Uncertainty in PMP and Recurrence Interval-Based Design Storm Estimates from Sparse Data

GENE BOSLEY, P.E.; PETE KOWALEWSKI, P.E.; AND MARYBETH DOWELL

PREPARED FOR:

PREPARED BY:

TIERRA GROUP INTERNATIONAL, LTD.

111 E. Broadway, Suite 220
Salt Lake City, UT 84111
P: 801.210.9600

February 26, 2014

This page left intentionally blank.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

LIST OF ACRONYMS

AEP	Annual Exceedance Probability
AMS	Annual Maximum Series
ARI	Average Recurrence Interval
CapEx	Capital Expenditures
CDA	Canadian Dam Association
GEV	Generalized Extreme Value
Gumbel EV	Gumbel Extreme Value
HMR	Hydrometeorological Report
LP3	Log-Pearson Type III
NCDC	National Climatic Data Center
NOAA	National Oceanic and Atmospheric Administration
NWS	National Weather Service
PDS	Partial Duration Series
PMP	Probable Maximum Precipitation
T	Time, Return Period (i.e. T-year precipitation depth), or Recurrence Interval

TSF	Tailings Storage Facility
USDA	U. S. Department of Agriculture
USGS	United States Geological Survey
WRSF	Waste Rock Storage Facility
WMO	World Meteorological Organization
WRCC	Western Regional Climate Center

LIST OF UNITS

amsl	above mean sea level
in	inch
km	kilometer
km^2	square kilometers
m	meter(s)
mi^2	square miles
mm	millimeter
%	percent

ABSTRACT

Dam safety and pollution control considerations require large, infrequent precipitation events (50-year up to Probable Maximum Precipitation (PMP)) to be determined for use in the design of tailings storage facilities (TSF), waste rock storage facilities (WRSF), heap leach facilities, and other process-related infrastructure. New, remote sites in particular lack data with a significant period of record to allow these relatively infrequent events to be determined. Designers often use regional data, attempt to correlate it to the site over short concurrent records, and correct for orographic factors. In some cases, designers adapt U.S. methods or use statistical methods such as Hershfield's (World Meteorological Organization (WMO), 1986) to determine the PMP. Bias, error, and uncertainty associated with such evaluations are seldom estimated or reported.

Using tropical, desert, and mountain site data, common methods for estimating the PMP and other rare events from limited data were investigated. Sites were selected based on the availability of long, concurrent records of local and regional data or relatively dense gage networks; enabling a comparison of results generated from nearby datasets and sampling shorter periods of record, emulating the results that would be obtained had the total record been shorter or obtained outside the project area. Statistically-derived PMP estimates for U.S. sites are compared to estimates obtained using standard published U.S. National Weather Service (NWS) methods.

The results obtained in this study demonstrate the need to reevaluate design storm estimates as data becomes available, not merely at the beginning of a project. Given the sensitivity of facility cost to freeboard and other runoff-related design criteria, revising design storm estimates during operations or approaching closure may achieve significant reductions in capital cost and/or environmental risk.

1.0 INTRODUCTION

Dam safety and pollution control considerations require that large, infrequent precipitation events (50-year average recurrence interval up to PMP) be used in mine waste containment facility design (Canadian Dam Association (CDA), 2007). New, remote sites in particular suffer from limited data with which to estimate such relatively rare events, potentially leading to bias, error, and uncertainty in design rainfall depths. Furthermore, many mines establish design storm estimates early in the project development or permitting process, and never revisit them after the initial mine design has been completed, even though additional data may become available later given the typical duration of the project development process and the typical operating life of a mine. Insufficiently conservative initial estimates of design rainfall can result in under-design of facilities and present environmental and safety risks; conversely, overly conservative designs may increase capital expenditure (CapEx) and misallocate resources.

In this paper, using tropical, desert, and mountain site data, we identify the uncertainty and potential shortcomings of common approaches to determining design storms from limited data, identify their appropriate limits of applicability, and discuss situations where reevaluation of design storms can potentially reduce risk or decrease cost. Design storm estimates are shown to vary according to the period of record, suggesting a need to revise the design storm estimates as sufficient new data becomes available. Revisiting design storm estimates during operations or as the site approaches closure may support a more accurate estimate of design storms to be used in the design of new facilities, expansion of existing facilities, or closure of facilities, reducing CapEx and environmental risk.

2.0 BACKGROUND – COMMON APPROACHES TO DESIGN STORM ESTIMATION

Available rainfall data typically includes regional data (sometimes far-removed from the site) and local data with a relatively short period of record. When regional datasets are available, designers attempt to correlate it to site data over short concurrent records, and correct for orographic factors, thus developing a synthetic data series for the site. Once local data is obtained or a synthetic dataset developed, site-specific estimates of recurrence interval-based design storms can then be made from statistical methods, and the PMP can be obtained using statistical methods or adaptations of standard U.S. NWS methods.

Whether regional data, local measurements, or synthetic time series are used, there are a few common approaches taken for determining design storm precipitation:

- Recurrence-interval (also called annual exceedance probability) design storms are determined by fitting a probability distribution (Log Pearson, Gumbel Extreme Value, etc.) to the data, and using the fitted distribution to estimate design storm depths at selected average recurrence intervals;
- Within the U.S. and its territories, the U.S. NWS's Hydrometeorological Report (HMR) series provides a method for determining the PMP from meterological principles; and
- Outside the U.S., designers most often resort to statistical methods, such as Hershfield's (WMO, 1986), for determining the PMP.

For the benefit of those unfamiliar with the methods, each of them is summarized below.

2.1 Probability Distributions

At its simplest, fitting a probability distribution to data requires only the data series (annual maximum or partial duration), textbook equations describing the distribution to be used, and a 'goodness-of-fit' assessment method – visual and/or mathematical. Typical distributions used in hydrology include Log-Pearson Type III (LP3), Gumbel Extreme Value (Gumbel EV), and Generalized Extreme Value (GEV). Gumbel EV and GEV typically provide a good fit to precipitation data, and National Oceanic and Atmospheric Administration (NOAA) selected GEV for analysis of Southwest U.S. data in Atlas 14 (Bonnin et al, 2006). LP3 is commonly applied to maximum streamflow statistics. Once a distribution is selected, a frequency factor (K) can be computed for each desired recurrence interval, and the design precipitation determined as:

$$P_T = X + K * S$$

Where:

$$P_T = Design\ storm\ precipitation, mm\ (in);$$

$$T = Average\ recurrence\ interval\ (ARI), years;$$

$$X = Mean\ maximum\ precipitation\ from\ the\ series, mm\ (in);$$

$$S = Standard\ deviation\ of\ the\ series, mm\ (in); and$$

$$K = Frequency\ factor\ (a\ function\ of\ the\ distribution\ and\ ARI).$$

The equation above is the general relationship used for most analyses; however, depending on the method used and the distribution selected, adjustments for skew and outliers may be applied at appropriate points in the analysis. A well-known example of such adjustments is found in the United States Geological Survey (USGS) Bulletin 17B (USGS, 1982) methodology used for peak streamflow analysis, where outliers may be identified and eliminated, and results corrected for station skew and regional (map) skew.

2.2 PMP from HMR Series (U.S. Sites)

The U.S. NWS HMR series of reports defines the PMP as "…the theoretically greatest depth of precipitation for a given duration that is physically possible over a particular drainage basin at a particular time of year." (American Meteorological Society, 1959, quoted in Hansen, et al 1984). To that end, each report examines the maximum recorded precipitation events within a given region, and seeks to maximize their rainfall potential up to the theoretical maximum. In the Southwest U.S., two PMP's are estimated (Hansen, et al 1984): the Local Storm, representative of short-duration, isolated thunderstorms, and the General Storm, representative of regional systems (typically decaying tropical storms). Estimation methods for both PMP types were developed from storms of record occurring within and near the region, obtained from sites with a record length of 20 years to over 100 years.

The Local Storm PMP (durations from 15 minutes up to 6 hours) is obtained from maps provided in the HMR report, showing the value for a 1-hour rainfall duration over a 1 square mile area. Mapped values of the Local PMP were developed by moisture maximization of storms of record. Adjustment factors are applied to correct for duration and watershed size.

The General Storm PMP (durations from 6 hours to 72 hours) consists of two components:

- The Convergence component is that rainfall due solely to atmospheric processes. It is obtained by moisture-maximization of storms of record, adjusted for barriers ("rain shadow") and elevations; and
- The Orographic component of PMP is that due to moist air forced upward by mountain slopes.

Each General Storm component is adjusted for watershed size and duration. The General Storm PMP is the sum of the Convergence and Orographic components. In practice, the HMR series provides maps and tables from which the user can obtain each component and correction factor for each month of the year. Computing the PMP is a matter of addition and taking ratios. Note that the most intense portion of a General Storm may in fact be a thunderstorm embedded within a regional rainfall pattern; however, among all possible thunderstorms the most intense, isolated storms will fall into the Local Storm category.

The HMR series has been widely criticized in the U.S. dam safety community, for a myriad of reasons. Key criticisms include:

- The series is outdated; the **reprint** of HMR-49 was issued 30 years ago. New data, methods, and computational capacity have become available since the series was developed;
- It is generally believed to provide conservative, hence safe but costly, estimates of PMP;
- As no recurrence interval is associated with the PMP, and the correspondence between PMP and low recurrence interval storms varies significantly and inconsistently by region, quantifying risk is impossible. This is especially an issue for facilities with short lifespans but high hazard classifications – say, a TSF; and
- Study areas, while logically delimited based on physiography and meteorology, are large, and local effects are likely to have been ignored as a simple consequence of scale. For example, HMR-49 covers the Southwest U.S. from the crest of the Sierras to the Continental Divide (crest of the Rockies), and from the southern extent of Columbia River drainage south to the Mexican border.

2.3 PMP from Statistical Methods (Hershfield)

Hershfield's method (World Meteorological Organization (WMO), 1986) enables estimation of the PMP using only the rainfall record from a site, without regard to the site's location and meteorology. It is thus an attractive option on international projects where standardized methods such as the HMR series have not been established, and where data may be limited. At its heart the method is similar to other statistical methods in that it estimates the PMP from an equation of a similar form ($P = X + K_m*S$) to that introduced above. Each component of the equation is subject to correction factors that address the characteristics of the time series. Without detailing the method, below is a summary of the corrections made to the data and moments of the series:

- The sample mean and standard deviation are computed for both the full series and the trimmed series excluding the highest value, which is assumed to be an outlier. No formal outlier test is employed;
- The coefficient K_m (equivalent to a frequency factor) is obtained from a graph, as a function of the mean annual maximum precipitation;
- The mean and standard deviation are each adjusted by factors that depend on the record length (10 to 50 years) and the ratio of the respective moments computed with and without excluding the maximum value in the series;
- The mean and standard deviation are each adjusted for the record length. Shorter records yield greater values for the adjustment factors;
- The overall estimate is adjusted for the number of observational units. For a 24-hour storm, the adjustment factor is 100.6% for hourly data (24 observations), and 113% for daily readings (a common practice particularly with older/historical records at remote sites); and

- Finally, the overall estimate is adjusted for the watershed area, if it exceeds 25 km^2 (10 mi^2).

From inspection of the method, sensitivity to outliers would be expected, and the validity of the correction factors across regions and storm types (cyclonic, convective, etc.) and climate regions is unknown.

3.0 TECHNICAL APPROACH

Using tropical, desert, and mountain site data, common methods for estimating the PMP and other rare events from limited data were investigated. Sites were selected for long, concurrent records of local and regional data or relatively dense gage networks; enabling a comparison of results generated from nearby datasets and sampling shorter periods of record, emulating the results that would be obtained had the total record been shorter or obtained outside the project area. Statistically-derived design storm estimates for U.S. sites are compared to estimates obtained using HMR-49 and NOAA Atlas 14. For all sites, the effects of record length are investigated by sampling shorter periods from long-term records, and comparing to the design storm estimated from the full record.

3.1 Selected Sites

Three sites were selected for the overall analysis, with the level of analysis at a given site determined by the available data. The sites are:

- Site 1 (Desert) - Walnut Gulch Experimental Watershed (U.S. Department of Agriculture (USDA), Agricultural Research Service, 2013), near Tombstone, Arizona;
- Site 2 (High desert) - Carlin Newmont Mine (Western Regional Climate Center (WRCC), 2013a), near Carlin, Nevada;
- Site 3 (Mountainous) - Brawley Peaks (WRCC, 2013b), near Brawley, Nevada; and
- Site 4 (Tropical) - Mine site in Central America (confidential client).

Multiple rain gages were available at Site 1, with long records in excess of 50 years. Nine rain gages well-distributed throughout the Walnut Gulch watershed, with long, continuous records were selected for analysis. Shorter, single records were available at the other three sites. Table 3.1 summarizes the site characteristics and available data.

TABLE 3.1: SITE SUMMARY

Site Name	Location	Physiography	Period of Record	Average Annual Precipitation		Storm of Record	
			Years	mm	in	mm	in
Walnut Gulch – 04	Arizona, USA	Desert	60	287.3	11.31	68.3	2.69
Walnut Gulch – 13	Arizona, USA	Desert	60	298.8	11.76	65.3	2.58
Walnut Gulch – 42	Arizona, USA	Desert	59	303.2	11.94	59.9	2.36
Walnut Gulch – 44	Arizona, USA	Desert	59	296.0	11.66	73.9	2.91
Walnut Gulch – 46	Arizona, USA	Desert	52	321.4	12.65	85.2	3.36

Walnut Gulch – 60	Arizona, USA	Desert	59	320.0	12.60	73.2	2.88
Walnut Gulch – 68	Arizona, USA	Desert	59	312.0	12.28	80.3	3.16
Walnut Gulch – 80	Arizona, USA	Desert	51	309.2	12.17	66.6	2.62
Walnut Gulch – 81	Arizona, USA	Desert	52	322.6	12.70	58. 7	2.31
Carlin Newmont Mine	Nevada, USA	High Desert	35	307.1	12.09	71.1	2.8
Brawley Peaks[1]	Nevada, USA	Mountainous	27	88.3	3.48	61.2	2.41
Central America	Central America	Tropical	42	2342.8	92.24	198.9	7.83

Note: [1]The Brawley Peaks average annual precipitation does not include snowfall. The average annual precipitation, including snowfall, at the nearest station (Bodie, California) is 12.75 inches.

3.2 Points of Comparison

Consideration of the various design storm calculation methods suggests several obvious sources of potential bias and uncertainty. While other uncertainties no doubt exist, this present study sought to compare effects of the following practices on estimates of PMP and T-year design storms:

- Use of annual maximum series (AMS) versus partial duration series (PDS);
- Use of daily data (recorded once daily at a fixed time) versus 24-hour data (where the maximum 24-hour period, rather than calendar day, is considered the peak value), including evaluation of Hershfield's correction factor for observation interval;
- Single-point statistical estimates of T-year storms (Gumbel EV) versus comprehensive regional studies (NOAA Atlas 14);
- Single-point statistical estimates of PMP (Hershfield) versus regional meteorological-based methods (HMR-49); and
- Effect of record length, estimated by analysis of partial records sampled from longer datasets.

3.3 Calculation Methods

For the baseline analysis at each site, the 24-hour storm depth was estimated for both recurrence interval-based and PMP design storms, based on the AMS for the full period of record. Hershfield's method was used to estimate the PMP depth, and the Gumbel EV method was used to estimate frequency storm depths. Calculations were performed using an Excel spreadsheet, with lookup tables and regression equations used in place of Hershfield's graphical methods. The base case estimate using the AMS and full record was taken as the "true" value, from which ratios were computed to enable comparisons.

To support comparisons, the analysis was repeated using the PDS at each site. For the nine Walnut Gulch sites, the AMS analysis was repeated using both 24-hour and daily datasets. NOAA Atlas 14 estimates were obtained online, and HMR-49 PMP estimates were developed for the U.S. sites. Several Walnut Gulch sites were situated close enough together as to be identical at the scale of mapping used in HMR-49; hence, aside from elevation corrections only three independent HMR-49 PMP estimates were made (one each at Carlin, Brawley Peaks, and Walnut Gulch).

To investigate the sensitivity of design storm estimates to the period of record, the available record was segmented in two ways, and each segment was analyzed for PMP and T-year storm depths:

- The most recent 10, 20, 30, 40, and 50 years (when available) of each record was analyzed as if it were the entire record, and the results compared to the value obtained using the full record, and
- Sub-samples of 10, 20, and 30 years in duration were taken at each possible start point within a given record. For example, a 35-year record would support 35 sub-samples, each starting on a different year. The sub-sample results are compared to the baseline value obtained from the full record.

4.0 RESULTS AND DISCUSSION

4.1 AMS versus PDS

While Hershfield's method requires the use of the AMS, recurrence-interval storms may be estimated using either the AMS or the PDS. Both results are published in NOAA Atlas 14, but generally differ by little. For the U.S. sites included in this study, NOAA Atlas 14's PDS results for recurrence interval storms less than 100-year were from 0 to 0.5 mm greater than those using the AMS, while 100-year through 1,000-year results were identical. This was not the case for the Gumbel EV analysis conducted for this study. Use of the PDS generally reduced design storm estimates, sometimes significantly.

Table 4.1 lists the ratios of design storm estimates using the PDS versus the AMS, with daily data. Use of the PDS generally reduced the computed design storm, by a modest 4% to 8% at the 50-year recurrence interval, trending upward for a reduction of up to 14% for the 1000-year recurrence interval. The effect on the PMP was even more pronounced, at up to 28% reduction. The PMP is included for comparison; however, it is improper to use the PDS for computing the PMP by Hershfield's method. One exception to the tendency to decrease the computed design storm is the Central American site – use of the PDS increased the design storm estimate by 4% (50-year), with the increase declining to zero at 1000-year recurrence.

TABLE 4.1: RATIO OF DESIGN STORM ESTIMATES USING PDS TO THOSE USING AMS – DAILY DATA

Site Name	50-Year	100-Year	200-Year	500-Year	1,000-Year	PMP[1]
Walnut Gulch – 04	0.95	0.94	0.93	0.91	0.91	0.82
Walnut Gulch – 13	0.95	0.93	0.91	0.90	0.89	0.78
Walnut Gulch – 42	0.92	0.90	0.89	0.87	0.86	0.74
Walnut Gulch – 44	0.93	0.91	0.90	0.88	0.87	0.73
Walnut Gulch – 46	0.94	0.92	0.91	0.89	0.88	0.73
Walnut Gulch – 60	0.96	0.94	0.93	0.91	0.90	0.80
Walnut Gulch – 68	0.96	0.95	0.94	0.92	0.91	0.83
Walnut Gulch – 80	0.93	0.91	0.90	0.88	0.86	0.72
Walnut Gulch – 81	0.94	0.92	0.91	0.89	0.88	0.76
Carlin Newmont Mine	0.95	0.93	0.92	0.91	0.90	0.80
Brawley Peaks	0.97	0.96	0.95	0.94	0.93	0.85
Central America	1.04	1.03	1.02	1.01	1.00	0.96
Minimum	**0.92**	**0.90**	**0.89**	**0.87**	**0.86**	**0.72**
Maximum	**1.04**	**1.03**	**1.02**	**1.01**	**1.00**	**0.96**

Average	**0.95**	**0.94**	**0.92**	**0.91**	**0.90**	**0.79**

Note: [1]Prior to adjustment for observational interval.

At all sites – especially desert sites with infrequent rainfall – use of the PDS compresses the lower end of the range of the data series, reducing variability (and hence standard deviation), while also increasing the mean of the series. Appendix A provides further discussion of this phenomenon, with contrasting examples from Walnut Gulch – 04 and the Central American site.

4.2 Daily versus 24-Hour Data

Operating mines commonly employ automated weather stations that record rainfall data at hourly or (more typically) smaller intervals, enabling determination of true maximum “24-hour” events. However, especially in developing countries, available historical data often consists of daily (calendar-day) values – either summarized from finer-scale data and reported by calendar date, or obtained from physically reading an accumulating gage once each day. Depending on rainfall characteristics at the site and the time of day readings are taken, design storm estimates made from once-daily readings are biased low due to overnight storm events or those that span the recording time. WMO (1986) recommends a correction factor be applied to correct for the number of observational units. For 24-hour storms, the respective corrections are 113% for a single observation and 100.6% for 24 (hourly) observations.

Table 4.2 compares the results of Gumbel EV and Hershfield design storm estimates for the Walnut Gulch gages, using both daily and 24-hour AMS data. 24-hour data were unavailable at the other sites. In an effort to facilitate comparison the PMP estimates in Table 4.4 were made without applying the correction factor for the number of observations. The actual PMP estimate from the full Hershfield method would be 13% higher for the daily data. There is a slight trend towards a decreasing ratio with increasing recurrence interval. Hershfield’s correction factor envelopes the PMP results, indicating that use of the correction factor is a valid, albeit conservative approach given the data. Trends were similar for the PDS-derived estimates, however the ratios are slightly lower (maximum 1.12 for the PMP).

TABLE 4.2: RATIO OF DESIGN STORM ESTIMATES (AMS, 24-HOUR RECORD VS. DAILY RECORD)

Site Name	50-Year	100-Year	200-Year	500-Year	1000-Year	PMP[1]
Walnut Gulch – 04	1.12	1.12	1.12	1.12	1.12	1.12
Walnut Gulch – 13	1.17	1.18	1.18	1.18	1.18	1.13
Walnut Gulch – 42	1.07	1.07	1.07	1.06	1.06	0.99
Walnut Gulch – 44	1.09	1.08	1.08	1.07	1.07	1.07
Walnut Gulch – 46	1.09	1.08	1.08	1.08	1.07	1.02
Walnut Gulch – 60	1.05	1.05	1.05	1.04	1.04	1.03
Walnut Gulch – 68	1.06	1.06	1.06	1.06	1.05	1.01
Walnut Gulch – 80	1.06	1.06	1.05	1.05	1.04	1.03
Walnut Gulch – 81	1.09	1.09	1.09	1.09	1.08	1.06
Minimum	**1.05**	**1.05**	**1.05**	**1.04**	**1.04**	**0.99**

Maximum	**1.17**	**1.18**	**1.18**	**1.18**	**1.18**	**1.13**
Average	**1.09**	**1.09**	**1.09**	**1.08**	**1.08**	**1.05**

Note: [1]Prior to adjustment for observational interval.

Table 4.3 summarizes the differences between the average and peak annual maxima, demonstrating the correction factor is not necessarily applicable to the series itself. It should be noted that annual maximum events at desert sites tend to be short-duration, evening thunderstorms, rather than overnight events. Thus, many annual maximum events will be identical whether derived from daily or 24-hour data, assuming a midnight or morning delimiter between data 'days'. A site favored by longer-duration, overnight storm events may produce higher 24-hour/daily ratios. Therefore, Hershfield's correction factor is valid and even conservative for the PMP at the desert site investigated here; however, it is unknown whether it is conservative for sites with considerably different climate. A higher value of approximately 1.18 would envelop the results for frequency storms, and 1.09 would be valid on-average for frequency storms.

TABLE 4.3: SUMMARY OF ANNUAL MAXIMA –DAILY AND 24-HOUR RECORDS

Site Name	Average of Max Precipitation					Highest Annual Max Precipitation				
	Daily		24-hr		Ratio	Daily		24-hr		Ratio
	mm	in	mm	in		mm	in	mm	in	
Walnut Gulch – 04	34.0	1.34	38.8	1.53	1.14	68.3	2.69	72.6	2.86	1.06
Walnut Gulch – 13	33.3	1.31	38.2	1.50	1.15	65.5	2.58	90.7	3.57	1.38
Walnut Gulch – 42	33.6	1.32	37.6	1.48	1.12	59.9	2.36	78.1	3.08	1.30
Walnut Gulch – 44	35.8	1.41	40.7	1.60	1.14	73.9	2.91	73.9	2.91	1.00
Walnut Gulch – 46	37.1	1.46	41.9	1.65	1.13	85.2	3.36	95.6	3.77	1.12
Walnut Gulch – 60	36.4	1.43	40.0	1.57	1.10	73.2	2.88	73.7	2.90	1.01
Walnut Gulch – 68	37.9	1.49	41.4	1.63	1.09	80.3	3.16	93.0	3.66	1.16
Walnut Gulch – 80	35.3	1.39	39.5	1.56	1.12	66.5	2.62	67.1	2.64	1.01
Walnut Gulch – 81	35.8	1.41	39.8	1.57	1.11	58.7	2.31	66.7	2.63	1.14
Minimum	**33.3**	**1.31**	**37.6**	**1.48**	**1.09**	**58.7**	**2.31**	**66.7**	**2.63**	**1.00**
Maximum	**37.9**	**1.49**	**41.9**	**1.65**	**1.15**	**85.2**	**3.36**	**95.6**	**3.77**	**1.38**
Average	**35.5**	**1.40**	**39.8**	**1.57**	**1.12**	**70.2**	**2.76**	**79.0**	**3.11**	**1.13**

4.3 Comparison to Regional Studies – Frequency Storms

Point estimates were made for frequency-based design storms using the Gumbel EV distribution, while NOAA Atlas 14 estimates used a comprehensive, regional approach and the GEV distribution. Table 4.4 compares the results of the present study with published NOAA Atlas 14 estimates. Results were mixed – the Walnut Gulch sites averaged approximately 75% of the corresponding NOAA Atlas 14 precipitation depth (outside of the lower 90% confidence interval), and Brawley Peaks data yielded design storm estimates of less than half the published values. The Carlin Newmont Mine record matched the NOAA Atlas 14 estimates well, no doubt due to the fact that the Carlin Newmont Mine gage was included in NOAA Atlas 14's analysis while the Brawley Peaks and Walnut Gulch sites were not.

TABLE 4.4: GUMBEL EV DESIGN STORM ESTIMATE - RATIO VS. NOAA ATLAS 14

Site Name	Period of Record	50-Year	100-Year	200-Year	500-Year	1000-Year
Walnut Gulch – 04	60	0.78	0.77	0.77	0.76	0.75
Walnut Gulch – 13	60	0.73	0.73	0.72	0.72	0.71
Walnut Gulch – 42	59	0.75	0.75	0.74	0.74	0.74
Walnut Gulch – 44	59	0.76	0.75	0.75	0.74	0.74
Walnut Gulch – 46	52	0.78	0.77	0.77	0.76	0.76
Walnut Gulch – 60	59	0.76	0.75	0.75	0.75	0.75
Walnut Gulch – 68	59	0.79	0.78	0.78	0.78	0.77
Walnut Gulch – 80	51	0.73	0.72	0.71	0.71	0.70
Walnut Gulch – 81	52	0.70	0.68	0.67	0.66	0.65
Average Walnut Gulch	**56.8**	**0.75**	**0.75**	**0.74**	**0.74**	**0.73**
Carlin Newmont Mine	35	1.04	1.03	1.02	1.01	1.00
Brawley Peaks	27	0.48	0.48	0.47	0.46	0.46

As shown in Figure 4.1, the nearest NOAA Atlas 14 gages to Walnut Gulch are Tombstone, Arizona (9 km SE, 101 years of record, 1405 m elevation) and Fairbank 1 S (6 km SW, 60 years of record, 1174 m elevation). The nearest station to Brawley Peaks was the Bodie station in California (Figure 4.2), approximately 20 km west and 88 m higher (47 years of record, 2551 m elevation). The Brawley Peaks station's shorter record and lack of cold-season data is a possible reason for its relatively poor predictive performance, underscoring the importance of data completeness and record length – an item to be investigated further in this paper. The Walnut Gulch results do not lend themselves to a simple explanation, as the NOAA Atlas 14 gages mostly bracket the elevations of the Walnut Gulch sites, and all have long records.

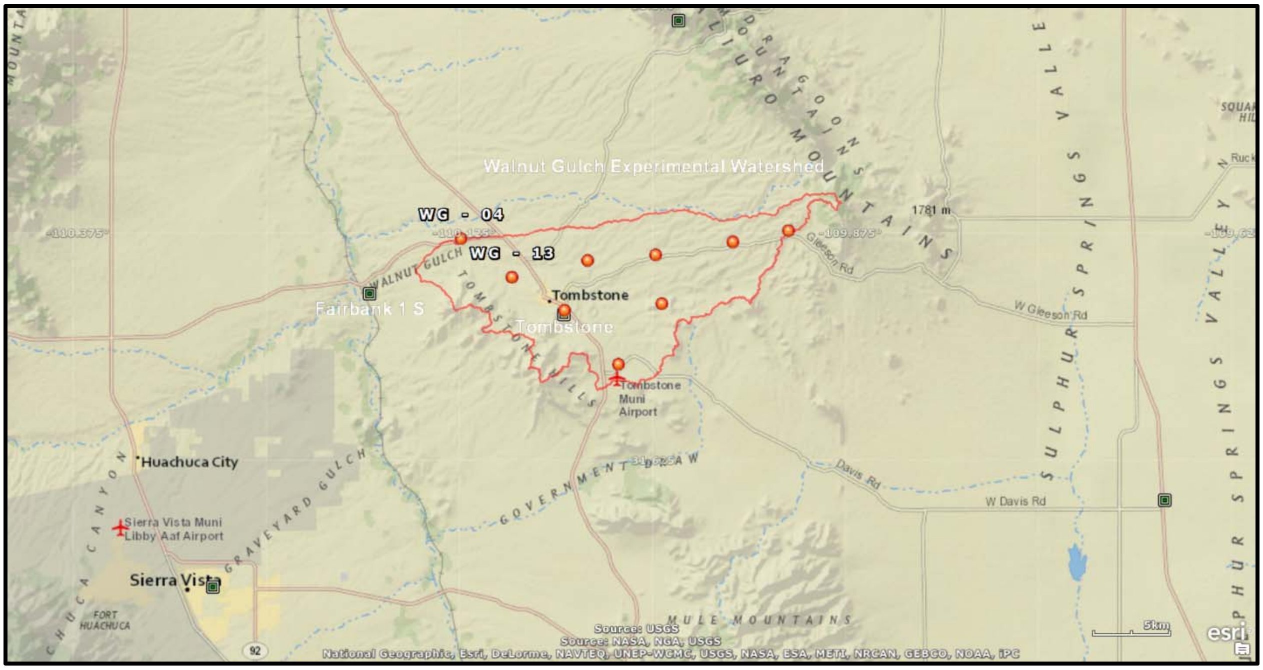

FIGURE 4.1: NOAA ATLAS 14 WEATHER STATIONS NEAR WALNUT GULCH

FIGURE 4.2: NOAA ATLAS 14 WEATHER STATIONS NEAR BRAWLEY PEAKS

4.4 Comparison to Regional Studies – PMP

Table 4.5 compares the statistically-derived PMP estimates with those computed using HMR-49 methodology. In both cases, point estimates are reported – i.e., no areal reduction factors are applied. Hershfield's method yielded significantly higher PMP estimates than HMR-49 at the two

Nevada sites, while HMR-49 yielded the higher estimate at Walnut Gulch. It should be noted that the 6-hour Local Storm exceeds the 24-hour General storm at the two Nevada sites, but Hershfield's result still exceeds both HMR-49 estimates at the Nevada sites.

TABLE 4.5: SUMMARY OF PMP ESTIMATES USING HMR-49 AND HERSHFIELD'S METHOD

Site Name	HMR-49 6-hr Local Storm		HMR-49 24-hr General Storm		Hershfield 24-hr PMP		Ratio of Hershfield vs. HMR-49 Storm	
	mm	in	mm	in	mm	in	Local	General
Walnut Gulch – All (Average)	373	14.67	391	15.38	317	12.48	0.85	0.81
Walnut Gulch – 81 (Least)	373	14.67	391	15.38	269	10.57	0.72	0.69
Walnut Gulch – 68 (Greatest)	370	14.56	391	15.38	371	14.61	1.00	0.95
Carlin Newmont Mine	227	8.94	185	7.30	325	12.80	1.43	1.75
Brawley Peaks	274	10.78	219	8.62	304	11.98	1.11	1.39

HMR-49 (Section 5.8, pp 140-142) included a comparison with PMP estimates using Hershfield's method, for sites with at least 50 years of record. In that comparison, though there was considerable scatter, HMR-49 estimates exceeded those of Hershfield. However, Hansen et al did not use Hershfield's correction factors, which would have served to increase the Hershfield estimates. In any case, use of HMR-49, application of similar methods, or other site-specific meteorological PMP studies should not be expected to provide results that agree with results from Hershfield's method. Detailed, site-specific meteorological studies may improve upon either HMR-style or Hershfield methods, but are generally cost-prohibitive and suffer from insufficient supporting data.

4.5 Effects of Record Length and Position within Record

As a first test of the effects of record length, each record was truncated several times, retaining only the last 10, 20, 30, 40, or 50 years of the record. Each analysis was re-run with the truncated record. Table 4.6 summarizes the results of the record truncation; Figures 4.3 and 4.4 show the numerical values and variability of the results for the PMP and 100-year events, respectively. All frequency storms followed a similar pattern; only the 100-year event is included for clarity.

TABLE 4.6: FRACTION OF FULL-RECORD ESTIMATE OBTAINED FROM TRIMMED RECORD USING MOST-RECENT 10 THROUGH 50 YEARS

Site Name	Period of Record	Ratio to Full Record, by Length of Record Segment (years)				
		10	20	30	40	50
PMP						
Walnut Gulch – 04	60	1.27	0.95	0.95	0.91	0.92
Carlin Newmont Mine	35	1.17	0.94	0.86	-	-
Brawley Peaks	27	0.86	1.17	-	-	-
Central America	42	1.43	1.09	1.08	1.00	-
100-Year						
Walnut Gulch – 04	60	1.04	0.93	0.95	0.92	0.94
Carlin Newmont Mine	35	1.15	0.99	0.93	-	-
Brawley Peaks	27	1.16	1.11	-	-	-
Central America	42	1.08	1.00	1.05	1.00	-

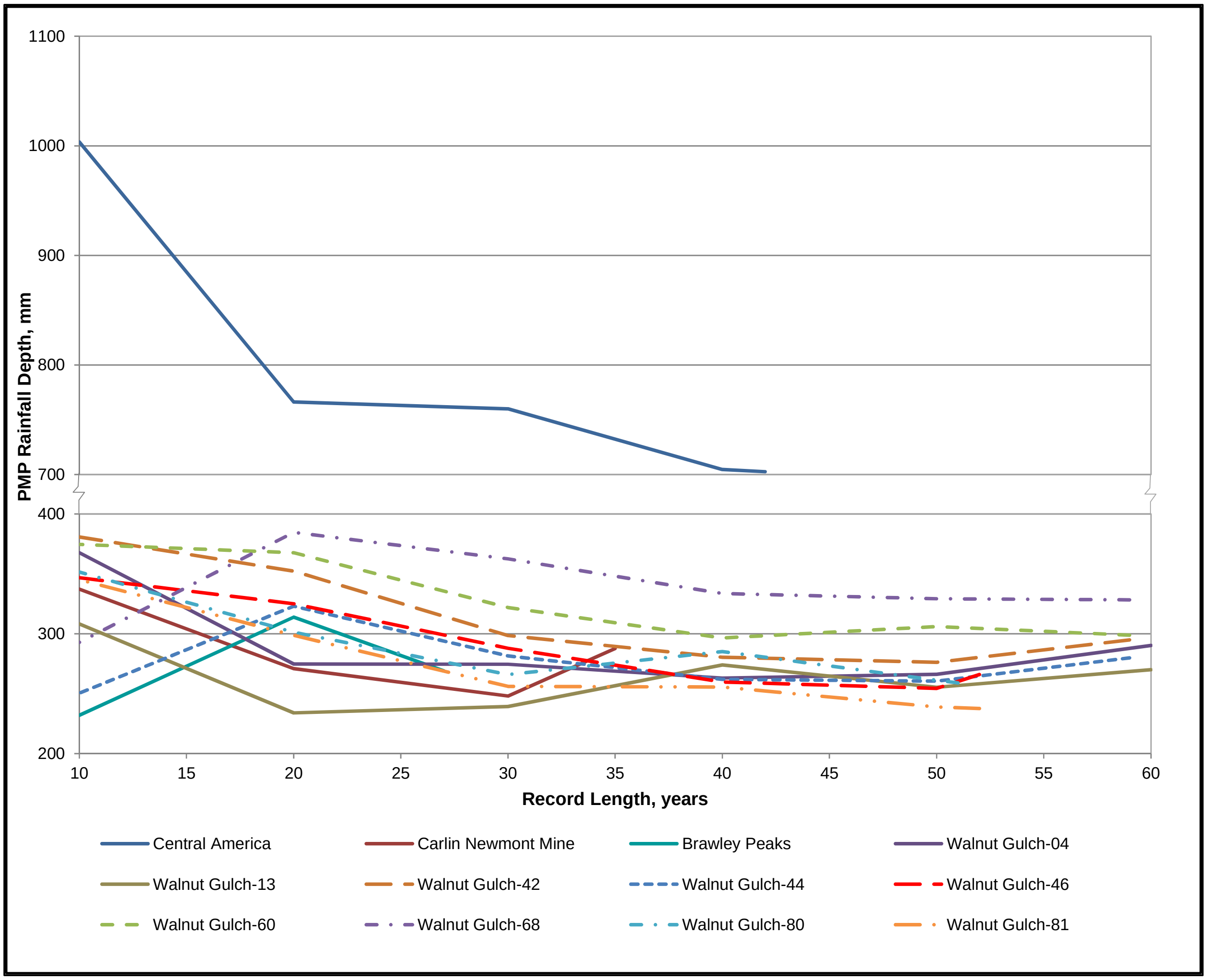

FIGURE 4.3: EFFECT OF RECORD TRUNCATION ON PMP ESTIMATE

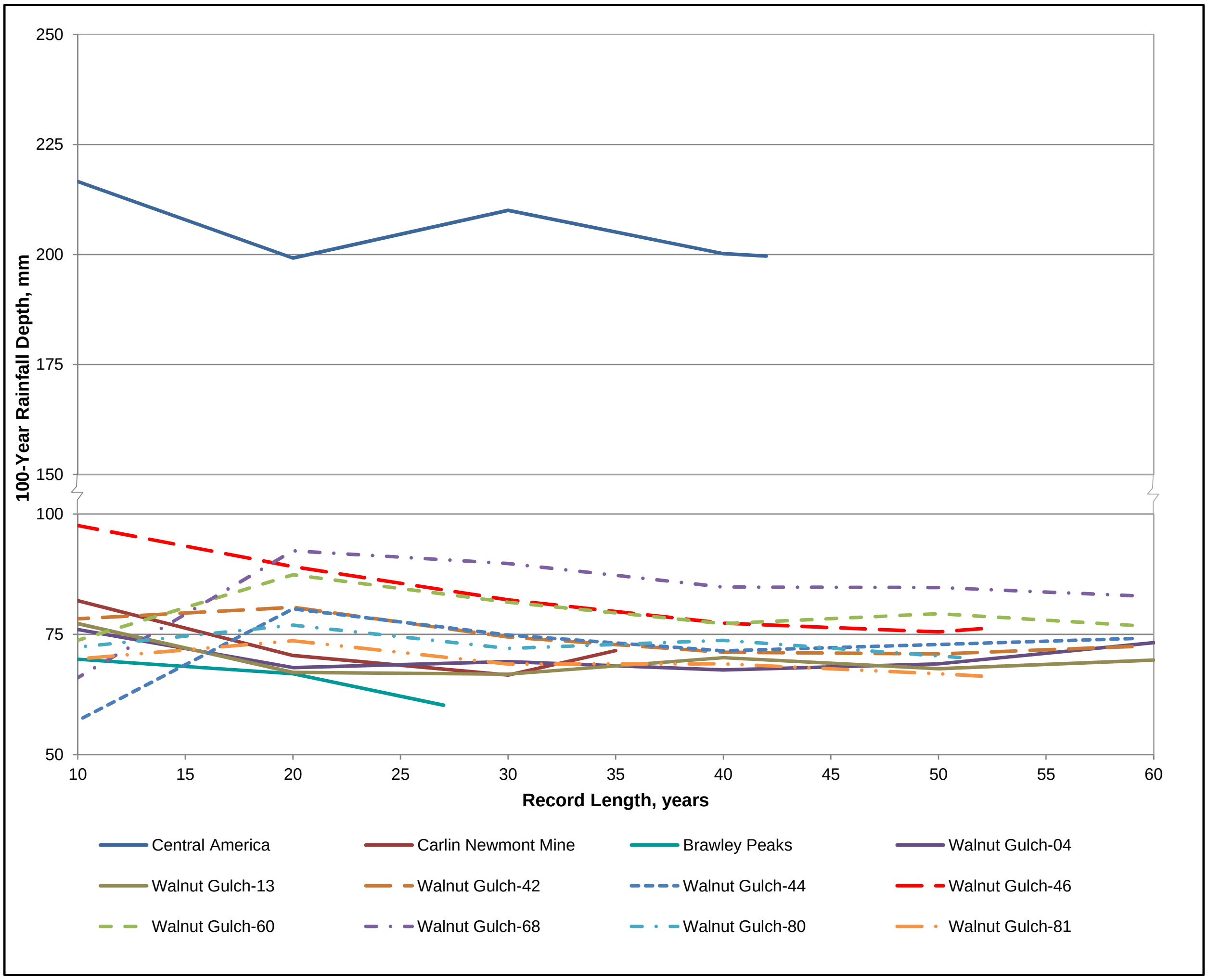

FIGURE 4.4: EFFECT OF RECORD TRUNCATION ON 100-YEAR FREQUENCY DESIGN STORM ESTIMATE

For both the PMP and the 100-year storms, there is an obvious trend towards convergence with the "full-record" estimate as the partial record length increases. Deviations from the general trend are explained by the position of outliers within the record. For example, sampling only the final segment of the Carlin Newmont Mine record eliminates the highest value in the record, which occurs at the beginning of the period of record. This effect is most pronounced for the 30-year sample.

The divergence from the "full record" value is positive for the smallest samples. The sharp uptick in PMP estimates for the 10-year samples is, in part, due to the period-of-record correction factors Hershfield developed – 130% correction to the standard deviation for a 10-year record, versus 108% for a 20-year record. The Gumbel EV estimates do not incorporate such a correction factor yet they too overestimate rainfall depth for the shortest records. This suggests that the use of short records, while not ideal, may err on the conservative side more often than not – especially for PMP estimates which are subject to a potentially overly large correction factor (using Hershfield's method).

To confirm the hypothesis that statistical approaches overestimate design storm depths for short records, the analysis was repeated with successive 10, 20, and 30-year sample windows, cycled through each record to provide a number of design storm estimates equal to the record length. The results identified the effect of the position of outliers within the record, and give an approximation of the uncertainty inherent in estimating design storms from short records. Figure 4.5(a) is a typical result, using successive 10-year samples of the Walnut Gulch -04 site. The PMP estimate fluctuates wildly, according to which outliers are sampled. The 100-year results are better behaved, but still show significant variation. Figures 4.5(b) and 4.5(c) show the results for 20-year and 30-year samples, respectively, indicating that variability continues, albeit with less amplitude, for record lengths commonly considered "long".

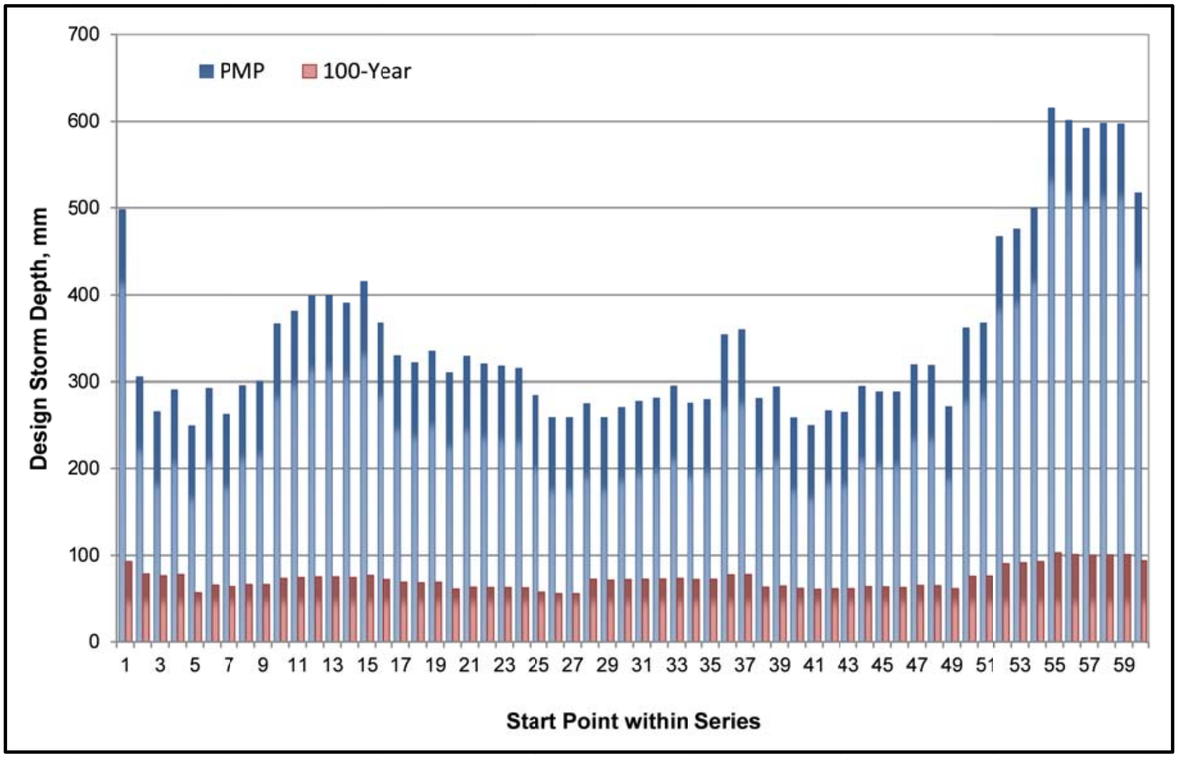

FIGURE 4.5(A): VARIATION IN PMP AND 100-YEAR ESTIMATES FOR WALNUT GULCH-04 RECORD RESAMPLED IN 10-YEAR SEGMENTS

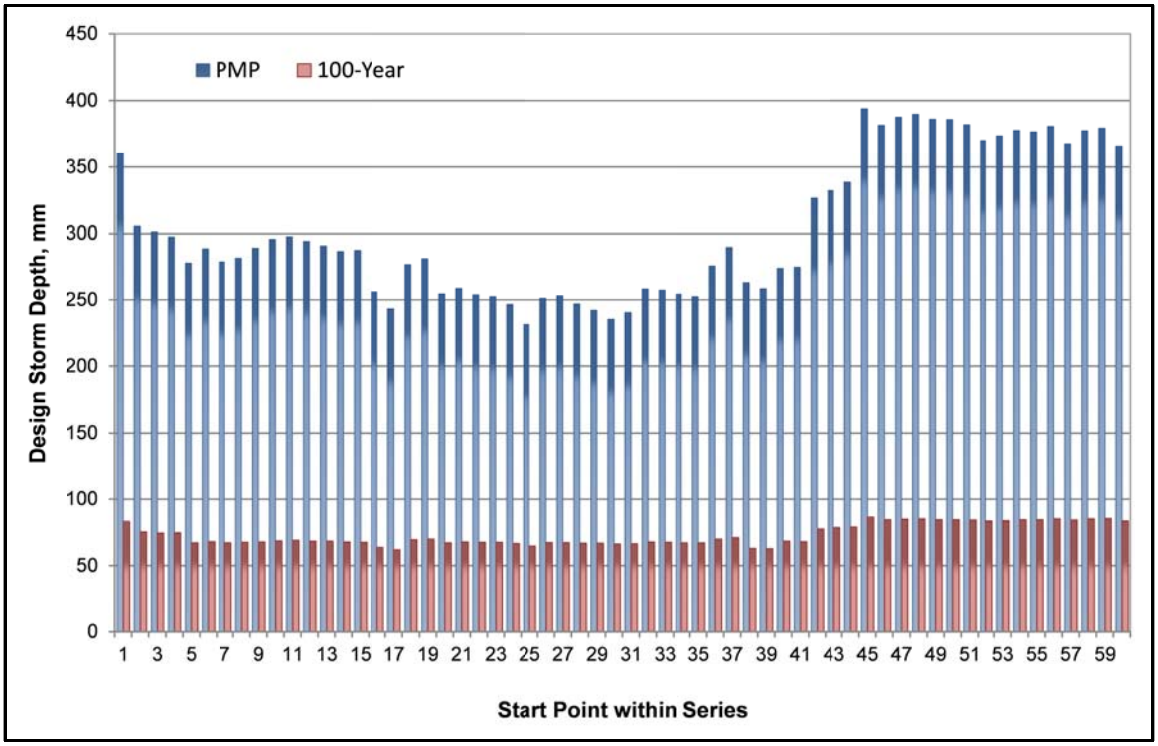

FIGURE 4.5(B): VARIATION IN PMP AND 100-YEAR ESTIMATES FOR WALNUT GULCH-04 RECORD RESAMPLED IN 20-YEAR SEGMENTS

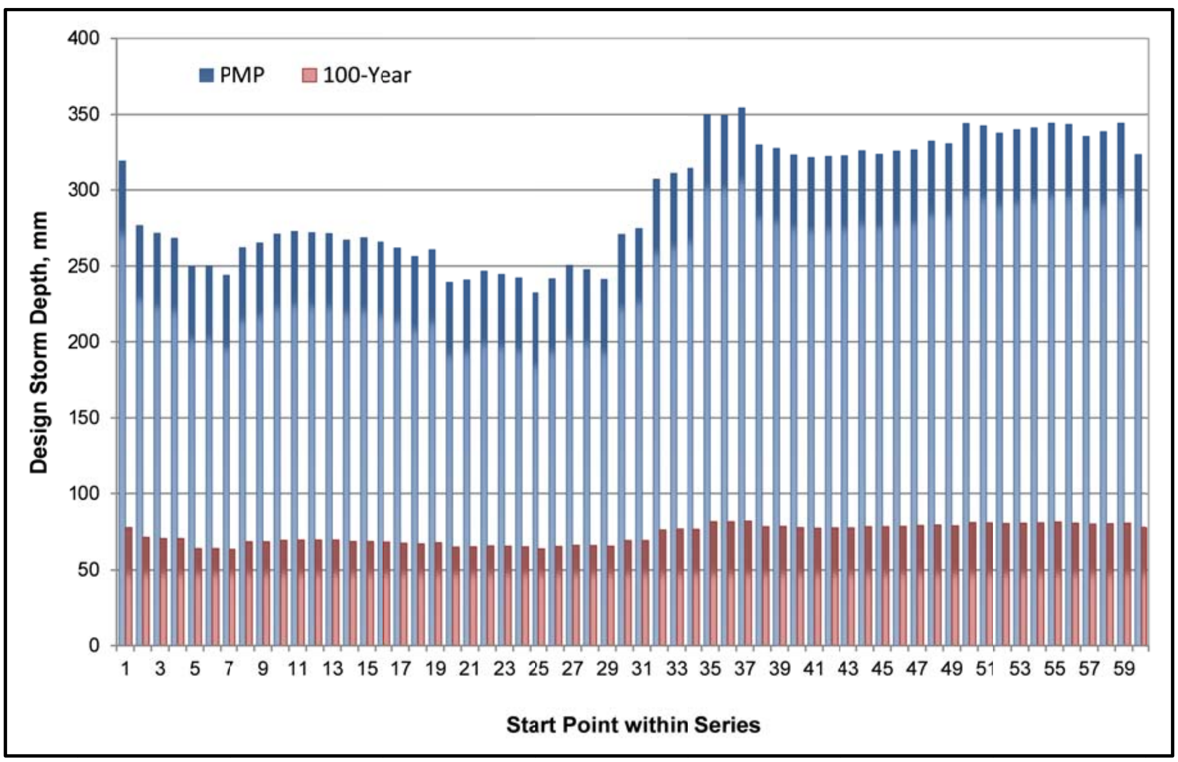

FIGURE 4.5(C): VARIATION IN PMP AND 100-YEAR ESTIMATES FOR WALNUT GULCH-04 RECORD RESAMPLED IN 30-YEAR SEGMENTS

Figures B-1 through B-4 in Appendix B present the results for Walnut Gulch-13, Carlin Newmont Mine, Brawley Peaks, and the Central American site. As would be expected, the amplitude of variation of the estimate decreases as the sample size increases.

Table 4.7 summarizes the results for the PMP, expressed as ratios of the full record-based estimate, which is considered as the “true” estimate. Similar to the case with truncated records, the Hershfield PMP estimate generally overestimates the PMP for short records. However, the range of the estimate can vary considerably for short records due to the effect of missing outliers. This effect is least noticeable for the tropical Central American site, which has lower variability in the underlying data. From Table 4.7 it can be seen that while, on average, the Hershfield method is biased high for short records, the variability induced by outliers can have considerable effects on the estimate, and reduction of the 10-year period of record correction factor is not advised.

TABLE 4.7: FRACTION OF FULL-RECORD PMP ESTIMATE OBTAINED FROM CYCLING 10- THROUGH 30-YEAR SAMPLES OF RECORD

Site Name	Period of Record	Statistic	Design Storm Estimate Ratio vs. Full Record, by Sample Length (years)		
			10	20	30
Walnut Gulch – 04	60	Average	1.19	1.04	1.02
		Minimum	0.55	0.80	0.80
		Maximum	1.94	1.36	1.22
Walnut Gulch – 13	60	Average	1.20	1.04	1.01
		Minimum	0.86	0.85	0.82
		Maximum	2.12	1.22	1.25
Carlin Newmont Mine	35	Average	1.11	1.01	0.99
		Minimum	0.56	0.73	0.78
		Maximum	1.98	1.29	1.09
Brawley Peaks	27	Average	1.05	0.98	-
		Minimum	0.40	0.56	-
		Maximum	1.86	1.19	-
Central America	42	Average	1.12	1.02	1.01
		Minimum	0.74	0.82	0.89
		Maximum	1.56	1.23	1.11

Table 4.8 summarizes the results for the 100-year event. Other frequency storms provide similar results. On average, use of a 10-year record will slightly underestimate the 100-year design storm, but may also significantly under- or overestimate the 100-year design storm, depending on the occurrence of outliers within the record. Estimates begin to converge as the sample size approaches 30 years, depending on the variability of the underlying data. More consistent datasets such as the Central American site allow reliable 100-year design storm estimates to be made from shorter records of approximately 20 years in length.

TABLE 4.8: FRACTION OF FULL-RECORD 100-YEAR ESTIMATE OBTAINED FROM CYCLING 10- THROUGH 30-YEAR SAMPLES OF RECORD

Site Name	Period of Record	Statistic	100-yr Design Storm Estimate Ratio vs. Full Record, by Sample Length (years)		
			10	20	30
Walnut Gulch – 04	60	Average	0.99	1.00	1.00
		Minimum	0.73	0.85	0.87
		Maximum	1.33	1.18	1.12
Walnut Gulch – 13	60	Average	0.99	1.00	1.00
		Minimum	0.76	0.86	0.86
		Maximum	1.40	1.11	1.12
Carlin Newmont Mine	35	Average	0.98	0.99	1.00
		Minimum	0.70	0.76	0.91
		Maximum	1.38	1.15	1.05
Brawley Peaks	27	Average	0.91	0.98	-
		Minimum	0.52	0.59	-
		Maximum	1.41	1.12	-
Central America	42	Average	0.99	1.00	1.00
		Minimum	0.77	0.84	0.94
		Maximum	1.16	1.11	1.06

Many mines establish design storm estimates early in the project development or permitting process, and never revisit them after the initial mine design has been completed. Given the duration of the project development process, and the operating life of a mine, significant periods of new data become available between the initial permitting effort and operations – especially at long-lived projects with multiple, sequential waste management facilities on site. Given the demonstrated variability in design storm estimates according to the period of record, design storms estimates should be revisited when sufficient data becomes available. The triggering amount is site-specific, depending on data variability and the period of record used for the initial estimate. In general, any record of less than 20 years duration should be augmented whenever at least 5 years of new data becomes available.

5.0 CONCLUSIONS

Using data from desert, tropical, and mountainous sites, common methods for estimating the PMP and other infrequent storm events from limited data were investigated. Key findings regarding the potential bias and limitations of the Hershfield, Gumbel EV, and U.S. NWS methods include:

- The use of the PDS instead of the AMS for frequency analysis of 50-year and larger events is biased 4 to 10% low on average, and as much as 14% low at the worst site investigated. The PDS is therefore not recommended for estimation of large design events from frequency analysis. The PDS should never be used in Hershfield's method for estimating the PMP.
- Hershfield's correction factor for observational intervals (1.13 for daily observations) is valid and sometimes conservative for the PMP at the desert site investigated here; however, it is unknown whether it is conservative for other sites. For the PMP, Hershfield's correction factors should be used as-published. A higher value of approximately 1.18 would envelop the results for frequency storms at the desert site, and 1.09 would be valid on-average for frequency storms.
- Gumbel EV frequency analysis results are often inconsistent with NOAA's Atlas 14 regional studies. The desert site averaged approximately 75% of the corresponding NOAA Atlas 14 estimate, and the mountain site yielded design storm estimates of less than half the published values, while the high desert site matched the published values within a few percent. When regional data is available, it can produce better estimates, but only when corrections for local-scale effects can be made.
- NWS HMR-49 PMP estimates are similarly inconsistent – two sites had HMR-49 estimates significantly lower than those from Hershfield's method, while another was slightly higher than the Hershfield estimate. At U.S. sites, the HMR series will generally be favored due to regulation. At international sites, adaptation of HMR-49, application of similar methods, or other site-specific meteorological PMP studies should not be expected to provide results that agree with results from Hershfield's method.
- PMP and frequency storm estimates are sensitive to record length, especially in the presence of highly variable desert data where outliers exert a large influence. The sensitivity persists for the PMP, albeit with less amplitude, for record lengths commonly considered "long" (~30 years).
- Hershfield's period-of-record correction factor (130% correction to the standard deviation for a 10-year record) tends to overcorrect PMP estimates for short (<20-year) records. As a result, on average, PMP estimates from 10-year records are biased 5% to 20% high; however, extreme values as much as 60% low may still occur. This suggests that

the use of short records, while not ideal, may err on the conservative side more often than not. The variability induced by outliers can have considerable effects on the estimate, and reduction of the Hershfield's period of record correction factor is not advised.

- On average, use of a 10-year record will slightly underestimate the 100-year design storm, but may significantly under- or overestimate the 100-year design storm, depending on the occurrence of outliers within the record. Estimates begin to converge as the sample size approaches 30 years, depending on the variability of the underlying data - consistent datasets such as the Central American site allow reliable 100-year design storm estimates to be made from shorter records of approximately 20 years in length.
- While compensating errors may exist, uncertainty in design precipitation can easily exceed 10% to 30% due to only one or two of the factors mentioned above. Because runoff production is non-linear with respect to rainfall, a +/- 30% change in design storm depth would typically equate to +/- 40 to 50% change in runoff volume or peak flow, affecting facility size (spillway width, TSF freeboard, etc.), cost, and design adequacy accordingly.

Many mines establish design storm estimates early in the project development or permitting process, and never revisit them after the initial mine design has been completed. Given the duration of the project development process, and the operating life of a mine, significant periods of new data become available between the initial permitting effort and operations – especially at long-lived projects with multiple, sequential waste management facilities on site. Given the demonstrated variability in design storm estimates according to the period of record, design storm estimates should be revisited when sufficient data becomes available. Revisiting design storm estimates during operations or as the site approaches closure may provide additional data to support a more accurate estimate of design storms to be used in the design of new facilities, expansion of existing facilities, or closure of facilities. Oftentimes, the capital cost of facility design is largely influenced by freeboard requirements. The triggering amount of data justifying reevaluation is site-specific, depending on data variability and the period of record used for the initial estimate. In general, any record of less than 20 years duration should be augmented whenever at least 5 years of new data becomes available.

6.0 REFERENCES

American Meteorological Society, 1959. *Glossary of Meteorology*, Boston, MA. 638 pp.

Bonnin, GM, D Martin, B Lin, T Parzybok, M Yekta, and D Riley, 2006. *NOAA Atlas 14 Precipitation-Frequency Atlas of the United States, Volume 1 Version 5.0: Semiarid Southwest (Arizona, Southeast California, Nevada, New Mexico, Utah).* National Weather Service, Silver Spring, MD, 2004. Revised 2006 and 2011.

Canadian Dam Association (CDA), 2007. *Dam Safety Guidelines*. Edmonton, Alberta, Canada.

Hansen, E.M., F.K. Schwarz, and J.T. Riedel, 1984. *Hydrometeorological Report No. 49: Probable Maximum Precipitation Estimates, Colorado River and Great Basin Drainages*. National Weather Service, Silver Spring, MD. October 1994.

United States Geological Survey (USGS), 1982. *Guidelines for Determining Flood Flow Frequency.* Bulletin #17B of the Hydrology Subcommittee, Interagency Advisory Committee on Water Data. Revised September 1981, Editorial Corrections March 1982.

U.S. Department of Agriculture (USDA), 2013. Southwest Watershed Research Center, Walnut Gulch Experimental Watershed Precipitation Data, various stations, accessed 05 November 2013 at: http://www.tucson.ars.ag.gov/dap/

Western Regional Climate Center (WRCC), 2013a. Western Regional Climate Center, Desert Research Institute, Period of Record Monthly Climate Summaries, Carlin Newmont Mine station, accessed at: http://www.wrcc.dri.edu/summary/Climsmnv.html

Western Regional Climate Center (WRCC), 2013b. Western Regional Climate Center, Desert Research Institute, Remote Automated Weather Station Climate Data, Brawley Peaks Station, accessed at: http://www.raws.dri.edu/nvF.html

World Meteorological Organization (WMO), 1986. Operational Hydrology Report No. 1: *Manual for Estimation of Probable Maximum Precipitation*, Second Edition. WMO – No. 332.

APPENDIX A

Supplemental Information – Annual Maximum vs. Partial Duration Series

At all sites – especially desert sites with infrequent rainfall – use of the PDS compresses the lower end of the range of the data series, reducing variability (and hence standard deviation), while also increasing the mean of the series. Table A.1 lists an example, from Walnut Gulch, having an approximately 10% swing in the mean and standard deviation between the AMS and PDS. The highest ten values are identical for both series, while the bottom 10 values are 26% to 65% higher for the PDS. The AMS contains 16 events containing less than 25 mm of rainfall, unsurprising given the localized and infrequent storms typical of the Arizona desert. Even the lowest value of the PDS exceeds 25 mm, also reasonable given that the PDS contains several events from each of a number of wet years, while the AMS by definition will contain the highest rainfall day for each dry year as well as each wet year.

TABLE A.1: PDS AND AMS EXCERPTS FOR WALNUT GULCH GAGE 04

AMS			PDS		
Year	Precipitation		Year	Precipitation	
	mm	In		mm	in
Top 10 Values in Series					
1954	68.33	2.69	1954	68.33	2.69
1957	66.80	2.63	1957	66.80	2.63
1990	61.72	2.43	1990	61.72	2.43
2012	59.82	2.355	2012	59.82	2.355
1972	53.85	2.12	1972	53.85	2.12
1968	52.07	2.05	1968	52.07	2.05
1998	50.29	1.98	1998	50.29	1.98
2008	49.53	1.95	2008	49.53	1.95
1977	45.47	1.79	1977	45.47	1.79
1966	44.70	1.76	2012	45.09	1.775
Bottom 10 Values in Series					
1991	22.61	0.89	1989	28.45	1.12
1974	21.84	0.86	2005	28.45	1.12
2013	20.83	0.82	2007	28.45	1.12
1970	19.81	0.78	2010	28.45	1.12
1981	19.81	0.78	1958	28.19	1.11
1959	19.56	0.77	1971	27.94	1.10
1956	19.30	0.76	1993	27.94	1.10
1978	19.05	0.75	2001	27.31	1.075
2009	17.02	0.67	1977	27.18	1.07
2004	16.51	0.65	1990	27.18	1.07

In contrast to desert sites, tropical sites experience the opposite effect – widening the range of the high end of the data series, as well as the low end, with multiple large storms from wet years exceeding the larger annual maxima. Table A.2 presents an example from Central America. For this case, the PDS has a 16% higher mean and 16% lower standard deviation than the AMS, with the net effect of increasing the design storm estimates, especially for lower recurrence intervals. While use of the PDS may be conservative for tropical sites, based on the majority of this analysis (at arid and semiarid sites) the use of the PDS is usually biased low and is therefore not recommended for estimation of large design events.

TABLE A.2: PDS AND AMS EXCERPTS FOR CENTRAL AMERICAN SITE

AMS			PDS		
Year	Precipitation		Year	Precipitation	
	mm	In		mm	in
Top 10 Values in Series					
1983	198.88	7.83	1983	198.88	7.83
2008	171.20	6.74	1983	190.50	7.50
2010	163.83	6.45	1983	187.96	7.40
1998	149.86	5.90	2008	171.20	6.74
1987	149.86	5.90	2010	163.83	6.45
1997	130.81	5.15	1987	149.86	5.90
1986	127.51	5.02	1998	149.86	5.90
1959	126.49	4.98	1997	130.81	5.15
1978	121.92	4.80	1986	127.51	5.02
1990	121.41	4.78	1986	127.51	5.02
Bottom 10 Values in Series					
2005	76.45	3.01	2010	98.30	3.87
1979	76.20	3.00	2009	97.28	3.83
2002	74.93	2.95	1998	95.25	3.75
1961	73.66	2.90	2008	95.00	3.74
2007	71.63	2.82	1994	93.47	3.68
1958	69.09	2.72	1986	92.46	3.64
2000	58.42	2.30	2008	92.20	3.63
1976	56.64	2.23	1995	90.93	3.58
2011	54.61	2.15	1982	90.17	3.55
1962	12.70	0.50	1980	89.41	3.52

APPENDIX B

Figures – Effects on PMP and 100-year storm of Resampling Records in 10-year, 20-year, and 30-Year Segments

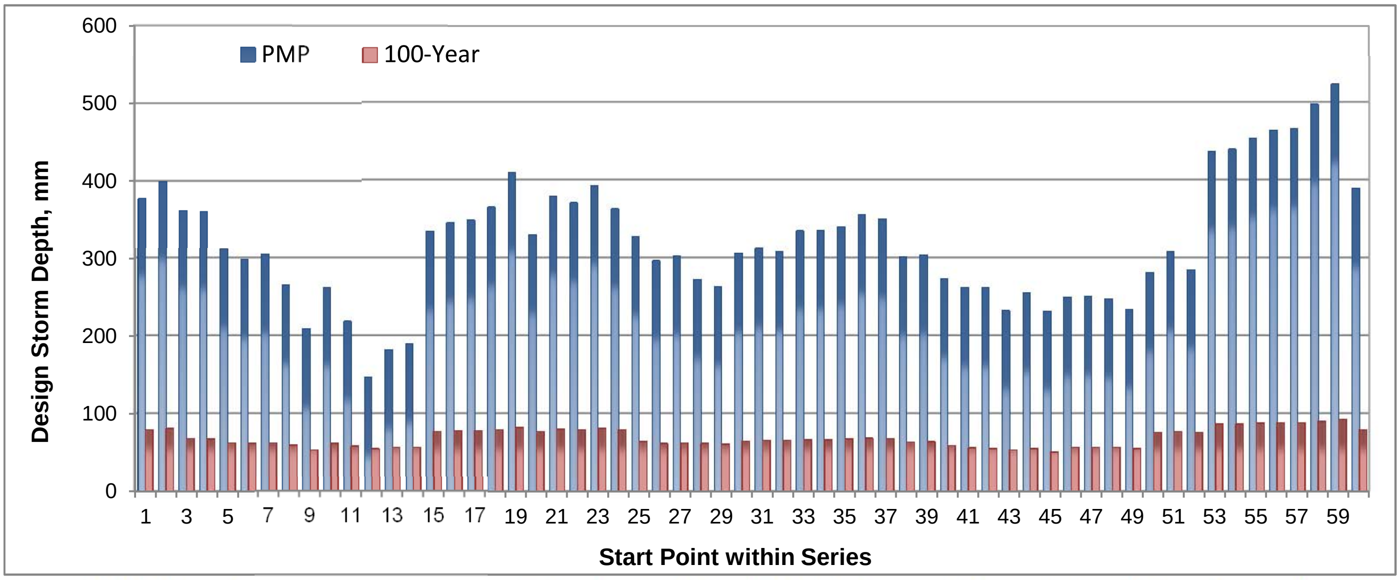

Figure B.1(a): Variation in PMP and 100-Year Estimates for Walnut Gulch-13 Record Resampled in 10-Year Segments

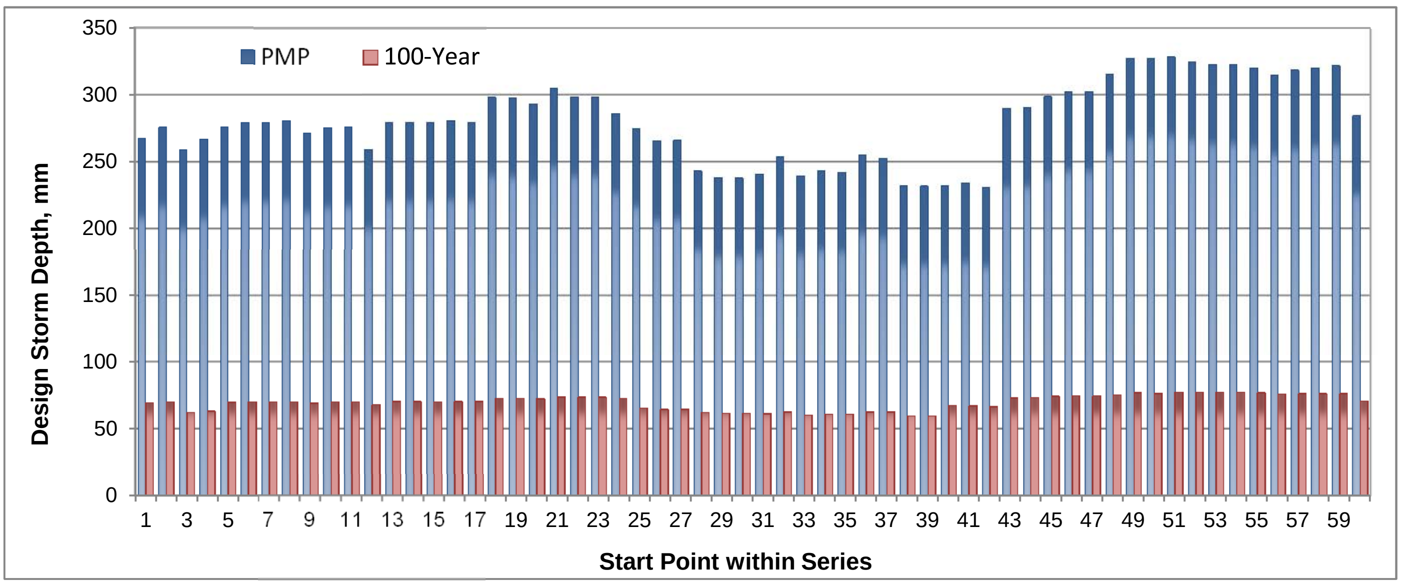

Figure B.1(b): Variation in PMP and 100-Year Estimates for Walnut Gulch-13 Record Resampled in 20-Year Segments

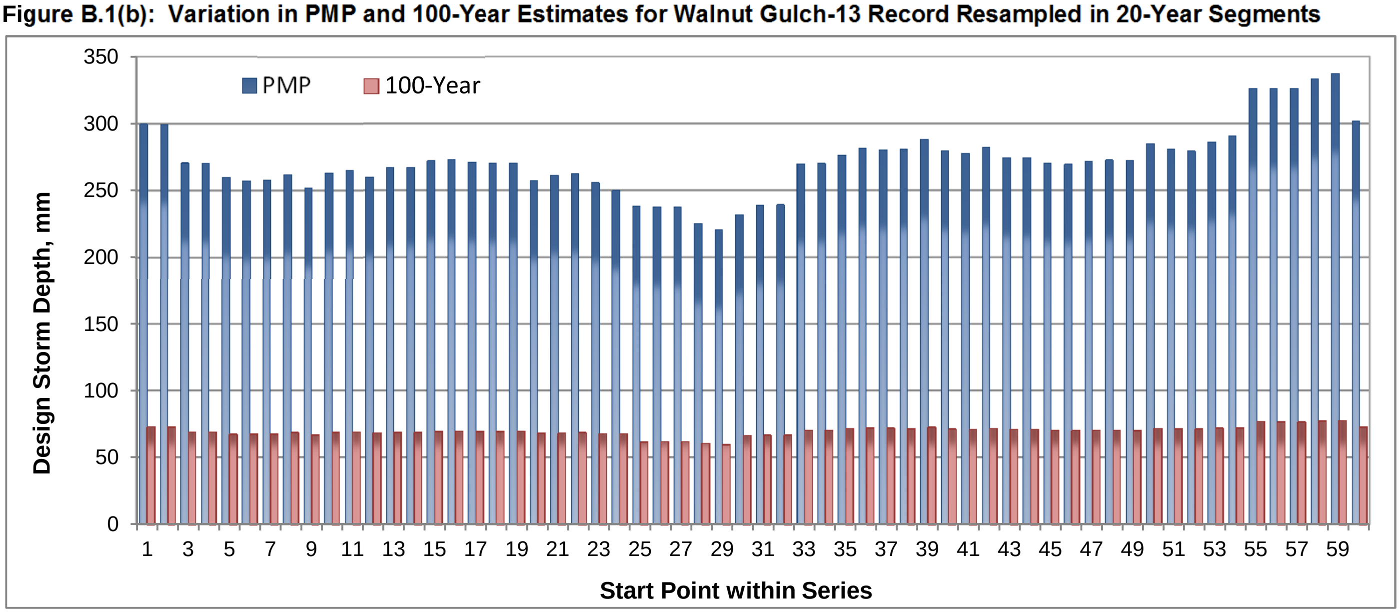

Figure B.1(c): Variation in PMP and 100-Year Estimates for Walnut Gulch-13 Record Resampled in 30-Year Segments

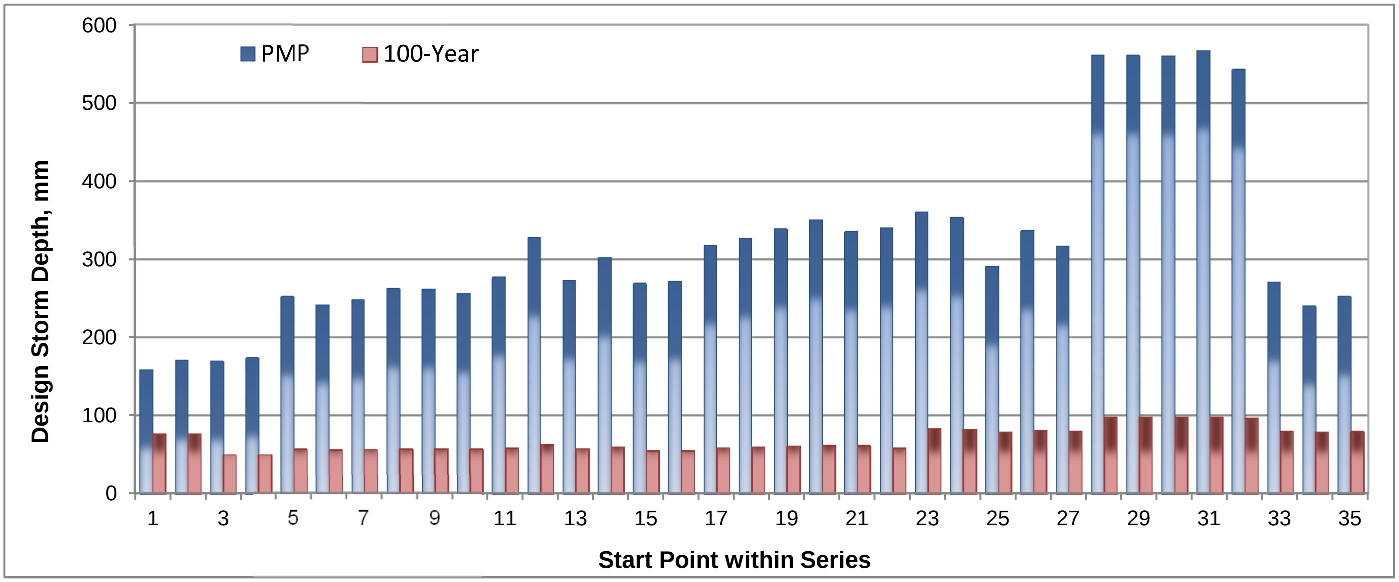

Figure B.2(a): Variation in PMP and 100-Year Estimates for Carlin Newmont Mine Record Resampled in 10-Year Segments

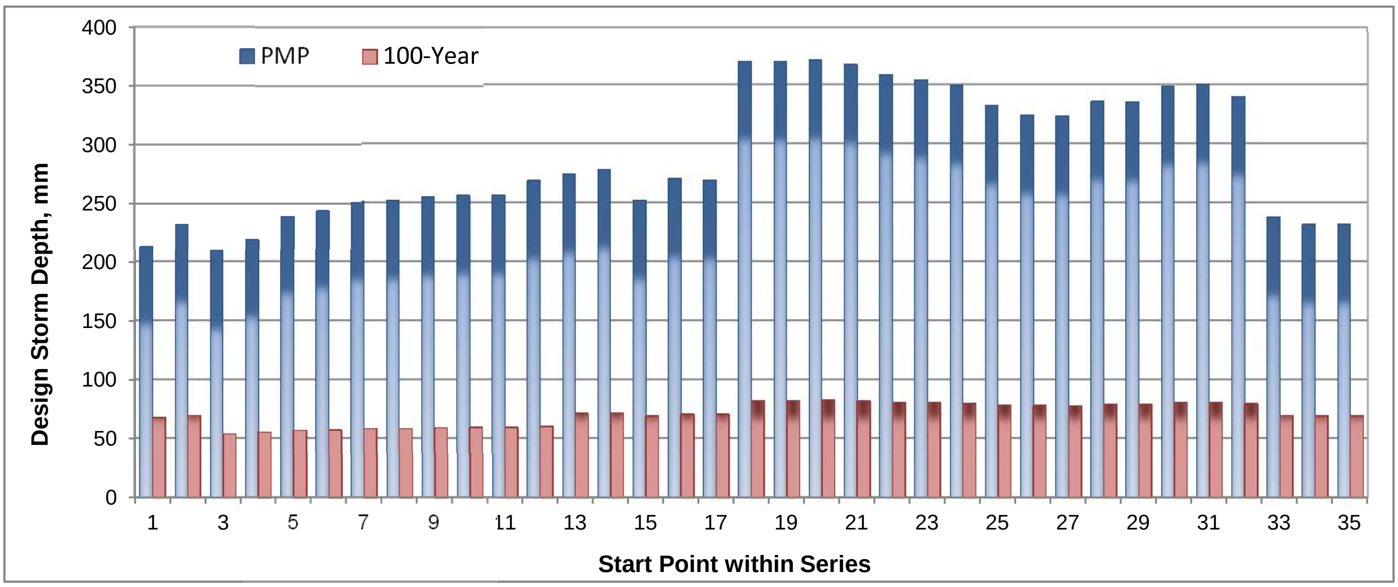

Figure B.2(b): Variation in PMP and 100-Year Estimates for Carlin Newmont Mine Record Resampled in 20-Year Segments

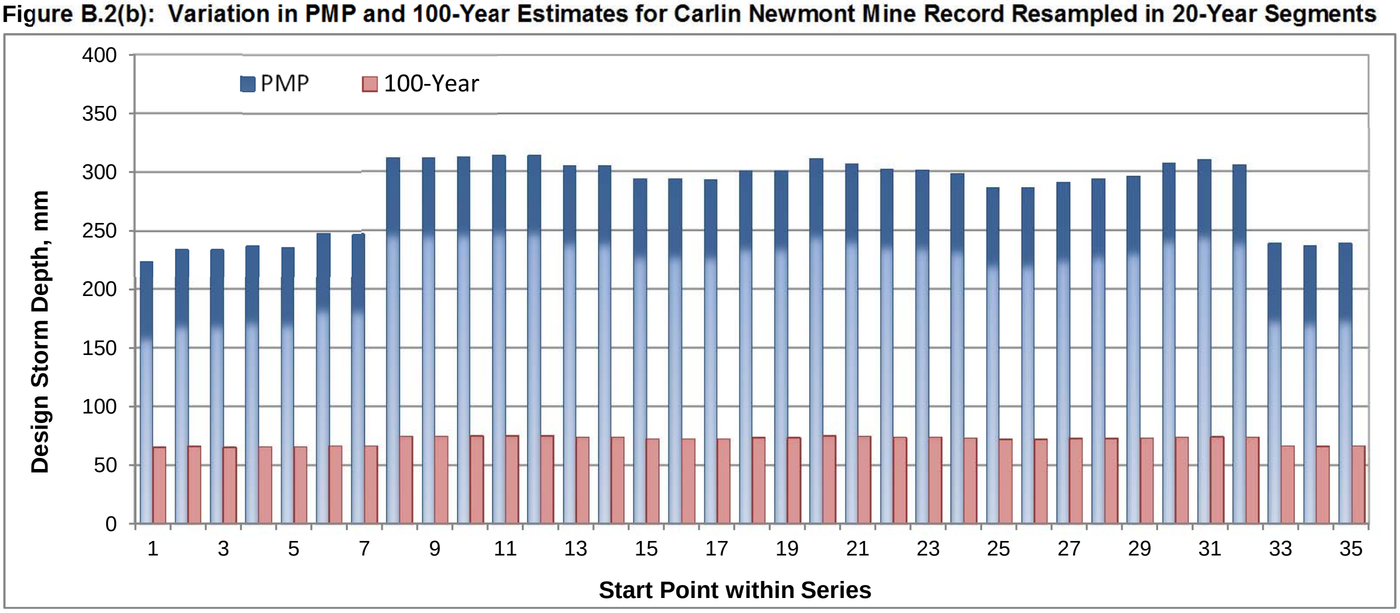

Figure B.2(c): Variation in PMP and 100-Year Estimates for Carlin Newmont Mine Record Resampled in 30-Year Segments

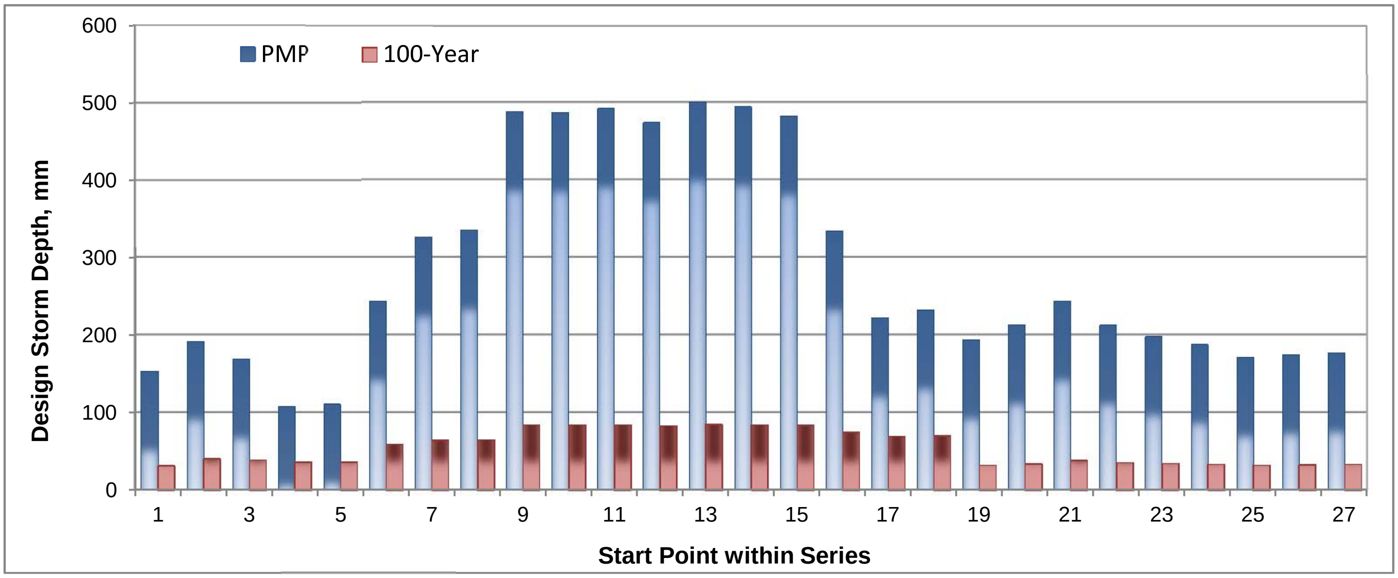

Figure B.3(a): Variation in PMP and 100-Year Estimates for Brawley Peaks Record Resampled in 10-Year Segments

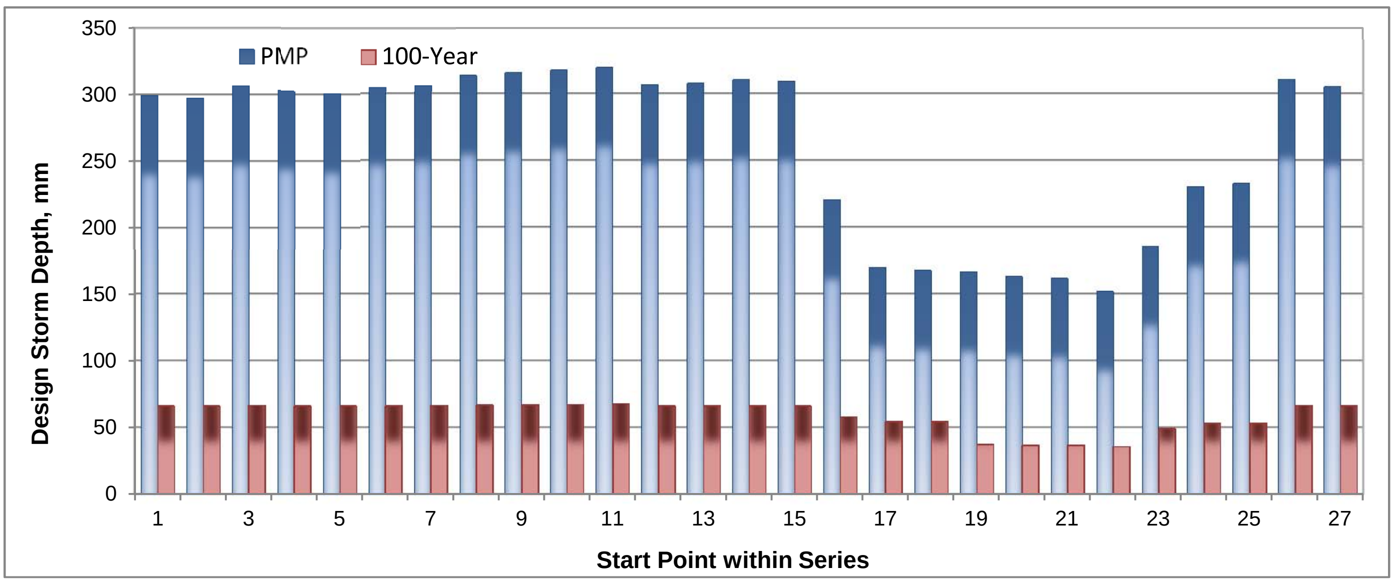

Figure B.3(b): Variation in PMP and 100-Year Estimates for Brawley Peaks Record Resampled in 20-Year Segments

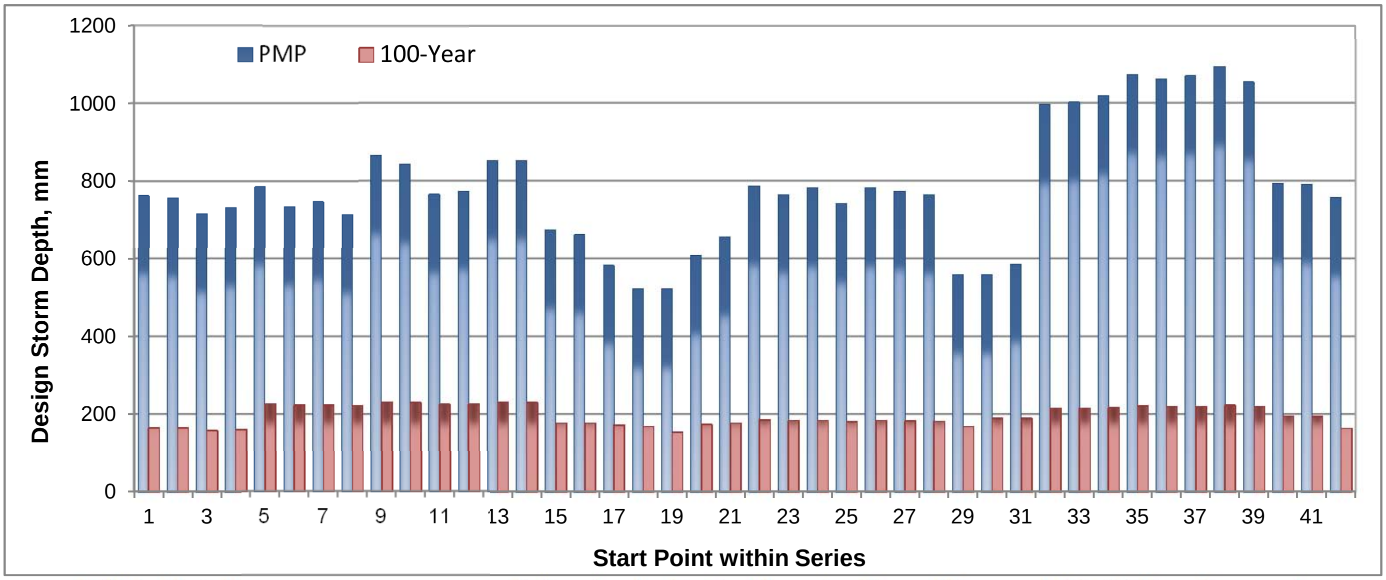

Figure B.4(a): Variation in PMP and 100-Year Estimates for Central America Record Resampled in 10-Year Segments

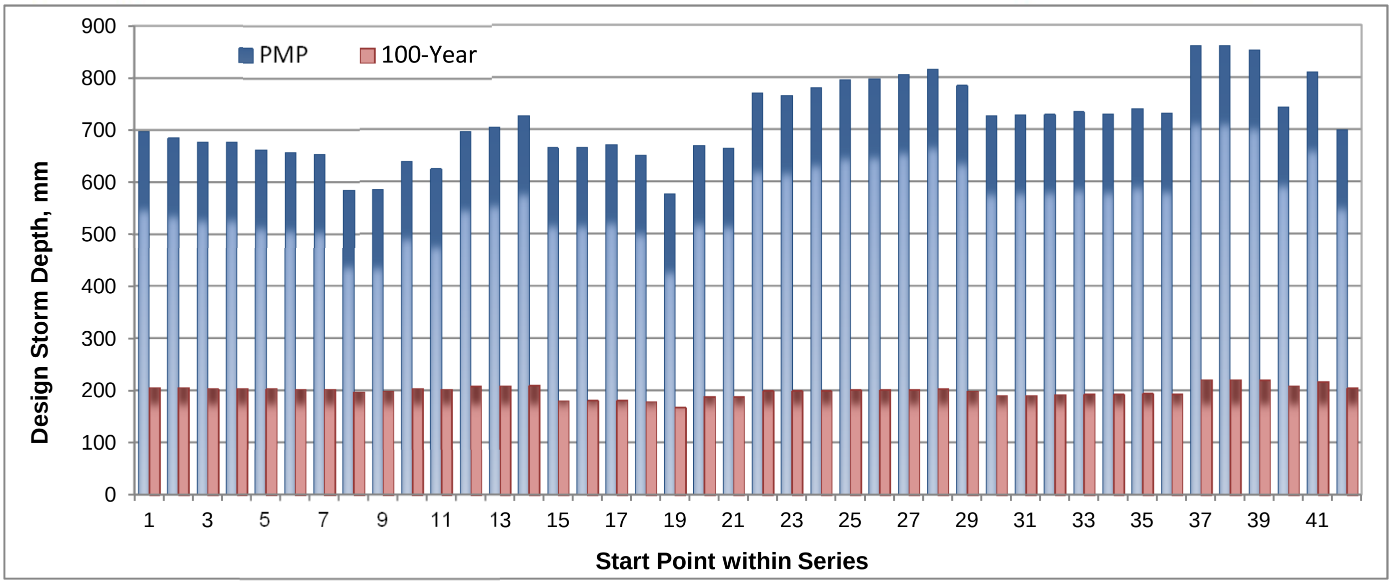

Figure B.4(b): Variation in PMP and 100-Year Estimates for Central America Record Resampled in 20-Year Segments

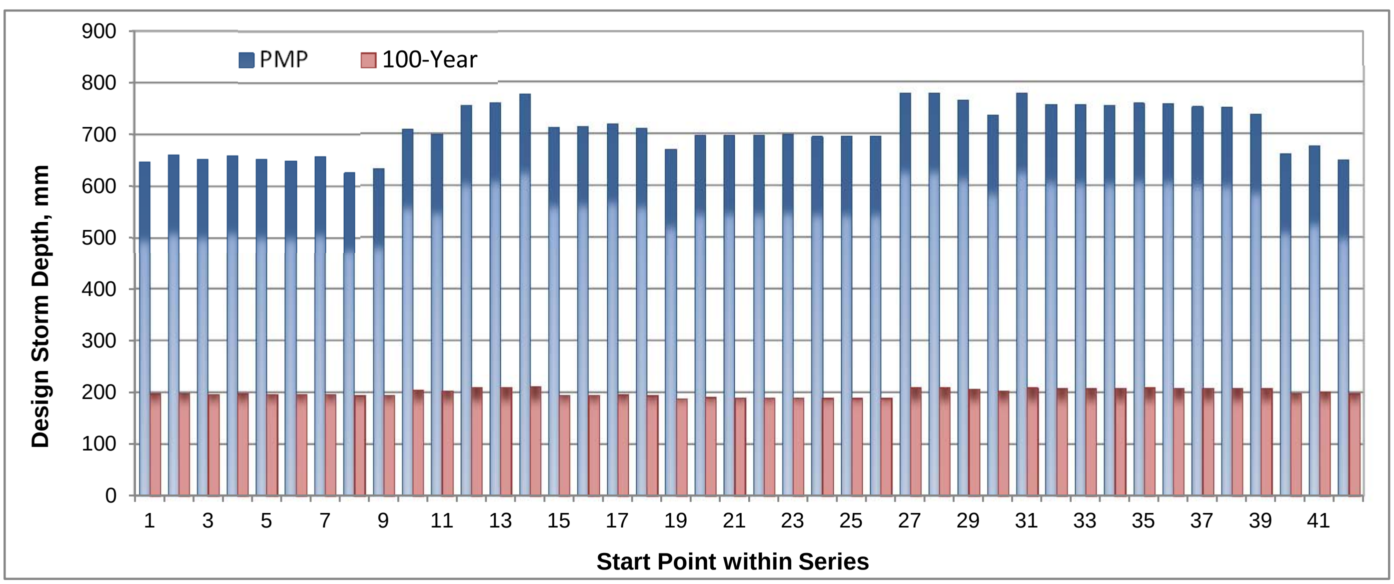

Figure B.4(c): Variation in PMP and 100-Year Estimates for Central America Record Resampled in 30-Year Segments