JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 106, NO. D12, PAGES 12,603–12,613, JUNE 27, 2001

A further study of the method for estimation of SAGE II opaque cloud occurrence

Pi-Huan Wang,[1] Robert E. Veiga,[2] Lelia B. Vann,[3]
Patrick Minnis,[3] and Geoffrey S. Kent[1]

Abstract. Information on vertical cloud distribution is important to atmospheric radiative calculation, general circulation modeling, and climate study. The method used for estimating the vertical structure of opaque cloud occurrence from the solar occultation observations obtained by the Stratospheric Aerosol and Gas Experiment (SAGE) II has been reviewed for further understanding of the nature of the derived cloud statistics. Most importantly, based on the SAGE II tropical observations (1985–1998), the present study illustrates that the derived opaque cloud occurrence at a given altitude is generally independent of the cloud occurrence at other altitudes, except for some anticorrelation between high-level (12.5 km) and low-level (1–3 km) clouds. This feature of the layer cloud frequency independence is also evident when regional data over the Pacific warm pool and the eastern Pacific are examined. The independent information of the layer cloud frequency is significant and makes it possible to use the derived vertical distribution of cloud occurrence to estimate the probability of multilayer clouds. The limitation is that it is difficult to determine how frequently the multilayer clouds are actually overlapping or how frequently thick cloud (>1 km) really occurs based on the SAGE II observations alone. A discussion of the SAGE II tropical opaque cloud occurrence in relation to the cloud climatology based on visual observations from surface stations and ships, the International Satellite Cloud Climatology Project data, and the cloud statistics using rawinsonde records is also provided.

1. Introduction

Satellite remote sensing employing the solar occultation technique is extremely sensitive to the presence of clouds. This high sensitivity is a consequence of the relatively long atmospheric limb tangent viewing path in contrast to the vertical path length of a nadir-viewing satellite instrument [e.g., *Wylie and Wang*, 1997]. The solar occultation observations from the Stratospheric Aerosol and Gas Experiment (SAGE) II have been used to infer the global distributions of two general groups of clouds, namely opaque and subvisual clouds [*Wang et al.*, 1996]. The presence of opaque clouds terminates the profiling of the SAGE II instrument [*McCormick et al.*, 1979]. Therefore the information about opaque cloud occurrence is imbedded implicitly in the SAGE II measurements. The SAGE II subvisual cloud extinction coefficients are within the measurement range of the SAGE II instrument. The capability of the SAGE II instrument for measuring subvisual clouds can be easily understood because aerosols and clouds are interrelated through microphysical processes [e.g., *Pruppacher and Klett*, 1978; *Wang et al.*, 1994].

When the atmosphere is not heavily influenced by volcanic aerosols, the termination of the profile of the SAGE II 1.02-μm extinction coefficient above the surface is clearly an indication of the presence of an opaque cloud. Because of frequent opaque cloud occurrence, the satellite solar occultation sampling events over a given area and time period are generally reduced with decreasing altitude. *Wang et al.* [1995] presented a method that can be used to derive the occurrence of SAGE II opaque clouds as a function of altitude on a statistical basis. The identification of subvisual clouds from the SAGE II observations is a more involved process. Perhaps the most reliable method for identification of subvisual clouds to date is the two-wavelength method developed by *Kent et al.* [1993]. On the basis of the distribution of the SAGE II extinction coefficients at 0.525 and 1.02 μm at a given altitude over a certain geographic region on a seasonal basis, the two-wavelength method reveals generally two distinct data groups corresponding separately to aerosols and clouds. In brief, the two-wavelength method separates aerosols from clouds based essentially on the magnitude of the measured extinction coefficient as well as the particle size.

As cloudiness is one of the most important climate-related parameters, and the solar occultation technique allows for highly resolved vertical observations (1-km resolution), the SAGE II derived global scale cloud occurrence can be used as input for atmospheric radiative and general circulation model investigations, as well as for climate change studies. The objective of the present study is to provide a thorough review of the method that has been used for deriving the vertical distribution of SAGE II opaque clouds and to further investigate and understand the nature of the derived cloud occurrence from the solar occultation satellite instrument. In particular, the independence of the derived cloud occurrence in a given layer relative to cloud occurrence in the other layers will be examined closely by using the SAGE II observations. Furthermore, the usefulness of the derived layer cloud occurrence for inferring multilayer cloud probability will be investigated. Section 2 contains a brief description of the relevant SAGE II observational features. The opaque cloud occurrence derivation is reviewed in section 3, followed by the observational data analysis of the derivation in section 4, and discussion in section 5. The summary is given in section 6.

[1]Science and Technology Corporation, Hampton, Virginia.
[2]Science Applications International Corporation, Hampton, Virginia.
[3]Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia.

Copyright 2001 by the American Geophysical Union.

Paper number 2001JD900138.
0148-0227/01/2001JD900138$09.00

2. SAGE II Features

The SAGE II satellite instrument is a seven-channel radiometer [*Mauldin et al.*, 1985; *McCormick*, 1987]. Using the solar occultation technique, the instrument measures the profile of attenuated solar radiation as a function of tangent height during spacecraft sunrises and sunsets and is capable of providing self-calibrated atmospheric limb transmission. The seven SAGE II channels are centered at 0.385-, 0.448-, 0.453-, 0.525-, 0.600-, 0.940-, and 1.02-μm wavelengths. SAGE II slant path transmission data are used to derive ozone, nitrogen dioxide, and water vapor vertical distributions, as well as the particulate extinction coefficient profile at 0.385, 0.453, 0.525, and 1.02 μm. *Chu et al.* [1989] have discussed the SAGE II data inversion algorithm in detail. The SAGE II sensor has been operating since October 1984 on the Earth Radiation Budget Satellite (ERBS) and is still providing measurements at the time this report was prepared.

The orbital characteristics of ERBS are such that the SAGE II sunrise measurement provides a latitudinal coverage of ~135° in ~1 month. The latitudinal extremes of the measurements vary with the season. With a 90-min orbiting period the 15 daily sunrise measurements are distributed almost evenly in longitude with a small latitudinal shift between successive events. These sampling features are also true for the sunset measurements. The instrument field of view (FOV) at the tangent point is specified by a coverage of 0.5 km in the vertical by 2.5 km in the horizontal. The SAGE II retrieval uses an onion-peeling type of inversion technique, assuming a local spherically symmetric atmosphere with 1-km-thick concentric shells [*Chu et al.*, 1989]. In such an atmospheric shell configuration the part of the limb viewing path that traverses the tangential shell (the shell centered at the tangent) is ~200 km in length, corresponding to a possible horizontal error of 100 km.

The upper limit of the SAGE II optical path length along an entire occultation tangential path is about 6 at 1.02-μm wavelength. The molecular scattering and particulates (aerosols and clouds) contribute to the observed optical path length. The molecular contribution is rather small at 1.02-μm wavelength. Thus, when the atmosphere is not severely disturbed by volcanic aerosols, clouds along the tangential path with a total optical path length greater than 6 would terminate the SAGE II sampling event. Unfortunately, the SAGE II observation does not specify the size and location of the detected opaque cloud along the line of sight. Assuming that the cloud is uniformly distributed in the 200-km tangential shell, the optical path length upper limit of 6 translates to an extinction coefficient of 0.03 km^{-1} at 1.02 μm. This value of the extinction coefficient corresponds to an optical depth of 0.03 for a 1-km-thick cloud. According to the *Sassen and Cho* [1992] cloud classification, this value is close to the optical depth separating subvisual from thin cirrus clouds. Thus the SAGE II opaque clouds generally include all types of clouds, except optically thin subvisual clouds [*Wang et al.*, 1996]. Even if the length of the opaque cloud is half the length of the 200-km tangential shell, the optical depth is still as small as 0.06. Clouds with such a small optical depth would be extremely difficult to detect by most nadir-viewing satellite instruments, from surface visual cloud observations, or by analysis of rawinsonde data.

It should be noted that there is also the possibility that the opaque cloud may not be located in the tangential shell implied by the retrieval. In other words, the cloud could be located in higher shells along the occultation tangential viewing path. The worst case would be a cloud located at an altitude of 17 km along the lowest tangential viewing path. Because the SAGE II observation does not specify the size and location of the detected cloud, the retrieval would give an impression that the cloud is centered at the tangential point, leading to a vertical uncertainty of 17 km and a horizontal uncertainty of ~450 km at the surface. For the same reason the SAGE II retrieval would underestimate the high-altitude cloud occurrence probability and overestimate low-altitude cloud occurrence probability. In the atmosphere, there are generally more low-altitude clouds than high-altitude clouds [*J. Wang et al.*, 2000, Figure 6]. Thus this type of uncertainty would decrease as altitude increases. *Wang et al.* [1994] have discussed the constraints of the SAGE II cloud observations in detail.

The SAGE II derived cloud occurrence has been compared with the surface weather observations (SWOBS) of *Warren et al.* [1986, 1988; *Wang et al.*, 1995], the cloud data product of the International Satellite Cloud Climatology Project (ISCCP) [*Liao et al.*, 1995a, 1995b], and more recently with the cloud measurements from the High-Resolution Infrared Radiometer Sounder (HIRS) [*Wylie and Wang*, 1997, 1999], and from the cryogenic limb array etalon spectrometer (CLAES) [*Mergenthaler et al.*, 1999]. In general, the SAGE II cloud frequency is greater than other climatologies because of its high sensitivity to cloud presence resulting mainly from the long occultational viewing path. The SAGE II cloud statistics in the tropopause region have been used to study the effect of heterogeneous reactions on ozone chemistry [*Solomon et al.*, 1997], and the controlling mechanisms of the formation of subvisual clouds [*Gierens et al.*, 2001]. Furthermore, the SAGE II multiyear cloud data set is an invaluable asset for studying the long-term behavior of clouds and for estimating changes in cloud outgoing longwave radiative forcing [*Wang et al.*, 1999; *P.-H. Wang et al.*, 2000]

3. Cloud Probability Density Function

For a given sampling period, let $N(z)$ denote the number of times for which the SAGE II solar occultation observation provides extinction coefficient (1.02-μm) measurements in a layer centered at altitude z. $N(z)$ is also known as the penetration count of the satellite measurement in the given layer and is a measure of the clear-sky frequency. Thus it can be used to estimate the probability of cloud occurrence below the given layer. Because opaque clouds block the FOV, no 1.02-μm measureable information is available below the layer if an opaque cloud is encountered in the next lower layer (centered at altitude $z-1$) during a given sampling event. It is noted that to estimate the cloud probability, the sample sizes must be considered carefully at each altitude layer.

To estimate the probability of the clear-sky frequency or no opaque cloud in a layer at altitude z, the ratio of penetration count of the layer at z to that of the next higher layer centered at $z+1$ can be used. By definition of $N(z)$ it is apparent that $0 \le N(z) \le N(z+1)$, because all information below altitude $z+1$ is unavailable once an opaque cloud is encountered in the layer at z. Symbolically, we write the probability of no opaque cloud for a layer at altitude z as

$$p(z)=\frac{N(z)}{N(z+1)}.$$

Thus $1-p(z)$ gives the probability that an opaque cloud occurs in a layer centered at altitude z, and can be written as

$$f(z)=\frac{N(z+1)-N(z)}{N(z+1)}. \quad (1)$$

The above expression of $f(z)$ is referred to as the layer cloud probability or layer cloud occurrence [*Wang et al.*, 1995]. Note the numerator of $f(z)$ is just the reduction in the clear-sky sampling number that occurs in the layer centered at z with respect to the sampling size of the layer centered at $z+1$. Obviously, this reduction in the clear-sky sampling size corresponds precisely to the number of the SAGE II measurements that encountered opaque clouds in the layer centered at z. Keep in mind that the opaque cloud probability is estimated by normalizing the cloud count of a

given layer to the number of samples reaching the layer, and that this concept of the cloud occurrence derivation can be applied to any thick atmospheric layers.

Wang et al. [1995] have also discussed the lower and upper bounds of the opaque cloud occurrence estimate. The expressions for the lower and upper bounds of opaque cloud probability are

$$f^l(z)=\frac{N(z+1)-N(z)}{N_t}, \quad (2)$$

$$f^u(z)=\frac{N_t-N(z)}{N_t}, \quad (3)$$

respectively. The superscripts l and u indicate the lower and upper limits, respectively. The numerator in $f^l(z)$ is the opaque cloud count, and is the same as in $f(z)$. However, to determine the cloud probability, the cloud count is normalized to the total overpass N_t of the satellite instrument, as if no cloud occurred above the layer of interest. As a result, $f^l(z)$ underestimates the cloud probability and is therefore a lower limit. The $f^u(z)$ value attributes all events that fail to reach the layer at z, i.e., $N_t - N(z)$, to the opaque clouds that occurred in the layer, even when no clouds occurred in the particular layer considered during the sampling period. Therefore $f^u(z)$ corresponds to a situation in which any opaque cloud encountered by SAGE II at an altitude above z would extend all the way down to the layer z under study, i.e., extremely thick opaque clouds. For this reason, $f^u(z)$ overestimates the cloud probability, especially at lower altitudes, and is thus an upper limit of opaque cloud probability.

In a sense, $f^u(z)$ provides a cumulative estimate of all opaque clouds that occur above z, because the numerator of $f^u(z)$ is the number of events that fail to reach the layer at z due to opaque clouds in all layers at z and above. At the surface, $f^u(z)$ would thus provide an estimate of total cloud probability, corresponding closely to surface observations. It can be shown that $f^u(z)$ is the sum of all the lower estimates given by $f^l(z)$ in that layer and the layers at higher altitudes. Because $f(z)$ gives greater cloud occurrence than the lower estimate of $f^l(z)$, the total cloud above a given altitude integrated from $f(z)$ is greater than the cumulative estimate of $f^u(z)$ at that altitude. We will return to discuss this feature in section 5. To distinguish $f^u(z)$ from $f(z)$, $f^u(z)$ will be referred to as the cumulative cloud probability. In light of the above discussion, $f(z)$ can be regarded as the best estimate of the opaque cloud occurrence [*Wang et al.*, 1995].

Interestingly, the SAGE II cloud count is equivalent to a binomial experiment with the frequency $f(z)$ as the binomial parameter. To determine the uncertainty of the derived frequency, the 95% confidence interval of the binomial parameter can be calculated by using quantities of the F distribution. The formula for the corresponding lower and upper bound of the confidence interval given by *Johnson and Kotz* [1969] is employed in the present study.

4. Data Analysis

To further explore and understand the nature of the cloud occurrence derivation, we utilize the SAGE II 1.02-μm multiyear observations from 1985 to 1998 in the tropics (20°S–20°N) on a seasonal basis. It is noted that data are excluded from this investigation if they were obtained during the period having high volcanic aerosol loading from the June 1991 Pinatubo volcanic eruption to the end of 1993. Without those data a total of 45 individual seasons with 15,295 SAGE II overpasses are available. The vertical distributions of the SAGE II sampling event for the 45 individual seasons are presented in Figure 1 and tabulated in Table 1. As anticipated the seasonal sampling count decreases as the altitude decreases (Figure 1). The 45 seasonal sample count data sets are then used to determine the seasonal vertical cloud occurrence distributions as well as the multiyear mean cloud probability according to $f(z)$, $f^l(z)$, and $f^u(z)$.

The derived multiyear mean vertical distribution of opaque cloud probability according to $f(z)$ along with its error estimate is given in Figure 2. The tropical layer cloud occurrence reveals a local maximum of 11% at ~12.5 km. It decreases rapidly with increasing altitude as it approaches the tropopause. Below the local maximum, the distribution shows a local minimum of 5% at an altitude of 8 to 9 km. In the lower troposphere below 8 km the cloud probability increases sharply toward the surface. Figure 2

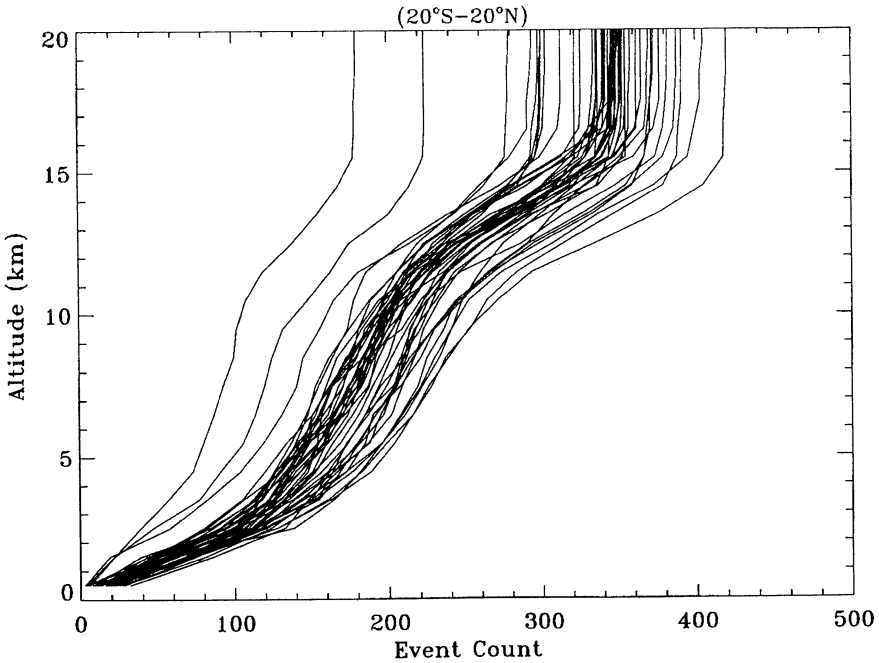

Figure 1. Vertical distribution of the SAGE II event count for 45 individual seasons between 1985 and 1998, except the period from June 1991 to November 1993.

Table 1. Seasonal SAGE II Tropical Event Count as a Function of Altitude (0.5–19.5 km)

DJF8485	10	60	102	122	138	149	163	171	177	181	194	218	229	252	279	311	340	346	346	346
MAM8585	11	50	90	114	129	145	151	163	177	187	202	215	241	271	306	347	361	365	365	365
JJA8585	10	72	119	155	169	181	192	206	210	217	231	243	268	294	329	355	355	355	355	355
SON8585	16	56	109	127	138	148	166	175	183	194	207	224	243	282	314	340	348	348	348	348
DJF8586	10	73	121	154	164	173	187	199	208	215	223	237	254	281	300	319	331	345	348	348
MAM8686	9	39	93	125	138	155	165	174	187	196	205	213	235	255	286	320	344	347	348	350
JJA8686	25	61	114	152	177	196	209	221	230	241	248	260	275	299	322	346	354	355	355	355
SON8686	11	51	101	122	134	145	158	171	176	189	197	211	223	253	290	316	336	340	340	340
DJF8687	19	70	111	132	148	154	170	183	191	192	202	222	235	254	285	319	334	336	336	336
MAM8787	23	59	106	129	151	160	169	175	187	195	211	222	244	277	312	339	348	352	352	352
JJA8787	21	71	114	143	165	180	192	204	209	221	229	238	260	287	318	335	343	344	344	344
SON8787	19	67	104	122	143	161	174	181	188	196	203	222	239	274	315	336	347	349	349	349
DJF8788	10	56	103	126	140	158	171	177	191	203	215	227	248	274	306	329	336	342	342	342
MAM8888	8	54	90	117	133	148	172	187	194	203	212	220	241	270	303	335	346	346	347	347
JJA8888	32	86	133	148	167	187	202	209	217	223	233	246	267	303	336	344	345	346	346	346
SON8888	19	67	113	135	153	164	177	188	192	198	206	217	245	284	313	336	344	345	345	346
DJF8889	16	54	86	113	127	142	170	182	189	200	208	221	240	266	308	329	340	341	341	341
MAM8989	16	47	80	105	120	134	151	162	175	187	198	215	237	285	312	334	349	351	351	351
JJA8989	27	81	124	150	162	182	197	204	212	218	226	239	259	290	321	335	337	337	337	337
SON8989	27	69	106	130	146	158	164	175	184	192	204	221	243	280	302	327	335	336	337	337
DJF8990	21	56	104	117	127	141	156	164	174	183	196	206	225	252	280	313	325	326	326	326
MAM9090	17	58	94	112	127	146	164	177	191	200	216	228	248	279	315	348	362	362	362	362
JJA9090	19	67	112	133	152	169	181	190	200	206	213	229	252	280	310	322	322	322	322	322
SON9090	22	58	107	127	143	155	164	180	186	209	221	239	261	292	337	353	357	357	358	358
DJF9091	18	62	100	115	126	137	147	155	164	174	178	186	209	238	272	299	311	313	313	313
MAM9191	17	63	101	130	146	155	168	180	185	193	205	225	243	279	315	347	352	353	353	353
JJA9191	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SON9191	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DJF9192	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAM9292	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JJA9292	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SON9292	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DJF9293	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAM9393	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
JJA9393	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
SON9393	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DJF9394	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
MAM9494	13	46	102	138	161	174	185	194	202	212	223	233	253	284	320	360	373	377	377	377
JJA9494	10	48	91	118	130	147	161	168	180	185	200	210	232	258	284	294	298	299	299	300
SON9494	9	55	98	135	157	174	185	195	211	214	225	247	287	315	348	374	379	381	382	382
DJF9495	12	43	95	118	130	140	148	153	162	175	188	206	219	241	259	279	291	291	293	294
MAM9595	19	62	117	145	161	174	185	203	214	229	246	269	295	330	358	379	386	387	388	388
JJA9595	28	75	114	150	173	196	208	219	225	236	244	266	292	326	355	367	371	371	373	373
SON9595	4	44	84	114	136	154	157	164	184	197	203	216	237	261	278	293	294	297	298	298
DJF9596	16	51	95	126	136	147	156	163	173	189	206	225	247	272	300	324	331	333	334	334
MAM9696	17	51	99	122	143	159	165	177	187	198	208	236	259	290	312	329	333	336	336	336
JJA9696	21	72	118	139	165	188	195	205	216	229	244	267	299	333	359	367	369	371	372	372
SON9696	5	49	82	106	131	143	153	160	167	181	189	206	224	249	277	296	299	300	300	300
DJF9697	10	55	112	145	173	185	190	208	221	230	248	265	291	308	328	351	367	369	369	370
MAM9797	11	49	93	123	135	145	155	164	179	190	202	212	232	253	282	296	301	302	303	303
JJA9797	21	74	128	161	183	199	216	228	237	253	264	284	315	347	379	388	389	391	391	391
SON9797	3	19	58	82	104	119	131	141	145	156	165	181	214	242	262	276	277	278	278	279
DJF9798	13	51	139	165	182	196	211	222	233	241	252	274	308	337	373	395	399	403	403	405
MAM9898	6	22	39	57	73	80	87	93	100	102	108	119	138	154	168	178	179	179	180	180
JJA9898	22	71	128	163	189	204	216	226	239	252	271	294	336	375	405	418	418	419	420	420
SON9898	4	22	47	77	91	106	114	120	125	132	148	163	175	200	214	223	224	224	224	224

shows the lower estimate of $f'(z)$, the multiyear mean cumulative cloud probability using $f''(z)$, and the normalized multiyear mean event count. The normalized mean event count represents basically the mean tropospheric measurement probability as a function of altitude in the tropics [*Wang*, 1994].

The vertical distribution of tropical opaque cloud probability derived from the layer cloud occurrence $f(z)$ is very interesting. The reason for the maximum at ~12.5 km and minimum at 8–9 km can be very involved, as the cloud formation is governed by many factors including vertical temperature and water vapor distributions, air mass cooling rate, stability, etc. Here we speculate that the decreasing tropospheric temperature with height favors cloud occurrence at high altitudes, while more water vapor near the surface favors cloud development at lower altitudes. As a result,

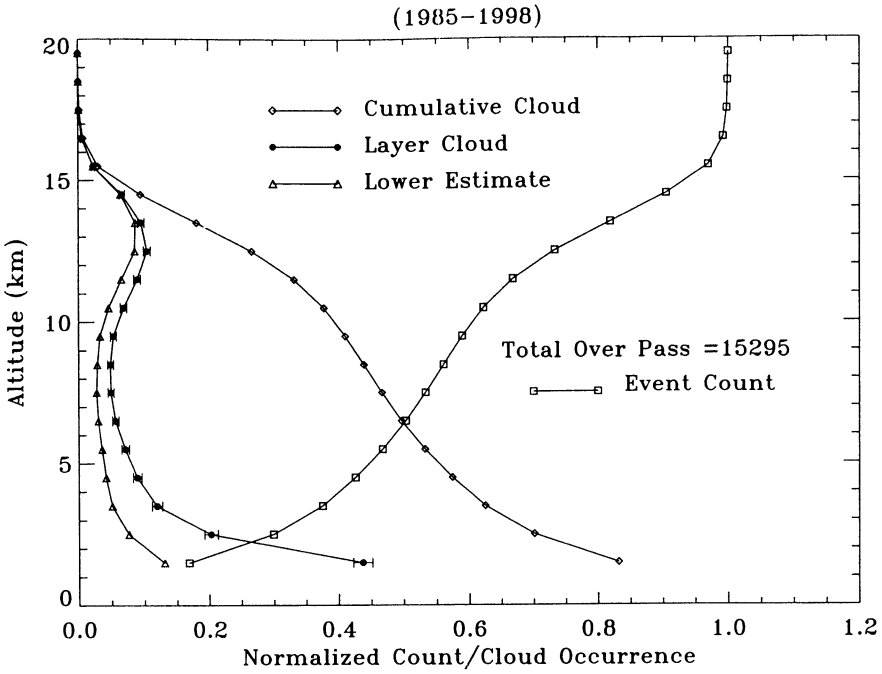

Figure 2. Normalized vertical distribution of multiyear mean SAGE II event count, layer cloud occurrence, cumulative cloud occurrence, and lower estimate of the cloud occurrence. The error bars represent the 95% confidence interval of the layer cloud frequency.

the cloud frequency is at a minimum somewhere in the middle troposphere. The reason for cloud decreases above 12.5 km is likely due to insufficient available water vapor even though the temperature is very cool there. It appears that the maximum at 12.5 km could be related to the frequent occurrence of high-altitude cirrus or cumulonimbus anvils developed through deep convective activities. The convection more often ends at 12.5 km resulting in the formation of more cirrus at that altitude than below. The cirrus is more long-lived than the vertical shaft of the convection, leading to more cloudiness around 12.5 km.

At this point, it is useful and informative to compare the SAGE II tropical cloud observations with other cloud data sets. In doing this we incorporate the SAGE II statistics into the recent study of *J. Wang et al.* [2000], who compared the cloud frequency of rawinsonde observations (RAOBS) with that of the International Satellite Cloud Climatology Project (ISCCP) [*Rossow et al.*, 1996], and the surface weather observations (SWOBS) [*Warren et al.*, 1986, 1988]. Here the SAGE II layer cloud frequency $f(z)$ is calculated for thick layers according to the high-level, middle-level, and low-level clouds defined by *Wang and Rossow* [1995]. The calculated results are then compared with the mean of the land and ocean cloud occurrence of *J. Wang et al.* [2000] in Table 2. One of the remarkable features of the results in Table 2 is the smallest cloud occurrence in the middle-level cloud category in all four cloud observational data sets, despite some detailed quantitative differences. Furthermore, all data sets show the highest occurrence frequency in the low-level cloud, except for the ISCCP data products. It should be noted that besides the different cloud observational techniques, their data periods used in the data intercomparison are also quite different (Table 2).

5. Discussion

The knowledge of the vertical distribution of layer cloud occurrence is important for understanding the interaction of clouds and radiation in the atmosphere, general circulation model validation, and studying cloud climate effects [e.g., *Fowler and Randall*, 1996; *Wang and Rossow*, 1998]. In this regard, the particular shape of tropical opaque cloud probability derived from the layer cloud occurrence $f(z)$ as featured by a local maximum at 12.5 km and a minimum at ~8 to 9 km (Figure 2) is very significant. Because the vertical opaque cloud occurrence distribution is statistically derived from an integrated set of terminated profile measurements, one important question is whether the information of the derived cloud frequency at a given altitude is independent of the cloud frequency at other altitudes. In this section the independence of the layer cloud frequency will be explored first. We then investigate the application of the layer cloud frequency to study the multilayer cloud probability. The regional dependence of the layer cloud frequency independence will be examined last.

5.1. Independence of Layer Cloud Occurrence

To answer the question of whether the information of the layer cloud occurrence $f(z)$ at a given altitude is independent of the cloud occurrence at all other altitudes, we correlate the seasonal cloud occurrence data set of a selected layer to the occurrence data of all layers included in the present study. The derived vertical distribution of the correlation coefficient for cloud frequency calculated by using the layer cloud probability $f(z)$ is displayed in Figure 3 for

Table 2. Probability of Tropical Cloud Occurrence[a]

	SAGE II 1985–1998	SWOBS[b] 1965–1980	RAOBS[b] 1976–1995	ISCCP[b] 1990–1992
High cloud ($z > 8$ km)	44	65	47	52
Middle cloud (3 km $< z <$ 8 km)	33	53	39	37
Low cloud ($z < 3$ km)	88	70	77	49

[a]Probability is in percent.
[b]*J. Wang et al.* [2000]

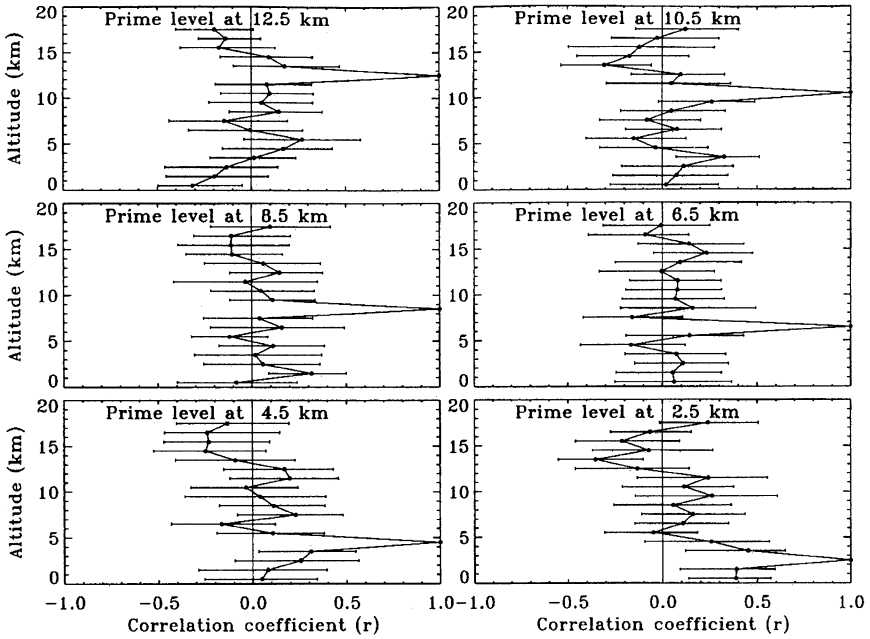

Figure 3. Correlation analysis for selected prime cloud layers for layer cloud occurrence $f(z)$. The error bars represent the 95% confidence interval of the correlation coefficient.

six selected prime cloud layers centered at altitudes between 2.5 and 12.5 km. To estimate the uncertainty of the calculated correlation coefficient, we have used the bootstrap method [*Efron*, 1985] to determine the 95% confidence interval. One of the distinct features of the resultant correlation profile is the sharp peak at the altitude of the prime cloud layer (Figure 3). In addition, the distribution of the correlation coefficient at all altitudes other than the prime layer appears to be random in nature, except for the negative correlation between the clouds at 12.5 km and 1–3 km. The correlation analysis based on the cumulative cloud probability $f^u(z)$ is presented in Figure 4. The results show a drastically different correlation coefficient profile. It reveals a pattern that is roughly similar to a normal distribution with a rather broad width.

The results given in Figures 3 and 4 indicate that the layer cloud occurrence determined according to $f(z)$ is very much independent of the cloud occurrence in other layers. In contrast, the cumulative cloud occurrence estimate derived from $f^u(z)$ correlates heavily with the cumulative cloud occurrence in almost all other layers. Therefore the cumulative cloud probability in a

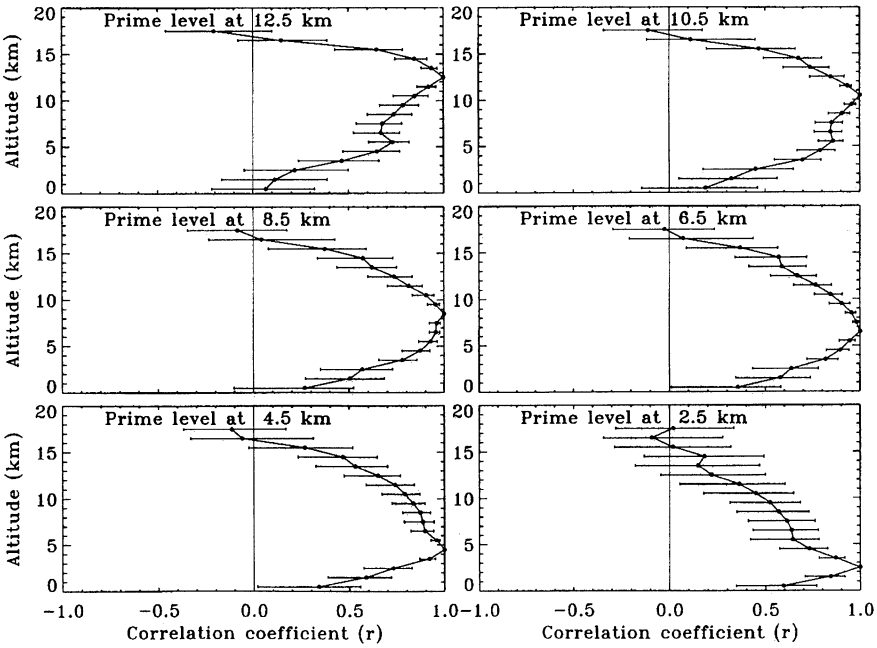

Figure 4. Same as Figure 3, except for the cumulative cloud occurrence $f^u(z)$.

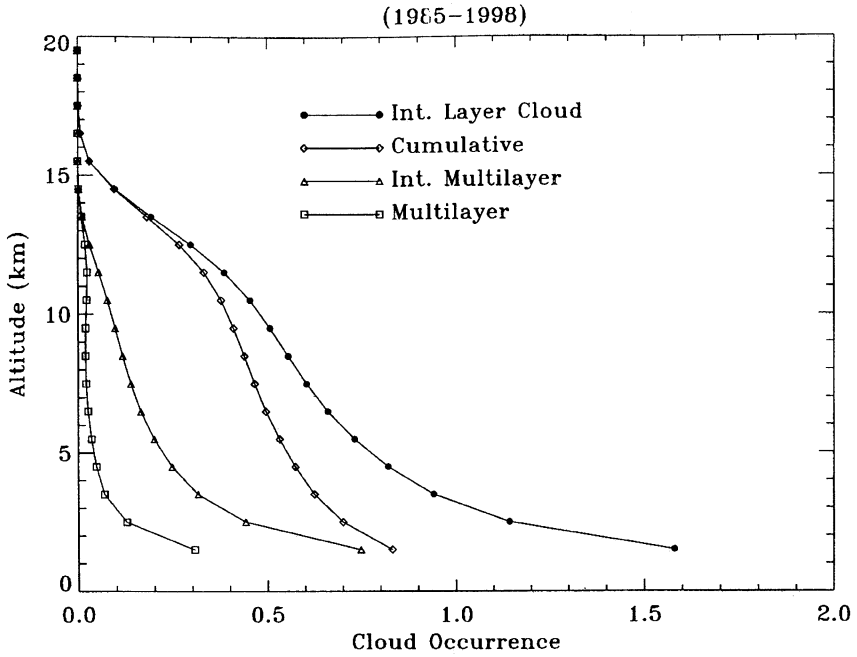

Figure 5. Vertical distribution of cumulative cloud occurrence, vertically integrated layer cloud occurrence, multilayer cloud occurrence, and vertically integrated multilayer cloud occurrence.

given layer derived from $f^u(z)$ contains basically no independent information. The reason for this feature is that the cumulative cloud probability in a layer at a lower altitude contains cloud counts of all individual layers at higher altitudes, whereas $f(z)$ counts the opaque cloud events in a given layer independently, as it is defined.

Note that $f^l(z)$ also counts the opaque cloud event in a layer independently. Thus $f^l(z)$ is similar to $f(z)$ regarding layer cloud independence. Because $f^l(z)$ estimates the probability for the case that no cloud is present above the layer, it provides cloud information suitable for estimating the outgoing longwave and shortwave radiation measured by satellite instruments such as the Earth Radiation Budget Experiment (ERBE). This suitability arises from the fact that both the outgoing longwave and shortwave radiation are affected predominantly by the uppermost layer of the opaque cloud systems.

In section 3, we referred to $f^u(z)$ as the cumulative cloud probability indicating cloud occurrence above a given altitude. Figure 5 compares $f^u(z)$ with the vertically integrated total cloud probability above a given altitude based on $f(z)$. The integrated result from $f(z)$ reveals a greater calculated total cloud occurrence than the cumulative cloud occurrence estimate. The total cloud occurrence above 1.5-km altitude integrated from the layer cloud probability is about a factor of 2 greater than the cumulative cloud occurrence. It is important to recognize that, in determining the layer cloud probability, the cloud count of a given layer is normalized to the number of the sampling event reaching the layer. Therefore the obtained opaque cloud probability of the given layer is not subject to the influence from cloud occurrence in the other layers. Consequently, $f(z)$ is capable of providing independent layer cloud statistics, resulting in a vertically integrated layer cloud frequency that contains composite information of multilayer clouds, greater than the cumulative cloud probability.

The apparent random nature of the correlation of the SAGE II tropical opaque cloud occurrence with different layers (Figure 3) implies that the cloud presence at different heights in the tropics is likely to be a random process. One exception to this apparent randomness is the negative correlation between clouds at 1–3 km and those at 12.5 km. Some tropical areas are dominated by marine boundary layer stratus-stratocumulus systems that occur in areas of subsiding air, precluding areas of deep convection or extensive cirrus clouds. In other regions the cloud vertical structure is much more complex because of a wide variation in stability, water vapor distribution, etc. Thus further investigation of the cloud vertical correlation of layer cloud occurrence on a global scale by using different data sets in the future is highly desirable.

5.2. Multilayer Clouds

As indicated in section 3, $f^l(z)$ assumes no opaque cloud above the layer of interest. Therefore the difference in cloud occurrence between $f(z)$ and $f^l(z)$ would provide an estimate for the occurrence of multilayer clouds. Symbolically, it can be shown that

$$f(z) - f^l(z) = f(z) \times f^u(z+1). \tag{4}$$

The right-hand side is just the probability for a cloud to occur in the layer at altitude z when there is a cloud above the given layer. The vertical distribution of the difference is given in Figure 5. As shown in the subsection 5.1, the integrated layer cloud occurrence above a given altitude is greater than the cumulative cloud occurrence at that altitude (Figure 5) because of multilayer clouds. Thus the difference between them at a given altitude is indicative of the integrated multilayer clouds above that altitude. In fact, the derivative of the difference is exactly the multilayer cloud probability (4).

It is understood that the difference in cloud probability between $f(z)$ and $f^l(z)$ contains composite information on the multilayer cloud occurrence above a given altitude. In fact, the profile of $f(z)$ can be used to obtain the detailed probability of multilayer cloud occurrence. For example, the probability of co-occurrence of clouds in two different layers can be determined by $f(z_i) \times f(z_j)$, where i and j are the layer indexes ($i \neq j$). The calculated results of the two-layer cloud probability using the SAGE II multiyear mean tropical observations shown in Figure 2 are presented in

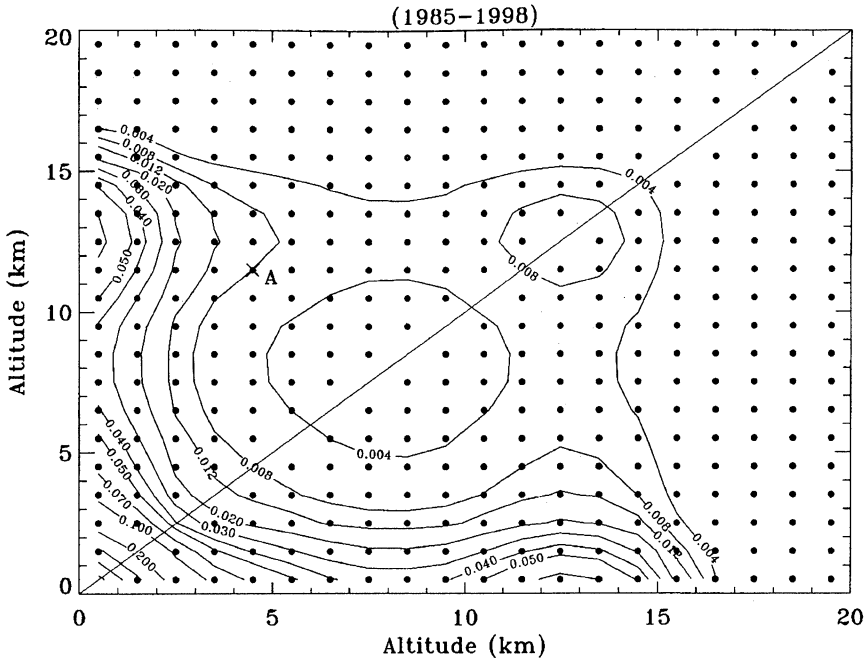

Figure 6. SAGE II statistics of tropical double layer clouds separated by various vertical distances. For an example, point A (marked by a cross) represents a frequency of 0.008 for the case of a cloud at 11.5 km when there is also a cloud at 4.5 km.

Figure 6. The results reveal many interesting features. Most of the tropical double-layer clouds occur at altitudes below ~5 to 6 km. In addition, the probability of the occurrence of double-layer clouds separated by 1 km at altitudes between 10 and 15 km is about 0.004 to 0.008. Because the occurrence of the high-level cloud is ~0.07 to 0.11 (Figure 2), the probability 0.004 to 0.008 translates to a relative frequency of the double-layer high cloud occurrence of ~6 to 8%. The case of a high-level cloud occurrence between 10 and 15 km when there is a low-level cloud between 2.5 and 4.5 km is also quite frequent. Figure 6 suggests that ~10% to 40% of opaque clouds at 12 km occur when there is a low-level cloud.

It should be emphasized that, in a single event, the SAGE II satellite instrument can only detect the top of the uppermost layer of the opaque cloud, regardless of whether it is a thick cloud (> 1 km) or there are any coexisting overlapping cloud(s) below the detected cloud. As a result, the SAGE II opaque cloud statistics at a given altitude reflect an ensemble of observations, including all types of possible multilayer clouds (at lower altitudes), as well as thick clouds (> 1 km), in addition to single layer cloud (≤ 1 km). Because of the mostly random nature of cloud occurrence, the independence of the derived layer cloud frequency allows us to infer multilayer cloud occurrence. The limitation is that it is unlikely to distinguish the occurrence of overlapping, non-overlapping multilayer clouds or thick clouds from the derived cloud statistics based on SAGE II measurements alone. Because of this limitation, the SAGE II statistics would miscategorize the thick single-layer cloud as multilayer clouds.

The vertically integrated multilayer cloud probability at 1.5-km altitude is ~75% (Figure 5). This probability is about the same magnitude as the multilayer cloud occurrence derived from tropical rawinsonde observations [*Wang and Rossow*, 1995, Figure 19]. This good agreement may be coincidental. As *Wang and Rossow* [1995] pointed out, their cloud identification technique might have missed the optically thinnest fraction of high-level clouds (see also *J. Wang et al.* [2000]). Additionally, the rawinsonde data set does not include some extensive areas of the tropical oceans. Furthermore, as indicated in the preceding paragraph, the SAGE II multilayer cloud statistics include nonoverlapping multilayer clouds and possible thick clouds. Thus further detailed multilayer cloud data intercomparison between the SAGE II and rawinsonde observations is desirable.

5.3. Regional Dependence of Layer Cloud Frequency Independence

The independent nature of the tropical layer cloud frequency from the correlation analysis presented in subsection 5.1 is significant. However, there appears to be some relationship between high (12.5 km) and low (1–3 km) clouds due to some regional differences. To determine if there is a regional variation of the independence of layer cloud frequency, we focus the analysis on two specific tropical areas, i.e., the Pacific warm pool (100°E–170°E) and the eastern Pacific (80°W–160°W). Meteorologically, these two areas are related through the so-called Walker circulation, with ascending motion over the Pacific warm pool and descending motion over the eastern Pacific [e.g., *Wang*, 1994]. Because of the different vertical air motion systems, the behavior of cloud occurrence is expected to be very different between those two tropical regions.

Over the Pacific warm pool, we have identified a total of 2969 SAGE II overpasses compared to 3462 over the eastern Pacific. The vertical distributions of the layer cloud occurrence over the Pacific warm pool and the eastern Pacific are shown in Figure 7. As anticipated, the results indicate that there are significantly more high-altitude clouds over the Pacific warm pool than over the eastern Pacific. The vertical distributions of the double-layer cloud occurrence are also quite different as shown in Figure 8. Double-layer clouds occur more often over the Pacific warm pool than over the eastern Pacific, particularly in the free troposphere.

The layer cloud correlation analyses for those two specific tropical regions are presented in Figures 9 and 10. The results indicate more independence of the layer cloud frequency than seen in Figure 3. The area dominated by marine stratocumulus has

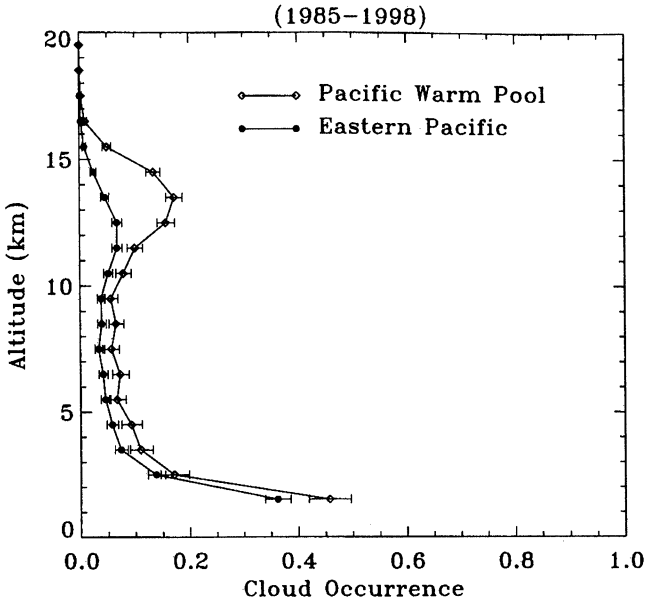

Figure 7. Vertical distributions of the layer cloud occurrence over the Pacific warm pool and the eastern Pacific. The error bars represent the 95% confidence interval of the layer cloud frequency.

been somewhat separated from the region dominated by convection. The strong anticorrelation between clouds at 1 km and the prime level at 12.5 km in Figure 3 is not evident in Figure 10 but remains in Figure 9. Similarly, the anticorrelation between clouds at 12.5 and 2.5 km in Figure 3 is still evident in Figure 9 but not in Figure 10. Thus, in a large area predominated by subsidence the occurrence of low-level clouds is independent of high-level clouds, while in areas with less atmospheric stability the low-level clouds are more likely to occur if no high-level clouds are present and vice versa.

6. Summary and Concluding Remarks

To enhance the understanding of the derived SAGE II cloud statistics, the method used to estimate the vertical distribution of opaque cloud occurrence from the SAGE II solar occultation observation has been reviewed thoroughly. Most importantly, based on SAGE II observations over the entire tropics from 1985 to 1998, the present study has illustrated that the derived information of cloud occurrence in a given layer is generally independent of the cloud occurrence in other layers, except for some anticorrelation between high-level (12.5 km) and low-level (1–3 km) clouds. This behavior of the layer cloud frequency is also evident when regional data over the Pacific warm pool and the eastern Pacific are examined. The anticorrelation occurs over the Pacific warm pool where deep convection and high-level clouds are prevalent but is not found in the eastern Pacific where high-level cloud occurrence is probably more random than in less stable areas.

The independent information of the layer cloud frequency makes the vertical distribution of the layer cloud occurrence useful for inferring the probability of multilayer clouds. The limitation is that it is difficult to determine how frequently the multilayer clouds are actually overlapping or how frequently thick cloud (>1 km) really occurs, based on SAGE II observations alone. Because the SAGE II statistics count a thick cloud (>1 km) more than once depending on the thickness of the cloud, the vertically integrated cloud frequency of a thick atmospheric layer would overestimate the cloud occurrence for the given thick layer. This possibility, together with the fact that the SAGE II statistics include non-overlapping clouds as well as overlapping clouds, makes the SAGE II cloud statistics likely to be an upper limit of the multilayer cloud occurrence.

A discussion of the SAGE II tropical cloud occurrence in relation to the cloud statistics using rawinsonde records, the ISCCP data products, and the surface weather observations is also provided in the present study. The results of this cloud data intercomparison indicate that all data sets reveal (1) the smallest cloud occurrence for the middle-level cloud, and (2) the largest cloud occurrence for the low-level cloud, except the ISCCP data set. Furthermore, both the SAGE II observations and the rawinsonde analysis reveal 75% total multilayer cloud occurrence in the tropics. Nevertheless, further detailed comparison of cloud data derived from different observational techniques is desirable.

Again, the primary focus of the present study is to review the method used to derive the SAGE II opaque cloud statistics for further understanding of the nature of the derived SAGE II cloud statistics. Previously, the SAGE II observations from 1985 to 1990 were used to determine a 6-year cloud climatology [*Wang et al.*, 1996]. With the current multiyear SAGE II observations we have extended that study to develop an advanced solar occultation

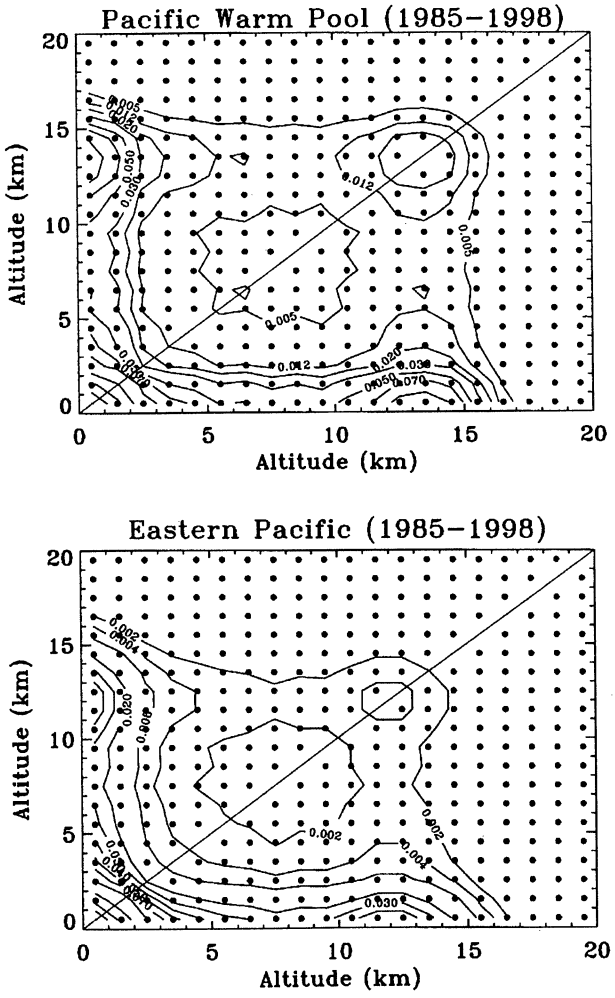

Figure 8. SAGE II statistics of tropical double layer clouds over the (top) Pacific warm pool and the (bottom) eastern Pacific separated by various vertical distances. Note that the contour levels for the Pacific warm pool are different from those for the eastern Pacific.

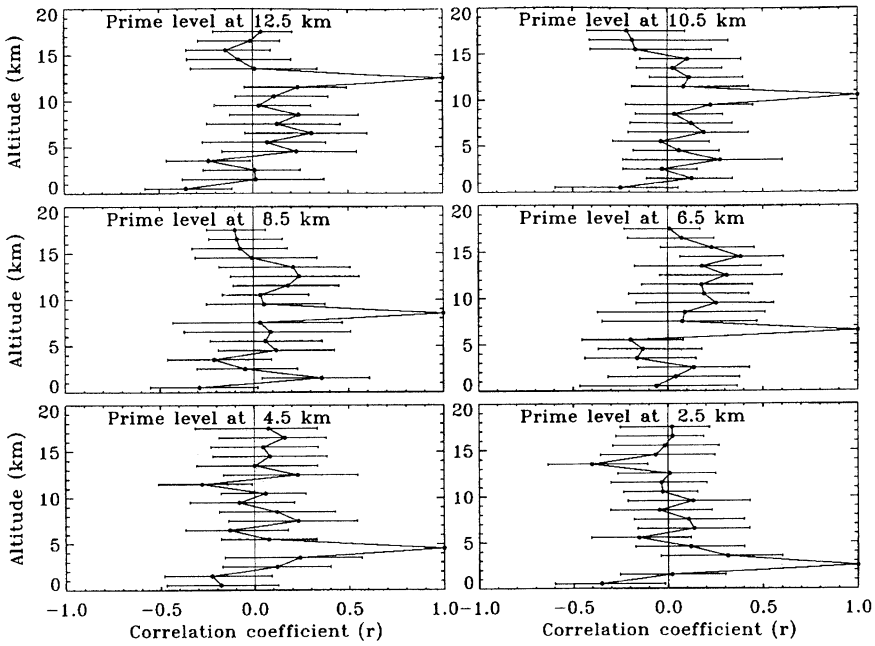

Figure 9. Correlation analysis for selected prime cloud layers for layer cloud occurrence $f(z)$, based on SAGE II observations over the Pacific warm pool. The error bars represent the 95% confidence interval of the correlation coefficient.

cloud climatology, including an assessment of possible cloud trend behavior important to climate change investigations. This latter work is in progress and will be reported separately.

Finally, the present study has related the SAGE II statistics to the surface-based observations and the ISCCP data record. This association, however, is admittedly far from comprehensive. It is clear that different cloud observational techniques may have respective unique characteristics, strengths, and weakness. Currently, several multiyear global cloud data sets exist, including data from visual cloud observations from surface and ships, rawinsonde analyses, the International Satellite Cloud Cimatology Project (ISCCP) data products, and satellite data derived from the High-Resolution Infrared Radiometer Sounder (HIRS), SAGE II, etc. It is therefore extremely important and highly desirable for the cloud observational community to conduct a comprehensive cloud data intercomparison for understanding the characterisics of

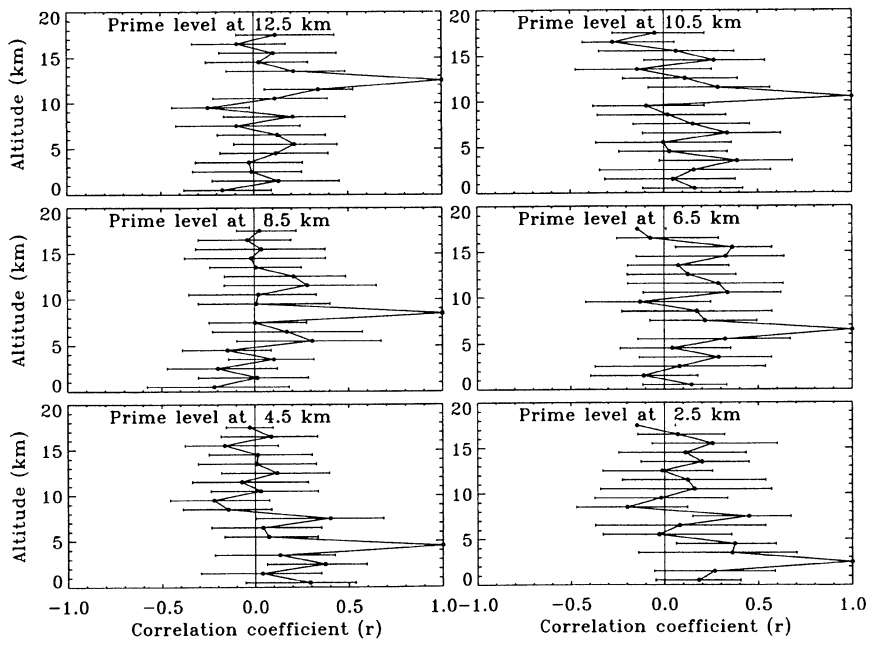

Figure 10. Correlation analysis for selected prime cloud layers for layer cloud occurrence $f(z)$, based on SAGE II observations over the eastern Pacific. The error bars represent the 95% confidence interval of the correlation coefficient.

different cloud observational techniques through cooperative effort.

Acknowledgments. The authors would like to thank Bruce Wielicki and William P. Chu of the NASA Langley Research Center for many useful discussions and comments. The authors are also grateful for the constructive and clarifying comments of two anonymous reviewers. This study is supported by NASA contracts NAS1-99129 and NAS1-19570 and by the NASA First ISCCP Regional Experiment (FIRE) program.

References

Chu, W. P., M. P. McCormick, J. Lenoble, C. Brogniez, and P. Pruvost, SAGE II inversion algorithm, *J. Geophys. Res., 94*, 8339–8351, 1989.

Efron, B., Bootstrap confidence intervals for a class of parametric problems, *Biometriks, 72*, 45–58, 1985.

Fowler, L., and D. A. Randall, Liquid and ice microphysics in the CSU general circulation model, II, Impact on cloudiness, the Earth's radiation budget, and the general circulation of the atmosphere, *J. Clim., 9*, 530–560, 1996.

Gierens, K., U. Schumann, M. Heten, H. Smit, and P.-H. Wang, Ice-supersaturated regions and subvisible cirrus in the northern midlatitude upper troposphere, *J. Geophys. Res.,* in press, 2001.

Johnson, N. L., and S. Kotz, *Discrete Distribution*, 328 pp, 1969, Houghton Mifflin, Boston, Mass.

Kent, G. S., D. M. Winker, M. T. Osborn, M. P. McCormick, and K. M. Skeens, A model for the separation of cloud and aerosol in SAGE II occultation data, *J. Geophys. Res., 98*, 20,725–20,735, 1993.

Liao, X., W. B. Rossow, and D. Rind, Comparison between SAGE II and ISCCP high-level clouds, 1, Global and zonal mean cloud amounts, *J. Geophys. Res., 100*, 1121–1136, 1995a.

Liao, X., W. B. Rossow, and D. Rind, Comparison between SAGE II and ISCCP high-level clouds, 2. Locating cloud tops, *J. Geophys. Res., 100*, 1137–1148, 1995b.

Mauldin, L. E., III, N. H. Zaun, M. P. McCormick, J. H. Guy, and W. R. Vaughn, Stratospheric Aerosol and Gas Experiment II instrument: A functional description, *Opt. Eng., 24*, 307–312, 1985.

McCormick, M. P., SAGE II: An overview, *Adv. Space Res., 7*(3), 219–226, 1987.

McCormick, M. P., P. Hamill, T. G. Pepin, W. P. Chu, T. J. Swissler, and L. R. McMaster, Satellite studies of the stratospheric aerosols, *Bull. Am. Meteorol. Soc., 60*, 1038–1049, 1979.

Mergenthaler, J. L., A. E. Roche, J. B. Kumer, and G. A. Ely, Cryogenic limb array etalon spectrometer observations of tropical cirrus, *J. Geophys. Res.*, 104, 22,183–22,194, 1999.

Pruppacher, H. R., and J. D. Klett, *Microphysics of Clouds and Precipitation*, 714 pp., D. Reidel, Norwell, Mass., 1978.

Rossow, W. B., A. W. Walker, D. E. Beuschel, and M. D. Roiter, International Satellite Cloud Cimatology Project (ISCCP) documentation of new datasets. *Tech. Doc. WMO/TD-737*, 115 pp., World Meteorol. Org., Geneva, Switzerland, 1996.

Sassen, K., and B. S. Cho, Subvisual-thin cirrus lidar dataset for satellite verification and climatological research, *J. Appl. Meteorol., 31*, 1275–1285, 1992.

Solomon, S., S. Borrmann, R. R. Garcia, R. Portmann, L. Thomason, L. R. Poole, D. Winker, and M. P. McCormick, Heterogeneous chlorine chemistry in the tropopause region, *J. Geophys. Res., 102*, 21,411–21,429, 1997.

Wang, J., and W. B. Rossow, Determination of cloud vertical structure from upper-air observations, *J. Appl. Meteorol., 34*, 12243–12258, 1995.

Wang, J., and W. B. Rossow, Effect of cloud vertical structure on atmospheric circulation in the GISS GCM, *J. Clim., 11*, 3010–3029, 1998.

Wang, J., W. B. Rossow, and Y. Zhang, Cloud vertical structure and its variations from a 20-yr global rawinsonde dataset, *J. Clim., 13*, 3041–3056, 2000.

Wang, P.-H., SAGE II tropospheric measurement frequency and its meteorological implication, reprint, in *Seventh Conference on Satellite Meteorology and Oceanography*, pp. J15–J18, Am. Meteorol. Soc., Boston, Mass., 1994.

Wang, P.-H., M. P. McCormick, L. R. Poole, W. P. Chu, G. K. Yue, G. S. Kent, and K. M. Skeens, Tropical high cloud characteristics derived from SAGE II extinction measurements, *Atmos. Res., 34*, 53–83, 1994.

Wang, P.-H., M. P. McCormick, P. Minnis, G. S. Kent, G. K. Yue, and K. M. Skeens, A method for estimating vertical distribution of the SAGE II opaque cloud frequency, *Geophys. Res. Lett., 22*, 243–246, 1995.

Wang, P.-H., P. Minnis, M. P. McCormick, G. S. Kent, and K. M. Skeens, A 6-year climatology of cloud occurrence frequency from SAGE II observations (1985–1990), *J. Geophys. Res., 101*, 29,407–29,429, 1996.

Wang, P.-H., P. Minnis, T. Wang, B. Wielicki, L. B. Vann, G. K. Yue, G. S. Kent, and P. L. Lucker, Satellite observations of tropical clouds and outgoing longwave radiation from 1985 to 1998: Preliminary results, *Eos Trans. AGU, 80*(46), Fall Meet. Suppl., F143, 1999.

Wang, P.-H., B. Wielicki, P. Minnis, T. Wang, An estimation of SAGE II clouds outgoing longwave radiative forcing between 1985 and 1998: Preliminary results, in *Abstracts, Current Problems in Atmospheric Radiation*, p.163, St.-Petersburg State Univ., Russia, 2000.

Warren, S. G., C. J. Hahn, J. London, R. M. Chervin, and R. L. Jenne, Global distribution of total cloud cover and cloud type amounts over land, *Tech. Note NCAR TN-273+STR*, NTIS number DE87-006903, Natl. Cent. Atmos. Res., Boulder, Colo., 1986.

Warren, S. G., C. J. Hahn, J. London, R. M. Chervin, and R. L. Jenne, Global distribution of total cloud cover and cloud type amounts over ocean, *Tech. Note NCAR TN-317+STR*, Natl. Cent. Atmos. Res., Boulder, Colo., 1988.

Wylie, D. P., and P.-H. Wang, Comparison of cloud frequency data from the high-resolution infrared radiometer sounder and the Stratospheric Aerosol and Gas Experiment II, *J. Geophys. Res., 102*, 29,893–29,900, 1997.

Wylie, D. P., and P.-H. Wang, Comparison of SAGE-II and HIRS co-located cloud height measurements, *Geophys. Res. Lett., 26*, 3373–3375, 1999.

G. S. Kent and P.-H. Wang, Science and Technology Corporation, Hampton, VA 23666. (e-mail: p.wang@larc.nasa.gov)

P. Minnis and L. B. Vann, Atmospheric Sciences Division, NASA Langley Research Center, Hampton, VA 23681.

R. E. Veiga, Science Applications International Corporation, Hampton, VA 23681.

(Received July 24, 2000; revised January 31, 2001;
accepted February 13, 2001.)