Earth Syst. Dynam. Discuss., 3, C120–C133, 2012
www.earth-syst-dynam-discuss.net/3/C120/2012/


Earth System Dynamics Discussions

# *Interactive comment on* "On the determination of the global cloud feedback from satellite measurements" *by* T. Masters

**T. Masters**

tmasters@ucla.edu

Received and published: 12 April 2012

> *As I have done more preparation for this response, I am becoming less and less inclined to recommend publication of M12 in anything close to its present form. The fundamental approach of M12 is to analyze several data sets and show that they don't agree âĂŤ and therefore conclude that we don't know anything about what the actual value of the cloud feedback is and that Dessler 2010 is wrong. However, as I show below, a strong argument can be made that some of these data sets are less reliable than others âĂŤ and that the more correct data agree with the results of Dessler 2010.*

I appreciate the detailed look that the reviewer has given M12. However, I believe this



last statement is simply incorrect. Even if one were to assume the CERES Terra and Aqua results are "less reliable", and added the +0.3 W/m^2/K for the D10 adjustments, using the full EBAF dataset still yields a likely negative cloud feedback estimate. Furthermore, the ERA-Interim clear-sky forecast fluxes themselves over the Aqua period yield an estimated cloud feedback that is less than half the "likely" value expressed in D10, and this value will fall below the likely short-term cloud feedbacks estimated for all GCMs presented in D10. In fact, over the Aqua period where the different datasets are in best agreement, despite the fact that the net cloud feedback will depend on adjustments and the flux data source used, all datasets tend to be in agreement that the diagnosed SW component is negative, compared to the positive LW component. This is in contrast to D10, which diagnoses both components as positive. As one can see from my last response to the reviewer (Fig 1), the magnitude of that positive feedback estimate in D10 is greatly sensitive to both the clear-sky flux dataset and the time period used.

I should add that it is not that the different clear-sky datasets are in poor agreement - indeed, for most of the time period they agree quite well. Rather, it is that despite this general agreement, the subtle differences greatly affect the resulting cloud feedback estimate when regressed against T_surface. This is why the sensitivity tests are performed.

> *1. 2-m air temperature vs. skin temperature: This is a red herring. In climate models, the feedback obtained using the skin temperature and the 2-m temperature are nearly identical âĂŤ as expected since these variables must track each other closely. In updating the Dessler 2010 calculation (covering 2/00-12/10), I find that the cloud feedback changes from +0.58 to +0.56 W/m2/K as one switches from skin temperature to 2-m temperature using MERRA and from +0.49 to +0.38 W/m2/K using ECMWF.*

My last response detailed that the M12 results were confirmed when using the reanalysis surface (2-m) temperatures - that they were not an artifact of polar extrapolation in GISS or NCDC, as the reviewer seemed to imply. It was only using the specific combination of skin temperature and ERA-Interim that I was able to reproduce the more positive results in the reviewer's table. Contrary to the reviewer's suggestion, I never claimed that this had a major impact on D10, but rather that it impacted the estimated feedbacks in the reviewer's table said to reveal "the same thing as was found by D10". Indeed, this fact is confirmed in the reviewer's latest table, which shows a negative influence of ~0.3 W/m^2/K when using the more appropriate surface-air temperatures.

*2. Clear-sky fluxes: I cannot reproduce the numbers in Table 1 of the author's response. I have therefore included an updated version of Table 1 from my original comment (as Fig. 1). It is clear that the calculations using Terra clear-sky fluxes lie at one end of the range, with EBAF and reanalysis agreeing closely at the other end, and with Aqua in between.*

Once again, this does not appear to be the case. As the revised manuscript shows, using EBAF clear-sky produces a negative cloud feedback over the full period, and the estimated feedback is similarly in the middle of SSF and the reanalysis in the Aqua period. Tables 1 and 2 below show the correlation matrices among the different sets, where it is clear that EBAF and SSF are in far better agreement than with ECMWF. Furthermore, as mentioned above, the Aqua period results are quite consistent among all datasets that the shortwave component is negative, suggesting that the positive ECMWF results over the early period are an abberation (similarly, the negative Terra SSF LW results are probably on the other extreme).

*M12 is entirely predicated on the assumption that we know nothing about which of the clear-sky data sets is best. However, it is possible to delve into the data to determine that the Terra CERES clear-sky fluxes likely have problems To begin, let's regress global average clear-sky longwave flux*



*anomalies vs. 2-m surface temperature anomalies for each data set. During the Aqua period, this yields: Terra: 1.71, Aqua: 1.79, EBAF: 1.91, AIRS: 1.99, ECMWF: 1.96 (all W/m2/K). This is an important result. In this type of analysis, the EBAF should inarguably be the best CERES clear-sky data set. EBAF has a more sophisticated clear-sky algorithm that is designed to capture information from partly cloud pixels, thus leading to a more accurate estimate. M12 argues that SSF1 has better long-term stability (which is true), but the results here and in Dessler 2010 come from a regression against surface temperature âĂŤ and warm and cool months occur throughout the data set. Thus, the impact of spurious trends on the results is small, as discussed in Dessler 2010. The AIRS clear-sky fluxes provide an independent confirmation from an entirely different method. Results are similar over the Terra period, although there are only three data sets to compare. These results suggest Terra CERES clear-sky fluxes should not be relied upon (and Aqua CERES may only be a little better). To investigate further, Fig. 2 shows the standard deviation of the clear-sky flux anomalies at each grid point over the entire time series for Terra, EBAF, and AIRS clear-sky fluxes. Both CERES data sets show large increases in variability just poleward of 60N in each hemisphere. This looks unphysical and there is no reason why clear-sky longwave flux variability should increase suddenly at this latitude, and indeed the AIRS data do not show it. However, there is a good explanation: the CERES data require a determination of whether the footprint is clear or not, and doing this accurately requires sunlight. As a result, it is much more difficult at high solar zenith angles or at night, conditions that frequently occur poleward of 60N. That would lead to exactly the variability pattern shown here. The AIRS, on the other hand, uses a cloud-clearing algorithm that does not require a clear/cloudy determination and, as a result, it does not show the same pattern. Between 60N and 60S, the EBAF agrees well with AIRS, while the Terra shows much higher variabil-*

*ity. This is as expected. The reanalyses, not shown, agree closely with the AIRS and 60N-60S EBAF. This confirms that the Terra CERES regressions should be considered less reliable.*

I disagree with the characterization that "M12 is entirely predicated on the assumption that we know nothing about which of the clear-sky data sets is best." Clearly, there are sensitivities to the time period of the regressions and temperature datasets as well as flux sources, with the D10 results on one extreme end, and the SSF1 degree Terra estimate on the other. That the ERA-Interim results are most sensitive to the start date (with the results becoming more and more negative as the start date is moved back) suggests that the singular result is far from "robust".

Elaborating on the point above, in Table 1 below I show the correlation matrix between the longwave clear-sky flux data over this Aqua period. The Terra_Aqua averaged flux is the one used over this period as described in M12, and any comparison against EBAF should use this, as EBAF uses data from both Aqua and Terra over this time period. The AIRS modelled clear-sky fluxes indeed provide an independent look, although these suffer from their own biases as discussed in M12 and the referenced Sun et al. (2011). There are several key points here:

a) From table 1, it is notable that the Terra, Aqua, and EBAF datasets all agree better with each other and even ERA-Interim than they do with AIRS, with Terra, Aqua, and Aqua-Terra in better agreement with AIRS than EBAF. It is indeed interesting that the modelled AIRS OLR shows higher correlation with the modelled ERA-Interim fluxes. However, the most important take-away point from Table 1, and from figure 2b of M12, is that the agreement between the OLR clear-sky flux anomalies over this time period is (as Reviewer 2 puts it) "remarkable". It is therefore questionable that analyzing the slight differences in the OLR component over this time period, when they make little difference in the global anomalies, would result in an accurate diagnosis of the cause of the major discrepancies in the SW component during the earlier period. Particu-



larly since the dry-bias and undetected thin clouds, to the degree that they affect the CERES product, would not have as much influence in the SW component, whereas the more questionable aspects of the reanalysis products (surface albedo, aerosols) would influence the SW portion.

b) As mentioned in the last response, the EBAF clear-sky product will be included in the revised paper, and using these results confirms the sensitivity described in M12 . Still, the claim that EBAF is definitively better for this purpose is dubious. EBAF indeed reduces much of the sampling error in cloudy regions by supplementing CERES observations with those from MODIS, but in doing so introduces a narrow-to-broadband error. As mentioned in the last response, this conversion caused issues in the LW clear-sky component until it was recently fixed in 2.6r of EBAF, where it was clear that the discrepancy between CERES SSF and EBAF deseasonalized clear-sky anomalies resulted from the EBAF clear-sky algorithm, not SSF. This spurious drift (over a short period) resulted in a difference of 0.41 W/m^2/K in the estimated cloud feedback between 2.6 and 2.6r when regressed against GISS, so the claim that the stability would not significantly affect the estimate because D10 simulated a single, long-term drift over the entire period would appear incorrect.

c) The reviewer consistently points out that the variability is higher in the grid points of Terra and Aqua products, which is to be expected in a) measured vs. modelled values and b) infrequent sampling in the cloudy regions, per the M12 discussion. However, there is no suggestion by the reviewer as to why this increased sampling noise, which produces little difference in the latter period, would be correlated with T in such a way as to create a bias in the global mean deseasonalized anomalies, leading to the substantially higher correlation. From table 2, it is also clear that EBAF and Terra are in much better agreement for this short-wave component than they are with ERA-Interim.

d) Figure 1 below shows the net CRF for Terra between 60N and 60S versus the global mean, as the reviewer has concerns about the polar regions. There is excellent agreement between the two, particularly in the first part of the record, where the ERA-Interim and Terra values diverge. Regressing this 60N to 60S flux against NCEP temperatures averaged over the same region yields an apparent negative feedback that is smaller in magnitude (-0.63 $W/m^2/K$) than with global means, with a slightly more negative LW feedback but a more positive (less negative) SW feedback to counteract this. The LW result demonstrates what I was discussing in point (d) above - that the greater variability in this polar region is not creating a bias in the estimate. A less negative result for the SW component is expected when using only this region, as the surface albedo bias in the CRF anomalies is almost entirely restricted to the polar regions, and excluding them would effectively remove the need to make any positive adjustment. On the other hand, the ERA-Interim result becomes more negative (albeit barely) when excluding the polar region, which should raise questions marks – if ERA-Interim were properly capturing the surface-albedo changes at the poles, this non-cloud influence on CRF should overwhelm the cloud changes in this region.

> *3. In my original comment, I pointed out that the M12's criticisms of the methodology (including "robustness") were essentially bald-faced speculations unsupported by any analysis. In response, the author simply restated his original speculations and shows no sign that he will fix this. Unless he can demonstrate real issues with the methodology, all of that should be removed from the paper. And I note that issues of short-term feedbacks versus long-term feedbacks are adequately discussed in Dessler 2010, so they need not be repeated here.*

As the reviewer notes, this ground was covered in the last response. I will expand, however, on why using the all-sky kernels in the adjustments to $r_{cloud}$, for surface albedo in particular, is highly questionable. Almost the entire surface albedo interannual variation / contributions to CRF come from the polar regions, with the majority of it coming from 75S to 60S. GFDL CM2.1 (from which the Soden kernels are derived), and indeed most models, have huge (~33%) errors in the modelled all-sky SW flux



over this region [See figure 8.4a in AR4]. I've included Figure 2 to show that there is a large bias (not simply error) in the GFDL CM2.1 outgoing SW over this region. Thus, you'd essentially be using the least reliable all-sky region to perform the cloud-masking, yielding problems even if the ERA-Interim forecast surface albedo were free of issues.

> *4. The author needs to review exactly what the rËĘ2 statistic tells us. rËĘ2 is proportional to slope: if the slope is zero, rËĘ2 must be zero; as the slope increases, rËĘ2 increases. Thus, the findings highlighted throughout M12 and the response that higher slopes have larger rËĘ2 does not mean what the author seems to think it does. The higher rËĘ2 does not mean the slopes are more reliable, nor does it mean that the CRF adjustment has injected information. The author needs to carefully scrub the paper to correct these misapprehensions.*

The $r^2$ value is affected by noise/extranous variables, and is quite is possible to have large slopes with low $r^2$ values, and small (non-zero) slopes with high $r^2$ values. If clouds were primarily responding to globally averaged surface temperatures during this period, then undoubtedly the $r^2$ value would be higher. Without repeating my last response, an example of this is the correlation between the CRF adjustment and Ts ($r^2$=6.9%) versus the correlation between CRF and Ts ($r^2$=0.3%) in D10, despite having slopes that are similar in magnitude. Supposing all else is equal with respect to noise/extraneous variables, then indeed a larger slope will explain more of the variance, and hence yield a higher $r^2$ - I agree with the reviewer on this point, although it is not particularly relevant. This is because if the larger slope were the result of this increased sampling noise (as the reviewer seems to imply), then both the slope and noise would increase, leaving the SNR approximately the same, unless this sampling specifically creates a bias with respect to $T_{surface}$.

However, this has little to do with the actual paper. As the reviewer does not point to anything specific in M12, it is not clear what changes are requested. M12 makes several references to the low overall correlation between CRF and T_s, noting that this means other factors are influencing CRF besides T_s, which is clearly the case and to my knowledge is not in dispute by the reviewer. The only direct references to r^2 are in sentencing beginning on Page 80, line 14 and Page 80, line 21, both of which simply report the values without further comment.

> *5. The author's argument that one can take an anomaly of an anomaly (because they are calculated over different time periods) is simply incorrect. I defy the author to find an example anywhere in the peer-reviewed literature where anyone else has ever done this.*

As we both agree, this has little effect on the results. Nonetheless, taking an "anomaly of an anomaly" simply has the effect of re-baselining, which is done quite frequently in peer-reviewed literature (particularly when comparing temperature series with different baselines). Showing that this should be done with a numerical example is trivial:

a) Simulate a random, 40-year time series of monthly values, X, with a seasonal cycle added (if desired).

b) Assign X to Y (obviously, X and Y are now perfectly correlated).

c) Calculate the anomaly of X relative to the entire period, THEN grab only the last 10 years.

d) Grab only the last 10 years for Y, THEN calculate the anomaly over this 10-year period.

e) Compare (c) against (d). Note there is no longer perfect correlation, due to taking anomalies relative to different periods (the monthly averages are different).

f) Now calculate the anomaly of (c) relative to this last 10 year period (i.e., take the "anomaly of anomaly" for X)



g) Compare (f) against (d), and the perfect correlation is restored. This is because we have removed the "noise" from the change in seasonal cycle.

A script for the example may be found here: http://dl.dropbox.com/u/9160367/Climate/RebaselineExample.R

> *6. I wanted to add one last comment. I strongly recommend that the author adjusts all of the dCRF/dTs values in his paper to obtain cloud feedbacks. I'm afraid that there'll be confusion by those reading the paper and that readers will mistake the lower values of the dCRF/dTs statistic for the actual value of the cloud feedback. It will also make the discussion simpler to follow and make comparisons to Dessler 2010 more obvious.*

I understand the reviewer's concern here, and in the revised manuscript have added several more caveats noting that dCRF/dTs is not adjusted for the non-cloud influences, and that D10 found approximately +0.3 W/m^2/K should be added. However, given the issues mentioned in M12 section 2.4, and in #3 above, along with the lack of observations of interannual surface albedo changes, I don't believe we are able to accurately remove these effects.

Table 1. Correlation matrix of OLR from 9/2002 - 6/2011

| | Aqua_Terra | Aqua | Terra | EBAF | ECMWF | AIRS |
|---|---|---|---|---|---|---|
| Aqua_Terra | 1.00 | 0.97 | 0.97 | 0.94 | 0.85 | 0.82 |
| Aqua | 0.97 | 1.00 | 0.87 | 0.85 | 0.79 | 0.79 |
| Terra | 0.97 | 0.87 | 1.00 | 0.97 | 0.85 | 0.80 |
| EBAF | 0.94 | 0.85 | 0.97 | 1.00 | 0.86 | 0.78 |
| ECMWF | 0.85 | 0.79 | 0.85 | 0.86 | 1.00 | 0.87 |
| AIRS | 0.82 | 0.79 | 0.80 | 0.78 | 0.87 | 1.00 |



Table 2. Correlation matrix of SW component from 3/2000 to 6/2011

| | Terra | EBAF | ECMWF |
|---|---|---|---|
| Terra | 1.00 | 0.79 | 0.39 |
| EBAF | 0.79 | 1.00 | 0.45 |
| ECMWF | 0.39 | 0.45 | 1.00 |

Interactive comment on Earth Syst. Dynam. Discuss., 3, 73, 2012.

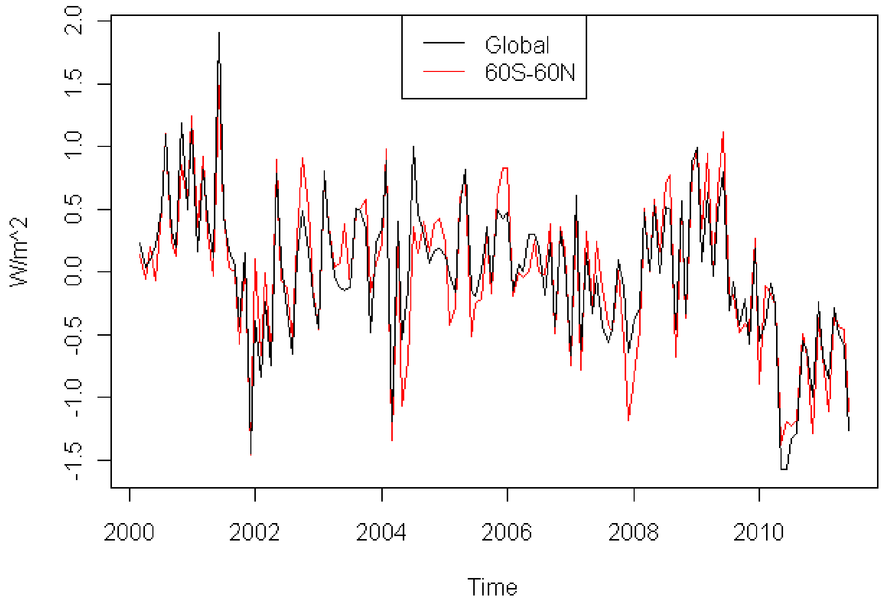


**Fig. 1.** CRF anomaly for CERES Terra (global vs. near-global)



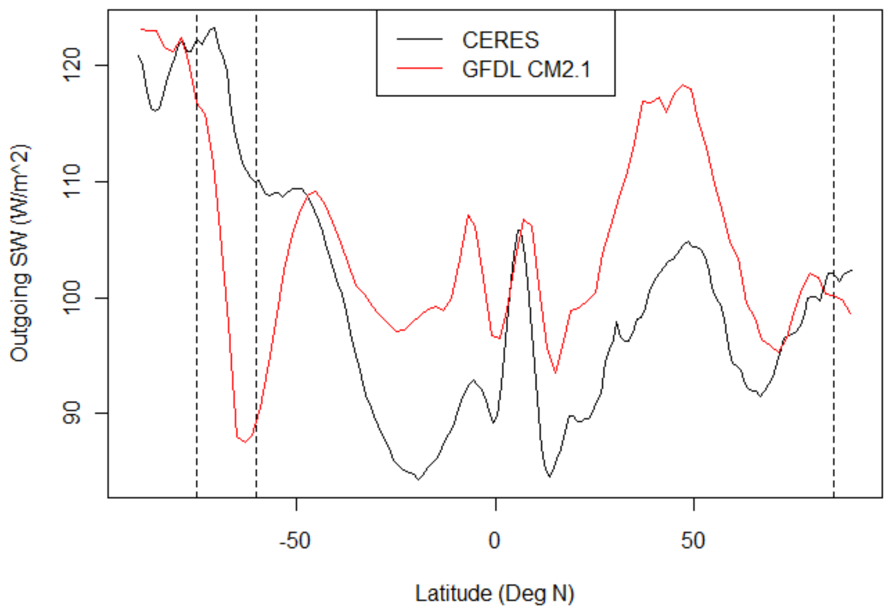


**Fig. 2.** All-sky outgoing SW flux by latitude