NOAA Technical Memorandum ERL ESG-25

THE ESTIMATION FROM SATELLITE IMAGERY OF SUMMERTIME RAINFALL OVER VARIED SPACE AND TIME SCALES

Cecilia Girz Griffith

Environmental Sciences Group
Boulder, Colorado
April 1987

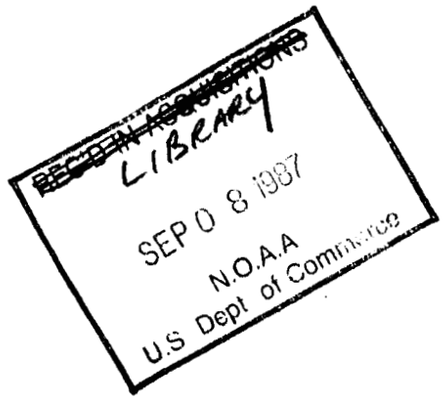
noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

NOAA Technical Memorandum ERL ESG-25

THE ESTIMATION FROM SATELLITE IMAGERY OF SUMMERTIME RAINFALL OVER VARIED SPACE AND TIME SCALES

Cecilia Girz Griffith
Weather Research Program

Environmental Sciences Group
Boulder, Colorado
April 1987



UNITED STATES DEPARTMENT OF COMMERCE

Malcolm Baldrige, Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION

Anthony J. Calio, Administrator

Environmental Research Laboratories

Vernon E. Derr, Director

National Oceanic and Atmospheric Administration
TIROS Satellites and Satellite Meteorology

ERRATA NOTICE

One or more conditions of the original document may affect the quality of the image, such as:

Discolored pages
Faded or light ink
Binding intrudes into the text

This has been a co-operative project between the NOAA Central Library and the Climate Database Modernization Program, National Climate Data Center (NCDC). To view the original document contact the NOAA Central Library in Silver Spring, MD at (301) 713-2607 x124 or Library.Reference@noaa.gov.

HOV Services
Imaging Contractor
12200 Kiln Court
Beltsville, MD 20704-1387
January 26, 2009

NOTICE

Mention of a commercial company or product does not constitute an endorsement by NOAA Environmental Research Laboratories. Use for publicity or advertising purposes of information from this publication concerning proprietary products or the tests of such products is not authorized.

For sale by the National Technical Information Service, 5285 Port Royal Road
Springfield, VA 22161

CONTENTS

LIST OF TABLES

LIST OF FIGURES

THE ESTIMATION FROM SATELLITE IMAGERY OF SUMMERTIME RAINFALL OVER VARIED SPACE AND TIME SCALES[1]

Cecilia Girz Griffith

ABSTRACT. Rainfall estimates for the central United States, inferred from the thermal infrared channel of the Geostationary Operational Environmental Satellite (GOES) for August 1979, are compared with gage rainfalls over hourly and daily time frames. Area-averaged amounts are computed for the region 90.25° to 108.25°W and 30.5° to 45.75°N, which extends roughly from the Rocky Mountains to the Mississippi River and from the North Dakota-South Dakota border into central Texas, covering 3.6×10^6 km^2. Point values within this region are also compared. Estimates are made with two versions of the satellite technique. One incorporates cloud life cycle information derived from a sequence of images; the other uses area and temperature information from a single image. Each version of the satellite technique uses relationships derived in Florida. Subsequently the rain estimates are corrected for environmental differences between Florida and the Great Plains by a one-dimensional cumulus model. In general the environmentally corrected satellite rainfalls tend to be smaller by 20-40% than the corresponding gage amounts. Root-mean-square errors are approximately 1 mm for daily and 0.1 mm for hourly area-averaged rainfalls, and 14 mm for daily and 5 mm for hourly point rainfalls. Because of the relatively coarse spatial resolution of the satellite data, comparisons with point gage rainfalls are not recommended; correlations are small (<0.2). Satellite-gage differences of daily (hourly) point values can be large, but 50% of the satellite amounts are within ±4 (±2) mm of the gage amount, and 90% are within ±20 (±8) mm. Timing of rain relative maxima is coincident or differs by one time period for the majority of the area-averaged cases, but the hourly area-averaged satellite data exhibit considerably fewer short-term fluctuations than the gage data. For the two satellite algorithms tested, the streamlined (single image) satellite technique requires 10% of the computation time needed by the life history technique and shows little difference in its performance as assessed by the gages. Image frequency greatly affects rainfall patterns, even in an analysis for the month. Computations based on 3-hourly or more frequent imagery result in the most realistic patterns.

1. INTRODUCTION

Water is vital to life and an essential resource for many human activities. The measurement of its movement throughout the water cycle is therefore very important. Still, the estimation of rainfall over the globe is a difficult problem. Unlike most meteorological parameters, rainfall is

---

[1] From a thesis submitted to the Academic Faculty of Colorado State University in partial fulfillment of the requirements for the Ph.D.

discontinuous in space and time and exhibits large natural variability. Currently deployed observing systems such as rain gages and radar are generally limited to the measurement of precipitation over land, and in these networks the density of gages and the spacing of radars varies across political boundaries and sometimes within nations. Yet the greater part of the globe is covered not by land but by ocean. Because these vast expanses exist with little or no permanent human activity, oceanic rainfall is more frequently extrapolated from other data than actually measured over seasonal or annual time frames. Few studies focus on daily oceanic rainfall.

Satellites have been touted as a means to circumvent some of the difficulties attendant in gage and radar measurement of rainfall. From the satellite platform large regions can be viewed simultaneously and, under certain conditions, frequently. For instance, the tropics and middle latitudes can be viewed frequently from geosynchronous orbit, and the high latitudes from a polar orbit. Further advantages of the satellite platform are that a number of meteorological satellites are now in orbit and, with computer processing of digital data, timely estimates can be made at homogeneous densities for large regions. In some regions the spatial resolution of the satellite sensor is finer than the resolution of present ground rainfall networks. One substantial drawback is that no satellite will ever carry sensors that directly measure rain, but a number of techniques have been developed to infer rainfall from visible or infrared data, or to relate it more physically to radiation emitted at microwave frequencies.

The literature on satellite precipitation estimation is too extensive to review in depth here. Barrett and Martin (1981) provided an excellent synthesis of the numerous rainfall techniques and related cloud studies developed during the first 20 years of meteorological satellites. However, every technique discussed in Barrett and Martin is designed to estimate rain from convective systems. New techniques continue to be developed for convective rainfall, such as the VIS and IR technique of Martin and Howland (1986) and the microwave work of Spencer (1986). In addition, work on snowfall and precipitation in cold-season extratropical cyclones has begun; Fenner (1982), Neil (1984), Scofield and Spayd (1984), and DelBeato and Barrell (1985) are examples.

There are a number of notable problems in satellite rain estimation. Some occur because of the limitations of the satellite sensors, some because of the meteorology. Other problems arise in the verification process or in the philosophy of the technique and its application. Often there is no clear-cut delineation among these categories. For instance, there is an interplay between the satellite limitations and the meteorology. Convective phenomena occur over a wide range of time and space scales. Depending on whether meso-α- or meso-β-scale convection is of interest, convection may be well or ill sampled within the constraints of the satellite orbit and sensor characteristics. The diurnal cycle is another example of this interplay. If one has access only to data from a polar-orbiting satellite to estimate precipitation in the tropics, one visible image per day (or two thermal infrared) at a fixed crossing time is certainly not sufficient to capture the diurnal cycle. A third example is the importance of time-change characteristics of clouds. Do they need to be incorporated into the technique? Are they sampled with the available data?

Likewise there is interaction between verification approaches and technique philosophy. The verification procedures should obviously be driven by the application for which the technique was derived, but they are also driven by the type of verification data available. Hourly satellite estimates cannot be verified with daily gage data or point estimates with radar data. Limitations of ground data and assumptions in derived quantities (gage-generated area-averaged amounts, for example) must be discussed.

Philosophy of technique construction involves many interesting questions. Should data of only one type be used or is it proper (and profitable) to mix data types? For example, is IR alone preferable to a technique that incorporates visible during the 12 or fewer hours that it is available? Should satellite data and ground data (i.e., gages or radar, or both) be incorporated into a technique? Are rainfall rates or amounts to be estimated and, if the latter, over what periods? Or will the technique be a classification scheme denoting rain and no rain; or a categorization scheme of no rain, light, moderate, and heavy rain? And what is the role of a human being in the technique? Should the scheme be totally automated or should it contain subjective components that will be satisfied through human input? Answers to all these questions can be given only in the context of the application at hand and the satellite data source.

Stated in this way these concerns appear confused and muddled. They can be sorted out, however, through the verification. The focus should be on rainfall quantity, rainfall timing, and rainfall patterning. Rainfall quantity means amount, rain rate, or rain category. Is the rainfall quantity similar in the satellite and ground data sets, and if not, how large are the differences? The timing of rainfall maxima and minima in the satellite technique and the conventional data is the second aspect of verification. Do the time series show major and minor rain events occurring concurrently? If not, is there a relative bias? Are there periodicities in the time series, are these meteorologically based, and do the time series from both sensors reveal them? Lastly, how comparable are the rainfall patterns derived from the ground data and from the satellite technique? Are the maxima and minima in the same location? Are there preferred regions of rainfall and of no rainfall in both the satellite and ground estimates? Is there a systematic shift between cloud top (which the satellite senses) and the ground or near ground (where the surface data observe) that is apparent? Can this shift be traced to the physics of the sensor or to the meteorology that produces the precipitation?

This paper focuses on the estimation of warm season convective rainfall from infrared satellite data with a view toward large-area, operational use. In particular the region of interest is the central third of the United States during August 1979, and the satellite is the Geostationary Operational Environmental Satellite (GOES). One emphasis is a comparison of GOES rain estimates with gage estimates; these comparisons are geared toward addressing the outstanding questions of verification. A second emphasis relates to the construction of the satellite technique first described in Griffith et al. (1978). Two aspects of this technique are tested: (1) the contribution of time-dependent terms, and (2) the feasibility of using the sparse, twice-daily operational radiosonde data to create an environmental correction over a large area at the resolution of the satellite data. Time dependency is assessed with two versions of the satellite technique. One requires a sequence of

images to capture each cloud's life cycle prior to starting a rainfall computation; the other operates on satellite parameters measurable from a single image.

The gage comparisons herein are in contrast with three current satellite techniques that operationally produce estimates of warm-season, convective rainfall over large areas and have done so for longer than one year. For lack of a better convention, these techniques are usually referred to by the names of their creators: the Scofield-Oliver technique, the Kilonsky-Ramage technique, and the Arkin technique. In each of these techniques for differing reasons, little verification has been done to date.

The Scofield-Oliver (Scofield and Oliver 1977) technique was originally devised as a subjective technique based on hard-copy visible and infrared imagery; it has subsequently been automated in part (Clark and Perkins, 1985). Its use is not for the routine estimation of rainfall, but for the identification of flash-flood-producing storms with an emphasis on point rainfalls. The technique is based on a decision tree in which point rainfall amounts are increased from a base value depending on the presence of a number of meteorological signatures assessed from satellite and conventional data sources. These signatures include approximately 40 items, such as cloud shape, cloud rate of change, cloud lifetimes, low-level inflow, and atmospheric moisture.

There are many unanswered questions with regard to this technique. Although it has been taught to forecasters and used operationally for close to a decade, nothing has yet appeared in the formal literature describing the derivation of the technique or its verification. It is not known, for instance, how the technique's rainfall amounts were derived, whether the inclusion of the visible imagery influences the accuracy of the estimates, if the technique exhibits a strong bias that is dependent on the analyst, or even how accurate satellite point rainfall estimates can be. A paper by Scofield (1987), now in review, is eagerly awaited for discussion of these issues.

The Kilonsky-Ramage and Arkin techniques aim at providing climatological rain estimates for the tropical oceans. The Kilonsky-Ramage technique (Kilonsky and Ramage, 1976) uses the visible picture from the NOAA polar-orbiting satellite to estimate monthly rainfall for 1° latitude by 1° longitude grid squares. The number of days in a month with highly reflective clouds covering each 1° by 1° square is correlated with monthly rainfall on island stations. In their 1976 paper Kilonsky and Ramage verified their results with estimates from other authors based in the main on surface observations. The verification consisted of comparisons of longitudinal plots of rainfall. Garcia (1981) tested the Kilonsky-Ramage technique in the Atlantic Ocean for the period of GATE (Global Atmospheric Research Program Atlantic Tropical Experiment) and compared the results with estimates of rainfall from the B-scale ship radars covering a 3° square region (Hudlow and Patterson, 1979) and with the estimates of Griffith et al. (1980). Differences with radar of -6% to +15% were found by Garcia for the three phases of GATE. Correlations over 1° square grids with the Griffith et al. data are high (in general greater than 0.75) and greater for the ocean-only comparisons than for the land-and-ocean comparisons. There are substantial differences along the coast of West Africa and in the interior. Coastal differences are due to a maximum in the diurnal cycle which occurs at night

there and which is not adequately sampled with one visible image per day. Woodley et al. (1980b) have shown that the rainfall over the African continent in Griffith et al. (1980) is too great and that an environmental correction must be applied to the rain estimates over the African interior. This is true for Garcia's data as well.

The Kilonsky-Ramage technique has the advantage that monthly ocean rainfall can be estimated from hard copy imagery. The disadvantages are that it is a labor-intensive, subjective scheme based on uncalibrated visible imagery. Although an oceanic rainfall atlas (Garcia, 1985) has been produced with this technique, the assessment of highly reflective clouds is very dependent on the analyst, and as a consequence some original analyses had to be repeated (O. Garcia, personal communication). Furthermore, this scheme has potential problems with the diurnal cycle, which has been shown by several researchers to have an amplitude over oceanic regions that is greater than initially thought (Gray and Jacobson, 1977; McGarry and Reed, 1978; Augustine, 1984; Albright et al., 1985). Lastly, the representativeness of island rainfall for the open ocean is not a completely solved question (Holle and McKay, 1975), although, on the basis of the GATE comparison this does not appear to be a major drawback.

Arkin's technique originated in an analysis of GATE radar and satellite data (Arkin, 1979). It is a simple automated technique that has been used operationally since December 1981 to compute monthly rain over the tropical oceans (see for example Kousky, 1986). Rainfall in 2.5° latitude by 2.5° longitude grid squares is a linear function of daily coverage of each grid square by cold cloud, sampled at 3-hour intervals (Meisner and Arkin, 1987). Little direct verification, other than that performed in the derivation of this technique with the GATE radars, is possible over the oceans; Meisner and Arkin are forced to verify with proxy data. Furthermore, although the technique is fast and can keep pace with the real-time data stream, it is devised for climatology rather than for day-to-day weather.

A major deficiency of these three techniques is in the area of verification. The work in the present paper addresses the problem of verification on two time and space scales. Questions on the importance of the diurnal cycle, the feasibility of making point rainfall estimates from satellite data, and the spatial and temporal frequency of satellite imagery required for the estimation of convective rainfall are discussed in a more complete and rigorous manner than has been done for any of the three other methods.

2. DATA AND QUALITY CONTROL

Three major data sets were processed in this study: the hourly thermal infrared imagery from the GOES-E (Geostationary Operational Environmental Satellite-East) VISSR (Visible and Infrared Spin Scan Radiometer), the hourly gage data available from the National Climate Data Center (NCDC) at Asheville, North Carolina, and the twice-daily upper-air data also archived at NCDC. There were 872 gage locations and 7084 satellite array elements within the region of interest (90.2°W - 108.6°W and 30.4°N - 45.8°N). The radiosonde network comprises the thirty-four upper-air stations within and surrounding this area. Area-averaged and point rainfall estimates were made with both the

satellite and gage data sets. The upper-air data were used to construct adjustment factors that take into account environmental differences between the site of this study (the U.S. Great Plains) and the region of the technique's derivation (south Florida). Computer routines prepared each data set for use on an hourly, 12-hourly or daily basis.

2.1 The Gage Data

Digitized hourly rainfall for the United States of America during August 1979 were obtained from NCDC (formerly, the National Climatic Center) in Asheville, North Carolina. In the area of interest there were 872 gages (Fig. 2.1) at a nominal spacing of approximately 1 gage per 4000 km^2. Of these an average of 860 gages were operating properly at any given hour. Rainfall amounts are accumulations for the 60-minute period ending on the hour.

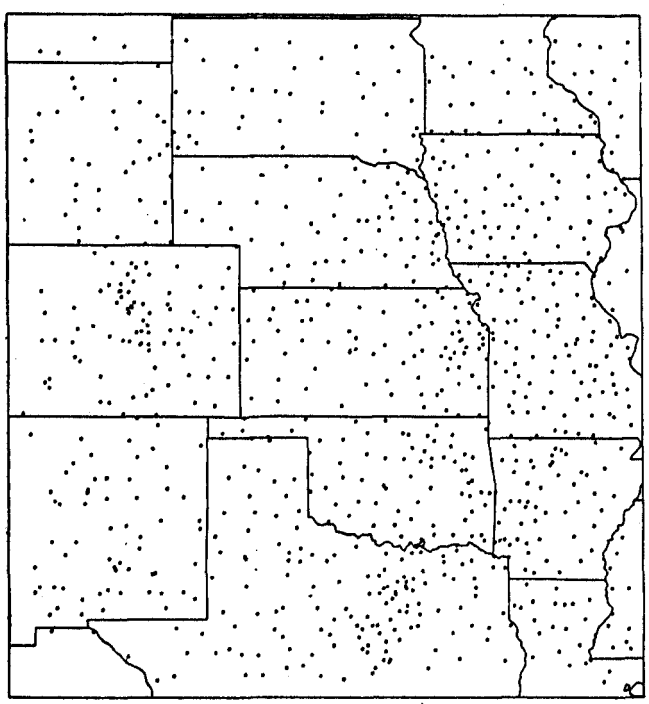


Fig. 2.1. Locations of the hourly gage stations during August 1979.

Two gage data sets were compiled from the archive tape for each day of August 1979--an hourly data set and a daily data set. The gage data are published in local standard time, and the area of study covered two local time zones (Central and Mountain), so it was convenient to convert to Greenwich Meridian Time. They are then also compatible with the time convention of the upper-air and satellite data.

The hourly and daily data sets are not derivatives of each other, so the hourly data set may not contain the same gages as the daily data set for any given day. For example, a gage that has a missing hour during a given day will not be included in the daily data set, but will be included in each of the hourly data sets except, of course, in the hour for which it is missing. Likewise, gages for which the hourly rainfall was an accumulation were not

included in the hourly data sets, but may have been included in the daily data if the accumulation both began and ended within the 24-hour period of the day in question.

The recording gages are of two types--those that record rainfall to 0.10 inch (Fisher-Porter gages) and those that record to 0.01 inch. Gage type was inferred from the rainfall record. Sixty-one gages reported no rain during the entire month of August 1979; it was assumed that these gages resolve rainfall to 0.01 inch. The Fisher-Porter gages were the most numerous--497 Fisher-Porter gages compared with 313 higher resolution gages.

The gage data set was hand-edited for obviously bad gage amounts. Two extreme rainfalls were found on the tape. In both cases the hourly amounts exceeded 5 inches (7.60 inches at Gageby 1 WNW, Texas on Aug. 17, 2100 GMT, and 6.50 inches at Moville, Iowa on Aug. 27, 1300 GMT) and were preceded and followed by hours in which no rain was recorded. Although these amounts hardly seemed physically possible (one would expect some rainfall, however small, to precede or follow such a downpour), as confirmation, these gages were compared with the MDR (Manually Digitized Radar) summary maps that are compiled at 35 minutes past the hour. These two cases were found to be in locations that showed no radar activity in the vicinity of the gage during the hour of extreme rainfall or in the preceding or following hours. These two points were then removed from the hourly data set.

The hand editing revealed an important feature of the gage rainfall data. Thirty-seven hourly rainfalls exceeding 0.99 inch were preceded and followed by hours in which no rain was recorded. With four exceptions (the two cases mentioned above and two cases where the monitoring radars were not available or were out for maintenance), these >0.99-inch rainfalls were under echoes on the MDR summary maps. However, these gages were frequently under echoes in the preceding and following hours, too. There are two possible explanations. If the radar rain did not actually reach the surface, the gages would correctly measure no rain. Alternatively, rainfalls smaller than the minimum required to trip the gage would accumulate and be incorrectly included in the amount for the following hour. This latter situation has adverse implications for the satellite-gage point comparisons discussed in section 3.

2.2 The Satellite Data

Hourly, digital, thermal infrared imagery from GOES-E were used to derive rain estimates over the central third of the United States for the 31 days in August 1979. Thirty-two of the possible 744 images during this period were missing (Table 2.1). The most frequently missed picture (9 out of 32) was the 0600 GMT image, with the remaining missing images being randomly distributed. The longest gaps were three hours long and there were two of these (Aug. 8, 0800-1100 GMT; Aug. 21, 0700-1000 GMT). Occasionally when the image starting on the hour was missing, a substitute was used, usually the image starting on the half hour. Table 2.2 enumerates these images. Such periods could not be used for the hourly comparisons, however, because of the resulting discrepancy between the accumulation period of the gage rainfall (one hour ending on the hour) and that of the satellite (either one-half or one hour, ending on the half hour).

Table 2.1. Missing GOES-E imagery

Date	Hour (GMT)	Date	Hour (GMT)	Date	Hour (GMT)
8/01	1200	8/12	1800	8/23	2000
8/04	0600	8/13	0600	8/24	1000
8/05	1700	8/14	0600	8/25	0700
8/06	1000	8/15	0600	8/26	1800
8/07	0600	8/17	0800	8/27	2200
8/08	0900	8/17	0900	8/28	1500
8/08	1000	8/18	0600	8/29	1500
8/09	0600	8/20	0400	8/30	1600
8/10	0600	8/20	1200	8/31	0500
8/11	0600	8/21	0500		
8/12	0200	8/22	1700		

Table 2.2. GOES-E images not starting on the hour

Date	Hour (GMT)	Date	Hour (GMT)
8/01	2248	8/05	1033
8/02	1433	8/05	1133
8/02	1533	8/21	1133
8/04	1533	8/22	1533
8/05	0733	8/26	0733
8/05	0833	8/26	1433
8/05	0933		

In the navigation program, all digital images were converted from the line/element coordinate system of the satellite to a rectilinear, latitude/longitude coordinate system on Earth. Because of machine memory limitations the satellite data were also degraded in the navigation program from a nominal spatial resolution of 8 km to 22.2 km (1/5° of latitude). These rectilinearized, Earth-located, spatially degraded data were the input data for the subsequent analysis routines.

Sixteen permutations of satellite rain estimates were produced as shown schematically in Fig. 2.2. Each of these divisions (life history versus streamlined; unadjusted versus adjusted; array versus point; hourly versus daily) is discussed.

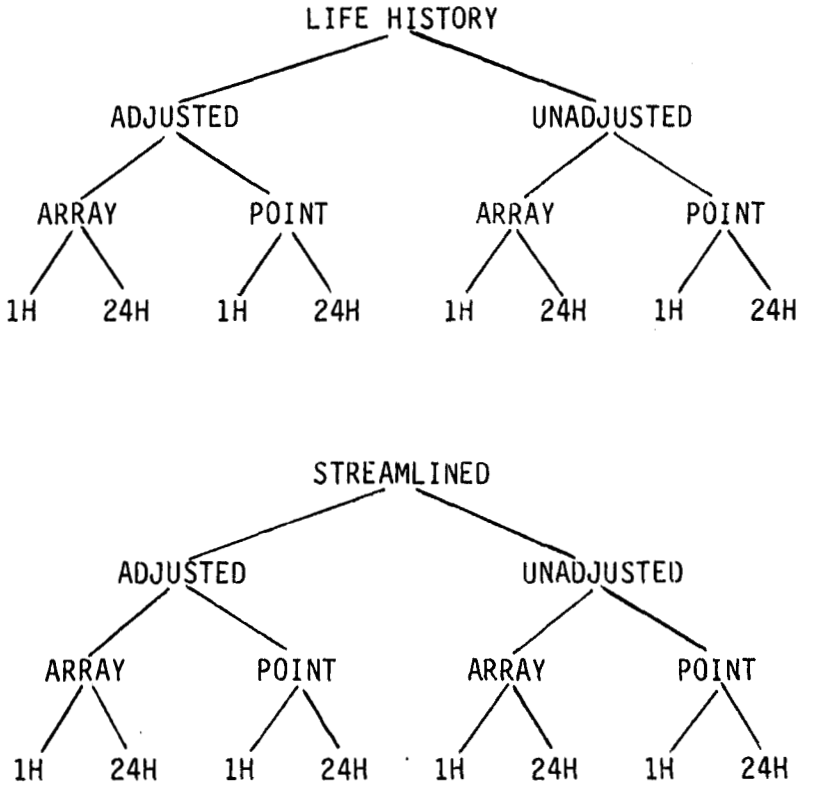


Fig. 2.2. Permutations of the satellite rain estimation technique.

2.2.1 The Life History Technique

All empirical satellite rain estimation techniques employing the visible and thermal infrared channels start from the assumptions that raining clouds can be differentiated in the satellite data from nonraining clouds, and that there are unique relationships between cloud characteristics (observable from the satellite platform) and a rainfall parameter, be it rain area, rain volume, rain rate, or rain occurrence. This study has used the empirical relationships of a technique that was derived in south Florida (Griffith et al., 1978) for summertime, tropical convection. The components of the scheme relate cloud area, as measured from the satellite image, with rainfall that reaches the ground.

In recent years researchers have accumulated worldwide evidence of power-law relationships between radar echo height and a number of radar precipitation characteristics. Pertinent to this study are the strong relationships between maximum echo height and total lifetime rain volume produced by the echo, and between maximum echo height and maximum echo area. Gagin et al. (1985) showed such relationships to exist in Florida and Woodley and Gagin (1986) found (or reported on) similar relation-ships in Texas, Israel and South Africa as well. The Texas rain-volume/echo-height relationship (Fig. 2.3) is typical. The Florida and Israel volume-height

relationships have greater rain volumes than the Texas relationships for the same top heights, and the South Africa clouds produce less rain than the Texas clouds of the same height. These radar relationships indicate a physical basis for the strong relationship seen in this satellite technique between satellite-measured cloud area and surface rainfall.

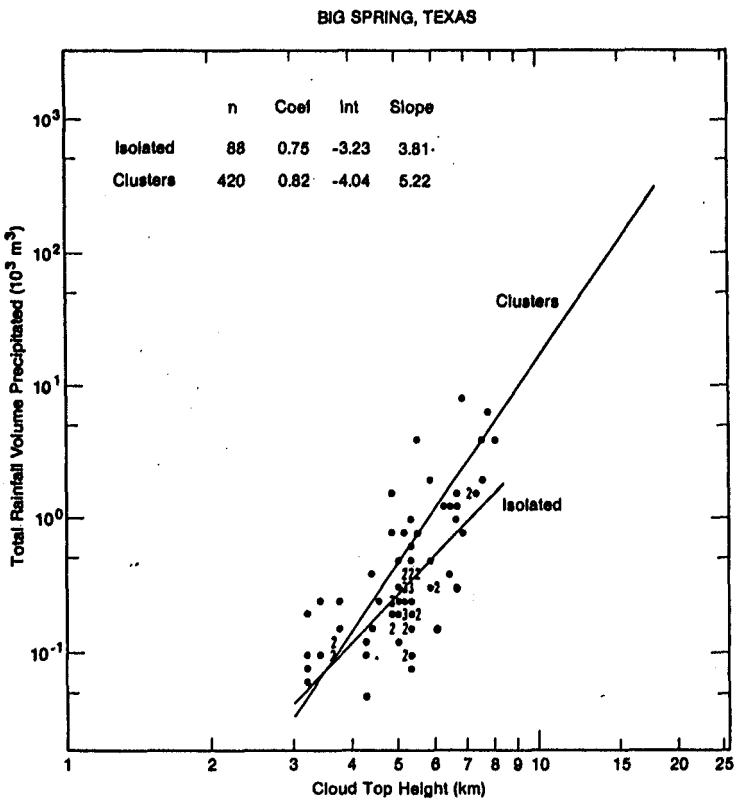


Fig. 2.3. Total rainfall volume precipitated versus maximum radar echo top height for isolated and clustered cells at Big Spring, Texas on six selected days during the Texas HIPLEX program. (From Woodley and Gagin, 1986.)

The life history technique described in Griffith et al. (1978) is based on the premise that, when viewed from the geostationary satellites, the area of a convective cloud undergoes a life cycle of growth and decay, and the rainfall associated with the convection undergoes a similar life cycle. This had been found to be true for radar echoes within the range of the Miami WSR-57 radar, and subsequent analysis showed that the relationship held in the satellite data for long-lived cloud systems, but not necessarily for individual clouds which sometimes merge before they die out.

The mechanics of the scheme are that raining clouds are identified in the IR data as those clouds that are as cold as or colder than -20°C. Cloud area (if any) is measured on each image, and from this area history the cloud's volumetric rain history is computed in a two-step process: (1) Normalized cloud area is used to infer an echo area, which is then related to (2) volumetric rain rate through a family of linear relationships embedded in an expression that includes additional satellite cloud characteristics. Cloud area is normalized by the maximum area achieved by the cloud during its lifetime, which thus permits the derivation and use of one relationship for

convective clouds that cover a range of sizes. A third step is required to convert the cloud's rain volume to isohyets.

The cloud-area/echo-area relationships for IR data are shown in Fig. 2.4. Normalized cloud area, A_C/A_M, is plotted along the horizontal axis and echo area normalized by maximum cloud area, A_E/A_M, is along the vertical axis, for A_C representing satellite cloud area, A_M representing maximum satellite cloud area during the cloud's lifetime, and A_E representing radar echo area. Both axes can also be thought of as time. Along the horizontal axis, for

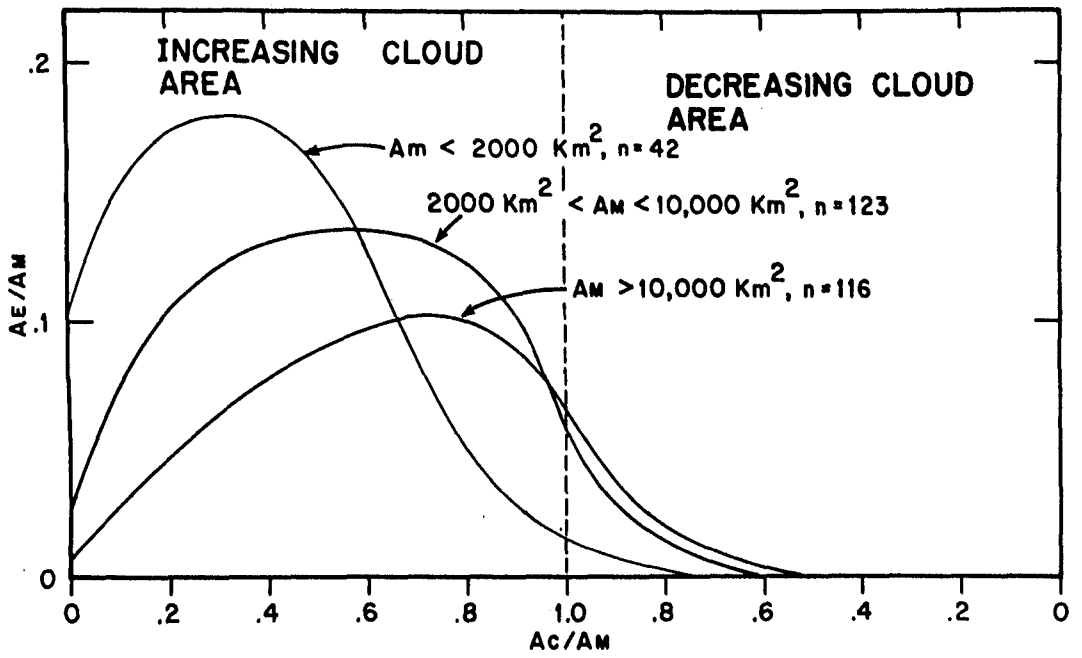


Fig. 2.4. The cloud area-echo area relationships derived in south Florida from thermal infrared, geostationary satellite data. Both cloud area (A_C) and echo area (A_E) have been normalized by maximum cloud area (A_M). The curves are subjective fits to mean data and have been stratified by maximum cloud area.

example, the cloud's life cycle begins at the left of the figure and moves toward the right as the cloud grows, decreases in area and eventually disappears. Echo area can also be seen to increase and then decrease, but echo area peaks and disappears long before cloud area does. This offset between echo life cycle and cloud life cycle is important because it compensates for contamination from cold but inactive cirrus.

There are three relationships, stratified by maximum cloud area, in Fig. 2.4. The smallest clouds are individual cumuli which have relatively short lifetimes, whereas the largest "clouds" are cumulus complexes, lasting hours. The inferred echo area may be that of one echo for the smallest clouds, but it is the total area of the several echoes embedded in the larger complexes.

Once the cloud area history has been measured and the corresponding inferred echo-area ratios (denoted as $\langle A_E/A_M \rangle$) have been obtained from Fig. 2.4, volumetric rainfall per cloud per image is computed from

$$R_V = I \times \langle A_E/A_M \rangle \times A_M \times \Delta t \times \Sigma a_i b_i \times 10\ m^3\ mm^{-1}\ km^{-2} \qquad (2\text{-}1)$$

where

R_V is rain volume (m^3),
I is rain rate (mm/h),
A_E is the inferred echo area (km^2),
A_M is the maximum cloud area during the cloud's life cycle (km^2),
Δt is the time between subsequent images (h),
a_i is the fractional coverage of the cloud by colder temperatures,
b_i is an empirical weighting coefficient for the colder temperatures, and
the factor of 10 converts the units from mm km^2 to m^3.

The angular brackets around the echo ratio term represent the average value for this term as shown by the curves in Fig. 2.4. The rain rate I is a function of echo life cycle and assumes the values listed in Table 2.3. The values of the weighting coefficients were empirically derived and are a function of cloud top temperature (T, expressed as digital count D) according to

$$b = \exp(0.02667 + 0.01547 \times D)/11.1249 \qquad 154 \leq D < 176 \qquad (2\text{-}2)$$
$$-20^0C \leq T < -31^0C$$
$$b = \exp(0.11537 + 0.01494 \times D)/11.1249 \qquad 176 \leq D < 255$$
$$-31^0C \leq T \leq -110^0C$$

Equation (2-1) is based on experience with radar rain measurement, and has been suitably modified to include satellite parameters. It was known for Florida convection that the volumetric rain rate (m^3/h) is linearly related to echo area (km^2) through the rain rates (mm/h) of Table 2.3. The left side of equation (2-1) and the first three terms on the right ($I \times \langle A_E/A_M \rangle \times A_M$) reflect this. (The product of the second and third terms on the right of (2-1) is echo area.) The fourth term (Δt) integrates the rainfall for an instantaneous image over the period between the current and the following image. The fifth and last term on the right ($\sum a_i b_i$) can best be thought of as a means of increasing the rainfall when colder tops are present. It is analogous to rainfall computed from radar echoes that contain inner cores. Consequently, for two clouds that cover the same area at -20°C, the cloud with the colder top will be assigned more rain than the cloud with the warmer top. The sum can run over all temperatures from -20°C to the coldest possible temperature on the image, but it typically is used for three temperature ranges defined at their colder ends by -20°C, -49°C and -72°C.

Rain volume is an inconvenient and unconventional parameter to display, so rain volume per cloud on each image is converted to rain rate through an apportionment scheme. Two apportionment schemes were tested prior to the formulation of the the scheme that is used here. Initially (Woodley et al., 1980a), the satellite-computed rain volume was apportioned over the entire cloud enclosed by the -20°C temperature contour. This resulted in maximum rain depths that were far too small and finite rain depths over too extensive an area when compared with radar rain depths. A second scheme apportioned the rain volume over an area equivalent to the echo area inferred from Fig. 2.4; this scheme, however, resulted in flood-producing rain depths that were concentrated over very small regions. The scheme used herein (Augustine et al., 1981a) is an empirical compromise between the two previous schemes. It places one-half of the calculated rain volume into the pixels that constitute

Table 2.3. Rain rate as a function of echo growth trend for south Florida echoes

Echo growth trend	Rain rate I ($\times 10^2$ mm/h)
Increasing echo area	
$0.00 \leq A_E/A_{EM} < 0.25$	13.3
$0.25 \leq A_E/A_{EM} < 0.50$	17.3
$0.50 \leq A_E/A_{EM} < 0.75$	21.1
$0.75 \leq A_E/A_{EM} < 1.00$	23.8
Maximum echo area	
$A_E/A_{EM} = 1.00$	20.7
Decreasing echo area	
$1.00 > A_E/A_{EM} \geq 0.75$	21.1
$0.75 > A_E/A_{EM} \geq 0.50$	16.7
$0.50 > A_E/A_{EM} \geq 0.25$	11.9
$0.25 > A_E/A_{EM} \geq 0.00$	8.2

Note: A_E refers to echo area (defined by the 1 mm/h rain rate) and A_{EM} refers to the maximum area an echo attains in its life cycle.

the coldest 10% of the cloud's area. The remaining half is apportioned over the pixels whose temperatures fall into the next warmer 40% of the cloud's area. Thus, only half of the cloud area that was used to compute the rain volume will contain a rain rate and the coldest cloud tops have the most rain.

The rain rate inferred for a particular pixel is a function of the pixel's temperature and grid size, and the cloud's rain volume according to

$$D_{ij} = \frac{(R_V/2) \times b_{ij}}{\frac{(\Sigma b)}{n} \times g_{ij}} \times \frac{\text{mm km}^2}{10^3\ \text{m}^3} \qquad (2\text{-}3)$$

where

D_{ij} is rain depth (mm) in the (i,j) pixel,

$R_V/2$ is one-half of the cloud's rain volume (m^3) for this image,

b_{ij} is the weighting coefficient for the (i,j) pixel,

$\sum b$ is the sum of the weighting coefficients with the index "n" as described below,

g_{ij} is the area (km^2) of the (i,j) pixel, and

10^{-3} converts the units from m^3/km^2 to mm.

Preliminary to the depth calculation, all pixels constituting the cloud are ranked from coldest to warmest. Three cutoffs are noted from this ranking. Each cutoff is chosen so that the area represented by the pixels that fall between two cutoffs equals a specified fraction of the cloud's total area. The cutoffs are at 1) the coldest pixel, 2) the pixel where the area represented by the pixels between it and the coldest pixel is 10% of the cloud's area, and 3) the pixel where the represented area between it and the coldest pixel is 50% of the cloud's area.

The indices of the sum over the weighting coefficients ($\sum b$) are determined by these cutoffs and by the temperature of the (i,j) pixel. If the pixel's temperature is ranked among the coldest 10% of the pixels constituting the cloud's area, then n runs from the coldest pixel to that pixel at cutoff 2, and the sum is over the corresponding b values for the coldest 10% of the cloud. If the pixel's temperature in not ranked in the coldest 10%, but is among the coldest half of the cloud, then the sum is over the b values for the coldest 50% of the cloud, excluding the coldest 10%, that is, n runs from the pixel at cutoff 2 to the pixel at cutoff 3. If the pixel's temperature is ranked among the warmest half of the cloud, D_{ij} is set to zero.

An example best illustrates the life history rain estimation scheme. Consider the 10 satellite samples (each sample is a "pixel" 144 km^2 in area) that make up the cloud in Fig. 2.5. Assume that this cloud is growing and that on this image it is four-tenths of its maximum size of 3600 km^2. From the middle curve of Fig. 2.4, the echo ratio is 0.12 and the echo growth trend is found to be at $A_E/A_{EM} = 1.00$. This implies a rain rate of 20.7×10^2 mm/h from Table 2.3. By using the -20°C (digital count 154), -49°C (digital count 194), and -72°C (digital count 217) thresholds for the summation term, there

Fig. 2.5. Digital count field containing a hypothetical cloud (the hatched squares).

j \ i	1	2	3	4	5	6
6	145	150	153	153	150	145
5	145	153	174	174	153	150
4	153	174	200	198	175	153
3	153	174	196	180	178	153
2	153	153	153	153	153	150
1	145	145	145	145	145	145

are seven pixels in the first temperature interval, three in the second, and none in the third. The rain volume calculation for this cloud is

$$R_V = 20.7 \times 10^2 \text{ mm/h} \times 0.12 \times 3600 \text{ km}^2 \times 1 \text{ h} \times$$
$$\times (0.7 \times 1.00 + 0.3 \times 1.83) \times 10 \text{ m}^3 \text{ mm}^{-1} \text{ km}^{-2}$$
$$= 11.17 \times 10^6 \text{ m}^3$$

The temperatures of these pixels have been ranked from coldest to warmest; their associated digital counts, (i,j) locations, and weighting coefficients are listed in Table 2.4 Under the apportionment, half of the rain volume will be placed in the coldest 10% of the cloud. In this example the coldest 10% amounts to one pixel, (3,4), for a depth of 38.8 mm (column five of Table 2.4). The remaining volume is apportioned into the next warmest 40% of the cloud (i.e., four pixels) as indicated in the same table.

Table 2.4. Sample pixel values and computed rain depths for the life history (L.H.) and streamlined (S.L.) techniques

T (°C)	DC	(i,j)	b	L.H. D_{ij} (mm)	S.L. D_{ij} (mm)
-55°	200	(3,4)	2.00	38.8	1.7
-53°	198	(4,4)	1.94	11.2	0.5
-51°	196	(3,3)	1.89	10.9	0.5
-35°	180	(4,3)	1.48	8.5	0.4
-33°	178	(5,3)	1.44	8.3	0.4
-31°	175	(5,4)	1.38	0.0	0.0
-30°	174	(2,3)	1.36	0.0	0.0
-30°	174	(2,4)	1.36	0.0	0.0
-30°	174	(3,5)	1.36	0.0	0.0
-30°	174	(4,5)	1.36	0.0	0.0

In summary, the life history technique produces an array of unadjusted rainfall estimates for the area of interest at the time and space scales of the navigated satellite digital imagery. Rain that is inferred with the life history scheme has the following properties:

(1) Raining clouds are those that are as cold as or colder than -20°C.
(2) Rainfall is directly proportional to cloud area on any given image.
(3) Rainfall is inversely proportional to cloud top temperature.
(4) Rainfall is a function of cloud life cycle such that more rain is inferred in the early stages of a cloud's history than in the later stages.

2.2.2 The Streamlined Technique

A major disadvantage of the life history technique is that it is not configured to make estimates of rainfall in real time. A cloud system must be followed through a sequence of images to the time where a maximum in areal extent can be identified before the rain calculation can begin. To circumvent this, a version of the technique that makes no use of the time change information was developed (Waters et al., 1977; Woodley et al., 1978). Referred to as the streamlined technique, this version is streamlined in the sense that the complexity of the technique (and consequently the computer code) is greatly reduced--rain estimates are made from single images rather than from the sequence of images that is required in the life history scheme. The streamlined technique recognizes (as do other satellite studies such as Stout et al., 1979; Lovejoy and Austin, 1979; and Negri et al., 1984, and as do analogous radar studies: Doneaud et al., 1981, Lopez et al., 1983 and Gagin et al., 1985) that the dependence of the inferred rainfall on the system's stage in its life cycle is small compared with the cloud area term. The advantages of the streamlined technique are that (1) rain estimates can be made as each satellite image is received, and (2) computation time, excluding image navigation, is one-tenth of that needed for the life history method.

The component relationships of the life history technique are still used but with the modification that cloud area on each image is assumed to be the maximum cloud area. This forces the value of the rain rate to be 16.7×10^2 mm/h, and the inferred echo ratio assumes the value where $A_C/A_M = 1.00$, that is, where the curve crosses the vertical dashed line shown in Fig. 2.4. This ratio is 0.067 for clouds larger that 10,000 km^2, 0.047 for clouds between 2000 and 10,000 km^2 in area, and 0.016 for clouds smaller than 2000 km^2. Equation (2-1) becomes

$$R_V = 16.7 \times 10^2 \text{ mm/h} \times \left\{ \begin{matrix} 0.016 \\ 0.047 \\ 0.067 \end{matrix} \right\} \times A_M \times \Delta t \times$$

$$\times \Sigma a_i b_i \times 10 \text{ m}^3 \text{ mm}^{-1} \text{ km}^{-2} \qquad (2\text{-}4)$$

where the symbols are as before and the inferred echo ratio term (shown in braces) is assigned on the basis of current cloud area.

A sample rain calculation for the cloud of Fig. 2.5 would yield a rain volume of

$$R_V = 16.7 \times 10^2 \text{ mm/h} \times 0.016 \times 1440 \text{ km}^2 \times 1 \text{ h} \times$$

$$\times (0.7 \times 1.00 + 0.3 \times 1.83) \times 10 \text{ m}^3 \text{ mm}^{-1} \text{ km}^{-2}$$

$$= 0.48 \times 10^6 \text{ m}^3.$$

Rain depths per pixel are calculated from (2-3) as before and are listed in the sixth column of Table 2.4. As can be seen, the streamlined depths are an order of magnitude smaller than the life history depths. Differences between area-averaged rainfalls for the streamlined and life history schemes are in general much smaller than this example would indicate. Computations for clouds in the smallest size interval are most affected, as this example shows.

In summary, the streamlined technique produces an array of unadjusted rainfall estimates for the area of interest at the time and space scales of the navigated satellite digital imagery. Rain that is inferred with the streamlined scheme has the following properties:

(1) Raining clouds are those which are as cold as or colder than -20°C.
(2) Rainfall is directly proportional to cloud area on any given image.
(3) Rainfall is inversely proportional to cloud top temperature.

2.2.3 The Environmental Adjustment

The empirical relationships used in this study were derived for south Florida. To use these relationships in the middle latitudes environmental differences between the subtropics and the middle latitudes must be accounted for. Large total liquid water content and low vertical wind shear are characteristic of air masses found over Florida. In the Great Plains regimes, however, shear and subcloud evaporation significantly affect cloud efficiency. The environmental adjustment attempts to account for these differences. Several adjustment factors were originally tested on 15 cases at three sites in Montana, Kansas, and Texas (Griffith et al., 1981). Five parameters were tested: precipitable water, subcloud evaporation, environmental shear-precipitation efficiency, a combination of these three, and a precipitation factor derived from a one-dimensional cumulus cloud model, similar to the environmental correction devised by Wylie (1979). Each of these five parameters can be derived from or with the upper-air data. The cumulus model adjustment factor was found to provide the best correction of the satellite estimates toward the ground measurements of rainfall. The derivation and comparison of the five adjustment factors are discussed in detail in Appendix A.

The model adjustment factor (MAF) is defined as

$$MAF = \frac{\overline{R}_H \times \sigma_H}{\overline{R}_F \times \sigma_F} \qquad (2\text{-}5)$$

where $\overline{R}$ is the mean model rain production for eight thermal bubble radii, σ is the standard deviation of the eight model rainfalls, and H and F refer to the High Plains and Florida, respectively. The values of $\overline{R}_F$ and σ_F computed from a typical Florida sounding are 10.210 and 4.195, respectively. Although the individual model adjustment factors calculated for this study ranged from 0.00 to 1.45, most of them were less than 1.0, as would be expected for a region that is drier than south Florida. At those times when the sounding was sampling air that had been modified by a mesoscale convective system, the model adjustment factor was set to 1.00. These cases are discussed in detail in Appendix B.

After the model adjustment factors had been computed for the stations discussed in section 2.3, model adjustment factor fields at the same spatial resolution as the satellite rainfalls were produced by Gaussian interpolation (see Appendix C). These fields multiplied the satellite arrays of life history and streamlined data to produce the adjusted life history or adjusted streamlined data sets.

2.2.4 Satellite "Point" Estimates

Both the life history and streamlined techniques produce an array of satellite-inferred rain depths (at 1/5° of latitude by 1/5° of longitude spatial resolution), which have been accumulated over one hour. In order to make point comparisons between satellite and gage rainfalls, one of these data sets must be interpolated to the resolution of the other. Because less computation is required to interpolate to a sparser data set, the approximately 8000 satellite points were interpolated to the approximately 900 gage locations. The interpolation that was used was a bilinear interpolation. All eight permutations of the satellite array rainfalls (life history and streamlined; unadjusted and adjusted; hourly and daily) were interpolated.

2.2.5 Temporal Resolution

Satellite rain depths accumulated for hourly periods were computed in the array format from the hourly satellite imagery. The appropriate satellite arrays were subsequently summed to produce daily accumulations in the satellite array format. For each period (1 hour or 24 hours) the accumulation starts at the time of the first image and all times are in GMT.

Hourly rain estimates for 31 days would result in 744 hours in the hourly data set. However the missing satellite imagery listed in Table 2.1 combined with periods when satellite images from 30 or 45 minutes past the hour were used to fill in for images missing on the hour (Table 2.2) resulted in only 662 hourly periods for the satellite data.

2.3. The Upper-Air Data

The 34 upper-air stations shown in Fig. 2.6 were the basis for the model

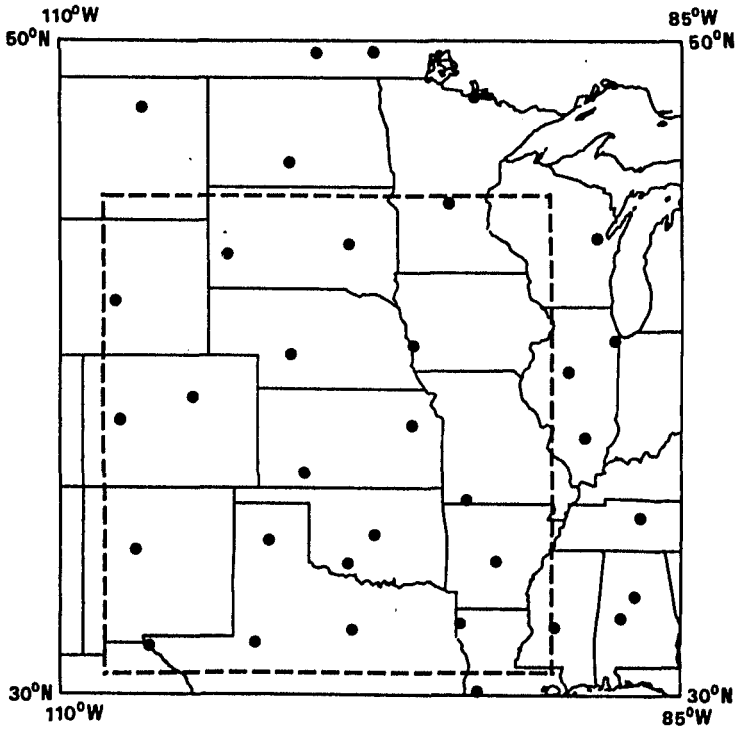


Fig. 2.6. Locations of the upper-air stations during August 1979.

adjustment calculations. The twice daily soundings at the synoptic times of 0000 and 1200 GMT were used. Of the 2108 possible soundings from these stations during August 1979, 327 were missing. The stations in Canada were most often missing. As noted in Appendix C, the Gaussian interpolation requires that these stations cover an area that is somewhat larger than the region of the rain estimates.

Several of these stations were under the canopy of an MCC at the synoptic time of the sounding. Because the cumulus model requires the input of a sounding that has sampled the air the produces the convection, soundings that had sampled air modified by the MCC would not result in accurately modeled convection for the particular time. Therefore soundings from 55 stations on 20 days that were determined to be under the influence of MCC-modified air were not used in the one-dimensional model and the corresponding model adjustment factors were set to 1.00 as detailed in Appendix B.

3. COMPARISONS OF SATELLITE AND GAGE AREA-AVERAGED RAINFALLS

Rainfall comparisons can be based on any of a number of rain measures; rainfall rate, cumulative rainfall, areal rainfall, point rainfall, rainfall time series and rain pattern are several common measures. This study focuses on comparisons of rainfall amount and on the timing of rainfall events. Point and areally-averaged rainfall amounts for hourly, daily, and monthly accumulations are presented in this and subsequent sections. "Events" will be defined by time series of hourly and daily area-averaged amounts; no time series of point rainfalls will be presented. Not only are comparisons between gage and satellite estimates shown; comparisons among the satellite permutations are also made. The relative performance of the life history versus the streamlined technique is of interest both scientifically and computationally.

In addition to the more familiar scatterplots and time series, graphical methods of data presentation developed in EDA (exploratory data analysis) (Tukey, 1977) are shown. These include box-and-whisker and stem-and-leaf plots. Many classic statistics assume the normal distribution. EDA techniques (see Appendix D) however make no such assumption, emphasizing medians, hinges (quartiles), and extremes, which more appropriately describe rainfall than do the parameters of the normal distribution.

Five difference measures (Table 3.1) are used extensively to describe the satellite-gage comparisons. The measure of whether the satellite estimates are larger or smaller than the gage estimates for the month is the ratio R_M. Similarly, R_P is an average period (either daily or hourly) ratio of satellite to gage rainfalls; an average for a daily period is indicated by R_D and for an hourly period by R_H. Contributions to R_P from ratios of S/G that are less than 1.0 can offset those values greater than 1.0. The measure E_R has eliminated that effect, for E_R is defined to be greater than 1.0. Thus, E_R is to R_P as $|x|$ is to x. The root-mean-square error (E_{RMS}) measures the absolute difference in rain depth for the satellite-gage pairs, whereas the normalized root mean-square-error (norm. E_{RMS}) measures the absolute difference as a fraction of the observed gage rainfall. The latter therefore distinguishes, for example, between a 3-mm difference in a total area-averaged gage rainfall of 10 mm, and a 3-mm difference in a total area-averaged gage rainfall of 2

Table 3.1. Measures of difference between the satellite (S) and gage (G) rainfall estimates for a sample of size N

Difference Measure		Definition
Monthly Ratio	(R_M)	$\Sigma S/\Sigma G$
Mean Period Ratio	(R_P)	$\frac{\Sigma(S/G)}{N}$
Factor of Difference	(E_R)	$\frac{\Sigma R}{N}$ for $R = S/G$ if $S<G$; $R = G/S$ if $G<S$
RMS Error	(E_{RMS})	$\left[\frac{\Sigma(S-G)^2}{N}\right]^{1/2}$
Normalized RMS Error	(Norm. E_{RMS})	$\left\{\frac{\Sigma[(S-G)/G]^2}{N}\right\}^{1/2}$

mm.

The comparisons to follow involve each of the 16 permutations of Fig. 2.2. Array estimates are used in section 3.1 to compute area-averaged rainfalls, and point estimates are discussed in section 3.2. The remaining three levels of permutations (life history vs. streamlined; adjusted vs. unadjusted; hourly vs. daily) provide a stratification for discussion.

Area-averaged rainfalls were computed from both the satellite array data and the gage point rainfalls. Area-averaged rainfalls are typically used to assess rainfall over a basin for hydrological purposes (Peck, 1980) or over a fixed experimental area (Woodley et al., 1975). A number of techniques have been devised to estimate areal precipitation from point measurements. For example, Thiessen (1911) discussed a manual technique that weights rainfall from randomly located gages by the area that the corresponding gage is defined to represent in the basin. More recently, Hatch (1976) described an interpolation scheme for gage data, and Bras and Rodriguez-Iturbe (1976) used multivariate estimation theory to produce area-averaged rainfalls from gages. In this study the simplest area-averaging technique is used, namely that of computing means. Rain depths for each element in the satellite array or for each gage in the area were summed for the period of interest (1 hour or 24 hours), and then divided by the number of rainfall locations that went into the sample. Areal averages of rainfall assume that the measuring system representatively samples both the area and the rainfall. The satellite array estimates are uniformly distributed over the region, but the gages are less so. Their minimum spacing ranges from 50 km^2 to 150,000 km^2. Likewise convective rainfall is distinguished by its sharp gradients and is not uniform either over the region or from one storm to the next.

3.1 Daily Rainfalls

The most striking feature of the daily unadjusted satellite rainfalls from both the life history (Fig. 3.1) and the streamlined (Fig. 3.2) techniques is that their values are much larger than the corresponding gage depths. Over half are more than double the gage amounts and approximately one-quarter of them are more than three times larger. Fewer than 10% of the satellite estimates are smaller than the gage results. Not surprisingly the environmental correction, applied to either the life history (Fig. 3.3) or the streamlined (Fig. 3.4) technique, greatly decreases the satellite-estimated amount, and also decreases the scatter of the estimates. Between the unadjusted and adjusted plots the correlation coefficient (Table 3.2) increases by about 10 points from roughly 0.6 to 0.7, and the intercept of the fit moves toward the origin, decreasing from about 1.7 to about 0.3 mm. The adjustment, however, leaves most of the satellite estimates with values somewhat smaller than the gage amounts for the slope changes from a value greater than 1.0 to a value on the order of 0.6.

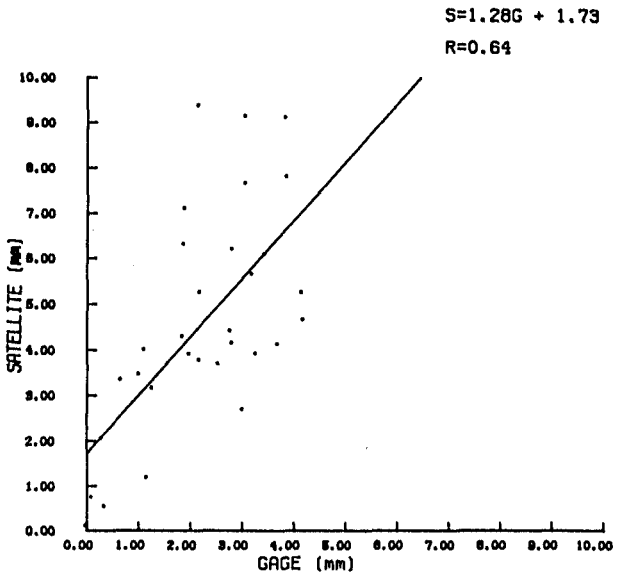


Fig. 3.1. Scatterplot of daily area-averaged satellite and gage rainfalls for the unadjusted life history technique. The regression line has been plotted; the slope, intercept, and correlation coefficient are noted.

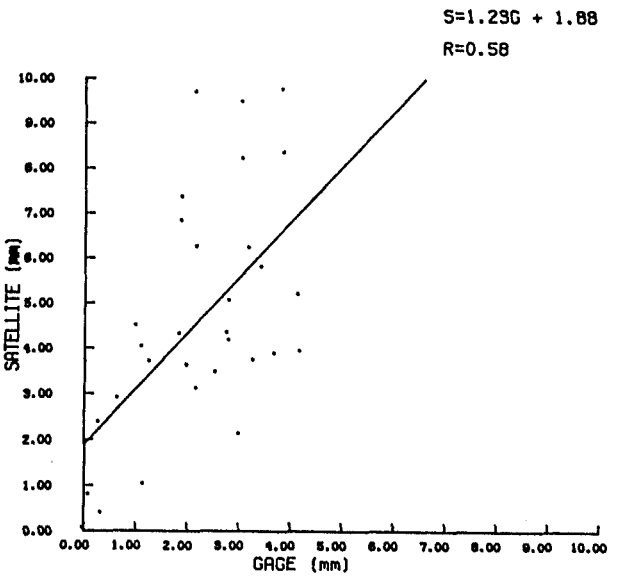


Fig. 3.2. Scatterplot of daily area-averaged satellite and gage rainfalls for the unadjusted streamlined technique as in Fig. 3.1.

Medians, means, and standard deviations (Table 3.3) of the unadjusted satellite samples are about double their respective gage values. After adjustment the satellite medians and means are approximately 20% smaller than the gage value, but standard deviations differ by less than 10%. Box-and-whisker plots (Fig. 3.5) highlight these differences in the sample distributions. The unadjusted satellite techniques have large interquartile

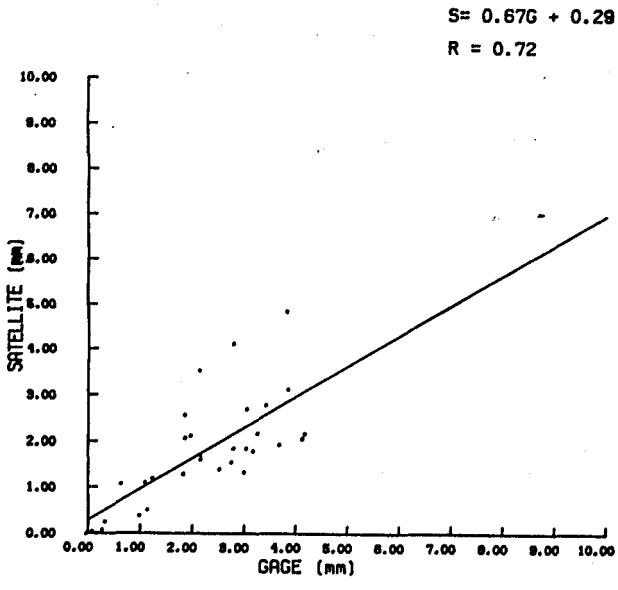


Fig. 3.3. Scatterplot of daily area-averaged satellite and gage rainfalls for the adjusted life history technique as in Fig. 3.1.

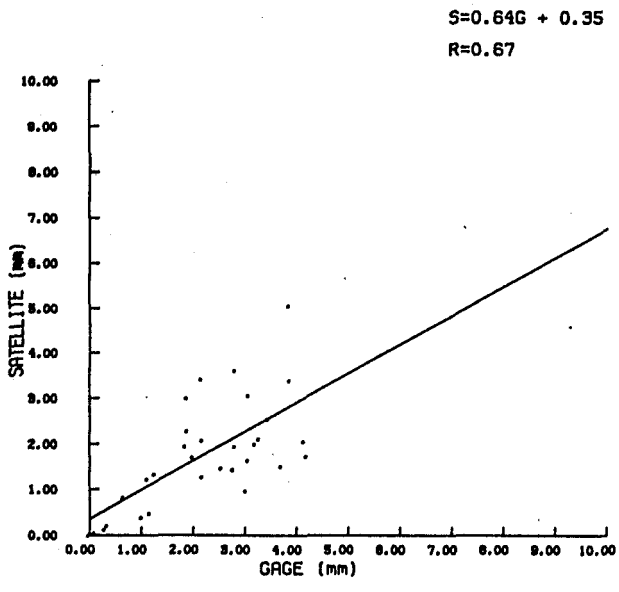


Fig. 3.4. Scatterplot of daily area-averaged satellite and gage rainfalls for the adjusted streamlined technique as in Fig. 3.1.

ranges and large total ranges. In contrast, the adjusted satellite data more closely resemble the gage data, particularly at the smaller values where the smallest values and the lower hinges are almost identical. However, the upper hinges occur at the value of the gage median and the maximum values exceed the largest gage value.

Table 3.2. Least-squares-fit parameters for satellite-raingage regressions of daily, area-averaged rainfall

Satellite Permutation	Least-Squares Linear Fit		
	ρ	Slope	Intercept (mm)
Unadjusted life history	0.64	1.28	1.73
Adjusted life history	0.72	0.67	0.29
Unadjusted streamlined	0.58	1.23	1.88
Adjusted streamlined	0.67	0.64	0.35
Perfect correspondence	1.00	1.00	0.00

Table 3.3. Statistics for the daily, area-averaged satellite and raingage samples

Data Set	No. of Samples	Median (mm)	Mean (mm)	Std. Dev. (mm)
Unadjusted life history	31	4.25	4.71	2.46
Adjusted life history	31	1.87	1.85	1.15
Unadjusted streamlined	31	4.27	4.76	2.65
Adjusted streamlined	31	1.77	1.83	1.18
Gage	31	2.27	2.34	1.24

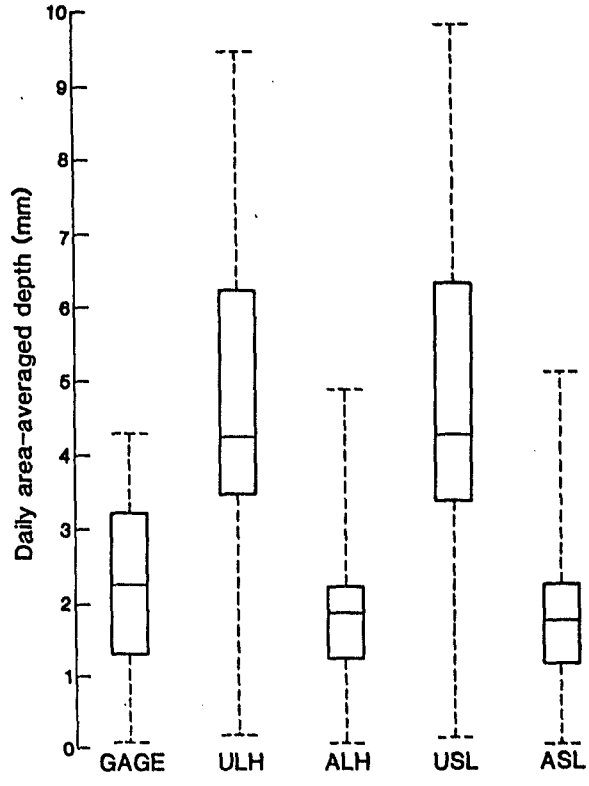


Fig. 3.5. Box-and-whisker plots of daily area-averaged rainfalls (mm), from gage data (GAGE), unadjusted life history (ULH), adjusted life history (ALH), unadjusted streamlined (USL), and adjusted streamlined satellite (ASL) data. All samples have 31 points.

The box-and-whisker plots confirm the numbers in Tables 3.2 and 3.3, which indicate little difference in gross performance between the streamlined and life history techniques. The unadjusted life history and unadjusted streamlined box-and-whisker plots are almost identical. Likewise, the adjusted life history and the adjusted streamlined box-and-whisker plots show similar ranges and medians.

In the box-and-whisker plots there appears to be a modicum of similarity between the two unadjusted permutations, between the two adjusted permutations, and among the gage and adjusted satellite data sets. However, under the closer scrutiny of stem-and-leaf plots (Fig. 3.6) the differences become more apparent. In these plots, the stems are integer amounts of daily area-averaged rainfall (mm) and the leaves are tenths of millimeters. The similar

Life History

GAGE	
0	0 1 3 4 7
1	0 2 2 3 9 9 9
2	0 2 2 2 6 8 8 9
3	1 1 1 2 3 5 7 9 9
4	2 2

ULH	
0	1 6 8
1	2
2	1 7
3	2 4 5 7 8
4	0 0 1 2 2 3 5 7
5	3 3 7
6	1 2 4
7	1 7 9
8	
9	2 2 4

ALH	
0	0 1 1 3 4 5
1	1 1 2 3 4 4 6 6 7 8 9 9
2	0 1 1 1 2 2 6 7 8
3	1 6
4	1 9

Streamlined

GAGE	
0	0 1 3 4 7
1	0 2 2 3 9 9 9
2	0 2 2 2 6 8 8 9
3	1 1 1 2 3 5 7 9 9
4	2 2

USL	
0	1 4 8
1	1
2	2 4 9
3	2 5 7 7 8 9
4	0 1 2 4 4 5
5	1 2 8
6	3 3 9
7	4
8	3 4
9	5 7 8

ASL	
0	0 1 1 2 4 5 8
1	0 2 3 3 5 5 5 7 7 7
2	0 0 0 1 1 1 3 6
3	0 1 4 4 6
4	
5	1

Fig. 3.6. Stem-and-leaf plots of daily area-averaged rainfalls (mm) from life history and streamlined satellite data (see caption, Fig. 3.5).

shapes of the two unadjusted sets and of the two adjusted sets are confirmed. Likewise, it can be seen that the adjusted satellite data sets more closely resemble the gage data than the unadjusted sets do. However, the satellite rainfall category with the most occurrences is 2 mm smaller than the gage category with the highest number of occurrences. This shift in the locations of the peak graphically depicts the reason for the smaller adjusted satellite median and mean. Since the adjusted satellite samples are more normally distributed than the gage sample, the standard deviations are also smaller. However, it is obvious that none of these samples constitutes a normal distribution.

Time series of satellite rainfalls with gage amounts superimposed are shown in Figs. 3.7 and 3.8. These time series serve two purposes. From them the differences between each day's rainfall amount as estimated by the satellite and by the gage can be immediately seen. Secondly, coincidence of rainfall maxima and minima during the month are obvious. The amount differences mirror the scatterplots and the EDA plots. That is, the unadjusted satellite permutations (Fig. 3.7) show rainfalls that are much greater than the gage rainfalls, while the adjusted satellite rainfalls (Fig. 3.8) correspond more closely to the gage data, albeit as underestimates of the daily area-averaged gage rainfalls. In all four sets the no-rain period of August 5-7 is captured by the satellite techniques. On these three days there were no clouds to speak of, so the correspondence is not surprising. Interestingly, in the adjusted satellite time series the closest correspondence in amount generally occurs when the gage time series exhibits a minimum.

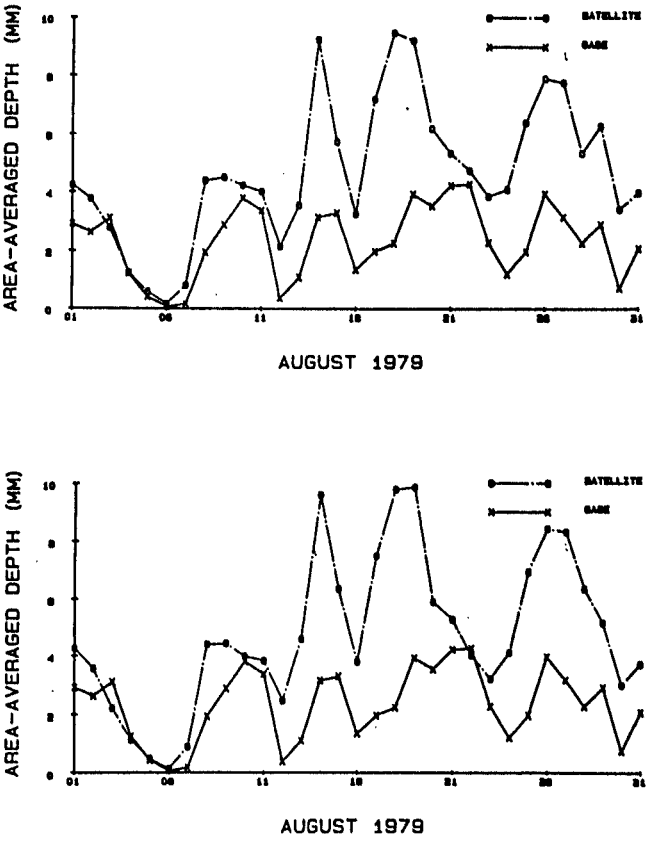


Fig. 3.7. Daily area-averaged unadjusted life history (top) and streamlined (bottom) rainfalls for August 1979, compared with daily area-averaged gage rainfalls.

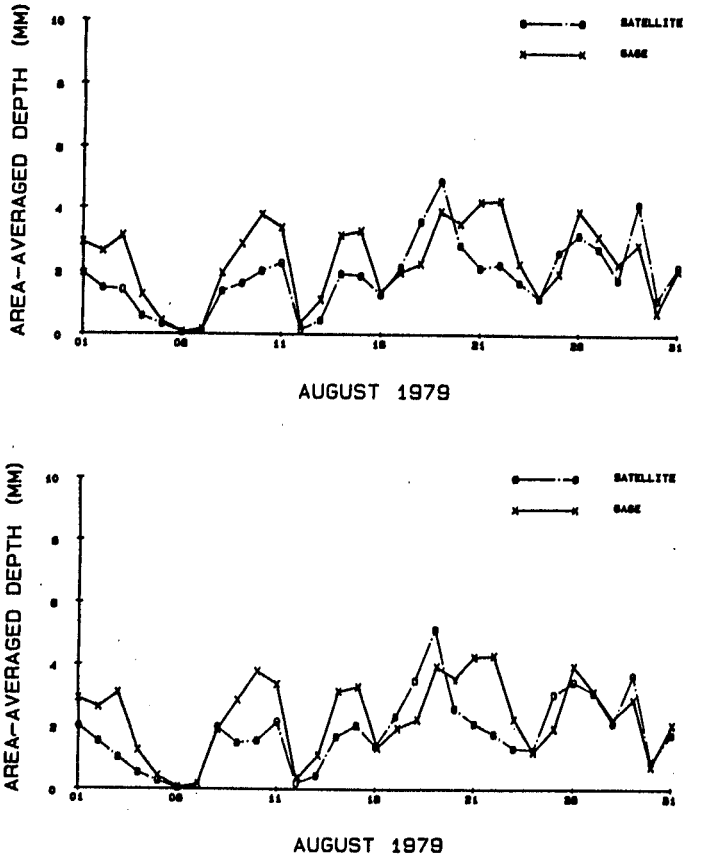


Fig. 3.8. Daily area-averaged adjusted life history (top) and streamlined (bottom) rainfalls for August 1979, compared with daily area-averaged gage rainfalls.

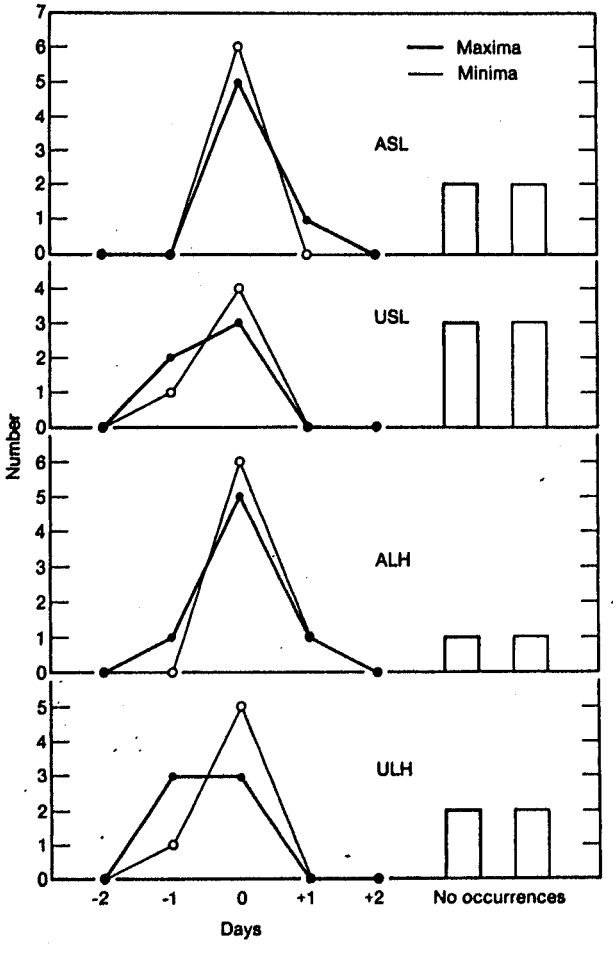


Fig. 3.9. The difference in timing of daily area-averaged rainfall maxima and minima for four satellite permutations. Gage data define the relative maxima. A negative day means that the satellite relative maximum precedes the gage. The number of times the satellite does not indicate a maxima or minima ("nonoccurrences") is also shown.

There is very good correspondence between the satellite and gage time series on the occurrence of rainfall maxima and minima events. If the gages are used as the standard, there are eight rainfall minima and nine maxima during the month. Fig. 3.9 summarizes the number of maxima and minima in each of the satellite permutations that correspond to these gage maxima and minima and their differences in timing. A negative day means that the satellite event occurred one (or two) days before the gage event. There were some gage events when the satellite indicated none ("nonoccurrences"). On the whole, the satellite events are coincident with more than 50% of the gage events. The adjusted data have more coincident events than the unadjusted data do. The noncoincident events in the unadjusted satellite techniques all occur before the gage event.

Table 3.4. Harmonics and their fractional variance for the two harmonics with the greatest power

Data set	Harmonic	Period (days)	Fractional Variance
Gage	5	6.2	0.33
	3	10.3	0.16
Unadjusted life history	3	10.3	0.10
	5	6.2	0.08
Adjusted life history	3	10.3	0.21
-	8	3.9	0.08
Unadjusted streamlined	5	6.2	0.11
	7	4.4	0.05
Adjusted streamlined	3	10.3	0.22
	15	2.1	0.05

Harmonic analyses were performed on these five time series after the mean of each series had been subtracted out. The periodogram for the daily gage data show the most power in the fifth harmonic (a period of 6.2 days). This frequency, however, accounts for only 33% of the variance. The third harmonic (representing a period of 10.3 days) accounts for about 15% of the variance. A total of 50% of the variance is explained by these two harmonics. (See Table 3.4.) For the satellite permutations, the two harmonics containing the most power explain 20-30% of the total variance. Again the third or fifth harmonics, singly or together, constitute the two harmonics showing the most power, but other harmonics appear as well, although with very little power.

Difference measures for the daily, area-averaged rain estimates are given in Table 3.5. For the month the unadjusted satellite estimates are larger

than the gage estimates by a little more than 100%, but the adjusted cases are smaller than the gage rainfalls by about 20%. (See column two). Similar results hold for the average daily ratio of satellite to gage estimates (column three). The factor of difference (column four) ranges from 1.6 for the adjusted to 2.5 for the unadjusted satellite data. The root-mean-square error (column five) is on the order of 3 mm for the unadjusted data, but drops to about 1 mm for the adjusted data. The root-mean-square error of the adjusted satellite data sets is comparable with the standard deviation of the data set. The normalized RMS error (column six) is on the order of 190% in the unadjusted satellite sets, but decreased drastically to a relative error of about 40% for the adjusted cases.

Table 3.5. Difference measures for 31 daily area-averaged rain samples

Satellite Permutation	Mean Monthly Ratio	Factor Daily Ratio	RMS of Diff.	Norm. Error (mm)	RMS Error
Unadj. L.H.	2.02	2.39	2.40	3.03	1.83
Adj. L.H.	0.79	0.81	1.55	1.01	0.38
Unadj. S.L.	2.04	2.43	2.46	3.24	1.97
Adj. S.L.	0.78	0.80	1.60	1.09	0.39
Perf. Corresp.	1.00	1.00	1.00	0.00	0.00

For certain applications in hydrometeorology it is important to know whether rainfall observations are within a given fraction of "true" rainfall, or within an absolute measure of "true" rainfall. Both assessments were made for the four permutations of daily area-averaged satellite rainfall. Daily differences of area-averaged rainfalls ($\bar{S}$-$\bar{G}$) are highly skewed around zero (Fig. 3.10). Most differences in the two unadjusted satellite data sets are positive, whereas the opposite is true in the adjusted satellite data sets. The adjusted satellite data sets indicate that half the differences lie between 0 and -1 mm and almost all the differences are between ±2 mm. Relative differences (Fig. 3.11) of daily area-averaged rainfalls $[(\bar{S}-\bar{G})/\bar{G}]$ show a much larger range in the unadjusted than in the adjusted data sets. In the former, fractional differences larger than 500% were found, whereas the fractional differences in the adjusted satellite data sets were within ±60%.

The figures and measures of this section indicate some similarities, but also some significant differences between the gage and satellite rainfalls. Multi-response permutation procedures (MRPP) (Mielke et al., 1981) were run to test whether the daily area-averaged satellite and gage rain estimates

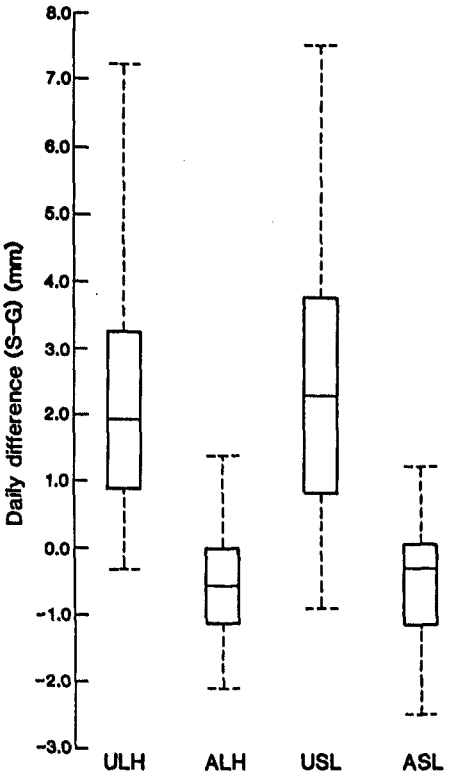


Fig. 3.10. Box-and-whisker plots of daily differences (S-G) in area-averaged rainfalls (mm). Four satellite permutations are represented: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL).

constituted one or more rainfall populations. Four MRPP runs were made with two groups in each run--the 31 daily area-averaged gage data and the 31 daily area-averaged data from one of the satellite permutations. The null hypothesis for all runs is that all partitions of the 64 rain estimates into

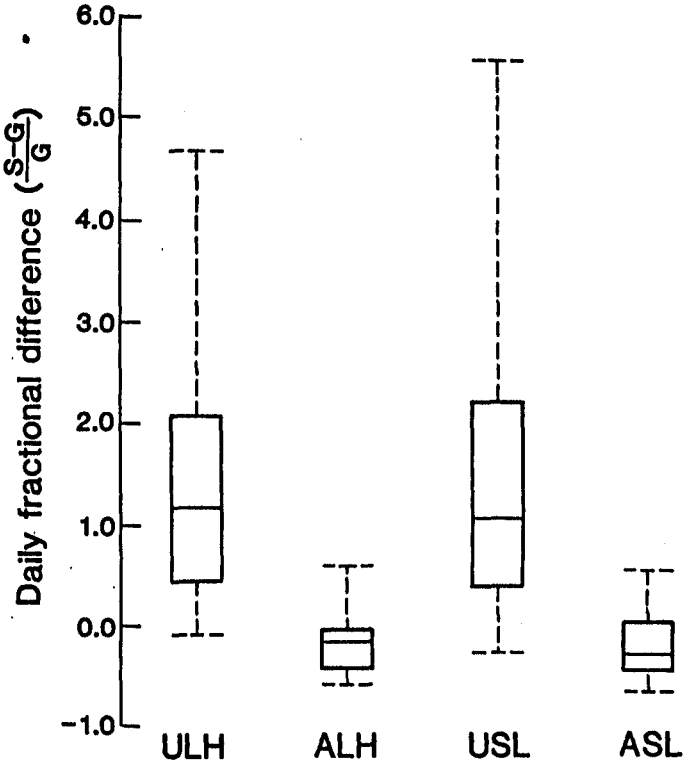


Fig. 3.11. Box-and-whisker plots of daily fractional differences [(S-G)/G] in area-averaged rainfall: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL) techniques.

two groups of 31 measurements could have occurred with equal chance. If the null hypothesis is proved (that is, if the MRPP test statistic has a large value relative to its value for all partitions possible with the data), then the satellite and gage data do not represent two groups. If the test statistic is small, then the satellite and gage data are indeed two different groups of data. The results of the four MRPP runs are that the gage data and each of the satellite data sets are two separate groups, significance being better than 1% for the unadjusted life history and streamlined techniques and better than 10% for the two adjusted schemes. The differences in these groups arise from the clustering of the data (Fig. 3.6). The adjusted satellite rainfalls occur most frequently in the 1-mm category, in contrast to the gage data which are most frequent at 3 mm.

From the foregoing tables and figures it can be seen that there is little difference in daily area-averaged rainfalls estimated from the life history technique or from the streamlined technique. The box-and-whisker plots and the medians are essentially identical. The other single-value statistics (e.g., percent of coincident relative maxima, correlation coefficient) indicate that the adjusted life history estimates have an edge over the adjusted streamlined estimates, but this difference is small. Relative to the factor-of-10 difference in the computational time required to make the life history versus the streamlined estimates, it is for most purposes inconsequential.

The MRPP tests do not support the hypothesis that satellite and gage estimates of rainfall for these 31 days are from the same group, and the other measures of comparison of the gage and satellite data show that the unadjusted rainfalls (that is, rain from the Florida relationships) are substantially greater than the gage data, and the environmentally adjusted rainfalls are 20% smaller than the gage data. The response function that results from the Gaussian interpolation of the adjustment factors at 34 radiosonde locations (shown in Fig. C.3) indicates that the interpolated field incorporates at most 73% of the original signal. A portion of the underestimation of the gage rainfall certainly arises from this source.

3.2 Hourly Rainfalls

The August data set has 744 hourly periods. However, because of missing satellite images, only 662 hourly periods in the life history data set (and 664 in the streamlined) coincide with an hourly gage estimate. Area-averaged rainfalls for these hourly periods were computed as for the daily data sets. Fewer than 4% of these hours had no rain anywhere in the region of study.

Smaller rain amounts and more scatter are to be expected in hourly versus daily comparisons of satellite and gage rainfalls. The former occurs because of the shorter accumulation period, and the latter because 24 satellite samples contribute to a daily rainfall estimate, but only one satellite image is used in the hourly estimate. (In contrast the gages sample continuously.) Scatterplots of hourly area- averaged amounts (Figs. 3.12 to 3.15) bear out these expectations. Amounts are smaller and, not surprisingly, the majority of the unadjusted satellite rainfalls (Figs. 3.12 and 3.13) exceed the corresponding gage amount. However, the adjusted satellite rainfalls (Figs. 3.14 and 3.15) appear to be more evenly split between amounts that are greater than and smaller than the appropriate gage value. Scatter has definitely increased

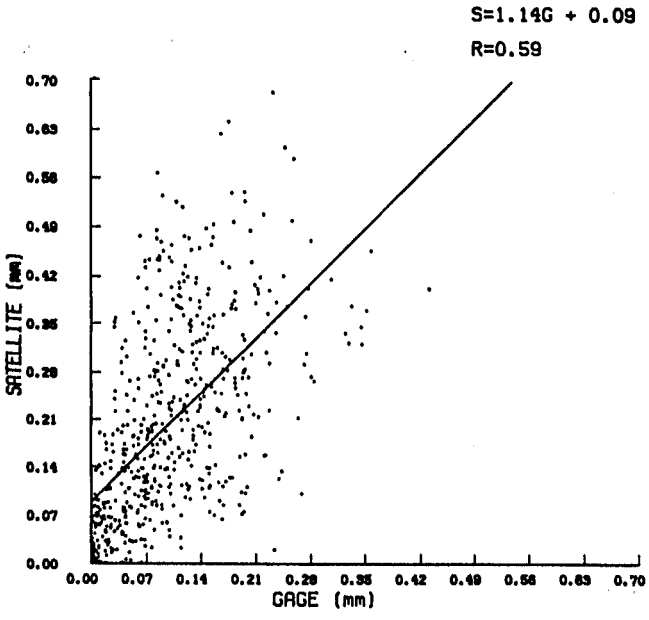


Fig. 3.12. Scatterplot of hourly area-averaged satellite and gage rainfalls for the unadjusted life history technique. The regression line has been plotted; the slope, intercept, and correlation coefficient are noted.

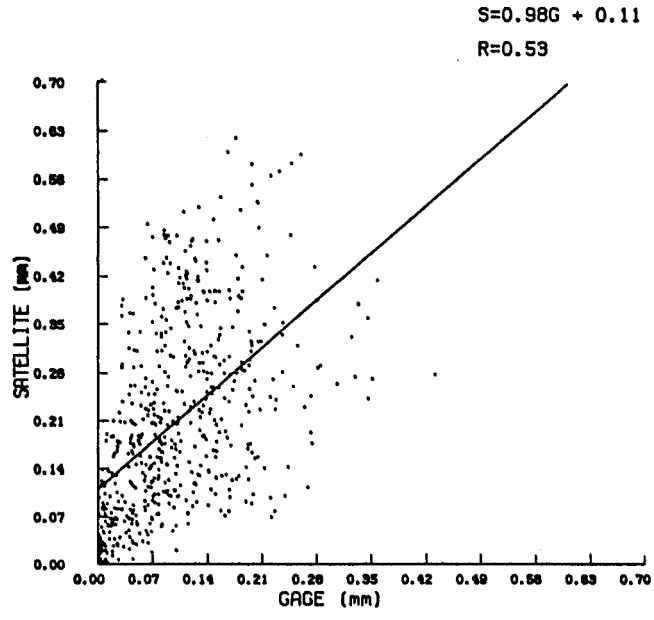


Fig. 3.13. Scatterplot of hourly area-averaged satellite and gage rainfalls for the unadjusted streamlined technique as in Fig. 3.12.

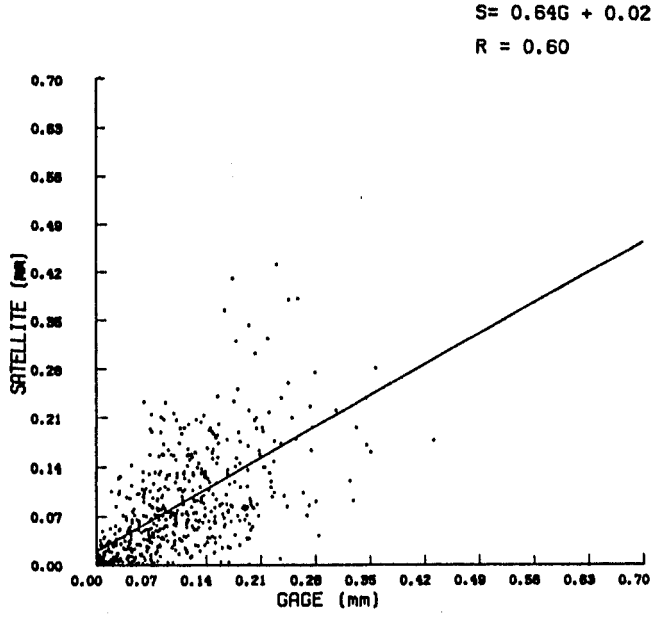


Fig. 3.14. Scatterplot of hourly area-averaged satellite and gage rainfalls for the adjusted life history technique as in Fig. 3.12.

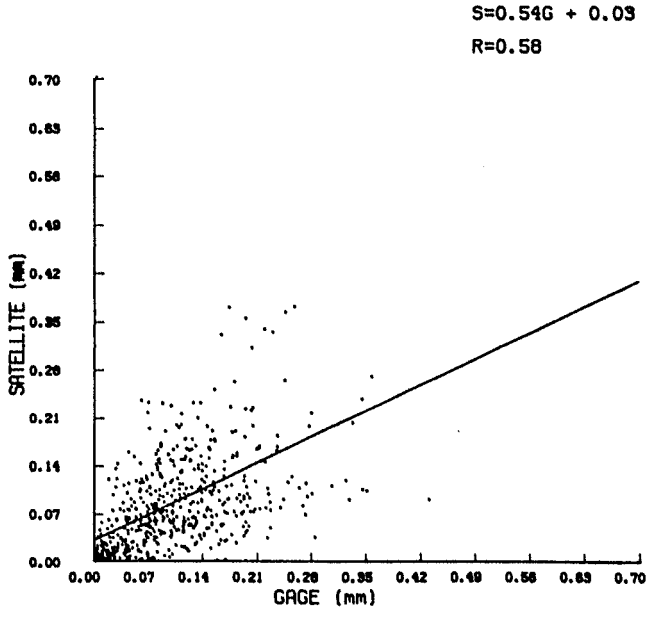


Fig. 3.15. Scatterplot of hourly area-averaged satellite and gage rainfalls for the adjusted streamlined technique as in Fig. 3.12.

and the correlation coefficients (Table 3.6) are smaller than their daily counterparts (Table 3.2) by 5 to 10 points. In these hourly data only 25-35% of the variance is accounted for. Slopes of the least squares fits are roughly comparable with their daily values and, when scaled by a factor of 24 (the difference in the accumulation period), the intercepts are comparable to their daily values, too. In contrast to the daily results, though, the environmental correction has only negligibly improved the correlation from the unadjusted to the adjusted data sets. As before, satellite values smaller than gage values indicate that the environmental adjustment is too large.

Table 3.6. Least-squares-fit parameters for satellite-raingage regressions of hourly area-averaged rainfall

Satellite Permutation	Least-Squares Linear Fit		
	ρ	Slope	Intercept (mm)
Unadjusted life history	0.59	1.14	0.09
Adjusted life history	0.60	0.64	0.02
Unadjusted streamlined	0.53	0.98	0.11
Adjusted streamlined	0.58	0.54	0.03
Perfect correspondence	1.00	1.00	0.00

The means and medians of the unadjusted and adjusted satellite data set (Table 3.7) reflect the utility of the environmental adjustment. The means, medians, and standard deviations of the unadjusted satellite data sets are all two to three times larger than the corresponding gage values. After adjustment these values differ only slightly from their counterparts in the gage data.

Box-and-whisker plots (Fig. 3.16) show the similarities and the differences among these five data sets. With the exception of the upper extreme values, the two unadjusted satellite permutations are almost indistinguishable from each other and obviously different from the three remaining sets. The gage and the two adjusted satellite data sets are fairly similar, but differences occur at the higher values. The upper quartile range of the satellite data, especially the life history results, is much larger than the corresponding gage range. The satellite data also have a slightly smaller interquartile range than the gage data show.

Stem-and-leaf plots (Fig. 3.17) highlight the details of the dissimilarities. (In these plots 0.43 mm, for example, is represented by a

Table 3.7. Statistics for hourly area-averaged satellite and raingage samples.

Data Set	N	Median (mm)	Mean (mm)	Std. Dev. (mm)
Unadjusted life history	662	.16	.20	.14
Adjusted life history	662	.06	.08	.08
Gage	662	.08	.09	.07
Unadjusted streamlined	664	.18	.20	.14
Adjusted streamlined	664	.06	.08	.07
Gage	664	.08	.09	.07

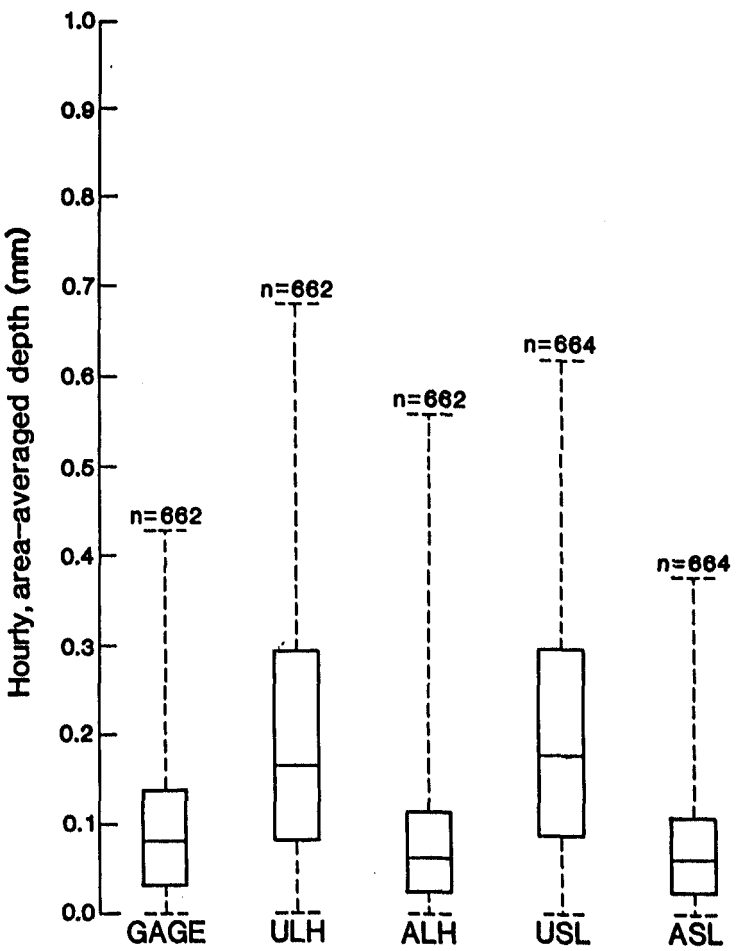


Fig. 3.16. Box-and-whisker plots for hourly area-averaged rainfalls (mm) for the gage data (GAGE), unadjusted life history (ULH), adjusted life history (ALH), unadjusted streamlined (USL), and adjusted streamlined (ASL) satellite data. The life history samples have 662 points and the streamlined samples have 664 points.

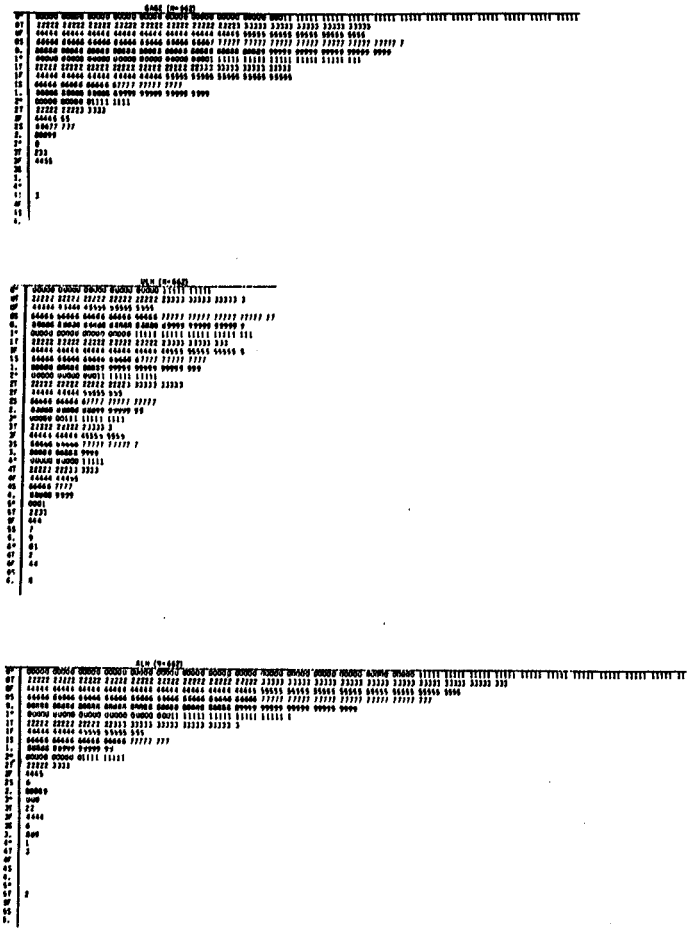


Fig. 3.17a. Stem-and-leaf plots of hourly area-averaged rainfalls of the gage (top), unadjusted life history (middle) and adjusted life history (bottom) data.

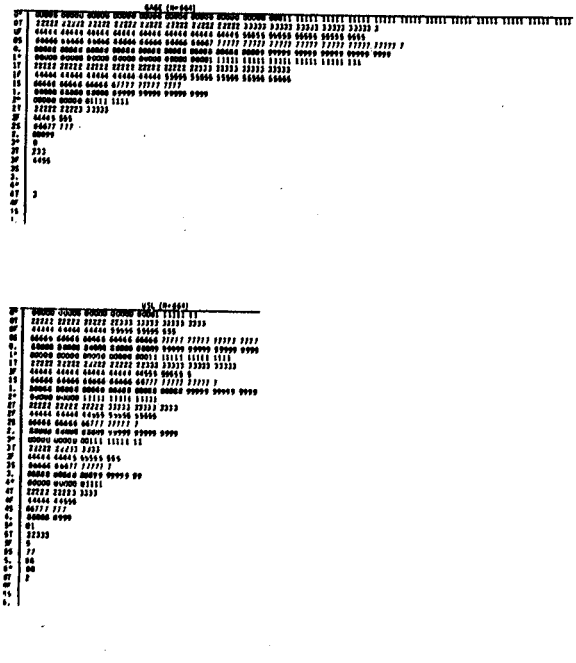


Fig. 3.17b. Stem-and-leaf plots of hourly area-averaged rainfalls of the gage (top), unadjusted streamlined (middle) and adjusted streamlined (bottom) data.

"stem" of 4 in the vertical and a "leaf" of 3 along the horizontal. Because of the number of observations, data for each tenths stem value are written over five horizontal lines.) The higher values of the unadjusted satellite distribution have been moved into smaller amounts by the environmental adjustment; the frequency of area-averaged rainfalls smaller than 0.10 mm has increased so that their numbers are larger than the gage data for the same interval. The gage data show two peaks at 0.00-0.01 mm and 0.06-0.07 mm. Only the adjusted streamlined data have captured this double peak.

Time series of the hourly gage and satellite rain amounts indicate the reason for the decrease in the correlation coefficient. In a representative, 4-day period of moderate gage rainfalls (Figs. 3.18 and 3.19), satellite estimates are both greater and smaller than the gage values. It is also evident that the gage data are more variable than the satellite data; the gage data show high-frequency fluctuations in the time series that are not evident in the satellite time series. Moreover, the gage data often show a peak where the satellite data indicate only a continuing trend of increasing or decreasing rainfall, and the streamlined time series have smoother signatures than their life history counterparts. It appears that satellite data do not capture the showery nature of convective rainfall when calculations are made with hourly imagery. The full time series (Appendix E) confirm these impressions.

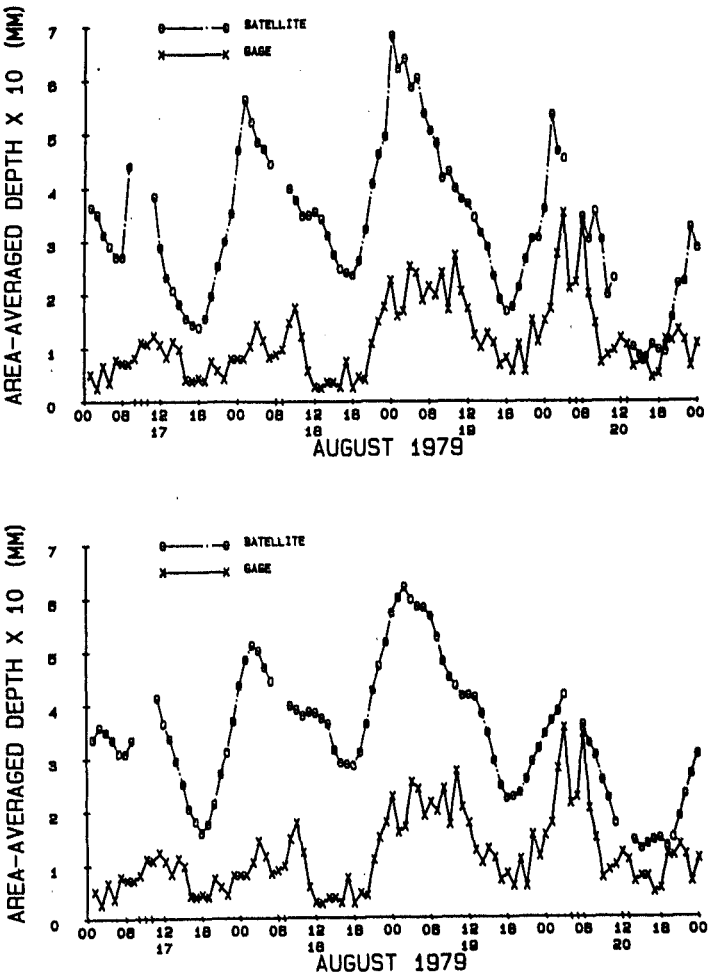


Fig. 3.18. Time series of hourly area-averaged rainfalls for August 17-20, 1979 from the unadjusted life history (top) and unadjusted streamlined (bottom) techniques.

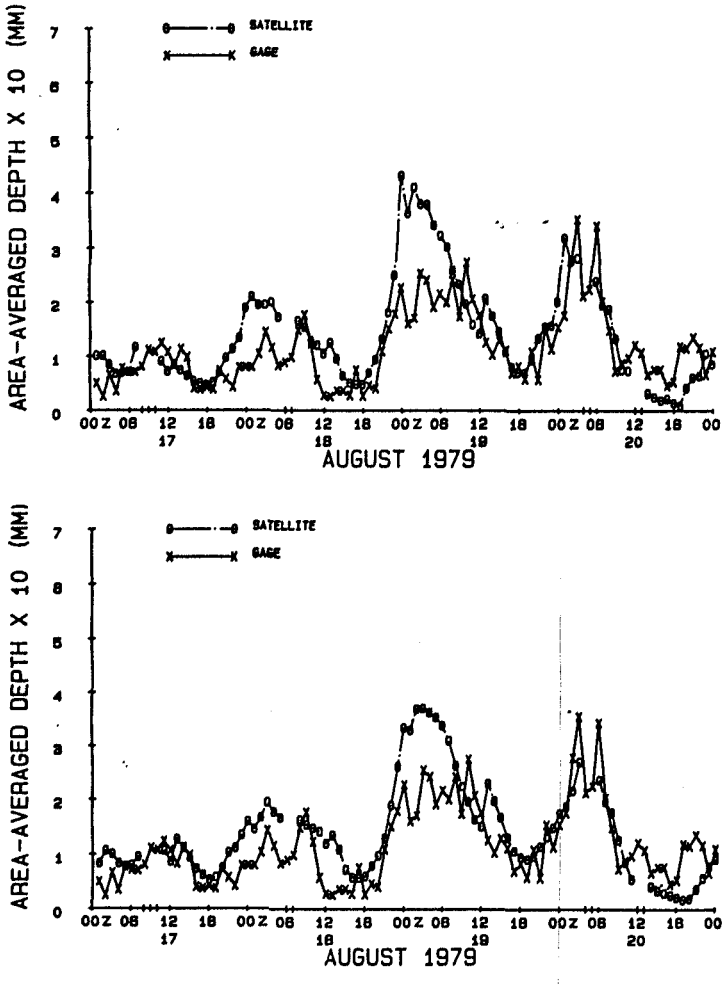


Fig. 3.19. Time series of hourly area-averaged rainfalls for August 17-20, 1979 from the adjusted life history (top) and adjusted streamlined (bottom) techniques.

Relative maxima and minima (Table 3.8) are as likely to occur 1 hour apart as to be coincident. Two- and three-hour differences were counted, but it becomes unreasonable to ascribe relative maxima to the same rain event when time differences are greater than 3 hours. Frequently there is no range 0-1 inches is blank. The shading of the NCC data is it becomes unreasonable to ascribe relative maxima to the same rain event when time differences are greater than 3 hours. Frequently there is no corresponding peak or valley in the satellite data, again emphasizing that hourly satellite data do not represent very well the convective showers that occur on 10-20 minute time frames.

A harmonic analysis was performed on satellite data sets in which missing hourly amounts were interpolated to form a continuous series. For a period with n missing hours, a first estimate of the missing rainfall is made by linearly interpolating between the preceding and following "good" hours of rainfall. The interpolated amounts, however, are constrained to equal the accumulated amounts calculated by the satellite for the n hours. Any difference is equally divided among the n hours and algebraically added to the first-guess rainfall.

Of the three harmonics showing the most power (Table 3.9), no single harmonic accounts for more than 20% of the variance, and the three greatest in any set barely account for 30% of the variance. Each of the five data sets shows the expected diurnal cycle. Interestingly, the gages and the two

Table 3.8. Coincidence of maxima and minima times in the satellite and gage samples of hourly area-averaged rainfall

Satellite Permutation	Differ by Coincident	No correspondence one hour	two hours	No. of gage points	
Unadjusted life history	36	50	15	67	168
45	43	10	73	171	
Adjusted life history	46	42	16	59	163
40	48	16	71	175	
Unadjusted streamlined	20	33	8	106	167
18	30	4	113	165	
Adjusted streamlined	28	33	8	95	164
33	38	11	82	164	

Note: The relative maxima as defined by the gage time series are counted. Rainfall maxima are tabulated on the top line and minima are on the bottom line of each pair.

adjusted satellite sets have indications of a 10.3-day period, while the gage and the two unadjusted satellite data sets show a 6-day period. The semidiurnal cycle also appears in all these data sets, but is not shown in this table because its explained variance is on the order of 1%.

All the difference measures of hourly rainfalls that involve ratios (Table 3.10) are much larger than those of daily rainfalls. For the month as a whole, the unadjusted satellite rainfalls are twice as large as the gage monthly rainfall, and the adjusted satellite amounts are 20% smaller. (See column two). On an hour-by-hour basis (column three), the satellite estimates are too high by a factor of 4 in the unadjusted and by 20-30% in the adjusted satellite techniques. Absolute factors of difference (column four) are about 5.0 for the life history and 3.5-4.0 for the streamlined data sets. The normalized RMS error (column six) is 3 to 5 times greater than its previous values for the daily results. RMS errors (column five), though, are almost simply scaled by the difference in the two accumulation periods. Rather than being a factor of 24 smaller, RMS errors are a factor of 15-20 smaller.

Despite the large values appearing in Table 3.10, a very different evaluation of the satellite rainfalls is apparent in the distribution of satellite and gage differences (Fig. 3.20). The two unadjusted satellite rainfalls have a positive bias ($\bar{S}>\bar{G}$) and relatively large ranges. The adjusted satellite results are more symmetrically distributed around zero, half the differences being within ±0.05 mm and 90% within 0.15 mm.

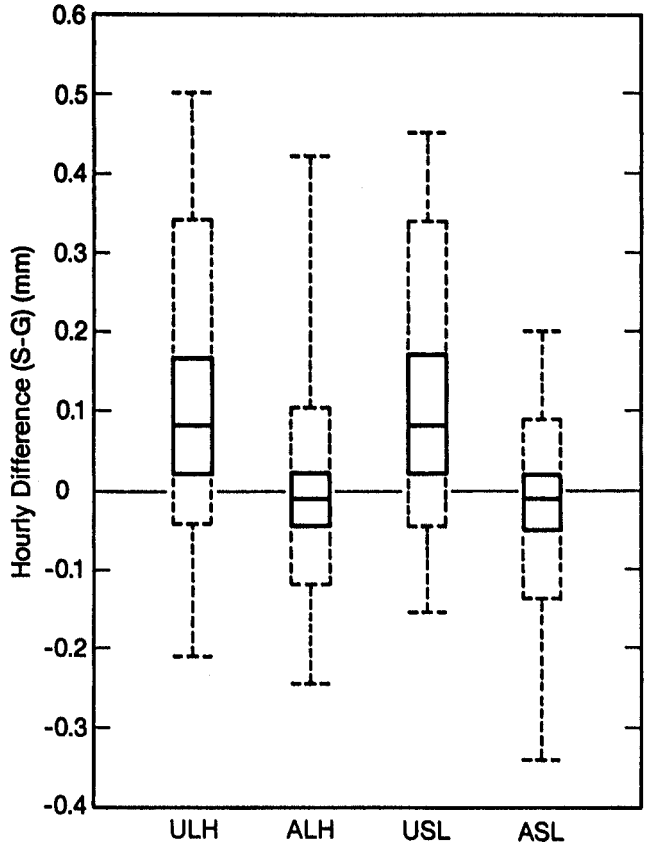
Table 3.9. Harmonics and their fractional variance for the three harmonics with the greatest power

Data set	Harmonic	Period (days)	Fractional Variance
Gage	5	6.2	.17
	3	10.3	.06
	31	1.0	.05
Unadjusted life history	31	1.0	.20
	5	6.2	.05
	7	4.4	.03
Adjusted life history	31	1.0	.16
	3	10.3	.08
	9	3.4	.04
Unadjusted streamlined	31	1.0	.16
	5	6.2	.08
	7	4.4	.04
Adjusted streamlined	31	1.0	.16
	3	10.3	.09
	9	3.4	.05

Table 3.10. Difference measures for hourly area-averaged rain samples

Satellite Permutation	Monthly Ratio	Mean Hourly Ratio	Factor of Diff.	RMS Error (mm)	Norm. RMS Error
Unadj. L.H.	2.12	3.97	5.10	0.16	7.04
Adj. L.H.	0.87	1.31	5.23	0.07	1.97
Unadj. S.L.	2.13	3.99	4.29	0.16	6.99
Adj. S.L.	0.81	1.20	3.40	0.07	1.47
Perf. Corresp.	1.00	1.00	1.00	0.00	0.00

Fig. 3.20. Modified box-and-whisker plots of hourly differences (S-G) in area-averaged rainfalls (mm). The dashed box encloses 90% of the sample. Four satellite permutations are represented: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL) techniques.



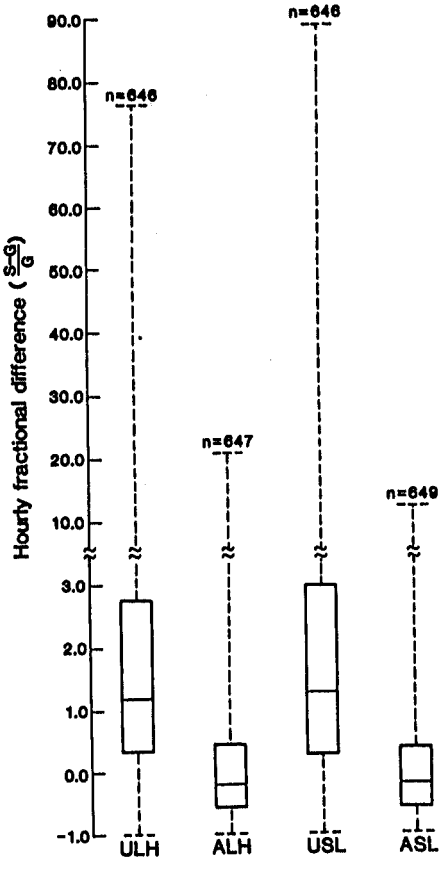


Fig. 3.21. Modified box-and-whisker plots of hourly fractional differences [(S-G)/G] in area-averaged rainfalls. The dashed box encloses 90% of the sample. Four satellite permutations are represented: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL) techniques.

Fractional differences $[(\bar{S}-\bar{G})/\bar{G}]$ are large again for the unadjusted and moderately large for the adjusted satellite rainfalls (Fig. 3.21). The large positive bias of the median in the former and smaller negative bias in the latter, noted in Table 3.10, are reflected here. Roughly half the adjusted rainfall differences are within ±50% of the gage amount, but 90% of the differences cover a large range from -90% to +300% of the gage amount.

4. COMPARISONS OF SATELLITE AND GAGE POINT RAINFALLS

Point rainfalls are the forte of gages. A standard gage with its 8-inch diameter is, with qualification, recording "true" rainfall at a location. Caveats include factors such as exposure and wind conditions, but gage errors in the measurement of convective rainfall are, by contrast, much smaller than for the measurement of snow (Peck, 1980). However, rainfall at a point may not be representative of the rain 3 km away because of the tremendous gradients that can occur in convective rainfalls (Woodley et al., 1975). The satellite resolution, on the other hand, is so coarse (and is further degraded in this study to ~500 km^2) that rainfall estimated for a single pixel is really an area-averaged rainfall. Point comparisons are made because some techniques (e.g., Scofield, 1987) routinely generate point rainfall amounts, albeit with the full resolution IR data. This section gives an idea of how far satellite data can be pushed.

Satellite "point" rainfalls are generated from the satellite rainfall arrays. A bilinear interpolation is used to compute satellite point rainfall from the four pixels adjacent to or at the gage position. The daily and hourly data sets are interpolated separately. The interpolation can be assessed by comparing area-averaged rainfalls computed from the interpolated and array data sets of satellite rainfalls. Some noise is introduced by the interpolation procedure, but the interpolated data sets do not produce area-averaged rainfalls that are very different from the array data. Correlations between interpolated and array rainfalls (Table 4.1) are high for both the

Table 4.1. Correlation between area-averaged rainfalls from array and interpolated satellite data sets

Satellite Permutation	Correlation Coefficient	
	Daily	Hourly
Unadj. life history	.931	.911
Adj. life history	.928	.925
Unadj. streamlined	.947	.916
Adj. streamlined	.927	.919

daily and hourly time frames; correlation coefficients are greater than 0.90 for all four satellite permutations. Distributions of area-averaged amounts (Figs. 4.1 and 4.2) are almost indistinguishable for the interpolated and array subsets of any permutation, exclusive of some striking exceptions at the highest values.

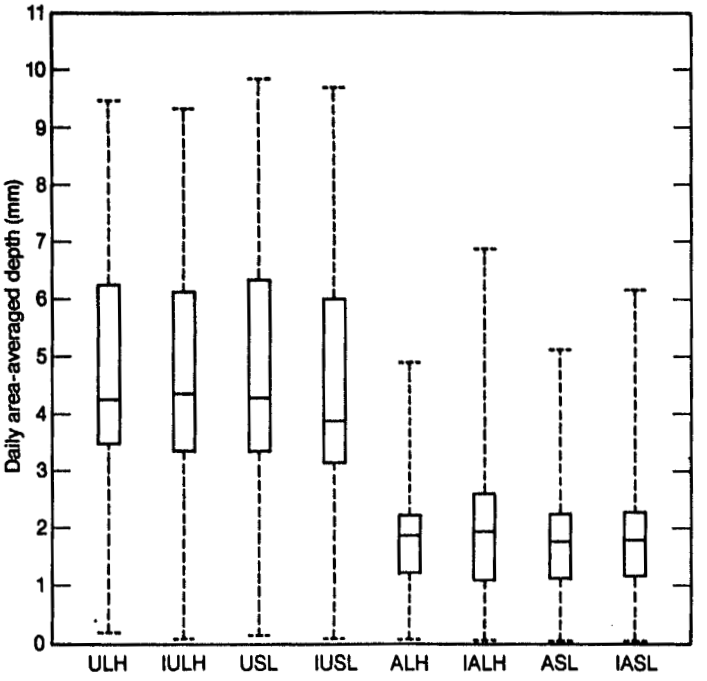


Fig. 4.1. Box-and-whisker plots of daily area-averaged satellite depths (mm) from array and interpolated data sets: Unadjusted life history pairs (ULH); adjusted life history pairs (ALH); unadjusted streamlined pairs (USL); adjusted streamlined (ASL) pairs. The prefix I denotes interpolated data sets.

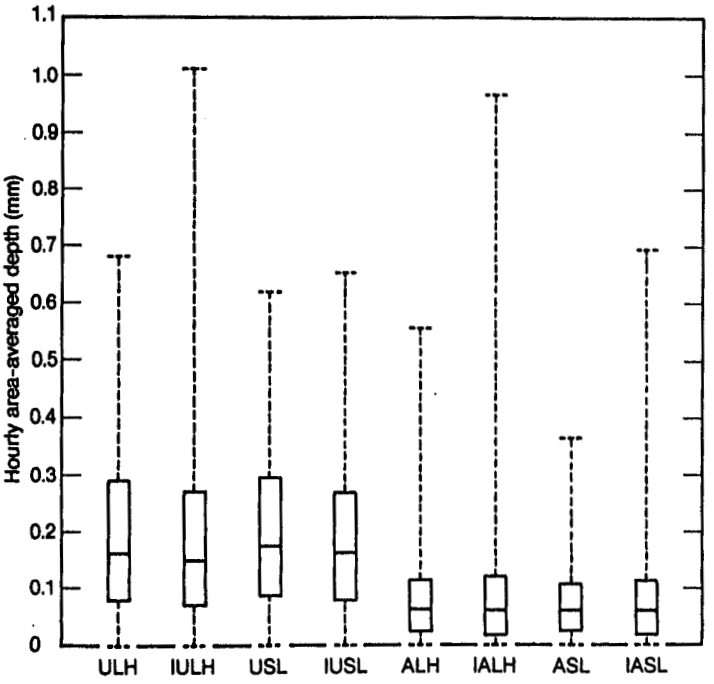


Fig. 4.2. Box-and-whisker plots of hourly area-averaged satellite depths: Unadjusted life history pairs (ULH); adjusted life history pairs (ALH); unadjusted streamlined pairs (USL); adjusted streamlined pairs (ASL). The prefix I denotes interpolated data sets.

4.1 Daily Rainfalls

Gross classification of rain versus no-rain conditions is possible with point data and is a useful, if low-level, assessment of any rain estimation technique. An "event" in this context occurs at a location during a specified time period and can be either rain or no rain. Identical classifications occur when the satellite and gage both indicate rain or both indicate no rain. The majority of the gage and satellite events is identically classified (Table 4.2). When nonidentical classifications occur, the satellite is more likely to indicate rain when the gage shows none than vice versa. The net effect of the environmental adjustment on the satellite estimates (compare the unadjusted with the adjusted permutation) is to move about 10% of the satellite rain sample into the no-rain class. This amounts to a slight (3-4 point) improvement in the number of identical classifications, but at the expense of a 4-point drop in the rain-rain class.

Table 4.2. Contingency tables of the percentage of occurrence of rain/no-rain events for satellite (S) and gage (G) daily point rainfalls

Unadjusted life history

S \ G	R	NR
R	15.03	19.81
NR	5.70	59.46

Adjusted life history

S \ G	R	NR
R	11.46	12.76
NR	9.27	66.51

Unadjusted streamlined

S \ G	R	NR
R	15.31	20.90
NR	5.42	58.37

Adjusted streamlined

S \ G	R	NR
R	11.05	12.39
NR	9.68	66.88

Note: There are 26,315 pairs in each sample.

It is not surprising that the most frequently occurring class is the no-rain category. It is clear to those living in this region that there are more days when it is not raining than when it is, during the summer months. Although it is just as important to distinguish rain from no-rain situations, suffice it to say that the satellite does an adequate job of identifying no-rain occurrences. Only nonzero rainfalls are considered in the quantitative analyses that follow.

A closer look at these daily rainfalls is provided by box-and-whisker plots (Fig. 4.3). These plots reveal that the highest rainfall amounts seen in the gage data are not achieved by the satellite data. This is to be expected, because the satellite estimate is an area average and uniformly high rainfall amounts do not generally occur over the entire region of any one satellite sample. The unadjusted satellite amounts almost replicate the gage distribution of the point rainfalls in the lower three quartiles, but the medians and hinges are somewhat higher than the gage values; median and mean values in Table 4.3 corroborate this. The adjusted satellite medians and hinges on the other hand are slightly smaller than the gage amounts in the lower three quartiles.

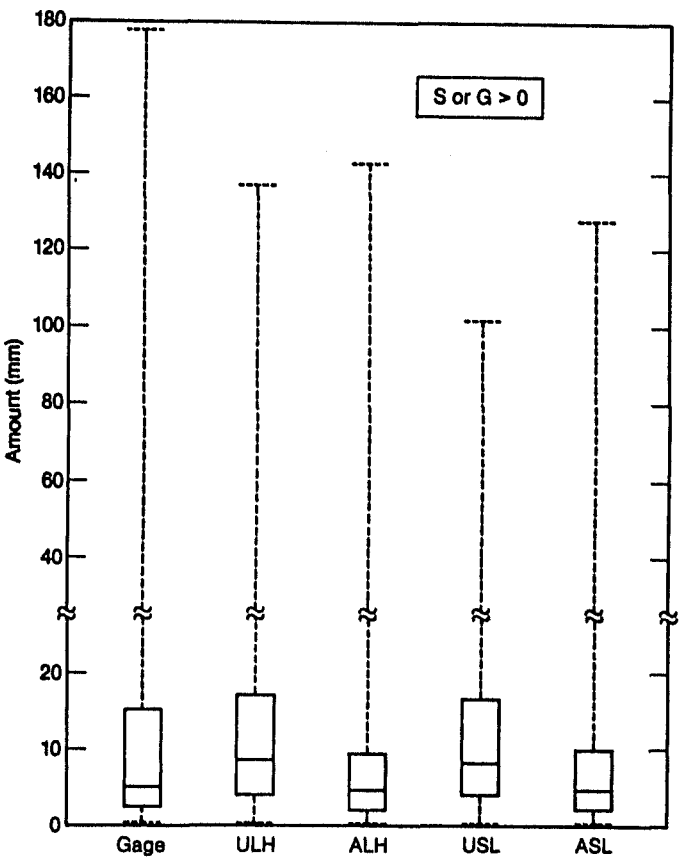


Fig. 4.3. Box-and whisker plots of daily point amounts (mm) for gage (GAGE) and interpolated satellite data sets. Four satellite permutations are represented: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL).

There is no significant correlation between the gage and satellite estimates of daily point rainfalls. (See Table 4.4, which excludes events when both the satellite and gage are zero. When the zero-zero pairs are included the correlation coefficients are increased by about 20 points.) The fact that the gage and satellite point rainfalls are uncorrelated most

Table 4.3. Statistics for nonzero daily point satellite and raingage samples

Data Set	No. of Samples	Median (mm)	Mean (mm)	Std. Dev. (mm)
Satellite-unadj. life history	9,164	8.64	12.93	13.49
Satellite-adj. life history	6,371	4.57	7.87	10.85
Satellite-unadj. streamlined	9,524	8.38	12.22	11.91
Satellite-adj. streamlined	6,165	4.83	7.92	10.36
Gage	5,448	5.08	11.07	13.46

certainly arises from the vast spatial differences between the satellite and gage point resolution. These findings concur with the Florida results of Meitin et al. (1981) that indicate low correlations for point rainfalls. Even with full-resolution (8 km x 8 km) GOES IR data, correlations for 6-hourly rainfalls (the longest accumulations discussed by Meitin et al.) are 0.4.

Difference measures (Table 4.5) also reflect the large scatter between gage and satellite point rainfalls. Compared with the daily area-averaged

Table 4.4. Least-squares fit parameters for satellite-raingage regressions of nonzero daily point rainfall

Satellite Permutation	No. of Samples	Least-Squares Linear Fit		
		ρ	Slope	Intercept (mm)
Unadj. life history	10,621	0.28	0.34	9.22
Adj. life history	8,765	0.22	0.18	4.50
Unadj. streamlined	10,899	0.28	0.30	9.02
Adj. streamlined	8,659	0.21	0.16	4.50
Perf. corresp.	-	1.00	1.00	0.00

Table 4.5. Difference measures for nonzero daily point rain samples

Satellite Permutation	Monthly Ratio	Mean Daily Ratio	Factor of Diff.	RMS Error	Norm. RMS Error
Unadj. L.H.	1.96	2.53	4.55	15.75	7.04
Adj. L.H.	0.83	1.01	4.91	13.72	3.55
Unadj. S.L.	1.93	2.55	4.54	14.73	7.25
Adj. S.L.	0.81	1.01	5.19	13.72	3.88
Perf. Corresp.	1.00	1.00	1.00	0.00	0.00

results (Table 3.6), monthly (R_M) and daily (R_D) biases are about the same, but the factor of difference ratio (E_R), the RMS error and the normalized RMS error are substantially (3 to 13 times) larger than before. The effect of area averaging on smoothing out differences in gage and satellite rainfalls is evident.

Absolute differences in point rainfalls (Fig. 4.4) are small enough that the estimation of convective rainfall from satellite data looks promising. Although the extreme daily differences can be as large as 160 mm, half the differences are no greater than 10 mm and 90% are within approximately ±20 mm for all four satellite permutations. Differences in the adjusted satellite sets are rather symmetrically distributed around zero, indicating little bias in contrast to the unadjusted sets. The environmental adjustment has resulted in longer tails in these adjusted data, though.

4.2 Hourly Rainfalls

The most extreme test in comparing gage and satellite data is that for hourly point rainfalls. Such comparisons push the satellite data to their limits, for the gage and satellite data are from two very different time and space scales. Differences in spatial scales amount to comparing a 3 km^2 (or smaller) gage area with a 500 km^2 satellite area, as discussed in section 4.1. In addition, hourly comparisons pit data from gages that sample almost continuously against satellite data that are nearly instantaneous images recorded once per hour. The final result is that small-scale convective showers that often last no longer than 20 minutes usually fall between the cracks of the satellite space and time resolutions.

Contingency tables for hourly point rainfalls (Table 4.6) show that better than 95% of the events are identically classified, but almost all of

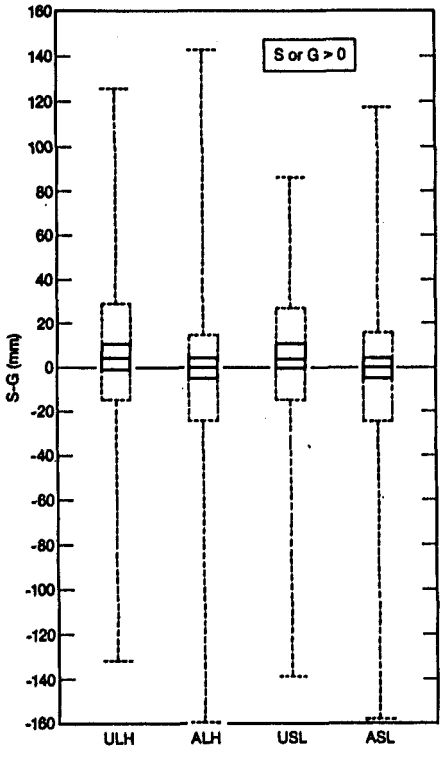


Fig. 4.4. Modified box-and-whisker plots of daily point differences (mm) of interpolated satellite depths less gage amounts. Four satellite permutations are represented: Unadjusted life history (ULH); adjusted life history (ALH); unadjusted streamlined (USL); adjusted streamlined (ASL). The dashed box indicates the range of 90% of the data.

Table 4.6. Contingency tables of the percentage of occurrence of rain/no-rain events for 569,426 satellite (S) and gage (G) hourly point rainfalls

Unadjusted life history

S \ G	R	NR
R	0.79	3.56
NR	1.46	94.19

Adjusted life history

S \ G	R	NR
R	0.50	1.90
NR	1.75	95.85

Unadjusted streamlined

S \ G	R	NR
R	0.78	3.58
NR	1.48	94.17

Adjusted streamlined

S \ G	R	NR
R	0.50	1.92
NR	1.75	95.83

these are the no-rain/no-rain cases. The nonidentical classifications occur about 4-5% of the time. In the unadjusted satellite data, the satellite is more likely to indicate rain when the gage shows none than vice versa. However, either of the two nonidentical classifications is equally likely in the adjusted satellite data set. Again, it is not unexpected that hours with no rainfall predominate in the hourly point estimates of rainfall. The statistics and analyses that follow are based on the nonzero data alone.

Satellite average values and standard deviations of hourly rainfall at point locations (Table 4.7) are smaller by 25-50% than the gage values in all permutations. Median values in the unadjusted and adjusted satellite data bracket the gage median, the unadjusted values being closer than the adjusted values. The box-and-whisker plots (Fig. 4.5) again indicate that the lower three quartiles are not very different in their distributions, but that large differences occur for values in the upper quartile. None of the peak values from any of the satellite permutations is as large as the biggest amounts measured by the gages. This probably arises from the bilinear interpolation rather than being inherent in the satellite techniques; higher rainfall values are more likely to be surrounded by lower values than vice versa.

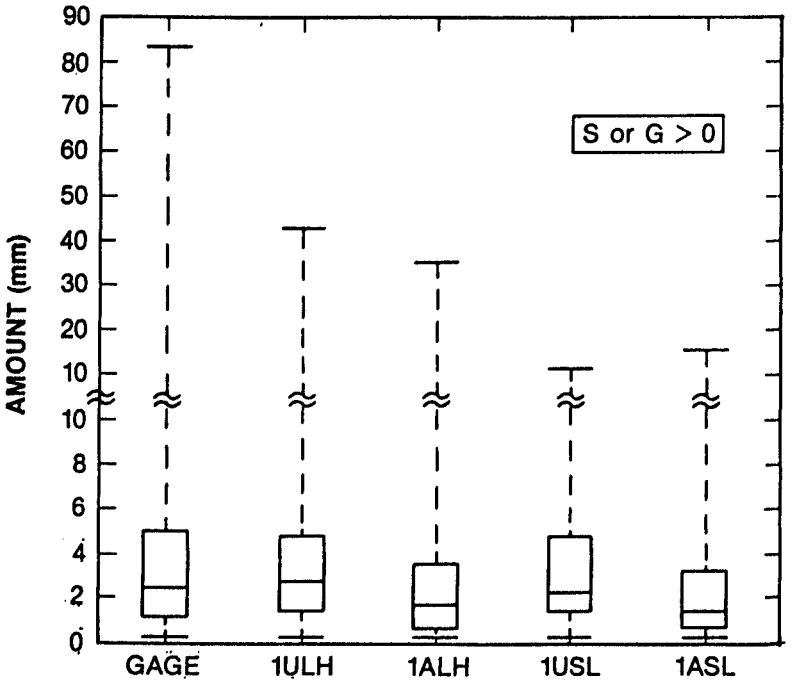
Table 4.7. Statistics for nonzero hourly point satellite and raingage samples

Data Set	No. of Samples	Median (mm)	Mean (mm)	Std. Dev. (mm)
Satellite-unadj. life history	24,732	2.79	3.61	3.30
Satellite-adj. life history	13,675	1.78	2.82	3.30
Satellite-unadj. life history	24,863	2.29	3.30	2.56
Satellite-adj. streamlined	13,818	1.52	2.34	2.41
Gage	12,884	2.54	4.14	5.69

Despite the large differences between satellite and gage rainfalls (Fig. 4.6) (and the largest differences occur when the satellite amount is smaller than the corresponding gage value), half the satellite estimates are within ±2 mm of the gage values and 90% are within ±8 mm. The Florida results of Meitin et al. (1981) for hourly point rainfalls show low correlations ($\rho \approx 0.1$), even with full resolution (8 km x 8 km) GOES IR data. Similarly, there is no correlation for the hourly point data (Table 4.8) and the fitted least squares lines are horizontal with some positive offset.

The difference measures (Table 4.9), though not as large as the values for the daily point estimates are among the largest computed in this study.

Fig. 4.5. Box-and-whisker plots of hourly point amounts (mm) for gage and interpolated satellite data sets. Four satellite permutations are represented: Unadjusted life history (1ULH); adjusted life history (1ALH); unadjusted streamlined (1USL); adjusted streamlined (1ASL). The prefix 1I denotes the hourly interpolated data sets.



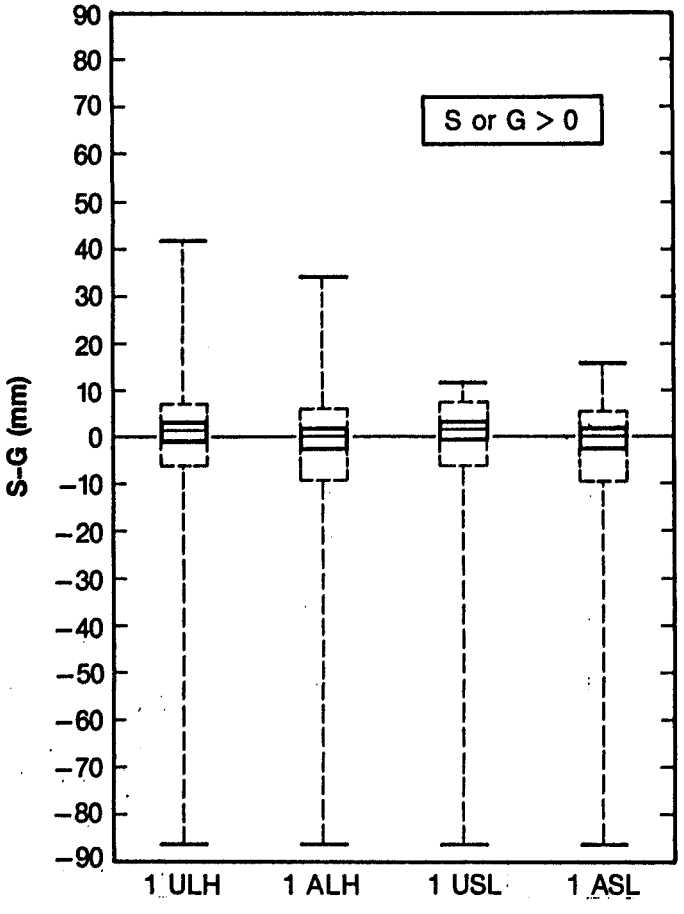


Fig. 4.6. Modified box-and-whisker plots of hourly point differences (mm) of interpolated satellite depths less gage amounts. Four satellite permutations are represented: Unadjusted life history (1ULH); adjusted life history (1ALH); unadjusted streamlined (1USL); adjusted streamlined (1ASL). The dashed box indicates the range of 90% of the data. The prefix 1I denotes the hourly interpolated data sets.

Table 4.8. Least-squares-fit parameters for satellite-raingage regressions of nonzero hourly point rainfall

Satellite Permutation	No. of Samples	Least-Squares Linear Fit		
		ρ	Slope	Intercept (mm)
Unadj. life history	33,066	-0.04	-0.03	2.74
Adj. life history	23,637	-0.05	-0.03	1.71
Unadj. streamlined	33,309	-0.07	-0.05	2.53
Adj. streamlined	23,836	-0.07	-0.03	1.43
Perf. corresp.	-	1.00	1.00	0.00

Like the area-averaged results, these hourly measures are smaller than their daily counterparts. But unlike the area-averaged results, the hourly measures are not simply scaled by a factor of 24, which is the difference in the accumulation period. The scaling factor varies, but is generally much smaller, indicating more scatter in the hourly point estimates than in the daily.

Table 4.9. Difference measures for nonzero hourly point rain samples

Satellite Permutation	Monthly Ratio	Mean Hourly Ratio	Factor of Diff	RMS Error	Norm. RMS Error
Unadj. L.H.	1.7	0.97	3.94	5.33	3.22
Adj. L.H.	0.73	0.46	4.38	5.59	2.33
Unadj. S.L.	1.5	0.91	3.96	5.08	2.82
Adj. S.L.	0.61	0.41	4.42	5.33	2.08
Perf. Corresp.	1.00	1.00	1.00	0.00	0.00

5. SPATIAL PATTERN STUDIES

The forte of the geostationary satellite platform is to view large areas relatively homogeneously in space and in time. This capability provides a potential for the satellite to fill in spatial information around gage data, to close gaps in radar coverage, and provide rainfall estimates in data-void regions. Thus, not only is the correspondence of satellite rainfalls with gage amounts and with gage temporal information important, but the correspondence of the spatial patterns is also an objective.

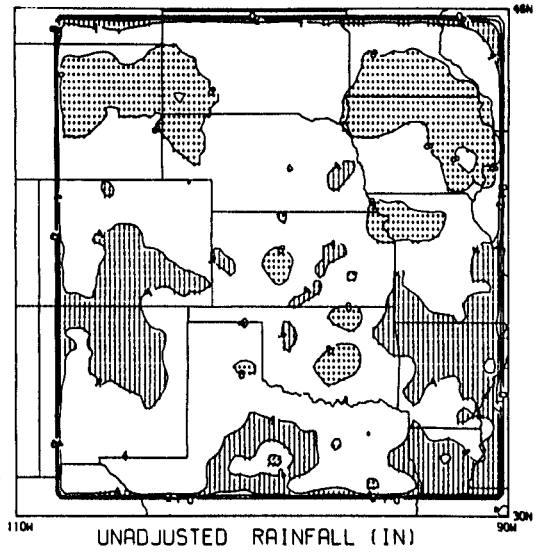
This section subjectively and qualitatively discusses spatial patterns for monthly rainfalls. The first section compares satellite and gage fields for the region of study. Comparisons are made for the unadjusted and adjusted life history and streamlined techniques. The remainder of the section explores the sensitivity of satellite isohyets to the temporal resolution of the imagery used for the rainfall computations.

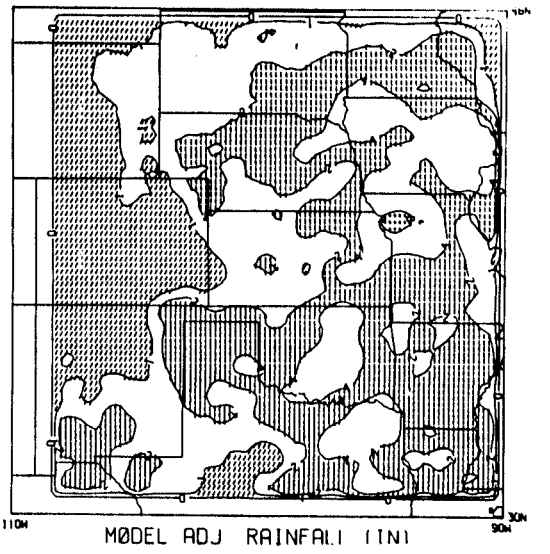
5.1 Monthly Rainfall Patterns

The National Climate Data Center (formerly NCC, the National Climatic Center) publishes a monthly rainfall analysis for the contiguous United States. The NCC analysis for August 1979 appears in Fig. 5.1, along with similar maps derived from the unadjusted and adjusted life history techniques. Per the NCDC convention, the units are inches and the contour lines are 1, 2, 4, and 8 inch. In both the satellite and gage plots, the noted in the figure legend, and the characters used in the shading of the satellite data are an attempt to reproduce the gage shading.

At first glance the fields in Fig. 5.1 appear to be from three different events. Each indicates a finite amount of rain over the entire region, but beyond that there is little overall agreement. There are regions of similarity, though. The unadjusted life history pattern with its 8-inch contour over Iowa and southern Minnesota covers the same states as an 8-inch isohyet in the gage data. The adjusted satellite and gage patterns exhibit a rough correspondence in that each has more rain in the eastern half of the array than in the western. Finally, 4-inch contours are indicated by the adjusted life history plot and the gage plot in the Texas panhandle and over central Oklahoma, although the locations do not exactly coincide.

The effect of the environmental adjustment is obvious between the two satellite plots. In general the unadjusted life history contours have higher values and are more extensive than the adjusted life history or gage estimates. With respect to the gage field, the impression is that the environmental adjustment creates too large a decrease in the satellite rainfall over the period of the month, particularly in the western half of the region. For instance, before adjustment the satellite shows Colorado and New Mexico to be receiving 2 inches and a substantial area receiving >4 inches. After adjustment, the satellite rains are less than 1 inch. The gages, on the other hand, indicate more than half of the area receiving 2 inches and embedded pockets receiving 4-inches. Similarly, both the unadjusted satellite data and the gage data show >8 inches over southern Minnesota and northern Iowa. After correction though, the satellite shows 2 to 4 inches for the same area.





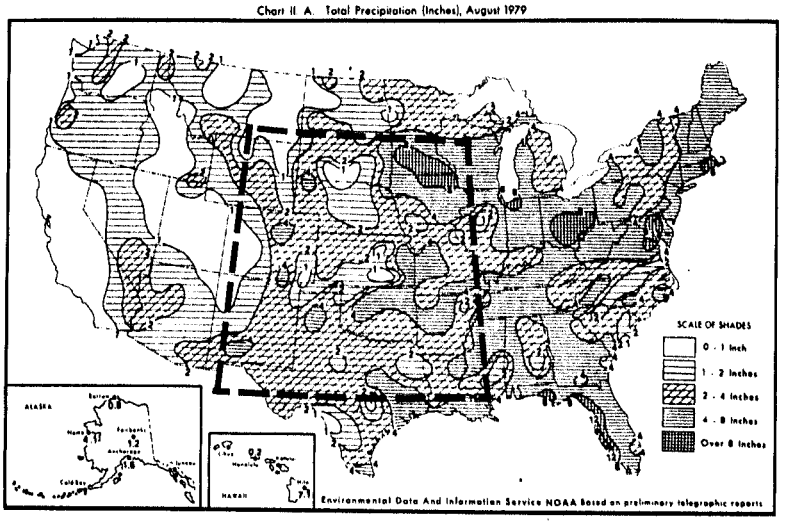


Fig. 5.1. Satellite rainfall inferred from Florida relationships and the life history technique (top left), satellite rainfall adjusted by the environmental correction (top right), and gage rainfall after NCDC (bottom) for August 1979.

Monthly rain volumes for this region support the contention that the environmental correction causes too large a decrease in the calculated rainfall. The volume of rain falling in this region is 4.03×10^{11} m^3 estimated by the unadjusted life history technique, 1.65×10^{11} m^3 by the model-adjusted life history technique, and 2.22×10^{11} m^3 from the gage data. (The gage isohyets were hand planimetered.) In other words, there is an 81% difference between the gage volume and the uncorrected life history estimate, and a 26% difference for the final satellite product. In the first instance the satellite amount is high compared with the gage data and in the second it is low.

5.2 Image Frequency Sensitivity Studies

Clearly, one requirement for an operational satellite rain estimation technique is that the computations keep abreast of the incoming data and be completed within the interval between advisories. If the computer on which a technique is run is a dedicated machine, there is no problem given today's 1 MIP machines. On a nondedicated VAX 11/780, for example, the author estimates that hourly rain amounts for the global tropics and subtropics (30°S to 30°N) can be completed in less than 3 minutes per image (clock time), using the technique discussed here and full-resolution GOES IR data.

Under certain conditions images less frequent than hourly are used to compute rainfall. For instance with a nondedicated machine, computation time may need to be minimized; calculations based on fewer images is one way to accomplish this. Alternately, operational data bases exist where only selected images are available. The Japanese GMS system with its 3-hourly cycle of images is an example. In these cases the effect of the temporal resolution of the data on the rain estimates must be determined.

The temporal resolution of the data can affect both rainfall amount and rainfall pattern. In section 4 hourly satellite rainfalls for the life history and streamlined techniques were assessed. This section addresses both the amount and pattern effects for rainfalls that are computed from successively less frequent images (3-, 6- and 12-hourly imagery) with the life history technique alone. The 1-hourly, 1/5° resolution data constitute the standard.

Over the course of the month, cumulative regional rain volumes computed from these four image frequencies differ by less than 10% (Table 5.1). From 1-hourly to 3-hourly to 6-hourly imagery, the rain volumes monotonically decrease by about 4%. With the 12-hourly imagery the trend reverses, and the rain estimates are 2% larger than the standard. Differences for the 12-hourly imagery may depend on what time the clouds are sampled during the diurnal cycle. In this case the 0000 and 1200 GMT images (roughly 0600 and 1800 LST) were used. However, it can be concluded that if one wishes to be within 10% or better of the 1-hourly rainfall standard, one can compute with images as infrequent as two per day.

The isohyets (Fig. 5.2), on the other hand, lead to a very different conclusion. There are minor changes in the isohyetal patterns from 1-hourly to 3-hourly data, but the patterns for the less frequent imagery show marked dissimilarities. For the 6-hourly calculations, the broad areas of >8 mm in the northern third of the array have increased in size and the region between

Table 5.1. Satellite rain volumes inferred from the life history technique for August 1979, as a function of image frequency

Image Frequency	Rain Volume ($m^3 \times 10^9$)	Difference From Standard (%)
1 h	438.74	0.0
3 h	421.01	-4.0
6 h	405.44	-7.6
12 h	448.56	+2.2

them has become cluttered with contour lines. The same changes are apparent in the 12-hourly data, but to a smaller degree. In the southern two-thirds of the array for both the 6- and 12-hourly computations, the broad areas of >4 mm along the borders have broken up, and the small regions of >8 mm over Kansas and Oklahoma are very much broadened.

The overwhelming impression from these four plots is that the isohyets have become increasingly cellular as the interval between images increases from 6 to 12 hours. This "cellularization" is easily explained on the basis of the rain calculation scheme. Longer intervals between images mean that 6 or 12 hours worth of rain are "dumped" at the location of the cloud, whereas for the shorter image intervals the isohyets are smoothed by cloud motion between images. The plots of Fig. 5.2 suggest 1- and 3-hourly imagery to be preferable to 6- and 12-hourly imagery for preserving monthly and shorter term rainfall patterns.

6. CONCLUSIONS

Comparisons of the kind discussed in sections 3 and 4 are vital for intelligent application of satellite rainfall methods. Until the day when more direct estimates of rainfall from satellite platforms are a reality, comparisons are the basis for judging whether the differences between satellite and ground estimates of rainfall are prohibitive or are tolerable in the application of interest.

In this study two versions of a satellite technique provided estimates of Great Plains rainfall during one month of the convective season. Satellite estimates, evaluated on the basis of rainfall amount, rainfall timing and rainfall patterns, were compared with gage estimates of rain on monthly, daily, and hourly time frames for area-averaged and point rainfalls. In addition, results from the two versions of the satellite technique (life history and streamlined) were contrasted.

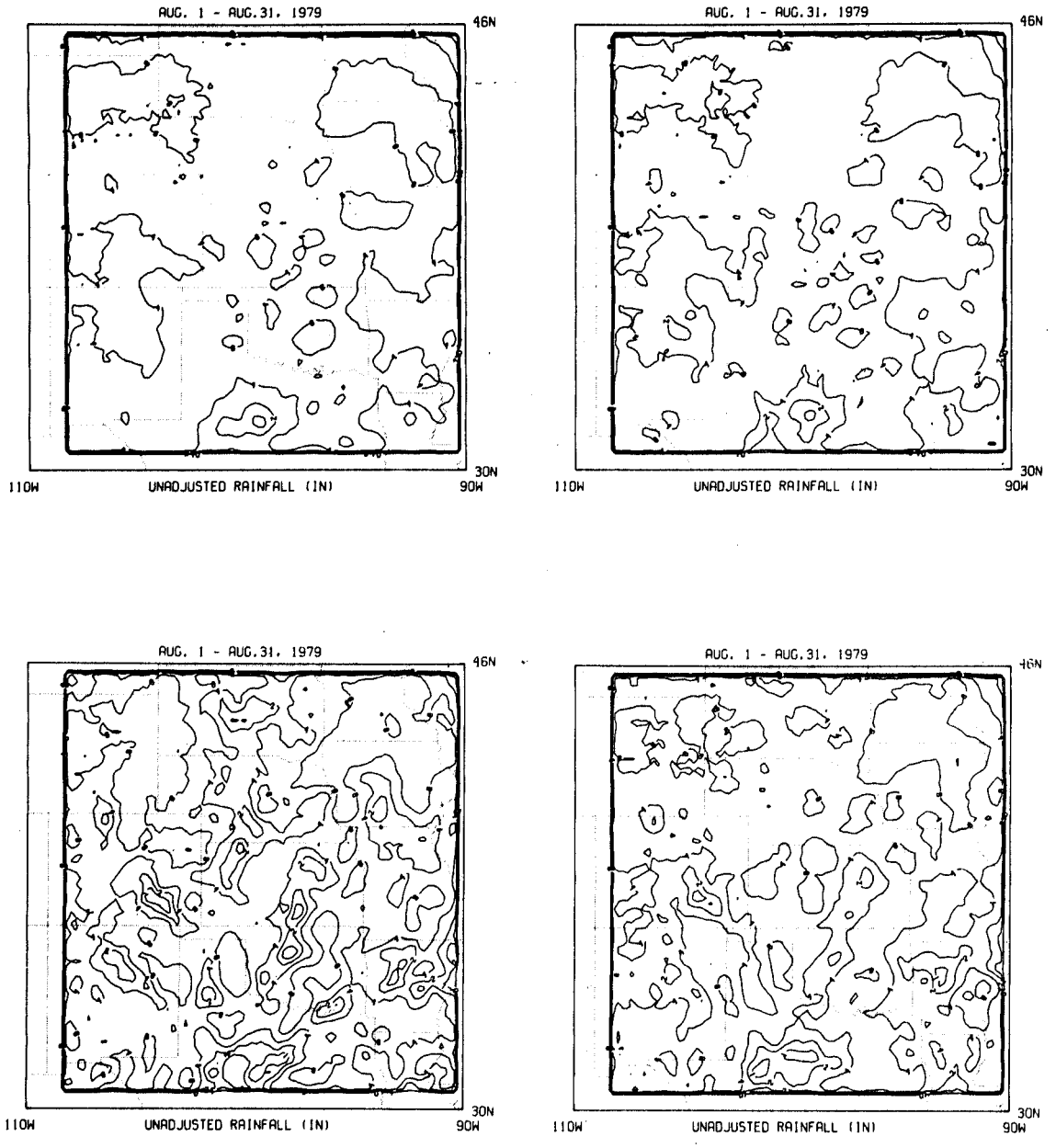


Fig. 5.2. Unadjusted life history satellite rainfall for August 1979. Starting from the upper left and rotating clockwise, the estimates have been calculated from 1-, 3-, 6-, and 12-hourly imagery.

The following sections discuss in detail the findings of this study. The salient results can be summarized as follows.

- Satellite estimates from single pixels (or over degraded grid squares) are area averages and not point estimates of rainfall.

- The smallest, shortest-lived convection is missed by hourly, 500-km^2 resolution GOES IR data, but their loss from the total rainfall may be small because most of the rain occurs from mesoscale convective complexes.

- Correlations between satellite and gage rainfalls increase with increasing space or time scales.

- Timing of relative maxima in satellite and gage time series are comparable when the satellite indicates a corresponding relative maxima, but shorter term fluctuations, evident in the gage data, are lost with hourly imagery.

- Using more than one satellite image to calculate rain for a given period improves the estimate. Better correspondence between satellite and gage estimates occur for daily than for hourly rainfalls; twenty-four images were used in the daily satellite estimate, but only one in the hourly.

- Rainfall patterns are highly dependent upon the temporal frequency of the satellite imagery. As the interval between images becomes longer than 3 hours, the satellite-derived isohyets show less resemblance to isohyetal patterns based on gage data.

- Little difference, by almost any measure, between results from the life history and streamlined techniques suggests that the additional computational complexity of the life history scheme is not necessary.

6.1 Satellite-Gage Comparisons

The feasibility of interpolating environmental correction factors (based on operational weather service radiosonde data and a one-dimensional cumulus cloud model) to an array covering the Great Plains was proved. An identical environmental correction has been previously used (Griffith et al., 1981) with single-station data correcting relatively small areas (dense gage networks operated during field projects) at intervals more frequent than are operationally available. A Gaussian interpolation scheme (Barnes, 1964), with appropriate values for the coefficients, provides a reasonable means of adjusting rainfall inferred from Florida relationships for environmental differences at middle latitude locations. Consequently, model-adjusted satellite rainfalls constitute a satellite estimate in extratropical locations.

The environmentally corrected satellite estimates result in smaller area-

averaged rainfalls, by about 20-40%, than those calculated with the gage data. A similar range with the same bias was found for 15 cases in the pilot study (Griffith et al., 1981). A portion of the shortage found here may be due to the timing of the radiosonde data. The 1200 GMT sounding samples the early morning, preconvective environment. But a surface inversion often exists at this time that could produce less active convection from the model. Although an attempt was made to decrease this effect by lifting the parcel with a layer-averaged temperature that is time dependent (see Appendix A), a better procedure would be to predict a maximum afternoon surface temperature to be used in the model. A second possibility is that the shortage may be a function of how MCCs are handled. Recall that model adjustment factors for stations under the canopy of an MCC are set to 1.0 to avoid contamination from soundings that are not measuring the thermodynamic environment that produced the MCC. The reasoning is that MCCs results in at least as much rain at the surface as convection in Florida. Recall also that the bulk of the rain in region of this study falls from MCCs. It is possible that the 20% rainfall shortage in the satellite estimates is due to model-adjustment factors that are too small near and under MCCs. Further research along these lines is outlined in section 6.3.

The satellite is, in general, able to discriminate between rain and no-rain cases for point events. When compared with the gage assessment of rain versus no-rain cases, the majority of the satellite estimates over both daily and hourly periods result in correct classifications. The fraction of incorrect classifications depends on the accumulation period. Daily satellite estimates are about one-third more likely to indicate rain when the gage shows none than vice versa. Hourly satellite estimates are equally likely for either incorrect classification category. Based on calculations for one month, daily satellite estimates have a 50% probability of being within ±5 mm of the gage estimates, and a 90% probability of being within ±16 mm; for hourly estimates these figures are ±2 mm and ±8 mm, respectively.

Area-averaged satellite rainfalls are correlated with area-averaged gage rainfalls, but "point" satellite amounts show no correlation with point gage amounts. The highest correlations are for daily area-averaged rainfalls (0.7), followed closely by correlations of 0.6 for hourly area-averaged rainfalls. Correlations for point amounts fall to 0.2 for daily periods and to 0.0 at hourly periods. In a Florida study, Meitin et al. (1981) also found low correlations (0.4 for 6-hourly and 0.1 for hourly point rainfalls) when using full-resolution (8 km x 8 km) GOES IR data.

A number of single value statistical parameters were calculated for comparison. Biases as previously noted range from 20% to 40%, the satellite bias being smaller than the gage bias. RMS errors are on the order of 1.0 mm and 0.1 mm for daily and hourly area-averaged estimates, respectively; for point estimates, RMS errors are 14 mm and 5 mm at daily and hourly periods, respectively. Smaller errors in daily point rainfalls are found when additional data (conventional meteorological data and human judgment) are also used (Scofield, 1987).

Box-and-whisker plots of rainfalls for any time or space permutation indicate a larger H-spread (interquartile range) in the gage sample than in the adjusted satellite samples. A portion of this greater range in the gage data is due to the abysmal, 0.1-inch resolution of the majority of the gages,

(as opposed to 0.01-inch resolution). This resolution also affects the timing of the rainfalls in the sense that if a rain event occurs over more than one hour, but less than 0.1 inch falls in a particular hour, that rainfall could be included in the rain amount for the succeeding hour.

The majority of satellite area-averaged rainfall maxima are coincident with or differ at most by one time period (i.e., 1 hour or 1 day) from corresponding maxima in the gage time series. For about one-third of the gage maxima in both the daily and hourly records, there are no corresponding satellite maxima.

Patterns of rainfall contoured from gage and from satellite data agree in some large-scale features, but do not agree in very many smaller scale details. The east-west increase in rainfall from the Rocky Mountains to the Mississippi River is captured, but absolute rain amounts are too small in the environmentally adjusted satellite data.

Rainfall patterns are also highly influenced by the temporal frequency of the satellite data. There is little difference in the patterns when satellite computations are based on 1-hourly or 3-hourly data, but 6-hourly and 12-hourly data result in vastly different rainfall patterns. Rain volumes for each of these four permutations, however, differ by less than 10%.

Overall, the best satellite and gage comparisons occur with the longest and largest time and space scales (daily area-averaged rainfalls), and the comparisons were worst at the shortest and smallest time and space scales (hourly point estimates). This suggests several things: (1) More frequent satellite samples are preferable to less frequent; that is, 24 samples for the daily estimates are better than the single satellite sample that went into an hourly estimate. (2) The degraded satellite resolution of 500 km^2 is too crude to adequately represent rain over the area sampled by a single gage (~1-3 km^2). Even the full-resolution GOES IR data at their present resolution of ~100 km^2 are still too crude. Therefore, satellite comparisons with single gage values are tricky at best. Area-averaged rainfalls are more reliable, simply because the highest satellite resolution represents an area average compared with a gage sample. (3) The phenomenon itself (i.e., convection) would mandate more frequent images than one per hour and better spatial resolution than 500 km^2, for cloud changes on the order of 1 km^2 in 15 minutes to be captured. (4) Although the one-dimensional cumulus model correction vastly improves the use of the empirical satellite rainfall relationships in the middle latitudes, fine tuning to better approximate the gage-derived rainfalls is needed.

6.2 Satellite Technique Comparisons

The finding that there is little difference, by almost any measure, between the life history technique and the streamlined technique suggests that the additional computational complexity of the life history scheme is not worth the effort when area-averaged or large-sample point rainfalls are required; the cloud area term alone explains the bulk of the variance in the gage-derived rainfalls. These results are consistent with those of a number of researchers who by using radar or satellite data have determined that adequate rainfall estimates can be made from areal information alone. Doneaud et al. (1981) have shown this to be so for radar data collected in the

Midwest, while López et al. (1983) recently proved a similar idea with radar data from south Florida. Stout et al. (1979) developed a linear model for satellite rain estimation that contains two terms--an area term and a time-rate-of-change-of-area term. They found that on the average the time change term is half the area term. Negri et al. (1984) calculated from one case study that the area term and the temperature-weighting term of the life history technique used here showed the highest correlations with the calculated rain volume (0.9 and 0.7, respectively) and that a low correlation (0.1) occurs for the rain rate term. The findings of Negri et al. and the results of this paper lend support to recent efforts to apply Doneaud's area-time integral (ATI) technique to satellite data (Doneaud et al. 1984; 1986).

6.3 Future Research

Many problems suggest themselves as profitable for future study in the field of satellite rain estimation. Additionally there are several analyses that are logical extensions of, or better approaches to, the work described here.

The time and space scales included here are extremes in the application of a satellite technique to the operational diagnosis of convective rainfall. A more complete study along the lines of Meitin et al. (1981) would be preferable. In that paper satellite rain estimates made at four time (1/2, 1-1/2, 3, and 6 h) and four space (55, 220, 2220 and 9350 km^2) scales were assessed. In the context of this study, it would be worthwhile to examine at least one additional space scale. Grid squares on the order of 10^5 km^2 would be of interest. This is in the range of the 2.5° x 2.5° grid square used for the climate diagnostics produced at the National Oceanic and Atmospheric Administration's Climate Analysis Center (Kousky, 1986). It is also the upper limit for dense raingage networks that are operated within field experiments such as FACE (Florida Area Cumulus Experiment) and OK PRE-STORM (Oklahoma-Kansas Preliminary Regional Experiment for STORM Central).

No assessment of the gage rainfalls was made. Work from the Soviet Union (WMO, 1972) for summertime convection suggests a large error in the assessment of daily area-averaged rainfall for a gage network of the density used here; the standard deviation of the error as a percentage of the mean total is on the order of 70% of the true area average. A rather simple method of determining the error present in gage-derived, area-averaged rainfalls was discussed by Horton (1923). By knowing mean rainfall and rainfall range from a long record and the number of gages within the area, both the average and maximum departure from the arithmetic mean can be computed. Horton's method is based on an extrapolation of results for the Derwent River Basin in England, and it does not require a dense gage network to assess these errors. More rigorous assessments have been outlined where dense gage data are recorded (Huff, 1970; Woodley et al., 1975; Schaake, 1979). Silverman et al. (1981) modeled the raincell structure of High Plains storms to define gaging requirements for detecting a 25% change in storm mean rainfall in order to evaluate seeding experiments. Their approach could be profitably modified to assess gaging requirements for the determination of areal mean rainfall as well. Augustine et al (1981b) used Florida data to determine the accuracy of satellite rain estimates in terms of an equivalent raingage density. This type of analysis would be extremely useful over the U.S. Great Plains, but would require a denser gage network than the HPD gages. The gage data from

the OK PRE-STORM mesonetwork (Cunning, 1986) are an obvious choice for assessing both HPD gage and satellite rainfall accuracies.

In addition to high density gage data, high-density/high-frequency soundings were taken during OK PRE-STORM. These data can be used to specify a better value for the model adjustment factors in the vicinity of MCCs. Soundings must be carefully selected, following the storm inflow/outflow model developed by Maddox et al. (1986).

There are a handful of days on which extreme differences occur between the area-averaged satellite and gage rainfalls. An investigation into the causes of these differences is necessary for understanding the performance of the satellite rain estimation technique in middle latitudes.

The comparison of the life history and streamlined techniques has not addressed the question of technique performance when rainfall from individual clouds or cloud systems is of interest. My suspicion is that the life history technique is superior to the streamlined technique in the estimation of rainfall from individual clouds or systems because it takes into account the evolution of the cloud. This question has obvious implications for the evaluation of extra-area seeding effects for weather modification projects.

Because machine memory is no longer a limitation, satellite estimates from full spatial resolution imagery can be computed. Comparisons with the results of this study can determine the loss of accuracy in satellite rainfalls that is associated with imagery of degraded spatial resolution, similar to the qualitative assessment of loss of pattern information related to variations in temporal resolution.

Quantitative comparisons of rainfall patterns were not made, but three methods would appear to be promising--orthogonal eigenfunctions, structure function analysis, and MRPP. Orthogonal eigenfunctions are an objective and compact means of quantitatively diagnosing rainfall patterns. They have been used productively to study diurnal (Balling, 1985), and three- and seven-day (Richman and Lamb, 1985) variations in gage rainfalls, but they have not been applied to comparisons of satellite- and gage-derived rainfall fields. Hillger and VonderHaar (1979) and Hillger et al. (1986) used structure function analysis to compare satellite infrared sounding data with conventional sounding temperature and moisture data, and to compute the noise level in satellite sounding data. The structure function has the advantage of computational simplicity, being the mean squared difference between measurements as a function of the vector separation between measurements. Recently Tucker (1986) has used MRPP (in the block formulation of Mielke and Iyer [1982]) to assess model-forecast fields against the observed analysis. A comparison of satellite- and gage-derived rainfall fields would be conceptually similar.

Two convective forecasting techniques could offer much if included in satellite rain estimation schemes; these are the convective equilibrium level and a lapse-rate/short-wave-advection technique. The convective equilibrium level has been used to identify the occurrence of severe weather (Meitin and Griffith, 1986; Griffith and Meitin, 1986). It may prove to be useful for the identification of heavy convective rainfall as well. Recently Caracena (Caracena et al., 1983; Doswell et al., 1985; Caracena and Flueck, 1986)

developed a scheme for forecasting downbursts in the lee of the Rocky Mountains. The components of the scheme are 700- to 500-mb lapse rate, 850-mb moisture, and the presence or absence of a short wave aloft. Caracena and Flueck found high correlations between the occurrence of microbursts and these parameters; the parameters are certainly related to rainfall also, and may present a simple scheme in conjunction with satellite data for the estimation of convective rainfall.

ACKNOWLEDGMENTS

Messers John A. Augustine and José G. Meitín, and Dr. William L. Woodley were involved in the preliminary studies and early stages of this work; acknowledgment is made to them for numerous discussions on many aspects of this work, and to the former two for computer code and for the initial computer processing of portions of the GOES data. Helpful discussions on the Gaussian interpolation were provided by Drs. Stanley L. Barnes, Charles A. Doswell, III, and Robert A. Maddox. Mr. Irv Watson pointed out a quirk in the processed raingage data set. Various parts of the satellite, gage, and upper-air data sets were prepared by Mr. Vaughn Martin, Mr. Myron Nealey, Ms. Jane Skradski, and Ms. Patricia Tait. Mr. Ben Baron did much of the coding for the interpolated data sets, and oversaw many of the production runs.

Part of this work was funded through NOAA AgRISTARS contracts, which were very ably and critically monitored by Dr. Rod Scofield. GOES data were provided by Dr. Thomas H. VonderHaar from the CSU Direct Readout Ground Station. Acknowledgment is made to the National Center for Atmospheric Research, which is sponsored by the National Science Foundation, for a portion of the computing time used in this research.

REFERENCES

Albright M.D., E.E. Recker, R.J. Reed, and R. Dang, 1985: The diurnal variation of deep convection and precipitation in the central tropical Pacific during January-February 1979. MON. WEA. REV., 113, 1663-1680.

Arkin, P.A., 1979: The relationship between fractional coverage of high cloud and rainfall accumulations during GATE over the B-scale array. MON. WEA. REV., 107, 1382-1387.

Auer, A.H., Jr., and J.D. Marwitz, 1968: Estimates of air and moisture flux into hailstorms on the High Plains. J. APPL. METEOR., 7, 196-198.

Augustine, J.A., 1984: The diurnal variation of large-scale inferred rainfall over the tropical Pacific Ocean during August 1979. MON. WEA. REV., 112, 1745-1751.

Augustine, J.A., C.G. Griffith, W.L. Woodley, and J.G. Meitín, 1981a: Insights into errors of SMS-inferred GATE convective rainfall. J. APPL. METEOR., 20, 509-520.

Augustine, J.A., J.G. Meitín, C.G. Griffith, and W.L. Woodley, 1981b: An objective evaluation of the Griffith/Woodley rain estimation technique. Preprints, 4th Conf. on Hydrometeorology, Amer. Meteor. Soc., Boston, Mass., 134-140.

Balling, R.C., Jr., 1985: Warm season nocturnal precipitation in the Great Plains of the United States. J. CLIM. APPL. METEOR., 24, 1383-1387.

Barnes, S.L., 1964: A technique for maximizing details in numerical weather map analysis. J. APPL. METEOR., 3, 396-409.

Barnes, S.L., 1973: Mesoscale objective map analysis using weighted time series observations. NOAA Tech. Memo. ERL NSSL-62, 60 pp. [NTIS COM-73-10781]

Barrett, E.C., and D.W. Martin, 1981: The Use of Satellite Data in Rainfall Monitoring. Academic Press, London, 340 pp.

Braham, R.R., Jr., 1952: The water and energy budgets of the thunderstorm and their relation to thunderstorm development. J. METEOR., 9, 227-242.

Bras, R.L., and I. Rodriguez-Iturbe, 1976: Network design for the estimation of areal mean of rainfall events. WATER RESOUR. RES., 12, 1185-1195.

Caracena, F.C., and J.A. Flueck, 1986: The classification and prediction of small-scale windshear events in a dry environment. Preprints, 33rd Annual Meeting Aerospace: Century XXI, American Astronautical Society, San Diego, Calif.

Caracena, F.C., J. McCarthy, and J.A. Flueck, 1983: Forecasting the likelihood of microbursts along the Front Range of Colorado. Preprints, 13th Conf. on Severe Local Storms, Amer. Meteor. Soc., Boston, Mass., 261-264.

Chisholm, A.J., 1970: Alberta hailstorms: a radar study and model. Ph.D. Thesis, Dept. of Meteorology, McGill University, Montreal, Canada, 237 pp.

Clark, D., and M. Perkins, 1985: The NOAA Interactive Flash Flood Analyzer. Preprints, International Conf. on Interactive Information and Processing Systems for Meteorology, Oceanography and Hydrology, Amer. Meteor. Soc., Boston, Mass., 255-259.

Cressman, G.P., 1959: An operational analysis system. MON. WEA. REV., 87, 367-374.

Cunning, J.B., 1986: The Oklahoma-Kansas Preliminary Regional Experiment for STORM-Central. BULL. AMER. METEOR. SOC., 12, 1478-1486.

DelBeato, R., and S.L. Barrell, 1985: Rain estimation in extratropical cyclones using GMS imagery. MON. WEA. REV., 113, 747-755.

Doneaud, A.A., P.L. Smith, A.S. Dennis, and S. Sengupta, 1981: A simple method for estimating the convective rain volume over an area. WATER RESOUR. RES., 17, 1676-1682.

Doneaud, A.A., J.R. Miller, L.R. Johnson, T.H. VonderHaar, and P. Laybe, 1984: An attempt to extend the ATI technique to estimate convective rain volumes using satellite data. Preprints, 22nd Conf. on Radar Meteorology, Amer. Meteor. Soc., Boston, Mass., 176-181.

Doneaud, A.A., J.R. Miller, Jr., and A. Makarau, 1986: A modified ATI technique for possible satellite applications. Preprints, Joint Sessions 23rd Conf. on Radar Meteorology and Conf. on Cloud Physics, Amer. Meteor. Soc., Boston, Mass., JP27-JP32.

Doswell, C.A., III, F. Caracena, and M. Magnano, 1985: Temporal evolution of 700-500 mb lapse rate as a forecasting tool--a case study. Preprints, 14th Conf. on Severe Local Storms, Amer. Meteor. Soc., Boston, Mass., 398-401.

Fankhauser, J.C., 1971: Thunderstorm-environment interactions determined from aircraft and radar observations. MON. WEA. REV., 99, 171-192.

Fenner, J.H., 1982: Using satellite imagery to estimate winter season rainfall in Europe. Preprints, 9th Conf. on Weather Forecasting and Analysis, Amer. Meteor. Soc., Boston, Mass., 245-249.

Foote, G.B., and J.C. Fankhauser, 1973: Airflow and moisture budget beneath a northeast Colorado hailstorm. J. APPL. METEOR., 12, 1330-1353.

Fujita, T., 1959: Precipitation and cold air production in mesoscale thunderstorm systems. J. APPL. METEOR., 16, 454-466.

Gagin, A., D. Rosenfeld, and R.E. López, 1985: The relationship between height and precipitation characteristics of summertime convective cells in south Florida. J. ATMOS. SCI., 42, 84-94.

Garcia, O., 1981: A comparison of two satellite rainfall estimates for GATE. J. APPL. METEOR., 20, 430-438.

Garcia, O., 1985: An atlas of highly reflective clouds for the global tropics, 1971-1983. U.S. Dept. Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, In press. [Available from Superintendent of Documents, Washington, D.C.]

Gray, W.M., and R.W. Jacobson, Jr., 1977: Diurnal variation of deep cumulus convection. MON. WEA. REV., 105, 1171-1188.

Griffith, C.G., and B.J. Meitín, 1986: A test of the convective, equilibrium-level temperature enhancement as a forecasting tool. Preprints, 11th Conf. on Weather Forecasting and Analysis, Amer. Meteor. Soc., Boston, Mass., 288-293.

Griffith, C.G., W.L. Woodley, P.G. Grube, D.W. Martin, J. Stout, and D.N. Sikdar, 1978: Rain estimation from geosynchronous satellite imagery--visible and infrared studies. MON. WEA. REV., 106, 1153-1171.

Griffith, C.G., W.L. Woodley, J.S. Griffin, and S.C. Stromatt, 1980: Satellite-derived precipitation atlas for GATE. U.S. Dept. Commerce, National Oceanic and Atmospheric Administration, Environmental Research Laboratories, 280 pp. [Available from Superintendent of Documents, Washington, D.C.]

Griffith, C.G., W.L. Woodley, and J.A. Augustine, 1981: Satellite rain estimation in the U.S. High Plains. J. APPL. METEOR., 20, 53-66.

Hartsell, C., 1970: Case study of a traveling hailstorm. Report 70-1, Institute of Atmos. Sci., South Dakota School of Mines and Tech., Rapid City, South Dakota, 72 pp.

Hatch, D.J., 1976: Estimation of local rainfall by sample gauging. WEATHER, 31, 79-83.

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

Hillger, D.W., J.F.W. Purdom, and T.H. VonderHaar, 1986: The use of structure function analysis to estimate noise levels and the potential resolution of satellite measurements. Preprints, 2nd Conf. on Satellite Meteorology/Remote Sensing and Applications, Amer. Meteor. Soc., Boston, Mass., 102-107.

Hoaglin, D.G., F. Mosteller, and J.W. Tukey, 1983: Understanding Robust and Exploratory Data Analysis. J. Wiley and Sons, Inc., New York, New York.

Holle, R.L., and S.A. MacKay, 1975: Tropical cloudiness from all-sky cameras on Barbados and adjacent Atlantic Ocean. J. APPL. METEOR., 14, 1437-1450.

Horton, R.E., 1923: Accuracy of areal rainfall estimates. MON. WEA. REV., 51, 348-353.

Hudlow, M.D., and V.L. Patterson, 1979: GATE radar rainfall atlas. NOAA Special Report, U.S. Dept. Commerce, 155 pp. [Available from Superintendent of Documents, Washington, D.C., Stock No. 003-019-00046-2.]

Huff, F.A., 1970: Sampling errors in measurement of mean precipitation. J. APPL. METEOR., 9, 35-44.

Jordan, C.L., 1958: Mean soundings for the West Indies area. J. METEOR., 15, 91-97.

Kilonsky, B.J., and C.S. Ramage, 1976: A technique for estimating tropical open-ocean rainfall from satellite observations. J. APPL. METEOR., 15, 972-975.

Kleiner, B., and T.E. Graedel, 1980: Exploratory data analysis in the geophysical sciences. REV. GEOPHYS. & SPACE PHYS., 18, 699-717.

Kousky, V.E., ed., 1986: Climate Diagnostics Bulletin. NOAA National Weather Service, National Meteorological Center, Climate Analysis Center, Washington, D.C. Published monthly.

López, R.E., J. Thomas, D.O. Blanchard, and R.L. Holle, 1983: Estimation of rainfall over an extended region using only measurements of the area covered by radar echoes. Preprints, 21st Conf. on Radar Meteorology, Amer. Meteor. Soc., Boston, Mass., 681-686.

Lovejoy, S., and G. L. Austin, 1979: The sources of error in rain amount estimating schemes from GOES visible and IR satellite data, MON. WEA. REV., 107, 1048-1054.

Maddox, R.A., 1980a: Mesoscale convective complexes. BULL. AMER. METEOR. SOC., 61, 1374-1387.

Maddox, R.A., 1980b: An objective technique for separating macroscale and mesoscale features in meteorological data. MON. WEA. REV., 108, 1108-1121.

Maddox, R.A., K.W. Howard, D.L. Bartels, and D.M. Rodgers, 1986: Mesoscale Convective Complexes in the Middle Latitudes. In Mesoscale Meteorology and Forecasting, Peter S. Ray, Ed., Amer. Meteor. Soc., Boston, p 399.

Martin, D.W., and M.R. Howland, 1986: Grid history: a geostationary satellite technique for estimating daily rainfall in the tropics. J. CLIM. APPL. METEOR., 25, 184-195.

Marwitz, J.D., 1972: Precipitation efficiency of thunderstorms on the High Plains. Preprints, 3rd Conf. on Weather Modification, Amer. Meteor.

Soc., Boston, 245-247.

McGarry, M.M., and R.J. Reed, 1978: Diurnal variations in convective activity and precipitation during Phases II and III of GATE. MON. WEA. REV., 106, 101-113.

Meisner, B.N., and P.A. Arkin, 1987: Spatial and annual variations in the diurnal cycle of large-scale tropical convective cloudiness and precipitation. In review MON. WEA. REV.

Meitín, B.J., and C.G. Griffith, 1986: An investigation of a GOES infrared enhancement based on the convective equilibrium-level temperature. Preprints, 11th Conf. on Weather Forecasting and Analysis, Amer. Meteor. Soc., Boston, Mass., 282-287.

Meitín, J.G., C.G. Griffith, J.A. Augustine, and W.L. Woodley, 1981: A standard verification for rainfall estimation from remote platforms. Precipitation Measurements from Space, Workshop Report, D. Atlas and O. Thiele, eds., NASA Goddard Space Flight Center, Greenbelt, Md., D-94 - D-97.

Mielke, P.W., Jr., and K. Iyer, 1982: Permutation techniques for analyzing multi-response data from randomized block experiments. COMM. STATIST.-THEOR. METH., 11, 1427-1237.

Mielke, P.W., Jr., K.J. Berry and G.W. Brier, 1981: Application of multi-response permutation procedures for examining seasonal changes in monthly mean sea-level pressure patterns. MON. WEA. REV., 109, 120-126.

Negri, A. J., R. F. Adler and P. J. Wetzel, 1984: Rain estimation from satellites: an examination of the Griffith-Woodley technique. J. CLIM. APPL. METEOR., 23, 102-116.

Neil, L.K., 1984: Estimating precipitation areas over the eastern Pacific using GOES satellite imagery. Preprints, Conf. on Satellite/Remote Sensing and Applications, Amer. Meteor. Soc., Boston, Mass., 121-126.

Newton, C., 1966: Circulations in large sheared cumulonimbus. Tellus, 18, 699-713.

Peck, E.L., 1980: Design of precipitation networks. BULL. AMER. METEOR. SOC., 61, 894-902.

Richman, M.B., and P.J. Lamb, 1985: Climatic pattern analysis of three- and seven-day summer rainfall in the central United States: Some methodological considerations and a regionalization. J. CLIM. APPL. METEOR., 24, 1325-1343.

Schaake, J.C., Jr., 1980: Accuracy of point and mean areal precipitation estimates from point precipitation data. NOAA Tech. Memo. NWS-HYDRO, Washington, D.C., 98 pp.

Scofield, R.A., 1987: The NESDIS operational convective precipitation estimation technique. In review MON. WEA. REV.

Scofield, R.A., and V.J. Oliver, 1977: A scheme for estimating convective rainfall from satellite imagery. NOAA Tech. Memo. NESS-86, Washington, D.C., 47 pp.

Scofield, R.A., and L.E. Spayd, Jr., 1984: A technique that uses satellite, radar and conventional data for analyzing precipitation from extratropical cyclones. NOAA Tech. Memo. NESDIS 8, U.S. Dept. Commerce, Washington, D.C. 20746, 51 pp.

Silverman, B.A., L.K. Rogers, and D. Dahl, 1981: On the sampling variance of raingage networks. J. APPL. METEOR., 20, 1468-1478.

Simpson, J., and V. Wiggert, 1969: Models of precipitating cumulus towers. MON. WEA. REV., 97, 471-489.

Simpson, J., and V. Wiggert, 1971: 1968 Florida cumulus seeding experiment: numerical model results. MON. WEA. REV., 99, 87-118.

Spencer, R.W., 1986: A satellite passive 37-GHz scattering-based method for measuring oceanic rain rates. J. CLIM. APPL. METEOR., 25, 754-766.

Stout, J., D. W. Martin, and D. N. Sikdar, 1979: Estimating GATE rainfall with geosynchronous satellite images. MON. WEA. REV., 107, 585-598.

Thiessen, A.H., 1911: Precipitation averages for large areas. MON. WEA. REV., 30, 1082-1084.

Tucker, D.F., 1987: Interaction of orographically induced winds within a mesoscale forecasting model system. Ph.D. Dissertation, Atmospheric Science Dept., Colorado State University, Ft. Collins, Colorado.

Tukey, J.W., 1977: Exploratory Data Analysis. Addison-Wesley, Reading, Mass, 688 pp.

Velleman, P.F., and D.C. Hoaglin, 1981: A,B,C's of Exploratory Data Analysis, Duxbury Press, Boston, Mass.

Waters, M.P., III, C.G. Griffith, and W.L. Woodley, 1977: Use of digital geostationary satellite imagery for real-time estimation of hurricane rain potential in landfalling storms. Volume of Conference Papers, 11th Technical Conf. on Hurricanes and Tropical Meteorology, Amer. Meteor. Soc., Boston, Mass., 198-203.

Woodley, W.L., and A. Gagin, 1986: Meteorological studies in the Texas High Plains. Final Report, Texas Water Commission, Austin Texas, 71 pp.

Woodley, W.L., A.R. Olsen, A. Herndon, and V. Wiggert, 1975: Comparison of gage and radar methods of convective rain measurement. J. APPL. METEOR., 14, 909-928.

Woodley, W.L., C.G. Griffith, J.S. Griffin, and J. Augustine, 1978: Satellite rain estimation in the Big Thompson and Johnstown flash floods. Preprints, Conf. on Flash Floods: Hydrometeorological Aspects, Amer. Meteor. Soc., Boston, Mass., 44-51.

Woodley, W.L., C.G. Griffith, J.S. Griffin, and S.C. Stromatt, 1980a: The inference of GATE convective rainfall from SMS-1 imagery. J. APPL. METEOR., 19, 388-408.

Woodley, W.L., J.A. Augustine, J.G. Meitín, and C.G. Griffith, 1980b: On the objective adjustment of satellite rain estimates in west Africa. Unpublished ms.

World Meteorological Organization, 1972: Casebook on hydrological network design practice. WMO-No. 324, Geneva, p I-1.1-10.

Wylie, D.P., 1979: An application of a geostationary satellite rain estimation technique to an extratropical area. J. APPL. METEOR., 18, 1640-1648.

APPENDIX A

ENVIRONMENTAL ADJUSTMENT FACTORS FOR THE GREAT PLAINS

Five factors were developed to account for the environmental differences between the subtropics and the Great Plains of the United States: precipitable water, subcloud evaporation, vertical wind shear, a combination of these three, and a cumulus model adjustment factor.

Precipitable Water

On the average, precipitable water over the U.S. Great Plains is about half that in the subtropics of the United States. The precipitable water correction (P), defined as

$$P = \frac{W_H}{W_F},$$

was calculated to account for this difference, where W is precipitable water computed from the surface upward, and H and F refer to the High Plains and Florida, respectively. The W_F value is fixed at 4.28 cm and was computed from the data of Jordan (1958). The value of W_H was computed on a case by case basis from the sounding closest in space and time.

Subcloud Evaporation

A manifestation of the decrease of precipitable water between Florida and the middle latitudes in the United States is the increase of subcloud evaporation at the higher latitudes compared with Florida. The problem of subcloud evaporation was addressed by Fujita (1959), who modified an in-cloud evaporation scheme by Braham (1952) to produce a relationship between the ratio of evaporated rain (R_e) to surface rain (R_s) and cloud base height above ground level for High Plains storms (Fig. A.1). This relationship shows that the ratio R_e/R_s increases with cloud base height in a nonlinear fashion. For cloud bases higher than 2.75 km (9000 ft) above the ground, subcloud evaporation exceeds surface rainfall. In Florida there is little, if any, subcloud evaporation, so that the rain that is inferred from the technique's relationships is the rain that actually reaches the surface. In the middle latitudes, however, the rain (R) estimated for the High Plains (subscript "H") from the Florida (subscript "F") relationships of this technique (R_{HF}) represents the sum of both the rainfall evaporated below cloud base and the rainfall that reaches the surface ($R_{HF} = R_e + R_s$). The rain actually caught by gages in the High Plains (R_H) is the rainfall that reaches the surface ($R_H = R_s$). It is this rainfall that the satellite technique attempts to estimate. The ratio of actual High Plains rainfall to High Plains rainfall inferred from the Florida relationships is

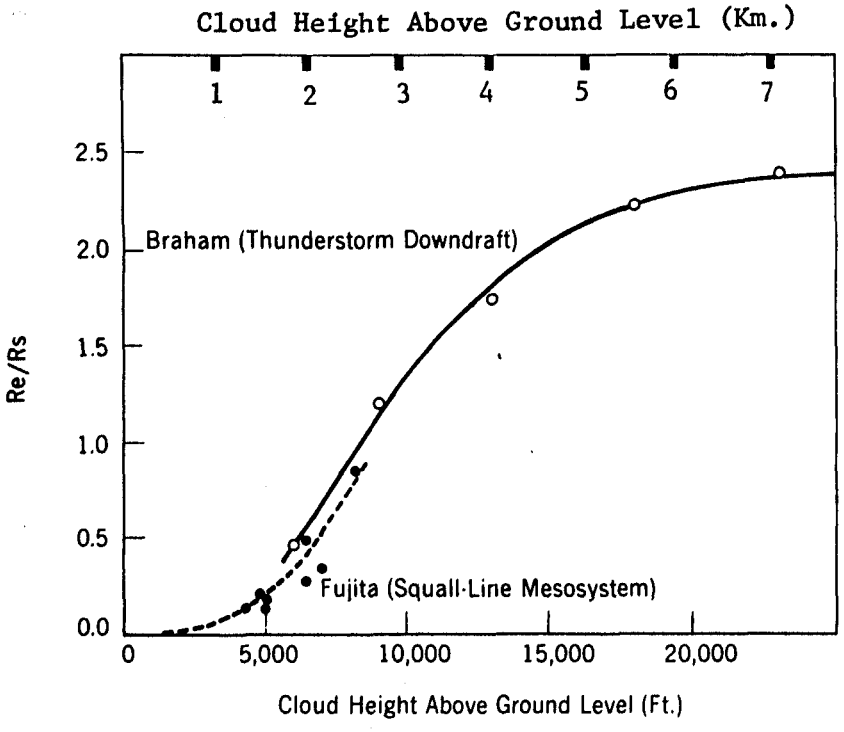


Fig. A.1. In-cloud evaporation computed by Braham from the water budgets of thunderstorm circulations (upper curve) and subcloud evaporation obtained by Fujita with the use of total excess mass and total surface rain accompanied by squall-line mesosystems. (After Fujita, 1959.)

$$\frac{R_H}{R_{HF}} = \frac{R_s}{R_e + R_s} . \qquad \text{(A-1)}$$

To derive an adjustment factor that accounts for subcloud evaporation, the left side of (A-1) is identified as the subcloud evaporation correction factor, E, and the right hand side is rewritten in terms of Fujita's ratio:

$$E = \frac{1}{(R_e/R_s) + 1} . \qquad \text{(A-2)}$$

To use this correction factor, cloud base is first computed from the local sounding. From this cloud base the value of the ratio R_e/R_s is read from Fig. A.1 and then entered into (A-2) to complete the computation of the subcloud evaporation correction factor.

Vertical Wind Shear

The third factor that was tested considers the difference between shear and precipitation efficiency in the two regions. Foote and Fankhauser (1973) and a number of others (Marwitz, 1972; Auer and Marwitz, 1968; Fankhauser, 1971; Chisholm, 1970; Hartsell, 1970; Newton, 1966) have independently studied the precipitation efficiency of selected Great Plains thunderstorms. Their composited results can be found in Fig. A.2. In these studies vertical wind shear was the vector difference in wind through the depth of the cloud, and precipitation efficiency was defined as the ratio of moisture influx at cloud base to the measured rain rate at the ground. From Fig. A.2 it can be seen that efficiency decreases nonlinearly as wind shear increases.

In applying these results to the satellite data, a cloud depth was determined from the temperatures of the infrared imagery and a local sounding. Cloud top was inferred from the coldest infrared temperature over the site during the period of the rain calculation. This temperature was

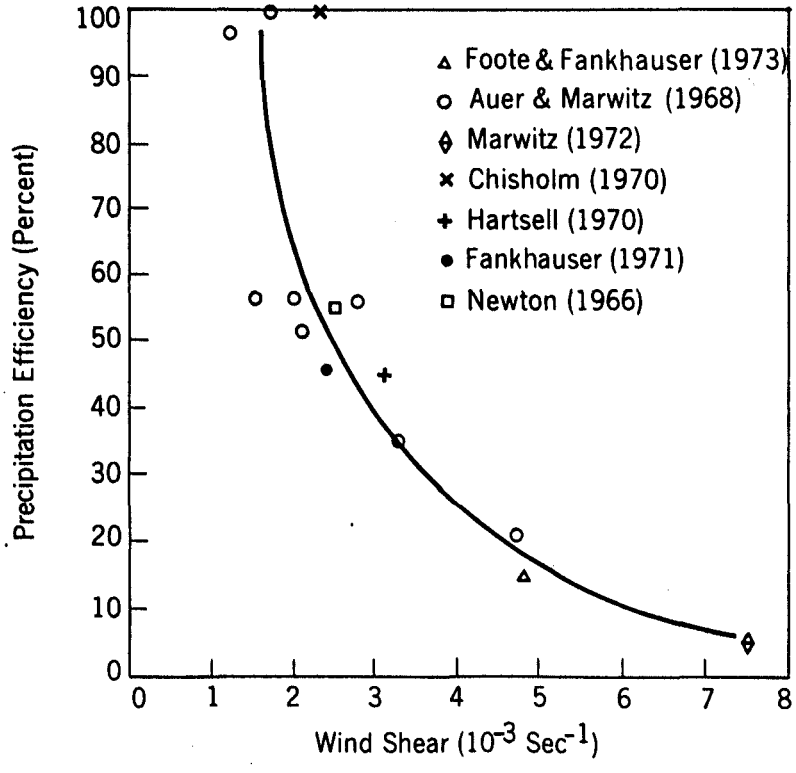


Fig. A.2. Scatter diagram of vertical wind shear versus precipitation efficiency for fourteen thunderstorms on the High Plains of North America. (From Foote and Fankhauser, 1973.)

converted to height with the local sounding. Cloud base was defined to be the mean convective condensation level (MCCL) determined from the sounding. The mean mixing layer of the MCCL was 5 kPa for the period after solar heating was established (1200 LST through 0100 LST on the following day), and 10 kPa at any other time. Cloud depth was the difference between the satellite cloud top and the MCCL cloud base. The vertical shear (V_s) of the sounding winds through this depth was used to determine the precipitation efficiency from Fig. A.2.

Combined Adjustment Factor

The combined adjustment factor (C) is the product of the precipitable water, subcloud evaporation, and vertical wind shear factors, $C = P \times E \times V_s$. It should be noted that C is not composed of independent terms. Precipitable water and subcloud evaporation, for instance, should be highly correlated.

Cumulus Model Adjustment

The first four correction factors can all be rather simply and quickly computed from upper-air data and satellite imagery. The final adjustment factor that was tested is more complicated, being based on the output of the one-dimensional cumulus cloud model of Simpson and Wiggert (1969; 1971) and is conceptually similar to an environmental correction described by Wylie (1979). The Simpson-Wiggert model simulates the active rising portion of an individual convective tower from the temperature and moisture information of the local sounding. The model considers static stability only and does not incorporate dynamics. Therefore forced convection, due to orography or generated by air mass movement, is not treated.

The Simpson-Wiggert model requires as input (1) a vertical sounding (pressure, temperature and a moisture parameter), minimally for every 5-kPa interval from the surface up to the 10-kPa pressure level; (2) a thermal bubble radius; and (3) cloud base. Eight fixed bubble radii, ranging from 500 to 3000 m, are used because there is currently no way of predicting cloud dimensions for a given day. The model has two options regarding cloud base: cloud base can be input as a fixed height above the surface or can be computed from a computer-calculated MCCL. The latter option was used, so that a mean mixing layer was input along with the sounding. As in the MCCL calculation for the vertical shear correction factor, a 10-kPa mixing layer is specified in the period prior to solar heating (0100 through 1200 LST) and a 5-kPa layer is specified once solar heating has been established (the hours between 1200 and 0100 LST). The model itself contains no constraints to ensure that the calculated cloud base represents a reasonable surface temperature.

A sample output of the cumulus model is given in Table A.1. The model adjustment factor is based on the precipitation fallout of column two. Simpson and Wiggert found that the prediction of the precipitation production is proportional to, but less than, the observed rainfall. For this reason and the fact that tower radius is unknown, the model adjustment factor (MAF) is not based on model precipitation for any particular thermal bubble, but is defined as

$$MAF = \frac{\bar{R}_H \times \sigma_H}{\bar{R}_F \times \sigma_F}, \qquad \text{(A-3)}$$

where $\bar{R}$ is the mean rain production of the eight radii, σ is the standard deviation of the rain production from the eight radii, and H and F refer to the High Plains and Florida, respectively. The standard deviations in (A-3) are a means of accounting for the range of the model rain production of the eight thermal radii. The values of $\bar{R}_F$ and σ_F are computed from a typical Florida sounding and are 10.210 and 4.195, respectively. If the atmosphere at the site of interest is similar to the mean Florida environment, the value of the model adjustment factor will be close to 1. For a region that is much drier than Florida, the MAF will be smaller than 1 and the satellite rainfall will be decreased. The converse is true for locations wetter than Florida.

Evaluation of Adjustment Factors

Each adjustment factor was tested on 15 cases at 3 locations in the High Plains (Griffith et al. 1981). At each site a dense gage network was used to evaluate the adjustment factors. The number of gages ranged from 40 to 90, and the networks covered relatively small areas ranging from 2,000 to 12,000 km^2. Gage densities ranged from one gage per 20 km^2 to one gage per 150 km^2. The periods of calculation of the satellite rainfalls for each case varied from 4.0 to 16.5 hours. Because of these differences, the satellite and gage amounts that were used in the comparison were accumulated over the period of calculation and then normalized by the period, as well as by the area, of the calculation resulting in rainrate (mm/h).

Four measures (defined as follows) were devised to evaluate the effect of the environmental adjustments, along with the three linear-fit parameters: correlation (ρ), slope, and intercept.

Table A.1. Sample output from the Simpson-Wiggert one-dimensional cumulus cloud model (without the seeding simulation). Data are for 8/19/79 (left block, first line), 0000 Greenwich Central Time (middle block) at station 24023, North Platte, Nebraska (abbreviated as 23 at right). MCCL cloud base is 994.8 m. Cloud top height (m) and precipitation fallout (g/kg) as a function of eight bubble radii are listed.

81979 0GCT 23

ASSUMED CLOUD BASE AT 994.8 METERS

RADIUS (METERS)	A UNSEEDED WATER TO -40C
500.0 M.	4294.8 M. (2.061 G/KG)
750.0 M.	4894.8 M. (3.659 G/KG)
1000.0 M.	5544.8 M. (5.189 G/KG)
1250.0 M.	8144.8 M. (7.853 G/KG
1500.0 M.	8844.8 M. (8.719 G/KG)
2000.0 M.	9944.8 M. (9.735 G/KG)
2500.0 M.	10944.8 M. (10.513 G/KG)
3000.0 M.	14244.8 M. (11.673 G/KG)

Mean event ratio: $$\overline{S/G} = \frac{\Sigma S}{\Sigma G} \quad \text{(A-4)}$$

Mean hourly ratio: $$B = \frac{\Sigma(S/G)}{N} \quad \text{(A-5)}$$

Factor of difference: $$E_R = \frac{\Sigma R}{N} \quad \text{(A-6)}$$

Normalized root-mean-

square error:

$$E_{RMS} = \left[\frac{\sum \left(\frac{S-G}{G} \right)^2}{N} \right]^{1/2} \qquad \text{(A-7)}$$

In the above definitions N is the number of cases (15) and R in (A-6) is the ratio formed from S and G such that $R \geq 1.00$. The definition of the root-mean-square error (E_{RMS}) differs from the usual by the factor of G in the denominator of the numerator. This factor scales the rain differences according to the amount of rain on the ground, because a difference of 1 mm is more significant in a total rainfall of 2 mm than in a total rainfall of 10 mm.

The unadjusted satellite rainfalls result in some of the largest values for these measures in Table A.2. With the exception of the factor of difference (E_R) for the combined factor (C) and the normalized root-mean-square error (E_{RMS}) for the vertical wind shear factor (V_S), each adjustment decreased the value of these performance measures. For the measures B, E_R and E_{RMS}, the model adjustment factor (MAF) resulted in values that are closest to those for perfect correspondence (PC). For the correlation coefficient (ρ), the slope, and intercept of the linear fit, and the mean event ratio ($\overline{S}/\overline{G}$), the model adjustment factor shows values among those that are closest to perfect correspondence. On the basis of these tests the model adjustment factor was chosen to adjust the satellite rain estimates for environmental differences between the site of interest and south Florida, the region of derivation of the technique's relationship.

Table A.2. Comparison of area-averaged and time-averaged gage (G) and satellite (S) rainfalls for five adjustment schemes.

Adjustment	ρ	Slope	Inter-cept	$\overline{S}/\overline{G}$	B	E_R	E_{RMS}
None	0.94	3.14	0.23	3.52	5.87	9.68	12.47
MAF	0.90	0.62	0.07	0.74	1.30	5.22	3.03
P	0.94	1.73	0.13	1.94	4.37	9.60	9.64
E	0.90	2.40	0.06	2.50	4.20	9.90	7.94
V_S	0.89	2.60	-0.03	2.56	4.93	9.27	13.32
C	0.82	1.22	-0.04	1.16	2.43	11.08	5.47
PC	1.00	1.00	0.00	1.00	1.00	1.00	0.00

APPENDIX B

MESOSCALE CONVECTIVE COMPLEXES AND THEIR IMPLICATIONS FOR SATELLITE RAIN ESTIMATION

The existence of mesoscale convective complexes was recently documented by Maddox (1980a). MCC's are large, long-lived convective systems that are circular in shape, in contrast to the linear squall line. Maddox describes them as "organized, meso-α scale convective weather systems over the central U.S." The characteristics of the systems are defined by their structure when viewed in thermal infrared imagery that has been enhanced by the NESDIS "MB" curve.

The size, shape, and duration criteria used by Maddox to define MCC's are detailed in Table B.1. The size criteria imply phenomena that are more than two orders of magnitude larger than individual thunderstorms and that are on the order of the size of tropical cyclones, or greater. The size and duration criteria also ensure that MCC's will be sampled by one or more upper-air

Table B.1. Definition of mesoscale convective complexes* (MCC) based on analyses of enhanced IR satellite imagery

Size†	A: Cloud shield with contiguously low IR temperature $\leq$ -32°C must have an area 100,000 km^2.
	B: Interior cold cloud region with temperatures $\leq$ -52°C must have an area $\geq$50,000 km^2.
Duration	Size definitions A and B must be met for a period of $\geq$6h.
Maximum extent	Contiguous cold-cloud shield (IR temperature $\leq$-32°C) reaches maximum size
Shape:	Eccentricity (minor axis/major axis) must be $\geq$0.7 at the time of maximum extent.

*After Maddox (1980a).
†Initiation occurs when size definitions A and B are first satisfied. Termination occurs when size definitions A and B are no longer satisfied.

soundings during their lifetimes. The shape criterion requires that these systems be circular or nearly so, thus implying little vertical wind shear aloft. A direct consequence of the temperature threshold used by Maddox is that rain will always be estimated for these systems by the satellite rainfall relationships described in this study.

The dynamics of the MCC drastically change the local environment. At middle levels in the atmosphere, potentially cool environmental air is entrained. This produces strong downdrafts, causing cold air outflows at the surface, manifested as gust fronts and individual outflows. Thus, at middle levels and at the surface the atmosphere is significantly different from the moist, unstable low-level air that fuels the complex. Any sounding sampling air modified by the dynamics of the MCC would exhibit little to no convective activity when processed through the one-dimensional model. This would result in a small value for the model adjustment factor (typically ~0.01), so little rain would be inferred by the satellite technique for the complex after adjustment. However, much precipitation is produced by an MCC, particularly in its mature stage. Without modeling a composite MCC, a more realistic estimate for the MAF under an MCC is a value of 1.00. Certainly the amount of rain produced by an MCC is closer to that produced by a Florida system of comparable size, rather than one one-hundredth of the rain from a Florida system.

The 6-hourly, MB-enhanced GOES hard-copy imagery for August 1979 were reviewed to identify MCC's within the region of this study. Occasionally the unenhanced, full-disk IR image was examined when the MB-enhanced IR picture was not available. The full-disk data were used mainly for continuity. Table B.2 lists the upper-air stations that were under MCC canopies at the synoptic times during the month of this study.

The model adjustment factor for the 55 stations in Table B.2 were set to 1.00. Undoubtedly, there were also convective systems present in the August data set that did not meet the MCC criteria yet modified the atmosphere to such an extent that soundings and model adjustment factors are contaminated by their effects. No effort, however, was made to identify all the soundings that may have been in the outflow of large convective systems or that had sampled in-cloud air of such systems.

Table B.2. Upper air data contaminated by mesoscale convective complexes, August 1979

Date	Time	Station	Date	Time	Station
8/02	12Z	Longview, TX	8/21	00Z	Jackson, MS
		Stephenville, TX			Midland, TX
8/04	12Z	St. Cloud, MN			Nashville, TN
8/05	12Z	Green Bay, WI			Grand Junction, CO
8/08	12Z	Green Bay, WI			Peoria, IL
		Huron, SD		12Z	Stephenville, TX
		St. Cloud, MN	8/22	12Z	St. Cloud, MN
		Rapid City, SD			Jackson, MS
		Bismark, ND	8/25	00Z	Oklahoma City, OK
8/12	00Z	Jackson, MS			Amarillo, TX
8/14	00Z	Albuquerque, NM	8/26	12Z	Nashville, TN
8/15	00Z	Longview, TX			Amarillo, TX
		Oklahoma City, OK			Dodge City, KS
8/17	12Z	Green Bay, WI	8/27	00Z	Amarillo, TX
8/18	12Z	Denver, CO		12Z	Amarillo, TX
		Peoria, IL			Albuquerque, NM
		North Platte, NE	8/28	12Z	Salem, IL
		Green Bay, WI			Huron, SD
8/19	00Z	Peoria, IL			St. Cloud, MN
		Huron, SD			International Falls, MN
		Salem, IL	8/29	12Z	Topeka, KS
8/19	12Z	Peoria, IL			Peoria, IL
		Omaha, NE			Omaha, NE
		Huron, SD			Green Bay, WI
		St. Cloud, MN	8/30	00Z	Topeka, KS
8/20	00Z	Topeka, KS	8/31	00Z	Nashville, TN
	12Z	Peoria, IL		12Z	Salem, IL
					International Falls, MN

APPENDIX C

GAUSSIAN INTERPOLATION OF METEOROLOGICAL DATA

An interpolation scheme is a process whereby a continuous function fills an underlying grid with values by operating on point observations. Two interpolation schemes frequently used in meteorology are those of Cressman (1959) and Barnes (1964; 1973). Both schemes are weighted-average methods in which the value of the variable at the grid point is the sum of weighted values of the individual observations within a fixed distance from the grid point. Cressman's scheme uses a weighting function that is scaled from 1 at the location of the grid point to 0 at some fixed distance N from the grid point. In the Barnes objective analysis technique the continuous function is a Gaussian and the value at the grid point is generally smaller than 1. This study has used the latter's scheme to interpolate the model adjustment factors calculated at point locations to a field of adjustments at the spatial resolution of the satellite data.

The average spacing of the original data determines which atmospheric waves can be uniquely specified and consequently the values of the analysis constants. If the spacing of the observations is Δx, then atmospheric phenomena of wavelengths no smaller than $2\Delta x$ can be sampled. In fact, the $2\Delta x$ wave is a lower limit on the sampling and, in practice, a more reasonable prospect is to assure adequate sampling of the $6\Delta x$ waves (S.L. Barnes, personal communication). The response function is a measure of the contribution that each wavelength makes to the interpolation. It is defined by

$$R = \exp\left(-\frac{4\pi^2}{\lambda^2} c\right) , \qquad \text{(C-1)}$$

where R is the response, λ is the wavelength (km) of the phenomenon of interest and c is the analysis constant (km^2).

Barnes' objective analysis scheme as described in his 1973 Technical Memorandum is a two-pass scheme. In the first pass, a Gaussian weighting function operates on the point observations to produce a first estimate of the interpolated field. In the second pass, a correction restores the amplitude of the wave(s) of interest to yield the final field.

In this study, the Cartesian coordinate formulation (Maddox, 1980b) rather than the polar coordinate formulation of Barnes' scheme has been used. Consider the point observations at locations (x,y) of a parameter f(x,y). These data are to be interpolated to values f_0 on an (i,j) grid. The first-guess interpolated values are computed from

$$f_o(i,j) = \sum_{n=1}^{N} w_n f_n(x,y) \; / \; \sum_{n=1}^{N} w_n \; , \qquad \text{(C-2)}$$

where N is the number of points in the original data that are allowed to influence the interpolated point, and w_n is the weight function. In Barnes' scheme the weight function is the Gaussian function

$$w_n = \exp(-d_n^2/4c) \; . \qquad \text{(C-3)}$$

The variable d_n is the distance from the (i,j) grid point to the observed datum f_n that is located at (x,y). The correction pass on the first-guess values, $f_o(i,j)$, computes the final interpolated values, f(i,j), by

$$f(i,j) = f_o(i,j) + \sum_{n=1}^{N} w_n' D_n \; / \; \sum_{n=1}^{N} w_n' \; . \qquad \text{(C-4)}$$

The function w_n' is a Gaussian in which the constant of (C-3) has been reduced:

$$w_n' = \exp(-d_n^2/4gc) \qquad \text{(C-5)}$$

for g between 0 and 1. D_n is the difference between the observed value and the initial guess value at the same point

$$D_n = f_n(x,y) - f_o(x,y) \; . \qquad \text{(C-6)}$$

The value $f_o(x,y)$ is estimated by a biquadratic interpolation between the $f_o(i,j)$ values at the four grid points closest to (x,y).

The extent of the grid that can be interpolated is determined by the spacing and location of the observations. The interpolated field must lie within the area bound by the outermost observations. Although the gages are roughly bound by 30°N, 50°W, 85°W, and 110°W, the boundaries of the maximum area that can be interpolated are 30.5°N, 46.75°N, 90.25°W, and 108.5°W.

In interpolating the model adjustment factors, the intent was to closely reproduce the original data. The first set of runs used the two-pass system described above. Four observations were allowed to influence each grid point, and the constants c and g were set to 2000 km^2 and 0.2, respectively. The response functions of the two passes and the resultant filter are shown in Fig. C.1. Response is plotted as a function of wavelength, and the responses from the first and second passes are the lower and upper curves, respectively. As can be seen from the figure, the second pass has considerably "sharpened up" the response at all wavelengths, so much so that the final response looks like a top hat.

For the radiosonde network, Δx is ~400 km. Fig. C.1 shows that wavelengths much smaller than 800 km have a high response; that is, the response is much greater than 0.1. This high response at small wavelengths means that noise from short wavelengths is included in the interpolation results. Fig. C.2 supports this contention. In this example the model adjustment field appears rather blocky and there are small wiggles on the angular contour lines. Incorrect values for the analysis constants were hypothesized as probable causes for the angularity. Truncation errors that are associated with the number of observations allowed to influence each grid point calcula

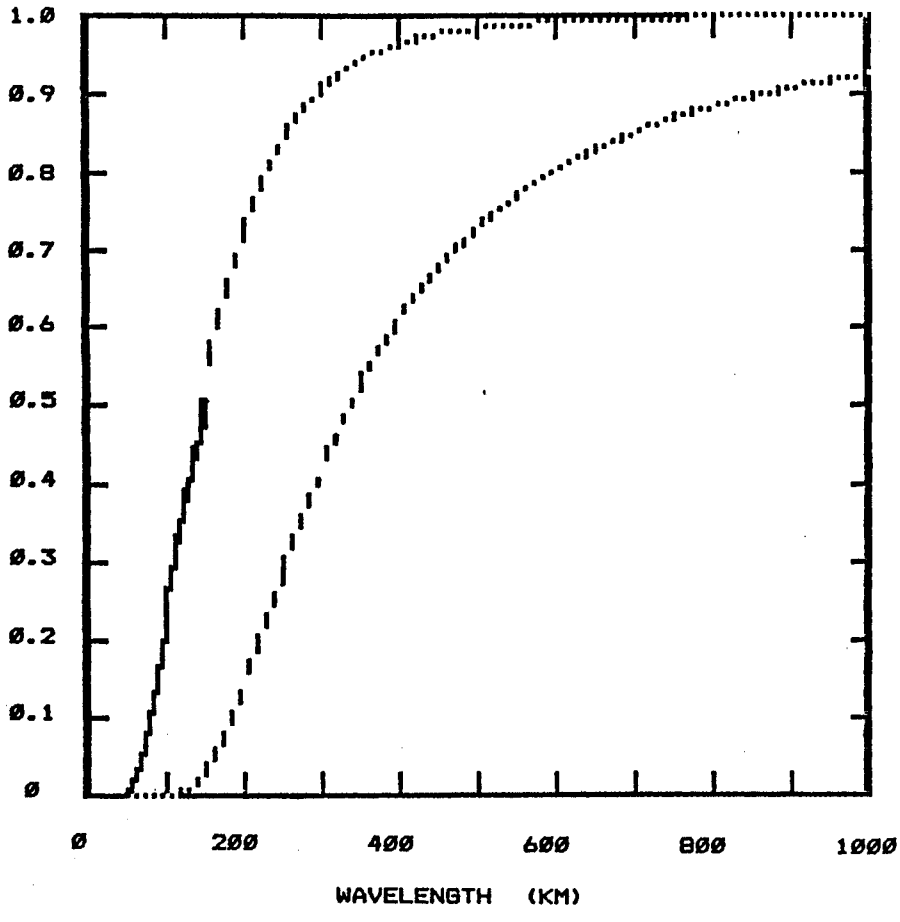


Fig. C.1. First- and second-pass response curves for the Gaussian weight function with the constants c = 2,000 km^2 and g = 0.2.

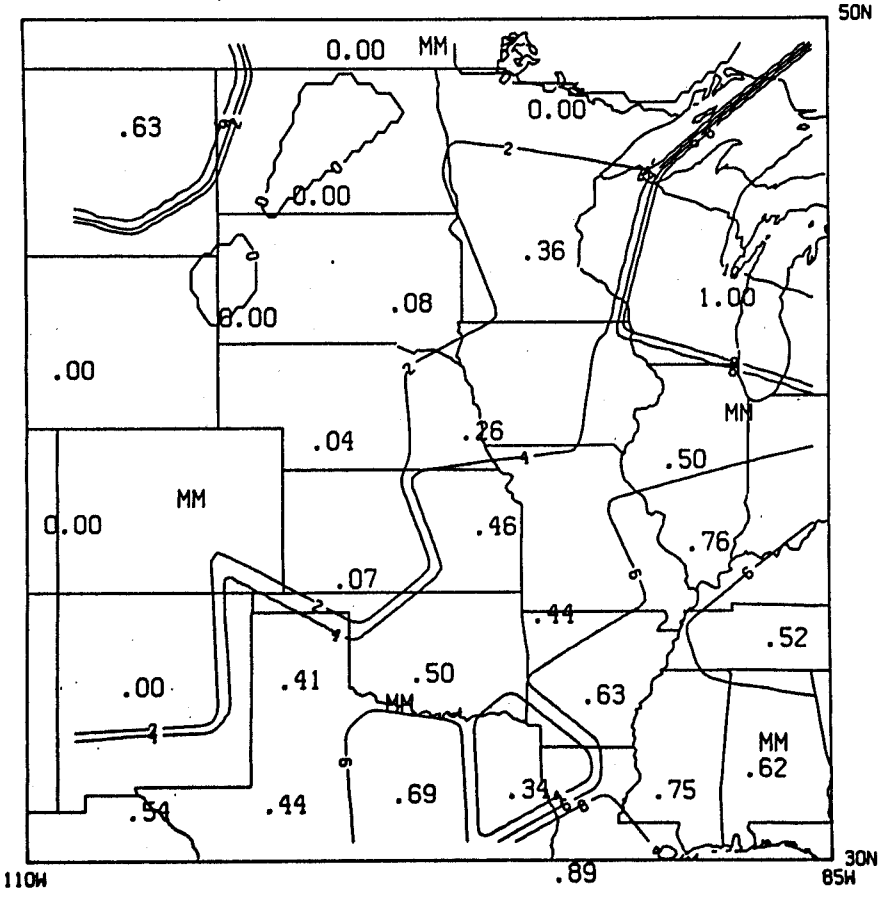


Fig. C.2. The model adjustment field interpolated by the Gaussian described in Fig. C.1 for 5 August 1979 at 1200 GMT. Four observations have been allowed to influence the calculation at each grid point. The model adjustment factors are also shown.

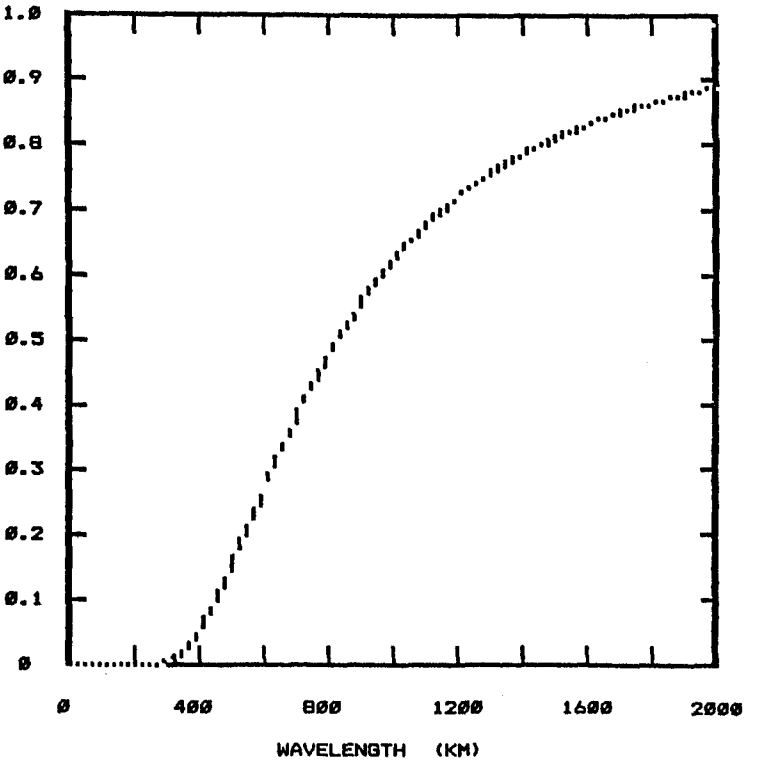


Fig. C.3. First-pass response curve for the Gaussian weight function with the constant c = 12,000 km^2.

tion were thought to be the source of the noise that produced the smaller scale wiggles.

Equation (C-1) was used to choose the value of the analysis constant "c" so that small wavelengths are excluded from the interpolation. Furthermore, it was advised that the second pass be omitted, thereby eliminating the "top hat" response (C.A. Doswell, III, personal communication). Fig. C.3 shows the response function for a value of c equal to 12,000 km^2. Only one curve appears in this figure because no second pass was used. The response at 800 km has dropped substantially to less than 0.1. At $4\Delta x$ (1600 km) the response is about 50%, and at $6\Delta x$ it is on the order of 70% of the value of the amplitude of the original. The resulting model adjustment factor field is shown in Fig. C.4. If judged simply by the smoothness of the field, this field is a decided improvement over the field of Fig. C.2, although there is still small-scale noise as evinced by the wiggles on the contours.

Dayton Vincent (personal communication) found that the optimum number of observations needed to influence each grid computation should be in the range of 6 to 10, in contrast to the four points used in the interpolations shown so far. Fig. C.5 is the model adjustment field that results from Barnes' scheme of one pass with c=12,000 km^2 (as in Fig. C.4), but with eight observations influencing each calculation. The smallest wiggles that result from truncation errors in the computation have disappeared. Consequently all model adjustment factor fields calculated for this study were generated under the specifications of a one-pass Barnes interpolation, with the analysis constant c equal to 12,000 km^2 and with eight observations being allowed to influence each calculation.

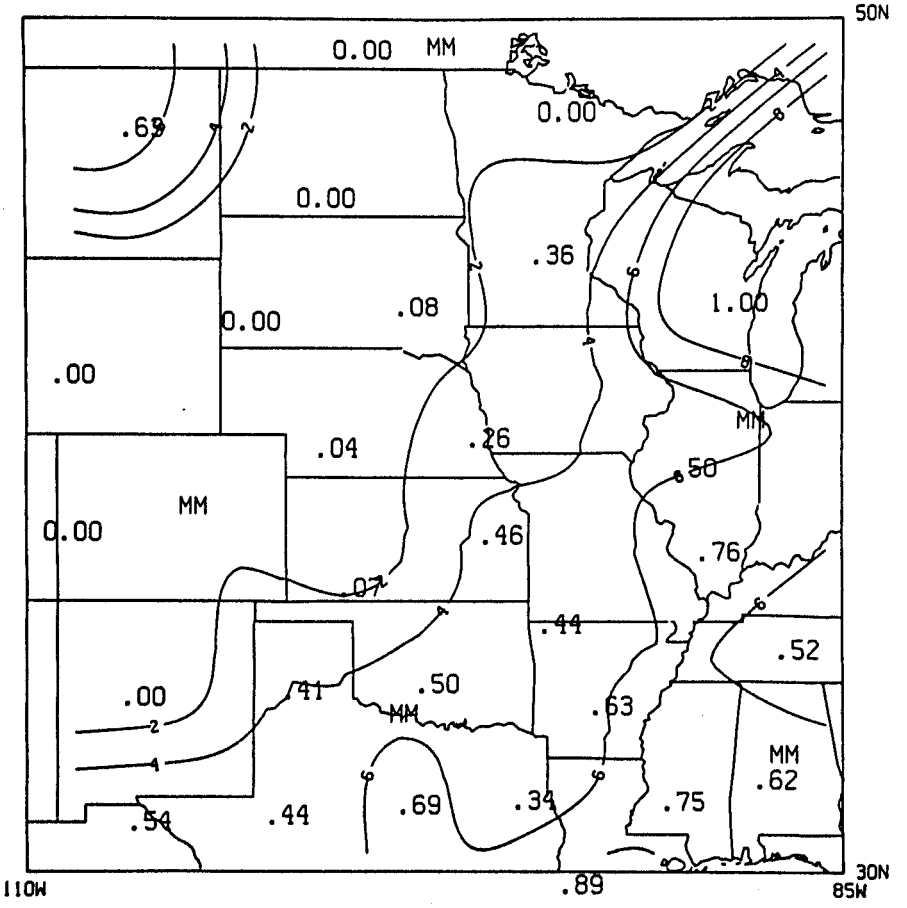


Fig. C.4. The model adjustment field interpolated by the Gaussian described in Fig. C.3 for 5 August 1979 at 1200 GMT. Four observations have been allowed to influence the calculation at each grid point. The model adjustment factors are also shown.

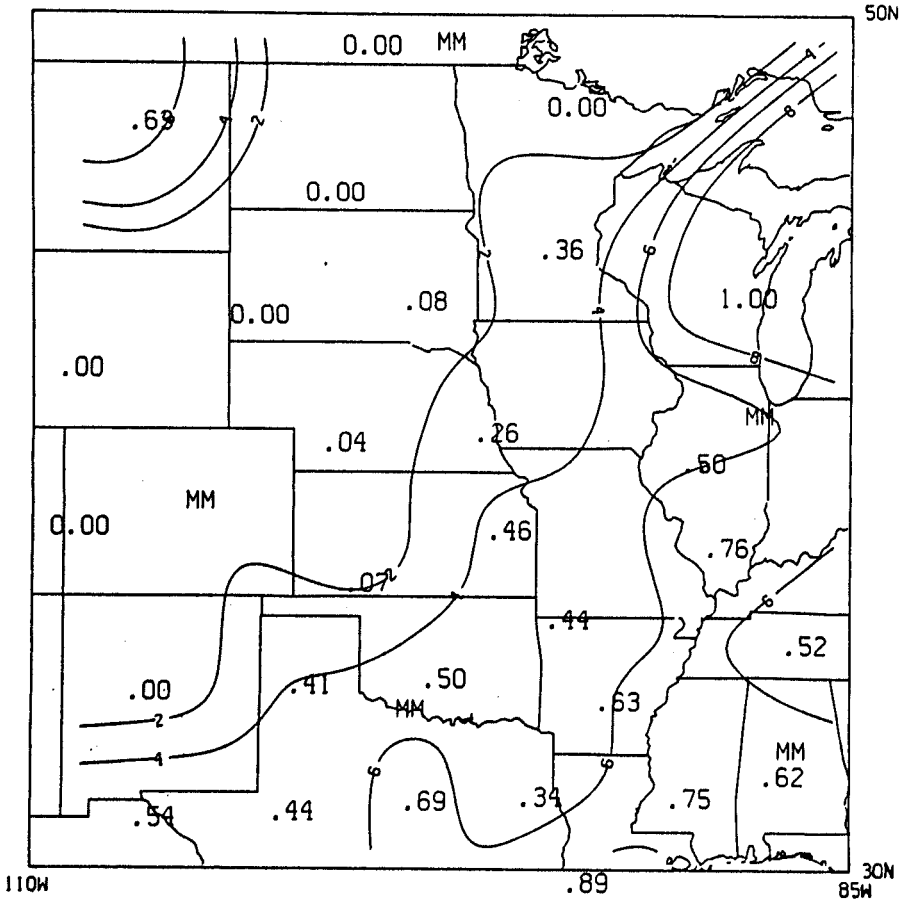


Fig. C.5. The model adjustment field, as Fig. C.4, but with eight observations allowed to influence each grid point.

APPENDIX D

EXPLORATORY DATA ANALYSIS TECHNIQUES

Observations of geophysical data are rarely normally distributed, and consequently the use of statistical measures that are based on the normal distribution can be misleading, if not in error. Likewise, there are often extreme and outlying values in the data that highly influence the computation of the mean of a sample or its least-squares regression. In recent years techniques of data presentation and analysis that are not based on the normal distribution and that are resistant to the effect of a small number of outliers have begun to appear in the literature under the title of exploratory data analysis (EDA). Kleiner and Graedel (1980) presented a brief tutorial on the subject with applications to the geophysical sciences. Tukey (1977), Velleman and Hoaglin (1981), and Hoaglin et al. (1983) were the primary reference sources used in this study.

EDA includes such familiar graphical displays as the histogram and such exotic sounding devices as the rootogram. The EDA tools used in this study are the box-and-whisker plot and the stem-and-leaf plot. The advantages of these tools are that they help to clarify the distribution of the data sample in more or less detail, allow impressions of single data sets that can be easily apprehended, and greatly facilitate comparisons among several data sets through the patterns found in these displays. Perhaps the greatest advantage of EDA is to draw the attention of the investigator to data values that do not fall within the bulk of the data set. It is these outliers that often raise the really interesting questions.

The most detailed of the two tools used here is the stem-and-leaf plot. In this device every observation in the sample is displayed in relation to the other observations. A sample stem-and-leaf plot of the daily, area-averaged satellite rainfalls (in millimeters and expressed to one decimal place) for the adjusted streamlined technique is found in Fig. D.1. In this example, the integer value of the area-averaged rainfalls forms the stem; these are the digits that are written to the left of the vertical line. The tenths part of the fractional portion of each measurement is the leaf, written to the right of the vertical line. There are 31 observations, 7 of which, for instance, are area-averaged rainfalls smaller than 1 mm (0.0, 0.1, 0.1, 0.2, 0.4, 0.5, and 0.8 mm) and 1 of which is larger than 5 mm (5.1 mm). There are no observations in the range of 4.0 to 4.9 mm. From the stem-and-leaf plot one can see that the sample is asymmetric, peaking around 1 mm and having a tail with one extreme value (5.1 mm) that is separated from the rest of the data. The digits on the extreme left are a running sum of the number of observations from each end to the interval in which the median occurs (which is indicated by the parentheses). The median, of course, is the value for which half the (ranked) sample is smaller and half is larger, and is defined in the usual way. That is, in a sample containing an odd number of observations, the

Number	Box-and-whisker plot	
7	0	0 1 1 2 4 5 8
(10)	1	0 2 3 3 5 5 5 7 7 7
14	2	0 0 0 1 1 1 3 6
6	3	0 1 4 4 6
1	4	
1	5	1

Fig. D.1. Stem-and-leaf plot of daily area-averaged rainfalls (mm) computed from the adjusted streamlined satellite technique for August 1979.

median is the single middle value. For a sample with an even number of observations, the median is the average of the two middle observations.

In addition to the median, other values that are used in EDA to characterize the sample distribution include the hinges (which are almost numerically equivalent to the upper and lower quartile values) and the extreme values. The hinges divide into halves the ranked sample between the extremes and the median. The hinges are defined by

$$d(H) = ([d(M)] + 1) / 2 , \qquad (D\text{-}1)$$

where, in the notation of Velleman and Hoaglin (1981),

H refers to the hinge,
M refers to the median,
d(H) is the depth of the hinge, that is, the ordinal position of the hinge in the ordered sample,
d(M) is the depth of the median, and
[ ] denotes the integer part of the enclosed quantity.

In this example the lower and upper hinges are 1.1 and 2.2 mm, respectively. The hinges lie closer to the median than the quartiles do and so differ slightly in value from them. The lower and upper quartiles for the data of Fig. D.1 have values of 1.0 and 2.3 mm, respectively. The extreme values of the sample are 0.0 and 5.1 mm.

In EDA a differentiation is made between the extreme values of the sample and outliers in the sample. The extreme values are the largest and smallest values in the sample. Outliers, on the other hand, have a specified relationship to the hinges, and two types of outliers are defined--outside and far outside outliers. Outliers have not been identified in this study, however, because their definition is based on parameters related to the normal distribution, and rainfall clearly does not follow a normal distribution.

Box-and-whisker plots are a simple graphical summary of the median, hinges, and extremes. The box-and-whisker plot for the data in Fig. D.1 is given in Fig. D.2. (This type of box-and-whisker plot is referred to by Tukey [1977] as a "schematic plot" and is called a boxplot by Velleman and Hoaglin [1981].) The hinges of the sample form the ends of the box and the median is the horizontal line in the middle of the box. The dashed lines are the

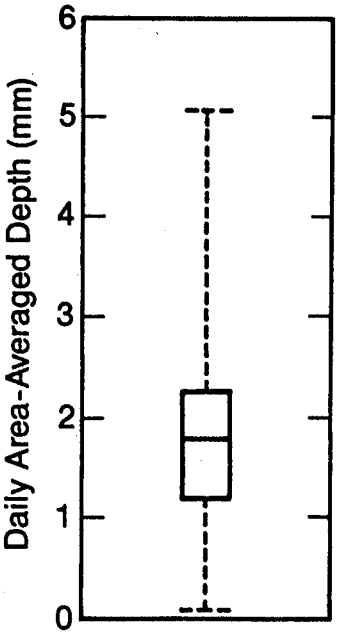


Fig. D.2. Box-and-whisker plot of daily area-averaged rainfalls (mm) computed from the adjusted streamlined satellite technique for August 1979.

whiskers and extend from the hinges to the extreme values in the sample.

A box-and-whisker plot is thus a compact summary of the data sample. The box emphasizes the middle 50% of the sample. The distances between the hinges and the median give some idea of the skewness of the sample. And the extremes indicate the existence of tails on the distribution. For the data of Fig. D.2, the middle 50% are symmetrically distributed around the median, but the upper extreme value causes a tail toward higher rain depths. As the modified box shows, the tail for the lower amounts is fairly evenly filled out (the 5% interval is close to the lower extreme), but the upper tail is not, as we have seen in Fig. D.1.

I have also devised a modified box-and-whisker plot that emphasizes the middle 90% of the data and delineates how much of either tail is due to 10% of the data. In Fig. D.3, the dashed box encloses the middle 90% and each tail contains 5%. The 5% and 95% intervals in this example occur at 0.1 and 3.6 mm, respectively.

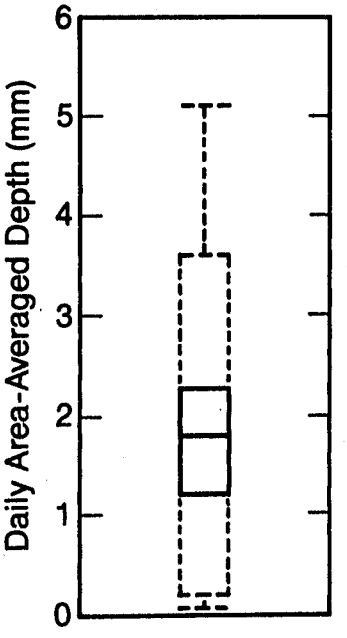


Fig. D.3. Modified box-and-whisker plot of daily area-averaged rainfalls (mm) computed from the adjusted streamlined satellite technique for August 1979. The dashed box encloses 90% of the data.

APPENDIX E

TIME SERIES OF HOURLY AREA-AVERAGED RAINFALL

Time series of hourly area-averaged rainfalls for August 1979 are presented here. Gage rainfalls with one satellite array permutation are plotted on each figure. The four satellite permutations are unadjusted life history (Figs. E.1-E.4), adjusted life history (Figs. E.5-E.8), unadjusted streamlined (Figs. E.9-E.12), and adjusted streamlined (Figs. E.13-E.16).

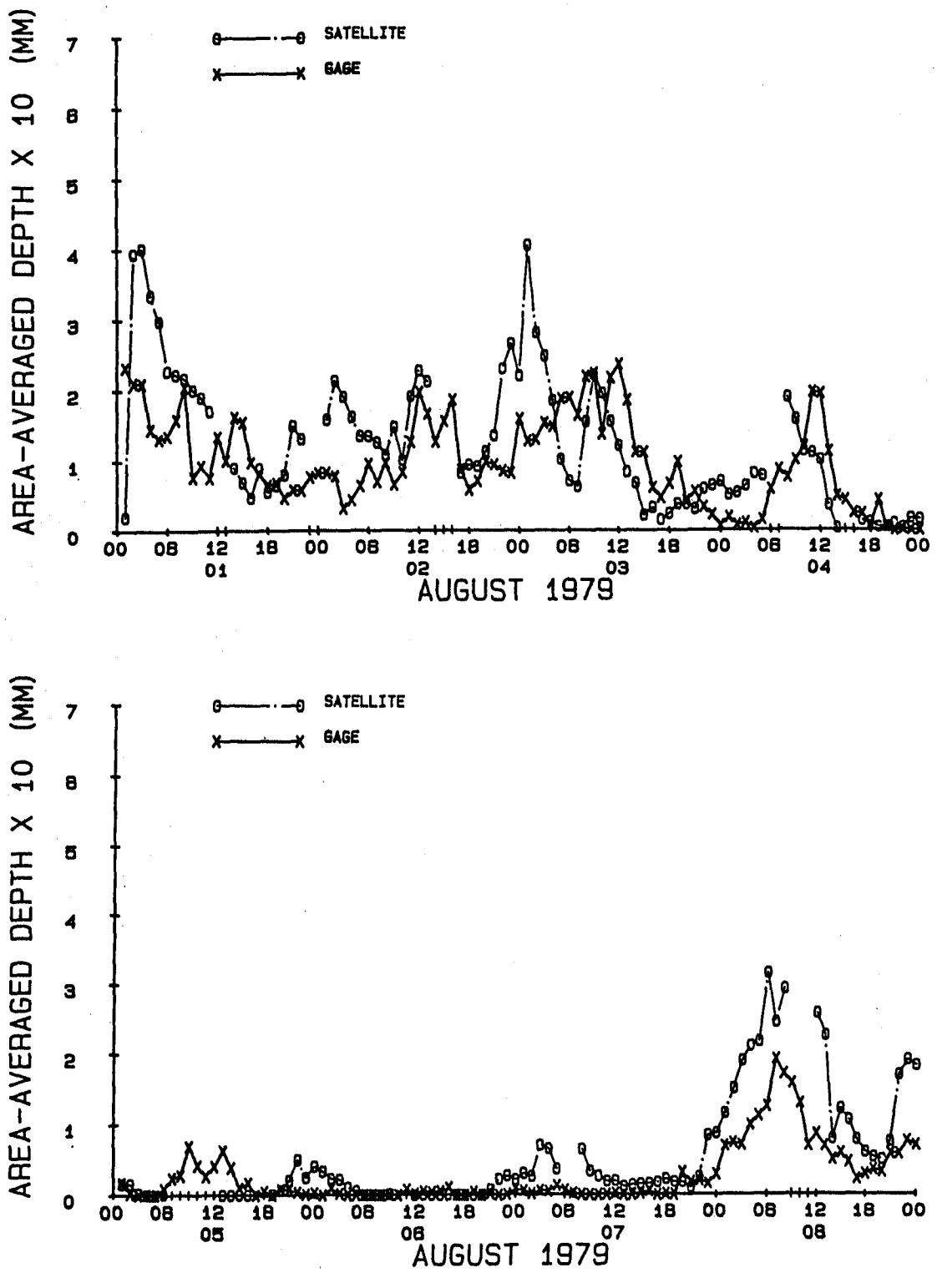


Fig. E.1. Time series of hourly area-averaged unadjusted life history and gage rainfalls for August 1-4 (top) and 5-8 (bottom), 1979.

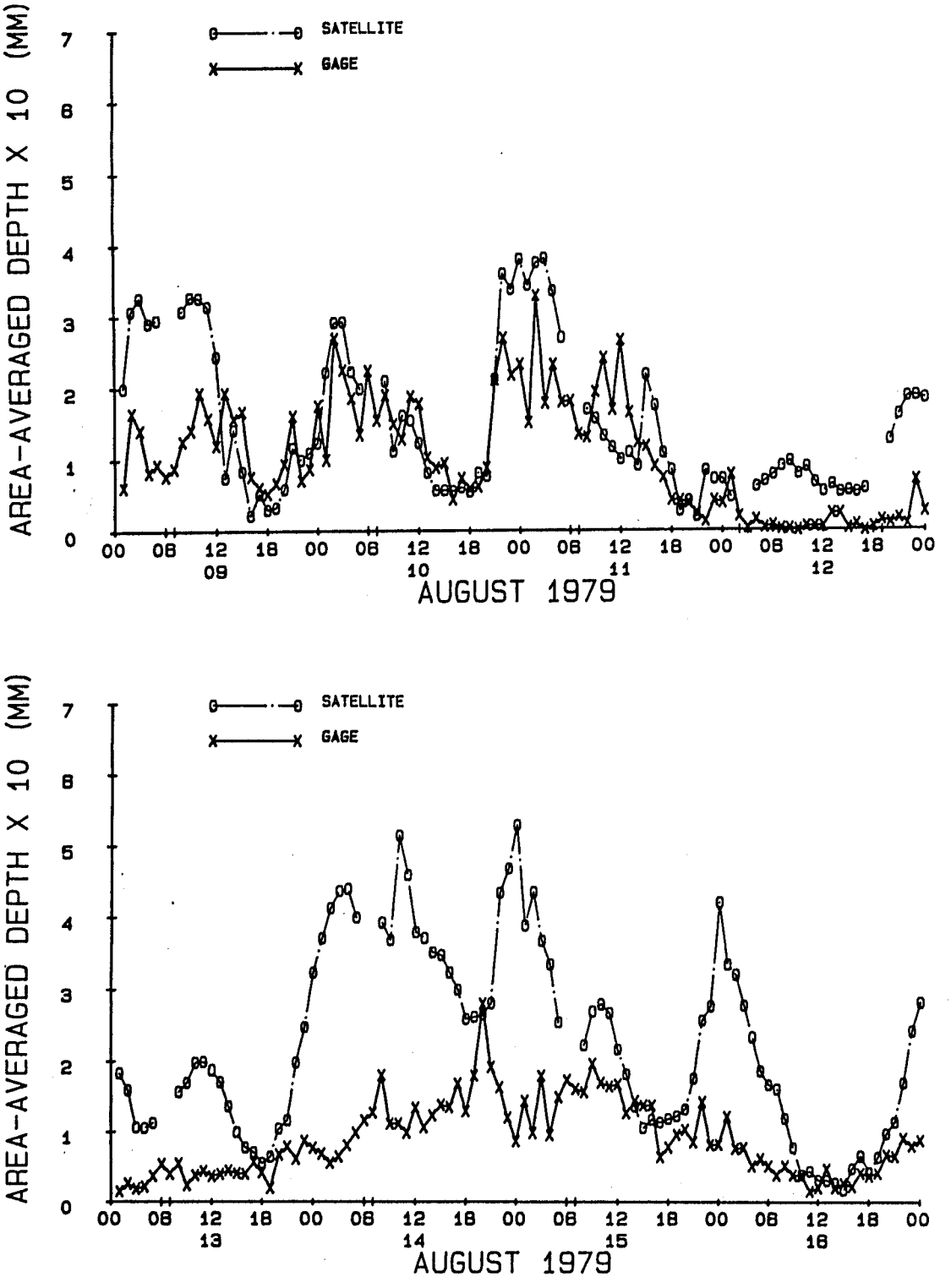


Fig. E.2. Time series of hourly area-averaged unadjusted life history and gage rainfalls for August 9-12 (top) and 13-16 (bottom), 1979.

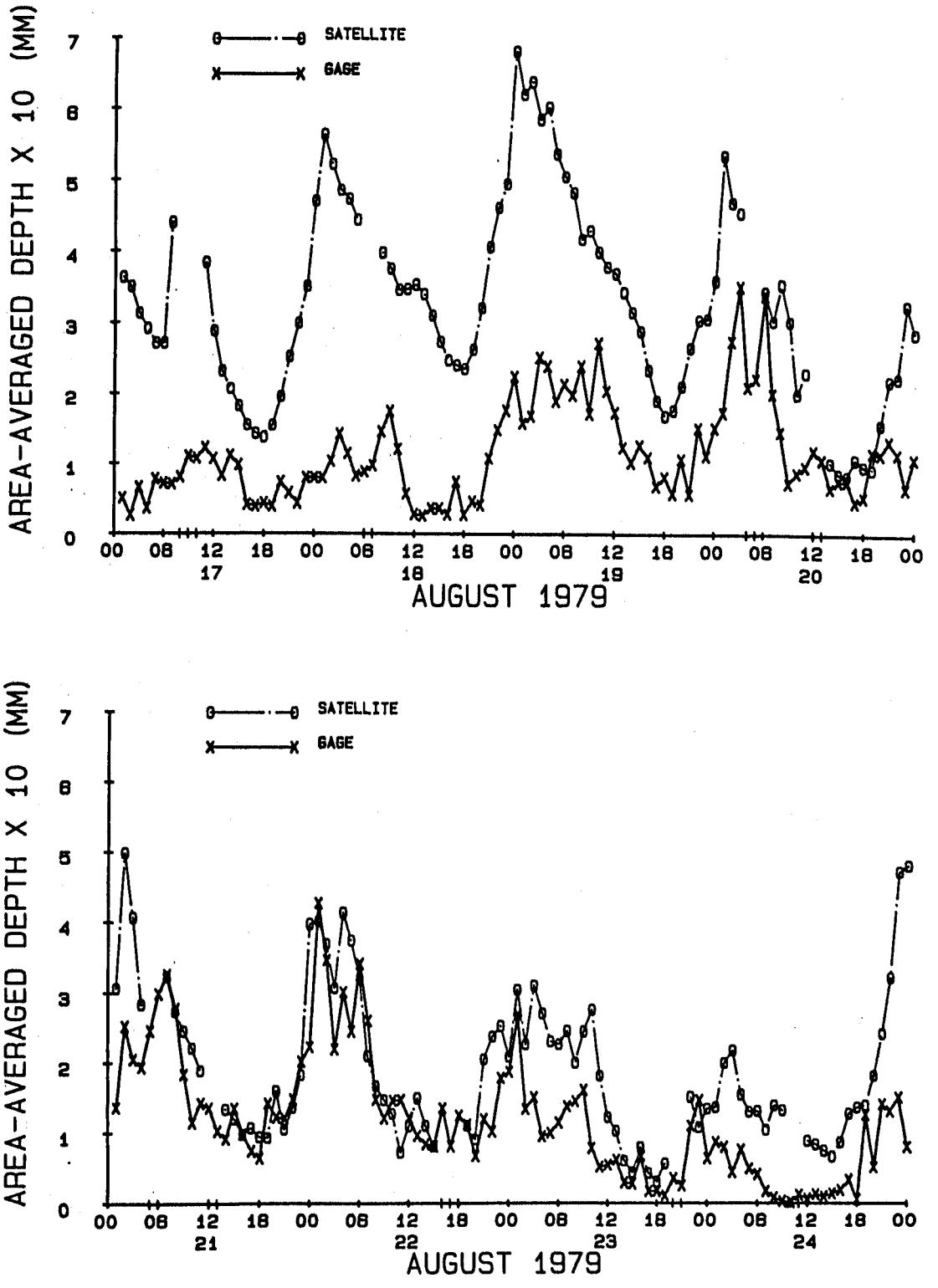


Fig. E.3. Time series of hourly area-averaged unadjusted life history and gage rainfalls for August 17-20 (top) and 21-24 (bottom), 1979.

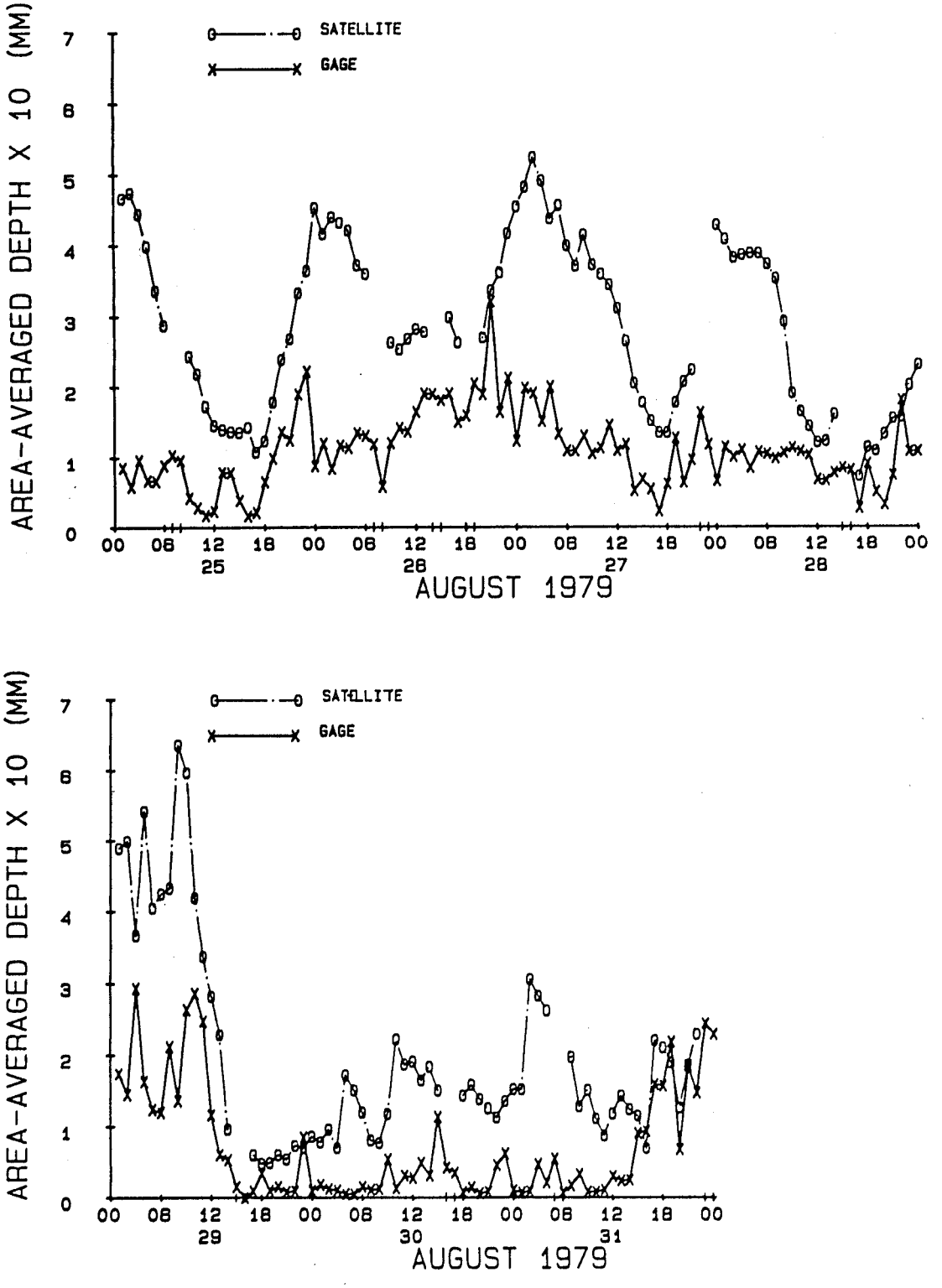


Fig. E.4. Time series of hourly area-averaged unadjusted life history and gage rainfalls for August 25-28 (top) and 29-31 (bottom), 1979.

Blank page retained for pagination

Blank page retained for pagination

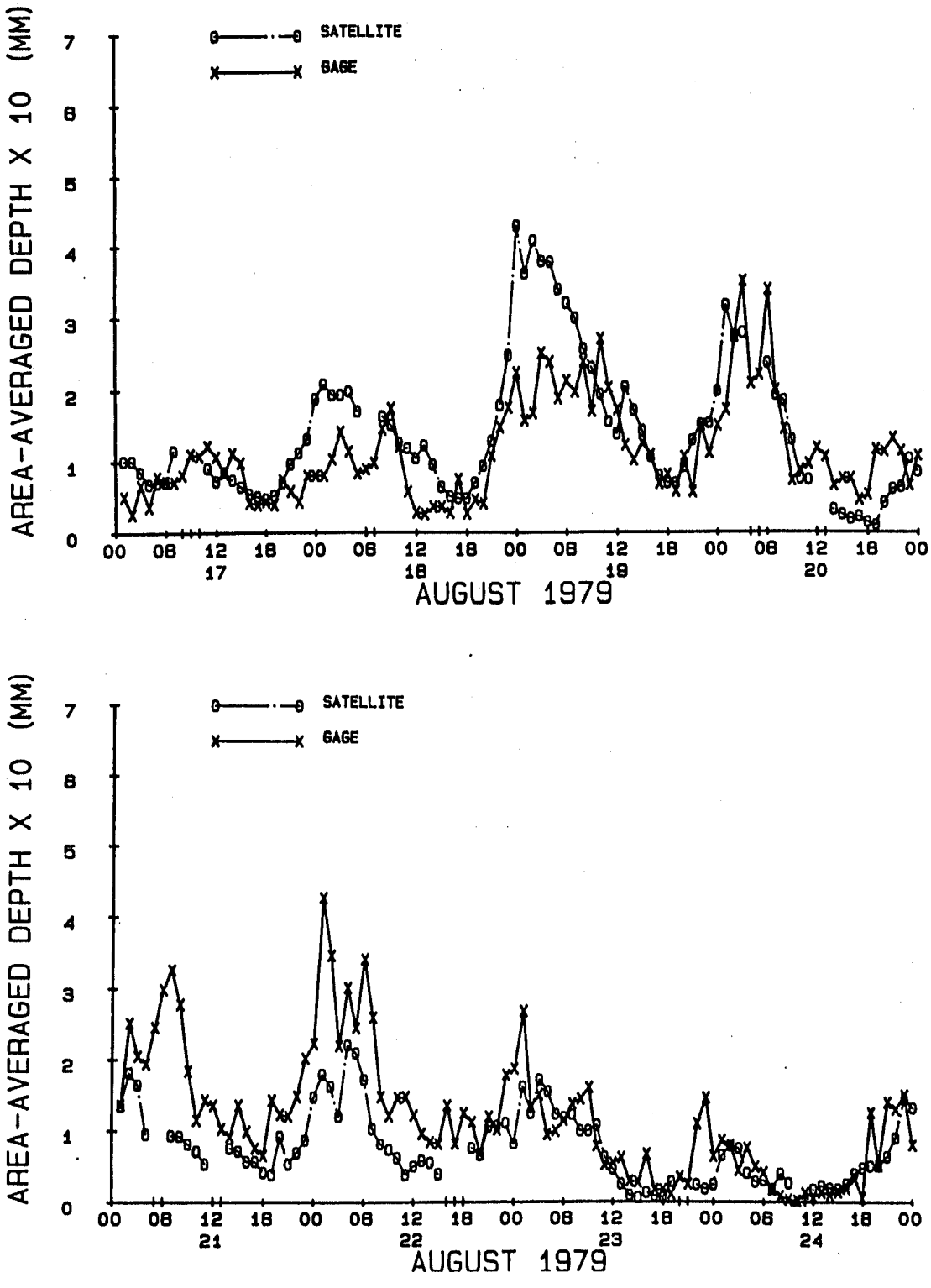


Fig. E.7. Time series of hourly area-averaged adjusted life history and gage rainfalls for August 17-20 (top) and 21-24 (bottom), 1979.

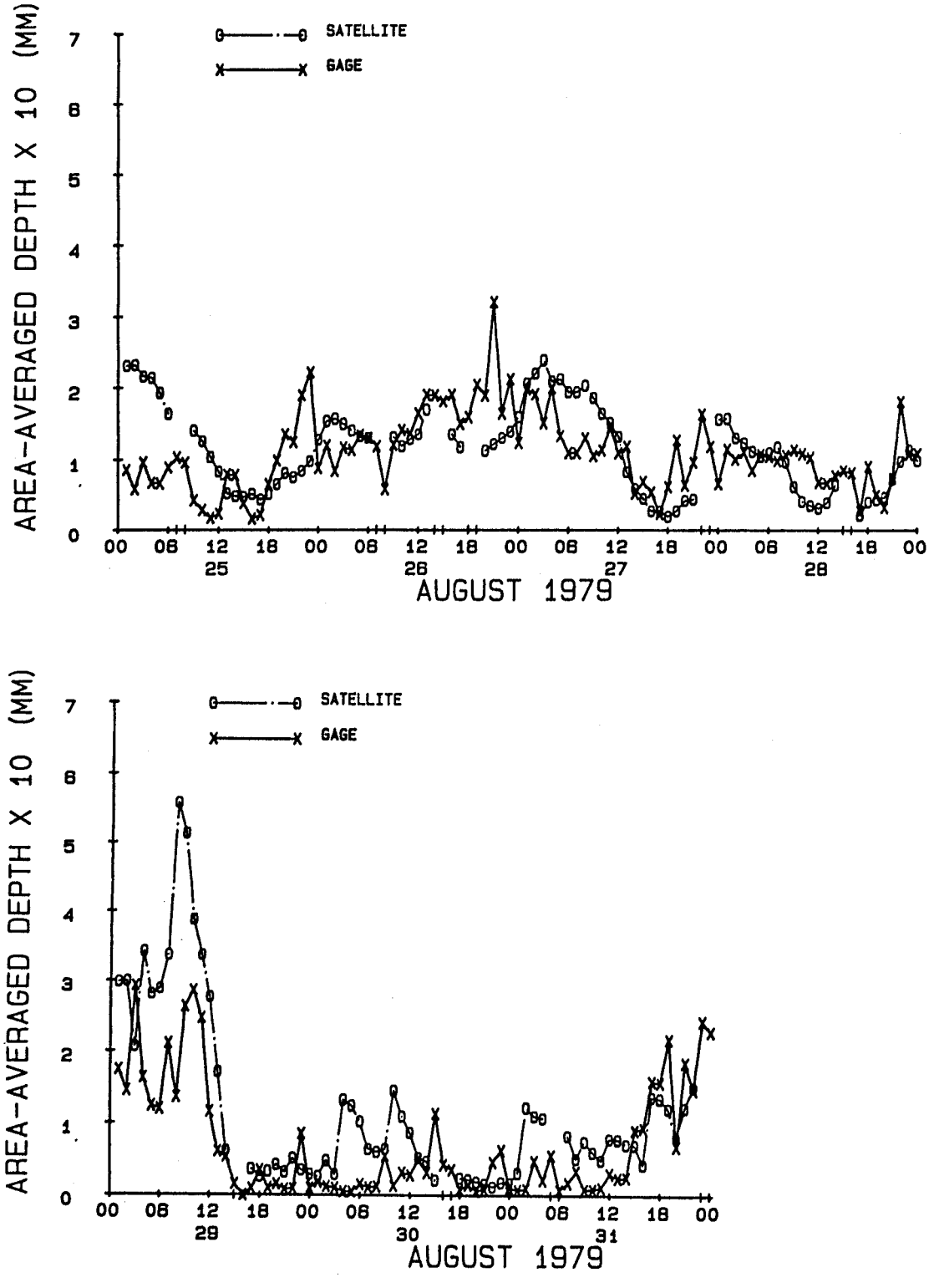


Fig. E.8. Time series of hourly area-averaged adjusted life history and gage rainfalls for August 25-28 (top) and 29-31 (bottom), 1979.

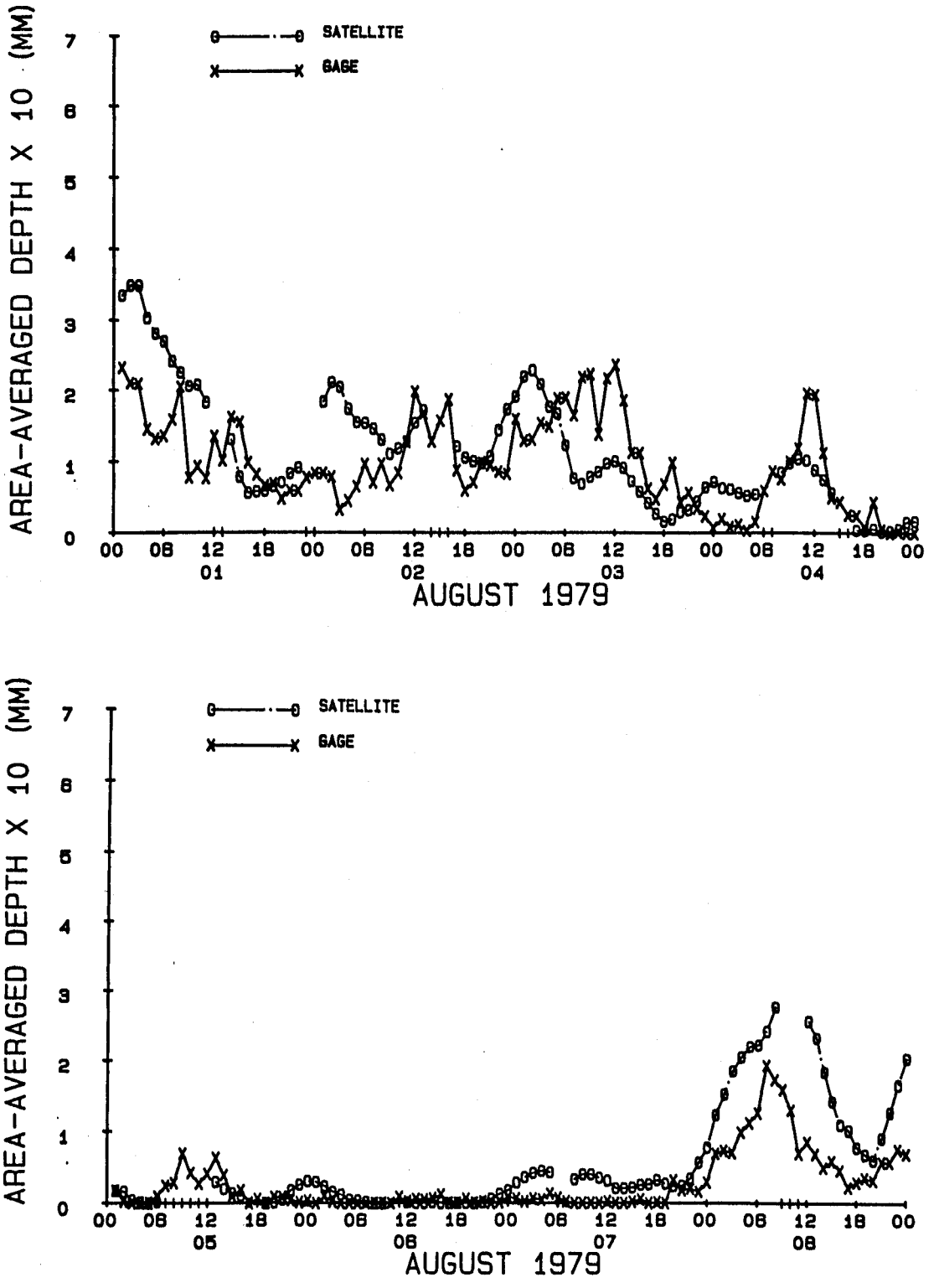


Fig. E.9. Time series of hourly area-averaged unadjusted streamlined and gage rainfalls for August 1-4 (top) and 5-8 (bottom), 1979.

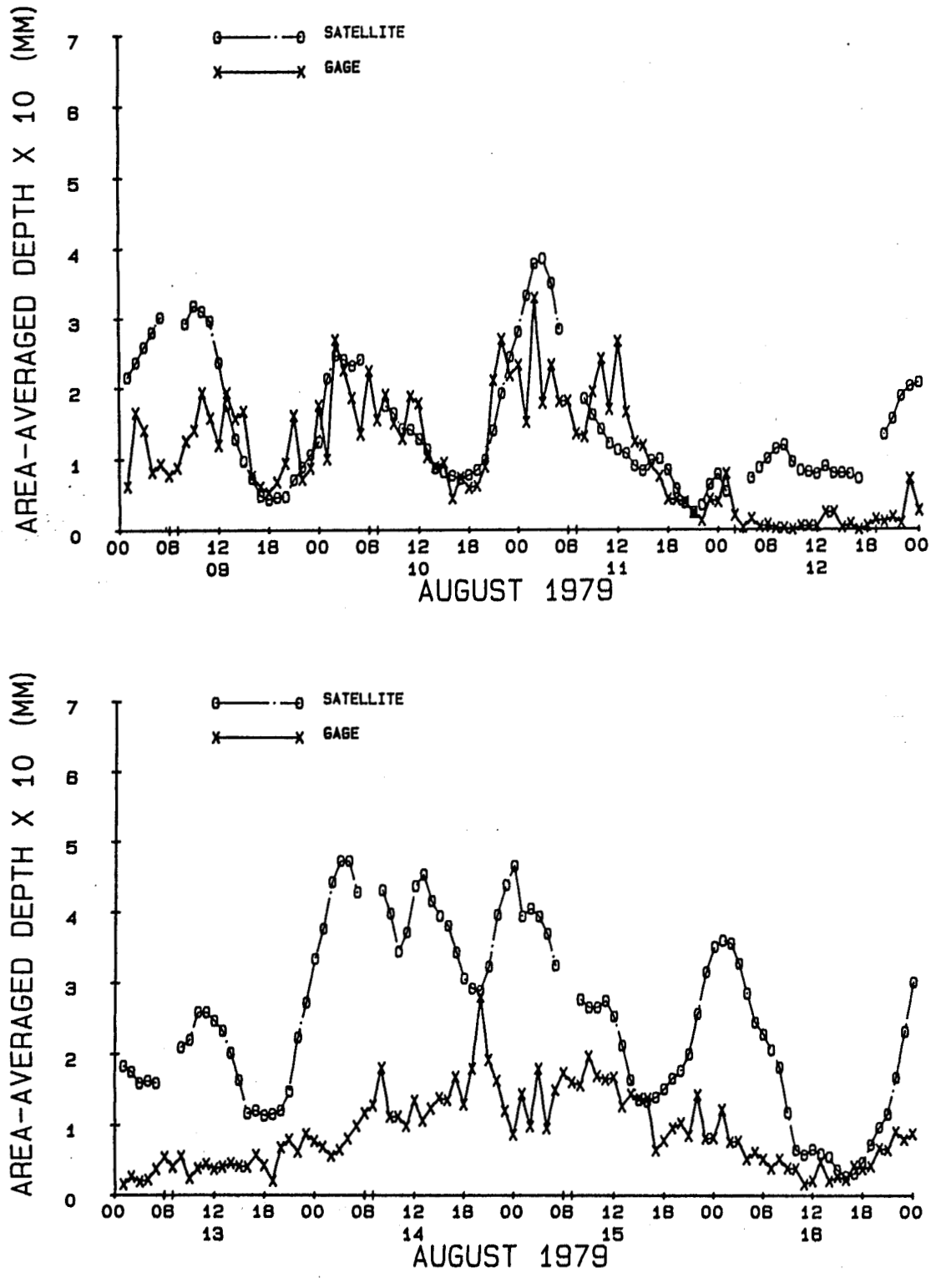


Fig. E.10. Time series of hourly area-averaged unadjusted streamlined and gage rainfalls for August 9-12 (top) and 13-16 (bottom), 1979.

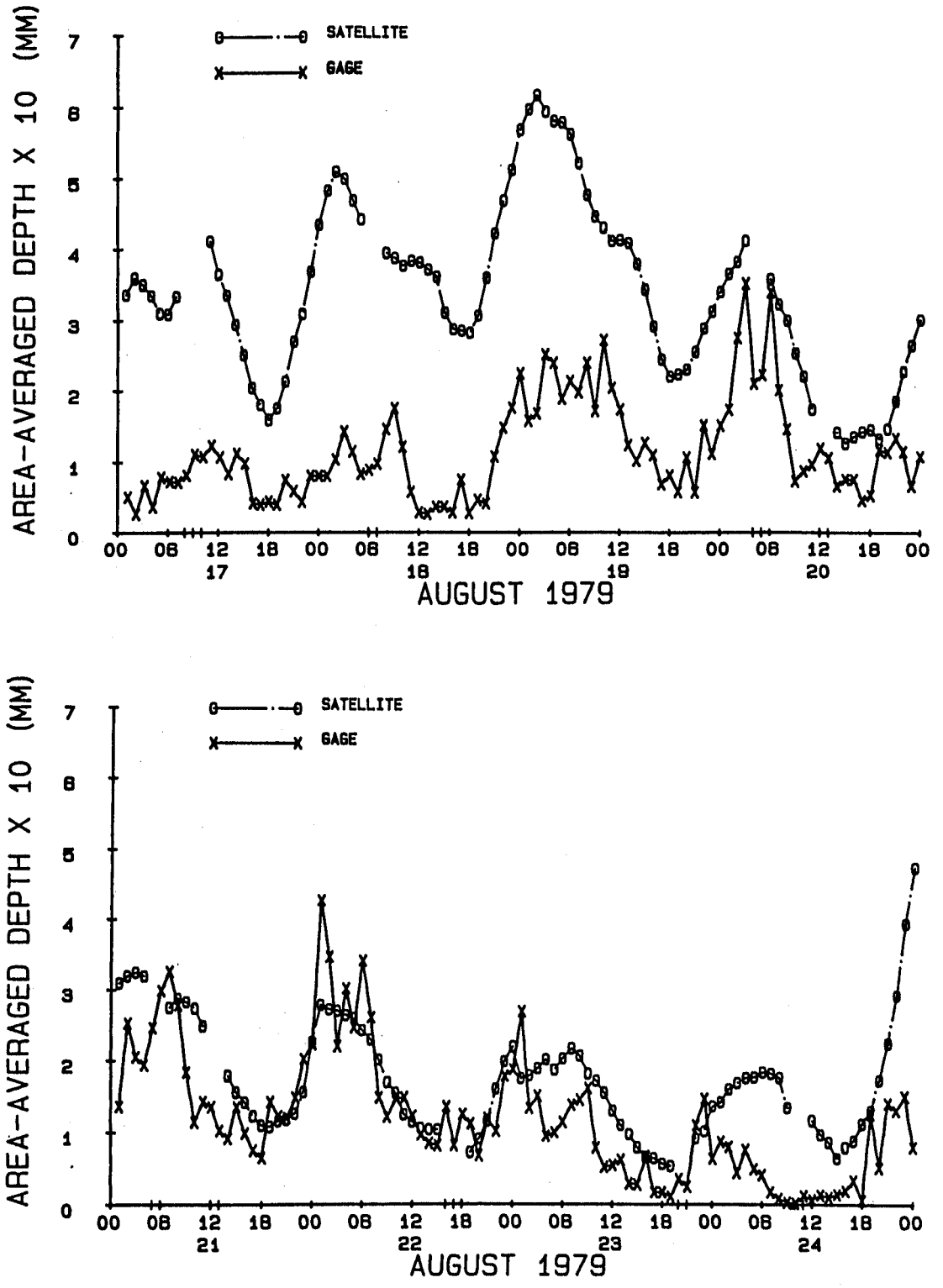


Fig. E.11. Time series of hourly area-averaged unadjusted streamlined and gage rainfalls for August 17-20 (top) and 21-24 (bottom), 1979.

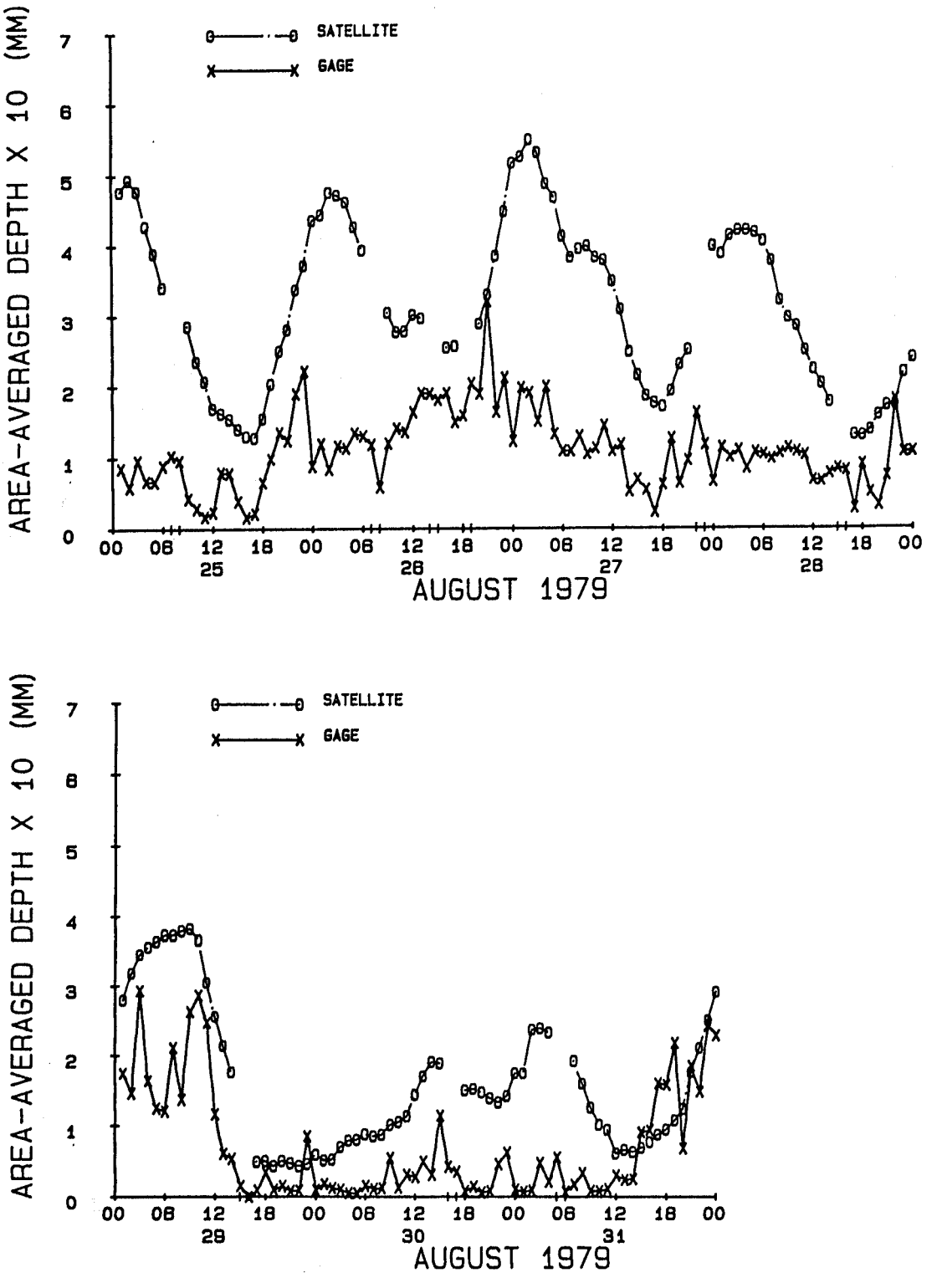


Fig. E.12. Time series of hourly area-averaged unadjusted streamlined and gage rainfalls for August 25-28 (top) and 29-31 (bottom), 1979.

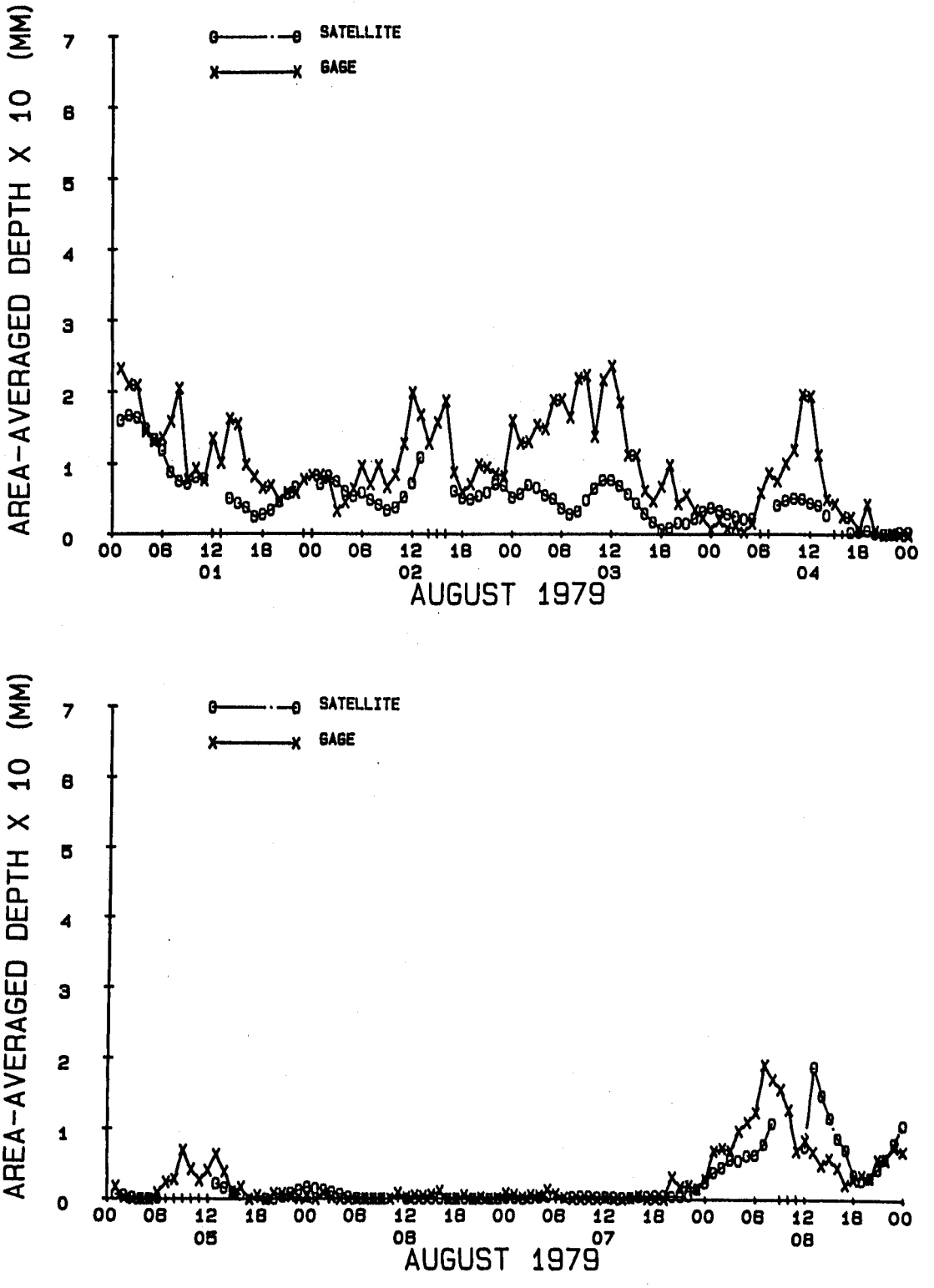


Fig. E.13. Time series of hourly area-averaged adjusted streamlined and gage rainfalls for August 1-4 (top) and 5-8 (bottom), 1979.

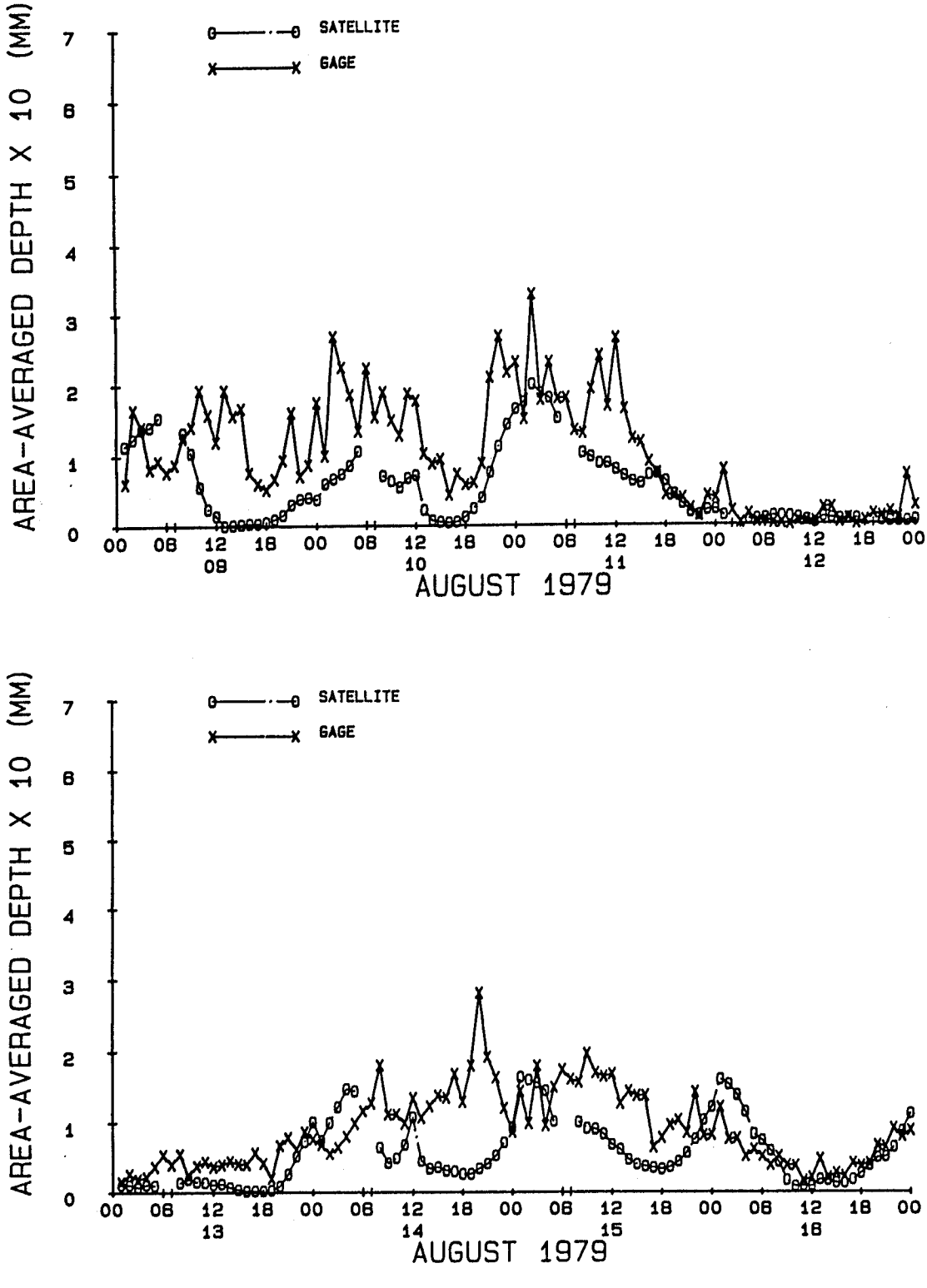


Fig. E.14. Time series of hourly area-averaged adjusted streamlined and gage rainfalls for August 9-12 (top) and 13-16 (bottom), 1979.

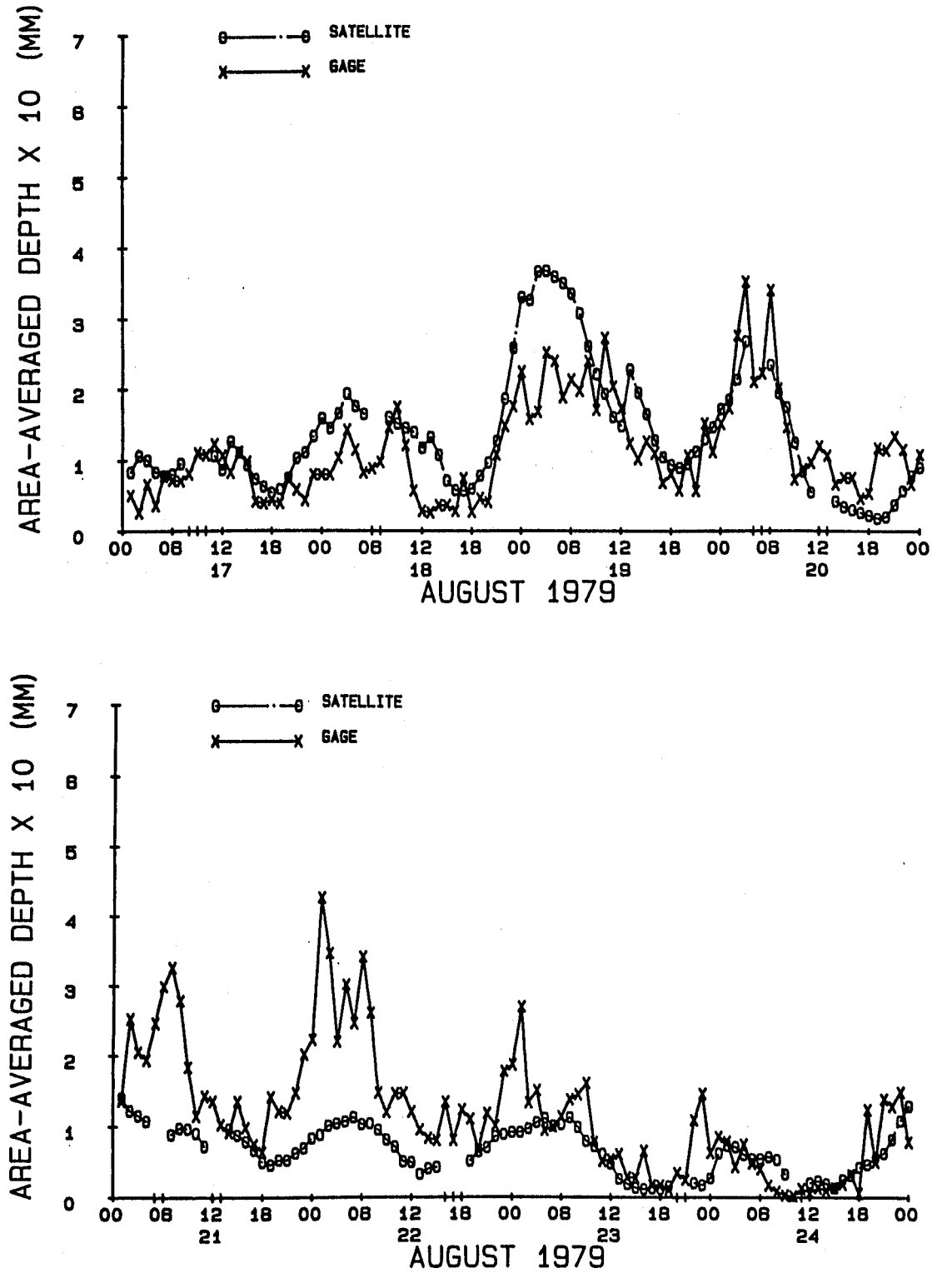


Fig. E.15. Time series of hourly area-averaged adjusted streamlined and gage rainfalls for August 17-20 (top) and 21-24 (bottom), 1979.

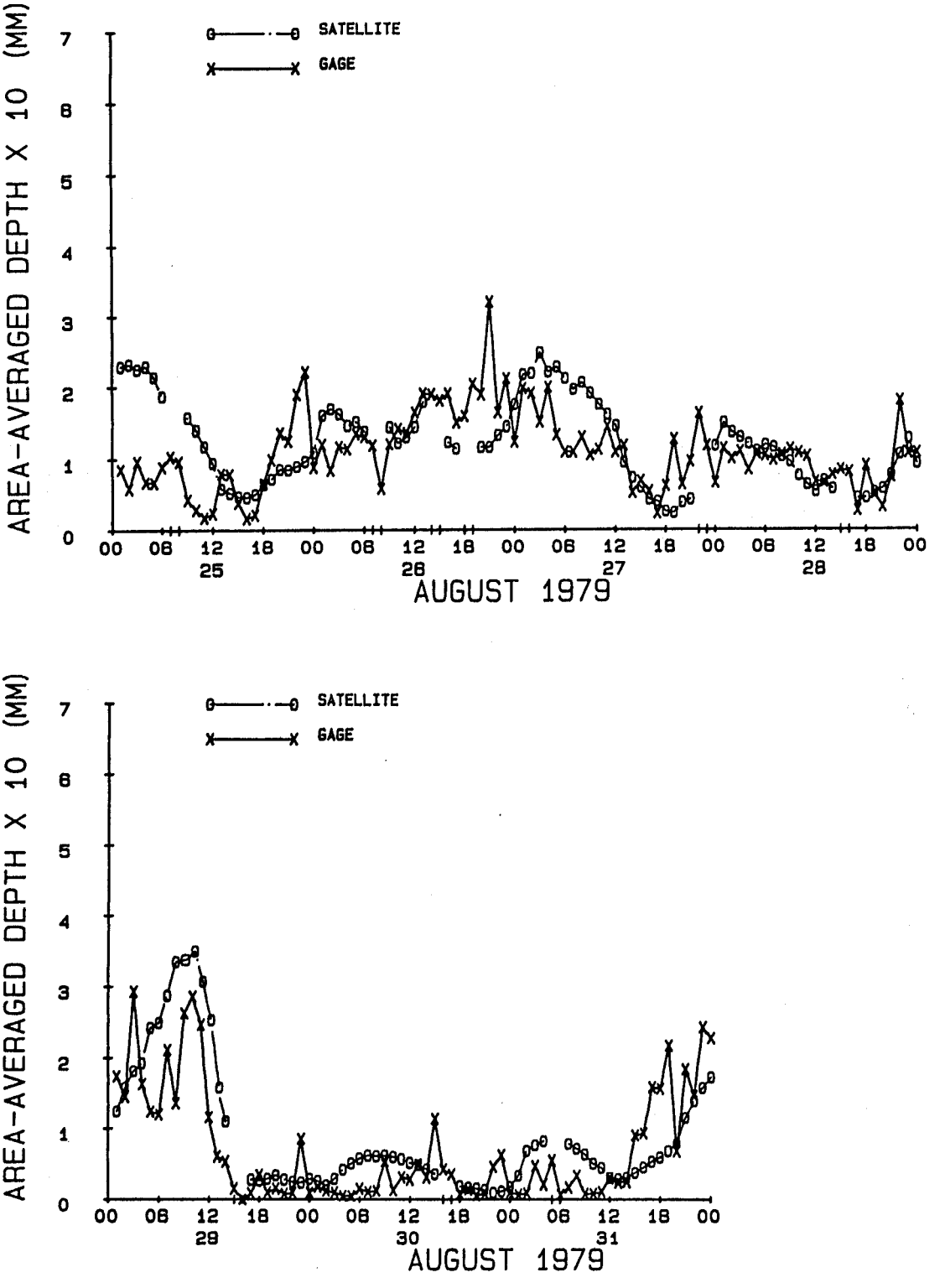


Fig. E.16. Time series of hourly area-averaged adjusted streamlined and gage rainfalls for August 25-28 (top) and 29-31 (bottom), 1979.