

## Application of an Artificial Neural Network Simulation for Top-of-Atmosphere Radiative Flux Estimation from CERES

KONSTANTIN LOUKACHINE

*Science Applications International Corporation, Hampton, Virginia*

NORMAN G. LOEB

*Center for Atmospheric Sciences, Hampton University, Hampton, Virginia*

(Manuscript received 13 February 2003, in final form 21 May 2003)

### ABSTRACT

The Clouds and the Earth's Radiant Energy System (CERES) provides top-of-atmosphere (TOA) radiative flux estimates from shortwave (SW) and longwave (LW) radiance measurements by applying empirical angular distribution models (ADMs) for scene types defined by coincident high-resolution imager-based cloud retrievals. In this study, CERES ADMs are simulated using a feed-forward error back-propagation (FFEB) artificial neural network (ANN) simulation to provide a means of estimating TOA SW and LW radiative fluxes for different scene types in the absence of imager radiance measurements. In all cases, the ANN-derived TOA fluxes deviate from CERES TOA fluxes by less than  $0.3 \text{ W m}^{-2}$ , on average, and show a smaller dependence on viewing geometry than TOA fluxes derived using ADMs from the Earth Radiation Budget Experiment (ERBE). The ANN-derived TOA SW and LW fluxes show a significant improvement in accuracy over the CERES ERBE-like fluxes when compared regionally.

### 1. Introduction

The objective of the Clouds and the Earth's Radiant Energy System (CERES) is to provide global radiative flux estimates at several levels from ground to the top of the atmosphere (TOA) together with coincident cloud and aerosol properties in order to improve our understanding of how clouds and aerosols affect climate (Wielicki et al. 1995). To achieve these goals, broadband radiance measurements from CERES are combined with radiances from a high-resolution, multispectral imager on the same spacecraft. To estimate reflected shortwave (SW) and emitted longwave (LW) radiative fluxes at the top of the atmosphere, CERES radiances are converted to TOA fluxes using empirical angular distribution models (ADMs) that account for the angular dependence in the radiation field. Since the anisotropy of earth scenes is highly variable, ADMs are defined as a function of scene type. Recently, Loeb et al. (2003) developed ADMs for over 200 scene types by combining CERES measurements with cloud and aerosol retrievals inferred from the Visible Infrared Scanner (VIRS) on the Tropical Rainfall Measuring Mission (TRMM) satellite. The new ADMs were shown to sig-

nificantly improve the accuracy in TOA fluxes compared to the ERBE ADMs (Loeb et al. 2002, hereafter LLM). ADMs for the Earth Radiation Budget Experiment (ERBE) (Barkstrom 1984) were defined for 12 scene types using a maximum likelihood estimation (MLE) technique (Wielicki and Green 1989). The ERBE MLE uses SW and LW ERBE radiance measurements together with a priori information from the cloud algorithm on *Nimbus-7* to determine scene type over a field of view (FOV).

In this study, the new CERES/TRMM ADMs are simulated using a feed-forward error back-propagation (FFEB) artificial neural network (ANN) simulation technique to provide TOA flux estimates directly from CERES measurements without using imager data. The idea is to use the ANN to relate the CERES SW and LW radiances and viewing geometry directly to the SW and LW anisotropic correction factors provided by the new CERES/TRMM ADMs. The motivation for this study is to provide an improved method for estimating TOA fluxes from instruments on spacecraft that have no imager data [e.g., the Earth Radiation Budget Satellite (ERBS)]. The methodology can also be used to infer TOA fluxes when the imager coverage over a footprint is insufficient for reliable scene identification, or if the imager flying alongside a broadband instrument fails prematurely during the mission.

In the following, the FFEB-ANN application is de-

---

*Corresponding author address:* Dr. K. Loukachine, SAIC, One Enterprise Parkway, Suite 300, Mail Stop 927, Hampton, VA 23666.  
E-mail: k.loukachine@larc.nasa.gov

scribed in detail. The technique is then validated by comparing ANN-derived TOA fluxes with TOA fluxes inferred from the new CERES/TRMM ADMs. In addition, various TOA flux consistency tests are applied to demonstrate the improvement in TOA flux accuracy from the ANN method compared to TOA fluxes obtained by applying the ERBE ADMs.

## 2. Observations

The CERES/TRMM satellite was launched on 27 November 1997 in a 350-km circular, precessing orbit with a 35° inclination angle. TRMM has a 46-day repeat cycle, so that a full range of solar zenith angles over a region is acquired every 46 days. On the TRMM satellite, CERES has a spatial resolution of approximately 10 km (equivalent diameter) and operates in three scan modes: cross-track, along-track, and rotating azimuth plane (RAP) mode. In RAP mode, the instrument scans in elevation as it rotates in azimuth, thus acquiring radiance measurements from a broad range of viewing configurations. An unprecedented level of calibration stability ( $\approx 0.25\%$ ) between in-orbit and ground calibration (Priestley et al. 1999) has been provided by CERES/TRMM. Unfortunately, the instrument suffered a voltage converter anomaly and only acquired 9 months of science data: 192 days of cross-track data, 9 days of along-track data, and 68 days in RAP scanning mode. All CERES/TRMM data are collected in the tropical region between 38°S and 38°N latitude.

The CERES/TRMM Single Scanner Footprint TOA/Surface Fluxes and Clouds (SSF) product with CERES in RAP mode is considered for building the ANN training sets. Nine months of CERES SSF data product, from January to August 1998 and March 2000 (all scanning modes), are used for validation of the ANN-derived SW and LW TOA fluxes. The SSF product combines CERES radiances and fluxes with scene identification information inferred from coincident high spatial and spectral resolution VIRS measurements (Kummerow et al. 1998), and meteorological fields based on European Centre for Medium-Range Weather Forecasts (ECMWF) data assimilation analysis (Rabier et al. 1998). A comprehensive description of all parameters appearing in CERES SSF is provided in the CERES collection guide (Geier et al. 2001). Only CERES footprints that at least partially lie within the VIRS swath and whose centroids can be located on the earth's surface are retained in the SSF product. Since VIRS scans in the cross-track direction to a maximum viewing zenith angle (VZA) of 49°, CERES footprints with VZAs greater than 49° appear in the SSF product when CERES scans in either RAP or along-track mode.

The ANN-derived TOA fluxes are compared with TOA fluxes of SSF and CERES "ERBE-like" products on a footprint-by-footprint basis. The CERES ERBE-like product is produced in order to extend the historical record of earth radiation budget observations by pro-

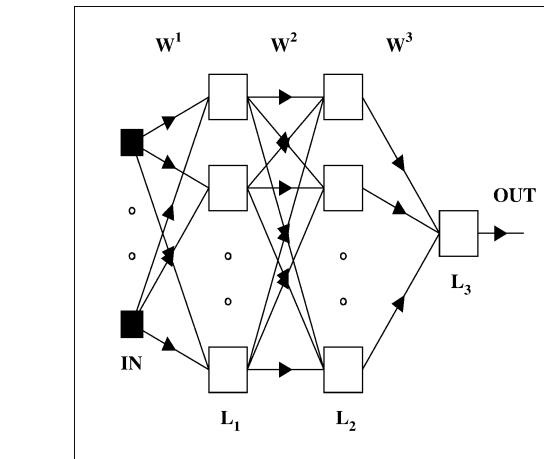


FIG. 1. Layout of feed-forward ANN with error back-propagation used in this study. Filled rectangles show the ANN normalized inputs  $v_1, \dots, v_n$  (IN), and open rectangles show the the ANN neurons, which form two hidden layers,  $L_1$  and  $L_2$ , and an output layer,  $L_3$ . The anisotropic correction factor is the ANN output (OUT).

cessing CERES measurements with algorithms developed during ERBE (Smith et al. 1986). Note that since the ERBE-like product is produced independent of VIRS, all CERES FOVs (including those outside of VIRS swath) are retained in this product. However, in order to compare TOA fluxes from the ERBE-like and SSF products, only CERES footprints common to both products are considered in this study.

## 3. Methodology

### a. Radiance conversion to flux

To obtain the ANN-derived SW and LW TOA fluxes from CERES measurements, we train ANN to reproduce anisotropic correction factors (ACFs) as a function of selected variables,  $v_1 \dots v_n$ , provided by the new CERES/TRMM SSF ADMs (Loeb et al. 2003). Knowing the ACF, we convert a CERES broadband radiance to a TOA flux as follows:

$$F_{\text{ANN}}(v_1, \dots, v_n) = \frac{\pi I}{R_{\text{ANN}}(v_1, \dots, v_n)}, \quad (1)$$

where  $F_{\text{ANN}}(v_1, \dots, v_n)$  denotes ANN-derived TOA flux, and  $I$  and  $R_{\text{ANN}}(v_1, \dots, v_n)$  are the measured broadband radiance and ANN-derived ACF, respectively.

### b. ANN structure

In this study, a simulated FFEB multilayer ANN is used to reproduce CERES/TRMM SSF anisotropic correction factors. This neural network configuration is generally applied as a highly nonlinear approximator (Hagan et al. 1996; Bose and Liang 1996). The general layout of the ANN simulation is shown in Fig. 1. For our purposes, an ANN is simulated with the normalized inputs,

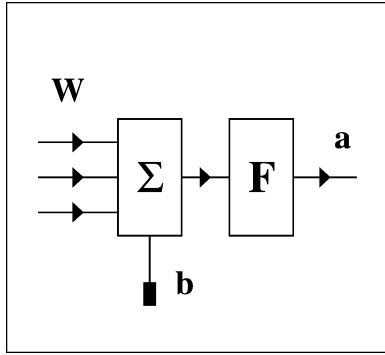


FIG. 2. An artificial neuron structure: weighted inputs ( $W$ ), integrating part ( $\Sigma$ ), activation function ( $F$ ), bias ( $b$ ), and output ( $a$ ).

$v_1, \dots, v_n$ ; two hidden neuron layers,  $L_1$  and  $L_2$ ; and one output neuron layer,  $L_3$ . It has been shown (Cybenko 1989) that any absolutely integrable function can be approximated by ANN having only one hidden sigmoidal layer, provided a sufficient number of neurons are needed. However, for ANN with only one hidden layer, the number of neurons can be impractically large (Bose and Liang 1996). Each neuron of the  $n$ th layer is connected with all neurons of the previous layer with an assigned weight. The structure of a simulated artificial neuron is shown in Fig. 2. Each neuron unit performs three simple operations: sums the inputs, adds the bias value, and applies the activation function to the result. Therefore, the feed-forward propagation of inputs through the network can be written in matrix form as

$$\mathbf{a}^n = \mathbf{F}^n(\mathbf{x}^n) = \mathbf{F}^n(\mathbf{W}^n \mathbf{a}^{n-1} + \mathbf{b}^n), \quad (2)$$

where  $\mathbf{a}^n$  is the vector of the  $n$ th layer output (in our case  $n = 0, 1, 2, 3$ , where  $\mathbf{a}^0$  represents the ANN input vector and  $\mathbf{a}^3$  is the ANN output),  $\mathbf{F}^n$  is the vector of the activation functions,  $\mathbf{W}^n$  is the weight matrix, and  $\mathbf{b}^n$  is the bias vector of the  $n$ th layer. Vectors are denoted by boldface serif, matrices are denoted by boldface sans serif, and their dimensions are provided in the appendix. For all neurons in the hidden layers ( $L_1$  and  $L_2$ ), a hyperbolic tangent is chosen as the activation function,

$$\mathbf{F}^{1,2}(\mathbf{x}^{1,2}) = \frac{e^{\mathbf{x}^{1,2}} - e^{-\mathbf{x}^{1,2}}}{e^{\mathbf{x}^{1,2}} + e^{-\mathbf{x}^{1,2}}}, \quad (3)$$

and a linear activation function,  $\mathbf{F}^3(\mathbf{x}^3) = \mathbf{x}^3$ , is chosen for the output neuron layer ( $L_3$ ). Both functions are defined over real space ( $\mathbf{x}^n$  are vectors of real numbers).

Generally, the FFEAN simulation includes three basic steps: 1) definition of the training sets, 2) supervised training of the network on the training sets to reproduce the target values within desirable error bounds, and 3) after successful training, freezing of the ANN parameters and testing on the data of interest.

### c. Training sets

The CERES/TRMM SSF dataset contains millions of CERES footprints, and it is impossible to use the data

TABLE 1. Variables, stratification intervals, and normalization factors for the SW ACF training set. Solar zenith angle is denoted as SZA, relative azimuth angle as RAZ, and longwave and shortwave CERES broadband radiance as LWR and SWR, respectively.

Variable	$N$ bins	Bin width	Normalization
VZA	7	$10^\circ$	$90^\circ$
SZA	9	$10^\circ$	$90^\circ$
RAZ	9	$20^\circ$	$180^\circ$
LWR	15	$10 \text{ W m}^{-2} \text{ sr}^{-1}$	$150 \text{ W m}^{-2} \text{ sr}^{-1}$
SWR	30	$10 \text{ W m}^{-2} \text{ sr}^{-1}$	$300 \text{ W m}^{-2} \text{ sr}^{-1}$

directly for the ANN training. Therefore, we need to create compact training sets. However, to allow a good degree of ANN generalization the training sets must represent the complexity of data well. This can be achieved by stratifying the data in the variable of interest and using the corresponding means.

Five input variables are selected for the SW ACF approximation: solar zenith angle, viewing zenith angle, relative azimuth angle, and CERES LW and SW broadband radiance measurements. To create the training sets, the 68 days of CERES RAP data are stratified by these variables using the intervals shown in Table 1. For every scene type and in every interval, the mean value of all five variables and the mean CERES SSF anisotropic correction factor are calculated. A minimum number of CERES footprints (sampling threshold) is required in every data bin. In order to avoid operations with large numbers in the process of network training, the means of input variables are normalized to their maximum allowed value in the training set as shown in Table 1. The scene types are defined as a function of surface type, in a manner consistent with the angular distribution models developed by Loeb et al. (2003): glint ocean, no-glint ocean, medium-high tree-shrub, low-medium tree-shrub, dark desert, and bright desert. The number of intervals, their widths, and sampling thresholds for every scene type are chosen by a compromise between reducing natural noise, keeping good representation of the data, and the required computer processing time for the training process. A training set with a reduced number of intervals does not represent the complexity of the data, and a training set with an increased number of intervals becomes noisy and requires a long time for the ANN training. The thresholds and number of points in the training sets for every scene type are shown in Table 2. Thus, each training set contains a number of input vectors of normalized mean variables, and a corresponding mean ACF target value.

The LW training set for all-sky ocean type is built in similar fashion as the shortwave sets. For the longwave ACF approximation, four input variables are used: viewing zenith angle, LW and SW broadband CERES radiance measurements, and surface skin temperature  $T_s$ , provided by the ECMWF. The stratification and normalization parameters of the longwave ACF training set for the all-sky ocean are shown in Table 3. This training

TABLE 2. Shortwave scene types, their corresponding sampling threshold for each training set, number of entry vectors in training set, and resulting error index per entry, mean bias, and standard deviation of the ANN output and training set ACF difference. The scene types are medium–high tree–shrub (MH/TS), low–medium tree–shrub (LM/TS), dark desert (DD), and bright desert (BD).

Scene type (all-sky)	Threshold		E/N ( $\times 10^{-3}$ )	Bias (%)	STD (%)
	(FOV/ bin)	$N$ entry vectors			
No-glint ocean	250	16 255	2.40	0.34	5.84
Glint ocean	75	15 263	6.90	0.34	6.40
MH/TS	50	17 013	1.61	0.14	4.05
LM/TS	75	18 012	1.42	0.14	3.88
DD	30	16 761	1.80	0.16	4.22
BD	15	15 685	0.89	0.08	3.14

set is built with a 30 FOV/bin sampling threshold, and includes 38 042 training vectors.

#### d. Training algorithm

The goal of the training process is to adjust the weight and bias values such that the ANN reproduces the target values of ACF for an entire training set with minimal error. The derivation and description of a standard training algorithm for FFEB multilayer ANN, the generalized delta rule, is given in Hagan et al. 1996. Taking into consideration the specifics of our task we apply it as follows.

The weight matrices and bias vectors are initialized with random numbers from 0 to 1, and the training proceeds iterations over an entire training set. At first, within an iteration  $k$ , every training input vector  $\mathbf{a}^0$  is propagated forward through the ANN layer by layer [Eq. (2)] producing the output  $\mathbf{a}^3$ . The output values of every neuron are stored during the feed-forward phase. Next, the gradient matrix of the activation functions  $\dot{\mathbf{F}}^n$  is numerically calculated for each neuron layer  $n$ :

$$\dot{\mathbf{F}}^n = \frac{\delta \mathbf{F}^n(\mathbf{x}^n)}{\delta \mathbf{x}^n}. \quad (4)$$

For the chosen activation functions,  $\dot{\mathbf{F}}^3 = 1$ , and for the first and second hidden neuron layers the derivative can be expressed as a function of the layer output at iteration  $k$ :

$$\dot{\mathbf{F}}^{1,2} = (1 + \mathbf{a}^{1,2})(1 - \mathbf{a}^{1,2}). \quad (5)$$

Then, the sensitivity to the error,  $t - \mathbf{a}^3$ , is calculated for the output layer  $L_3$  as

$$\mathbf{S}^3 = -2\dot{\mathbf{F}}^3(t - \mathbf{a}^3), \quad (6)$$

where  $t$  denotes the corresponding ACF target value, and then it is back-propagated from  $L_3$  to the first hidden layer  $L_1$  by using

$$\mathbf{S}^{n-1} = \dot{\mathbf{F}}^{n-1} \mathbf{W}^n \mathbf{S}^n, \quad (7)$$

where  $\mathbf{S}^n$  denotes the sensitivity vector of neuron layer  $n$  ( $n = 3, 2$ ). For every entry in the training set, the

TABLE 3. Variables, stratification intervals and normalization factors for the all-sky ocean scenes LW ACF training set. Longwave and shortwave CERES broadband radiances are denoted as LWR and SWR, respectively.

Variable	$N$ bins	Bin width	Normalization
VZA	28	2.5°	90°
$T_s$	12	2.5 K	310 K
LWR	24	5 W m <sup>-2</sup> sr <sup>-1</sup>	120 W m <sup>-2</sup> sr <sup>-1</sup>
SWR	20	15 W m <sup>-2</sup> sr <sup>-1</sup>	300 W m <sup>-2</sup> sr <sup>-1</sup>

gradients of every weight and every bias parameter are calculated as

$$\Delta \mathbf{W}^n(k+1) = \gamma \Delta \mathbf{W}^n(k) - (1 - \gamma) \alpha \mathbf{S}^n(k) (\mathbf{a}^{n-1})^T, \quad (8)$$

$$\Delta \mathbf{b}^n(k+1) = \gamma \Delta \mathbf{b}^n(k) - (1 - \gamma) \alpha \mathbf{S}^n(k), \quad (9)$$

where  $\Delta \mathbf{W}^n$  and  $\Delta \mathbf{b}^n$  are gradients of weight and bias values, respectively;  $\alpha$  is learning rate;  $\gamma$  is learning momentum; and  $k$  denotes iteration number (at the first iteration  $\alpha = 0.1$  and  $\gamma = 0.6$ ). Then, the gradients of the weight and bias values are averaged over the training set to obtain the mean  $\overline{\Delta \mathbf{W}^n}$  and  $\overline{\Delta \mathbf{b}^n}$ . The error index at iteration  $k$  is calculated as

$$E(k) = \sum_{i=1}^N [t - \mathbf{a}^3(k)]^2, \quad (10)$$

where  $N$  is the number of entry vectors in the training set (see Table 2), and  $t$  is the ACF target value. At this point, to evaluate an iteration, the new weight and bias values are used for the ANN test:

$$\mathbf{W}_{\text{test}}^n = \mathbf{W}^n(k) + \overline{\Delta \mathbf{W}^n}(k+1), \quad (11)$$

$$\mathbf{b}_{\text{test}}^n = \mathbf{b}^n(k) + \overline{\Delta \mathbf{b}^n}(k+1). \quad (12)$$

The entire training set is propagated forward again [Eq. (2)], and  $E_{\text{test}}$  is calculated using the ANN test output  $\mathbf{a}_{\text{test}}^3$  [Eq. (10)]. The decision about the next iteration is based on the error index gradient,  $\Delta E_k = E(k) - E_{\text{test}}$ , for each training iteration. If the error index decreases,  $\Delta E_k \geq 0$ , the new weight, bias, and error index values are kept for the next iteration. The learning rate  $\alpha$  gets a small increase (0.001) until it reaches the maximum allowed value (0.5). If the error index increases,  $\Delta E_k < 0$ , the ANN proceeds to the next iteration with unchanged error index weight, and bias values, and the learning rate is set to a low value (0.05). The training continues until the error index over the training set reaches an acceptable value. At this point, the ANN weight and bias values are frozen and stored, and the network is ready for testing.

#### e. ANN testing

The ANN performance for all SW and LW scene types is tested on the entire CERES/TRMM SSF dataset (described in section 2), including all CERES scanning modes. First, input variables are normalized to the factors used to create the training sets (Table 1 and 3).



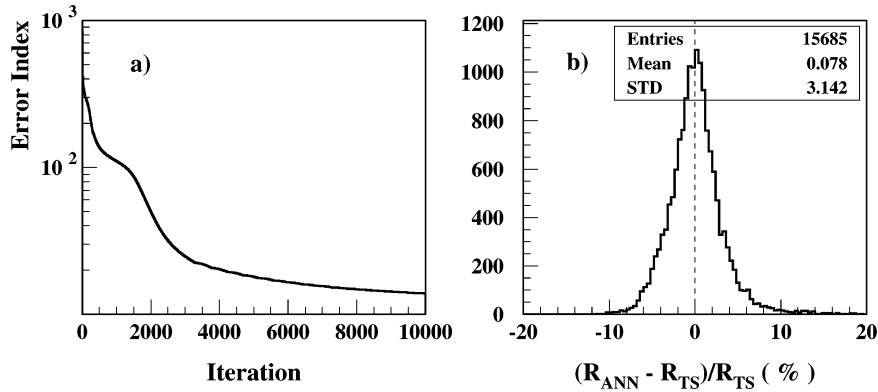


FIG. 3. Training results for all-sky bright desert scene type: (a) reduction of the error index over the training set with iteration number; (b) frequency distribution of the relative difference of the ANN and training set SW ACF,  $(R_{ANN} - R_{TS})/R_{TS}$ , where the numbers denote the size of the training set, mean, and standard deviation (STD) of the distribution.

Then, the normalized input data are propagated through the trained ANN [Eq. (2)] using stored parameters. The network output ACF value is used to obtain the ANN-derived TOA flux by applying Eq. (1).

#### 4. Results

##### a. Shortwave TOA flux

To approximate CERES SSF SW anisotropic factors, the ANN configuration involves 11 neurons in the first hidden neuron layer and 7 neurons in the second hidden neuron layer. The ANN training with a variable learning rate and constant momentum,  $\gamma = 0.6$ , terminates after  $10^4$  training iterations for all scene types. The value of the error index normalized to the number of vectors in the training set is shown in Table 2. Scenes over bright desert have the smallest training error, while ocean scenes that are affected by sun glint have the largest training error. The process of error index reduction with training iterations is illustrated in Fig. 3a for scenes over bright desert. In this case, the error index monotonically decreases with iteration number. Figure 3b shows the

frequency distribution of the relative difference between the target and ANN-derived SW anisotropic factors for the bright desert training set after  $10^4$  iterations. The mean bias of the distribution is 0.078%, and the standard deviation (STD) is 3.14%. The relative mean bias and STD for the other surface types are summarized in Table 2. As expected, the all-sky glint ocean case has the largest STD value.

Figures 4a, 5a,b and 6 compare SW TOA fluxes inferred from CERES radiances for all-sky conditions from the ANN approach, the new CERES/TRMM ADMs, and the ERBE ADMs for all of the surface types in Table 2. Since LLM have demonstrated that TOA fluxes based on CERES/TRMM ADMs are consistent with model-independent fluxes obtained by direct integration of the measured radiances, in this study we assume that the CERES SSF fluxes are representative of the true flux values. Figure 4a shows the mean all-sky SW TOA flux stratified by CERES VZA. Since TOA flux is an integral over all viewing directions, TOA flux should not depend on VZA. The ANN-derived SW TOA flux in Fig. 4a is closer to the CERES SSF fluxes

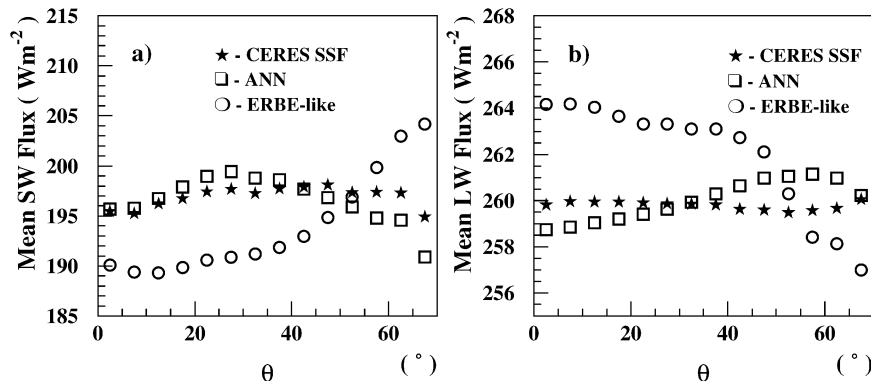


FIG. 4. (a) SW TOA and (b) LW TOA all-sky mean flux stratified in viewing zenith angle for CERES SSF (stars), ERBE-like (open circles), and ANN-derived (open squares) fluxes.

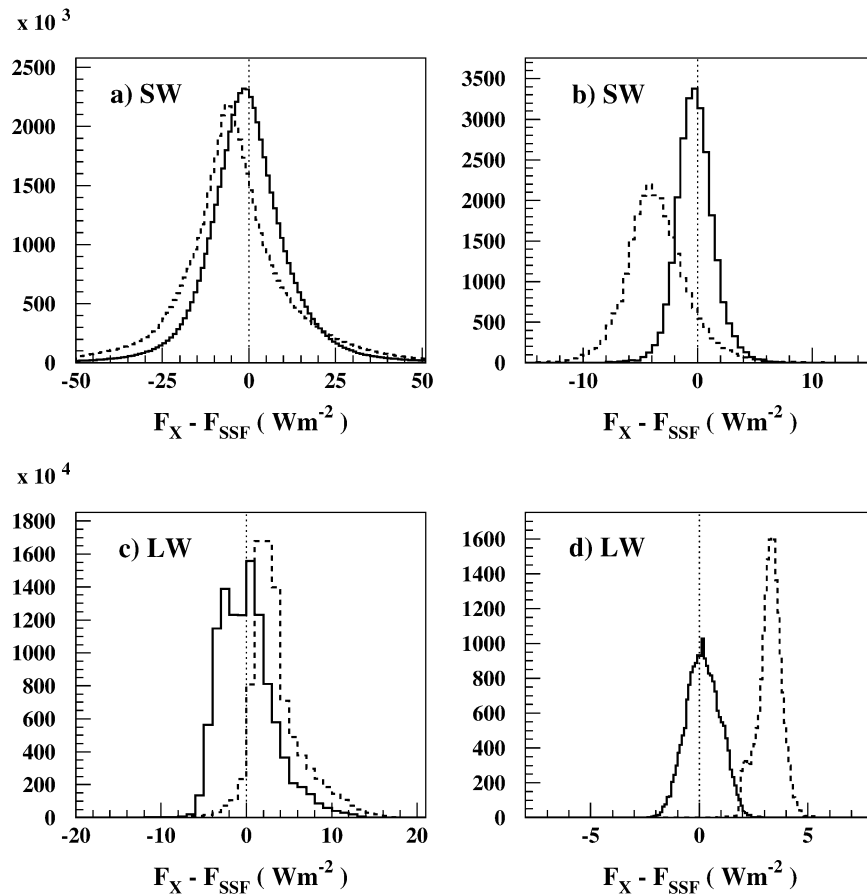


FIG. 5. (a) Frequency distribution of all-sky all-scene *instantaneous* SW TOA flux differences  $F_{ERBE} - F_{SSF}$  (dashed line) and  $F_{ANN} - F_{SSF}$  (solid line); (b) frequency distribution of all-sky all-scene *regional mean* SW TOA flux differences  $F_{ERBE} - F_{SSF}$  (dashed line) and  $F_{ANN} - F_{SSF}$  (solid line); (c) frequency distribution of all-sky ocean *instantaneous* LW TOA flux differences  $F_{ERBE} - F_{SSF}$  (dashed line) and  $F_{ANN} - F_{SSF}$  (solid line); (d) frequency distribution of all-sky ocean *regional mean* LW TOA flux differences  $F_{ERBE} - F_{SSF}$  (dashed line) and  $F_{ANN} - F_{SSF}$  (solid line). Regional resolution is  $1^\circ$  in lat and lon in both the SW and LW cases.

than ERBE-like, which shows a large dependence on viewing geometry. Clearly, the ANN result is a significant improvement over the ERBE ADMs.

Figure 5a provides the frequency distribution of instantaneous SW TOA flux differences between ANN-derived and CERES SSF fluxes (solid line) and between ERBE-like and CERES SSF fluxes (dashed line). On average, the ANN TOA fluxes differ by  $-0.3 W m^{-2}$  from the CERES SSF fluxes, whereas the ERBE-like TOA fluxes differ from the CERES SSF fluxes by  $-3.6 W m^{-2}$ . The standard deviations in the differences are  $14.2$  and  $17.1 W m^{-2}$  for ANN and ERBE-like, respectively. Figures 6a,b show the differences of the all-sky ocean TOA SW mean regional fluxes in the Tropics with a  $1^\circ$  latitude  $\times$   $1^\circ$  longitude grid. The frequency distribution of  $1^\circ$ -mean regional SW TOA flux differences is shown in Fig. 5b. On average, the ANN-derived TOA fluxes differ by  $-0.2 W m^{-2}$  from the CERES SSF fluxes, whereas the ERBE-like TOA

fluxes differ from the CERES SSF fluxes by  $-3.5 W m^{-2}$ . The standard deviations in the differences are  $1.7$  and  $2.8 W m^{-2}$  for ANN and ERBE-like, respectively. In both cases, instantaneous and regional means, the ANN-derived fluxes are more accurate than the ERBE-like. We point out that the CERES SSF dataset is dominated by a VZA less than  $49^\circ$  due to the requirement that VIRS data are available over every CERES FOV. Based on Fig. 4, this means biases at a VZA less than  $49^\circ$  dominate. The actual ERBE-like product includes all CERES footprints, so that errors at large VZAs compensate for the errors at small VZAs in that dataset. The regional plots in Fig. 6 indicate a cloud pattern for both ANN-derived and ERBE-like flux differences with SSF flux. The ANN method tends to overestimate SW TOA flux in the intertropical convergence zone and underestimate it in stratus regions off Peru, Angola, and California. The regional trend for ERBE-like SW fluxes is the opposite.

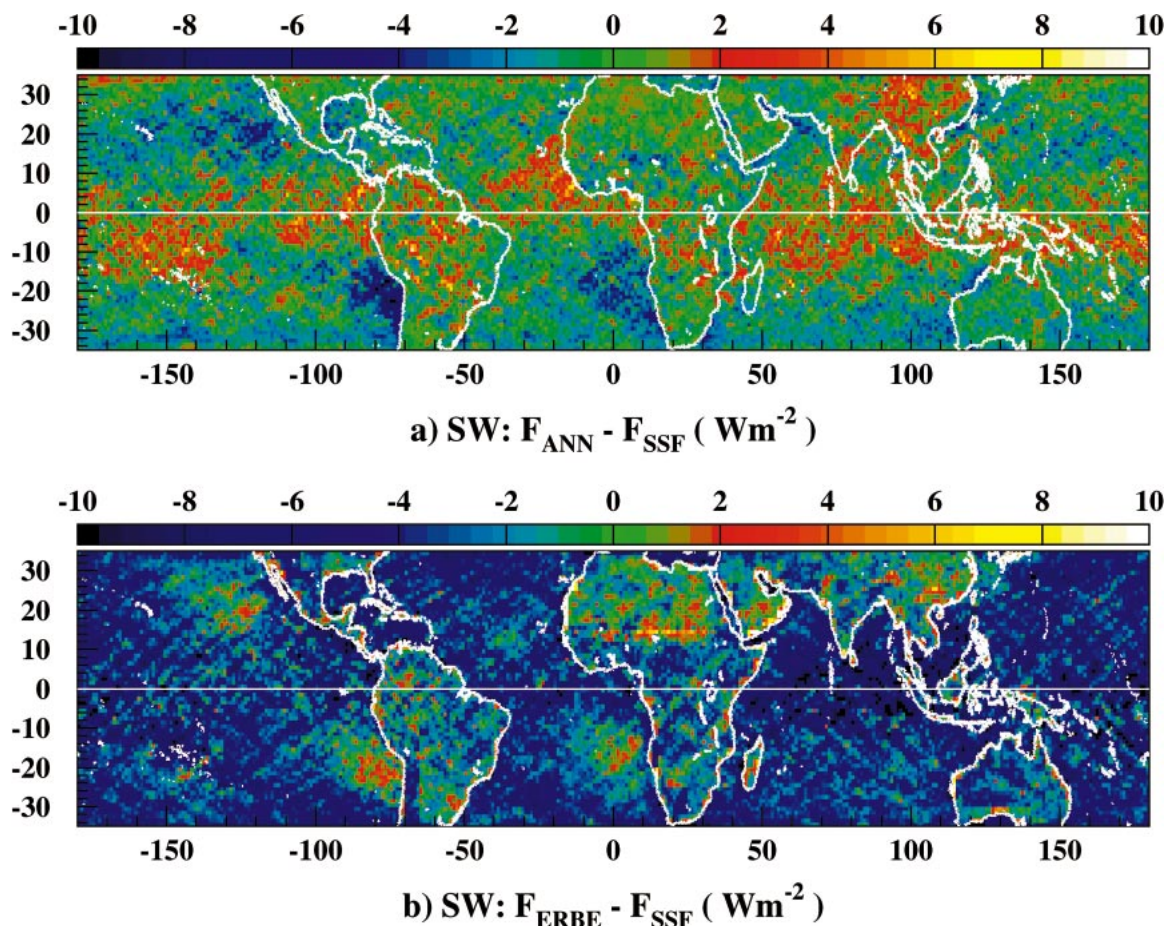


FIG. 6. All-sky mean regional TOA SW flux difference for (a) ANN-derived and SSF fluxes,  $F_{ANN} - F_{SSF}$ , and (b) CERES ERBE-like and SSF fluxes,  $F_{ERBE} - F_{SSF}$ . Regional resolution is  $1^\circ$  in lat and lon.

### b. Longwave TOA flux

We illustrate the ANN-based LW TOA flux retrievals using all-sky ocean CERES footprints. To approximate CERES SSF LW anisotropic factors, the ANN configuration involves 4 neurons in the first hidden neuron layer and 13 neurons in the second hidden neuron layer. The ANN training with a variable learning rate and constant momentum,  $\gamma = 0.6$ , terminates after  $10^3$  training iterations. When ANN-derived LW anisotropic factors are compared with those for the CERES/TRMM ADMs, the mean relative difference is close to zero and the STD of the difference is 1.3%.

Figure 4b compares all-sky ocean mean LW TOA fluxes stratified by VZA for the ANN approach, the new CERES/TRMM ADMs, and the ERBE ADMs. The mean ANN LW flux is within  $2 W m^{-2}$  of the mean CERES SSF LW flux at all VZAs, whereas the ERBE-like LW TOA flux deviates from the CERES SSF LW flux by as much as  $4 W m^{-2}$  at small VZAs.

Frequency distributions of the instantaneous and mean regional LW TOA flux differences,  $F_{ANN} - F_{SSF}$  (solid line) and  $F_{ERBE} - F_{SSF}$  (dashed line), are shown

in Figs. 5c,d. Although the distributions have comparable STDs, the ANN-derived LW mean fluxes show a significant improvement in accuracy over the ERBE-like. Figures 7a,b show the differences in the all-sky ocean TOA LW regional mean fluxes in the Tropics with a  $1^\circ$  latitude  $\times$   $1^\circ$  longitude grid. The overall mean of the difference is  $0.18 W m^{-2}$  for the ANN-derived and SSF fluxes, and  $3.24 W m^{-2}$  for the ERBE-like and SSF fluxes. As in the SW case, the regional cloud patterns can be observed for ANN-derived and ERBE-like LW TOA fluxes.

## 5. Summary and conclusions

We simulated and trained a feed-forward error back-propagation artificial neural network using CERES broadband measurements and CERES/TRMM ADMs, and applied it to radiative flux retrievals using the CERES measurements alone. The trained ANN reproduced the original CERES SSF short- and longwave fluxes with very small mean bias for different scene types. This indicates that constructed training sets represent



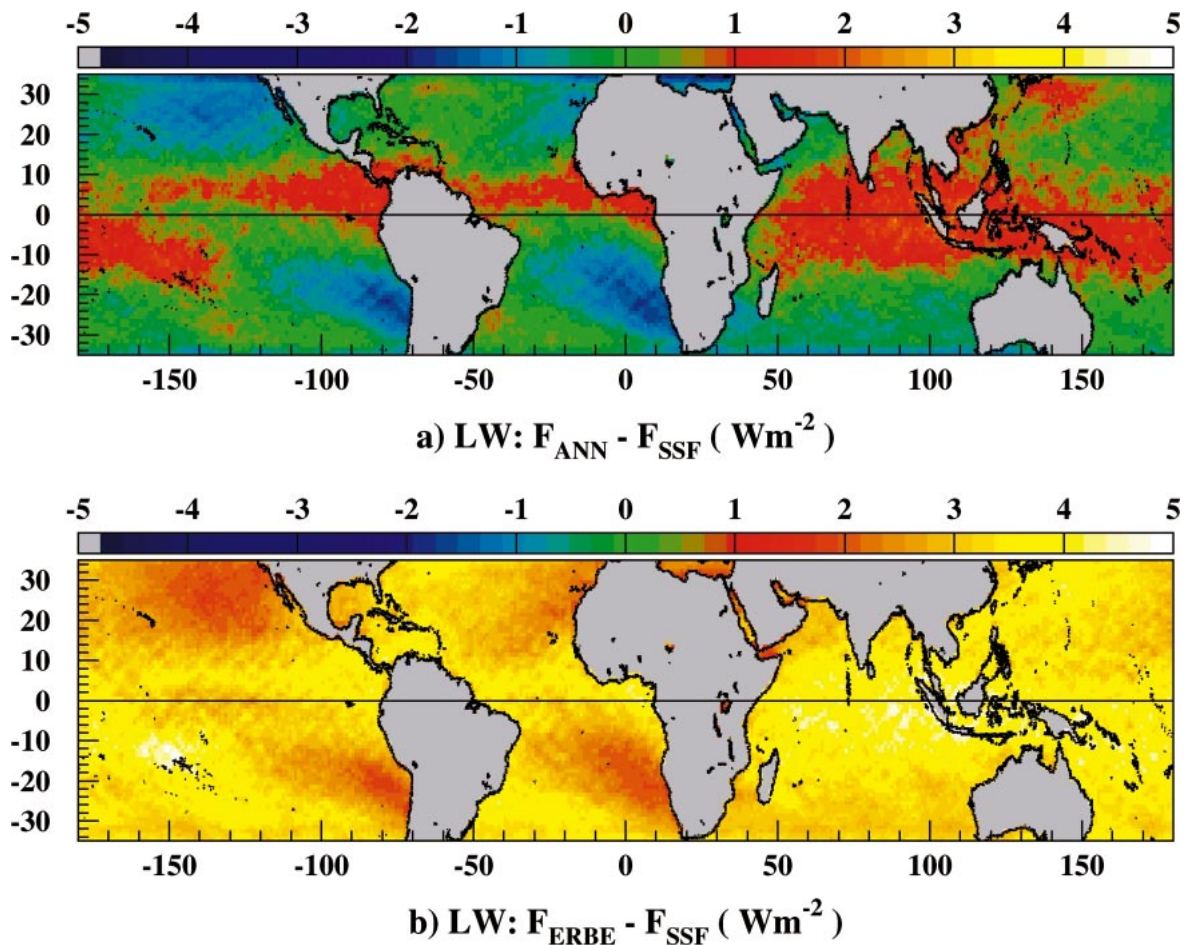


FIG. 7. All-sky ocean mean regional TOA LW flux difference for (a) ANN-derived and SSF fluxes,  $F_{\text{ANN}} - F_{\text{SSF}}$ , and (b) CERES ERBE-like and SSF fluxes,  $F_{\text{ERBE}} - F_{\text{SSF}}$ . Regional resolution is  $1^\circ$  in lat and lon.

the data well and allow a good generalization of the ANN. The ANN-based retrievals provide more accurate SW and LW CERES TOA flux estimates from only broadband measurements than MLE-based ERBE fluxes for the range of VZAs considered. This result can be explained by the ability of neural networks to reproduce the high degree of nonlinearity of CERES/TRMM ADMs and an indirect use of additional (in comparison with ERBE) imager information during the neural network training phase. The advantages of ANN simulation algorithms can be successfully incorporated in the next generation of the ADMs for the *Terra* and *Aqua* spacecraft, when CERES measurement is not sufficiently provided with imager information. The observed regional cloud patterns in the ANN-derived flux retrievals suggest a possibility of further improvement by using additional variables for ANN training, such as longitude and latitude, or data provided by the ECMWF.

The ANN simulations proved to be a very useful and flexible data analysis tool. These techniques generally can be of great interest in the remote sensing field: the ANN, pretrained on the available high-precision data-

sets, can be applied to satellite data without coincident imager information or complex scene identification. Similarly, old existing data can be reanalyzed with better accuracy.

*Acknowledgments.* The authors would like to thank Sandra K. Nolan of Science Applications International Corporation for preparing the CERES/TRMM SSF dataset. This research was funded by the Clouds and the Earth's Radiant Energy System (CERES) project under NASA Grant NAG-1-2318.

## APPENDIX

### Dimensions

The dimension of vectors  $\mathbf{x}^{1,2,3}$ ,  $\mathbf{a}^{1,2,3}$ ,  $\mathbf{b}^{1,2,3}$ , and  $\mathbf{S}^{1,2,3}$  is equal to the number of neurons in the corresponding neuron layer (vectors of output neuron layer  $L_3$  have scalar dimensionality). The dimension of the input vector  $\mathbf{a}^0$  is equal to the numbers of variables chosen for the approximation, 5 and 4 for the SW and LW cases,



respectively. Weight matrices  $\mathbf{W}^n$  have a number of rows equal to the number of neurons in the corresponding neuron layer and a number of columns equal to the number of inputs to the same neuron layer. The gradient matrices  $\mathbf{F}^n$  are square diagonal matrices with dimensions equal to the number of neurons in the corresponding layer.

## REFERENCES

- Barkstrom, B. R., 1984: The Earth Radiation Budget Experiment. *Bull. Amer. Meteor. Soc.*, **65**, 1170–1186.
- Bose, N. K., and P. Liang, 1996: *Neural Network Fundamentals with Graphs, Algorithms, and Applications*. McGraw-Hill Electrical and Computer Engineering Series, McGraw-Hill, 478 pp.
- Cybenko, G., 1989: Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.*, **2**, 304–314.
- Geier, E. B., R. N. Green, D. P. Kratz, P. Minnis, W. F. Miller, S. K. Nolan, and C. B. Franklin, 2001: Single satellite footprint TOA/surface fluxes and clouds (SSF) collection guide document. CERES Release 2, 243 pp. [Available online at [http://asd-www.larc.nasa.gov/ceres/collect\\_guide/](http://asd-www.larc.nasa.gov/ceres/collect_guide/).]
- Hagan, M. T., H. B. Demuth, and M. H. Beale, 1996: *Neural Network Design*. Brooks/Cole, 736 pp.
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, and J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) sensor package. *J. Atmos. Oceanic Technol.*, **15**, 809–817.
- Loeb, N. G., N. Manalo-Smith, S. Kato, W. F. Miller, S. Gupta, P. Minnis, and B. Wielicki, 2003: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Tropical Rainfall Measuring Mission satellite. Part I: Methodology. *J. Appl. Meteor.*, **42**, 240–265.
- , K. Loukachine, and N. Manalo-Smith, 2003: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Tropical Rainfall Measuring Mission satellite. Part II: Validation. *J. Appl. Meteor.*, **42**, 1748–1769.
- Priestley, K. J., R. B. Lee III, R. N. Green, S. Thomas, and R. S. Wilson, 1999: Radiometric performance of the Clouds and the Earth's Radiant Energy System (CERES) proto-flight model on the Tropical Rainfall Measuring mission (TRMM) spacecraft for 1998. Preprints, *10th Conf. on Atmospheric Radiation*, Madison, WI, Amer. Meteor. Soc., 33–36.
- Rabier, F., J.-N. Thepaut, and P. Courtier, 1998: Extended assimilation and forecast experiments with a four-dimensional variational assimilation. *Quart. J. Roy. Meteor. Soc.*, **124**, 1861–1887.
- Smith, G. L., R. N. Green, E. Raschke, L. M. Avis, J. T. Suttles, B. A. Wielicki, and R. Davis, 1986: Inversion methods for satellite studies of the earth's radiation budget: Development of algorithms for the ERBE mission. *Rev. Geophys.*, **24**, 407–421.
- Wielicki, B. A., and R. N. Green, 1989: Cloud identification for ERBE radiation flux retrieval. *J. Appl. Meteor.*, **28**, 1133–1146.
- , R. D. Cess, M. D. King, D. A. Randall, and E. F. Harrison, 1995: Mission to planet Earth: Role of clouds and radiation in climate. *Bull. Amer. Meteor. Soc.*, **76**, 2125–2153.