Preprint Volume, Fifth Conference on Applied Climatology, AMS March 9-13 1987, Baltimore Maryland.

Summertime Cloud Climatologies Constructed from Satellite Imagery

Frank P. Kelly, John F. Weaver, Paul W. Mielke, Jr., Thomas H. Vonder Haar

Colorado State University
Fort Collins, Colorado

1. INTRODUCTION

A previous paper (Klitch, et al., 1985) described the results of a study in which full resolution GOES data over Colorado were computer averaged. Results included two different types of averages (also called "composites"). The first consisted of a simple PIXEL by PIXEL average of visual imagery from coincident times of day during July and August 1982. The "average images" were constructed for four times -- 1700, 1900, 2100 and 2300 GMT. Additionally, four other VIS composites, consisting of images from severe weather days only, were made and compared for the same times. The second type of composite was an average of interactively-altered imagery. In this case, information from both VIS and IR data were combined to identify regions having either 1) deep convection (defined as $<-30\,^{o}C$ in the IR and subjectively determined "bright" cloudiness in the VIS) versus clear regions versus all other cloudiness. A new image was created in which the deep convection was colored white, the clear ground black, and the "other cloudiness" regions a deep gray. These images were then averaged together as were the VIS images.

There were several interesting results in the study. The composites illustrated the potential for developing a highly detailed climatology of topographically controlled convection. The data furnished a much more dense observation network than is typically available. For example, Colorado (CO) has nearly 200 precipitation reporting stations statewide. However, since Colorado covers an area of roughly 2.67 (10)5 km^2, the average coverage is one observation every 1,335 km^2. GOES pixels range from 2 km^2 (for 1/2 mi VIS) to 80 km^2 (for IR) at the latitude of CO. Consequently, the satellite coverage has between 17 and 668 times more dense than the statewide precipitation reporting network. This is not to imply that satellite can measure the same variables as surface networks; only that, for what they do measure (i.e., total cloud cover, cloud type, location of deep convection, etc.), the sensor yields satellite measurements with very high resolutions.

By allowing three categories in the bi-spectrally altered imagery, an ambiguity was introduced which created a problem in making definitive statements. Klitch, et al. (1985) suggested future work should create binary composites (e.g., cloud/no cloud, convection/no convection, etc.) to avoid these problems. It was stated the summer 1982 study was merely a first step of what they hoped would become a sequence of larger based climatological studies over different regions of the country (especially where geographic features force repetitive convective cycles).

Following the 1982 project, a study utilizing similar averaging was completed by another group at Colorado State University (McQueen and Pielke, 1985). That study described deep convective cloud patterns over southern Florida on synoptically undisturbed days. The authors subdivided their data into four common wind flow regions. Composite results were compared to a three dimensional mesoscale model (Pielke, 1974; Pielke and Mahrer, 1978) which incorporates sea breeze forcing, synoptic flow patterns, and ground characteristics. The composited data helped interpret model results in relation to general convective behavior patterns.

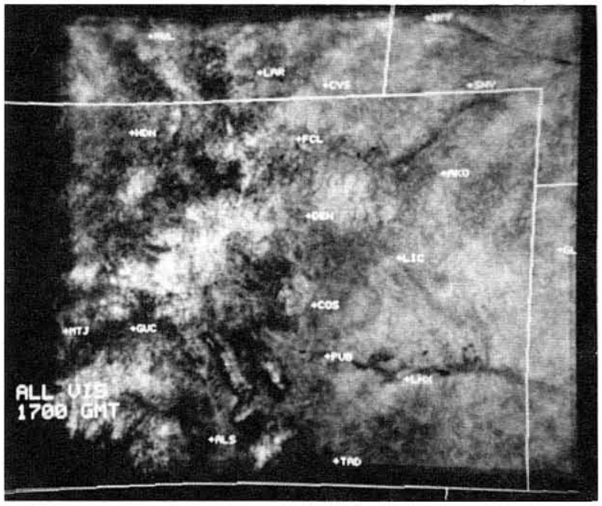
The summer of 1985 marked the beginning of a new phase to the cloud climatology project. The current plan includes utilizing data collected during July and August 1985 to design new composite types for implementation during June through August of 1986 and 1987.

The first phase of the project is now complete, and the 1985 composites are available. This paper discusses the compositing methodology and qualitative climatological aspects of the composites. Quantitative techniques for satellite composite analysis based on metric distance functions are discussed.

2. New Composites

One of the lessons learned in the first compositing study was ambiguities were easy to come by. Thus, one of the primary concerns was to formulate determinate averages. This was accomplished by altering the data to a binary form whenever possible. A second concern was to present the averaged data in a precise, clear format. That is, even though the averaged arrays give an accurate, point by point frequency count, the subjective appearance of the image is often confusing. It may not look like a satellite image, and may not, therefore, be the best product from the user standpoint. To solve this problem, a secondary set of images was built. For these, a completely cloud-free satellite image of the study area was collected. The interactive, image processing system described in Klitch, et al, (1985) was then used to replace pixels in the clear image wherever the frequency exceeded a chosen threshold. The combination of the frequency image and the clear image made the frequency data appear as clouds as in a normal satellite image.

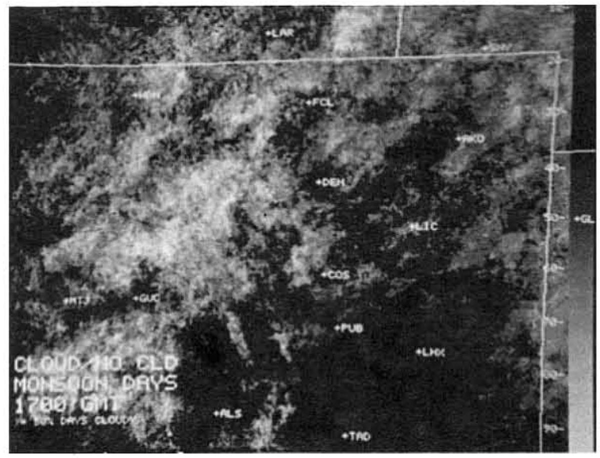
Three types of composites are illustrated here. The first is a simple average of all of the VIS images for the entire data collection period. It encompasses all brightnesses, for whatever cause, and may include regions of repetitively bright ground as well as those regions having frequent clouds. The resulting composite for 1700 GMT is shown in figure 1. It is purposely left in its original format to illustrate the potential for confusion. For example, the bright areas in eastern CO are actually "mostly clear" areas where the ground has a high albedo, while the bright regions in western and central CO result from frequent cloudiness. Consequently, the interpretation is ambiguous, and may mislead an operational user.



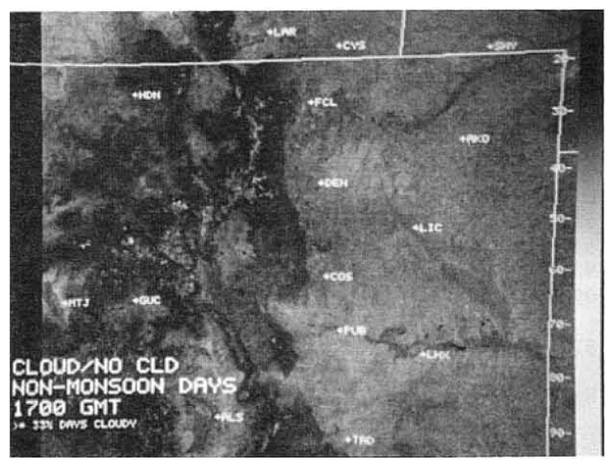
1. Computer average of all VIS imagery collected at 1700 GMT, daily, between 13 July 1985 and 23 August 1985.

One of the several weather stratifications used for this pilot study separates "monsoon" from "non-monsoon" cases. CO receives its moisture for summer thunderstorms from at least two sources--easterly upslope Gulf moisture (or evaporated rain), or advection of Pacific moisture from the west or southwest. The latter has been labeled "monsoon flow" (after Hales, 1974). A monsoon situation normally results from several days of southwesterly 700 mb through 500 mb geostrophic flow. In most years, monsoon flow accounts for about a third of the convective days in CO. For the 1985 data period the ratio was nearly 50 percent. Thus, the decision to separate cases of monsoon/non-monsoon flow days. A "monsoon day" was defined as days having a moist southwesterly flow at 700 mb and/or 500 mb.

Figure 2 is a composite of monsoon day images from 1700 GMT. Before averaging, each image was interactively altered by transforming each cloudy pixel to absolute white and each non-cloudy pixel to absolute black. The average of all such images represents the frequency of cloudiness, at each point, over the entire sample period. We have shown only points that had cloudiness 50 percent of the time or more. For the rest of the image, the "clear ground" image was substituted as described earlier. Note the sharp contrast between figure 2 and figure 3 (which is the same type of composite for non-monsoon days). Monsoon days generally appear to have more moisture. Careful scrutiny of cloud streaks in relation to the wind flow reveals a relationship of cloudiness over the plains in eastern CO and WY, to terrain features (figure 4) in the mountains.

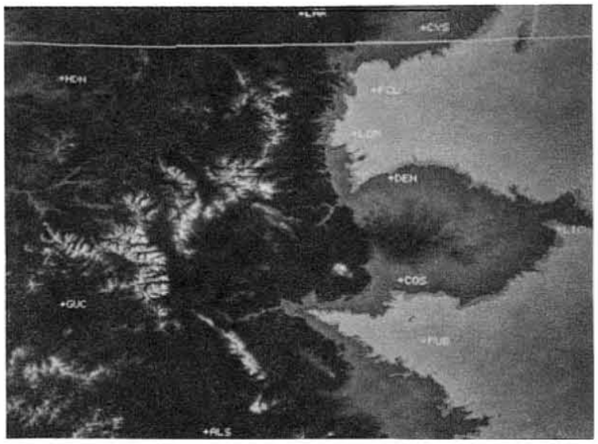


2. Computer average of interactively altered VIS images collected at 1700 GMT, daily, between 13 July 1985 and 23 August 1985 on "monsoon" days (see text). Before averaging, cloudy pixels were turned pure white, and non-cloudy pixels black.



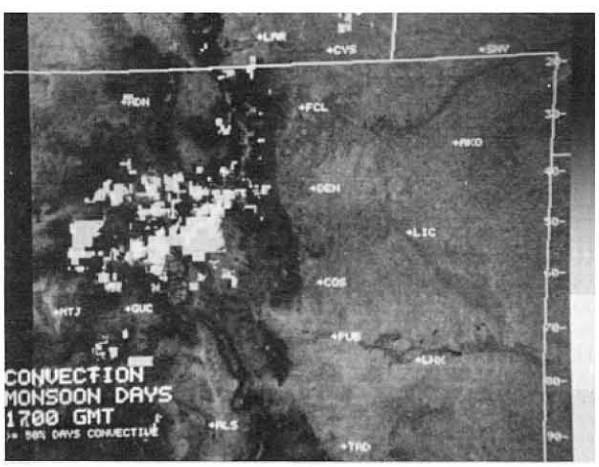
3. Same as figure 2, except for non-monsoon days.

The lower cloud frequency on non-monsoon days in northern Colorado raises an intriguing question in light of findings in the earlier paper. Klitch et al. (1985) showed Colorado severe thunderstorm days differed from non-severe days by the cloud frequency in northern Colorado and southeastern Wyoming. The results showed higher cloud frequencies in those areas on severe days. In the current study, enhancement to the north is found only on "monsoon" days. This result indicates a relationship between mid-level monsoon flow, and severe weather on the high plains. While significant weather reports (Storm Data, 1985) don't fully support this, there was enough of a tendency to warrant further investigation. For example, defining a "significant thunderstorm" as one producing tornadoes, hail of 0.5", damaging winds, or 15 min. rain rates of 1.5"/hr, we found significant storms occurred on 88 percent of the monsoon days, but only 28 percent of the non-monsoon cases. Similarly, 56 percent of the monsoon case days produced severe thunderstorms (tornadoes, hail of 0.75", or winds > 50 kts), while only 24 percent of the non-monsoon had severe thunderstorms. These statistics corroborate the stated relationship but the data set is small and strong inferences are not possible.



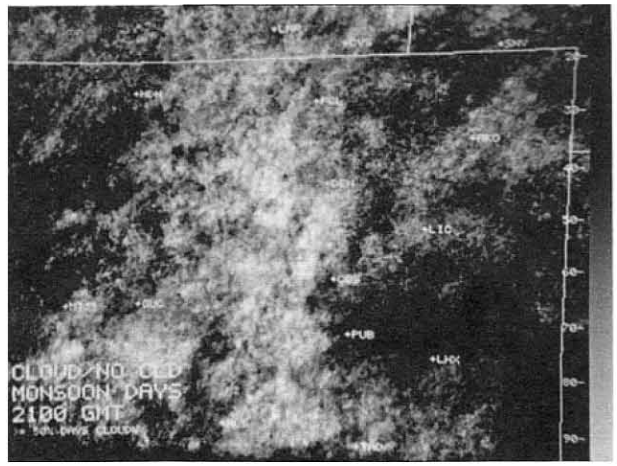
4. Colorado topographic image. The brightest regions are the high ranges of the Rocky Mountains.

By defining "convection" as points having bright clouds on the VIS, and an IR temperature of less than -12C (about 20,000 feet in summer in CO), we were able to construct black/white images representing convection/no convection. Figure 5 is a "convection/no convection composite" of images for the cloud/no cloud data shown in figure 2. Note what a small percentage of the cloudiness at 1700 GMT is actually convection.

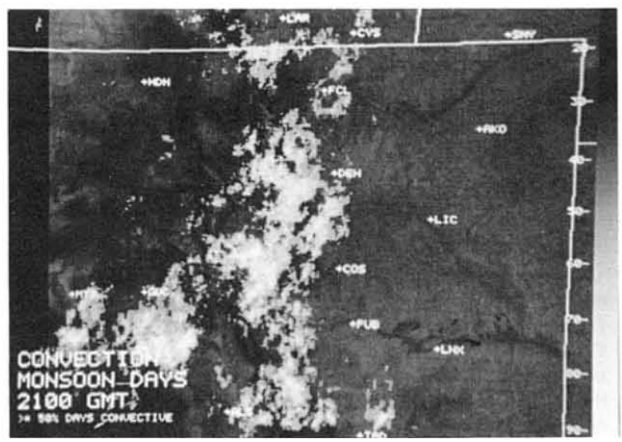


5. Computer average of interactively altered satellite imagery collected at 1700 GMT, daily, between 13 July 1985 and 23 August 1985 on "monsoon" days. Before averaging, convective pixels (see text) were turned pure white, and non-covective pixels black.

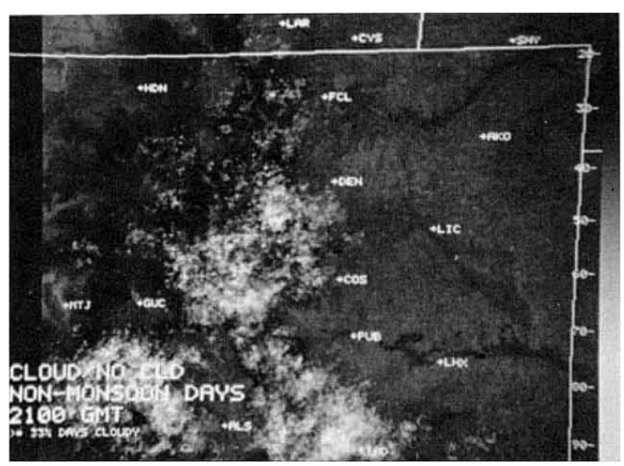
For completeness, figures 6-9 are included to show the subsequent evolution of the monsoon/non-monsoon days.



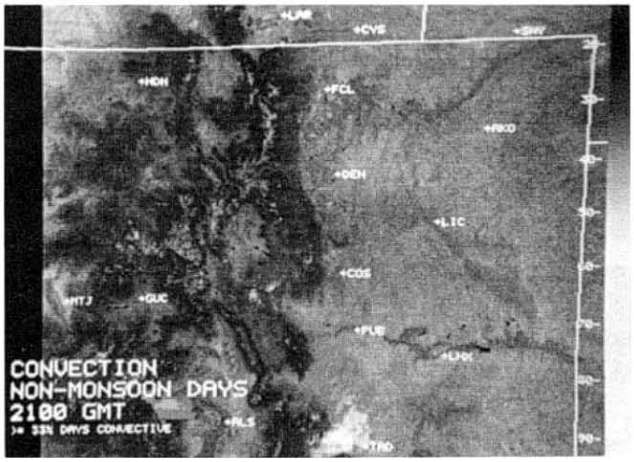
6. Same as figure 2, except composite is for 2100 GMT.



7. Same as figure 5, except composite is for 2100 GMT.



8. Same as figure 3, except composite is for 2100 GMT.



9. Same as figure 5, except composite is for non-monsoon cases at 2100 GMT.

3. Quantitative Analysis Using Metric Distance Functions: A Discussion

Mielke (1984, 1986) describes an analysis technique based on metric distance functions (MDF) whose characteristics can enhance our ability to analyze the raw cloud data. The advantage of the MDF over, for example, a classical structure function (STR) (Gandin 1963) is the simplicity of assumptions required by the MDF. Previous investigations have applied the STR analysis to satellite data (Hillger et al, 1979) with the assumptional constraints inherent in the STR method. The STR scheme requires a normally distributed, homogeneous, and isotropic data field. These conditions rarely, if ever, exist in a raw cloud data field. In contrast, the MDF analysis, since it is based upon appropriate permutations of the raw data (Mielke 1984), only requires that the data be obtained from a random sample. This distribution-free character provides a powerful combination tool for comparison of data fields e.g. VIS and IR imagery.

The MDF analysis also maintains the principle of congruence between the data and analysis spaces (Mielke 1986). Simply, this maintains a physically meaningful correspondence between the raw data and the results of analyses. This is in contrast with the STR analysis which uses mean square differencing in the analysis space resulting in incongruent units between the data and analysis spaces. Consequently, the maintenance of the congruence principle in the MDF analysis case provides for straightforward, realistic physical interpretation of the analyzed data.

The essence of the MDF analysis scheme is the way it provides quantitative statistical information directly applicable to inferential statements. Mielke et al. (1981) describe the basic method to determine the probability structure of the analysis space. Quite simply, the MDF analysis results in a delta-function (δ) which allows for partitioning of data into groups or "clumps". Intuitively, a small value of δ implies the concentration of measurements within a certain sub-group of the total data space. This gives a deterministic value for the probability of existence of that group. (Mielke et al., 1981).

The application of the MDF analysis considered here is first concerned with cloud classification and second with the temporal behavior of the cloud fields. The MDF analysis, with its deterministic δ value, can determine the statistical significance of a particular data grouping. This suggests applying the concept to the classification of bi-spectral satellite cloud observations. Since the underlying data type distribution is not an impediment to VIS and IR comparisons (though data resolution differences must be considered) using MDF ideas, quantitative δ values are determinable for individual bi-spectral cloud groupings. The δ values can then be applied as threshold values on an image processing system to assist in the construction of binary composite fields.

MDF analysis is also applicable to the temporal structure of the derived composite fields. The temporal existence of significant cloud "clumps" as determined through the δ value concept can result in quantitative cloud frequency information distributed over time. This points toward development of a climatological persistence index with the potential for assisting in applying the composited data to estimation models.

4. Summary and Plans for Future Research

The cloud compositing project will continue and include data from several orographically and non-orographically forced regions. The targeted areas will include data from the Great Lakes, Colorado, Alabama, southern Florida/Caribbean Islands, Gulf of Mexico/Central America, and the southeastern US centered on Georgia. Presently, the primary stratification categories planned are 1) averages of all data, 2) a monsoon/non-monsoon breakdown, and 3) severe versus non-severe thunderstorm days. Subsets planned for each stratification will include a) cloud/no cloud images, b) convection/no convection, and c) stratiform/fog versus "other". Additionally, half minute resolution, topographic data over the study area will be remapped and matched for direct comparison.

The generation of composites for these areas will assist in the development of short range cloud forecasting techniques. The composites will provide a basis for the first guess estimation of cloud development and motion. Empirical studies relating the derived composites to the dynamics of the cloud field will assist in the forecasting of the subsequent cloud field through the de-weighting of parameters used to drive dynamic estimation models.

The goal of the MDF analysis concept is to provide a logical, physically realistic method for data compositing. The permutation structure of the analysis provides for the direct quantitative inter-comparison of diverse types of meteorological and climatological data. The application to cloud compositing efforts is ongoing. The software and hardware required to meet these goals is still in flux. Specific results, to date, are not definitive but are encouraging. Future efforts will streamline the compositing process by using MDF based automated cloud classification schemes. The consequent application to climatological estimation products is a foreseeable result of this research.

5. Acknowledgements

The authors would like to thank Debra Lubich of CIRA for her assistance in streamlining much of the software used in this study. The research is funded in part by NOAA/ERL Grant No. NA84AA-D-00017, NA-85-RAH-05045, and The United States Air Force AFIT Program.

6. References

Gandin, L.S., 1963: Objective Analysis of Meteorological Fields. Leningrad, Hydrometeor. Publ. House. (English Translation, Israel Program for Scientific Translations, 1965-242pp.)

Hales, J.E., Jr., 1974: Southwestern United States summer monsoon source- Gulf of Mexico or Pacific Ocean? J. Appl. Meteor., 13, 331-342.

Hillger, D.W. and Thomas H. VonderHaar, 1979: An analysis of satellite infrared soundings at the mesoscale using statistical structure and correlation functions. J. Atmos. Sci., 36, 287-305.

Klitch, M.A., J.F. Weaver, F.P. Kelly, and T.H. Vonder Haar, 1985: Convective cloud climatologies from satellite imagery, Mon. Wea. Rev., 113, 326-337.

McQueen, J.T., and R.A. Pielke, 1985: A numerical and climatological investigation of deep convective cloud patterns in South Florida. Atmos. Sci. Paper No. 389. Dept. of Atmos. Sci., Colorado State University, Fort Collins, CO 177pp.

Mielke, P.W., Jr., 1986: Non-metric statistical analyses: Some metric alternatives, J. Statis. Plann. Inference 13, 377-387.

------, 1984: Meteorological applications of permutation techniques based on distance functions. In: P.R. Krishnaiah and P.K. Sen, Eds, Handbook of Statistics, Vol. 4. North-Holland, Amsterdam, 813-830.

------, K.J. Berry and G.W. Brier, 1981: Application of multi-response permutation procedures for examining seasonal changes in monthy mean sea-level pressure patterns. Mon. Wea. Rev., 109, 120-126.

Pielke, R.A., 1974: A three-dimensional numerical model of the sea breezes over South Florida. Mon. Wea. Rev., 102, 115-139.

------, and R. Mahrer, 1978: Verification analysis of the University of Virginia three-dimensional mesoscale model prediction over South Florida for 1 July 1973. Mon. Wea. Rev., 106, 1568-1589.

Storm Data, 1985: NOAA/NESDIS (ISSN 0039-1972), Vol. 7, NR. 7,8.