

Estimation of solar radiation using a neural network based on radiative transfer

Hideaki Takenaka,¹ Takashi Y. Nakajima,² Akiko Higurashi,³ Atsushi Higuchi,¹ Tamio Takamura,¹ Rachiel T. Pinker,⁴ and Teruyuki Nakajima⁵

Received 12 October 2009; revised 10 December 2010; accepted 3 January 2011; published 30 April 2011.

[1] This study developed an algorithm for estimating solar radiation from space using a neural network (NN) with an improved learning algorithm to approximate radiative transfer code. The NN solver for the solar radiation budget is based on radiative transfer calculations. All data sets for testing and training the NN were generated from radiative transfer code. Thus the NN traces the radiative transfer calculation that is approximated by a learning algorithm. To demonstrate the effectiveness of the NN approach for high-speed estimation and multiparameter problems, the NN was applied to data from a geostationary satellite and a Sun-synchronous subrecurrent orbit satellite. The developed algorithm was applied to data from the Multi-functional Transport Satellite-1 Replacement (MTSAT-1R) geostationary satellite, and estimations were validated against in situ observations for March 2006 at four SKYNET sites. Byproducts of the algorithm include UVA, UVB, and photosynthetically active radiation (PAR) fluxes as well as direct and diffuse components. The NN approach enables semi-real-time analysis of these products by high-speed calculation. In addition, the NN allows for consideration of detailed particle optical parameters in the solar radiation budget without the need for a massive database. The method was also applied to observations from the Advanced Earth Observing Satellite-II/Global Imager (ADEOS-II/GLI) for May 2003. The results showed trends in the direct and diffuse components of downward solar radiation over the North Pacific Ocean. This report outlines the construction of the NN for radiation budget estimation and demonstrates the effectiveness of the NN approach.

Citation: Takenaka, H., T. Y. Nakajima, A. Higurashi, A. Higuchi, T. Takamura, R. T. Pinker, and T. Nakajima (2011), Estimation of solar radiation using a neural network based on radiative transfer, *J. Geophys. Res.*, 116, D08215, doi:10.1029/2009JD013337.

1. Introduction

[2] Clouds strongly influence the solar energy budget and the Earth's climate [Manabe and Wetherald, 1967; Schneider, 1972]. Clouds can cool the Earth by reflecting solar radiation but also maintain warmth by absorbing and emitting terrestrial radiation. Furthermore, clouds play important roles in the energy balance at the Earth's surface and at the top of the atmosphere (TOA) and have complex connections to the Earth system and feedback processes.

Szilder and Lozowski [1995] and Szilder *et al.* [1998] discussed cloud feedback processes using zero-dimensional time-dependent climate models including three climate feedbacks. They showed that change in cloud water amount led to drastic environmental change. However, their model included uncertainty in the optical properties of clouds. Wetherald and Manabe [1988] discussed cloud feedback process using a general circulation model (GCM). A modeling study by Tsushima and Manabe [2001] tested cloud feedback sensitivity to global mean surface temperature based on explicit definitions of cloud feedback processes. The GCM result showed strong cloud feedback sensitivity to global mean surface temperature, but the observation-based global mean surface temperature did not show a relationship to cloud feedback process. This result highlighted our low understanding of the influence of clouds on the Earth system, and the need for accurate observation-based estimations of the influence of clouds on the radiation budget. Moreover, because clouds and aerosols are closely related in the climate system [Twomey, 1977; Twomey *et al.*, 1984], the influences of other particles must also be examined. The complex internal/external mixing of certain par-

¹Center for Environmental Remote Sensing, Chiba University, Chiba, Japan.

²Research and Information Center, Tokai University, Tokyo, Japan.

³Atmospheric Environment Division, National Institute for Environmental Studies, Tsukuba, Japan.

⁴Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland, USA.

⁵Center for Earth Surface System Dynamics, University of Tokyo, Chiba, Japan.

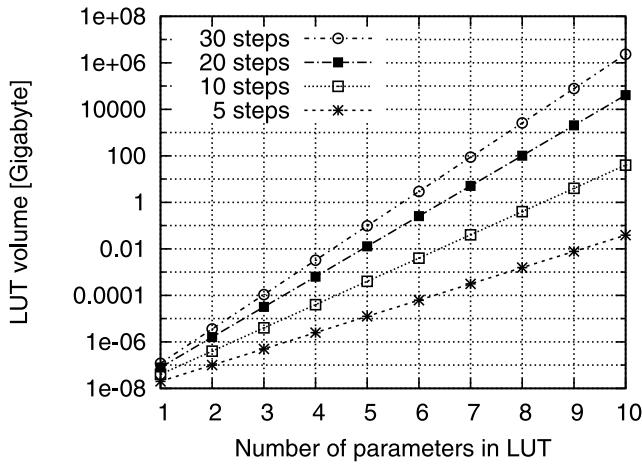


Figure 1. Relationship between the volume of the lookup table (LUT) and the number of parameters. The LUT was assumed to be floating points of 4 bytes, and each variable was divided in arbitrary steps. The four lines depict the difference in number of steps (5, 10, 20, and 30 steps).

ticles may alter the nominal optical properties that influence the radiation budget [Lesins *et al.*, 2002]. Accurate estimation of large-scale radiation budgets is crucial for improving model parameterization and better understanding the Earth system. Thus, a powerful method capable of handling numerous parameters efficiently is needed.

[3] Geostationary satellite observations are useful for estimating upward and downward solar radiation at the surface and at the TOA over wide regions and at high temporal resolution. For example, Pinker *et al.* [2002, 2003], as part of the Global Energy and Water Cycle Experiment (GEWEX) Continental-scale International Project (GCIP) for modeling the water cycle, estimated the surface radiation budget using Geostationary Operational Environment Satellite (GOES-8/Imager) data with auxiliary inputs from the National Centers for Environmental Prediction (NCEP) operational regional Numerical Weather Prediction (NWP) model. Zhang *et al.* [2004] estimated TOA and surface radiative fluxes using a radiative transfer model with data from the International Satellite Cloud Climatology Project (ISCCP); from the obtained radiation budget, they discussed the distribution of cross-zonal heating and cooling.

[4] In this study, we developed an accurate, high-speed, neural network (NN)-based algorithm to estimate solar radiation. Previous meteorological and solar radiation studies have used the NN concept for a variety of applications. For example, Lopez *et al.* [2001] estimated hourly global photosynthetically active radiation (PAR) using a NN based on ground-observed data for Almeria, Spain (90% of the observation data was used to test the NN, with the remaining 10% used for training). Dorylo *et al.* [2002] estimated solar radiation using a radial basis function network trained with long-term data from six observation sites (Masirah, Salalah, Seeb, Sur, Fahud, and Sohar) in Oman. Zarzalejo *et al.* [2005] estimated hourly global irradiance using a NN trained with data from 15 Spanish observation stations with

a satellite-derived cloud index. However, although these empirical estimation models used ground-observed solar radiation data for training, they were not based on radiative transfer calculations. Key and Schweiger [1998] proposed a method to speed up the radiative transfer code “STREAMER” by using a NN with Rumelhart’s back-propagation (BP) algorithm [Rumelhart *et al.*, 1986]. STREAMER includes two radiative transfer solvers employing discrete ordinate and two-stream methods. It also includes five aerosol optical models and cloud containing principal gases (H_2O , O_3 , CO_2). A BP algorithm is one of the most general learning algorithms of a NN. The NN is built by minimizing the error margin between the training data (radiative transfer output) and the network output. The trained NN behaves like a high-speed radiative transfer solver.

[5] Our aim was to achieve accurate estimations and advanced analysis capabilities by utilizing a NN with an improved learning algorithm. The main advantages of the NN approach are computation speed and the ability to consider multiple parameters, as large data sets are not required. Highly accurate radiative transfer codes can calculate solar radiation, including the effect of gasses and particles. However, researchers have had to deal with the trade-off between calculation speed and accuracy. In general, satellite-based estimation methods have used lookup tables (LUTs). Because precalculated values are used, the LUT method is effective when processing large amounts of data. However, if the effects of absorbing gasses and the optical properties of particles are incorporated precisely, the LUT volume becomes too large and slows computations. Figure 1 illustrates the relationship between parameters and LUT volume. Including many values for numerous parameters not only increases the LUT volume but also requires complex interpolation/extrapolation procedures [Sauer and Yuan, 1995; Gasca and Sauer, 2000]. Calculation speed falls markedly when a complex interpolation/extrapolation procedure such as Lagrangian interpolation is needed. The NN approach is one solution to these problems, allowing for high-speed calculation and multiparameter problems. The simple structure of a NN and that fact that it needs neither a large database nor complex interpolation/extrapolation techniques make NNs appropriate for approximating radiative transfer code. High-speed calculation allows for more rapid analysis, and the ability to handle numerous parameters will allow more options for advanced analysis of atmospheric radiation.

[6] The rest of this paper is arranged as follows. Section 2 presents the construction of the NN, the improved learning algorithm, and application of the NN. Sections 3 and 4 describe the data used and the estimation of cloud optical properties. Section 5 presents results from tests of the developed algorithm. The algorithm was applied to MTSAT-1R data and validated against SKYNET ground observations. The results indicate that real-time analysis of UVA, UVB, and PAR flux byproducts is possible. In addition, the NN approach allows for advanced analysis of detailed optical parameters, exceeding limits of the LUT method. The NN approach was also applied to the updated standard product of the Advanced Earth Observing Satellite-II/Global Imager (ADEOS-II/GLI) for May 2003.

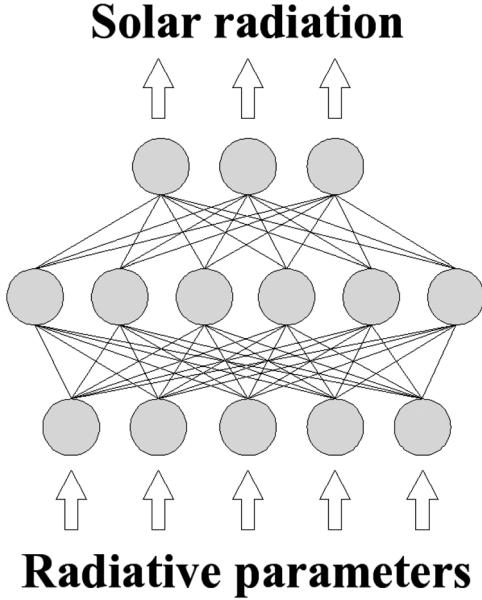


Figure 2. Schematic illustration of the solver concept. Radiative parameters are input to the NN, which then calculates solar radiation. Various optical parameters (e.g., particle optical properties and gas absorptions) are included for solar flux, and the NN behaves like the radiative transfer calculation.

The results demonstrate the effectiveness of the NN-based algorithm.

2. Construction of the Neural Network Based on Radiative Transfer

[7] A NN is a numerical model of the brain neuron network with some characteristics resembling those of a biological brain. Its most important feature is the ability to be “trained.” In this study, the solver for solar radiation of the NN was based on radiative transfer calculations. Figure 2 illustrates the solver concept. The NN traces the radiative transfer calculation by a function approximation that is achieved by the learning algorithm. This provides a comprehensive solver that utilizes the NN characteristics for estimation of solar radiation. In this section, we describe the construction of the neurolink network solver (NN solver).

2.1. Selection of Radiative Fluxes and Parameters for the Neural Network Method

[8] First, to optimize the NN design, we must decide what elements require NN support for estimation of solar radiation. Blind learning disturbs and confuses the NN. This section describes the estimation of solar radiation. Downward and upward shortwave (SW) fluxes at the surface and the TOA are derived, with incoming solar radiation at the TOA expressed as follows:

$$R_{TOA}^{\downarrow} = \frac{\mu_0 S_0}{r_{dst}^2}. \quad (1)$$

Downward solar radiation at the TOA is determined by the Sun-Earth distance and the solar zenith angle, where μ_0 is the $\cos(SolarZenithAngle)$, r_{dst} is a Sun-Earth distance [AU], and S_0 is integrated solar radiation at wavelength $S_{0\lambda}$ ($\text{W/m}^2/\mu\text{m}$) ($\int S_{0\lambda} d\lambda = S_0$), using the solar constant (1367 [W/m^2]). Downward solar radiation at the surface is a more complicated function due to absorption and diffusion by gases and particles:

$$R_{SFC}^{\downarrow} = \frac{1}{r_{dst}^2} \int F_{sfc}^{\downarrow} d\lambda, \quad (2)$$

where $\int F_{sfc}^{\downarrow} d\lambda$ is downward surface radiation, given from accurate atmospheric radiative transfer calculations with integration of wavelength [Nakajima and Tanaka, 1986, 1988]. Solar flux at each wavelength, F_{sfc}^{\downarrow} , is given as follows:

$$F_{sfc}^{\downarrow} = \int_0^1 \int_0^{2\pi} I_{\lambda}(\tau_{atm}; \mu, \phi) \mu d\mu d\phi, \quad (3)$$

where τ_{atm} is the total optical thickness of the atmosphere that includes Rayleigh and particle scattering and absorption by gasses ($\tau_{atm} = \tau_{ray} + \tau_{par} + \tau_{gas}$), I_{λ} is the spectral radiance, and ϕ is the azimuthal angle. F_{sfc}^{\downarrow} includes the influence of diffuse radiation by particles and absorbing gases.

Upward solar radiation at the surface is given by

$$R_{SFC}^{\uparrow} = A_G \times R_{SFC}^{\downarrow}, \quad (4)$$

where A_G is the surface albedo for global flux. Upward SW flux at the surface depends on the albedo. When surface albedo is given, it is provided briefly. Upward solar radiation at the TOA, which is strongly affected by the condition of the atmosphere, is given as follows:

$$R_{TOA}^{\uparrow} = \frac{1}{r_{dst}^2} \int F_{toa}^{\uparrow} d\lambda \quad (5)$$

$$F_{toa}^{\uparrow} = \int_{-1}^0 \int_0^{2\pi} I_{\lambda}(\tau_{atm}; \mu, \phi) \mu d\mu d\phi. \quad (6)$$

Clouds, in particular, change the reflectivity of the atmosphere. The function approximation by the NN is applicable to the R_{SFC}^{\downarrow} and R_{TOA}^{\uparrow} . Downward SW flux at the surface can be divided into diffuse and direct components (R_{dif}^{\downarrow} and R_{dir}^{\downarrow}) as follows:

$$R_{SFC}^{\downarrow} = R_{dif}^{\downarrow} + R_{dir}^{\downarrow}, \quad (7)$$

where

$$\begin{aligned} R_{dif}^{\downarrow} &= \frac{1}{r_{dst}^2} \int F_{dif}^{\downarrow} d\lambda \\ &= \frac{1}{r_{dst}^2} \int \left\{ 2\pi \int_0^1 I_{\lambda}(\tau_{atm}; \mu) \mu d\mu \right\} d\lambda \end{aligned} \quad (8)$$

$$\begin{aligned} R_{dir}^{\downarrow} &= \frac{1}{r_{dst}^2} \int F_{dir}^{\downarrow} d\lambda \\ &= \frac{1}{r_{dst}^2} \int \left\{ \mu_0 \pi S_{0\lambda} \exp\left(-\frac{\tau_{atm}}{\mu_0}\right) \right\} d\lambda. \end{aligned} \quad (9)$$

Table 1. Input Parameters of NN Solver for Geostationary Satellite MTSAT-1R/JAMI

Parameter	Symbol	Unit
Solar zenith angle	θ_0	deg
Water cloud optical thickness	τ_{cW}	none
Water cloud effective radius	r_{cW}	μm
Ice cloud optical thickness	τ_{cI}	none
Ice cloud effective radius	r_{cI}	μm
Cloud top pressure	P_c	hpa
Surface albedo	A_g	none
Surface pressure	P_0	hpa
Total column ozone	OZN	DU
Total column water vapor	PWV	g/m^2

Here, we applied three components, R_{dir}^\downarrow , R_{dif}^\downarrow , and R_{TOA}^\uparrow , for high-speed calculation of the NN and to avoid the need for a LUT. Because R_{TOA}^\uparrow and R_{SFC}^\downarrow can be obtained by simple calculation without a large data set, they do not need the support of the NN. The Sun-Earth distance correction $1/r_{dst}^2$ was given from outside of NN.

[9] To accurately approximate a function, the function's form and relation to the variables should be known. The solver for cloudy and clear sky conditions is explained in this section (and applied to data of the geostationary satellite MTSAT-1R in section 5). The parameters applied to the radiative transfer calculation and NN are shown in Table 1. The solar zenith angle is a vital parameter of solar irradiance. The water and ice cloud optical thickness and effective radius affect the cloud droplet scattering (here, droplets were assumed to be spheres). The Rayleigh scattering optical thickness is determined by the atmospheric pressure from the TOA to the surface. Cloud top height is converted to cloud top pressure by the vertical pressure profile. This determines the Rayleigh scattering and gaseous absorption from the TOA to the cloud top. Surface albedo includes the surface reflected flux and multiple reflection between the surface and atmosphere. Total column water vapor and ozone are included, as these are effective absorbing gases of SW radiation. Twenty-nine other kinds of gas are also included, based on U.S. Standard Atmosphere. The vertical distribution is generated using these parameters and input to the radiative transfer model for the training data set. The vertical profile can be input directly to the radiative transfer model and the NN. However, detailed changes in the vertical profiles were not the focus of this study. The training and test data sets were given by the radiative transfer code RSTAR5b [Nakajima and Tanaka, 1986, 1988]. Each parameter was normalized (0.0–1.0) and applied to the NN for learning. Generally, slow changes and smooth functions are desirable for approximation. Ozone and water vapor caused a decline in the SW flux by absorption, while particulate matter produced complicated absorption and scattering. Typical examples are shown in Figures 3 and 4. The direct component has a simple shape (Figures 3a and 4a), but the diffuse component peaks in the lower optical thickness range (Figures 3b and 4b). In addition, the range of optical thickness is narrow in the direct component. Upward solar radiation at the TOA is stable in the higher optical thickness range (Figures 3c and 4c). The effect of different albedos on multiple reflection is shown in Figure 5a. Moreover, the effect of albedo appears in a lower

optical thickness range in the upward flux at the TOA (Figure 5b). An important variability occurs in the low optical thickness range. Therefore, normalized parameters are input to the NN through the primary transfer function

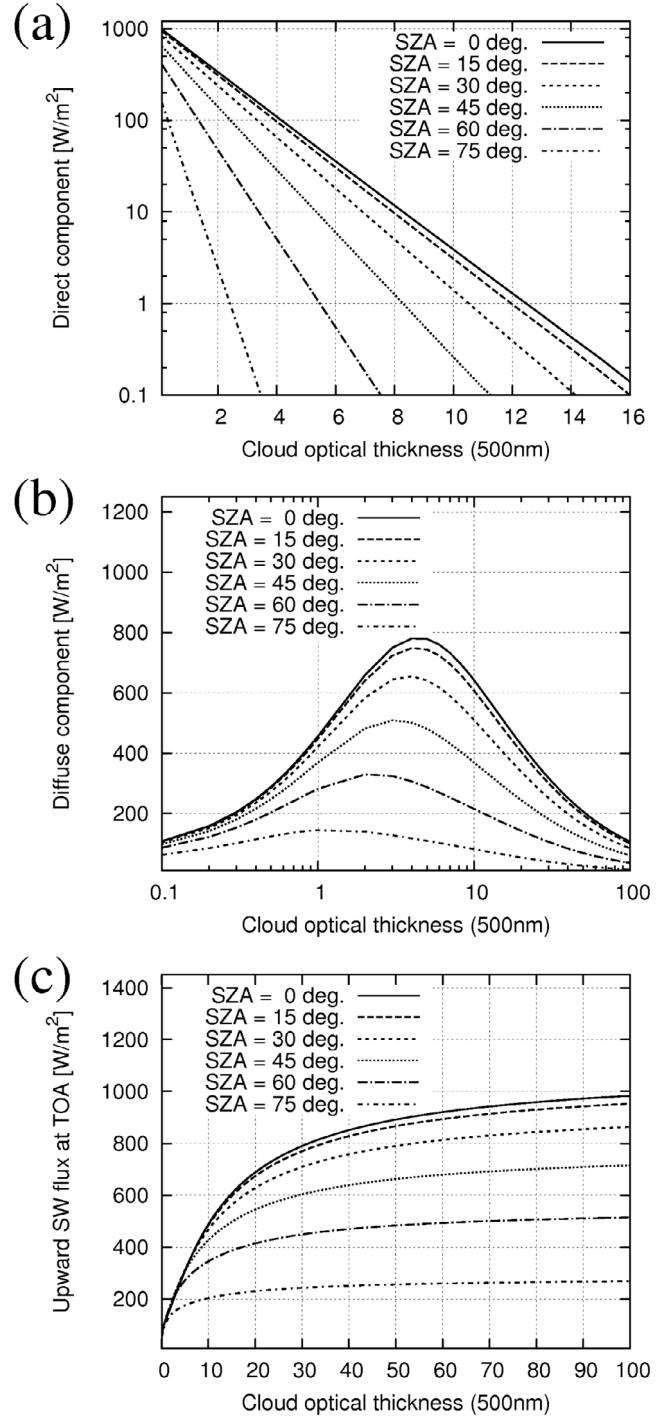


Figure 3. Influence of clouds on SW fluxes, based on classification of the solar zenith angle (SZA). (a) Direct component of downward SW flux at the surface. (b) Diffuse component of downward SW flux at the surface. (c) Upward SW flux at the TOA. (U.S. Standard Atmosphere, water particle effective radius = $10.0 \mu\text{m}$, cloud top pressure = 540 hpa, and surface albedo = 0.03.)

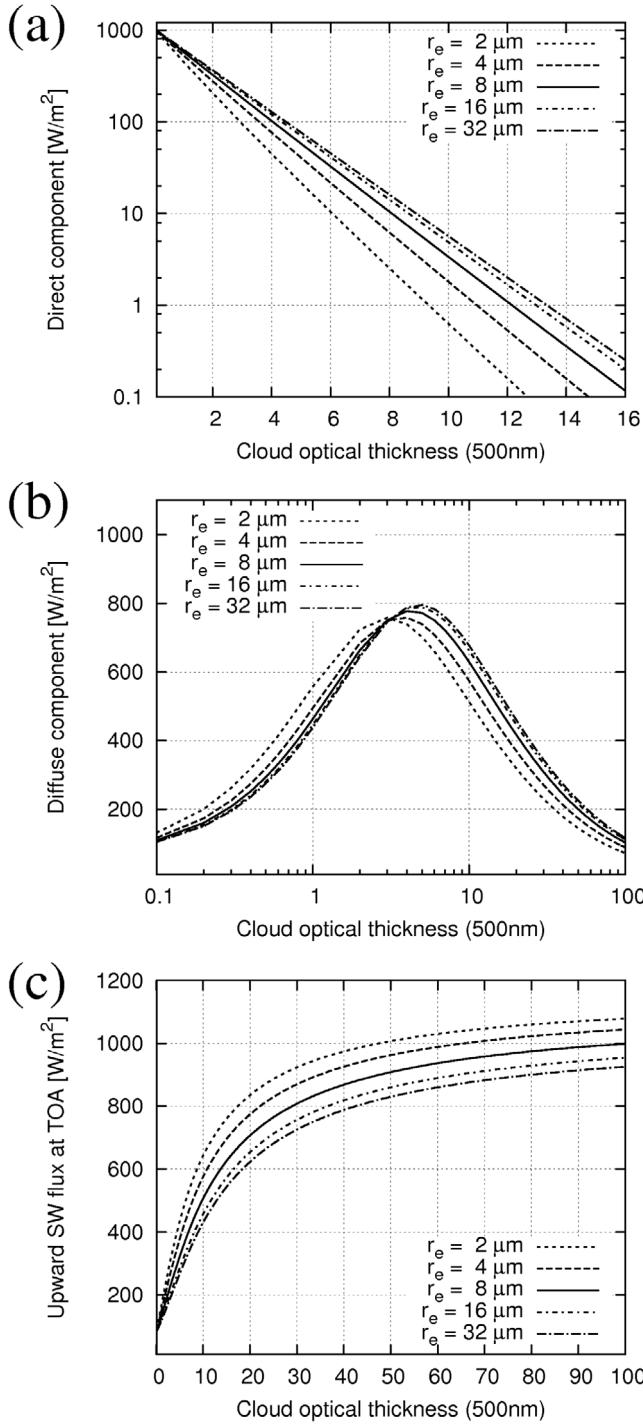


Figure 4. Same as Figure 3 but for classification of cloud droplet effective radius r_e . (Solar zenith angle = 0.0. U.S. Standard Atmosphere, water particle effective radius = $10.0 \mu\text{m}$, cloud top pressure = 540 hpa, and surface albedo = 0.03.)

$X = f_p(x)$, which is the square root. Because $f_p(x)$ is weighted to the low optical thickness range, it improved the accuracy.

2.2. Neural Network Architecture

[10] The three-layer NN has a simple and compact structure in which four equations give the functions of all

neurons. The neuron layers are called the “input layer,” “hidden layer,” and “output layer” (Figure 6). Details on the functions of individual neurons and the three-layer NN are presented in Appendix A1.

$$T_j = \sum_{i=1}^n X_i W_{ji}, \quad (10)$$

$$H_j = f(T_j), \quad (11)$$

$$U_k = \sum_{j=1}^m H_j W_{kj}, \quad (12)$$

$$O_k = f(U_k). \quad (13)$$

Each layer has the neurons of n , m , and l pieces of the input, hidden, and output layers. Here X stands for the input parameter, W_{kj} and W_{ji} are the output layer and hidden layer neuron synaptic weights, respectively, T and H are the net value and output of the hidden layer neuron, respectively,

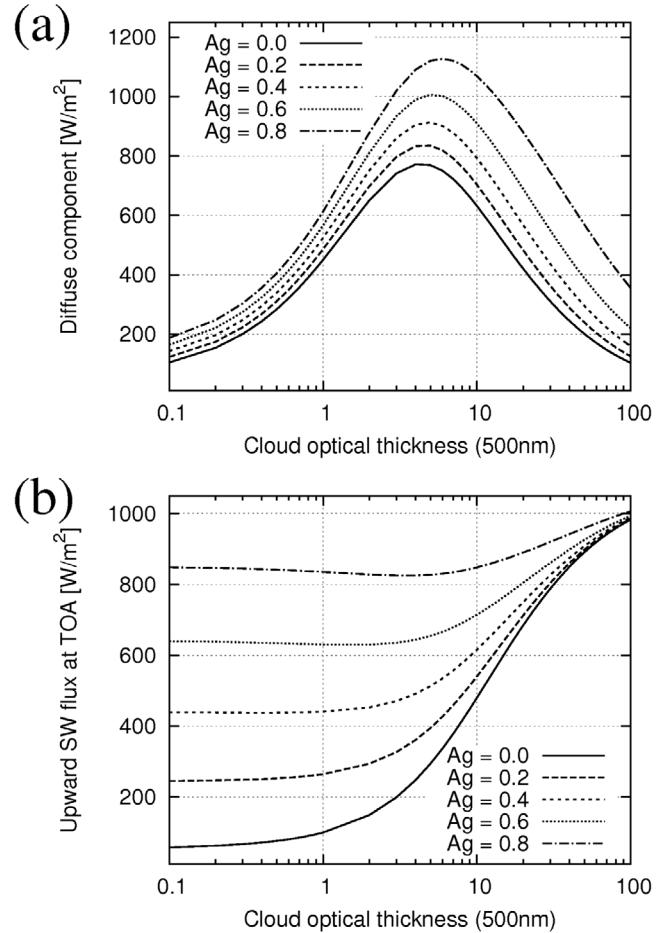


Figure 5. Influence of clouds on SW fluxes, based on the classification of surface albedo Ag . (a) Diffuse component of downward SW flux at the surface. (b) Upward SW flux at the TOA (solar zenith angle = 0.0. U.S. Standard Atmosphere, water particle effective radius = $10.0 \mu\text{m}$, and cloud top pressure = 540 hpa).

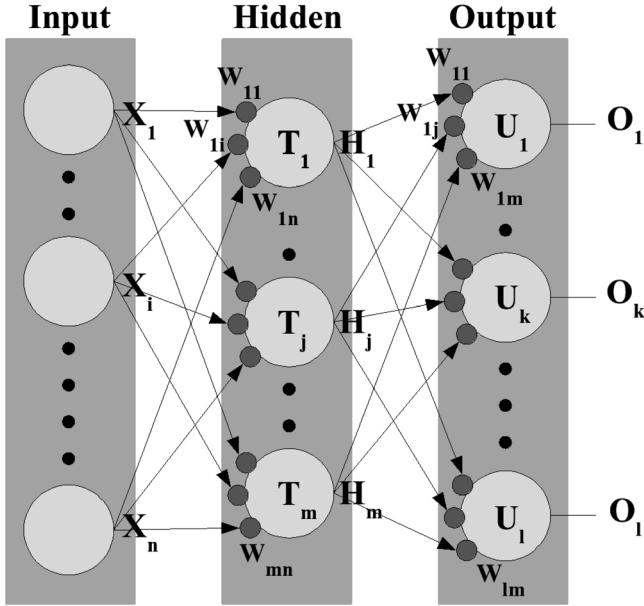


Figure 6. Schematic illustration of the three-layer network structure. Each layer neuron is connected by a synaptic weight. A complex system is composed of sets of simple neurons.

and U and O are the net value in the output layer neuron and the output of the network, respectively. The $f(T)$ and $f(U)$ are kernel functions as a transfer function of a neuron. The input of radiative transfer code is related to X of the input layer and O of the output layer indicates radiation. Each radiative parameters are input using the primary transfer function $X = \sqrt{x}$. The hidden layer is concealed from radiative transfer; however, the activity of each neuron in each layer achieves the function of radiative transfer by superposition of each neuron output. In each layer, because each neuron is independent, we may calculate it from any neuron, and the calculations can be parallel. Parallel calculation is one of the advantages of the NN approach in that it enables high-speed calculation. Moreover, because four equations can be achieved by a simple computational program, the NN is compact. Computer resources (e.g., data storage and memory space) are not occupied by the NN, and thus some NNs can also calculate in parallel.

[11] For an untrained NN, a solver function must be added, which should be given by a learning process that uses training data. In the BP learning algorithm [Rumelhart *et al.*, 1986], the neuron synaptic weights of each layer are adjusted so as to minimize the error margin between the network output O_k and the training data (true value) t_k . Details of the BP learning algorithm are presented in Appendix A2. In this section, we discuss the major concepts behind the learning algorithm and the results. The Rumelhart BP method optimizes the neuron synaptic weights based on the gradient of the error margin (steepest descent method). The synaptic weight is corrected as follows:

$$\Delta W = -\eta \nabla E = -\eta \frac{\partial E}{\partial W} \quad (14)$$

$$W^{(s+1)} = W^{(s)} + \Delta W^{(s)}, \quad (15)$$

where E is a nominal error margin, ∇E is a gradient of the error in the NN, s is the nominal iteration number, and η is the learning rate. Here, the error E is given as follows:

$$E = \sum_{q=1}^N \sum_{k=1}^l \frac{1}{2} (t_{k|q} - O_{k|q})^2, \quad (16)$$

where N is the number of data pairs in the training data set, which optimizes W to a low error margin from the gradient that is applied to each neuron synaptic weight in iterations based on the training data. The high-convergence method uses the Laplacian $\nabla^2 E$ [Hagan and Menhaj, 1994; Hagan *et al.*, 1996]. This approach, referred to as the Levenberg-Marquardt (LM) BP method [Levenberg, 1944; Marquardt, 1963], works as follows:

$$\begin{aligned} \Delta W &= -[\nabla^2 E]^{-1} \nabla E \\ &= -[H]^{-1} g \\ &\approx -[J^T J + \eta I]^{-1} J^T e. \end{aligned} \quad (17)$$

Equation (14) is replaced by equation (17), where J and J^T represent the Jacobian and transposed Jacobian, respectively. $[H]^{-1}$ and $J^T e (= g)$ are the inverse Hessian and gradient, respectively, and I is the identity matrix. Laplacian of the error $\nabla^2 E$ is represented by a data constellation in this equation. Thus, the method is basically a batch optimization. The error e and Jacobian J are given as follows:

$$e = \begin{bmatrix} e_{1|1} \\ e_{2|1} \\ \vdots \\ e_{1|2} \\ e_{2|2} \\ \vdots \\ e_{k|q} \\ \vdots \\ e_{l|N} \end{bmatrix} = \begin{bmatrix} t_{1|1} - O_{1|1} \\ t_{2|1} - O_{2|1} \\ \vdots \\ t_{1|2} - O_{1|2} \\ t_{2|2} - O_{2|2} \\ \vdots \\ t_{k|q} - O_{k|q} \\ \vdots \\ t_{l|N} - O_{l|N} \end{bmatrix} \quad (18)$$

$$J = \begin{bmatrix} \frac{\partial e_{1|1}}{\partial W_1} & \frac{\partial e_{1|1}}{\partial W_2} & \frac{\partial e_{1|1}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{1|1}}{\partial W_n} \\ \frac{\partial e_{2|1}}{\partial W_1} & \frac{\partial e_{2|1}}{\partial W_2} & \frac{\partial e_{2|1}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{2|1}}{\partial W_n} \\ \vdots & \vdots & \vdots & & & \vdots \\ \frac{\partial e_{1|2}}{\partial W_1} & \frac{\partial e_{1|2}}{\partial W_2} & \frac{\partial e_{1|2}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{1|2}}{\partial W_n} \\ \frac{\partial e_{2|2}}{\partial W_1} & \frac{\partial e_{2|2}}{\partial W_2} & \frac{\partial e_{2|2}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{2|2}}{\partial W_n} \\ \vdots & \vdots & \vdots & & & \vdots \\ \frac{\partial e_{k|q}}{\partial W_1} & \frac{\partial e_{k|q}}{\partial W_2} & \frac{\partial e_{k|q}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{k|q}}{\partial W_n} \\ \vdots & \vdots & \vdots & & & \vdots \\ \frac{\partial e_{l|N}}{\partial W_1} & \frac{\partial e_{l|N}}{\partial W_2} & \frac{\partial e_{l|N}}{\partial W_3} & \cdots & \cdots & \frac{\partial e_{l|N}}{\partial W_n} \end{bmatrix}. \quad (19)$$

The array of W [$W_1, W_2, W_3, \dots, W_n$] shows all the neuron synaptic weights in the developed network. The term ηI indicates the Levenberg-Marquardt modification. If $J^T J$ is effective, it has quadratic convergence. When it is unstable, equation (17) is approximated to equation (14) by ηI .

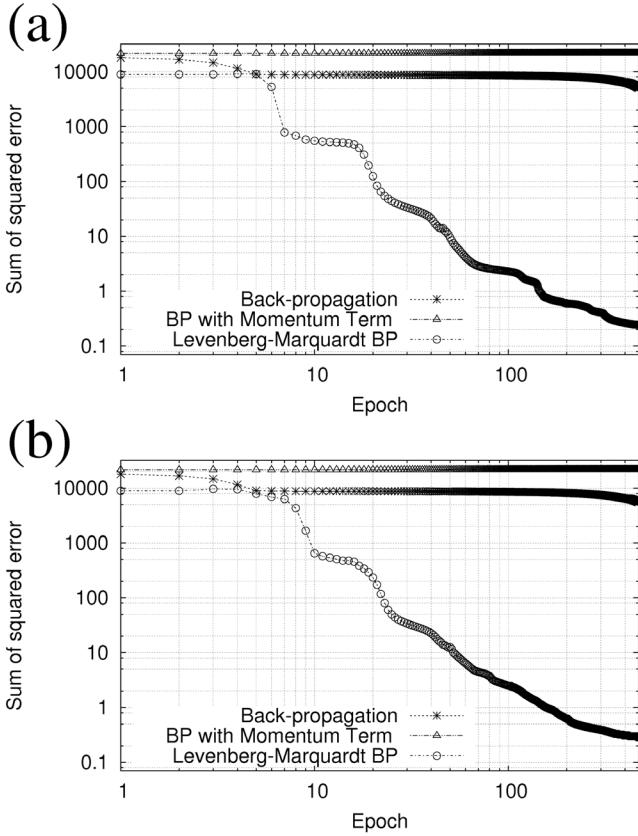


Figure 7. Sum of squared error in batch optimization for output of a test calculation of the neural network between the network output (O_k) and the true value (t_k). (a) Typical trends of batch optimization. (b) Average of ten patterns (based on different initial weight). The same training data set was used in Figures 7a and 7b. Sum of squared error = samples $\sum_{test=1}^n \frac{1}{2} (t_k - O_k)^2$.

Generally, the LMBP method proposes the best convergence. Typical results of test calculations are shown in Figure 7. The training data sets were given by the radiative transfer code RSTAR5b [Nakajima and Tanaka, 1986, 1988] based on section 2.1. Each parameter in the training data set was generated by a random number. The BP method has smooth and very slow convergence. Addition of the momentum term is proposed to improve convergence; this addition slightly accelerates the BP method, but the acceleration appears negligible in Figure 7. Compared to the BP method, the LMBP method had greater convergence. Because the LMBP is unstable in this study, it showed the same performance as the BP method in the first few steps due to the influence of ηl . Overall, the LMBP method appears to have a high convergence rate and low error.

[12] An online BP method is also presented. The batch method calculates ΔW using the entire training data, while the online method calculates ΔW using each training data point. The BP method involves the first derivative ∇ and is applicable to each set of training data. Learning at each point gradient might bring about convergence of any straying

data, but this is a nonfatal problem that often leads to a better result. In this study, an online BP method is proposed with some improvements and applied to the function approximation. The neuron survival rule is applied, and an antilocal minimum is included as one of the improvements. If each weight takes a value extremely close to 0 for a long time in the iteration of the learning process, it is replaced with 0 as the online value. As the weights of a specific neuron become 0 or near 0, that neuron gradually dies off because it does not contribute to the network function (Figure 8a). Other necessary neurons survive. Likewise, a neuron in the hidden layer can die suddenly when the weight of the output layer neuron becomes 0 (Figure 8b). By this process, negligible neurons are removed from the network. This study did not use a weight-decay technique, but it did apply the following expressions to replace equation (14):

$$\Delta W = -\eta \nabla E = -\eta \frac{\partial E}{\partial W} \quad (20)$$

$$\Delta V = -\zeta \nabla E = -\zeta \frac{\partial E}{\partial V}. \quad (21)$$

Here, V should be an independent variable that is a neuron parameter. If W and V are independent, the probability of satisfying $\Delta W = 0$ and $\Delta V = 0$ at the same time becomes

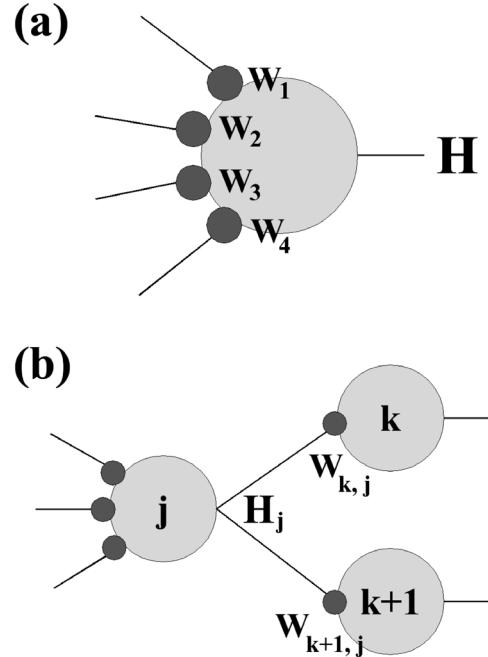


Figure 8. Schematic illustration of neuron survival. (a) When all weights of W_1 to W_4 become 0 by learning iterations, the activity of neuron output H is lost. Therefore, the neuron faces gradual death in the learning process. (b) When weights $W_{k,j}$ and $W_{k+1,j}$ (at the output neuron k and $k + 1$) become 0 by learning iteration, output of the j neuron is nullified. Nullification does not depend on output H_j of the j neuron; the neuron is nullified regardless of the activity of the j neuron. Therefore, the neuron dies suddenly.

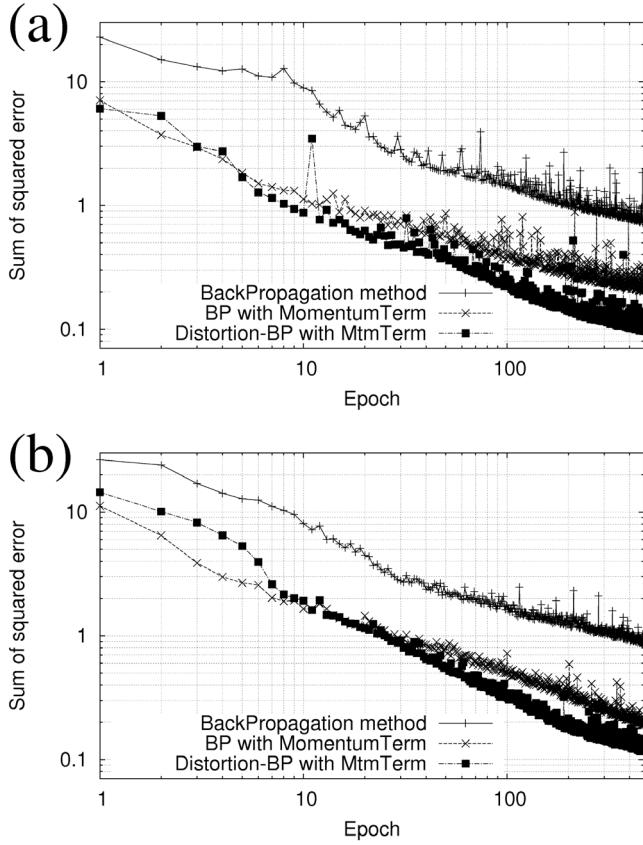


Figure 9. Same as Figure 7 but for the online method. The training data set was the same as that used in Figure 7.

lower in iterations of the learning process. When $\Delta W = 0$ and $\Delta V \neq 0$, the network is modified by ΔV . Because ΔV distorts the condition of network function, it promotes the escape of ΔW from the local minimum. While ΔV and ΔW behave like noise to each other, they are different from noise because they converge at the learning process. In this study, variable V is defined as the gain of the sigmoidal function, which is corrected as follows:

$$f(x) = \frac{1}{1 + e^{-\frac{x}{a}}} \quad (22)$$

where,

$$V = \frac{1}{a}. \quad (23)$$

Therefore, ΔV can be obtained from the learning process. An initial number of neurons is generated by random numbers. The number of neurons in the network is decided by the gradual and sudden deaths of neurons in the learning process. Therefore, the learning process will automatically decide parameters other than the layer structure, and the learning rate η , ζ will decrease linearly with iterations. This “distortion-BP” method was effective in this study, but its stability depends on the initial values of the network. The process stabilizes by applying empirical weights to ΔW and

ΔV and will not stabilize without the weights. The improved feed-forward operation is as follows:

$$T_j = \sum_{i=0}^n X_i W_{ji} \quad (24)$$

$$H_j = f(T_j, V_j) \quad (25)$$

$$U_k = \sum_{j=0}^m H_j W_{kj} \quad (26)$$

$$O_k = f(U_k, V_k). \quad (27)$$

The above equations are also applicable for a multilayered network (more than three layers). Figure 9 shows typical results of test calculations. The BP with the momentum term was more effective than the original BP, and the distortion BP with the momentum term was more effective than both former methods. This indicates that both improvements for the local minimum were effective. The LMBP method showed the most effective convergence rate, but the error margin of the online optimization was smaller than that of batch optimization in this study. The online optimization seems to be appropriate for this case. Figure 10 presents each calculation time. Calculation of the inverse matrix $[J^T J + \mu I]^{-1}$ led to large differences in calculation time. Previous studies have proposed ways to correct the LMBP method to speed computation. *Dias et al. [2005]* introduced online training for a time-dependent system by LMBP optimization, and *Ranga Suri et al. [2002]* applied an improved LMBP parallel calculation method for Linux clusters. However, we cannot disregard consideration of the inverse matrix. For real problems, the learning algorithm has both strong and weak points. A versatile learning algorithm does not exist. It is important to select the learning algorithm that fits best to the problem for generalization of a NN. In this study, the distortion BP with the momentum term was

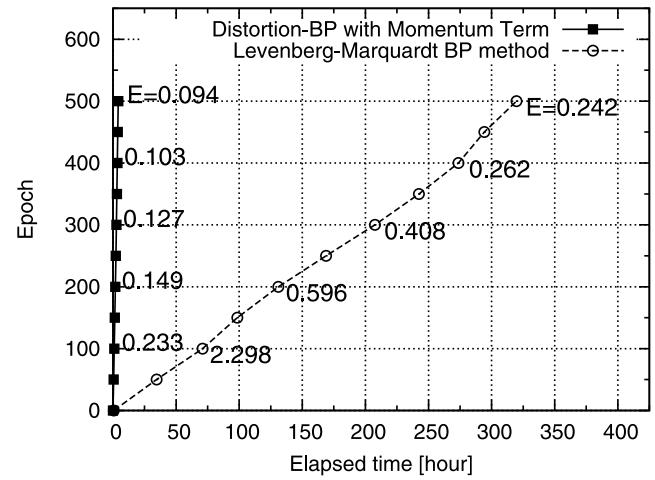
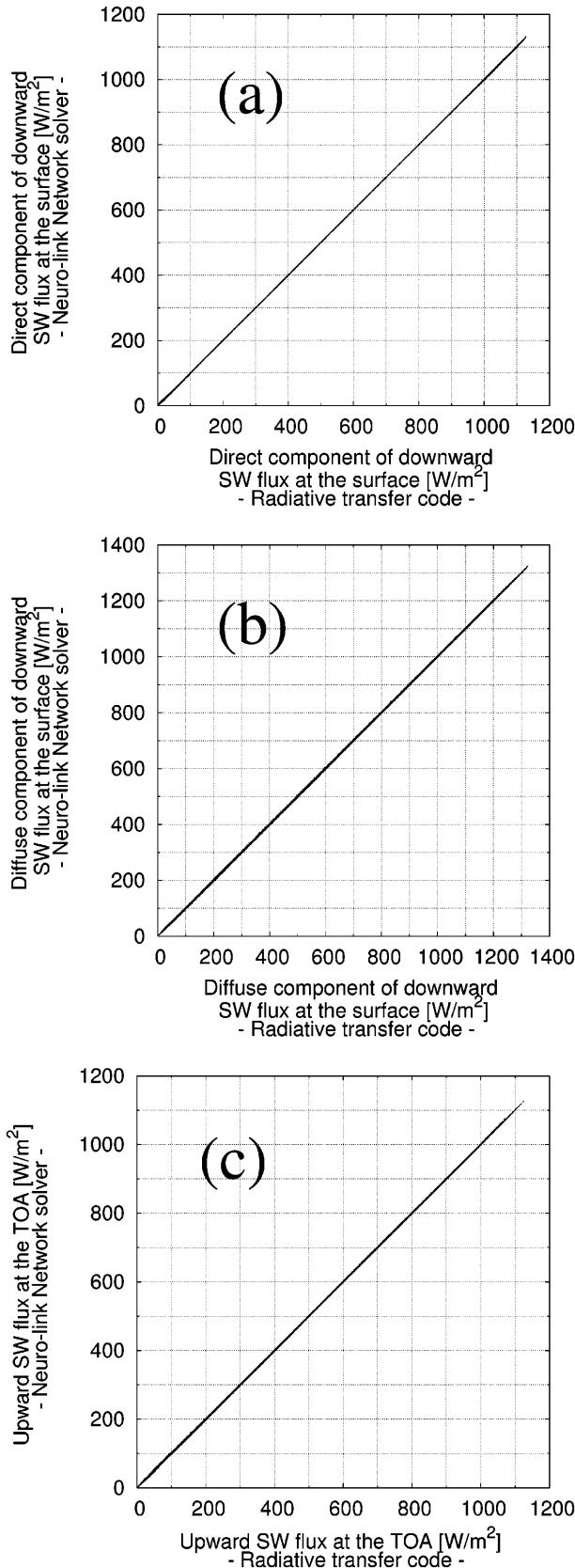


Figure 10. Relationship between the elapsed time and the attained epoch. The difference in calculation time is about 77 times.



the most effective approach and was employed to build the NN.

[13] To discuss the influence of clouds on solar radiation, Figure 11 presents a typical example of the learning results. The neurolink network solver (NN solver) was verified by test data for two million points (test data \neq training data). Figures 11a and 11b give the direct and diffuse components of SW flux at the surface, respectively, and Figure 11c shows upward SW flux at the TOA. Neither a fatal bias nor a critical failure was confirmed, and the accuracy was sufficient for estimation of solar radiation from satellite measurements.

[14] As noted above, an advantage of a NN is that it can include numerous parameters. However, blindly increasing the number of parameters is not essential. The learning process might be disturbed by irrelevant parameters, which will harm the accuracy of the NN. For accurate estimation, a new parameter must be chosen logically based on reconciliation between in situ observations and radiative transfer theory.

3. Data

[15] This section describes the satellite data used. Multi-functional Transport Satellite-1 Replacement (MTSAT-1R) data were used to estimate solar radiation in this study and to demonstrate the effectiveness of the NN for high-temporal resolution and semi-real-time analysis. In addition, the Advanced Earth Observing Satellite-II/Global Imager (ADEOS-II/GLI) particle product was used for test analysis. Estimation of solar fluxes including strict particle optical properties demonstrated the NN's advantageous ability to handle multiple parameters. Additionally, atmospheric profiles from the Japanese Re-Analysis 25-year (JRA-25) and Japan Meteorological Agency (JMA) Climate Data Assimilation System (JCDAS) products were used as supplementary data to estimate the cloud optical properties by the inversion process. Water vapor profiles from JRA-25 and JCDAS were also used to estimate the solar radiation. Furthermore, for increased ozone accuracy, the JRA-25 and JCDAS ozone data were replaced with Total Ozone Mapping Spectrometer (TOMS) and Ozone Monitoring Instrument (OMI) data.

3.1. MTSAT-1R/JAMI

[16] The MTSAT-1R/Japanese Advanced Meteorological Imager (MTSAT-1R/JAMI) is a geostationary satellite launched in February 2005. It has four infrared (IR) channels and one visible (VIS) channel. The five channels are centered at 0.67, 11.0, 12.0, 6.7, and 3.7 μm . The 3.7 μm near-infrared channel is a new addition to the IR bands on MTSAT-1R. The other channels are similar to those of the Geostationary Meteorology Satellite-5/Visible and Infrared Spin-Scan

Figure 11. Scatter diagram for comparison of calculation results between radiative transfer code and the NN solver (test data \neq training data). (a) Direct component of downward SW flux at the surface. (b) Diffuse component of downward SW flux at the surface. (c) Upward SW flux at the TOA. There is no regression line in the diagrams. The convergence of the two million points simply looks like a regression line.

Radiometer (GMS-5/VISSR). The water-absorbing channel ($3.7\text{ }\mu\text{m}$) allows for retrieval of the cloud droplet effective radius.

[17] In this study, the IR1 and IR2 channels were used for cloud classification [Inoue, 1987], and the VIS channel was used to retrieve cloud optical properties. The split-window method is suitable because the IR1 and IR2 response functions are well separated, which also results in better cloud classification. The visible channel covers a broad band. Gas absorption and aerosol/cloud scattering are included in one channel. Thus, accurate radiative transfer calculation is important. Because the MTSAT-1R/JAMI visible channel was not calibrated, a temporary calibration value obtained by vicarious calibration was used in this study. An accurate vicarious calibration scheme must be developed for improved retrievals. The received data were used at the merged 4 km resolution.

3.2. ADEOS-II/GLI Standard Product (Atmosphere)

[18] The GLI has 36 channels that can observe the land, ocean, and atmosphere [Nakajima et al., 1999]. It can retrieve detailed particle optical properties in the atmosphere. ADEOS-II/GLI products are used to examine the influence of the detailed optical properties of atmospheric particles. The latest aerosol, cloud, and surface albedo products were used as a test version (<http://suzaku.eorc.jaxa.jp/GLI/index.html>).

3.3. JRA-25 and JCDAS

[19] JRA-25 is a reanalysis data set for the period 1979 to 2004. The JRA-25 project was undertaken to build a long-term global data set based on observational data with data assimilation [Onogi et al., 2007]. The reanalysis period began with the start of satellite observations in 1979. Some satellite data were used for the reanalysis. Brightness temperature of the Television Infrared Observation Satellite (TIROS) Operational Vertical Sounder (TOVS)/Advanced TOVS (ATOVS) Level 1C data was directly used by three-dimensional variational data assimilation. Some satellite data were incorporated from the middle of the reanalysis period, i.e., SSM/I snow cover and precipitable water vapor retrieval and GMS and METEOSAT recalculated wind properties. The JMA Climate Data Assimilation System (JCDAS) uses the same systems as JRA-25 to produce operational analysis. In this study, atmospheric profile data were used for cloud and radiation analysis.

3.4. TOMS and OMI Ozone Data

[20] TOMS and OMI retrievals provide total column ozone. TOMS observations were conducted from 1978 to 2005 (Nimbus-7: 1978–1993, Meteor-3: 1992–1994, and Earth Probe: 1996–2005), while those of OMI are available from 2004 to the present (Earth Probe). These instruments retrieve ozone using two wavelengths (weakly and strongly absorbing UV channels). The JRA ozone data were replaced by the TOMS and OMI ozone data.

4. Retrieval of Cloud Optical Properties

[21] We used the Comprehensive Analysis Program for Cloud Optical Measurement (CAPCOM) [Nakajima and Nakajima, 1995; Kawamoto et al., 2001] to retrieve cloud

optical thickness from MTSAT-1R/JAMI. This algorithm was also adopted as one of the standard algorithms for the Advanced Earth Observing Satellite II/Global Imager (ADEOS-II/GLI) products [Nakajima et al., 1999]. The original CAPCOM estimated the cloud optical thickness, effective particle radius, and cloud-top temperature from visible, near-infrared, and thermal infrared channels using a LUT calculated from radiative transfer [Nakajima and Tanaka, 1986, 1988] under a plane-parallel and single-layer cloud model. However, we applied CAPCOM to MTSAT-1R/JAMI data to retrieve only the cloud optical thickness using the VIS channel because the observed radiance of the MTSAT-1R/JAMI NIR channel will have cross-talk components and must be characterized more carefully.

[22] Briefly, cloud optical thickness is obtained as follows. First, each pixel is classified using the IR1 and IR2 split-window technique. Second, cloud optical thickness is retrieved from the MTSAT-1R VIS channel with corresponding atmospheric model profile data and scan geometries at the observation point, such as satellite and solar zenith angles and their azimuth angles. In the inversion process, multiple scattering components due to bright surfaces are corrected by the following:

$$L_{obs}(\tau; \mu, \mu_0, \phi) = L(\tau; \mu, \mu_0, \phi) + t(\tau; \mu) \frac{A_g}{1 - \bar{r}(\tau)A_g} t(\tau; \mu_0) \frac{\mu_0 F_0}{\pi}, \quad (28)$$

where τ is cloud optical thickness, μ and μ_0 are cos(satellite viewing angle) and cos(solar zenith angle), respectively, ϕ is the Sun-satellite azimuthal angle, \bar{r} is spherical albedo in the channel, and A_g is surface albedo in the channel. F_0 is solar irradiance outside the atmosphere weighted by the sensor response function. The function of t is atmospheric transmittance. The first term on the right side of the formula is cloud scattering, and the second term is interaction at the ground surface. A Lambert surface was assumed, and albedo was estimated from MTSAT-1R data by a simple statistical method with a clear-sky condition.

5. Results and Discussion

[23] Section 2 described the learning algorithm and construction of the NN solver. We applied the extreme speed and approximation module multiple drive system (EXAM SYSTEM) that controls NN solvers by multithreading. EXAM SYSTEM can drive NN solvers at the same time according to the number of central processing unit (CPU) cores and the capacity. An advantage of EXAM SYSTEM for geostationary satellite applications is its high calculation speed. We applied it to the MTSAT-1R data and validated the results using in situ observation data. In addition, we tested our solar flux estimations using the standard product of ADEOS-II/GLI. We discuss the ability of the NN to handle multiple parameters of particle optical properties.

5.1. Estimation of Shortwave Flux and Validation by in Situ “SKYNET” Observations of Surface Downward Radiation

[24] EXAM SYSTEM with NN solvers was applied to MTSAT-1R/JAMI data (Figure 12) to estimate the solar

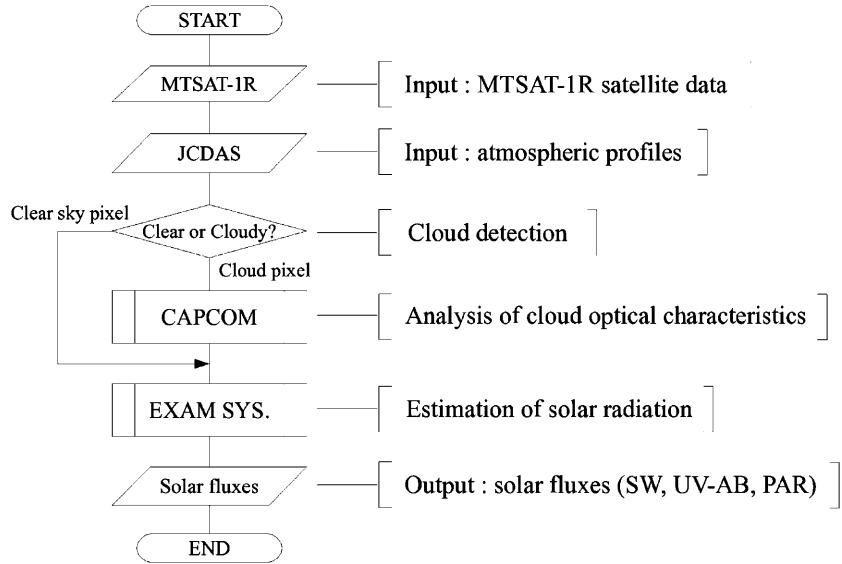


Figure 12. Flowchart for estimation of solar radiation using MTSAT-1R data. First, MTSAT-1R data are input with the retrieved surface albedo. The albedo data are for a clear-sky condition, and a Lambert surface is assumed. The corresponding JCDAS data sets are also input. Second, cloud detection is conducted based on a split-window method and reflectivity, and cloud analysis is executed by CAPCOM. Third, the solar radiation is estimated by EXAM SYSTEM, which outputs the SW, UV, and PAR fluxes.

radiation at the TOA and the surface via high-speed computation of temporal data. The study area covered 60°S to 60°N latitude and 80°E to 160°W longitude. EXAM SYSTEM logically sliced the area into eight regions and calculated the fluxes using eight CPU cores at the same time and multiple threads. Figure 13 shows the parallel performance of the NN algorithm. The computing time has been compressed to about one sixth compared to a single thread. Higher performance will be enabled using a processor optimized for parallel calculation by many cores (for example: application of general purpose computing on graphics processing units, GPGPU). Figure 14 presents an example set of flux images. The umbrella effect of clouds can be seen at the surface (Figure 14a) and the TOA (Figure 14b). Interception of solar radiation by clouds is remarkable in the direct component (Figure 14d). However, the strong scattering in the diffuse component is attributable to thin cloud (Figure 14c).

[25] For validation, the estimated solar radiation was compared with in situ ground observations by the radiation observation network SKYNET. SKYNET is an atmospheric radiation and weather observation network developed to understand aerosol-cloud-radiation interactions in the atmosphere (<http://atmos.cr.chiba-u.ac.jp/>). Downward SW flux observed by a CM21 Kipp and Zonen pyranometer at the Chiba/SKYNET site (35.62°N, 140.10°E) was used for validation of the downward SW flux at the surface. The CM21 pyranometer is classified as a World Meteorological Organization (WMO) “secondary standard” (i.e., highest quality) device, with nominal sensitivity from 7 to 17 $\mu\text{V}/\text{W m}^{-2}$. Some typical examples of observations are shown in Figure 15. Good correlation between the estimations and observations was found under the clear-sky condition (Figure 15a), and reasonable agreement was obtained under the cloudy condition (Figure 15b), showing that thick

cloud has a strong decremental effect on solar radiation. Figure 15c represents the coexistence of cloudy and clear-sky conditions within a day. In the condition of broken clouds (Figure 15d), accuracy was poor because satellite sensors cannot yet capture phenomena at very small sub-pixel levels. Figure 16 presents the corresponding cloud conditions. The error of each input parameter that contributes to the estimated flux should be considered. The solar zenith angle and Sun-Earth distance are calculated based on the satellite scan time, and the errors are suffi-

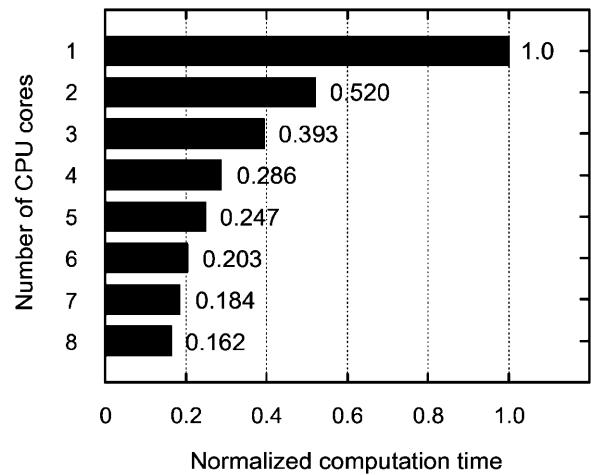
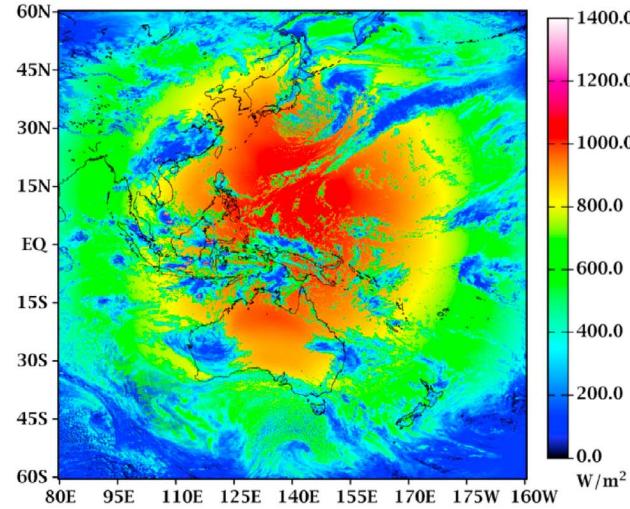
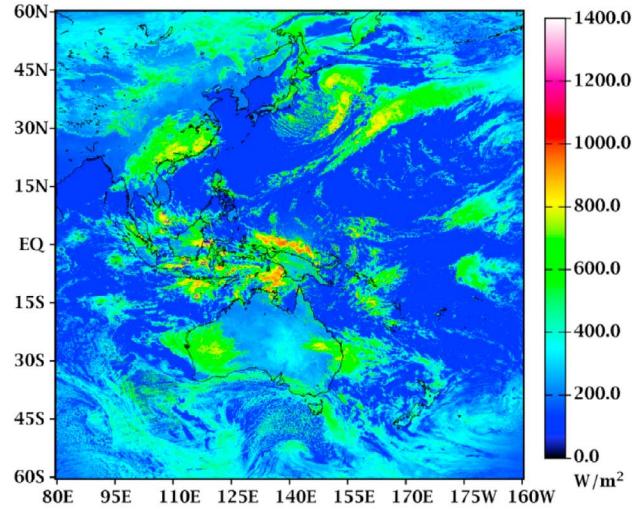


Figure 13. Parallel performance of the NN algorithm. Each computing time was normalized by a single thread condition. One thread was executed in one core $\text{thread/core} = 1$. The computing time was compressed to one sixth by the parallel computation of eight threads using eight CPU cores.

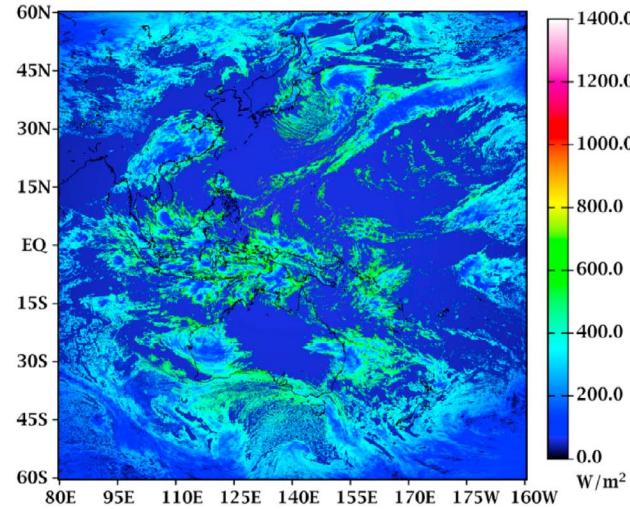
(a) Downward SW flux at the surface.



(b) Upward SW flux at the TOA.



(c) Diffuse component of downward SW flux at the surface.



(d) Direct component of downward SW flux at the surface.

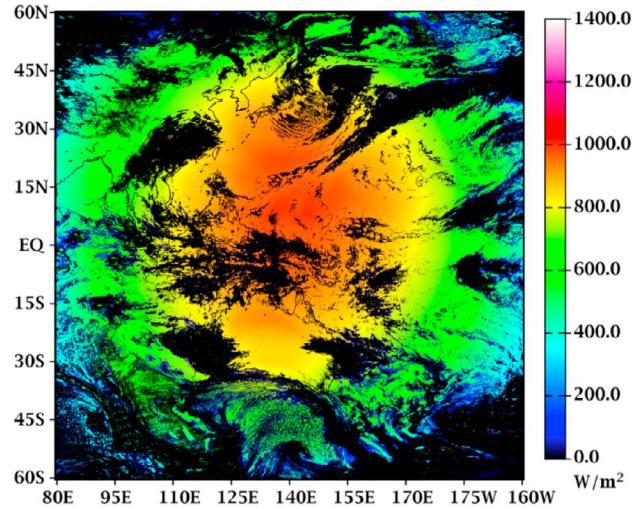


Figure 14. Estimated shortwave radiation (0300 UTC, 31 March 2006). (a) Downward SW flux at the surface. (b) Upward SW flux at the TOA. (c) Diffuse component of downward SW flux at the surface. (d) Direct component of downward SW flux at the surface.

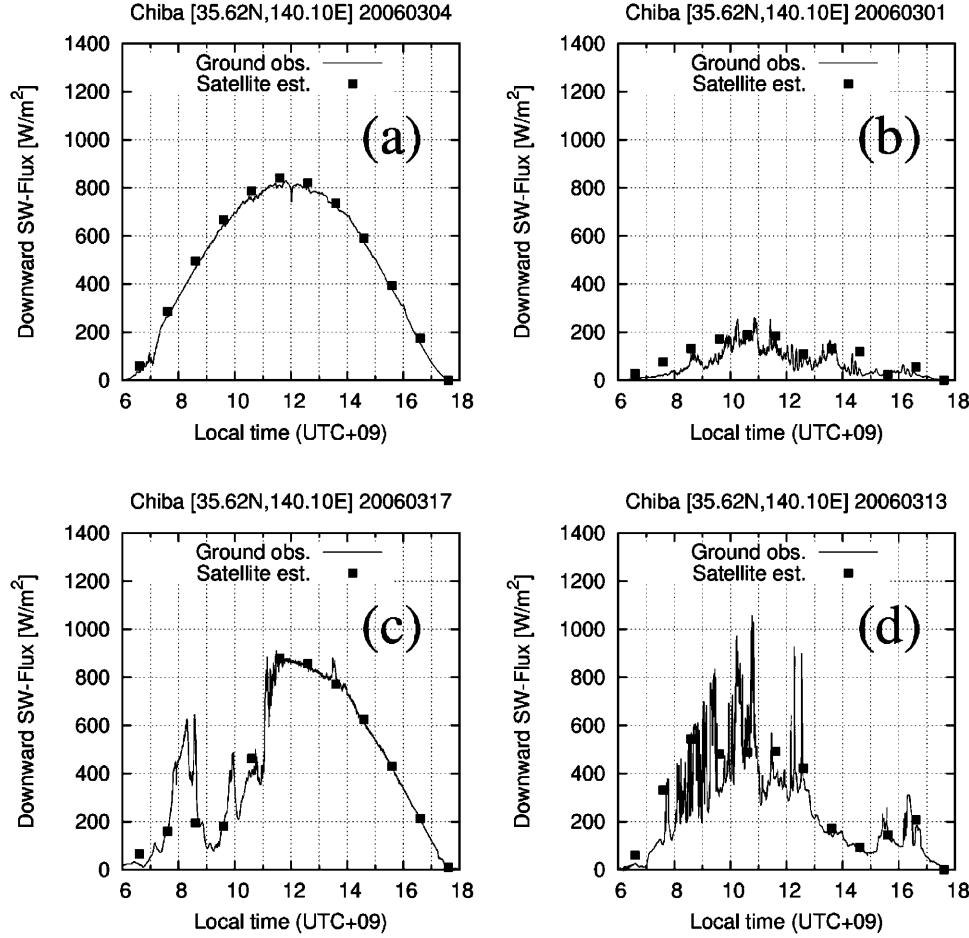


Figure 15. Typical example of downward SW flux at the surface observed at the Chiba/SKYNET site (35.62°N , 140.10°E). The x axis gives local time. (a) Clear-sky condition, 4 March 2006. (b) Cloudy condition, 1 March 2006. (c) Coexistence condition, 17 March 2006. (d) Broken cloud condition, 13 March 2006.

ciently small. The accuracy of the parameters that decide the total optical thickness of the atmosphere directly controls the estimation accuracy of the flux. The sensitivity of atmospheric parameters to the downward SW flux is shown in Figure 17. The error margin is as follows:

$$R_{SFC}^{\downarrow} = q(x_1, x_2, x_3, \dots, x_n) \quad (29)$$

$$\begin{aligned} \text{Error} &= \frac{\Delta R_{SFC}^{\downarrow}}{R_{SFC}^{\downarrow}} \times 100[\%] \\ &= \frac{1}{R_{SFC}^{\downarrow}} \left\{ \frac{\partial R_{SFC}^{\downarrow}}{\partial \tau_{atm}} \frac{\partial \tau_{atm}}{\partial x_i} \right\} \Delta x_i \times 100[\%], \end{aligned} \quad (30)$$

$x_1, x_2, x_3, \dots, x_i, \dots, x_n$ represents atmospheric radiative parameters, and equation (29) shows the radiative transfer calculation (simple form of equation (2)). Equation (30) gives the error margin of R_{SFC}^{\downarrow} based on the error of the atmospheric parameter Δx_i , which is decided by a random number. The errors caused by absorption gases (water vapor, ozone) are small (Figures 17c and 17d). The surface pressure depends mainly on the geographical altitude and is based on reanalysis data. Thus, the error in the surface

pressure is small (less than 5%) excluding complex topography. The difference in cloud microphysical properties causes the biggest error, as shown in Figure 17a. The retrieval error of the CAPCOM algorithm is smaller than 5% in optical thickness and smaller than 10% in effective radius [Nakajima and Nakajima, 1995]. The behaviors of calibration error and stray light in the sensor are complicated and may increase the errors in the optical properties. The cloud optical thickness is the most sensitive parameter in the algorithm, to which the most attention should be paid. One of the biggest issues is the nonuniformity of the target in the sensor field of view. The assumption of the plane parallel atmosphere is not appropriate to nonuniform condition. Very small cloud fragments and/or cloud holes less than the pixel resolution can cause errors larger than those shown in Figure 17. Future technical breakthroughs are expected to enable satellite sensors to make accurate observations even when broken clouds exist (for instance, by increasing the apparent resolution by application of super-resolution techniques [Park et al., 2003; Ling et al., 2005] based on very high temporal (several seconds) resolution or simultaneous multiangle scanning). Such advances will help improve the accuracy of the solar radiation estimation at the

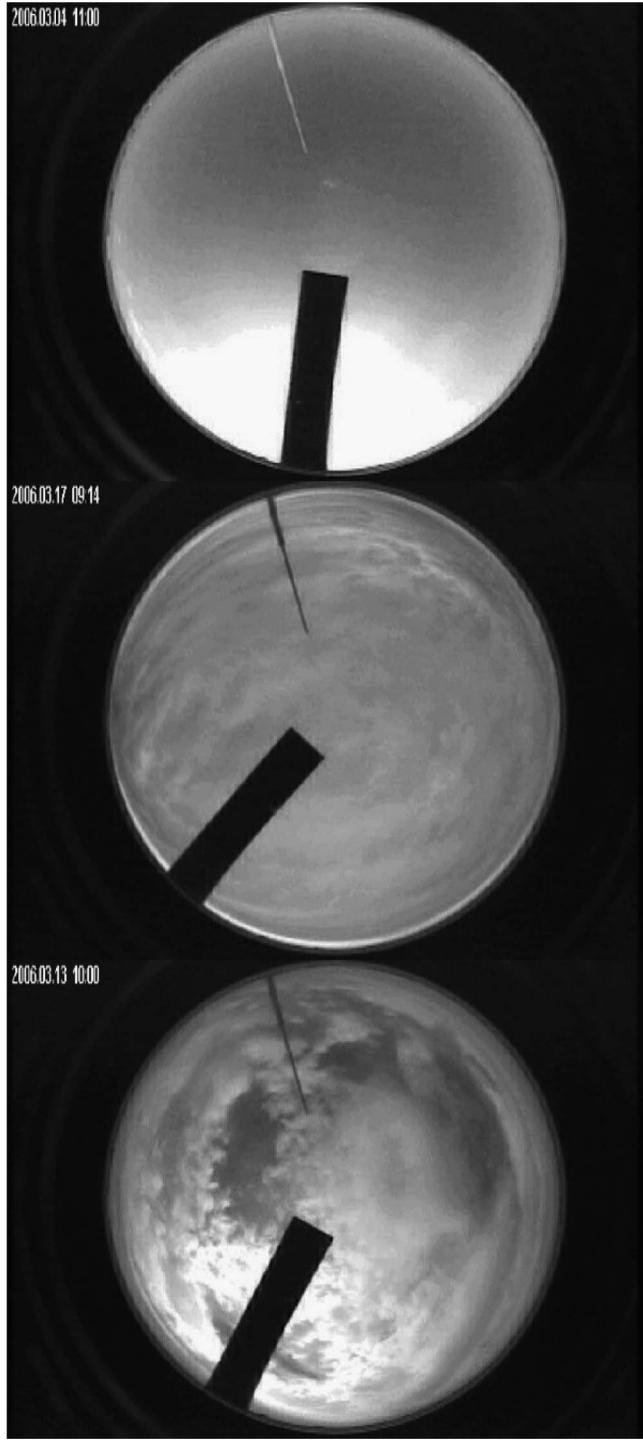


Figure 16. Sky condition corresponding to that in Figure 15, photographed by a camera with a CCD fisheye lens. (top) Clear-sky condition at 1100 local time (LT), 4 March 2006. (middle) Cloudy condition at 0914 LT, 17 March 2006. (bottom) Broken cloud condition at 1000 LT, 13 March 2006.

expense of an increased number of parameters (optical properties and spatial distribution of the cloud) incorporated in the calculation using a three-dimensional radiative transfer model. In this case, the NN solver would be a useful

tool because, unlike LUT approaches, the NN does not require large databases (memory space) and complex interpolation/extrapolation.

[26] The scatter diagram in Figure 18 compares downward SW flux between the satellite estimation and ground observation for March 2006 at the Chiba/SKYNET site and three other SKYNET sites (Cape Hedo, Fukue-jima,

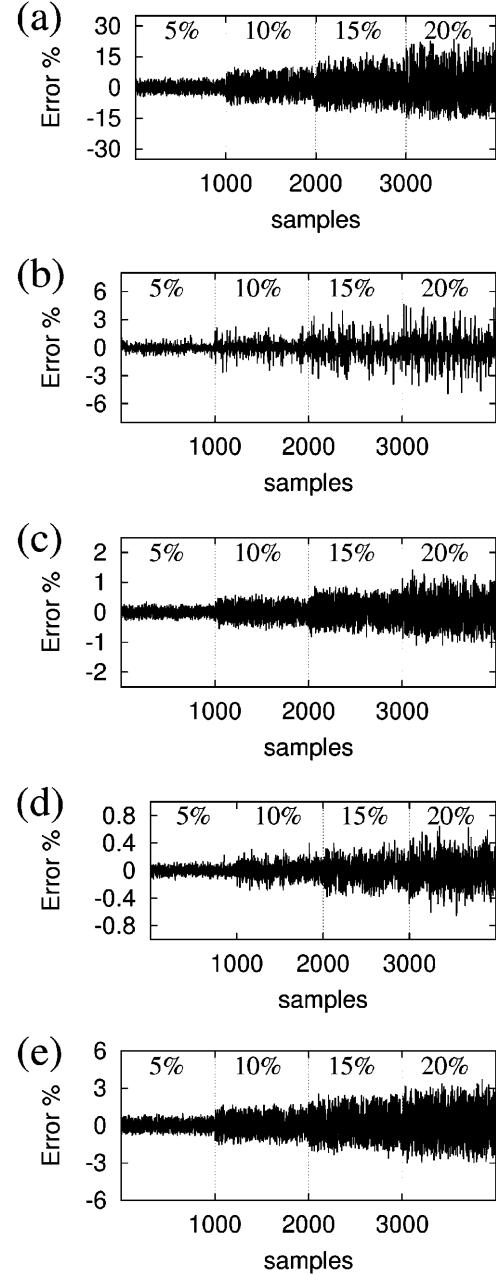


Figure 17. Sensitivity analysis of the atmospheric parameters in the downward SW flux at the surface. The x axis gives the calculation samples. The calculation conditions were generated by random numbers, and the random error (5%, 10%, 15%, 20%) was added to each parameter: (a) cloud optical thickness, (b) cloud droplet effective radius, (c) column water vapor, (d) column ozone, and (e) surface pressure. The y axis shows the error margin in downward SW flux at the surface.

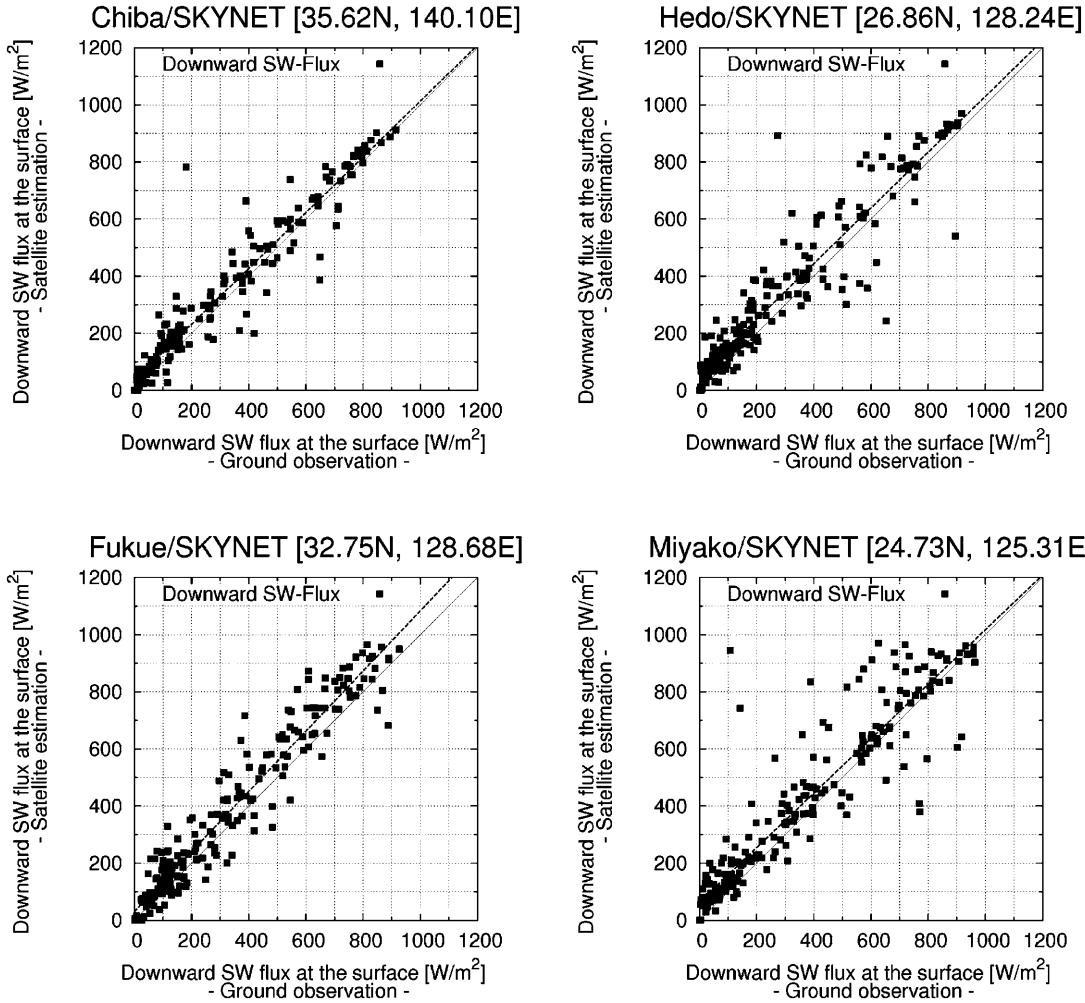


Figure 18. Comparison of downward shortwave flux at the surface by ground observation and the satellite estimation for the SKYNET sites. The x axis gives the ground-observed data, while the y axis shows the satellite-based estimation.

Miyako-jima). The results are shown in Table 2. The bias was ascribed mainly to the effect of aerosols, as aerosol effects were not considered in this case. Except for data points showing large deviations from the 1:1 line, which were probably due to broken clouds, each result indicates that the estimation algorithm yielded reasonable values.

5.2. Application of the NN Solver for UVA, UVB, and PAR Flux

[27] This section discusses solar radiation bands that should be considered when applying the NN solver. The UVA and UVB bands classify ultraviolet radiation, with UVA covering the wavelength region of 0.315 to 0.4 μm and UVB including the region from 0.28 to 0.315 μm . UV fluxes can have large impacts on living beings by damaging the collagen forming the skin [de Gruyl, 1999]. UVB can cause eye damage as well as sunburn, and it is a trigger of skin cancer. UVA makes up a large amount the energy that passes through the atmosphere; this radiation can deeply infiltrate the skin, causing tanning and damage. Figure 19 shows examples of the UVA and UVB analysis. Cloud does not readily influence UVB, although UVB energy is less than UVA energy. The weighting ratios of direct and

diffuse components are reversed between UVA and UVB. Few past studies have estimated the contribution of ultra-violet rays within the Earth system. However, given the above-mentioned biological effects of these rays, the ability to estimate and monitor them is important and can be provided by the NN solver.

[28] Photosynthetically active radiation (PAR) is related to the growth of plants and directly affects carbon uptake in ecosystems. The direct and diffuse components of PAR fluxes provide information for better understanding the carbon cycle and predicting plant growth. PAR is observed

Table 2. SKYNET Sites and Validation Results

SKYNET Site		Validation Result			
Name	Location	Linear Regression ^a		Correlation	RMS
		a	b		
Chiba Univ	35.62°N, 140.10°E	0.992	33.54	0.93	79.41
Cape Hedo	26.86°N, 128.24°E	0.975	53.50	0.89	100.26
Fukue-jima	32.75°N, 128.68°E	1.054	29.80	0.93	88.27
Miyako-jima	24.73°N, 125.31°E	0.956	55.76	0.87	102.28

^aLinear regression: $y = ax + b$.

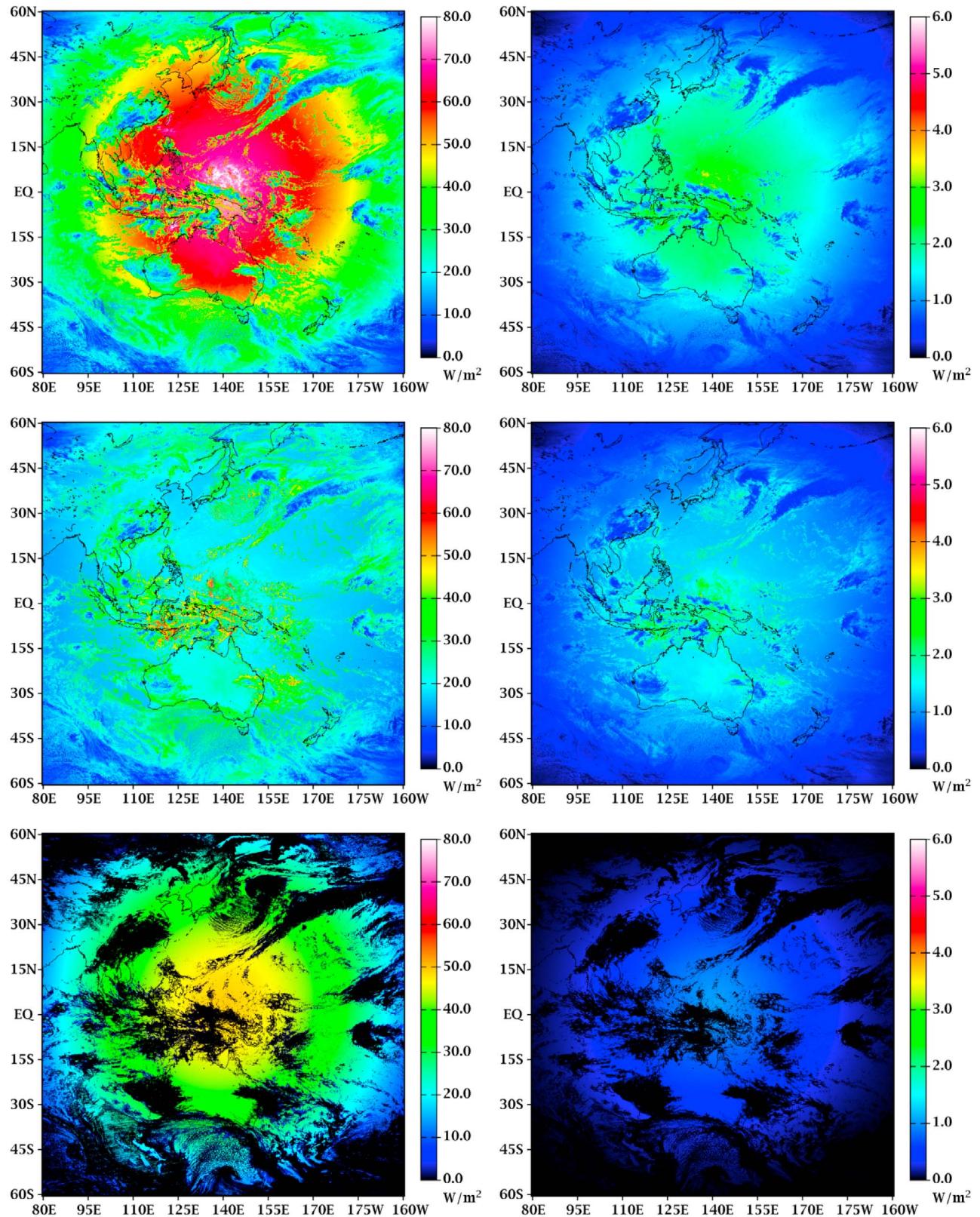


Figure 19. Downward UV flux at the surface (0300 UTC, 31 March 2006). (left) UVA. (right) UVB. (top) Global downward UV flux. (middle) Diffuse component. (bottom) Direct component.

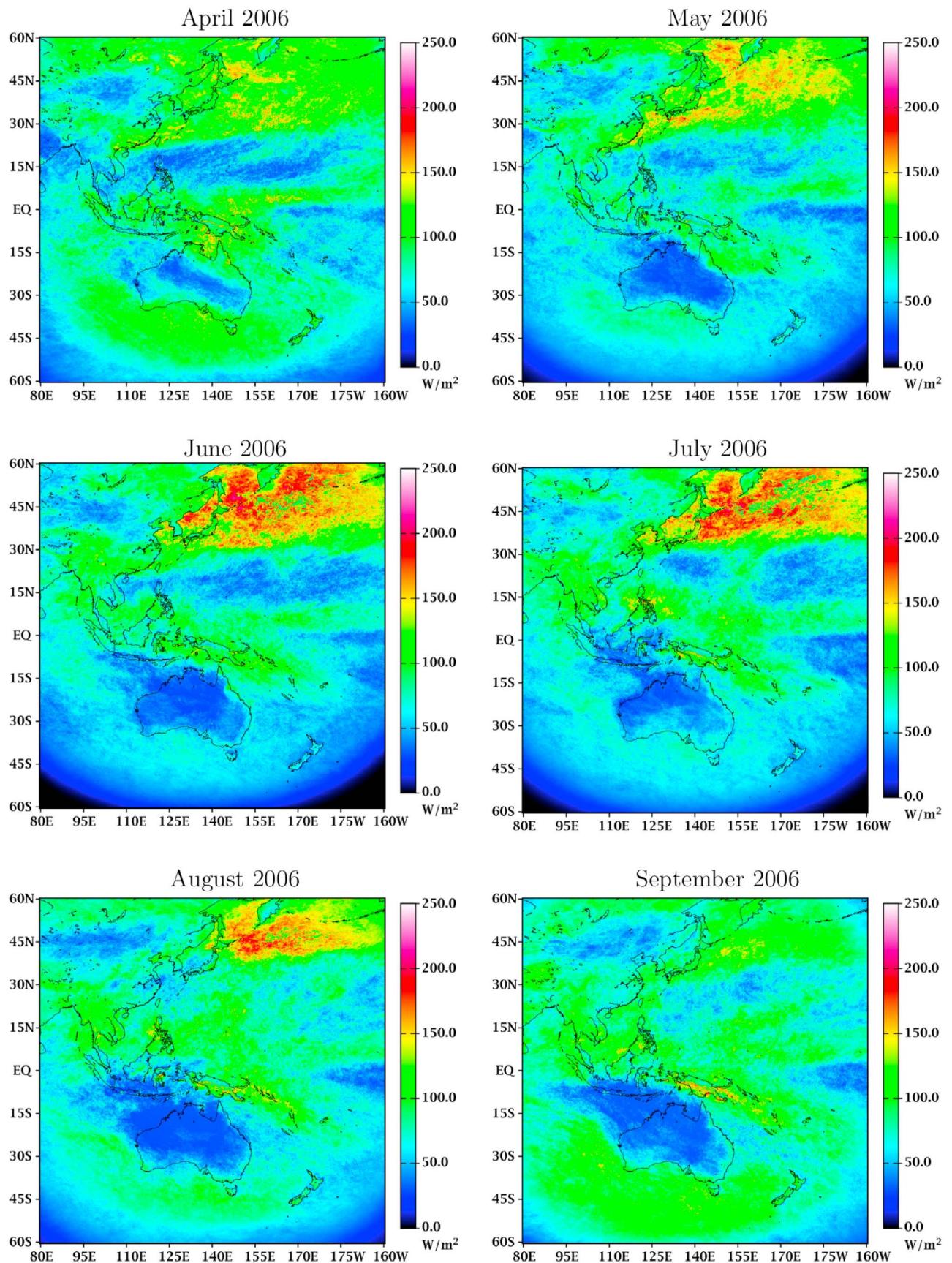


Figure 20. Downward PAR flux at the surface, diffuse component. Monthly average at 0300 UTC.

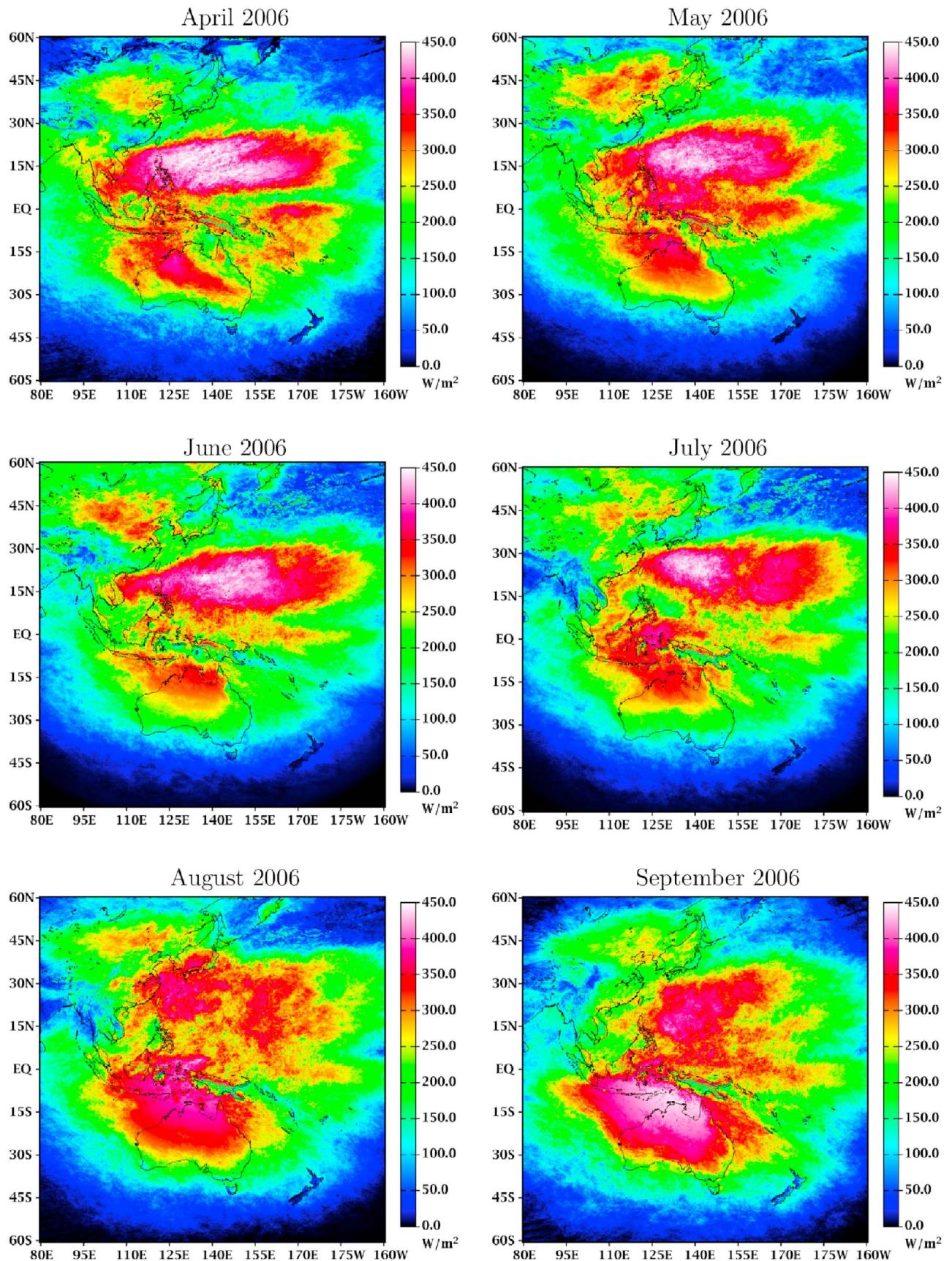


Figure 21. Same as Figure 20 but for the direct component.

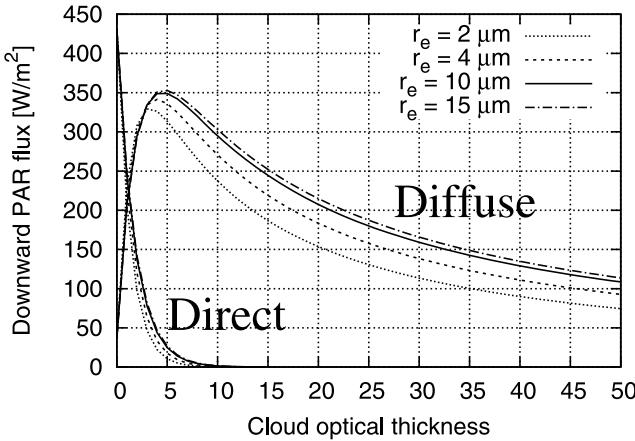


Figure 22. Influence of cloud on downward PAR flux at the surface, based on classification of the cloud droplet effective radius. The lines indicate different cloud droplet effective radius values. (Solar zenith angle = 20.0 degrees, U.S. Standard Atmosphere, water particle cloud top pressure = 540 hpa, and surface albedo = 0.01.)

in the wavelength range of 0.395 to 0.715 μm ; the influence of gases is not detected, except for very weak absorption of water vapor and ozone. Because gas absorption is negligible, the PAR range is good for evaluating the influence of particles such as cloud particles. Figures 20 and 21 present monthly average PAR flux of diffuse and direct components, respectively, at 0300 UTC from April to September 2006. The direct component is dominant over Australia, where the cloud fraction is small. Features similar to those over Australia are also found over the subtropical Western Pacific around 140°E, 15°N. The diffuse component is predominant in the midlatitude region. From East Asia to the North Pacific Ocean, the diffuse component increases until June and thereafter seems to decrease. An inverse correlation was found for the direct component. These results indicate that the interaction between the cloud and the solar zenith angle changes with the seasonal march, a trend caused by the cloud radiative effect. In Figure 22, the diffuse component is stronger for thin cloud than for thick cloud, with an inverse correlation between direct and diffuse components in the range of $\tau = 0 \sim 5$. In evaluating the diffuse component, the detection and distribution of clouds with low optical thickness is important to the radiative energy distribution. The estimation algorithm provided by the NN solver would be effective for monitoring the distribution of direct and diffuse components at high temporal resolution. These operations were applied using semi-real-time analysis (<http://atmos.cr.chiba-u.ac.jp/takenaka/en/>).

5.3. Application of the NN to Various Optical Properties

[29] In this section, we discuss the advantages of the NN approach over the LUT method for multiparameter estimations of solar radiation. We present an example of solar flux estimation including the detailed optical properties of cloud and aerosol retrieved by the ADEOS-II satellite. ADEOS-II is a polar orbit (Sun-synchronous subrecurrent orbit) satellite that performs about once-daily sampling; recurrence period = 4 days. Estimation of the solar flux, including information

on strict optical properties, will be helpful for better understanding the Earth's energy budget. In this study, we examined the solar radiation, including the cloud [Nakajima and Nakajima, 1995; Kawamoto et al., 2001] and aerosol [Higurashi and Nakajima, 1999, 2002] optical properties retrieved by the GLI (Figure 23). A cloud flag was used to categorize each pixel as cloudy or clear sky. The pixels included cloud and aerosol optical properties. The parameters for the cloudy condition were the same as presented in section 5.1. For clear sky, aerosol optical thickness and the imaginary part of the refractive index, in addition to the size distribution corrected by the Ångstrom exponent, were applied. The aerosol bimodal lognormal volume spectrum was

$$\frac{dV}{d \ln r} = \sum_{n=1}^2 C_n \exp \left[-\frac{1}{2} \left(\frac{\ln r - \ln r_{mn}}{\ln S_n} \right)^2 \right] \quad (31)$$

where $r_{m1} = 0.17 \mu\text{m}$, $r_{m2} = 3.44 \mu\text{m}$, $S_1 = 1.96$, and $S_2 = 2.37$. This particle size distribution is based on the GLI aerosol product [Higurashi and Nakajima, 1999, 2002], where V and r_m are the volume density and mode radius and $\ln S_n$ is the standard deviation of $\ln r$. The peak ratio of the bimodal size distribution $\gamma = C_2/C_1$ was added to the parameter list. The imaginary part of the refractive index was calculated based on the ratio of absorption and non-absorption aerosols. Eight parameters were included in the aerosol conditions (solar zenith angle, surface albedo, surface pressure, ozone, water vapor, and three aerosol optical properties). EXAM SYSTEM integrated these atmospheric radiation parameters into the solar fluxes. Typical examples by aerosol parameters are shown in Figures 24 and 25. The effect of the imaginary part of the refractive index was small in the SW flux direct component (Figure 24a) but showed an absorption effect in the diffuse component (Figure 24b). The effect of the ratio of the bimodal size distribution was

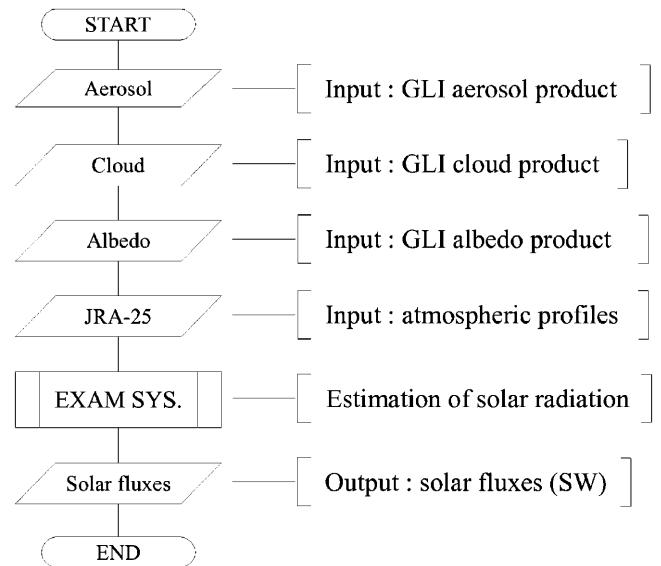


Figure 23. Flowchart of solar radiation estimation using GLI products. The GLI products (cloud and aerosol optical properties and surface albedo) and JRA-25 data set are input as the first step. The solar radiation is estimated based on the input data, and SW fluxes are obtained as output.

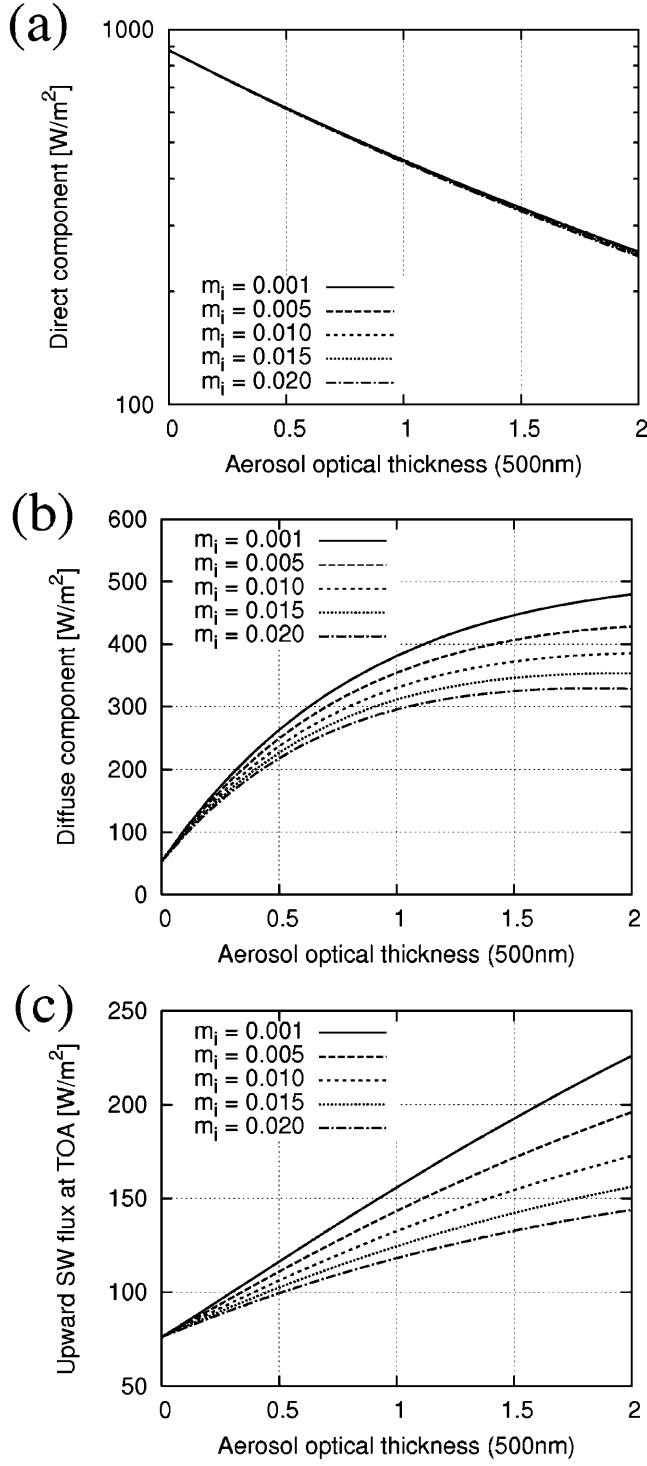


Figure 24. Influence of aerosols on the SW fluxes, based on classification of the imaginary part of the refractive index m_i ($m = m_r - im_i$; $m_r = 1.5$). (a) Direct component of downward SW flux at the surface. (b) Diffuse component of the downward SW flux at the surface. (c) Upward SW flux at the TOA. (Solar zenith angle = 30.0 degrees, U.S. Standard Atmosphere, $m_i = 0.005$, and surface albedo = 0.03).

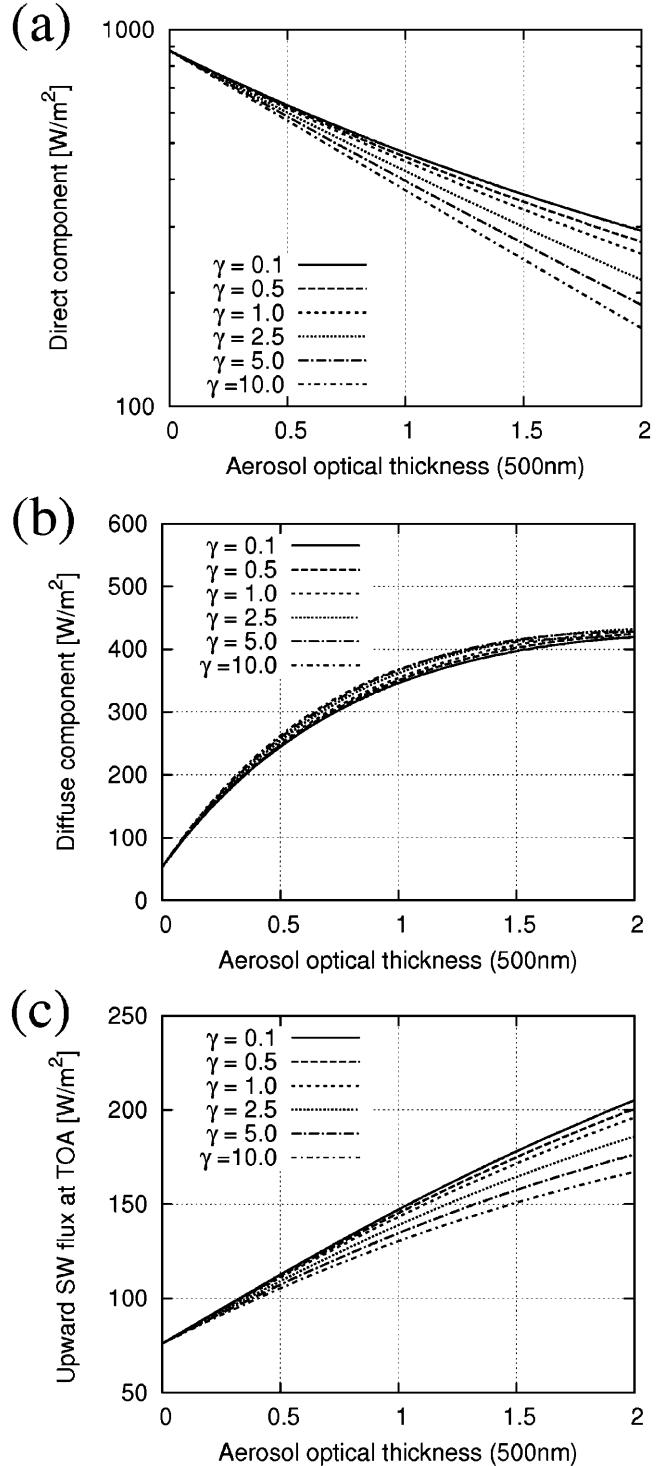


Figure 25. Same as Figure 24 but for classification of the bimodal size distribution $\gamma = C_2/C_1$. (Solar zenith angle = 30.0 degrees, U.S. Standard Atmosphere, $m_i = 0.005$, and surface albedo = 0.03.)

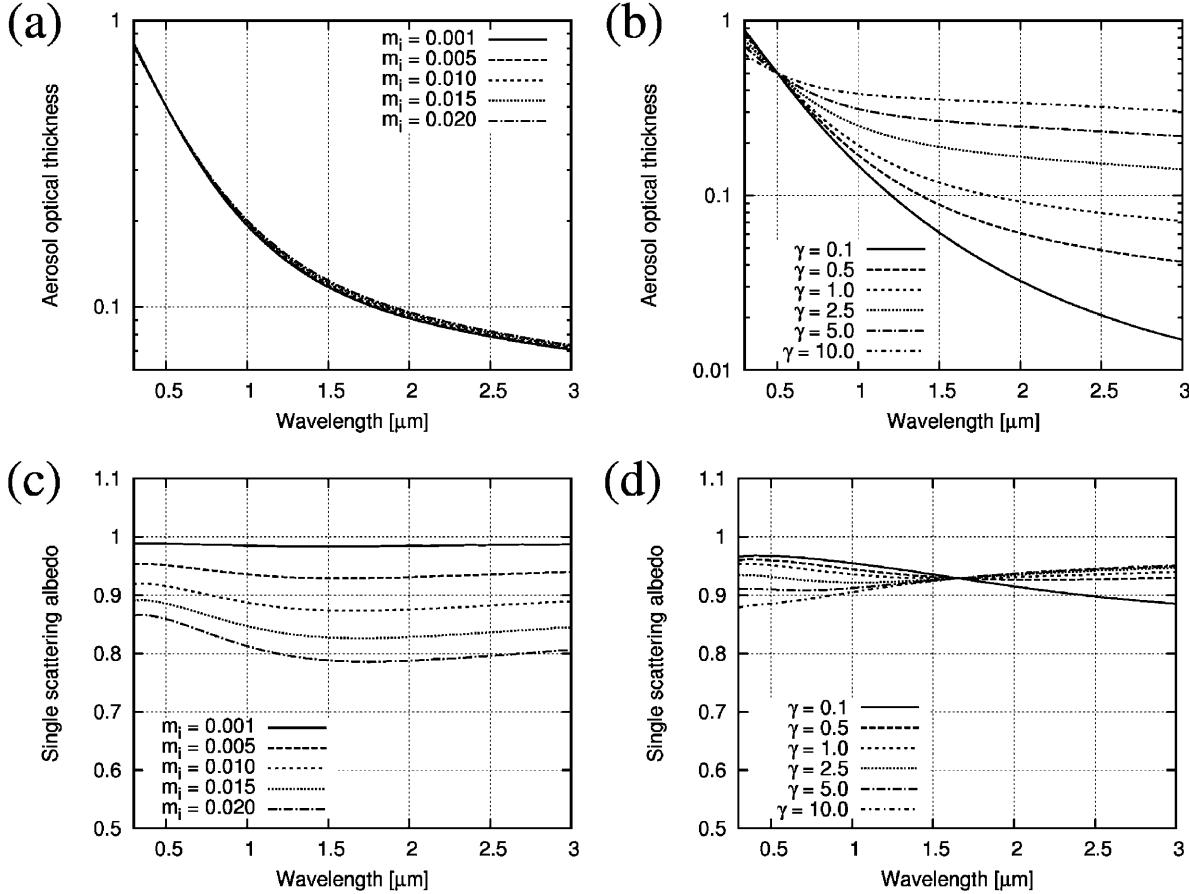


Figure 26. Influence of optical parameters m_i and γ on τ and SSA in each wavelength. Optical thickness: (a) Classification of the imaginary part of the refractive index m_i and (b) classification of the ratio of the bimodal size distribution γ . Single scattering albedo: (c) classification of the imaginary part of refractive index m_i and (d) classification of the peak ratio of the bimodal size distribution γ . (Aerosol optical thickness at 500 nm = 0.5.)

small (Figure 25b), and the SW direct component declined (Figure 25a). Aerosols had a cooling effect on the ocean (dark target) in the upward flux at the TOA (Figures 24c and 25c). Both the m_i and γ parameters caused the TOA upward flux to decrease. These trends were based on the wavelength dependency (Figure 26). All parameters are shown in Table 3.

[30] Figure 27 shows an example of the analysis results. Strong global flux was distributed at low latitudes (around the equator). Global flux was weaker at the equator line, indicating the influence of clouds of the Intertropical Convergence Zone (ITCZ). Overall, the diffuse component predominated in the Northern Hemisphere side. This trend was remarkably apparent over the North Pacific and North Atlantic oceans and may have related closely to cloud and aerosol activity in the Northern and Southern hemispheres. In the North Pacific and North Atlantic oceans, thin cloud and aerosol caused strong scattering. Thus, Northern Hemisphere midlatitude regions of the Pacific and Atlantic oceans were kept bright by the diffuse component, and blocking of the direct component was a distinct trend. This trend was similar to the PAR result presented in section 5.2, especially over East Asia to the North Pacific. These findings may provide fragments of information about aerosol-cloud-

radiation interactions. However, all the orbit paths were for local noon because ADEOS-II has a Sun-synchronous orbit. Consideration of the relationship between time and place is necessary to understand the phenomenon in detail. A more detailed discussion will be attempted after the new GLI product for land-surface aerosol is completed. The NN approach will enable analysis even though aerosols over the

Table 3. Input Parameters of NN Solver for Polar Orbit Satellite ADEOS-II/GLI

Parameter	Symbol	Unit
Solar zenith angle	θ_0	deg
Aerosol optical thickness	τ_a	none
Peak ratio of bimodal size distribution	γ	none
Imaginary part of refractive index	m_i	none
Water cloud optical thickness	τ_{ew}	none
Water cloud effective radius	r_{ew}	μm
Ice cloud optical thickness	τ_{ei}	none
Ice cloud effective radius	r_{ei}	μm
Cloud top pressure	P_c	hpa
Surface albedo	A_g	none
Surface pressure	P_0	hpa
Total column ozone	OZN	DU
Total column water vapor	PWV	g/m^2

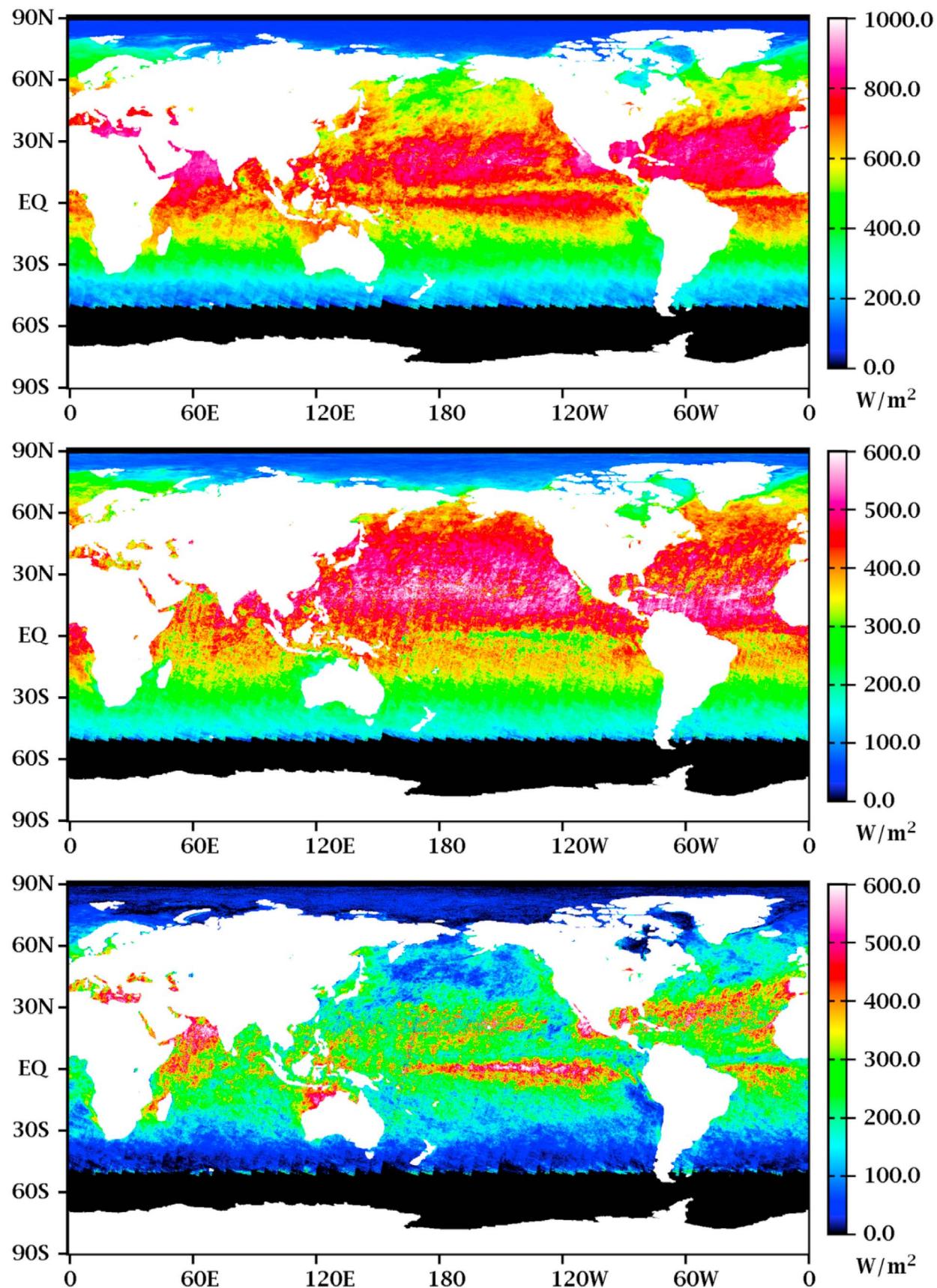


Figure 27. Monthly average downward SW flux at the ocean surface (May 2003). (top) Global downward SW flux. (middle) Diffuse component. (bottom) Direct component.

land region will also require many parameters of optical properties.

[31] MTSAT-1R cannot obtain detailed particle optical properties because the JAMI sensor does not have important channels for retrieval. If the algorithm and technical feedback of the ADEOS-II/GLI are improved in geostationary satellites, such detailed analysis can be achieved by the NN approach with high temporal resolution.

6. Conclusion

[32] This study developed an algorithm for solar radiation estimation using a neural network (NN). The NN has a simple structure that enables high-speed calculation and does not require large amounts of data as required by a lookup table (LUT) approach. Thus, advantages of the NN include high-speed estimation and the ability to handle multiple parameters. The NN solver with the improved learning algorithm produced accurate function approximations. EXAM SYSTEM with NN solvers was applied to MTSAT-1R data and validated by in situ, ground-observed solar irradiance data from the SKYNET sites. The results showed reasonable performance, except in the case of broken clouds; for the all-sky condition in March 2006. This study verified the algorithm at only four observation sites. A more wide-ranging validation at several field sites will be required. Additionally, specific bands, namely UVA, UVB, and PAR fluxes, were used to demonstrate the method. Monthly averaged PAR fluxes showed that clouds cause a strong diffuse component and blocking of the direct component from East Asia to the North Pacific Ocean according to the seasonal march. The high-speed calculation capabilities of the NN allow for semi-real-time analysis. In addition, the NN was applied to the particle product derived from ADEOS-II/GLI for the estimation of solar radiation. The results demonstrated another advantage of the NN, namely its ability to handle multiple parameters. As a test of the method, solar fluxes were estimated by including aerosol and cloud optical properties. The results showed that the diffuse component predominated in the downward SW flux at the surface of the North Pacific and North Atlantic oceans, with blocking of the direct component. Over the North Pacific, a trend similar to that of PAR flux was found. The midlatitude region of the North Pacific remained bright due to aerosol and cloud particles, possibly representing a regional characteristic of surface radiation based on the influence of aerosol and cloud activity. These results demonstrate the effectiveness of the NN approach for analysis of radiation budgets.

[33] In the future, the NN approach including more parameters can be applied to estimate three-dimensional radiative transfer for clouds. Moreover, this approach will also be effective for estimating radiation budgets that include some particle mixing (cloud and aerosol internal/external mixing or mixing of aerosol types, nonspherical shape). The NN approach allows for examination of more new parameters (for instance, the spatial distribution of cloud and optical properties in three dimensional radiative transfer) than does the LUT method because it does not require the large database spaces of a LUT. The LUT method does not allow for advanced analyses of numerous parameters; the ability to handle numerous and new para-

meters is an important advantage of the NN approach. Strict NN-based analysis of various optical properties will promote better understanding of the Earth's radiation budget. Development of the NN algorithm for estimating solar flux is the first step toward these efforts.

Appendix A: Neural Network Operation and Derivation of the Learning Algorithm

[34] This appendix provides details of the operation of the NN and the derivation of its learning algorithm. First, operation of the single-neuron and three-layer networks by the back-propagation (BP) algorithm is described. The BP algorithm is then extended by the distortion-BP approach.

A1. Operation of the Feed-Forward Network

[35] Figure A1 shows a simple model of a single neuron with one connection. In this simple model, data flows from the left side as an input to the right side as an output through the single neuron. The net value T is calculated as follows:

$$\begin{aligned} T &= \text{input} \times \text{weight} \\ &= XW, \end{aligned} \quad (\text{A1})$$

where variables X and W are an input and a synaptic weight value, respectively. Variable H is an output of the neuron through the kernel function $f(T)$. The sigmoidal function is usually used as a kernel function as follows:

$$H = f(T), \quad (\text{A2})$$

where

$$f(T) = \frac{1}{1 + e^{-T}}. \quad (\text{A3})$$

The single neuron model with a single synapse can be extended into multiple synapses with different weights. In this case, equation (A1) is replaced by the following expression:

$$\begin{aligned} T &= \text{input}_1 \times \text{weight}_1 + \text{input}_2 \times \text{weight}_2 + \dots + \text{input}_n \times \text{weight}_n \\ &= X_1 W_1 + X_2 W_2 + \dots + X_n W_n \\ &= \sum_{i=1}^n X_i W_i. \end{aligned} \quad (\text{A4})$$

Moreover, a multineuron model is introduced corresponding to the multiple inputs, and in parallel, this model can be applied to a three-layer network, as shown in Figure 6. Forward operation of the three-layer network is conducted by equations (10)–(13). The network consists of an input layer, a hidden layer, and an output layer, which have neurons n , m , and l , respectively. The neuron in the input layer generates X and propagates to the next layer. The hidden layer processes X based on equations (10) and (11) and then propagates H to the next layer. Similarly, the output layer outputs O based on equations (12) and (13). The numbers of n and l neurons are given a priori for the targeted task, but the number of m neurons is decided empirically. The NN works based on dependencies on the weighting W (neuron synaptic weight) and kernel function in each layer. The optimized variable W is the most important factor in the NN.

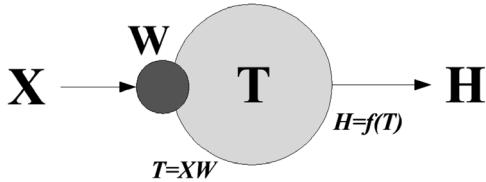


Figure A1. Schematic diagram of a single neuron function. X and W are the input parameter and synaptic weight value, respectively. T is the neuron net value. H is output by the kernel function $f(T)$.

A2. Error Back-Propagation Method

[36] The NN method depends strongly on the retrieval of neuron synaptic weight. The error BP method minimizes the error margin between network output and training data by propagating the error backward through the network. Error E is defined by equation (16). In the training process, the most suitable value of W is retrieved to minimize the value of E by the steepest descent method. The steepest descent method is applied by equations (14) and (15) as follows:

$$W^{(s+1)} = W^{(s)} - \eta \frac{\partial E}{\partial W} \Big|_{W=W^{(s)}} \quad (\text{A5})$$

where s is a nominal iteration number, and η is a learning rate that can control convergence in the learning process (usually set to $0.0 < \eta < 1.0$). In equation (A5), the synaptic weight correction at the output and hidden layers in the three-layer network is written as follows:

$$\begin{aligned} \Delta W_{kj} &= -\eta \frac{\partial E}{\partial W_{kj}} = -\eta \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial U_k} \frac{\partial U_k}{\partial W_{kj}} \\ &= -\eta \{ -(t_k - O_k) O_k (1 - O_k) H_j \} \end{aligned} \quad (\text{A6})$$

$$\begin{aligned} \Delta W_{ji} &= -\eta \frac{\partial E}{\partial W_{ji}} = -\eta \left\{ \sum_{k=1}^l \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial U_k} \frac{\partial U_k}{\partial H_j} \right\} \frac{\partial H_j}{\partial T_j} \frac{\partial T_j}{\partial W_{ji}} \\ &= -\eta \left\{ \sum_{k=1}^l -(t_k - O_k) O_k (1 - O_k) W_{kj} \right\} H_j (1 - H_j) X_i. \end{aligned} \quad (\text{A7})$$

The synaptic weights in each layer were retrieved correctly by equations (A6) and (A7) with iterations (Figure A2). A momentum term was added that requires antilocal minima. Thus, a moment parameter α is introduced:

$$\Delta W^{(s+1)} = -\eta \frac{\partial E}{\partial W} \Big|_{W=W^{(s)}} + \alpha \Delta W^{(s)}, \quad (\text{A8})$$

where $\alpha \Delta W^{(s)}$ gives a moment of inertia (generally, $\alpha = 0.9$) which contains the history of synaptic weight correction.

A3. Distortion Error Back-Propagating Method (Distortion BP)

[37] Here, we introduce the neuron survival rule and secondary parameter V . The neuron survival rule is described in section 2. In this section, we derive equations (20) and (21). Parameter V should be an independent variable

and neuron parameter. It is defined as a gain of the sigmoidal function in this study. Thus, the kernel function is extended in equations (22) and (23), and equations (20) and (21) are applied to the three-layer network (Figure 6). Correction of V is as follows:

$$\Delta V_k = -\zeta \frac{\partial E}{\partial V_k} = -\zeta \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial V_k} \quad (\text{A9})$$

$$\Delta V_j = -\zeta \frac{\partial E}{\partial V_j} = -\zeta \left\{ \sum_{k=1}^l \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial U_k} \frac{\partial U_k}{\partial H_j} \right\} \frac{\partial H_j}{\partial V_j}, \quad (\text{A10})$$

where

$$\frac{\partial E}{\partial O_k} = \frac{\partial}{\partial O_k} \left\{ \frac{1}{2} (t_k - O_k)^2 \right\} = -(t_k - O_k) \quad (\text{A11})$$

$$\begin{aligned} \frac{\partial O_k}{\partial V_k} &= \frac{\partial}{\partial V_k} \left\{ \frac{1}{1 + \exp(-U_k V_k)} \right\} = -\frac{-U_k \times \exp(-U_k V_k)}{(1 + \exp(-U_k V_k))^2} \\ &= \frac{U_k}{1 + \exp(-U_k V_k)} \left\{ 1 - \frac{1}{1 + \exp(-U_k V_k)} \right\} = O_k (1 - O_k) U_k \end{aligned} \quad (\text{A12})$$

$$\begin{aligned} \frac{\partial O_k}{\partial U_k} &= \frac{\partial}{\partial U_k} \left\{ \frac{1}{1 + \exp(-U_k V_k)} \right\} = -\frac{-V_k \times \exp(-U_k V_k)}{(1 + \exp(-U_k V_k))^2} \\ &= \frac{V_k}{1 + \exp(-U_k V_k)} \left\{ 1 - \frac{1}{1 + \exp(-U_k V_k)} \right\} = O_k (1 - O_k) V_k \end{aligned} \quad (\text{A13})$$

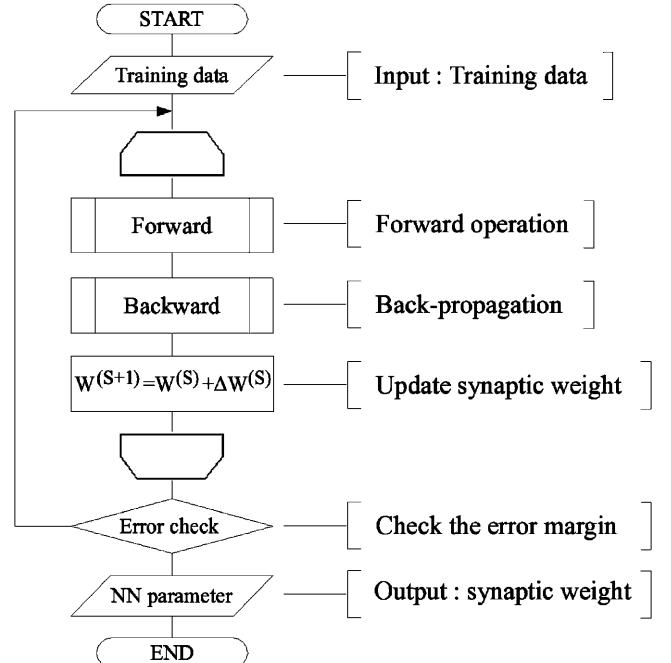


Figure A2. Flowchart of the back-propagation learning algorithm. As the first step, training data are loaded. The forward operation is executed based on the nominal synaptic weight, and the weight correction ΔW is calculated by back propagation. W is updated, and these processes are conducted in iterations. When the error margin becomes small, better synaptic weight is obtained.

$$\frac{\partial U_k}{\partial H_j} = \frac{\partial}{\partial H_j} \{H_1 W_{k1} + \dots + H_j W_{kj} + \dots + H_m W_{km}\} = W_{kj} \quad (\text{A14})$$

$$\begin{aligned} \frac{\partial H_j}{\partial V_j} &= \frac{\partial}{\partial V_j} \left\{ \frac{1}{1 + \exp(-T_j V_j)} \right\} = -\frac{-T_j \times \exp(-T_j V_j)}{(1 + \exp(-T_j V_j))^2} \\ &= \frac{T_j}{1 + \exp(-T_j V_j)} \left\{ 1 - \frac{1}{1 + \exp(-T_j V_j)} \right\} = H_j(1 - H_j) T_j. \end{aligned} \quad (\text{A15})$$

Thereby,

$$\Delta V_k = -\zeta \{-(t_k - O_k) O_k (1 - O_k) U_k\} \quad (\text{A16})$$

$$\Delta V_j = -\zeta \left\{ \sum_{k=1}^l -(t_k - O_k) O_k (1 - O_k) V_k W_{kj} \right\} H_j(1 - H_j) T_j. \quad (\text{A17})$$

Correction of W is changed by the modified kernel function. The modified ΔW is as follows:

$$\Delta W_{kj} = -\eta \frac{\partial E}{\partial W_{kj}} = -\eta \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial U_k} \frac{\partial U_k}{\partial W_{kj}} \quad (\text{A18})$$

$$\Delta W_{ji} = -\eta \frac{\partial E}{\partial W_{ji}} = -\eta \left\{ \sum_{k=1}^l \frac{\partial E}{\partial O_k} \frac{\partial O_k}{\partial U_k} \frac{\partial U_k}{\partial H_j} \right\} \frac{\partial H_j}{\partial T_j} \frac{\partial T_j}{\partial W_{ji}}, \quad (\text{A19})$$

where

$$\frac{\partial E}{\partial O_k} = \frac{\partial}{\partial O_k} \left\{ \frac{1}{2} (t_k - O_k)^2 \right\} = -(t_k - O_k) \quad (\text{A20})$$

$$\begin{aligned} \frac{\partial O_k}{\partial U_k} &= \frac{\partial}{\partial U_k} \left\{ \frac{1}{1 + \exp(-U_k V_k)} \right\} = -\frac{-V_k \times \exp(-U_k V_k)}{(1 + \exp(-U_k V_k))^2} \\ &= \frac{V_k}{1 + \exp(-U_k V_k)} \left\{ 1 - \frac{1}{1 + \exp(-U_k V_k)} \right\} = O_k(1 - O_k) V_k \end{aligned} \quad (\text{A21})$$

$$\frac{\partial U_k}{\partial W_{kj}} = \frac{\partial}{\partial W_{kj}} \{H_1 W_{k1} + \dots + H_j W_{kj} + \dots + H_m W_{km}\} = H_j \quad (\text{A22})$$

$$\frac{\partial U_k}{\partial H_j} = \frac{\partial}{\partial H_j} \{H_1 W_{k1} + \dots + H_j W_{kj} + \dots + H_m W_{km}\} = W_{kj} \quad (\text{A23})$$

$$\begin{aligned} \frac{\partial H_j}{\partial T_j} &= \frac{\partial}{\partial T_j} \left\{ \frac{1}{1 + \exp(-T_j V_j)} \right\} = -\frac{-V_j \times \exp(-T_j V_j)}{(1 + \exp(-T_j V_j))^2} \\ &= \frac{V_j}{1 + \exp(-T_j V_j)} \left\{ 1 - \frac{1}{1 + \exp(-T_j V_j)} \right\} = H_j(1 - H_j) V_j \end{aligned} \quad (\text{A24})$$

$$\frac{\partial T_j}{\partial W_{ji}} = \frac{\partial}{\partial W_{ji}} \{X_1 W_{j1} + \dots + X_i W_{ji} + \dots + X_n W_{jn}\} = X_i. \quad (\text{A25})$$

Pursuantly,

$$\Delta W_{kj} = -\eta \{-(t_k - O_k) O_k (1 - O_k) V_k H_j\} \quad (\text{A26})$$

$$\Delta W_{ji} = -\eta \left\{ \sum_{k=1}^l -(t_k - O_k) O_k (1 - O_k) V_k W_{kj} \right\} H_j(1 - H_j) V_j X_i. \quad (\text{A27})$$

The network is optimized by ΔW and ΔV in iterations based on the neuron survival rule.

[38] **Acknowledgments.** This study was partly supported by the Virtual Laboratory for the Earth's Climate Diagnostics of the Center for Climate System Research (CCSR), the Hydropheric Atmospheric Research Center (HyARC), the Center for Environmental Remote Sensing (CEReS), and the Center for Atmospheric and Oceanic Studies (CAOS). The vicarious calibration of MTSAT-1R was done as part of a joint research study with the Meteorological Satellite Center/Japan Meteorological Agency (MSC/JMA) to develop vicarious calibration and global synthesis techniques for geostationary satellites. Algorithm development of the CAPCOM and the REAP was partly supported by ADEOS-II Global Imager science mission of the Japan Aerospace Exploration Agency (JAXA). The real-time received data from MTSAT-1R were provided by Weathernews Inc. (WNI), H. Ishibashi, and D. Abe. Preprocessing of the received satellite data and archive was conducted by M. K. Yamamoto, Y. Kato, and the CEReS database committee. The author appreciates the insightful comments and suggestions about learning algorithm by I. Matsuba (Faculty of Engineering, Chiba University). Special thanks are devoted to all persons involved in SKYNET observations. Also, there were many useful comments and suggestions by the reviewers.

References

- de Grujil, F. R. (1999), Skin cancer and solar UV radiation, *Eur. J. Cancer*, 35(14), 2003–2009.
- Dias, F. M., A. Antunes, J. Vieira, and A. M. Mota (2005), On-line training of neural networks: A sliding window approach for the Levenberg-Marquardt algorithm, in *Artificial Intelligence and Knowledge Engineering Applications: A Bioinspired Approach*, pp. 577–585, Springer, New York.
- Dorvlo, A. S. S., J. A. Jervase, and A. Al-Lawati (2002), Solar radiation estimation using artificial neural networks, *Appl. Energy*, 71, 307–319.
- Gasca, M., and T. Sauer (2000), Polynomial interpolation in several variables, *Adv. Comput. Math.*, 12, 377–410.
- Hagan, M. T., and M. Menhaj (1994), Training feedforward networks with Marquardt algorithm, *IEEE Trans. Neural Networks*, 5(6), 989–993.
- Hagan, M. T., H. B. Demuth, and M. Beal (1996), *Neural Network Design*, PWS Publ., Boston, Mass.
- Higurashi, A., and T. Nakajima (1999), Development of a two channel aerosol retrieval algorithm on global scale using NOAA/AVHRR, *J. Atmos. Sci.*, 56, 924–941.
- Higurashi, A., and T. Nakajima (2002), Detection of aerosol types over the East China Sea near Japan from four-channel satellite data, *Geophys. Res. Lett.*, 29(17), 1836, doi:10.1029/2002GL015357.
- Inoue, T. (1987), A cloud type classification with NOAA-7 split-window measurements, *J. Geophys. Res.*, 92, 3991–4000, doi:10.1029/JD092iD04p03991.
- Kawamoto, K., T. Nakajima, and T. Y. Nakajima (2001), A global determination of cloud microphysics with AVHRR remote sensing, *J. Clim.*, 14, 2054–2068.
- Key, J. R., and A. J. Schweiger (1998), Tools for atmospheric radiative transfer: STREAMER and FLUXNET, *Comput. Geosci.*, 24, 443–451.
- Lesins, G., P. Chylek, and U. Lohmann (2002), A study of internal and external mixing scenarios and its effect on aerosol optical properties and direct radiative forcing, *J. Geophys. Res.*, 107(D10), 4094, doi:10.1029/2001JD000973.
- Levenberg, K. (1944), A method for the solution of certain nonlinear problems in least squares, *Q. Appl. Math.*, 2, 164–168.
- Ling, D., H. Y. Hsu, G. Lin, and S. H. Lee (2005), Enhanced image-based coordinate measurement using a super-resolution method, *J. Rob. Comput. Integr. Manuf.*, 21(6), 579–588, doi:10.1016/j.rcim.2004.12.004.
- Lopez, G., M. A. Rubio, M. Martinez, and F. J. Battles (2001), Estimation of hourly global photosynthetically active radiation using artificial neural network models, *Agric. For. Meteorol.*, 107, 279–291.

- Manabe, S., and R. T. Wetherald (1967), Thermal equilibrium of the atmosphere with a given distribution of relative humidity, *J. Atmos. Sci.*, **24**, 241–259.
- Marquardt, D. W. (1963), An algorithm for least-squares estimation of non-linear parameters, *SIAM J. Appl. Math.*, **11**, 431–441.
- Nakajima, T. Y., and T. Nakajima (1995), Wide-area determination of cloud microphysical properties from NOAA AVHRR measurements for FIRE and ASTEX regions, *J. Atmos. Sci.*, **52**, 4043–4059.
- Nakajima, T., and M. Tanaka (1986), Matrix formulations for the transfer of solar radiation in a plane-parallel scattering atmosphere, *J. Quant. Spectrosc. Radiat. Transfer*, **35**, 13–21.
- Nakajima, T., and M. Tanaka (1988), Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation, *J. Quant. Spectrosc. Radiat. Transfer*, **40**, 51–69.
- Nakajima, T., T. Y. Nakajima, M. Nakajima, and the GLI Algorithm Integration Team (GAIT) (1999), Development of ADEOS-II/GLI operational algorithm for earth observation, *Proc. SPIE Int. Soc. Opt. Eng.*, **3870**, 314–322.
- Onogi, K., et al. (2007), The JRA-25 reanalysis, *J. Meteorol. Soc. Jpn.*, **85**(3), 369–432.
- Park, S. C., M. K. Park, and M. G. Kang (2003), Super-resolution image reconstruction: A technical overview, *IEEE Signal Proc. Mag.*, **20**, 21–36.
- Pinker, R. T., I. Laszlo, J. D. Tarpley, and K. Mitchel (2002), Geostationary satellite parameters for surface energy balance, *Adv. Space Res.*, **30**(11), 2427–2432.
- Pinker, R. T., et al. (2003), Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) Project, *J. Geophys. Res.*, **108**(D22), 8844, doi:10.1029/2002JD003301.
- Ranga Suri, N. N. R., D. Deodhare, and P. Nagabhushan (2002), Parallel Levenberg-Marquardt-based neural network training on linux clusters—A case study, paper presented at 3rd Indian Conference on Computer Vision, Graphics and Image Processing, IUPRAI, Chennai, India.
- Rumelhart, D. E., G. E. Hinton, and R. J. Williams (1986), Learning representations by backpropagating errors, *Nature*, **323**, 533–536.
- Sauer, T., and X. Yuan (1995), On multivariate Lagrange interpolation, *Math. Comput.*, **64**, 1147–1170.
- Schneider, S. H. (1972), Cloudiness as a global climate feedback mechanism: The effects on the radiation balance and surface temperature of variations in cloudiness, *J. Atmos. Sci.*, **29**, 1413–1422.
- Szilder, K., and E. P. Lozowski (1995), Cloud feedback examined using a two-component time-dependent climate model, *Beitr. Phys. Atmos.*, **68**, 43–57.
- Szilder, K., E. P. Lozowski, and G. W. Reuter (1998), A stochastic model of global atmospheric response to enhanced greenhouse warming with cloud feedback, *Atmos. Res.*, **47**–**48**, 475–489.
- Tsushima, Y., and S. Manabe (2001), Influence of cloud feedback on annual variation of global mean surface temperature, *J. Geophys. Res.*, **106**(D19), 22,635–22,646, doi:10.1029/2000JD000235.
- Twomey, S. (1977), The influence of pollution on the shortwave albedo of clouds, *J. Atmos. Sci.*, **34**, 1149–1152.
- Twomey, S., M. Piepgrass, and T. L. Wolfe (1984), An assessment of the impact of pollution on global cloud albedo, *Tellus, Ser. B*, **36**, 356–366.
- Wetherald, R. T., and S. Manabe (1988), Cloud feedback processes in a general circulation model, *J. Atmos. Sci.*, **45**, 1397–1415.
- Zarzalejo, L. F., L. Ramirez, and J. Polo (2005), Artificial intelligence techniques applied to hourly global irradiance estimation from satellite-derived cloud index, *Energy*, **30**, 1685–1697.
- Zhang, Y., W. B. Rossow, A. A. Lacis, V. Oinas, and M. I. Mishchenko (2004), Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data, *J. Geophys. Res.*, **109**, D19105, doi:10.1029/2003JD004457.
-
- A. Higuchi, T. Takamura, and H. Takenaka, Center for Environmental Remote Sensing, Chiba University, 1-33 Yayoi-cho, Inage-ku Chiba 263-8522, Japan. (takenaka@ceres.cr.chiba-u.ac.jp)
- A. Higurashi, Atmospheric Environment Division, National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, Ibaraki 305-8506, Japan.
- T. Nakajima, Research and Information Center, Tokai University, Tokyo 151 0063, Japan.
- T. Nakajima, Center for Earth Surface System Dynamics, University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa, Chiba 277-8564, Japan.
- R. T. Pinker, Department of Atmospheric and Oceanic Science, University of Maryland, College Park, MD 20742, USA.