Supplemental Documentation Volume 5

Precipitation Analysis

Updated: September 1, 2009

Table of Contents

Volume 5: Precipitation Analyses v5-1

v5.1 METEOROLOGIC CHARACTERIZATION v5-1

v5.1.1 Long-Term Meteorologic Conditions v5-1

v5.1.2 Local Meteorologic Conditions v5-5

v5.1.3 Identifying a Representative 12-month Period in Precipitation Record v5-5

v5.1.4 Modifying the Selected 12-month Precipitation Record v5-12

v5.2 NORMALIZING RAIN GAGE NETWORK BIASES USING CALIBRATED RADAR RAINFALL ESTIMATES v5-21

v5.2.1 Introduction v5-21

v5.2.1.1 Homogeneity of Rain Gauge Station Records v5-21

v5.2.1.2 Normalizing Rain Gage Network Biases v5-22

v5.2.1.3 Background v5-22

v5.2.2 Homogeneity of PWD Rain Gauge Station Records v5-23

v5.2.2.1 Data Set v5-23

v5.2.2.2 Double Mass Regression and Cumulative Residual Time Series Analysis v5-24

v5.2.2.3 Adjusting Heterogeneous Gauge Records v5-26

v5.2.3 Normalizing Spatial Biases in Rain Gage Data Using Calibrated Radar Rainfall Estimates v5-28

v5.2.3.1 Data Set v5-28

v5.2.3.2 Bias Adjustment Using Double Mass Regression v5-29

v5.2.3.3 Results v5-32

v5.3 INVERSE DISTANCE SQUARED WEIGHTING AND BASIN AVERAGE RAINFALL CALCULATIONS v5-39

VOLUME 5 PRECIPITATION ANALYSES

v5.1 METEOROLOGIC CHARACTERIZATION

The United States Environmental Protection Agency (US EPA) CSO Control Policy (1994) requires the characterization of the CSS area and evaluation of control measure performance in terms of system-wide average annual hydrologic conditions. The identification of an average annual precipitation record, therefore, is critical for the evaluation of CSS performance.

v5.1.1 Long-Term Meteorologic Conditions

The hydrologic conditions over the Philadelphia CSS area are characterized using the long-term historic hourly precipitation record, 59-year period (1948-2006), for the National Weather Service Cooperative Station located at the Philadelphia International Airport (PIA) (WBAN#13739). Statistical analyses of the long-term record are performed to determine the average frequency, volume, and peak intensity of rainfall events.

Identification of long-term average hydrologic conditions over the CSS is based primarily upon average annual and monthly precipitation volumes determined from the long-term record at the PIA. Comparisons are made between the individual annual precipitation volumes and the long-term average to identify relatively 'wet' and 'dry' years.

Figure v5-1 presents total annual precipitation volumes at the PIA for the years 1948-2006 along with one standard deviation from the mean. By this measure, 1983 and 1922 are the wettest and driest years on record, respectively. Furthermore, it is seen that during the past 15-years (since 1990) one year, 1996, is characterized as being wet and five individual years are characterized as being dry by having a total annual precipitation volume greater than one standard deviation from the mean.

Figure v5-2 shows the average monthly precipitation volumes relative to a range of plus and minus one standard deviation from the mean based upon the PIA historical record. Table v5-1 presents accompanying historical monthly precipitation volume statistics. Long term seasonal variation in monthly precipitation volumes can readily be seen between summer and winter.

The PIA long-term empirical cumulative distribution function of hourly rainfall intensity is presented in Figure v5-3.

Event Based Precipitation Analyses

Event based analysis of the long term precipitation record is used to best represent average annual CSO frequency and volume statistics needed for presumptive measurement of collection system performance. These event statistics are specific for a given minimum inter-event time (MIT) used for event definition.

A minimum inter-event time (MIT) is chosen for event definition so that the coefficient of variation (the ratio of the standard deviation to the mean) of inter-event times most closely approximates unity. This follows an exponential distribution of inter-event times for which the mean equals the standard deviation, and is based on the results of National Urban Runoff Program (EPA 1993). A six-hour minimum inter-event time is selected on this basis for the PIA using hourly precipitation data for the period 1948-2006 as seen in Table v5-2.

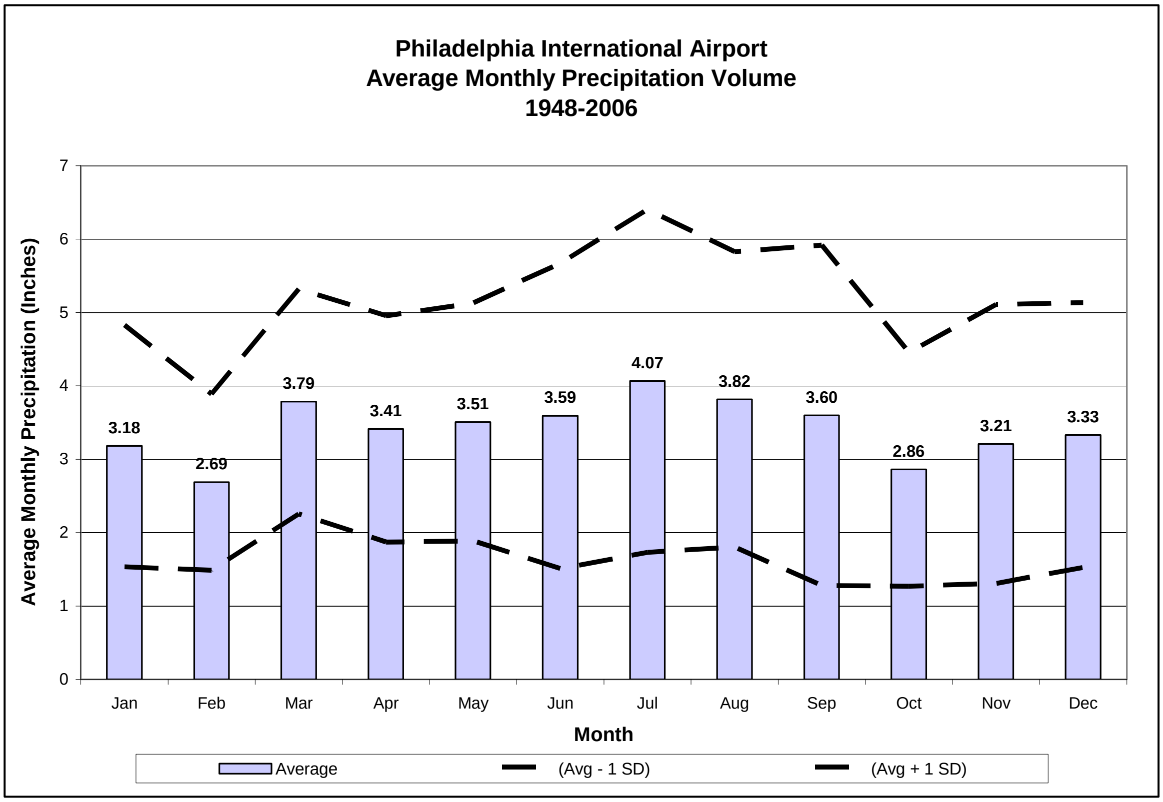

Figure v5-1 PIA Total Annual Precipitation Volume (1948-2006)

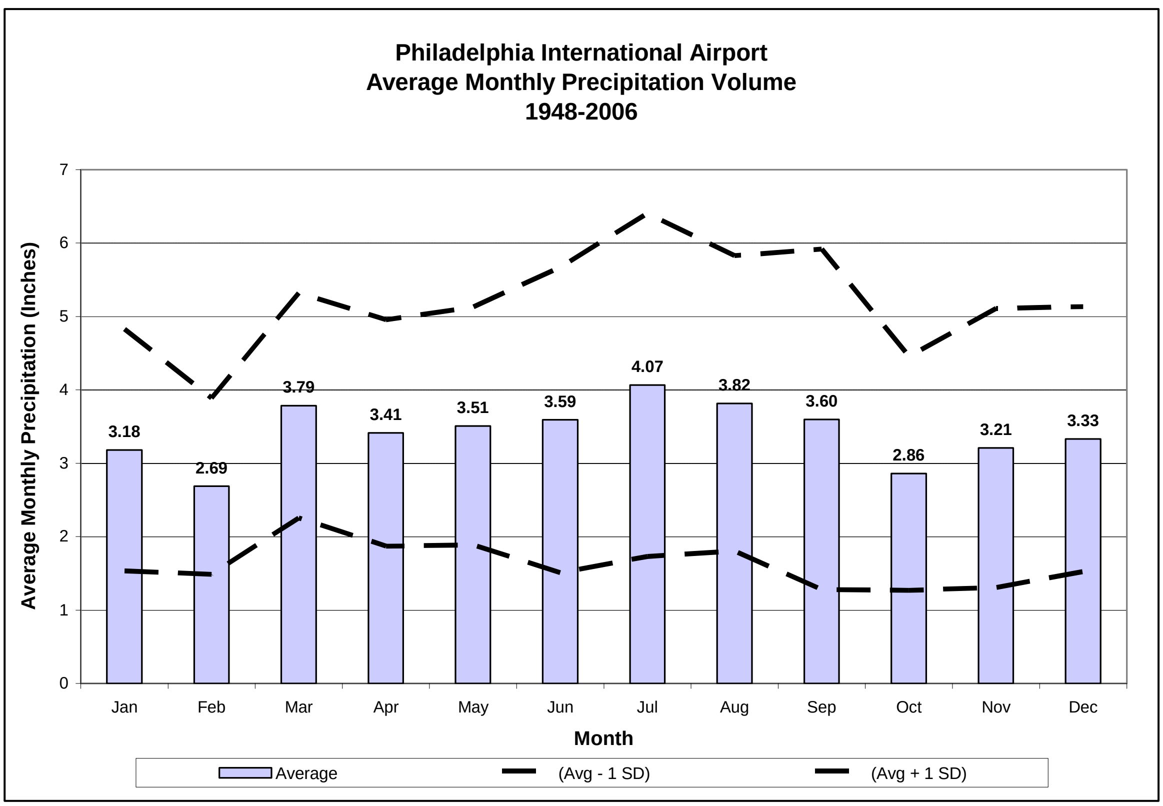

Figure v5-2 PIA Average Monthly Precipitation Volume (1948-2006).

Table v5-1 Monthly Precipitation Inches Statistics for PIA Historical Record (1948-2006)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual
Average	3.18	2.69	3.79	3.41	3.51	3.59	4.07	3.82	3.60	2.86	3.21	3.33	41.05
Avg +1SD	4.83	3.89	5.32	4.95	5.13	5.67	6.40	5.83	5.92	4.46	5.11	5.14	47.71
Avg - 1SD	1.54	1.49	2.26	1.87	1.89	1.51	1.73	1.80	1.28	1.27	1.31	1.53	34.39
Std. Dev.	1.65	1.20	1.53	1.54	1.62	2.08	2.34	2.01	2.32	1.59	1.90	1.80	6.66
Maximum	8.86	6.44	6.89	8.12	7.03	8.08	10.42	9.70	13.07	8.68	9.05	8.09	54.41
Minimum	0.45	0.46	0.69	0.61	0.48	0.11	0.37	0.49	0.21	0.09	0.32	0.25	29.34

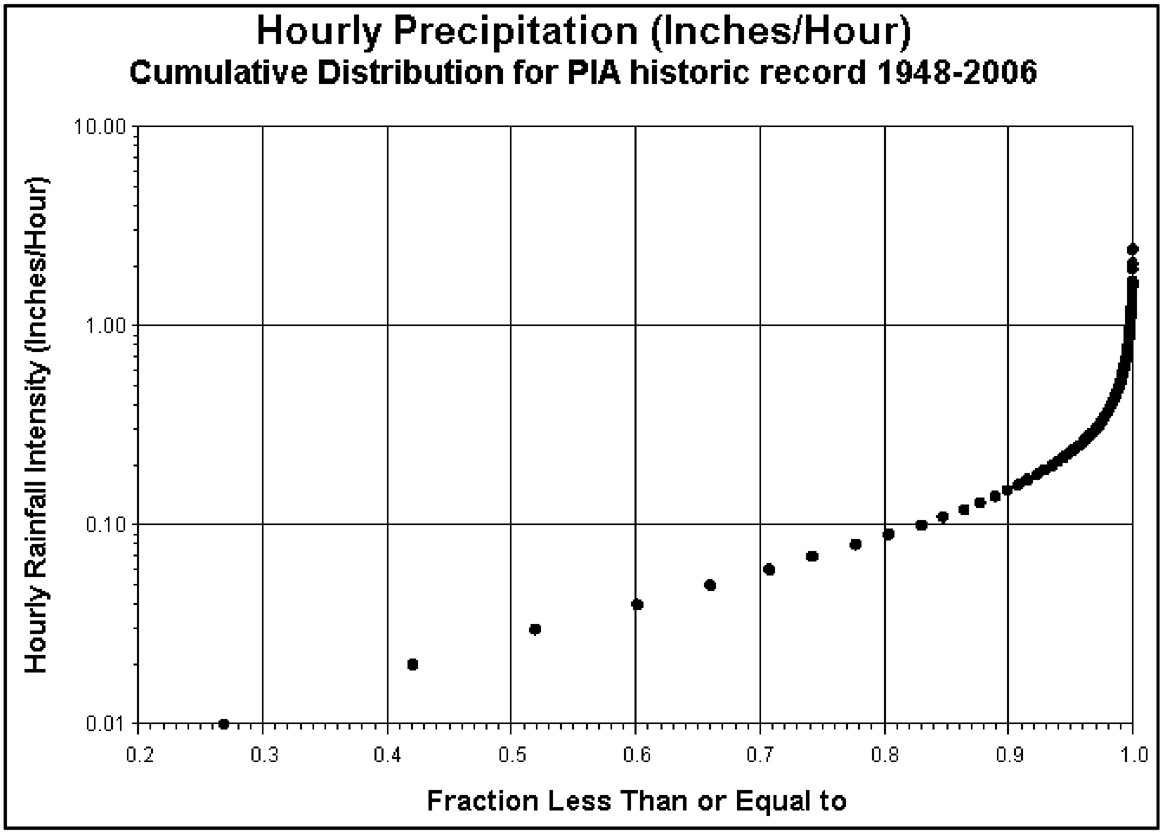

Figure v5-3 PIA Empirical Cumulative Distribution Function of Hourly Rainfall Intensity (1948-2006)

A minimum total event volume of 0.10 inches is selected as the minimum storm depth needed for precipitation events to significantly increase wastewater flows potentially contributing to CSO discharges. Table v5-3 presents event-based summary statistics for the PIA long-term precipitation record.

Table v5-2 Inter-event Time (IET) Statistics Determined for a Range of Minimum Inter-Event Times (MIT) using PIA Hourly Precipitation (1948-2006)

MIT (Hours)	Mean IET (Hours)	Std. Dev.IET (Hours)	CV IET
2	48.2	70.7	146.5
4	66.2	76.2	115.1
6	75.5	77.5	**102.7**
8	81.4	78.0	95.8
10	85.6	78.2	91.3
12	89.5	78.2	87.4
14	92.7	78.2	84.4
16	95.2	78.2	82.1
18	97.5	78.1	80.1
20	99.5	78.1	78.4
22	101.8	78.0	76.6
24	104.0	77.9	74.9

Table v5-3 Philadelphia International Airport Average Annual Wet Weather Event Statistics (1948-2006)

Month	Event Size Class	Average Number of Events	Average Total Rainfall (Inches)	Average Event Peak Hourly Intensity (In / hour)	Average Event Duration (hours)	Average Inter-Event Time (hours)
1	>= 0.05 in	6.4	3.04	0.11	11.2	83.2
2	>= 0.05 in	5.9	2.66	0.11	11.1	82.0
3	>= 0.05 in	7.1	3.81	0.14	10.9	83.6
4	>= 0.05 in	7.1	3.27	0.15	9.4	66.5
5	>= 0.05 in	7.6	3.46	0.18	7.9	73.5
6	>= 0.05 in	7.3	3.51	0.25	5.8	79.5
7	>= 0.05 in	7.2	4.02	0.29	5.6	83.7
8	>= 0.05 in	6.7	3.77	0.32	6.0	90.3
9	>= 0.05 in	5.7	3.58	0.26	8.1	95.7
10	>= 0.05 in	4.9	2.82	0.19	9.3	115.1
11	>= 0.05 in	5.7	3.16	0.16	9.9	100.1
12	>= 0.05 in	6.0	3.31	0.13	11.9	89.4
All	>= 0.05 in	77.6	40.39	0.19	8.7	77.1
All	< 0.05 in	30.3	0.62	0.02	1.7	74.6
All	**All**	**107.9**	**41.05**	**0.14**	**6.7**	**76.4**
* Events defined based on 6 hour Minimum Inter-Event Time (MIT)						

v5.1.2 Local Meteorologic Conditions

The average spatial distribution of precipitation over the CSS areas is characterized using the 17-year rainfall record for the PWD 24-raingage network collected over the period 1990-2006, along with 15 months of gage calibrated radar rainfall data. Extensive analyses of non-climatic gage biases based on inter-gage comparison and radar rainfall data are performed leading to the creation of a bias adjusted rainfall dataset for the PWD 24-raingage network over the 17-year period of record (1990-2006). The detailed analyses are presented in Section v5.2 Normalizing Rain Gage Network Biases Using Calibrated Radar Rainfall Estimates.

v5.1.3 Identifying a Representative 12-Month Period in Precipitation Record

The characterization of long-term system-wide average hydrologic conditions across the CSS is necessary in order to identify a continuous short–term period contained within the PWD 24-gage fifteen-minute rainfall record (1990-present) that simulates long-term average annual CSO statistics needed for performance evaluation of CSO control measures.

CSO occurrence is considered to be a complex function of storm-event characteristics such as total volume, duration, peak intensity, and length of antecedent dry period or inter-event time (IET). In order to identify short-term continuous periods likely to generate CSO statistics representative of the long-term record, continuous 12-month periods selected from the recent 17-year PWD 24-raingage record (1990-2006) were evaluated against the long-term record based on the following storm-event characteristics:

- Annual number of storm events
- Total annual rainfall volume
- Best fit cumulative distribution function (CDF) plot of event peak hourly rainfall intensity

Manipulation of the rainfall data is performed using Statistical Analysis System (SAS) code. SAS is a high-level programming language that is particularly well suited to processing large amounts of data with relatively simple programming code.

The first step in the analyses is to parse the bias adjusted and inverse distance-squared weight (IDW) filled 17-year precipitation record (1990-2006) for each gage location into continuous 12-month periods beginning with January 1, 1990 and progressing with 1-month increments to the final continuous 12-month period beginning on January 1, 2006. Each continuous 12-month period is thereby identified by the starting year and month.

Next, event statistics, including total volume, average duration, average peak intensity, and average inter-event time (IET) are determined, based on a minimum inter-event time (MIT) of 6 hours, for each 12-month period. Small events, defined as events with total volumes less than 0.05 inches, are removed from further analysis. Similarly, average annual event statistics are determined for each gage location over the s17-year period of record (1990-2006). The differences in average event statistics between each continuous 12-month period and the period of record are determined for each gage location. The absolute value of the average difference across all gage locations is then determined for each continuous 12-month period, and the result is then ranked in order of ascending magnitude as a measure of goodness of fit to the long term average for each event statistic.

The cumulative frequency distribution of event peak rainfall intensity is considered to be a critical measure for identifying rainfall periods that produce average long-term CSO statistics. Event peak hourly rainfall intensities are ranked and a left-continuous empirical CDF is generated with fractional ranks computed by dividing each rank by the denominator n+1, where n is the number of events. In this manner empirical cumulative distribution functions are generated for each continuous 12-month period and the 17-year period of record for each rain gage. The sum of the absolute differences in peak hourly rainfall intensity between the 12-month period and the 17-year period of record, determined for each event within the 12-month period based on its fractional rank, is used to measure the goodness of fit for each rain gage and 12-month period. This measure, referred to here as the total deviation, is averaged across all rain gages for each 12-month period and ranked in ascending order.

A final ranking is performed based on total deviation of peak hourly rainfall intensity, average annual rainfall volume, and average annual number of rain events. The top results from this ranking are presented in Table v5-4. The 12-month period beginning January 1, 2005 is chosen to represent long-term average hydrologic conditions for Long Tern Control Plan CSO performance evaluations based on the additional criteria that it is a recent calendar year.

Table v5-4 Ranking of Recent Representative Continuous 12-month Periods Based on the Best-Fit Distribution of Event Peak Hourly Rainfall Intensity, Average Annual Rainfall Volume, and Annual Number of Events

Start Year	Start Month	Average Event Peak (in/hr)	Annual Rainfall Volume (in)	Average IET (hrs)	Average Event Duration (hrs)	Annual Number of Events	Rank
1990 - 2006		0.11	42.80	76.54	9.74	76	
1992	3	0.10	43.57	74.48	10.33	79	1
1996	11	0.09	41.66	81.47	10.57	69	2
1992	8	0.10	44.08	76.52	9.95	78	3
1995	7	0.10	41.45	77.67	9.06	80	4
2004	12	0.11	43.51	77.27	10.63	76	5
1992	2	0.10	42.32	76.16	10.04	81	5
1992	1	0.10	41.36	78.87	10.15	79	5
2003	4	0.10	43.17	62.09	9.98	83	8
2005	3	0.11	43.02	82.31	10.00	75	9
2005	**1**	**0.11**	**44.06**	**78.76**	**10.26**	**79**	**10**
2000	8	0.11	41.43	77.65	9.34	73	11
1998	11	0.09	43.20	83.60	9.42	69	12
2002	7	0.08	41.67	72.39	11.85	76	13
1998	12	0.09	43.93	79.20	9.43	70	13
1997	7	0.09	43.09	71.39	11.92	80	15
1994	3	0.12	43.11	73.77	9.01	79	16
1997	8	0.09	41.87	71.42	11.77	80	17

* Only continuous periods with annual rainfall volumes within +/- 1.5 inches of the 17-year average annual rainfall volume (41.30 to 44.30 inches) were considered. Only rainfall events with total volume >= 0.05 inches, based on a 6-hour M.I.T. are included in the analysis

Table v5-5 compares selected PIA precipitation event statistics for the calendar year 2005 to PIA long-term historic median values. Events with total volumes less than 0.05 inches were excluded from the analysis because they are not expected to significantly influence CSO statistics.

Table v5-5 Seasonal Precipitation Event Statistics Comparing Long-Term Historic Record to Calendar Year 2005*

Statistic	2005		1948-2006	
	Recreation Season **	Annual	Recreation Season ** Median	Annual Median
Number of Events	41	78	40	76
Mean Event Volume (in)	0.53	0.53	0.52	0.53
Maximum Event Volume (in)	6.05	6.05	2.54	3.07
Mean Event duration (hr)	7.28	8.58	6.69	8.58
Mean Event Average Rainfall Intensity (in/hr)	0.09	0.08	0.10	0.07
Mean Event Peak Rainfall Intensity (in/hr)	0.24	0.19	0.25	0.19
Std. Dev. of the Mean Event Peak Rainfall Intensity (in/hr)	0.27	0.22	0.26	0.21
Maximum Event Peak Rainfall Intensity (in/hr)	1.55	1.55	1.17	1.17
Total Rainfall (in)	21.80	41.69	20.94	40.50

* Only rainfall events with total volume >= 0.05 inches, based on a 6-hour M.I.T. are included in the analysis

** Recreation season includes months May - October

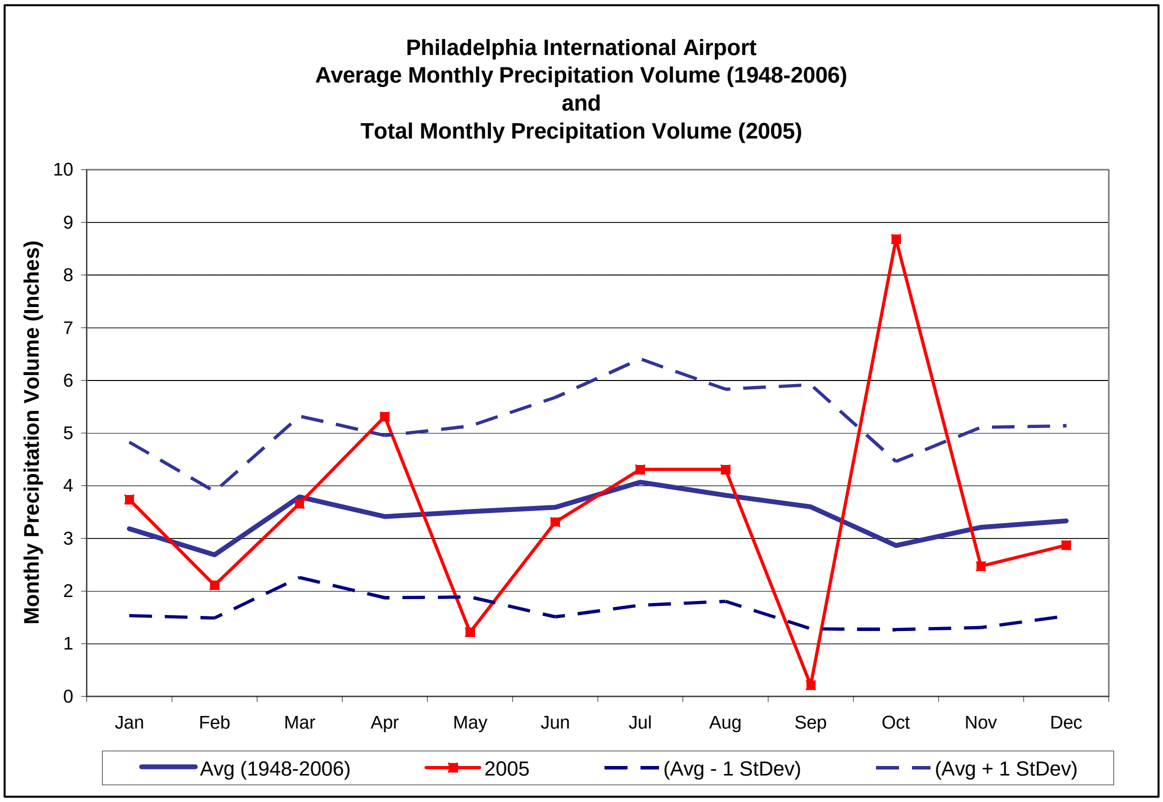

Figure v5-4 PIA Average Monthly Precipitation Volume Comparing the Long-Term Record (1948 – 2006) and Calendar year 2005

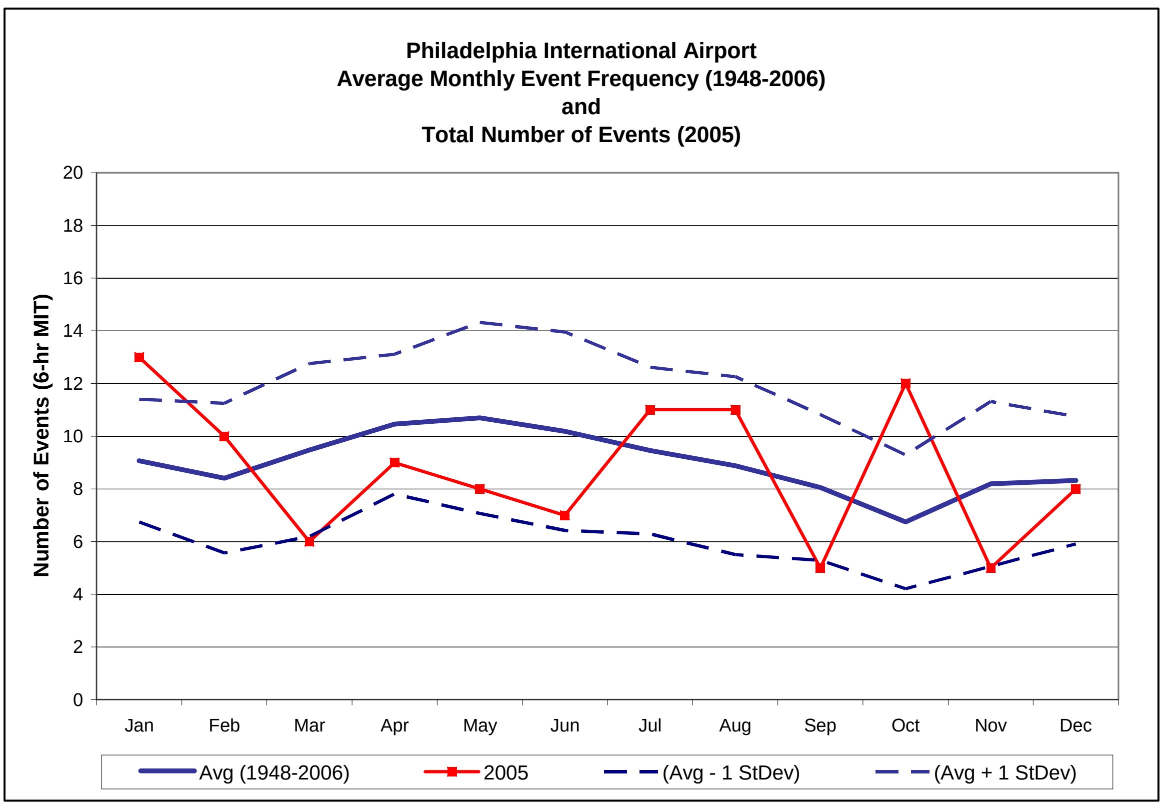

Figure v5-5 PIA Average Monthly Number of Events Comparing the Long-Term Record (1948 – 2006) and Calendar year 2005

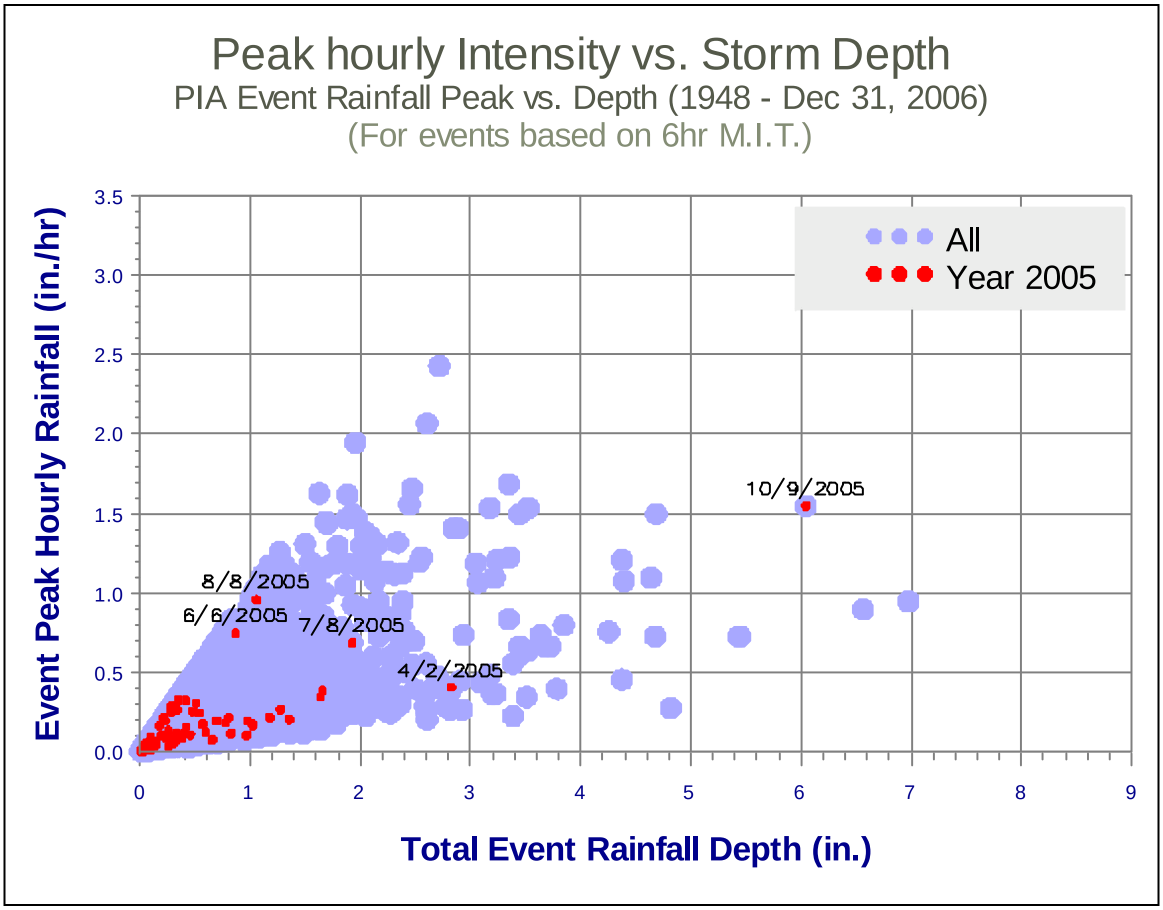

Figure v5-6 PIA Event Peak Hourly Rainfall Intensity Plotted and Rainfall Volume Comparing the Long-Term Record (1948 – 2006) and Calendar Year 2005

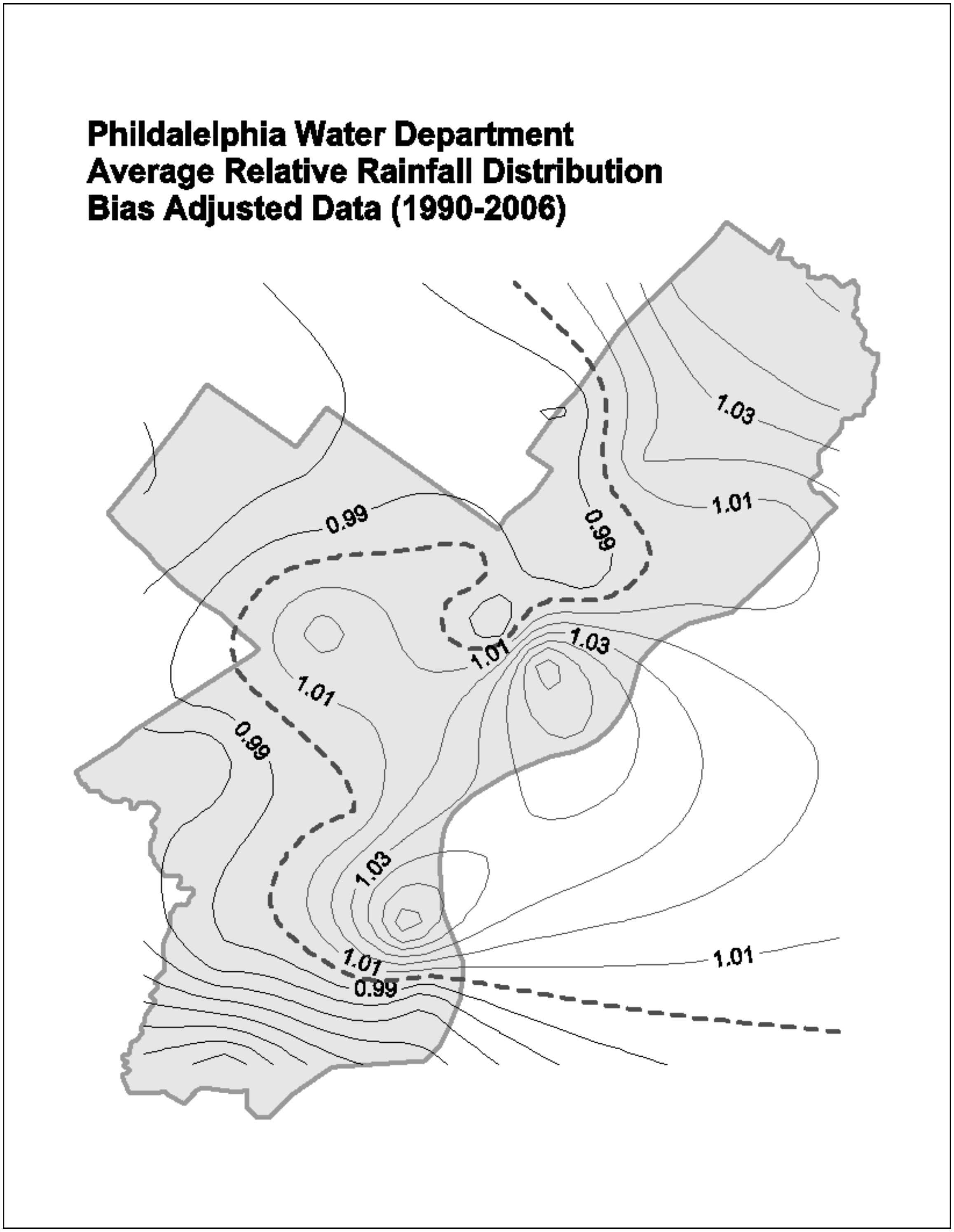

Figure v5-7 Relative Rainfall Distribution Map for PWD 24-raingage Network Bias Adjusted Data for the 17-year Period (1990-2006)

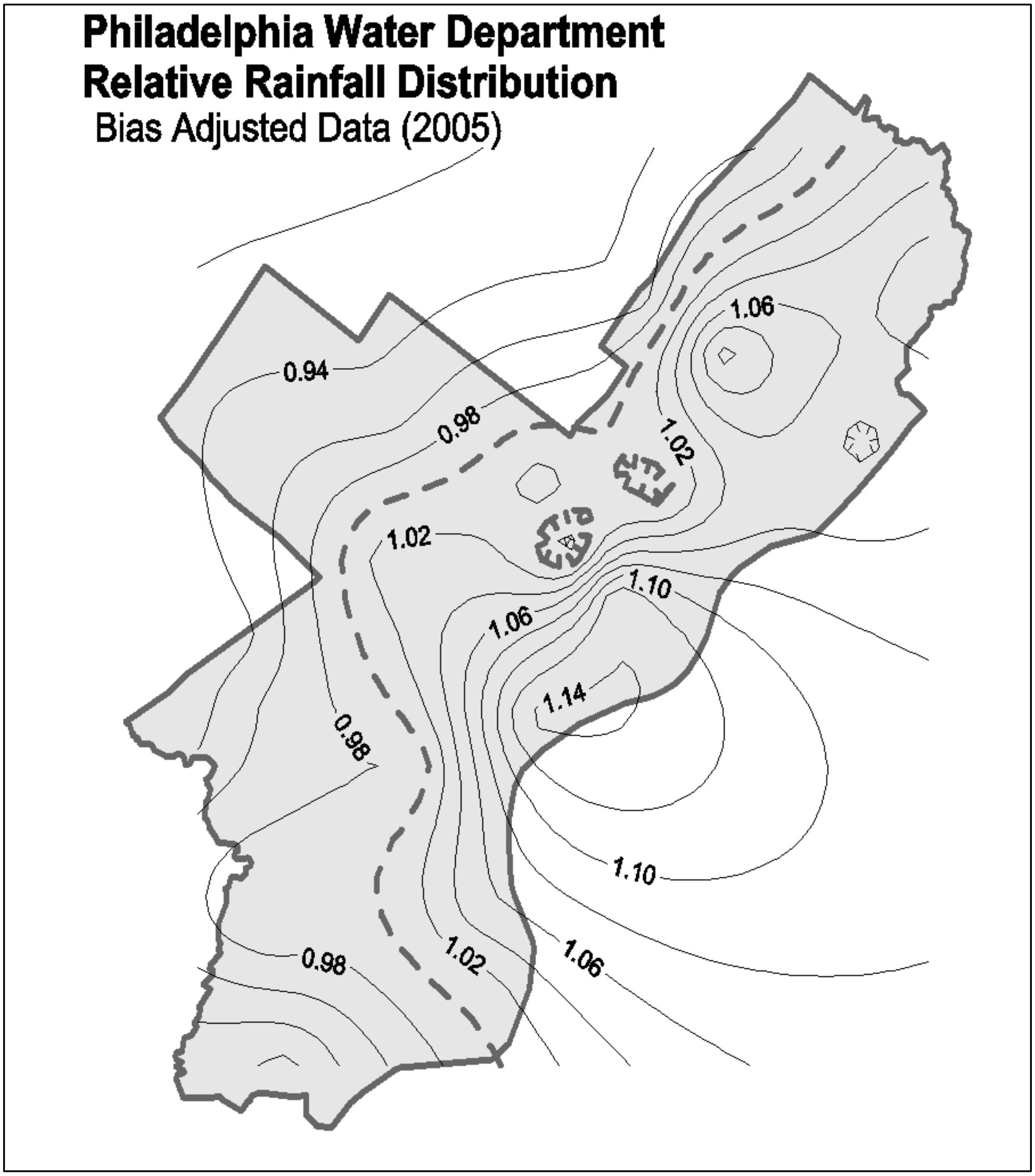

Figure v5-8 Relative Rainfall Distribution Map for PWD 24-raingage Network Bias Adjusted Data for Calendar Year 2005

v5.1.4 Modifying the Selected 12-month Precipitation Record

Initial selection of the calendar year 2005 to represent long-term average hydrologic conditions for CSO LTCP project evaluations was based on the annual number of storm events, the total annual rainfall volume, and the best fit CDF plot of event peak hourly rainfall intensity, with preference given to more recent calendar years to better represent current conditions.

The calendar year 2005, however, contains the extreme event of October 8, 2005 which recorded an average rainfall volume across the PWD 24-gage network of 5.40 inches between October 7 at 12:15 PM and October 9 at 8:45 AM. This rainfall event has the third largest annual peak rainfall volume recorded at the Philadelphia International Airport (PIA) station over the long-term period of 1948-2006. Because the extreme rainfall event of October 8, 2005 accounts for a disproportionately large fraction of the total annual overflow volume the results of CSO LTCP project evaluations may be unintentionally skewed to minimize the long-term effectiveness of certain alternatives in favor of others.

In response to these concerns, a decision was made to adjust the rainfall record for the calendar year 2005 to better represent long-term average hydrologic conditions by scaling down the October 8 rainfall event so that the average rainfall volume across the PWD 24-gage network for this event is equal to the median peak annual rainfall volume estimated for the network over the long-term period of 1948-2006.

After scaling down the October 8 event, several other events are selected to be scaled up so that the average total rainfall volume across the PWD 24-raingage network is equal to the long-term average annual rainfall volume estimated across the network for the long-term period 1948-2006.

Time Series Modification Procedures

Median Peak Annual Rainfall Volume

The median peak annual rainfall volume is estimated for the PWD 24-raingage network over the long-term period of 1948-2006 by scaling the PWD median peak annual rainfall event volume for the 17-year period of 1990-2006 by the ratio of PIA median peak annual rainfall event volume for the long-term period of 1948-2006 to that for the period 1990-2006. The result is an estimated average peak annual rainfall event volume across the PWD 24-gage network of 3.40 inches based on a 6-hr Minimum Inter-event Time (MIT) as presented in Table v5-6.

Table v5-6 Median Annual Peak Rainfall Event Volumes (MIT = 6 hrs)

	1948-2006	1990-2006
	(inches)	(inches)
PIA	3.18	2.78
PWD	**3.40**	2.97

October 8, 2005 Rainfall Event Scaling

The October 8, 2005 extreme rainfall event is scaled down by multiplying the time series data for each of PWD rain gages by the factor 0.630. This scaling factor is determined as the ratio of the average median peak annual rainfall volume estimated for PWD 24-raingage network over the long-term period of 1948-2006 (3.40 inches), to the average rainfall volume for PWD 24-raingage network during the October 8, 2005 event (5.40 inches).

Average Annual Rainfall Volume

The average annual rainfall volume across the PWD 24-raingage network for the recent period 1990-2006 is scaled up by a factor of 1.03 to estimate the average for the long-term period (1948-2006) based on the ratio of averages at the PIA for these periods. Simply stated, the PIA long-term average annual rainfall (1948-2006) is 3% greater than that for the more recent period 1990-2006.

Average annual rainfall volumes for the PIA and the PWD 24-raingage network are presented for each time period in Table v5-7. The long-term average annual rainfall volume across the PWD network is estimated to be approximately 44.79 inches.

Table v5-7 Average Annual Rainfall Volume Comparison

Rainfall Dataset	Average Annual Rainfall (inches)
PIA (1948-2006)	41.05
PIA (1990-2006)	39.84
PWD (1990-2006)	43.48
PWD (1948-2006) estimated	**44.79**
PWD 2005	44.53
PWD 2005 w/ Oct 8 Scaled Down	42.53
PWD Oct 8 Event	5.40
PWD Oct 8 Event Scaled Down	3.40

Scaling Representative Year to Match Long-Term Average Annual Rainfall Volume

The average annual rainfall volume across the PWD network for the year 2005, after scaling down the October 8 event, is 42.533 inches - approximately 2.25 inches less than the long-term average annual rainfall estimated for the PWD network of 44.785 inches. The 2.25 inches of rainfall are distributed back into the annual time series by selecting events to scale up based on the CDF of total event rainfall volumes.

The CDF plots of total event rainfall volume for each gage are considered in order to identify a range of event frequencies that have lower event volumes in the calendar year than in the long-term record (17-yrs).

The first step in this process is to generate a master event list based on the sum of the rainfall from all 24 PWD raingages using a 6-hr MIT. This allows the rainfall time series data for all rain gages to be scaled within the same selected set of event boundaries. The 18 events selected for distributing the 2.25 inches of rainfall needed on average to match the long-term average annual rainfall volume are presented in Table v5-8.

Table v5-8 Event Boundaries Selected for Scaling Up

Master Event No.	Start Time	End Time
2189	1/7/05 21:30	1/9/05 0:00
2190	1/11/05 15:00	1/12/05 1:15
2195	1/25/05 13:00	1/25/05 18:15
2204	2/4/05 9:00	2/4/05 16:00
2214	3/1/05 11:15	3/1/05 18:30
2215	3/8/05 4:15	3/8/05 14:15
2217	3/20/05 3:00	3/21/05 4:15
2228	4/23/05 6:30	4/24/05 1:45
2252	6/10/05 11:15	6/10/05 16:15
2257	7/1/05 16:00	7/1/05 19:00
2259	7/5/05 16:15	7/5/05 18:00
2280	8/8/05 6:30	8/9/05 11:15
2283	8/16/05 12:30	8/17/05 8:15
2287	8/27/05 8:15	8/27/05 19:30
2307	10/12/05 23:00	10/15/05 7:15
2328	12/9/05 7:45	12/10/05 0:00
2335	12/25/05 13:00	12/26/05 12:00
2337	12/29/05 7:00	12/29/05 23:30

A scaling factor is determined so that when multiplied by the time series data within the selected event boundaries for each gage then the average annual rainfall volume across all gages is equal to the long-term average. This factor for each gage is one plus the ratio of the total volume being added to the sum of the volumes of all events being selected for scaling. The total volume to be added for each gage is determined as the total annual volume for the gage excluding the volume for the October 8 event multiplied by the ratio of the average volume added to the average annual volume excluding the average volume for the October 8 event. The ratio of average volume added to average annual volume excluding the October 8th event is shown to be equal to (2.25 inches) / (44.53 inches – 5.4 inches) = 0.0576. Therefore, the scaling factor for each gage is determined by the formula:

$$\left[1 + 0.0576 \times \left(\mathrm{V}_{\mathrm{Annual}}^{RG} - \mathrm{V}_{Oct\,8}^{RG}\right) \middle/ \sum \mathrm{V}_{SelectedEvents}^{RG}\right]$$

The scaling factors applied to the selected events are presented for each gage in Table v5-9.

<u>Results</u>

The final results of the modification of the calendar year 2005 rainfall record is illustrated through the CDF plots of event rainfall volume produced for PWD RG-5 comparing the PWD period of record (1990-2006) and one of the following: calendar year 2005; modified calendar the year 2005. The two CDF plots produced for PWD RG-5 are presented in Figure v5-9 and Figure v5-10.

Table v5-9 Factors for Scaling Selected Rainfall Events for Each Raingage

RG	Factor	RG	Factor
1	1.32	13	1.348
2	1.304	14	1.35
3	1.362	15	1.337
4	1.331	16	1.347
5	1.334	17	1.364
6	1.329	18	1.348
7	1.354	19	1.372
8	1.359	20	1.351
9	1.313	21	1.384
10	1.356	22	1.325
11	1.344	23	1.301
12	1.335	24	1.383

In addition, Philadelphia International Airport (PIA) hourly rainfall data were used to generate CDF plots of total event rainfall (for events greater than or equal to 0.05 inches with MIT = 6hrs). These plots are presented in Figure v5-11 through Figure v5-13 comparing the following three time periods: 59-year period (1948-2006); 17-year period (1990-2006); calendar year 2005.

Comparing the CDF plots for each PWD gage indicates that event volumes corresponding to percentiles between 40% and 70% are generally lower for the calendar year 2005 before modification than for the 17-year average. Although these event volumes are increased after modification, they appear to be generally lower than the 17-year average. The same general relationship between calendar year 2005 and the 17-year average is seen at the PIA in Figure-4. Furthermore, PWD rain gage event volumes corresponding to percentiles between 70% and 90% are generally higher than the 17-year average before modification, and are increased further above the average after modification. A similar pattern is observed for the PIA.

Figure v5-14 and Figure v5-15 present cumulative frequency distribution plots of 15-minute rainfall intensities pooled for all 24 PWD rain gages over the period 1990-2006 for frequencies of occurrence less than or equal to 50% and greater than or equal to 50%, respectively .

Figure v5-16 presents a relative rainfall distribution map based on Inverse Distance Squared (IDS) weighting of 1-km square grid cells from bias adjusted PWD 24-raingage Network data for the modified representative year 2005. This IDS grid rainfall record is basin averaged and used as input for all hydrologic models as part of the LTCPU as described in Section v5.3.

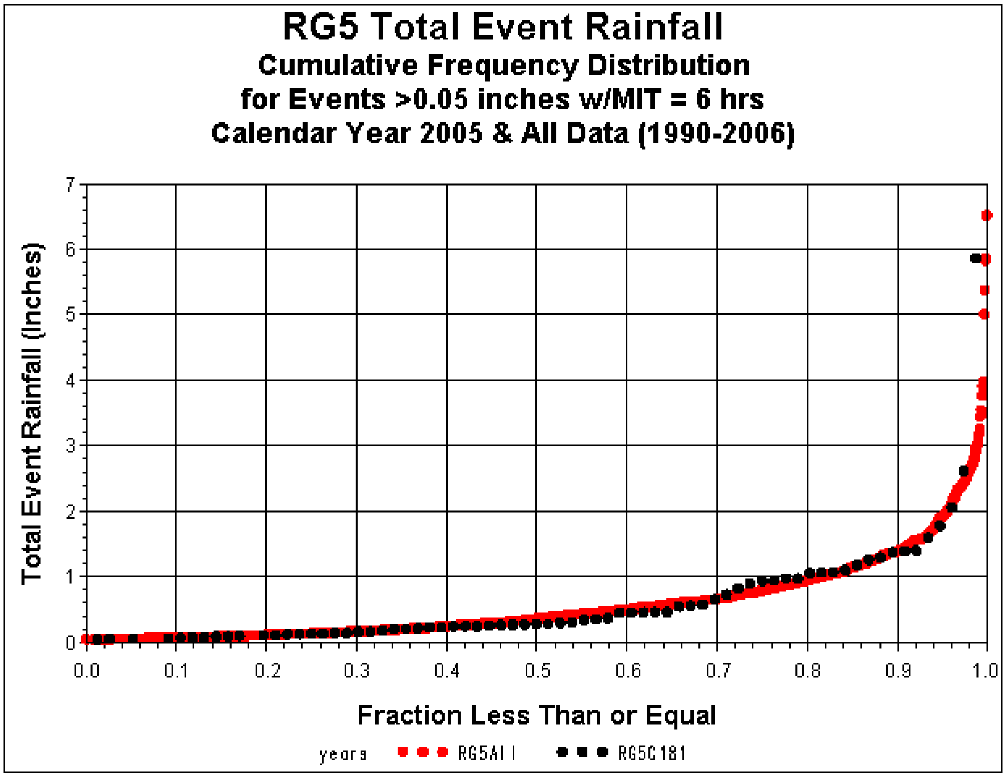

Figure v5-9 CDF Plot of RG-5 Rainfall Event Volumes Comparing the 17-year Period (1990-2006) and Calendar Year 2005

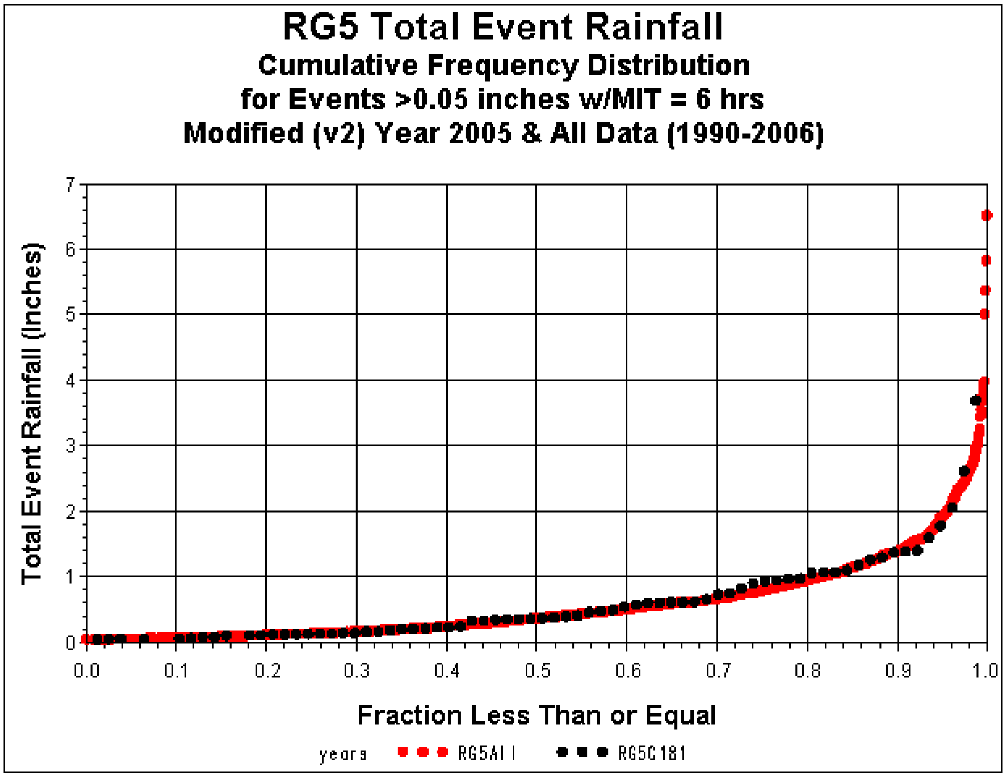

Figure v5-10 CDF Plot of RG-5 Rainfall Event Volumes Comparing the 17-year Period (1990-2006) and Modified Calendar Year 2005

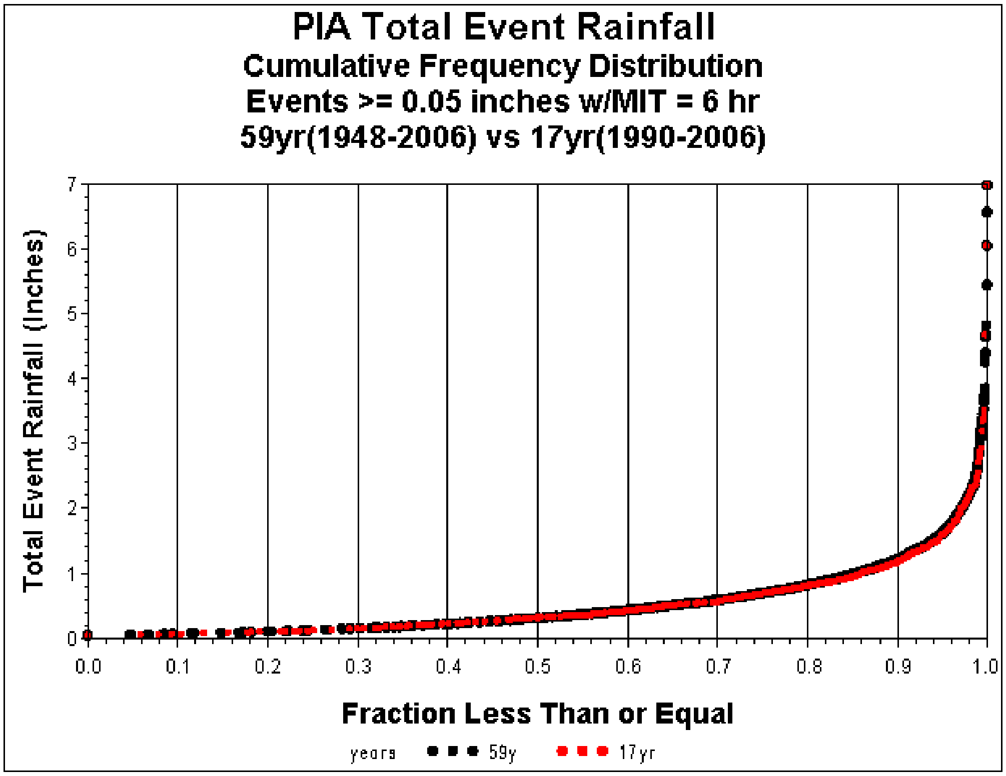

Figure v5-11 CDF Plot of Event Rainfall Volume at the PIA Comparing the 17-year Period (1990-2006) to the 59-year Period (1948-2006)

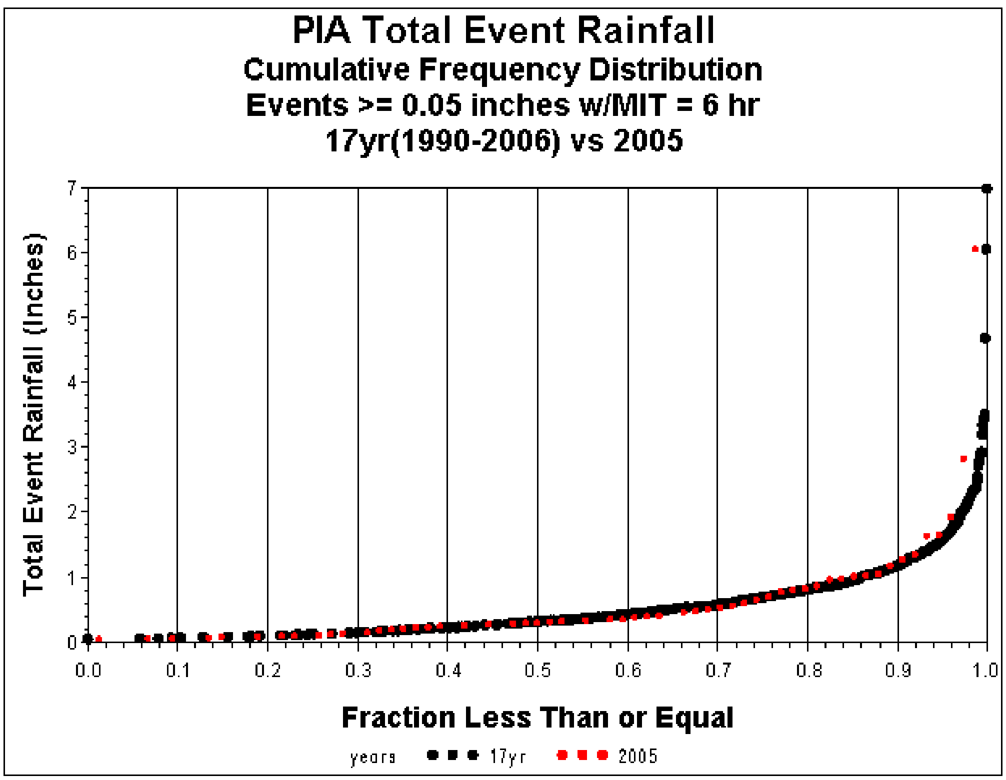

Figure v5-12 CDF Plot of Event Rainfall Volume at the PIA Comparing the 17-year Period (1990-2006) to the Calendar Year 2005

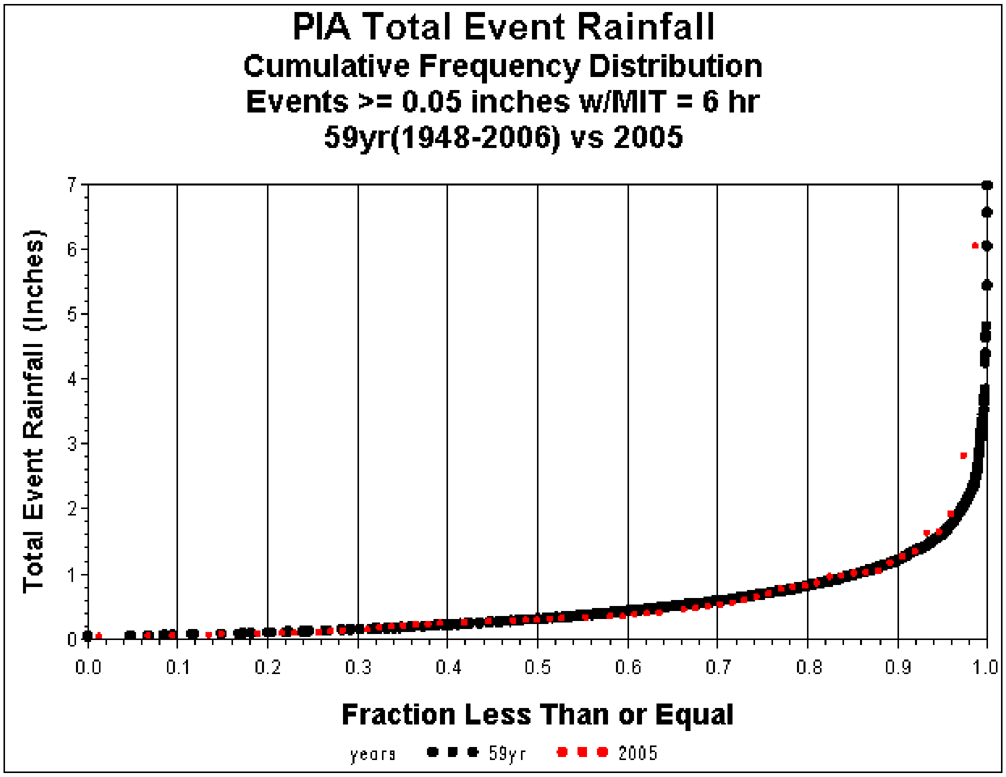

Figure v5-13 CDF Plot of Event Rainfall Volume at the PIA Comparing the 59-year Period (1948-2006) to the Calendar Year 2005

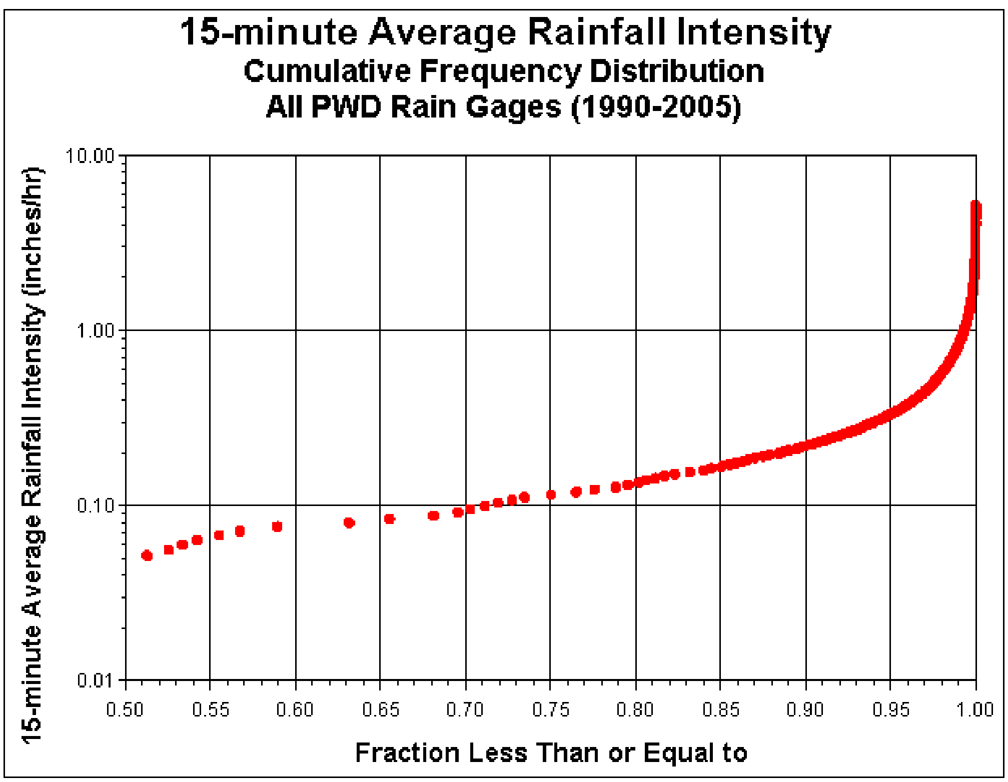

Figure v5-14 CDF Plot of 15-minute Rainfall Intensity for All 24 PWD Rain Gages (1990-2005) Less than or Equal to 50%

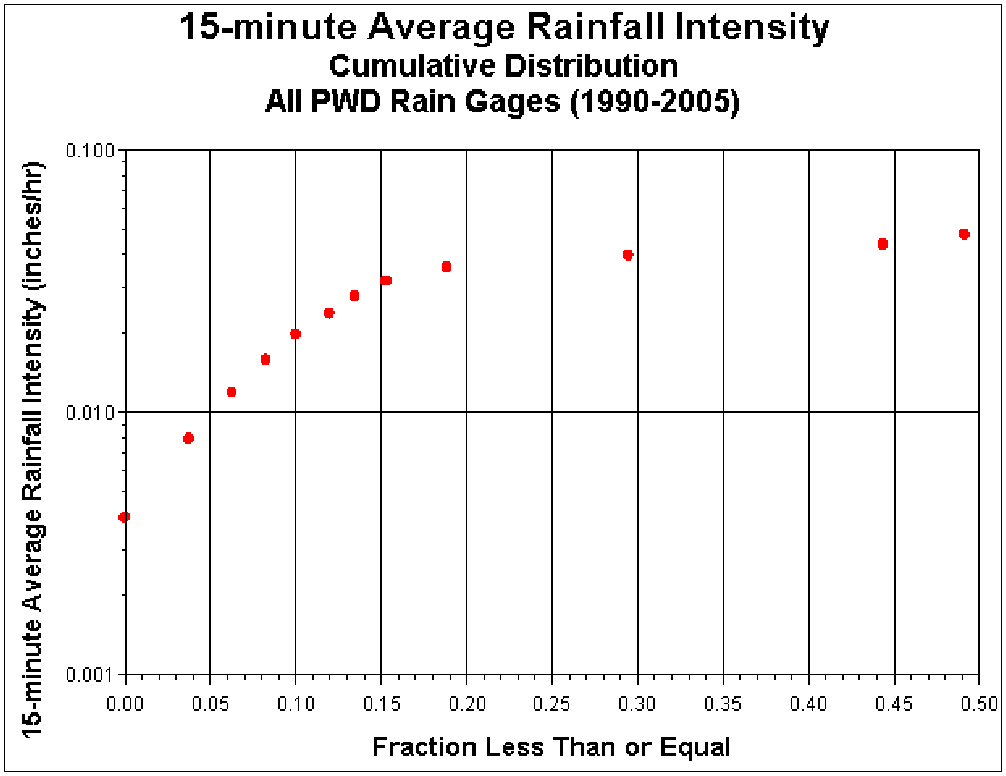

Figure v5-15 CDF Plot of 15-minute Rainfall Intensity for All 24 PWD Rain Gages (1990-2005) Greater than or Equal to 50%

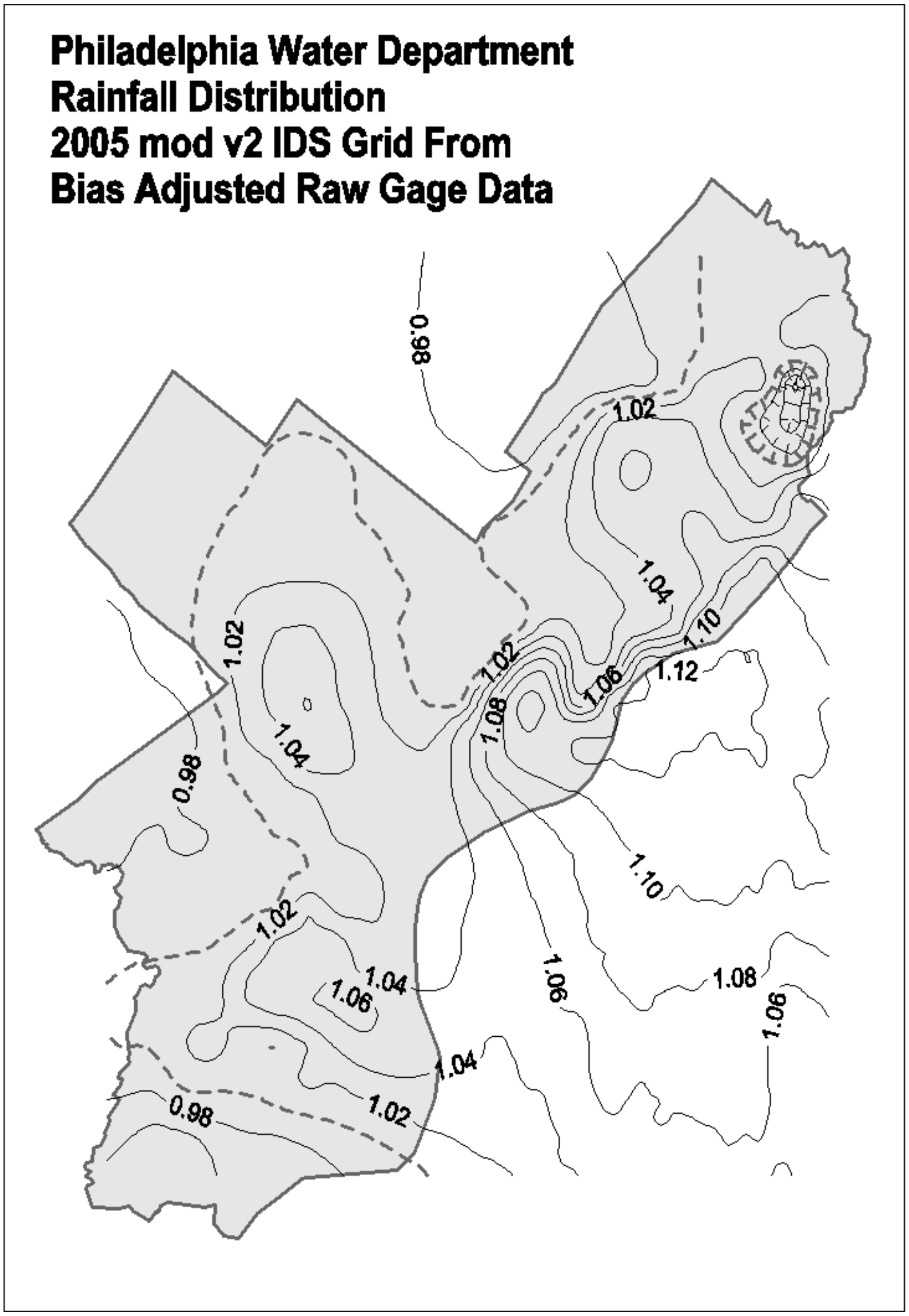

Figure v5-16 Relative Rainfall Distribution Map of Inverse Distance Squared (IDS) Weighting 1-km square Grid Bias Adjusted PWD Rain Gage Network Data for Modified Year 2005

v5.2 Normalizing Rain Gage Network Biases Using Calibrated Radar Rainfall Estimates

The identification and adjustment of precipitation time series data for non-climatic changes in recording bias among rain gages can be instrumental in controlling uncertainty in hydrologic models. Hydrologic models depend upon the reliability of precipitation and flow monitoring data sets used for calibration and simulation. Consistent precipitation and flow monitoring measurements clearly can be important when attempting to characterize rainfall runoff relationships over time. Hydrologic models require rain gage networks to represent the spatial distribution of precipitation across a drainage basin and benefit from the normalization of relative rain gage biases across the network.

Calibration of large urban sewer system models, using a moderately-dense basin-wide rain gage network and continuous flow monitoring data, is improved by creating continuous homogeneous rainfall records with normalized spatial biases.

Double-mass regression and cumulative residual time series analysis techniques are used to evaluate and adjust historical rain gage network data to correct for non-homogeneity of individual rainfall records and to normalize spatial bias across the network. Homogeneity of rainfall time series data is evaluated and adjusted by comparison to the rain gage network mean over a 13-year period of record. Spatial bias across the network, then, is normalized by comparison to continuous calibrated radar rainfall estimates obtained over a 15-month period. Cumulative residual time series analysis techniques also are applied to evaluate the homogeneity of flow monitoring data used in model calibration. The benefits of normalizing the rain gage network biases to model calibration are illustrated by comparing model results using gage data with and without bias correction.

v5.2.1 Introduction

v5.2.1.1 Homogeneity of Rain Gage Station Records

Hydrologic model calibration of lumped runoff parameter estimates depends on consistent rather than precise absolute precipitation and flow monitoring measurement over time. Homogeneity of a rain gage record refers to the consistency of non-climatic bias in precipitation measurements at a gage location over its period of record. Changes in the method of measurement, location of the gage, or conditions immediately surrounding the gage, can cause the readings to differ systematically from prior readings and are indications of a need for correction (Easterling *et al.*, 1995). Homogeneity adjustment of rain gage data is performed to create a consistently scaled rainfall record at each gage location.

Adjustment of gage data to form homogeneous time series depends upon the ability to identify times when changes in measurement conditions may have occurred (Alexandersson 1986). Meta-data, a gage history record documenting changes in equipment and site conditions, often is used as the primary means for identifying changes in rain gage measurement conditions (Guttman 1998). Meta-data alone, however, is often insufficient for identifying non-homogeneity of gage records. Major reported equipment or station location changes can have little if any effect on the gage record, whereas seemingly minor adjustments and undocumented changes in site conditions may result in profound changes in the rainfall record as identified by analytic methods (Peterson *et al.*, 1998).

Time series analysis methods used for evaluating homogeneity and adjusting rain gage records depend upon comparison of gage data to a homogeneous reference time series. An appropriate reference time series is created by averaging measurements from several highly correlated nearby gages (Guttman 1998, Peterson *et al.*, 1998). It is also found valuable to include gages with short and incomplete data series in the reference value (Alexandersson 1986).

v5.2.1.2 Normalizing Rain Gage Network Biases

Representing precipitation spatially across a water or sewer-shed depends upon consistent recording among gages in a rain gage network. Once homogeneous rain gage records are created, it is important that all gages in the network be scaled to combine data for use in filling missing records. The goal is to develop a continuous rainfall record for each gage location, and to determine spatial bias adjustment factors to consistently represent the spatial distribution of rainfall across the network.

Normalizing rain gage network biases in this manner depends upon a reliable reference precipitation data set that represents spatial variation across the region with a uniform bias over a sufficiently long period of record. Calibrated radar rainfall estimates are used for this purpose.

v5.2.1.3 Background

The Philadelphia Water Department (PWD), as part of the City of Philadelphia's combined sewer overflow (CSO) permit compliance program, developed system hydraulic models of its separate sanitary and combined sewer systems that contribute flows to each of its three water pollution control plants, draining nearly 140 square miles of the city. The City maintains a network of 24 tipping bucket rain gages as part of this program. In addition, the City has obtained 18 months of largely continuous historical gage calibrated radar rainfall estimates provided by NEXRAIN Corporation, in order to further refine calibration of its large complex hydraulic system models. A map of Philadelphia showing approximate locations of PWD rain gages, radar rainfall grid, as well as, the combined and sanitary sewer service areas is presented in Figure v5-17.

Comparison of long term rainfall accumulations at neighboring gages revealed potentially significant systematic differences in non-climatic biases. Double mass and cumulative residual analyses of gage station records against the gage network mean value have further revealed non-homogeneity of station records due to changes in equipment operation or site conditions. Adjustment of rain gage data, therefore, was applied to create a consistently scaled precipitation record at each gage location. In addition to creating consistent (homogeneous) gage records over time, it also is important to scale all the gages consistently within the network to one another in order to combine data from different gages, for use in filling missing records, and representing spatial variation.

The goals of this investigation are to develop procedures to evaluate and adjust the historic PWD rain gage network record to produce homogeneous rain gage records at each gage location and to normalize rain gage network biases using radar rainfall estimates.

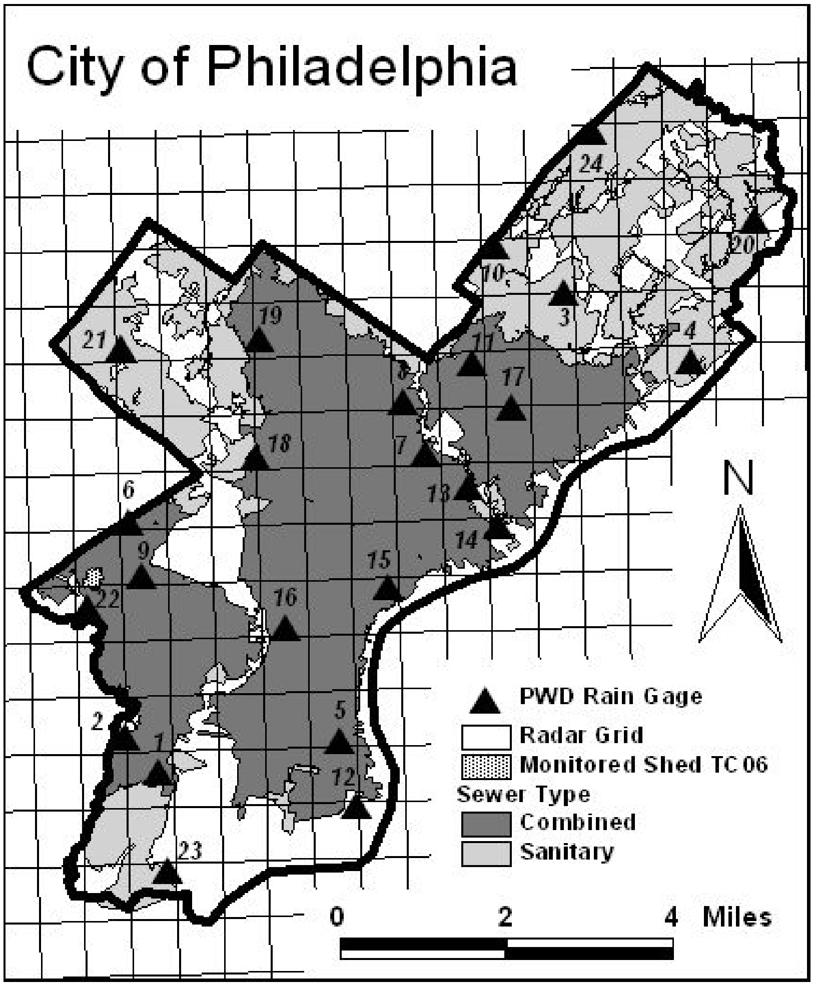

Figure v5-17 City of Philadelphia Showing the Approximate Locations of PWD Rain Gages, Radar Rainfall Grid, as Well as in City Combined and Sanitary Sewer Service Areas

v5.2.2 Homogeneity of PWD Rain Gage Station Records

This section describes the methods used to evaluate and adjust PWD rain gage records to create continuous homogeneous rainfall records at each location in the network for use as hydrologic model input.

v5.2.2.1 Data Set

The PWD maintains a database of 15-minute accumulated precipitation totals collected from its 24 tipping bucket rain gage network for the period 1990 to the present. The uncorrected, 2.5-minute accumulated, 0.01 inch tip count, rain gage data are subjected to preliminary quality assurance and quality control procedures. Identification and flagging of bad or missing data are performed for each rainfall event on a monthly basis by visual inspection of 15-minute accumulated data comparing measurements at nearby gages and looking for patterns of obvious gage failures, including plugged gages and erratic tipping. Flagged data for each gage subsequently are filled with data from the five nearest gages using inverse distance squared weighting. Neighboring gage data that are flagged are removed prior to weighting.

Daily rainfall volumes were totaled for each gage. Daily gage totals containing any filled data were removed from further analysis. A daily mean total was calculated for each gage as the average of all daily gage totals excluding the total at the gage itself. The resulting dataset consisted of daily gage totals, and daily mean totals (mean of all the other gages) for each gage.

v5.2.2.2 Double Mass Regression and Cumulative Residual Time Series Analysis

Double mass regression and cumulative residual time series analysis methods were used for evaluating the homogeneity and adjusting PWD rain gage records. These methods, like other reliable analytical methods of homogenizing rainfall time series, depend upon comparison of gage data to a homogeneous reference series. A reference series can be created using the average of a collection of nearby gages, or a homogeneous record at a single nearby gage (Allen *et al.*, 1998). The rain gage network mean value, calculated as described above, was selected as the reference series for homogenization of PWD rain gage data.

Evaluating the homogeneity of PWD rain gage records, and identifying dates when apparent changes in measurement conditions may have occurred, was performed using double mass and cumulative residual time series analysis techniques. A series of graphs was produced comparing gage to mean daily rainfall totals for each gage in the network. An example of the output generated is presented for PWD rain gage 22 in Figures v5-18 and v5-19.

A double mass plot of gage to mean cumulative daily rainfall totals was produced for each gage. The slope of the linear regression line passing through the origin is referred to as the double mass regression slope, DMRS, as shown in Figure v5-18. Potential heterogeneities are identified by visual inspection of the double mass plot as seen by systematic departures from the trend line. These departures can be identified more easily by plotting the accumulated residual from the simple linear regression of gage against mean daily rainfall totals over time (Craddock 1979) as shown in the cumulative residual plot in Figure v5-19.

Evaluation of potentially significant gage record non-homogeneity was aided by the addition of an objective graphic analytic tool to the cumulative residual plot. An ellipse was drawn on the plot to contain the residual of a homogeneous time series for a given probability of the standard normal variate (Allen *et al.*, 1998, Henriques et al 1999). The 80% probability level, commonly used by others according to Allen *et al.*, 1998, was chosen for this data evaluation program. Because the cumulative residual time series plot in Figure v5-19 is not contained within the ellipse, we reject at the 80% confidence level the hypothesis that the rainfall record at PWD rain gage 22 is homogeneous with respect to the mean. The parametric equation defining the probability ellipse is given by (Allen *et al.*, 1998)

$$x = \alpha Cos(\theta) + \alpha$$
$$y = \beta Sin(\theta)$$

with

$$\alpha = n/2$$
$$\beta = nZ_p S_{y,x} / \sqrt{(n-1)}$$
$$S_{y,x} = S_y \sqrt{(1-r^2)}$$

where:

n = the number of observations
S_y = the sample standard deviation
r = the Pearson correlation coefficient
Z_p = the standard normal variate at 80% probability
θ = an angle in radians varying from 0 to 2π

Cumulative residual analysis reveals even subtle change in gage bias often undetected by routine inspection of precipitation data and gage history records. Furthermore, subjective evaluation of the cumulative residual plot enables effective identification of the approximate dates abrupt changes in the relationship between the gage and its neighbors occur (Craddock 1979). In this manner, a set of adjustment periods are determined that contain continuous, relatively homogeneous segments of each gage record. Objective statistical methods are used by others to identify significant break points in gage record homogeneity (Peterson *et al.*, 1998). These methods have not been used here, however, like the subjective method of defining homogeneous periods used in this data homogenization program, they rely on comparative evaluation of time series data from a moderately dense and highly correlated gage network.

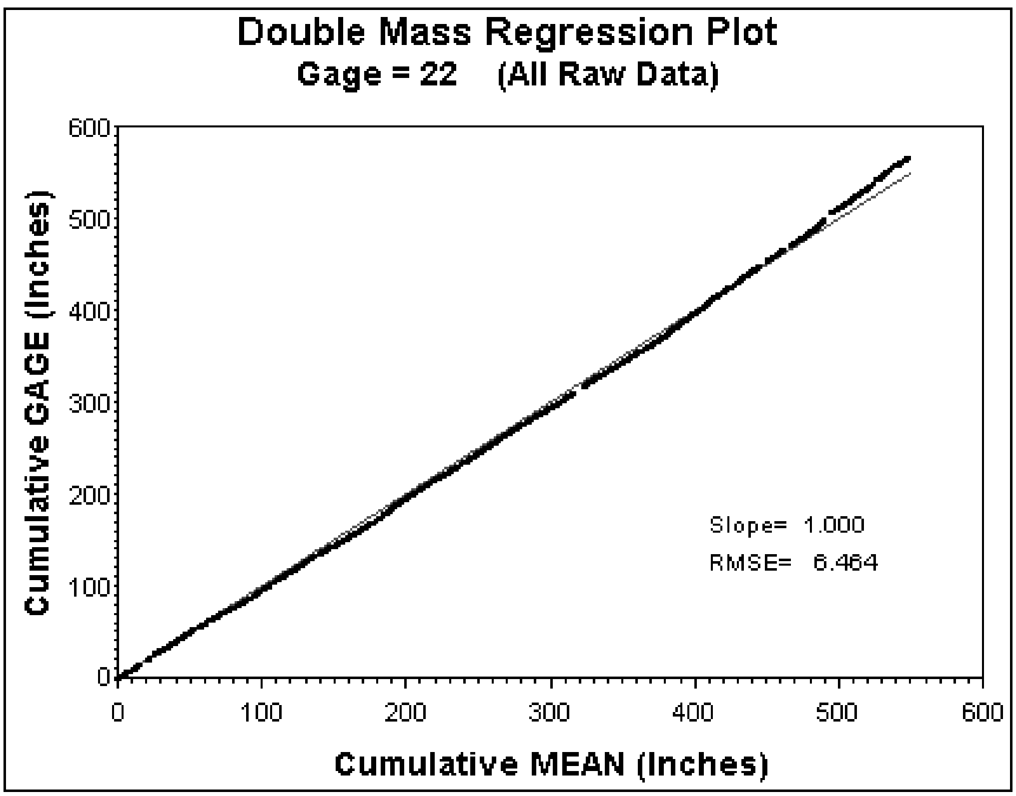

Figure v5-18 Double Mass Regression Plot of Cumulative Daily Rainfall at PWD Rain Gage 22 Against Mean Using All Raw Data for the 1990-2006 Period of Record

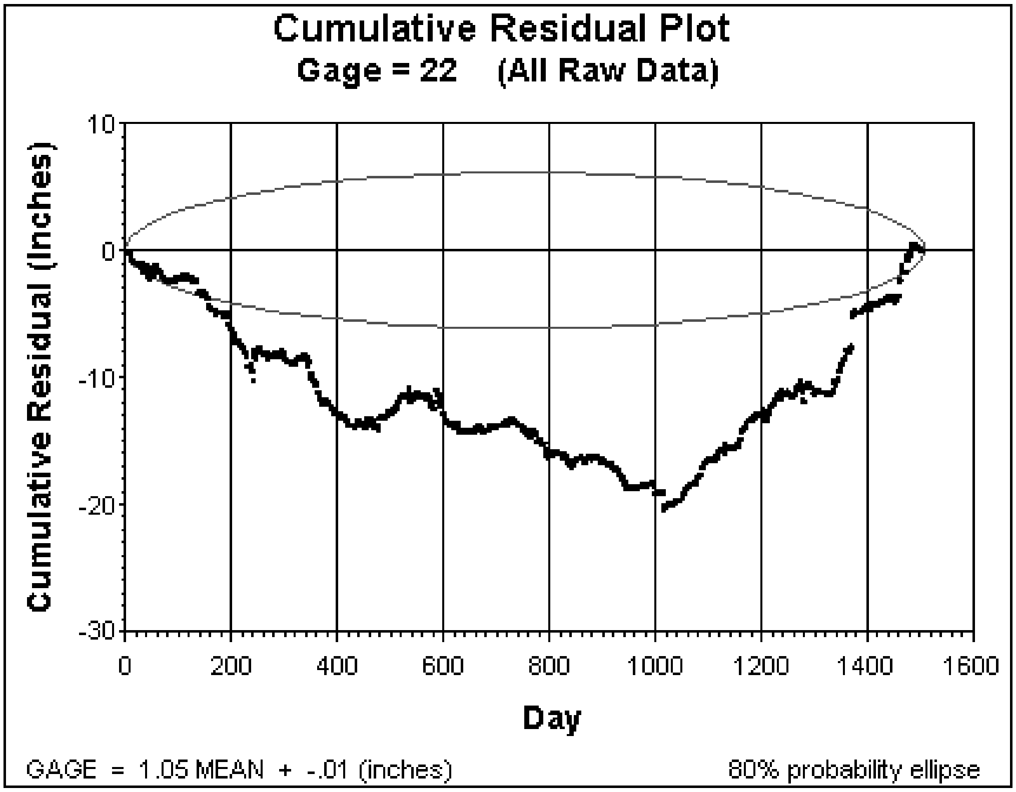

Figure v5-19 Cumulative Residual Time Series Plot from the Linear Regression of Daily Gage Against Mean Rainfall Totals at PWD Rain Gage 22 Using All Raw Data for the Period 1990-2006

v5.2.2.3 Adjusting Heterogeneous Gage Records

Once significant non-homogeneity of the gage record is determined, and the limits of homogeneous adjustment periods have been identified, adjustment factors are computed for these periods to form a homogeneous data record.

Homogeneity adjustment factors are viewed as the ratio of the average gage biases between a reference period and the period to be adjusted (Guttman 1998). Several methods of calculating adjustment factors are investigated. Each method employs a different form of expressing the average bias at the gage relative to the mean. Three methods of estimating average biases for a gage period considered are:

1. Average Daily Ratio of rain gage value to the mean (high influence of small event outliers)
2. Double mass linear regression slope (most stable with respect to outliers)
3. Linear regression slope (high influence of large event outliers)

The double mass linear regression slope with y-intercept = 0 was found to be the most stable with respect to outliers and was chosen for this study to determine homogeneity adjustment factors for selected periods of the rain gage record.

Rain gage record homogeneity adjustment factors were calculated for each interval by dividing the double mass regression slope (DMRS) of the entire unadjusted data record (Figure v5-18) by the DMRS for the adjustment period. This calculation is performed for each adjustment period identified and all raw data within the adjustment periods then is multiplied by these factors to generate the corrected rain gage record.

The results of homogeneity adjustment are presented for PWD rain gage 22 with the double mass regression plot in Figure v5-20 and the cumulative residual plot in Figure v5-21. Comparison of these plots to those presented in Figures v5-18 and v5-19 for the raw data reveal a significant improvement in homogeneity of rain gage bias relative to the network mean over the period of record for this gage.

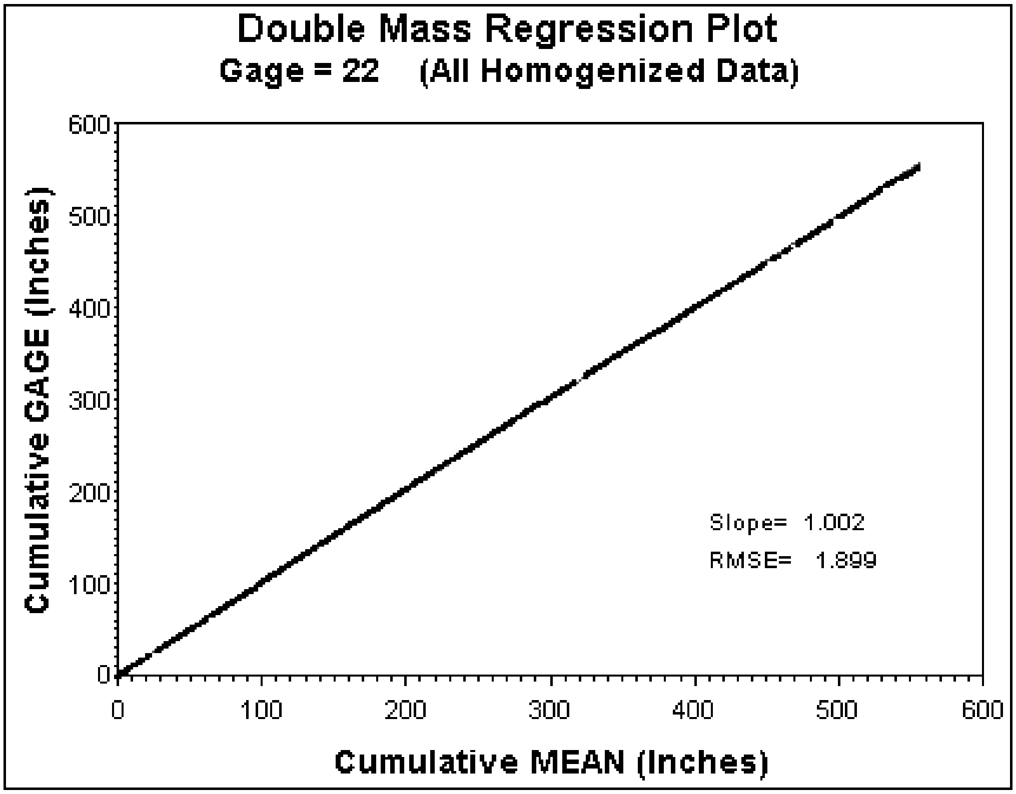

Figure v5-20 Double Mass Regression Plot of Cumulative Daily Rainfall at PWD Rain Gage 22 Against Mean Using Homogenized Data for the 1990-2006 Period of Record

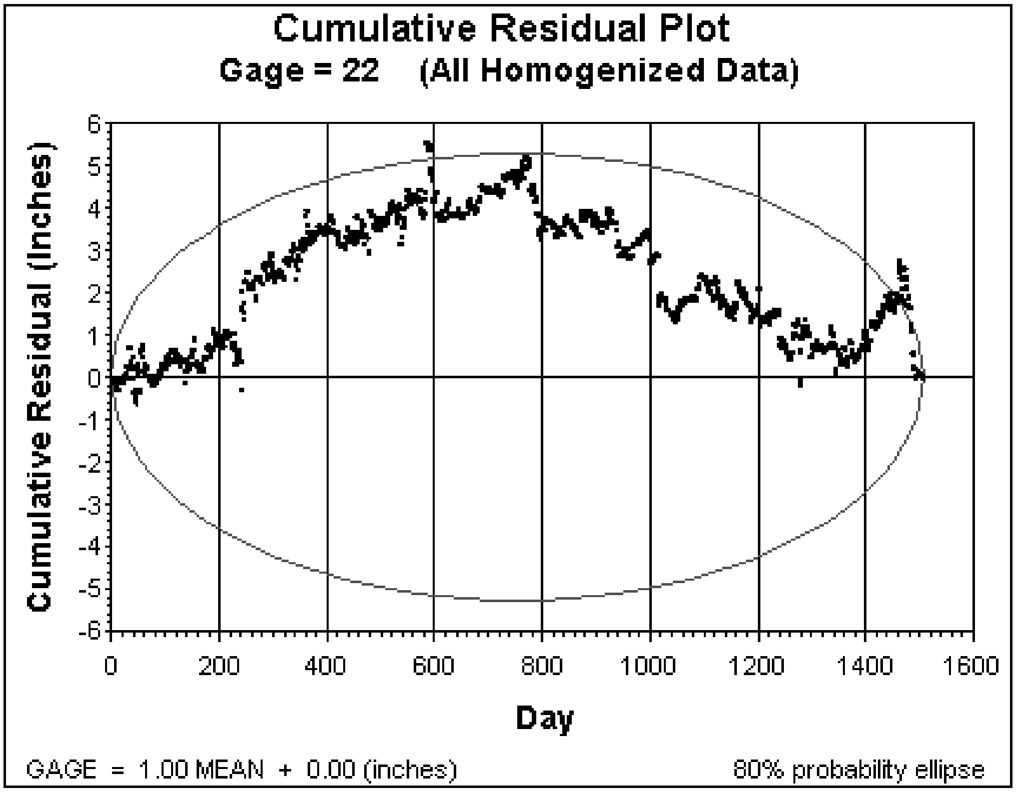

Figure v5-21 Time Series Plot of the Cumulative Residual from the Linear Regression of Daily Gage Against Mean Rainfall Totals at PWD Rain Gage 22 Using Homogenized Data for the Period 1990-2006

v5.2.3 Normalizing Spatial Biases in Rain Gage Data Using Calibrated Radar Rainfall Estimates

Once homogeneous gage records are created, all gages then are adjusted for consistent net systematic biases resulting from differences in gage equipment, gage site conditions, and previous homogeneity adjustments.

The second major goal of this investigation is to adjust all the gages in the PWD network to the same average bias, so the gage network more reliably represents spatial variation across the region. To achieve this goal a reliable reference series is needed to represent spatial variation of rainfall across region with a uniform bias, over a sufficiently long and homogeneous period of record. Calibrated radar rainfall estimates are used for this purpose.

v5.2.3.1 Data Set

Radar rainfall estimates provided by NEXRAIN Corporation are derived from a 2km x 2km National Weather Service level 3 radar mosaic product, corrected for ground clutter and other anomalies, and calibrated to the PWD 24 rain gage network using a mean field bias adjustment. The 15-minute calibrated radar rainfall estimates for a 15-month period including two relatively recent intervals: October, 1999 through August, 2000 and March, 2002 through June, 2002 are used for this analysis. The radar rainfall estimates are calibrated to the PWD rain gage network using a mean field bias adjustment method where the mean event accumulation for the radar pixels containing the

PWD rain gages is set equal to the mean event accumulation measured at the gages. In this way the total volume of rainfall reported within the network is conserved, while the spatial variation represented by radar data is retained.

v5.2.3.2 Bias Adjustment Using Double Mass Regression

Double mass regression analysis is used to correlate PWD rain gage measurements to calibrated radar rainfall data. In this way, overall spatial bias adjustment factors are determined that best represent the spatial distribution of rainfall over the full period of record for each rain gage in the network.

Before determining the overall bias of a rain gage record relative to the calibrated radar rainfall estimates, the homogeneity of the data sets is verified. The verification is performed by examining the time series plots of the cumulative residual from the linear regression of daily radar against rain gage rainfall totals, as shown for PWD rain gage 22 in Figure v5-22. Once an acceptable degree of homogeneity between the datasets is determined for the 15-month radar study period, the spatial bias adjustment factor is calculated for the complete gage record. The program developed for determining overall site bias factors at each gage is a two-part process using the same techniques developed for homogenization of the gage record.

The first step in determining the overall spatial bias adjustment factor, once the homogeneity of all datasets is established, is to determine the bias at each gage using the double mass regression slope,of radar to rain gage daily rainfall totals for the 15-month radar study period as presented in Figure v5-23 for PWD rain gage 22.

Next, the gage bias for the radar study period is related to the overall bias of the gage record to determine spatial bias adjustment factors that are applied to adjust the entire period of record for the gage, not just the 15-month radar period. This is done by determining the average gage bias for the 15-month radar period using the DMRS of the daily gage versus the mean rainfall for this period as shown by Figure v5-24. The DMRS for the entire period of record was previously determined using all homogenized data as shown by Figure v5-20. Then the ratio of the DMRS for the entire period of record (all data) to that of the radar study period (15-month) is calculated. This ratio is multiplied by the DMRS radar to rain gage bias from Figure v5-23, to yield the overall spatial bias adjustment factor for each gage.

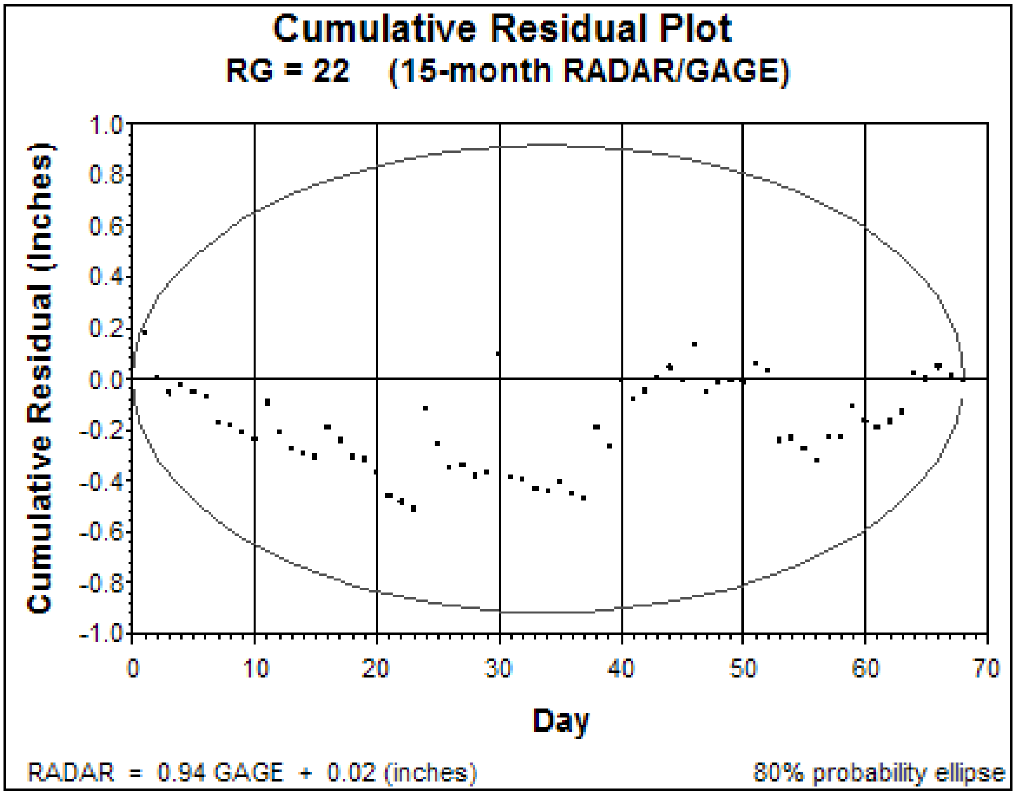

Figure v5-22 Cumulative Residuals: Linear Regression of Daily Radar Rainfall Estimates Against Rain Gage Totals at PWD Rain Gage 22 for the 15-Month Radar Study Period

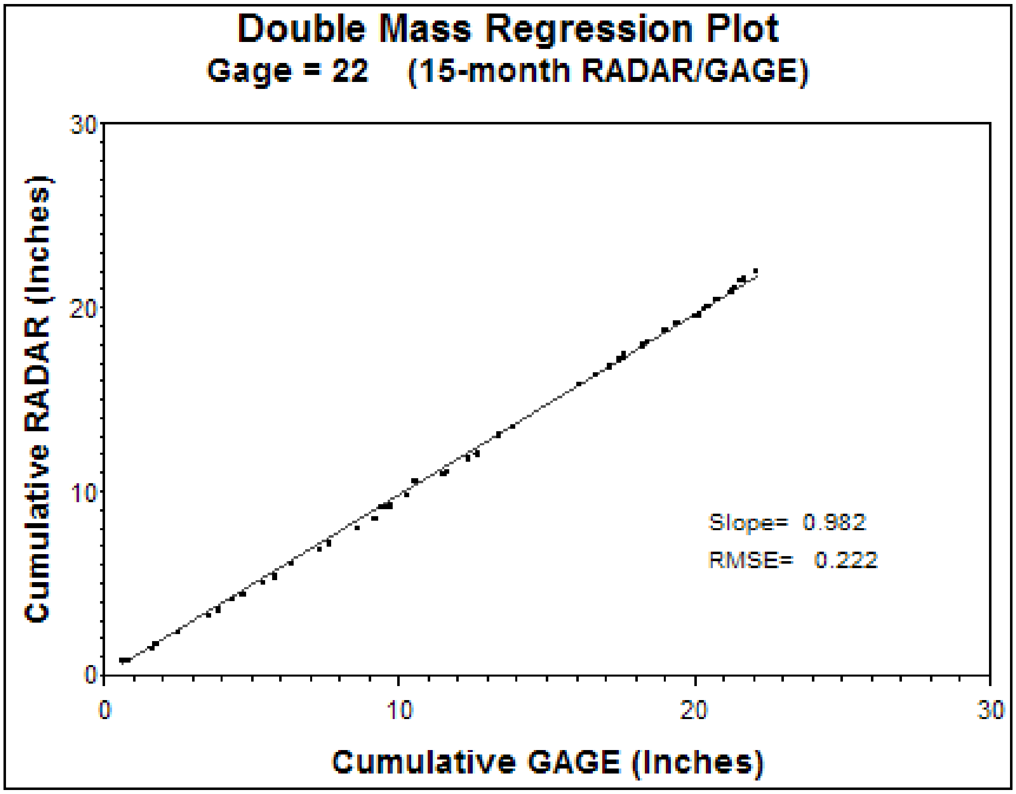

Figure v5-23 Double Mass Regression Plot of Cumulative Daily Radar Against Gage Rainfall Totals at PWD Rain Gage 22 for the 15-Month Radar Study Period

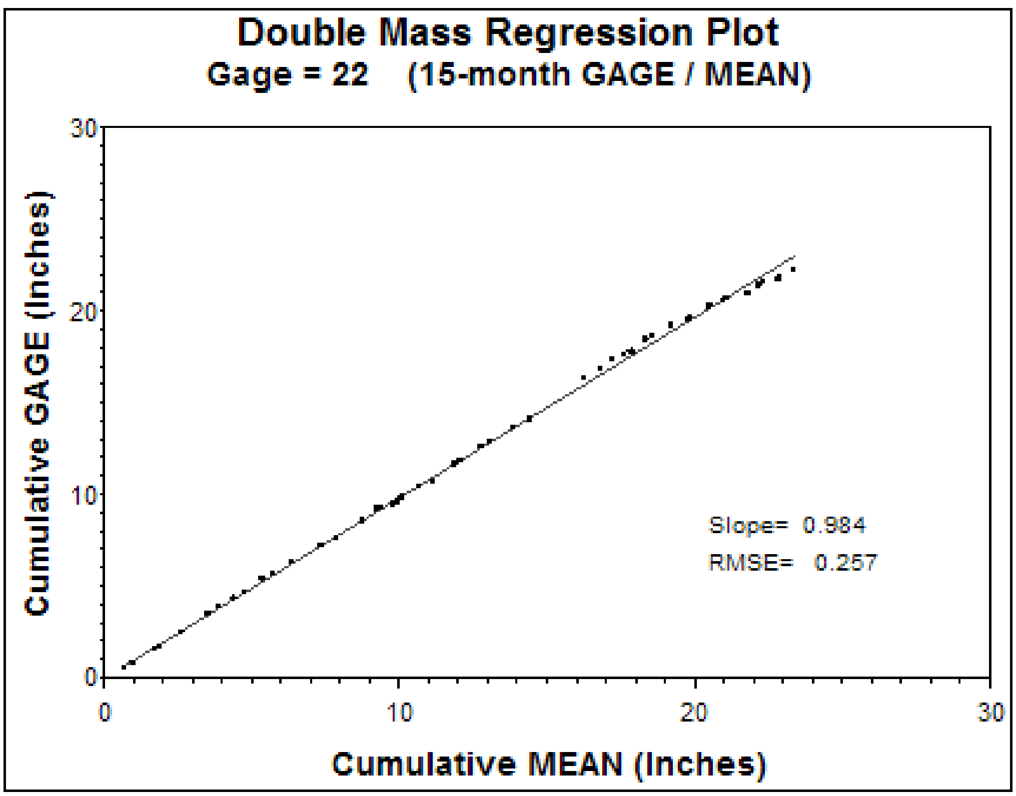

Figure v5-24 Double Mass Regression Plot of Cumulative Gage to Mean Daily Rainfall Totals at PWD Rain Gage 22 for the 15-Month Radar Study period

v5.2.3.3 Results

The overall spatial bias adjustment factors determined for each of the PWD rain gages is presented in column 8 of Table v5-10. PWD rain gage 14 is removed from the analysis because it lacks a sufficient number of daily measurements corresponding to the 15-month radar rainfall study period.

Inspection of column 3 in Table v5-10 reveals that across the network PWD rain gages 18 and 21 exhibit the greatest overall biases in raw gage data relative to the mean with overall biases of negative 8% and positive 9% respectively. These two rain gages are located among five PWD stations covering an approximately four square mile region of North West Philadelphia. This represents an average difference in rainfall between these two gages of approximately 15% over approximately a two mile distance.

The spatial distribution represented by radar rainfall data in column 10 of Table v5-10, however, reveals an average bias relative to the mean at gages 18 and 21 of positive and negative 3%, respectively. The spatial bias adjustment factors for PWD gages 18 and 21 are the greatest in the network. The final results of homogenization and spatial bias adjustment, given by the gage to mean double mass regression slopes presented in column 9 of Table v5-10, represent the long term average relative rainfall distribution across the region.

To better visualize the spatial distribution of rainfall represented by the final and intermediate results presented in Table v5-10, surface contour plots were generated using the DMRS for all rain gage locations relative to the network mean. The long term average distribution of rainfall over the Philadelphia area illustrated in Figure v5-25 is determined from the double mass linear regression slope, column 3 of Table v5-10, using all raw data for the 1990-2006 period of record.

Isometric contours lines are generated from this data using the kriging method of spatial interpolation provided by Surfer™ for Windows Notes V6 ©Golden Software Incorporated, 1993-97.

The average rainfall distribution relative to the network mean over the 15-month radar study period is presented in Figure v5-26 for raw, homogenized but not spatial bias adjusted, rain gage data. This figure reveals gage locations with significant long and short term differences in rainfall from nearby gages, as well as the regional average.

The final overall spatial bias adjustment factors for all homogenized PWD rain gage data over the 1990-2006 period of record are given in column 8 of Table v5-10 expressed as differences from the mean. These factors are graphically presented in Figure v5-27. Note that bias adjustment factors range from plus-to-minus seven percent.

The average relative distribution of rainfall observed from the radar rainfall data over the 15-month study period (Figure v5-28) compares favorably to that of the final adjusted PWD rain gage data determined over the 1990-2006 period of record (Figure v5-7).

Table v5-10 Final Spatial Bias Adjustment of PWD 24-Raingage Network Data with NEXRAIN Calibrated Radar Rainfall Data

1	2	3	4	5	6 = 5 / 4	7	8 = 6 x 7	9	10
PWD Rain Gage	Days with Radar Data	DMRS (Gage / Mean) Raw Data	DMRS (Gage / Mean) Homogenized Data		Ratio DMRS 15-month to DMRS All Data	DMRS (Radar / gage)	Overall Spatial Bias Adjustment Factor (Radar/All Data)	DMRS Final Adjusted Data	DMRS (Gage / Mean) Radar Data
			All Data	15-month					
1	136	1.00	0.99	1.00	1.01	0.96	0.97	0.99	0.98
2	136	0.98	0.99	1.00	1.01	0.97	0.98	0.98	0.97
3	120	0.96	0.97	0.98	1.01	1.02	1.02	1.02	1.01
4	46	0.98	1.00	1.00	1.00	0.98	0.98	1.00	0.96
5	123	0.99	1.00	1.03	1.03	1.01	1.04	1.07	1.06
6	151	1.03	1.03	1.03	1.00	0.95	0.94	0.99	0.98
7	132	1.04	1.04	1.04	1.00	0.92	0.92	0.97	0.98
8	75	0.97	0.98	0.96	0.98	1.03	1.01	1.01	0.98
9	122	1.03	1.03	1.07	1.04	0.89	0.92	0.97	0.98
10	106	1.02	1.02	1.01	0.99	0.95	0.94	0.98	0.99
11	37	1.01	1.01	1.06	1.05	0.91	0.95	0.98	1.00
12	41	0.91	0.92	0.93	1.01	1.03	1.04	0.97	0.98
13	13	0.96	0.96	1.19	1.24	0.88	1.09	1.07	1.04
14	NA	NA	NA	NA	NA	NA	NA	NA	NA
15	147	1.02	1.01	1.00	0.99	1.00	0.99	1.02	1.02
16	57	1.00	1.00	0.97	0.97	1.00	0.98	1.00	1.00
17	144	1.04	1.03	1.03	1.00	0.93	0.94	0.98	0.97
18	147	0.92	0.92	0.95	1.03	1.05	1.08	1.03	1.03
19	99	0.99	0.99	1.01	1.02	0.96	0.97	0.98	0.98
20	112	1.01	0.99	0.94	0.95	1.09	1.04	1.05	1.06
21	70	1.09	1.09	1.07	0.98	0.89	0.88	0.97	0.97
22	68	1.00	1.00	0.98	0.98	0.98	0.96	0.98	0.97
23	44	0.95	0.93	0.96	1.03	0.93	0.96	0.91	0.91
24	43	0.97	0.99	0.97	0.99	1.05	1.04	1.04	1.06

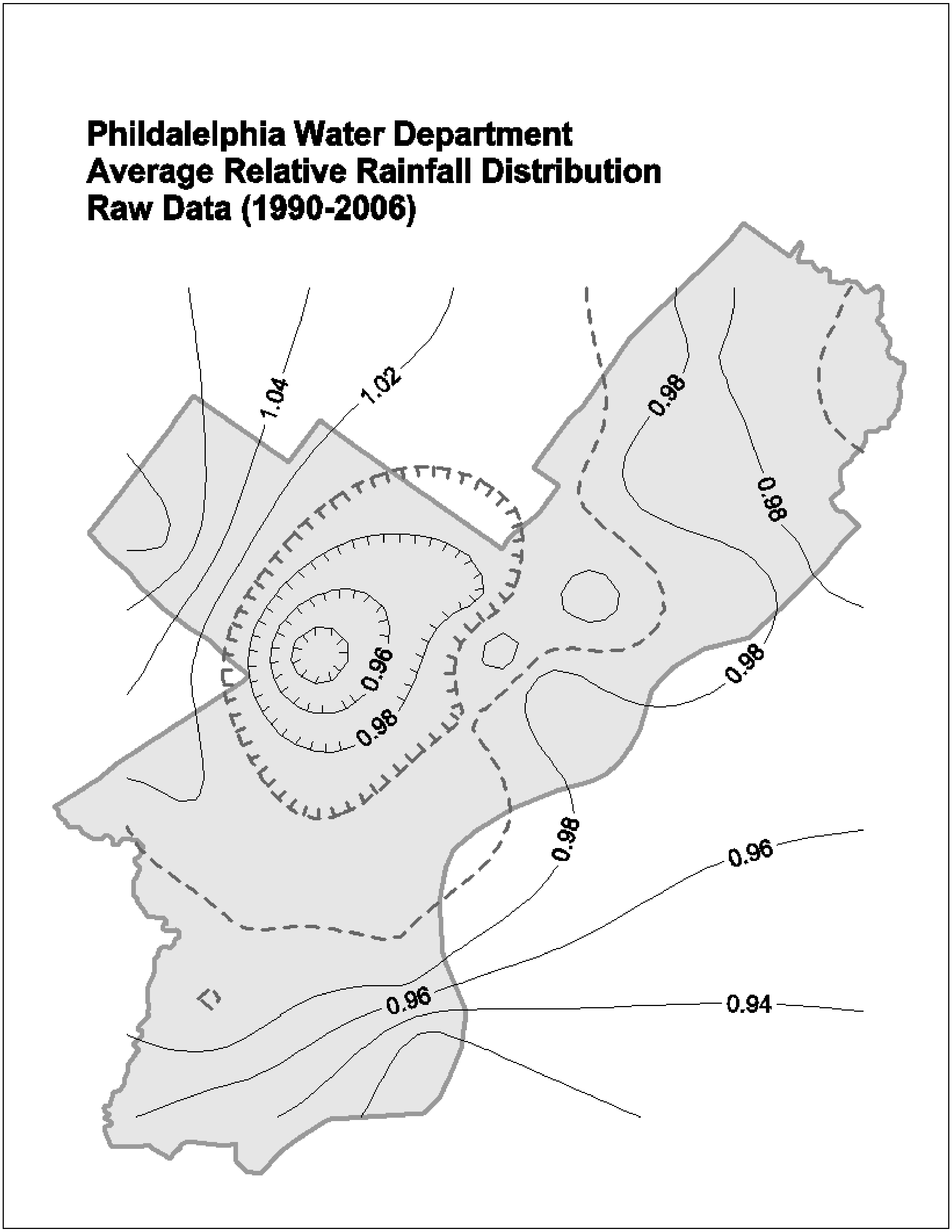

Figure v5-25 Relative Rainfall Distribution Map of the Philadelphia Area Showing Double Mass Regression Slopes of PWD Gage to Mean Daily Rainfall Totals Using Raw Data for the Period 1990-2006

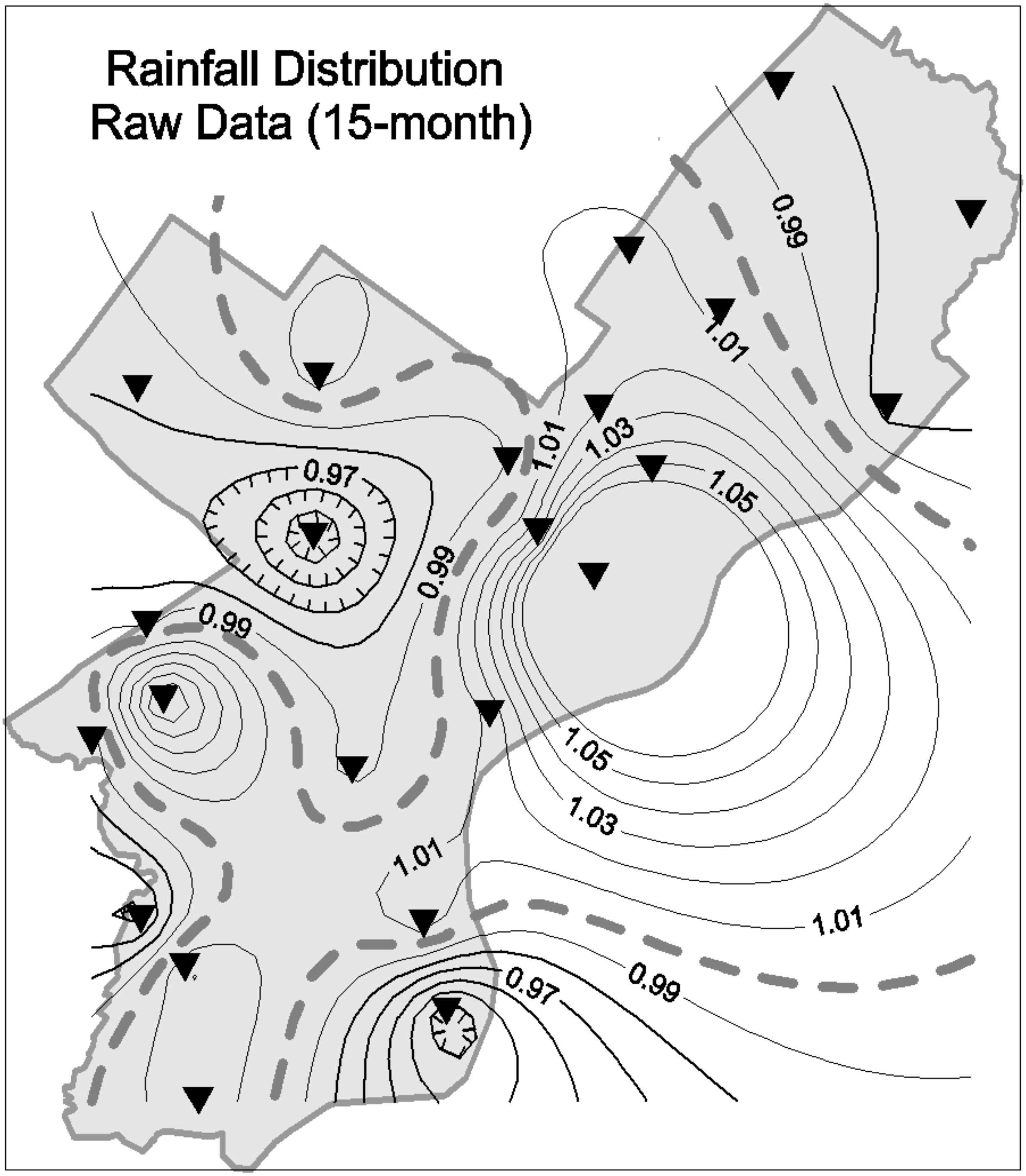

Figure v5-26 Relative Rainfall Distribution Map of the Philadelphia Area Showing Double Mass Regression Slopes of PWD Gage to Mean Daily Rainfall Totals for the 15-Month Radar Study Period

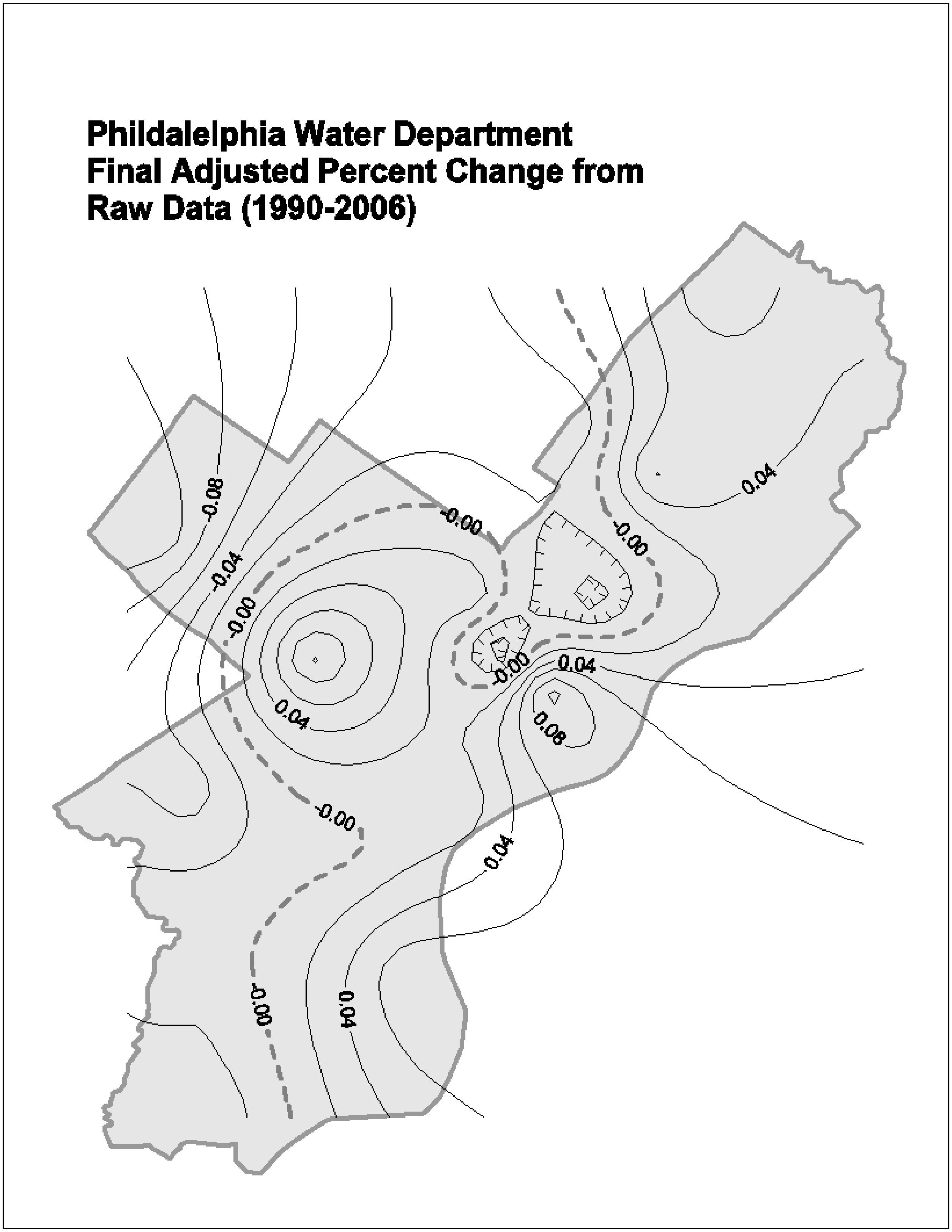

Figure v5-27 Map of Philadelphia Area Showing Final Bias Adjusted Data Percent Change from Raw Data for the Period 1990-2006

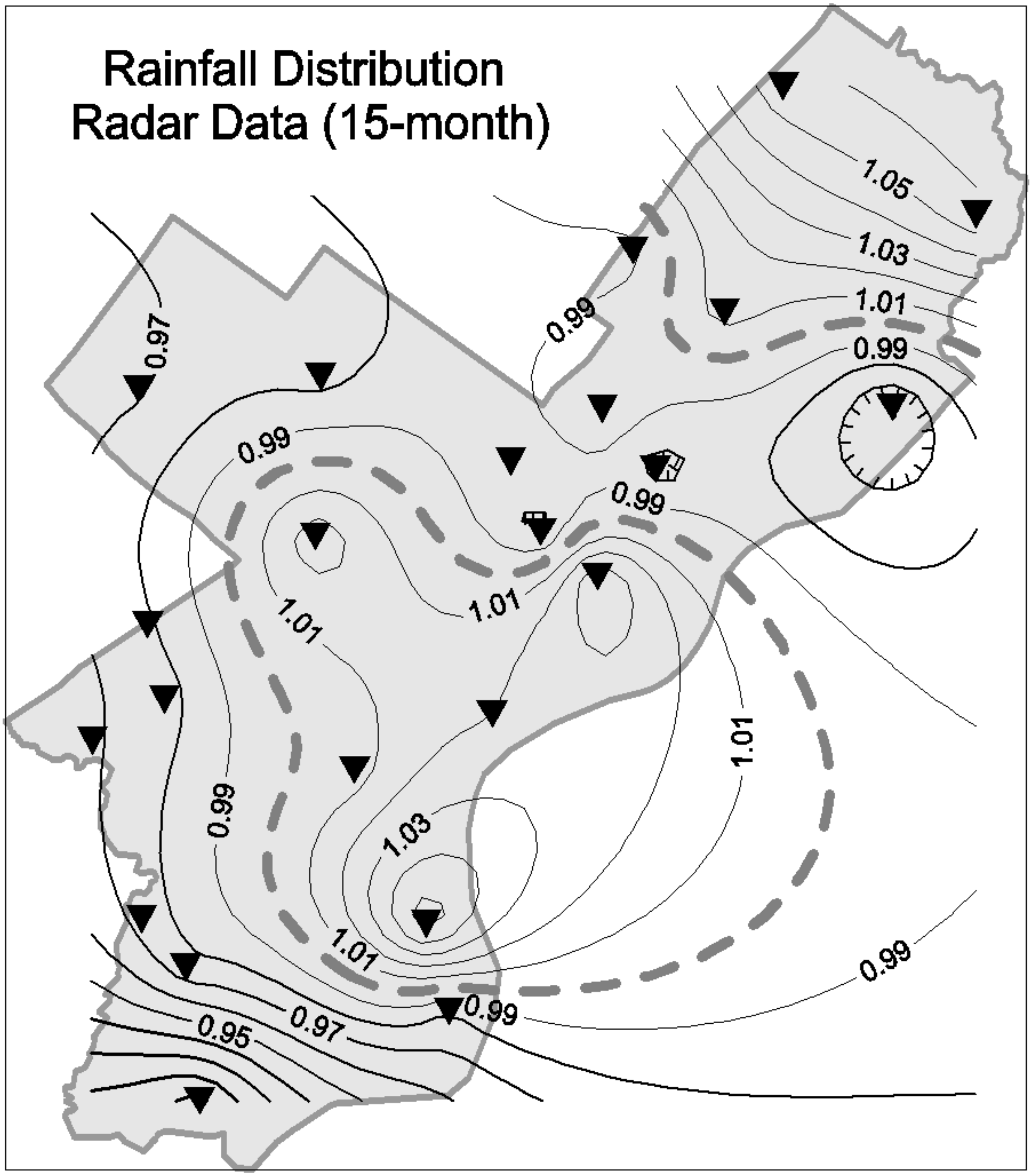

Figure v5-28 Rainfall Distribution Map of the Philadelphia Area Using Radar Data Over the 15-Month Study Period. Contours are Double Mass Regression Slopes at Gage Locations Relative to the Mean

References

Allen, Richard G. and Pereira, Luiss with contributions by Teixeira J.L., 1998: Crop evapotranspiration – Guidelines for computing crop water requirements – FAO irrigation and drainage paper 56. Food and Agriculture Organization of the United Nations, Rome, Annex 4. Statistical analysis of weather data sets.

Alexandersson, Hans, 1986: A homogeneity test applied to precipitation data. J. Climate, 6: 661-675.

Craddock, J.M. 1979: Methods of comparing annual rainfall records for climate purposes. Weather 34: 332-345.

Dai, Aiguo and Fung, Inez Y. and Del Genio, Anthony D., 1997: Surface observed global land precipitation variations during 1900-88, J. Climate, 10: 2943-2962.

Easterling, David R. and Peterson, Thomas C. and Karl Thomas R., 1995: Notes and correspondence: On the development and use of homogenized climate datasets. J. Climate, 10: 1429-1434.

Guttman, Nathaniel B., 1998: Hogeneity, Data Adjustments and Climatic Normals, National Climatic Data Center, 151 Patton Avenue, Ashville, NC 28801-5001.

Henriques, Raquel and Santos, Maria Joao, 1999: Analysis of the European annual precipitation series, Technical report to the ARIDE project No.3: Water Institute, DSRH – Av. Alimrante Gago Coutinho 30, 1049-066 Lisbon, Portugal

Peterson, T.C. and Easterling, David R. and Karl, Thomas R. and Groisman, Pavel and Nicholls, Neville and Plummer, Neil and Torok, Simon and Auer, Ingeborg and Boehm, Reinhard andGullett, Donald and Vincent, Lucie and Heino, Raino and Tuomenvirta, Heikki and Mestre, Olivier and Alexandersson, Hans and Jones, Philip and Parker, David, 1998: Homogeneity adjustments of in situ atmospheric climate data: A review. Intl. J. Climatol. 18: 1493-1517.

Vincent, Lucie A., 1998: Techniques for the identification of inhomogeneities in Canadian temperature series. J. Climate, 11: 1094-1104.

v5.3 Inverse Distance Squared Weighting and Basin Average Rainfall Calculations

Much of the uncertainty in a carefully constructed hydrologic and hydraulic model is derived from uncertainty in the rainfall record. Therefore, increasing the level of detail of the rainfall input, both spatially and temporally, increases the accuracy and precision of the model results. Careful attention to rainfall collection and analysis is critical to the modeling effort.

The Runoff module of US EPA's Storm Water Management Model (SWMM) is used to simulate the hydrology of the separate and combined sewersheds in the service area. The service area is subdivided into a number of smaller sewersheds that each drain to a particular point in the collection system. Input data, including imperviousness, slope, and precipitation data, are entered for each sewershed. A rainfall value is required for each sewershed at each date and time for which a simulation will be run. These values must be derived from some combination of rain gage rainfall data and a method of estimating rainfall at points where no gages exist.

Bias adjusted 15-minute accumulated rainfall data for the PWD 24-raingage network are used for the weighting analysis. No filling of missing data is performed prior to the inverse-distance-square weighting. Bias adjustments are performed on the data as documented in Section v5.2.

There are a number of methods of estimating rainfall in areas between rainfall gages, including Thiesson polygons and inverse distance-squared weighting. An inverse distance-squared weighting procedure is chosen as described below.

A one-square-kilometer grid is imposed over the PWD service area, and the results of the weighting calculations are applied to this grid. Thus, each of the cells on the grid is assigned a rainfall value at each date and time. These grid values are later used to provide area-weighted average 15-minute rainfall values for each individual sewershed.

Manipulation of the rainfall data is performed by Statistical Analysis System (SAS) code. SAS is a high-level programming language that is particularly well suited to processing large amounts of data with relatively simple programming code. The algorithm includes five steps that apply to each date and time. These steps are listed and discussed in further detail below.

1. Read in gage rainfall data, areas, and coordinates
2. Populate the grid center points with rainfall values
3. Area-weight cell data to create a rainfall value for each sewershed
4. Output the results

Step 1: Read in raw data, areas, and coordinates.
Data input to the program include the following:

- Raw (bias-adjusted) rainfall data
- Results of GIS intersect between grid cells and sewersheds
- Rain gages assigned to each grid cell by earlier GIS analysis
- State plane coordinates of rain gage locations and grid cell centroids

Rain gages are assigned using the following logic:

- For grid cells other than those close to the edge of the service area, the three closest gages surrounding the cell are identified, forming a triangle that contains the grid cell centroid
- For some grid cells close to the edge of the service area, only two gages are assigned. For example, cell centroids contained in triangular polygon "010203" in Figure v5-29 are assigned gages 1, 2, and 3; cell centroids in irregular polygon "0102" are assigned gages 1 and 2
- Each assigned rain gage is assigned a backup gage to be used in cases where measurements at the primary gage do not pass quality assurance. In this case, data from the backup gage replace data from the primary gage. However, since the backup gage is further from the cell centroid, its data ultimately get a lower weight than data from the primary gage would have

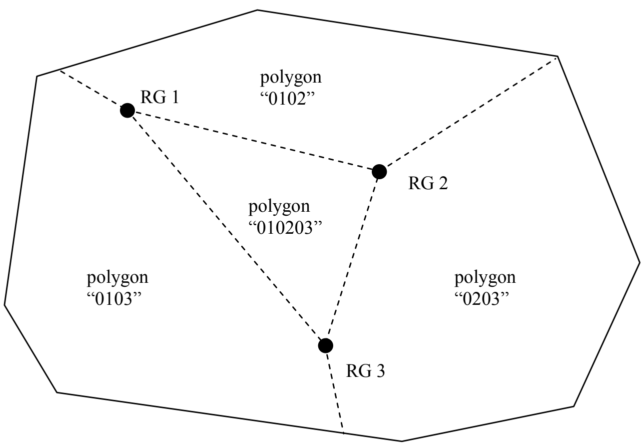

Figure v5-29 Schematic Diagram of Gauge Assignment Process

Step 2: Populate the grid center points with rainfall values.
After reading in the necessary data, the program uses it to populate the radar grid with rainfall values. Each point on the grid is assigned a value by inverse distance-squared weighting of 2 to 3 nearby rain gages. This process is depicted in Figure v5-30.

Rainfall for cell *j*, at a particular date and time, is given by the following equation:

$$P_j = \sum_{i=1}^{n} (f_{i,j} P_i)$$

Where P_j is the calculated precipitation at cell j,
j is the cell number,
i is the rain gage number,
n is the number of rain gages assigned to the cell (3 in the example),
$f_{i,j}$ is the rainfall weighting factor give by

$$f_{i,j} = \frac{(D_{i,j})^{-2}}{\sum_{i=1}^{n} (D_{i,j})^{-2}}$$

$D_{i,j}$ is the distance between gage i and cell j (by the Pythagorean Theorem), and P_i is the measured precipitation at rain gage i

Rules for missing data are as follows:

- A careful distinction is made between zero values and missing data due to quality control. Zero values are treated as zeros in the mathematical equations
- When a value is flagged as missing due to quality assurance, a value from a backup gage is substituted. If the backup gage data is also flagged, data assigned to the grid cell are based on data from the remaining 1-2 assigned gages, with backup gage values substituted as necessary. There were no instances in which none of the primary or secondary gages assigned had quality-assured data

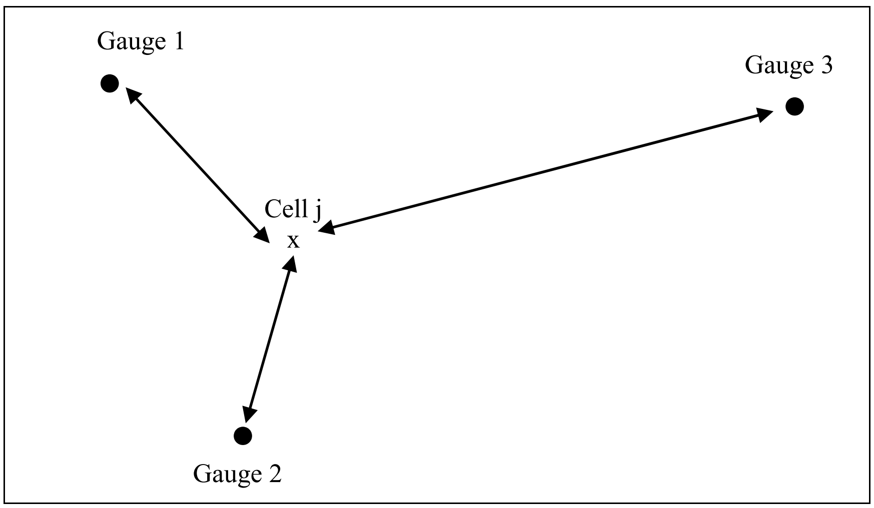

Figure v5-30 Conceptual Diagram Showing Gage Cell Assignments

Step 3: Area-weight cell data to create a rainfall value for each sewershed.
The final step in the calculations is the area-weighting of cell data to derive rainfall values for all sewersheds at all dates and times. This process is described by Figure v5-31.

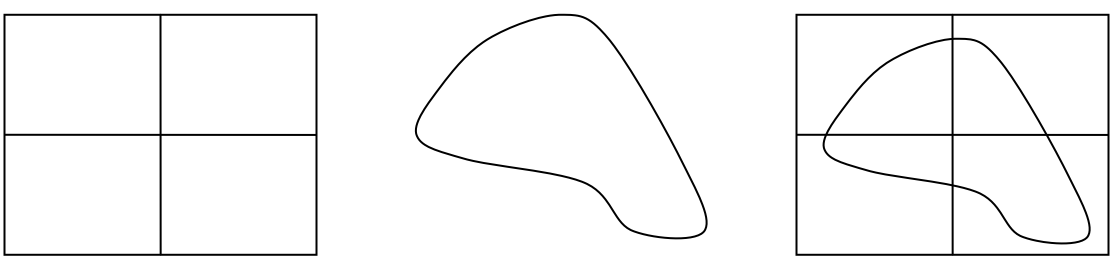

Figure v5-31 Conceptual Diagram of the Cell-Sewershed Relationship

At a given date and time, the precipitation for a sewershed is given by the following equation:

$$P_k = \sum_{j=1}^{n} \left(\frac{A_{j,k} P_j}{\sum_{j=1}^{n} A_{j,k}} \right)$$

Where P_k is the calculated rainfall for sewershed k,
j is the cell number,
n is the number of cells contained all or partially within sewershed k,
$A_{j,k}$ is the area of cell j within sewershed k, and
P_j is the calculated rainfall at cell j.

Step 4: Output the results.
The results of the calculations, consisting of rainfall at every sewershed, date, and time, are output in format that can readily be read by a SWMM model.

Quality assurance consists of verification of input parameters and verification of output with spreadsheet calculations. Scenarios for verification include some chosen at random and some chosen as special cases. Special cases include situations in which only one to two assigned gages provide usable data at a particular sewershed and time. Errors found through this process have been corrected, and it has been repeated until the project team has a high level of confidence that the results are accurate.