Bull. Disas. Prev. Res. Inst., Kyoto Univ., Vol. 40, Part 2, No. 347, June 1990

Nephanalysis of the GMS Imagery Data

By Pingping XIE

(Manuscript received on May 23, 1990)

Abstract

Techniques estimating cloud amount and classifying cloud type from GMS IR and/or VIS imagery data have been developed and evaluated. Cloud amount is estimated by a Two-Threshold-Method (TTM) which takes account of the partially cloud-covered pixels. Cloud type is classified by a discriminant analysis technique. Fairly good agreement is obtained between the satellite estimated and the surface observed cloud amounts. The comparison between the satellite classified and ground observed cloud types shows correctness ratios of 52.9% and 47.7% for the daytime and nighttime respectively, but the distinction between cumulus and middle clouds seems insufficient. The techniques were applied to produce the time series of cloud amounts and types during the period from 3 to 11 Sept. 1980, and the results were acceptable.

1. Introduction

Cloud information, such as cloud amount, cloud height and cloud type have been considered as important and essential parameters in meteorology and climate. Clouds take part in the water circulation process by condensing water vapor converged from the environment and then precipitating it out of the clouds. Clouds also affect the energy budget by reflecting solar radiation, absorbing the long wave radiation from the earth and lower atmosphere and releasing the latent heat of condensation. Cloud information can be looked on as parameters for rainfall information which can be used for rainfall estimation from satellites.

Cloud observations have been made routinely from surface observational stations. But they can not provide accurate information about middle and high clouds when the low cloud cover is significant. Their space representiveness, especially over the ocean, is insufficient because of sparse island-located meteorological observational stations and ships.

Since the launch of the first meteorological satellite TIROS-1, satellite imagery data have been used broadly to derive meteorological information. The observation made by a satellite has the advantages of nearly uniform space resolution and wide scope. By use of the network comprised of polar-orbited and geostationary satellites, nearly global observation can been made at 30 minute intervals. However, there has been no method, until now, to derive cloud amount and type with acceptable accuracy and on a completly objective base, perhaps because of the characteristics of satellite observation and the enormous amount of data to be processed.

Most techniques for cloud amount estimation with meteorological satellite imagery data use IR blackbody temperature and/or albedo histograms for the area of interest. Koffler, *et al.*[1] investigated method for estimating high, middle and low cloud amounts with TIROS IR imagery data. Every pixel is assigned to clear sky,

low cloud, middle cloud or high cloud according to T_{BB}, being the surface temperature T_{SFC}–5k, 700 mb temperature (T_{700}), 400 mb temperature (T_{400}) or lower. The high, middle and low cloud amounts for a 32 line $\times$32 pixel area are then defined as the ratioes of high, middle and low cloud pixel numbers to the total pixel number of the area, respectively. Meteorological Satellite Center of Japan calculates high and low cloud amounts routinely using T_{SFC}–5k and T_{400} as thresholds.

The cloud amount estimation methods mentioned above define a pixel as being clear sky or totally cloud covered from observation and by a single threshold value (hereafter referred to as Single Threshold Method, STM). It is thought that the estimations by STM may have some error arising from the use of a fixed threshold and the existence of partially cloud covered pixels, and that this error may be relatively large in estimation from large size pixels such as infrared data of geostationary satellites.

Meanwhile, the cloud type is classified from satellite imagery data operationally by trained meteorologists at present. It is clear that it is difficult for such a subjective method to make a fine cloud map (nephanalysis map) for a wide area (e.g. the whole globe) on a real time base. Aiming to solve this problem, objective cloud type classification methods have been investigated by using threshold, clustering and discriminant analysis techniques.

The threshold technique, as described by Koffler *et al.*[1] and Liljas[2], classifies cloud type according to the cloud top temperature or albedo by use of T_{BB} and/or albedo thresholds. Clustering technique, as described by Desbois *et al.*[3] and Seze and Desbois[4], classifies "similar" pixels in an interest area into several clusters. Relatively long computation time is consumed for repeating the calculation and the resulted clusters must be related to real cloud types.

The discriminant analysis technique, as investigated by Harris and Barrett[5], Parikh[6], Parikh[7] and Parikh and Ball[8], classifies an observed satellite parameter $\boldsymbol{X}$ for interest area by comparing the probabilities of $\boldsymbol{X}$ being produced by each cloud type. It is considered as a relative appropriate method for objective cloud type classification, because only simple computation is needed to classify cloud type after the cloud types to be classified are determined, the training cases for every type were selected based upon the ground truth and the discriminant equations were established.

The present paper describes a two threshold method (hereafter referred to as TTM) to estimate total cloud amount taking account of the spectral peaks due to the ground surface as well as the pixels partially covered by cloud. We will also describe a discriminant method to classify four cloud types, in relation to the future rainfall estimation, employing the data of the geostationary meteorological satellite GMS.

2. Data

In order to develop the cloud amount estimation and cloud type classification

methods, six flat areas of about 10^4 km^2 in Japan were selected as the testing areas. The IR and VIS imagery data of GMS for these six areas were collected, and the simultaneous cloud information (total cloud amount and cloud amounts for each cloud form) of the meteorological stations were collected and used as the ground truth for the periods 3 to 10 Sept. 1980, 23 July 1982 and 22 to 23 July and 7 to 9 Aug. 1983.

Fig. 1 shows the six testing areas: Hokkaido, Kanto, Kinki, Setonaikai, Sanin and Kyushu, the lengths (NS direction) and widths (EW direction) of which were respectively 80 and 110 km on average. There are four or five meteorological stations in each area. The pixel numbers are about 550 and 4400 for IR and VIS in each area.

GMS observes the earth's atmosphere by visible and infrared spin scan radiometer (VISSR) which has two observation channels of VIS (0.5–0.75 μm) and IR (10.5–12.5 μm). It had been observing routinely at an interval of 3 hours until 1986. In order to compare the satellite data with the ground observations which are made at intervals of 6 hours for most meteorological stations, only the data sets for 00Z, 06Z and 12Z could be used in this study.

VIS and IR imagery data of the six testing areas were extracted from the GMS original MT. The line and pixel numbers in the GMS imagery frame corresponding to the latitudes and longitudes of four corners of the testing area were calculated by applying a navigation method which makes a correction with the earth's edge in the image frame (Takahashi[9]).

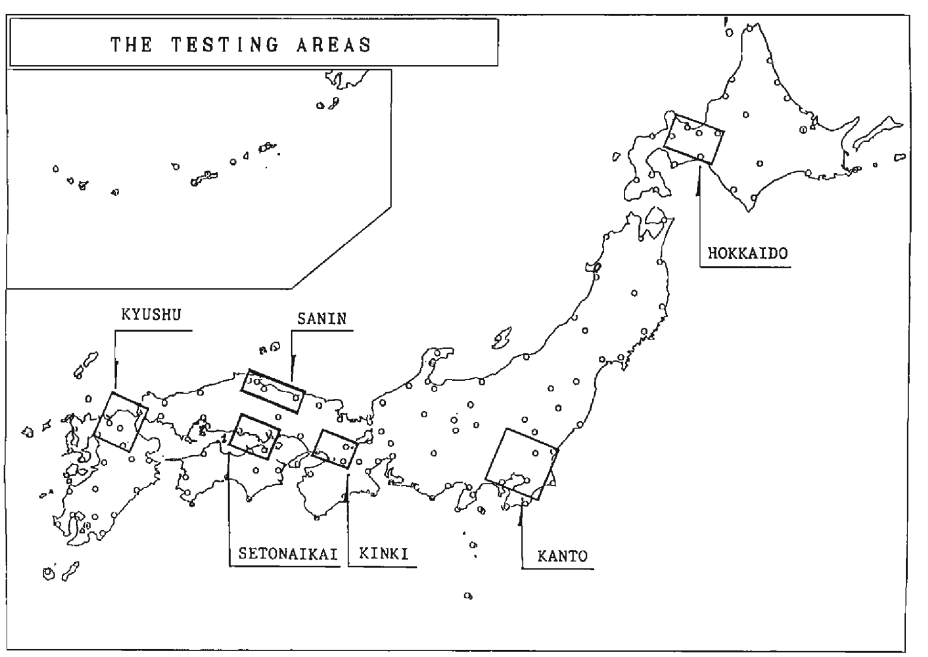


Fig. 1. The testing areas.

The extracted VIS and IR data are converted to albedo, A and blackbody temperature, T_{BB}. The conversion from IR digital count (DC) to T_{BB} can be accomplished using the calibration table which is recorded on the top of the original VISSR MT. As for the method converting VIS DC to albedo, the calibration table recorded on the MT is for the illumination condition that the sun and the satellite are both right above the reflector (e.g. clouds), it is necessary to normalize the observed albedo, a considering the solar and the satellite zenith angles, as follows:

$$A = \frac{a}{\cos Z \cdot \cos \theta} \tag{1}$$

$$\cos Z = \sin \phi \cdot \sin \delta \cdot + \cos \phi \cdot \cos \delta \cdot \cos t \tag{2}$$

$$\cos \theta = \frac{1}{(1+\tan^2 \theta)^{1/2}} \tag{3}$$

$$\tan \theta = \frac{\sin \delta}{\cos \delta - \dfrac{R}{R+H}} \tag{4}$$

$$\cos \delta = \cos \phi \cdot \cos (\lambda_s - \lambda) \tag{5}$$

where, A is normalized albedo, Z solar zenith angle, θ satellite zenith angle, ϕ and λ latitude and longitude of the observed point, δ solar declination, t hour angle, λ_s latitude of the sub-satellite point, R radius of the earth and H being the height of the satellite.

The ground truth of total cloud amounts and cloud amounts of each cloud form of the testing areas were obtained as the area weighted mean of the observed values at each station in the areas.

Cloud observation from the surface has some limitations: 1) observation extent is restricted by obstacles (e.g. mountains, high buildings) in the field of view; 2) only the clouds within a range of about 20–30 km can be observed even when there is no obstacle; 3) the visual size varies with the angle of view; 4) upper clouds can not be observed when the lower clouds are prominent and 5) the reliability declines when illumination conditions are bad (e.g. at night). In the present study, because averages of 4–5 station observations are used as the ground truths for each area of about 100 km × 100 km on plain, 1), 2) and 3) can be disregarded. Such averages are expected to give relatively reliable ground truth for cloud information.

Included among the data used in this study are, 3–11 Sept. 1980 for typhoon T8013, 23 July. 1982 for heavy rain in Nagasaki, 1–3 Aug. 1982 and 7–9 Aug. 1983 for typhoons.

3. Cloud Amount Estimation from Satellite Data

Shenk and Salomoson[10] simulated the effects of spatial resolution of the sensor on the cloud amount estimation, and showed that, in order to estimate cloud amount

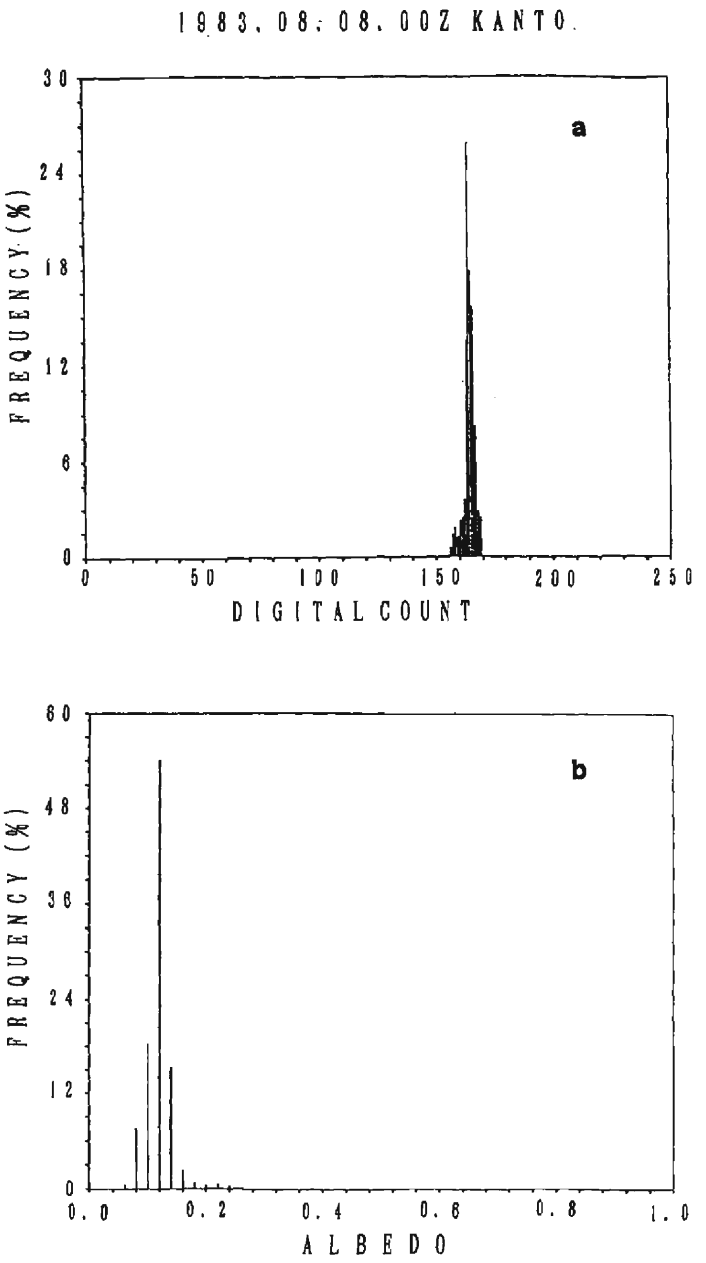


Fig. 2. The example of frequency distribution of a) T_{BB} and b) albedo, for the case of clear sky. T_{BB} 220k, 240k, 260k, 280k and 300k correspond to digital counts 46, 68, 96, 132 and 176, respectively in a). The albedo values have been normalized by considering the solar and satellite's zenith angle.

within an accuracy of 10%, the cloud size to pixel size ratio should be greater than 100 when STM is used. In the case of GMS, as the pixel size for IR is about 7 km near Japan and cloud size may be in the order of 10 km, it is difficult to satisfy this ratio. Therefore, TTM was introduced as stated before. In this chapter, after a simple analysis of T_{BB} and A histograms, the TTM, two-threshold cloud amount estimation method, was developed using satellite and surface observation data of 00Z and 06Z.

3.1 Examples of Histograms

Among the total of 156 cases of 00Z and 06Z data used here, 6 cases are with no ground observed cloud amount (clear sky case). **Fig. 2** shows an example of T_{BB} and albedo histograms for clear sky. The T_{BB} and albedo histograms have shape mono-mode distributions and concentrate in warm (T_{BB}) or dark (albedo) side. The investigation of the histograms for the six clear sky cases shows that, while the modes of T_{BB} differ from each other because of the different surface temperatures, the variation in the modes of albedo is not so large. The standard deviation of T_{BB} in a testing area ranges from 1.0°C to 3.2°C with an average of 2.0°C, and the standard deviation of albedo ranges from 0.02 to 0.05 with an average of 0.04.

Fig. 3 shows an example of T_{BB} and albedo histograms for the case of ground-observed cloud amount of 5. The ground observation shows the existence of middle clouds. The T_{BB} and albedo histograms have a warm or dark peak related to the surface while other points distribute to the cold and bright side showing the second peak corresponding to the middle clouds. When the area is overcast with clouds, the T_{BB} and albedo histograms tend to extend to the cold and bright sides, and the distribution pattern differs with cloud type. **Fig. 4** shows an example of T_{BB} and albedo histograms for the convective cloud case with total cloud amount of 10.

3.2. Cloud Amount Estimation by TTM

Based upon the above results, a two-threshold cloud amount estimation method (TTM), taking account of the spectral peaks due to the ground surface and of the pixels partially covered by clouds, was developed for daytime observations (00, 06Z). A simple schematic explanation of TTM is given in **Fig. 5**.

(a) IR Histogram

IR imagery data are the only data that can be obtained throughout the day but they have the shortcoming of poor spatial resolution. When the IR 1-dimensional histogram of the area is constructed as in **Fig. 5**, the two thresholds that discriminate the no-cloud, partial cloud and total clouds pixels were determined as follows.

If there is a peak on the warm side (T_{BB} near normal surface air temperature) of the IR histogram, and the peak temperature is T_G, the peak is assumed to be yielded by ground surface. In case no peak can be found on the warm side, the

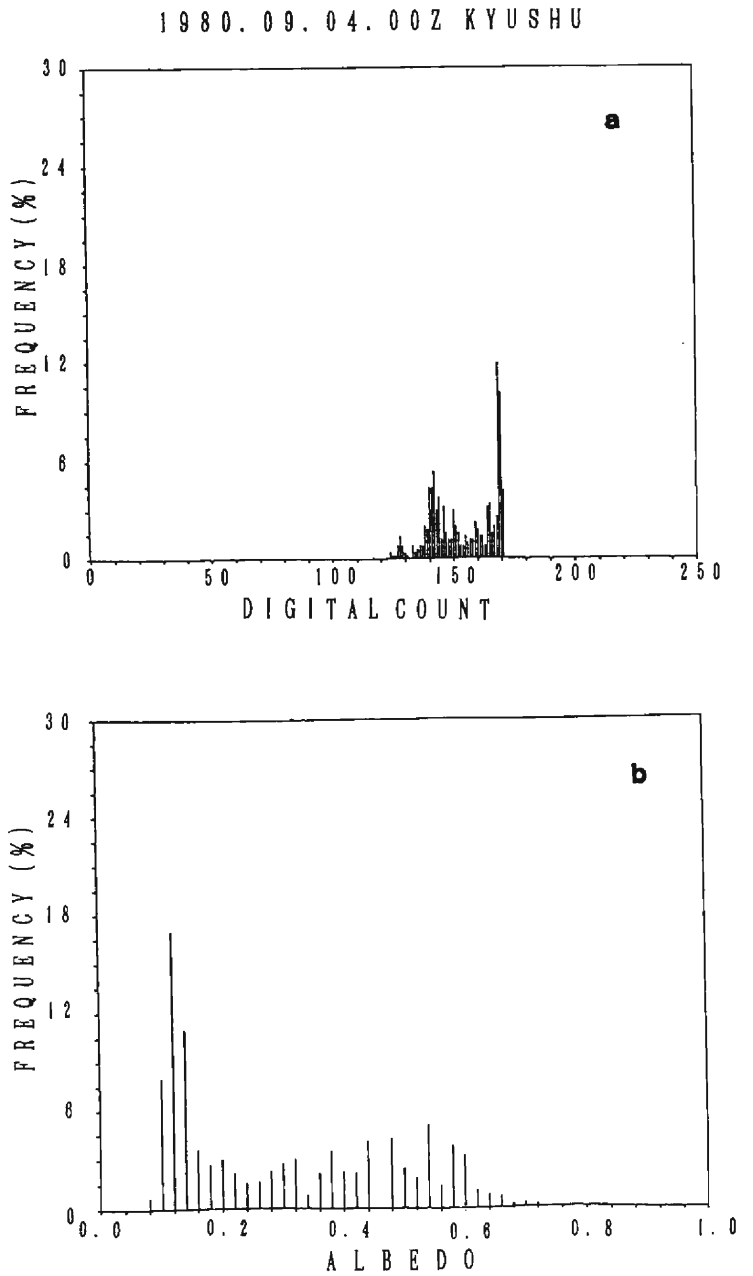


Fig. 3. As in **Fig. 2**, except for the case of total cloud amount 5. T_{BB} 220k, 240k, 260k, 280k and 300k correspond to digital counts 51, 74, 105, 143 and 189, respectively in a).

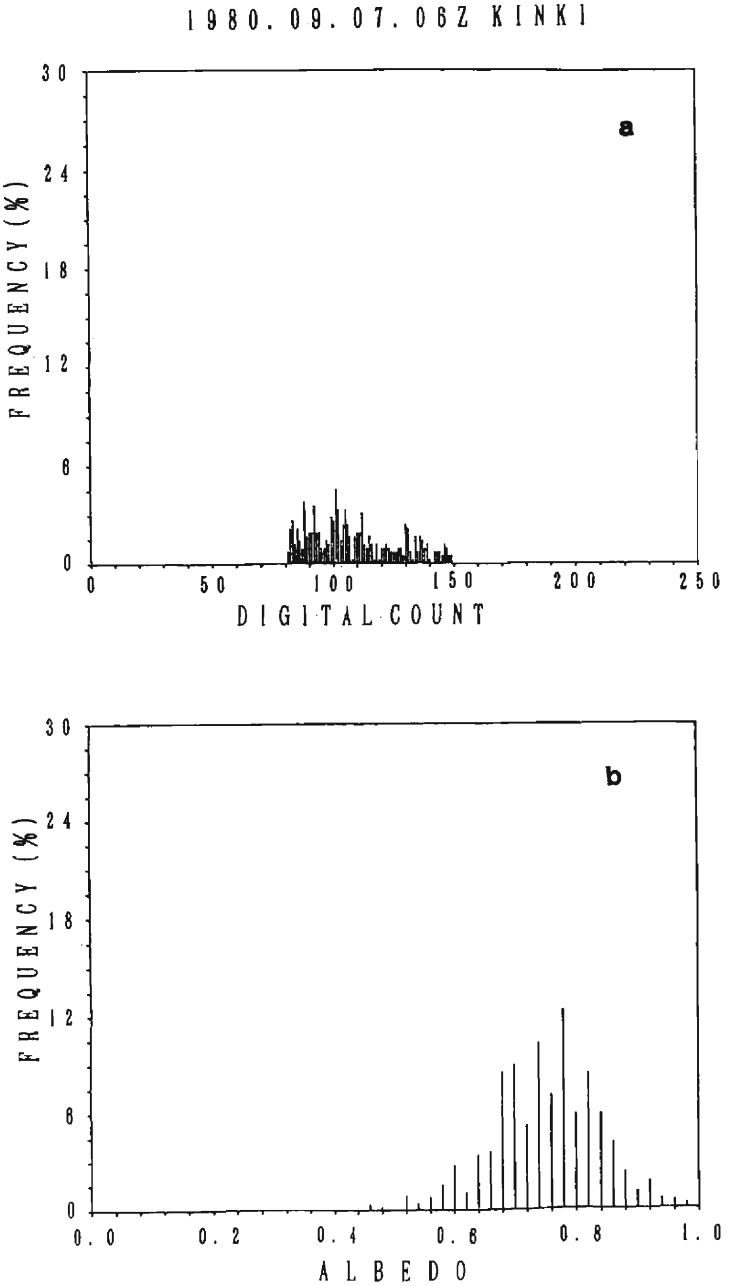


Fig. 4. As in **Fig. 2**, except for the case of total cloud amount 10 (convective cloud). T_{BB} 220k, 240k, 260k, 280k and 300k correspond to digital count 51, 74, 105, 143 and 189, respectively in a).

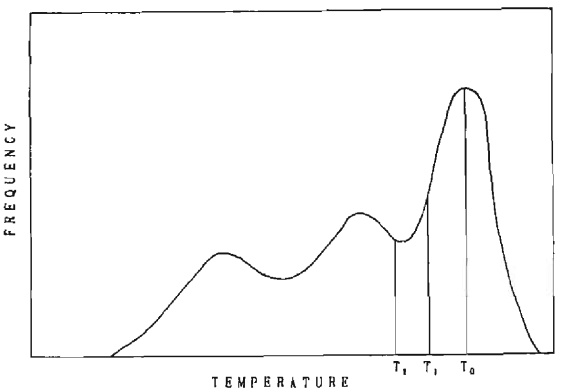


Fig. 5. Schematic figure of the two-threshold cloud amount estimation method (TTM) from T_{BB} histogram.

average of T_G of those cases with a ground peak at the same hour at the same place is used. As described above, the standard deviations of T_{BB} in the investigation areas for 6 clear sky cases have an average of 2.0k. When the area is partially covered by clouds, T_{BB} due to the no-cloud part (ground surface) are assumed to distribute around T_G with standard deviation the same as that of clear sky cases and a pixel with temperature warmer than $T_1 = T_G - 2.0$ k is defined as a no-cloud pixel.

If the warmest pixel covered totally by clouds is assumed to be with a T_{BB} of T_2, a pixel with a T_{BB} between T_1 and T_2 is considered to be covered partially by clouds. The ratio of clouds can be assumed to be linearly proportional to $\Delta T = T_1 - T_2$. That is, the satellite estimation of cloudiness, FC for an area is calculated by the following equations;

$$\mathrm{FC} = \frac{1}{N} \sum f_i \tag{6}$$

where

$$f_i = \begin{cases} 1 & T_i \leqq T_2 \\ \dfrac{T_1 - T_i}{T_1 - T_2} & T_2 < T_i \leqq T_1 \\ 0 & T_1 < T_i \end{cases} \tag{7}$$

where T_i is T_{BB} of a pixel, N is the total pixel number of the area.

Because the second threshold T_2 can't be determined directly from IR histograms, it is selected by an indirect way. First, FC's for all cases are calculated from IR histograms by assuming various values of T_2. Regressions between the surface observed cloud amounts (GC's) and FC's for various T_2 are made and T_2 with the highest correlation coefficient is selected as the second threshold which discriminates the partial cloud and total cloud pixels. The correlation coefficients vary from 0.810 to 0.861 when $\Delta T = T_1 - T_2$ varies from 1.0k to 10.0k, with the

highest value at $\Delta T=1.0$k. The correlation coefficient becomes 0.853 for $\Delta T=0.0$, namely the STM case. Although the difference between correlation coefficients for TTM and STM cases was not so large, it is considered that the small difference results from the dominant cases of overcast and the cloud amount estimation is improved for the partial cloudy case. The two threshold cloud amount estimation method developed by Parikh and Ball[8] for GATE uses the warmest peak temperature as T_1, the second warmest peak temperature as T_2, and the difference between the so defined T_1 and T_2 has an average of about 2.5k, which is close to the corresponding value of 3.0k ($\Delta T+2.0$k) in the present method. Although ΔT is calculated here from limited data and the value can be different for other areas or seasons, the method to determine it could be applicable for any case.

(b) VIS Histogram

Although VIS data can be used only in daytime and complicated normalization has to be performed, they are still considered as an important information source because of their fine spatial resolution. Similarly for IR data, when a VIS 1-dimensional histogram is used to estimate cloud amount, first of all the ground surface related peak is detected in the dark side of the histogram and in the peak albedo A_G. The first threshold $A_1=A_G+0.04$ is then determined in the same way as IR data. Assuming the second threshold has a value of A_2, the satellite estimation can be calculated by using eqs. (6) and (8).

$$f_i=\begin{cases} 1 & A_i\geqq A_2 \\ \dfrac{A_i-A_1}{A_2-A_1} & A_2>A_i\geqq A_1 \\ 0 & A_1>A_i \end{cases} \tag{8}$$

Here A_i is the normalized albedo of a pixel. And in the same way as IR, the second threshold value A_2 is determined as $A_2=A_1-0.02$ with a correlation coefficient of 0.825.

(c) IR·VIS 2-Dimensional Histogram

Although the estimation from IR or VIS histograms has an advantage of simplicity, a relative larger error will result, because the very low clouds and thin cirrus can hardly be detected from IR or VIS data. Takeda and Hattori[11] indicated that when IR data of NOAA with spatial resolution of 1 km are used to estimate cloud amount over the ocean, the low level clouds can be detected by incorporating spatial distribution information of T_{BB} (such as standard deviation around a pixel). But no obvious improvement can be achieved when their method is applied to the IR data of GMS with a spatial resolution of 5 km. Meanwhile, as was shown by Xie and Mitsuta[12] it is possible to detect low level clouds and thin cirrus by using IR·VIS 2-dimensional histograms.

The procedures for the 2-dimensional histogram method are the same as those for the 1-dimensional one described above, except that 1) the corresponding 2 thre-

sholds are constituted by 2 perpendicular lines in a IR·VIS domain, 2) the satellite estimation FC is calculated by eqs. (6) and (9).

$$f_i = \begin{cases} 1 & T_i \leqq T_2 \text{ or } A_i \geqq A_2 \\ \dfrac{T_1 - T_i}{T_1 - T_2} & T_2 < T_i \leqq T_1,\ A_1 > A_i \\ \dfrac{A_i - A_1}{A_2 - A_1} & T_1 < T_i,\quad A_2 > A_i \geqq A_1 \\ \mathrm{MAX}\left[\dfrac{T_1 - T_i}{T_1 - T_2}, \dfrac{A_i - A_1}{A_2 - A_1}\right] & T_2 < T_i \leqq T_1,\ A_2 > A_i \geqq A_1 \\ 0 & T_1 < T_i,\quad A_1 > A_i \end{cases} \tag{9}$$

Here, T_i and A_i are the T_{BB} and albedo of pixel, $T_1 = T_G - 2.0\text{k}$ and $A_1 = A_G + 0.04$ are the first thresholds to discriminate clear sky and partial cloud pixels. The statistical results showed that the highest correlation coefficient of 0.889 is obtained when $T_2 = T_1 - 2.0\text{k}$ and $A_2 = A_1 - 0.04$.

3.3 Comparison of Satellite Estimated Cloud Amounts and Ground Observations

The cloud amount estimation methods described above were tested and compared with the corresponding ground observations. Comparisons were made for daytime (00Z, 06Z) data which were used in the development of the method, and nighttime (12Z) with the same methods, respectively.

The satellite estimated FC's were calculated from IR, VIS 1-dimensional histograms and IR-VIS 2-dimensional histograms separately by using thresholds determined in the preceeding section. **Fig. 6** gives a comparison between IR estimated FC's and ground observed GC's for the daytime. The points were scattered when ground cloud amounts were 4–6. The correlation coefficient between FC's and GC's is 0.861 and the RMS error for FC's was 0.155. Kubota and Endo[13] estimated the cloud amounts for a $1° \times 1°$ area by using GMS IR imagery data in a one threshold method, defining the cloud pixels as 5°C colder than the surface temperatures defined by the GMSSA climate data. Their comparison between the estimated values and the ground observations showed a correlation coefficient of 0.438 and RMS error of 0.331 for overland cases. Although no direct comparison can be made between the results in their and the present study because different data are used, significant difference in corelation coefficients in the two studies implies that TTM developed here did improve cloud amount estimation.

The comparisons between VIS estimated FC's and GC's, IR-VIS estimated FC's and GC's were also conducted for the daytime. The correlation coefficients and RMS errors were 0.825 and 0.167 for VIS estimations, and 0.889 and 0.140 for IR-VIS two dimensional estimations. It was found from the comparisons that the best correspondence between FC's and GC's is obtained when both IR and VIS data are

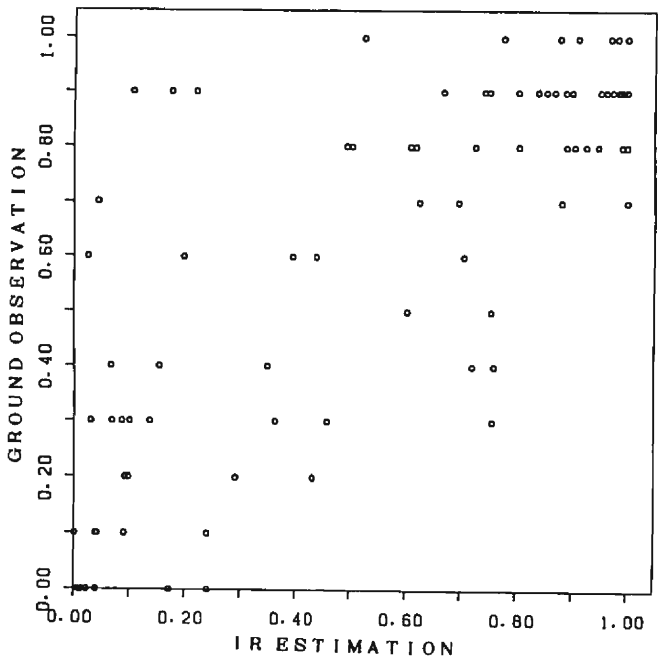


Fig. 6. The comparison between ground observed cloud amounts GC and IR derived fractional coverages (FC) by use of the two-threshold cloud amount estimation method (TTM) for 156 daytime cases. GC and FC are indicated in the range of 0–1.

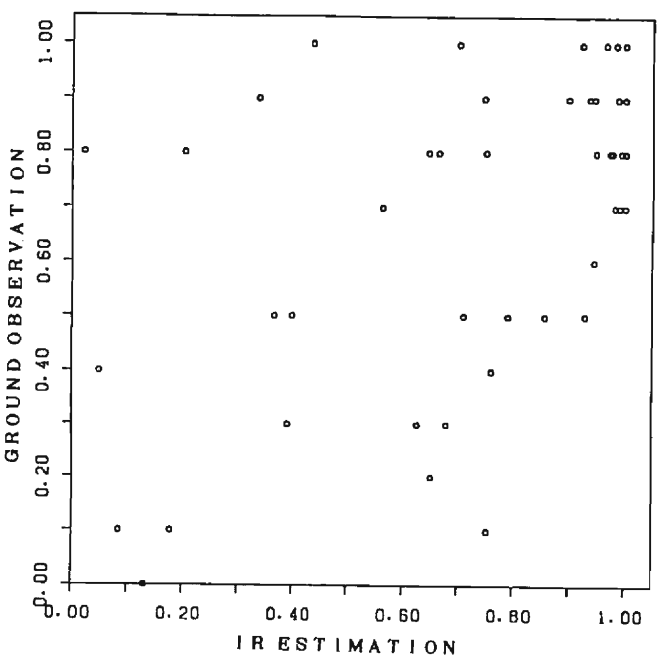


Fig. 7. As in **Fig. 6,** except for 84 nighttime cases.

used, but the differences from other methods are not so large.

The IR cloud amount estimation method developed with the daytime data was applied to estimate the nighttime cloud amounts with the same methods and the results were compared with the surface observations for 84 cases of 12Z. **Fig. 7** shows the scatter diagram of IR estimated and surface observed cloud amounts. In general, the IR estimations were overestimated compared to the surface observations, the correlation coefficient was 0.659, and lower than that for the daytime (0.861). It is considered that the reduction of the coefficient results from an error of ground temperature T_G when cirrus clouds present and errors in observation of clouds from the surface at night.

4. Cloud Type Classification

As described in the introduction, discriminant analysis technique is considered as the most appropriate one for objective cloud type classification among the three methods. In this section, such a method is developed to classify the cloud type for the 6 test areas by using GMS imagery data. Although similar methods have been investigated by several authors (e.g. Parikh and Ball[8]), it is still necessary to apply and test the technique to classify cloud types defined for rainfall estimation, and to attempt to get the best results by trying to select the most effective parameters from the various satellite parameters used and unused in previous studies.

As described by Okuno *et al.*[14] and Sugiyama[15], discriminant analysis technique classifies cloud types as follows: assuming there is a total number of p parameters which characterize cloud type i out of n, and the observed vector $\boldsymbol{X}$ of cloud type parameters is obtained, the probability to obtain $\boldsymbol{X}$ when cloud type is i, $Pr(\boldsymbol{X}|i)$, becomes as follows, assuming the p-th degree normal distribution for $Pr(\boldsymbol{X}|i)$;

$$Pr(\boldsymbol{X}|i) = \frac{1}{(2\pi)^{p/2}|\boldsymbol{\Sigma}|^{1/2}} \times \exp\left[-\frac{1}{2}(\boldsymbol{X}-\boldsymbol{M}_i)^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M}_i)\right] \tag{10}$$

where $\boldsymbol{M}_i$ is the mean vector for cloud type i and the covariance matrix $\boldsymbol{\Sigma}$ is assumed to be equal for every cloud type. The discriminant analysis technique assigns a satellite observation with parameter vector $\boldsymbol{X}$ to cloud type j, when

$$Pr(\boldsymbol{X}|j) > Pr(\boldsymbol{X}|i), \quad \text{for all} \quad i \neq j \tag{11}$$

A cloud type classification method was thus developed by using GMS IR and/or VIS imagery data for 00Z and 06Z and the classification results were evaluated by the comparison with the cloud types obtained from the ground.

4.1 Determination of the Cloud Types

The cloud types to be classified are different case by case according to the purpose of the study. Koffler *et al.*[1] classified 3 cloud types based on the cloud top temperatures. Liljas[2] identified 6 cloud types according to the temperatures and

Table 1. Criteria for selection of the cloud type classification training cases

Cloud Type		Number of Selected Case	Selection Criteria
Sign	Name		
A	Cumulus	18	NT=10, Cu>7, No Cb
B	Cumulonimbus	9	NT=10, Cb>5
C	Middle Cloud	28	NT=10, As+Ac>7, No Cb
D	High Cloud	15	NT=10, Ci+Cs+Cc>7, No lower cloud

albedos of the clouds. Harris and Barrett[5] classified surface and 4 cloud types: stratiform, stratocumulus, mixed and cumulus clouds. Pairkh[6] discriminated 4 types of low clouds, mixed clouds, cumulonimbus and cirrus. He later[7] divided cirrus clouds into cirrus unmixed with other clouds and cirrus mixed with other clouds. Parikh and Ball[8] designed a method for classifying 5 types: low clouds only, middle clouds with no significant high clouds, high clouds with no significant lower clouds, high clouds with significant lower clouds and Cb.

In the present study, 4 cloud types as defined in **Table 1** were chosen for classification, considering the future application of the discriminant results to the rainfall estimation technique. In **Table 1**, type A, B, C and D denote cumulus, cumulonimbus, middle clouds and high clouds, and NT, Cu, Cb etc. are total cloud amount and cloud amounts for every cloud form observed from the surface, respectively. 70 cases (training cases) that satisfy the selection criteria described in **Table 1** were selected from the 156 cases of 00Z and 06Z based upon the ground observed cloud types, and used to develop the cloud type classification method. Among the 70 selected cases, type A, B, C and D include 18, 9, 28 and 15 cases, respectively.

4.2 Method of Classification

264 parameters for classifying the cloud types were selected and calculated from the total histograms, quadrant histograms, difference histograms[6),7)] and Roberts Gradient histograms[8)] of IR and VIS imagery data. 80 of them are spectral parameters reflecting ground or cloud information such as mode, mean, maximum and minimum, and 184 of them are textural parameters reflecting ground or cloud smoothness information.

The best combinations for classifying the 4 cloud types were then selected from the calculated 264 parameters by using the forward selection technique described in Okuno *et al.*[14)] and Sugiyama[15)]. As a result, it was revealed that 1) the best classification results are obtained when both IR and VIS parameters were used, and the worst results appeared when only IR parameters were used, 2) better classification resulted using more parameters but the improvement becomes less significant when more than 4 parameters were used. The selected best 4 parameter combination for IR data is given in **Table 2** and $TBB_{0\%}$, $TBB_{10\%}$, $TBB_{50\%}$ and $TBB_{90\%}$ are TBB values at 0%, 10%, 50% and 90% points in a cumulated TBB histogram,

Table 2. Combination of selected IR parameters

No.	Parameter Definition
1	Coefficient of Variation (S.D./Mean)
2	$TBB_{90\%}$—$TBB_{10\%}$
3	$TBB_{50\%}$—$TBB_{0\%}$
4	90% Value in Roberts Gradient Histogram

respectively.

The 4 parameters selected for IR, VIS and IR and VIS were used to classify the 4 cloud types. On the assumption that the observed parameters follow the 4-dimensional normal distributions and have an equal covariance matrix, discrimination of cloud type was made consequently by comparing the values, D_i for every cloud type. $\boldsymbol{X}$ was then assigned to the cloud type j with the maximum D_i,

$$D_i = \frac{1}{2}\boldsymbol{X}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_i + \frac{1}{2}\boldsymbol{M}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{X} - \frac{1}{2}\boldsymbol{M}_i^T\boldsymbol{\Sigma}^{-1}\boldsymbol{M}_i \qquad (12)$$

From the values of $\boldsymbol{\Sigma}$ and $\boldsymbol{M}_i$ calculated for the parameters of the selected 70 training cases of 00Z and 06Z, the coefficients on the right side of eq. 12 (classification equations) were determined. Cloud type classification was then made simply by substituting the observed parameter values into the equations.

4.3 Results of Classification

The cloud type classification method was then applied to classify cloud types for all cases of 00Z and 06Z, including the cases in which cloud type was not clearly defined.

The additional cases, in which ground observed cloud type was not clearly defined as shown in **Table 1**, were classified into additional categories as shown in **Table 3**. Type S (clear sky) are those with cloud amounts smaller than 0.3, Type F (fraction) are those with cloud amounts GC 0.4–0.6, and Type MA (Mixed A), MB (Mixed B), MC (Mixed C) and MD (Mixed D) are those with cloud amounts over 0.7 and covered mainly by the clouds of Type A, B, C and D but not defined as A, B, C and D in **Table 1**.

The satellite cloud type classification was performed in two steps. At first, the cloud amount was estimated by using TTM described in Chapter 3. The cases having cloud amount estimations smaller than 0.3 and 0.7 were then classified as Type S and F, respectively. Secondly, the classification equations described in (4.2) for 4 types were then applied to the cases with cloud amounts not smaller than 0.7.

Table 3 shows the comparison of the results of satellite classified cloud types using IR parameters with the ground observed data. Among the 70 training cases, 5 cases were classified as Type F, 37 cases were classified correctly, giving a correctness ratio of 52.9%. Among the 50 mixed cloud cases, 5 MA cases, 2 MB cases,

Table 3. Cloud type classification result from IR data for daytime

Type	Discriminated Type						
	S	F	A	B	C	D	Total
S	20 (83.3)	2 (8.3)	0 (0.0)	0 (0.0)	2 (8.3)	0 (0.0)	24
F	4 (33.3)	4 (33.3)	0 (0.0)	0 (0.0)	4 (33.3)	0 (0.0)	12
A	0 (0.0)	0 (0.0)	11 (61.1)	0 (0.0)	7 (38.9)	0 (0.0)	18
B	0 (0.0)	1 (11.1)	2 (22.2)	4 (36.4)	1 (11.1)	1 (11.1)	9
C	0 (0.0)	2 (7.1)	15 (53.6)	0 (0.0)	11 (39.3)	0 (0.0)	28
D	0 (0.0)	2 (13.3)	0 (0.0)	0 (0.0)	2 (13.3)	11 (73.3)	15
MA	2 (14.3)	3 (21.4)	5 (35.7)	1 (7.1)	1 (7.1)	2 (14.3)	14
MB	1 (33.3)	0 (0.0)	0 (0.0)	2 (66.7)	0 (0.0)	0 (0.0)	3
MC	1 (5.9)	0 (0.0)	1 (5.9)	5 (29.4)	8 (47.1)	2 (11.8)	17
MD	0 (0.0)	3 (18.8)	2 (12.5)	2 (12.5)	0 (0.0)	9 (56.3)	16
Total	28 (17.9)	17 (10.9)	36 (23.1)	14 (9.0)	36 (23.1)	25 (16.0)	156 (100.0)

8 MC cases and 9 MD cases were classified as Type A, B, C and D, respectively. In general, classifications of type B (Cb) and D (high clouds) matched comparatively well, but misclassification between types A (Cu) and C (middle clouds) was significant. Misclassification may result from the similarility in appearances of the two types of clouds when observed from a satellite.

The cloud type classification method developed for the daytime data were also applied to classify the cloud types for the nighttime (12Z) cases which the surface cloud observations are less reliable. The comparison between IR classified and ground observed cloud types are shown in **Table 4**. Results of classification were similar to that for the daytime cases. Among the 44 cases of cloudy condition with single cloud type, 21 cases were classified correctly and the correct ratio decreased to 47.7%, among the 23 mixed cloud cases, 3 MA cases, 2 MC caess and 2 MD cases were classified as Type A, C and D, respectively.

As seen above, cloud classification by GMS IR imagery data was acceptably good and even in mixed cloud cases, the main cloud form was detectable except for recognition of middle clouds from cumulus clouds, and vice versa.

Table 4. Cloud type classification result from IR data for nighttime

Type	Discriminated Type						
	S	F	A	B	C	D	Total
S	3 (37.5)	0 (0.0)	1 (12.5)	0 (0.0)	4 (50.0)	0 (0.0)	8
F	1 (11.1)	0 (0.0)	2 (22.2)	0 (0.0)	5 (55.6)	1 (11.1)	9
A	2 (10.0)	0 (0.0)	5 (25.0)	7 (35.0)	6 (30.0)	0 (0.0)	20
B	0 (0.0)	0 (0.0)	1 (25.0)	1 (25.0)	1 (25.0)	1 (25.0)	4
C	0 (0.0)	0 (0.0)	2 (13.3)	2 (13.3)	11 (73.3)	0 (0.0)	15
D	0 (0.0)	0 (0.0)	0 (0.0)	1 (20.0)	0 (0.0)	4 (80.0)	5
MA	0 (0.0)	0 (0.0)	3 (25.0)	3 (25.0)	6 (50.0)	0 (0.0)	12
MB	0 (—)	0 (—)	0 (—)	0 (—)	0 (—)	0 (—)	0
MC	0 (0.0)	0 (0.0)	1 (20.0)	2 (40.0)	2 (40.0)	0 (0.0)	5
MD	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	4 (66.7)	2 (33.3)	6
Total	6 (7.1)	0 (0.0)	15 (17.9)	16 (19.0)	39 (46.4)	8 (9.5)	84 (100.0)

5. Applications of the Cloud Amount Estimation and Cloud Type Classification Techniques

The techniques to estimate cloud amounts and to classify cloud types from GMS IR imagery data discussed in the previous sections were tested and applied to a time series of data of six testing areas for the period from 00Z Sept. 3, 1980 to 12Z Sept. 11, 1980.

The results are shown in **Fig. 8 a-f** together with ground observed data at 00, 06 and 12Z. As explained before, the satellite data were observed in 3 hour interval, instead of hourly as at present. During this period, a stationary front extended over almost the entire Japan islands from Chishima island in the direction of northeast-southwest, on Sept. 8 and Typhoon T8013 passed Kyushu Island on Sept. 10.

The cloud amount estimated from the satellite data showed good agreement with the ground observed data, if considered a significant difference of observation times between both methods. The GMS scans the vicinity of the Japanese islands about 20 min earlier than the surface observation.

The cloud type classification from satellite may be significantly in agreement with the ground observed data, except for the confusion of type A and C in satellite classifications. We could not find any significant difference in estimation score in

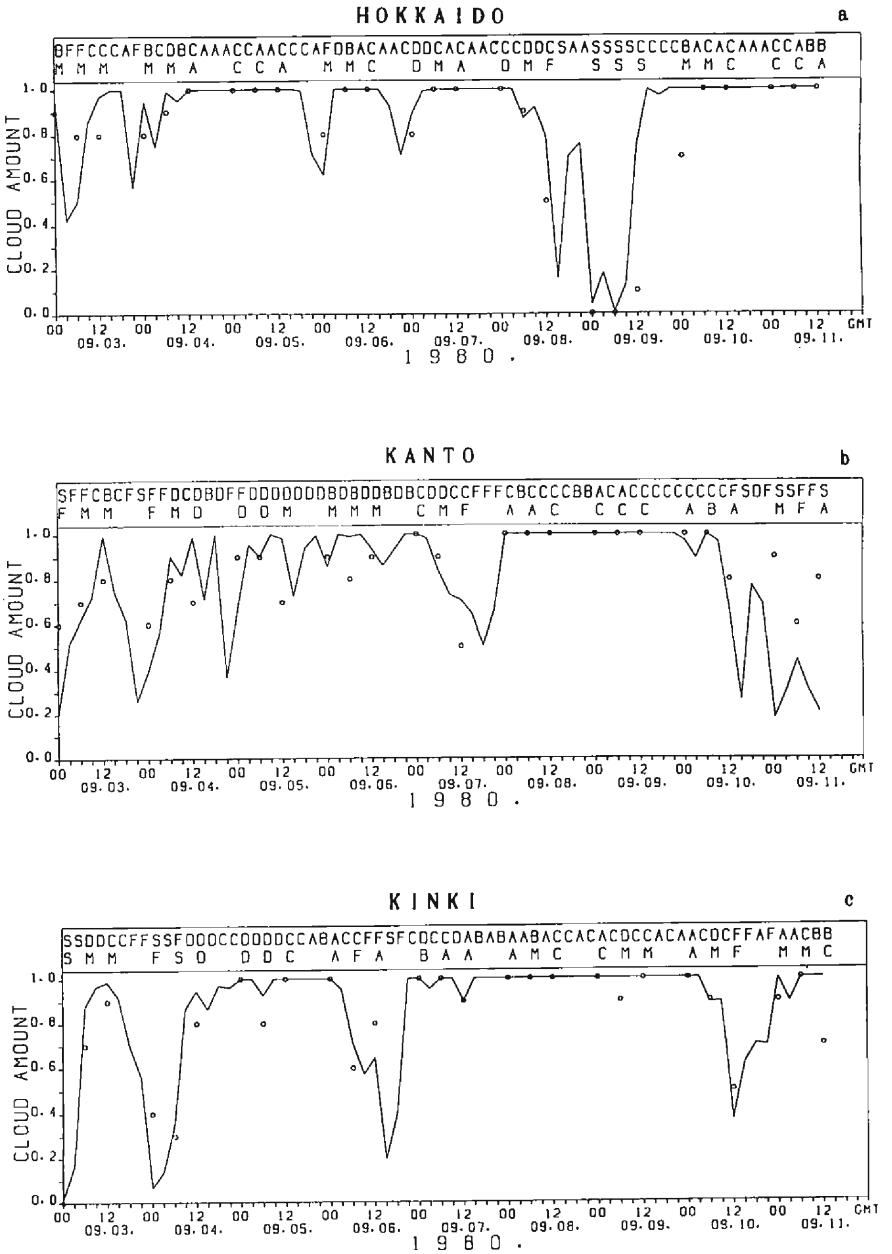
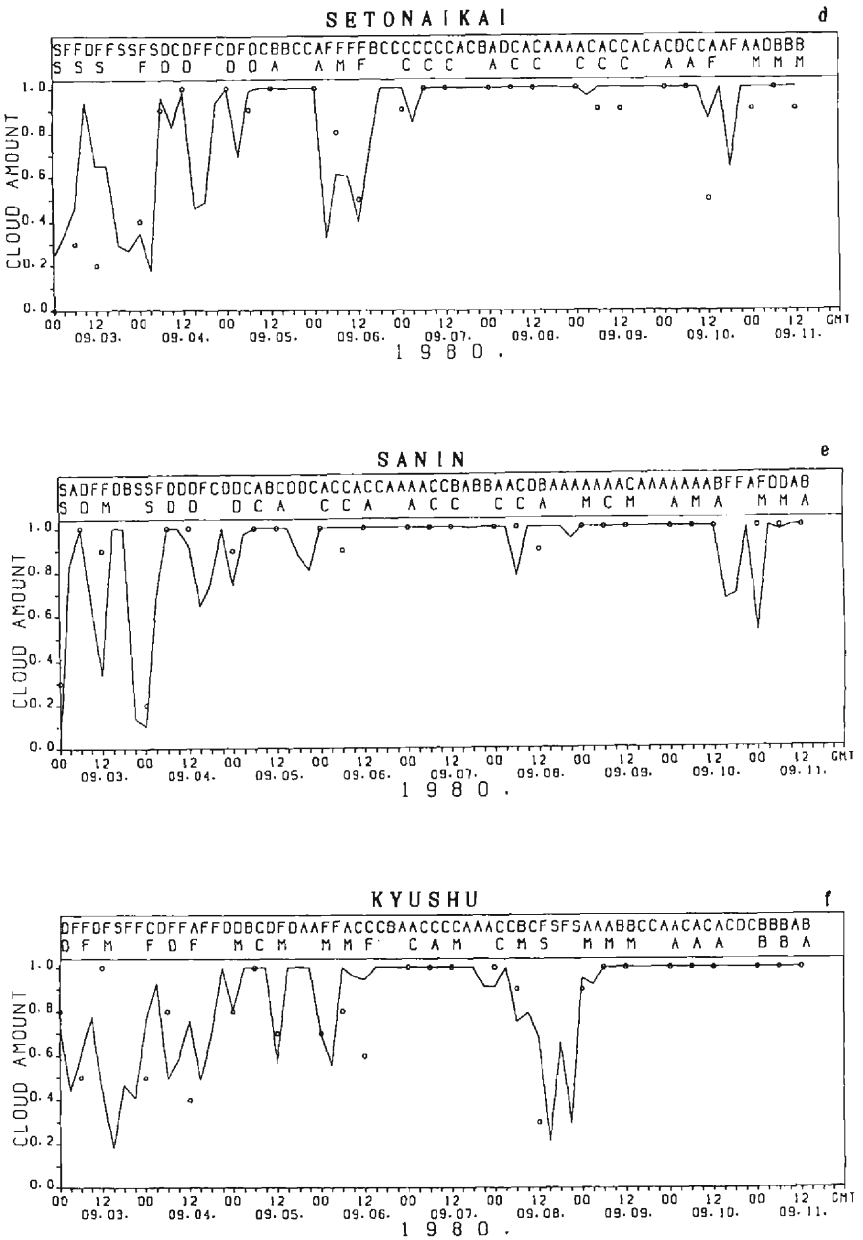


Fig. 8. Time series of IR estimated and ground observed cloud amounts and cloud types for the period of 00Z Sept. 3 to 12Z Sept. 11, 1980, for a) Hokkaido, b) Kanto, c) Kinki, d) Setonaikai, e) Sanin and f) Kyushu. The cloud amounts of ground observation and IR estimation are ploted in circles and solid line,



respectively. The cloud types derived from ground observation and IR data are listed in the lower and upper lines above the cloud amounts charts. Signs of S, F, A, B, C and D denote Clear Sky, Fraction, Type A, B, C and D, respectively. Sign M includes MA, MB, MC and MD.

areal difference from Hokkaido to Kyushu.

6. Conclusions

Methods to estimate cloud amount and classify cloud type for an area of about 10^4 km^2 have been investigated by using IR and/or VIS imagery data of GMS and the surface cloud observation data as reference.

A Two-Threshold-Method (TTM) which takes account of the ground related spectral peak and of the partially cloud-covered pixels has been developed to estimate the total cloud amount from GMS IR and/or VIS imagery data. The satellite estimated and surface observed cloud amounts agree apprecicably with correlation coefficients of 0.861 and 0.659 for the daytime and nighttime respectively when IR data were used.

A discriminant analysis technique was applied to classify 4 types of clouds: cumulus, cumulonimbus, middle clouds and high clouds. Combinations of four parameters are selected to classify the cloud types. The comparisons between the cloud types classified by IR data and ground observed data show correctness ratios of 52.9% and 47.7% for the daytime and nighttime. Confusions of cumulus and middle cloud are the largest cause of misclassification.

The methods were applied to produce the time series of cloud amounts and cloud types for the period of a week by use of IR data. The satellite derived cloud amounts and cloud types showed variations corresponding to the passages of weather systems and their variation trends agree relatively well with those indicated by the surface observations.

To have more reliable classification, the methods should be improved in the introduction of time changes of images in hourly observations, as well as diminishing the effects of cirrus and distinguishing cumulus from middle cloud.

Acknowledgements

The author would like to express his thanks to Prof. Y. Mitsuta of the Disaster Prevention Research Institute (DPRI), for his continuous guidance and encouragements. The author is also grateful to Prof. M. Yamada and Mr. M. Horiguchi of DPRI, for their valuable discussions and encouragement. The computations were performed by FACOM M-730 computer at the Information Processing Center of Disaster Prevention Studies of DPRI of Kyoto University.

References

1) Koffler, R., A. G. Decotiis and K. Rao: A Procedure for Estimating Cloud Amount and Height from Satellite Infrared Radiation Data, Mon. Wea. Rev., Vol. 101, 1973, pp. 240–243.
2) Liljas, E.: Automated Technique for the Analysis of Satellite Cloud Imagery, In "Nowcasting", 1982, pp. 167–176, Acadmedic Press.
3) Desbois, M., G. Seze and G. Szejwach: Automatic Classification of Clouds on METEOSAT

Imagery: Application to High-Level Clouds, J. Appl. Met., Vol. 21, 1982, pp. 401–412.
4) Seze, G. and M. Desbois: Cloud Cover Analysis from Satellite Imagery Using Spatial and Temporal Characteristics of the Data, J. Climate Appl. Met., Vol. 26, 1987, pp. 287–303.
5) Harris, R. and C. Barrett: Toward an Objective Nephanalysis, J. Appl. Met., Vol. 17, 1978, pp. 1258–1266.
6) Parikh, J.: A Comparative Study of Cloud Classification Techniques, Remote Sensing of Environment, Vol. 6, 1977, pp. 67–81.
7) Parikh, J.: Cloud Classification from Visible and Infrared SMS-1 Data, Remote Sensing of Environment, Vol. 7, 1978, pp. 85–92.
8) Parikh, J. and J. T. Ball: Analysis of Cloud Type and Cloud Amount during GATE from SMS Infrared Data, Remote Sensing of Environment, Vol. 9, 1980, pp. 225–245.
9) Takahashi, D.: Navigation of VISSR Imagery by Detecting the Earth Edge, Meteorologiacl Satellite Center Technical Report, Vol. 3, 1980, pp. 55–68 (in Japanese).
10) Shenk, W. E. and V. V. Salomonson: A Simulation Study for Exploring the Effects of Sensors Spatial Resolution, J. Appl. Met., Vol. 11, 1972, pp. 214–220.
11) Takeda, T. and F. Hattori: Analysis of Low Clouds over Sea Using NOAA Data, Preceeding of 1988 Spring Conference of JMS, 1988, pp. 30 (in Japanese).
12) Xie, P. and Y. Mistsuta: Rainfall Estimation from GMS Infrared and Visible Imagery Data, Annuals of Disas. Prev. Res. Inst., Kyoto Univ., No. 32B-1, 1988, pp. 201–217 (in Japanese).
13) Kubota, K. and A. Endo: Digitalization of Cloud Distribution, Meteorological Satellite Data and Numerical Forcasting, Report of Computer Division, No. 26, 1980, pp. 30–33 (in Japanese).
14) Okuno, M. et al.: Multivariate Analysis, Nichikagiren, 1976, pp. 299 (in Japanese).
15) Sugiyama, T.: Introduction to Multivariate Analysis, Asakurashoten, 1983, pp. 173 (in Japanese).