COMPARISON OF RAIN ESTIMATES USING TRMM, NEXRAD AND RAIN GAUGES OVER PUERTO RICO

by

Jorge Alberto Briceño Murillo

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

COMPUTER ENGINEERING
(Software Engineering)

UNIVERSITY OF PUERTO RICO
MAYAGÜEZ CAMPUS

2002

Approved by:

_____	_____
Néstor Josué Rodríguez, Ph.D. Member, Graduate Committee	Date
_____	_____
José A. Borges, Ph.D. Member, Graduate Committee	Date
_____	_____
Edgar Acuña Fernández, Ph.D. Member, Graduate Committee	Date
_____	_____
Henrick Mario Ierkic V., Ph.D. President, Graduate Committee	Date
_____	_____
Carlos Pabón, Ph.D. Representative of Graduate Studies	Date
_____	_____
Héctor Monroy Ayala, M.S. Director of the Department	Date
_____	_____
L. Antonio Estévez, Ph.D. Director of Graduate Studies	Date

ABSTRACT

A Rainfall Processing (RAIN_PROC) computational interactive environment to analyze, compare, and visualize data of precipitation is presented. Visualization techniques are employed to display observations of precipitation from the Next Generation Radar (NEXRAD) system, rain gauges, and the Tropical Rainfall Measuring Mission - Precipitation Radar (TRMM-PR) satellite. The algorithms to develop the statistical analysis, the estimation of the Z-R relation using linear regression, to operate, assimilate, and integrate the data were developed. Data corresponding to fall and winter seasons of 1998 and 1999 respectively from NEXRAD and rain gauges are processed using RAIN_PROC. The analyses produce Z-R relationships that are classified by geographic regions of Puerto Rico (Carter and Elsner, 1997). This environment was implemented to run in the UNIX or Windows operation systems using the Interactive Data Language (IDL) numerical and graphical software package and the MINITAB statistical software program.

RESUMEN

Se presenta el ambiente computacional interactivo RAIN_PROC (Procesamiento de lluvias por sus siglas en inglés) para el análisis, comparación y visualización de datos de precipitación. Se emplearon técnicas de visualización para representar las observaciones del sistema NEXRAD, de medidores de lluvia y del satelite TRMM-PR. Se desarrollaron algoritmos para análisis estadísticos, estimación de la relación Z-R usando regresiones lineales, para operar, asimilar e integrar los datos. Se han procesado datos correspondientes al otoño y verano de 1998 y 1999, respectivamente, de NEXRAD y de medidores de lluvia usando RAIN_PROC. El análisis produjo ralaciones Z-R que son clasificadas por regiones geográficas de Puerto Rico (Carter y Elsner, 1997). Este ambiente se implementó para correr en los sistemas operativos UNIX y Windows utilizando el paquete gráfico y númerico IDL, y el programa para análisis estadístico MINITAB.

ACKNOWLEDGEMNTS

I wish to thank to Dr. Steve Goodman (NASA), Israel Matos (National Weather Service in San Juan), Pedro Diaz and Ana V. Sanchez (USGS) for all their support and interest in our research. Thanks to my advisor, Dr. Mario Ierkic who motivated me to continue studies in this area. I also want to thank Dr. M. Petitdidier and Dr. Carl Ulbrich. Thanks to Dr. Edgar Acuña for helping me in the completion of this thesis. This work was supported by NASA's Thunderstorm Observation and Research (ThOR) subcontract agreement NAS 59734.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER I

INTRODUCTION

Rainfall rates R, and reflectivity factor Z relationships for rain and snow, are commonly used to estimate the rate of precipitation. For many years radar meteorologists have attempted to find a useful formula that relates R to the reflectivity factor Z; but, unfortunately, there is no universal relationship connecting these parameters.

There is no doubt that a sufficiently dense network of rain gauges can measure rainfall better than radar. There may be large errors in rain depth at any one gauge representing the areal average; hydrologists have resorted to a network of rain gauges and to radar to improve areal average rainfall estimates. More recent work attempts to measure precipitation from space.

In fact, rain gauge measurements are accepted as the standard against which other measurement techniques are compared. Yet no matter how accurate gauges may be for point measurements (errors are typically 5-10%), their accuracy for areal average is a function of the gauge density and the spatial variability of rainfall (Huff, 1970; Doviak, 1983).

Radar has the advantage of being able to make millions of measurements in minutes and to survey remotely vast areas. The cost of a gauge network to match these capabilities in spatial continuity and to send data to a central location would be

prohibitive. Meteorologists have taken advantage from the best of each technique (the spatial coverage of radar data and the accuracy of gauge data) and they have combined those characteristics to calibrate radar measurements.

Since the parameters of an assumed drop-size distribution vary considerably in space and time from point to point, therefore we may also expect the R, Z relation to vary too. If only reflectivity factor measurements are available, one needs to choose an appropriate R, Z relation. On the other hand, for areal averages of rainfall, there is suggestive evidence that these parameters can be appropriately chosen to produce an R, Z relation that in the mean predicts the average rainfall measured by a network of gauges.

In this work, we expect to give support to satellites and in the future, with reliable radar results, the satellite measurements could be assimilated together with ground radar, rain gauges data and with numerical models. NASA has an operational precipitation radar (PR), which is one of five instruments on the TRMM satellite, and has made a commitment to study precipitation using lightning results. Our work can be of help to calibrate NASA's airborne radar measurements.

This thesis is organized as follows: **Chapter II** is a review of the literature about Doppler radar and its application in meteorological studies. **Chapter III** describes the fundamentals in radar, rain gauges, and TRMM-PR techniques. **Chapter IV** explains the graphical environment for the program used to read, and process the data. The data analysis is described in **Chapter V**. Results and discussion for data collected on this study (several days for the months of September and October of 1998, and for November and December of 1999) is described in **Chapter VI**. Finally, conclusions and recommendations are presented in **Chapter VII**.

CHAPTER II

LITERATURE REVIEW

In order to understand the precipitation, we must know how precipitation forms. Before explaining, we need a couple of facts. First, cloud droplets are very small, averaging under 20 micrometers in diameter (human hair is about 75 micrometers in diameter). Second, clouds consist of many billions of these droplets. Thus, for precipitation to form, millions of cloud droplets must somehow coalesce (join together) into larger drops enough that can fall and that will not evaporate completely before reaching the ground during their descent. There are two mechanisms that explain this phenomenon: the Bergeron process and the collision-coalescence process.

Bergeron Process

"The Bergeron Process is named after its discoverer, the Swedish meteorologist, Tor Bergeron. It relies on two properties of water. First, cloud droplets do not freeze at 0ºC as expected (water suspended in air does not freeze until it reaches a temperature of nearly -40ºC, water in the liquid state below 0ºC is referred to as supercooled). Second, the saturation vapor pressure above ice crystals is lower than above supercooled liquid droplets. With these properties, we can now explain how the Bergeron process produces precipitation. Now, imagine a cloud at -10ºC, where each ice crystal is surrounded by many thousands of liquid droplets.

Because the air was initially saturated (100% relative humidity) with respect to liquid water, it will be supersaturated (over 100%) with respect to the newly formed ice crystals. As a result of this condition, the ice crystals collect more water molecules than they lose by sublimation. Thus, continued evaporation from the liquid drops provides a source of water vapor to feed the growth of ice crystals.

Because the level of supersaturation with respect to ice can be large, the growth of snow crystals is generally sufficiently rapid to generate crystals large enough to fall. During their descent, these crystals enlarge as they intercept cloud drops that freeze on them. Air movement will sometimes break up these crystals, and the fragments will serve as freezing nuclei for other liquid droplets. A chain reaction develops and produces many snow crystals, which, by accretion, will form into larger masses called snowflakes. When the surface temperature is above 4ºC, snowflakes usually melt before they reach the ground and continue their descent as rain (Lutgens and Tarbuck, 1998).

Collition-Coalescence Process

“Before, meteorologists believed that the Bergeron process was responsible for the formation of most precipitation except for light drizzle. Later, it was discovered that copious rainfall is often associated with clouds located well below the freezing level (called warm clouds), especially in the tropics. Clearly a second mechanism also must trigger precipitation. Research has shown that clouds made entirely of liquid droplets must contain droplets larger than 20 micrometers if precipitation is to form. These large droplets form when giant condensation nuclei are present, or when hygroscopic particles exist (such as sea salt).

As the larger droplets fall through a cloud, they collide with the smaller, slower droplets and coalesce. Becoming larger in the process, they fall even more rapidly and increase their chances of collision and rate of growth. After collecting a million or so cloud droplets, they are large enough to fall to the surface without evaporating. Raindrops

can grow to a maximum of 5 millimeters when they fall at the rate of 33 kilometers per hour. At this size and speed, the water's surface tension, which holds the drop together, is surpassed by the friction drag of the air. Friction therefore pulls apart drops larger than 5 millimeters. The resulting breakup of a large raindrop produces numerous small drops that begin anew the task of sweeping up cloud droplets (Lutgens and Tarbuck, 1998)".

Measurements of drop-size distributions have been made around the globe under different climatic conditions. Battan (1973) lists around 69 different R, Z relations (see Figure 2.1), Atlas and Chmela (1957) report considerable variability in the R, Z relations. Their data show that the same reflectivity factor Z could be associated with different values of R. For example, a value of reflectivity factor Z could be associated with either R = 33 mm/h or R = 11 mm/h, a possible 300% error depending on which measured drop-size distribution is used (Doviak and Zrnic, 1993).

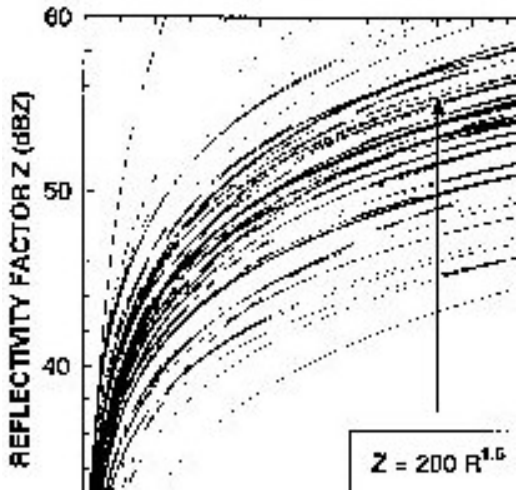

Figure 2.1: Sixty-nine R, Z relationship (adapted from Battan, 1973).

The island of Puerto Rico has a horizontal scale of 100 km. The horizontal scale of summertime rain phenomena is sufficiently large to encompass multiple stations. The average distance of proximal stations is only 10.7 Km and is reasonable to suggest that several stations may share a common rainfall variance signature. With this suggestion Carter and Elsner (1997) identified regions of similar mesoscale rainfall variance (smaller

regions having similar rainfall patterns) employing the variances technique of factor analysis (Myers, 1990). Factors are intangible entities that may not be observed, but represent regions in which several variables are highly correlated with each other, and uncorrelated with other variables. Each of the factors represents a rainfall region, and each of the variables represents a station (Carter and Elsner, 1995). A complete description of the factor analysis regionalization is given in Carter and Elsner (1997). The regionalization using six common factors is shown in Figure 2.2.

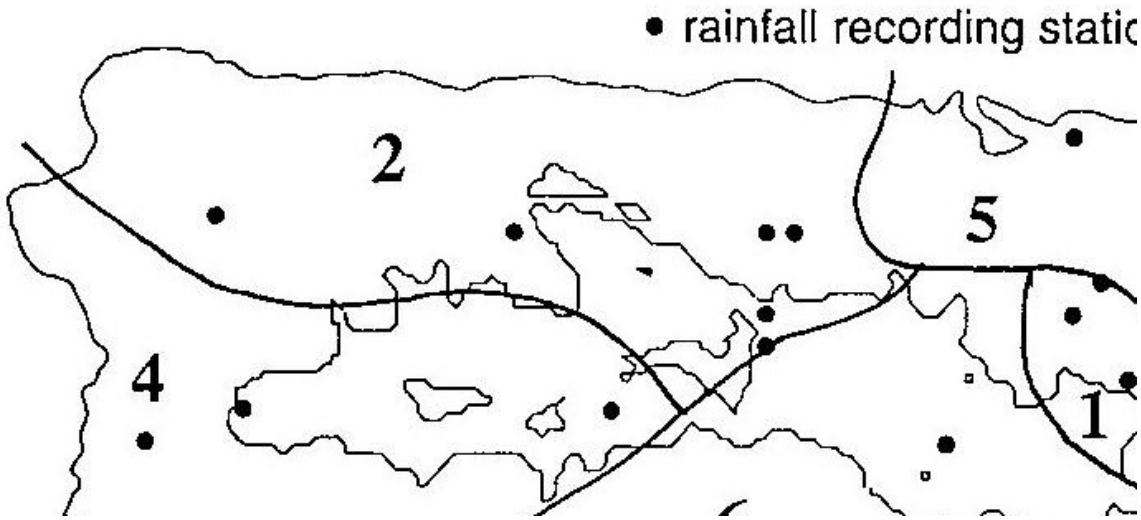

Figure 2.2: Geographic regionalization of Puerto Rico based on a factor analysis of summertime convective rainfall (adapted from Carter and Elsner, 1997).

CHAPTER III

ANALYSIS TECHNIQUE

This chapter describes some fundamental concepts of Doppler radar, rain gauges, and TRMM Precipitation Radar, emphasizing theirs application to the study of precipitation. Section 3.1 gives a general description of radar fundamentals. Fundamental concepts of rain gauges are summarized in section 3.2. Finally, section 3.3 gives a general description of TRMM-PR fundamentals. Every section deals with the data set from each instrument and the structure of the data file format (HDF Format) for each data set.

3.1 Radar Fundamentals

Radar is an electromagnetic system for the location and detection of reflecting objects such as aircraft, ships, spacecraft, people, vehicles, and the natural environment. Radar operates by radiating electromagnetic energy into space and detecting the echo signal reflected from a target, or object.

The reflected energy that is returned to the radar indicates the presence of a target and, by comparing the signal that was transmitted with the received echo signal, its location can be determined along with other target-related information.

3.1.1 Principles of Meteorological Doppler Radar

Scattering

Scattering is the re-radiation of incident E-M energy by targets in many different directions. Backscatter is that portion of the radar's transmitted energy that is intercepted by a target and scattered back toward the radar (see Figure 3.1). The degree of backscatter from a target is associated with the size (diameter), state (liquid, ice, or a mixture), shape (round, oblate, flat, etc.), and concentration (number per unit volume) of the target.

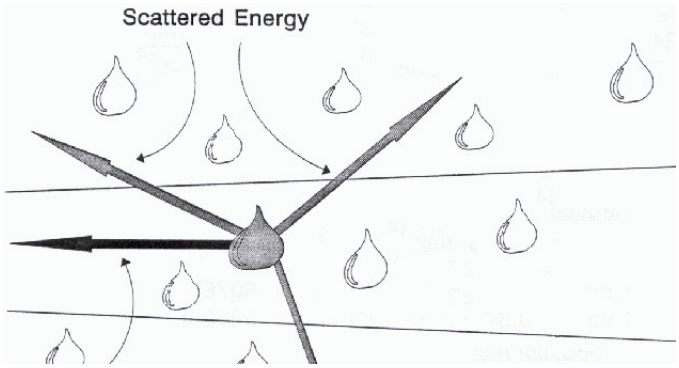

Figure 3.1: Scattering.

Rayleigh and Non-Rayleigh scattering

Rayleigh scattering is the scattering of particles whose diameter (D) is small compared to the wavelength (λ) of the incident radiation ($D < 0.3\lambda$; ~ 3cm for the WSR-88D). The Rayleigh scattering it is assumed that has approximately a backscatter cross section of D^6, droplets are assumed spherical by the radar equation, and the maximum observed drop diameter seldom exceed 7 mm (0.7 cm), so all drops are Rayleigh scatterers. Large hailstones ($D > 3.31$ cm) are non-Rayleigh scatters for the WSR-88D with λ of 10 – 11 cm. Big diameter particles are highly reflective, but actually generate more scatter away from the radar than towards it.

3.1.2 Precipitation Estimation

Reflectivity Factor Z

Radar meteorologists need to relate the reflectivity Z, which is general radar terminology for the backscattering cross section per unit volume, to factors that have meteorological significance. If water drops are spherical and have diameters that are small compared to the wavelength, we can obtain the reflectivity factor of water spheres with the following formula:

$$Z = \int_0^\infty N(D) D^6 dD \tag{3.1}$$

where,

Z = reflectivity factor

D = drop diameter

N(D) = number of drops for a given diameter per cubic meter

So, Z is defined as the sixth power of the hydrometeor diameter summed over all hydrometeors in a unit volume in term of the hydrometeor size distribution.

Because values of Z commonly encountered in weather observations span many orders of magnitude, radar meteorologists use a logarithmic scale $10\log_{10}Z$, where Z has units of mm^6/m^3 and the scale $Z(dBZ)=10\log_{10}Z$ is in dBZ units. Precipitation produces Z ranging from near 0 dBZ in cumulus congestus clouds to values somewhat larger than 60 dBZ in regions of heavy rainfall; and hail. For example, in one cubic meter,

729 drops with D = 1mm

1 drop with D = 3mm

$Z = (729 \text{ drops/m}^3)(1\text{mm})^6 + (1 \text{ drop/m}^3)(3\text{mm})^6 = 1458 \text{ mm}^6/\text{m}^3$

for $Z = 1458 \text{ mm}^6/\text{m}^3$, dBZ = 32

The WSR-88D can detect reflectivity as low as –32 dBZ near the radar site in "clear air" VCPs, and as high as 95 dBZ. If $0mm^6/m^3 < Z < 1mm^6/m^3$, then the dBZ value will be negative and if $Z = 109mm^6/m^3$, then dBZ = 90.

Estimation of Z

A direct measurement of Z is not possible, and Z is estimated via the following radar equation:

$$Ze = \frac{\Pr R^2}{C} \tag{3.2}$$

where:

Pr = returned power

R = range

C = radar constant

Some key points affecting the reflectivity factor Z:

- Z is dependent on dropsize distribution.
- Radars cannot measure drop-size distribution.
- Radars estimate Z from returned power.
- Different drop-size distribution can generate identical Z's, affecting accuracy of rainfall estimation.

These factors can introduce important errors in the Z estimation:

- Ground clutter (overestimate or underestimate depending on level of clutter suppression).
- Anomalous propagation (overestimate in super-refractive conditions).
- Partial beam filling (underestimate at long range with expanded areal coverage).
- Wet randome (underestimate if R+, ZR/ZL onsite).
- Incorrect hardware calibration (WSR-88D is self calibrating).

Rainfall Rate R

When we know the drop diameter spectrum, we can directly compute the rainfall rate in still air using the following formula:

$$R = (\pi / 6)\int_0^\infty N(D)D^3 W_t(D)dD \tag{3.3}$$

where,

R = rainfall rate
D = drop diameter
N(D) = number of drops for a given diameter per cubic meter
Wt(D) = fall velocity for a given diameter

The rainfall rate R is usually measured as depth of water per unit time. For example, in one cubic meter,

600 drops with D = 1mm
fall velocity, Wt(D) = 4m/s
$R = \pi/6\ (600\ \text{drops/m}^3)(1\text{mm})^3\ (4\text{m/s})$

After converting units, R = 0.18 inch/hr

R, Z Relations for Rain

R, Z relationships for rain are commonly used to estimate the rate of precipitation. For many years radar meteorologists have attempted to find a useful formula that relates R to reflectivity factor Z. Unfortunately, there is no universal relationship connecting these parameters. Measurements of drop-size distribution around the globe under different climatic conditions have been made, and Battan (1973) lists no fewer than 69 different R, Z relations. Here we show the two most frequently used Z-R relationships, namely: the default or WSR-88D convective relationship (for

summer deep convective), and the Rosenfeld Tropical relationship (for tropical convective systems). These relations are shown in the following equations:

Default or WSR-88D convective relationship:

$$Z = 300R^{1.4} \tag{3.4}$$

Rosenfeld Tropical relationship:

$$Z = 250R^{1.2} \tag{3.5}$$

Z and R do not have a unique relationship: for a given R, many Z's are possible. The reasons being that there might be different raindrop distribution sizes, different physical states of water (water vapor, liquid water, ice), different spatial in homogeneities and temporal variability.

Advantages and Disadvantages of Precipitation Estimation Methods

Radar data only

Advantages:

- Areal distribution of precipitation is continuous.
- Network will offer nearly 100% coverage.

Disadvantages:

- Estimates are not ground based.
- Precipitation may be missed between scans.
- Quality control is required.

Radar estimates most favorable when:

- Error sources are minimized.
- Gauge network is sparse.

Rain Gauge data only

Advantages:

- Sampling is continuous at each point.
- Estimate is ground based.

Disadvantages:

- Points can miss significant precipitation.
- Real time retrieval is insufficient.
- Many different gauge types are in the field.

Rain gauge estimates most favorable when:

- Radar error sources are significant.
- Precipitation areal coverage is extensive.

Combination of Radar and Rain Gauge Data

Advantages:

- Strengths of both can be merged.
- Data is updated hourly.

Disadvantage:

- Significant quality control is required.
- Merging data is a difficult task.

3.1.3 Signal Processing and Database Generation

Doppler Effect

Doppler effect is the change in frequency with which energy reaches a receiver when the receiver and the energy source are in relative motion. WSR-88D receives from targets a very small portion of the energy transmitted towards those targets. Radar location is fixed, so relative motion can occur only if target location changes. Frequency change is proportional to target motion. However, the WSR-88D does not

measure frequency changes. For example, for the WSR-88D, where the source (target) is moving 50 knots towards or away from a receiver:

Frequency of transmitted wave = 2850 MHz (2,850,000,000 Hz)
Frequency shift due to motion = +/- 487 Hz

The Doppler shift is about .00002 % of the original wave frequency. The WSR-88D is not designed to measure these minute frequency changes.

WSR-88D Velocity detection method

WSR-88D is designed to detect the phase of the returned signal and compute pulse-to-pulse phase changes (phase is a particular angular stage or point of advancement in a cycle. It is the fractional part of the angular period through which a wave has advanced, measured from the phase reference). This technique is known as Pulse-Pair Processing. The WSR-88D can only detect motion towards or away from it. WSR-88D uses the pulse-to-pulse phase change of the returned signal to obtain velocity information. Radial velocity is equal to the actual target velocity when target motion is directly towards or away from the radar. If the target is stationary, or is moving perpendicular to the radar beam, its radial velocity is equal to zero (0). The Mathematical expression for the radial velocity is:

$$|Vr| = |Va| x Cos\beta \tag{3.6}$$

where:

Vr = radial velocity
Va = actual velocity
Cos β = cosine of the smaller of two angles between Va and radar radial.

The WSR-88D estimates the magnitude of radial velocity by computing the mean pulse-to-pulse phase change. Returned signal is processed into two components, I (in-

phase) and Q (quadrature), in order to unambiguously determine the phase. Both I and Q components of the signal are needed to determine the direction of motion and the radial velocity. Radial velocity (Vr) will never be greater than the true velocity. Vr will be a maximum when the target is moving parallel to the radar radial. Target speed |Vr| can be computed if Vmax and pulse-to-pulse phase shift are known.

$$\frac{|Vr|}{V\max} = \frac{pulse - pair\_phase\_shift}{180^{\circ}} \tag{3.7}$$

Returned power (Pr) estimate

Radar cannot determine drop-size distribution, so it cannot directly compute Z and dBZ. However, the radar can measure returned power (Pr). Returned power is measured directly from the returned signal by squaring the mean amplitude.

$$P = I^2 + Q^2 \tag{3.8}$$

If Pr is known, Ze (equivalent reflectivity factor) can be derived via the radar equation. dBZe, a range normalized field, is displayed by the WSR-88D. The Average Pr from 4 successive 250 meters range gates is used to derive dBZe at its best available range resolution, 1 km.

$$\mathrm{Pr}_{(0.54nmi)} = \frac{\mathrm{Pr}_1 + \mathrm{Pr}_2 + \mathrm{Pr}_3 + \mathrm{Pr}_4}{4} \tag{3.9}$$

Since drop-size distribution is unknown, radar uses power estimate to solve for Ze. Ze is converted to dBZ by:

$$dBZ = 10\log Ze \tag{3.10}$$

Spectrum width estimate

Spectrum width is a measure of velocity dispersion within the radar sample volume. Spectrum width is generated by statistical auto-correlation. Statistical auto-correlation measures the similarity of the returned signal from successive pulses. The method assumes that the Doppler power has a bell shaped dependence with frequency.

Features/conditions typically associated with relatively high spectrum widths include: boundaries (e.g. weather fronts), thunderstorms, shear regions, and regions where side-lobe returns mixed with main-lobe returns. Contributors to spectrum width are turbulence, wind shear, antenna motion, different fall speeds for different sized hydro-meteors, and internal WSR-88D noise.

In summary, the WSR-88D generates the following database:

- Base Reflectivity data derived from returned signal's power.
- Base Velocity data derived via pulse-pair processing.
- Spectrum Width data derived via statistical auto-correlation.

3.1.4 Level II Data Set from WSR-88D or NEXRAD

Weather **S**urveillance **R**adar - 19**88** **D**oppler (WSR-88D), or **NEX**t Generation **RAD**ar (NEXRAD), Level II data are the base digital data produced by the signal processor (mean radial velocity, reflectivity, and spectrum width) at the full spatial and temporal resolution of the radar. The Level II data is recorded in 8mm tapes. Each tape records approximately 10 hours of Volume Coverage Pattern 11 (VCP 11, see Appendix E), 18 hours of VCP 21, or 40 hours of VCP 31 or 32 using the EXABYTE 8200 mode. Using the EXABYTE 8500 mode doubles, both the storage capacity and number of hours of data possible per 8mm tape. For more information of Level II data refer to http://www.ncdc.noaa.gov/ol/radar/leveliidoc.html site.

The analysis will be done using the software WSR-88D Algorithm Testing and Display System (WATADS). This software requires rawinsonde data (see Appendix A) to establish boundary conditions. Here I only present a graph with precipitation results using the standard Z—R relation, namely $Z = 300R^{1.4}$. WATADS is also capable of providing various other products including Velocity Azimuth Displays (VAD), Radial Velocity, Spectrum Width, and Reflectivity (dBZ).

The Figure 3.2 displays total accumulated precipitation for one-hour time interval. The precipitation image is created using outputs from WATADS.

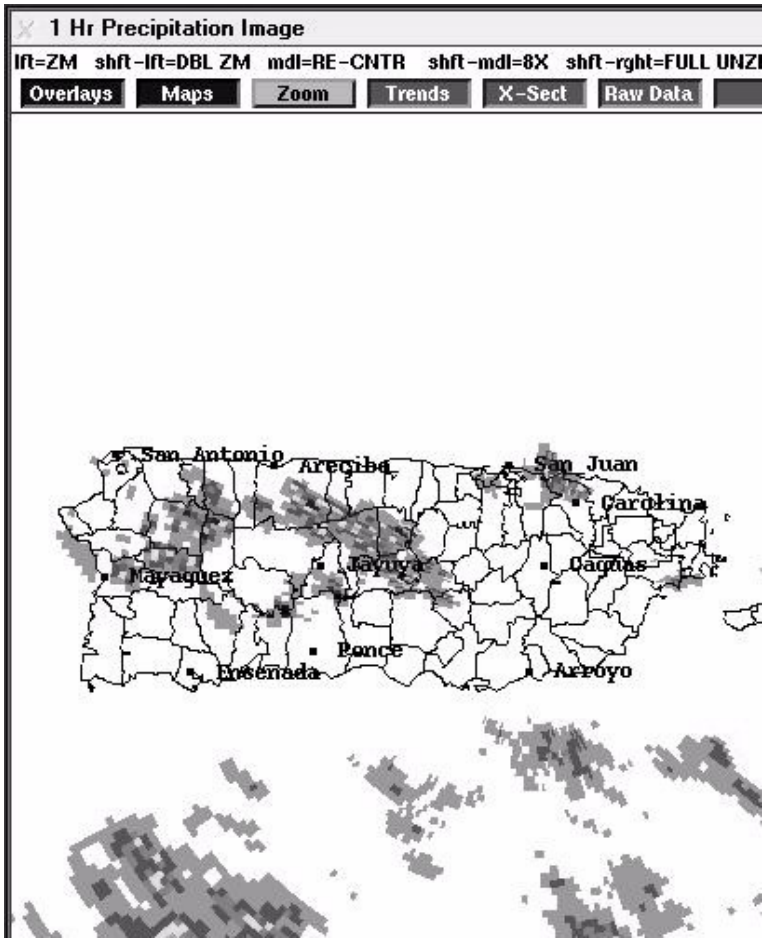

Figure 3.2: Precipitation accumulation for one-hour time interval (from WATADS)

3.1.5 WSR-88D Algorithm Testing And Display System (WATADS 10.0)

WATADS is a product of the National Severe Storms Laboratory (NSSL) and consists of many special meteorological computer programs to process, analyze, and display WSR-88D data. It was designed to run on SUN and Hewlett-Packard Unix workstations; so, WATADS utilizes the X11 windowing system. WATADS reads WSR-88D Level II data from 8mm data tapes, and process it with programs that create special disk files for use by NSSL and National Weather Service meteorological algorithms. These special computer algorithms detect, predict and analyze severe weather events. Users may adjust various algorithm parameters, to focus on regional weather conditions, in successive algorithm runs using WATADS.

WATADS has the following components:

- WSR-88D Level II Data Processing. WATADS extracts base data from 8 mm radar data tapes (in level II format) and converts them to RADS format, which produces high-resolution images.
- Setting of Algorithm Parameters. WATADS allows a friendly point-and-click interface, to change NSSL and NWS algorithm, parameters.
- WSR-88D Meteorological Algorithms. Algorithms from the National Weather Service are included in WATADS.
- NSSL Meteorological Algorithms. These algorithms are a complex suite of meteorological algorithms, created and maintained at NSSL, which processes WSR-88D data.
- Radar Analysis and Display System (RADS). RADS is the display and user interface portion of WATADS, used for visualization of data, products, trends, and derived images.

For more information about the WATADS components refer to the “WSR-88D Algorithm Testing and Display System Reference Guide.

Using WATADS on New Data

These processes must be completed for a new data case:

- Information gathering.
- Set system environmental parameters.
- Create data directory for this case or tape.
- Set algorithm parameters (including sounding data).
- Index data tape (optional).
- Process data from Level II tape to disk.
- Run NSSL and NWS meteorological algorithms on data to create overlay products, trend information and derived images.
- Run RADS to view base data, derived images, and products.

Each of these processes is discussed in detail in WATADS reference guide and WSR-88D Algorithm Testing and Display System (WATADS), Version 10.0, Algorithm Reference Guide.

The Hybrid Scan Reflectivity file is an intermediate file produced by the WATADS Precipitation Processing Subsystem. The file should be generated for each volume scan of radar data. Each file will be titled HYB_YYMMDD_HHMM where YYMMDD is the year, month, and day and HHMM is the hours and minutes since midnight (UTC). The Hybrid files are each 166,024 bytes. The first 400 bytes contain the date/time and the adaptation data. The next 165,600 bytes contain the Hybrid Scan Reflectivity information in coded dBZ (360 degree by 230 kilometers by 2). The final 24 bytes are trailing information, the first 8 bytes contain the date and time information (4 bytes each). The date is a Julian date based on the number of days since 1/1/70. The time is the number of seconds since midnight (UTC).

The coded dBZ is a value from 0 to 255. Zero (0) is used for data below threshold (SNR of approximately -5dB). One (1) is reserved for range ambiguous data and is not used in the reflectivity domain. Therefor the valid coded dBZ values that can be converted to dBZ range from 2 to 255. We used the following formula to convert from coded dBZ to dBZ:

$$dBZ = (coded\_dBZ / 2) - 33 \tag{3.11}$$

Therefore the coded dBZ is half dBZ increments and ranges from -32 dBZ to +94.5 dBZ.

For the purpose of this study, the data was recorded in HDF format, the file name has the syntax HYB.YYYYMMDD.1.NRPR where YYYYMMDD is year/month/day, the number 1 is for version one (this was made with the IDL program NR_FORMAT_ALL.PRO). Figure 3.3 shows the structure for a Hybrid Scan Reflectivity data file in terms of the component objects, in this case for the date 1998/09/30.

Where:

Version: The version of the structure. This version is 1.0.0b.
Date: Date when the file was created.
Author: Author's name of this data structure.
Email address: My e-mail address.
Time: The time when the data was recorded.
50010500, 50010800, etc: The recorded data for every station.

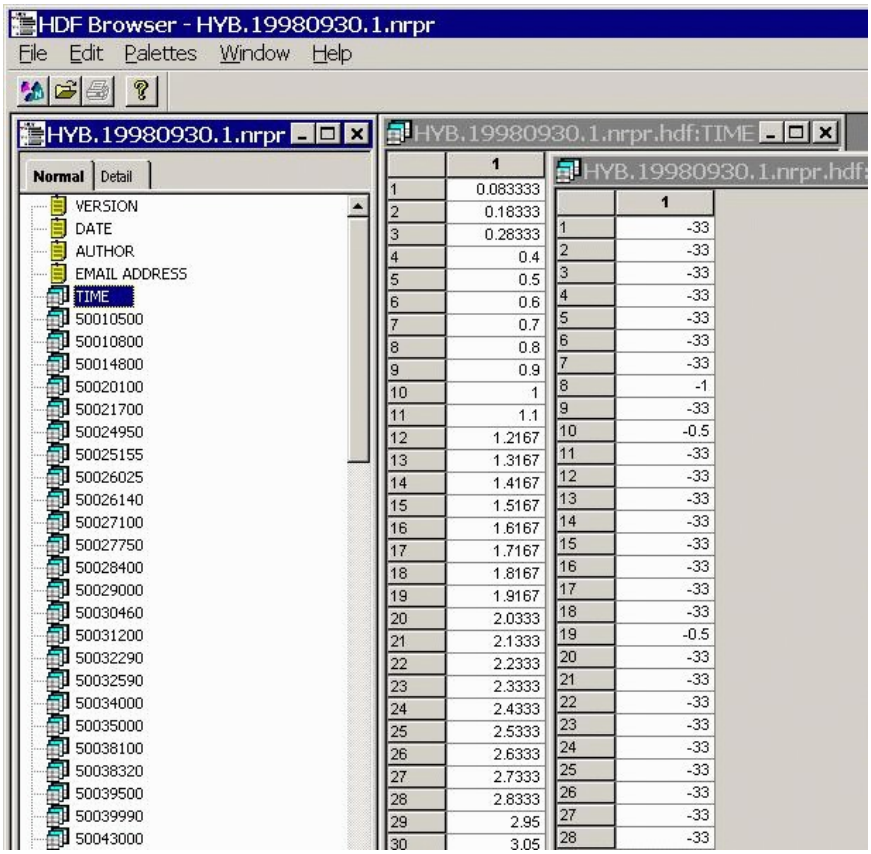

Figure 3.3: File structure for a Hybrid Scan Reflectivity data file.

WATADS 10.0 System Requirements

SUN Workstation

Solaris 1 or 2, color monitor, 8mm tape drive, at least 32 MB of RAM, and at least 500 MB free hard drive space.

HP Workstation

HP-UX 10.2 or better, color monitor, 8mm tape drive, at least 32 MB of RAM, and at least 500 MB free hard drive space.

Support contact information

e-mail: watads@nsslsun.nssl.noaa.gov

Web Page: http://www.nssl.noaa.gov/~watads

WATADS installation instructions are available in the Appendix E of the "WATADS Reference Guide".

3.2 Rain Gauges Fundamentals

The desired estimator for rainfall amounts is an area average estimate, rather then a single-point observation. Over anything other than a very small area, several gauges are usually necessary to obtain a reasonable estimate. While some specialised investigations may use a high density of gauges, even the best meteorological network rarely has more than one gauge for every 25 km^2. With a funnel diameter of 12.7 cm (5 inches) the area is 126.7 cm^2. This means that the area in which rain is collected is around 10^{-9} of the area for which this rainfall is to be an areal estimated. In most parts of the world, the rain gauge density is much smaller than this, and over the seas almost non-existent. While some argue that in a situation like this it is not worth measuring rain very carefully, the contrary is true, for without accurate point measurements, the spatial variability of rainfall cannot be established, because it will be swamped by rain gauges error. Also, a rain gauge with a systematic error will always bias the areal estimates. It is very important, therefore, to try to get the point reading as accurate as possible, through a careful choice of the site, carefully positioning the gauge at the chosen site and through the use of the best rain gauge possible.

The total precipitation is the sum of all the liquid collected, expressed as the depth it would cover in a flat surface assuming no losses due to evaporation. While inches have been used in the past to measure precipitation, and still are in some countries, millimeters

are the more usual unit today (0.01 inches is about equivalent to 0.25 millimeters) (Strangeways, 2000).

3.2.1 Rain Gauges Stations

Precipitation is estimated with 108 gauges distributed throughout Puerto Rico and using satellite links. The USGS instruments geographic distribution is shown in the Figure 3.4. The rain gauges spatial coverage is reasonably dense.

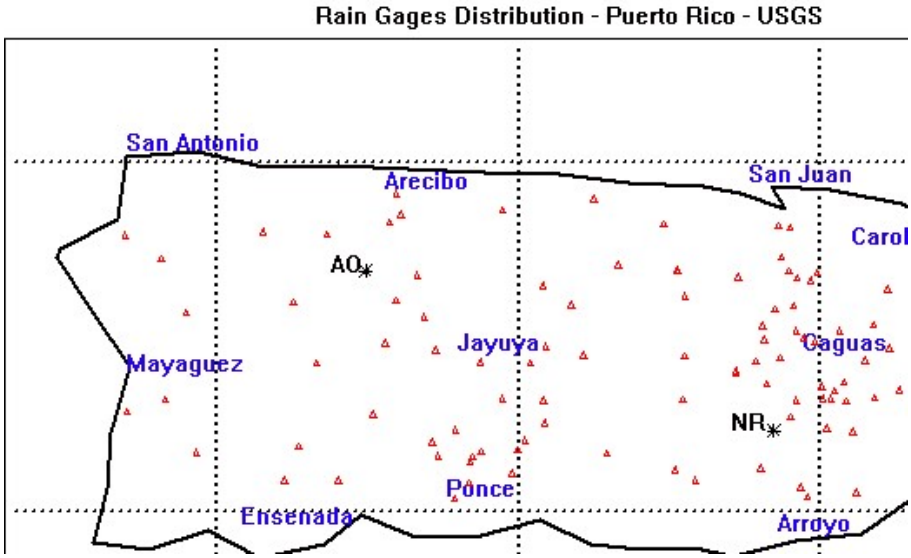

Figure 3.4: Rain gauges distributed throughout Puerto Rico, from USGS.

3.2.2 Data Set from Rain Gauges Stations

The Figure 3.5 represent observations of instantaneous precipitation measured by the rain gauge located at GUAJATACA river at LARES, PR for September 1998. The integration time for these estimates is in general five minutes.

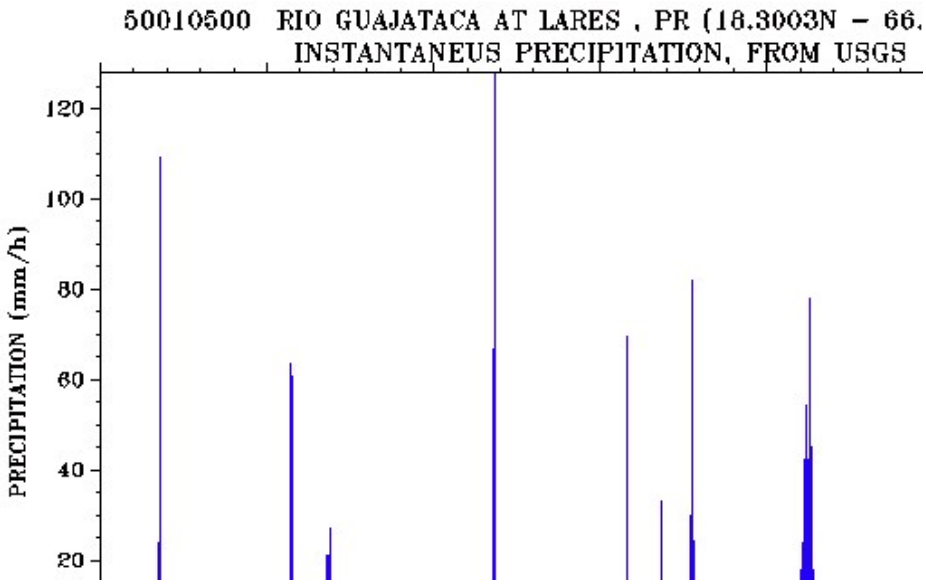

Figure 3.5: Instantaneous precipitation, data from USGS.

Originally, the rain gauge data was recorded in an Excel file format for every rain gauge station. But now, for the purpose of this study, I recorded the data en HDF format, the file name has the syntax ID.YEAR.1.RGPR where. Figure 3.6 shows the structure for a rain gauge data file in terms of the component objects where:

Version: The version of the structure. This version is 1.0.0b.

Date: Date when the file was created.

Author: Author's name of this data structure.

Email address: My e-mail address.

Gauge Id: The Id number for a specific rain gauge station, in this case 50010500.

Gauge Name: The name of the station, in this case Guajata river at Lares.

Gauge Latitude: The latitude position of the station.

Gauge Longitude: The longitude position of the station.

Sep, Oct: Time and data values for September and October months. The first column is the observation time in hour fraction and the second column is the observed values.

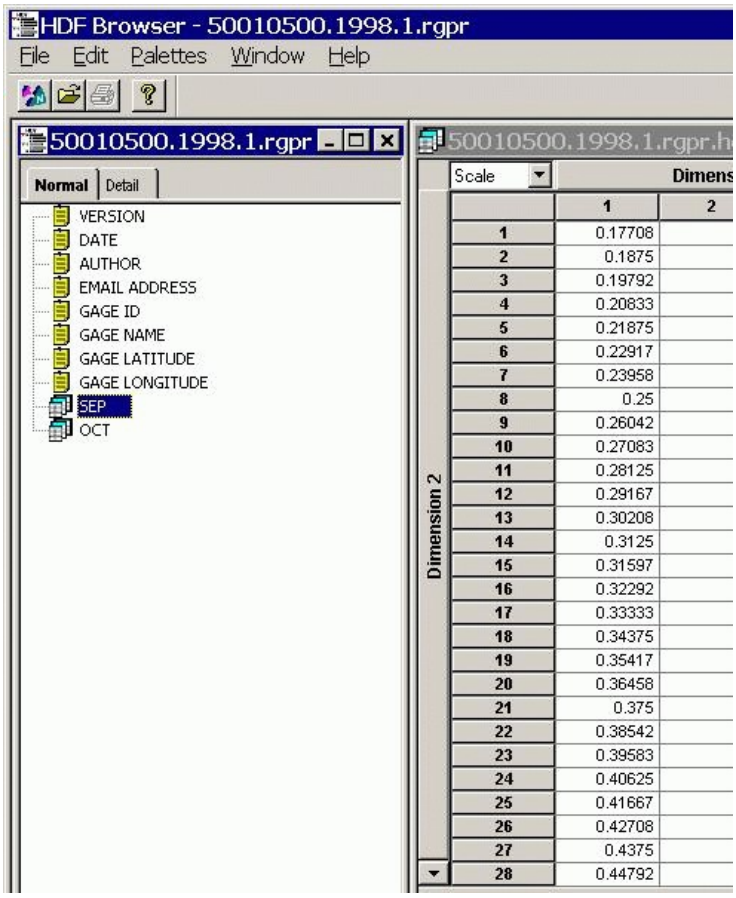

Figure 3.6: File structure for a rain gauge data file.

3.3 Tropical Rainfall Measuring Mission - Precipitation Radar (TRMM – PR) Fundamentals

3.3.1 TRMM – PR

The Tropical Rainfall Measuring Mission (TRMM) is an integral part of the National Aeronautics and Space Administration (NASA) Earth Science Enterprise. The TRMM observatory was launched in 1997, into a near circular orbit of approximately 350 kilometers altitude, and a period of 91.5 minutes (15.7 orbits per day). Figure 3.7 shows the diagram of the TRMM spacecraft with all its sensors, and Table 3.1 shows the TRMM sensor summary.

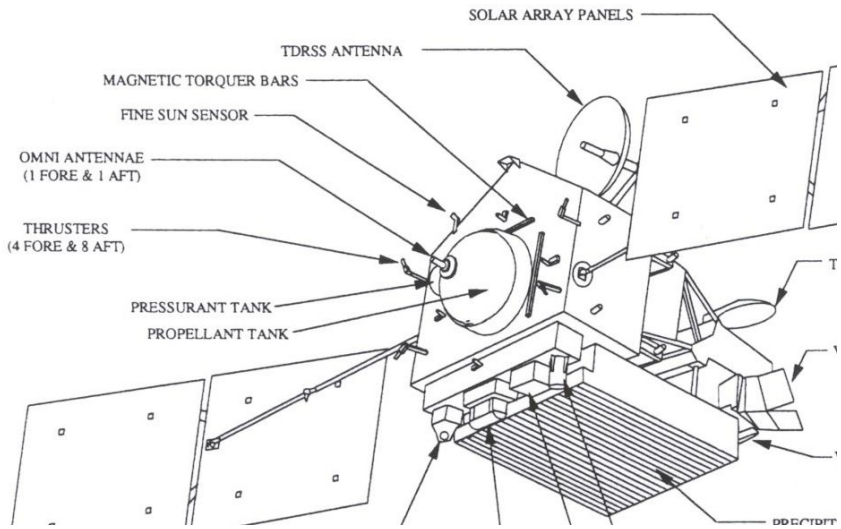

Figure 3.7: TRMM Observatory Configuration (Simpson, 1996)

The TRMM Science Data and Information System (TSDIS) is the data processing and science information system for TRMM. The Earth Observation System Data and Information System (EOSDIS) serves as the long-term archive for all TRMM data products.

Table 3.1: TRMM Sensor Summary - Rain Package.

Microwave radiometer (TMI)	Precipitation Radar (PR)	Visible/Infrared radiometer (VIRS)
10, 19, 21, 37, 85.5 GHz (dual polarized) 21 km resolution 760 km swath	13.8 GHz 4 km footprint 250 m range resolution 220 km swath	0.63 μm, and 10 μm also 1.6, 3.75, and 12 μm 2.2 km resolution 750 km swath
Additional (eos) instruments: one ceres (cloud and earth radiant system) and one lis (lightning imaging sensor), 35º inclination, 350 km altitude, high resolution, rapid precession, extensive validation by 10-12 surface radar sites required		

3.3.2 Instrument and Scan Geometry

The Precipitation Radar (PR), one of five instruments on the TRMM satellite and the first in space, will define the layer depth of the precipitation and provide information about the rainfall reaching the surface, a key parameter to determine the latent heat input to the atmosphere. A unique feature is that it will permit the measurement of rain over land where passive microwave channels have more difficulty. The PR is an electronically scanning radar operating at 13.8 GHz with horizontal polarization using a 128-slotted waveguide antenna and solid power amplifiers to develop an active phased array. The horizontal resolution is 4.3 km at nadir, the range resolution is 250 m and the scanning swath width is 220 km.

The Figure 3.8 depicts reflectivity (near surface Z) measured by the TRMM satellite over the Caribbean Region for November 16, 1999. Unfortunately, we do not have satellite coverage over Puerto Rico coinciding with the study dates at all time.

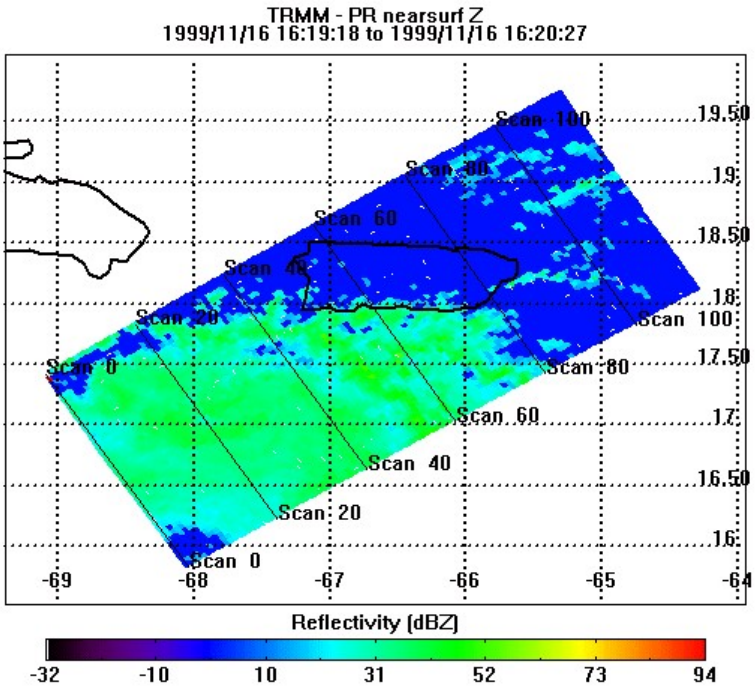

Figure 3.8: Reflectivity measured by the TRMM satellite.

For more information about the TRMM-PR system refer to the Tropical Rainfall Measuring Mission (TRMM), Science and Information System, Real time System and File Specifications.

3.3.3 HDF Format

Hierarchical Data Format

HDF is a multi-object file format for the transfer of numerical and graphical data between heterogeneous machines. HDF Format was developed by the National Center for Supercomputing Applications (NCSA) at the University of Illinois at Urbana-Champaign and is the selection for the common data format of EOSDIS.

HDF format provides the following advantages to algorithm designers and data users:

- **Data abstraction:** The user does not have to be concerned with the physical layout of the file in order to either read or write an HDF file.
- **Extensible file format:** HDF currently supports six different data models and new data models will be incorporated in the future. HDF can easily accommodate new data models from HDF users.
- **Portable file format:** HDF file can be created on one system (e.g., WINDOWS) to be used on a different system (e.g., UNIX).
- **Self-describing:** HDF permits the inclusion of a data index (metadata) inside the data file, which describes the contents of the file. This simplifies data sharing among diverse users.
- **Versatile file format:** HDF support six different data models (8-bit raster model, palette model, 24-bit raster model, scientific data (SDS) model, annotation model, and the virtual data (Vdata) model).

HDF manual and software may be obtained via anonymous ftp at ftp.ncsa.uiuc.edu

Relevant HDF Data Models

The EOSDIS data structures are composed of V groups and simple HDF models that are well tested and commonly accepted within the science community. A V group is like a node in a tree structure. The members of a V group can be data models or more V groups. The TSDIS file specifications use the following models:

1. Scientific Data Set – Is an HDF structure that stores rectangular gridded arrays of data with information about the data. An SDS can contain data that are multi-dimensional, but must be of a singular data type and composed of the following elements:

    - The actual data values.
    - The number of dimensions (rank), and the size of each dimension.
    - The type of the data (e.g., floating point or integer).
    - A fill value for representing missing data in a data set (optional).
    - A range, attributing maximum and minimum values for the data set (optional).
    - Calibration information including an offset and scale factor (optional).
    - Format specifications to be used when displaying values for the dimensions and for the data (optional).
    - Labels for all dimensions and for the data (optional).
    - Scales to be used along the different dimensions when interpreting or displaying the data (optional).
    - The coordinate system to be used when interpreting or displaying the data (optional).
    - Units for all dimensions and for the data (optional).

2. VData – This model provides a framework for storing related data in a table-like structure in HDF files. Each table is comprised of a collection of similar records called VData records, whose values are stored in fixed-length fields. Each field can support its own data types, however, every VData record in a table must contain the same types of fields.

3.3.4 2A-25 PR Profile Data Set from TRMM and Data Structure

Originally, one file of 2A-25 (Level 2A, derived physical parameters in full resolution) PR Profile contained vertical rainfall rate profile for one orbit. I reduced the file to contain only rainfall data over Puerto Rico. Figure 3.9 shows the structure of the 2A-25 new product version in terms of the component objects.

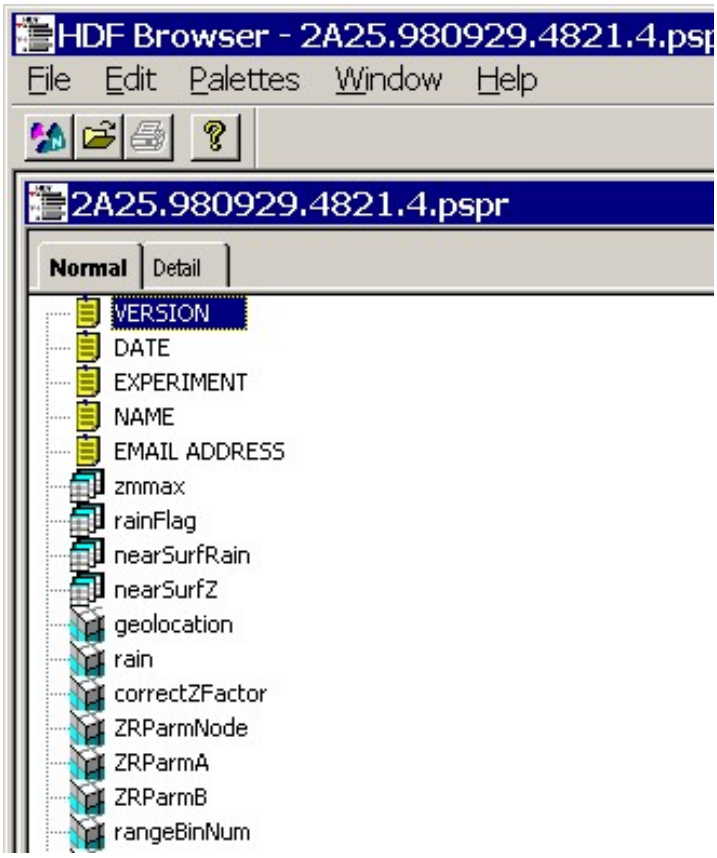

Figure 3.9: File structure for a TRMM data file

Version: The version of the structure. This version is 1.0.0b.

Date: Date when the file was created.

Experiment: In this case is 2A25.

Name: Author of this structure data.

Email address: Author's e-mail address.

Zmmax (SDS, array size 49 x 116, 4-byte float): This is the maximum value of measured reflectivity factor at each IFOV. It ranges from 0.0 to 100.0 dBZ.

RainFlag (SDS, array size 49 x 116, 2-byte integer): The Rain Flag indicates rain or no rain status and the rain type assumed in rain rate retrieval. The default value is 0 (no rain). Bit 0 is the least significant bit. The following meanings are assigned to each bit in the 16-bit integer if the bit = 1.

Bit 0 rain possible
Bit 1 rain certain
Bit 2 zeta^beta > 0.5
Bit 3 large attenuation (PIA larger than 10 dB)
Bit 4 stratiform
Bit 5 convective
Bit 6 bright band exists
Bit 7 warm rain
Bit 8 rain bottom above 2 km
Bit 9 rain bottom above 4 km
Bit 14 data missing between rain top and bottom
Bit 10, 11, 12, 13, 15 not used.

NearSurtRain (SDS, array size 49 x 116, 4-byte float): Rainfall rate near the surface. The range is 0 to 3000 mm/hr.

NearSurfZ (SDS, array size 49 x 116, 4-byte float): Reflectivity near the surface. The range is 0.0 to 100.0 dBZ.

Geolocation: (SDS, array size 2 x 49 x 116, 4-byte float): The first dimension is North latitude and West longitude, in that order. The next dimensions are ray and scan.

Rain (SDS, array size 80 x 49 x 116, 2-byte integer): This is the estimate of rain rate at the radar range gates from 0 to 20 km. It ranges from 0.0 to 300.0 mm/hr and is multiplied by 100 and stored as a 2-byte integer. A value of –88.88 mm/hr, it means ground clutter (stored as –889).

CorrectZFactor (SDS, array size 80 x 49 x 116, 2-byte integer): This is the attenuation corrected Z-factor at the radar range gates from 0 to 20 km. It ranges from 0.1 to 80.0 dB of mm^6 m^{-3} and is multiplied by 100 and stored as a 2-byte integer. A value of –88.88 dB (stored as –889), it means ground clutter. A value of –77.77 dB (stored as –778) means Z was less than 0 dBZ.

ZRParmNode (SDS, array size 80 x 49 x 116, 2-byte integer): The Z-R Parameter Node gives the range bin numbers of the nodes at which the Z-R parameters "a" and "b" are given (see below). The values of a and b between the nodes are linearly interpolated. This variable ranges from 0 and 79 and is unitless.

ZRParmA (SDS, array size 80 x 49 x 116, 2-byte integer): Parameter a for Z-R relationship ($R=aZ^b$). a is computed at 80 radar range gates for each ray. It ranges from 0.0050 to 0.2000 and is multiplied by 10^4 and stored as a 2-byte integer.

ZRParmB (SDS, array size 80 x 49 x 116, 2-byte integer): Parameter b for Z-R relationship ($R=aZ^b$). b is computed at 80 radar range gates for each ray. It ranges from 0.500 to 1.000 and is multiplied by 10^3 and stored as a 2-byte integer.

RangeBinNum (SDS, array size 6 x 49 x 116, 2-byte integer): This array gives the Range Bin Number of various quantities for each ray in every scan. The definitions are:

- top range bin number of the interval that is processed as meaningful data in 2A-25
- bottom range bin number of the interval that is processed as meaningful data in 2A-25
- actual surface range bin number
- range bin number of the bright band if it exits
- range bin number at which the path-integrated Z-factor first exceeds the given threshold
- range bin number at which the measured Z-factor is maximum

RainAve (SDS, array size 2 x 49 x 116, 2-byte integer): There are two kinds of Average Rain Rate. The first one is the average rain rate for each ray between the two predefined heights of 2 and 4 km. It ranges from 0.0 to 3000. mm/hr and is multiplied by 10 and stored as a 2-byte integer. The second one is the integral of rain rate from

rain top to rain bottom. It ranges from 0.0 to 3000 mm/hr km and is multiplied by 10 and stored as a 2-byte integer.

Pr_navigation (SDS, array size 9 x 116, float).

Scan_time: (VData table, record size 8 bytes, 116 records): A time associated with the scan. It is expressed as the UTC seconds of the day. The exact relation between Scan Time and the time of each IFOV is described in ICS Volume 3, section 3.

CHAPTER IV

THE GRAPHICAL ENVIRONMENT

In this chapter, we will describe the graphical environment of the RAIN_PROC program. First, we will consider the main window characteristics. The principal features of this environment such as push buttons and window display will be explained in detail. The source code for these programs can be accessed from a CD in Appendix G. The reader is referred to the "IDL Programmer's Guide" for further reading about the widget toolboxes.

4.1 RAIN_PROC Graphical Environment

The graphical environment for RAIN_PROC program was developed using IDL language. The graphical environment allows the user to exchange information in multiple ways: textual and graphically, through images and list of tables. The program provides the capacity to analyze new data from different sources (gauges, NEXRAD, and TRMM) and integrate those data. The data are contained in files with HDF (Hierarchical data format) format, the structure of this format is explained in Chapter III. The RAIN_PROC program defines default settings for the window display and for the parameters needed at the beginning.

The graphical environment is divided in three sections, see Figure 4.1. The first one (left side) is the parameters and procedures section where the user can set parameters and choose the procedure desired. The second section (center side) is the graphical section, where the program presents the results and graphics. The last section (right side) is the text section and shows the data in text mode through two lists, the first list (upper corner) is the Reflectivity data and the second list (lower corner) is the Instantaneous Precipitation data. Each section of the main window will be discussed later. This graphical environment can be manipulated using the right mouse button only. In some cases, as in the Minimum dBZ and the Maximum dBZ, the values are set with the keyboard.

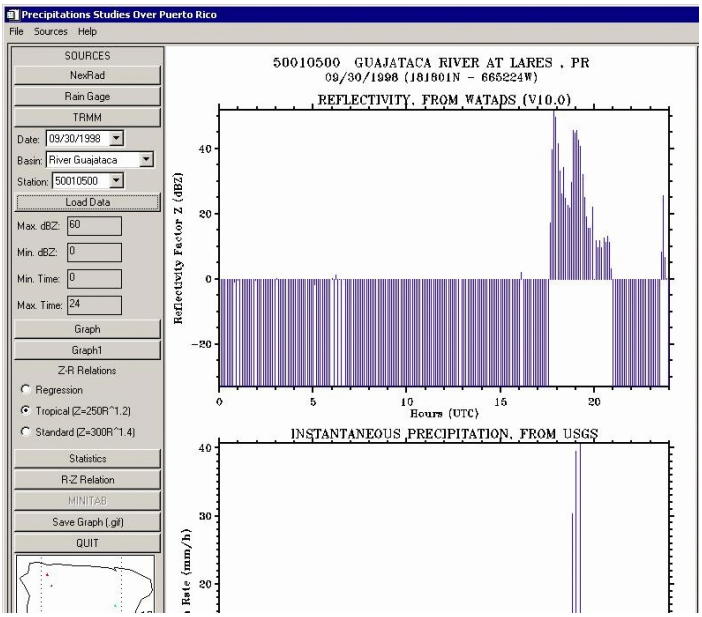

Figure 4.1: Main Window Display.

4.2 Parameters and Procedures Section

The parameters and procedures section is located at the left side of the main window. This section is divided basically in four subsections. The Sources section, the Parameters section, the Procedures section, and the Map section.

4.2.1 The Sources section

This section provides the user with three buttons to select data sources for data exploration and visualization, based on data set from the NEXRAD, Rain gauges, and TRMM sources. When the user selects the NEXRAD button, the program opens a new window (see Figure 4.2) to manipulate and visualize the database containing data from NEXRAD. In this window we have the options to select the Date and Time for the data desired. Then there are two buttons, the first one is the Load Data button, to load the data from the disk after the date and time were set. The second button is the Animation button that allows to animate the data for the date selected.

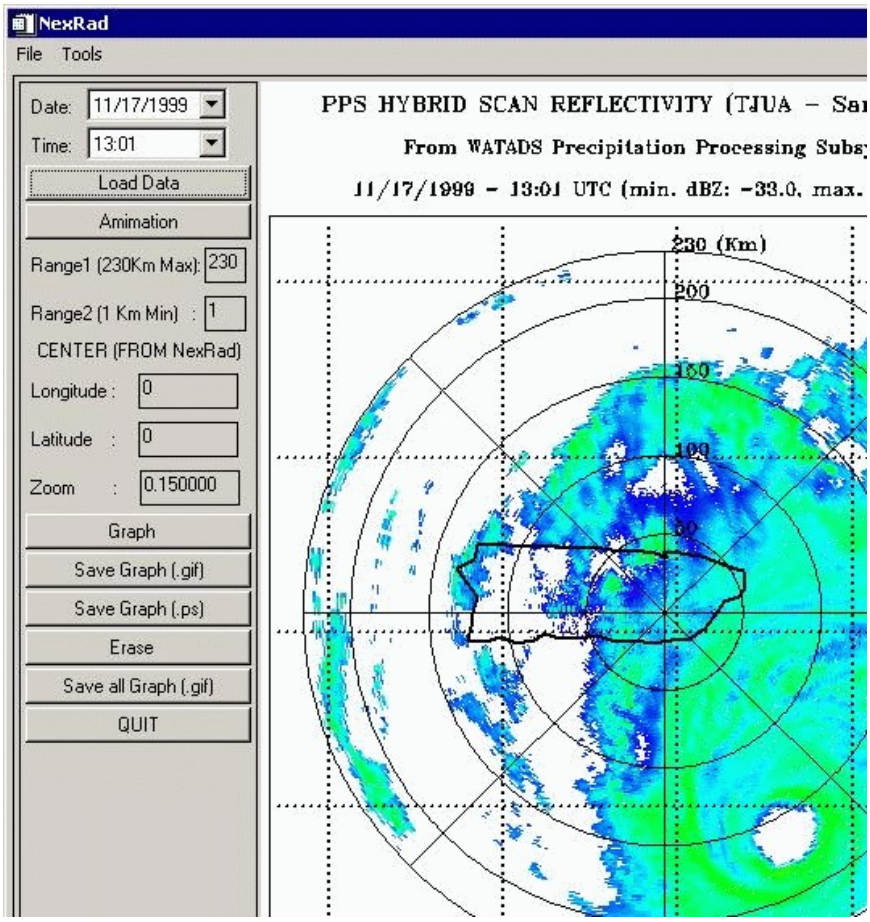

Figure 4.2: NEXRAD window (default parameters).

The next parameters are Range1 and Range2, used to set the ranges for which to visualize the data. Figure 4.3 shows a graph with the values for Range1 and Range2 at 150 Km and 1 Km respectively. In this graph we can see the reflectivity from 1 Km to 150 Km distance, relative to the radar position. This data corresponds to November 17, 1999 at 13:01 hours.

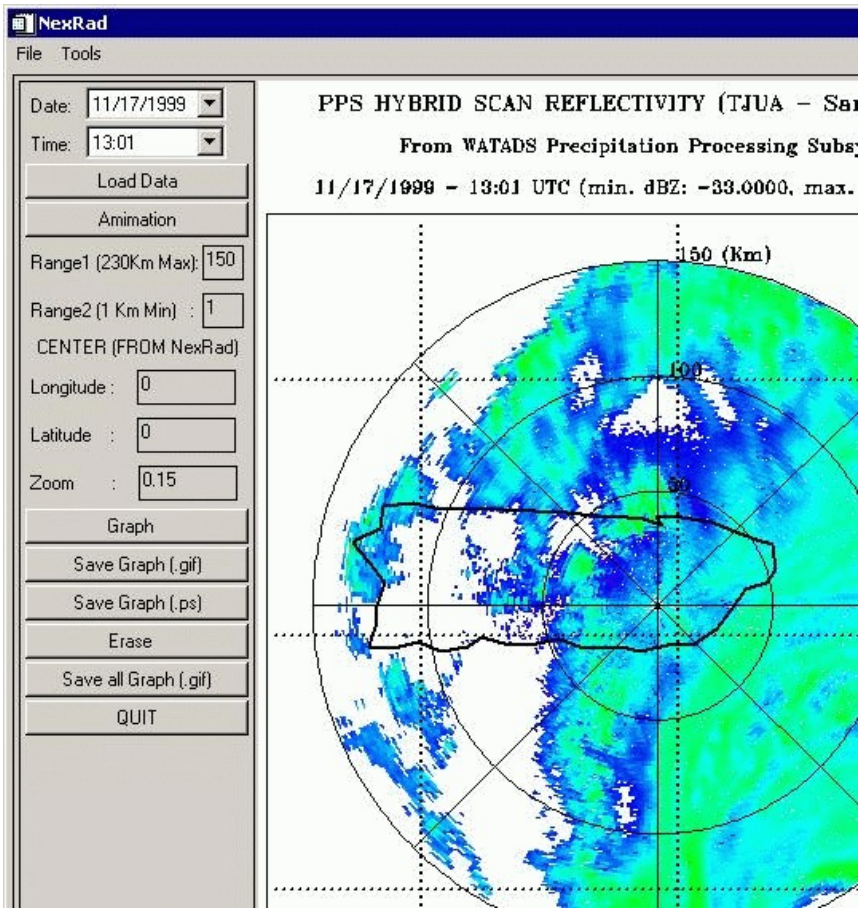

Figure 4.3: NEXRAD window (range parameters changed).

The three next parameters are Longitude, Latitude, and Zoom. These parameters allow the user to change the central position of the graph with respect to the NEXRAD geographical position. Figure 4.4 shows the graph for the same data but Longitude, Latitude, and Zoom parameters with the -0.5, -0.2, and -0.95 values, respectively.

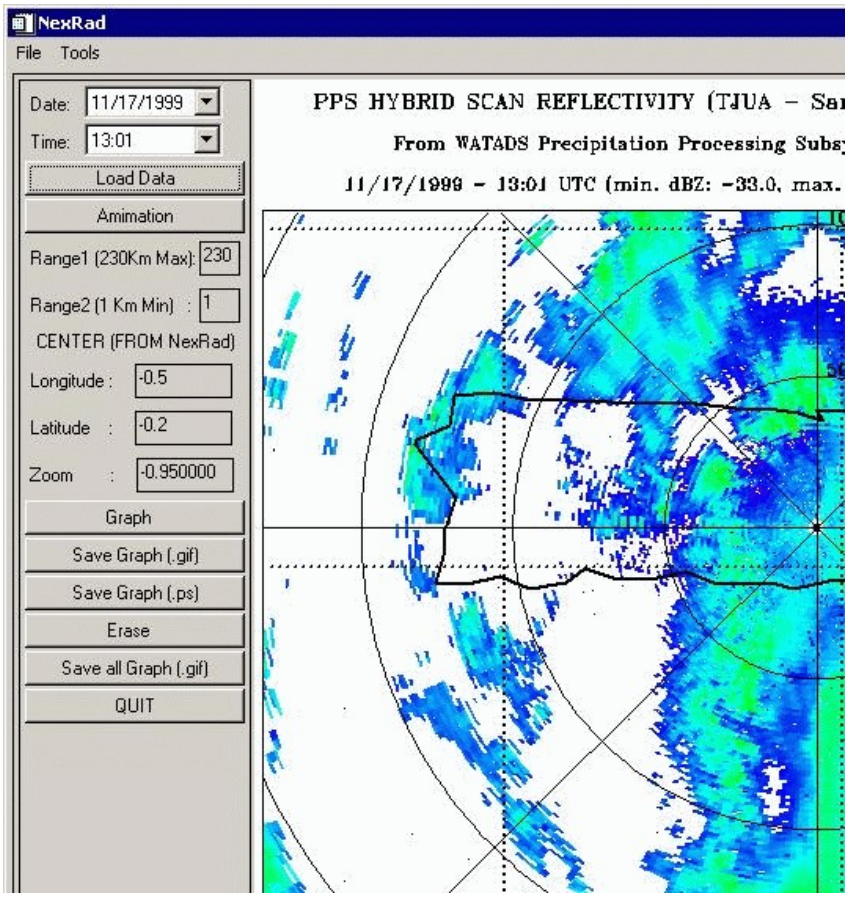

Figure 4.4: NEXRAD window (long., lat., and zoom parameters changed).

When the user selects the Rain Gauge button, the program opens a new window (see Figure 4.5) to manipulate and visualize the database containing data from the rain gauges network. In this window we have the options to set the Date, Basin, and Station for the data desired.

Then, there are four buttons; the first one is the Load Data button, to load the data from disk, after the parameters were set. The second button allows to modify the original data format from USGS to HDF format based on the parameter selected. The third button will show a map of Puerto Rico and the location of each rain gauge station for the basin selected (see Figure 4.5) and the

five zones in which Puerto Rico was divided. Finally, the last button allows to exit and to close the Rain gauge window.

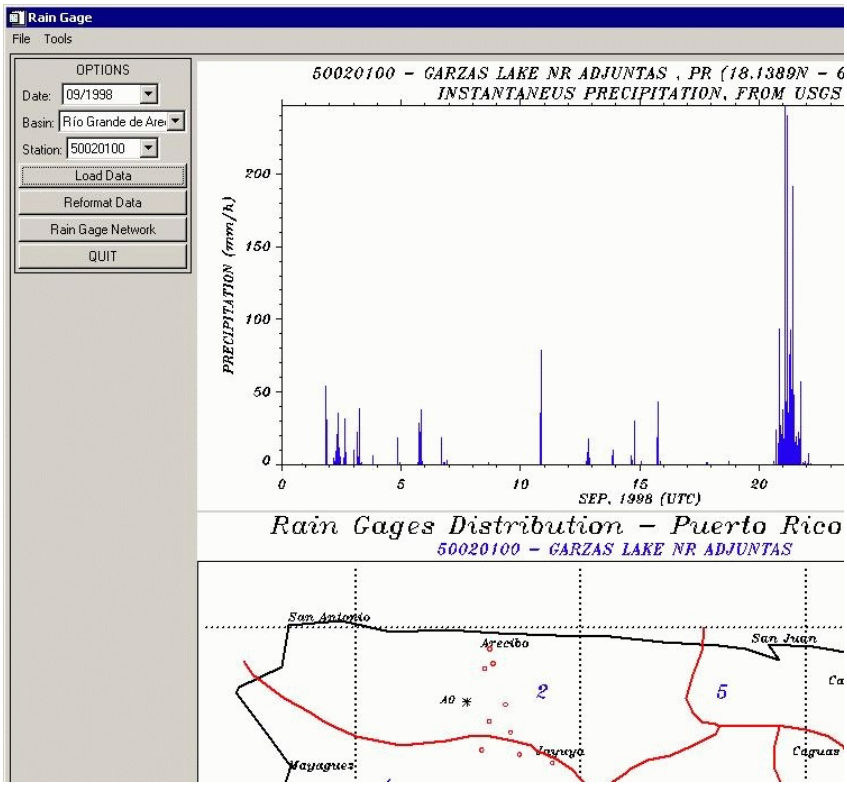

Figure 4.5: Rain gauge window.

Finally, when the TRMM button is selected, the program opens a new window (see Figure 4.6) to manipulate and visualize the database from the TRMM Precipitation Radar. There are four sections in this window, three for graphics, and the last one has controls and a button for data manipulation. The first graph section (upper left corner) depicts reflectivity (near surface Z) measured by the TRMM satellite over the Caribbean Region. The second graph (upper right corner) depicts a vertical cross section across the orbit scan. The last graph (lower left corner) depicts a vertical cross section along an orbit track cross section for a TRMM orbit ray (see Figure 4.6).

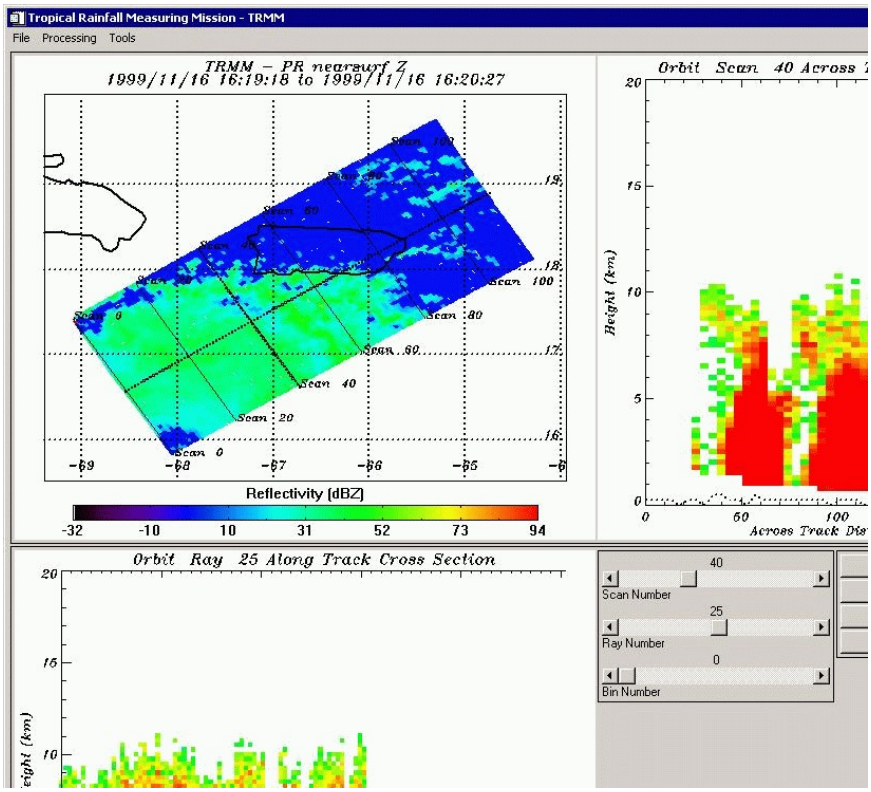

Figure 4.6: TRMM window.

The control section has a set of controls and buttons to trigger the following operations:

The Scan Number control allows to select the number of the orbit scan to visualize in the Orbit scan across section. The Ray Number control allows to select the number of the orbit ray to visualize, in the Orbit ray along track cross section.

The Open File button allows to open a file for a specified date and/or hour. The Save Graph and Save As PostScript File buttons allows to generate a graphical file in GIF and PostScript formats, respectively. The last button exits and closes the windows.

4.2.2 The Parameters section

This section provides a button and the options to set Date, Basin, Station, Maximum dBZ, Minimum dBZ, Minimum Time, and Maximum Time parameters. The Date parameter allows to select the date of the data to manipulate. The Basin parameter allows the selection of the name of the basin where the desired station is located. The Station parameter selects the station. The data is loaded into memory with the Load Data button. Maximum dBZ and Minimum dBZ parameters are used to set the maximum and minimum values for the reflectivity data (used to filter input data) which the user want to manipulate. Minimum and Maximum Time parameters set the time range for the desired data.

4.2.3 The Procedures section

This section provides seven buttons that perform the following operations:

The Graph button shows a plot of the original reflectivity (Reflectivity Factor Z in dBZ) and the instantaneous precipitation (Rain rate in mm/h) data. The plot for the reflectivity data is shown in the upper side of the graphical section of the main window display, and in the lower side the plot of the instantaneous precipitation data. The values of reflectivity are shown considering the Maximum dBZ and Minimum dBZ parameters. Both plots are shown within the time constraints selected in the Minimum and Maximum time parameters (see Figure 4.1). Two files are generated: the R_Data.txt and Z_Data.txt files, containing the values of the R and Z variables based on the parameter set.

The Graph1 button shows a plot of the instantaneous precipitation and the reflectivity, transformed to a rain rate using one of the Z-R relationships selected from the Z-R Relations options. These are, Regression (equation generated with linear a regression), Tropical and Standard relationships. Figure 4.7 shows a graph

using the tropical Z-R relation, the solid line is the original rain gauge data and dashed line is the rain rate calculated from reflectivity data using the tropical Z-R relationship.

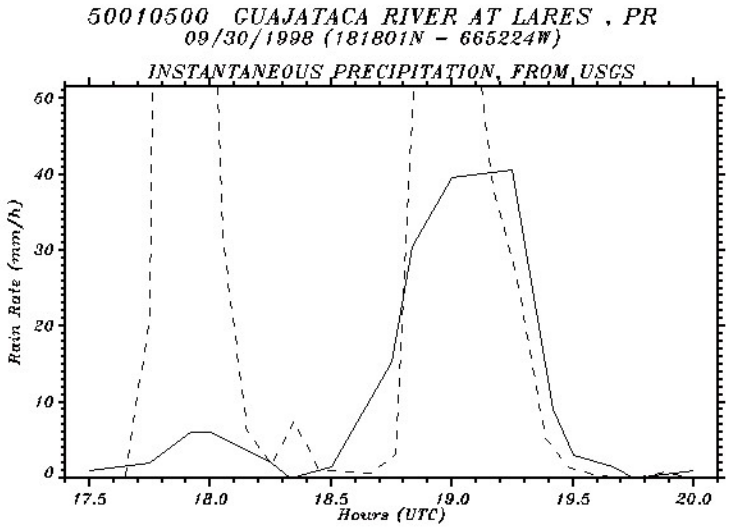

Figure 4.7: Rain Rate graph using the tropical Z-R relation.

The statistics button allows some statistical calculations and to graph the result in the graphical section of the main window. This operation generated a web page file including all of relative values calculated in this process. Figure 4.8 shows a sample of the result for a data set.

The R-Z Relation button generates a graph including the plots for the Tropical, Standard, and Generated Z-R relation based on a data set. This process generates two files, the Id_Date_ZR.txt file that includes the values of Z and R selected from the data set, and the ID_DATE_STAT.MAC file (a macro file to be used with the MINITAB program).

In Chapter V we will explain this macro in detail. See Appendix D to view an example of the macro file. Table 4.1 shows a typical file generated by this procedure.

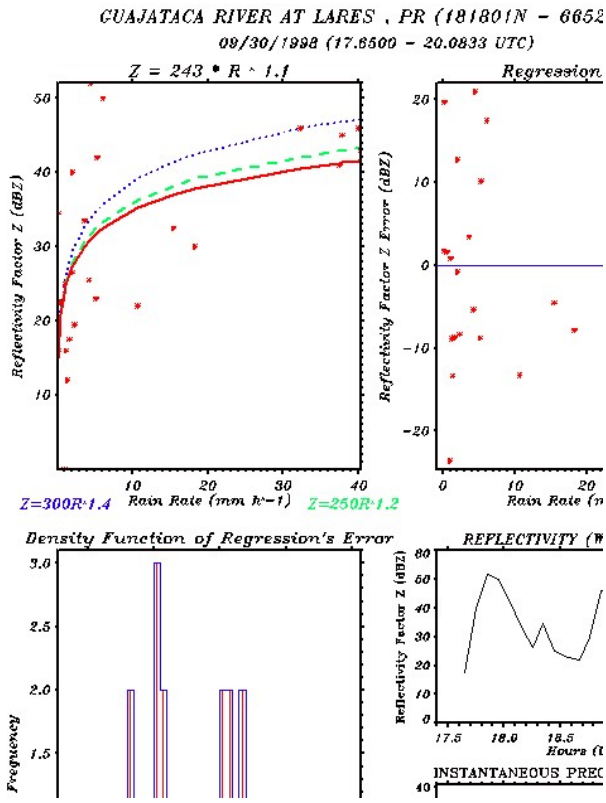

Figure 4.8: Statistical results for a data set.

Table 4.1: Typical values in Id_Date_ZR.txt file.

52.00	4.47
40.00	2.03
50.00	6.09
42.00	5.28
33.50	3.65
26.50	2.03
34.50	0.15
25.00	1.06
23.00	5.18
22.00	10.66
30.00	18.28
46.00	32.30
45.00	37.79
46.00	39.89
43.00	40.30
41.00	37.48
32.50	15.43

The MINITAB button allows to execute the MINITAB program and to run the macro file generated by R-Z Relation procedure. The Save Graph button saves the graph shown in the graphical section in GIF format. This procedure generates different files for every type of graph generated. These files have the syntax: ID_DATE_TYPE.GIF, where Type can be RZ, STAT or GRAPH. Finally, the Quit button allows to exit the program and close the main window.

4.2.4 The Map section

This section shows a map of Puerto Rico with geographical location of every rain gauge station for the selected basin. The selected rain gauge stations are shown in red color, and the others in blue color. There is a green color location point for the NEXRAD geographical location.

4.3 Graphical Section

The Graphical Section is a part of the main window used to present graphical results for every process executed in the program. Each process generates different types of graphs, corresponding to different types of calculations. Those graphical results were explained before. See section 4.2 to review details about the graphical section.

4.4 Text Section

The Text Section shows the data in text mode in two lists. The first list (upper right corner) is the Reflectivity data and the second list (lower right corner) is the Instantaneous Precipitation data. Every list has two columns. The first column lists the observation times for the data. The data appears in the second column. The number of rows can change depending on the range of data as indicated with the parameters to manipulate. The ranges can be adjusted in time

and in magnitude, Maximum dBZ, Minimum dBZ, Minimum Time, and Maximum Time. When the number pairs of values is greater than 20, a scroll bar allows to see different sections of the data set.

4.5 Main Menu Bar

The main menu bar has three sub-menus, File, Sources, and Help sub-menus. Each sub-menu has a few possibilities to select from. The File sub-menu only has the Exit option to quit the program.

The Sources sub-menu has the NEXRAD, Rain Gauge, and TRMM option, similar to the Source section in the parameters and procedures section, from the main window. Finally, the Help sub-menu has the Rain Proc Info option to show a brief explanation about the RAIN_PROC program.

4.6 Message Area

The message area displays information for the user. This area is similar to the one used by IDL. In this area the program provides messages about the procedures and the possible errors generated during the execution.

CHAPTER V

DATA ANALYSIS

The analysis of data will be discussed in this chapter. The first section reviews data fitting to Z-R model. The next section addresses the calculations of the coefficient of determination (R^2). Finally, the statistical analysis and the bootstrap technique will be explained. Calculations were made using the IDL language and the code for these programs can be accessed from a CD in Appendix G. Statistical analysis were made using the MINITAB program.

5.1 Fitting Data to Z-R model

We consider the problem of fitting a set of N data points to the Z-R model. In order to apply the linear regression (which describes the linear relationship between two variables, say, x and y) we must transform the equation $Z = a_1 R^{b_1}$ to a linear equivalent. For this purpose we apply the logarithm to the above Z-R relation as follows:

$$\log_{10} Z = \log_{10} a_1 + b_1 \log_{10} R \tag{5.1}$$

Then, we replace the equation as follows:

$$y = \log_{10} Z, \quad x = \log_{10} R, \quad a = \log_{10} a_1, \quad b = b_1 \tag{5.2}$$

Conventionally the symbol "x" is used for the "independent", or predictor variable, and the symbol "y" is used for the "dependent", or predictand variable.

From equation C.11 in Appendix C we have the following solution for the best-fit model parameters a and b:

$$\Delta \equiv SS_{xx} - (S_x)^2$$

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta} \tag{5.3}$$

$$b = \frac{SS_{xy} - S_x S_y}{\Delta}$$

where:

$$S \equiv \sum_{i=1}^{N} 1 = n \qquad S_x \equiv \sum_{i=1}^{N} x_i$$

$$S_y \equiv \sum_{i=1}^{N} y_i \tag{5.4}$$

$$S_{xx} \equiv \sum_{i=1}^{N} x_i^2 \qquad S_{xy} \equiv \sum_{i=1}^{N} x_i y_i$$

Replacing "x" and "y" from equation 5.2:

$$S \equiv n \qquad S_x \equiv \sum_{i=1}^{N} (\log_{10} R_i)$$

$$S_y \equiv \sum_{i=1}^{N} (\log_{10} Z_i) \tag{5.5}$$

$$S_{xx} \equiv \sum_{i=1}^{N} (\log_{10} R_i)^2 \qquad S_{xy} \equiv \sum_{i=1}^{N} [(\log_{10} R_i)(\log_{10} Z_i)]$$

A customized IDL procedure, LINEAR_REGRESSION.PRO, was implemented according to the formulas expressed by equation 5.5. For example, the data presented in Table 5.1 were analyzed and the resulting *a,* and *b* parameters can be seen in Figure 5.1.

The data in Table 5.1 are from the rain gauge station at Rio Grande de Loiza at Quebrada Arenas, PR (50050900). These were taken on November 15, 1999 from 21.0 to 22.1 hours (UTC), and with a minimum reflectivity value of 20.0 dBZ (Minimum db).

Table 5.1: Data set for one rainfall event.

Reflectivity (dBZ)	**Precipitation (mm/h)**
33.00	5.07
36.33	9.14
38.33	30.47
34.99	18.28
34.30	10.66
32.16	12.19
28.50	10.66
27.75	15.23
22.74	3.04
23.16	3.04
24.00	3.04

A brief explanation of Figure 5.1 follows. In the upper left side, we see curves for the Z-R relationship of the Tropical, Standard, and Calculated relations. The dotted line is for the Standard relationship ($Z=300\ R^{1.4}$), the dashed line is for The Tropical relationship ($Z=250\ R^{1.2}$), and the solid line is for the calculated relationship ($Z=75.4\ R^{1.26}$).

Graphs on the upper right side and lower left side are for the regression's error and the density function of regression's error respectively. Finally, the time series of the radar and rain gauges measurements can be see in the graph at the lower right side.

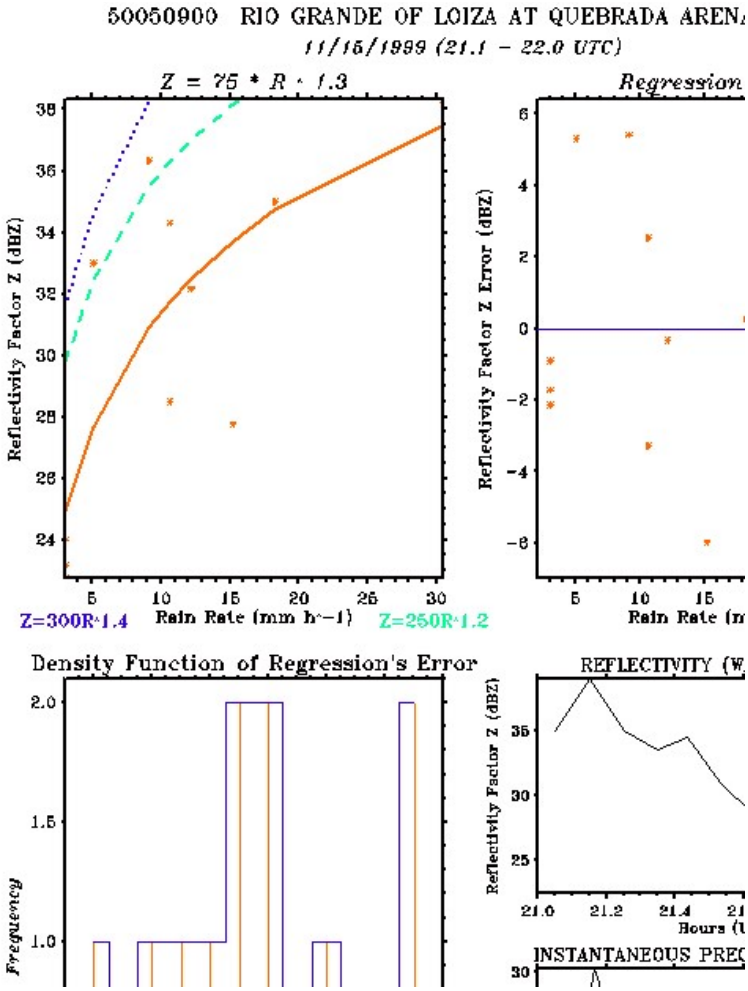

Figure 5.1: Data fit to the Z-R model (a = 75.4, b = 1.3).

5.2 Coefficient of determination (R^2)

The Coefficient of determination R^2 is interpreted as a measure of the strength of the linear relationship between the predictor variable "x" and the response variable "y", and is calculated with the following equation (Hoel, 1971):

$$R^2 = \frac{(N\sum_{i=1}^{N} x_i y_i - \sum_{i=1}^{N} x_i \sum_{i=1}^{N} y_i)^2}{[N\sum_{i=1}^{N} x_i^2 - (\sum_{i=1}^{N} x_i)^2][N\sum_{i=1}^{N} y_i^2 - (\sum_{i=1}^{N} y_i)^2]} \tag{5.6}$$

Note that in equation 5.6, the value of value of R^2 will be equal to 1 if, and only if, the data points lie on a straight line. Applying equation 5.6 to the data set shown in Table 5.1 we obtain the result $R^2 = 0.61$. This value of R^2 shows that there is an acceptable tendency for values of "x" to be associated with values of "y" in a straight line. We defer, for Chapter VI, the discussion of another cases.

5.3 Statistical analysis

Further analysis usually follows the fitting of any type of model, such as regression. But, the most important part is the plotting of residuals. Many analyses rely on the assumption that residuals are normally distributed with constant variance, independent of the value of the response variable.

The residuals are often used to detect and assess the degree of discrepancy between the assumed model (including the assumptions made) and the observed data. In order to verify these assumptions, the MINITAB program provides subcommands to graphically illustrate whether or not these assumptions have been met. In this part of the analysis, we will explain statistical results obtained with the MINITAB program for the data set shown in the Table 5.1.

Figure 5.2 shows another angle to statistical characterization for the data. The box-and-whiskers plot shows the main features of a batch of data, Figure 5.2(a) and Figure 5.2(b). The median of the data is marked with a line dividing the box. The line in the left of the box is the lower hinge (essentially the first quartile) and the line in the right of the box is the upper hinge (essentially the third quartile). The box represents the middle half of the data. The extent of the data and the location of unusual observations are indicated with special symbols.

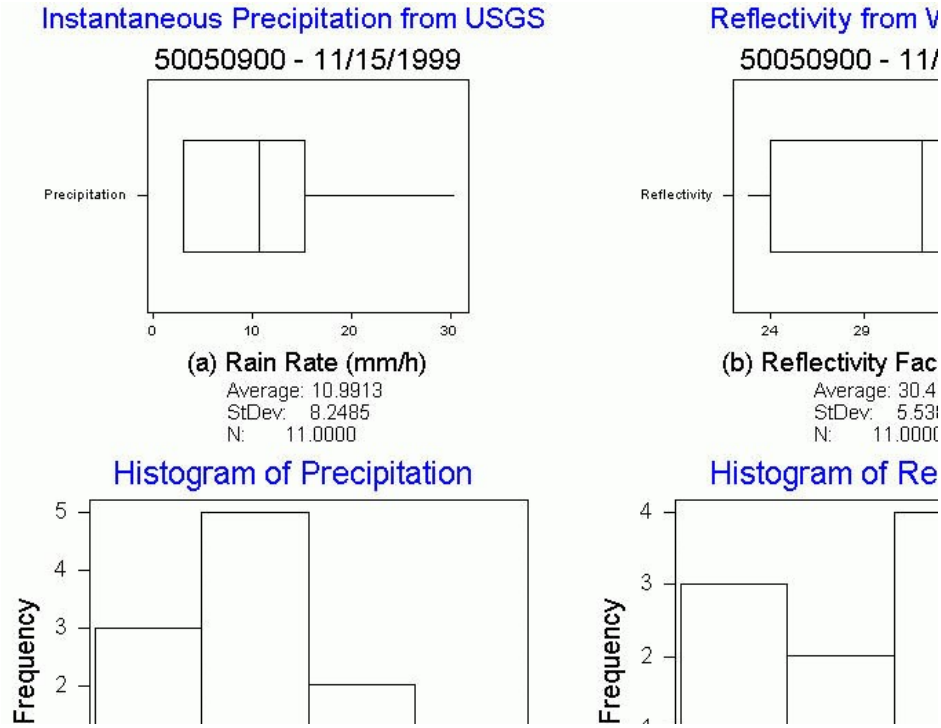

Figure 5.2: Statistical results from MINITAB program.

For this example, the box-and-whiskers plots show that both precipitation and reflectivity are asymmetrical and that there are no outliers. There is not a lot of variability for precipitation but there is some for reflectivity. The average, the standard deviation, and the number of elements for each set of data are indicated below.

Figures 5.2(c) and 5.2(d) present still another characterization of the data. The histogram casts the data into appropriate intervals (bins) on the x axis, and for each interval draws a bar whose height is the number of observations (or frequency) within that interval. This graphic, of course, agrees with the box-and-whiskers plots.

Normality of the errors is required for the validity of the inferences drawn from the fitted model. Figure 5.3(a) shows a normal probability plot of the errors. In the horizontal axis appear the percentiles of the normal distribution; the vertical axis contains the residuals.

A deviation from a straight line can result from a model miss-specification about normality. The validity of the normal error plots depends on correct model specification. In this case, as is shown in Figure 5.3(a), the data roughly satisfy the assumption of normal distribution of the errors.

The regression analyst sometimes has to cope with individual data points that do not fit the trend. These suspect point named outliers, may arise for a number of reason, most of which are classified into the failure of some assumptions. As a result is natural and convenient to treat the problem of outliers with analysis of residuals. The size of the residuals is crucial to diagnostic.

Figure 5.3(b) draws a control chart for individual standardized residual. The estimate SL of the standard deviation is based on the average of the moving range divided by an unbiased constant. By default, the moving range is of length 2, since consecutive values have the greatest chance of being alike. In this case, the average of the moving range is the average of the difference of consecutive values.

We can see in Figure 5.3(c) the histogram of the standardized residuals supporting the roughly normality that is shown by Figure 5.3(a). The Figure 5.3(d) is a classical appearance of such a standardized residual plot. The figure indicates a random pattern around zero with no detectable trend. This is a condition to the linear model.

The Figure 5.3(e) is reasonably clear. The funnel effect indicates that as the response variable gets larger, the deviations of the $10^{\text{Residuals}}$-1 from zero become shorter. Hence, we detect a condition whereby the error variance is not constant but decreases as the measured response increases. This means that the model has a multiplicative regression error (see Appendix C). Then, we can say that the linear transformation is valid.

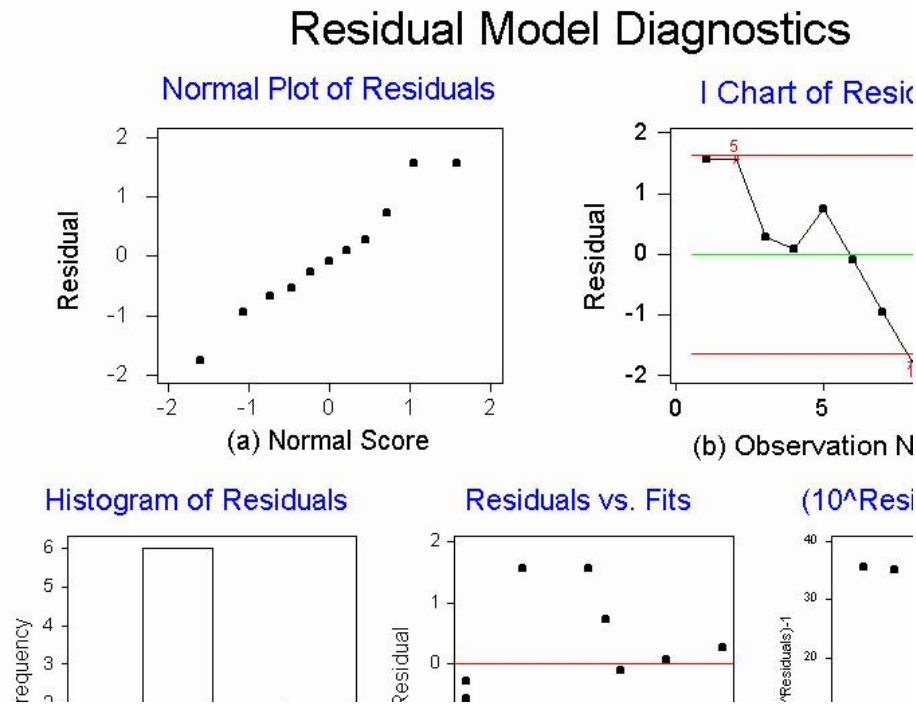

Figure 5.3: Residual analysis and regression diagnostic results from MINITAB program.

5.4 Bootstrap

The Bootstrap is a non-parametric statistical technique very useful in single-sample situations. The basic idea is to build up a collection of artificial data sets from a given collection of real data, by re-sampling the observations. Sometimes such methods are also known as re-randomization test or Monte-Carlo test. It is described in Appendix C.

In general, re-sampling techniques have two very appealing advantages. First, no assumptions regarding an underlying theoretical distribution for the data are necessary, because the procedures consist entirely of operation on the data themselves. Second, any statistic that may be suggested as important by the physical nature of the problem can form the basis of the procedure, so long as it can be computed from the data.

As a one-sample procedure, the bootstrap is often used to estimate confidence intervals around observed values of a test statistic. Because we do not know the analytical form of its sampling distribution, the procedure can be applied to any statistic, regardless of how mathematically complicated it may be. The usual confidence interval requires a normality distribution of the population from the sample drawn. Very often, this does not happen. We have used bootstrapping to compute approximately 95% confidence intervals for the *a,* and *b* parameters of our model.

The bootstrap operates by constructing the artificial data batches using sampling with replacement from the original data. This process is repeated a large number, say n_B times, yielding n_B bootstrap samples of size n. A particular member of the original data set may be drawn more than once, or not at all, in a given bootstrap sample.

Fundamental to this process are algorithms called random-number generators, or more precisely, uniform [0, 1] random number generator. Such algorithms produce a stream of numbers that resemble independent values drawn randomly from the uniform [0, 1] PDF $f(u) = 1$, $0 \leq u \leq 1$. This distribution is also known as the rectangular distribution. As the name and density function imply, these random numbers are equally likely to take on any value between 0 and 1.

Bootstrap samples are generated by assigning numbers; from 1 to n, to each of the n observed data values to be resampled. The output of the random-number generator is then used to form the bootstrap sample by repeatedly choosing among the n values with equal probability (Efrom and Tibshirani, 1993).

Figure 5.4 shows the distribution of 10,000 bootstrap samples computed for the *a,* and *b* parameters from the data shown in Table 5.1. The Table 5.2 shows the bootstrap confidence intervals using the percentiles of 2.5% and 97.5% for the bootstrap distribution. The estimated 95% confidence interval for the parameter *a* is between $ta_{2.5} = 35.43$ and $ta_{97.5} = 500.92$, and for the *b* parameter is between $tb_{2.5} = 0.49$ and $tb_{97.5} = 1.67$.

From the Figure 5.4 we can see that the distribution curve for the *a* parameter is asymmetric, this is reasonable because we made the bootstrap from the antilogarithm parameter calculated for *a*. See equation 5.2.

Table 5.2: Bootstrap confidence intervals for *a*, and *b* parameters.

Station Id	**N**	**$ta_{2.5}$**	**$Ta_{97.5}$**	**$tb_{2.5}$**	**$tb_{97.5}$**
50050900	11	35.43	500.92	0.49	1.67

The Estimation Error (EE) of a and b parameters are calculated as follow:

$EE(a) = (ta_{97.5} - ta_{2.5})/2 = (500.92 - 35.43)/2 = 232.745$

$EE(b) = (tb_{97.5} - tb_{2.5})/2 = (1.67 - 0.49)/2 = 0.59$

The estimation error of the *a* parameters is large due to a small sample size and to the asymmetrical distribution.

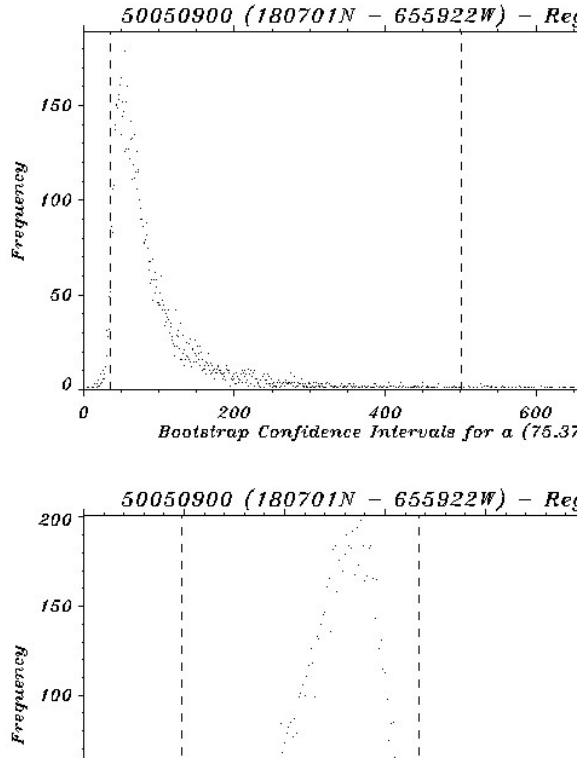

Figure 5.4: Bootstrap Confidence Intervals for *a* and *b* parameters.

CHAPTER VI

RESULTS AND DISCUSSION

In this chapter we present some important results for selected rainfall events. The RAIN_PROC program was applied to a specific data set collected with NEXRAD and the rain gauges network. From these data sets we have originally selected 68 events by considering their duration, intensity, and geographical regions (see Appendix F). Note that in the analysis of rainfall for individual events described here, emphasis will be placed on the six events with the best Z-R relationships, one per region. Table 6.1 shows the six selected events, Table 6.2 shows their calculated parameters, and Table 6.3 shows the bootstrap confidence intervals for those parameters. After analyzing the data with the fitting algorithm, further comparisons are made using data from TRMM.

Table 6.1: The six selected rainfall events with the best Z-R relationships by region.

Station Id.	Date	Initial Time (UT)	Final Time (UT)	Hours	Region
50051180	19991117	03.75	08.00	4.25	1
50039500	19991202	15.40	21.20	5.80	2
50065500	19991203	02.55	04.70	2.15	3
50021700	19991202	18.80	21.40	2.60	4
50081000	19991202	19.80	22.00	2.20	5
50055100	19991202	18.30	21.80	3.50	6

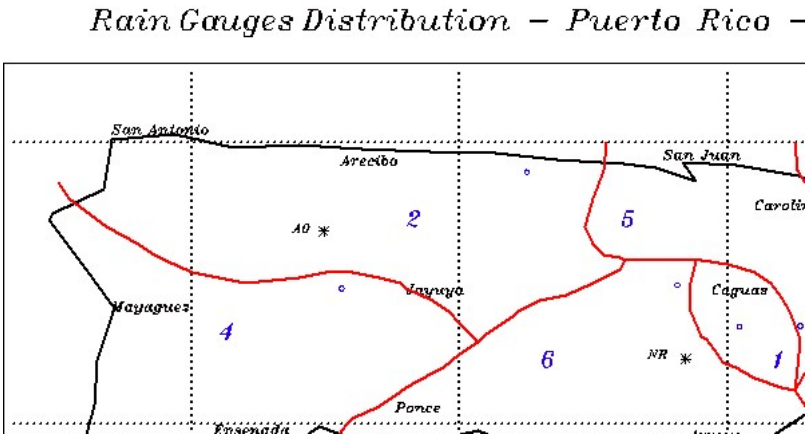

Figure 6.1: The six selected rainfall event locations by region.

Figure 6.1 shows the geographical position of the six stations where the six selected events used in this analysis were registered. The stations are marked in small circles. In Table 6.1 we can see the station identification, the date, initial time, final time, and the geographical region classification. The station identification is a number assigned for each station in the network.

Table 6.2 shows both calculated *a,* and *b* parameters for each selected event applying the linear regression. We can see values for the *a* parameter between 56.65 and 244.23 and for the *b* parameter between 0.75 and 1.12. The R^2 column indicated the coefficient of determination and the N column is the length of the time series. Also included are the region, station identification, date and the minimum reflectivity value employed in the regression procedure. The minimum reflectivity values are between 17.0 and 29.0 dBZ.

Table 6.2: Parameters calculated for the six selected rainfall events.

a	b	R^2	N	Region	Station Id.	Date	Min. dBZ
145.91	1.12	0.57	18	1	50051180	19991117	17.0
138.83	0.79	0.32	37	2	50039500	19991202	20.0
56.65	0.80	0.39	12	3	50065500	19991203	20.0
237.61	0.75	0.61	16	4	50021700	19991202	25.0
221.47	0.81	0.43	17	5	50081000	19991202	29.0
244.23	0.91	0.43	15	6	50055100	19991202	24.0

Table 6.3: Bootstrap confidence intervals for *a,* and *b* parameters of the selected events.

Station Id.	N	$ta_{2.5}$	$ta_{97.5}$	EE(a)	$tb_{2.5}$	$tb_{97.5}$	EE(b)
50051180	18	40.93	370.68	164.87	0.66	1.91	0.62
50039500	37	67.28	322.98	127.85	0.38	1.13	0.37
50065500	12	14.36	908.25	446.94	-0.55	1.28	0.91
50021700	16	110.16	491.85	190.84	0.48	1.06	0.29
50081000	17	63.14	919.42	428.14	0.28	1.18	0.45
50055100	15	55.25	2171.73	1058.24	0.22	1.44	0.61

In Table 6.3 we have the station identification, the number of values collected per event and the percentiles of 2.5% and 97.5% for the bootstrap confidence intervals for the *a* and *b* parameters. To continue with the second tie of the selection process, we disregard events with poor coefficients of determination (less that 30%), too small sizes, and those events with values for the *a,* and *b* parameters far away from the Tropical Z-R relationship parameters (see Appendix F).

6.1 Selected event for the region 1

The selected event for the region 1 was registered in the rain gauge station at Quebrada Salvatierra NR San Lorenzo, PR (50051180). The data were taken on November 17, 1999 from 3.75 to 8.0 hours (UTC), and with minimum reflectivity value of 17.0 dBZ (Minimum db). The data fit to Z-R model and the bootstrap estimate results are shown in Figure 6.2. In the upper left side, we see curves for the time series of the radar and rain gauges measurements. The dotted line is for the reflectivity in dBZ, and the solid line is for precipitation in mm/h.

In the graph on the upper right side we see curves for the Z-R relationship of the Tropical, Standard, and Calculated specific relations. The dotted line is for the Standard relationship ($Z=300\ R^{1.4}$), the dashed line is for the Tropical relationship ($Z=250\ R^{1.2}$), and the solid line is for the Calculated relationship ($Z=145.91\ R^{1.117}$). In this graph we can see that the curve of the Calculated relationship is near to Tropical relationship. The coefficient of determination is 0.57; this is a low but acceptable value. Graphs on the

lower right side and lower left side are for the histogram of 10,000 bootstrap samples computed for *a* and *b* parameters respectively for the data shown in Table 6.4. In Table 6.3 we can see the bootstrap confidence intervals for this case (in the second line). The estimated 95% confidence interval for the *a* parameter is between $ta_{2.5} = 40.93$ and $ta_{97.5} = 370.68$, and for *b* is between $tb_{2.5} = 0.66$ and $tb_{97.5} = 1.91$. We can see the bootstrap intervals delimited by two vertical dashed lines. As in the case in the previous chapter, the distribution curve is asymmetric for *a* parameter. The distribution curve is a little asymmetric for *b* because of the outliers in the precipitation data.

The outliers are shown in Figure 6.4(a). The box-and-whiskers plots in Figure 6.4 show that both precipitation and reflectivity are asymmetrical in agreement with the results in Figures 6.4(c) and 6.4(d). The residual diagnostic graph, Figure 6.5, shows that there is no normality in the errors, as we can see in Figures 6.5(a) and 6.5(c). There is not enough confidence in the conclusion of no-normality for this case. There are no outliers for residuals, and as was mentioned in Chapter V, the model has a multiplicative regression error.

Table 6.4: Data set for selected rainfall event for the region 1.

Reflectivity (dBZ)	**Precipitation (mm/h)**
25.34	6.09
27.59	3.05
30.40	3.05
28.79	2.03
25.19	3.05
20.99	2.03
25.40	2.03
17.49	2.03
38.29	17.27
30.69	5.08
30.00	6.09
31.09	4.06
32.80	8.13
30.80	4.06
31.09	9.14
34.79	45.72
35.69	8.13
37.19	11.18

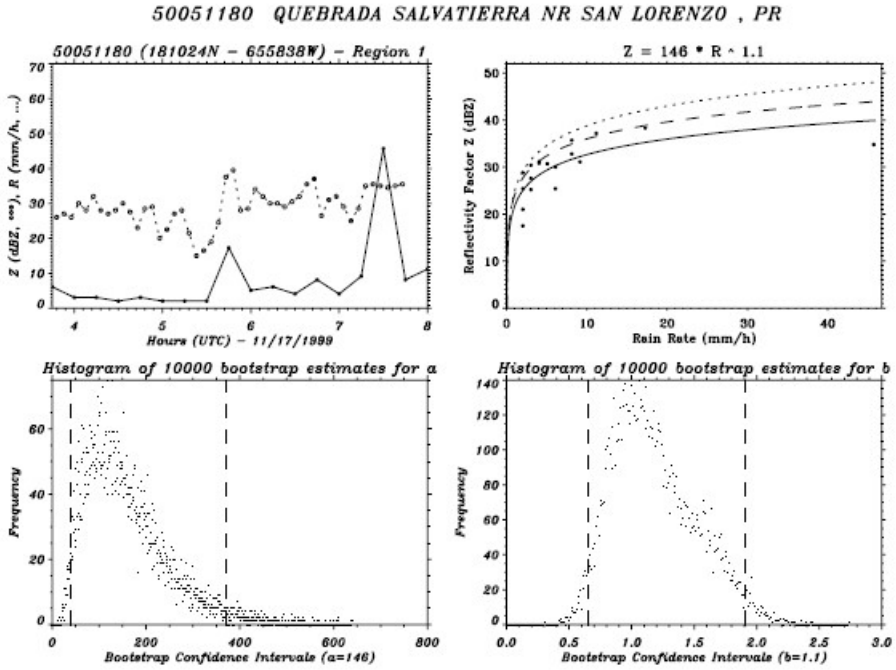

Figure 6.1: Data fit to Z−R model and bootstrap estimates in the region 1.

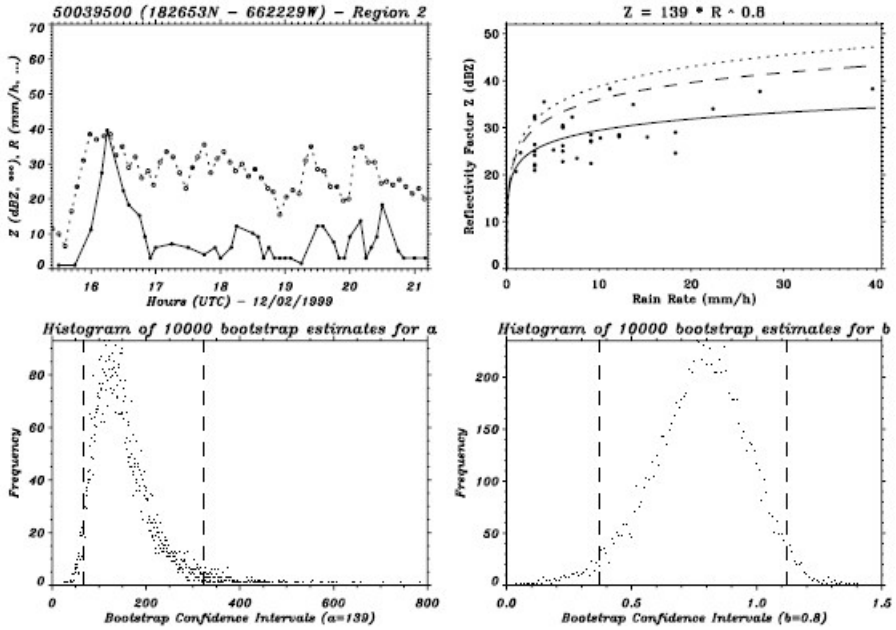

Figure 6.2: Data fit to Z−R model and bootstrap estimates in the region 2.

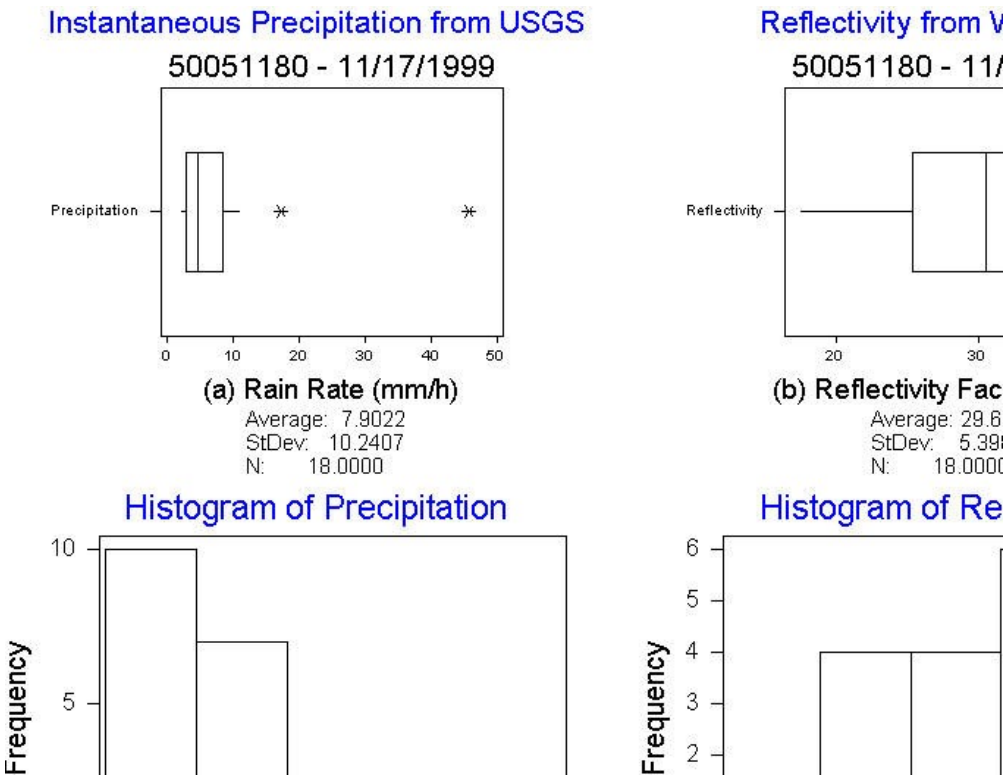

Figure 6.4: Statistical results obtained with MINITAB for region 1.

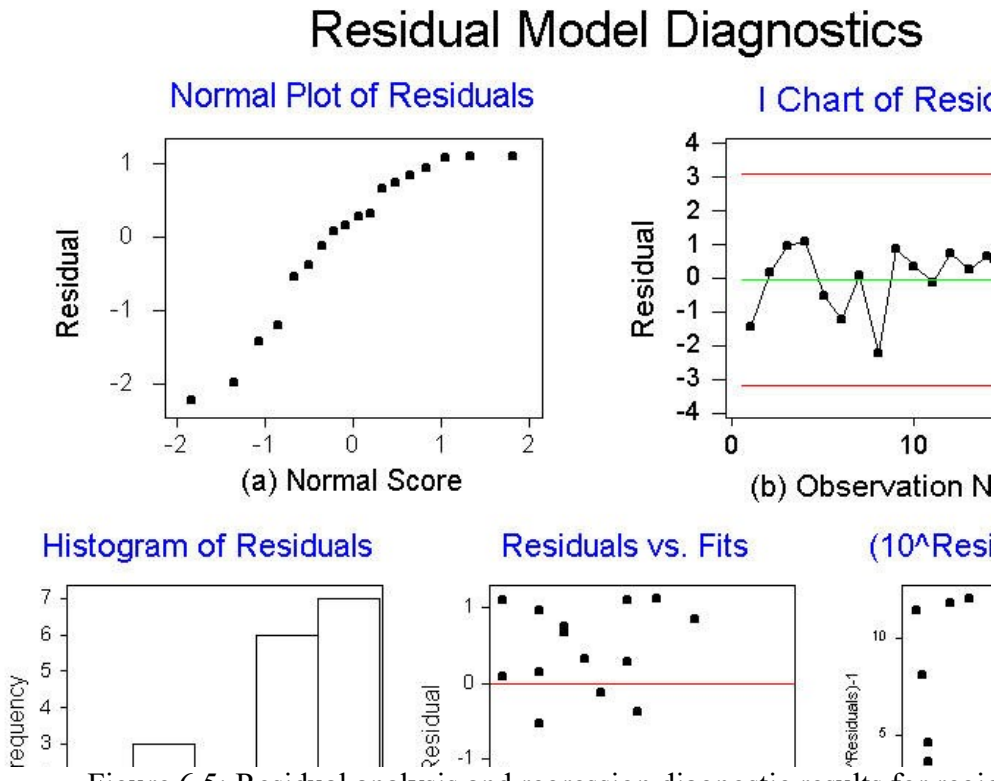

Figure 6.5: Residual analysis and regression diagnostic results for region 1.

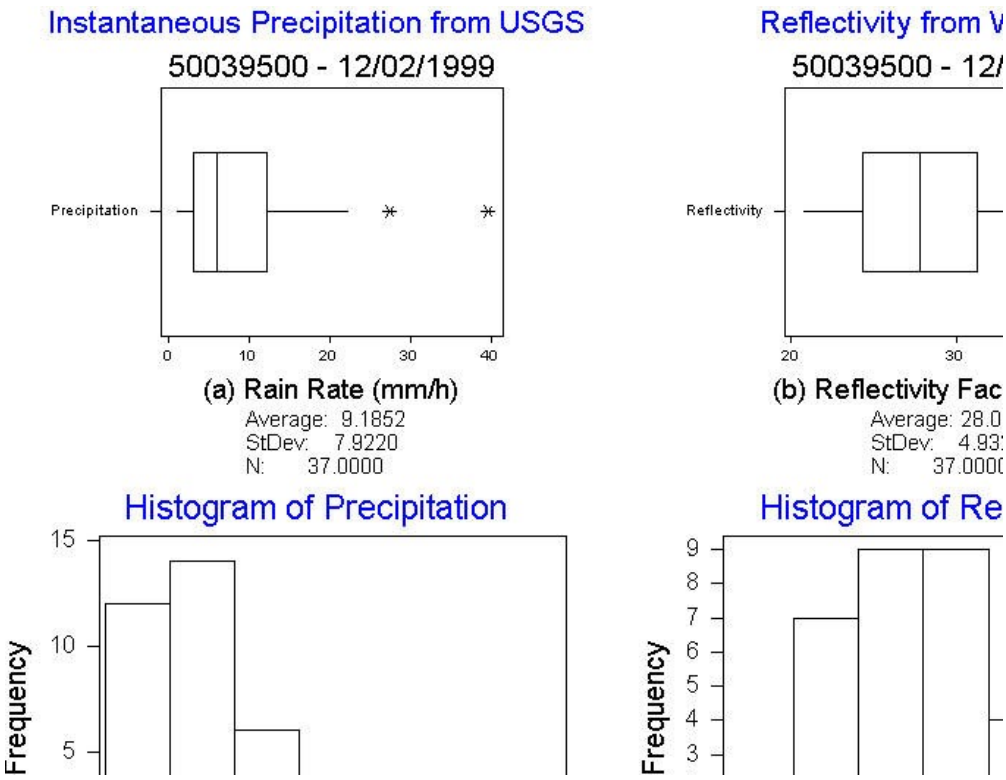

Figure 6.6: Statistical results obtained with MINITAB for region 2.

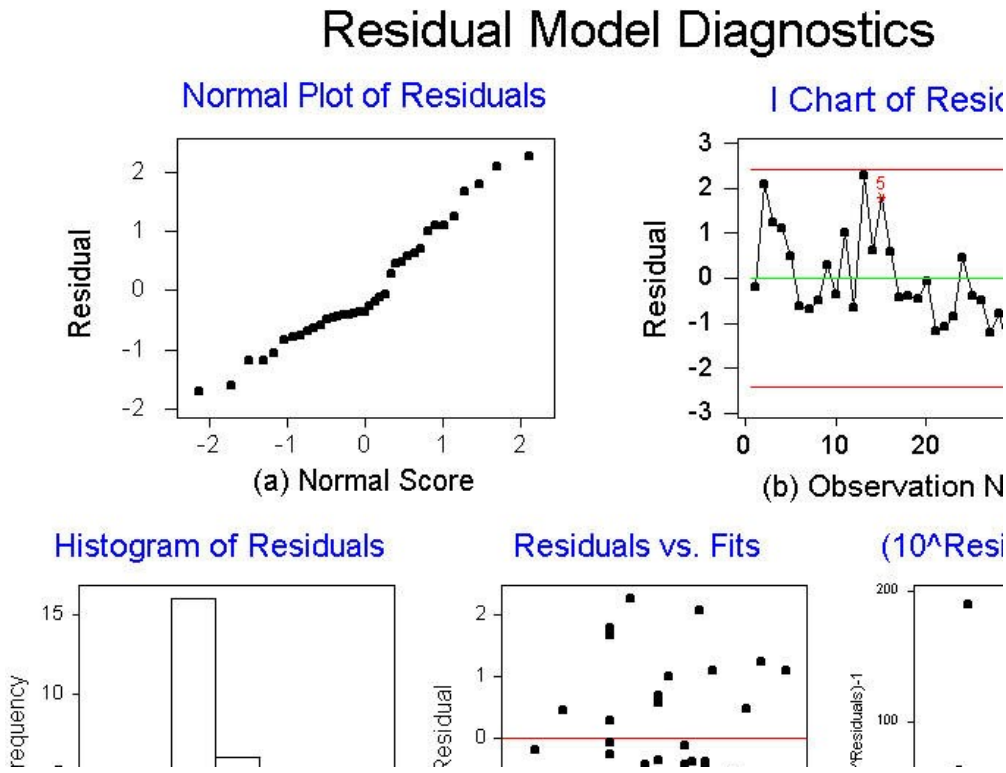

Figure 6.7: Residual analysis and regression diagnostic results for region 2.

6.2 Selected event for the region 2

The selected event for the region 2 was registered in the rain gauge station from Cibuco river at Vega Baja, PR (50039500). The data were taken on December 2, 1999 from 15.4 to 21.2 hours (UTC) with minimum reflectivity value of 20.0 dBZ. Table 6.5 shows the data for this event. As shown in Figure 6.3, the Calculated relationship is $Z=138.83\ R^{0.79}$ (the solid line) and is more similar to the Tropical relationship too. The coefficient of determination is 0.32; this is a poor value. The estimated 95% confidence interval for *a* is between $ta_{2.5} = 67.28$ and $ta_{97.5} = 322.98$, and for *b* is between $tb_{2.5} = 0.38$ and $tb_{97.5} = 1.13$. Curve forms of the bootstrap are similar to the previous case.

Table 6.5: Data set for selected rainfall event for the region 2.

Reflectivity (dBZ)	Precipitation (mm/h)
20.69	1.02
38.25	11.17
37.71	27.43
38.25	39.62
34.00	22.35
29.00	18.29
27.99	15.24
26.99	9.14
26.40	3.05
26.17	6.09
32.25	7.11
24.99	6.09
35.50	4.06
30.17	6.09
32.49	3.05
30.00	6.09
28.33	12.19
27.83	10.14
27.25	9.14
24.99	3.05
22.80	6.09
20.92	3.05
21.83	3.05
24.67	1.52
28.50	12.19

Table 6.5: Data set for selected rainfall event for the region 2 (cont.).

Reflectivity (dBZ)	Precipitation (mm/h)
28.08	12.19
23.50	7.619
22.17	3.05
22.42	9.14
34.92	13.72
32.00	3.05
30.50	6.09
28.50	9.14
24.59	18.29
25.25	5.08
24.17	3.05
21.99	3.05

The Figure 6.6(a) shows the outliers in the precipitation data. The box-and-whiskers plots indicated that both precipitation and reflectivity are asymmetrical and the Figures 6.6(c) and 6.6(d) agree with the box-and-whiskers plots. Figures 6.7(a) and 6.7(c) illustrate that there is some normality in the errors. There are no outliers for the residuals and the model has a multiplicative regression error, as is shown in the Figure 6.7(e).

6.3 Selected event for the region 3

The selected event for the region 3 was registered in the rain gauge station at Mameyes river NR Sabana, PR (50065500). The data were taken on December 3, 1999 from 2.55 to 4.7 hours (UTC), and with minimum reflectivity value of 20.0 dBZ (Minimum db). Table 6.6 shows the data for this event.

The Calculated relationship for this event is $Z=56.65\ R^{0.80}$ (the solid line) and is far from the Tropical relationship as is shown in the Figure 6.8. The coefficient of determination is 0.39; this is a low, barely acceptable value. The estimated 95% confidence interval for *a* is between $ta_{2.5} = 14.36$ and $ta_{97.5} = 908.25$, and for *b* is between $tb_{2.5} = -0.55$ and $tb_{97.5} = 1.28$. The shape of the bootstrapped distribution curve for *a* is similar to the region 2 and the line for the upper interval is not shown because it is out of

the abscissa range; but for *b* it is bi-modal because of the outlier in the reflectivity data. When the outliers are in the vertical direction, they produce bigger effects in the resampling. There are two reasons for this. The first one is that methods for fitting that are resistant to outliers are usually not very efficient, and may behave badly under resampling. The second reason is that outliers can be disruptive of resampling analysis for methods such as least squares that are not resistant to outliers (Davison and Hinkley, 1997).

Table 6.6: Data set for selected rainfall event for the region 3.

Reflectivity (dBZ)	Precipitation (mm/h)
22.50	6.09
22.50	7.62
20.59	12.19
25.75	24.38
26.49	18.29
37.09	67.06
25.33	9.14
29.49	6.09
25.33	6.09
27.17	6.09
20.67	8.13
22.99	3.05

Figure 6.10 shows outliers in both the precipitation data and reflectivity data. The box-and-whiskers plots show that both precipitation and reflectivity are asymmetrical and the Figures 6.10(c) and 6.10(d) agree with the box-and-whiskers plots. Figures 6.11(a) and 6.11(b) shows that there is some degree of normality in the errors. There are no outliers for residuals and the model has a multiplicative regression error, as shown in Figure 6.11(e).

6.4 Selected event for the region 4

The selected event for the region 4 was registered in the rain gauge station at Rio Grande de Arecibo ABV Utuado, PR (50021700). The data were taken on December 2, 1999 from 18.8 to 21.4 hours (UTC), and with minimum reflectivity value of 25.0 dBZ

(Minimum db). Table 6.7 shows the data for this event. The Calculated relationship for this event is $Z=237.6\ R^{0.75}$ and it is similar to the Tropical relationship, see Figure 6.9. Does not add support to Ulbrich's suggestion that the rain characteristic near the center of the island obeys a continental relationship (Ulbrich, Petitididier, and Campos, 1999). The coefficient of determination is 0.61; this is a fairly good value. The estimated 95% confidence interval for *a* is between $ta_{2.5} = 110.16$ and $ta_{97.5} = 491.85$, and for *b* is between $tb_{2.5} = 0.48$ and $tb_{97.5} = 1.06$. The curve for the bootstrap distribution for *a* is similar to case of region 1. For *b*, in this case, the curve is nearly symmetric.

Figure 6.12 shown that there are no outliers in either the precipitation data and reflectivity data. The box-and-whiskers plots show that both precipitation and reflectivity are asymmetrical, and the Figures 6.12(c) and 6.12(d) agree with the box-and-whiskers plots. Figures 6.13(a) and 6.13(b), shows that there is some normality in the errors. There are no outliers for residuals and the Figure 6.13(e) we can see that the model has a multiplicative regression error.

Table 6.7: Data set for selected rainfall event for the region 4.

Reflectivity (dBZ)	**Precipitation (mm/h)**
25.17	2.03
34.49	16.26
35.89	15.24
36.75	18.29
33.00	24.38
36.75	33.53
33.83	33.53
32.00	27.43
31.00	15.24
29.49	3.05
30.75	12.19
27.00	6.09
30.33	6.09
31.99	6.09
30.25	6.09
26.00	6.09

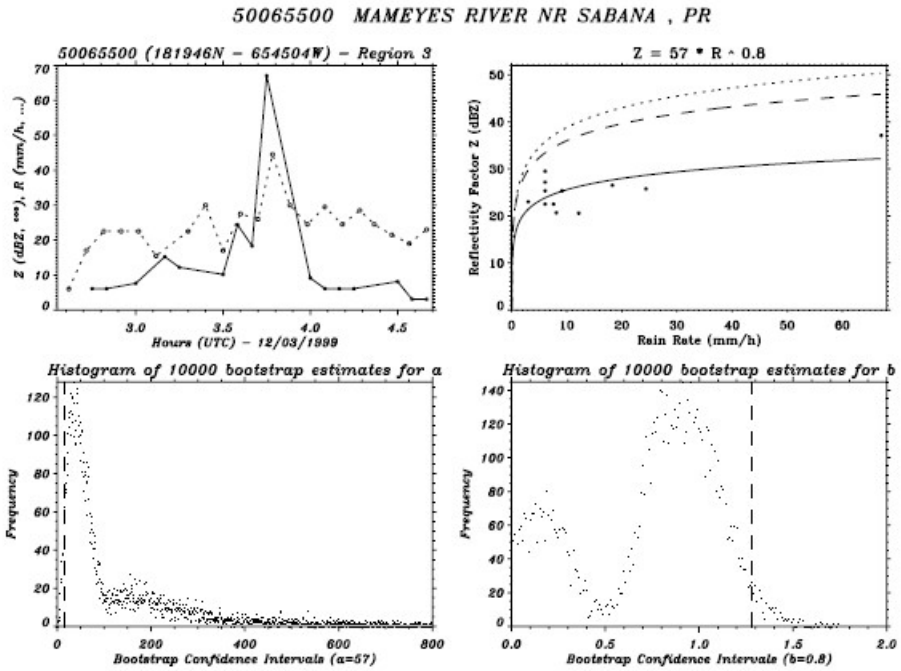

Figure 6.7: Data fit to Z−R model and bootstrap estimates in the region 3.

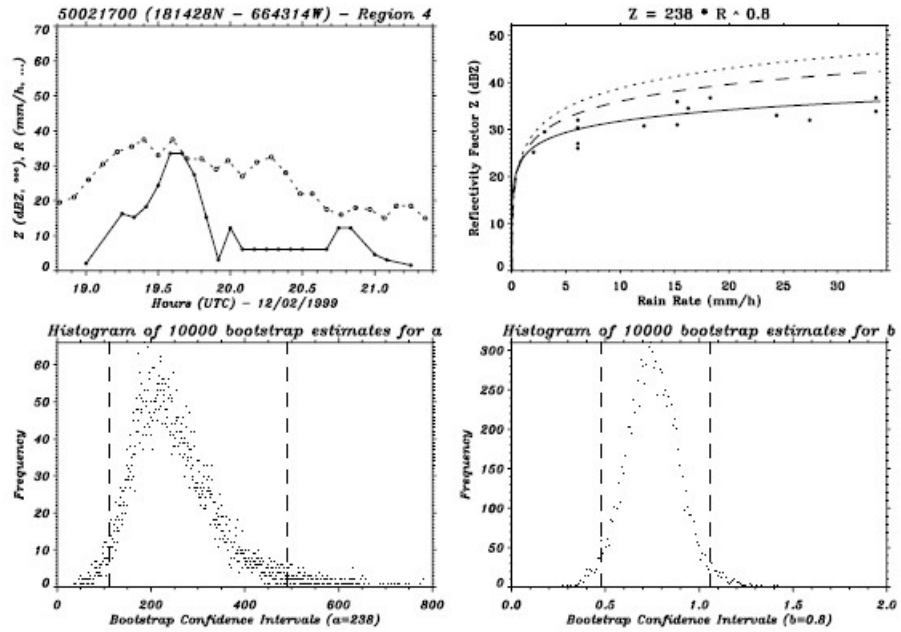

Figure 6.8: Data fit to Z−R model and bootstrap estimates in the region 4.

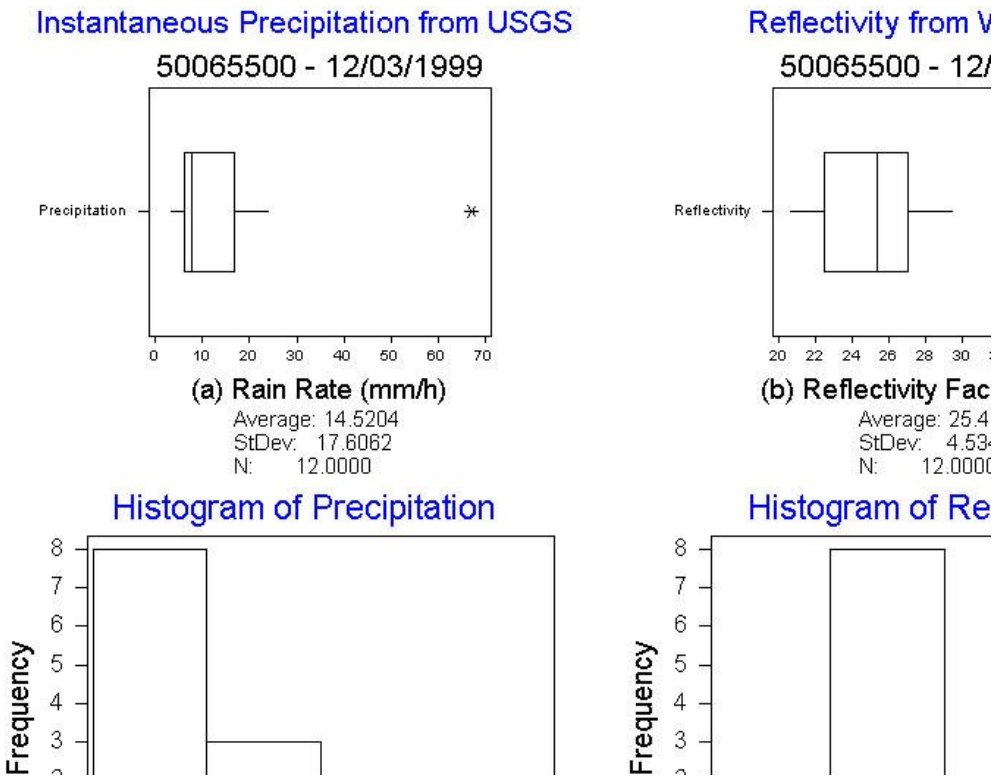

Figure 6.10: Statistical results obtained with MINITAB for region 3.

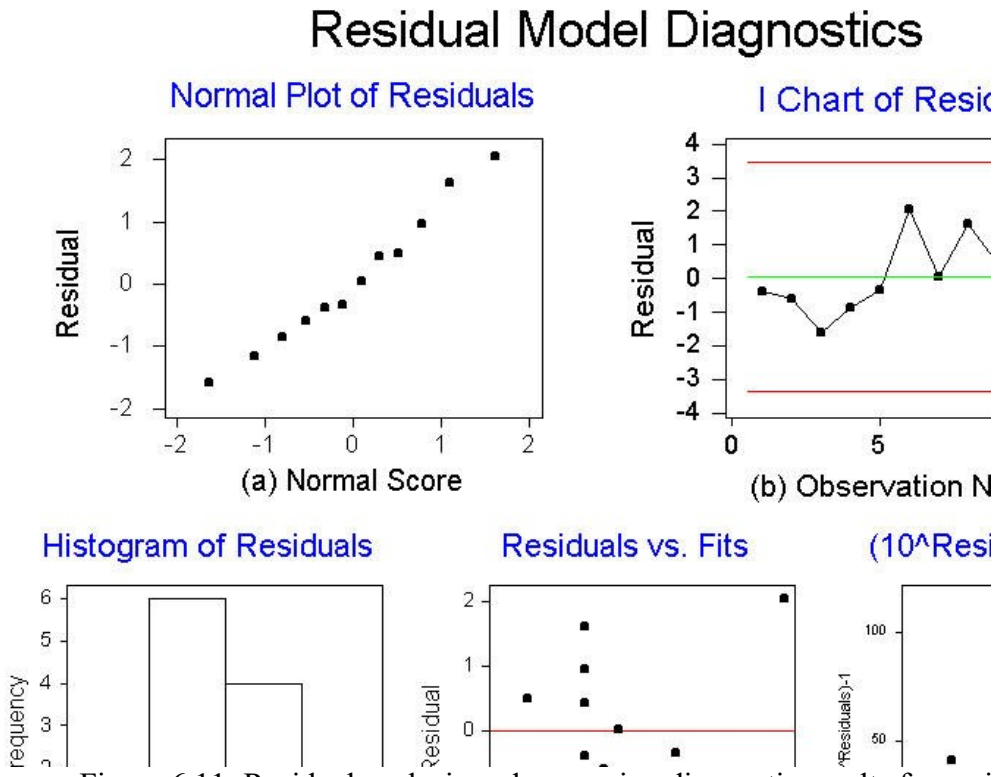

Figure 6.11: Residual analysis and regression diagnostic results for region 3.

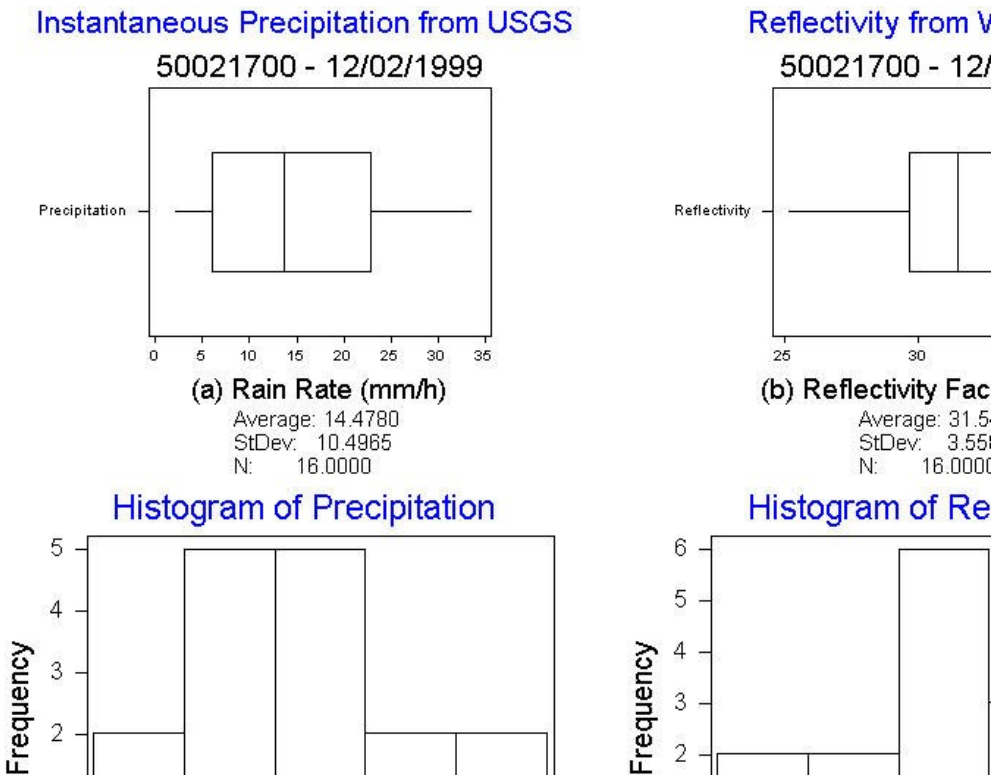

Figure 6.12: Statistical results obtained with MINITAB for region 4.

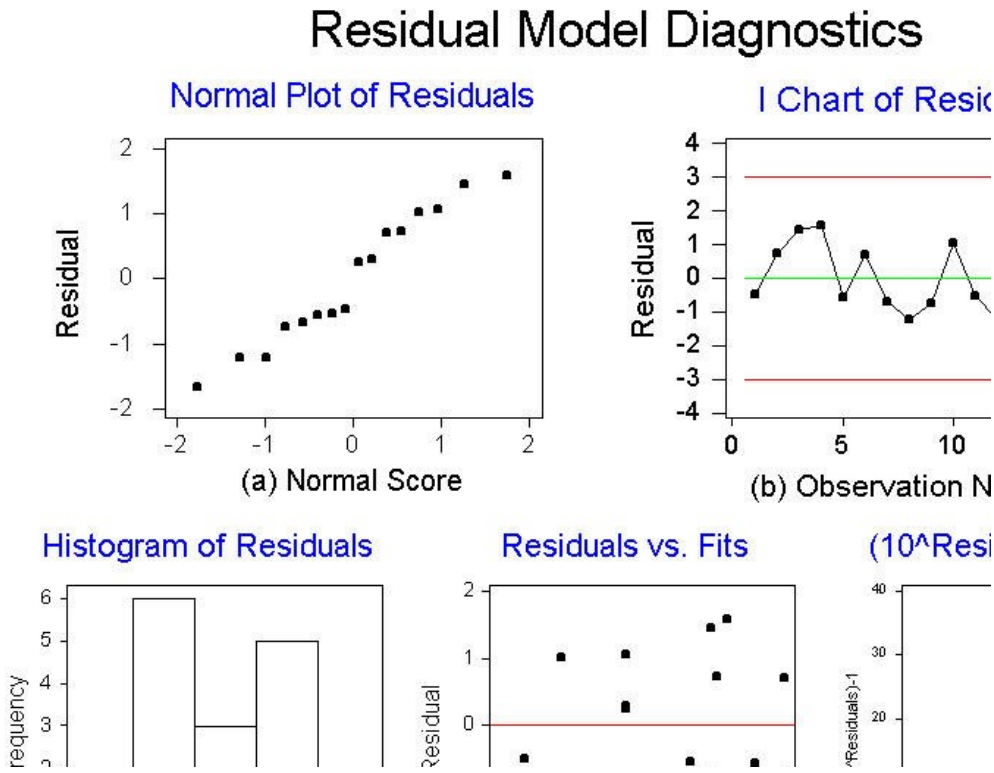

Figure 6.13: Residual analysis and regression diagnostic results for region 4.

6.5 Selected event for the region 5

The selected event for the region 5 was registered in the rain gauge station from Humacao river at Las Piedras, PR (50081000). The data were taken on December 2, 1999 from 19.8 to 22.0 hours (UTC), and with minimum reflectivity value of 29.0 dBZ (Minimum db). The data for this event are shown in Table 6.8.

Table 6.8: Data set for selected rainfall event for the region 5.

Reflectivity (dBZ)	Precipitation (mm/h)
29.25	4.06
32.99	24.38
32.17	15.24
40.00	53.34
38.67	57.91
37.42	41.15
38.50	18.29
37.50	12.19
32.90	15.24
33.50	12.19
32.33	9.14
32.25	15.24
31.33	21.34
30.42	19.81
30.00	18.29
30.42	15.24
30.83	12.19

The Calculated relationship for this event is $Z=221.6\ R^{0.81}$ and is similar to the Tropical relationship, see Figure 6.14. The coefficient of determination is 0.42, which is an acceptable value. The estimated 95% confidence interval for *a* parameter is between $ta_{2.5} = 63.14$ and $ta_{97.5} = 919.42$, for *b* parameter is between $tb_{2.5} = 0.28$ and $tb_{97.5} = 1.18$. The curve of the bootstrap for the *a* parameter is similar to the first case, and the line for the upper side of the interval is not showing because is out range. For the *b* parameter, in this case, the curve is a little asymmetric, this because of the outliers in the precipitation data.

The Figure 6.16 shows that there are outliers in precipitation data and not in the reflectivity data. The box-and-whiskers plots are showing that both precipitation and reflectivity are asymmetrical and the Figures 6.16(c) and 6.16(d) agree the box-and-whiskers plots. Figure 6.17 shows that there is no normality in the errors, as are shown in Figures 6.17(a) and 6.17(c). There are no outliers for residuals and Figure 6.17(e) shows that the model has a multiplicative regression error.

6.6 Selected event for the region 6

The selected event for the region 6 was registered in the rain gauge station at Caguitas river NR Aguas Buenas, PR (50055100). The data were taken on December 2, 1999 from 18.3 to 21.8 hours (UTC), and with minimum reflectivity value of 24.0 dBZ (Minimum db). We can see the data for this event in Table 6.9.

The equation $Z=244.23\ R^{0.90}$ is the Calculated relationship for this event and is similar to the Tropical relationship, as shown in Figure 6.15. The coefficient of determination for this event is 0.43, which is an acceptable value. The estimated 95% confidence interval for *a* parameter is between $ta_{2.5} = 55.25$ and $ta_{97.5} = 2171.73$, for *b* parameter is between $tb_{2.5} = 0.22$ and $tb_{97.5} = 1.44$. The curve of the bootstrap for *a* parameter is similar to the first case and the line for the upper side of the interval is not showing because is out of the x range, similar to the case on the region 5. For the *b* parameter, in this case, the curve is a little asymmetric and wide; this because there is a lot of variability of the estimates because of the variability of the vertical direction (reflectivity), as is shown in the Figure 6.18(b).

Figure 6.19 shows that there are no outliers in either precipitation data or reflectivity data. The box-and-whiskers plots are showing that both precipitation and reflectivity are asymmetrical and they have a lot of variability, Figures 6.16(c) and 6.16(d) agree with the box-and-whiskers plots. Figure 6.19, shows that there is no

normality in the errors, as is shown in Figures 6.19(a) and 6.19(c). There are outliers for residuals and the Figure 6.19(e) shows that the model has a multiplicative regression error.

Table 6.9: Data set for selected rainfall event for the region 6.

Reflectivity (dBZ)	**Precipitation (mm/h)**
42.17	3.05
41.50	30.48
41.00	39.62
41.90	36.58
41.58	42.67
38.67	45.72
33.50	39.62
30.99	33.53
33.19	18.29
35.99	10.67
27.25	6.09
27.00	4.57
26.67	3.05
24.33	3.05
24.25	3.05

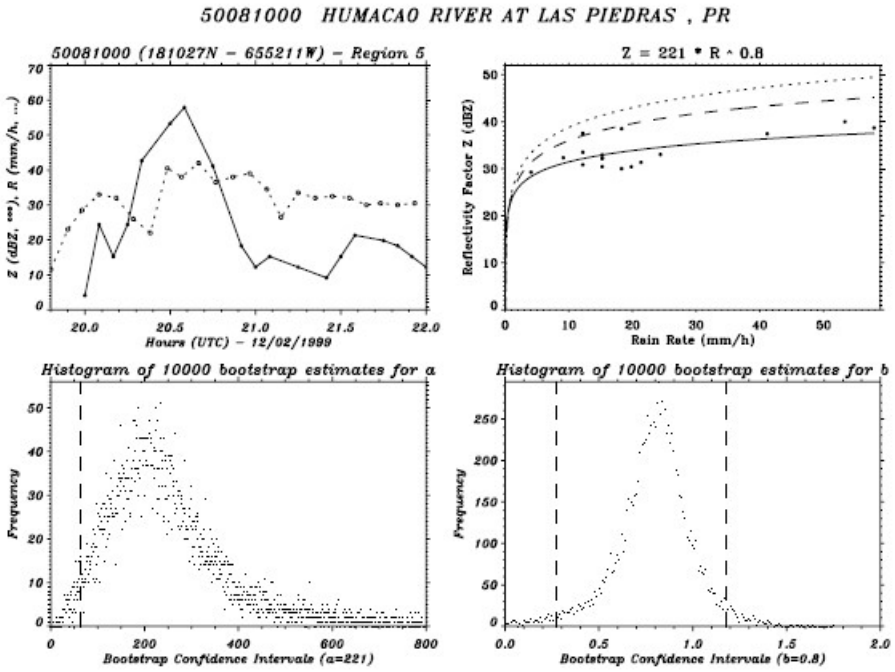

Figure 6.13: Data fit to Z−R model and bootstrap estimates in the region 5.

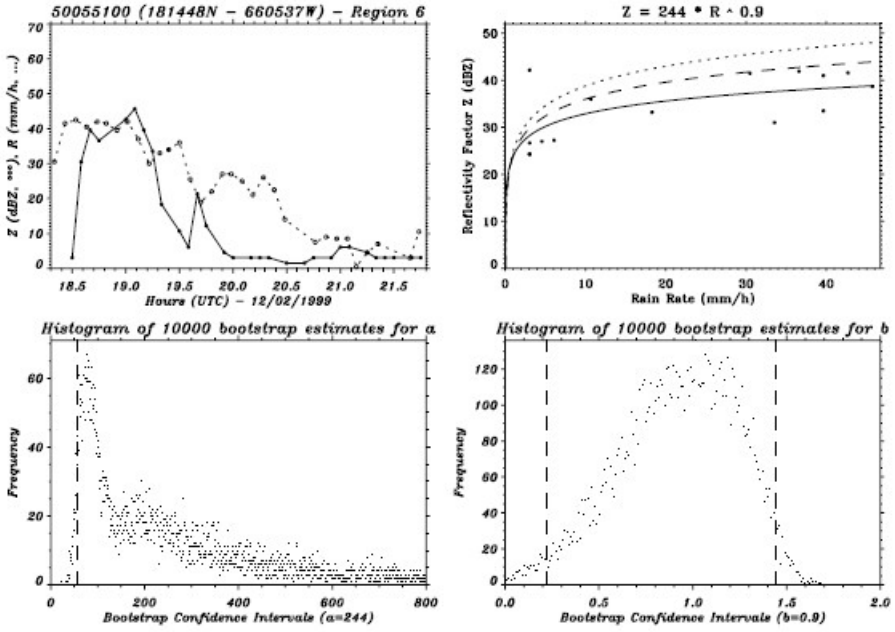

Figure 6.14: Data fit to Z−R model and bootstrap estimates in the region 6.

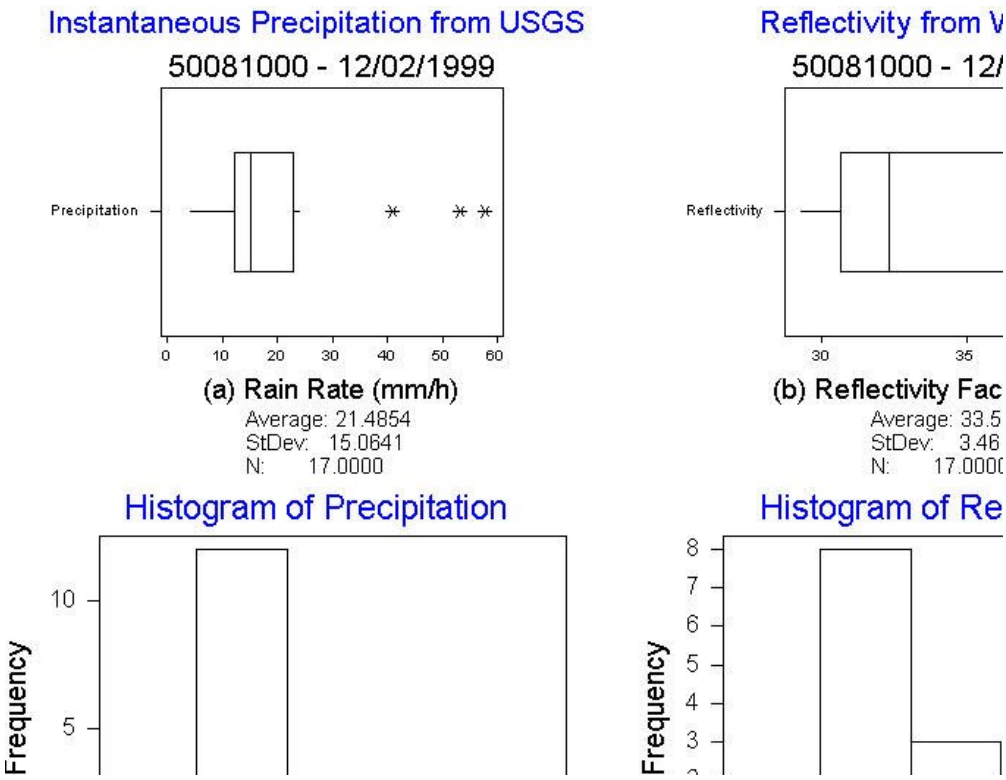

Figure 6.16: Statistical results obtained with MINITAB for region 5.

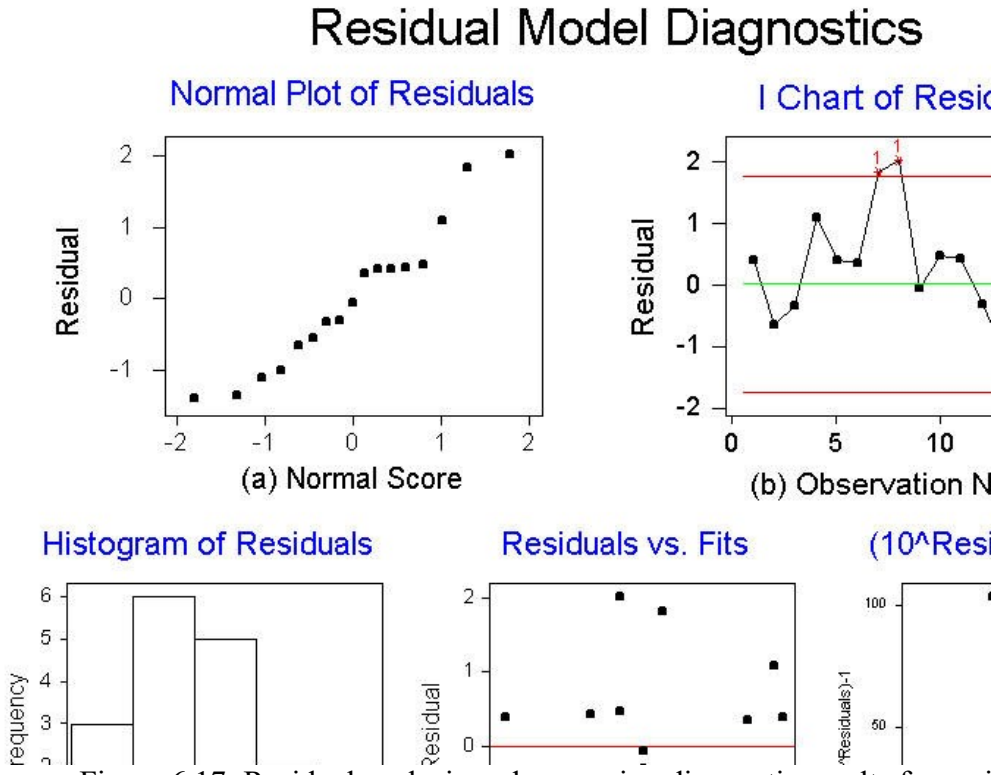

Figure 6.17: Residual analysis and regression diagnostic results for region 5.

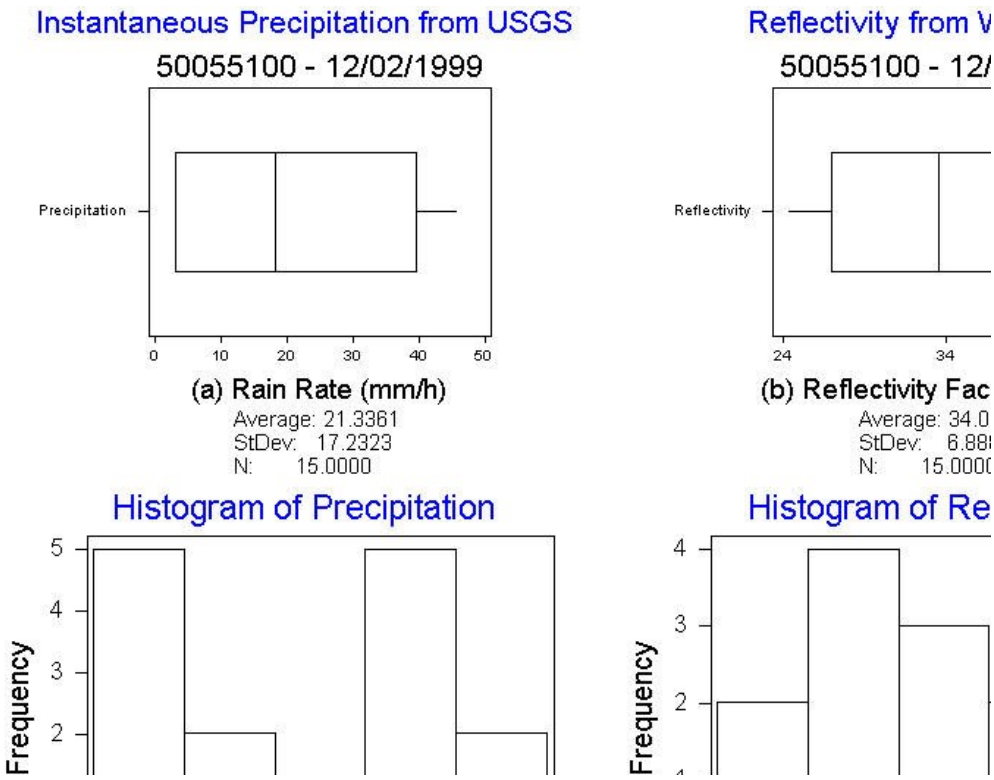

Figure 6.18: Statistical results obtained with MINITAB for region 6.

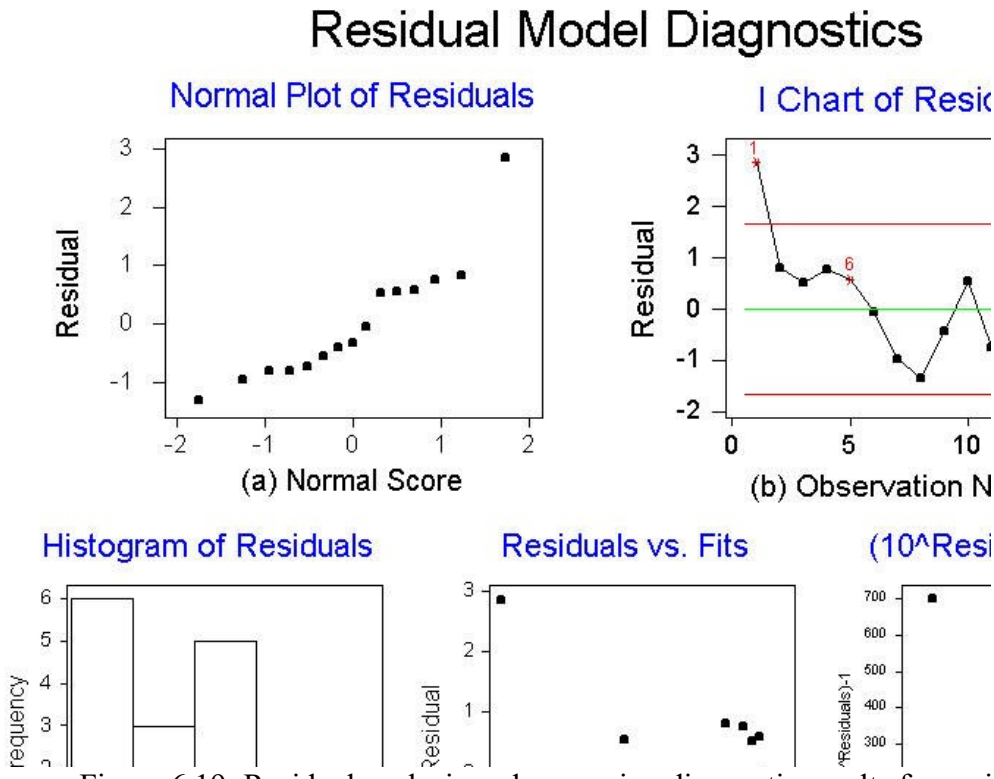

Figure 6.19: Residual analysis and regression diagnostic results for region 6.

6.7 Comparison of NEXRAD and TRMM data

In this section we show further comparisons between data from NEXRAD and TRMM-PR (PR here stand for Precipitation Radar). We found a coincidence in both NEXRAD and TRMM measurements as shown in Figures 6.20 and 6.21. Figure 6.20 depicts reflectivity data derived from WATADS precipitation processing subsystem. The data were registered December 02, 1999 at 16:23 hours UTC by the NEXRAD system. Figure 6.21 depicts reflectivity data registered on December 02, 1999 from 16:20:32 to 16:21:41 hours UTC by the TRMM-PR satellite system over the Caribbean Region. As expected, there are few satellite passes over Puerto Rico coinciding with our study dates.

In the graph we can see the similarity between the NEXRAD and TRMM radars measurements; in both, we can see the reflectivity signals from rain along the North coast of Puerto Rico. The values of the reflectivity are similar in both cases, and they lie between 30 and 60 dBz. In Figure 6.20 we can see clearly the clutter around the radar position. Both figures show the rain delimiting the boundary of the mountains.

In Figure 6.21, we can see reflectivity in from of the North-West of the island, but in the Figure 6.20 is no present, this is because it is outside of the actual target range from the NEXRAD.

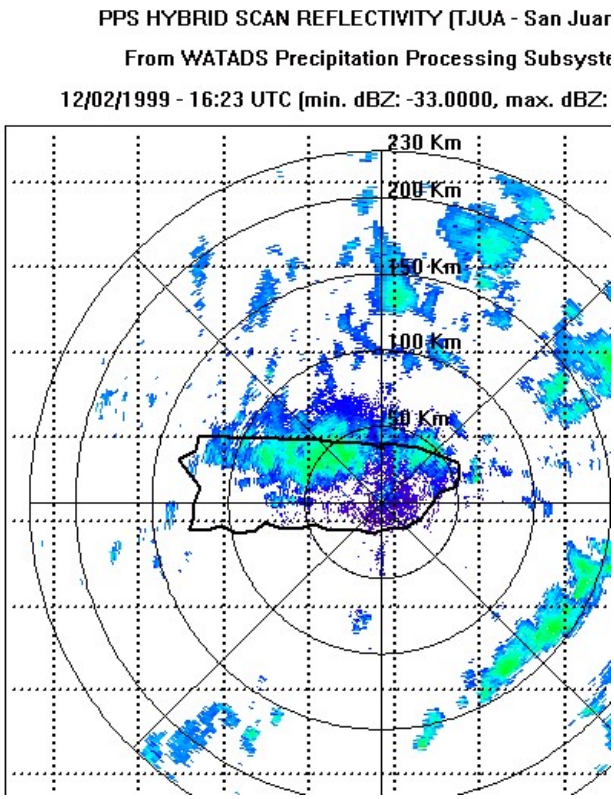

Figure 6.20: Scan reflectivity (Z) from NEXRAD.

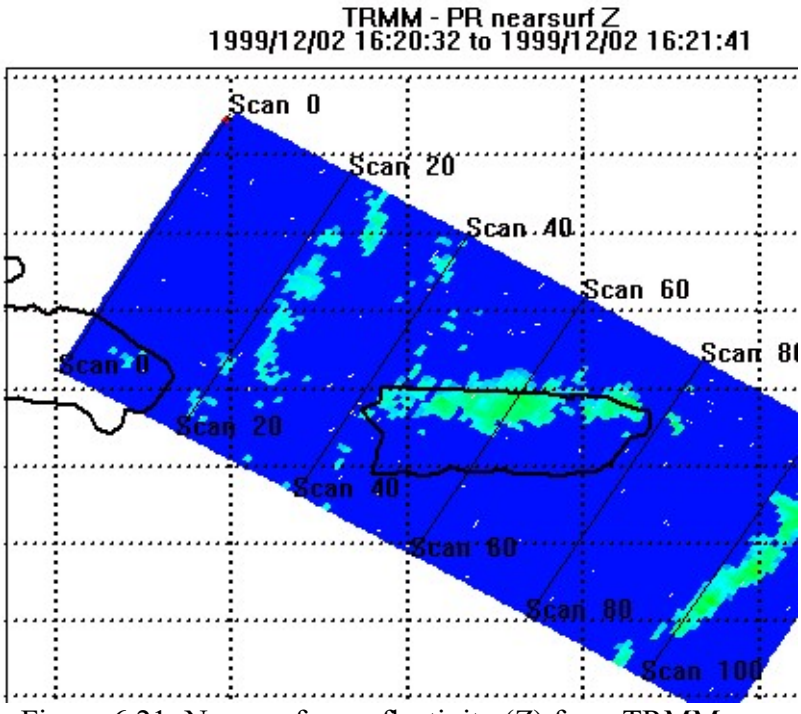

Figure 6.21: Near surface reflectivity (Z) from TRMM.

CHAPTER VII

CONCLUSIONS

In the present work, we have developed a Rainfall Processing environment (RAIN_PROC) using the IDL numerical and graphical computation software package. This computer environment is a useful tool to analyze, compare, and visualize observational data of precipitation from the NEXRAD system in Cayey, Puerto Rico (NWS and FAA), rain gauges operated by the USGS on the island, and from the TRMM-PR satellite (NASA). This computing environment is capable of manipulating, and assimilating data from these sources and is ready to be used with a more extensive database. We have transformed the original data formats to HDF format in order to have a standard format for all data sets form each source and made the access faster and direct.

RAIN_PROC uses statistical methods (e.g. bootstrap) making this software package a useful tool for statistical analyses and for the estimation of the Z-R relation using linear regression. Our software makes it possible look at the rainfall data in two ways, textual and graphical through maps, images, and lists. We can exchange information with the MINITAB statistical program to perform more complex statistical analyses.

Table 7.1 shows that the fitting process applied to the six selected events - one by region - provided values for the *a* parameter between 56.65 and 244.23, and for the *b*

parameter between 0.75 and 1.12. These values were calculated considering minimum values of reflectivity in the application of the regression technique. According to our results we conclude that the parameter values are suggestive of a Tropical relationship, except in the region 3 case. A more definitive conclusion cannot be arrived at without bringing in a more extensive data set.

Table 7.1: Parameters calculated for the six selected rainfall events.

a	B	R^2	N	Region	Station Id.	Date	Min. dBZ
145.91	1.12	0.57	18	1	50051180	19991117	17.0
138.83	0.79	0.32	37	2	50039500	19991202	20.0
56.65	0.80	0.39	12	3	50065500	19991203	20.0
237.61	0.75	0.61	16	4	50021700	19991202	25.0
221.47	0.81	0.43	17	5	50081000	19991202	29.0
244.23	0.91	0.43	15	6	50055100	19991202	24.0

In Table 7.2 we have the percentiles of 2.5% and 97.5% for the bootstrap confidence intervals for *a*, and *b* parameters calculated. We have the observed bootstrap values and their estimation errors (EE). The table shows that the values of EE are large in some cases; this is due to the small number of samples collected per event.

Table 7.2: Bootstrap confidence intervals for *a,* and *b* parameters of the selected events and their estimation errors (EE).

Region	Station Id.	N	$ta_{2.5}$	$ta_{97.5}$	EE(a)	$tb_{2.5}$	$tb_{97.5}$	EE(b)
1	50051180	18	40.93	370.68	164.87	0.66	1.91	0.62
2	50039500	37	67.28	322.98	127.85	0.38	1.13	0.37
3	50065500	12	14.36	908.25	446.94	-0.55	1.28	0.91
4	50021700	16	110.16	491.85	190.84	0.48	1.06	0.29
5	50081000	17	63.14	919.42	428.14	0.28	1.18	0.45
6	50055100	15	55.25	2171.73	1058.24	0.22	1.44	0.61

The Calculated relationship for the event in the Region 4 is $Z=237.6\ R^{0.75}$ and it is similar to the Tropical relationship, see Table 7.1. These numbers do not agree with

Ulbrich's suggestion that the rain characteristic near the center of the island obeys a continental relationship (Ulbrich, Petitididier, and Campos, 1999). The coefficient of determination is 0.61; this is a fairly good value. The estimated 95% confidence interval for the *a* parameter is between 110.16 ($ta_{2.5}$) and 491.85 ($ta_{97.5}$), and for the *b* parameter is between 0.48 ($tb_{2.5}$) and 1.06 ($tb_{97.5}$).

From the regression errors plot, shown in the graph for $10^{Residuals}$-1 versus the expected value of reflectivity (E[y]), we can see the "funnel effect" indicating that as the response variable gets larger, the deviation of the regression error become smaller. This means that the model has a multiplicative regression error (see Appendix C). Also, for small values of precipitation the reflectivity values have large deviations, but for large values of precipitation, the variability of the reflectivity becomes smaller. Hence, the error variance is not constant, but it decreases when the precipitation values increase.

We have looked at our data to determine if any dependence exists between radar rainfall estimates and the distance from the NEXRAD. No systematic trend was found in contrast with Vásquez and Roche (private communication, 2001). The letter work obtains the Z-R relationship from storm totals. This is, in our view, the weak part of their work.

CHAPTER VIII

RECOMMENDATIONS

The approach outlined in this work should give effective results when it is applied to a more comprehensive data set with coincident observations. We suggest observation campaigns considering other NEXRAD modes of operation and longer periods of time that cover seasonal behavior, types of rain, its strength, etc.

An extensive investigation of the NEXRAD rainfall estimation process will require considering other parameters like turbulence of the wind, and the effects of the topography in order to improve the Z-R relationship because of the discrepancies between the radar and rain gauges measurements. The stochastic character of the rainfall has to be recognized and accounted for (Jameson and Kostinski, 2001). Differences of course arise due to the volumetric estimation with the radar and the point estimation with the rain gauges.

It is important to evaluate the radar performance versus distances. We suggest to look at new, and larger data sets, and look for this effect. We think that it is very important to effect NEXRAD observations on a campaign basis, with a well calibrated radar. We also recommend to collect satellite images of clouds coverage others to improve the data analyses and interpretation.

RAIN_PROC could be modified and made it more user friendly. For example, it can include selection with the mouse of the rain events by time and intensity. Additional graphical buttons to adjust for logarithmic, normalized, or linear scales could be included. RAIN_PROC capabilities could be enhance to include comparison not only over the rain gauge position but over cells in the neighborhood.

BIBLIOGRAPHY

Atlas, D., and A. Chmela. 1957. Physical-synoptic variations of drop size parameters. Proc. Weather Radar Conf. pp. 21-30.

Battan, L. J. 1973. Radar Observation of the Atmosphere. Univ. of Chicago Press, Chicago, Illinois.

Carter, M. M., and J. Elsner. 1997. A Statistical Method for Forecasting Rainfall over Puerto Rico. American Meteorological Society.

Carter, M. M. and J. Elsner. 1995. Monthly Rainfall Climatology for Puerto Rico. 21st Annual Hurricane and Tropical Meteorology Conference.

Davison, A. V., and D. Hinkley. 1997. Bootstrap Methods and their Application. Cambridge University Press, Cambridge, New York.

Doviak, R. J., and D. Zrnic. 1993. Doppler Radar and Weather Observations. Academic Press, Inc., Second Edition.

Doviak, R. J. 1983. A survey of radar rain measurement techniques, J. Appl. Meteorol. 22, pp. 832-849.

Efrom, B., and R. Tibshirani. 1993. An Introduction to the Bootstrap, Chapman & Hall.

Hoel, P. G. 1971. Introduction to Mathematical Statistics. John Wiley & Sons, Inc., Fourth Edition.

Huff, F. A. 1970. Sampling errors in measurement of mean precipitation, J. Appl. Meteorol. 9, pp. 35-44.

Jameson, A. R., and A. B. Kostinski. 2001. What is a raindrop size distribution?. Bull. Amer. Meteor. Soc., 82, 1169-1177.

Lutgens, F. K., and E. Tarbuck. 1998. The Atmosphere. An Introduction to Meteorology. Prentice-Hall, Inc. Seventh Edition.

Myers, R. H. 1990. Classical and Modern Regression with Applications. Brooks/Cole. Second Edition.

Strangeways, I. 2000. Measuring the Natural Environment. Cambridge University Press.

Tropical Rainfall Measuring Mission (TRMM), Science and Information System, Real time System File Specification, Release 1.05, Prepared for: National Aeronautics and Space Administration, Goddard Space Flight Center, May 28, 1999.

Tropical Rainfall Measuring Mission (TRMM), Science and Information System, Tutorial for Reading Tropical Rainfall Measuring Mission (TRMM) Data Products, Release 1.1, Prepared for: National Aeronautics and Space Administration, Goddard Space Flight Center, December 4, 1998.

Tropical Rainfall Measuring Mission (TRMM). Science and Information System. Interface Control Specification Between the Tropical Rainfall Measuring Mission Data, and Information System (TSDIS) and the TSDIS Science User (TSU), TSDIS-P907, Volume 4: File Specifications for TRMM Products – Level 2 and Level 3, Release 5.07, Prepared for: National Aeronautics and Space Administration, Goddard Space Flight Center, April 14, 2000.

Ulbrich, C., M. Petitididier, and E. Campos. 1999. Radar Properties of Tropical Rain found from Disdrometer data at Arecibo, PR. 29th Radar Meteorology. American Meteorological Society, pp. 676-679.

Vetterling, W. T., S. Teukolsky, W. Press, and B. Flannery. 1997. Numerical Recipes in C. The Art of Scientific Computing. Cambridge University Press. Second Edition. Cambridge, New York.

WSR-88D Algorithm Testing and Display System (WATADS), Version 10.0, Reference Guide.

WSR-88D Algorithm Testing and Display System (WATADS), Version 10.0, Algorithm Reference Guide, a publication for applications branch, February 1998.

APPENDICES

APPENDIX A

SOUNDING DATA

In this appendix we have the Radiosonde data from Puerto Rico station at San Juan (JSJ). The geographical position of the station is 18.43º North Latitude, and 66º West Longitude. In the tables we can see values for wind direction, wind speed, and high in both meters and knots.

There are data for the following dates and hours (UTC):

- October 15, 1998 – 00:00pm
- October 15, 1998 – 12:00m
- September 30, 1998 – 00:00pm
- November 4, 1999 – 12:00m
- November 7, 1999 – 00:00pm
- November 15, 1999 – 00:00pm
- November 17, 1999 – 06:00am
- November 23, 1999 – 12:00m
- December 2, 1999 – 12:00
- November 19, 1999 – 06:00am
- November 21, 1999 – 00:00pm

For more information about the data and the station refer to the following web page: http://www.fsl.noaa.gov/docs/data/fsl-data.html.

Table A.1: October 15, 1998 – 00:00pm

High-Meters	Direction	Speed	High-knots
10	160	8	32.8
413	160	12	1354.64
1000	155	17	3280
2000	160	16	6560
2665	160	14	8741.2
3000	160	13	9840
4000	160	11	13120
5062	125	5	16603.36
6000	120	4	19680
7000	155	6	22960
8000	150	9	26240
9000	140	8	29520
10430	135	6	34210.4
12000	195	3	39360
14000	260	3	45920
16000	225	2	52480
19300	195	8	63304
20000	200	7	65600
24870	155	8	81573.6
25000	155	8	82000
30000	30	2	98400
31725	35	2	104058
35000	325	7	114800
35860	320	5	117620.8
40650	50	7	133332
46000	250	11	150880
46523	265	12	152595.44
48000	270	13	157440
50000	250	10	164000
52000	215	11	170560

Table A.2: October 15, 1998 – 12:00m

High-Meters	Direction	Speed	High-knots
456	100	9	1495.68
1000	95	14	3280
2000	90	13	6560
2700	105	13	8856
3000	105	13	9840
4000	110	12	13120
5000	120	11	16400
5095	130	10	16711.6
6000	135	9	19680
7000	125	8	22960
8000	120	6	26240
9000	125	8	29520
10456	110	9	34295.68
12000	90	10	39360
14000	85	4	45920
16000	175	2	52480
19324	185	7	63382.72
20000	175	8	65600
24900	25	6	81672
25000	25	7	82000
30000	15	9	98400
31000	30	11	101680
31760	30	11	104172.8
35000	10	12	114800
35892	10	11	117725.76
36000	10	11	118080
38000	340	14	124640
40683	350	12	133440.24
42000	355	11	137760
43000	300	11	141040

Table A.3: September 30, 1998 – 00:00pm

High-Meters	Direction	Speed	High-knots
10	210	4	32.8
3000	175	7	9840
4000	165	7	13120
5070	170	6	16629.6
6000	195	6	19680
7000	210	5	22960
8000	200	4	26240
9000	160	4	29520
10430	160	5	34210.4
12000	220	7	39360
14000	210	8	45920
16000	235	8	52480
19325	320	5	63386
20000	335	5	65600
24900	335	1	81672
25000	325	1	82000
30000	15	7	98400
31725	5	7	104058
35000	335	1	114800
35860	300	3	117620.8
40682	335	8	133436.96
46522	45	11	152592.16
47000	45	11	154160
50000	335	5	164000
52000	25	12	170560
54000	40	17	177120
54265	40	17	177989.2
55000	50	17	180400
57000	105	20	186960
58000	115	20	190240

Table A.4: November 4, 1999 – 12:00m

High-Meters	Direction	Speed	High-knots
3	95	5	9.84
138	100	13	452.64
304	110	20	997.12
609	110	20	1997.52
823	110	20	2699.44
914	110	19	2997.92
1219	105	18	3998.32
1552	100	19	5090.56
1828	100	19	5995.84
2133	105	16	6996.24
2438	110	12	7996.64
2743	105	10	8997.04
3186	95	12	10450.08
3352	105	11	10994.56
3657	95	8	11994.96
4267	50	12	13995.76
4876	20	13	15993.28
5880	10	17	19286.4
6096	10	19	19994.88
6400	5	25	20992
7010	20	35	22992.8
7580	15	38	24862.4
7620	15	39	24993.6
8229	360	40	26991.12
9144	360	39	29992.32
9670	5	43	31717.6
10668	5	50	34991.04
10930	5	54	35850.4
11277	5	57	36988.56
12390	10	48	40639.2

Table A.5: November 7, 1999–00:00pm

High-Meters	Direction	Speed	High-knots
3	90	6	9.84
122	85	9	400.16
304	80	14	997.12
609	70	14	1997.52
804	90	14	2637.12
914	95	16	2997.92
1219	110	21	3998.32
1530	105	20	5018.4
1828	100	17	5995.84
2133	115	14	6996.24
2438	120	14	7996.64
2743	85	13	8997.04
3157	85	14	10354.96
3657	80	13	11994.96
4267	125	16	13995.76
4572	130	18	14996.16
4876	115	19	15993.28
5486	125	16	17994.08
5791	100	11	18994.48
5860	100	10	19220.8
6096	100	9	19994.88
7315	125	11	23993.2
7570	120	10	24829.6
7620	120	10	24993.6
8229	85	13	26991.12
9144	110	19	29992.32
9448	120	20	30989.44
9650	120	19	31652
10363	125	16	33990.64
10668	140	10	34991.04

Table A.6: November 15, 1999 – 00:00pm

High-Meters	Direction	Speed	High-knots
3	90	6	9.84
122	85	9	400.16
304	80	14	997.12
609	70	14	1997.52
804	90	14	2637.12
914	95	16	2997.92
1219	110	21	3998.32
1530	105	20	5018.4
1828	100	17	5995.84
2133	115	14	6996.24
2438	120	14	7996.64
2743	85	13	8997.04
3157	85	14	10354.96
3657	80	13	11994.96
4267	125	16	13995.76
4572	130	18	14996.16
4876	115	19	15993.28
5486	125	16	17994.08
5791	100	11	18994.48
5860	100	10	19220.8
6096	100	9	19994.88
7315	125	11	23993.2
7570	120	10	24829.6
7620	120	10	24993.6
8229	85	13	26991.12
9144	110	19	29992.32
9448	120	20	30989.44
9650	120	19	31652
10363	125	16	33990.64
10668	140	10	34991.04

Table A.7: November 17, 1999 – 06:00am

High-Meters	Direction	Speed	High-knots
3	90	10	9.84
30	90	12	98.4
304	85	28	997.12
609	95	30	1997.52
711	100	30	2332.08
914	110	28	2997.92
1219	125	31	3998.32
1439	125	38	4719.92
1828	135	44	5995.84
2133	135	44	6996.24
2438	135	43	7996.64
2743	135	35	8997.04
3076	135	30	10089.28
3657	135	31	11994.96
3962	140	25	12995.36
4267	135	26	13995.76
4876	125	19	15993.28
5486	100	14	17994.08
5800	105	17	19024
6096	105	20	19994.88
7530	90	18	24698.4
7620	95	18	24993.6
7924	95	18	25990.72
9144	100	15	29992.32
9650	100	18	31652
10058	115	28	32990.24
10363	140	14	33990.64
10668	150	20	34991.04
10930	140	28	35850.4
12410	150	39	40704.8

Table A.8: November 19, 1999 – 06:00am

High-Meters	Direction	Speed	High-knots
3	10	10	9.84
76	10	13	249.28
304	10	20	997.12
609	25	16	1997.52
759	25	16	2489.52
914	20	16	2997.92
1219	10	21	3998.32
1487	15	23	4877.36
1828	15	18	5995.84
2133	10	18	6996.24
2438	20	21	7996.64
2743	25	22	8997.04
3048	10	28	9997.44
3128	10	28	10259.84
3657	20	30	11994.96
3962	25	31	12995.36
4267	25	24	13995.76
4876	35	18	15993.28
5791	35	21	18994.48
5870	35	21	19253.6
6096	30	23	19994.88
7580	335	32	24862.4
7620	330	30	24993.6
8839	340	36	28991.92
9144	340	42	29992.32
9680	330	44	31750.4
9753	330	36	31989.84
10668	280	20	34991.04
10950	290	22	35916
10972	295	23	35988.16

Table A.9: November 21, 1999 – 00:00pm

High-Meters	Direction	Speed	High-knots
3	330	10	9.84
53	330	10	173.84
304	325	12	997.12
609	315	12	1997.52
737	320	12	2417.36
914	320	11	2997.92
1219	325	11	3998.32
1466	325	11	4808.48
1828	345	11	5995.84
2133	355	11	6996.24
2438	345	14	7996.64
2743	340	15	8997.04
3048	320	20	9997.44
3096	320	21	10154.88
3657	320	33	11994.96
4267	325	27	13995.76
4876	320	24	15993.28
5486	310	22	17994.08
5810	310	20	19056.8
6096	325	18	19994.88
6400	340	20	20992
7520	310	41	24665.6
7620	310	37	24993.6
8534	330	33	27991.52
9144	315	36	29992.32
9610	300	32	31520.8
10058	295	32	32990.24
10668	295	26	34991.04
10880	295	27	35686.4
12192	270	40	39989.76

Table A.10: November 23, 1999 – 12:00m

High-Meters	Direction	Speed	High-knots
3	0	0	9.84
77	335	5	252.56
304	275	18	997.12
609	285	16	1997.52
760	275	13	2492.8
914	280	12	2997.92
1219	280	12	3998.32
1488	280	12	4880.64
1524	280	12	4998.72
1828	265	11	5995.84
2133	255	11	6996.24
2438	235	10	7996.64
2743	245	12	8997.04
3113	255	10	10210.64
3657	245	13	11994.96
4267	285	22	13995.76
4876	260	28	15993.28
5820	255	34	19089.6
6096	250	36	19994.88
7520	265	36	24665.6
7620	265	37	24993.6
8534	255	35	27991.52
9144	295	44	29992.32
9448	290	46	30989.44
9620	290	44	31553.6
10668	270	35	34991.04
10890	260	33	35719.2
10972	255	32	35988.16
11582	250	44	37988.96
11887	270	30	38989.36

Table A.11: December 2, 1999 – 12:00m

High-Meters	Direction	Speed	High-knots
3	0	0	9.84
121	45	2	396.88
304	90	4	997.12
609	85	4	1997.52
800	100	5	2624
914	105	6	2997.92
1219	135	7	3998.32
1526	165	5	5005.28
1828	185	3	5995.84
2133	155	3	6996.24
2438	100	1	7996.64
2743	200	1	8997.04
3154	185	8	10345.12
3657	190	15	11994.96
3962	205	16	12995.36
4267	225	14	13995.76
4876	280	11	15993.28
5791	320	11	18994.48
5850	315	10	19188
6096	260	7	19994.88
6400	150	14	20992
7550	200	24	24764
7620	205	25	24993.6
7924	210	27	25990.72
9144	225	36	29992.32
9448	230	38	30989.44
9640	230	40	31619.2
10668	225	47	34991.04
10890	220	48	35719.2
11277	220	47	36988.56

APPENDIX B

RAIN GAGES LIST

ID	STATION NAME	LAT	LON .
50010500	GUAJATACA RIVER AT LARES	18º18'01"	66º52'24"
50010800	GUAJATACA LAKE AT DAMSITE NR QUEBRADILLAS	18º24'02"	66º55'25"
50014800	CAMUY RIVER NR BAYANEY	18º23'48"	66º49'03"
50020100	GARZAS LAKE NR ADJUNTAS	18º08'20"	66º44'29"
50021700	RIO GRANDE OF ARECIBO ABV UTUADO	18º14'28"	66º43'14"
50024950	RIO GRANDE OF ARECIBO BLW UTUADO	18º18'07"	66º42'15"
50025155	SALIENTE RIVER AT COABEY NR JAYUYA	18º12'48"	66º33'49"
50026025	CAONILLAS RIVER AT PASO PALMA	18º13'53"	66º38'14"
50026140	CAONILLAS LAKE AT DAMSITE NR UTUADO	18º16'43"	66º39'24"
50027100	DOS BOCAS LAKE AT DAMSITE	18º20'16"	66º40'05"
50027750	RIO GRANDE OF ARECIBO ABV ARECIBO	18º25'29"	66º41'44"
50028400	TANAMA RIVER AT CHARCO HONDO	18º24'52"	66º42'52"
50029000	RIO GRANDE OF ARECIBO AT CENTRAL CAMBALACHE	18º27'20"	66º42'10"
50030460	OROCOVIS RIVER NR OROCOVIS	18º13'25"	66º23'34"
50031200	RIO GRANDE OF MANATI NR MOROVIS	18º17'45"	66º24'47"
50032290	EL GUINEO LAKE AT DAMSITE NR VILLALBA	18º09'41"	66º31'38"
50032590	LAKE OF MATRULLAS AT DAMSITE NR OROCOVIS	18º12'45"	66º28'50"
50034000	BAUTA RIVER NR OROCOVIS	18º14'10"	66º27'18"
50035000	RIO GRANDE OF MANATI AT CIALES	18º19'26"	66º27'36"
50038100	RIO GRANDE OF MANATI AT HWY 2 NR MANATI	18º25'52"	66º31'37"
50038320	CIBUCO RIVER BLW COROZAL	18º21'13"	66º20'07"
50039500	CIBUCO RIVER AT VEGA BAJA	18º26'53"	66º22'29"
50039990	CARITE LAKE AT GATE TOWER NR CAYEY	18º03'46"	66º05'58"
50043000	DE LA PLATA RIVER AT PROYECTO LA PLATA	18º09'37"	66º13'44"
50043800	DE LA PLATA RIVER AT COMERIO	18º13'23"	66º13'30"
50044830	GUADIANA RIVER AT GUADIANA	18º18'30"	66º13'29"
50045000	LA PLATA LAKE AT DAMSITE NR TOA ALTA	18º20'40"	66º14'10"
50045010	DE LA PLATA RIVER BLW LA PLATA DAM	18º20'45"	66º14'12"

ID	STATION NAME	LAT	LON .
50046000	DE LA PLATA RIVER AT HWY 2 AT TOA ALTA	18º24'41"	66º15'39"
50047550	CIDRA LAKE AT DAMSITE NR CIDRA	18º11'57"	66º08'29"
50047560	RIVER OF BAYAMON BLW LAKE OF CIDRA	18º12'04"	66º08'26"
50047850	RIVER OF BAYAMON NR BAYAMON	18º20'08"	66º08'13"
50048680	LAS CURIAS LAKE AT DAMSITE NR PIEDRAS RIVER	18º24'23"	66º03'05"
50048690	QUEBRADA LAS CURIAS BLW LAS CURIAS OUTFALL	18º20'42"	66º03'08"
50048770	PIEDRAS RIVER AT EL SENORIAL	18º21'51"	66º03'56"
50049100	PIEDRAS RIVER AT HATO REY	18º24'34"	66º04'10"
50050900	RIO GRANDE OF LOIZA AT QUEBRADA ARENAS	18º07'10"	65º59'22"
50051150	QUEBRADA BLANCA AT EL JAGUAL	18º09'40"	65º58'58"
50051180	QUEBRADA SALVATIERRA NR SAN LORENZO	18º10'24"	65º58'38"
50051310	CAYAGUAS RIVER AT CERRO GORDO	18º09'27"	65º57'29"
50051800	RIO GRANDE LOIZA AT HWY 183 AT SAN LORENZO	18º11'09"	65º57'42"
50053025	TURABO RIVER ABOVE BORINQUEN	18º09'35"	66º02'26"
50055000	RIO GRANDE OF LOIZA AT CAGUAS	18º14'33"	66º00'34"
50055100	CAGUITAS RIVER NR AGUAS BUENAS	18º14'48"	66º05'37"
50055225	CAGUITAS RIVER AT VILLA BLANCA	18º14'55"	66º01'40"
50055390	BAIROA RIVER AT BAIROA	18º15'32"	66º02'24"
50055750	GURABO RIVER BLW EL MANGO	18º14'02"	65º53'07"
50056400	VALENCIANO RIVER NR JUNCOS	18º12'58"	65º55'34"
50057000	GURABO RIVER AT GURABO	18º15'30"	65º58'05"
50058350	CANAS RIVER AT CANAS RIVER	18º17'41"	66º02'44"
50059000	LOIZA LAKE AT DAMSITE NR TRUJILLO ALTO	18º19'49"	66º01'00"
50059050	RIO GRANDE OF LOIZA BLW DAMSITE	18º20'33"	66º00'20"
50061800	CANOVANAS RIVER NR CAMPO RICO	18º19'08"	65º53'21"
50063800	ESPIRITU SANTO RIVER NR EL VERDE	18º21'37"	65º48'49"
50064200	RIO GRANDE NR EL VERDE	18º20'42"	65º50'30"
50065500	MAMEYES RIVER NR SABANA	18º19'46"	65º45'04"
50065549	BISLEY METEOROLOGICAL STATION	18º18'16"	65º44'46"
50066000	MAMEYES RIVER AT MAMEYES	18º22'27"	65º45'50"
50067000	SABANA RIVER AT SABANA	18º19'52"	65º43'52"
50070500	FAJARDO RIVER ABV FAJARDO	18º15'21"	65º43'18"
50071000	FAJARDO RIVER NR FAJARDO	18º17'56"	65º41'42"
50074950	QUEBRADA GUABA NR NAGUABO	18º17'02"	65º47'20"
50075000	ICACOS RIVER NR NAGUABO	18º16'38"	65º47'09"
50075001	METEOROLOGICAL STA AT ICACOS RIVER NR NAGUABO	18º16'08"	65º47'12"
50081000	HUMACAO RIVER AT LAS PIEDRAS	18º10'27"	65º52'11"
50090500	MAUNABO RIVER AT LIZAS	18º01'38"	65º56'24"
50092000	RIO GRANDE OF PATILLAS NR PATILLAS	18º02'04"	66º01'58"
50093045	PATILLAS LAKE AT DAMSITENR PATILLAS	18º01'15"	66º01'19"
50100200	LAPAS RIVER NR RABO DEL BUEY	18º03'36"	66º14'28"
50100450	MAJADA RIVER AT LA PLENA	18º02'40"	66º12'27"
50106100	COAMO RIVER AT COAMO	18º05'00"	66º21'16"
50110900	TOA VACA RIVER ABV TOA VACA LAKE	18º07'36"	66º27'25"

ID	STATION NAME	LAT	LON .
50111210	TOA VACA LAKE AT DAMSITE	18º06’07”	66º29’23”
50111300	GUAYABAL LAKE AT DAMSITE NR JUANA DIAZ	18º05’17”	66º30’09”
50111500	JACAGUAS RIVER AT JUANA DIAZ	18º03’16”	66º30’40”
50112500	INABON RIVER AT REAL ABAJO	18º05’10”	66º33’46”
50113800	CERRILLOS RIVER ABV CERRILLOS LAKE NR PONCE	18º07’01”	66º36’17”
50113950	CERRILLOS LAKE AT DAMSITE NR PONCE	18º04’41”	66º34’38”
50114000	CERRILLOS RIVER NR PONCE	18º04’15”	66º34’51”
50114390	BUCANA RIVER AT HWY 14 BRIGE NR PONCE	18º02’29”	66º34’58”
50114900	PORTUGUES RIVER AT TIBES	18º06’00”	66º38’34”
50115900	PORTUGUES RIVER AT HWY 14 AT PONCE	18º01’09”	66º36’26”
50124200	GUAYANILLA RIVER NR GUAYANILLA	18º02’40”	66º47’53”
50125780	LUCHETTI LAKE AT DAMSITE NR YAUCO	18º05’37”	66º51’54”
50128900	LOCO LAKE AT DAMSITE NR YAUCO	18º02’41”	66º53’17”
50131990	GUANAJIBO RIVER AT HWY 119AT SAN GERMAN	18º05’06”	67º02’02”
50136400	ROSARIO RIVER NR HORMIGUEROS	18º09’36”	67º05’08”
50138000	GUANAJIBO RIVER NR HORMIGUEROS	18º08’36”	67º08’57”
50141500	GUAYO LAKE AT DAMSITE NR CASTANER	18º12’46”	66º50’06”
50144000	RIO GRANDE OF ANASCO NR SAN SEBASTIAN	18º17’05”	67º03’05”
50147800	CULEBRINAS RIVER AT HWY 40 NR MOCA	18º21’42”	67º05’33”
50148890	CULEBRINAS RIVER AT REPRESA MARGARITA NR AGUADA	18º23’45”	67º09’07”
50999954	QUEBRADA SALVATIERRA RAINGAGE, SAN LORENZO	18º10’46”	65º59’54”
50999956	QUEBRADA BLANCA RAINGAGE, SAN LORENZO	18º09’43”	65º59’47”
50999958	PUEBLITO DEL RIO RAINGAGE, LAS PIEDRAS	18º14’54”	65º49’56”
50999959	GURABO ABAJO RAINGAGE, GURABO	18º16’02”	65º54’45”
50999960	QUEBRADA ARENAS RAINGAGE, SAN LORENZO	18º06’50”	65º56’49”
50999961	LA PLAZA RAINGAGE, CAGUAS	18º08’08”	66º03’00”
50999962	CANABONCITO RAINGAGE, AGUAS BUENAS	18º12’53”	66º06’25”
50999963	JAGUEYES ABAJO RAINGAGE, AGUAS BUENAS	18º17’21”	66º04’34”
50999964	BAIROA ARRIBA RAINGAGE, AGUAS BUENAS	18º15’57”	66º05’45”
50999965	VAQUERIA EL MINO RAINGAGE, CAGUAS	18º13’11”	66º04’03”
50999966	BO. BEATRIZ RAINGAGE, CAGUAS	18º11’00”	66º05’22”
50999967	BO. MONTONES RAINGAGE LAS PIEDRAS	18º09’48”	65º54’39”
50999968	LAS PIEDRAS CONSTRUCTION RAINGAGE LAS PIEDRAS	18º12’16”	65º50’27”
50999970	BARRIO APEADERO RAINGAGE, VILLALBA	18º09’31”	66º27’34”
50999971	LAS CURIAS RAINGAGE NR CUPEY	18º20’05”	66º02’23”
50115000	PORTUGUES RIVER NR PONCE	18º04’45”	66º38’01”

APPENDIX C

SIMPLE LINEAR REGRESSION

We consider the problem of fitting the set of N data points (x_i, y_i) to a straight-line model

$$y(x) = y(x; a, b) = a + bx \tag{C.1}$$

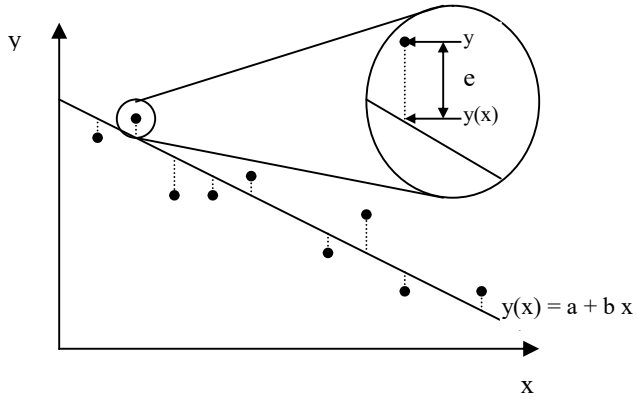

Figure C.1: Schematic illustration of simple linear regression.

The regression procedure chooses that line producing the least error for prediction of y given observations of x. Figure C.1 illustrates the situation. Given a data set of (x, y) pairs, the problem is to find the particular straight line that minimizes the squared vertical distances (dotted line) between it and the data points. These vertical distances, also called residuals errors, are defined as

$$e_i = y_i - y(x_i) \tag{C.2}$$

The sign convention for the residual is unimportant, since it is the minimization of the sum of squared residuals that defines the best-fitting line. Combining equations (C.1) and (C.2) yields the regression equation

$$y_i = y(x_i) + e_i = a + bx_i + e_i \tag{C.3}$$

This indicates that the true value of the predictand is the sum of the predicted value (equation C.1) and the residual. In order to minimize the sum of squared residuals

$$\sum_{i=1}^{N}(e_i)^2 = \sum_{i=1}^{N}[y_i - y(x_i)]^2 = \sum_{i=1}^{N}(y_i - [a + bx_i])^2 \tag{C.4}$$

it is only necessary to set the derivatives of equation (C.4) with respect to the parameters a and b to zero and solve. These derivatives are

$$\frac{\partial}{\partial a}\sum_{i=1}^{N}(e_i)^2 = \frac{\partial}{\partial a}\sum_{i=1}^{N}(y_i - a - bx_i)^2 = -2\sum_{i=1}^{N}(y_i - a - bx_i) = 0$$

(C.5)

and

$$\frac{\partial}{\partial b}\sum_{i=1}^{N}(e_i)^2 = \frac{\partial}{\partial b}\sum_{i=1}^{N}(y_i - a - bx_i)^2 = -2\sum_{i=1}^{N}[x_i(y_i - a - bx_i)] = 0 \tag{C.6}$$

Rearrangement of equations (C.5) and (C.6) leads to the so-called normal equations:

$$\sum_{i=1}^{N} y_i = \sum_{i=1}^{N} a + b\sum_{i=1}^{N} x_i = na + b\sum_{i=1}^{N} x_i \tag{C.7}$$

and

$$\sum_{i=1}^{N} x_i y_i = a\sum_{i=1}^{N} x_i + b\sum_{i=1}^{N} (x_i)^2 \tag{C.8}$$

These equations can be rewritten in a convenient form by defining the following sums:

$$\begin{aligned} S &\equiv \sum_{i=1}^{N} 1 = n \qquad S_x \equiv \sum_{i=1}^{N} x_i \\ & \qquad\qquad\qquad\qquad S_y \equiv \sum_{i=1}^{N} y_i \\ S_{xx} &\equiv \sum_{i=1}^{N} x_i^2 \qquad S_{xy} \equiv \sum_{i=1}^{N} x_i y_i \end{aligned} \tag{C.9}$$

With these definitions equations (C.7) and (C.8) become

$$\begin{aligned} aS + bS_x &= S_y \\ aS_x + bS_{xx} &= S_{xy} \end{aligned} \tag{C.10}$$

The solution to these two equations is calculated as

$$\Delta \equiv SS_{xx} - (S_x)^2$$

$$a = \frac{S_{xx}S_y - S_x S_{xy}}{\Delta} \tag{C.11}$$

$$b = \frac{SS_{xy} - S_x S_y}{\Delta}$$

Equation (C.11) gives the solution for the best-fit model parameters a and b (Vetterling, , Teukolsky, Press, and Flannery, 1997).

Suppose that a plot of y against x reveals a curvature as shown in Figure C.2. One may assume that a power function of the form $y = \alpha x^{\beta}$ generated the data (apart from random error) and that a model of the type from equation C.12 should be fit.

$$\log y_i = \beta_0 + \beta_1 \log x_i + \varepsilon_i \qquad \text{(C.12)}$$

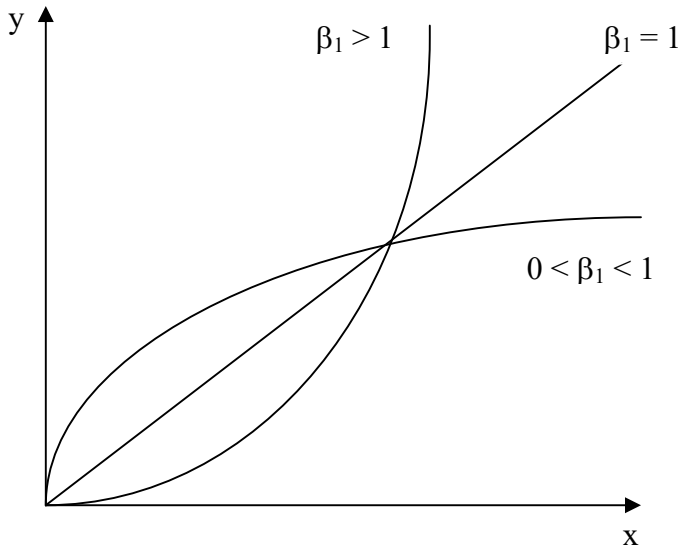

Figure C.2: Power function.

The transformation works because the true mechanism is the power function. If we ignore errors in the model (which truly we cannot), then, by taking the logarithm of the power function, we are led to $\log y_i = \beta_0 + \beta_1 \log x_i$. However, if we allocate the usual additive error terms to the two relationships, a log transformation does not produce the equation C.12.

We can obtain the equation C.13 by beginning with the model

$$y = \alpha x^{\beta}(1 + \varepsilon^*) \tag{C.13}$$

We might assume that that ε* has the usual properties, E(ε*) = 0 and homogeneous variance. If these assumptions hold, then Var(y) = $\sigma^2[E(y)]^2$, which varies with E(y). We say, in this case, that there is a multiplicative error structure.

Under these conditions, a logarithmic transformation of the equation C.13 results in the following equation:

$$\log y = \log\alpha + \beta \log x + \log(1 + \varepsilon^*) \tag{C.14}$$

Suppose E[ln(1+ε*)] = α*; then we can write

$$\log(y) = (\log\alpha + \alpha^*) + \beta \log x + \varepsilon \tag{C.15}$$

where E(ε*) = 0 and Var ε = σ^2, say, is homogeneous, not varying with E(y). Now, the model of equation C.15 is of the form of the equation C.12. As a result, a transformation that produces the equation C.12 truly begins with the multiplicative error assumption of the model in the equation C.13 (Myers, 1990).

APPENDIX D

MINITAB MACRO FILE EXAMPLE

```
GMACRO
MYMACRO
  NAME C1='Reflectivity' C2='Precipitation'
  NAME C3='LOG10(Z)' C4='LOG10(R)'
  NAME C5='Residuals' C6='Coeffs' C7='(10^Residuals)-1'
  NAME C8='FITS' C9='PROB' C10='E[Y]' C11='e**' C15='PARAM'
  READ C1-C2;
    FILE
"D:\USER\JORGE\CURSOS\TESIS\RAIN_PROC\50021700_19991202_ZR.TXT".
  DIVI C1 10.0 C3
  LOGTEN C2 C4
  REGRESS C3 1 C4;
    SRESIDUALS C5;
    COEFFICIENTS C6;
    FITS C8.
  LET C6(1) = 10**C6(1)
  LET C7 = (10**C5) - 1
  LET C10 = C6(1)*(C2**C6(2))
  LET C11 = C10 * C7
  MEAN C1 C20
  LET C12(1) = C20(1)
  STDEV C1 C20
  LET C12(2) = C20(1)
  N C1 C20
  LET C12(3) = C20(1)
  MEAN C2 C20
  LET C12(4) = C20(1)
  STDEV C2 C20
  LET C12(5) = C20(1)
  N C2 C20
  LET C12(6) = C20(1)
  LET K1 = ROUND(SQRT(C20)+1.0)
```

```
  NTOA C12 C14
  READ C13;
    FORMAT (3x,a9).
  Average:
  StDev:
  N:
  Average:
  StDev:
  N:
END

  READ C16-C17;
    FORMAT (F4.2,1x,F4.2).
2.28 2.80
2.18 2.80
2.08 2.80
2.28 0.75
2.18 0.75
2.08 0.75
END

  CONCATENATE C13-C14 C15
  CALL STATISTICS
  CALL RESIDUALDIAG
ENDMACRO
#=============================================================
GMACRO
STATISTICS
  LAYOUT;
    wtitle 'Statistics'.

    BOXPLOT C2;
      Transpose;
      Box;
        Type 0 0;
        Color 1 0;
      Symbol;
        Outlier;
      Title "Instantaneous Precipitation from USGS";
        TSize 2.0;
        TCol 4;
      Title '50021700 - 12/02/1999';
        TSize 2.0;
        offs .025 -0.105;
      Overlay;
```

```
  ScFrame;
  ScAnnotation;
  axis 2;
    label "(a) Rain Rate (mm/h)";
    tsiz 1.75;
  data .25 .8 .2 .8;
  figure .0 .5 .57 1; # place plot in upper right
    etype 0.

BOXPLOT C1;
  Transpose;
  Box;
    Type 0 0;
    Color 1 0;
  Symbol;
    Outlier;
  Title 'Reflectivity from WATADS';
    TSize 2.0;
    TCol 4;
  Title '50021700 - 12/02/1999';
    TSize 2.0;
    offs .025 -0.105;
  Overlay;
  ScFrame;
  ScAnnotation;
  axis 2;
    label '(b) Reflectivity Factor Z (dBZ)';
    tsiz 1.75;
  data .25 .8 .2 .8;
  figure .5 1 .57 1; # place plot in upper left
    etype 0.

HIST C2;
  title 'Histogram of Precipitation';
    tsize 2;
    tcol 4;
  axis 1;
    label '(c) Rain Rate (mm/h)';
    tsiz 1.75;
  axis 2;
    tsiz 1.75;
  tick 1;
    tsiz 1.5;
  tick 2;
    tsiz 1.5;
```

```
    NInterval K1;
    data .15 .9 .2 .9;
    figure .0 .5 0 .5; # place plot in lower left
      etype 0.

  HIST C1;
    title 'Histogram of Reflectivity';
      tsize 2;
      tcol 4;
    axis 1;
      label '(d) Reflectivity Factor Z (dBZ)';
      tsiz 1.75;
    axis 2;
      tsiz 1.75;
    tick 1;
      tsiz 1.5;
    tick 2;
      tsiz 1.5;
    NInterval K1;
    data .15 .9 .2 .9;
    figure .5 1 0 .5; # place plot in lower left
      etype 0.

  PLOT C16*C17;
    data 0.0 1.0 0.0 1.0;
    etyp 0;
    figure 0.0 1.0 0.0 1.0;
    nobrush;
    max 1 4;
    min 1 0;
    max 2 4;
    min 2 0;
    symb;
      type 0;
      labe C15;
      plac 1 0;
      tsiz 0.65;
      offs 0.0 0.0;
      size 2.5;
    nofr.

  ENDLAYOUT
ENDMACRO
#=============================================================
GMACRO
```

```
RESIDUALDIAG
# C5
  nscores C5 C9
  LAYOUT;
    wtitle 'Residual Model Diagnostics';
    title 'Residual Model Diagnostics';
      tsiz 1.5.

    hist C5;
      title 'Histogram of Residuals';
        tsize 2.4;
        tcol 4;
        offs .025 -.075;
      axis 1;
        label '(c) Residual';
        tsiz 2.2;
      axis 2;
        tsiz 1.75;
      tick 1;
        tsiz 1.5;
      tick 2;
        tsiz 1.5;
      NInterval K1;
      data .18 .95 .2 .8;
      figure 0 .33 0 .485; # place plot in lower left
        etype 0.

    brief 2

    ichart C5;    # residuals in RES
      title 'I Chart of Residuals';
        tsize 2;
        tcol 4;
        offs .025 -.075;
      symbol;
        type 6;
        size .5;
      axis 1;
        label '(b) Observation Number';
        tsiz 1.75;
      axis 2;
        label 'Residual';
        tsiz 1.75;
      tick 1;
        tsiz 1.5;
```

```
  tick 2;
    tsiz 1.5;
  test 1:8;
  data .25 .8 .2 .8;
  figure .45 1 .465 .95; # place plot in upper left
    etype 0.

brief 0

plot C5*C8; # residuals in RES vs. fits in FIT
  title 'Residuals vs. Fits';
    tsize 2.4;
    tcol 4;
    offs .025 -.075;
  symbol;
    type 6;
    size .5;
  axis 1;
    label '(d) Fit';
    tsiz 2.2;
  axis 2;
    label 'Residual';
    tsiz 1.75;
  tick 1;
    tsiz 1.5;
  tick 2;
    tsiz 1.5;
  refe 2 0;
    type 1;
    colo 2;
  data .18 .95 .2 .8;
  figure .34 .66 0 .485; # place plot in upper left
    etype 0.

plot C5*C9; # residuals in RES vs. nscores in PROB
  axis 1;
    side 2;
    labe 'Normal Plot of Residuals';
    tsize 1;
    tcol 4;
    offs 0 .025;
  symbol;
    type 6;
    size .5;
  axis 1;
```

```
        label '(a) Normal Score';
        tsiz 0.875;
        offs 0 -.05;
      axis 2;
        label 'Residual';
        tsiz 0.875;
      tick 1;
        leng .01;
        tsiz 0.75;
        offs 0 -.015;
      tick 2;
        leng .01;
        tsiz 0.75;
      data .1375 .44 .562 .853. # place plot in upper left

    Plot C7*C10;
      title '(10^Residuals)-1 vs. E[y]';
        tsize 2.4;
        tcol 4;
        offs .03 -.075;
      Symbol;
        type 6;
        size .5;
      axis 1;
        label '(e) E[y]';
        tsiz 2.2;
      ScFrame;
      ScAnnotation;
      data .18 .95 .2 .8;
      figure .67 1 0 .485; # place plot in upper left
        etype 0.

  ENDLAYOUT
  brief 2
ENDMACRO

#============================================================
#          0  White       8  Dark Red
# (default) 1  Black       9  Dark Green
#          2  Red        10  Dark Blue
#          3  Green      11  Dark Cyan
#          4  Blue       12  Dark Magenta
#          5  Cyan       13  Dark Yellow
#          6  Magenta    14  Dark Gray
#          7  Yellow     15  Light Gray
```

APPENDIX E

WSR-88D OPERATIONS

Scan strategy

The scan strategy tells the radar how many elevations angles will be used during a single volume scan (a volume scan is the completion of a sequence of elevation angles), and the amount of time it will take to complete that sequence of elevation angles. Scan strategies are made up of a set constant elevation cuts, each one being a single full rotation of the antenna's 1º pencil beam (in azimuth) at a discrete elevation angle. Three scan strategies are used by the WSR-88D:

- Scan strategy #1: This scan strategy positions the antenna at fourteen separate elevation angles in a 5 minutes period (14/5).
- Scan strategy #2: Specifies that the antenna is to be positioned at nine separates elevation angles in six minutes (9/6).
- Scan strategy #3: The antenna is positioned at five separate elevation angles in ten minutes (5/10).

Volume Coverage Pattern (VCP)

While the scan strategy determines how many angles will be used in a volume scan, the VCP specifies the actual elevation angles the antenna will be positioned in during the volume scan (i.e. 0.5º, 1.5º, etc.) The WSR-88D is capable of 20 different VCPs, four of which have been defined:

- VCP 11 (scan strategy #1, version 1): specified the actual angles of the 14 elevations the antenna will be positioned in during the 5 minutes volume scan (see Figure E.1).

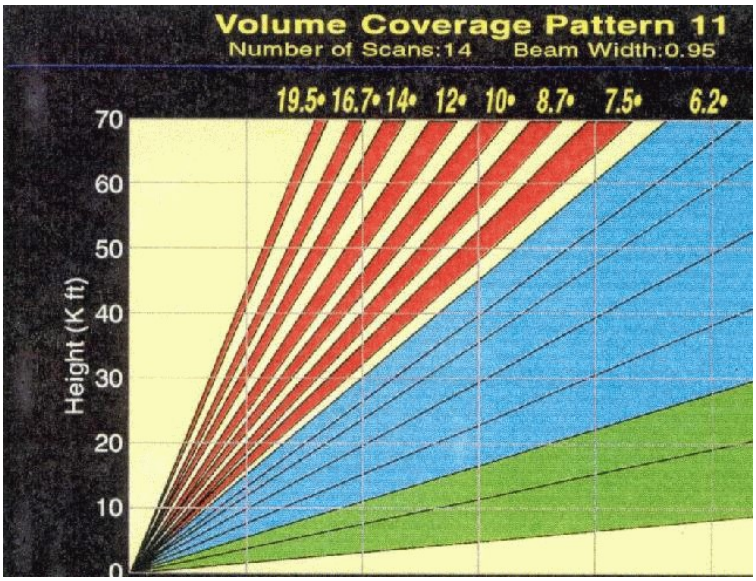

Figure E.1: Volume Coverage Pattern 11.

- VCP 21 (scan strategy #2, version 1): determines the actual angles of the 9 elevations the antenna will be positioned in during the 6 minutes volume scan (see Figure E.2).

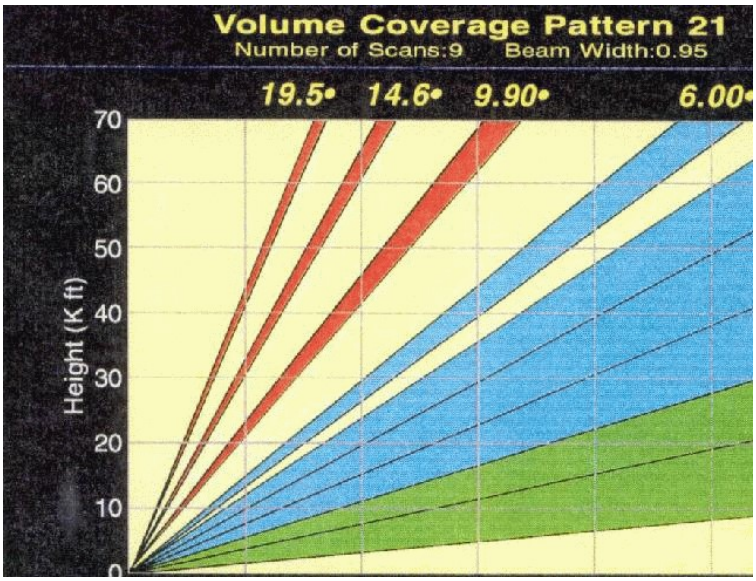

Figure E.2: Volume Coverage Pattern 21.

- VCP 31 (scan strategy #3, version 1): provides the antenna with the 5 specific angles at which it will be positioned during the 10 minutes volume scan (see Figure E.3).

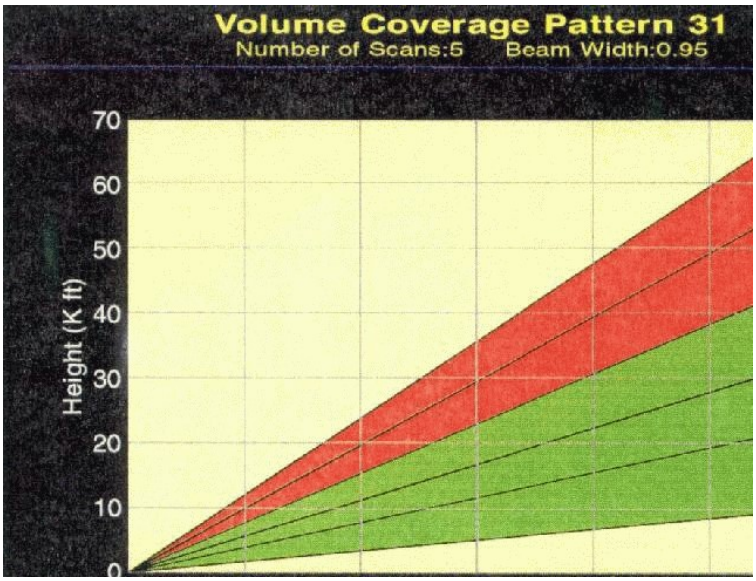

Figure E.3: Volume Coverage Pattern 31.

- VCP 32 (scan strategy #3, version 2): specifies the same elevation angles to be used during the 5 minutes volume scan as VCP 31. VCP 31 employs long radar pulse (4.7useg) and VCP 32 uses short pulse (1.57useg), see Figure E.4.

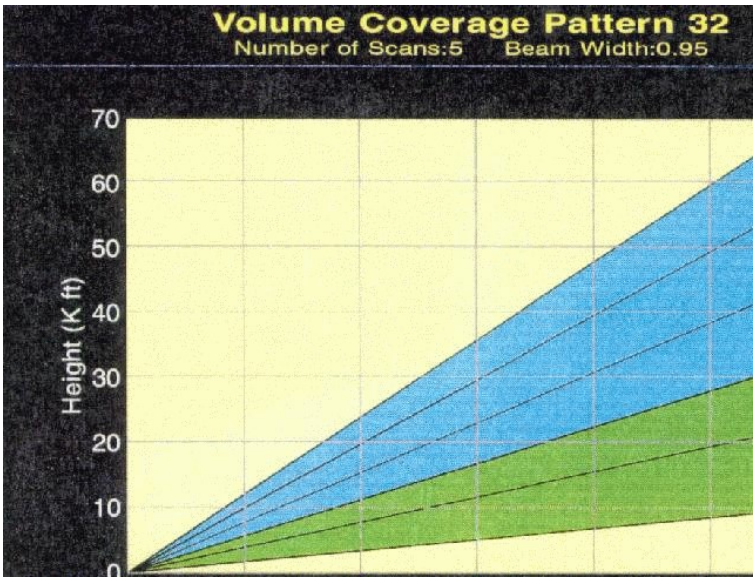

Figure E.4: Volume Coverage Pattern 32.

The WSR-88D also has two operational modes which play a part in antenna positioning: precipitation mode (mode A) and clear air mode (mode B). The mode in which the radar is operating determines which of the VCPs can be used, which in turn, determines the scan strategy being employed:

- Mode A: Precipitation mode uses either VCP 11 or 21, for which the scan strategies of 14/5 and 9/6 are employed respectively.
- Mode B: Clear air mode uses either VCP 31 or 32, both of which use the 5/10 scan strategy.

The WSR-88D component group is made up of four primary sub-components (transmitter, antenna, receiver, and signal processor) which generate and radiate radio frequency pulses, receive reflected energy from those pulses, and process this received energy into digital base data. The first step in acquiring radar data is the transmission of an RF signal. This is accomplished with a klystron amplifier (as opposed to magnetrons used by older radars) which generates a high-power (750 kw peak), very stable, ten centimeter wave-length RF pulse. The antenna is the sub-component that broadcasts the RF signal out into the atmosphere and intercepts returning energy. The antenna has a positioning range of 0.5 degrees to 20 degrees in elevation and one degree wide beam (see Figure E.5).

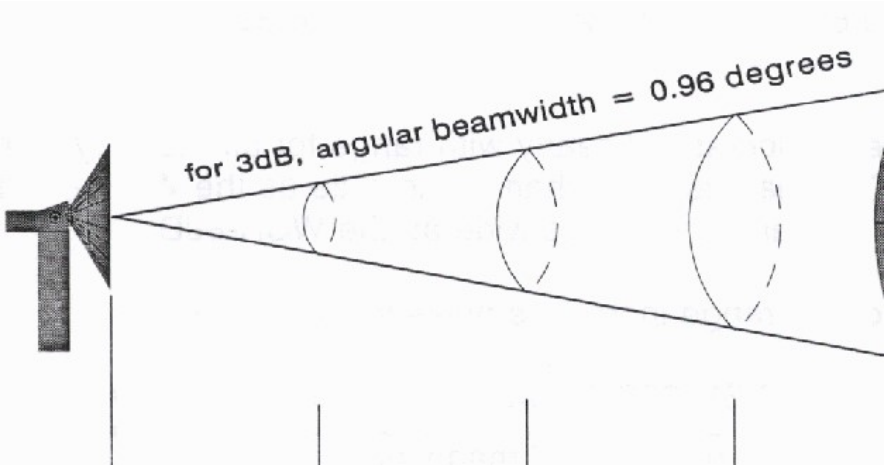

Figure E.5: WSR-88D Beam Geometry.

APPENDIX F

ORIGINAL SELECTED EVENTS BY REGIONS

Table F.1: Original selected events for Region 1.

Station Id	Date	Initial Time	Final Time
50050900	19991115	21.00	22.10
50051150	19991115	21.40	22.10
50056400	19991115	21.45	22.90
50999967	19991115	21.00	22.60
50050900	19991117	06.00	08.00
50051150	19991117	05.60	07.80
50051180	19991117	03.75	08.00
50051310	19991117	05.75	08.55
50053025	19991117	05.70	08.60
50999954	19991117	03.30	08.80
50999956	19991117	00.00	07.20
50051150	19991202	19.40	21.10
50051310	19991202	19.20	23.30
50051800	19991202	19.40	22.45
50055225	19991202	18.80	20.55
50056400	19991202	19.60	22.80
50057000	19991202	19.45	21.95
50999954	19991202	19.20	21.30
50999956	19991202	19.40	23.30
50999967	19991202	19.60	23.20

Table F.2: Original selected events for Region 2.

Station Id	Date	Initial Time	Final Time
50010500	19980930	18.40	19.70
50010800	19980930	19.05	19.80
50038100	19980930	18.30	19.05
50010500	19981015	18.60	19.80
50010500	19991104	18.70	20.05
50046000	19991115	19.70	21.40
50034000	19991117	05.45	08.30
50010500	19991202	16.80	20.60
50014800	19991202	15.80	20.90
50026140	19991202	19.60	21.90
50027100	19991202	16.10	17.60
50031200	19991202	16.30	20.40
50034000	19991202	16.55	22.20
50035000	19991202	16.60	22.10
50038100	19991202	17.45	21.30
50039500	19991202	15.40	21.20

Table F.3: Original selected events for Region 3.

Station Id	Date	Initial Time	Final Time
50065500	19991202	18.80	23.90
50066000	19991202	18.95	22.20
50070500	19991202	17.80	23.99
50071000	19991202	19.55	22.30
50065500	19991203	02.55	04.70
50070500	19991203	03.30	05.45
50071000	19991203	02.25	05.30

Table F.4: Original selected events for Region 4.

Station Id	Date	Initial Time	Final Time
50025155	19981015	18.10	19.40
50144000	19991104	18.30	19.30
50125780	19991105	16.80	18.40
50138000	19991105	17.80	19.55
50138000	19991106	18.80	20.70
50113800	19991115	18.30	21.40
50021700	19991202	18.80	21.40

Table F.5: Original selected events for Region 5.

Station Id	Date	Initial Time	Final Time
50044830	19991115	20.30	22.10
50045000	19991115	19.30	20.60
50047850	19991115	20.35	21.10
50081000	19991115	20.80	21.90
50044830	19991202	14.80	19.30
50045000	19991202	16.90	19.20
50048680	19991202	18.30	21.55
50061800	19991202	19.70	23.55
50081000	19991202	19.80	22.00
50999959	19991202	19.10	22.90
50999968	19991202	20.10	23.70
50999971	19991202	18.00	20.70
50063800	19991203	02.20	04.55
50999971	19991203	01.10	03.40

Table F.6: Original selected events for Region 6.

Station Id	Date	Initial Time	Final Time
50110900	19991116	16.70	18.40
50090500	19991117	05.55	08.10
50055100	19991202	18.30	21.80
50999963	19991202	17.30	22.90

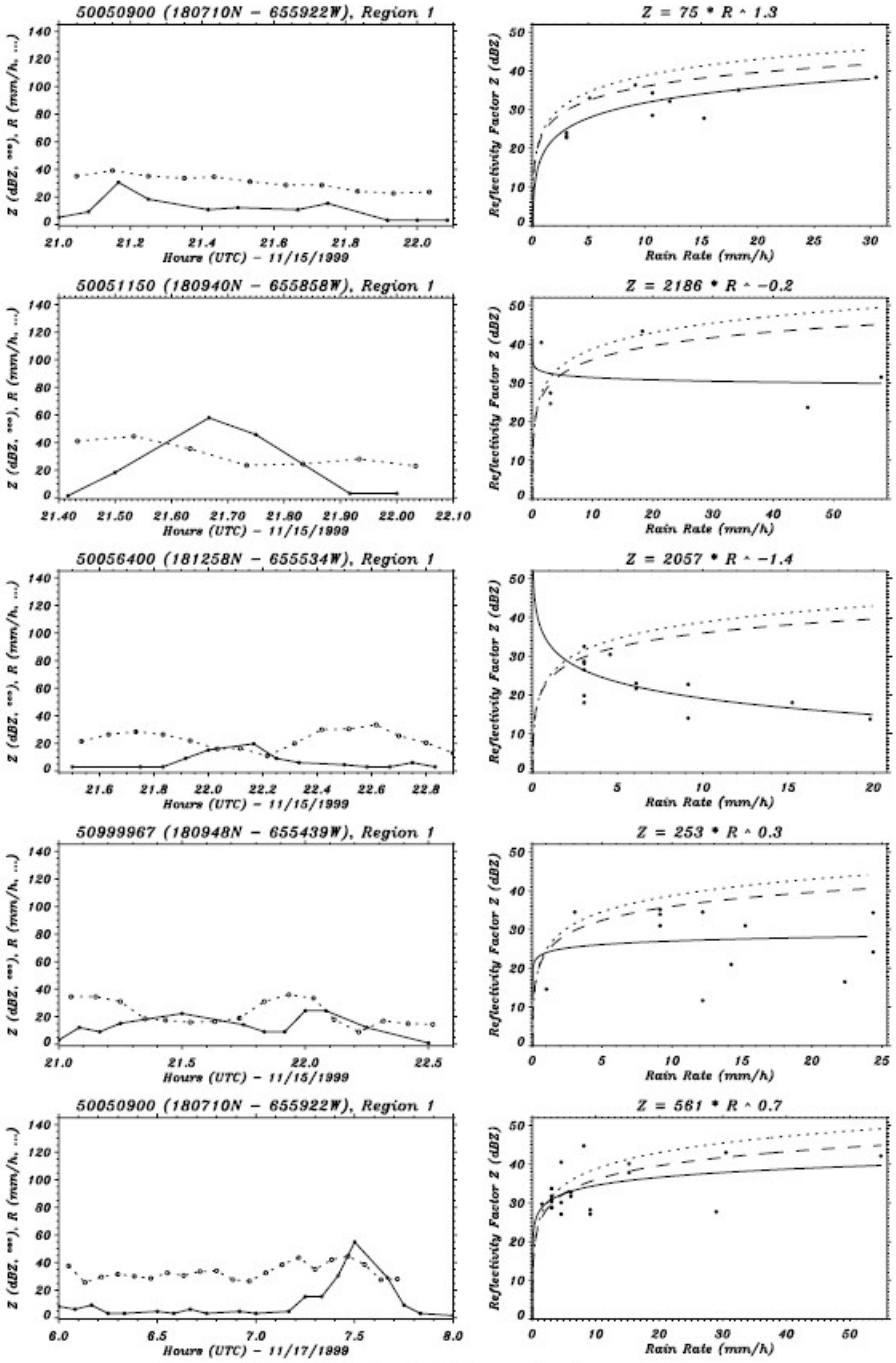

Figure F.1: Original data sets for region 1.

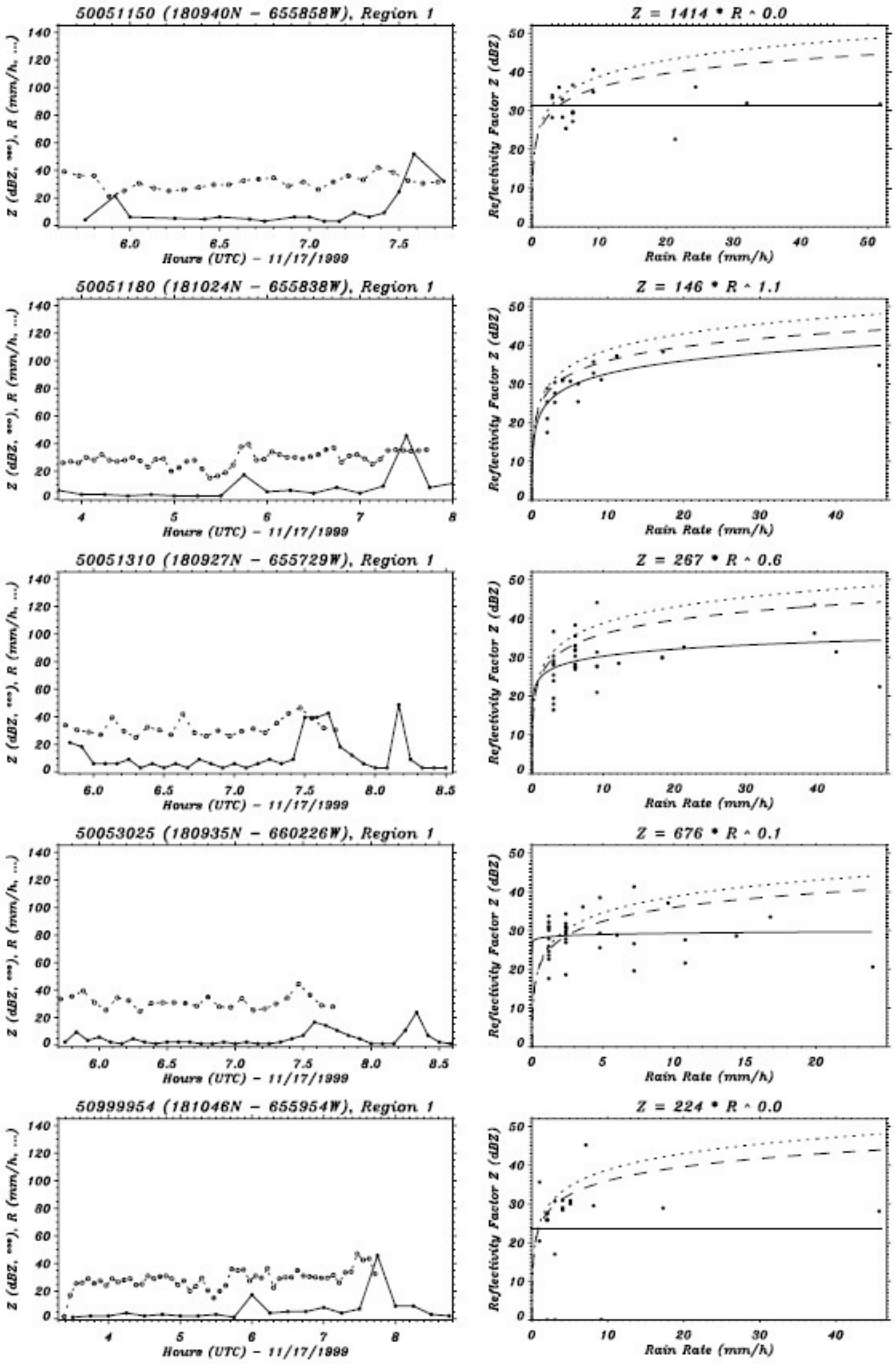

Figure F.1: Original data sets for region 1. (cont.)

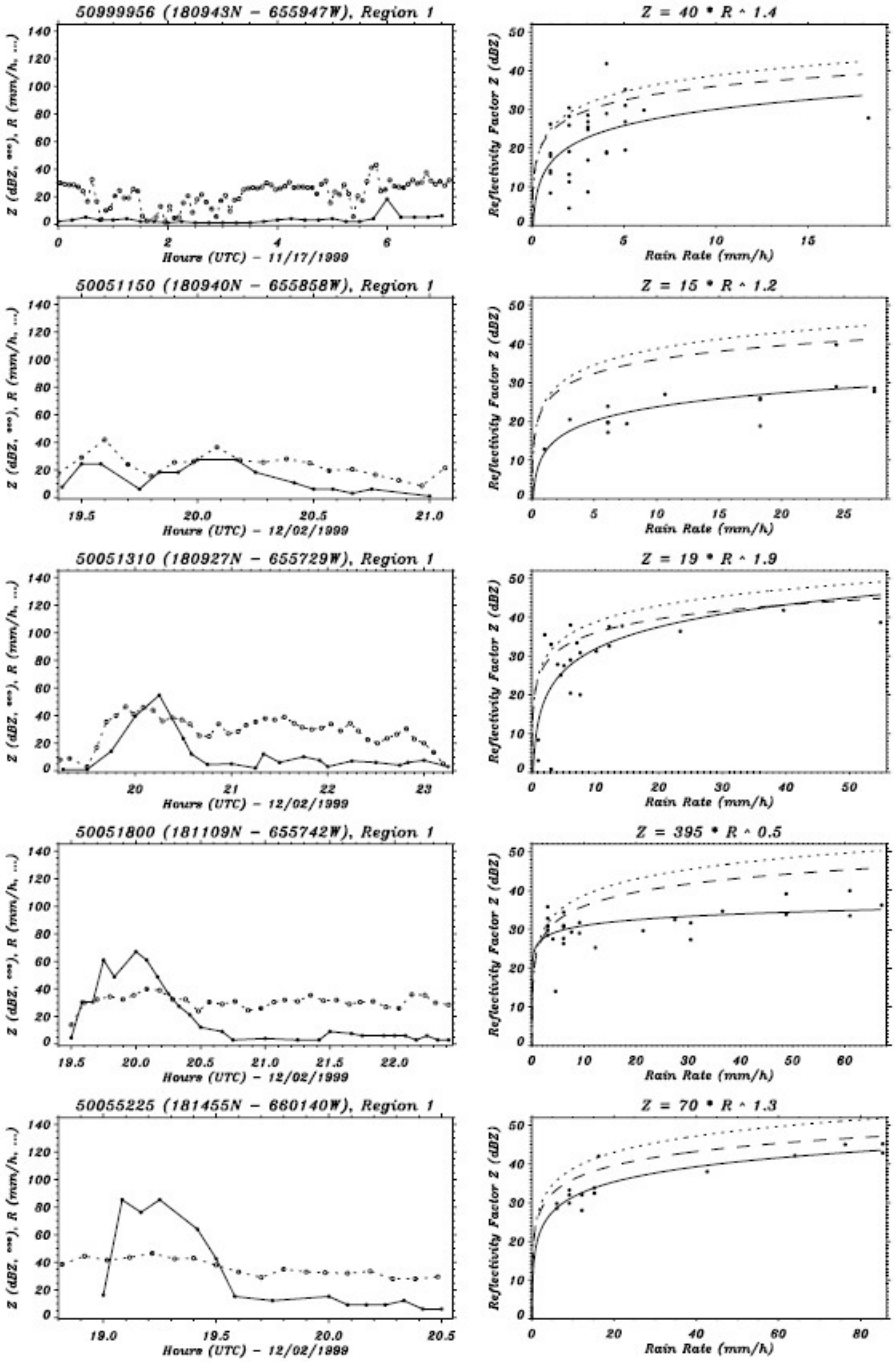

Figure F.1: Original data sets for region 1. (cont.)

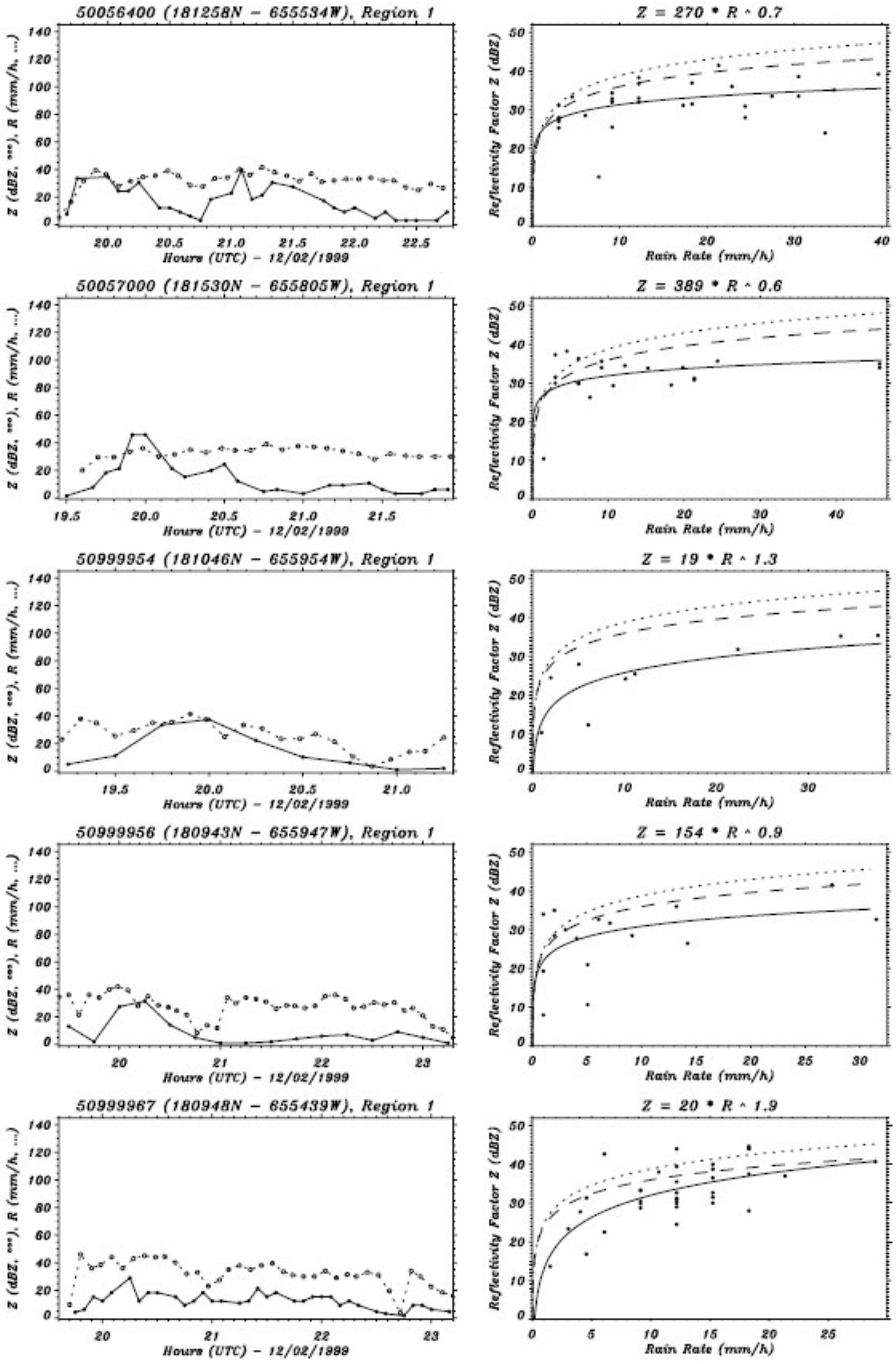

Figure F.1: Original data sets for region 1. (cont.)

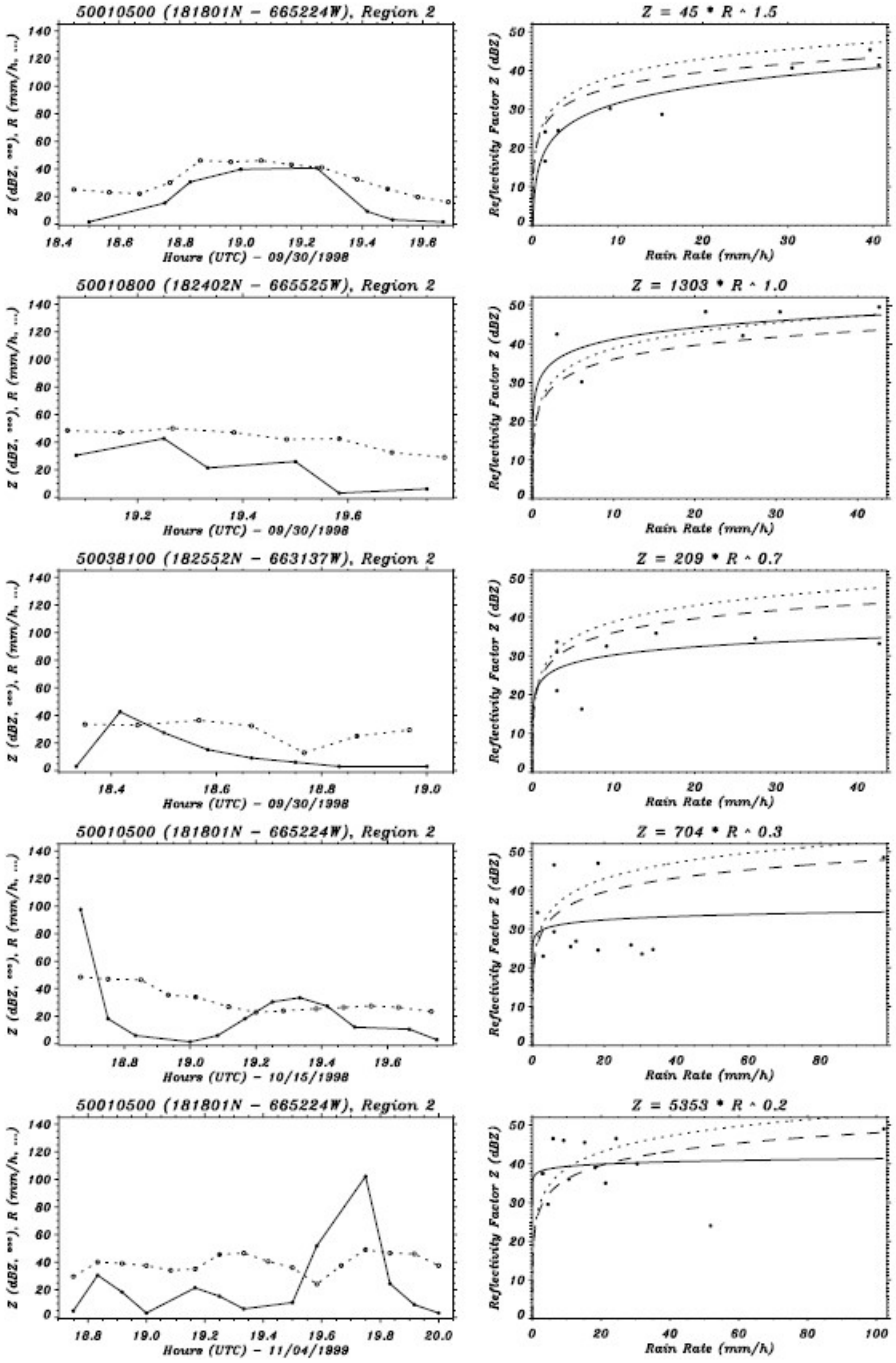

Figure F.2: Original data sets for region 2.

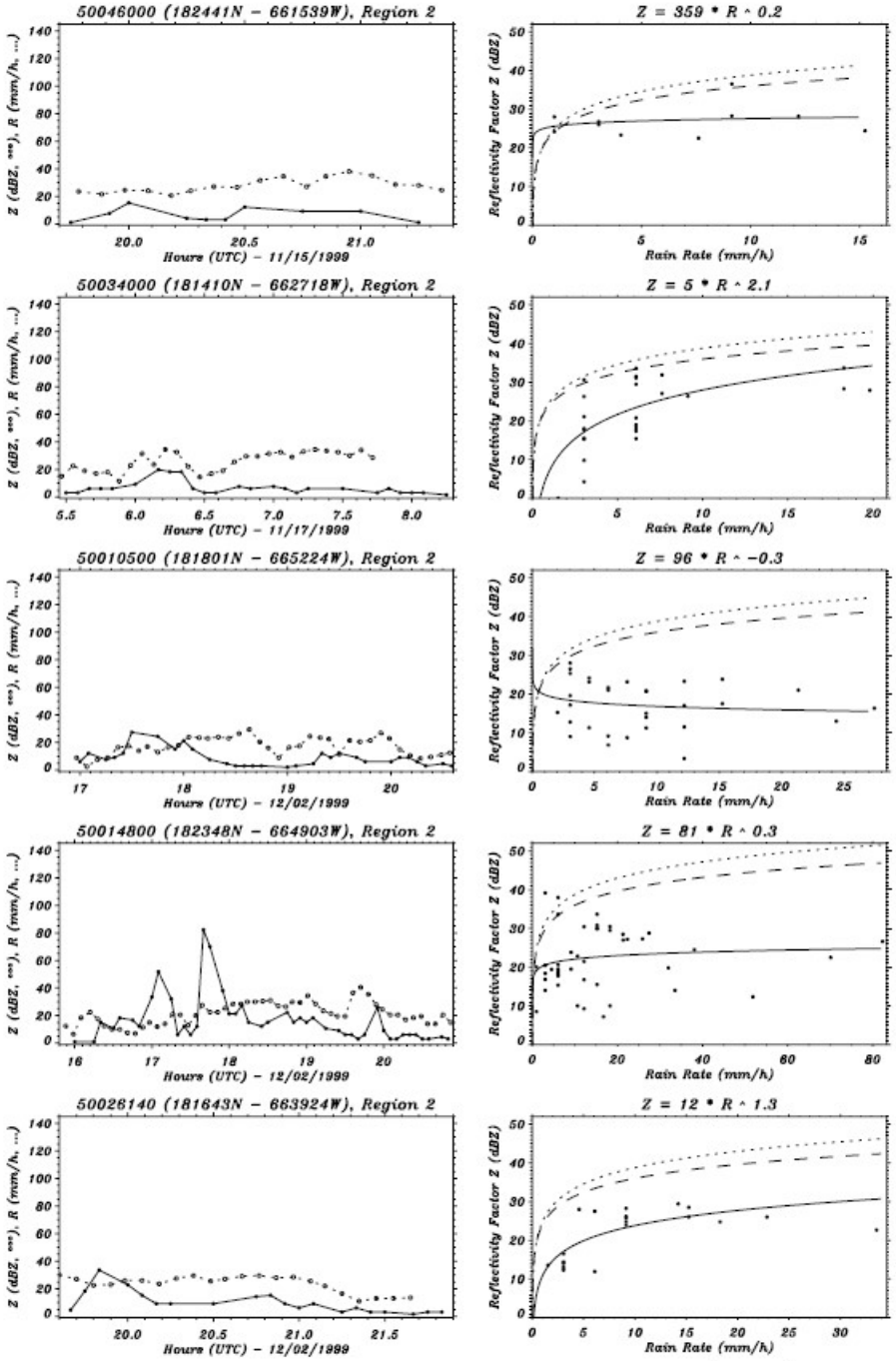

Figure F.2: Original data sets for region 2. (cont.)

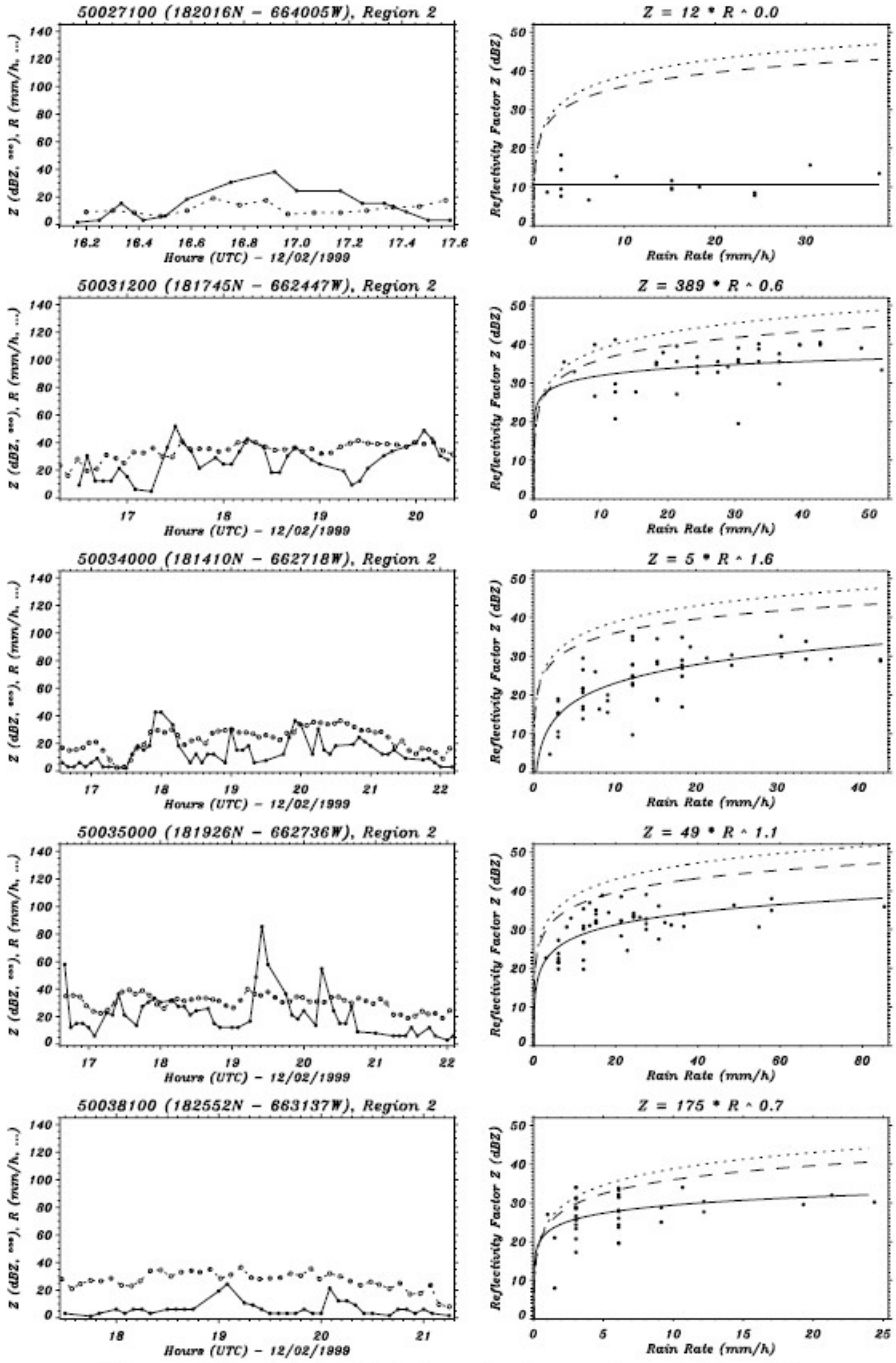

Figure F.2: Original data sets for region 2. (cont.)

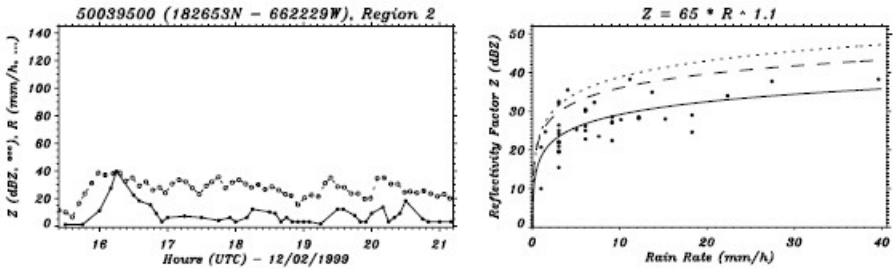

Figure F.2: Original data sets for region 2. (cont.)

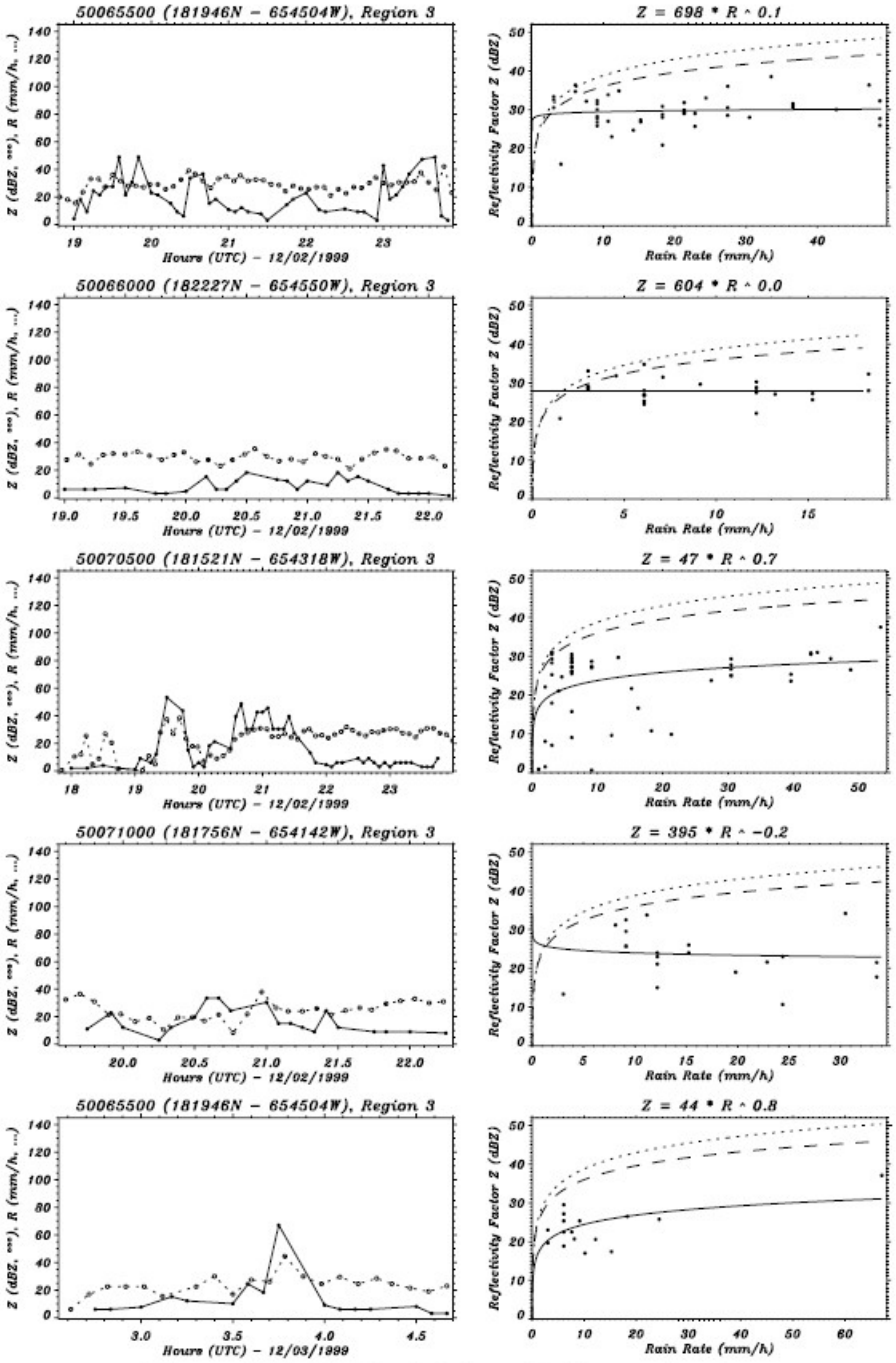

Figure F.3: Original data sets for region 3.

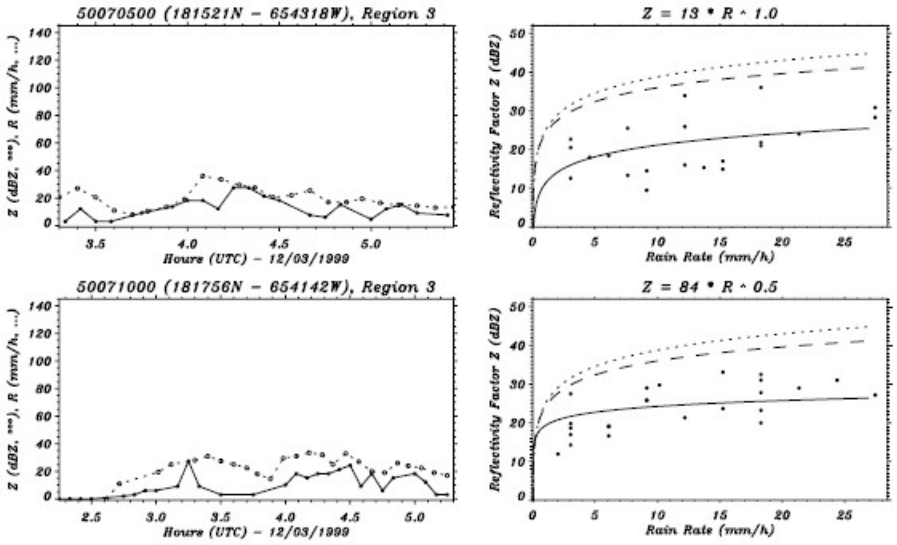

Figure F.3: Original data sets for region 3. (cont.)

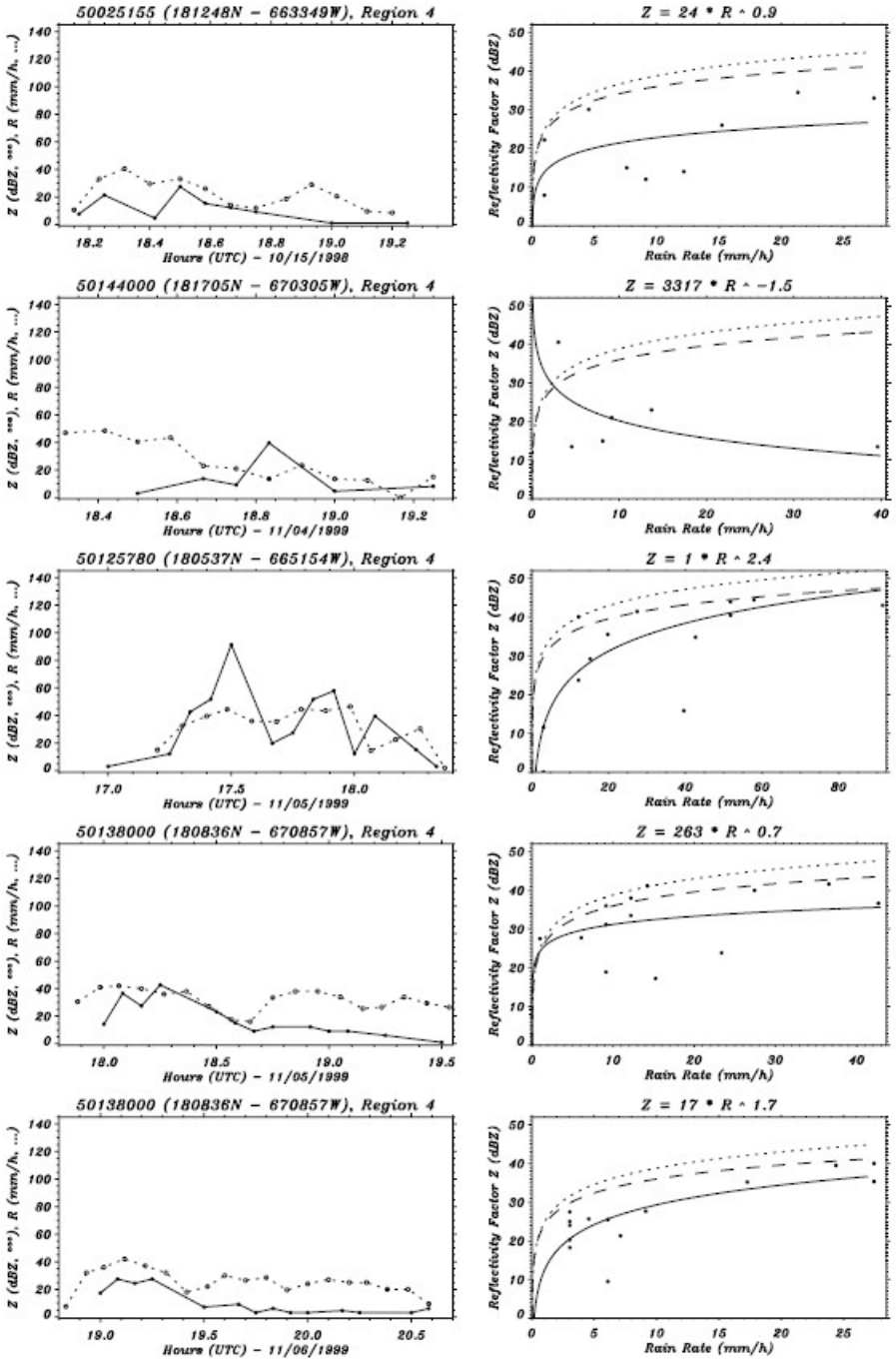

Figure F.4: Original data sets for region 4.

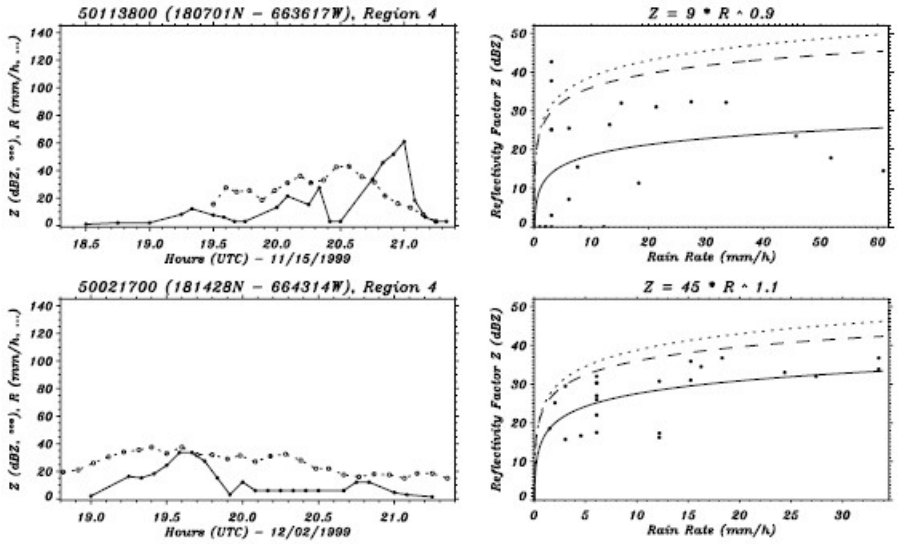

Figure F.4: Original data sets for region 4. (cont.)

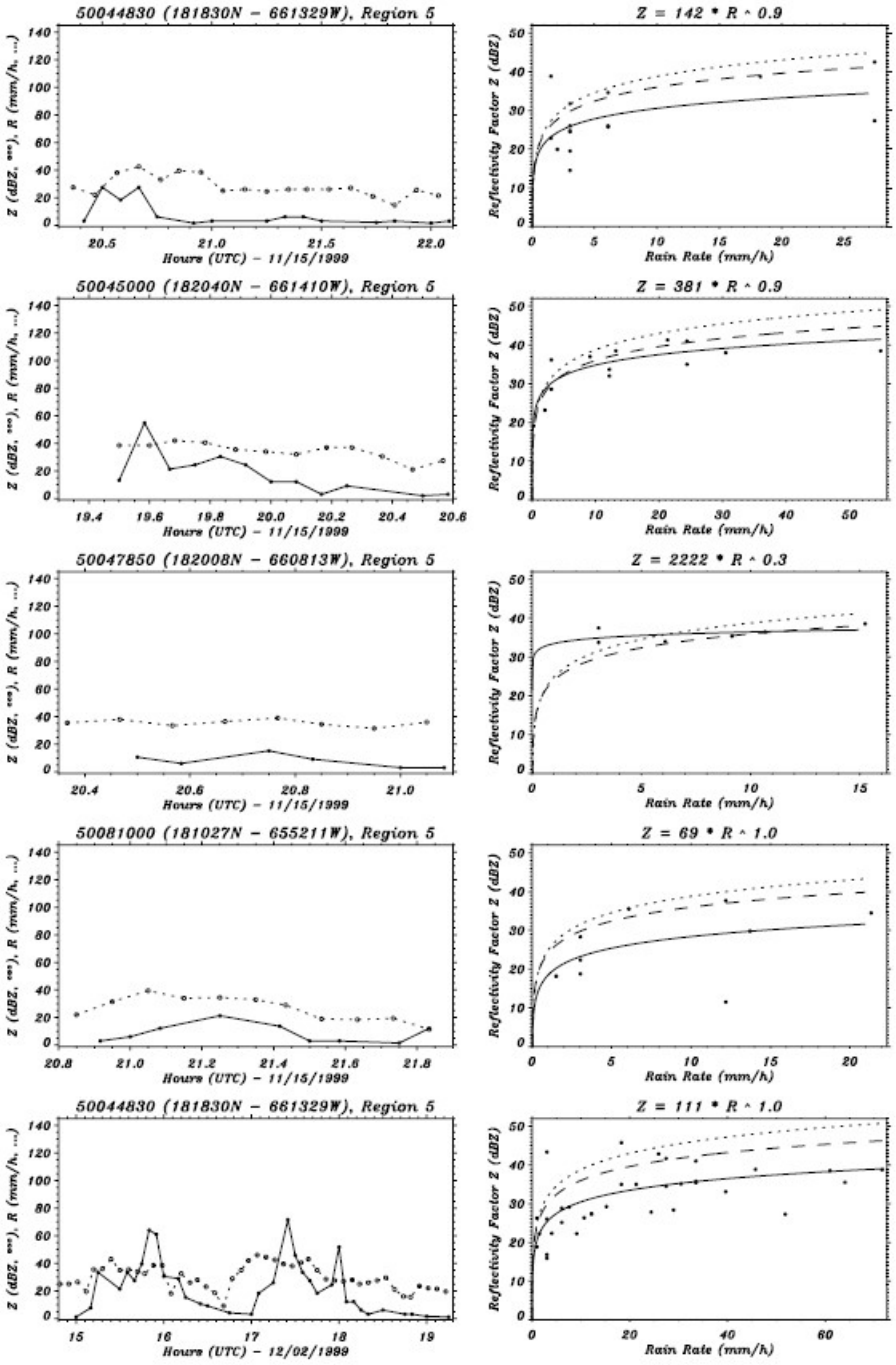

Figure F.5: Original data sets for region 5.

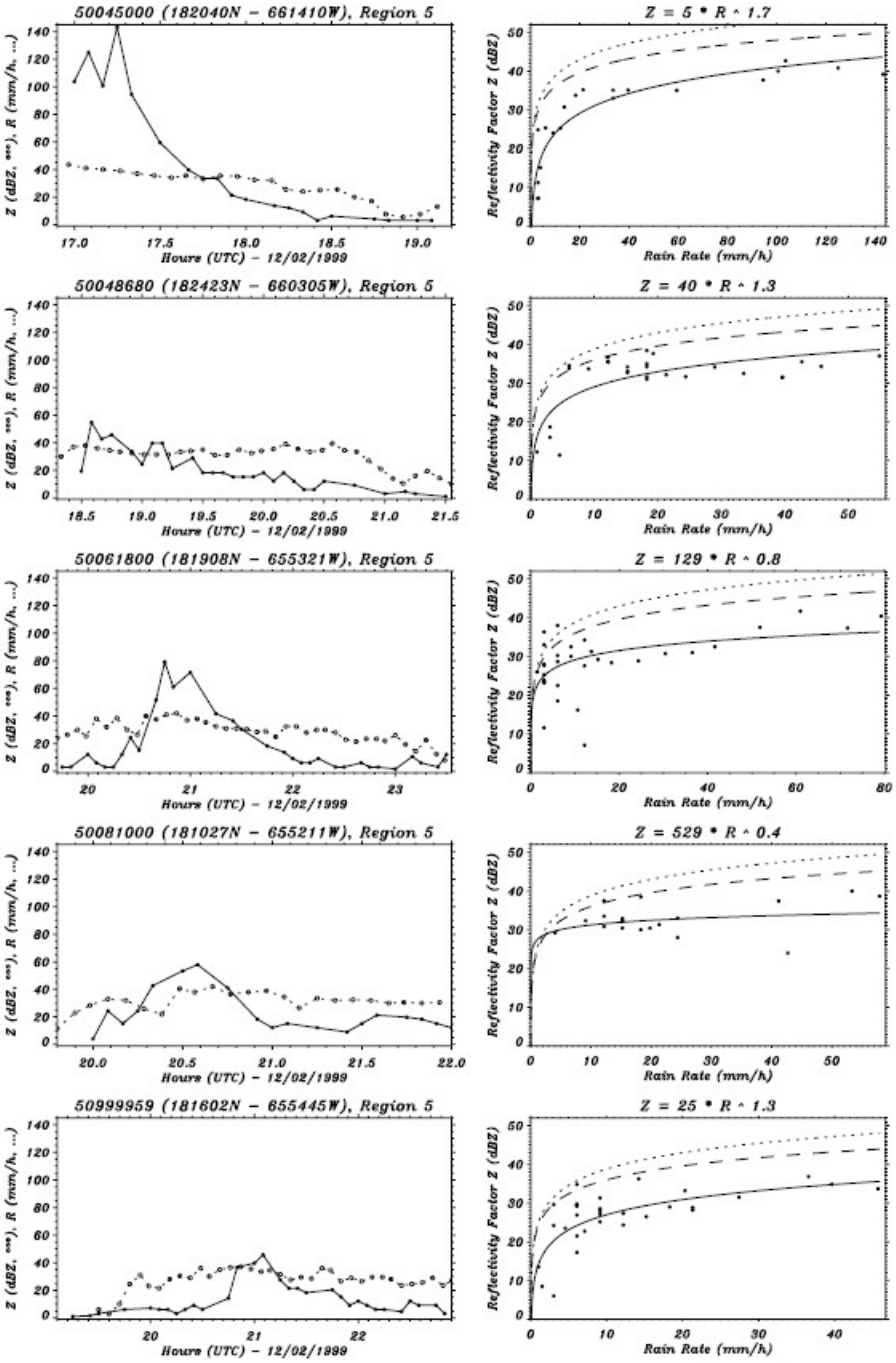

Figure F.5: Original data sets for region 5. (cont.)

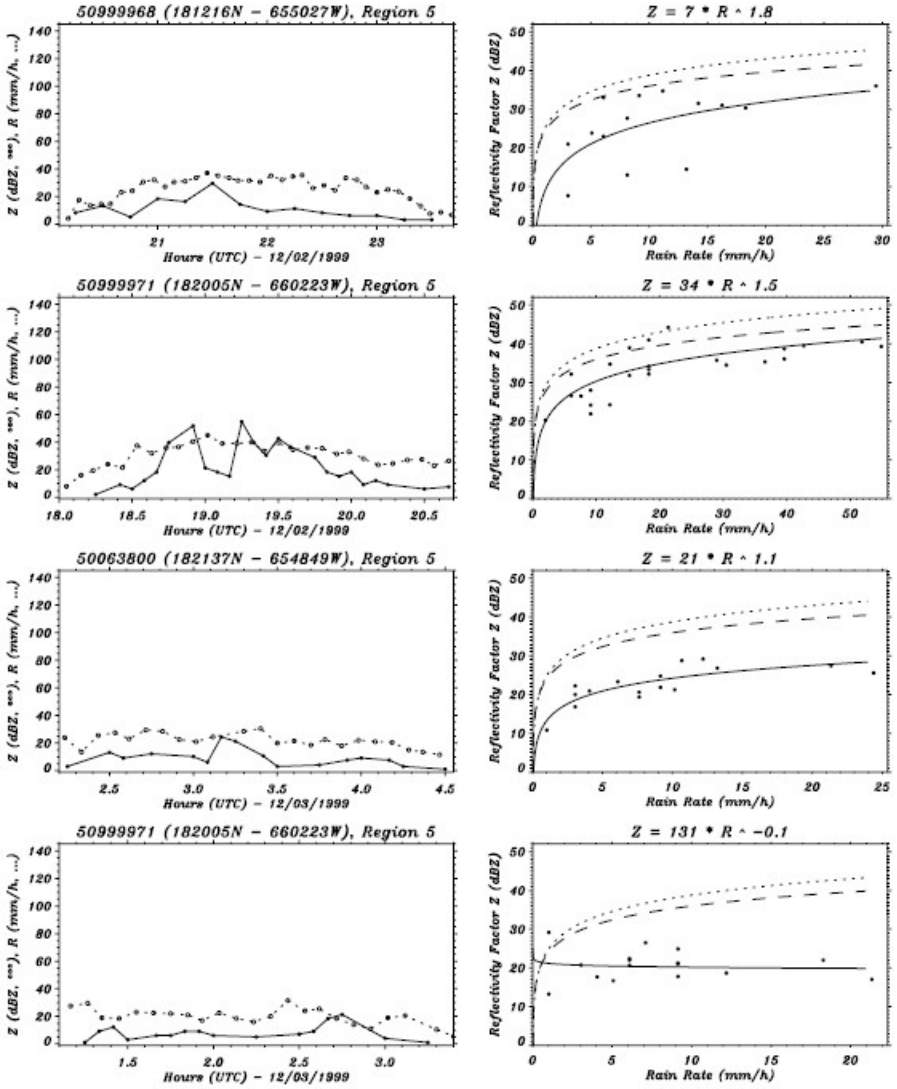

Figure F.5: Original data sets for region 5. (cont.)

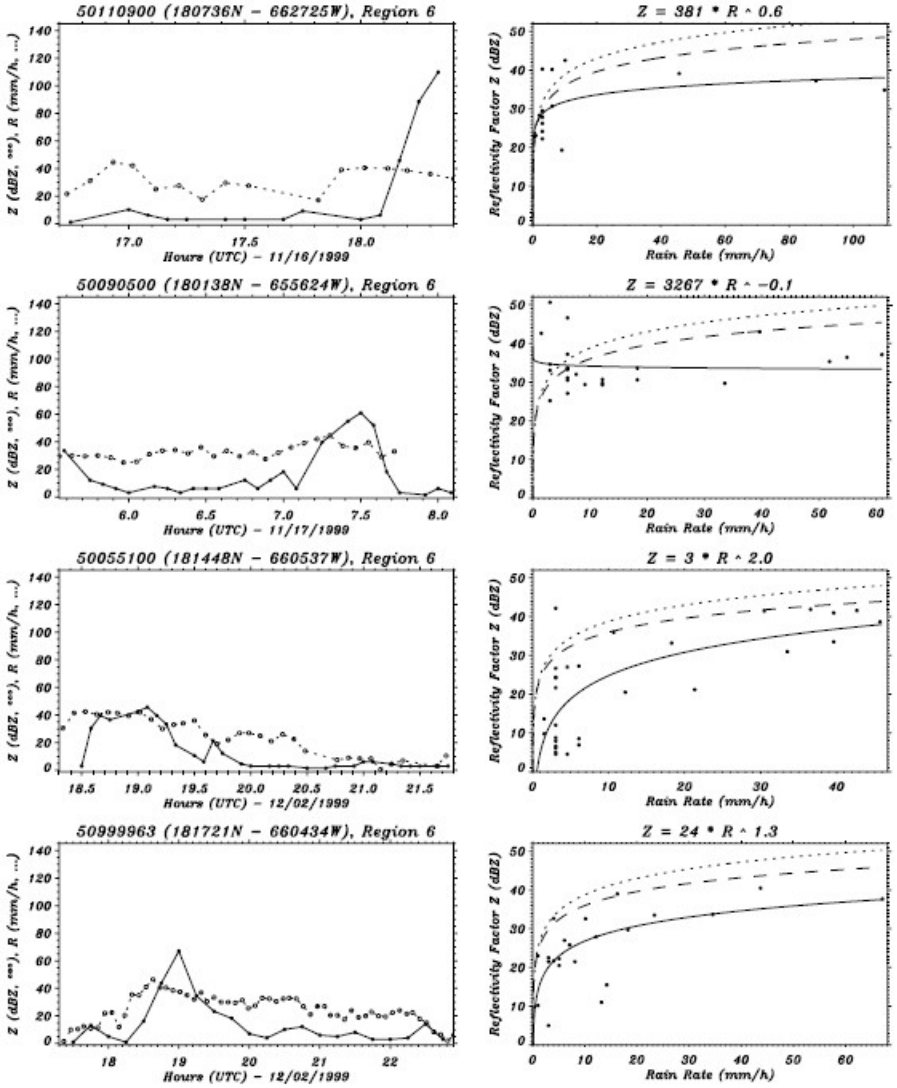

Figure F.6: Original data sets for region 6.

APPENDIX G

PROGRAMS AND DATA SETS (CD)