biases.

1) Top of the atmosphere

Our first observation is that the high all-sky reflected SW flux in the MERRA-2 global mean stems entirely from the oceanic regions: the TOA albedo over ocean areas is 11% higher in MERRA-2 than EBAF, producing an excess of about 10 W m\(^{-2}\) in the reflected flux, while the value over land is slightly low. The TOA SWCRE is correspondingly large over the ocean, implying strong excess cloudiness in MERRA-2, especially over the tropical oceans, as evident in Fig. 4. TOA ASWUP over land is somewhat low for both MERRA and MERRA-2. The slightly high all-sky outgoing LW flux in MERRA seems most closely related to an excess over the oceans under clear-sky conditions. TOA CLWUP was not pictured in Fig. 5, but has overall less contrast between areas of excess or insufficient flux than the surface CLWUP as well as positive biases of up to 10 W m\(^{-2}\) over most of the ocean. The fact that similar biases do not appear in the surface CLWUP suggests the presence of differences in the atmosphere above the ocean surface.

Breaking out the fluxes in this way reveals other land–ocean differences that are not visible in the global mean. For example, the all-sky reflected SW flux in MERRA is about 5 W m\(^{-2}\) low over land, while the MERRA-2 value is close to EBAF. Since the clear-sky error is similar in MERRA and MERRA-2, this indicates that the representation of clouds over land improved somewhat in the reanalysis upgrade.

2) Surface

The large negative bias in the global mean downward LW flux at the surface for both reanalyses manifests itself nearly equally over the land and ocean regions. At the same time, the biases in the corresponding clear-sky term are small, leading to large LWCRE errors. This supports the conclusion that weaknesses in the treatment of cloud processes in both MERRA and MERRA-2 worldwide cause the deficient surface ALWDN terms.

The situation with ASWDN is more complicated. In the global mean, this term is too large in MERRA but not in MERRA-2. It turns out that ASWDN is quite
high over land for both MERRA ($-12 \text{ W m}^{-2}$) and MERRA-2 ($-10 \text{ W m}^{-2}$); however, it is only high over the ocean in MERRA. In contrast, the mean ASWDN over the ocean is low by about 5 $\text{ W m}^{-2}$ in MERRA-2. The resulting cancellation of errors gives MERRA-2 a global mean ASWDN close to the value in EBAF. For both reanalyses and surface types, the CSWDN is somewhat (4–7 $\text{ W m}^{-2}$) too high, leading to SCRE errors that are small or moderate, except over the ocean in MERRA-2, where it is 7 $\text{ W m}^{-2}$ low. Together, these values point to overly transparent atmospheres globally in both models plus excessive cloudiness over the ocean in MERRA-2.

Other items of interest include differences in the biases in surface total net flux between the land and ocean areas. Under all-sky conditions over land, the high biases in ASWDN exceed the low biases in ALWDN, leading to positive total net flux errors. Over the ocean, the ASWDN errors are either smaller than or of the same sign as the ALWDN errors, summing to negative errors in ATOTNET. These negative errors exceed the positive errors over the land, yielding the observed low total net fluxes in the global mean. This occurs in both MERRA and MERRA-2, although the large negative error over the oceans in MERRA-2 gives the bias in the global mean a larger magnitude. A similar situation is present in both MERRA and MERRA-2 under clear-sky conditions. The positive biases in CSWDN over land are reinforced by modest negative biases in CLWUP, leading to large positive biases in the total net clear-sky flux. This is largely averaged out by the better agreement over ocean areas, so that the positive biases in the global mean CTOTNET are small. Finally, the surface albedos over ocean areas, this brings the flux more into line with the EBAF values throughout the tropics and the southern midlatitudes. The largest deviations have a magnitude of about 10 $\text{ W m}^{-2}$, again in MERRA-2.

2) SURFACE

The surface SW downward flux (Fig. 6c) has a bell shape due to the higher solar input over the equator than the poles, although a local minimum occurs due to clouds at the ITCZ. There is an additional dip around $-60^\circ$ where the Ferrell and polar cells meet, causing extensive cloudiness over the Southern Ocean. The shapes of all three curves are similar, but the MERRA flux is high except close to the poles, which leads to a high bias in MERRA’s global mean. The corresponding MERRA-2 flux is somewhat lower than the MERRA flux from the ITCZ down to the South Pole. In some areas, this brings the flux more into line with the EBAF values, but in certain areas, most notably around 60°S, the decrease overshoots the EBAF values. MERRA-2’s surface SW positive and negative differences relative to EBAF average out in the global mean. The maximum difference for MERRA is about 17 $\text{ W m}^{-2}$, close to 35°N/S, while for MERRA-2 it is $-22 \text{ W m}^{-2}$ at 60°S.

Surface ASWUP (Fig. 6e) agrees well among the three datasets. This is in part because upward SW fluxes are small. Poorer agreement is found beyond about 60° to the south and north. Based on the global maps of surface ASWUP differences (not shown), these differences occur mainly over the Arctic and the Antarctic marginal sea ice zone.

The global mean biases given in Table 3 are largest for the all-sky downward surface LW flux (Fig. 6d). ALWDN in both MERRA and MERRA-2 is consistently smaller than the EBAF fluxes over all latitudes. Although ALWDN fluxes are nearly equal close to the equator, they deviate by as much as 30 $\text{ W m}^{-2}$ toward the South Pole (27 $\text{ W m}^{-2}$ in MERRA-2). The ITCZ is not prominent in the surface ALWDN differences. Surface LW flux agreement is much better for the all-sky
Fig. 6. Comparisons between zonal mean all-sky TOA and surface radiative fluxes from MERRA, MERRA-2, and CERES EBAF over the period January 2001–December 2015. (a) TOA reflected ASWUP, (b) TOA outgoing ALWUP, (c) surface ASWDN, (d) surface ALWDN, (e) surface ASWUP, and (f) surface ALWUP. Note differences in the scales for SW and LW fluxes.
upward term (Fig. 6f). MERRA and MERRA-2 track each other nearly perfectly and match EBAF everywhere, except close to the poles. Their maximum deviation in surface ALWUP from EBAF is 13 W m$^{-2}$ close to the North Pole.

While in general the greatest differences between the reanalyses and EBAF occur for MERRA-2, in several cases (surface ASWUP, ALWDN, and ALWUP), the curves for MERRA and MERRA-2 are nearly the same. In addition, MERRA shows poor agreement over a wider latitude range than MERRA-2 for TOA ALWUP.

4. Discussion

There are several recurring themes in this analysis. Many of the comparisons provide evidence of excess clouds in MERRA-2. As also observed by Bosilovich et al. (2015), the excess clouds primarily take the form of too much high cloud over the tropical warm pool and other tropical oceans, but also occur as low clouds over the Southern Ocean. This leads to a positive bias of 7 W m$^{-2}$ in MERRA-2’s TOA reflected SW flux. While MERRA also exhibits too much cloud over the tropical oceans, it is to a lesser degree. Overall, MERRA has a more even balance of regions with positive and negative TOA ASWUP biases, leading to good agreement with EBAF in the long-term global mean. Errors in the cloud amount in the tropics are not surprising, as they are common in reanalyses (e.g., Dolinar et al. 2016). We note that a second typical error, insufficient cloudiness in the marine stratocumulus areas, is also found in the reanalyses, although to a lesser degree. Small areas of low TOA SW reflectivity are seen off the southwest coasts of North America and Africa in the upper plots of Fig. 4. However, corresponding features are not apparent in the outgoing LW flux, indicating that the missing clouds are from close to the surface. The magnitudes of the biases in these areas can exceed 30 W m$^{-2}$ in both reanalyses.

Another result identified by several of the analyses is that clouds are generally the biggest contributors to disagreement between the TOA outgoing SW and LW fluxes from the reanalyses and the EBAF data. In both spectral bands, differences in CRE are much larger than differences that may be due to aerosol or albedo errors. This is also not an unknown phenomenon in atmospheric models. The fact that clouds have large optical depths compared to aerosols and water vapor means that cloud errors will naturally have large radiative impacts.

The atmosphere in both reanalyses is found to be too transparent. CSWDN at the surface is too high in MERRA in regions of high biomass burning and sulfate emission, although simultaneously too low in dusty regions. While MERRA-2 improves CSWDN in the dusty and polluted areas, aerosol from biomass burning still appears to be lacking. On average, MERRA-2’s atmosphere is even more transparent than MERRA’s globally.

Weaknesses of the SW radiative transfer model used in the reanalyses likely contribute to this excess transparency. The Chou and Suarez (1999; SW) and Chou et al. (2001; LW) routines used in both MERRA and MERRA-2 performed relatively poorly in a test of GCM radiative transfer codes against line-by-line computations (Oreopoulos et al. 2012). For all of the clear-sky cases, TOA upward and surface downward SW fluxes were too high, indicating that the Chou and Suarez (1999) routine is both too reflective and too transmissive. The results of Oreopoulos et al. (2012) also showed high TOA upward LW fluxes under all conditions, while surface LW flux errors depended on precipitable water vapor (PWV), with the differences ranging from large and negative for an Arctic case with PWV of 0.32 cm to small and positive for a case with PWV of 4.85 cm. However, some of the errors noted by Oreopoulos et al. (2012) are only observed here for either MERRA or MERRA-2, even though the radiation codes remain the same. In addition, these observed biases are small, compared to regional biases found in this study. We can thus conclude that inaccuracies in the radiative transfer codes are only one factor contributing to the discrepancies reported here.

The difference between the methods of determining TOA clear-sky fluxes in EBAF and the reanalyses could have some impact on these comparisons. In particular, Sohn et al. (2010) have found that atmospheric columns that are truly clear generally have lower water vapor content than cloudy columns. Thus, TOA CLWUP from computations where clouds are artificially “removed” from the columns tends to be smaller than that measured for cloud-free columns. Their study showed that modeled TOA CLWUP can be as much as 10 W m$^{-2}$ lower in areas of deep convection, with corresponding impact on computed TOA LCWCRE. However, the spatial patterns of these differences shown by Sohn et al. (2010) do not closely match the patterns of differences between the MERRA reanalyses and EBAF in Fig. 5, so we conclude that although the computational differences may increase uncertainty, they do not dominate our results.

Numerous differences between the reanalyses and EBAF were noted in latitudes beyond 60°N/S. This may be caused by differences in environmental factors such as cloud amount, temperature, or water vapor in the Arctic and Antarctic. In particular, differences in the
assignment of snow and ice cover to surface grid cells are expected. Sea ice cover information for EBAF is taken from the National Snow and Ice Data Center’s Near-Real-Time SSM/I-SSMIS EASE-Grid Daily Global Ice Concentration and Snow Extent product, which has 25-km spatial and daily temporal resolution (Rutan et al. 2015). On the other hand, the sea ice concentrations used in MERRA are based on the 1° weekly sea surface temperature product of Reynolds et al. (2002), linearly interpolated to the model time step. In addition, MERRA-2 sea ice is taken from the 1/4° daily product of Reynolds et al. (2007) through 2006 and the 1/20° daily OSTIA dataset from Donlon et al. (2012) thereafter. Use of different sea ice data may contribute to flux differences in the sea ice margin of Antarctica, as suggested by the ragged edge of the surface CSWUP differences below −60° in Fig. 4.

Aspects of the determination of EBAF radiative fluxes in the polar areas can also affect agreement with the reanalyses. First, errors in the EBAF flux values can be caused by poor discrimination between clouds and high-albedo ice or snow surfaces, a problem that has been noted with respect to many satellite retrievals (e.g., Radkevich et al. 2013). Because clouds occur less frequently when snow or sea ice is present, the average surface albedo is higher under true clear-sky conditions (Kato et al. 2013). This could lead to an apparent low bias in the reanalysis surface CSWUP fluxes, as observed in Fig. 4. Additionally, the geostationary satellite observations used to fill out the cloud information between Terra and Aqua overpasses are not available beyond ±60° (Doelling et al. 2013). While the Terra and Aqua orbits maintain adequate temporal sampling here, this affects cloud retrieval because different algorithms are used for the narrowband channels of MODIS than the broadband imagers of the geostationary satellites. These inconsistencies can lead to discontinuities at 60°N/S, such as seen in the TOA LWCRE comparisons, particularly for MERRA.

The CERES team recently released EBAF Edition 4.0 (Loeb et al. 2018; Kato et al. 2018). Many modifications to the EBAF algorithms were made in this transition, such as increasing the temporal resolution of the underlying observations from 3-hourly to hourly; improving the estimation of nighttime cloud properties; and adding bias corrections for upper-tropospheric temperature, low-level cloud fraction, and overall cloud fraction to the derivation of surface fluxes. The interplay among these modifications is complex: a quick check shows that agreement between MERRA and MERRA-2 and the new EBAF flux products improves for some terms and locations and worsens for others. Thus, a thorough re-evaluation of the radiation budgets of the MERRA reanalyses in the context of this updated satellite dataset is in order.

5. Conclusions

This analysis provides an evaluation of the radiative energy budgets of NASA’s MERRA and MERRA-2 reanalyses and the differences between them. Fifteen years of radiative flux values from the CERES EBAF dataset are used as the comparison reference. Several developments are observed going from MERRA to MERRA-2. Whereas the TOA long-term global mean energy budget is nearly balanced in MERRA (bias of −0.5 W m⁻²), excess cloudiness over the tropical oceans leads to a large bias (6.7 W m⁻²) in reflected TOA all-sky upward SW flux in MERRA-2 that, in summation with a small negative bias in ALWUP, translates to a TOA imbalance close to −5 W m⁻². Since the atmosphere is too transmissive in both reanalyses, the extra TOA reflected SW flux in MERRA-2 provides compensation, such that the SW flux arriving at the surface is correspondingly improved (bias of ~7 W m⁻² reduced to ~1 W m⁻²). At the same time, the deficiencies in all of the long-term global mean TOA and surface LW flux terms are slightly reduced going from MERRA to MERRA-2. Nevertheless, the surface total net flux bias in MERRA-2 (~8.3 W m⁻²) is also worse than that in MERRA (~3.3 W m⁻²) due to loss of cancellation of errors. Generally speaking, MERRA-2 better represents the variability and trends of the global mean radiative fluxes over time, particularly by eliminating spurious multiyear trends seen in MERRA.

Regionally, the evolution from MERRA to MERRA-2 leads to mixed results. In many cases, the biases shift in the same direction over most of the globe, with varying effect on local agreement. For example, TOA all-sky reflected SW flux and CRE are more positive in MERRA-2 than in MERRA. This improves agreement over Northern Hemisphere land, but results in large positive biases over the tropical oceans and Southern Ocean, most likely due to increased cloudiness. Similarly, a general increase in atmospheric transmittance in the Northern Hemisphere that likely originates in MERRA-2’s assimilation of aerosol observations improves surface CSWDN over northern Africa, much of Asia and Europe, and the Mediterranean Sea, but increases the global mean bias. In short, the results of this analysis do not show that the radiative fluxes in MERRA-2 are clearly better than those in MERRA since both improvements and deterioration are found. Since clouds are clearly a source of significant flux errors in both MERRA and MERRA-2, a close examination of cloud properties in the reanalyses is warranted. Such
an analysis should benefit from the addition of a MODIS cloud simulator (Pincus et al. 2012) to MERRA-2. Further attention should also be given to improving both the satellite retrieval and model simulation of cloud properties and radiative fluxes in the polar regions.

Acknowledgments. Funding for this project was provided by NASA CloudSat and CALIPSO Science Team Grant NNX14AB27G. Global Modeling and Assimilation Office (GMAO) MERRA and MERRA-2 data were obtained from the NASA Goddard Earth Sciences Data and Information Services Center. CERES EBAF-Surface_Ed2.8 and CERES EBAF-TOA_Ed2.8 data were obtained from the Atmospheric Science Data Center at NASA Langley Research Center. Guidance were obtained from the Atmospheric Science Data and Information Services Center. CERES EBAF-TOA_Ed2.8 summary (March 19, 2014). CERES Tech. Rep., 52 pp., https://ceres.larc.nasa.gov/documents/DQ_summaries/CERES_EBAF-TOA_Ed2.8_DQS.pdf.


