



1 Identification of ice-over-water multilayer clouds using multispectral 2 satellite data in an artificial neural network

3 Sunny Sun-Mack¹, Patrick Minnis¹, Yan Chen¹, Gang Hong¹, and William L. Smith, Jr.²

4 ¹Analytical Mechanics Associates, Inc., Hampton, VA, USA 23666

5 ²NASA Langley Research Center, Hampton, VA, USA 23681

6 *Correspondence to:* Sunny Sun-Mack (szedung.sun-mack-1@nasa.gov)

7 **Abstract.** An artificial neural network (ANN) algorithm, employing several Aqua MODerate-resolution Imaging
8 Spectroradiometer (MODIS) channels, the retrieved cloud phase and total cloud visible optical depth, and temperature and
9 humidity vertical profiles is trained to detect multilayer (ML) ice-over-water cloud systems identified by matched 2008
10 CloudSat and CALIPSO (CC) data. The trained MLANN was applied to 2009 MODIS data resulting in combined ML and
11 single layer detection accuracies of 87% (89%) and 86% (89%) for snow-free (snow-covered) regions during the day and
12 night, respectively. Overall, it detects 55% and ~30% of the CC ML clouds over snow-free and snow-covered surfaces,
13 respectively, and has a relatively low false alarm rate. The net gain in accuracy, which is the difference between the true and
14 false ML fractions, is 7.5% and ~2.0% over snow-free and snow/ice-covered surfaces. Overall, the MLANN is more accurate
15 than most currently available methods. When corrected for the viewing-zenith-angle dependence of each parameter, the ML
16 fraction detected is relatively invariant across the swath. Compared to the CC ML variability, the MLANN is robust seasonally
17 and interannually, and produces similar distribution patterns over the globe, except in the polar regions. Additional research is
18 needed to conclusively evaluate the VZA dependence and further improve the MLANN accuracy. This approach should
19 greatly improve the monitoring of cloud vertical structure using operational passive sensors.

20 1 Introduction

21 Passive remote sensing with polar-orbiting and geostationary passive imagers is currently the only approach suitable for nearly
22 continuous monitoring of clouds day and night around the globe. While cloud remote sensing is well established and the
23 methodologies are abundant (e.g., Stubenrauch et al., 2013), detecting and characterizing multilayered clouds remains a
24 continuing challenge. Typically, algorithms employed to retrieve properties such as cloud optical depth, effective hydrometeor
25 size, top height, and phase, treat the radiances for a given cloudy imager pixel as emanating from a single plane-parallel cloud
26 sheet. Rarely, if ever, will an actual cloud entirely satisfy the plane-parallel assumptions. Instead, the sizes and densities of the
27 hydrometeors vary vertically and horizontally within the atmospheric column corresponding to a cloudy imager pixel. The top
28 and side surfaces of clouds, even stratus, typically have bumps and troughs that deviate to various degrees from the uniformity



29 implicit in the plane-parallel model (e.g., Loeb and Coakley, 1998). Simply having vertical extent can cancel the plane-parallel
30 assumption when viewing parts of a cloud side (e.g., Liou and Ou, 1979). Multilayered clouds also violate the model.

31 Accounting for variations in single-layer-cloud morphology with a non-plane-parallel type of model is too complex for
32 use in operational retrieval algorithms and likely requires information that is currently unavailable in most imager radiance
33 datasets. Moreover, radiative transfer calculations used in weather and climate models are based on the same plane-parallel
34 premise, although methods are being developed to account for some 3D structure (e.g., Schäfer et al., 2016). For single-layer
35 (SL) clouds, the nonuniform geometry is the predominant deviation from the plane-parallel ideal. It mostly affects retrievals
36 of cloud optical depth (COD) and particle effective radius (CER), but, less so, cloud-top phase and height (CTH). The presence
37 of two water phases and separation of the upper and lower layers in ice-over-water multilayer (ML) systems can also produce
38 large errors in COD and CER, and significantly diminish the accuracies of thermodynamic phase and cloud-top height (CTH)
39 retrievals (e.g., Minnis et al., 2007; Yost et al., 2023). Reducing uncertainties in the retrievals due to nonconformance with the
40 SL plane-parallel ideal, particularly for ML clouds, is critical to increasing the value of imager cloud retrievals for a variety of
41 applications.

42 Reliable determination of cloud characteristics is critical to the Clouds and the Earth's Radiant Energy System (CERES,
43 Loeb et al., 2016) for converting broadband radiance measurements to reliable shortwave and longwave fluxes at the top of
44 the atmosphere, within the atmosphere, and at the surface. Cloud properties extracted from satellite imagery are also exploited
45 in a wide variety of applications. These include, among others, verifying climate model cloud parameters (e.g., Zhang et al.,
46 2010; Stanfield et al., 2014), enhancing aviation safety (Mecikalski et al., 2007; Smith et al., 2012), and improving short-term
47 weather forecasts (e.g., Kurzock et al., 2019; Benjamin et al., 2021) and estimating surface radiative fluxes (e.g., Rutan et al.,
48 2015; Ryu et al., 2018). All of these applications and others (e.g., Chen and Zhang, 2000; Morcrette and Christian, 2000) will
49 benefit from more accurate cloud properties, especially for ML systems.

50 Active instruments such as the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) lidar on the Cloud-Aerosol
51 Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite (Winker et al., 2009) and the Cloud Penetrating
52 Radar (CPR) on CloudSat (Stephens et al., 2008) together have produced quite accurate depictions of cloud vertical structure.
53 These satellites are part of the A-Train, the Afternoon Constellation of sun-synchronous orbiters that, for years, flew nearly
54 the same tracks (13:30 Equatorial crossing time) separated by only a few minutes. Other satellites with imagers, particularly
55 Aqua with the MODerate-resolution Imaging Spectroradiometer (MODIS), are also members of the A-Train. The CALIOP
56 and the CPR are both near-nadir viewing instruments that generate profiles of atmospheric particles only in a narrow curtain
57 along the satellite track. Those profiles, which often include overlapping clouds, are valuable for many uses, especially when
58 combined with other instruments on the A-Train. Because they sample only a tiny fraction of the globe at two local times each
59 day, however, the current active instruments have limited utility for many of the applications served by operational satellite
60 imager products.



61 Efforts to accurately identify and unscramble ML clouds from passive imagery have yielded a variety of methods that
62 have either been demonstrated as efficacious or are being applied routinely. They are based on interpreting radiances from
63 either multiple instruments or from a single instrument with multiple channels. To identify ML clouds, Lin et al. (1998)
64 matched microwave radiometer (MWR) retrievals of cloud liquid water path (LWP) from polar-orbiting satellites with
65 retrievals of COD from geostationary satellites. Minnis et al. (2007) used MWR retrievals of LWP matched with imager
66 retrievals of COD and CER to detect and retrieve ML cloud properties over water surfaces. By combining Aqua MODIS CTP
67 and COD with the optical centroid cloud pressure retrieved from the Ozone Monitoring Instrument on the Aura satellite, *Joiner*
68 *et al.* [2010] discriminated between vertically extended and ML clouds.

69 The single-instrument approach, which is more viable for monitoring ML clouds from a greater number of satellites, often
70 relies on discrepancies between the visible (VIS, $\sim 0.65 \mu\text{m}$) channel COD and that determined from other channels. For
71 example, COD derived from infrared radiances is limited to values of less than ~ 5 because the usable signal diminishes for
72 thicker clouds. Thus, the two COD retrievals can be used to detect thin cirrus over a thicker lower cloud. Pavolonis and
73 Heidinger (2004) identified ML clouds by comparing the COD retrieved from the brightness temperature differences from the
74 11 and 12- μm channels, BTD_{1112} , with the VIS COD. They suggested that the MODIS 1.38- μm and 1.63- μm reflectances
75 could be combined with BTD_{1112} to retrieve the cirrus optical depths for comparison with the VIS COD. The MODIS CO_2
76 channels were used by *Chang and Li* [2005] to retrieve IR COD for ML cloud detection. Their method was simplified by
77 *Chang et al.* [2010] to employ brightness temperatures from CO_2 and 11- μm channels to identify high clouds and ultimately
78 to detect ML clouds in the CERES MODIS Edition 4 products (CM4; Minnis et al., 2021). Similarly, *Wind et al.* [2010]
79 contrasted the cloud top pressure (CTP) derived with a CO_2 method to that based on absorption in the MODIS 0.94- μm channel
80 along with other tests to identify ML clouds in MODIS pixels. To further improve ML detection, this technique was enhanced
81 with additional tests, including the Pavolonis and Heidinger (2002) method (Platnick et al., 2017; Marchant et al., 2020).

82 Desmons et al. (2017) exploited multi-angle polarized spectral reflectances and two different retrievals of CTP from the
83 Polarization and Directionality of the Earth's Reflectance (POLDER) instrument on another A-Train satellite. Instead of using
84 COD retrievals, Wang et al. (2019) utilized a series of tests applied to three spectral reflectances and two infrared brightness
85 temperatures measured by the Visible Infrared Imaging Radiometer Suite (VIIRS) to classify clouds as single-layer ice or
86 water, multilayer, probable multilayer, or uncertain phase and layering. Their technique yielded results similar to those from
87 Platnick et al. (2017). For detecting ML clouds in VIIRS data, CERES replaced the CO_2 channel with the 12- μm channel in
88 the Chang et al. (2010) approach. It found fewer ML clouds than the method using the CO_2 channel (Minnis et al. 2023).

89 CERES is focused on accurately characterizing the radiation budget at the top of the atmosphere, at the surface, and within
90 the atmosphere, so it is important to reliably identify all clouds and retrieve their properties, ideally at all levels within the
91 atmosphere. To improve the CERES ML detection, Sun-Mack et al. (2017) began the development of an artificial neural
92 network (ANN) to distinguish between SL and ML clouds using MODIS radiance data matched to CALIPSO and CloudSat
93 vertical profiles of clouds. Minnis et al. (2019) further enhanced this multilayer ANN (MLANN) by including more input



94 parameters and additional output variables such as upper layer CTH, COD, and cloud-base height (CBH). They also used only
95 high-confidence CloudSat and CALIPSO data for training. They found that, for nonpolar clouds, the MLANN correctly
96 identified ML and SL clouds together 80.4% and 77.1% of the time during day and night, respectively, using CALIPSO data
97 averaged over an 80-km distance. While those results are quite encouraging, the approach needs further refinement and global
98 coverage. Tan et al. (2022) applied several machine-learning methods to a geostationary satellite dataset and found even greater
99 accuracy with a random-forest technique. It is clear that artificial intelligence techniques are the most promising approach to
100 detecting ML clouds from passive imagery.

101 This paper reports on continued development of the MLANN to detect ML clouds. Revisions to the previous training are
102 made using a newer version of CALIPSO products with constrained horizontal resolution. In addition to the separate day/night
103 training used in previous versions, the MLANN herein is trained separately for CERES ice and water cloud pixels separately
104 for snow-free and snow/ice-covered surfaces. Finally, because the MLANN is trained with near-nadir data, its utility for full
105 swath MODIS data is examined.

106 2 Artificial neural network

107 A neural network is a form of artificial intelligence used to identify a relationship between input and output variables by
108 learning from a set of observed or simulated data (e.g., Karayiannis and Venetsanopoulos, 1993). Neural networks are
109 increasingly used in cloud remote sensing using satellite imagery. Kox et al. (2014) and Strandgren et al. (2017) employed an
110 ANN to determine cirrus COD and CTH, while Cerdena et al. (2007) used it to estimate liquid water cloud COD and CER and
111 Taravat et al. (2015) detected the presence of clouds with it. Using an ANN, Minnis et al. (2016) retrieved thick ice cloud
112 COD at night and Håkansson et al. (2018) accurately determined CTP and CTH. Stengel et al. (2020), Wang et al. (2020), and
113 White et al. (2021) use ANNs for cloud detection and phase discrimination. These and other examples demonstrate that ANNs
114 have significant potential for improving the characterization of global cloudiness. This potential appears to derive primarily
115 from the ability of the method to better interpret several subtle, but often ambiguous radiative signals that are difficult to
116 reconcile in physically based retrievals. This is particularly true for thick clouds at night (e.g., Minnis et al. 2016) or for ML
117 clouds. The frequency of ambiguous signals in the radiance fields spurred development of the MLANN by Sun-Mack et al.
118 (2017).

119 The MLANN is used to decide if a pixel contains a SL or ML cloud by relating a vector of input data to output (see
120 schematic in Fig. 1). Each input neuron, the variable, x_i , is linked to each member of a hidden layer through the function, $u_j =$
121 $f(g_j)$, 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 Fig. 1. The output
122 layer is uniquely linked to the input through the hidden layer by the weights w_j and the constant c . The number of neurons j ,
123 N_j , is selected to optimize the output accuracy. Likewise, the number of input variables or i neurons is also adjusted to minimize
124 the output uncertainty. In this study, 50 -70 neurons are used for the hidden layer. The output is a probability between 0.0 and
125 1.0.



126 The Levenberg-Marquadt optimization method is employed to calculate w_{ij} in the MLANN training. The entire process to
127 solve for the weights consists of three components: training, testing, and validation. A majority of the data is used to compute
128 the weights, while smaller portions are used for testing and validating the MLANN. Because the testing and training are
129 interwoven, the testing results force the fitting or training process to stop or to further adjust the weight values. When the
130 process stops (i.e., no further improvement is gained from adjusting the weights), the training ends, yielding the final weights.
131 The validation data are then used to determine if their output yields error statistics that are within the maximum stipulated
132 error. Validation vectors are used to stop training early if the MLANN performance fails to improve. The formulation of this
133 approach is given in more detail by Minnis et al. (2016, 2019) and Sun-Mack et al. (2017) along with the references found
134 therein.

135 **3 Data and methodology**

136 The MLANN is trained with Aqua MODIS imager data and cloud products, and numerical weather model analyses as input
137 and active sensor data as the output. Training is performed using 2008 data sampled to accommodate computer memory
138 capacity and the percentage of ML clouds. The results are assessed using 2008, 2009, and 2013 data. According to Kato et al.
139 (2010), approximately 51% of cloud systems identified by the CPR and CALIOP consist of two or more layers separated by
140 at least 200 m. Of these, 57, 28, 10, 4, and 1% have 2, 3, 4, 5, and 6+ layers, respectively. Those statistics include liquid-over-
141 liquid, ice-over-ice, water-over-ice as well as ice-over-water cloud overlap. Unscrambling this variety of layering is a daunting
142 task.

143 To simplify ML detection and later retrievals, only those clouds having the greatest differences in properties are assumed
144 to be multilayered. Thus, only systems having ice-over-liquid clouds are considered because they differ in phase, scattering
145 properties, and altitude, and are more common than liquid-over-ice clouds. Thus, multilayered clouds are defined here as any
146 combination of ice-cloud layers above one or more water cloud layers with the constraint that the top of the uppermost water
147 layer must be at least 1 km below the bottom of the lowest ice cloud layer. All ice cloud layers together are considered to
148 constitute a single cloud layer. Similarly, all liquid layers are considered together as a single layer. Selection of 1 km as the
149 minimum separation distance is based on the need to ensure complete separation between the ice and water layers and to
150 maximize the number of detected ML clouds. Sun-Mack et al. (2017) and Minnis et al. (2019) found that a larger separation
151 distance can result in greater accuracy but at the expense of missing a significant number of actual ML clouds. Tan et al. (2022)
152 have demonstrated that detection accuracy is significantly compromised when the separation distance is less than 1 km.

153 **3.1 Data**

154 MODIS on Aqua, the CALIOP, and the CPR, took measurements continuously within ± 3 min of each other over a given
155 location until 2011, when CloudSat suffered battery problems and thereafter only collected data during the day. CloudSat



156 exited the A-Train during February 2018. The CALIOP and the CPR were aligned to view nearly the same area along their
157 respective orbits. Because their flight tracks are typically close to the Aqua nadir path, MODIS scans the same scene at viewing
158 zenith angles, $VZA < 18^\circ$. Vertical profiles of clouds from the CALIPSO Version 4 (Vaughan et al., 2016) and the CloudSat
159 2B-CLDCLASS P_R04 (Sassen and Wang, 2008) datasets were matched with 1-km Aqua MODIS Collection 6.1 radiances
160 and CERES-retrieved cloud properties in an updated CloudSat, CALIPSO, CERES, and MODIS (C3M) product of Kato et al.
161 (2010). The CERES MODIS cloud properties of interest, cloud-top phase, cloud particle effective radius, R_{CM} , and COD, τ_{CM} ,
162 were retrieved from MODIS radiances using an interim CERES Edition 5 algorithm, CM4+. It uses the CM4 algorithms with
163 two changes. In CM4+, τ_{CM} is retrieved over snow and ice using a combination of 1.61- μm and 1.24- μm reflectances, as
164 suggested by Minnis et al. (2021). The former channel is used for thinner clouds ($COD < 8$ for ice clouds, $COD < 32$ for water
165 clouds) and the latter for thicker clouds. The Aqua MODIS 1.6- μm channel consists of 10 sensors. Of those ten, only six
166 operate properly. To obtain full 1.6- μm imagery, the 1.6- μm reflectance for each bad pixel was replaced with that from the
167 nearest good detector. That method was applied to the 101-pixel-wide MODIS swath of the C3M. The other change for CM4+
168 is the use of the two-habit ice crystal model of Loeb et al. (2018) for retrieving ice cloud properties.

169 A separate dataset is the full-swath Aqua CM4 cloud product and the Aqua MODIS Collection 6.1 radiances, which are
170 sampled every other scan line and every fourth pixel. Note that the 1.6- μm channel is only available from 3 of the five sensors
171 on the CERES sampled Aqua MODIS Collection 6.1 data. The corrections applied to faulty 1.6- μm sensors for C3M have not
172 yet been implemented for CERES Aqua MODIS data prior to 2019. Therefore, the sampling is reduced and all data from the
173 faulty sensors is not included in the full-swath 2009 and 2013 data.

174 To produce a complete vertical profile of cloud-filled layers, the C3M converts the CloudSat CLDCLASS high-confidence
175 cloud profiles from 240-m to 60-m vertical resolution, then merges them with the CALIPSO cloud profile and vertical feature
176 mask. The nominal horizontal footprint of a CALIOP shot at the Earth's surface is 330 m in diameter. To detect fainter clouds,
177 the CALIPSO processing system computes horizontal averages (HA) of the lidar signals from multiple shots corresponding to
178 increasing distances along the track: 1, 5, 20, and 80 km.

179 Trepte et al. (2019) established that the daytime CM4 cloud detection rate over ocean dropped from nearly 80% for the
180 optically thinnest clouds from CALIPSO, those having $COD < 0.1$, to 50% or less when the clouds detected by the CALIOP
181 with $HA > 20$ and 80-km were added to those clouds found using $HA \leq 5$ km. This indicates that few cirrus clouds discernible
182 only at the longer CALIPSO averaging distances can be detected and are less likely to be identified in ML conditions. In the
183 previous MLANN formulation (Minnis et al. 2019), CALIPSO data with $HA \leq 80$ km were used, no τ_{CM} minimum value was
184 assigned, and vertical profiles of humidity were not included in the input. Thus, extremely thin cirrus clouds were included in
185 that earlier training. Only CALIPSO data with $HA \leq 5$ km are used here to define the cloud profiles. Additionally, in the
186 previous analysis, all three CALIPSO 0.33-km pixels matched to a given MODIS pixel were required to be cloudy after the
187 horizontal averaging was performed. Here, only two out of three 0.33-km pixels are required to be cloudy and any cloud having
188 $\tau_{CM} < 0.5$ is assumed to be single-layered. The latter constraint assumes that the ML signal from such optically thin clouds is



189 negligible, and any retrieval attempt will yield upper and lower cloud-layer properties that are, at most, highly uncertain. To
190 cover all seasons for snow-free surfaces and facilitate computer processing for training, the C3M data were sampled every
191 fourth pixel of 2008 for snow/ice-free (SF) areas, while all pixels were used for the snow/ice-covered (SC) scenes. This full-
192 year training set is more comprehensive than the 1-month dataset of Minnis et al. (2019).

193 The C3M data were merged with the relevant surface skin temperature and vertical profiles of relative humidity taken
194 from the CERES Meteorology, Ozone, and Aerosol (MOA) product (Gupta et al., 1997). The latest MOA product is the result
195 of regridding and interpolating spatially and temporally the version 5.4 reanalysis produced by the Global Modeling
196 Assimilation Office Global Earth Observing System (GEOS-5.4), an update of the versions described by Rienecker et al.
197 (2008). These are the same data employed in the CM4 retrievals. To perform additional analyses, the CC COD, τ_{CC} , was
198 computed for each ice-layer pixel identified as ML in the CALIPSO-CloudSat profiles. The value of τ_{CC} is equal to the
199 CALIPSO ice COD, when the CALIOP signal shows a return from the lower layer cloud, otherwise it is equal to the combined
200 CALIPSO-CloudSat COD.

201 3.2 Input and Output Layers

202 The MODIS input variables are listed in Table 1. They include latitude, longitude, surface elevation, brightness temperatures
203 T_{λ} and brightness temperature differences, $BTD_{\lambda_1\lambda_2} = T_{\lambda_1} - T_{\lambda_2}$, where λ is the wavelength in μm abbreviated to the first two
204 digits. Here, brightness temperatures at 3.7, 6.7, 8.5, 11, 12, and 13.3 μm are used together with BTD_{3711} , BTD_{6711} , BTD_{8511} ,
205 BTD_{1112} , and BTD_{1113} . The parameters involving the 13.3- μm channel were not used over snow-covered surfaces because of
206 striping in the 13.3- μm images over Greenland and Antarctica. The GEOS-5.4 input data consist of the surface skin temperature
207 and relative humidity at the surface and at 850, 700, 500, 400, 300, 200, and 100 hPa. During the day, solar zenith angle (SZA)
208 and the reflectances, ρ_{λ} , from the MODIS 1.38, 1.61, and 2.13 μm channels and the reflectance differences between 1.6 and
209 2.1 μm . CM4+ retrievals of τ_{CM} and R_{CM} are also employed for the MLANN formulation. The reflectances and R_{CM} were not
210 used by Minnis et al. (2019).

211 A cloudy pixel is considered to be ML, if the MLANN output probability is ≥ 0.5 . Otherwise, it is identified as SL. Three
212 components are used in the training process: training, testing, and validation. The sampled 2008 data are split with 60%, 20%,
213 and 20% each for the respective components. To provide an independent validation of the trained MLANN, a fully sampled
214 Aqua MODIS dataset from 2009 is analyzed and the results compared with the matching CC data.

215 A total of eight training sets for each parameter of interest are developed based on the surface condition (snow-free or
216 snow-covered), time of day and the CM4+ cloud-top phase selection. Night is defined as times when $SZA > 82^\circ$, while daytime
217 corresponds to $SZA \leq 82^\circ$. Each training set yields the weights, w_{ij} and w_j , along with the constants b and c . Training is
218 performed using the minimum ice-liquid cloud separation threshold, $\Delta Z_{\min} = 1$, where the separation ΔZ is the difference in
219 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
220 is assumed to be present in the profile, if



- 221 (1) CALIPSO VFM cloud phase is either ice clouds or mixed phase clouds, or
222 (2) at least, one layer occurs at a height above the altitude corresponding to 253 K and no temperature
223 inversion exists in the atmospheric layer between the altitudes corresponding to 273 K and 253 K.
224

225 All C3M data having an ice cloud layer by that definition are used as output in the MLANN training. A given profile is assumed
226 to contain a ML cloud system if a cloudy layer exists at least 1 km below the lowest ice cloud layer and the base of that cloud
227 layer is below the 253-K level. Ice-over-water layer systems are considered to be contiguous if they are closer than 1 km. This
228 separation ensures that the two sensors have actually detected a ML cloud system.

229 Figure 2 breaks down the CC classification according to the CM phase selection, which presumably corresponds to the
230 cloud top. Over SF surfaces, 68% of the MODIS pixels classified as ice during the day (Fig. 2a) are SL ice according to CC,
231 while ~25% correspond to ML clouds. The CM phase is incorrect for ~7% of the cases. At night, more ML clouds are evident
232 and nearly 10% of the ice clouds are classified as liquid. For SF liquid clouds from MODIS (Fig. 2b), the day and night results
233 are comparable with ~15% of the clouds actually being SL ice according to the active instruments. A smaller fraction of the
234 ML clouds correspond to the CM liquid phase than for the ice phase. Yet overall, more ML clouds are found for the CM water
235 phase because the liquid pixels comprise 63% of the 9.7×10^6 SF observations during the day and 58% of the 11.1×10^6 SF
236 pixels at night. The breakdown is different for SC scenes. During the day (Fig. 2c), 80% of the ice pixels are SL ice and only
237 12% are ML. The phase accuracy is reduced at night, but the ML cloud fraction is the same. For CM liquid clouds (Fig. 2d),
238 multi-layer cloud occurrence is slightly greater than for the ice clouds, while phase accuracy is the worst for any of the four
239 categories. Around 30% of the SL ice clouds are mistaken as liquid.

240 4 Results

241 The results presented here consist of comparisons of the MLANN and corresponding CC parameters for the 2008 training
242 dataset along with data from 2009 to ensure robustness of the estimates. Weights and constants were determined by training
243 for each category and parameter and then applied to the independent datasets. The MLANN was trained using the C3M data
244 for the four categories in Fig. 2 to obtain four sets of weights and constants for each surface type.

245 4.1 Multi-layer cloud detection

246 Figure 3 plots the CC cloud profiles retrieved over six areas during 25 December, 2009. The CC layering classification uses
247 gray for SL and blue for ML. In addition to the cloud profiles, the MLANN selection of SL (gray) and ML (blue) is indicated
248 by the lines of dots across the top section of each panel. The surface elevation is denoted by the black dashed line in each
249 panel. On the left are daytime observations over the tropical Atlantic (Fig. 3a), eastern Europe (Fig. 3b), and eastern Antarctica
250 (Fig. 3c). Nocturnal profiles are given on the right for passes over the tropical Indian Ocean (Fig. 3d), the south Pacific (Fig.



251 3e), and northern Russia (Fig. 3f). In the tropical overpasses, the MLANN detects a large fraction of the ML clouds, but also
 252 misses a noticeable number of ML pixels. For example, only a few of the intermittent ML clouds between 2.5°N and 6.1°N
 253 are identified by the MLANN in Fig. 3a and a segment of continuous ML clouds near 12°S in Fig. 3d is missed by the MLANN.
 254 Similar results are seen for the mid-latitude SF areas, where a few ML clouds are missed around 59°N in Fig. 3b and near 35°S
 255 in Fig. 3e.

256 Some false ML clouds are also found in these panels. In Fig. 3b, for example, false ML clouds are evident at 51°N and
 257 also in two areas between 56 and 58°N. In the latter case, there are ML clouds in the profile but they were not classified as
 258 such by the CC constraints, possibly due to the 253 K liquid cloud temperature threshold. Thus, some of the false ML may
 259 actually be ML clouds. The detection rates for the two SC profiles are much reduced. During the daytime case (Fig. 3c), only
 260 one stretch of ML clouds is detected, while even fewer ML clouds are detected at night around 76°N (Fig. 3f). The unidentified
 261 ML clouds are more common in both cases.

262 To summarize the results for all of the data, a confusion matrix was constructed for each category. Referring to Table 2,
 263 agreements between the MLANN and CC SL and ML classifications are denoted as SS and MM, respectively, while false SL
 264 and ML pixels are given by SM and MS, respectively. These are used to define the following parameters.

265 Fraction correct (FC) or Accuracy: $ACC = FC = SS + MM$. (1)

266 Real Risk: $RR = SM + MS = 1 - ACC$. (2)

267 False ML rate: $FM = MS / (MS + MM)$. (3)

268 Precision: $PR = 1 - FM$. (4)

269 Recall: $RC = MM / (SM + MM)$. (5)

270 Single-layer Confidence: $CoS = 1 - SS / (SS + SM)$. (6)

271 Net Gain of Accuracy: $NGA = MM - MS$. (7)

272 These parameters facilitate the reporting and discussion of the results and the comparisons with other algorithms.

273 The training results in Table 3 include the confusion matrices for all eight categories with FC in bold along with the
 274 number of CC SL and ML pixels and their sum. During the day, FC is 84.1% for CM4 ice clouds over SF areas, with the
 275 fraction of ML correctly identified, i.e., $RC = 49\%$. The classification during the day is a bit better for CM4 liquid clouds: FC
 276 $= 88.7\%$ and $RC = 63\%$. The real risk for ice clouds is 15.9% compared to 11.3% for the liquid clouds. At night the results are
 277 similar, although a little worse for ice clouds, with $FC = 81.3\%$. However, the ice clouds yield a larger fraction, $PR = 55\%$, of
 278 true ML pixels than during the day. Fewer ML clouds are found for liquid clouds at night. More of the ML clouds are classified
 279 as ice because the nocturnal cloud temperature retrieval is based strictly on infrared radiances (Yost et al. 2021, 2023).
 280 Nevertheless, FC is the same for both times of day for liquid phase clouds. At night, RR for ice clouds increases to 18.7% and
 281 drops slightly for water clouds. A total of ~12 million pixels was used in the SF training.

282 As suggested by Fig. 3, the efficacies of the MLANN for detecting ML clouds over snow-covered areas are considerably
 283 reduced relative to their snow-free counterparts. While the FC values are actually greater than those during the day (Table 3),
 284 recall drops to 35% and 45% for ice and liquid clouds, respectively, during the day. The fraction detected, $CoM \sim 22\%$, is



285 even lower at night. Nevertheless, because fewer pixels qualify as ML clouds over SC surfaces, according to the definition
286 used here, the MLANN RR values are smaller than those over SF surfaces. It is notable that, for both SF and SC surfaces, the
287 ML false alarm rate is less than 55%.

288 4.2 Independent evaluation

289 The results from the training are encouraging, but they are not based on an independent dataset. To evaluate the robustness of
290 the MLANN, all 2009 Aqua MODIS data were processed with the trained algorithm. In general, the statistics for the eight
291 categories are all very similar to those in Table 3. To summarize the effectiveness of the MLANN, the 2009 ice and liquid ML
292 results for the SF/SC and day/night categories are combined in Table 4. Over SF surfaces, FC is 87.0% and 85.6% for day and
293 night, respectively, while the corresponding values over SC surfaces are 89.3% and 88.7%. Despite the large FC values, the
294 MLANN underestimates the ML fraction over SF surfaces by 5.8% and 4.8% during night and day, respectively, for the
295 matched CC and MODIS cloudy pixels. A total of 80 million SF pixels were processed, compared to 26 million for SC surfaces.
296 Over SF areas, RR is 13% for day and 15% at night. Real risks drop to 11% for SC regions. It should be noted that the ML
297 fractions reported here are for the number of multi-layer MODIS pixels divided by the total number of cloudy matched CC
298 and MODIS pixels. Since the CERES MODIS mean cloud fraction is ~ 0.66 , the actual fraction of MODIS pixels that are
299 classified as ML would need to be multiplied by 0.66.

300 The net gain of accuracy relative to the SL assumption is an important parameter to consider in any ML detection scheme.
301 Using the SL assumption in cloud retrievals, $FC = SS + SM$. Introducing a multilayer detection method yields both false and
302 true ML pixels. Thus, a new source of error comes with the additional information. The net gain of accuracy is not simply
303 equal to $FC - SS$, it must account for the newly introduced error, represented in MS. The falsely detected ML clouds are a
304 potentially worse source of error than the SL assumption for ML clouds because in retrieving ML properties, an additional
305 cloud layer has to be created to satisfy the ML constraint. Thus, including a ML detection algorithm in a retrieval may not be
306 reasonable if FM is too large. Based on Table 3, the MLANN NGA = 7.6% and 7.3% during day and night, respectively, over
307 SF surfaces. The corresponding values over snow-covered surfaces are 3.4% and 1.0%. While the MLANN provides a nearly
308 negligible amount of information over SC at night, elsewhere it clearly represents an improvement over simply assuming that
309 all clouds are single-layered.

310 Global distributions of the mean 2009 ML cloud fraction from the validation results are plotted in Fig. 4. Compared to the
311 average CC ML fractions (Fig. 4a), the daytime MLANN means (Fig. 4b) are distributed similarly, but are noticeably lower.
312 The fractions are smaller in nearly every location. At night (bottom), the patterns are much like those during the daytime,
313 except in the polar regions. More CC pixels (Fig. 4c) are classified as ML in the tropics than during the daytime. A comparable
314 increase occurs in the MLANN nocturnal results (Fig. 4d), which have ML fractions over much of the Amazon Basin and
315 central Africa that are comparable to their CC counterparts, although they are smaller elsewhere. As expected from Table 2,
316 the MLANN ML fractions in the polar regions are relatively small during the day and negligible at night.



317 The latitudinal variations of the mean ML fractions are plotted in Fig. 5. As expected from Fig. 4, the zonal patterns
318 are much the same with the MLANN values (triangles) being consistently less than their CC counterparts (circles). In the
319 tropics, the daytime differences generally fall between -0.06 to -0.04 and drop to as low as -0.08 in the polar regions. During
320 the night, the minimum of -0.10 is found over the polar regions, but the differences are comparable to the daytime values
321 between 45°S and 45°N. The MLANN is clearly less effective during the night over snow and ice-covered areas, especially at
322 night. Overall, the MLANN underestimates the 2009 ML cloud amount by 0.05 and 0.06 relative to the CCML cloud fraction
323 during day and night, respectively.

324 5 Discussion

325 These results represent a significant improvement over the previous MLANN formulation. Much of the increased accuracy is
326 due to the shorter horizontal averaging distances used here. By employing CALIPSO averages over distances up to 80 km,
327 Minnis et al. (2019) attempted to detect ML cloud systems that included many cirrus clouds having optical depths smaller than
328 0.20. Such clouds are difficult to detect with passive remote sensing even when they are single-layered. According to Yost et
329 al. (2021), systems having $\tau_{CC} < 0.2$ account for ~42% of all ML clouds for CALIPSO data using $HA \leq 80$ km compared to
330 only 18% for $HA \leq 5$ km during the day. Similar differences were observed at night. By using the smaller averaging distance,
331 the fraction of available ML clouds for this study dropped by almost 25% relative to that used in Minnis et al. (2019). Other
332 sources for the improvement arise from using additional input parameters, including those based on the 13.3- μm channel and
333 the reflectances at 1.38, 1.61, and 2.13 μm , as well as the vertical profiles of relative humidity. Additionally, the assumption
334 that all pixels having $\tau_{CM} < 0.5$ are automatically SL, regardless of the CALIPSO classification, probably removed some
335 difficult but less important cases.

336 5.1. Dependence on cloud properties

337 Much like other retrievals, the MLANN is sensitive to various cloud conditions, such as the altitudes of the two layers and
338 their respective optical depths. Because the MLANN uses a minimum separation distance of 1 km between the ice and liquid
339 cloud layers, the dependence on separation distance is not explicitly considered here. Its impact on MLANN, examined by
340 Sun-Mack et al. (2017) and Minnis et al. (2019), is similar to that from other studies. Tan et al. (2022), for example, found that
341 the probability of ML detection using a random forest method was greatest for separation differences of 3 km or more and that
342 it dropped from values exceeding 0.8 to less than 0.5 for cloud gaps smaller than 1 km. Greater discrepancies in altitude
343 between the upper and lower clouds increase the differences in the layer temperatures yielding stronger signals in the thermal
344 channels. It is assumed that this type of dependency, found in the aforementioned MLANN studies, is operative for this version
345 of the MLANN. Despite the apparent increase in accuracy using wider separation in the training, Minnis et al. (2019) found
346 that NGA was 60% greater for 1-km separation compared to the 3-km separation dataset. Moreover, the smaller separation
347 yielded nearly 50% more true ML clouds than the greater separation. The increase in apparent accuracy in the dataset using a



348 minimum 3-km gap relative to its 1-km counterpart is primarily due to assuming that a significant fraction of the ice-over-
349 water systems is single-layered, even though there is separation and two different phases in the column.

350 As formulated here, the MLANN assumes that all clouds with $\tau_{CC} < 0.5$ are SL. To examine this assumption, the MLANN
351 was also trained without any minimum COD limit. On average, FC dropped by 1.2% and the total fraction of ML clouds from
352 CC increased by 1.8%. Despite the drop in FC, NGA rose by 0.1%. Thus, the net impact is small and the downstream task of
353 unscrambling the upper and lower cloud properties from a cloud system with such a small COD will be eased somewhat.

354 For the radiation budget, some of the most important factors are the CODs associated with the detected and missed ML
355 systems. To determine the efficacy of the MLANN as a function of COD, the MLANN recall is plotted in Fig. 6 for each (τ_{CC} ,
356 τ_{CM}) bin for 2009. In the plots, τ_{CC} is the ice COD for ML clouds, i.e., the upper layer COD. Irregular scales are used for the
357 axes to provide more detail for the lower COD values. Because of large uncertainties and reduced sampling, bins having τ_{CC}
358 > 20 are not reliable. During the day, RC is greatest ($\sim 90\%$) for the bins having $\tau_{CC} \sim 1.7$ and $\tau_{CM} \sim 11$ for both ice (Fig. 6a) and
359 water clouds (Fig. 6b). RC exceeds 0.5 for ice clouds having $\tau_{CM} > 3$ and $0.3 < \tau_{CC} < 5$. When $1.5 < \tau_{CM} < 3$ and $0.3 < \tau_{CC} <$
360 1.3 , RC remains above the halfway mark. The shape of the 50th percentile envelope for the water clouds differs from the ice
361 clouds as a result of more upper-cloud CODs being smaller than for ML clouds identified as ice (Yost et al. 2021). Thus, the
362 training for liquid clouds produces better ML detection when the ice clouds have small CODs.

363 At night, τ_{CM} is based only on thermal channels and, therefore, is mostly constrained to values of 8 or less. Default values
364 of 8, 16, and 32 are employed whenever the cloud is assumed to be optically thick. The particular default value depends on
365 the circumstances (Minnis et al., 2021). Sometimes, the CM4 and CM4+ analytical COD retrievals produce a value exceeding
366 8. Typically, τ_{CM} is closer to the upper-cloud COD at night, being influenced little by the lower cloud when the separation
367 distance is large. Ignoring the high τ_{CC} bins, the nocturnal RC maxima are found around bins (1.1, 2.0) and (1.3, 4.5) for ice
368 (Fig. 6c) and water clouds (Fig. 6d), respectively. True ML clouds are found more often than false SL clouds for $1 < \tau_{CM} < 5$,
369 when the phase is ice and $0.1 < \tau_{CM} < 3$. The halfway COD bounds narrow to $0.2 < \tau_{CC} < 0.4$ for greater values of τ_{CM} . For
370 water-phase clouds, $RC > 50\%$ occurs mostly for $1.5 < \tau_{CM} < 8$ and $1.5 < \tau_{CC} < 4$. It is clear that the thermal channel method
371 is sensitive to thinner upper clouds compared to the daytime methods when the solar channel signal is overwhelmed by the
372 lower cloud reflectances. Conversely, the daytime method detects more ML clouds when $\tau_{CC} > 3$ or so.

373 This is more evident in Fig. 7, which shows histograms of the matrix parameters as a function of upper-layer τ_{CC} for SM
374 and MM and τ_{CM} for SS and MS over SF surfaces. For water phase clouds (Fig. 7a) the relative frequency of true ML pixels
375 (MM), shown as solid lines, is greater at night than during the day when $\tau_{CC} < 0.5$, but the occurrence of daytime MM pixels
376 exceeds their nighttime counterparts when $\tau_{CC} > 0.9$. Similar behavior is seen for the ice phase pixels (Fig. 7b), but the
377 thresholds shift from 0.5 to 1.4 and from 0.9 to 1.9. The false SL or missed ML (SM) clouds, shown as dashed lines, vary
378 differently. For the ice phase pixels (Fig. 7b), the SM pixel frequency rises with increasing τ_{CC} up to $\sim 8\%$ at $\tau_{CC} = 3.5$ before



379 decreasing to 5-7% , then dropping toward zero at $\tau_{CC} = 25$. This peak for the thick ice clouds reflects the difficulty of inferring
380 a lower layer under a nearly opaque cloud. For water-phase clouds (Fig. 7a), SM is most common for $\tau_{CC} < 0.3$ and diminishes
381 steadily to near zero around $\tau_{CC} = 30$. As τ_{CC} increases, the ML system is more likely to be identified as ice phase, so fewer
382 cases of ML systems having large upper-layer COD will be included in this population. In both cases, the night and day SM
383 frequencies track each other relatively closely with τ_{CC} . Similar variations are found over the SC surfaces (not shown).
384 Cumulative probability distribution functions based on the SF and SC results, presented in Fig. S1, show that 50% of the
385 missed ML clouds have $\tau_{CC} < 0.25$ for SF water clouds and < 0.5 for SC water clouds.

386 Figures 7c and 7d show the frequency histograms of SS and MS for CC SL liquid and ice clouds, respectively, as a
387 function of τ_{CM} . As expected, the peak SS (true SL) frequency occurs for $\tau_{CM} < 0.5$ for both phases, day and night. During the
388 day, a secondary true SL maximum is found around $\tau_{CM} \approx 25$ for ice and water clouds. At night, that secondary peak is around
389 $\tau_{CM} \approx 14$ for ice pixels and near $\tau_{CM} \approx 9$ for liquid clouds. Nocturnal false ML (MS) clouds are found mostly between CODs
390 of 1 and 5 at night for ice pixels and between 2 and 6 for water clouds. During the day, MS occurs most often for $\tau_{CM} \approx 14$ for
391 water clouds. In fact, the MS frequency seems to follow the SS values, except for $\tau_{CM} < 0.5$. The daytime MS occurrence is
392 relatively flat for ice clouds with $\tau_{CM} > 1.0$.

393 Another factor that can influence ML detection is the assumption that the lower cloud layer is composed of liquid water
394 whenever the cloud temperature is less than 253 K. While that is true for most clouds, a small fraction of ice clouds have top
395 temperatures above 253 K (e.g., Hu et al., 2010). In those instances, the ML signal would likely be reduced because of
396 similarities in the optical properties of the two layers. Mixed-phase clouds, which often occur in the supercooled temperature
397 range, would have a similar effect, but to a smaller degree depending on the amount of ice in the cloud. On the other hand,
398 supercooled clouds globally account for about half of the clouds having an infrared CTT between 243 K and 253 K. If only
399 snow and ice surfaces are considered, the range is 239 K to 242 K (see Fig. 6 of Hu et al., 2010). Thus, some systems with
400 cold (CTT < 233 K) ice clouds over supercooled liquid clouds with CTT < 253 K could be identified as SL ice by the definition
401 used here. These complementary effects due to supercooled clouds could produce some confusion in the training of the
402 MLANN, particularly in polar regions.

403 The CODs used in the training would not be the same as those determined using the standard CM4 algorithms employed
404 for the 2009 retrievals because the CM4+ algorithms used a different ice crystal model and a new method for retrieving COD
405 over snow. This change in COD retrieval apparently had minimal impact on the detection as the 2008 training and 2009
406 validation results are nearly identical.

407 5.2 Comparisons with other results

408 As noted earlier, multilayered cloud detection has been the subject of many different algorithmic studies, so it is important to
409 better understand how the current approach compares to those other algorithms. Direct comparisons are not straightforward



410 because each algorithm was developed with its own specific constraints and ML definitions. The CERES Ed4 ML algorithm
411 (Chang et al., 2010a, Minnis et al., 2021) was applied only when a cloud with pressure below 500 hPa was detected using a
412 CO₂-absorption method (Chang et al., 2010b). The MODIS science team algorithm (Wind et al. 2010) was applied to a 5-km
413 cloud product and was only used when the MODIS optical depth exceeded 4. The latest version, MYD06 C6.1 (Platnick et al.,
414 2017), adds the BTD₁₁₁₂ technique developed by Pavolonis and Hiedinger (2004). Desmons et al. (2017) used data from the
415 Polarization and Directionality of the Earth's Reflectance (POLDER) to detect ML clouds of all types, but only for $\tau > 5$. Ice-
416 over-water multilayered clouds were detected by Wang et al. (2019) during daytime using Suomi-NPP (SNPP) Visible Infrared
417 Imager Radiometer Suite (VIIRS) data in a thresholding method. Tan et al. (2022) placed no restrictions on either τ or the
418 number of layers, but they applied their random forest algorithm and other machine learning techniques only to geostationary
419 Himawari-8 data. Because of its orbit, the Himawari-8 observations are taken over a full range of VZA when matched with
420 the CC profiles, but the VZA is constant for a given location. Other published methods have either not produced extended
421 datasets or performed only case-study evaluations with objective data. Despite the sampling disparities, it is informative to
422 compare some of the statistics to provide some context to the performance of the MLANN. These comparisons are summarized
423 in Table 5.

424 Comparing with CC data, Desmons et al. (2017) found that for overcast clouds with $\tau > 5$, FC = 70% and CoS = 74%.
425 For the same conditions, they determined that MYD06 C6.1 yields FC = 67% and CoS = 73%. Additional parameters computed
426 from their Table 4 are listed in Table 5. Precision and recall from MYD06 are 54% and 47%, respectively, while they are 58%
427 and 47% from POLDER. These can be compared to the MLANN daytime validation results (Table 5), which combine the SC
428 and SF daytime data in Table 4 weighted by 0.13 and 0.87, fractions that roughly correspond to the areal coverage of the
429 respective surface types (e.g., Yost et al., 2023). All of the MLANN parameter values exceed their restricted MYD06 and
430 POLDER counterparts. Wang et al. (2019) only reported validation results in terms of percent of CALIOP ML and SL. Thus,
431 only RC and CoS could be determined from their results. For $\tau > 1$, the recall is about the same as the daytime MLANN, if the
432 ML and probably ML categories from their algorithm are combined. Similarly, their CoS is around 10% smaller than the
433 MLANN value. If clouds with $\tau < 1$ are included, both CoS and RC drop substantially. Note that Wang et al. (2019) did not
434 include CloudSat retrievals in their evaluation, so ML clouds with an optically thick upper cloud are not included in the
435 statistics.

436 Although no value for FC was given, the values of certain parameters can be estimated for all clouds with an unrestricted
437 optical depth from the figures in Desmons et al. (2017). From their Fig. 8, RR \approx 38%, so FC = 62%. The value of CoS is the
438 same for restricted and all clouds. The MODIS parameters change only negligibly for all clouds compared to the restricted
439 case because the MYD06 algorithm only uses clouds with $\tau > 4$. Marchant et al. (2020) also compared the MYD06 to the 2B-
440 CLDCLASS-lidar products and found that for clouds with $\tau > 4$, FC = 63% with the Pavolonis and Heidinger (2004) algorithm
441 and 65% without it. If it assumed that all clouds with $\tau < 4$ are SL, then FC jumps to 80% and 81% for the two algorithm
442 options. But that assumption excludes 45% of the ML clouds as defined by Marchant et al. (2020).



443 Except for the definition of what constitutes a ML cloud (ice over water, water over water, etc.), NGA is the one parameter
444 that is not too dependent on cloud optical depth assumptions. From Desmond et al. (2017), the daytime POLDER and MYD06
445 C6.1 cases yield NGA = 4.4% and 2.2%, respectively. Presumably, some of the POLDER results include water-over-water
446 clouds. Nevertheless, the POLDER algorithm yields a net gain in information. The results of the Marchant et al. (2017) analysis
447 yield slightly lower numbers for the MODIS C6.1 NGA, 0.2% and 1.4%, with and without the Pavolonis and Heidinger (2004)
448 method. In either case, the MLANN daytime NGA exceeds those of the MYD06 and POLDER techniques. Moreover, it greatly
449 exceeds the CERES Ed4 ML results (not shown).

450 The random forest results from Tan et al. (2022), confined to 60°S - 60°N and 80°E and 160°W, were trained with 1-km
451 matched 2B-CLDCLASS-LiDAR profiles using the product's layer flag to determine if a given pixel is SL or has more than
452 one layer, regardless of layer phase. That training dataset produced ACC = 85% and 79%, respectively, for the daytime and
453 all-hours algorithms. The latter method included no reflectance input from solar channels so it can be used for both day and
454 night conditions. It is included in the bottom section of Table 5 for comparing to the MLANN night version. For this technique,
455 PR = 81% and 73% for day and all-hours, respectively, with corresponding RC values of 72% and 64%. While ACC is less
456 than that found with the MLANN for both day and night, the random forest PR and RC results are greater than their MLANN
457 counterparts. At night, the MLANN PR is nearly equal to the Himawari-All value. The Himawari CoS and NGA daytime
458 values were deduced from the values in their Table V and their equations (1) - (3). The MLANN CoS values exceed their
459 Himawari counterparts, but the MLANN global NGAs are less than half of those from the random forest training results. Those
460 larger values arise, in part, from including many more types of ML clouds in the random forest training than used for the
461 MLANN.

462 A fairer comparison would use independent validation sets from both algorithms, but a complete summary of the
463 validation comparisons was not provided in Tan et al. (2022). However, several parameters can be determined from their Fig.
464 5, which utilized a dataset independent of the training data. The resulting values of PR are 70% and 64% for day and all-hours,
465 respectively, while CoS = 89% and 85%. The MLANN SL confidences are slightly greater at 90% and 88% and its PR values
466 exceed the Himawari validation results, especially for night/all-hours. Without further information it is not possible to
467 determine the values of ACC and RC for the geostationary validation dataset. However, because CoS is the same or larger for
468 the validation dataset and PR dropped by 11 points from the training results, it can be inferred that the fraction of false ML
469 clouds increased considerably. This would reduce ACC and substantially diminish NGA.

470 Interestingly, the best results from the Tan et al. (2022) validation analysis are for ice-over-liquid and ice-over-mixed
471 clouds. The former corresponds to conditions that the MLANN was developed to detect, while at least some of the latter were
472 included in the MLANN analysis. Approximately 30% of the actual ML clouds detected in the Tan et al. (2022) validation
473 analysis are for single phase or upper-layer mixed phase ML clouds that MLANN was not designed to identify. Assuming that
474 the portion of the ice-over-water/mixed is the same for the training dataset, the correctly detected ice-over-water cloud amount



475 is 0.10. Adding the ice-over-mixed would yield 0.17. Reducing that by the ratio of PRs from the validation and training sets
476 would drop the range to 0.09 - 0.15, which bounds the correct ML fraction from the SF cases in Table 4.

477 **5.3 Full-swath detection**

478 The MLANN training is based on near-nadir measurements from both the CC and MODIS instruments. Increasing optical path
479 lengths due to increasing VZA modify the radiances emanating from a given location through absorption and scattering. This
480 is particularly true for radiances at solar wavelengths. Thus, the near-nadir-based MLANN coefficients are not necessarily
481 valid for observations taken at other VZAs. For operational use with Aqua MODIS data, the MLANN must be reliable across
482 all viewing angles.

483 **5.3.1 Angular dependence**

484 The VZA dependency is examined by first computing the mean radiances for each viewing angle across the full scan for data
485 taken during JAJ0 2019. It was found that the radiance VZA dependence is sensitive to the forward or backward portion of
486 the scan cycle. The former view is toward the sun, while the latter is directed away from the sun. Figure 8 plots the reflectance
487 averages for each VZA bin in the forward (positive) and backward (negative) directions. From near-nadir to 65°, the 1.60- μm
488 reflectance (solid lines) for water cloud pixels increases by 11% and 25% in forward and back directions, respectively. For ice
489 clouds, the corresponding increases are 22% and 37%. Similar changes are seen for the 2.13 μm reflectances (dashed lines).
490 The 1.38- μm reflectance for ice behaves in much the same manner (Fig. S2), but is nearly constant with VZA for water clouds.
491 The daytime 3.75- μm radiances (Fig. S3) are relatively flat in the back direction, but increase with VZA for liquid clouds. At
492 night, the radiances show the classic limb-darkening behavior of thermal radiation. This can be seen in Fig. 9. During the day
493 (solid lines), the water-cloud 10.8- μm radiances are relatively flat in the forward direction and drop a little at the higher VZAs
494 in the back direction. Ice cloud radiances decrease in both directions, but more so in the forward view where the 10.8- μm
495 radiances are lower than their back-direction counterparts. The forward scan views more shadowed areas that could affect the
496 thin cloud and partly cloudy scenes over land (Minnis et al., 2004). At night (dashed lines), the limb-darkening is more
497 apparent. No back and forward differences are considered at night. Similar variations in radiance are seen at 8.55- μm (Fig. S5)
498 and 11.90- μm (Fig. S6). There are only minor radiance differences between the forward and back directions during the day
499 for the 6.70- μm channel (Fig. S4), presumably because it is mostly unaffected by the layers below the cloud. Additional plots
500 of radiances as a function of VZA (Figs. S6 -S14) are provided in the Supplemental Material.

501 To adjust the MODIS radiances, ice and water correction factors were determined for each waveband, day and night,
502 separately over SF and SC surfaces. For daytime, the correction factors were computed for both forward and back scans. These
503 factors were developed for both the channel radiances and reflectances and each of the BTM parameters. The correction factor
504 is simply the ratio of the mean radiance for VZA between -3° and -18° divided by the mean radiance at a given VZA. Thus,
505 the observed radiance is adjusted to the near-nadir view of MODIS by simply multiplying it by the correction factor.



506 To test the impact of these factors on the retrievals, the MLANN was applied to the uncorrected and corrected full-swath
507 MODIS data for April 2009. Figure 10 shows the variation of mean ML fraction as a function of VZA for SF ice and water
508 clouds, day and night. During the day (Fig. 10a), the uncorrected and corrected ML fractions are nearly identical suggesting
509 that the correction factors have minimal effect on the radiances. While this is not surprising, given the relatively flat daytime
510 curves in Fig. 9 and for other thermal channels, it is actually due to a cancelation of the ML decrease in the back direction by
511 an increase in the forward direction (not shown). In contrast to the daytime results, the nocturnal ML fractions have a
512 nonmonotonic variation with VZA for the uncorrected radiances and a significant steady decrease to a value near zero for the
513 corrected case. The uncorrected radiances for ice clouds (Fig. 10b) yield a rise in ML detection during the day and a drop at
514 night. When the correction factors are applied to the radiances, the ML amounts are relatively constant with VZA. To obtain
515 the most consistent product across the swath, the adjustments are applied to all of the radiances, except for water cloud pixels
516 observed during the night.

517 5.3.2 Example images

518 Figure 11 shows the results of applying the MLANN with VZA correction to an Aqua MODIS image taken over the Southern
519 Ocean centered near 57°S, 165°E at ~3:50 UTC, 16 April 2019. The pseudo color RGB image (Fig. 11b) shows an extensive
520 area of stratocumulus on the left side that is apparently overlaid with thin cirrus that blurs the view of the underlying clouds.
521 A second extensive liquid cloud deck appears near the top center that might overlay some StCu clouds, but is itself covered
522 by thicker ice clouds. The CM4 cloud phase results (Fig. 11a) highlight those contiguous dense ice clouds, which likely obscure
523 lower clouds. Thin cirrus also appear to overlies parts of the second liquid deck. The MLANN (Fig. 11c) determines that a large
524 portion of the image consists of ice-over-water clouds. In general, the outline of the cloud effective heights (CEH) above 1 km
525 correspond to the ML pixels (Fig. 11d) except where the ice cloud is very thick, or perhaps, the ice cloud is in close proximity
526 to or contiguous with the lower deck, as over parts of the white clouds in top center part of the image. Other higher, SL liquid
527 clouds are seen near the top left corner and bottom right of center. Cloud phase is very mixed over the thin cirrus areas, yet
528 the MLANN determines most of the pixels as ML.

529 Multilayered clouds detected with the nighttime MLANN are shown in Fig. 12 for a MODIS image taken around 4:45
530 UTC the same day over the North Atlantic. The scene (Fig. 12b) contains extensive but variable cirrus coverage (white) and
531 broken stratus clouds typically between 1 and 3 km (Fig. 12d). Thicker cirrus clouds are identified as ice (Fig. 12a) while many
532 of faintest ones, primarily those over the low clouds, are classified as liquid. Denser ice clouds and those over open water
533 appear to be at altitudes between 9 and 14 km, while the thin Ci over St range from 3 to 7 km, which is expected, given the SL
534 cirrus altitudes. The MLANN appears to identify many of those ML clouds (Fig. 12c), but tends to miss those with overlying
535 thick cirrus. There may be some false ML clouds in the upper right, but it is difficult to tell because the thinnest Ci is not
536 always discernible in the RGB image.



537 The final example shown here (Fig. 13) is taken the same day around 01:50 UTC over the polar ice cap centered near
538 80°N, 155°E. Snow and ice cover provide the scarlet background (Fig. 13b), which is overlaid with low clouds in various
539 shades of white to gray and a slightly higher deck in the center with thin Ci covering much of the top half of the image. That
540 cirrus appears as blurry pinkish gray and identified as ice or liquid depending on the thickness (Fig. 13a), while most of the
541 cirrus over the deck in the middle is designated as liquid. Those clouds are identified as ML by the MLANN along with the
542 small area at the bottom and in the upper left (Fig. 13c). Only a few parts of the cloud left of center are classified as ML, while
543 it appears more ML clouds should have been detected. Most of the SL ice cloud CEHs are only between 3 and 6 km (Fig.
544 13d), while those over the middle deck are less than 3 km. The low CEH values are likely due to overestimation of the COD
545 by the CM4 retrieval for SL clouds and to the presence of the thick lower cloud for the ML retrievals. Detection of the ML
546 clouds will allow reclassification of the cloud tops as ice and recalculation of the cloud properties, when the components of a
547 two-layer retrieval system are in place.

548 These three examples and the two additional cases shown in Figs. S15 and S16 demonstrate that MLANN performs
549 reasonably well across the full swath. No wild false ML clouds are evident although some recognizable misses are seen, as
550 expected from the analyses above. Quantifying the accuracy of the correction-factor approach to full swath application of the
551 MLANN would ultimately require using the method on similar data taken by a different satellite, such as SNPP VIIRS, that
552 overlapped with the CC data at various VZAs. Using VIIRS data, Wang et al. (2019) found that the ML clouds they detected
553 showed minimal changes with VZA. That result is similar to the daytime curves in Fig. 10b. Developing and analyzing a
554 comparable VIIRS-CC dataset is beyond the scope of this paper, but is planned for future research.

555 5.3.3 Assessment of full swath results

556 In the interim, more indirect approaches are available. For example, the off-nadir and near-nadir results should be spatially
557 consistent if the swath approach is working properly. To examine this aspect, Figure 14 shows the distributions of ML cloud
558 fractions averaged over the months of January, April, July, and October (JAJO) 2009 from three different data sources. These
559 include daytime retrievals from all CC data (Fig. 14a), MLANN applied to Aqua MODIS radiances observed at the reference
560 near-nadir ($-3^\circ < \text{VZA} < 18^\circ$) angles (Fig. 14b), and to Aqua MODIS data taken at all VZAs (Fig. 14c). The corresponding
561 nocturnal results are plotted in Figs. 14d-f. These results are noisier than those in Fig. 4 because they are based on only 4
562 months of data and they include all observations, not just those having good CC and C3M cloudy pixel matches. The MODIS
563 results include both false and partially cloudy pixels.

564 As in Fig. 4, the CC and near-nadir patterns are comparable although the MLANN means are often smaller than their CC
565 counterparts. The areas with minimal ML amount in the near-nadir results (Fig. 14b) are in the same locations as those from
566 the CC retrievals, but are more pronounced. Some CC maxima are reproduced by the MLANN, but the MLANN fractions
567 near the maxima drop off more precipitously than their CC counterparts. For example, the maximum off the southern Chilean
568 coast in Fig. 14b is nearly identical that in Fig. 14a, but the MLANN fractions in the surrounding areas are generally smaller



569 than the CC values. While much smoother, the non-polar patterns in both all-VZA cases (Fig. 14c) are similar to those from
570 the near-nadir results, but the Chilean maximum is diminished somewhat.

571 Distributions of ML fractions from the same datasets are even more consistent at night. The maxima over northern South
572 America, central Africa, and Indonesia are well defined in all three maps. Like the daytime results, the non-polar minima are
573 much better delineated in Figs. 14e and 14f than in the CC data (Fig. 14d). Overall, the distributions in Fig. 16 demonstrate
574 that the full-swath MLANN does not yield spurious ML clouds in areas where they are not expected to occur and generally
575 produce results similar to the near-nadir values.

576 Another measure of robustness of the algorithm is its ability to reproduce the seasonal cycle. This is examined by
577 computing the monthly mean ML anomaly, which is defined as the monthly mean minus the annual average divided by the
578 annual average. It is clear that the SC results over snow miss many ML clouds, especially at night. Thus, to minimize the
579 influence of SC regions on the seasonal cycle, only nonpolar ($60^{\circ}\text{S} - 60^{\circ}\text{N}$) data are considered. Figure 15 plots the ML fraction
580 anomaly for each month of 2009 from CC and the MLANN applied to full-swath Aqua MODIS data. The MLANN day and
581 night anomalies track their CC counterparts remarkably well, within a few percent in most cases.

582 To further examine the reliability of the MLANN on longer time scales, it was applied to January, April, July, and October
583 (JAJO) 2013 Aqua MODIS full swath data. The global distributions of the 2009 and 2013 results (Fig. S16) are similar, but
584 reveal shifts in the locations of the maxima. Table 6 presents the global mean JAJO 2009 and 2013 ML fractions along with
585 the land-ocean ratio, L/O, which is the global average ML fraction over land divided by that over water surfaces. Overall, ML
586 fractions for all CC data are 3 - 5% greater than their MODIS-matched counterparts, a result comparable to the differences in
587 Fig. 4. The 2009 MLANN near-nadir values are 0.01 smaller than those for all VZAs. ML fractions in Table 6 are all less than
588 their counterparts in Fig. 4. This is due to the fact that the CC data in Table 6 include all cloudy pixels that the CERES cloud
589 mask classified as clear and the MLANN results include many partly cloudy pixels that are not likely to be ML. The clouds
590 detected by CALIPSO, but missed by CERES are mostly SL thin cirrus and SL low clouds (Yost et al. 2021, 2023), which
591 would dilute the ML fraction determined using all of the CC data. The differences between the CC and near-nadir MLANN
592 are reduced by ~2% compared to those using only the matched data. During daytime, the MLANN mean ML fractions from
593 2013 are 0.5% greater than those in 2009, while at night the 2013 averages exceed their 2009 counterparts by 0.2% near nadir
594 and 0.6% across the full swath. For both years, the nocturnal near-nadir values are ~1% less than for data taken at all VZAs.

595 The CC land-ocean ratios, L/O, in Table 6 reveal that fewer ML clouds occur over land than over water surfaces. For CC,
596 L/O is between 0.77 and 0.84, while it varies from 0.64 to 0.77 for the MLANN results, indicating that the MLANN is less
597 efficient at detecting ML clouds over land than over water bodies. Together with the similarity of the CC and ML seasonal
598 cycles, the consistency of the near-nadir and full-swath L/O values and small differences in ML amounts during both years
599 are quite encouraging for using the MLANN on an operational basis.



600 5.3.4 Operational considerations

601 CERES is a long-term project that utilizes many different satellites and imagers to characterize cloud properties. The MODIS
602 on Aqua and Terra and the VIIRS on SNPP and NOAA-20 are coincident with the CERES broadband radiometers and observe
603 non-polar regions at fixed times each day. Any system designed to detect ML clouds should be applicable to both the VIIRS
604 and MODIS imagers and, ideally, to the geostationary imagers that are used to help assess the radiation budget at other times
605 of day. Because the latter have had widely varying spectral channel complements since 2000, use of MLANN with them is
606 beyond the scope of this discussion. The VIIRS lacks certain channels used here (13.3 μm and 6.7 μm) and the channels
607 common to MODIS and VIIRS differ in spectral coverage and filtering. Additionally, the VIIRS is a higher resolution
608 instrument. Thus, it may be necessary to train VIIRS with CC data to obtain a consistent ML result. Another approach would
609 require careful inter calibration of the VIIRS and MODIS channels using spectral corrections (e.g., Scarino et al., 2016) and
610 the addition of radiances from the missing channels determined from a process that fuses data from VIIRS and the Crosstrack
611 Infrared Sounder (e.g., Weisz et al., 2017). If those are not available, then the MLANN would need to be retrained with fewer
612 input radiances, likely at the expense of accuracy. To that end, initial training tests indicate that without those channels, ACC
613 decreases 87.0% to 86.4% during the day and from 85.6% to 84.3% during the night over SF surfaces. During the day, NGA
614 drops from 7.6 to 7.1%, while at night NGA goes from 7.3 % to 6.2%. NGA is relatively unaffected by the loss of the 13.3-
615 μm channel; almost all of the diminished accuracy is due to the absence of the 6.7- μm channel, particularly at night. Even with
616 the loss of those channels, the resulting detection capability would still represent a significant advancement over previous
617 efforts.

618 As in all retrievals, reliable and consistent calibration across platforms is essential to providing an accurate ML product.
619 It may be even more important for the MLANN because the neural network relies on subtle radiance differences that may be
620 lost in the noise of a physical retrieval. Thus, any small trend in the calibration of one channel may introduce a growing bias
621 in the ML fraction. Similarly, inter platform calibration differences could cause a similar effect. Updated retrieval algorithms
622 and input data are introduced into the CERES data processing whenever major improvements are developed and errors
623 diminished. Since the MLANN relies on a few retrieval inputs such as COD and cloud phase, it would need to be retrained
624 whenever a new CERES cloud algorithm edition is introduced.

625 Further improvement of the MLANN itself, particularly over snow-covered areas, might be gained by using additional
626 parameters or spatial context. For example, Tan et al. (2022) found that radiances from the 7.3- μm channel comprise a highly
627 ranked predictor of ML clouds in their random forest approach. The MODIS equivalent channel was not considered here, but
628 would have to be created for VIIRS using the fusion process noted above. Information about the pixels surrounding the pixel
629 of interest increased the accuracy of ice water path retrieved from a Meteosat imager with a convolutional neural network
630 (Amell et al., 2022). Including selected radiances or BTDs from surrounding pixels might also enhance the MLANN.
631 Additional partitioning of the training categories might also raise ML detectability as it did when the original MLANN (Sun-



632 Mack et al, 2017) was divided into ice and water phase categories (Minnis et al. 2019). These and other approaches could lead
633 to greater accuracies than found here.

634 6. Summary and Conclusions

635 An artificial neural network method has been enhanced to more accurately identify ice-over-water ML cloud systems from
636 multispectral MODIS observations. The algorithm requires as input a variety of radiances, brightness temperature differences,
637 atmospheric profiles of temperature and humidity, and the CERES Edition 4 cloud phase and optical depths. Based on the
638 definitions of single and multilayer clouds used here for CALIPSO-CloudSat profiles, the MLANN correctly identifies SL and
639 ML clouds together 87.0% and 85.6% of the time over surfaces free of ice and snow during day and night, respectively. Over
640 ice or snow-covered areas, the corresponding correct identifications are 89.3% and 88.7%. Despite the good overall agreement,
641 the MLANN only detects 55% of the CC ML clouds over SF regions and only 40% and 20% over SC areas during day and
642 night, respectively. The majority of the missed SF ML clouds are those having an upper-cloud COD < 0.3 (water) or COD ≥
643 3 (ice), although ~35% and 20% of the water and ice-phase detected ML clouds meet those conditions. Over SC surfaces, the
644 undetected ML pixels mainly have an upper-cloud COD < 0.5 or COD > 2.

645 Despite its shortcomings, the MLANN, unlike many other techniques, yields a significant net gain in layering
646 identification accuracy because the number of false ML pixels is substantially less than that for true ML pixels. Overall, the
647 daytime MLANN evaluation metrics are more favorable than those based on physical retrievals or decision tree algorithms,
648 even with the differences in sampling, ML cloud definitions, and optical depth constraints. Few methods have been developed
649 for nocturnal application. Comparisons with results from a machine learning algorithm applied to geostationary satellite data
650 have yielded a more ambiguous assessment. The accuracy and SL confidence from MLANN are greater than those from the
651 Tan et al. (2022) random forest training results for day and night. Yet, the MLANN precision, recall, and NGA values are
652 smaller. If the validation results from Tan et al. (2022) are considered, the MLANN precision values are greater. It is not
653 known how much the MLANN recall and NGA would fare relative to the random forest validation results. Even if it were
654 known, the relative merits of the two methods would be difficult to quantify without accounting for the discrepancies in ML
655 definition and sampling areas and time periods. However, it can be concluded from the comparisons that the MLANN is among
656 the most capable of current ML detection methods.

657 Operationally, the MLANN, trained with near-nadir MODIS views, must be applicable to all the MODIS viewing angles.
658 To account for the variation of radiances with viewing zenith angle, the MODIS-based input parameters are normalized to the
659 nadir view using empirical correction factors. The adjustments yield ML cloud amounts that are mostly invariant with VZA
660 and produce visually reasonable ML detection across the MODIS swath. Spatial distributions of ML cloud fractions from full-
661 swath results are consistent with the near-nadir results and manifest similar detection efficiencies over land and water surfaces
662 that are the same as their near-nadir counterparts. Temporally, the MLANN produces the same seasonal cycle in ML clouds



663 as the active sensor data, albeit with the noted bias. Moreover, the results are similar in magnitude and distribution for different
664 years with shifts in maxima. While more detailed pixel-to-pixel comparisons should be performed using CC data matched to
665 imagery taken at off-nadir VZAs, the analyses performed here indicate that the MLANN should be as successful off of nadir
666 as it is in the near-nadir mode.

667 Applying the MLANN to other imagers should be performed with caution as sensors on other satellites can differ
668 spectrally and spatially (e.g., VIIRS) or may observe at other times of day (e.g., Terra MODIS). Platforms that are not in Sun-
669 Synchronous orbits, for example geostationary satellites, will observe a given scene at times of day and at viewing and
670 illumination angles that are not seen by Aqua MODIS and hence not in the training complement. Adapting the MLANN to
671 different types of orbits or times of day presents a challenge as there are few options for global training and validation. Current
672 and future cloud radar and lidar combinations are confined to afternoon Sun-synchronous satellites (e.g., Heliere et al. 2017).
673 Lidars that can be used for cloud detection have flown on the International Space Station (e.g., Pauly et al., 2019) in a
674 precessing orbit and on Aeolus in a sunrise/sunset Sun-synchronous orbit (Straume et al., 2020). CALIPSO has been slowly
675 moving away from its 1330 LT orbit covering several more hours of the day since 2018. Without the cloud radar, any and all
676 of those lidars could be used to define ML clouds to some extent, depending on their penetration depths, and may be of value
677 for training and validating ML clouds for geostationary imager data. Regardless of the particular satellite, the MLANN would
678 need to be retrained or the spectral channels normalized to MODIS.

679 With layer detection accuracies below 90%, there is clearly room for future improvement, especially over polar regions
680 covered with snow and ice. Use of additional channels or subsets of the current training categories may add a few more points
681 to the overall accuracy. Combining physical retrievals with the neural network may also be the means for detecting more ML
682 pixels. The definition of ML clouds used here is rather restrictive in that it is nominally confined to ice over liquid water
683 clouds. It is also somewhat ambiguous because 253 K serves as the threshold between ice and water clouds for the underlying
684 cloud deck. In lieu of any better information to define the lower cloud phase, the threshold should be altered to account for
685 variability of the 50th percentile ice phase in the supercooled temperature range. Other cloud combinations such as liquid over
686 liquid could be included in the MLANN but they might reduce the accuracy and would probably be more resolvable if treated
687 separately from the ice over water clouds.

688 Detecting multilayer clouds is a first step toward improving the characterizations of global vertical cloud structure using
689 passive sensors. Once identified, the upper and lower layer cloud properties need to be estimated. A number of approaches
690 have been suggested for estimating the top heights of the upper and lower clouds. These include physical retrievals (e.g., Chang
691 et al. 2010) and machine learning (e.g., Minnis et al. 2019). Similarly cloud optical depth and particle effective size could be
692 derived with a physical retrieval (e.g., Chang et al. 2010), a neural network (e.g., Cerdeña et al. 2007), or an optimal estimation
693 method that requires the cloud heights (e.g., Sourdevall et al., 2016). Having a reliable detection method, like the MLANN,
694 should serve as motivation for formulating a robust technique for unscrambling the upper and lower cloud layer properties in
695 future research.



696

697 *Data availability.*

698 MLANN training data used in the paper can be obtained from CERES Ordering Tool: <https://ceres.larc.nasa.gov/data/>

699

700 *Author contributions.* S. Sun-Mack, P. Minnis, and G. Hong developed the detection method. S. Sun-Mack implemented the

701 technique. Y. Chen, S. Sun-Mack, and P. Minnis performed the data analyses. The paper was first drafted by P. Minnis and

702 revised by S. Sun-Mack and W. Smith, Jr. The project was supervised by P. Minnis and W. Smith, Jr.

703

704 *Competing interests.* The authors declare no competing interests.



705

706 *Acknowledgments.* This research is supported by the NASA CERES Project.

707 **References**

- 708 Amell, A., Eriksson, P., and Pfreundschuh, S.: Ice water path retrievals from Meteosat-9 using quantile regression neural
709 networks, *Atmos. Meas. Tech.*, 15, 5701-5717, doi:10.5194/amt-15-5701-2022, 2022.
- 710 Benjamin, S. G., James, E. P., Hu, M., Alexander, C. R., Ludwig, T. T., Brown, J. M., Weygandt, S. S., Turner, D. D., Minnis,
711 P., Smith, Jr., W. L., and Heidinger, A. K.: Stratiform cloud hydro-meteor assimilation for HRRR and RAP model short-
712 range weather prediction, *Mon. Wea. Rev.*, 149, 2581-2598. doi:10.1175/MWR-D-20-0319.1, 2021.
- 713 Cerdeña, A., Gonzalez, A., and Perez, J. C.: Remote sensing of water cloud parameters using neural networks, *J. Atmos.*
714 *Oceanic Technol.*, 24, 52-63. doi:10.1175/JTECH1943.1, 2007.
- 715 Chang, F.-L., Minnis, P., Sun-Mack, S., Nyugen, L., and Chen, Y.: On the satellite determination of multi-layered multi-phase
716 cloud properties, *Proc. AMS 13th Conf. Atmos. Rad. and Cloud Phys.*, JP1.10, 2010.
- 717 Chang, F.-L., and Li, Z.: A new method for detection of cirrus overlapping water clouds and determination of their optical
718 properties, *J. Atmos. Sci.*, 62, 3993-4009, doi:10.1175/JAS3578.1, 2005.
- 719 Chen, T., and Zhang Y. C.: Sensitivity of atmospheric radiative heating rate profiles to variations of cloud layer overlap, *J.*
720 *Climate*, 13, 2941-2959, 2000.
- 721 Desmons, M., Ferlay, N., Riedl, J., and Theuleux, F.: A global multilayer cloud identification with POLDER/Parasol, *J. Appl.*
722 *Meteor. Climatol.*, 56, 1121-1139, 2017.
- 723 Gupta, S. K., Ritchey, N. A., Rose, F. G., Alberta, T. L., Charlock, T. P., and Coleman, L. H.: Regrid humidity and temperature
724 fields (system 12.0). CERES algorithm theoretical basis document release 2.2, NASA, Hampton, VA, NASA RP 1376.
725 [Online]. Available: https://ceres.larc.nasa.gov/documents/ATBD/pdf/r2_2/ceres-atbd2.2-s12.0.pdf, 1997.
- 726 Hélière, A., Gelsthorpe, R., Le Hors, L., and Toulemont, Y.: ATLID, the Atmospheric Lidar on board the EarthCARE
727 Satellite, *Proc. SPIE 10564, International Conference on Space Optics – ICSO 2012, 105642D* (20 November 2017),
728 <https://doi.org/10.1117/12.2309095>, 2017.
- 729 Håkansson, N., Adok, C., Thoss, A., Scheirer, R., and Hörnquist, S.: Neural network cloud top pressure and height for MODIS,
730 *Atmos. Meas. Tech.*, 11, 3177-3196, doi:10.5194/amt-11-3177-2018, 2018.
- 731 Hu, Y., S. Rodier, S., Xu, K.-M., Sun, W., Huang, J., Lin, B., Zhai, P., and Josset, D.: Occurrence, liquid water content, and
732 fraction of supercooled water clouds from combined CALIOP/IIR/MODIS measurements, *J. Geophys. Res.*, 115, D00H34.
733 doi:10.10292009JD012384, 2010.
- 734 Joiner J., Vasilkov, A. P., Bhartia, P. K., Wind, G., Platnick, S., and Menzel, W. P.: Detection of multi-layer and vertically
735 extended clouds using A-Train sensors, *Atmos. Meas. Tech.*, 3, 233-247. doi:10.5194/amt-3-233-2010, 2010.



- 736 Karayiannis, N. B., and Venetsanopoulos A. N.: Efficient learning algorithms for neural networks (ELEANNE), IEEE Trans.
737 Systems, Man, Cybernetics, 23, 1372-1383, 1993.
- 738 Kato, S., Rose, F. G., Ham, S.-H., Rutan, D. A., Radkevich, A., Caldwell, T., Sun-Mack, S., Miller, W. F., and Chen, Y.,:
739 Radiative heating rates computed with clouds derived from satellite-based passive and active sensors and their effects on
740 generation of available potential energy, J. Geophys. Res., 124, 1720-1740, doi:10.1029/2018JD028878, 2019.
- 741 Kato, S., Sun-Mack, S., Miller, W. F., Rose, F. G., Chen, Y., Minnis, P., and Wielicki, B. A.: Relationships among cloud
742 occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles,
743 J. Geophys. Res. 115, D00H28, doi:10.1029/2009JD012277, 2010.
- 744 Cox, S., Bugliaro, L., and Ostler, A.: Retrieval of cloud optical thickness and top altitude from geostationary remote sensing,
745 Atmos. Meas. Tech., 7, 3233-3246, doi:10.5194/amt-7-3233-2014, 2014.
- 746 Kurzrock, F., Nguyen, H., Sauer, J., Ming, F. C., Cros, S., Smith, Jr., W. L., Minnis, P., Palikonda, R., Jones, T. A., Lallemand,
747 C., Linguet, L., and Lajoie, G.: Evaluation of WRF-DART multi-phase cloud water path assimilation for short-term solar
748 irradiance forecasting in a tropical environment. *Geosci. Model Dev.*, 12, 3939-3954, doi:10.5194/gmd-12-3939-2019,
749 2019.
- 750 Li, J., Yi, Y., Minnis, P., Huang, J., Yan, H., Ma, Y., Wang, W., and Ayers, J. K.: Radiative effect differences between multi-
751 layered and single-layer clouds derived from CERES, CALIPSO, and CloudSat data, J. Quant. Spectrosc. Radiat. Transfer
752 112, 361-375, 2011.
- 753 Lin, B., P. Minnis, P., Wielicki, B. A., Doelling, D. R., Palikonda, R., Young, D. F., and Uttal, T.: Estimation of water cloud
754 properties from satellite microwave and optical measurements in oceanic environments. II: Results, J. Geophys. Res., 103,
755 3887-3905, 1998.
- 756 Liou, K.-N., and Ou, S.: Infrared radiative transfer in finite cloud layers, J. Atmos. Sci., 36, 1985-1996.
757 [https://doi.org/10.1175/1520-0469\(1979\)036<1985:IRTIFC>2.0.CO;2](https://doi.org/10.1175/1520-0469(1979)036<1985:IRTIFC>2.0.CO;2), 1979.
- 758 Loeb, N. G. and Coakley, Jr., J. A.: Influence of marine stratus cloud optical depths from satellite measurements: Does 1D
759 theory apply?, J. Climate, 11, 215-233, 1998.
- 760 Loeb, N. G., Su, W., Doelling, D. R., Wong, T., Minnis, P., Thomas, S., and Miller, W. F.: Earth's top-of-atmosphere radiation
761 budget. *Reference Module in Earth Systems and Environmental Sciences*, Elsevier Ltd, Oxford, UK, doi:10.1016/B978-0-
762 12-409548.9.10367-7, 2016.
- 763 Loeb, N. G., Yang, P., Rose, F. G., Hong, G., Sun-Mack, S., Minnis, P., Kato, S., Ham, S.-H., Smith, W. L., Jr., Hiroki, S.,
764 and Tang, G.: Impact of ice cloud microphysics on satellite cloud retrievals and broadband flux radiative transfer model
765 calculations, J. Climate, 31, 1851-1864, doi:10.1175/JCLI-17-0426.1, 2018.
- 766 Marchant, B., Platnick, S., Meyer, K., and Wind, G.: Evaluation of the MODIS collection 6 multilayer cloud detection
767 algorithm through comparisons with CloudSat cloud profiling radar and CALIPSO CALIOP products, Atmos. Meas. Tech.,
768 13, 3263-3275, 2020.



- 769 Mecikalski, J. R., Feltz, W. F., Murray, J. J., Johnson, D. B., Bedka, K. M., Bedka, S. M., Wimmers, A. J., Pavolonis, M.,
770 Berendes, T. A., Haggerty, J., Minnis, P., Bernstein, B., and Williams, E.: Aviation applications for satellite-based
771 observations of cloud properties, convection initiation, in-flight icing, turbulence and volcanic ash, *Bull. Amer. Meteor.*
772 *Soc.*, 88, 1589-1607, 2007.
- 773 Minnis, P., Ghambeer, A. V., and Doelling, D. R.: Azimuthal anisotropy of long wave and infrared window radiances from
774 the Clouds and the Earth's Radiant Energy System on the Tropical Rainfall Measuring Mission and Terra satellites. *J.*
775 *Geophys. Res.*, 109, D08202. doi:10.1029/2003JD004471, 2004.
- 776 Minnis, P., Hong, G., Sun-Mack, S., Smith, Jr., W. L., Chen, Y., and Miller, S.: Estimation of nocturnal opaque ice cloud
777 optical depth from MODIS multispectral infrared radiances using a neural network method, *J. Geophys. Res.*, 121,
778 doi:10.1002/2015JD024456, 2016.
- 779 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-
780 over-water cloud systems using TRMM VIRS and TMI data, *J. Geophys. Res.*, 112, D06206, doi:10.1029/2006JD007626,
781 2007.
- 782 Minnis, P., Sun-Mack, S., Yost, C. R., Chen, Y., Smith, Jr., W. L., Chang, F.-L., Heck, P. W., Arduini, R. F., Treppe, Q.
783 Z., Ayers, K., Bedka, K., Bedka, S., Brown, R. R., Heckert, E., Hong, G., Jin, Z., Palikonda, R., Smith, R., Scarino, B.,
784 Spangenberg, D. A., Yang, P., Xie, Y., and Yi, Y.: CERES MODIS cloud product retrievals for Edition 4, Part I: Algorithm
785 changes to CERES MODIS, *IEEE Trans. Geosci. Remote Sens.*, 58, doi:10.1109/TGRS.2020.3008866, 2021.
- 786 Morcrette J. J., and Christian J.: The response of the ECMWF model to changes in the cloud overlap assumption, *Mon. Wea.*
787 *Rev.* 128, 1707–1732, 2000.
- 788 Pauly, R. M., Yorks, J. E., Hlavka, D. L., McGill, M. J., Amiridis, V., Palm, S. P., Rodier, S. D., Vaughan, M. A., Selmer, P.
789 A., Kupchock, A. W., and Baars, H.: Cloud-Aerosol Transport System (CATS) 1064 nm calibration and validation, *Atmos.*
790 *Meas. Tech.*, 12, 6241–6258. doi:10.5194/amt-12-6241-2019, 2019.
- 791 Pavolonis, M. J., and Heidinger, A. K.: Daytime cloud overlap detection from AVHRR and VIIRS, *J. Appl. Meteor.*, 43, 762–
792 778, doi:10.1175/2099.1, 2004.
- 793 Rienecker, M. M., Suarez, M. J., Todling, R., Bacmeister, J., Takacs, L., Liu, H.-C., Gu, W., Sienkiewicz, M., Koster, R. D.,
794 Gelaro, R., Stajner, I., and Nielsen, J. E.: The GEOS-5 Data Assimilation System - Documentation of Versions 5.0.1, 5.1.0,
795 and 5.2.0, *Technical Report Series on Global Modeling and Data Assimilation*, vol. 27, 118 pp., 2008.
- 796 Rutan D. A., S. Kato, D. R. Doelling, F. G. Rose, L. T. Nguyen, T. E. Caldwell, and Loeb, N. G.: CERES synoptic product:
797 Methodology and validation of surface radiant flux, *J. Atmos. Oceanic Tech.*, 32, 1121–1143, 2015.
- 798 Ryu, Y.-H., Hodzic, A., Barre, J., Descombes, G., and Minnis, P.: Quantifying errors in surface ozone predictions associated
799 with clouds over CONUS: A WRF-Chem modeling study using satellite cloud retrievals, *Atmos. Chem. Phys.*, 18, 7509–
800 7525, <https://doi.org/10.5194/acp-18-7509-2018>, 2018.
- 801 Sassen, K. and Wang, Z.: Classifying clouds around the globe with the CloudSat radar: 1-year of results, *Geophys. Res. Lett.*,
802 35, L04805, doi:10.1029/2007GL032591, 2008.



- 803 Scarino, B. R., Doelling, D. R., Minnis, P., Gopalan, A., Chee, T., Bhatt, R., and Lukashin, C.: A web-based interface for
804 calculating spectral band difference adjustment factors derived from SCIAMACHY data, *IEEE Trans. Geosci. Remote*
805 *Sens.*, 54, 2529-2542, doi:10.1109/TGRS.2015.2502904, 2016.
- 806 Schäfer, S. A. K., Hogan, R. J., Klinger, C., Chiu, J. C., and Mayer, B.: Representing 3-D cloud radiation effects in two-
807 stream models: 1. Longwave consideration and effective cloud edge length, *J. Geophys. Res.*, 121, 8567-8582,
808 doi:10.1002/2016JD024876, 2016.
- 809 Smith, W. L., Jr., P. Minnis, P., Fleege, C., Spangenberg, D., Palikonda, R., and Nguyen, L.: Determining the flight icing
810 threat to aircraft using single-layer cloud parameters derived from operational satellite data, *J. Appl. Meteorol. Climatol.*,
811 51, 1794-1810, doi:10.1175/JAMC-D-12-057.1, 2012.
- 812 Sourdeval, O., Labonnote, C. J., Baran, A. J., Mülmenstädt, J. And Brogniez, G.: A methodology for simultaneous retrieval
813 of ice and liquid water cloud properties. Part 2: Near-global retrievals and evaluation against A-Train products. *Q. J. R.*
814 *Meteorol. Soc.*, 142, 3063-3081, doi:10.1002/qj.2889, 2016.
- 815 Stanfield, R. E., Dong, X., Xi, B., Del Genio, A. D., Minnis, P., and Jiang, J.: Assessment of NASA GISS CMIP5 and post-
816 CMIP5 simulated clouds and TOA radiation budgets using satellite observations. Part I: Cloud fraction and properties. *J.*
817 *Climate*, 27, 4189-4208, doi: 10.1175/JCLI-D-13-00558.1, 2014.
- 818 Stephens, G. L., et al., "CloudSat mission: Performance and early science after the first year of operation," *J. Geophys. Res.*,
819 vol. 113, Dec. 2008, Art. no. D00A18, doi: [10.1029/2008JD009982](https://doi.org/10.1029/2008JD009982).
- 820 Stengel, M., Stapelberg, S., Sus, O., Finkensieper, S., Würzler, B., Philipp, D., Hollmann, R., Poulsen, C., Christensen, M.,
821 and McGarragh, G.: Cloud_cci Advanced Very High Resolution Radiometer postmeridiem (AVHRR-PM) dataset version
822 3: 35-year climatology of global cloud and radiation properties. *Earth Sys. Sci. Data*, 12, 41-60,
823 <https://doi.org/10.5194/essd-12-41-2020>, 2020.
- 824 Straume, A. G., Rennie, M., Isaksen, L., de Kloe, J., Marseille, G.-J., Stoffelen, A., Flament, T., Stieglitz, H., Dabas, A.,
825 Huber, D., Reitebuch, O., Lemmerz, C., Lux, O., Marksteiner, U., Weiler, F., Witschas, B., Meringer, M., Schmidt, K.,
826 Nikolaus, I., Geiss, A., Flamant, P., Kanitz, T., Wernham, D., von Bismarck, J., Bley, S., Fehr, T., Floberghagen, R., and
827 Parrinello, T.: ESA's Space-Based Doppler Wind Lidar Mission Aeolus – First Wind and Aerosol Product Assessment
828 Results, *EPJ Web Conf.* 237, 01007, <https://doi.org/10.1051/epjconf/202023701007>, 2020.
- 829 Stubenrauch, C., Rossow, W. B., Kinne, S., Ackerman, S., Cesana, G., Chepfer, H., Getzewich, B., DiGirolamo, L., Guignard,
830 A., Heidinger, A., Maddux, B., Menzel, P., Minnis, P., Pearl, C., Platnick, S., Poulsen, C., Riedi, J., Sun-Mack, S., Walther,
831 A., Winker, D., Zeng, S., and Zhao, G.: Assessment of global cloud datasets from satellites: Project and database initiated
832 by the GEWEX Radiation Panel. *Bull. Am. Meteorol. Soc.*, 94, 1031-1049, doi:10.1175/BAMS-D-12-00117, 2013.
- 833 Su, W., Corbett, J., Eitzen, Z., and Liang, L.: Next-generation angular distribution models for top-of-atmosphere radiative
834 flux calculation from CERES instruments: validation, *Atmos. Meas. Tech.*, 8, 3297-3313, doi: [10.5194/amt-8-3297-2015](https://doi.org/10.5194/amt-8-3297-2015),
835 2015.



- 836 Sun-Mack, S., Minnis, P., Smith, W. L., Hong, G., and Chen, Y.: Detection of single and multilayer clouds in an artificial
837 neural network approach, *Proc. SPIE Conf. Remote Sens. Clouds and the Atmos. XXII*, Warsaw, Poland, 11-14, 10424-7,
838 12 pp., doi: 10.1117/12.2277397, 2017.
- 839 Tan, Z., Liu, C., Ma, S., Wang, X., Shang, J., Wang, J., Ai, W., and Yan, W.: Detecting multilayer clouds from the
840 geostationary Advanced Himawari Imager using machine learning techniques, *IEEE Trans. Geosci. Remote Sens.*, 60, doi:
841 10.1109/TGRS.2021.3087714, 2022.
- 842 Taravat, A., Proud, S., Peronaci, S., Del Frate, F., and Oppelt, N.: Multilayer perceptron neural networks model for meteosat
843 second generation SEVIRI daytime cloud masking, *Remote Sens.*, 7, 1529–1539, 2015.
- 844 Trepte, Q. Z., Minnis, P., Sun-Mack, S., Yost, C.R., Chen, Y. Jin, Z., Chang, F.-L., Smith, Jr., W. L., Bedka, K. M., and Chee,
845 T. L., 2019: Global cloud detection for CERES Edition 4 using Terra and Aqua MODIS data. *IEEE Trans. Geosci. Remote*
846 *Sens.*, 57, 9410-9449, doi:10.1109/TGRS.2019.2926620, 2019.
- 847 Vaughan, M. A., Pitts, M. Trepte, C., Winker, D., Detweiler, P., Garnier, A., Getzewich, B., Hunt, W., Lambeth, J., Lee, K.-
848 P., Lucker, P., Murray, T., Rodier, S., Tremas, T., Bazureau, A., and Pelone, J.: Cloud-Aerosol LIDAR Infrared Pathfinder
849 Satellite Observations (CALIPSO) data management system data products catalog, Release 4.10, NASA Langley Research
850 Center Document PC-SCI-503, Hampton, Va., USA, 2016.
- 851 Wang, C., Platnick, S., Meyer, K., Zhang, Z., and Zhou, Y.: A machine-learning-based cloud detection and thermodynamic-
852 phase classification algorithm using passive spectral observations, *Atmos. Meas. Tech.*, 13, 2257–2277, 2020.
- 853 Wang, J., Liu, C., Yao, B., Min, M., Letu, H., Yin, Y., and Yung, Y. L.: A multilayer cloud detection algorithm for the Suomi-
854 NPP Visible Infrared Imager Radiometer Suite (VIIRS). *Remote Sen. Environ.*, 227. doi:10.1016/j.rse.2019.02.024, 2019.
- 855 Watts, P. D., Bennartz, R., and Fell, F.: Retrieval of two-layer cloud properties from multispectral observations using optimal
856 estimation, *J. Geophys. Res.* 116, D16203, 22 pp., doi:10.1029/2011JD015883, 2011.
- 857 Weisz, E., Baum, B. A., and Menzel, P. W.: Fusion of satellite-based imager and sounder data to construct supplementary
858 high spatial resolution narrowband IR radiances, *J. Appl. Remote Sens.*, 11, doi:10.1117/1.JRS.11.036022, 2017.
- 859 White, C. H., Heidinger, A. K., and Ackerman, S. A. Evaluation of Visible Infrared Imaging Radiometer Suite (VIIRS) neural
860 network cloud detection against current operational cloud masks, *Atmos. Meas. Tech.*, 14, 3371–3394, 2021.
- 861 Wind G., Platnick, S., King, M. D., Hubanks, P. A., Pavolonis, M. J., Heidinger, A. K., Yang, P., and Baum, B. A.: Multilayer
862 cloud detection with the MODIS near-infrared water vapor absorption band, *J. Appl. Meteor. Climatol.* 49, 2315–2333,
863 doi:10.1175/2010JAMC2364.1, 2010.
- 864 Winker, M., Vaughan, M. A., Omar, A., Hu, Y., Powell, K. A., Liu, Z., Hunt, W., and Young, S. A.: Overview of the CALIPSO
865 mission and CALIOP data processing algorithms, *J. Atmos. Oceanic Tech.* 26, 2310-2323,
866 doi:10.1175/2009JTECHA1281.1, 2009.
- 867 Yost, C. R., Minnis, P., Sun-Mack, S., Chen, Y., and Smith, Jr., W. L.: CERES MODIS cloud product retrievals for Edition
868 4, Part II: Comparisons to CloudSat and CALIPSO, *IEEE Trans. Geosci. Remote Sens.*, 58,
869 doi:10.1109/TGRS.2020.3015155, 2021.



870 Yost, C. R., Minnis, P., Sun-Mack, S., Smith, Jr., W. L., Trepte, Q. Z., and Chen, Y.: VIIRS Edition 1 cloud properties for
 871 CERES. Part 2: Evaluation with CALIPSO, Remote Sens., 15, 1349, doi:10.3390/rs15051349, 2023.

872 Zhang, M. H., Lin, W. Y., Klein, S. A., Bacmeister, J. T., Bony, S., Cederwall, R. T., Del Genio, A. D., Hack, J. J., Loeb, N.
 873 G., Lohmann, U., Minnis, P., Musat, I., Pincus, R., Stier, P., Suarez, M. J., Webb, M. J., Wu, J. B., Xie, S. C., Yao, M.-
 874 S., and Zhang, J. H.: Comparing clouds and their seasonal variations in 10 atmospheric general circulation models with
 875 satellite measurements, J. Geophys. Res., 110, 10.1029/2004JD005021, 2005.

876

877

Tables

878

879

880

Table 1. Input parameters for MLANN.

Regional Parameters	GEOS-5.4	MODIS Thermal Data (K)	MODIS Solar Data [†]
Latitude (°), Longitude (°)	Surface skin temp (K)	T ₃₇ , T ₆₇ , T ₈₅ , T ₁₁ , T ₁₂ , T ₁₃ [@]	R _{CM} , τ _{CM}
Surface type, elevation	Relative humidity (%) at 8 levels*	BTD ₃₇₁₁ , BTD ₆₇₁₁ , BTD ₈₅₁₁ , BTD ₁₁₁₂ , BTD ₁₁₁₃ [@]	ρ _{1.38} , ρ _{1.61} , ρ _{2.13}
Solar zenith angle (°) [†]	Temperature (K) at 8 levels		ρ _{1.61} - ρ _{2.13}

881

882

883 * levels: surface, 850, 700, 500, 400, 300, 200, 100 hPa

884 [†] day only

885 [@], snow/ice free only

886

887



888

889

Table 2. Confusion matrix definition.

	CC Single	CC Multi	Total
MLANN Single	SS	SM	SS + SM
MLANN Multi	MS	MM	MS + MM
Total	SS + MS	SM + MM	SS + MM



892

893 Table 3. Confusion matrices (each bounded by dashed lines) for MLANN applied to Aqua MODIS
 894 relative to layer identification from CloudSat-CALIPSO, 2008, from the training set. The bold numbers
 895 indicate the percent correct for each matrix.

<u>CloudSat and CALIPSO</u>												
	<u>Snow-free, Day</u>			<u>Snow-free, Night</u>			<u>Snow-cover, Day *</u>			<u>Snow-cover, Night*</u>		
<u>MLANN</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>	<u>SL</u>	<u>ML</u>	<u>Total</u>
Ice SL,%	73.2	11.3	84.5	65.9	12.7	78.6	86.6	7.8	69.5	86.4	9.4	95.8
Ice ML,%	4.6	10.9	15.5	6.0	15.4	21.4	1.8	3.9	30.5	1.5	2.7	4.2
Total, %	77.8	22.2	84.1	71.9	28.1	81.3	88.4	11.7	90.4	87.9	12.1	89.1
# pixels x 10 ³	3,748	1,070	4,818	4,097	1,599	5,696	2,549	390	2,884	11,411	1,500	12,911
Liquid SL, %	76.3	7.4	83.7	79.5	7.9	87.4	82.7	8.2	71.3	87.3	9.0	96.3
Liquid ML, %	3.9	12.4	16.3	3.3	9.3	12.6	2.5	6.6	28.7	1.3	2.4	3.7
Total, %	80.2	19.8	88.7	82.8	17.2	88.8	85.2	14.8	89.3	88.6	11.4	89.7
# pixels x 10 ³	4,502	1,112	5,614	5,297	1,100	6,397	5,647	996	6,643	844	319	3,941

896

897

898

899



900

901

Table 4. Same as Table 3, but for combined liquid and ice results from applying MLANN to the 2009 validation dataset.

<u>CloudSat and CALIPSO</u>												
	<u>Snow-free, Day</u>			<u>Snow-free, Night</u>			<u>Snow-cover, Day</u>			<u>Snow-cover, Night</u>		
<u>MLANN</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>	<u>SL</u>	<u>ML</u>	<u>Total</u>
SL, %	75.3	8.9	84.2	74.0	10.1	84.1	83.6	8.4	92.0	86.2	9.8	96.0
ML, %	4.1	11.7	15.8	4.3	11.6	15.9	2.3	5.7	8.0	1.5	2.5	4.0
Total, %	79.4	20.6	87.0	77.3	21.7	85.6	85.9	14.1	89.3	87.7	12.3	88.7
# pixels x 10 ³	28,883	7,493	36,376	33,739	9,908	43,647	8,235	1,352	9,587	14,277	2,002	16,279

902

903

904

905

906



Algorithm	FC/ACC	PR	RC	CoS	NGA
MYD06 C6.1, $\tau > 5$	67	54	46	73	2.2
POLDER 95, $\tau > 5$	70	58	47	74	4.4
VIIRS, $\tau > 1$ ($\tau < 1$)	-	65 (53)	65 (53)	79 (64)	-
Himawari Training Day	85	81	72	87	18.3
Himawari Validation Day	-	70	-	89	-
MLANN day	87	74	55	90	7.6
Himawari Training All	79	73	64	82	14.3
Himawari Validation All	-	64	-	85	-
MLANN Night	86	72	52	88	6.5

907

908 **Table 5.** Confusion matrix metrics in % for various multilayer algorithms. MYD06 and POLDER 95 are based on Table 4 of
 909 Desmons et al. (2017). Himawari results are based on Tan et al. (2022) random forest results. MLANN results based on the
 910 2009 validation parameters in Table 4. Bold numbers denote the largest value of a given parameter for each section
 911 (separated by dotted line).

912

913



914

915

916

Table 6. Average ML fractions from CC and Aqua MODIS MLANN for JAJO.

Time	CC 2009	MLANN near-nadir, 2009	MLANN all VZA, 2009	MLANN near-nadir, 2013	MLANN all VZA, 2013
Day, ML (%)	15.4	12.1	11.8	12.6	12.3
Night, ML (%)	17.7	12.6	13.5	12.8	14.1
Day, L/O	0.77	0.64	0.65	0.63	0.64
Night, L/O	0.84	0.74	0.75	0.76	0.77

917

918

919

920

921

922

923

924

925 **Figures**

926

927

928

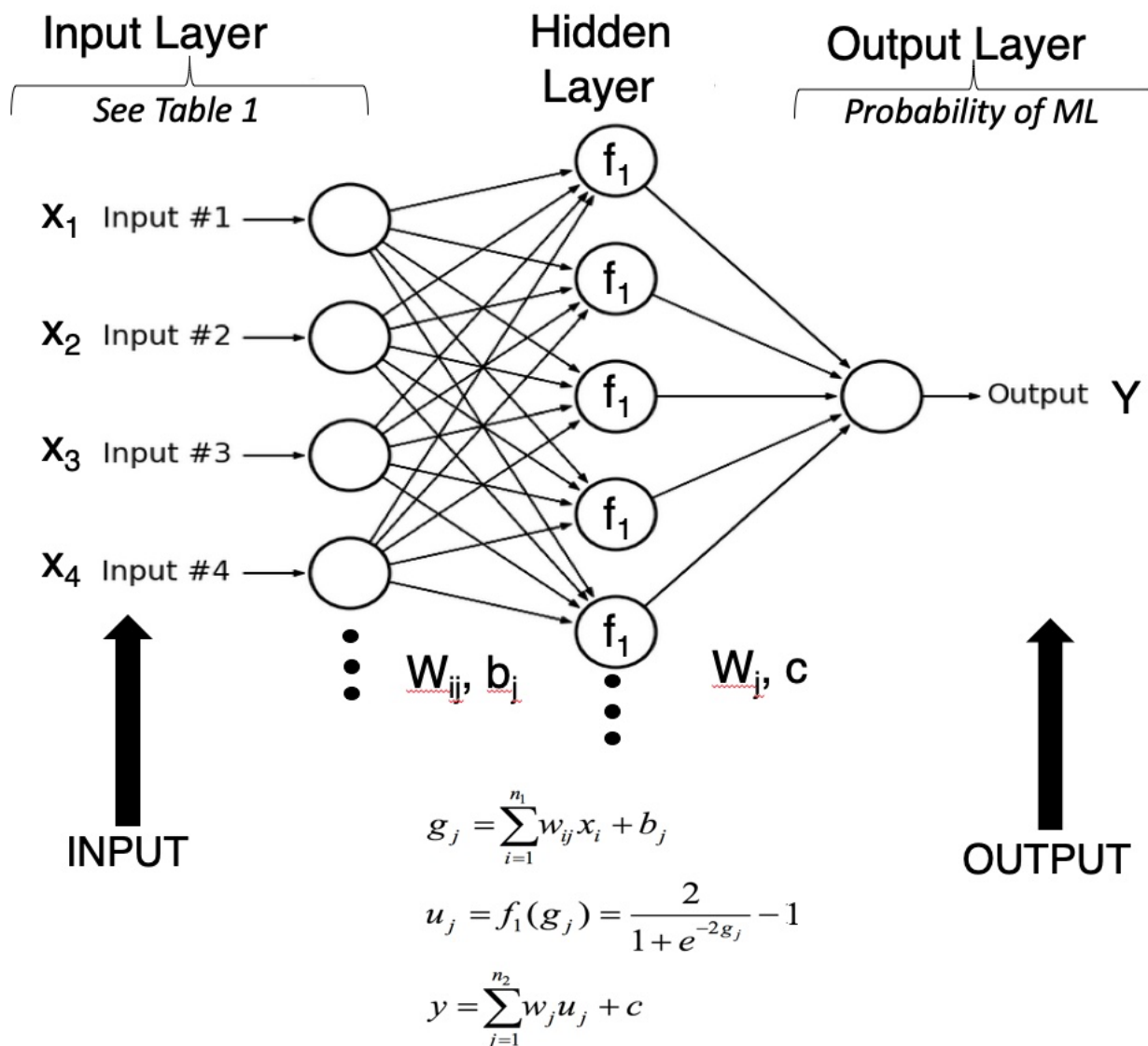
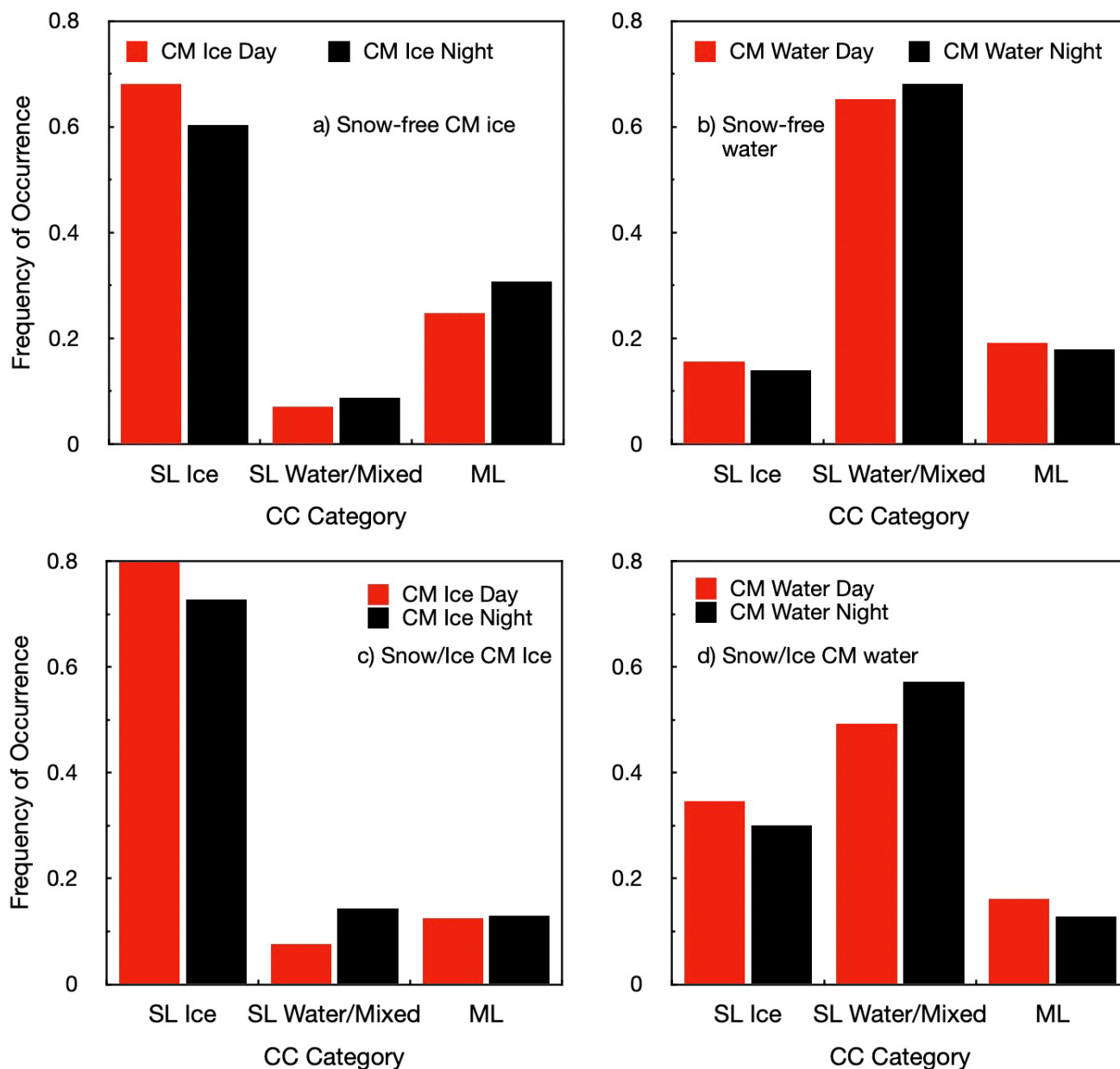


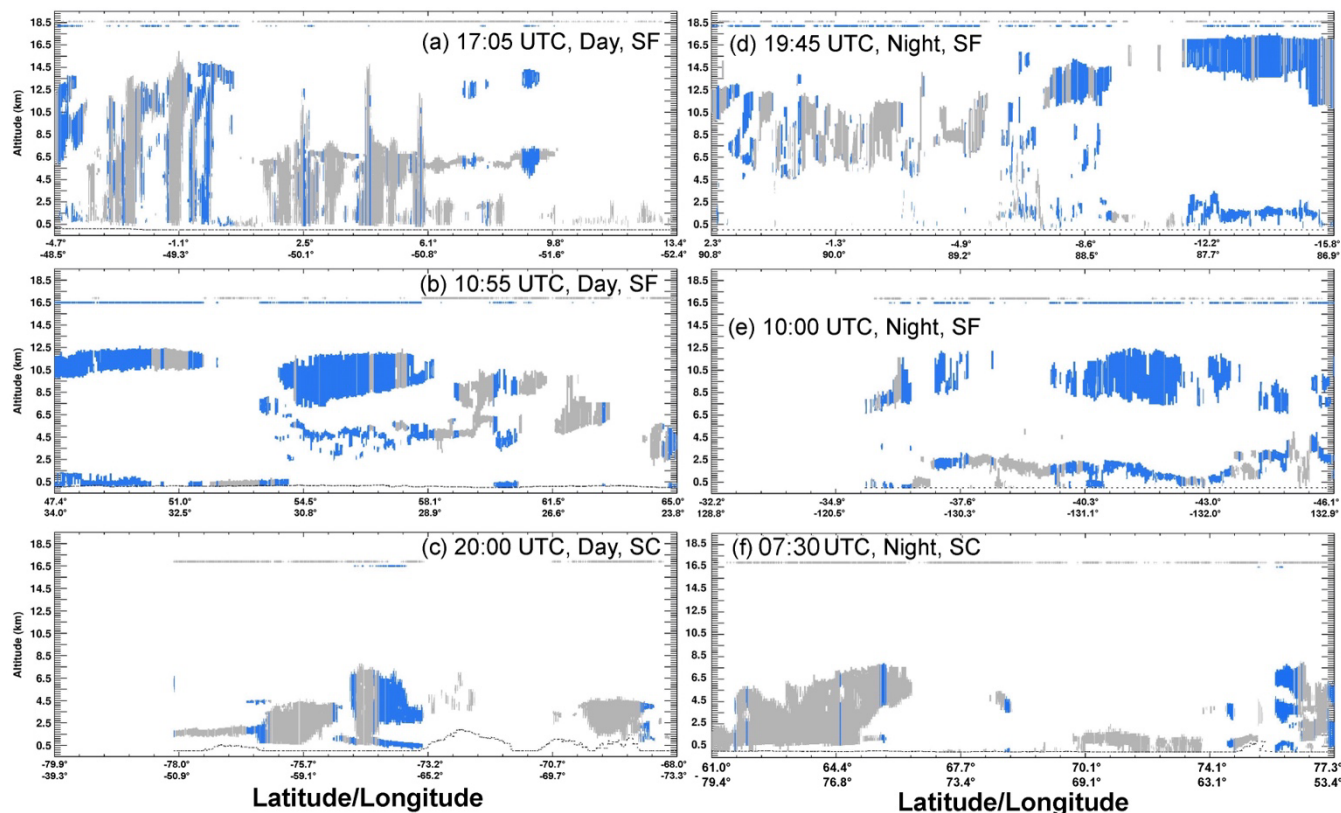
Figure 1. Schematic of artificial neural network used here.

929
 930
 931
 932
 933



934
 935
 936
 937
 938

Figure 2. CALIPSO-CloudSat classification frequency of occurrence for matched 2008 CERES-MODIS cloud phase selection, ice (left) and liquid (right) for snow-free surfaces (top) and snow/ice-covered surfaces (bottom).



939

940

941

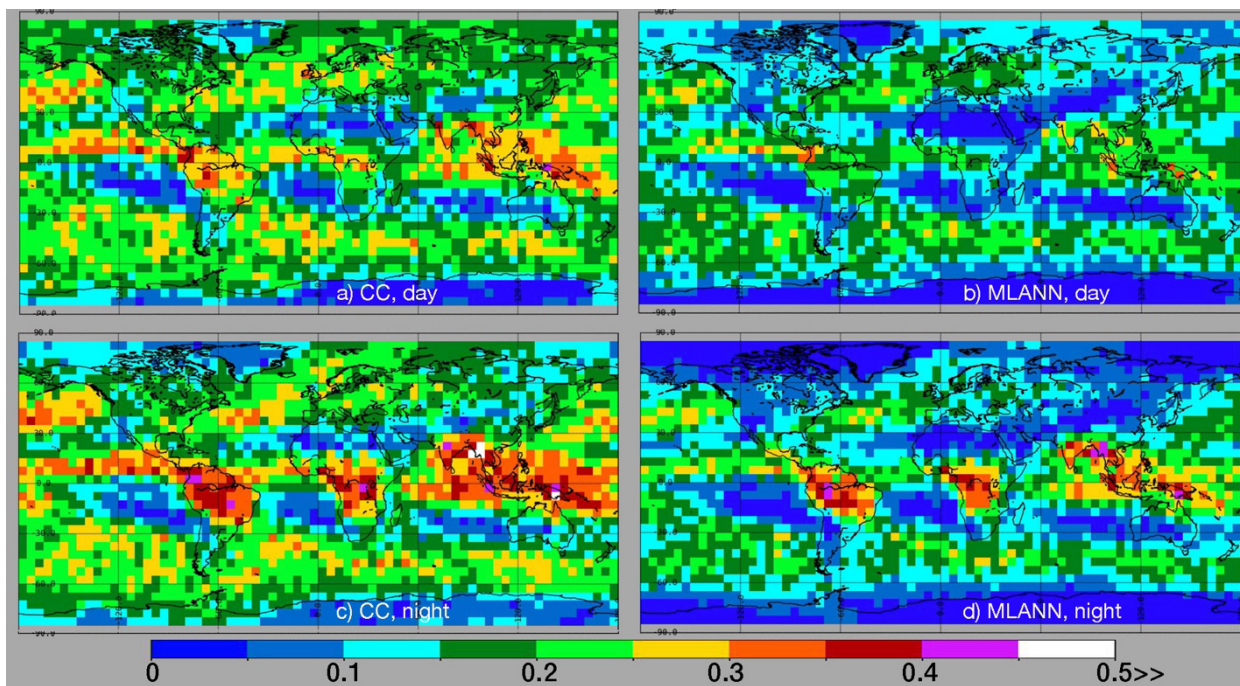
942

943

944

945

Figure 3. CALIPSO-CloudSat cloud profiles from C3M for 25 December 2009 with CC ML clouds indicated in blue and CC SL denoted in gray. The MLANN ML identification for each profile is indicated as a blue dot at the top of each figure. MLANN SL clouds are indicated with a gray dot. Surface elevation is given as the dotted line at the bottom of each panel. Tropical, midlatitude, and polar cloud profiles are given in the top, middle, and bottom profiles, respectively.



946

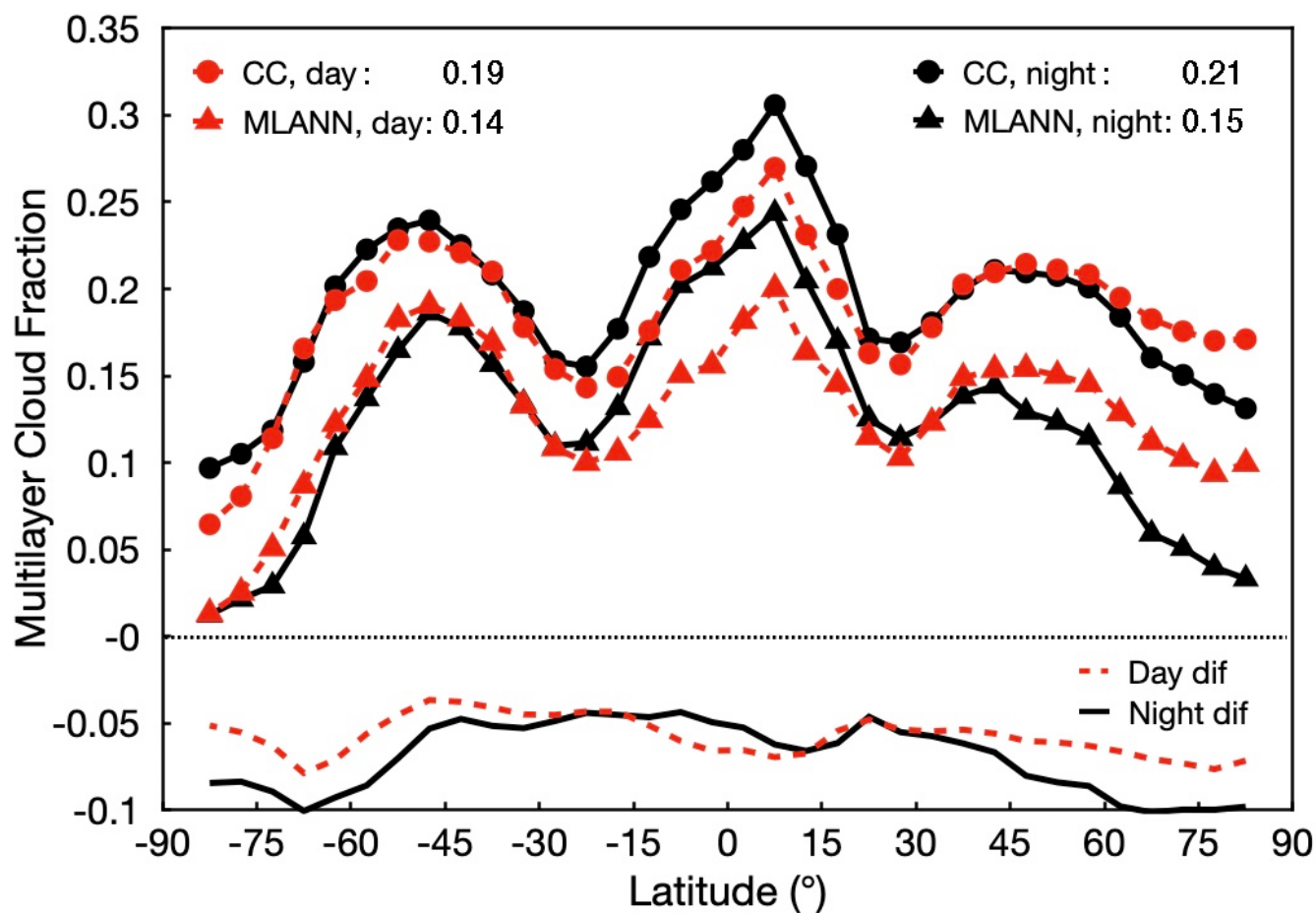
947

948

949

950

Figure 4. Fraction of matched 2009 CC and Aqua MODIS pixels classified as multilayer clouds. CALIPSO-CloudSat and Aqua MODIS ML classifications on left and right, respectively. Day and night pixels on top and bottom, respectively.

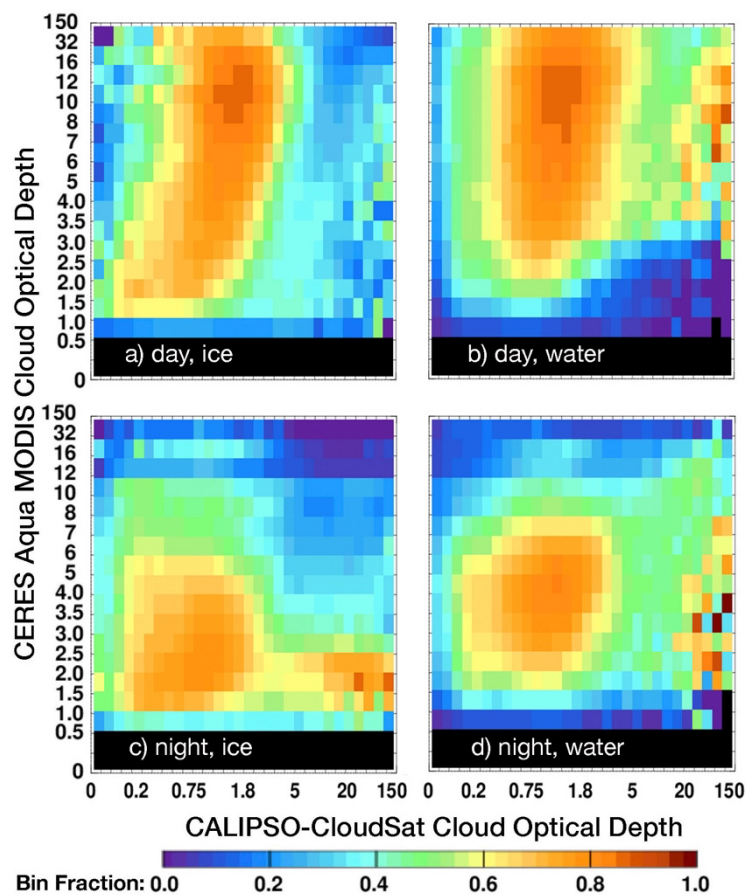


951

952 **Figure 5.** Zonal mean 2009 ML cloud fraction from matched CALIPSO-CloudSat and Aqua MODIS as in Fig. 4. Zonal differences,
953 MLANN - CC, are also plotted. Global averages are indicated in the legend.

954

955



956

957

958

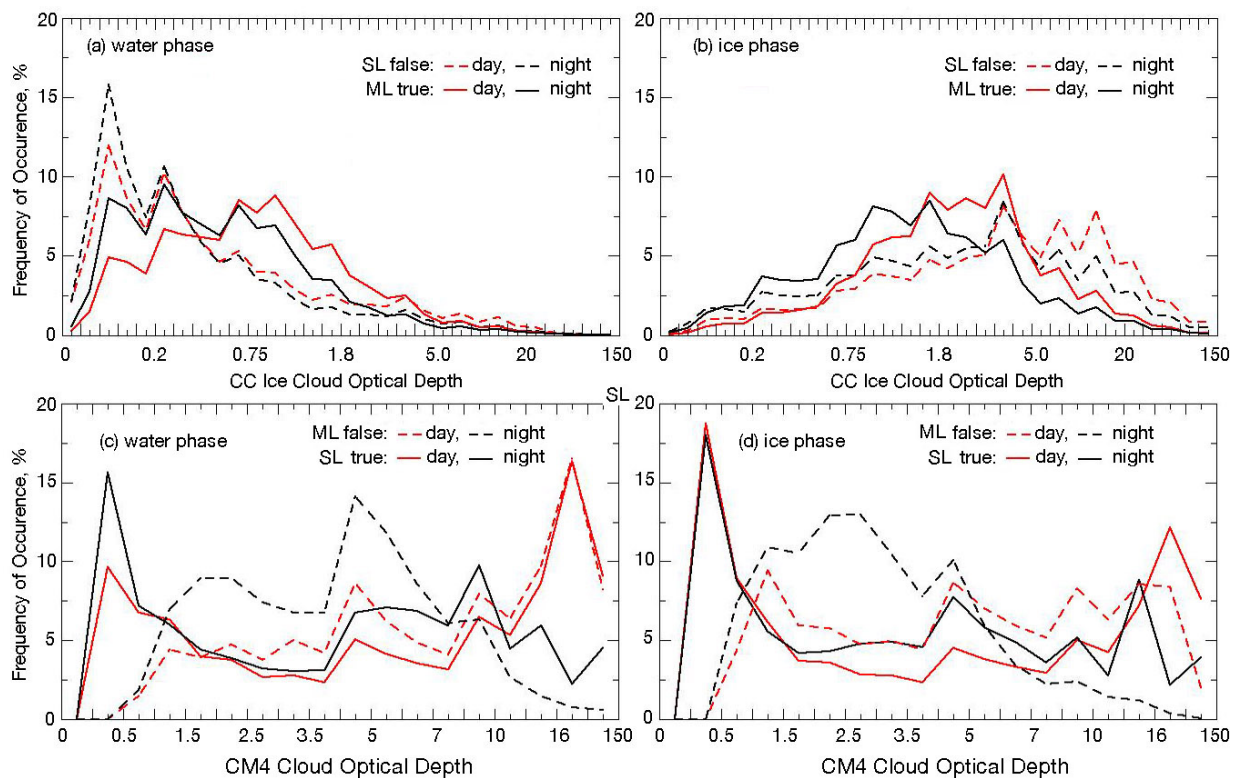
959

960

961

962

Figure 6. Recall or fraction of ML clouds detected within a given CC and CM cloud optical depth bin, 2009. Note, the irregular axis scales. The tick marks for the x-axis are 0, 0.0025, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 0.9, 1.1, 1.3, 1.5, 1.8, 2.1, 2.5, 3, 4, 5, 6, 8, 10, 15, 20, 30, 40, 60, 80, and 150.



963

964

Figure 7. Probability distributions of 2009 false SL and true ML clouds from MLANN as functions of upper-layer cloud optical

965

depth over SF surfaces for MODIS (a) water phase and (b) ice phase. Probability distributions of 2009 false ML and true SL

966

clouds from MLANN as functions of total column cloud optical depth over SF surfaces for MODIS (c) water phase and (d) ice

967

phase. The major tick marks for the x-axes on the top panels are 0, 0.0025, 0.05, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.75, 0.9, 1.1,

968

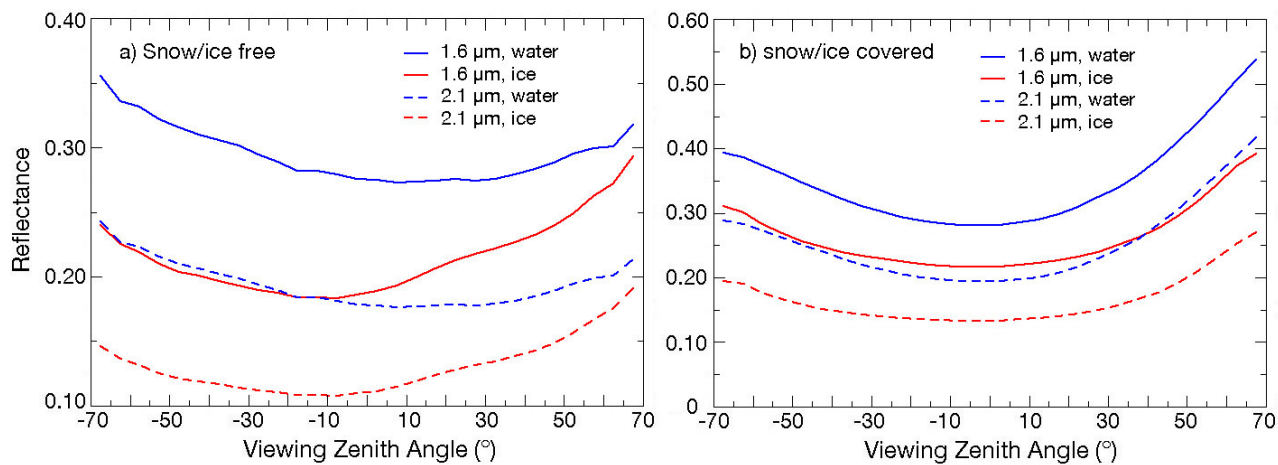
1.3, 1.5, 1.8, 2.1, 2.5, 3, 4, 5, 6, 8, 10, 15, 20, 30, 40, 60, 80, and 150. The major tick marks for the x-axes on the bottom panels

969

are 0, 0.025, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4, 5, 6, 7, 8, 10, 12, 16, 32, and 150.

970

971



972

973

Figure 8. Mean reflectance from Aqua MODIS as a function of VZA for CERES water and ice-phase clouds, JAJO 2019.

974

975

976

977

978

979

980

981

982

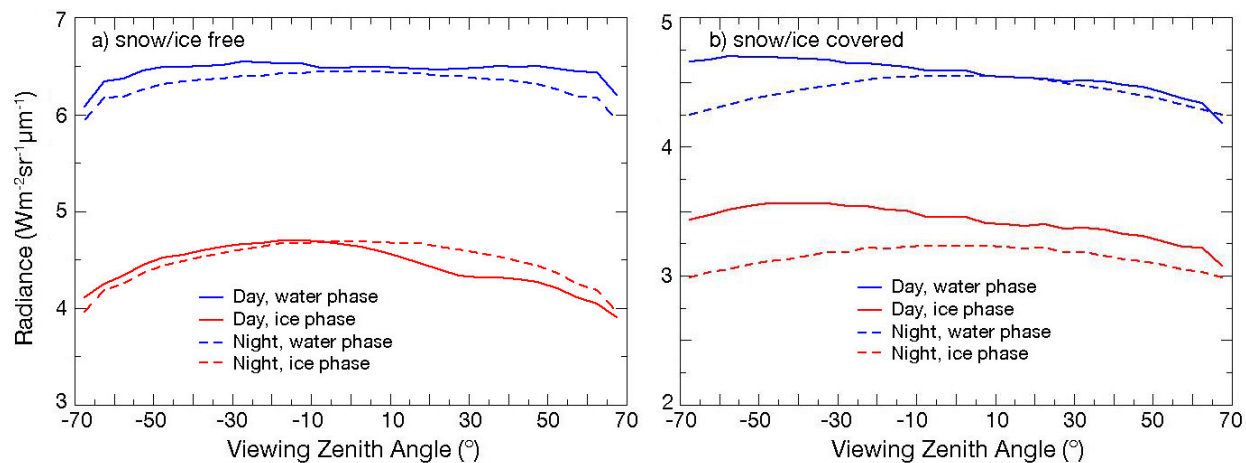
983

984

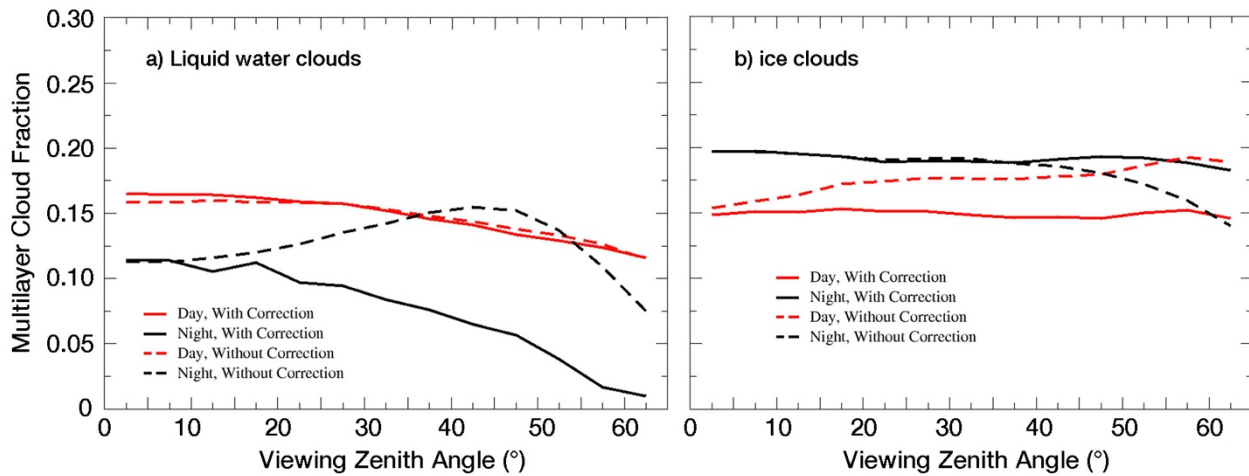
985

986

987



988
989 **Figure 9.** Mean 10.8- μm radiance from Aqua MODIS as a function of VZA for CERES water and ice-phase clouds, JAJO 2019.
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003



1004
1005 **Figure 10.** Mean MLANN multilayer cloud fraction from Aqua MODIS as a function viewing zenith angle, April 2009. MLANN
1006 was run with the MODIS data as observed (without correction) and after applying a VZA correction (with correction).

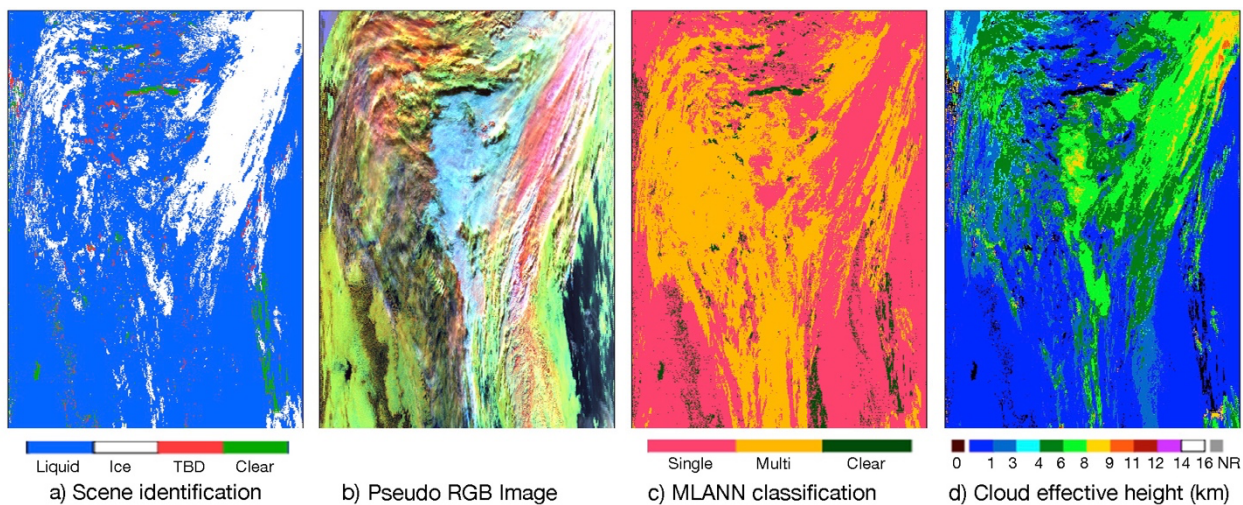


Figure 11. Cloud parameters derived from Aqua MODIS data between 62°S (top) and 52°S (bottom) around 165°E, at ~3:50 UTC, 16 April 2019. (a) CM4 pixel scene classification, (b) Pseudocolor RGB image, red: 0.64 μm reflectance, green: BT₃₇, green; blue: reverse BT₁₁. (c) MLANN classification, and (d) CM4 cloud effective height.

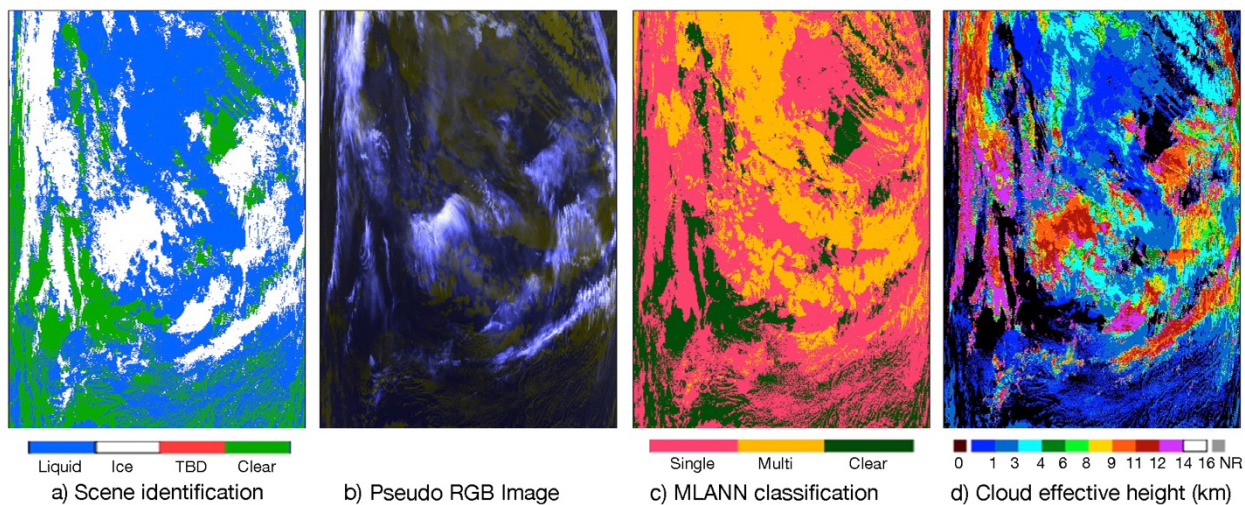


Figure 12. Cloud parameters derived from Aqua MODIS data between 42°N (top) and 24°N (bottom) around 50°W, at ~4:45 UTC, 16 April 2019. (a) CM4 pixel scene classification, (b) Pseudocolor RGB image, red: reverse BT₁₁, green: reverse BT₁₂, green; blue: BTD₃₇₁₁. (c) MLANN classification, and (d) CM4 cloud effective height.

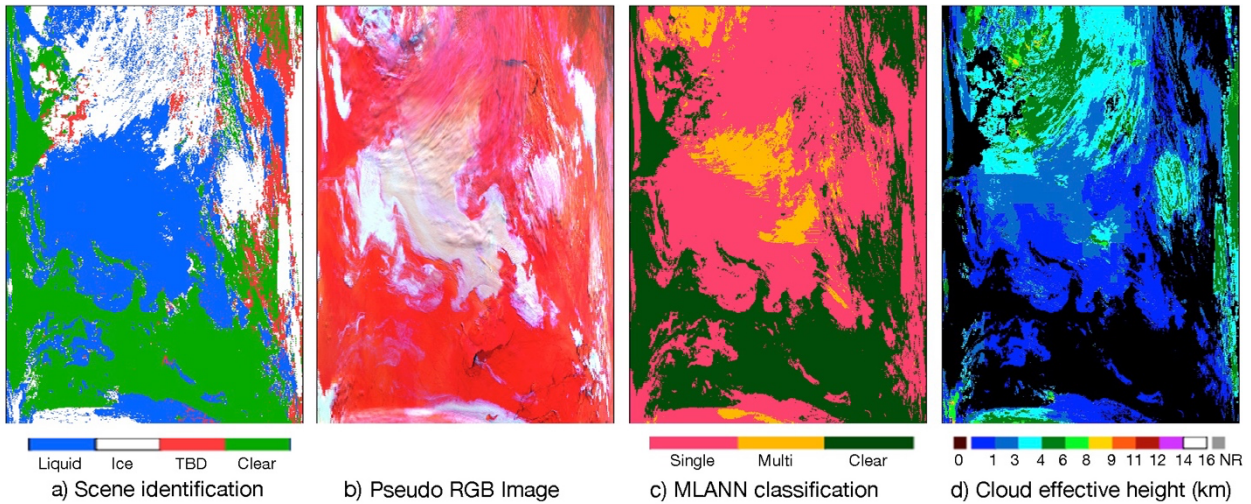
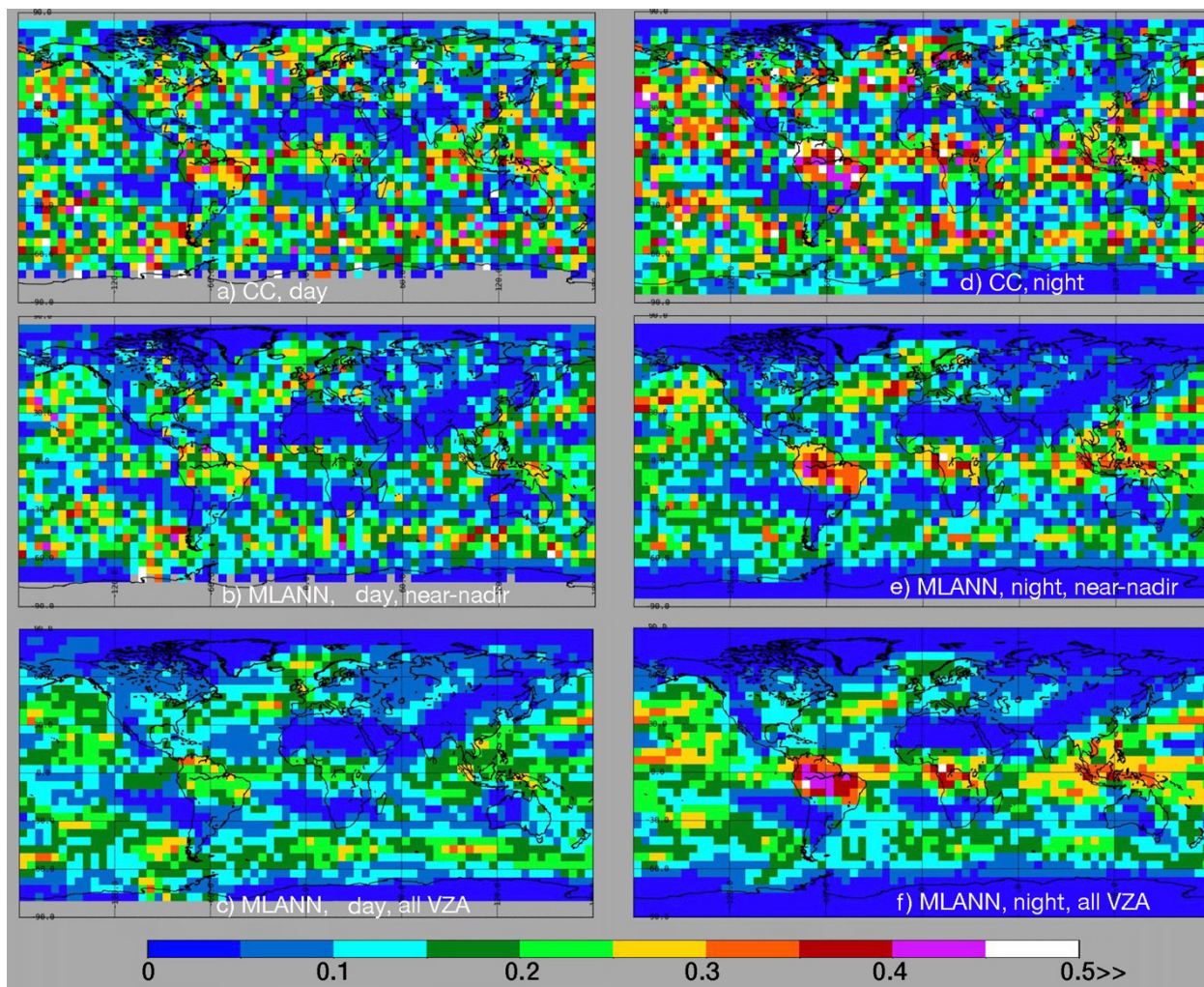
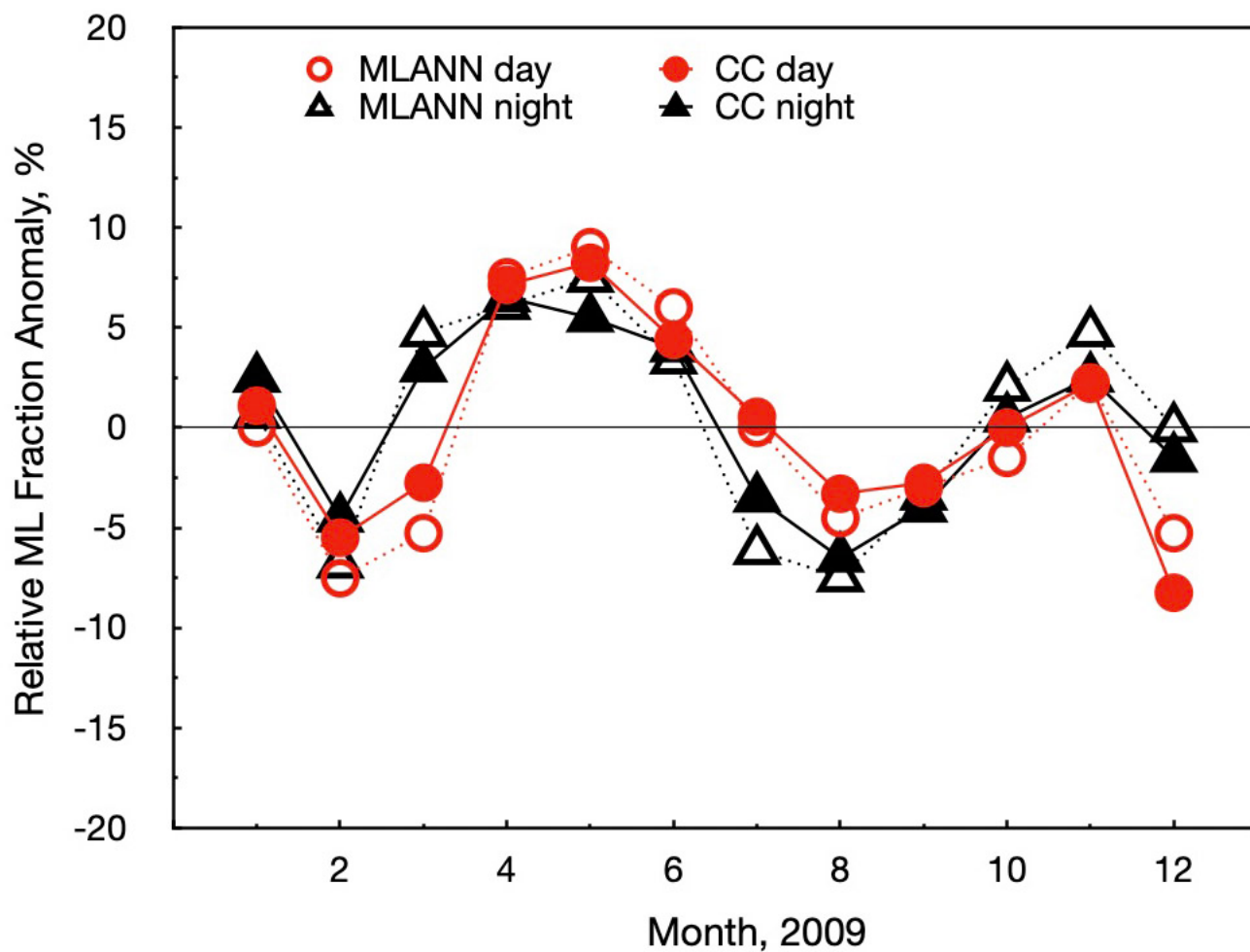


Figure 13. Cloud parameters derived from Aqua MODIS data between 77°N (top) and 83°N (bottom) around 155°E, at ~1:50 UTC, 16 April 2019. (a) CM4 pixel scene classification, (b) Pseudocolor RGB image, red: 0.64 μm reflectance, green: BT₃₇, green; blue: reverse BT₁₁. (c) MLANN classification, and (d) CM4 cloud effective height.



1050
1051
1052
1053
1054

Figure 14. Multilayer fraction of total cloud cover for JAJO 2009 using all CC data from 2009 (top), and using Aqua MODIS MLANN retrievals (middle) at near-nadir ($-18^\circ < \text{VZA} < 3^\circ$), and (bottom) for all VZAs. Daytime on left, nighttime on right.



1055

1056

1057

Figure 15. Monthly mean anomaly of multilayer fraction relative to total cloud cover for 2009 using all CC data and full-swath MODIS data.