Analysis of the atmospheric energy budget:  
A consistency study of available data sets

Rucong Yu,1 Minghua Zhang, and Robert D. Cess
Institute for Terrestrial and Planetary Atmospheres, Marine Sciences Research Center
State University of New York at Stony Brook

Abstract. This study examines the following question: Are the recently available climate data sets consistent with each other in describing the atmospheric energy budget? If they are not, what is the sign of the systematic bias and how large is it? The atmospheric radiative cooling in this budget analysis is derived from satellite measurements of radiative fluxes at the top of the atmosphere and satellite-algorithm products at the surface. Atmospheric transports are derived from the National Centers for Environmental Protection/National Center for Atmospheric Research and the European Centre for Medium-Range Weather Forecasts reanalysis products. Surface turbulent heat fluxes are derived from products on the basis of Comprehensive Ocean-Atmosphere Data Sets and reanalysis. Spatial and temporal averaging is performed to reduce random uncertainties in each data set. It is found that these data sets result in an unbalanced atmospheric budget with a residual of 20 W m\(^{-2}\) when averaged from 50°N to 50°S. The sign and magnitude of this systematic bias are shown to be consistent with the recently debated insufficient absorption of solar radiation in the atmosphere.

1. Introduction

Within the last 10 years, there have been several major efforts to construct and reconstruct comprehensive climate data sets. These independently derived data sets provide the necessary information to examine all energy budget components within the atmosphere. The focus of this study is to combine these data sets and examine whether they are consistent with each other in forming a closed atmospheric energy budget.

Because of errors in instruments and measurements as well as algorithms, all data sets contain significant uncertainties. Gleckler and Weare [1997] provide a recent overview of uncertainties in the oceanic surface flux data set of Oberhuber [1988]. Uncertainties caused by random factors such as poor sampling and random noise can be reduced by averaging data in space and in time. Therefore the most serious uncertainty in climatological data is the systematic bias. It is often even not apparent whether there is a systematic bias in a given climatological data set because if it were known, it would have been corrected in the construction process of the data. A budget consistency check offers an independent way to reveal and quantify the systematic biases in the data sets, making the unknown systematic bias a known systematic bias.

Several previous studies have examined the global atmospheric energy budget using either historic or contemporary data. Recent overviews of these studies are given by Kiehl and Trenberth [1997] and Trenberth [1997]. In a typical study of the energy budget, balance requirement of the energy components is often used to either constrain some of the uncertain components [e.g., Kiehl and Trenberth, 1997] or to derive an unknown component [e.g., Ramanathan et al., 1995]. The thrust of the present study differs from these studies in that we do not constrain the data sets. Instead, we analyze the consistencies or inconsistencies of the widely used data sets. It is found that atmospheric energy components in the recently available data sets, whatever combination is used, do not balance with each other. On average, the atmosphere needs an additional energy source of about 20 W m\(^{-2}\) to balance its budget.

The paper is organized as follows: Section 2 describes the data sets and the individual budget components. This section provides the background for the discussion of our main results in section 3. Section 4 contains a brief summary.

2. Data and Component Analysis

The zonally averaged atmospheric energy budget equation, when vertically averaged, can be written as

\[
(R_{\text{TOA,SW}}^{-1} - R_{\text{TOA,LW}}^{-1}) - (R_{\text{SRF,SW}}^{-1} - R_{\text{SRF,LW}}^{-1}) + LH + SH + \frac{1}{a \cos \phi} \int [v(e_p T + gz + Lq)] \cos \phi \, d\phi = 0
\]

where \(R\) denotes net radiative flux, TOA denotes the top of the atmosphere, SRF denotes the surface, SW denotes shortwave, LW denotes longwave, and LH and SH are upward latent and sensible heat fluxes across the surface. The overbar in the last term denotes time-zonal average, and the bracket denotes vertical integration. Other variables are as conventionally used. The last term represents energy transport within the atmosphere across the latitudes. It disappears if it is integrated over the entire globe. In the present study we employ data to calculate each of the seven terms on the left side of (1). Since some data sets do not extend to high latitudes, we restrict our discussions to the region from 50°N to 50°S. We present results both for the zonally averaged energy budget and for the budget averaged from 50°N to 50°S.
In the following discussions, energy components are first described in the order that they appear in (1). For each energy component, multiple data sets are available from different sources. For example, latent heat fluxes over the oceans have been available from assimilation products [Kalnay et al., 1996], from ship measurements of meteorological variables [Oberhuber, 1988], and from indirect satellite measurements [Liu et al., 1979]. We adopt the following principle in selecting the representative data sets: (1) Radiative fluxes at TOA and surface are those derived from satellite measurements. (2) Atmospheric energy transports are derived from the reanalysis products. (3) Surface latent and sensible heat fluxes over the oceans are those most recently derived from ship measurements. This choice is not only due to our belief that these data sets are currently the best in representing the individual energy components but also because they are often used to evaluate the performance of general circulation models. Even with these selection criteria, for most of the components, more than one data set is available. In that case, we employ all available data sets that meet our selection criteria. For example, there are three data sets of surface shortwave radiation that were derived from satellite measurements. The Working Group on Existing Flux Estimates from the World Climate Research Program (WCRP) Air-Flux Workshop described their accuracy as “probably comparable” [World Climate Research Program, 1996]. In our budget analysis we use the mean values as well as the extremes of the multiple data sets. The extremes are used to indicate the data ranges.

2.1. TOA SW and LW

Shortwave and longwave radiative fluxes at TOA are taken from the Earth Radiation Budget Experiment (ERBE). The data set is the S4 product that has combined measurements from three satellites: ERBS, NOAA9, and NOAA10. The data are available from February 1985 to December 1989 at a resolution of 2.5° by 2.5°. Annual mean and long-term climatology of the fluxes are derived from the monthly climatologies. The 5-year zonal means of the downward shortwave and upward longwave fluxes are shown in Figures 1a and 1b. The net TOA radiation represents an energy source for the Earth-atmospheric system between 35°N to 35°S, and then it becomes an energy sink at other latitudes.

Uncertainties in the ERBE data arise mainly from instrument errors, sampling biases, and errors from the conversion of radiance to irradiance using bidirectional models. Barkstrom et al. [1989] gave an error bound of 5 W m⁻² for the monthly regional fluxes. A more detailed analysis of the uncertainties in ERBE data has been made by Rieland and Raschke [1991]. Part of the errors, such as sampling errors and the bidirectional model errors, is random. If the magnitude of random error for variable \( x \) is \( \sigma_x \), then after averaging over \( N \) grids the magnitude of random error would become \( \sigma_x / \sqrt{N} \). For a zonally averaged quantity over the period of the ERBE S4 data, the total number of monthly data points at each latitude is \( N = 59 \times 144 = 8496 \). Even with a random error of \( \sigma_x = 100 \) W m⁻² for a 2.5° by 2.5° grid box, \( \sigma_x / \sqrt{N} < 1.5 \) W m⁻². Thus the impact of random errors on the climatology is negligible.

2.2. Surface SW and LW

The difficulty of constructing surface radiation climatology from direct measurements owes to the fact that the surface radiation changes dramatically with clouds and that the cloud impact is solar-zenith-angle dependent. As a result, measurements at one station cannot be easily extrapolated to other
stations or to other times. Thus a measurement network with high time and space resolutions is needed to construct a reliable climatology. Accurate surface radiometers, however, are hard to maintain. The most comprehensive data set of available direct measurements of surface insolation is the Global Energy Balance Archive (GEBA) [Ohmura and Gilgen, 1991, 1993]. Even these data have little coverage over the oceans and in the Southern Hemisphere [Whitlock et al., 1995]. Furthermore, one needs to know the upward reflected shortwave radiation in addition to insolation to obtain the net surface shortwave radiation. This is lacking over most of the GEBA stations. The surface radiative fluxes derived from ship measurements [e.g., Da Silva et al., 1994] are based on visual cloud amount observations and empirical algorithms and therefore are not direct radiation measurements.

We therefore resort to products that are derived from extrapolating satellite measurements from the TOA to the surface using radiation algorithms. Among this category, the first data set is the Surface Radiation Budget (SRB) from the NASA Langley Research Center (LaRC) [Whitlock et al., 1995, Darnell et al., 1996]. NASA LaRC has two SRB products on archive, one for a 46-month period referred to as the WCRP SRB product and the other from July 1983 to June 1991. The longer data set is an extension of the 46-month data set, with one difference being that the multiyear monthly averages of ERBE-derived clear-sky surface albedos are used for the months beyond the ERBE period. In addition, the longer data set contains the surface longwave radiation. We will therefore use the longer data set in our study. The surface net shortwave and longwave fluxes were derived from the Staylor algorithm [Darnell et al., 1988, 1992; Gupta et al., 1992]. The algorithm was constructed from a radiation transfer model. Input variables to the algorithm are atmospheric column precipitable water from the TIROS operational vertical sounder (TOVS), and cloud amount and cloud liquid and ice water paths from the International Satellite Cloud Climatology Project (ISCCP). The algorithm product has been compared with the available GOSAT measurements of Whitlock et al. [1995]. It was reported that over the GEBA stations, surface downward shortwave radiation in the SRB product is about 10 W m\(^{-2}\) larger than direct measurements.

The second data set of shortwave and longwave surface radiative fluxes is the Global Atmospheric Radiative Fluxes from the NASA Goddard Institute for Space Studies (GISS) [Zhang et al., 1995]. This product was also derived from a radiation model with inputs of atmospheric temperature, water vapor, and cloud structures from the ISCCP project. This product differs from the SRB in that it employs the radiation code in the GISS general circulation model (GCM) [Lacis and Oinas, 1991]. Zhang et al. [1995] have compared this product with surface measurements at some stations and have considered the data as reasonable, even though no quantitative estimation of uncertainty is given. This data set is available from April 1985 to January 1989 on the ISCCP equal-area grids.

The third data set is for shortwave only. It is the Surface Shortwave Radiation Budget (SSRB) data set from the Canada Center for Remote Sensing (CCRS) [Li and Leighton, 1993]. It was derived from an algorithm described by Li and Leighton [1993]. Input variables to the algorithm are the TOA albedo from ERBE, atmospheric precipitable water from the European Centre for Medium-Range Weather Forecasts (ECMWF), and the solar zenith angle. The algorithm itself is based on numerical results from a radiation code. This algorithm can be interpreted to calculate surface shortwave radiation from a radiation model that produces the ERBE TOA shortwave flux. The monthly data are available at the same resolution and same time period as those for the ERBE data.

It should be emphasized that all these surface radiative products were derived from radiation models using satellite TOA measurements. They are not necessarily equivalent to direct surface radiation measurements.

Figure 1c shows the zonally averaged climatology of the downward net surface shortwave radiation from the three data sets. The GISS ISCCP product gives consistently larger values than those in the Langley SRB product, while the CCRS product has the smallest surface net flux. When averaged from 50°N to 50°S, the GISS ISCCP, the SRB, and the CCRS data sets give 193.4, 188.3, and 185.2 W m\(^{-2}\), respectively. When the three data sets are averaged, the averaged values are very close to those in the LaRC SRB product. The averaged data are also plotted in Figure 1c. Uncertainty estimations of these data sets over all latitudes from 50°N to 50°S are not available.

Figure 1d shows the net upward longwave radiative fluxes at surface. The GISS ISCCP data show larger longwave fluxes in the subtropics and smaller values at high and low latitudes than those in the SRB product. Partly because of this latitudinal distribution, the averaged values from 50°N to 50°S in the two data sets are very close: 48.8 W m\(^{-2}\) in SRB and 51.0 W m\(^{-2}\) in GISS ISCCP.

The net total downward flux at the surface is calculated in two ways: one from using the means of all the data sets of longwave and shortwave fluxes, the other from the GISS ISCCP and the SRB alone. Theoretically, the net downward radiative flux at the surface can be constructed by using a combination of any shortwave radiation data set in Figure 1c with any longwave data set in Figure 1d. Since in both the SRB and the GISS ISCCP products shortwave and longwave are calculated using the same specifications of clouds, we consider it more reasonable not to match the different data sets for deriving the net radiative flux.

### 2.3. Sensible and Latent Heat Fluxes

We combine surface products derived from Comprehensive Ocean Atmospheric Data Sets (COADS) over oceans with the reanalysis surface products over land to construct the surface latent and sensible heat fluxes over the globe. This hybrid approach is based on the fact that surface fluxes over the oceans in the assimilation products are typically severely biased [e.g., Weller and Anderson, 1996], partly because of little observational input to the assimilation system over the oceans. On the other hand, no satisfactory direct or indirect flux measurements are available over land, so the assimilation products become a necessary choice.

One widely used oceanic sensible and latent heat flux data set is from Oberhuber [1988]. This data set was derived using the monthly mean sea surface temperature (SST), surface air temperature, and wind observations together with the Large and Pond [1982] formulation of the bulk formulas. Oberhuber also derived the shortwave radiation component in his data set on the basis of an empirical formula and visual observation of cloud amount. Oberhuber further tuned both the shortwave radiative flux and the turbulent heat fluxes in his data set so that the climatological net surface heat flux would agree with the oceanic heat transport. Specifically, the shortwave radiative flux was tuned down by 10%, and the Charnock’s constant in the bulk calculation was tuned up from 0.0064 to 0.032. This
tuning of the latent heat flux was justified to compensate for the underestimation of winds by using monthly mean winds in the bulk formulae (the often-called “classical method”). The Oberhuber monthly data set is available at a resolution of 2° by 2° from 1950 to 1979.

The second oceanic data set, also based on COADS, is from Da Silva et al. (1994). This data set differs from the Oberhuber (1988) data set in that the “sampling method,” instead of the “classical method,” has been used in the calculations of fluxes, in which meteorological measurements are used to calculate the fluxes before monthly averaging is performed. The original eddy transfer coefficients of Large and Pond (1982) were used and corrections have been made to the COADS winds caused by a bias in the erroneous World Meteorological Organization (WMO) Code 1100 Beaufort equivalent wind scale. The latent and sensible turbulent heat fluxes in the data set were not tuned, even though the total net heat flux in the Da Silva et al. data set was tuned to oceanic heat transport, which is not used in this study. The data set is available at a resolution of 1° by 1°, spanning from 1945 to 1989.

Figures 2a and 2b show the zonally averaged climatological turbulent heat fluxes over the oceans from the two data sets. The two data sets generally agree with each other. Averaged over 50°N to 50°S, the two data sets give latent heat fluxes of 101.3 and 99.5 W m⁻², respectively, and sensible heat fluxes of 7.3 versus 9.5 W m⁻².

The latent and sensible heat fluxes over land in the reanalysis products have been categorized as class “C” variables by Kalnay et al. (1996), which are highly model dependent. When we compare the oceanic turbulent heat fluxes in the reanalysis products with the Oberhuber (1988) and Da Silva et al. (1994) data sets, the reanalysis products give values about 1.2 times as large as the COADS-derived values. Since there is no reason to assume this same ratio to be valid over land, we decided to use directly the land surface heat fluxes in the reanalysis products as the available climatology. Figures 2c and 2d show the land surface turbulent heat fluxes in the two models. For the latent heat flux both products exhibit a well-defined peak at the low latitudes and a value of about 40 W m⁻² in mid-latitudes. The National Centers for Environmental Protection/National Center for Atmospheric Research (NCEP/NCAR) latent heat flux over the tropics is larger than
that in the ECMWF reanalysis, with a maximum difference of 30 W m$^{-2}$. Land occupies about 20% of the area at these latitudes; this difference contributes about 6 W m$^{-2}$ difference to the zonally averaged climatology at the latitude of maximum difference. For the sensible heat flux the NCEP/NCAR product gives values larger than that of the ECMWF reanalysis. The largest difference is in the subtropics, where the maximum difference reaches 30 W m$^{-2}$.

The zonally averaged surface turbulent heat fluxes over the globe can be constructed by combining either the Oburhuber [1988] oceanic data sets or the Da Silva et al. [1994] data set with either the NCEP/NCAR or the ECMWF reanalysis product. We have plotted all the combinations in Figures 2e and 2f for latent and sensible heat fluxes, respectively.

The mean is also plotted. Averaged from 50$^\circ$N to 50$^\circ$S, the four combinations give the latent heat fluxes 87.5, 88.5, 88.9, and 89.9 W m$^{-2}$ and the sensible heat fluxes: 14.0, 15.5, 16.4, and 17.9 W m$^{-2}$. While the differences among the data sets could reach about 10 W m$^{-2}$ at some latitudes, the spatial averaging has substantially reduced the differences.

2.4. Atmospheric Energy Transport

The atmospheric winds, temperature, and humidity in the NCEP/NCAR and ECMWF reanalyses are used to calculate the vertically integrated convergence of energy due to atmospheric transport. The two reanalyses are both available at a resolution of 2.5$^\circ$ by 2.5$^\circ$ with a vertical resolution of 17 pressure levels. High temporal resolution of 4 times per day is used in our calculation. The data used span from January 1985 to December 1989.

Several cautions need to be made in the transport calculation as discussed by Trenberth and Solomon [1994] and by Trenberth [1997]. One is that the reanalysis product generally does not conserve mass, so the spurious mass divergence needs to be removed. This is done by deriving the zonally averaged, vertically integrated climatological meridional wind and subtracting it from the daily meridional wind. The second is the treatment of topography. Since the reanalysis product is available at constant pressure levels, actual surface pressure needs to be used to calculate the vertically integrated quantities.

Figure 3 shows the transport convergence of atmospheric moist static energy. The two reanalysis products give very consistent results. This consistency is not particularly surprising in view of the consistency of atmospheric energy transports in a large number of general circulation models [Gleckler et al., 1995] when they are forced by observed sea surface temperatures. The atmospheric transport acts as an energy sink from 40$^\circ$N to 40$^\circ$S, with an averaged loss of about 20 W m$^{-2}$. At higher latitudes the transport acts as an energy source.

3. Budget Results

Without the actual knowledge about which data set is closer to the truth, we use the mean values of all data sets for each component to examine whether the energy components of the atmosphere meet the balance requirement described by (1). Figure 4a shows the zonal distributions of atmospheric radiative cooling, the surface turbulent heat fluxes, and the atmo-
must be explained by systematic biases in the data sets. For example, differences in the heat fluxes between the NCEP/NCAR or ECMWF reanalysis and the four combinations of the land and oceanic surface latent heat fluxes in the employed data sets will be used as one measure of data uncertainty. It is seen that even with the extreme combination of the data sets, the atmospheric energy budget still cannot be closed. The minimum required energy is about 20 W m⁻². This must be explained by systematic biases in the data sets.

As stated before, the range of data sets in each component was used as one measure of data uncertainty. For example, the four combinations of the land and oceanic surface latent heat fluxes between the NCEP/NCAR or ECMWF reanalysis and the Oberhuber [1988] or Da Silva et al. [1994] data set gave a range for the surface latent heat flux. These ranges are shown as error bars in Figure 4b. The error bars in the components propagate to the final column that describes the required energy source. It is seen that even with the extreme combination of the data sets, the atmospheric energy budget still cannot be closed. The minimum required energy is about 20 W m⁻². This must be explained by systematic biases in the data sets.

This required energy source would be reduced if the atmospheric energy loss in the data sets is smaller, that is, smaller radiative cooling or less heat transported out of the 50°N to 50°S domain. It could also be reduced if the energy gain in the current data sets is larger, namely, larger latent heat flux or sensible heat flux. In view of the magnitude of the atmospheric transport and the sensible heat flux, it is unlikely that systematic biases in these two components can entirely account for the magnitude of the residual. The most likely components that could explain the systematic bias would then be the latent heat and the atmospheric radiative cooling.

If the unclosed budget is caused by a systematic bias in the latent heat flux, then it must be an underestimation of latent heat flux in the current data sets. A 25 to 30% increase in surface latent heat flux is needed to close the energy budget. Latent heat flux can be independently examined by using the moisture budget of the atmosphere. On a long time average the difference between latent heat flux and precipitation should be equal to the atmospheric moisture transport using the two reanalysis products to calculate the atmospheric moisture transport as we did for the transport of moist static energy, and using the monthly precipitation from Xie and Arkin [1997] and Spencer [1993], we are able to obtain Figure 5, which describes the consistency of the moisture budget among the data sets. It is seen that averaged from 50°N to 50°S, the residual in the calculated moisture balance is quite small.

As stated before, the range of data sets in each component will be used as one measure of data uncertainty. For example, the four combinations of the land and oceanic surface latent heat fluxes between the NCEP/NCAR or ECMWF reanalysis and the Oberhuber [1988] or Da Silva et al. [1994] data set gave a range for the surface latent heat flux. These ranges are shown as error bars in Figure 4b. The error bars in the components propagate to the final column that describes the required energy source. It is seen that even with the extreme combination of the data sets, the atmospheric energy budget still cannot be closed. The minimum required energy is about 20 W m⁻². This must be explained by systematic biases in the data sets.

The most likely candidate of systematic bias is therefore in the atmospheric radiative cooling. Since little information is available about the uncertainty of surface longwave radiation, it is not clear whether the bias is in the longwave or shortwave. We point out, however, that several recent studies have suggested a significant underestimation of atmospheric shortwave absorption in current radiation models, either in clouds [Cess et al., 1995; Ramanathan et al., 1995; Pilowskie and Valero, 1995; Valero et al., 1997; Charlock et al., 1996] or in clear skies [Charlock et al., 1996; Arking, 1996]. When averaged over the globe, the underestimation from these studies is of the order of 20 W m⁻². Because the employed surface shortwave data sets were all derived from radiation models, correction to these models would have enhanced the atmospheric shortwave absorption and would have resulted in a budget that is very close to being in balance.

Figure 5. The same as Figure 4 except for the budget of atmospheric moisture.
4. Summary

We have calculated the energy budget components of the atmosphere by using the following recently available data sets: satellite measurements at the top of the atmosphere and satellite-derived surface radiative fluxes, the NCEP/NCAR and the ECMWF reanalysis products, and the surface turbulent heat fluxes derived from COADS. It is found that these data sets give an unbalanced energy budget for the atmosphere. Averaged from 50°N to 50°S, the atmosphere requires an additional energy source of about 20 W m⁻² to close its budget. Various combinations of multiple data sets are used, and the unbalanced budget applies to all combinations of these data sets.

We have also shown that the atmospheric water vapor budget can be reasonably closed using the existing data sets, and therefore the most likely candidate of systematic bias is the atmospheric radiative cooling rate, which could be related with insufficient absorption of solar radiation in the atmosphere. The sign and the magnitude of the residual shown here provide a basis to constrain the future revisions of the existing data sets.

That is, the existing data sets should represent the upper bounds for radiative cooling and atmospheric transport, and lower bounds for latent and sensible heat fluxes.

Acknowledgments. The authors wish to thank one of the anonymous reviewers, whose detailed and insightful comments have led to great improvements of the original manuscript. We thank colleagues at the NASA Langley Research Center for making the SRB data products available to us. Other data sets were obtained through the Scientific Computing Division at the National Center for Atmospheric Research. We also thank W. L. Gates and Peter Gleckler at the Lawrence Livermore National Laboratory and W.-C. Wang for useful discussions. This work was supported by the U.S. Department of Energy under grant DEFG0285ER00314 and by NASA under grant NAG53135.

References


(Received September 17, 1998; revised January 26, 1999; accepted February 10, 1999.)