

Advances in Neural Network Detection and Retrieval of Multilayer Clouds for CERES Using Multispectral Satellite Data

Patrick Minnis^{*1}, Sunny Sun-Mack¹, William L. Smith, Jr.², Gang Hong¹ and Yan Chen¹

¹SSAI, One Enterprise Parkway, Suite 200, Hampton, VA, USA 23666; ²NASA Langley Research Center, Hampton, VA, USA 23681

ABSTRACT

An artificial neural network (ANN) algorithm, employing several Aqua MODIS infrared channels, the retrieved total cloud visible optical depth, and vertical humidity profiles is trained to detect multilayer (ML) ice-over-water cloud systems as identified by matched CloudSat and CALIPSO (CC) data. The multilayer ANN, or MLANN, algorithm is also trained to retrieve the optical depth and the top and base heights of the upper-layer ice clouds in ML systems. The trained MLANN was applied to independent MODIS data resulting in a combined ML and single layer hit rate of 80% (77%) for nonpolar regions during the day (night). The results are more accurate than currently available methods and the previous version of the MLANN. Upper-layer cloud top and base heights are accurate to ± 1.2 km and ± 1.6 km, respectively, while the uncertainty in optical depth is ± 0.457 and ± 0.556 during day and night, respectively. Areas of further improvement and development are identified and will be addressed in future versions of the MLANN.

Keywords: CERES, CloudSat, CALIPSO, MODIS, Neural Network, Multilayer clouds

1. INTRODUCTION

Clouds are critical to the atmospheric energy system, particularly the radiative balance throughout the troposphere. The vertical distribution of cloud particles and phase determines the heating rates of atmospheric layers, the outgoing radiant flux at the top of atmosphere, and the radiation balance at the surface [1,2,3,4]. Satellite remote sensing with passive imagers is currently the only approach suitable for nearly continuous monitoring of clouds day and night around the globe. Passive satellite retrievals of cloud properties typically rely on interpreting observed radiances as emanating from a single, plane-parallel layer. Yet, a significant percentage of cloud systems comprise multiple layers of cloud formations, frequently with ice clouds overlaying liquid water clouds. For multilayered (ML) cloud systems, the retrieved parameters often have significant errors and give an inaccurate characterization of the cloud vertical structure. They have been identified as the largest source of cloud phase misidentification and cloud-top height errors in at least one passive retrieval algorithm using the Aqua Moderate-resolution Imaging Spectroradiometer (MODIS) [5].

Active satellite systems such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite [6] and the Cloud Penetrating Radar (CPR) on CloudSat can provide much more accurate depictions of the cloud vertical structure, but they are near-nadir viewing instruments that only produce profiles in a narrow curtain along the orbital track. To cover broad areas and make the retrievals more useful, it is necessary to employ passive satellite imagers that cover a broad swath or a large fraction of a hemisphere. A number of multispectral techniques have been developed to detect ML systems with imager radiances and have had varying degrees of success (e.g., [7,8,9]). Other methods have attempted to identify ML clouds and retrieve the properties of the bottom and top layers using multiple instruments (e.g., [10-11]) or multiple channels on the same instrument (e.g., [12,13,14]).

The Clouds and Earth's Radiant Energy System (CERES) Project Edition 4 processing system included, as a supplement, a CO₂-slicing based ML detection and retrieval algorithm as part of its comprehensive cloud identification and retrieval analysis package [15], but it has proved to be too limited in its applicability. To achieve the CERES goal of accurately characterizing the radiation budget at the top of the atmosphere, at the surface, and within the atmosphere, it is important to reliably identify ML clouds and retrieve their properties. To that end, Sun-Mack et al. [16] began the development of an artificial neural network (ANN) to discriminate between single-layer (SL) and ML clouds using April 2009 MODIS radiance data matched to CALIPSO and CloudSat vertical profiles of clouds. They were able to

*p.minnis@nasa.gov phone 1 757 880-9259 fax 1 757 951-1902

successfully identify ice-over-water ML clouds and SL clouds 75% and 72% of the time for day and night, respectively. While this hit rate is as good or better than previously reported values for other algorithms, the ML artificial neural network (MLANN) only detected 43% of the ML clouds during the day and 46% at night. In this paper, the MLANN is further developed to include more input parameters and additional output variables, as well as using only high-confidence CloudSat and CALIPSO data, to provide improved detection and to begin the retrieval of the ML system components.

2. NEURAL NETWORK

Neural networks are finding increased value in remote sensing of clouds. They have been used to determine cirrus optical depth and height [17], thick ice cloud optical depth at night [18], and cloud top pressure and altitude [19]. Figure 1 provides a schematic of the MLANN to estimate several parameters and decide whether a pixel contains a SL or ML cloud. The input neurons comprise a layer consisting of a set of input variables, x_i , that are each linked to each member of a hidden layer consisting of the function, $u_j = f_1(g_j)$, where g_j is the sum of the weights w_{ij} applied to the input variable x_i plus the constant b_j , as indicated in the figure. Likewise, the output layer is uniquely linked to the input layer through the hidden layer by the weights w_j and constant c to the input neurons. The number of j neurons, N_j , is selected to optimize the accuracy of the output. Similarly, the number of input variables or i neurons is also adjusted to minimize the estimate uncertainty. In this study, 50 -70 neurons are used for the hidden layer.

Levenberg-Marquadt optimization is used in the MLANN training. Calculation of the MLANN weights, or training of the MLANN, consists of three components: training, testing, and validation. Thus, for a given dataset, a major portion of the data is used as training data to make estimates of the weights, and smaller portions of the data are used for the testing and validation of the MLANN. The testing and training are interwoven such that the testing results force the fitting or training process to stop or adjust weights. When the training-testing is satisfied, the resulting weights are applied to the validation data to determine if the same error statistics are within the maximum stipulated error. Validation vectors are used to stop training early if the MLANN performance fails to improve. The formulation of this approach is given in more detail by [16] and [18] and the references found therein.

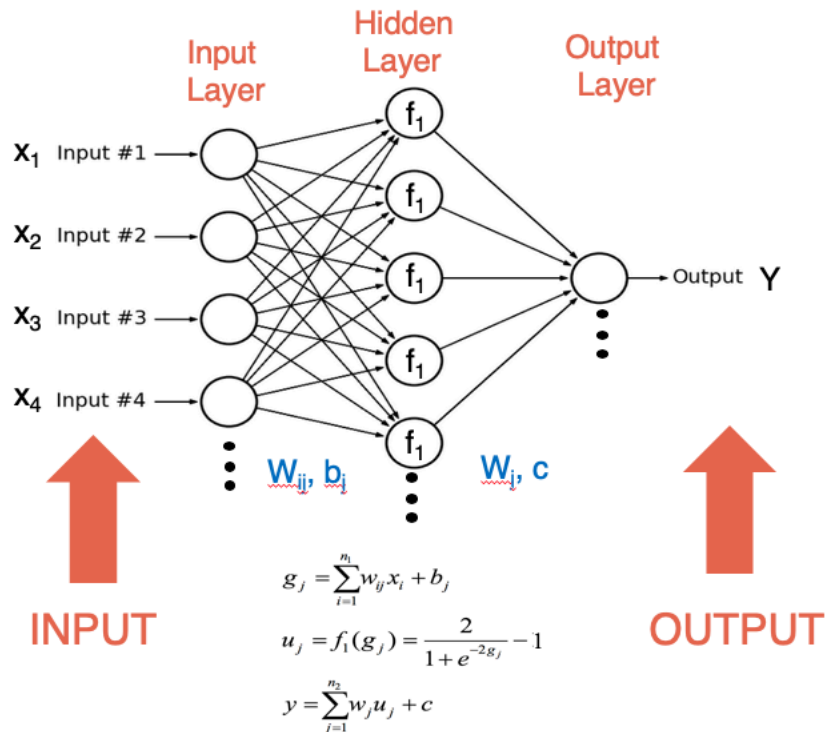


Figure 1. Schematic of artificial neural network used here.

3. DATA AND METHODOLOGY

The data for training the MLANN consist of passive imager data, numerical weather model analyses, and active sensor data. Data from July 2008 and April, July, and September 2009 are employed for training and validation as indicated in the appropriate section. Only data taken between 60°N and 60°S over snow-free surfaces are analyzed here. Multilayered clouds are defined in this study as any combination of ice-cloud layers over the top of one or more water cloud layers with the constraint that the top of the water layer must be at least 1 km below the bottom of the lowest ice cloud layer. All ice cloud layers together are considered to constitute only one cloud layer. Similarly, all liquid-phase layers are considered as a single layer. To examine sensitivity to this ML definition, a separation minimum of 3 km is also used.

3.1 Input Data

Aqua MODerate-resolution Imaging Spectroradiometer (MODIS) Collection 6.1 1-km brightness temperatures (BT) and cloud-top phase and cloud optical depths τ_M are retrieved using the Clouds and Earth's Radiant Energy System (CERES) Edition 4 algorithms [15] and are referred to as CERES-MODIS or CM products. Reanalyses from version 5.4 of the Global Modeling Assimilation Office Global Earth Observing System GEOS-5.4, an update of the versions described by [20], provide surface skin temperature and vertical profiles of relative humidity.

3.2 Active Sensor Data

CloudSat, CALIPSO, and Aqua are part of the A-Train satellite constellation and take measurements continuously with time differences of less than three minutes. CALIOP and the CloudSat radar are near-nadir viewing and aligned so that they view roughly the same area along their respective flight tracks. Their flight tracks are typically near the Aqua nadir path, so that MODIS views the same scene at viewing zenith angles less than 18°. MLANN is trained and validated using CALIPSO Version 4 [21] and CloudSat R04 [22] vertical profiles of clouds matched with 1-km Aqua MODIS Collection 6.1 radiances in an update of the CloudSat, CALIPSO, CERES, and MODIS (C3M) product of [23]. The C3M first converts the CloudSat CLDCLASS high-confidence cloud profiles from 240-m to 60-m vertical resolution, then merges them with the CALIPSO cloud profile (CPRO) and vertical feature mask (VFM) to produce a complete vertical profile of cloud-filled layers. All CALIPSO averaging resolutions, 0.33 – 80 km, are used to define the cloud profiles. The CALIPSO-CloudSat (CC) data are merged with the corresponding Aqua MODIS radiances and CERES Edition 4 cloud property retrievals, which include τ_M . In addition to defining the cloud layers, the CALIPSO-CloudSat optical depth for the ice layer τ_{CC} was computed for each pixel identified as ML by the CALIPSO-CloudSat profiles to perform the additional analyses. The value of τ_{CC} is equal to the CALIPSO ice cloud optical depth, when the CALIOP signal shows a return from the lower layer cloud, otherwise it is equal to the combined CALIPSO-CloudSat optical depth. Similarly, the top and base heights of the ice cloud layer, Z_{TCC} and Z_{BCC} , respectively, are also determined for ML cloud pixels.

3.3 Input Layer

The input parameters from MODIS include latitude, longitude, brightness temperatures $BT(\lambda)$ and brightness temperature differences $BTD(\lambda_1-\lambda_2) = BT(\lambda_1) - BT(\lambda_2)$, where λ is the wavelength in μm , and the visible-channel optical depth τ_M . Specifically, BT values at 3.7, 6.7, 8.5, 11, and 12 μm are used here along with $BTD(3.7-11)$, $BTD(6.7-11)$, $BTD(8.5-11)$, and $BTD(11-12)$. The GEOS input data consist of the surface temperature and relative humidity at the surface and at 850, 700, 500, 400, 300, 200, and 100 hPa. During the day, the solar zenith angle (SZA) in degrees is also used in the input. Inclusion of the GEOS humidity profiles represents a marked change in the input from that of [16].

3.4 Output Layer: Multilayer Identification

The MLANN as formulated here has four training sets for each parameter of interest: night ($SZA > 82^\circ$) and day ($SZA \leq 82^\circ$) with CM cloud-top phase of ice or water, so that each set yields the weights, w_{ij} and w_j , along with the constants b and c . Training is performed using the minimum cloud separation threshold $\Delta Z_{\min} = 1$ and 3 km, where the separation ΔZ is the difference in altitude between the bottom of the lowest ice layer and the top of the highest water layer in the CC profile. An ice-cloud layer is assumed to be present in the profile, if at least one layer occurs at a height above the altitude corresponding to 253 K and no temperature inversion exists in atmospheric layer between the altitudes corresponding to 273 K and 253 K. All C3M data having an ice cloud layer by that definition are used as output in the MLANN training. A given profile is assumed to contain a ML cloud system if a cloudy layer exists at least ΔZ_{\min} below the lowest ice cloud layer and the base of that cloud layer is below the 253-K level. Some ice-over-water layer systems

will be considered as SL if they are closer than ΔZ_{\min} . ML clouds defined by $\Delta Z_{\min} = 1$ km are classified as SL clouds when $\Delta Z_{\min} = 3$ km. The output consists of a probability. If the probability < 0.5 , the classification is SL, otherwise it is ML. For this study, all qualifying data from April 2009 are used in the training process with 60% for training and 20% each for testing and validation.

3.5 Output Layer: Other Parameter Estimation

To estimate the upper cloud optical depth τ_{ul} , the MLANN is trained with the same input variables in the same sets, but the output variable is τ_{cc} . Values for the weights are also computed for upper layer cloud top and cloud base heights, Z_{ul} and Z_{bul} , respectively, using the top of the highest and bottom of the lowest ice cloud in the CC ML profile. Qualifying July and September 2009 data are used for training these three parameter networks using the same fractions of the data for training, testing, and validating as employed for the ML detection.

3.6 Neural Network Retrievals

Once all of the weights are computed in the training process, they can then be used in the MLANN to produce output values for other input datasets having the same variables. When using the MLANN, the CERES Edition 4 cloud phase and SZA are used to select which set of weights will be used in the MLANN to determine the layering classification.

4. RESULTS AND DISCUSSION

The results presented here consist of comparisons of the MLANN and corresponding CC parameters for the training dataset along with data from a different period to ensure robustness of the estimate. Weights and constants were determined by training for each category and parameter and then applied to the independent datasets: July 2009 for ML identification, April 2009 for optical depth, and July 2008 for cloud top and base heights. Not all results are reported here.

4.1 Multilayer Clouds

Figure 2 plots the CC cloud profiles determined over a midlatitude area at night around 07 UTC, 2 July 2009. The CC layering classification, gray for SL and orange for ML, is based on $\Delta Z_{\min} = 3$ km. The corresponding MLANN ML detection is shown in orange along the line at 15.4 km. At first glance, it appears that the MLANN detects the ML clouds quite well with gaps mostly corresponding to the gray areas. For the SL areas on the right side of the plot, however, the gaps are noticeably larger than the gray segments. If the CC ML criterion were $\Delta Z_{\min} = 1$ km, the gray area indicated by the oval would have been included in the CC ML category.

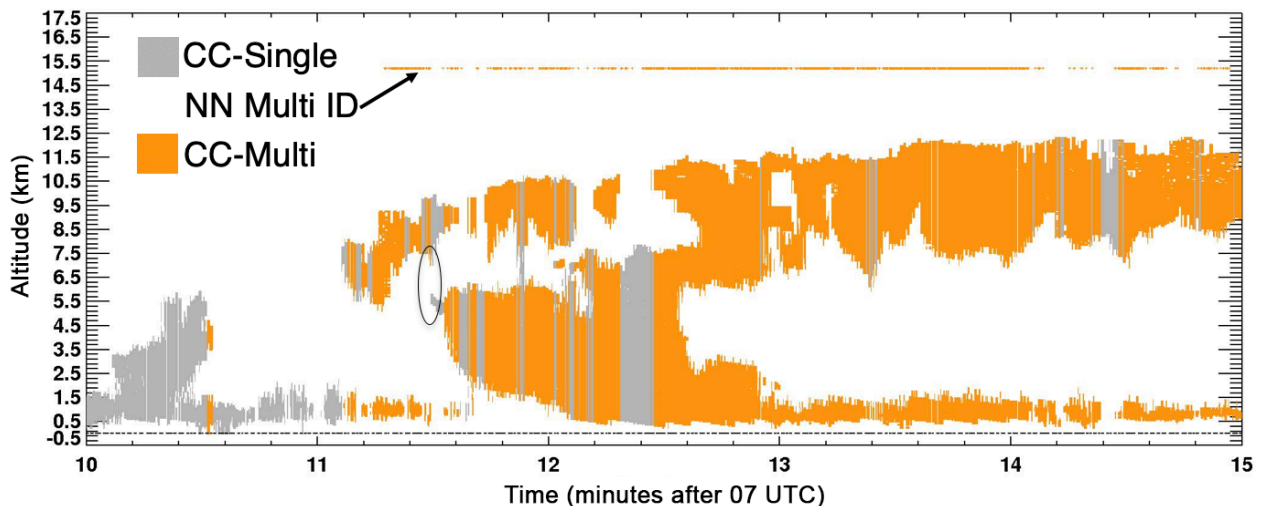


Figure 2. CALIPSO-CloudSat cloud profile from C3M for 2 July 2008 with CC ML clouds having $\Delta Z_{\min} = 3$ km indicated in orange and CC SL denoted in gray. The MLANN ML identification for each profile is indicated as an orange dot at 15.4 km. CC profiles that would be considered ML, if $\Delta Z_{\min} = 1$ km, are highlighted with the oval.

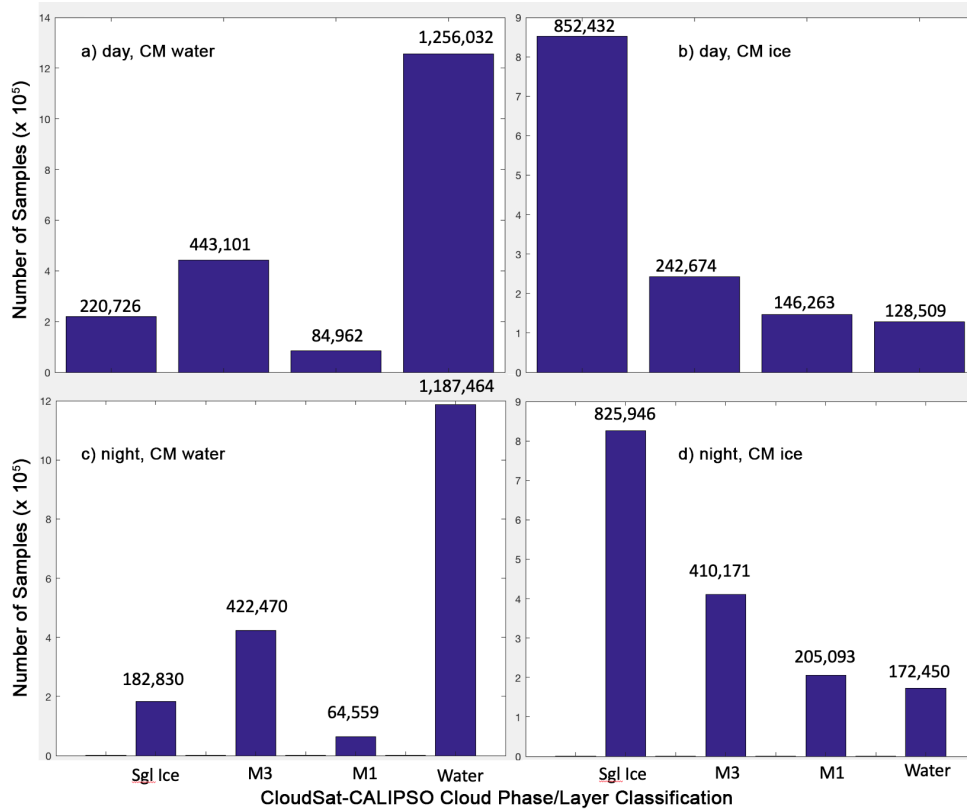


Figure 3. Layering classifications for a CloudSat-CALIPSO data matched CERES Aqua MODIS data for April 2009. Distributions of the various layering categories are shown for the different CERES-MODIS classifications. (a) Daytime, CM liquid, (b) Daytime, CM ice, (c) Night, CM liquid, and (d) Night, CM ice. M3: $\Delta Z_{\min} = 3$ km, M1: $\Delta Z_{\min} = 1$ km

In general, the use of $\Delta Z_{\min} = 3$ km captures a significant portion of the ice-over-water ML clouds as seen in Figure 3, which shows histograms of the CC cloud classifications for April 2009 for each of the four CM categories. During the day for CM clouds categorized as liquid (Figure 3a), $\Delta Z_{\min} = 3$ km accounts for 84% of the ML clouds, assuming that clouds with $\Delta Z < 1$ km are SL. At night (Figure 3c), that fraction rises to 87%. The contribution of clouds having $\Delta Z > 3$ km drops to 62 and 67% during the day (Figure 3b) and night (Figure 3d), respectively, when CM identifies the ML clouds as ice. Ideally, the goal of the MLANN should be to detect the most ML clouds as possible while minimizing false ML cloud identification.

Statistics within 3×3 confusion matrices are employed to help assess the MLANN SL-ML discrimination accuracy and reliability. Following the nomenclature of [9], the first column of each matrix lists the correct SL percentage, ST, with the false positive ML, MF, below. The sum of ST + MF, S the total percentage of true SL clouds, is at bottom of the column. The false positive SL, SF, is shown at the top of the middle column, the total percentage of positive ML clouds, MT, sits at the center of the matrix. At the bottom of the column resides the total CC ML fraction M, the sum, SF + MT. The estimated SL and ML fractions, ES and EM, respectively, occupy the top and middle cells of the third column. Normally, the lower right corner would contain the total fraction, but since it is always 100% in this study, the hit rate HR, or total fraction correct, ST + MT, is placed there. The real risk or chance of a misclassification is $RR = SF + MF$. The confidence in the estimates of SL and ML clouds are $C_S = ST/ES$ and $C_M = MT/EM$, respectively. The number of correctly identified ML pixels is $N_M = MT \cdot N/100$, where N is the total number of pixels comprising the matrix.

The MLANN was trained using the CM data for each of the four categories in Figure 3 to obtain a four sets of weights and constants for the two separation distances. Results of applying those coefficients to July 2009 data are summarized in the confusion matrices in Table 1 for $\Delta Z_{\min} = 1$ and 3 km on the left and right halves of the table, respectively. During the day, the 1-km hit rate HR, or percentage correct, is 75.5% for CM ice clouds, with the fraction of ML correctly identified being 0.42 and $N_M \sim 171,000$. The layering discrimination during the day is a bit better for CM liquid clouds:

Table 1. Confusion matrices (each bounded by dashed lines) for MLANN applied to Aqua MODIS relative to layer identification from CloudSat-CALIPSO, July 2009. The bold numbers indicate the percent correct for each matrix.

CloudSat and CALIPSO												
MLANN	$\Delta Z_{\min} \geq 1 \text{ km, Day}$			$\Delta Z_{\min} \geq 1 \text{ km, Night}$			$\Delta Z_{\min} \geq 3 \text{ km, Day}$			$\Delta Z_{\min} \geq 3 \text{ km, Night}$		
	<u>SL</u>	<u>ML</u>	<u>Total</u>	<u>SL</u>	<u>ML</u>	<u>Total</u>	<u>SL</u>	<u>ML</u>	<u>Total</u>	<u>SL</u>	<u>ML</u>	<u>Total</u>
Ice SL,%	62.0	18.8	80.8	49.4	17.2	66.6	77.3	11.9	89.1	62.6	14.7	77.3
Ice ML,%	5.7	13.5	19.2	10.3	23.2	33.4	3.3	7.6	10.9	7.2	15.5	22.7
Total, %	67.7	32.3	75.5	59.7	40.3	72.5	80.5	19.5	84.9	69.8	30.2	78.1
# pixels x 10 ³	859	410	1,269	1,072	724	1,796	1,022	247	1,269	1,254	542	1,796
Liquid SL, %	69.8	11.3	81.0	68.7	12.8	81.5	74.2	11.3	85.5	72.9	12.9	85.8
Liquid ML, %	5.5	13.5	19.0	6.1	12.4	18.5	4.9	9.6	14.5	4.8	9.4	14.2
Total, %	75.2	24.8	83.3	74.8	25.2	81.1	79.1	20.9	83.8	77.7	24.3	82.3
# pixels x 10 ³	1,643	540	2,183	1,624	549	2,173	1726	457	2183	1689	484	2173

HR = 83.3% and $N_M \sim 295,000$. For $\Delta Z_{\min} = 3 \text{ km}$ during the day, HR is better overall than its 1-km counterparts and the real risk is 15.2% for ice clouds compared to 24.5% for the 1-km case. However, MT is considerably reduced from its 1-km value and many more ML clouds with separation distances of 1 – 3 km are now identified as SL clouds. RR is nearly the same, ~16%, for both 1 and 3-km liquid clouds during the day. For the daytime 1-km data, CM ~71% compared to ~67% for 3-km data. At night, RR increases noticeably for ice clouds to 27.5% and 21.9% for 1 and 3-km data, respectively, while the hit rates drop from their daytime values, even though MT rises dramatically. For liquid clouds, MT drops slightly along with HR. Except for daytime 3-km data, the hit rate for ice clouds is significantly smaller for ice clouds than for liquid clouds. This may be due to the overall larger optical depths of the upper layer clouds for systems identified as CM ice clouds [5]. The larger ice cloud optical depths may muddle the signals distinguishing between pure ice clouds and ML systems. This is especially true during the day.

Overall, the CM ice and water clouds must be considered together, so the results from the two phases were combined for the 1-km data and are listed in Table 2. The 3-km data are not considered anymore since it is clear from Table 1 that maximizing the ML detection is best conducted using the smaller separation distance. The 1-km MLANN detects 1.15 million ML pixels compared to 0.79 million ML pixels by the 3-km MLANN, while suffering minimally in increased false detection. The overall 1-km HRs are 80.4% and 77.1% for day and night, respectively, while the corresponding RRs are 19.6% and 22.8%. During the day, CS and CM are 82.6% and 70.6% compared to 84.7% and 68.5% at night.

These results represent a significant advance from the initial MLANN formulation [16], which did not consider the CM phase categories separately and did not include atmospheric humidity profiles. The changes increase the gap between the various metrics determined from the MLANN and those from other methods reported by [16]. As it is applicable both day and night, it should lead to a more accurate rendering of cloud vertical structure from satellite imagers.

Table 2. Same as Table 1, but includes combined liquid and icer results for upper and lower layer separation $\geq 1 \text{ km}$ only.

CloudSat and CALIPSO						
MLANN	<u>Day</u>			<u>Night</u>		
	<u>SL</u>	<u>ML</u>	<u>Total</u>	<u>SL</u>	<u>ML</u>	<u>Total</u>
SL,%	66.9	14.0	80.9	59.7	14.8	70.5
ML,%	5.6	13.5	19.1	8.0	17.4	25.4
Total, %	72.5	27.5	80.4	59.7	40.3	77.1
# pixels x 10 ³	2503	949	3,452	2,622	1,248	3,870

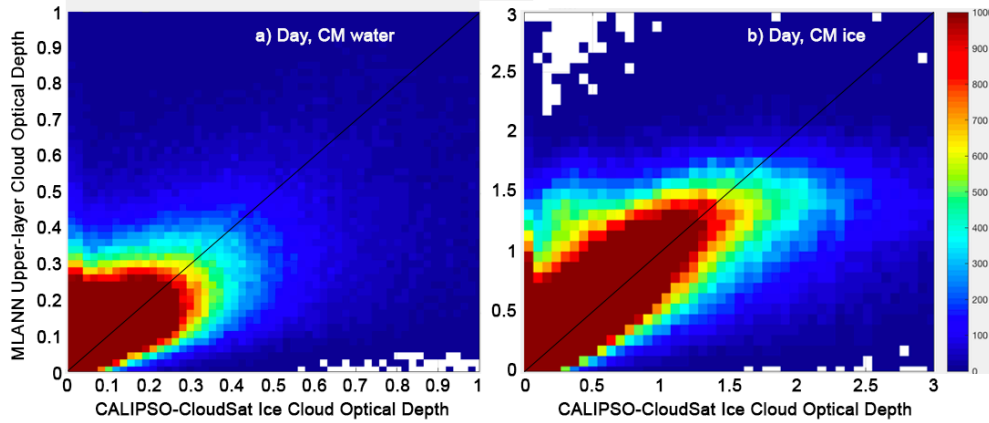


Figure 5. Scatterplot of MLANN upper layer optical depth versus CC optical depth for the daytime training, July and September 2009.

4.2 Upper-layer cloud optical depth

Figure 5 shows the scatterplots of MLANN τ_{ul} versus the CALIPSO-CloudSat optical depths for the daytime training. When the CM pixel is classified as a water cloud (Figure 5a), the resulting values show some correlation but the origin of any linear fit would be near (0,0.1) and the slope would be fairly horizontal. That holds true for the CM ice clouds (Figure 5b), but there appears to be better correlation. There is a significant difference between the range of optical depths for the two CM phases. For CM water, τ_{CC} is mainly less than 0.4, while for CM ice the significant range extends to $\tau_{CC} = 2$. As noted above, this discrepancy is due to the fact that as the upper-layer cloud optical depth increases, the CM algorithm increasingly selects ice as the cloud phase [5].

The distributions of the daytime optical depths and their differences are plotted in Figure 6. It is clear in Figures 6a and 6c that the MLANN is unable to capture the frequency of pixels having very small optical depths, $\tau_{CC} < 0.06$, for both

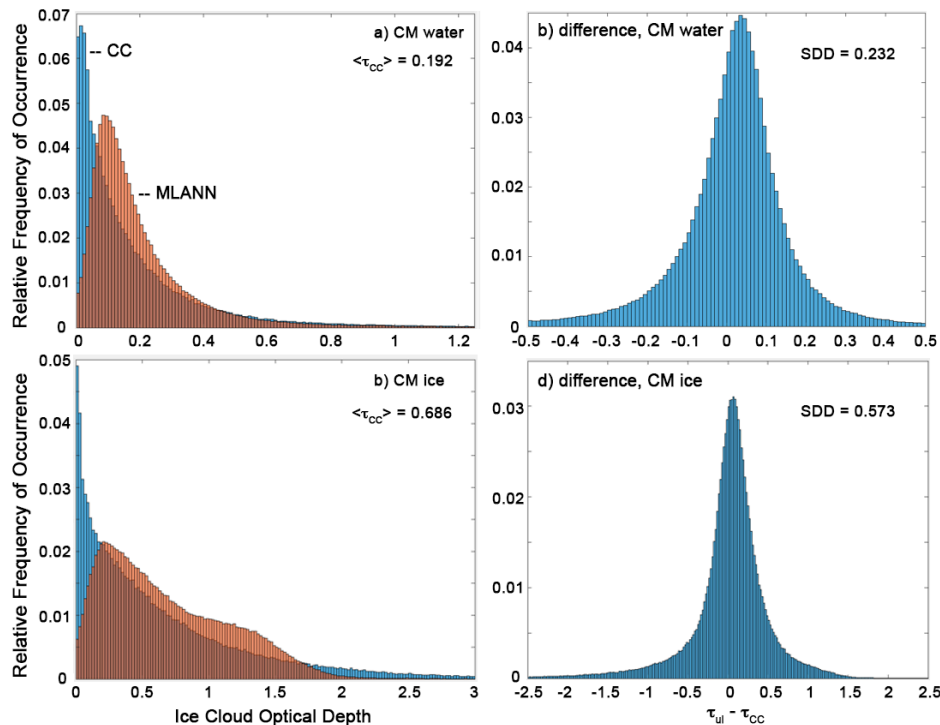


Figure 6. Ice cloud optical depth and difference probability distributions for upper layer cloud of ML ice-over-water systems from CERES-MODIS (τ_{ul}) and CALIPSO CloudSat (τ_{CC}), July and September 2009.

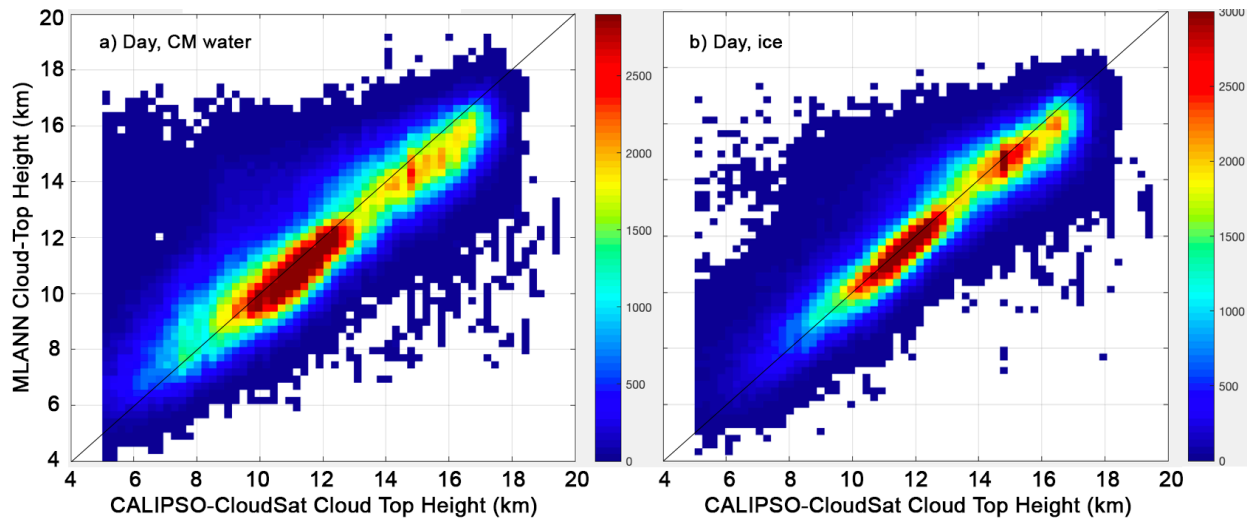


Figure 7. Daytime cloud-top height scatter plots for matched CM-MLANN and CC upper-layer clouds in ML cloud systems, July and September 2009, for CM (a) water and (b) ice clouds.

ML water and ice clouds. This inability to capture extreme values in cloud optical depth seems to be typical as it also appears in the results of [15] and [17]. Although the average CC optical depth, $\langle \tau_{CC} \rangle$, is the same as its τ_{ul} counterpart, the standard deviation of the differences, SDD, is roughly equivalent to 90% of the overall mean. The median differences for CM water and ice in Figures 6b and 6d are ~ 0.04 and 0.01 , respectively, while the difference histograms are highly skewed to negative values. The nighttime results are similar, but $\langle \tau_{CC} \rangle$ is less than its daytime counterpart.

4.3 Upper-layer Cloud Top Height

Daytime cloud-top heights determined by the MLANN for the training data are compared to their CC counterparts in Figure 7. The results are well correlated with the CC values for both the CM liquid (Figure 7a) and ice (Figure 7b) clouds. More extreme outliers are evident in the water cloud plot, which apparently balance a cluster of points that is below the 1:1 line at the high end of the range. The ice results are better behaved. Figure 8 shows the histograms of the heights and their differences from Figure 7. Again, the extrema are not well represented in the probability distributions in Figures 8a and 8c, while the number of points in the middle of the distributions are overestimated. The differences for daytime CM water clouds (Figure 8b) are skewed to positive values although the median difference is roughly -0.3 km. For CM ice clouds (Figure 8d) the distribution is less skewed and the median value is close to -0.1 km. The mean differences are 0.00 km and the SDDs are 1.38 and 1.03 km, respectively, for CM liquid and ice clouds. At night, the corresponding SDDs are 1.21 and 1.18 km.

The results here far exceed the accuracy of the single-layer cloud-top height retrievals of most standard algorithms (e.g., [5]). They are also comparable to those from a different neural network approach [19] and may be somewhat larger than those from a dedicated cirrus cloud top height analysis [17]. The exact differences, however, are difficult to determine because of sampling and data selection differences. Nevertheless, it is clear that the skill in determining Z_{Tul} is sufficient to confidently begin the process of retrieving the properties of the upper and lower layers in ML conditions.

4.4 Upper-layer Cloud Base Height

The MLANN was also trained to estimate the upper-layer cloud base height Z_{Bul} using the same sets of input parameters. The preliminary results for nighttime are shown as scatterplots in Figure 9 for July 2009. Concentrations of the points along the 1:1 line are not as dense as those seen in Figure 7, but the correlation appears to be as good with most of points scattered about the agreement line. The larger SDD values, 1.6 km for CM water (Figure 9a) and 1.7 km for CM ice (Figure 9b), are not surprising, given that the cloud base might correspond to a single layer of ice or multiple ice layers with clear air between them. Nevertheless, the SDDs are smaller than those found for SL cirrus clouds using more conventional techniques [5]. Results for the daytime analyses are similar.

With the upper-layer cloud top and base estimates, it should be possible define the upper-layer cloud boundaries for scenes identified as ML cloud systems. An example of applying the MLANN to characterize the cloud vertical extent is

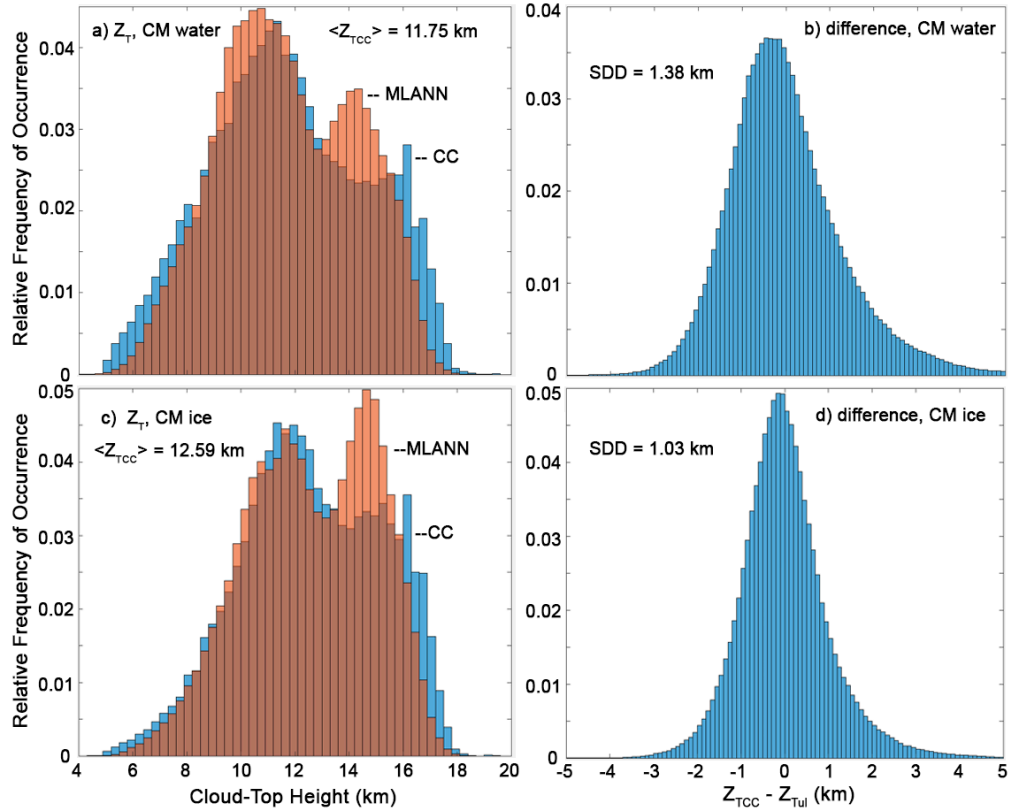


Figure 8. Upper-layer cloud top height (Z_T) and difference probability distributions for for ML ice-over-water systems from CERES-MODIS and CALIPSO-CloudSat (Z_{TCC}), July and September 2009.

shown in Figure 10 for a cloud profile taken by CALIPSO-CloudSat over a midlatitude ocean at night, centered at 11:28 UTC, 8 April 2009. The top panel (Figure 10a) shows the cloud-top height retrieved using the standard SL assumption overlaid in blue over the CC profiles. It is clear that few of the CM-retrieved Z_T values coincide with Z_{TCC} for any of the ML clouds. Where the upper cloud is very thin, the retrieved top is close to that of the lower cloud, while in many instances, it falls between the layers and occasionally exceeds Z_{TCC} . The lower panel Figure (10b) shows the MLANN values of Z_{Tul} (black) and Z_{Bul} (blue) overlaid on the same CC cloud profile. For the ML clouds, Z_{Tul} tracks Z_{TCC} quite well. Some exceptions are seen at minutes 26.7 and 29.9, where the two top heights diverge by 1-2 km. Overall, the

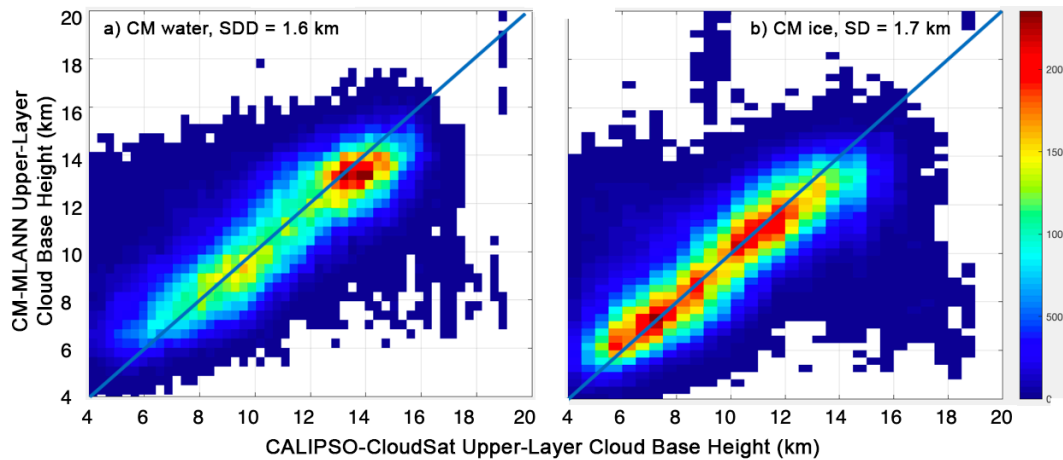


Figure 9. Scatterplots of nocturnal upper-layer cloud base heights from CALIPSO-CloudSat and matched MLANN Aqua MODIS retrievals, July 2009.

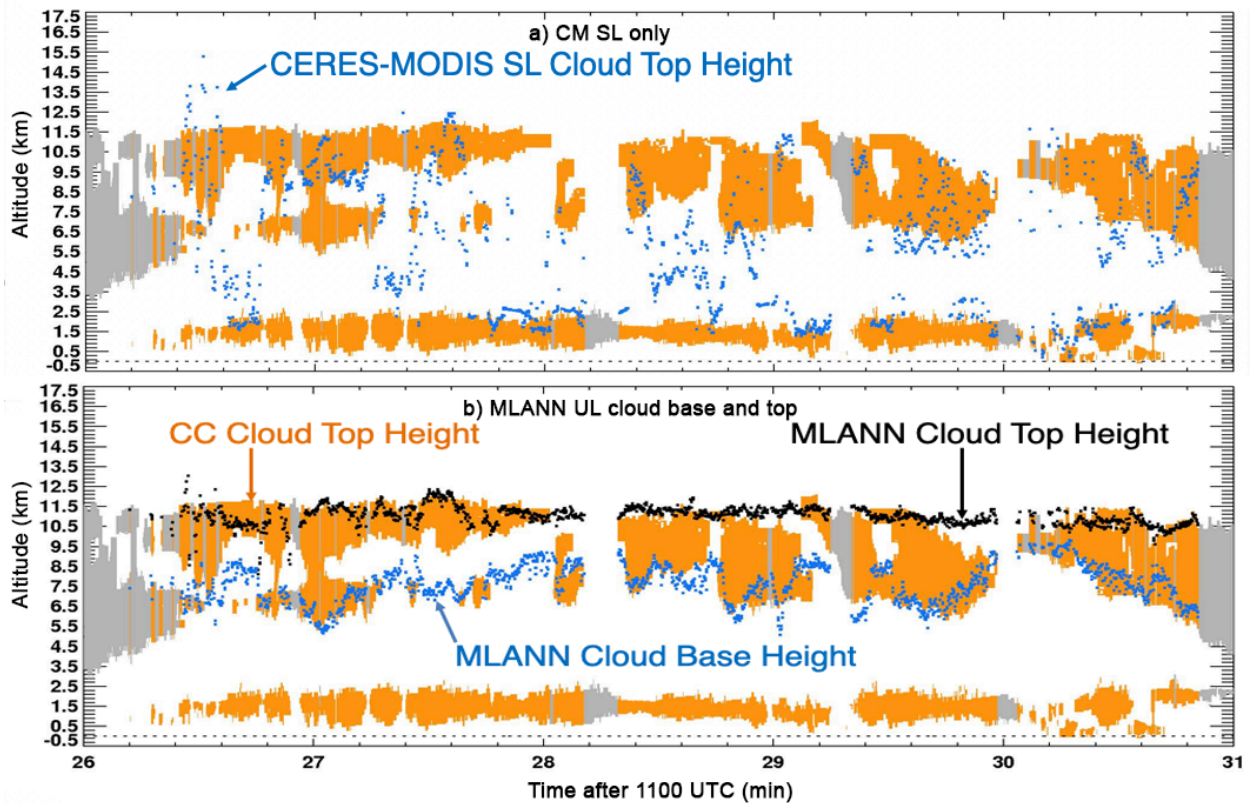


Figure 10. Comparison of nocturnal cloud height parameters determined from CALIPSO-CloudSat data and from Aqua MODIS over midlatitude ocean area using the (a) CERES-MODIS SL approximation and (b) the MLANN coefficients determined here for Z_{Tul} and Z_{Bul} , 8 April 2009.

discrepancies are nothing like those seen in the top panel. Cloud base heights from the MLANN also coincide remarkably well with their CC counterparts. The greatest deviations occur around minutes 26.7, 27.8, and 29.2. Nevertheless, the base heights are better behaved than Z_T from the CM retrievals.

5. CONCLUDING REMARKS

Development of a multilayer cloud detection and retrieval system, the MLANN, that depends heavily on artificial neural networks has continued with the addition of new input parameters and output variables that constitute a state-of-the-art ML detection method using channels that are common to many current satellite imaging systems. Yet, the method still only detects slightly more than 50% of the ML clouds as defined. Thus, more analyses of the ML clouds that are missed, as well as the false detections, are still warranted to make further advances in reliability. Additional months of data should be trained and analyzed, while all regions, including those with snow or ice cover, should be incorporated into the training. The viewing zenith angle dependence of the retrieval should be explored by applying the trained MLANN to full swath MODIS data and by training the MLANN with other datasets aligned with the CC overpasses. This paper has established a means to define the upper cloud layer, but retrieving the properties of the lower cloud remains for future studies. These could use the CC results to perform additional neural network training or use a physical retrieval employing ML radiative transfer lookup tables to determine the lower-layer properties. These and other analyses will be attempted in future research to build the MLANN into a reliable multilayer cloud retrieval system.

Acknowledgments. This research is supported by the NASA CERES Project and the NASA Modeling, Analysis, and Prediction Program.

REFERENCES

- [1] Chen T., and Zhang Y.C., “Sensitivity of atmospheric radiative heating rate profiles to variations of cloud layer overlap,” *J. Climate* 13, 2941–2959 (2000).
- [2] Morcrette J. J., and Christian J., “The response of the ECMWF model to changes in the cloud overlap assumption,” *Mon. Wea. Rev.* 128, 1707–1732 (2000).
- [3] Li, J., Yi, Y., Minnis, P., Huang, J., Yan, H., Ma, Y., Wang, W., and Ayers, J. K., “Radiative effect differences between multi-layered and single-layer clouds derived from CERES, CALIPSO, and CloudSat data,” *J. Quant. Spectrosc. Radiat. Transfer* 112, 361-375 (2011).
- [4] Kato, S., Rose, F. G., Ham, S.-H., Rutan, D. A., Radkevich, A., Caldwell, T., Sun-Mack, S., Miller, W. F., and Chen, Y., “Radiative heating rates computed with clouds derived from satellite-based passive and active sensors and their effects on generation of available potential energy,” *J. Geophys. Res.* 124, 1720-1740, doi:10.1029/2018JD028878 (2019).
- [5] Yost, C., Minnis, P., Sun-Mack, S., Chen, Y., and Smith, W. L., Jr., “CERES MODIS cloud product retrievals for Edition 4, Part II: Comparisons to CloudSat and CALIPSO,” *IEEE Trans. Geosci. Remote Sens.*, submitted, (2019).
- [6] M. Winker, M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W., and Young, S. A. , “Overview of the CALIPSO mission and CALIOP data processing algorithms,” *J. Atmos. Oceanic Tech.* 26, 2310-2323, doi:10.1175/2009JTECHA1281.1, (2009).
- [7] Pavolonis, M. J., and Heidinger, A. K., “Daytime cloud overlap detection from AVHRR and VIIRS,” *J. Appl. Meteor.* 43, 762–778, doi:10.1175/2099.1 (2004).
- [8] Wind G., Platnick, S., King, M. D., Hubanks, P. A., Pavolonis, M. J., Heidinger, A. K., Yang, P., and Baum, B. A., “Multilayer cloud detection with the MODIS near-infrared water vapor absorption band,” *J. Appl. Meteor. Climatol.* 49, 2315–2333, doi:10.1175/2010JAMC2364.1 (2010).
- [9] Desmons, M., Ferlay, N., Riedl, J., and Theuleux, F., “A global multilayer cloud identification with POLDER/Parasol,” *J. Appl. Meteor. Climatol.* 56, 1121-1139 (2017).
- [10] Lin, B., P. Minnis, P., Wielicki, B. A., Doelling, D. R., Palikonda, R., Young, D. F., and Uttal, T. “Estimation of water cloud properties from satellite microwave and optical measurements in oceanic environments. II: Results,” *J. Geophys. Res.* 103, 3887-3905, (1998).
- [11] Minnis, P., Huang, J., Lin, B., Yi, Y., Arduini, R. F., Fan, T.-F., Ayers, J. K., and Mace, G. G., “Ice cloud properties in ice-over-water cloud systems using TRMM VIRS and TMI data,” *J. Geophys. Res.* 112, D06206, doi:10.1029/2006JD007626 (2007).
- [12] Chang, F.-L., and Li, Z., “A new method for detection of cirrus overlapping water clouds and determination of their optical properties,” *J. Atmos. Sci.* 62, 3993–4009, doi:10.1175/JAS3578.1 (2005).
- [13] Watts, P. D., Bennartz, R., and Fell, F., “Retrieval of two-layer cloud properties from multispectral observations using optimal estimation,” *J. Geophys. Res.* 116, D16203, 22 pp., doi:10.1029/2011JD015883 (2011)
- [14] Chang, F.-L., Minnis, P., Sun-Mack, S. Nyugen, L., and Chen, Y., “On the satellite determination of multi-layered multi-phase cloud properties,” *Proc. AMS 13th Conf. Atmos. Rad. and Cloud Phys.*, JP1.10 (2010).
- [15] Minnis, P., S. Sun-Mack, S., Yost, C. R., Chen, Y., Smith, W. L., Jr., Chang, F.-L., Heck, P. W., Arduini, R. F., Trepte, Q. Z., Ayers, K., Bedka, K., Bedka, S., Brown, R. R., Heckert, E., Hong, G., Jin, Z., Palikonda, R., Smith, R., Scarino, B., Spangenberg, D. A., Yang, P., Xie, Y., and Yi, Y., “CERES MODIS cloud product retrievals for Edition 4, Part I: Algorithm changes to CERES MODIS,” *IEEE Trans. Geosci. Remote Sens.*, submitted (2019).
- [16] Sun-Mack, S., Minnis, P., Smith, W. L., Hong, G., and Chen, Y., “Detection of single and multilayer clouds in an artificial neural network approach,” *Proc. SPIE Conf. Remote Sens. Clouds and the Atmos. XXII, Warsaw, Poland*, 11-14, 10424-7, 12 pp., doi: 10.1117/12.2277397 (2017).
- [17] Kox, S., L. Bugliaro, L., and A. Ostler, A., “Retrieval of cloud optical thickness and top altitude from geostationary remote sensing,” *Atmos. Meas. Tech.* 7, 3233-3246, doi:10.5194/amt-7-3233-2014, (2014).
- [18] Minnis, P., Hong, G., Sun-Mack, S., Smith, Jr., W. L. , Chen, Y., and Miller, S., “Estimation of nocturnal opaque ice cloud optical depth from MODIS multispectral infrared radiances using a neural network method,” *J. Geophys. Res.* 121, doi:10.1002/2015JD024456 (2016).
- [19] Håkansson, N., Adok, C., Thoss, A., Scheirer, R., and Hörnquist, S., “Neural network cloud top pressure and height for MODIS,” *Atmos. Meas. Tech.* 11, 3177-3196, doi:10.5194/amt-11-3177-2018 (2018)

- [20] M. M. Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. D., Gelaro, R., Stajner, I., and Nielsen, J. E., "The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0, and 5.2.0," *Technical Report Series on Global Modeling and Data Assimilation*, vol. 27, 118 pp., (2008).
- [21] Vaughan, M. A., Pitts, M. Trepte, C., Winker, D., Detweiler, P., Garnier, A., Getzewich, B., Hunt, W., Lambeth, J., Lee, K.-P., Lucker, P., Murray, T., Rodier, S., Tremas, T., Bazureau, A., and Pelone, J., "Cloud-Aerosol LIDAR Infrared Pathfinder Satellite Observations (CALIPSO) data management system data products catalog, Release 4.10," NASA Langley Research Center Document PC-SCI-503, Hampton, Va., USA (2016).
- [22] Sassen, K. and Z. Wang, "Classifying clouds around the globe with the CloudSat radar: 1-year of results," *Geophys. Res. Lett.*, 35, L04805, doi:10.1029/2007GL032591 (2008).
- [23] Kato, S., S. Sun-Mack, W. F. Miller, F. G. Rose, Y. Chen, P. Minnis, and B. A. Wielicki, "Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles," *J. Geophys. Res.* 115, D00H28, doi:10.1029/2009JD012277 (2010).