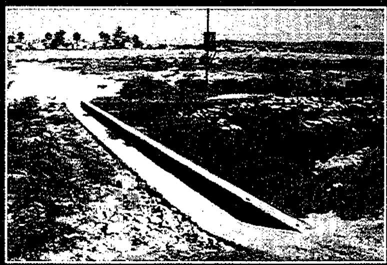


ANALYSIS OF EXTREME PEAK FLOWS FOR THE MAIN TIJUANA RIVER, SAN DIEGO, CALIFORNIA

FINAL REPORT

TASK ORDER #38
DOC ID# CSD-RT-12-URS38-01.F

JUNE 15, 2012

CITY OF SAN DIEGO

PREPARED BY URS CORPORATION

3.5.1 → 3.5.3

This page intentionally left blank

Table of Contents

List of Tables, Figures, and Appendices

Tables

Figures

Appendices

List of Acronyms and Abbreviations

USIBWC	United States International Boundary and Water Commission
SDRWQCB	San Diego Regional Water Quality Control Board
FEMA	Federal Emergency Management Agency
km	kilometer
yr	Year
m^3/s	cubic meter per second
cfs	cubic feet per second
USIACWD	United States Interagency Advisory Committee on Water Data
USGS	United States Geological Survey

This page intentionally left blank

EXECUTIVE SUMMARY

A critical component in understanding the primary sources of the sediment and trash impairments within the Tijuana River, is to understand the hydrology of the main Tijuana River and its final tributaries downstream of the United States and Mexico border. This study provides the base hydrology for the main Tijuana River that will be used in concurrent and future studies to optimize strategies for reducing the impacts of sediment and trash as they relate to flood control, pollutant reduction, ecological function, and added benefits of recreation and education activities.

The first step to evaluate the hydrologic characteristics of the main Tijuana River is to establish the peak flow for a given storm event return period that may be used for the design of water quality Best Management Practices (BMPs) and other activities. This study considers reviews of previous hydrology studies and data available and uses an alternative statistical approach to typical hydrologic analysis to estimate the peak flows for various storm event return periods.

A review of existing studies indicated that flows in the Tijuana River are largely controlled by dams. Approximately 73 percent of the Tijuana River watershed area is upstream of a dam. The Rodriguez Dam in Mexico, located about 18 km (11 miles) upstream of the border, has a significant impact on the runoff volume in the Tijuana River as it controls about 56% of the watershed. The Rodriguez Dam rarely releases water because it is used as water supply for a large portion of the City of Tijuana.

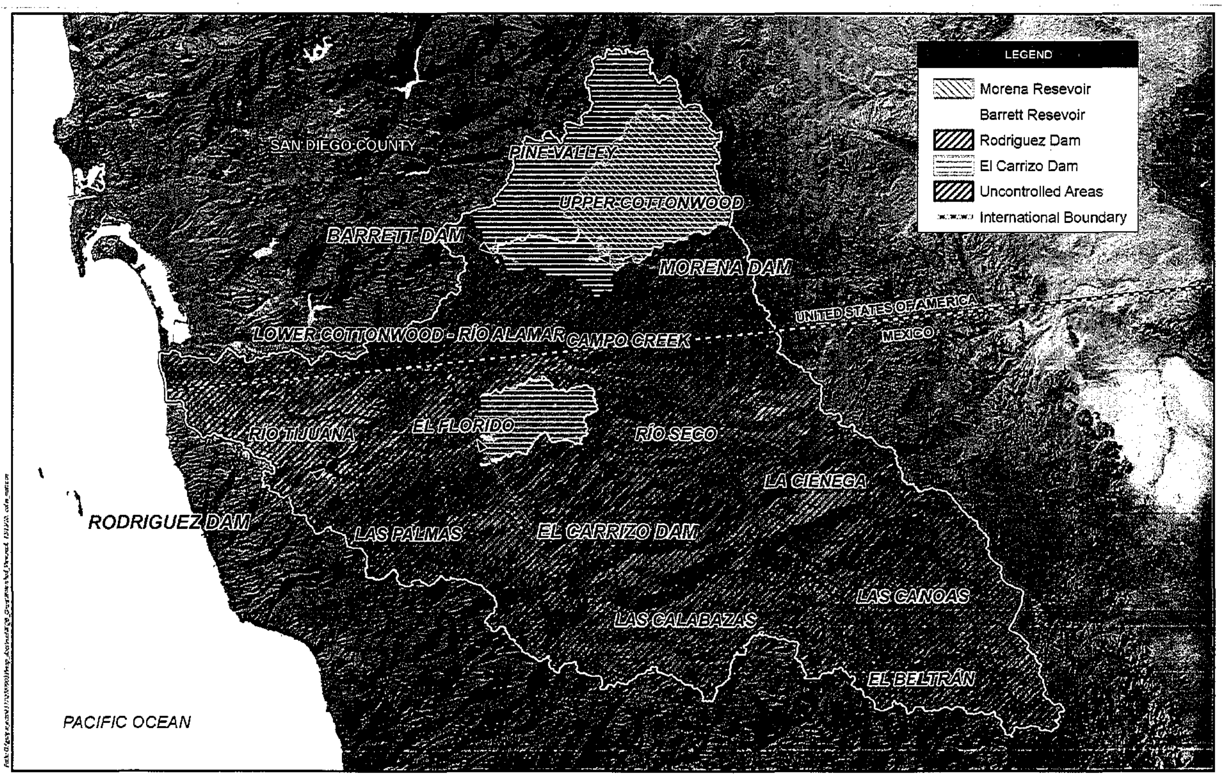


Figure ES-1. Tijuana River Watershed

The extreme peak flows for the Tijuana River were estimated in this study based on daily measurements made by the United States International Boundary and Water Commission (IBWC) during the last 48 years. The IBWC flow measurement station is located near the entrance of the Tijuana River [1] in U.S. territory. In addition, measurements made by the United States Geological Survey (USGS) from October 1936 until December 1981 for a total of 73 years of record were included in the analysis. The influence of historical flooding events [2] was also taken into account in the peak flow calculations.

Peak flows of the main Tijuana River were obtained and analyzed for a range of return periods. The return periods were selected based on the typical frequency values analyzed in the U.S. and Mexico for design purposes. The 2-, 5-, 10-, 25-, 50- and 100-year storm events are typical return periods associated with the Hydrology Manual of San Diego County [3]. A 200-year storm event is associated with a lower risk which is currently encouraged for urban level flood protection by the State of California Department of Water Resources [4]. The 500-year storm event is the largest peak flow analyzed by the Federal Emergency Management Agency (FEMA) for some river basins, and the 333-year and 1000-year storm events are typical design values for spillways and other structures used by Mexican agencies.

While the typical return periods were analyzed, a new statistical approach for estimating the peak flows for various return periods was utilized in this study. For record lengths of 50 years or more, a statistical analysis of flood frequency is the preferred method for peak flow estimation in lieu of a classical hydrology analysis or a stochastic-deterministic study [5, 6]. The length of the IBWC & USGS periods of record and quality of available data, allowed a detailed statistical-hydrologic frequency analysis to be conducted in order to characterize the extreme events in terms of peak flow, volume of runoff, and duration of the event. Typical return frequency analyses utilize statistical analyses of historic stream gage data utilizing a series of annual maximum peak floods to estimate return flood frequency values. The annual maxima method allows the use of only one peak flow per year. The current analysis uses an annual exceedance series which allows for analysis of multiple peak floods in one year. The use of annual exceedance extreme values rather than annual maxima values is advantageous as there are often more than one large peak flows during wet years. For the weather patterns over the past 80 years in the Tijuana River and southern California, the use of an annual maxima flood series underestimates the most frequent events (2-year and 5-year floods) because there may be multiple 'peak' floods within a given rainy season. Additionally, the largest 73 peak flows over the past 73 years have occurred in only 26 rainy seasons.

The major findings of this study indicate that:

- Peak flows for small return periods (5 years or less) are larger than in previous peak flood flow studies and that peak flows for large return periods (25 years and more) are lower than in previous studies (see Table 7-1). The use of a larger data set is also responsible for the modification of the peak flows at different return periods;

- The majority of the runoff volume occurs over relatively short periods of time. Only six months of data (February 1980, March 1983, January and February 1993, March 1980, and March 1995, from higher to lower runoff volume) equivalent to 1 percent of the time over 48 years of data account for nearly 53 percent of the total runoff that has been measured. Accordingly, BMPs and other flood control activities can be designed to address typical "average" storms and associated sediment/pollutant loading. However, larger storms that account for the majority of the runoff volume will still cause significant pollutant load inputs and flooding conditions in the Tijuana River Valley, because the size of the BMPs required to manage these large floods is infeasible from multiple perspectives.

The flow rates for the various return periods evaluated in this study should be considered in future analysis and design of stormwater quality BMPs for the main Tijuana River as part of the Tijuana River Valley Restoration Project. The next steps recommended include:

- performing similar hydrologic analyses for Smuggler's Gulch and Goat Canyon; and
- conducting hydraulic and sediment transport modeling using the results of this study.

Additional recommended future work includes:

Tijuana River

- Analysis of the trash load as a combination of floatables, suspended trash and bed load trash;
- One- and two- dimensional model of the sediment transport and floodplain;
- Feasibility Analysis of a trap for floatables in the main Tijuana River;
- Comparison of costs between source control in Mexico (trash collection in non-wet periods) versus implementation of structural trash capture BMPs in the United States;
- Evaluation of dam operations (Barrett and Morena) on contribution to sediment load;
- Evaluation of sediment basins in the United States upstream of the City of Tijuana and/or sediment basins or other BMPs in Mexico on trash and sediment loads; and
- Estimation of sediment loads of main United States tributaries draining into Mexico (Cottonwood Creek, others).

Smuggler's Gulch

- Continuous Hydrologic Simulation Model Analysis of Smuggler's Gulch;
- Refined sediment yield analysis: Trash-Sediment Proportion and Volumes Analysis;
- Determination of the sediment load under current conditions and under restored conditions to establish a cost effective target of sediment basins; and
- BMP efficiency and sensitivity analysis for sediment basins in Smuggler's Gulch.

Goat Canyon

- BMP efficiency evaluation of the existing Goat Canyon trash capture device and sediment basin; and
- Evaluation of additional upstream BMPs to further reduce sediment and trash loads to protect downstream habitat and reduce flooding.

This page intentionally left blank

SECTION 1 INTRODUCTION

The deposition of trash and sediment have been a chronic issue that adversely affects the physical and biological qualities of the riparian habitat and downstream estuary of the Tijuana River on the United States side of the international border. The analysis presented in this study is the first step in the process to analyze and optimize strategies to reduce the presence of sediment and trash in these areas and help begin the process to restore the hydrologic functions of the river. This report details the information available as well as documents the process used to estimate the hydrology basis for concurrent and future stormwater quality Best Management Practices (BMP) studies.

1.1 STUDY PURPOSE

The purpose of this study is to provide the Tijuana River Valley Restoration Project with a hydrologic analysis for the peak flows in the main Tijuana River for use in future stormwater quality studies. The Tijuana River watershed is approximately 1,724 square miles and over 70% of the watershed is in Mexico. Many studies have been completed in the Tijuana River watershed and were reviewed during this study; however, there is no study that provides sufficient hydrologic data analysis for the future stormwater quality BMP studies. Precipitation generates runoff but it is not the only factor to consider in the estimation of the peak flows. Initial watershed conditions and precipitation distribution play a large part in the volume and duration of the runoff and also in the sediment and trash transport abilities. This study establishes the 2-, 5-, 10-, 25-, 50- and 100-year storm return period flow rate results that will be used in future analysis and design of stormwater management project along the main Tijuana River.

1.2 BACKGROUND OF HYDROLOGIC AND HYDRAULIC STUDIES

Many studies have been conducted over the last 50 years concerning the peak flows for the Tijuana River. Unfortunately, most of those hydrologic results were not deemed appropriate for the hydrologic basis for the future BMP studies to be conducted. Six different studies were conducted by the US Army Corps of Engineers, Federal Emergency Management Agency (FEMA), and others since 1964 to try to estimate various aspects of the river's hydrologic features. The latest study was conducted in 1999 by FEMA to update the flood maps for the area. Additional changes to the physical characteristics of the watershed have occurred over the last 50 years including the construction of dams which now control approximately 73% of the watershed and channelization of the river in portions of the watershed which will impact the validity of past reports and provide useful information for this analysis. The results of these previous studies as well as over 70 years of historic flood data were compared to the results of this study validate the use of these peak flows in future studies.

1.3 BACKGROUND OF AVAILABLE DATA

Two sources of data were used in the daily peak flow time series analysis performed in this study. The United States Geological Survey (USGS) collected data from October 1936 through December 1981 and the United States International Boundary and Water Commission (USIBWC) collected data used for this study was dated from January 1962 through December 2009. While there were some inconsistencies between the data sets during the same time period, the two data sets were highly correlated. There were data gaps on days when stormwater runoff occurred and a linear interpolation was performed to account for the missing data. The number of missing days of data on accounted for 0.3 percent of total number of days. Section 3 details the analysis and selection process of the peak flow data used in this study to calculate the seventy-three year of peak flows.

1.4 PEAK FLOW ANALYSIS METHODOLOGY

The first step in the peak flow analysis was to estimate the characteristics of the runoff from the period of available data. Changes to the watershed such as construction of dams, channelization of the river, and development of the watershed were considered as well to account for physical watershed changes that may impact the results. The main statistical characteristic of the runoff data was reviewed and analyzed to understand the correlation between the rainfall and stream flow. Southern California is subject to cluster precipitation, flash flooding and extended droughts that result in a highly skewed runoff distribution when do not relate well to the typical water quality percentage of volume treated methodology. Using statistical analysis, different distributions were tested to evaluate the best statistical analysis method(s) to calculate the peak flows for the 2-, 5-, 10-, 25-, 50- and 100-year storm return period flow rates. The peak flow estimation methodology used in this study is detailed such that this statistical analysis may be duplicated in the future, as additional peak flow data becomes available.

SECTION 2 STATISTICAL MODEL JUSTIFICATION

According to the U.S. Interagency Advisory Committee on Water Data (USIACWD) [5], three types of analyses can be used in frequency studies of a river system: (1) a statistical analysis of the data, (2) a deterministic hydrological-hydraulic model of the watershed used to analyze extreme flood events from precipitation (classical hydrologic analysis), or (3) a comparison of data from similar catchments where a longer period of records exists. Also, an additional calibrated stochastic-deterministic model of the watershed can be constructed, where the objective is to reproduce the measured time series of the data through calibration of the hydrological model in order to predict possible scenarios of extreme events.

This report focuses only on the statistical analysis of the main Tijuana River data, from which a more than 99 percent complete time series of 73 years of average daily peak flows exists [1], and from which historic flooding events have been approximately estimated since 1885 [2]. A time series of this duration and the additional historical records can provide information in terms of peak flows, while the data can additionally provide information related to runoff volume and runoff duration.

A deterministic model of the watershed (classical hydrologic analysis) utilizing a 24-hr precipitation extreme event, land use, topology of the watershed, dam operation, infiltration assumptions and routing analysis was not conducted for this study because the duration of the time series data is such that better results can be obtained from a statistical analysis of the daily runoff. A classic hydrologic analysis of the entire Tijuana river watershed would have to rely on information obtained for each watershed sub-basin. The bi-national nature and size of the Tijuana River watershed limit the availability of sub-basin specific data such as:

- precipitation (quantity, duration and temporal and spatial distribution),
- topology of the watershed (network of tributaries),
- topography,
- land use,
- infiltration (in the watershed as a whole, and in the main channel during transmission), and
- geometry of the main water courses (including lateral and longitudinal storage of the main channels) and routing methodology.

Also, flow is controlled in more than 70% of the watershed by four dams: (1) Rodriguez Dam[1] in Mexico, which controls approximately 56% of the watershed; (2) El Carrizo Dam in Mexico, which controls 1.3% of the watershed; (3) Barrett Dam in the U.S. (controlling approximately 14% of the watershed) and (4) Morena Dam in the U.S., upstream of Barrett. Therefore, an understanding of the initial level of the reservoirs and the operation of the dams is crucial to estimating the resulting extreme event hydrographs

[1] The Abelardo Rodriguez Dam was designed in 1928, finished in 1936 (when the reservoir started to be filled) and inaugurated in 1940. It was originally constructed to satisfy the water needs of small City of Tijuana at the time. Currently Rodriguez Dam cannot satisfy the water demand of the City of Tijuana and it uses most of the water produced by the upstream portion of the watershed, except in very wet years (1978, 1980, 1983, 1993, 1995) when the dam was full and excess of water was discharged by the spillway. In some cases, a very wet year may find a very low initial level of the reservoir and no water release was considered by Mexican authorities (as in 1998 and 2005).

downstream in the Tijuana River valley on the U.S. side of the international border. However, appropriate data is not available currently for each of the dams. While the classical hydrological analysis requires this data, the statistical analysis used in this study does not require this level of information. Therefore, a classical hydrological analysis was not undertaken for this study.

A more complex stochastic-deterministic model was also not possible as it relies on the same data described above (e.g., detailed land use, precipitation data upstream of Rodriguez Dam) as well as other information not currently available. In order to reproduce the runoff time series in the stochastic-deterministic model, the following additional data would be needed to evaluate the effect on runoff: (1) evolution of the land use in the watershed in the past 50 years; (2) water use of the dams and discharges from the dams in the last 50 years (estimated through releasing modified hydrographs for extreme events, and using most of the runoff of smaller events); and (3) long-term simulation of the infiltration and subsurface water processes.

Following USIACWD [5, 6] for record lengths of 50 years or more, a statistical analysis of flood frequency was the preferred method for peak flow estimation. Therefore, for a record of 73 years, the statistical analysis of flood frequency remains the most applicable method to obtain estimate of peak flows.

SECTION 3 CHARACTERISTICS OF THE PEAK FLOW TIME SERIES

This section provides the characteristics of the peak flow time series including: 1) an analysis of the completeness of the data including period of record and data gaps associated with flow data from USGS and IBWC; 2) statistical analysis of the runoff time series data and summary of findings; 3) correlation between rainfall and runoff; and 4) significance of extreme flood events.

3.1 COMPLETENESS OF THE DATA

The data regarding the daily peak flow time series analyzed in this study dates from October 1, 1936 to December 31, 2009. In order to work with the same number of values for each year and eliminate seasonal variations, data was analyzed from January 1, 1937 to December 31, 2009. The data was collected by USGS from 10/1/1936 to 12/31/1981 and from IBWC from 1/1/1962 to 12/31/2009. Therefore, there were two sources of information from 1/1/1962 to 12/31/1981. For those years, data is extremely similar but IBWC data was preferred. Refer to Section 3.1.1 for the discussion about the selection of the IBWC during the overlapping time period. For peak analysis purposes, 73 years of data are analyzed from 1937 to 1961 (USGS) and from 1962 to 2009 (IBWC).

The USGS data set is shown graphically in Figures A.1-1 to A.1-13 (Appendix A). The IBWC data set is shown in Figures A.1-14 to A.1-37, while the graphic comparison between the data from 1/1/62 to 12/31/81 is shown in figures A.1-38 to A.1-42. The time series is largely complete. However, Table 3-1 provides a summary of those days where data was not collected or has been lost.

Table 3-1. Gaps in the Study Data Gathered

Time Period of Missing Data	Total Number of Days
January 1 to February 6, 1992	37
February 22 to March 2, 1992	10
February 3 to February 10, 1998 (data reconstructed with information from Winckell and LePage Study [8]).	8
March 20 to April 1, 1998	13
April 10 to April 15, 1998	6
Total Days of Missing Data	**74**
Total Days of Data to be Interpolated	**66**

The number of missing days is very small compared to the total number of days in the time series (74 out of 26,754 or 0.28 percent of missing days, of which eight have been reconstructed from data in reference [8]). However, during those missing days, runoff was taking place; therefore, the daily runoff calculations evaluated for this study do not include the actual runoff that occurred over 66 days of the total time period. Regarding those 66 days, it is known that during those days no releases of discharges from Barrett Dam [7] and Rodriguez Dam occurred, according to the maximum monthly levels in Barrett [7] and the years at which Rodriguez Dam released water [Daniel Sosa of CONAGUA in Mexico, personal communication]. Consequently, the runoff is associated with the uncontrolled areas of the watershed, and the occurrence of a large peak flow enhanced by the release of a dam is discarded.

Continuing with the analysis of the missing data, even though it is not possible to estimate the peak flows for the missing days, a linear interpolation was performed to account for these data points, with the exception of the February 3 to February 10, 1998 period. This period coincides with an extreme precipitation event measured in the city of Tijuana, and in South San Diego County. The results of a study conducted by Winckell and LePage [8] show the peak flows of the subbasins downstream of Rodriguez Dam during this time period. As it is known that Rodriguez Dam did not release water during the runoff event, an approximate daily average runoff hydrograph was constructed for those eight days according to the total runoff studied in [8] for the area. As a result, an additional significant peak flow was added to the series, improving its accuracy. It is also important to mention that according to Winckell and LePage [8], the event of February 3 - February 8, 1998 was the most important of that year. Therefore, the importance of the missing data of the other two periods in 1998 is less relevant than the importance of the data obtained according to [8].

A final analysis was made regarding the missing data: precipitation records at three locations (Rodriguez Dam, Chula Vista and Campo) were checked to verify that the precipitation in January, February and March of 1992 was not extraordinary. The average z value (normalized precipitation variable) for the three stations was $z = 0.1$, 1.5 and 1.4 respectively. Although February and March were more than one standard deviation above the average, they were not necessarily extreme, and more likely any possible peak flow lost in the missing data was not significant.

3.1.1 Selection of IBWC Data over USGS Data for the 1962-1981 Period

A comparison analysis between the IBWC and USGS data was performed for the 1962-1981 period (see comparison of data sets on figures A.1.-38 to A.1-42 of Appendix A. The data from IBWC was considered more accurate for the following reasons:

1) IBWC data does not display a calendar year or a season year where the total runoff was zero, which implies that the measurement of low flows seems more sensitive. The USGS data shows 1964 and 1971 with zero runoff. For season years, the 10/1/63-9/30/64 season also had zero runoff according to the USGS data.

2) The two data sets are highly correlated (correlation coefficient R=0.989). However, USGS data has an unexplained drop in runoff after 9/30/1978 which is the last day of the 77-78 water season. It seems that the USGS data at the end of the 77-78 season was measured against a different datum and then corrected at the beginning of the following season.

3) The USGS data from 1/1/1978 to 9/30/1978 also seems to be about 28 days displaced when compared with the IBWC data but the shape of the hydrographs is almost the same. Moving the data from the 77-78 season back 28 days (which deletes 28 zeroes) and completing the zeroes at the end of the season year, the correlation coefficient increases to 0.994.

4) The wettest year on record (1980) is almost identical for both data sets, which is an indication that the quality of the data is high for large events and almost identical for both measurements.

3.2 CHARACTERISTICS OF THE RUNOFF

The main statistical characteristics of the runoff are presented in Table 3-2.

Table 3-2. Statistical Characteristics of the Daily Runoff Data (1/1/1962-12/31/2009)

Characteristic of Runoff	All Days Considered from the IBWC Data	Only non-zero Days considered
Average $\mu^{(1)}$	1.96 m^3/s (69 cfs)	3.87 m^3/s (137 cfs)
Median M	0.01 m^3/s (0.35 cfs)	0.42 m^3/s (14.8 cfs)
Standard Deviation, $\sigma^{(1)}$	19.6 m^3/s (690 cfs)	27.4 m^3/s (968 cfs)
Skewness, γ	25.3	18.1
Coefficient of variation, $Cv = \sigma/\mu$	10.0	7.1
% of days when runoff is 0.	49.5 %	–

(1): In this work μ and σ are the population values and not the true values of those parameters.

A hydrological/statistical analysis of the IBWC runoff time series was performed. For this analysis USGS data was not considered since some precipitation records used for comparison purposes do not go back in time that long. The analysis revealed the following important findings.

1. The average daily peak flow (1.96 m^3/s or 69 cfs) is significantly different than the median peak flow for all days (0.01 m^3/s or 0.35 cfs). The average peak flow is only exceeded 9 percent of the time, which underscores the highly skewed nature of the Tijuana River discharges.

2. A significant amount of zero daily runoff existed for the Tijuana River, especially considering the magnitude of the watershed. On approximately 49.5 percent of the daily record, the average peak flow is 0 m^3/s.

3. On a significant number of days, runoff was lower than the threshold used to identify storms to be measured under the San Diego County Hydromodification Permit [9].
   - Using the threshold of 0.002 cfs/acre (equivalent to 0.0140 m^3/s / km^2), and ignoring the contributing area upstream of each dam for small runoff (to reduce even further the threshold peak flow), the hydromodication threshold peak flow was found to be 593 cfs (16.7 m^3/s). The uncontrolled area of the river is 465 mi^2, while the area upstream of dams is about 1259 mi^2, for a total area of the watershed of 1724 mi^2. If the total area of the watershed is used in the calculations, the peak flow threshold would be so high (2007 cfs or 62.5 m^3/s) that only 22 peaks can be found in the 48 year time series, rendering the analysis impossible.
   - On only 1.76 percent of the days, the runoff is actually higher than the hydromodification threshold using only the area uncontrolled by dams.

4. Runoff was significantly skewed for extreme events, even more so than precipitation.

- Regarding runoff, only six months of data (February 1980, March 1983, January and February 1993, March 1980 and March 1995, from higher to lower runoff volume)—equivalent to 1.04 percent of the time over 48 years—account for more than 50 percent of the total runoff that has been measured (52.8 percent to be exact).
- Regarding precipitation, the 55 wettest months (or 9.75 percent of the time) generated 50 percent of the total precipitation, using monthly precipitation data gathered in Mexico at the Rodriguez Dam for a similar time period (January 1962 to December 2008).

5. The runoff is even more skewed on a daily level than on a monthly level (note: daily precipitation data for Rodriguez Dam was not available for the daily comparison).
   - The 76 wettest days (0.44 percent of the time) account for 50 percent of the runoff volume.
   - As a comparison, the 237 wettest days at San Diego Lindberg Field Airport (1.38 percent of the time or 3 times more days than in the case of runoff) account for 50 percent of the precipitation volume.
6. Southern California runoff generation is subject to flash flooding and clusters of precipitation while at the same time extended drought periods occur. The drought phenomenon in the Tijuana River is represented in the time series by noticing that the number of months needed to account for 1 percent of the runoff volume is greater than the number of months needed to account for 1 percent of the precipitation. The larger skewness of the runoff compared with precipitation is remarkable, even with the understanding that extreme runoff events have a long tail causing runoff to continue after precipitation stops, and even considering the large number of months where no rain fell in the Tijuana River Watershed area.
   - From driest to wettest, the 335 driest months of runoff (59 percent of the time) accounted for 1 percent of the total runoff, while the driest 251 months of precipitation in Rodriguez Dam (44 percent of the time) accounted for 1 percent of the total precipitation.
   - These aspects of the data emphasize the importance of infiltration in dry climates for long-term runoff simulation: the number of months at which precipitation is recorded but no runoff is observed (months with low values of precipitation and/or dry antecedent moisture conditions) is greater than the number of months in which runoff is observed and precipitation is not recorded (months where the runoff is still occurring after precipitation has ceased).
7. A poor to moderate correlation exists between the dimensionless monthly precipitation (monthly precipitation divided by average monthly precipitation) and the dimensionless monthly runoff (monthly runoff divided by the average monthly runoff). The correlation coefficient between both series is only 0.46. This suggests that additional analysis may be needed to better understand the relationship between rainfall and runoff in the watershed. The average monthly dimensionless time series can be seen in Figures A.2-1 and A.2-2 (Appendix A). The average correlation plot can be seen in Figures A.3-1 to A.3-4 (Appendix A).
8. Based on observations made over the time series, four blocks of runoff (separated using a very low threshold of 0.0001 cfs/acre, equivalent to 0.84 m^3/s) correspond to clusters of precipitation

that account for more than 70 percent of the total runoff that has occurred in the Tijuana River Watershed in the last 48 yrs.

- These blocks are from January 29, 1980 to August 13, 1980 (24.6 % of runoff volume), from December 27, 1992 to June 25, 1993 (19.8 % of runoff volume), from January 24, 1983 to July 17, 1983 (18.8 % of runoff volume) and from January 1, 1995 to July 5, 1995 (7.2 % of runoff volume).
- Those four blocks are significantly different than the expected 24-hr 100-yr, 24-hr 50-yr, 24-hr 25-yr and 24-hr 10-yr extreme events calculated with SCS or similar methods using synthetic hydrographs, as the peak flows are not as high as the classical hydrology theory would suggest, but the duration and total volume of runoff is much larger than expected (see Figure 3-1 and Figure 3-2)
- Those four runoff periods coincide with 4 of the 5 times the Rodriguez Dam has released water in the last 48 years.
- The only other time the Rodriguez Dam has discharged water (1978) corresponds to the 7th largest runoff episode, after the very wet season of 2004/2005 and the very wet season of 1997/1998.

Rodriguez Dam has a significant effect in the runoff volume measured in the IBWC station; without the dam, the wet seasons of 1997/1998 and 2004/2005 would have been considerable larger in terms of runoff to the Tijuana River.

At this point, it is important to mention that wet years have occurred in the watershed even before the development and growth of the City of Tijuana. If we analyze the total runoff of the last 73 water years (October 1st to September 30) the seven wettest water years have been 1979-80, 1992-1993, 1982-1983, 1940-1941, 1994-1995, 1943-1944 and 1997-1998. It must be pointed out that for purposes of runoff volume analysis in the region, water years are a better indicator than calendar years. The use of calendar years may break a runoff event in two different periods as the end of December and the beginning of January are embedded in the middle of the rainy season.

Also, long-term drought can affect hydrologic analysis and perceptions in the river. From the 1944-1945 water year to the 1976-1977 water year, a 33 year drought period generated less runoff in those 33 years than the runoff from the previous wet year of 1943 to 1944. The drought was broken in 1978, and a wet period associated with El Niño followed. It is clear that urbanization was not responsible for the changes in hydrology, as very wet years occurred in the forties.

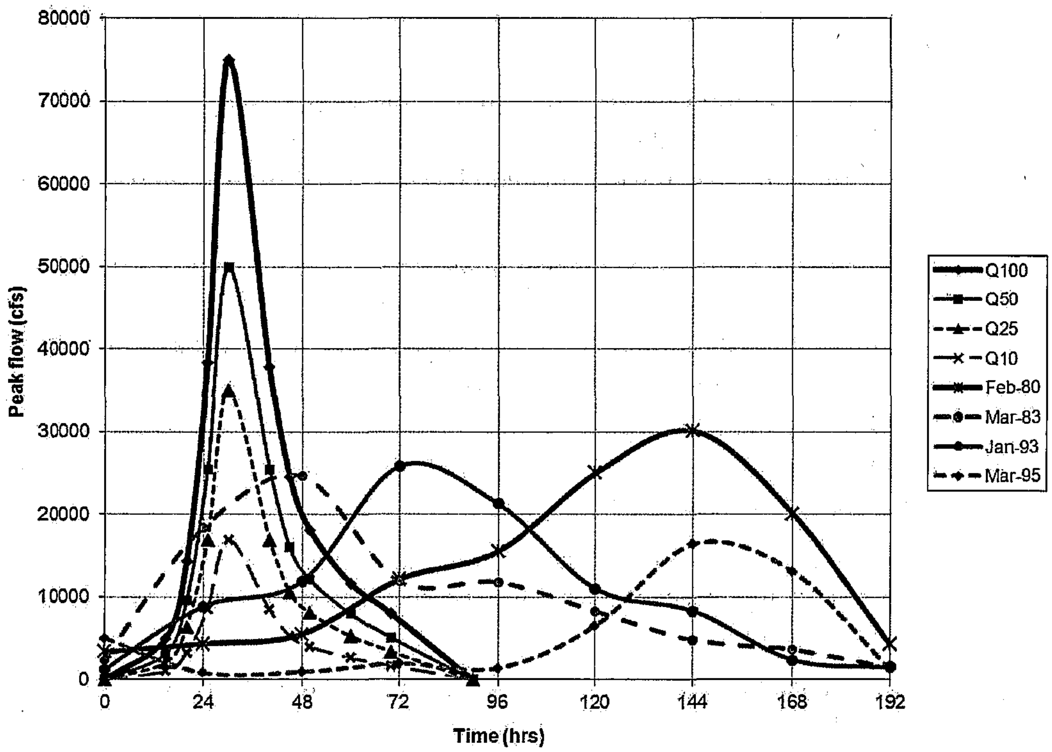


Figure 3-1. Theoretical[2] vs Measured Runoff (Most Intense 8 days for Largest 4 Runoff Events)

[2] Theoretical Runoff event results from the application of classic hydrologic methods (SCS in this case) as shown in the November 1994 report prepared by BSI consultant.

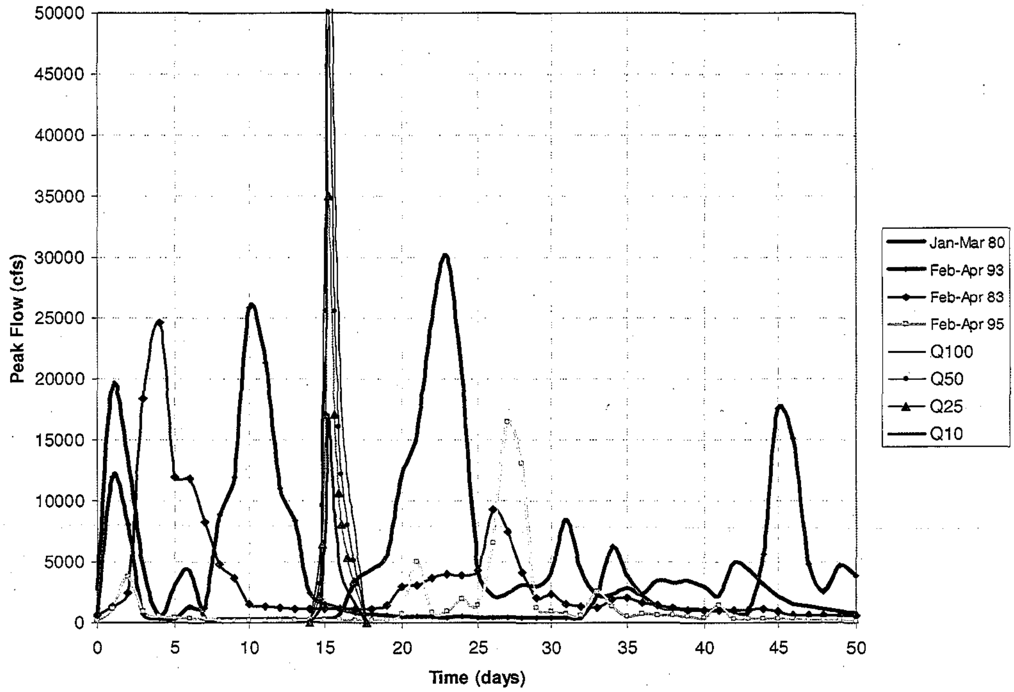


Figure 3-2. Theoretical[3] vs. Measured Runoff (Most Intense 50 days for Largest 4 Runoff Events)

3.2.1 Correlation between Rainfall and Runoff

The relatively low correlation between the dimensionless rainfall and the dimensionless runoff and the importance of the largest four events (Points 7 and 8 in the previous section) deserve some additional analysis. Although clear evidence exists of high runoff after high precipitation in the dimensionless comparison, the low correlation is tied to the lack of runoff between 1962 and 1977, compared to the rainfall and the significance of the extreme events of 1980, 1992-1993, 1983 and 1995, and the relatively low significance of the 1998 in terms of runoff but large significance in terms of precipitation.

Five contributing factors may explain the anomaly regarding the very low runoff during the period from 1962 to 1977. First, there is a tendency of moderate to extreme La Niña Years from 1962 to 1977 and it is well known that La Niña has a tendency to reduce the precipitation in the Southwest of the U.S. Second, dry patterns before 1977 may be related not only to relatively dry years but also to the period before the concrete channel was finished[4] and natural infiltration in the channel was occurring. Third, many isolated areas in the city of Tijuana were not connected to the natural Tijuana River but were later connected to the channel. Forth, urbanization in Tijuana during the 1960's and 1970's was significantly smaller than

[3] The horizontal scale change in Figure 3-2 was performed to demonstrate the influence of the runoff duration in the real scenario.

[4] The Tijuana River was channelized in Mexico to protect the City of Tijuana for flooding events and to limit the extension of the banks of the river to increase the area to be developed in the city. The first phase was completed in 1976, and successive phases were completed in the 80's. Currently the channel is 15.5 km long and extends from 2.5 km downstream of Rodriguez Dam up to the U.S.-Mexico border.

the 1980's to the present; therefore, fewer impervious areas existed. Fifth, discharges from Rodriguez Dam did not occur from 1962 to 1977, which could explain the lack of intermediate events during those years, as the portion of those events associated with the upper watershed were retained upstream of the dam.

The extreme runoff events of 1980, 1983, 1993 and 1995 are associated with the releases from Rodriguez Dam (those events would still be present without the dam), but the small volume of many other significant precipitation events is tied to the fact that Rodriguez Dam did not release water in any other year (except briefly in 1978). In particular, the extreme event of 1998 was significantly smaller in runoff importance than in precipitation significance, as probably a significant portion of the runoff was retained upstream of Rodriguez Dam, and unfortunately valuable runoff data is incomplete during February and March of 1998. It must be emphasized that Rodriguez Dam was not the cause of the extreme events, as the dam actually reduced the runoff volume from the extreme events most of the time; however, the dam was responsible for reducing the peak and volume of many intermediate events that otherwise would have reached the Tijuana Valley without the dam. As stated above, it is worth mentioning that a small release from Rodriguez Dam also occurred in 1978. This release may have been related to testing of the concrete channel for relatively high flows as the first construction phase of the channel was completed a few months before the release.

The extreme runoff events mentioned before are also significant in terms of duration and volume of runoff, and are not associated with a 24-hr storm event, but more likely with a cluster of many short-duration storms which could had occurred when water soil conditions were near saturation levels. In particular, the initial level of the dams could have been higher than normal. Figure 3-1 and Figure 3-2 show a brief comparison of those events and the 24-hr extreme events calculated in reference [2] for the Tijuana River watershed.

Finally, the accelerated urbanization of the City of Tijuana in the last 20 to 30 years and the inauguration of El Carrizo Dam in 1978 (which currently adds about 50,000 to 70,000 acre-ft/yr [10] to the Tijuana watershed as water consumption in the City of Tijuana) may help to explain why extremely low flow periods, such as the one from 1970 to 1977, have otherwise not been observed. Urbanization has resulted in increased runoff in the Tijuana area for three main reasons: (1) a large amount of water imported from the Colorado River watershed (and controlled by El Carrizo Dam since 1978) has been used for human consumption and added to the base flow of the river; (2) the concrete channel has reduced infiltration in the main channel; and (3) impervious areas of the City have reduced natural infiltration at the lower end of the watershed.

As rainfall is not as sensitive to anthropogenic effects as runoff is, runoff has a poor to moderate correlation with precipitation, due to the factors previously explained, that may be summarized as follows:

- The long drought from 1962 to 1977 is correlated to moderate to strong occurrences of La Niña during the same period (observations of USGS records indicate that the drought started as far back as 1945).
- Infiltration occurring in the main channel reduced baseflow runoff prior to 1976.

- Urbanization of Tijuana increased sharply starting in the 1980's after the channel was completed and El Carrizo Dam was in operation. Also, many areas of the city unconnected to the main river were connected to the concrete channel completed in 1976.
- Additional baseflow runoff added to the river from water consumption supplied by the Colorado River through El Carrizo Dam has increased since 1978.
- Additional peak flow runoff added as the imperviousness of the lower watershed has increased with urbanization.
- Rodriguez Dam retained runoff on intermediate event, and only releases runoff few times during extreme events (and never from 1962 to 1977).
- Clusters of events in favorable initial conditions for runoff (high level in dams, especially Rodriguez Dam, high levels of water content in the soils) generate a long runoff response during the four most significant runoff periods in 1980, 1983, 1993 and 1995.
- Operation problems in Rodriguez Dam could also contribute to increase the levels of runoff from an already high condition (as occurred in 1980 with the spillway gates that were unable to operate until the Dam was almost overtopped and they finally open and released the volume accumulated. Evidence of this situation can be found in [11]). It is important to mention that in 1980 the opening of Rodriguez Dam was coincident with the opening of the gates at Barrett Dam.

3.2.2 Percentiles, Volume of Runoff and Significance of Extreme Events

The highly skewed nature of the runoff distribution of the Tijuana River made meaningless the typical definitions of 85th percentile, 95th percentile and similar percentile analysis, as the peak flows will carry most of the runoff, and consequently trash and sediment load, for this watershed. The following section describes the theory of percentiles in order to understand runoff distribution in the Tijuana River Valley. Percentile analysis is only going to be applied to the last 48 years of records (IBWC Data).

Percentile Theory

The Zth percentile peak flow is the flow that exceeds Z% of the daily flows, and therefore is exceeded 100 minus Z percent of the time (typical values of Zth are 85th and 95th percentile). As there is not a clear definition of percentiles for runoff measurements (typical percentile analysis was developed for rainfall analysis), if the thousands of zero values in the time series are included, then the Zth percentile is the flow that is exceeded (100-Z) % of the time. If zeroes are not included, then the Xth percentile represents the flow that is exceeded (100 –X) % of the time when runoff is measured (runoff is reported in the Tijuana River when an average daily peak flow of at least 0.01 m^3/s or 0.35 cfs is measured). As the percentage of zero values is known (P_Z = 49.46%), the Zth percentile when zeroes are included can be related with the Xth percentile when zeroes are not included according to:

$$X^{th} = \frac{100}{100 - P_Z}\left(Z^{th} - P_Z\right) \tag{1}$$

If the peak flows are sorted from high to low, some relationships are important from the water quality point of view:

1. The magnitude of the percentile events (median or 50th percentile, upper third or 66.7th percentile, upper quartile or 75th percentile, 80th percentile, 85th percentile, 90th percentile and 95th percentile are typical values that could be needed). Those percentiles could be needed with or without consideration of the zero values.
2. The percentage of runoff that occurs when the given daily peak flow Q is larger than a given percentile peak flow (for example, the percentage of runoff occurring when the peak flow exceeds the 85th percentile value Q_{85}). This percentage volume can be denoted as $\mathbf{V\%_{Q>Qz}}$
3. The percentage of runoff occurring in excess of a given peak flow. This value represents the percentage of runoff by-passed by a flow-based BMP and it must not be confused with the previous definition as it is smaller. It represents the percentage of runoff occurring as an excess over the given percentile peak flow. This volume percentage can be denoted as $\mathbf{V\%_{Q-Qz}}$

Let Q_Z be the peak flow in m^3/s with a Z^{th} percentile value (remember that Z^{th} include the zero values). Let $Q_{ave,z}$ be the daily average peak flow for all days (1.96 m^3/s in our case).

From the mathematical point of view, only two of the previous three concepts (Q_Z, $V\%_{Q>Qz}$, $V\%_{Q-Qz}$) are independent. The following relation is valid per the definition of the variables involved:

$$V\%_{Q-Qz} + \frac{Q_Z}{Q_{ave,Z}}\left(100 - Z^{th}\right) = V\%_{Q>Qz} \qquad (2.a)$$

As Q_Z has also a X^{th} percentile defined by (1) and consequently can also be called Q_X, the following relation is also valid (with $V\%_{Q-Qz} = V\%_{Q-Qx}$ and $V\%_{Q>Qz} = V\%_{Q>Qx}$ as Q_Z and Q_X are the same):

$$V\%_{Q-Qx} + \frac{Q_X}{Q_{ave,X}}\left(100 - X^{th}\right) = V\%_{Q>Qx} \qquad (2.b)$$

The daily average without including zeroes $Q_{ave,X}$ is related to $Q_{ave,Z}$ by:

$$Q_{ave,X} = \frac{100}{100 - P_Z} Q_{ave,Z} \qquad (3)$$

Table 3-3 displays the relationship between peak flows, percentiles, percentage of runoff larger than a given peak flow, and percentage of runoff in excess of a given peak flow.

Table 3-3. Peak flows for the Typical Zth Percentiles and Xth Percentiles and Percentage of Volume Larger Than or in Excess of Those Peaks

Peak Flow "Q_i" (m^3/s)	Zth Percentile (includes zeroes)	Xth Percentile (only values $\neq 0$)	$V\%_{Q>Qi}$	$V\%_{Q-Qi}$
10.8	97.47	***95.00***	75.92	62.03
4.41	***95.00***	90.10	84.55	73.28
4.3	94.95	***90.00***	84.65	73.56
2.49	92.42	***85.00***	88.85	79.23
1.73	***90.00***	80.21	91.43	82.61
1.69	89.89	***80.00***	91.5	82.81
1.2	87.37	***75.00***	93.4	85.66
0.92	***85.00***	70.32	94.65	87.61
0.79	83.16	***66.67***	95.44	88.67
0.58	***80.00***	60.42	96.55	90.64
0.42	***75.00***	50.54	97.81	92.47
0.42	74.73	***50.00***	97.81	92.47
0.23	***66.67***	34.05	99.20	95.30
0.01	***50.00***	1.06	99.99	99.74

The concepts explained in equations (1) to (3) are better illustrated with an example using the runoff data of the Tijuana River. For a given value of the daily peak flow $Q_i = 10$ m^3/s, approximately 469 values are larger than it. As there are 17532 days of measurements, if zeroes are included that value represents the $Z^{th} = 97.3^{th}$ percentile, or the $X^{th} = 94.7^{th}$ percentile if zeroes are not included (there are $0.5034 \cdot 17532 = 8861$ days when some runoff was measured; X^{th} could also be determined with equation (1)). The percentage of runoff that has occurred with daily average peak flows that are larger than 10 m^3/s is 76.8 % ($V\%_{Q>Qi}$). Finally, the percentage of runoff that has occurred *in excess of* 10 m^3/s ($V\%_{Q-Qi}$) is 63.1% of the total runoff which by definition has to be smaller than the previous value of 76.8% because 10 m^3/s has been subtracted from each flow larger than such value. Even though the 63.1% value was obtained from the data, it could have been determined with equation (2.a) or (2.b):

(2.a): $V\%_{Q-Qz} = 76.8 - (10/1.96)\cdot(100-97.3) = 63.0\%$ (difference due to rounding).

(2.b): $V\%_{Q-Qx} = 76.8 - (10/3.87)\cdot(100-94.7) = 63.1\%$

In addition to the values presented in Table 3-3, one could be interested in determining the peak flow value such that a given round percentage of runoff is larger than it (or a given round percentage of runoff has occurred *in excess of* it). Table 3-4 displays the peak flow values such that a given percentage of runoff is larger than the calculated peak flow, or a given percentage of runoff occurs in excess of a given peak flow. For example, if one is interested in knowing the peak flow value such that 20% of the volume occurs under peaks larger than it, we would look at $V\%_{Q>Qi} = 20$ in Table 3-4, and read the value $Q_i = 443$

m^3/s to the left. This value has a 99.94th percentile if zeroes are included (99.88th percentile if only measurements when $Q \neq 0$ are included). If someone is interested in the peak value such that 33.3% of the volume occurs in excess of it, he/she would look at $V\%_{Q>Qj} = 33.3$, and find the Q_j value of 76.2 m^3/s which have a Zth and Xth percentiles of 99.60 and 99.20 respectively. The four columns on the left side of Table 3-4 correspond to percentage of volume larger than a given peak, and the four columns on the right side correspond to the percentage of volume of runoff in excess of a given peak flow.

Table 3-4. Peak Flows for Typical Values of Volume Percentages Larger than or In Excess of the Given Peak Flow

Peak Flow "Q_i" (m^3/s)	Zth of Q_i	Xth of Q_i	$V\%_{Q>Qi}$ (% of volume > Q_i)	Peak Flow "Q_j" (m^3/s)	Zth of Q_j	Xth of Q_j	$V\%_{Q-Qj}$ (% of volume in excess of Q_j)
710	99.99	99.98	5	447	99.94	99.89	5
570	99.97	99.94	10	322	99.88	99.77	10
520	99.95	99.91	15	242	99.87	99.74	15
430	99.94	99.88	20	180	99.83	99.66	20
340	99.90	99.81	25	128	99.77	99.55	25
230	99.85	99.71	33.3	76.2	99.60	99.20	33.3
69	99.57	99.15	50	24.5	98.82	97.66	50
21.4	98.65	97.33	66.7	7.64	96.66	93.40	66.7
12.2	97.64	95.34	75	3.77	94.36	88.84	75
7.44	96.58	93.24	80	2.30	91.85	83.88	80
4.13	94.79	89.69	85	1.30	87.92	76.09	85
2.15	91.49	83.17	90	0.64	81.14	62.68	90
0.86	84.21	68.76	95	0.25	67.96	36.60	95
0.26	68.27	37.22	99	0.04	55.18	11.31	99

From the analysis of the previous tables, it is evident that extreme events are paramount for understanding the behavior of the Tijuana River, and how typical water quality peak flows (like twice the 85th percentile peak flow, for example) have little meaning in the percentage of volume treated. If the 85th percentile peak flow (without including zeroes) is 2.49 m^3/s from Table 3-3, twice that value is 4.98 m^3/s. Such peak flow has a percentage of volume larger than the peak flow of about 84% and a percentage of volume in excess of the peak flow of about 72% (from interpolations of values in Table 3-3 or Table 3-4). Therefore, a flow-based BMP designed with a peak flow of twice the 85th percentile value will bypass 72% of the runoff (and 62% bypassed for a 95th percentile peak flow).

Further analysis will be discussed in the report *"Estimate of Sediment Yield of the Tijuana River"*. As the extreme events are important in terms of the percentage of runoff volume they carry, they are even more important in terms of bed load sediment transport.

SECTION 4 EXTREME PEAK FLOW DETERMINATION

From the time series analysis of the data, values of the 2-yr, 5-yr, 10-yr, 25-yr, 50-yr, 100-yr, 200-yr, 333-yr, 500-yr and 1000-yr peak flows were estimated using frequency analysis techniques. The "T-yr" peak flow corresponds to the peak flow that, on average, is equaled or exceeded only once every T years. The definition does not mean that events will occur on a clockwise basis every "T" years, but, in the statistical sense, it means that in an infinitely large time series there will be *on average* one "T-yr" extreme event every T years.

Two series of extremes are useful in determining peak flows: (1) the series of maximum annual floods (annual maxima series), assuming that such floods are independent, and (2) the series of the "n" largest independent values in an "n" year duration time series (annual exceedances series). In a weather pattern like the one that exists in the Tijuana River watershed, a series of annual maxima greatly underestimates the most frequent extreme events (2-yr and 5-yr events) as many independent events that occur in a wet year are not analyzed, even though they may be much larger than a maximum annual value of a dry year. For example, the largest 48 peak flow values in the IBWC time series have occurred in about 18 calendar years (or 16 seasons) depending on the threshold used to define independent events. Similarly, the largest 73 peak flows (including the USGS series) have occurred in 27 years (26 seasons).

Similar behavior has been observed in other watersheds in southern California, even in watersheds with little anthropogenic intervention. Thirty years of reliable runoff data [12] (1953 to 1967 and 1994 to 2009) exists in the San Mateo Creek watershed in Northern San Diego County (this watershed is one of the last remaining pristine watersheds in southern California). The 30 largest peak flows have occurred in only 12 calendar years (11 seasons), depending on the criteria used to separate independent events. Although additional data would be needed to generalize runoff behavior in southern California, the "n" largest values in a daily runoff series of n years of duration appear to occur during 30 percent to 40 percent of the total years.

Also, from the theoretical point of view and for independent events, as the number of years of data increases, it becomes more and more unlikely that the largest "n" events are evenly distributed in the n years. For example, assume that in 3 years the three largest independent peak flows are going to be distributed. Twenty seven combinations exists to accommodate the 3 values (3 events in 3 years: 3^3 total possibilities) but only 6 combinations occur with an event each year (3! possibilities). Therefore, the probability of the 3 independent events occurring in three years is $3!/3^3 = 0.222$ or 22.2 percent. As the number of years increases, it becomes more and more unlikely to have the biggest "n" independent events in exactly "n" different years as the probability P of such occurrence can be written as $P = n!/n^n$. For n larger than 6, the probability is less than 1 percent, and for n larger than 8, the probability is less than 0.1 percent. For n = 48 (48 years of data for the main Tijuana River), the probability of the 48 largest independent events occurring in 48 different years is practically zero: $48!/48^{48} = 2.48\text{x}10^{-20}$ percent. Therefore, even in a theoretically independent example, there is no possibility of having the largest n events in n years for a large value of n, which suggests not using annual maxima series but rather annual exceedances series. This conclusion is also suggested in [13].

It should be pointed out that, theoretically, the distribution of the largest n independent events in n years has a tendency to follow a normal distribution with mean around 0.63n. This means that if events are truly independent, the n events should be expected to occur around 50 percent to 75 percent of the total "n" years. As this is not the case in Southern California (and is not the case for most runoff time series), it is clear that in terms of occurrence, extreme events are not completely independent but linked to the existing auto-correlation memory in the time series. Therefore, long-term persistence is present and complete independence among extreme events does not exist, which is one of the characteristics of fractal time series [14].

In any case, extreme events are more adequately calculated using the n maximum independent values instead of the n annual maxima series. Table 4-1 shows the extreme peak flows used for this analysis under the different methods: the first five columns are annual exceedance series, the last column displays seasonal maxima, and the second to last column displays calendar year maxima.

4.1 ANNUAL MAXIMA DATA

To prove the difference between the annual maxima and the n largest events, the 73 annual maxima values are shown in Table 4-1, both in terms of calendar year and seasonal year (i.e., starting on October 1st and ending on September 30th of the next calendar year). Those values will not be used for the purpose of extreme event calculation, but are shown in Table 4-1 to indicate the significant difference between the annual peak flows and the maximum peak flow of the largest 73 independent events. In particular, Table 4-1 demonstrates how the mean value of the largest events is much smaller for calendar or seasonal extremes than for the largest n event obtained with different methods (this will be explained in detail in the following sections). It is also significant to note that in the annual maxima series, 11 seasons out of 73 are not included in the list of the "n" annual larger values, because the extreme event of a given year "x" occurs at the beginning of the year and the extreme event of the next year occurs at the end of the year "x+1" and consequently the season occurring from 10/1 of year "x" to 9/30 of year "x+1" (season "x/(x+1)") does not contribute to the maxima annual series. Similarly, in the seasonal maxima series, 8 years out of 73 are not included in the list of the "n" seasonal larger values (see Table 4-1). This occurs when the season's maximum value "(x-1)/x" belongs to the year "x-1" and the maximum value of the following season "x/(x+1)" belongs to the year "x+1", leaving the calendar year "x" out of the maxima seasonal series.

4.2 INDEPENDENCE OF EVENTS IN AN ANNUAL EXCEEDANCE SERIES

This study has shown that the use of the annual exceedance series (the series that selects the largest n events in n years, regardless of the selection of those events in fewer than n years) is preferred over the annual maxima series, but it is still not clear how to define independence of the events. When analyzing annual maxima, it is assumed that the maximum annual (or seasonal) peak flows are independent and no additional analysis is needed. The only potential problem is the unlikely scenario where the maximum flow in a given year happens to occur in the last few days of the year, and the maximum flow of the following year happens to occur during the first days of the next year (therefore, both values could be related to the same event). However, when calculating the annual exceedence series if the maximum "n" values of an n year time series are selected without using any criteria (by simply sorting the data), in most instances some of those extreme values are not independent and belong to few of the largest events. Therefore, for a daily series, the largest "n" values of a time series are not equal to the largest "n" independent values and some sort of criteria must be established to guarantee the independence of the selected values.

In runoff time series, an established general criterion of independence does not exist. Usually, two methods can be used to separate the largest peaks: (1) a peak flow threshold can be used to separate storms; only the largest peak at each block of runoff is determined regardless of how many peaks are present; (2) the mathematical criteria of maxima ($dy/dx = 0$) and a threshold of duration can be used to ensure that the value is not due to an oscillation within a given storm. Both criteria will be discussed below.

4.2.1 Peak Flow Threshold

The San Diego Hydromodification Permit requires the use of the peak flow threshold criterion [9]. In an hourly time series, a storm event is considered separate from the next event once the peak flow falls below a threshold value for at least 24 hours, and the threshold value is associated with the contributing area of the watershed at the point where the measurements are taken. The typical threshold is 0.002 cfs/acre in a 24-hr period. Therefore, if the contributing area is 100 acres, for example, once the peak flow goes below 0.2 cfs for 24 hrs or more, a new storm event is defined once the peak goes above 0.2 cfs.

The Tijuana River data shows that this criterion is relatively arbitrary. Assuming that the area upstream of the four existing dams was not considered part of the contributing area at the measurement station (because otherwise the peak flow threshold would be too large to be useful), the contributing area was about 1,200 square kilometers (km^2), and the threshold to separate storms was 16.8 m^3/s (593 cfs). It can be seen in Figures A.1-1 to A.1-37 (Appendix A) that many runoff events were left out of the analysis since they were not large enough to reach the required threshold. Also, and more importantly, many storm blocks (for example, from March 24 to April 5, 1979) were divided arbitrarily into many events, while other clearly independent storm events (for example, February 27 to April 16, 1983) were treated as one, as the peak flow never dropped below the threshold because the initial storm was too large and the second arrived before such a small discharge was reached. The 71 largest peak flows obtained using the Hydromodification Permit criteria are shown in Table 4-1. The use of the Hydromodification threshold criteria does not allow to find 73 peaks but only 71 meaning that a smaller threshold needs to be used. Results obtained using this method are for illustration purposes only, and will not be used in the determination of independent events due to the problems explained in this section and that can be seen in the time series graphics of appendix A.

Table 4-1. List of Extreme Daily Peak Flows for the Methods Analyzed.

Date	Threshold Duration:				Hydro-Mod	Calendar	Annual
	4 days	3 days	2 days	1 day	(0.002 cfs/acre)**	year	Season*
	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)
02/07/1937	275	275	275	275	275	275	275
02/15/1937	37.4	37.4	37.4	37.4	37.4		
03/23/1937	20.6	20.6			20.6		
03/03/1938	140	140	140	140	140	140	140
03/13/1938	43.6	43.6	43.6	43.6	43.6		
02/04/1939	30.6	30.6	30.6		30.6	30.6	30.6
02/05/1940							2.44
12/24/1940	50.7	50.7	50.7	50.7	50.7	50.7	
02/22/1941	111.0	111.0	111.0	111.0	111.0		
03/01/1941				47.86			
03/06/1941	130	130	130	130	130		
03/15/1941	194	194	194	194	194		
03/18/1941			117	117			
04/02/1941	195	195	195	195			
04/12/1941	295	295	295	295	295	295	295
04/22/1941			40.8	40.8			
04/28/1941	115.8	115.8	115.8	115.8			
05/03/1941		33.7	33.7	33.7	33.7		
03/17/1942	72.5	72.5	72.5	72.5	72.5	72.5	72.5
03/05/1943	22.9	22.9			22.9	22.9	22.9
02/24/1944	286	286	286	286	286	286	286
02/27/1944				132			
03/02/1944			53.5	53.5			
03/07/1944	60.6	60.6	60.6	60.6	60.6		
03/20/1944	54.1	54.1	54.1	54.1	54.1		
03/24/1945							7.84
12/23/1945	28.9	28.9	28.9		28.9	28.9	
01/06/1946						4.36	
11/14/1946							0.79
12/05/1947						4.87	4.87
02/06/1948						0.85	0.85
01/13/1949	32.3	32.3	32.3	32.3	32.3	32.3	32.3
02/07/1950						0.31	0.31
50-1 season							0
Year: 1951						0	
01/18/1952	24.9	24.9	24.9		24.9		
03/17/1952	51.0	51.0	51.0	51.0	51.0	51.0	51.0
52-3 season							0
Year: 1953						0	
03/23/1954						11.78	11.78
54-5 season							0
Year: 1955						0	
55-6 season							0
Year: 1956						0	
01/30/1957						0.28	0.28
04/08/1958						10.48	10.48
02/21/1959						1.02	1.02
01/15/1960						0.51	0.51
60-1 season							0
Year: 1961						0	

Date	Threshold Duration: 4 days	Threshold Duration: 3 days	Threshold Duration: 2 days	Threshold Duration: 1 day	Hydro-Mod (0.002 cfs/acre)**	Calendar year	Annual Season*
	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)
01/21/1962						1.98	
03/17/1963						1.78	1.78
11/20/1963							0.8
01/22/1964						0.5	
04/08/1965							1.43
11/23/1965						8.61	8.61
12/07/1966	29.7	29.7	29.7		29.7	29.7	29.7
12/19/1967						3.77	3.77
03/08/1968						1.76	
02/26/1969						12.8	12.8
03/01/1970							1.64
12/21/1970						2.1	2.1
12/27/1971						0.48	0.48
12/04/1972						1.34	
03/12/1973						2.74	2.74
03/08/1974						1.06	1.06
04/09/1975						2.81	2.81
02/09/1976						2.83	2.83
01/06/1977						2.73	2.73
01/15/1978					27.1		
01/17/1978	45	45	45	45	45		
02/13/1978					18.7		
03/01/1978	92.6	92.6	92.6	92.6	92.6	92.6	92.6
03/05/1978				41.1			
03/12/1978	28.9	28.9	28.9		28.9		
03/24/1979					17.2		
03/29/1979	37.4	37.4	37.4	37.4	37.4	37.4	37.4
04/05/1979					18.5		
01/30/1980	547	547	547	547	547		
02/04/1980			120	120	120		
02/21/1980	852	852	852	852	852	852	852
02/26/1980				86.9			
02/29/1980	241	241	241	241			
03/04/1980				80.1			
03/06/1980				97.1			
03/08/1980			98	98			
03/11/1980	140	140	140	140			
03/30/1980	73.9	73.9	73.9	73.9			
04/18/1980	20.7	20.7					
04/26/1980		22.4					
04/30/1980	24.1	24.1	24.1		24.1		
03/02/1981						6.4	6.4
03/18/1982	25.7	25.7	25.7		25.7	25.7	25.7
01/29/1983					20.4		
02/03/1983	41.9	41.9	41.9	41.9	41.9		
02/10/1983	41.9	41.9	41.9	41.9	41.9		
03/03/1983	697	697	697	697	697	697	697
03/22/1983				114			
03/25/1983	263	263	263	263			
03/29/1983				65.1			
04/03/1983	58.3	58.3	58.3	58.3			
04/12/1983	31.4	31.4	31.4				

Date	Threshold Duration:				Hydro-Mod	Calendar	Annual
	4 days	3 days	2 days	1 day	(0.002 cfs/acre)**	year	Season*
	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)
04/23/1983	38.8	38.8	38.8	38.8			
05/01/1983	43.9	43.9	43.9	43.9	43.9		
08/15/1983	22.7	22.7			22.7		22.7
12/27/1984						14.4	14.4
11/25/1985	20.4				20.4	20.4	
02/15/1986	26.2	26.2	26.2		26.2	26.2	26.2
10/10/1986							9.51
12/17/1987	25.0	25.0	25.0		25.0	25.0	
01/18/1988	47.3	47.3	47.3	47.3	47.3	47.3	47.3
02/03/1988	37.7	37.7	37.7	37.7	37.7		
12/25/1988							11.6
03/25/1989						6.14	
01/17/1990							5.30
12/20/1990						7.19	
03/01/1991	23.3	23.3	23.3		23.3	23.3	23.3
03/21/1991					17.5		
02/13/1992	20.1				20.1	20.1	20.1
01/07/1993	337	337	337	337	337		
01/12/1993				38.8			
01/16/1993	731	731	731	731	731	731	731
02/09/1993	176	176	176	176			
02/20/1993	496	496	496	496	496		
02/24/1993				134			
02/27/1993				63.4			
03/02/1993				45.6			
03/05/1993	68.8	68.8	68.8	68.8			
02/07/1994	26.9	26.9	26.9		26.9	26.9	26.9
01/05/1995					18.5		
01/26/1995	34.3	34.3	34.3	34.3	34.3		
02/15/1995	106	106	106	106	106		
03/06/1995	141	141	141	141			
03/09/1995				54.9			
03/12/1995	464	464	464	464	464	464	464
03/18/1995	73.6	73.6	73.6	73.6			
03/21/1995					21.2		
03/26/1995	39.9	39.9	39.9	39.9	39.9		
03/14/1996						4.47	4.47
01/26/1997							3.99
12/07/1997						5.37	
02/08/1998	112	112	112	112	112	112	112
02/23/1998	25.0	25.0	25.0		25.0		
03/07/1998	33.5	33.5	33.5	33.5	33.5		
04/06/1998	45.3	45.3	45.3	45.3	45.3		
01/27/1999						1.51	1.51
02/21/2000						3.88	3.88
01/11/2001						10	10
12/22/2001							1.54
12/20/2002						3.05	
03/16/2003						13.6	13.6
02/22/2004							6.29
12/29/2004	42.7	42.7	42.7	42.7	42.7	42.7	
01/04/2005	31.9	31.9	31.9	31.9	31.9		

Date	Threshold Duration:				Hydro-Mod (0.002 cfs/acre)**	Calendar year	Annual Season*
	4 days	3 days	2 days	1 day			
	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)	Q (m^3/s)
01/11/2005	52.1	52.1	52.1	52.1	52.1		
02/23/2005	207	207	207	207	207	207	207
03/11/2006						5.99	5.99
02/19/2007						10.3	10.3
01/07/2008							10.3
12/15/2008					30.4		
12/17/2008	133	133	133	133	133	133	133
12/07/2009	378	378	378	378	378	378	378
12/07/2009					19.9		

# years	31	29	28	20	20	73	65
# seasons	27	26	25	18	18	62	73
Average	132	132	136	145	118	74.7	73.1
St-dev	177	176	175	171	181	170	170

*: Annual Season: from 10/1/xxxx to 9/30/xxxx+1

**: The suggested threshold in the Hydro-Mod permit does not allow obtaining 73 peak flows but only 71.

4.2.2 Mathematical Maxima and Threshold of Duration

The second way to obtain maxima peak flows is to determine when a value in the time series is a maximum value. The value Q_i that occurs at a time t_i is maximum when it is the largest value in the portion of the time series belonging to the interval $(t_{i-j\cdot\Delta t}, t_{i+j\cdot\Delta t})$ [15]. The previous definition corresponds with Q_i being the largest value among $Q_{i-j}, Q_{i-j+1},\ldots, Q_{i-1}, Q_i, Q_{i+1}, \ldots Q_{i+j-1}, Q_{i+j}$, where j is the threshold of duration in terms of the time interval Δt (or, in the case of the Tijuana River, the number of days before and after the event at which we want a given value to be maximum). In a mathematical sense, Q_i is a maximum as soon as $Q_i>Q_{i+1}$ and $Q_i>Q_{i-1}$. However, the threshold of duration may be increased from more than one time interval to ensure that the time series is not oscillating up and down at each successive Δt. As the time series analyzed is a daily time series, the value j can be as low as one day or as high as desired. In this study, the maximum value assumed for j is four, because it seems clear that if a peak is the largest in a nine-day period (four days before, the day it occurs, and four days after) then it is a peak worth considering for an extreme analysis. Once the peaks are determined for a given threshold, the largest 73 are selected. It can be observed in Figure A.1 that, for a small threshold of one day, some extreme peaks seems to be part of a unique storm that is oscillating (see, for example, the runoff events of February 28 to March 7, 1978 and January 11 to March 11, 1993 in Appendix A). Also, the analysis revealed that once the threshold increases over 2 days or so, few changes in the highest 73 peaks are observed (as a matter of fact, the largest 73 values with a threshold of four days are almost the same as those for a threshold of three days, as shown in Table 4-1). As the selection of the peaks changes very little after a certain threshold, and as the threshold of 1 day is not considered appropriate due to potential oscillations in the data, the values obtained with thresholds of 2 days, 3 days and 4 days will be used for the determination of the 1 day average extreme events.

4.2.3 Data Selection for Frequency Analysis

For the purposes of a sensitivity analysis, the two-day, three-day and four-day threshold 73 maximum values were selected to estimate the extreme peak flows measured in the Tijuana River. The annual maxima data, both seasonal or in a calendar year, were not considered appropriate for this analysis. Similarly, the data derived from the peak-flow threshold, in accordance with the hydromodification permit assumption, or the data gathered from a threshold of one day were discarded for further analysis.

4.3 FINDING A DISTRIBUTION FOR THE DATA

In order to estimate the 2-yr, 5-yr, 10-yr, 25-yr, 50-yr, 100-yr, 200-yr, 333-yr, 500-yr, 1000-yr and any other desired return periods for a given data set, the distribution of the data needs to be adjusted in order to extrapolate the data for determination of a theoretical extreme event. In this section, one common hydrology method (plotting position) and 8 common distributions (normal, log-normal, 3 parameter log-normal, Generalized Extreme Value GEV, Gumbel, Weibull, Pearson Type III and Log-Pearson Type III) are used to calculate the most likely values of the extreme events as a function of the return period from the data gathered.

4.3.1 Plotting Position

For the more frequent events (2-yr, 5-yr, 10-yr, and up to a 25- to 100-yr return period), a plotting position method can be used to estimate the peak flow with a given return period.

For a series of n extreme values, sorted from largest to smallest, the return period T for the extreme value in position "i" (i can vary from 1 to n) can be obtained as [5, 6, 13]:

$$T = \frac{n+1-2a}{i-a} \tag{4}$$

The constant "a" depends on the plotting position method used. Hazen proposed a methodology equivalent to a = 0.5. Gumbel rejected this formula, in part because it assigned a return period of 2n years to the largest observation (i = 1). Weibull proposed the use of a = 0 to achieve unbiased exceedance probabilities for all distributions. Cunnane recommended a = 0.4 because the Weibull formula does not provide an estimate of the cumulative density function F such that the expected value E(F) equals the theoretical value for the i^{th} largest out of n total samples for any underlying distribution other than the uniform, and, therefore, all of the distributions commonly used for flood frequency analysis are excluded [13]. As the theoretical value of a is 0.375 for the normal and log-normal distributions, and 0.44 for the Gumbel distribution, and generally variable between those values for Pearson and Log-Pearson distributions (depending on the parameters obtained), Cunnane proposed a = 0.4 as a compromise [13]. The Cunnane recommendation has been adopted by the San Diego County Hydromodification Permit to calculate the 2-yr and 10-yr peak flows [9], and also has been considered as appropriate in this assessment.

Extreme events up to the 100-yr storm event have been calculated using the Cunnane plotting position method, and are displayed in Table 4-2. The maximum extreme event that can be obtained from the plotting position methodology corresponds with T = 2n in the Hazen formula, T = n+1 in the Weibull formula, or T = (1⅔) n + ⅓ in the Cunnane formula. In other words, the maximum observed peak flow in a 73-yr time series corresponds to a return period of 122 yrs using Cunnane (74 yrs according to Weibull, and 146 yrs according to the Hazen formulas). Therefore, the use of the plotting position methodology is not a viable method to estimate the 200-yr peak flow or higher peaks as it is limited by the duration of the data series. It is also worth noticing that the plotting position is highly sensitive to the occurrence of an extreme event at the end of the spectrum (for i=1 and i=2 in equation (4), and therefore, for peak flows between the 46-yr and 122-yr recurrence interval according to the Cunnane formula).

Extrapolation of a plotting position curve can be done by adjusting the plotting position curve into a polynomial function, either in normal paper or in log-log paper. This method is not explained in classical hydrology handbooks [6, 13, 16] and, therefore, is not considered here.

The plotting position has the disadvantage that only the two largest values are considered in the interpolation of extreme events. It is difficult to tie the largest observation to a specific period of return, as the period of return of the largest value among "n" values is a random variable. When using n=73 and i=1, the largest observation is expected to represent the peak flow with a period of return of 122 years, and the second largest observation (i=2) is expected to represent the peak flow with a period of return of 45.75 years. Therefore, the 50 year storm and the 100 year storm are only a function of the two largest values. As the distributions studied in the following sections take into account each value to determine its parameters, a lower weight will be assigned to the plotter position method for periods of return of 50 and 100 years (see section 5).

4.3.2 Normal Distribution

Many physical processes can be conceptualized as the sum of individual processes. Under very general conditions, the central limit theorem states that as the number of variables in the sum becomes large, the distribution of the sum of a large number of random variables will approach the Normal Distribution, regardless of the underlying distribution. Therefore, the Normal Distribution is often the first distribution at which random data is fitted.

Runoff distribution, however, is not normally distributed. The Normal Distribution is symmetrical with respect to the mean, and the runoff data for the Tijuana River is highly skewed to the left, with most values being slightly smaller than the mean, and few values being many times larger than the mean.

Extreme events under Normal Distribution are determined as:

$$Q_T = Q_{ave} + z_T \cdot \sigma_Q \qquad (5)$$

where:

Q_T : Peak flow with a return period T

Q_{ave} : Average peak flow of the exceedance series of flows selected

σ_Q : Standard deviation of the exceedance series of flows

z_T : Standard normal variate with a value such that the Cumulative Distribution Function (CDF) of z is related to the Return Period T in such a way that:

$$F(z_T) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-0.5u^2} du = 1 - \frac{1}{T} \tag{5.1}$$

Although equation (2.1) cannot be solved analytically, tables exist to obtain $F(z_T)$ vs z_T. As $F(z_T)$ is known and equal to (T-1)/T, z_T can be obtained with standard normal distribution tables. Values of Q_T at different periods of return have been calculated and are shown in Table 4-2 of this report.

4.3.3 Log Normal Distribution

The Log-Normal Distribution corresponds to the application of the normal distribution to the logarithmic value of the data. Consider a hypothetical runoff calculation in which runoff is equal to the product of functions of several random factors such as rainfall, contributing area, loss coefficient, evapotranspiration, evaporation, infiltration, etc. In general, if a random variable x results from the product of a large number of other random variables then the distribution of the logarithm of x will approach normality.

Extreme events following a log-normal distribution can be calculated using the same procedure than for the Normal Distribution (equations (5) and (5.1)), but working with logarithms of the data instead of the regular data. Extreme events have been calculated and are shown in Table 4-2.

4.3.4 Three Parameter Log-Normal Distribution

In many cases the logarithms of a variable X are not quite normally distributed, but subtracting a lower bound parameter ξ before taking logarithms may resolve the problem. Thus, Y = ln (X- ξ) is modeled as having a Normal Distribution [13].

A simple and efficient estimator of ξ is the quantile-lower-bound estimator [13] given by:

$$\xi = \frac{x_1 \cdot x_n - x_{median}^2}{x_1 + x_n - 2x_{median}} = \frac{x_1 \cdot x_{73} - x_{37}{}^2/2}{x_1 + x_{73} - 2 \cdot x_{37}} \tag{6}$$

where:

x_i represents the value in position "i" from sorting the data from highest to lowest. As there are 73 values, the median value x_{median} can be calculated as x_{37}.

The estimation of ξ could be improved by choosing the value of ξ that optimize the adjustment of the data to the three parameter log-normal distribution by minimizing the variable A^{*2} in the Anderson-Darling test.

4.3.5 Generalized Extreme Value Distribution (GEV Distribution)

In probability theory and statistics, the generalized extreme value (GEV) Distribution is a family of continuous probability distributions developed within extreme value theory to combine the Gumbel, Frechet and Weibull families also known as type I, II and III extreme value distributions. By the extreme value theorem (similar to the central limit teorem but for extreme values) the GEV Distribution is the limit distribution of properly normalized maxima of a sequence of independent and identically distributed random variables. Because of this, the GEV Distribution is used as an approximation to model the maxima of long (finite) sequences of random variables.

The GEV is a 3 parameter distribution whose CDF can be written as:

$$F_x(x) = \exp\left\{-\left[1-\frac{\kappa(x-\xi)}{\alpha}\right]^{1/k}\right\} = 1-\frac{1}{T} \tag{7}$$

In the previous equation, the parameter κ is smaller than 0 for extreme positive events, the minimum peak flow considered is $x = \xi + \alpha/\kappa$ and ξ, α, κ are related to the mean Q_{ave}, standard deviation σ_Q and skewness coefficient C_s through the equations displayed in page 18.13 of reference [13]. Once the parameters ξ, α, κ are found the peak flow x can be found with equation (4) for a given return period T. Extreme values obtained with the GEV Distribution are given in Table 4-2.

Regarding C_s, the typical skewness coefficient C of a series of N values can be obtained by the classical definition given by statistical books and also incorporated in spreadsheets and normal programs, and then can be corrected for hydrologic applications [17] to obtain the corrected skewness coefficient C_s to be applied for al distributions that use C_s as an estimator (GEV, Pearson and Log Pearson):

$$C_s = C\left(1+\frac{6}{N}\right) = \frac{N}{(N-1)(N-2)}\frac{\sum_{i=1}^{N}(Q_i-Q_{ave})^3}{\sigma_Q^{3/2}}\left(1+\frac{6}{N}\right) \tag{7.1}$$

4.3.6 Weibull Distribution

Weibull Distribution belongs to the family of extreme events and is related logarithmically with the Gumbel Distribution. If a given variable X has a Weibull Distribution, then Y=ln[X] has a Gumbel distribution.

The Weibull Distribution is a 2 parameter distribution whose CDF can be written as:

$$F_x(x) = 1-\exp\left[-\left(\frac{x}{\alpha}\right)^k\right] = 1-\frac{1}{T}; \quad or \quad x = \alpha(\ln[T])^{1/k} \tag{8}$$

In the previous equation, the parameter k is larger than 0 for extreme positive events, the minimum peak flow considered is x=0 and α, k are related to the mean Q_{ave} and standard deviation σ_Q of the data through the equations displayed in page 18.13 of reference [13] among many other statistical references. Once the parameters α, k are found the peak flow x can be found with equation (8) for a given return period T. Extreme values obtained with the Weibull Distribution are given in Table 4-2.

4.3.7 Gumbel Distribution

The Gumbel Distribution is a simple two parameter distribution for extreme events that arises from the theory of extremes. The Gumbel Distribution is unbounded to the left (negative peak flows are possible), but it is applied to the right side, for extreme events with a return period of at least 2 years.

In order to obtain the extreme events from the Gumbel Distribution as a function of the return period T, the following equation is used [16]:

$$Q_T = Q_{ave} - 0.7797\left[0.5772 + \ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right]\sigma_Q \qquad (9)$$

where:

- Q_T : Peak flow with return period T
- Q_{ave} : Average peak flow of the exceedance series of flows selected
- σ_Q : Standard deviation of the exceedance series of flows

Extreme values from the Gumbel distribution are shown on Table 4-2.

4.3.8 Pearson Type III (Gamma) Distribution

The Gamma Distribution (Pearson Type III) is very useful in hydrology because of its shape and mathematical properties. This three-parameter distribution is bounded to the left and has skewness different than zero (which is different than the normal distribution which has 0 skewness).

Extreme events from the Pearson Type III Distribution can be obtained using the following equation [6]:

$$Q_T = Q_{ave} + K(C_s, T)\cdot\sigma_Q \qquad (10)$$

where the only undefined parameter is the frequency factor K which is a function of two parameters: the corrected skewness coefficient C_s of the exceedance data, and the return period T.

Similarly to Normal and Log-Normal Distributions, K cannot be obtained analytically and Tables have been established to obtain K as a function of C_s and T. The frequency factor K can also be calculated with the gamma-inverse function of the 2010 Excel spreadsheet program.

Results from calculations from the Pearson Type III Distribution are shown in Table 4-2.

4.3.9 Log-Pearson Distribution

The Log-Pearson Distribution is identical to the Pearson Distribution but is applied to the logarithmic value of the data instead of to the data themselves. The results from the application of the Log-Pearson Distribution were obtained as explained for the Pearson Distribution, using equations (10) and (7.1) but applying them to the logarithmic values of the data. Results can be seen in Table 4-2.

4.4 TESTING THE DIFFERENT DISTRIBUTIONS

Many statistical procedures can be used to test a data set for normality (and log-normality) and some of those tests are general enough so they can be used in other distributions. A general procedure to test if data belongs to a distribution is to build a Q-Q plot of the data and see if the plot results in a straight line. Another method is to construct a histogram of the data and observe whether it fits the Probability Density Function (PDF) of the distribution. Graphical tests are more intuitively appealing but subjective at the same time, as they rely on varying human judgment to accept or reject the selected distribution.

An additional and more analytical method is to determine the confidence intervals of the data and see if they are outside the confidence limits, using the approximate Kolmogorov-Smirnov goodness-of-fit statistic to plot confidence intervals [13] or the Anderson-Darling Test [18] which is more rigorous as it is applied to all values in the data set and not only to the estimators (mean, variance and/or skewness). It has been suggested by Stephens [18], [19] and others [20] that the Anderson-Darling Test is a more powerful estimator of normality as the Kolgomorov-Smirnov test fails to discard distributions most of the time. However, the Anderson-Darling test is so robust that it has the opposite problem: it frequently discards a given distribution for a set of values (especially if the set is larger than 25 values) unless the data fits very well to the selected distribution. As many distributions have been analyzed in this study, the Anderson-Darling test (which is explained in the following section) will be used to discard distributions in this study in lieu of graphical methods or the confidence interval method.

Table 4-2. Results of the Analysis for Extreme Events Peak Flows (m^3/s)

Method		Peaks flow at a Return Period given by the Sub-index of Q (m^3/s)									
		Q_2	Q_5	Q_{10}	Q_{25}	Q_{50}	Q_{100}	Q_{200}	Q_{333}	Q_{500}	Q_{1000}
Threshold* = 2 days	**Plotting Position (Cunnane)**	**54.1**	**195**	**347**	**637**	**738^**	**817^**	**-**	**-**	**-**	**-**
	Normal	136	283	360	442	495	543	586	616	639	676
	Log-Normal	79.0	180	278	440	592	774	989	1170	1330	1638
	Log-Normal 3 Parameter Distr.	**62.2**	**172**	**319**	**641**	**1015**	**1543**	**2268**	**2960**	**3624**	**5038**
	GEV	96.4	235	341	493	622	764	923	1050	1159	1362
	Weibull	74.9	218	343	525	672	826	987	1109	1208	1382
	Gumbel	108	262	364	493	589	684	779	848	904	998
	Pearson Type III	66.9	208	341	534	689	847	1010	1131	1229	1397
	Log-Pearson III Distribution	**64.9**	**161**	**294**	**617**	**1053**	**1771**	**2946**	**4262**	**5397**	**9309**
Threshold* = 3 days	**Plotting Position (Cunnane)**	**50.7**	**195**	**347**	**637**	**738^**	**817^**	**-**	**-**	**-**	**-**
	Normal	132	281	358	441	494	543	587	617	640	677
	Log-Normal	72.4	172	270	437	597	790	1021	1219	1394	1733
	Log-Normal 3 Parameter Distr.	**56.7**	**162**	**307**	**626**	**1003**	**1539**	**2281**	**2994**	**3681**	**5154**
	GEV	91.6	232	339	493	623	766	926	1054	1164	1368
	Weibull	69.0	210	336	523	676	838	1009	1139	1245	1431
	Gumbel	103	259	362	493	589	686	781	852	908	1003
	Pearson Type III	61.1	201	336	534	692	856	1023	1148	1249	1422
	Log-Pearson III Distribution	**59.0**	**153**	**287**	**622**	**1088**	**1875**	**3195**	**4695**	**6358**	**10621**
Threshold* = 4 days	**Plotting Position (Cunnane)**	**50.7**	**195**	**347**	**637**	**738^**	**817^**	**-**	**-**	**-**	**-**
	Normal	132	280	358	441	494	543	587	617	640	678
	Log-Normal	71.8	172	271	440	603	800	1036	1238	1417	1765
	Log-Normal 3 Parameter Distr.	**56.1**	**163**	**312**	**642**	**1034**	**1593**	**2370**	**3120**	**3845**	**5402**
	GEV	91.4	231	339	493	623	766	926	1054	1164	1368
	Weibull	68.7	209	336	523	676	839	1010	1140	1247	1434
	Gumbel	103	259	362	493	590	686	782	852	908	1003
	Pearson Type III	60.9	202	337	534	692	856	1023	1148	1248	1421
	Log-Pearson III Distribution	**58.6**	**153**	**288**	**520**	**1092**	**1880**	**3195**	**4693**	**6349**	**10581**
Geometric mean of used distributions (2 digit precision).		**57**	**170**	**320**	**620**	**980**	**1500**	**2700**	**3700**	**4700**	**7300**

*: A threshold of n days means that the n largest values were selected in the time series in such a way that they were the 73 largest values that satisfied the condition of being maxima in a time interval centered in the value with a duration 2n + 1 days (n days before, the day of the value and n days after).

^: Plotting Position Method has a weight of 0.5 in the geometric mean calculation for 50-yr and 100-yr daily peak flow calculation while Three-Parameter Log-Normal and Log-Pearson Type III have a weight of 1.

Table 4-3. Results of the Analysis for Extreme Events Peak Flows (cfs)

Method		Peaks flow at a Return Period given by the Sub-index of Q (cfs)									
		Q_2	Q_5	Q_{10}	Q_{25}	Q_{50}	Q_{100}	Q_{200}	Q_{333}	Q_{500}	Q_{1000}
Threshold* = 2 days	Plotting Position (Cunnane)	**1910**	**6880**	**12300**	**22500**	**26100^**	**28900^**	-	-	-	-
	Normal	4820	10000	12700	15600	17500	19200	20700	21800	22600	23900
	Log-Normal	2790	6370	9810	15500	20900	27300	34900	41300	47000	57800
	Log-Normal 3 Parameter Distr.	**2200**	**6060**	**11300**	**22600**	**35900**	**54500**	**80100**	**105000**	**128000**	**178000**
	GEV	3400	8300	12000	17400	22000	27000	32600	37100	40900	48100
	Weibull	2640	7690	12100	18500	23700	29200	34900	39200	42700	48800
	Gumbel	3800	9250	12900	17400	20800	24200	27500	30000	31900	35300
	Pearson Type III	2360	7340	12000	18900	24300	29900	35700	40000	43400	49300
	Log-Pearson III Distribution	**2290**	**5700**	**10400**	**21800**	**37200**	**62500**	**104000**	**151000**	**201000**	**329000**
Threshold* = 3 days	Plotting Position (Cunnane)	**1790**	**6880**	**12300**	**22500**	**26100^**	**28900^**	-	-	-	-
	Normal	4660	9910	12600	15600	17500	19200	20700	21800	22600	23900
	Log-Normal	2560	6070	9540	15400	21100	27900	36100	43000	49200	61200
	Log-Normal 3 Parameter Distr.	**2000**	**5720**	**10800**	**22100**	**35400**	**54300**	**80600**	**106000**	**130000**	**182000**
	GEV	3240	8200	12000	17400	22000	27100	32700	37200	41100	48300
	Weibull	2440	7400	11900	18500	23900	29600	35600	40200	44000	50500
	Gumbel	3640	9150	12800	17400	20800	24200	27600	30100	32000	35400
	Pearson Type III	2160	7120	11900	18900	24400	30200	36100	40600	44100	50200
	Log-Pearson III Distribution	**2080**	**5410**	**10100**	**22000**	**38400**	**66200**	**113000**	**166000**	**225000**	**375000**
Threshold* = 4 days	Plotting Position (Cunnane)	**1790**	**6880**	**12300**	**22500**	**26100^**	**28900^**	-	-	-	-
	Normal	4660	9900	12600	15600	17500	19200	20700	21800	22600	23900
	Log-Normal	2530	6060	9560	15600	21300	28200	36600	43700	50000	62300
	Log-Normal 3 Parameter Distr.	**1980**	**5760**	**11000**	**22700**	**36500**	**56200**	**83700**	**110000**	**136000**	**191000**
	GEV	3230	8200	12000	17400	22000	27100	32700	37200	41100	48300
	Weibull	2430	7390	11900	18500	23900	29600	35700	40300	44000	50600
	Gumbel	3630	9140	12800	17400	20800	24200	27600	30100	32100	35400
	Pearson Type III	2150	7120	11900	18900	24400	30200	36100	40500	44100	50200
	Log-Pearson III Distribution	**2070**	**5410**	**10200**	**22100**	**38600**	**66400**	**113000**	**166000**	**224000**	**374000**
Geometric mean of used distributions (2 digit precision).		**2000**	**6000**	**11000**	**22000**	**34000**	**52000**	**95000**	**130000**	**170000**	**260000**

*: A threshold of n days means that the n largest values were selected in the time series in such a way that they were the 73 largest values that satisfied the condition of being maxima in a time interval centered in the value with a duration 2n + 1 days (n days before, the day of the value and n days after).

^: Plotting Position Method has a weight of 0.5 in the geometric mean calculation for 50-yr and 100-yr daily peak flow calculation while Three-Parameter Log-Normal and Log-Pearson Type III have a weight of 1.

4.4.1 Anderson-Darling Test

Let the values $Q_1, Q_2, \ldots, Q_{i-1}, Q_i, \ldots, Q_{n-1}, Q_n$ represent the "n" extreme independent values sorted from lower to higher. Let the function $\Phi(Q_i)$ represents the CDF value of a given Q_i or the probability that a given Q_i value is lower or equal than Q_i (for a flow value Q_x with a return period T, for example, the function $\Phi(Q_x)$ is equal to 1-1/T).

The Anderson-Darling test consists in obtaining the empirical A^{*2} value that satisfies the following equation for a set of data:

$$A^{*2} = \left(1 + \frac{4}{n} - \frac{25}{n^2}\right)\left(-n - \frac{1}{n}\sum_{i=1}^{n}\left[(2i-1)\ln(\Phi(Q_i)) + (2(n-i)+1)\ln(1-\Phi(Q_i))\right]\right) \quad (11)$$

If the CDF $\Phi(Q_i)$ corresponds with the normal distribution, then the hypothesis of normality is rejected when A^{*2} exceeds 0.631 for a 10% significance level, 0.751 for a 5% significance level test, or is rejected if A^{*2} exceeds 1.029 for a 1% significance level. For a Log-Normal Distribution, the data is transformed using logarithms and the previous test is performed, and for a Three- Parameter Log-Normal Distribution, the parameter ξ is subtracted from the data, logarithms are taken and the previous test is performed.

For other distributions (Pearson, Log-Pearson, Gumbel, Weibull, GEV) the function $\Phi(Q_i)$ is the CDF of the tested distribution. Therefore, the corresponding parameters must be obtained in each case and the value of A^{*2} must be calculated for each distribution.

Although values of A^{*2} to reject a distribution at a given level are different than the values given for the normal distribution (which will be the case to determine if a not normal can be rejected), it is known that the lower the parameter A^{*2} is, the better the data fits the given distribution. Therefore, only those distributions whose A^{*2} parameter is less than 1 will be considered for further analysis of peak flows. Values of A^{*2} are presented in Table 4-4 for the distributions analyzed and for both data sets: the data set obtained with a threshold of three days, and the data set obtained with a threshold of two days. It is clear from Table 4-4 that the Three-Parameter Log-Normal cannot be rejected at a 5% level test. Regarding the Log-Pearson distribution, values of A^{*2} to reject the distribution at a given level test have not been found in the literature review, and as it is the second best among the eight distributions studied, it will be also considered for peak flow determination. At this point, it should be emphasized that according to the U.S. Interagency Advisory Committee on Water Data [5], the Log-Pearson Type III distribution is usually the preferred option for flood analysis, and the Anderson-Darling test is not conclusive to reject this distribution.

It must be pointed out that the Plotting Position Method will always pass any statistical test as the CDF is not represented by a mathematical function. In a plotting position analysis, each sorted peak has a CDF that is independent of its value and only a function of its sorted position.

Table 4-4. A^{*2} for Studied Distributions

DISTRIBUTION(1)	A^{*2} (2-day threshold data)	A^{*2} (3-day threshold data)	A^{*2} (4-day threshold data)	REMARKS
Normal	9.46	9.87	9.81	Worst distribution to fit the data
Log Normal	2.26	2.62	2.47	4th best, but 3 Parameter Log-Normal Preferred
Log-Normal -3	**0.34**	**0.41**	**0.36**	**Best Distribution, with very low values of A^{*2}**
GEV	5.16	5.74	5.67	Poorly adjusted to the data
Weibull	3.92	3.83	3.68	Poorly adjusted to the data
Gumbel	6.68	7.16	7.10	Very high value of A^{*2}. Worst after normal.
Pearson Type III	1.80	2.08	1.94	Poor adjustment related to low values. Log-Pearson III preferred
Log Pearson Type III	**0.82**	**0.93**	**0.86**	**Values of A^{*2} relatively high, but second best adjustment**

(1) : Bold values represent the distributions selected to study the data

This page intentionally left blank

SECTION 5 SUMMARY OF RESULTS – 73 YEARS OF DAILY RECORDS

The Three Parameter Log-Normal Distribution and the Log-Pearson Type III Distribution were the only distributions that were not discarded by the Anderson-Darling mathematical test, with the former being better than the later. As both distributions correspond to the analysis of the logarithms of the data, it is considered more appropriate to use the geometric mean rather than the arithmetic mean to estimate the final peak flows in this technical memorandum.

Peak flows for the Tijuana River were calculated as stated below.

1. For 2-yr, 5-yr, 10-yr, and 25-yr events, the geometric mean was used among the values obtained using the plotting position method, Log-Pearson distribution and Three Parameter Log-Normal Distribution. The peaks were calculated with a threshold of 2 days, 3 days and 4 days to separate storms (geometric mean among nine values, equivalent to the arithmetic mean or average of the logarithms of the values).
2. For 50-yr, and 100-yr events, the weighted geometric mean was used among the values obtained using the plotting position method, the Log-Pearson Type III and the Three Parameter Log-Normal Distribution. The peaks were calculated with a threshold of 2 days, 3 days and 4 days to separate storms. Weights were assigned as 0.5 for the plotting position and 1 for the distributions. The weighted geometric mean is equivalent to the weighted average of the logarithms of the nine values considered.
3. For 200-yr, 333-yr, 500-yr and 1000-yr events, the geometric mean between the 3 Parameter Log-Normal and the Log-Pearson distribution was used, calculated with a 2-day, 3-day and 4-day threshold to define maxima (geometric mean among four values). Table 4-2 displays the results of the analysis.

If a different return period is desired, the logarithms of the values in the last file of Table 7-1 can be interpolated vs. the logarithm of the return period. If a peak flow Q for a return period T is needed, the values Q_1, T_1 and Q_2, T_2 are obtained from Table 4-2, such that $T_1 < T < T_2$. Equation (12) provides the logarithmic interpolation:

$$Q = \exp\left[\ln(Q_1) + \frac{\ln(T/T_1)}{\ln(T_2/T_1)}\ln(Q_2/Q_1)\right] \tag{12}$$

For example, the 4 yr return period peak flow will be:

$$Q_4 = \exp\left[\ln(57) + \frac{\ln(4/2)}{\ln(5/2)}\ln(170/57)\right] = 130 \text{ m}^3/\text{s}$$

This page intentionally left blank

SECTION 6 ADDITONAL ASPECTS RELATED TO EXTREME EVENTS: HISTORICAL RECORDS AND PEAK FLOWS

In the previous sections, daily average peak flows have been analyzed with 73 years of daily records. There is additional information related to historical records prior to 10/1/1936 and also additional information related with annual peak flows from the USGS data set (1937 to 1982) that will be discussed in the following to sections to understand the relevance of those aspects in this study.

It is important to make a distinction between peak flows and daily average peak flows: peak flows are related with an instantaneous measurement, associated with the maximum level at a gauge which is translated into an instantaneous peak flow whose duration is unknown but which is smaller than 1 day. The average daily peak flow represents the average value of the flow at a given day. The IBWC data displays the average daily peak flow since 1/1/1962, and the USGS data displays the average daily peak flow from 10/1/1936 to 9/30/1982. Also, the USGS web page has information related to the maximum peak flow for each calendar year from 1937 to 1981. Differences between both measurements will be discussed in section 6.2.

6.1 HISTORICAL RECORDS

There is significant historical evidence of extreme events occurring in the Tijuana Estuary since the 1880's and there have been estimates of extreme events prior to 1936. Although these peak flows should be seen as an approximation of the extreme conditions that occurred in the past, they could be an important tool to measure highly unlikely events and to establish the confidence in those events with a return period of 100 yrs of larger.

Table 6-1 shows the events that have occurred since 1884 [2, 21]. Although 1941 was an extremely wet year from the point of view of runoff volume (as a matter of fact, according to the USGS runoff records the total runoff in 1941 represents 48% of the total runoff from 1937-1977 and more than 3 times than the second largest runoff year of the period), it seems that the largest peak flow measured that year was only 13,800 cfs [23] and even less if the average daily peak flow is considered from the USGS time series (10,400 cfs on 4/12/1941). Therefore, no peak from 1941 is included in Table 6-1. Rodriguez Dam, finished at the end of 1936, retained and detained a significant volume of runoff which may explain the relatively low peak flows from that year. It is known from historical accounts that Rodriguez Dam was full for the first time in 1940 and the operation of Rodriguez Dam in 1941 more likely prevented larger peak flows at the entrance of the Valley.

Another important aspect worth mentioning regarding historic peak flows is the fact that peak flows before 1912 were not controlled by dams. This may partially explain why peak flows exceeding 20,000 cfs were more common before 1920. The different dams entered in operation as follows: Morena Dam 1912 (contributing area equal to 6.9 % of the watershed); Barrett Dam, 1922 (13.6 %, which includes the contributing area of Morena); Rodriguez Dam, 1937 (56.9 % of the watershed) and El Carrizo Dam, 1974 (2.6 % of the watershed). From the previous values, it is evident that Rodriguez Dam and to a lesser extent Barrett Dam have played an important role in the peak flows measured since 1937.

Table 6-1. Historical Extreme Events

DATE	Discharge (cfs)	Discharge (m^3/s)	RANK(5)
February, 1884	50,000(1)	1400(1)	2
December, 1889	20,000(1)	570(1)	8-9
February, 1991	20,000(1)	570(1)	8-9
January, 1895	38,000(1)	1100(1)	3
January, 1916	75,000(1)	2100(1)	1
February, 1927	25,000(1)	710(1)	6
February 7th, 1937	*17,700(2)*	*500(2)*	11
January 30th, 1980	31,000(1) **19,500(4)**	880(1) **547(4)**	10
February 21st, 1980	33,500(2, 3) **30,000(4)**	950(2, 3) **852(4)**	4
March 3rd, 1983	27,700(3) **24,500(4)**	780(3) **697(4)**	7
January 16th, 1993	33,000(3) **26,000(4)**	934(3) **731(4)**	5
February 20th, 1993	**17,500(4)**	**496(4)**	12
March 12th, 1995	**16,500(4)**	**464(4)**	13

Notes:
(1): Estimations of past floods (made originally in cfs). Published on [2]. Values assumed to be peak flows.
(2): Peak flow measurements from the USGS Nestor Gauge (in cfs) according to [1a]
(3): Measurements according to References [2, 21]
(4): Measurements per published values of the IBWC [1]. Measurements published in m^3/s
(5): Rank and statistical properties obtained with bold values when two values exist.

As a consequence of Rodriguez and Barrett controlling 70.5 % of the watershed, extreme events prior to 1937 are not considered statistically comparable to those after that date. Also, although El Carrizo Dam affected hydrologic records since 1974, its contributing area is comparatively small. Consequently, it is assumed that El Carrizo does not change the hydrology of the peak discharges significantly, and therefore the peak flow records after El Carrizo started operation are statistically similar to those before that time.

As a conclusion, although the six events prior to 1937 (and four prior to 1900) are a good indication of the extreme peak flows that the Tijuana River watershed can generate, they will not be included in the statistical analysis as (1) the dams in the watershed have altered the hydrology and (2) those peaks are estimations and not measurements associated with a gauged station. Both aspects made those peaks unreliable for statistical purposes.

It must be pointed out that some extreme events are different when measured by IBWC [1] or when measured by the USGS Nestor Gauge as published by [2]. In order to be consistent in the statistical analysis, daily measurements by IBWC will be selected in those cases where two measurements exist.

6.2 PEAK FLOWS VS DAILY AVERAGE PEAK FLOWS

There are differences between daily measurements from IBWC Data and USGS Data from 1/1/1962 to 9/30/1982 that have been discussed in section 3.1.1. However, there are also differences between peak flows and daily average peak flows that affect the statistical analysis of the data, and that will be discussed here.

There are 73 independent daily peak flows that represent the 73 largest events from the data series of 1937 to 2009 (see Table 4-1). Of those, 14 have an associated peak flow taken from the maximum annual peak flow data available from the USGS webpage. Table 6-2 shows the difference between the peak flows and the daily average peak flows.

Table 6-2. Differences between Peak Flow and Average Daily Peak Flow (cfs)

Peak flow Date	Peak Stream flow (USGS)	Average Daily flow (USGS)	Peak flow Date	Peak Stream flow (USGS)	Average Daily flow (USGS)	Average Daily flow (IBWC)
2/7/1937	17,700	9,710	2/24/1944	11,100	10,100	-
3/3/1938	6,760	4,950	12/23/1945	2,100	1,020	-
2/4/1939	1,730	1,080	1/13/1949	2,600	1,140	-
2/22/1941[1]			3/16/1952	2,460	1,330[3]	
	13,800	3,920	3/17/1952	-	1,800[3]	-
4/12/1941[1]	-	10,400	12/7/1966	2,020	907	*1049*
3/17/1942	2,770	2,560	3/1/1978	6,370	1,720[4]	*3,270*
1/27/1943[2]	1,060	473	3/29/1979	1,610	1,460	*1,321*
3/5/1943[2]	-	808	2/21/1980	33,500	30,200	*30,088*

Notes:
(1): Maximum peak flow on 1941 occurred on 2/22/41 according to USGS web site but maximum daily average occurred on 4/12/41 from the same website. As the peak flow on 2/22/41 is 3.5 times larger than daily average, and only 33% larger than the 4/12/1941 daily average, the date of the peak could have been reported incorrectly.
(2): Maximum peak flow on 1943 occurred on 1/27/43 according to USGS web site but maximum daily average occurred on 3/5/43 from the same website. As the peak flow on 1/27/43 is 2.2 times larger than daily average, and only 31% larger than the 3/5/1943 daily average, the date of the peak could have been reported incorrectly.
(3): Maximum peak flow reported on 3/16/1952 but maximum daily average reported on 3/17/1952. Peak could have occurred at late hours of 3/16/1952 so data does not look suspicious in this case.
(4): Peak flow reported on 3/1/1978, the same day that the IBWC Data displays the maximum daily average. However, USGS daily data displays the peak flow on 3/29/1978 or 28 days later. Analysis on section 3.1.1 already mentioned that the data was probably moved a month and probably the daily peaks may not be correct from the USGS Data in 1978.

From the analysis of Table 6-2 it is clear that peak flows are larger than the daily average peak flows. For very large peak flows (10,000 cfs or more) that are considered reliable, the difference is usually in the order of 10% (2/21/1980; 2/24/1944), except for the sudden storm event of 2/7/1937 where the average daily peak flow increased almost 3 orders of magnitude in one day (15 cfs on 2/6/1937 and 9,710 cfs on

2/7/1937). From smaller peak flows (less than 10,000 cfs) the increase can be relatively minor (3/17/1942, 3/29/1979) or more than twice even for reliable peak flows (12/23/1945, 1/13/1949, 12/7/1966).

However, there is not enough information for most of the 73 largest daily average peak flows as to make the analysis for the instantaneous stream flow peak. Consequently, the results from this study should be interpreted as the period of return for the average daily peak flow. Instantaneous peak flows will be calculated once the shape, volume and duration of different period of return hydrographs are established.

SECTION 7 COMPARISON OF PEAK FLOWS WITH PREVIOUS STUDIES

Table 7-1 presents different peak flows for the Tijuana River estimated from many studies in the last 50 years, from the analysis of the IBWC Data and from the results of the historical data. Values of peak flow as a function of the return period recommended in this study are included in the final column.

Table 7-1. Comparison of Peak Flows for Different Studies

T (yrs)	Peak Flow Estimates (Thousands of cfs)							
	1964 ACoE (1)	1973 FEMA (2)	1987 SLA (3)	1994 TVCWD (4)	1994 BSI (5)	1999 FEMA (6)	Historic Floods 1884-2009(7)	Results Daily data(8) (1937-2009)
1000	-	-	-	-	-	-	137	**260**
500	-	-	-	-	-	150	102	**170**
333	-	-	-	-	-	-	86	**130**
MPE	135	135	-	100	-	-	-	**-**
200	110	-	160	85	-	-	68	**95**
100	77	75	90	65	75	75	50	**52**
50	51	50	50	50	-	50	36.5	**34**
25	33	-	30	35	35	-	26	**22**
20	30	30	20	30	-	-	23.5	***19***(9)
10	16	-	6	17	-	17	13.6	**11**
5	4.3	5	1.9	4	-	-	-	**6.0**
4	2	1.5	1.3	2	-	-	-	***4.6***(10)
2	-	-	0.2	-	-	-	-	**2.0**

Notes:
(1): US Army Corps of Engineers. Tijuana River Basin in CA and Baja, May 1964, Appendix 5, Plate 1
(2): City of San Diego (FEMA Flow Rates) Tijuana River Valley Land Use and Flood Control Alternatives May 73
(3): Simons, Li & Associates Inc. Coast of California Storm and Tidal Wave Study, July 1st, 1987
(4): Tijuana Valley County Water District. Flood Control for the TRV, Jan 1994. Appendix C2
(5): TRV Two Alternatives Report: Flood Control and Infrastructure Study. BSI Consultants, Nov. 1994
(6): FEMA: Flood Insurance Study for San Diego County, CA. June 1999.
(7): Historic Floods analysis in this study, from data obtained at (5) and IBWC Daily Data and UC Berkley Study at http://escholarship.org/uc/item/11r5p44p
(8): Results from this study from daily runoff data gathered by the IBWC (1-1-62 to 12-31-09) and the USGS (10/1/1936 to 12/31/1981). Values recommended for design.
(9): Value interpolated using equation 9. T = 20 yrs
(10): Value interpolated using equation 9. T = 4 yrs

The following is worth noticing from Table 7-1:

- Studies made before 1980 do not include data measured after 1979 and therefore should not be longer consider as updated values.
- There is significant data accumulation not considered in many studies. Peak flow methodologies have been documented in detail in this study and each distribution has been statistically tested. Therefore, the data analysis of this study is considered more reliable.
- Common peak flows (5-yr, 4-yr, and 2-yr peak flows) have a tendency to occur in clusters or more than once per season and the use of annual maxima series will significantly underestimate the most common peak flows. As many of the years in the data are drought years, the use of annual maxima is not recommended for Southern California, and instead the use of annual exceedance time series is preferred. This is the reason for the much larger magnitude of the 5-yr, 4-yr and 2-yr peak flow events in this study. It must be noticed that the clustering effect of the data makes the occurrence of a 2 yr storm event or larger to happen in less than half of the calendar years or in less than half of the seasons.
- The use of an exceedance time series increases common peak flows while reduces extreme peak flows (25-yr return period and longer). The statistical explanation is the following: the exceedance calculation of the highest n values has a higher mean and smaller standard deviation than the selection of the n largest peaks. Therefore, as smaller values are more influenced by the mean, smaller periods of return are higher. For extreme events (25-yr, 50-yr, 100-yr, 200-yr and higher) the influence of the standard deviation is more important than the influence of the mean, and that's the reason why annual maxima generates larger extreme events. Intermediate events (10-yr) have a tendency to be relatively close between the exceedance analysis and the annual analysis. As explained before, exceedance analysis is preferred especially in the dry weather conditions of Southern California.
- Highly extreme events (500-yr and 1000-yr) are very unlikely and imply extreme extrapolation of typical statistical distributions based on very short records compared to the return period desired. Therefore, those results are statistical outliers rather than physical facts. However, it is worth noting that the concrete channel in the City of Tijuana was designed to handle a peak flow of 135,000 cfs which is approximately the maximum capacity of the spillway at Rodriguez Dam (which controls 56% of the watershed).

The values shown in the last column of Table 7-1 are the values recommended for design or analysis purposes for the Tijuana Estuary.

SECTION 8 CONCLUSIONS AND RECOMMENDATIONS

The peak flows of the main Tijuana River as calculated by the statistical analysis of 73 years of data at periods of return of 2, 5, 10, 25, 50, 100, 200, 333, 500 and 1000 years correspond to the values given in Table 8-1.

Table 8-1. Extreme Events as a Function of the Return Period

Return Period (years)	Peak Flow Rate (m^3/s)	Peak Flow Rate (cfs)
2	57	2,000
5	170	6,000
10	320	11,000
25	620	22,000
50	980	34,000
100	1,500	52,000
200	2,700	95,000
333	3,700	130,000
500	4,700	170,000
1,000	7,300	260,000

The peak flows from the daily IBWC runoff data adjust better to the Three-Parameter Log-Normal Distribution than to the Log-Pearson Type III Distribution. However, both distributions were considered to estimate the extreme peak flows when the IBWC Data was analyzed.

The largest thirteen (13) historic peak flows have not been considered in the analysis for two reasons: (1) there is a difference between maximum peak flow and maximum daily average flow which explains the difference in extreme peak flows in some years, and (2) peak flows before 1930 are not controlled by Rodriguez Dam and peak flows before 1912 are not controlled by any dam. The three existing dams have changed the hydrology of the watershed and the statistical consideration of some of the historical peak flows is not appropriate.

The peak flows obtained in this analysis do not include flows from the four Canyons draining to the Estuary downstream of the measurement point from upstream to downstream: Cañón del Sol (unnamed in English), Smuggler's Gulch (Cañón del Matadero), Goat Canyon (Cañón Los Laureles) and Yogurt Canyon (Cañón Los Sauces). Also, the peak flows do not include the contributing area of the estuary itself downstream of the measurement point. The percentage of the area not contributing to the IBWC measurements is about 1.6% of the total area of the watershed.

Peak flows for small return periods (5 years or less) are larger than in previous studies, while peak flows for large return periods (25 years and more) have a tendency to be lower than in previous studies (see Table 7-1). The use of exceedance extreme values rather than annual extreme values is preferred, and explains the previous observation. The use of a larger data set is also responsible for the modification of the peak flows at different return periods.

The peak flows obtained in this study are a function of the existing daily flow data. It is recommended that this study be updated at least every 10 years, or after the occurrence of a significant extreme event (with a return period of 10 yrs or more) as the estimation of the return period of a given peak flow changes as the historical data length increases or the number of extreme events analyzed increases.

The most representative peak flows for the purpose of the Tijuana River and estuary (the 100-yr and the 25-yr peak flows) calculated using this statistical analysis are lower than previous studies with the statistical analysis results. The best estimate for the 100-yr peak flow is 52,000 cfs (instead of the current 75,000 cfs) and the best estimate for the 25-yr peak flow is 22,000 cfs (instead of the current 35,000 cfs).

It should be noted that in the daily records, no peak has reached the previous value established for the 25-yr peak flow as the maximum daily peak flow measured has been 30,100 cfs on 2/21/1980 and the maximum instantaneous peak measured has been 33,500 cfs the same day. Therefore, there is strong statistical evidence that extreme peak flows are smaller than currently thought. It is worth noting that according to the Cunnane plotting position method (a preferred method by the San Diego County Hydromodification permit) the maximum peak flow in a 73 year time series should have a period of return of approximately 122 years.

This technical report focuses only in the determination of the daily average peak flow of extreme events. The volume and duration of the hydrograph associated with those events is analyzed in a separate technical report and will serve for future modeling of the sediment and trash transport into the Tijuana River.

SECTION 9 REFERENCES

[1] International Boundary and Water Commission web page: Data of Tijuana River. Available at: http://www.ibwc.gov/wad/ddqtjrib.htm

[2] BSI Consultants. Inc (1994). "Two Alternatives Report. Flood Control and Infrastructure Study". Prepared for the City of San Diego and the Tijuana River Valley Task Force.

[3] County of San Diego Hydrology Manual. Available at: http://www.co.san-diego.ca.us/dpw/floodcontrol/hydrologymanual.html

[4] State of California Department of Water Resources. "Implementing California Flood Legislation into Local Land Use Planning: A Handbook for Local Communities". October 2010. Available at: http://www.water.ca.gov/floodsafe/docs/Implementing_CA_Flod_Leg-A_Handbook.pdf

[5] U.S. Interagency Advisory Committee on Water Data, Hydrology Subcommittee. "Guidelines for Determining Flood Flow Frequency". Bulletin No 17B, issued 1981, revised 1983, Reston, Virginia.

[6] Ponce, Victor. "Engineering Hydrology. Principles and Practices". 1989, Prentice Hall.

[7] Flood Hydrology of the Binational Cottonwook Creek – Arroyo Alamar, California and Baja California. Victor Ponce, San Diego State University, 2005. Available at: http://alamar.sdsu.edu/

[8] Winckell & Le Page. "El Niño in the City: Vulnerability Assessment in Tijuana during the Winter of 1997-1998". Geofisica Internacional, july-september, año/vol 42, numero 003. Universidad Autonoma de Mexico, pp 553-565. Also available at: http://redalyc.uaemex.mx/pdf/568/56842328.pdf

[9] Final Hydromodification Management Plan. Prepared for the County of San Diego by Brown and Cadwell. October 18 2010 version. Also available at: http://www.projectcleanwater.org/pdf/susmp/final%20hydromodification%20management%20plan.pdf

[10] Potential for Renewable Energy in the San Diego Region. Chapter 7: Renewable Energy Resources in Baja California. http://www.renewablesg.org/docs/Web/Baja.pdf

[11] Padilla-Corona, Antonio: "La Presa Abelardo Rodriguez, Modelo de Ingeniería Hidráulica". Article about the history of Abelardo Rodriguez Dam (in Spanish). Abailable at: http://www.tijuana.gob.mx/ciudad/CiudadLaPresa4.asp

[12] USGS Data for the San Mateo Creek near San Clemente, CA. Available at: http://waterdata.usgs.gov/nwis/uv?11046300

[13] Maidment, David. "Handbook of Hydrology". 1992, McGraw Hill.

[14] Rodriguez-Iturbe, Ignacio & Rinaldo, Andrea. "Fractal River Basins, Chance and Self-Organization". Cambridge University Press, 1997.

[15] Bras, Rafael & Rodriguez-Iturbe, Ignacio. "Random Functions and Hydrology". 1993, Dover Publications, Inc.

[16] Bedient, Philip & Huber, Wayne. "Hydrology and Floodplain Analysis". 1992, 2nd Edition, Addison-Wesley.

[17] Tasker, G. and Stedinger, J. "Regional Skew with Weighted LS Regression". Journal of Water Resources Planning and Management, ASCE, vol 112, no 2, April 1986.

[18] Stephens, M. A. (1974). "EDF Statistics for Goodness of Fit and Some Comparisons". Journal of the American Statistical Association 69: 730–737.

[19] Stephens, M.A. (1979). "Test of fit for the Logistic Distribution Based on the Empirical Distribution Function". Biometrika, 66(3), 591-5.

[20] Shapiro, S.S. (1980) How to test normality and other distributional assumptions. In: The ASQC basic references in quality control: statistical techniques 3, pp. 1–78.

[21] Friedman, N (2008). "River Restoration for a Socially and Ecologically Devastated Border City". University of Berkeley. http://escholarship.org/uc/item/11r5p44p

[22] Zedler, Koenigs and Magdych (1984). "Freshwater Release and Southern California Coastal Wetlands: Streamflow for the San Diego and Tijuana Rivers". San Diego Association of Governments.

Appendix A Tijuana River Average Daily Peak Flow Data (1937-1981)

List of Figures

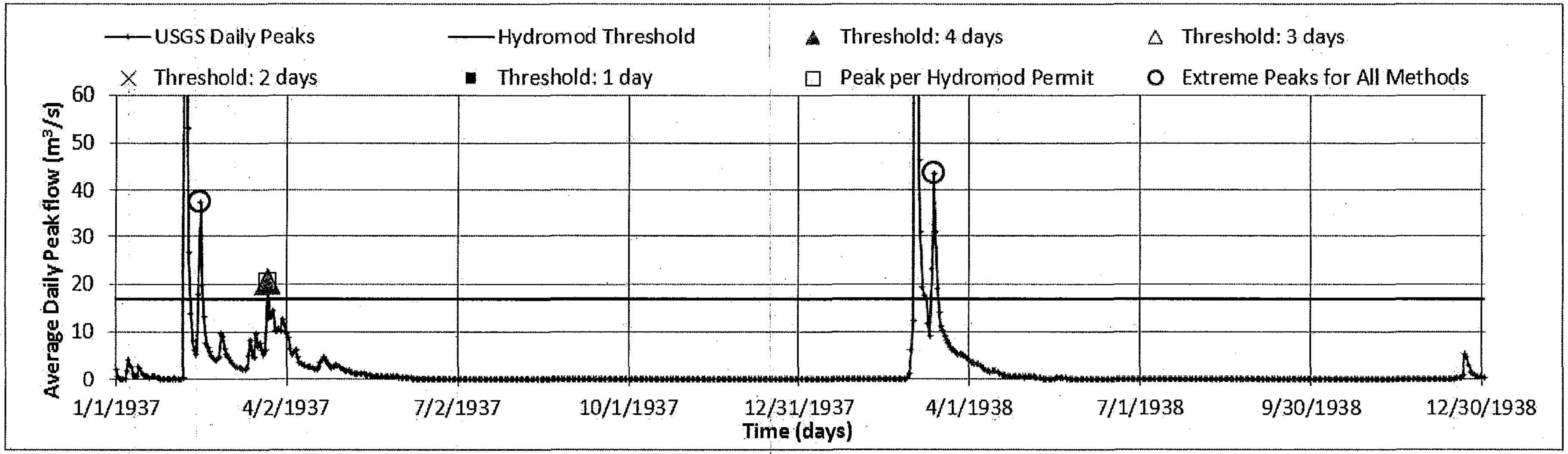


Figure A.1-1a. Peak Flow Time Series. 01/01/37 to 12/31/38.

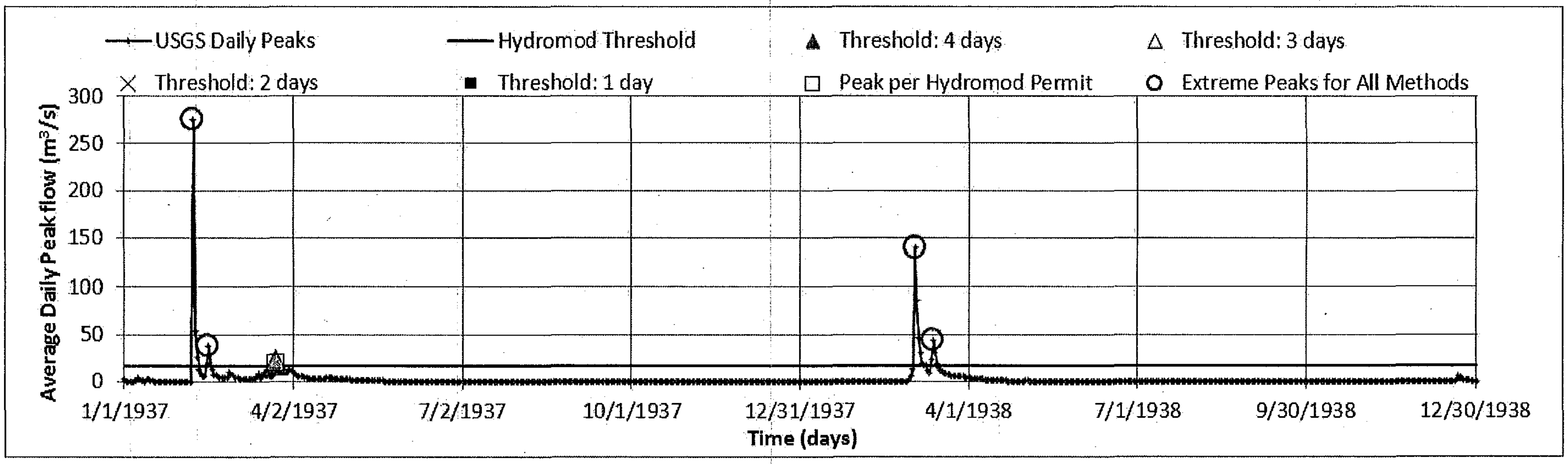


Figure A.1-1b. Peak Flow Time Series. 01/01/37 to 12/31/38. (Notice the change on vertical scale to capture extreme events).

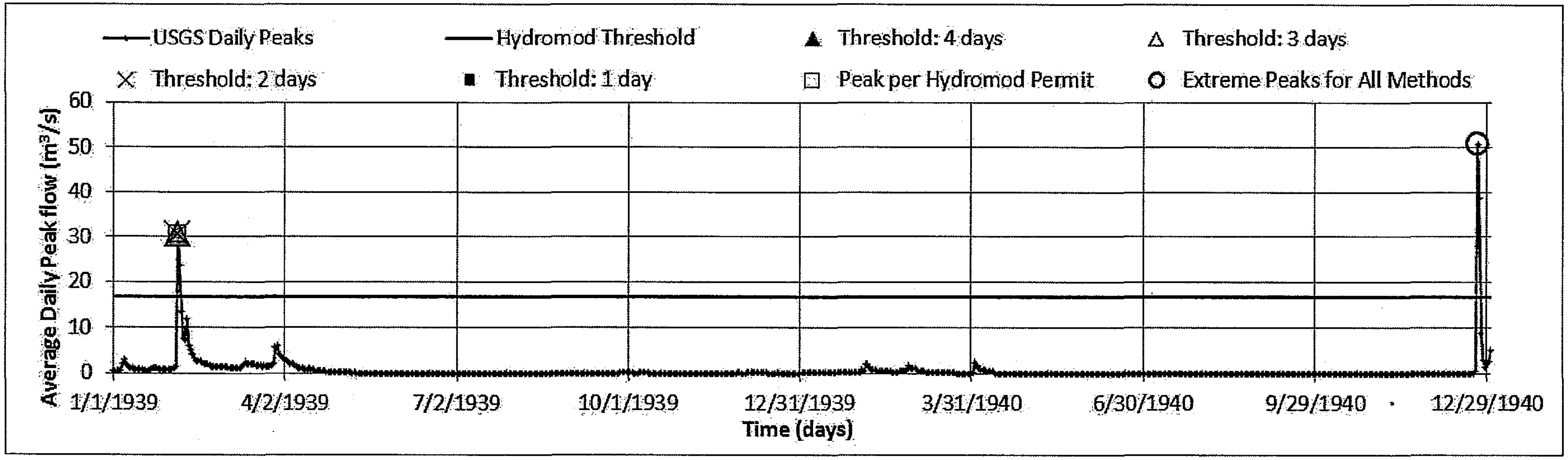


Figure A.1-2. Peak Flow Time Series. 01/01/39 to 12/31/40.

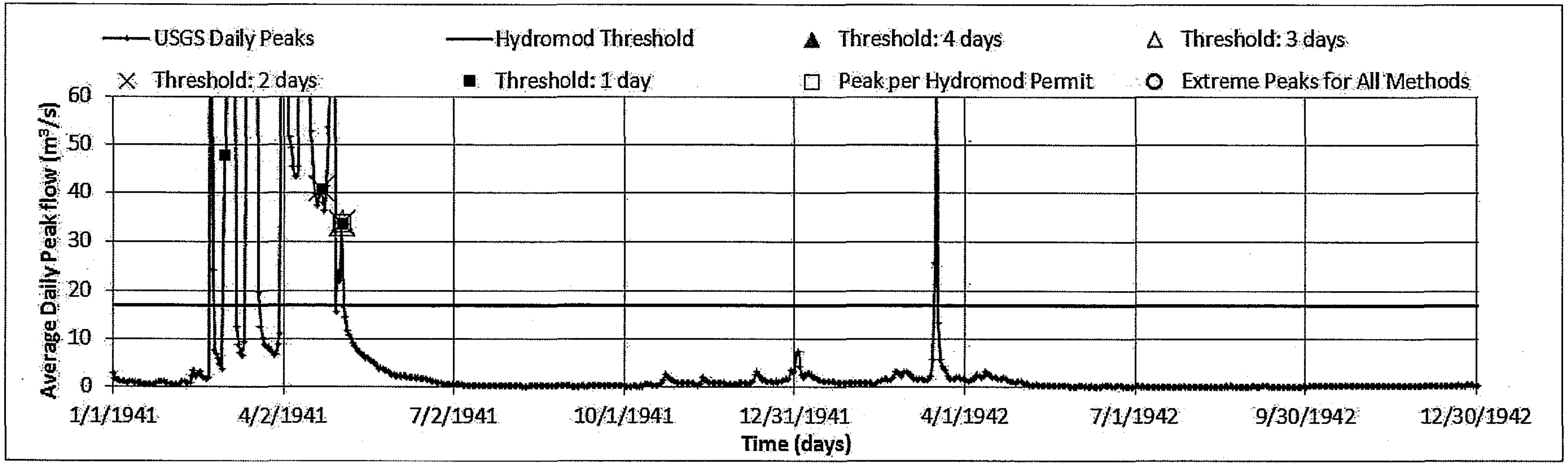


Figure A.1-3a. Peak Flow Time Series. 01/01/41 to 12/31/42.

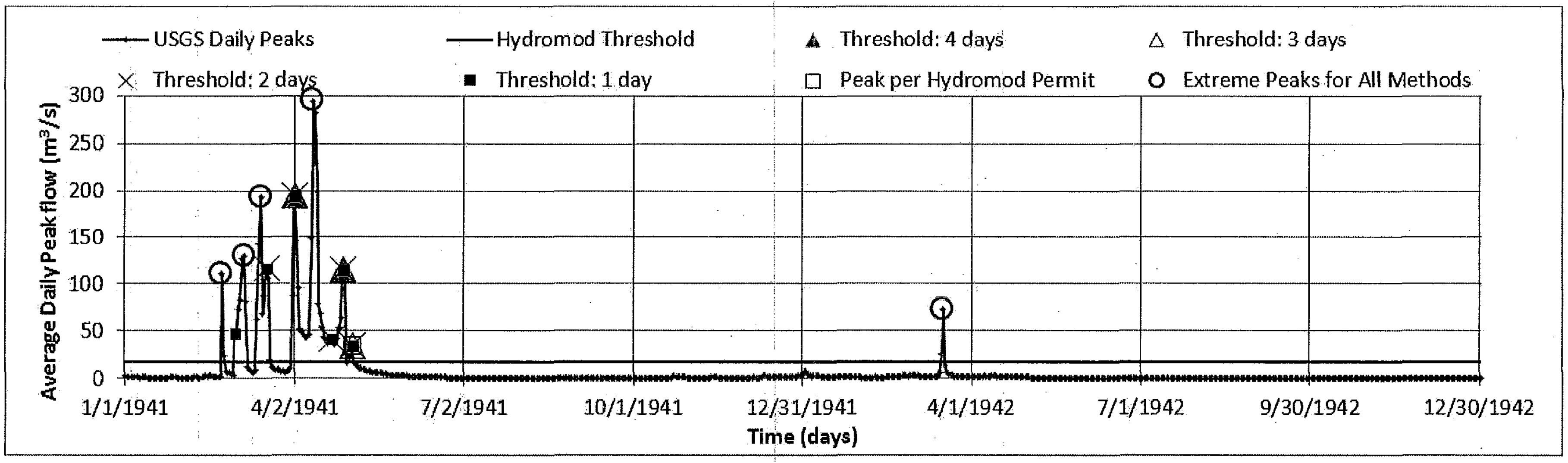


Figure A.1-3b. Peak Flow Time Series. 01/01/41 to 12/31/42. (Notice the change on vertical scale to capture extreme events).

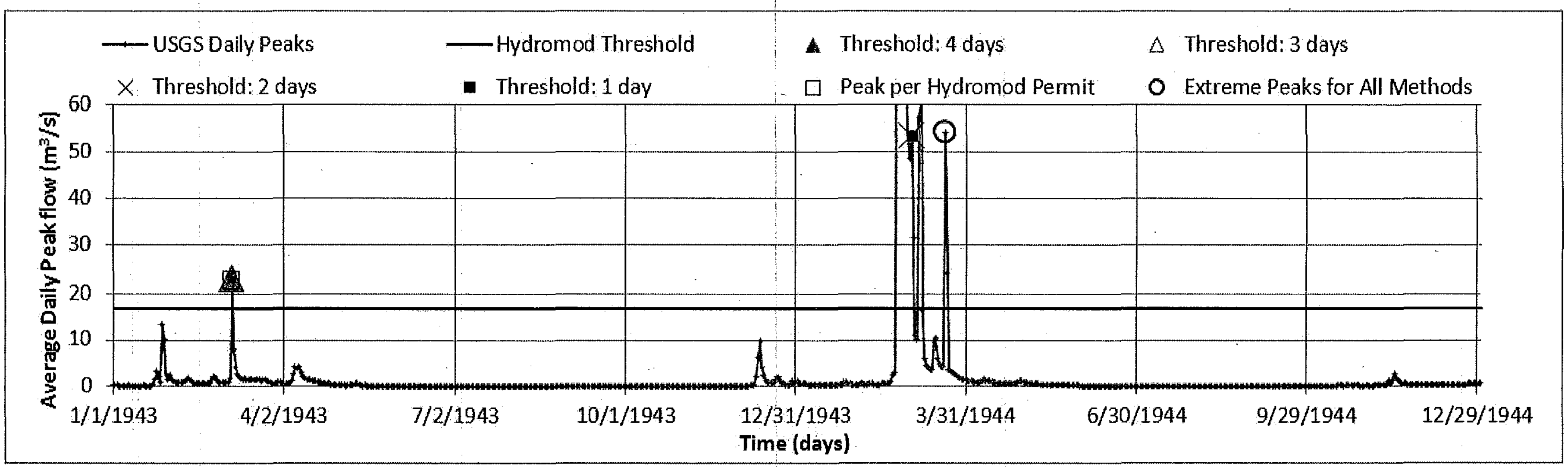


Figure A.1-4a. Peak Flow Time Series. 01/01/43 to 12/31/44.

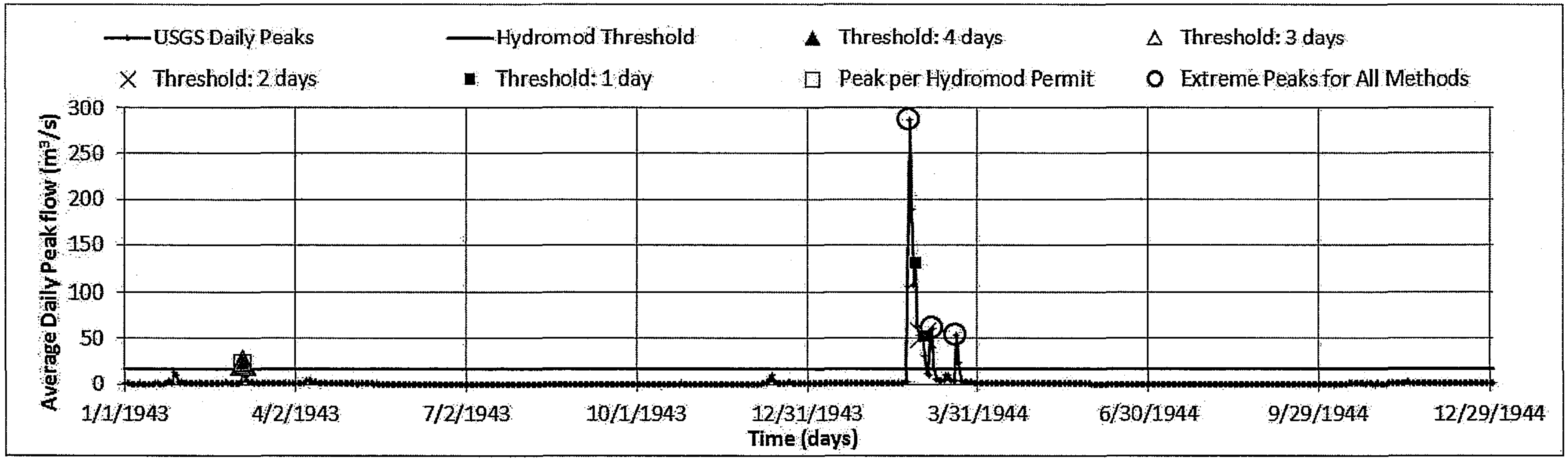


Figure A.1-4b. Peak Flow Time Series. 01/01/43 to 12/31/44. (Notice the change on vertical scale to capture extreme events).

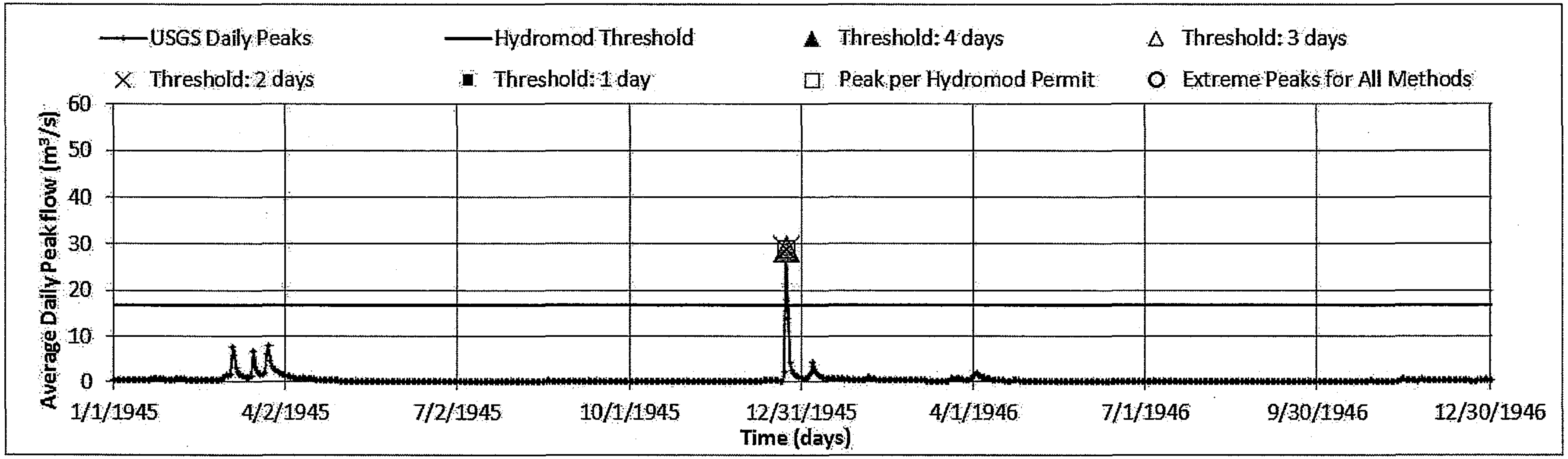


Figure A.1-5. Peak Flow Time Series. 01/01/45 to 12/31/46.

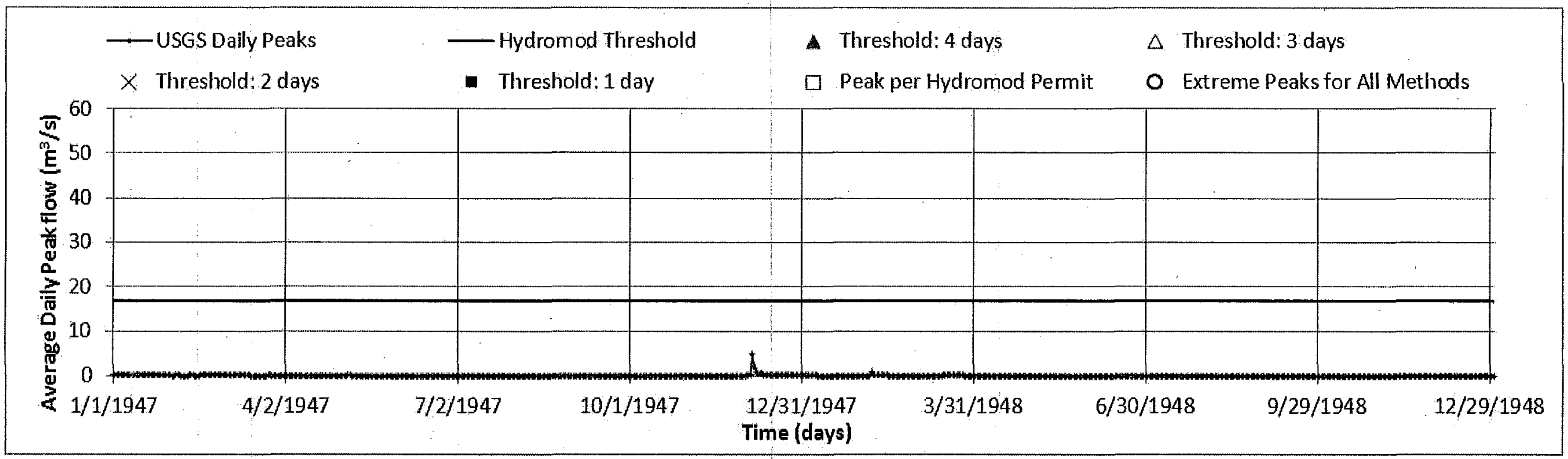


Figure A.1-6. Peak Flow Time Series. 01/01/47 to 12/31/48.

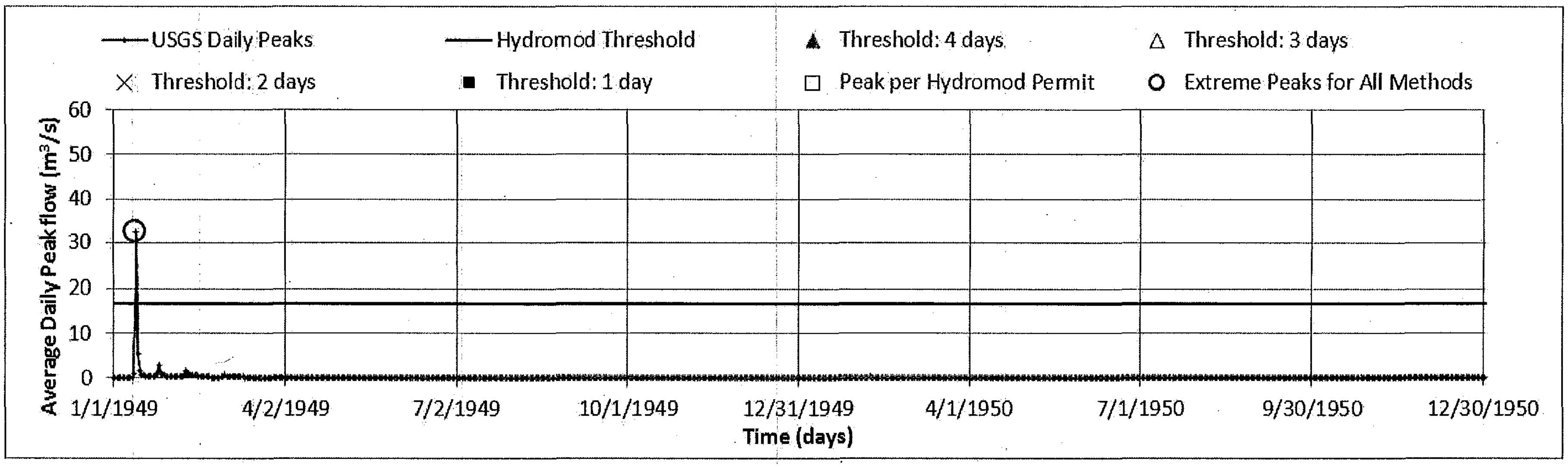


Figure A.1-7. Peak Flow Time Series. 01/01/49 to 12/31/50.

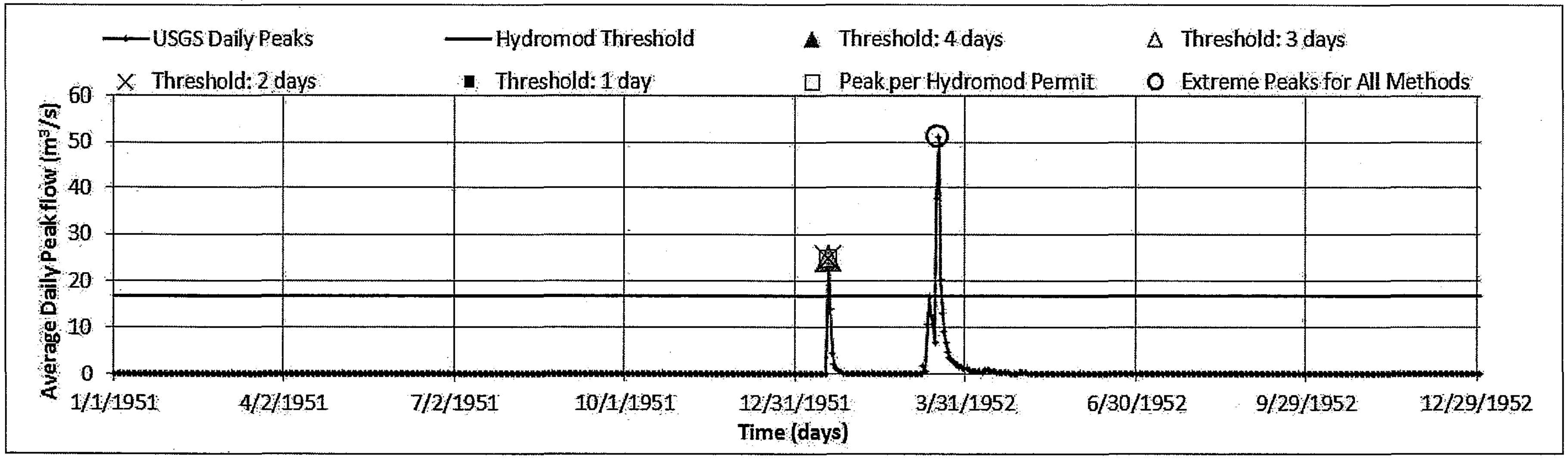


Figure A.1-8. Peak Flow Time Series. 01/01/51 to 12/31/52.

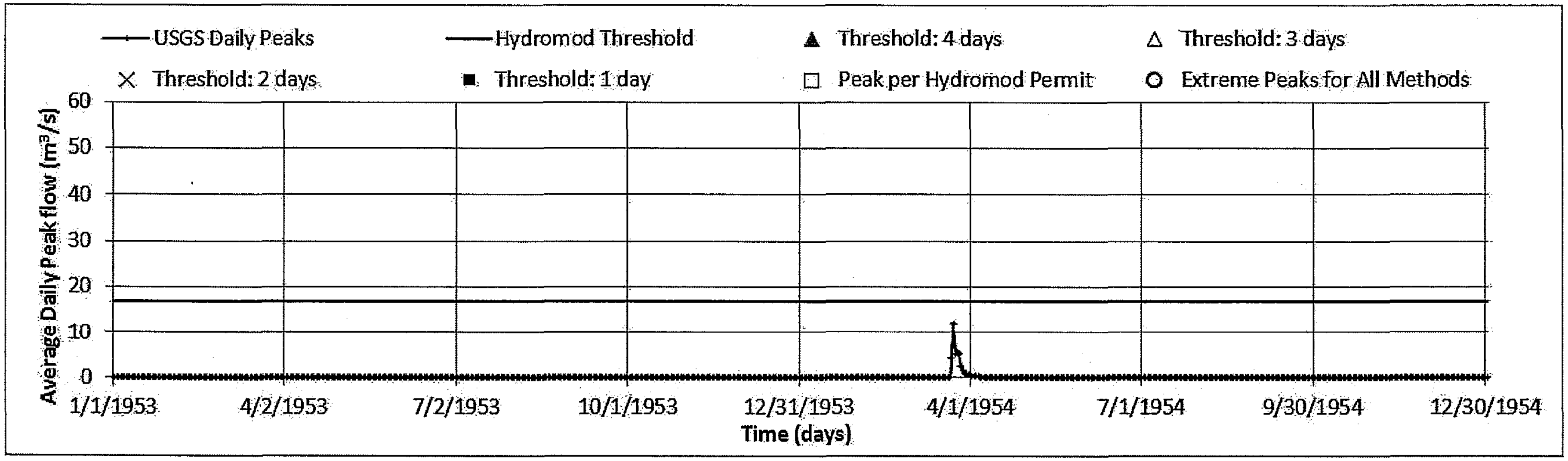


Figure A.1-9. Peak Flow Time Series. 01/01/53 to 12/31/54.

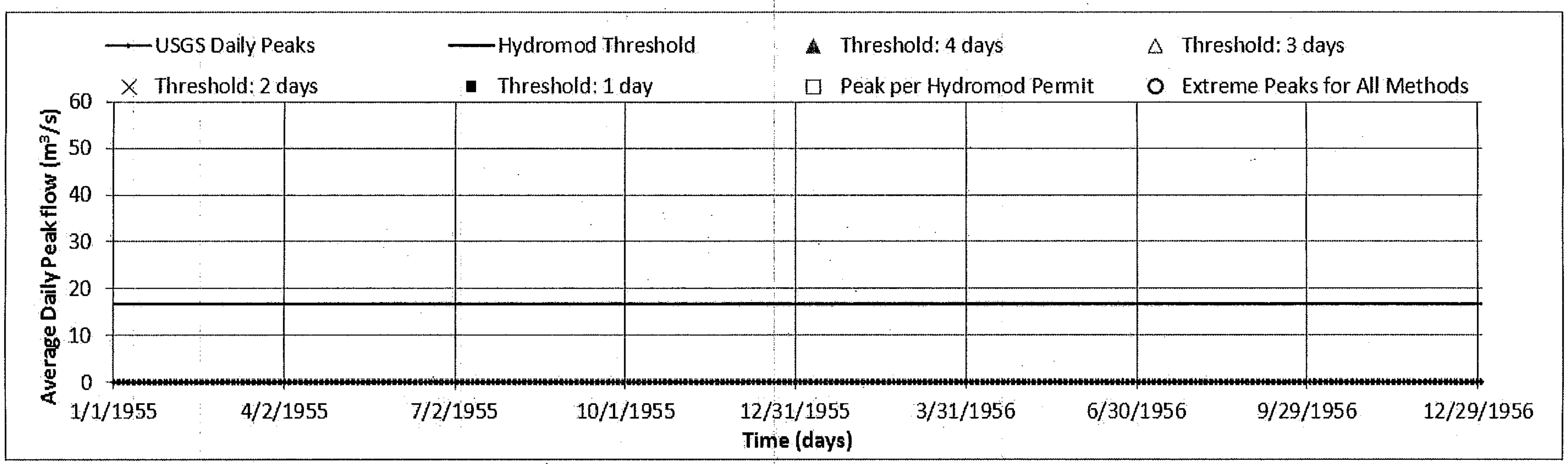


Figure A.1-10. Peak Flow Time Series. 01/01/55 to 12/31/56.

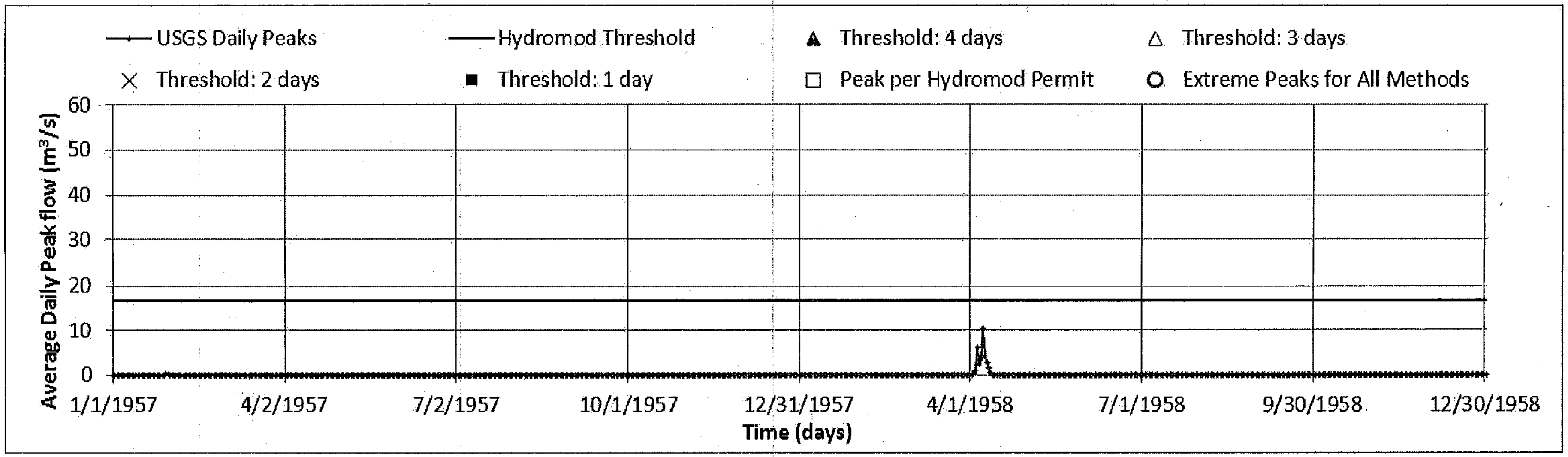


Figure A.1-11. Peak Flow Time Series. 01/01/57 to 12/31/58.

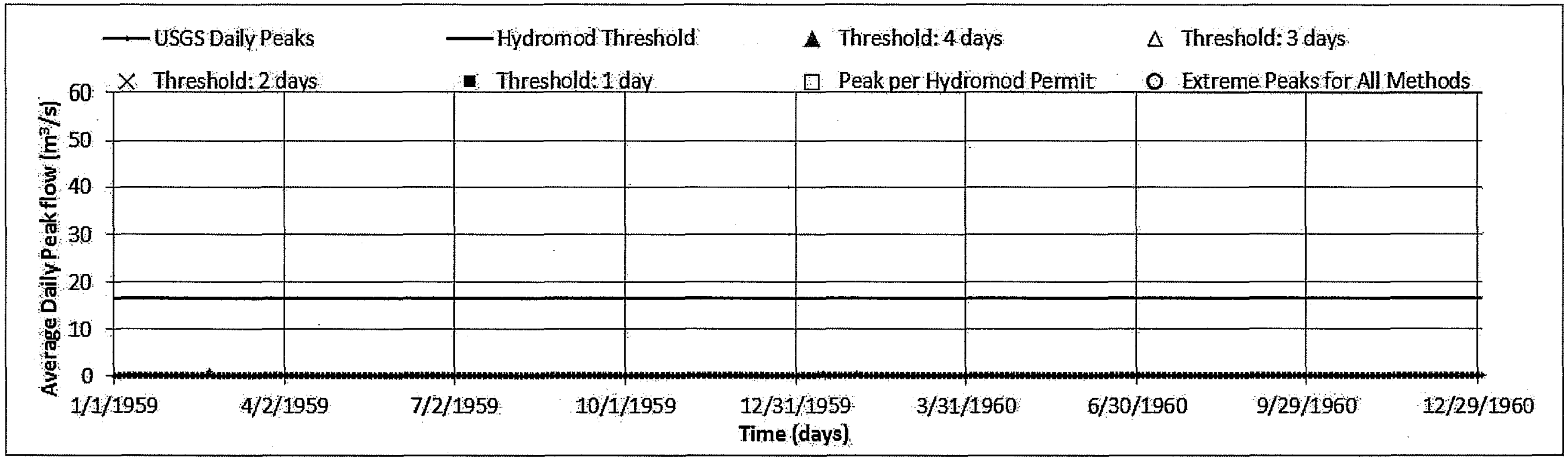


Figure A.1-12. Peak Flow Time Series. 01/01/59 to 12/31/60.

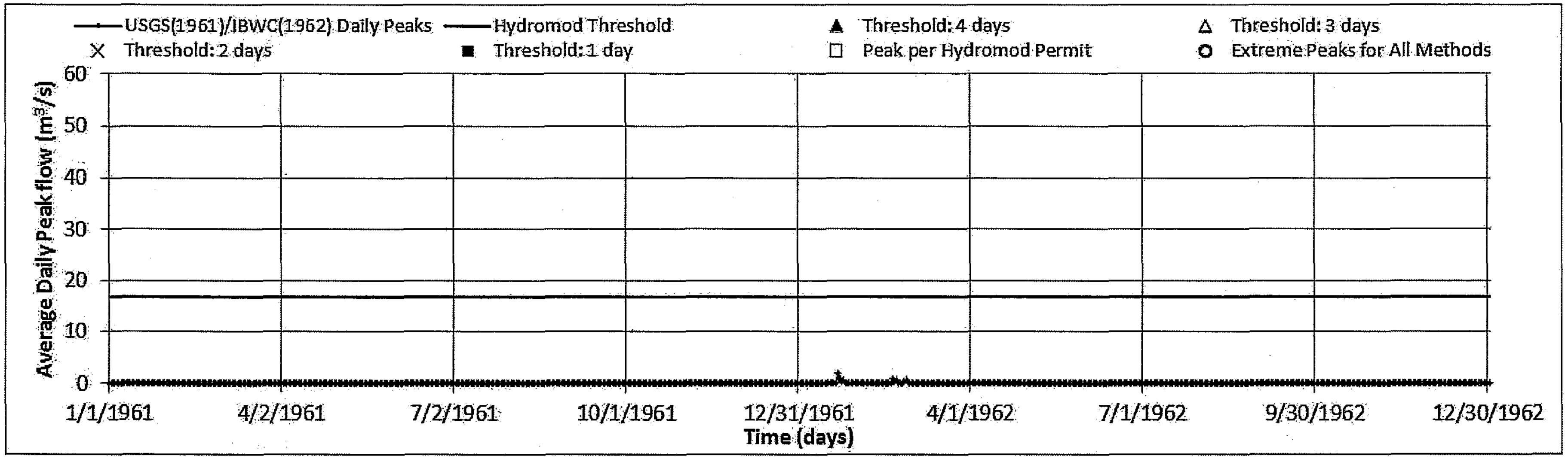


Figure A.1-13. Peak Flow Time Series. 01/01/61 to 12/31/62.

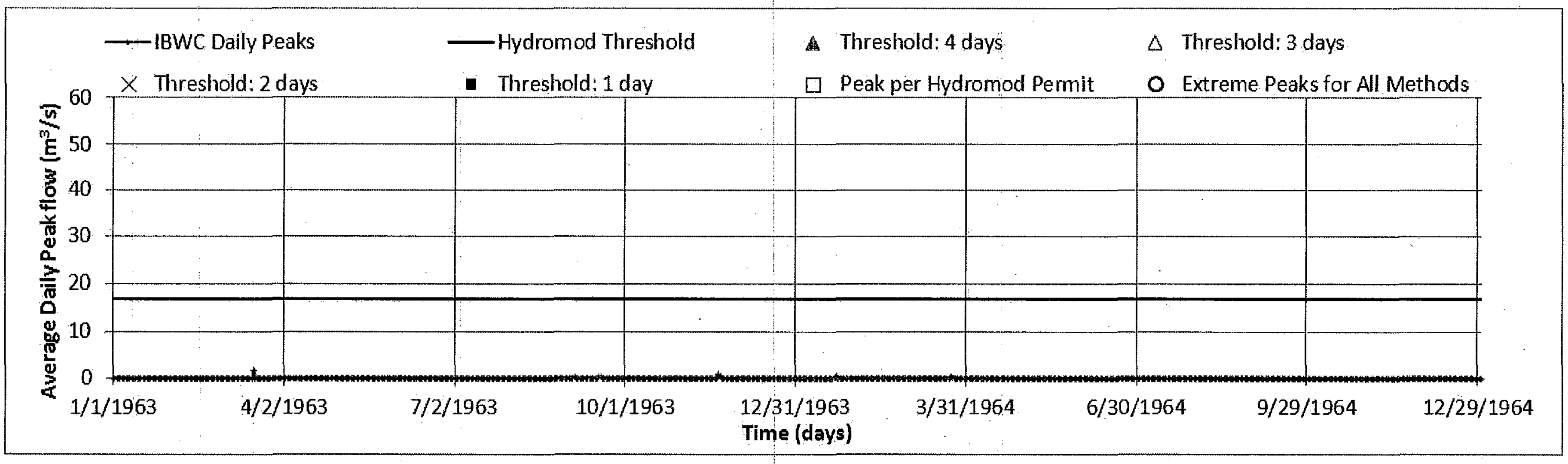


Figure A.1-14. Peak Flow Time Series. 01/01/63 to 12/31/64.

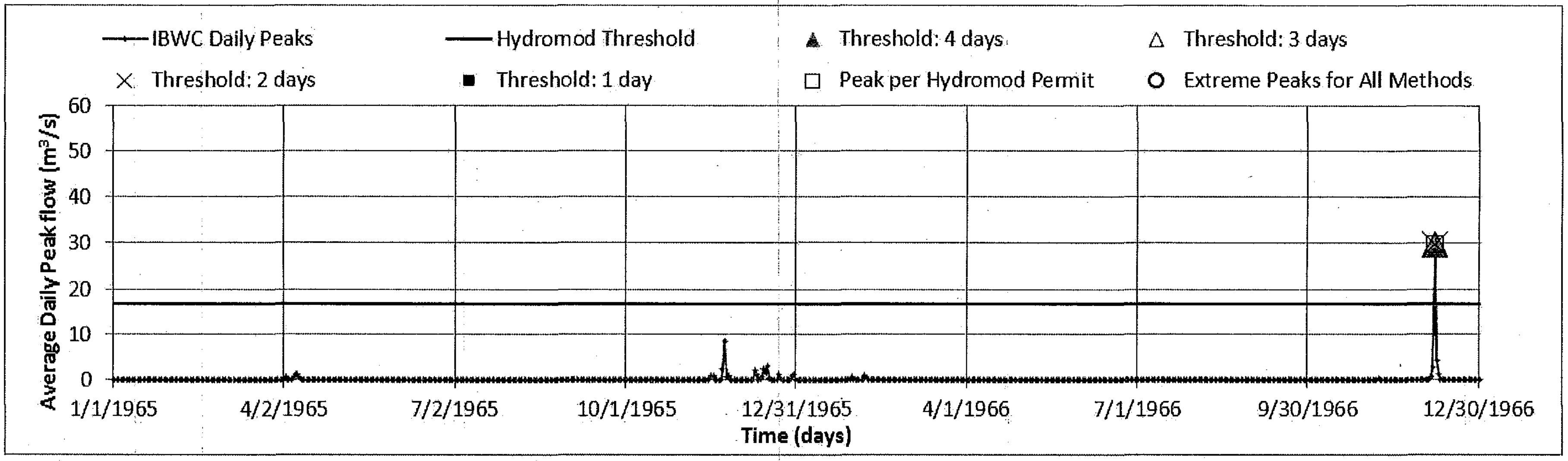


Figure A.1-15. Peak Flow Time Series. 01/01/65 to 12/31/66.

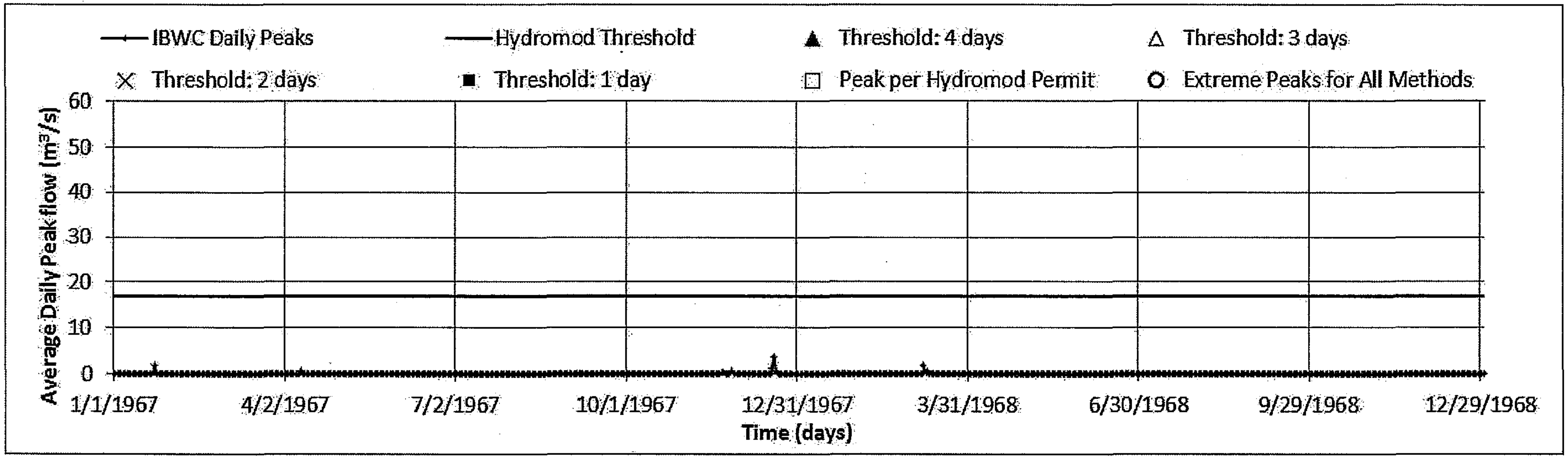


Figure A.1-16. Peak Flow Time Series. 01/01/67 to 12/31/68.

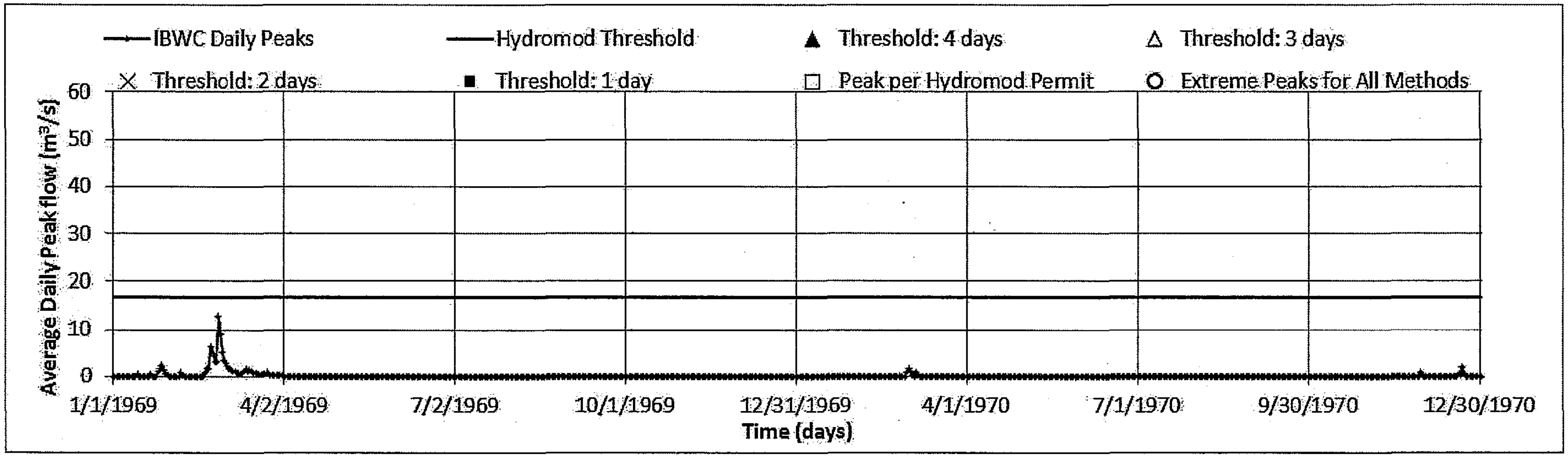


Figure A.1-17. Peak Flow Time Series. 01/01/69 to 12/31/70.

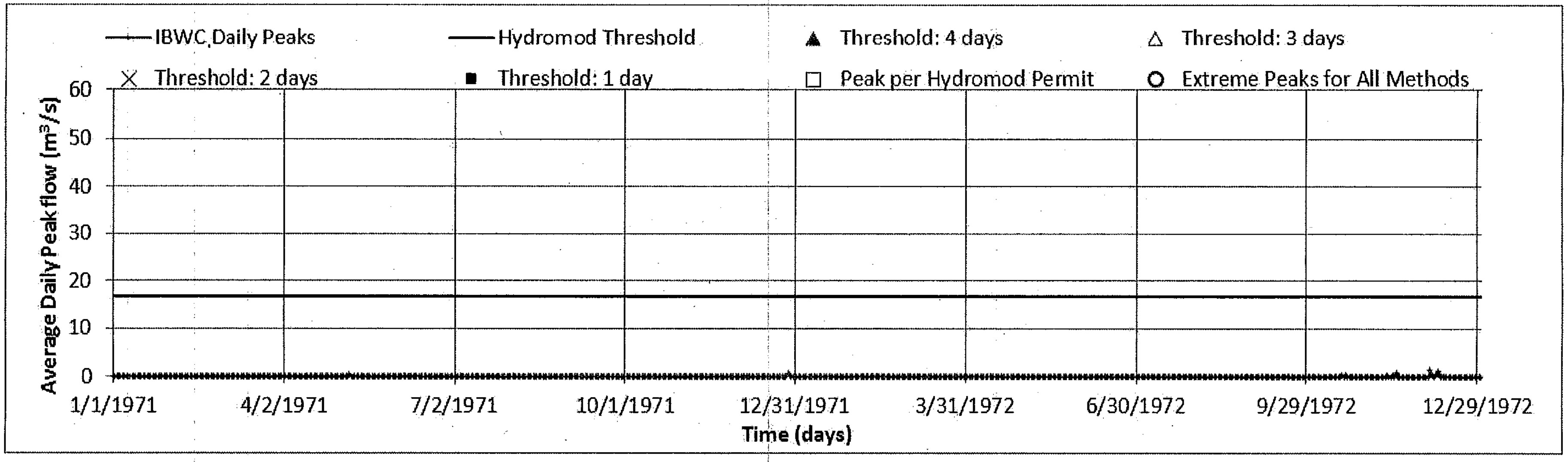


Figure A.1-18. Peak Flow Time Series. 01/01/71 to 12/31/72.

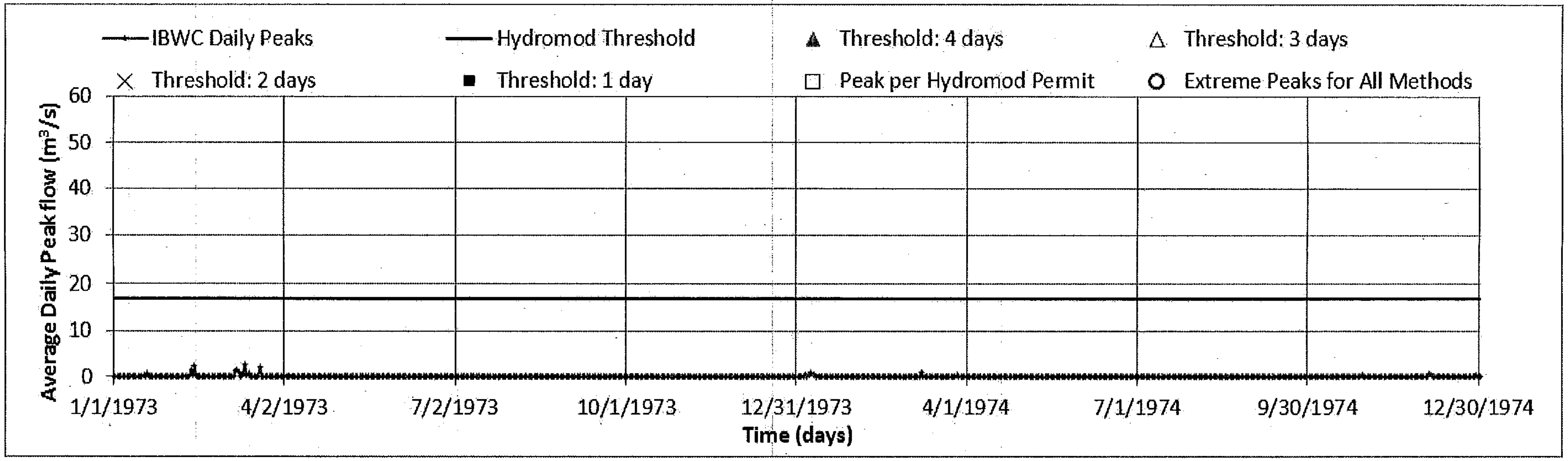


Figure A.1-19. Peak Flow Time Series. 01/01/73 to 12/31/74.

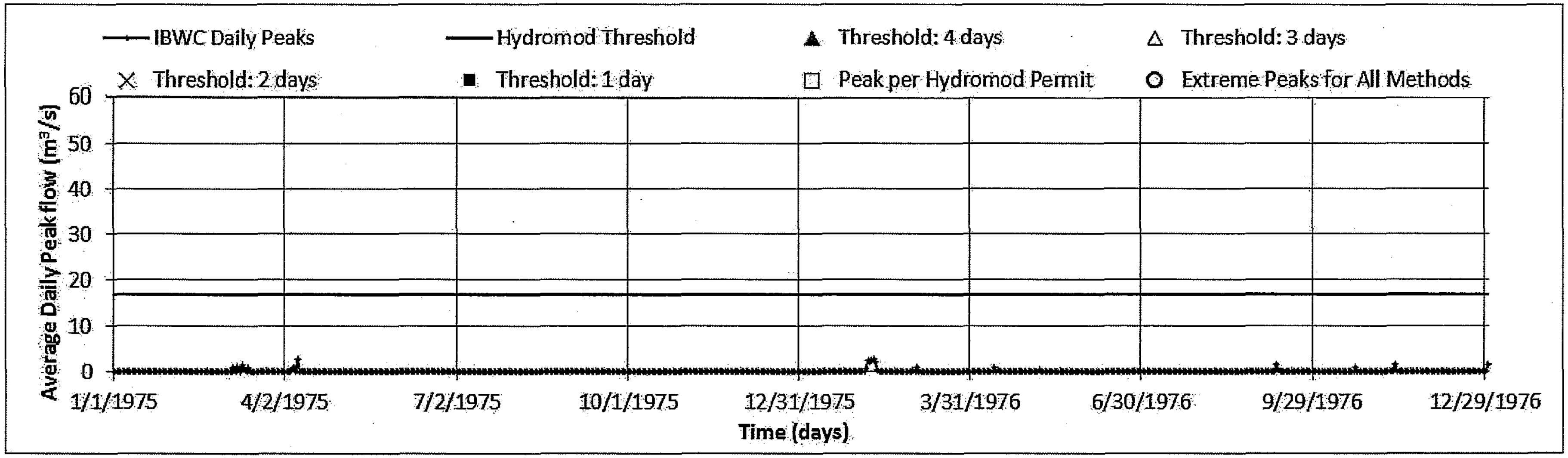


Figure A.1-20. Peak Flow Time Series. 01/01/75 to 12/31/76.

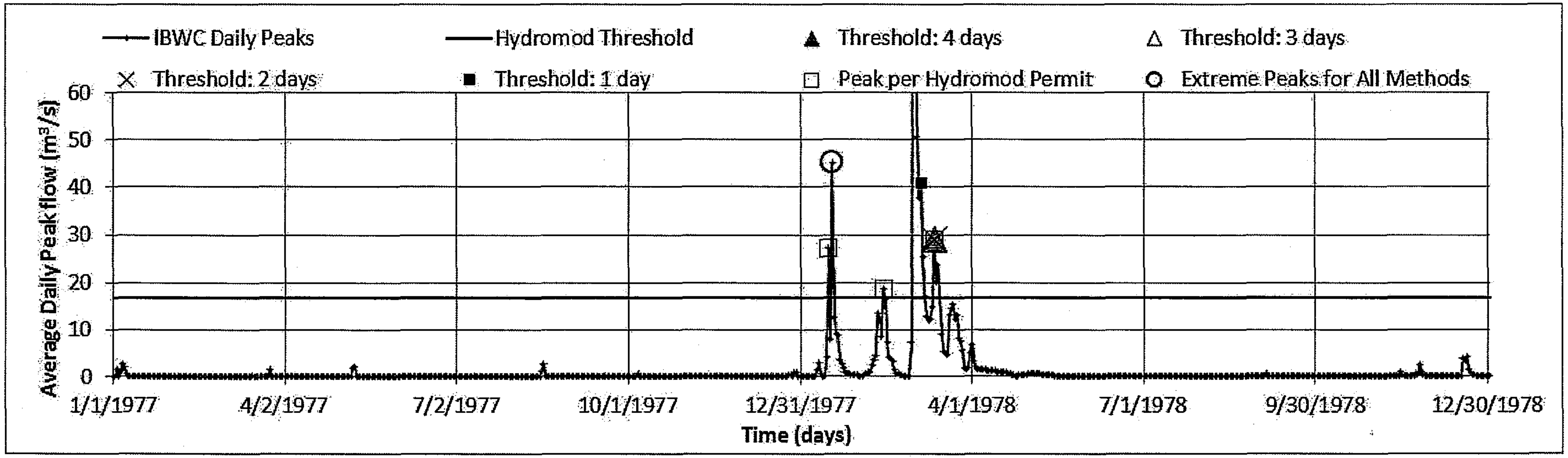


Figure A.1-21a. Peak Flow Time Series. 01/01/77 to 12/31/78.

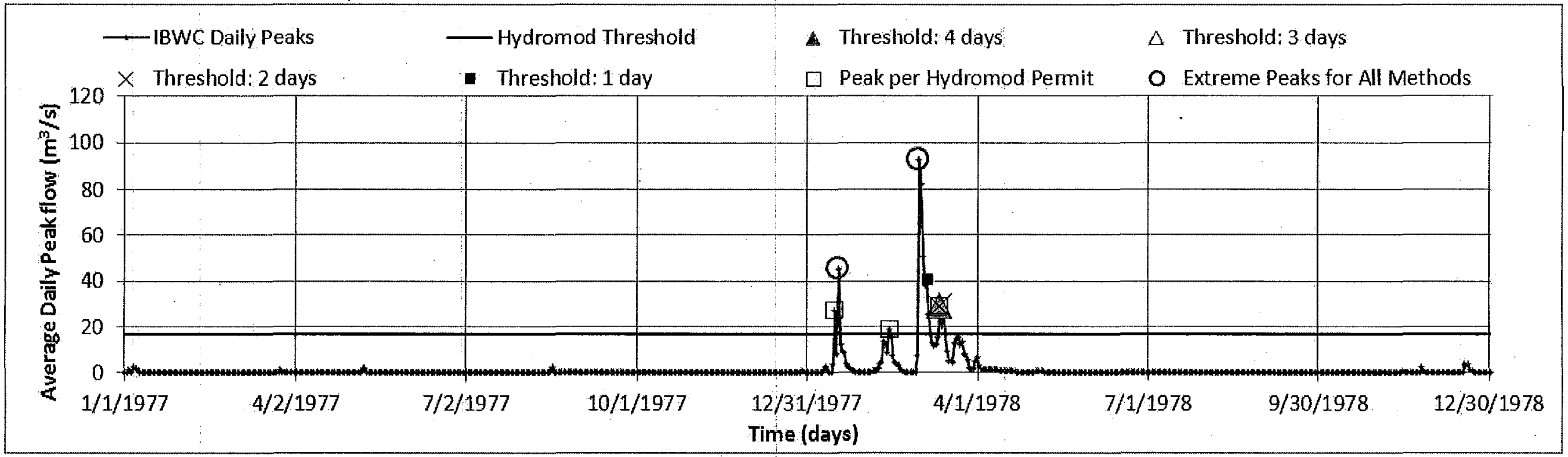


Figure A.1-21b. Peak Flow Time Series. 01/01/77 to 12/31/78. (Notice the change on vertical scale to capture extreme events).

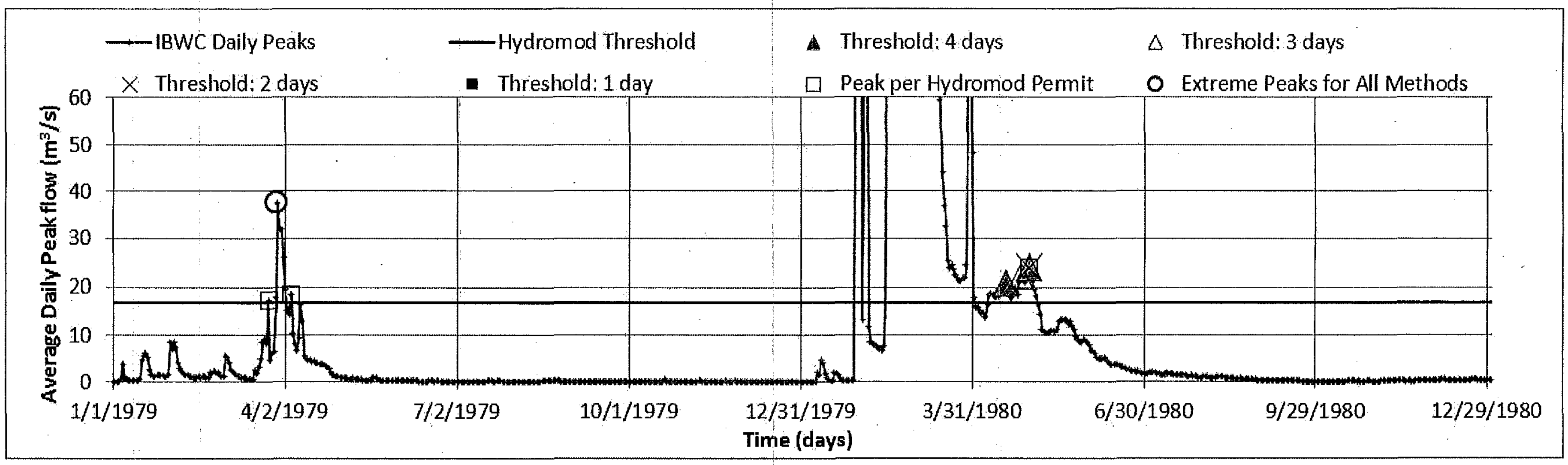


Figure A.1-22a. Peak Flow Time Series. 01/01/79 to 12/31/80.

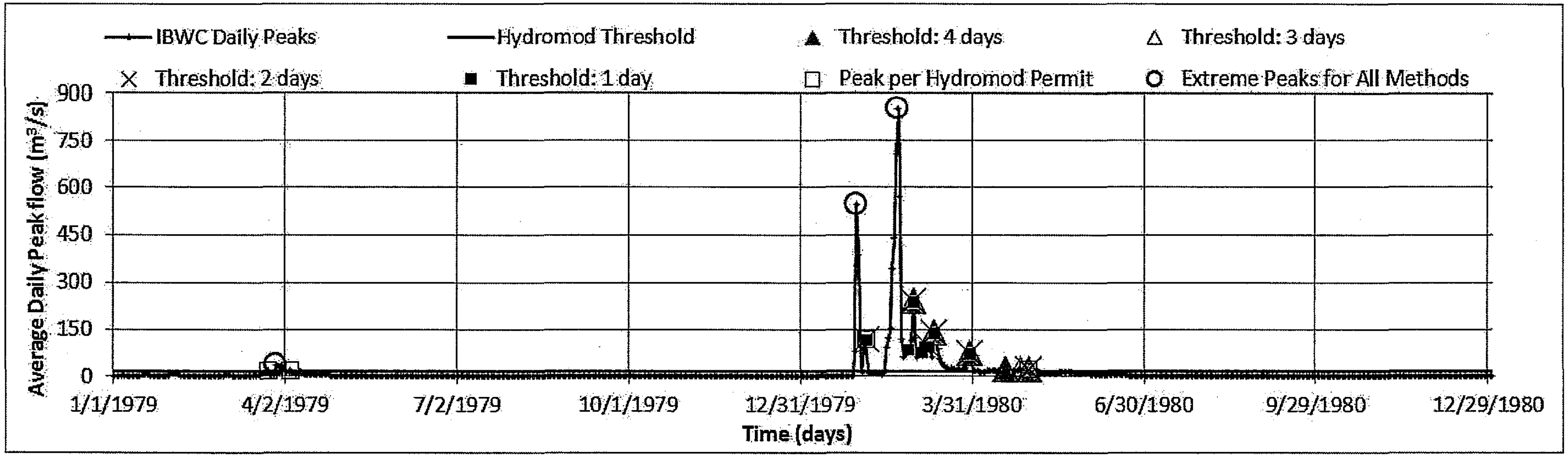


Figure A.1-22b. Peak Flow Time Series. 01/01/79 to 12/31/80. (Notice the change on vertical scale to capture extreme events).

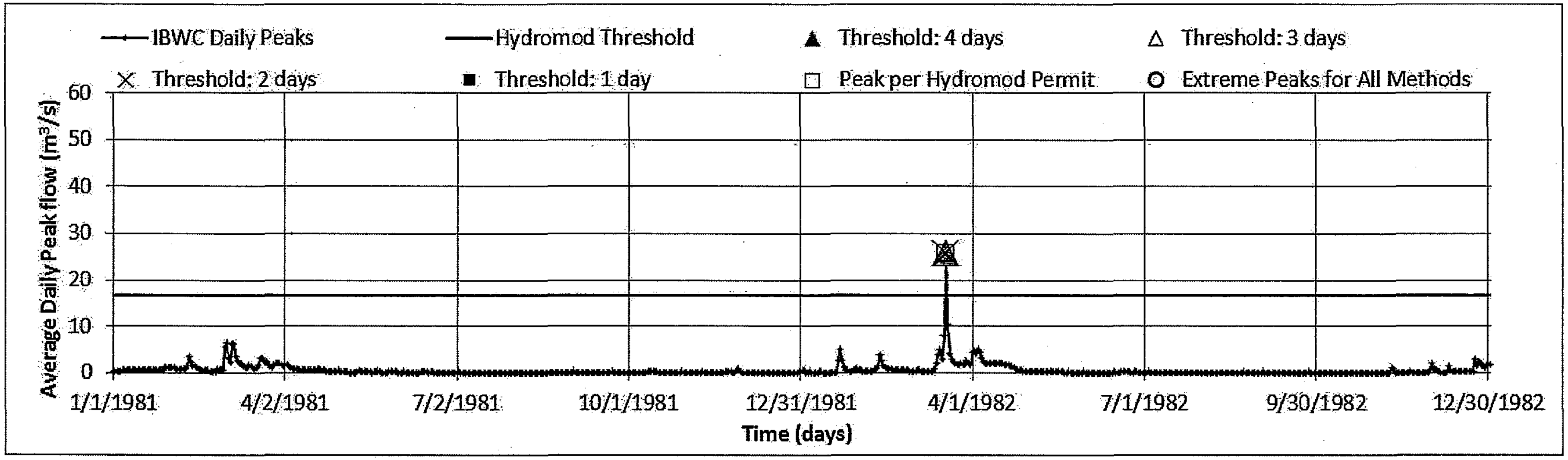


Figure A.1-23. Peak Flow Time Series. 01/01/81 to 12/31/82.

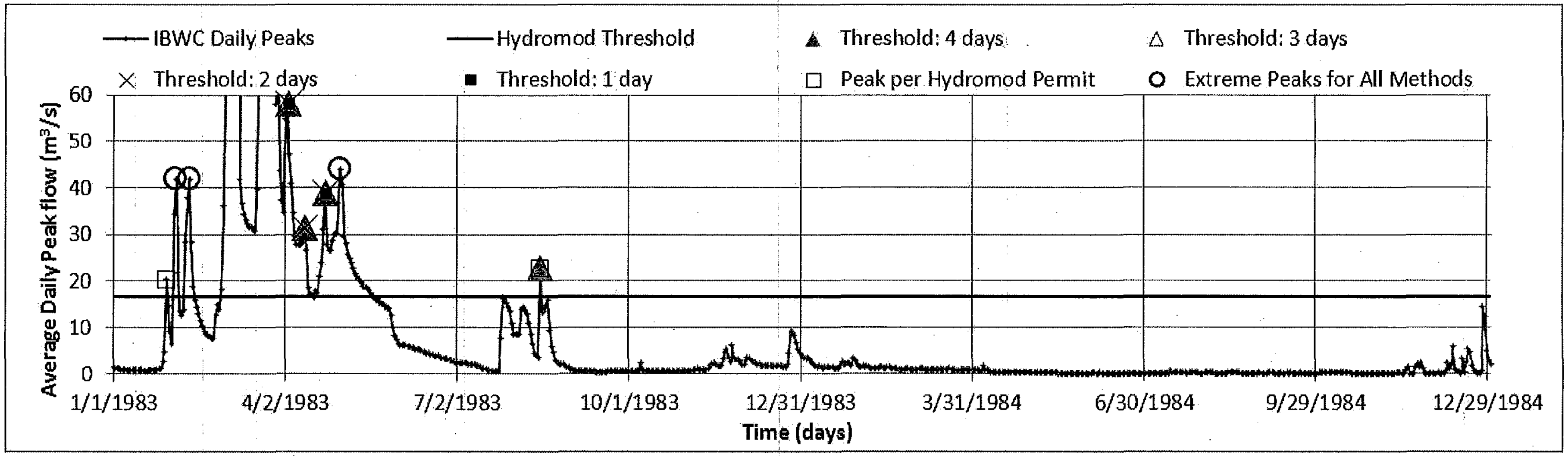


Figure A.1-24a. Peak Flow Time Series. 01/01/83 to 12/31/84.

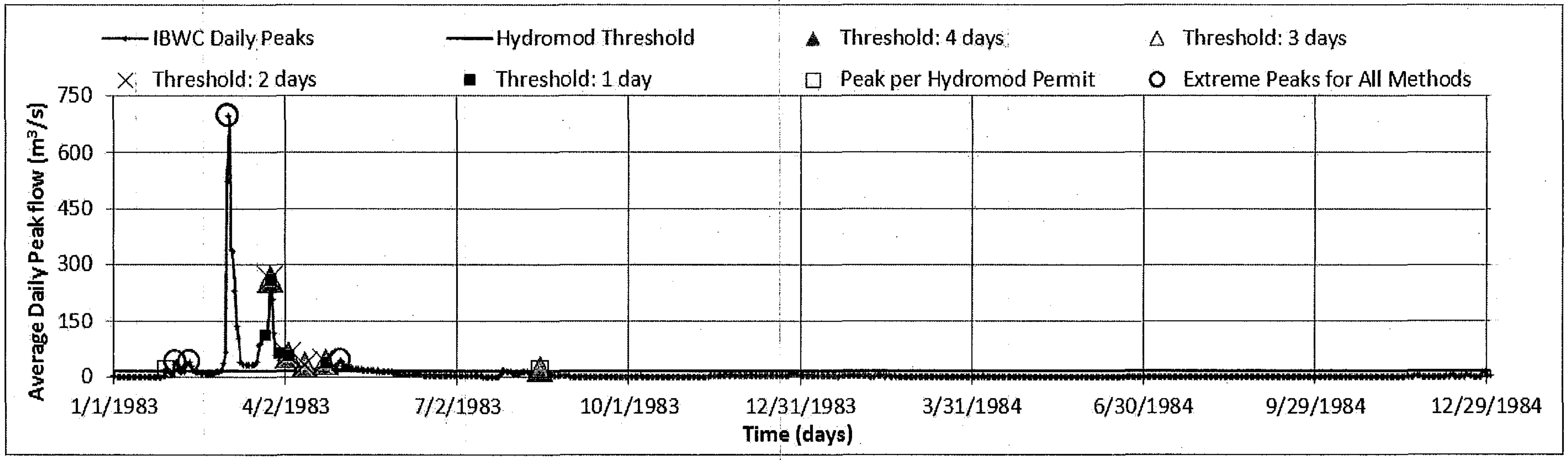


Figure A.1-24b. Peak Flow Time Series. 01/01/83 to 12/31/84. (Notice the change on vertical scale to capture extreme events).

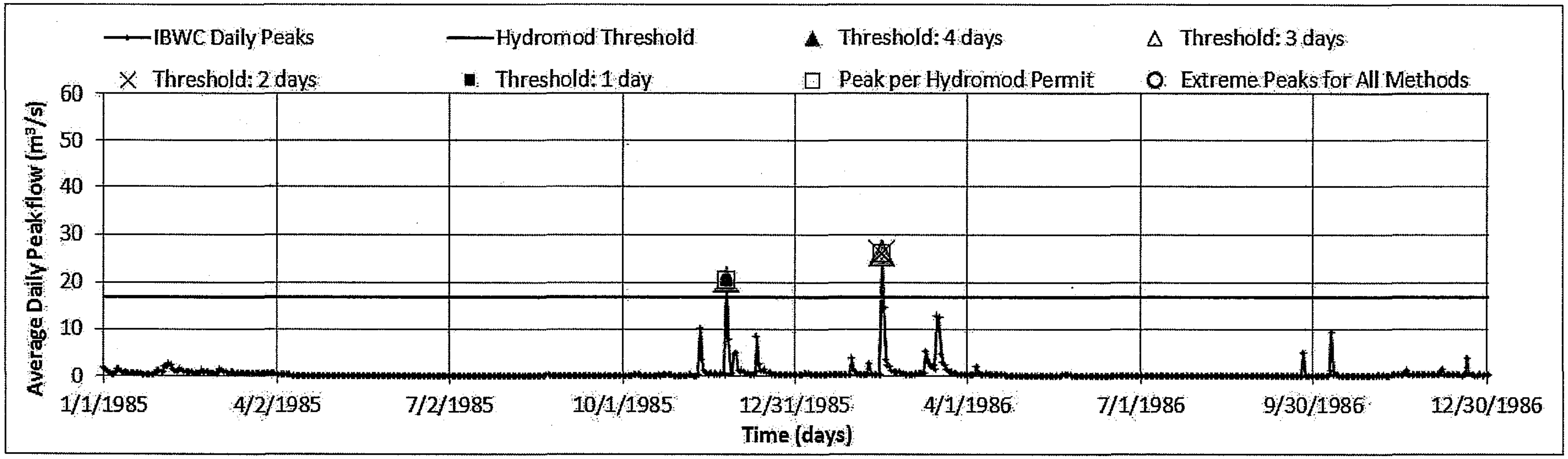


Figure A.1-25. Peak Flow Time Series. 01/01/85 to 12/31/86.

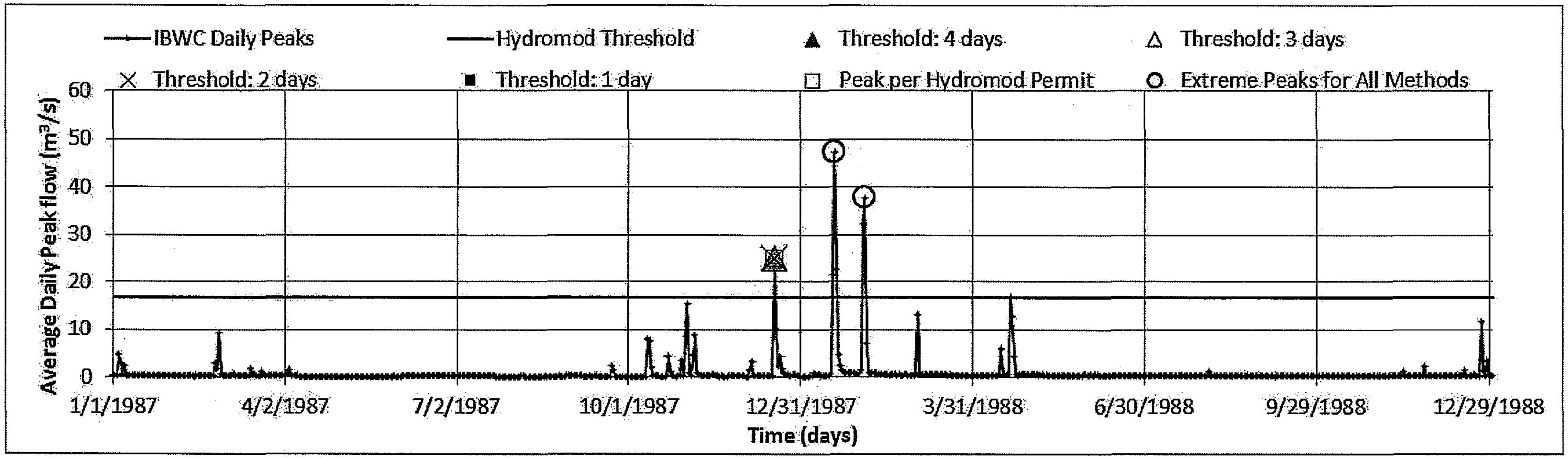


Figure A.1-26. Peak Flow Time Series. 01/01/87 to 12/31/88.

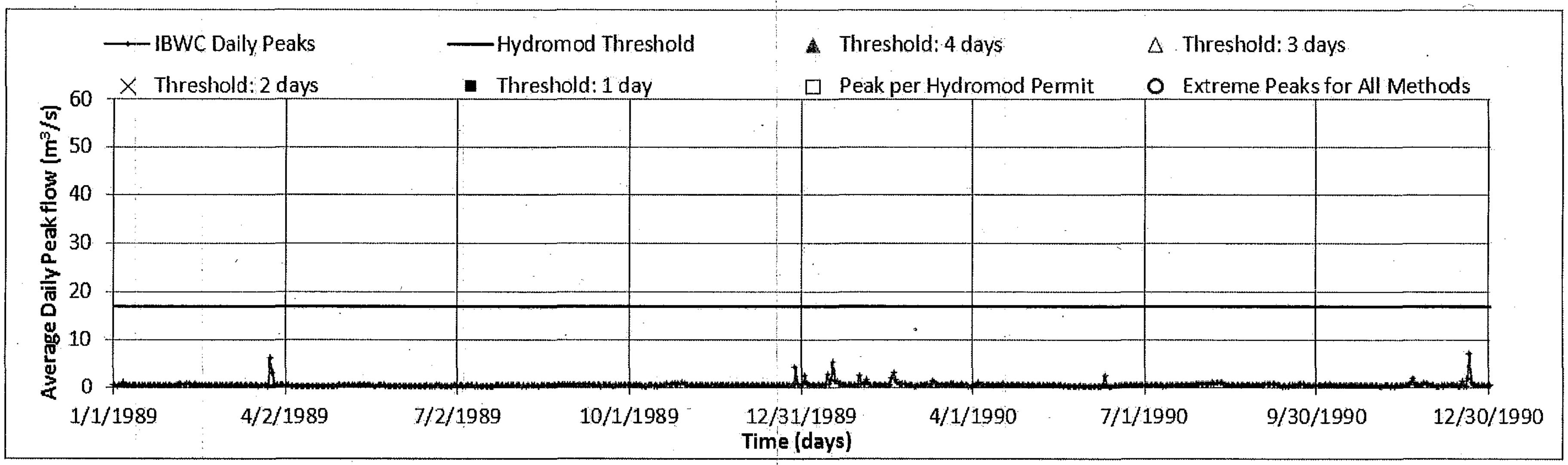


Figure A.1-27. Peak Flow Time Series. 01/01/89 to 12/31/90.

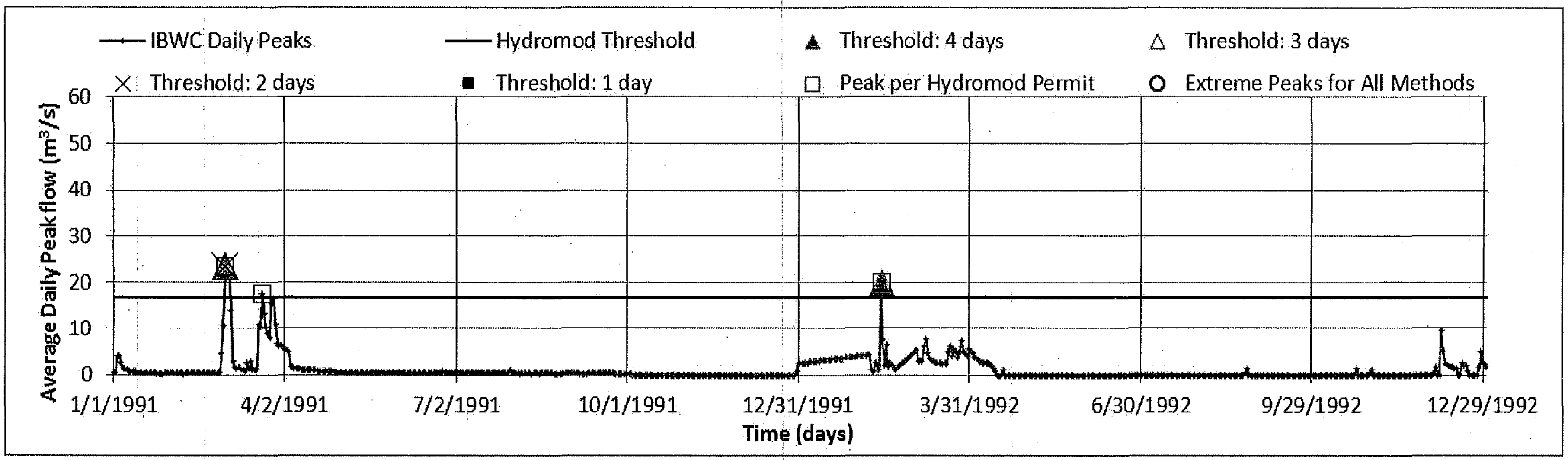


Figure A.1-28. Peak Flow Time Series. 01/01/91 to 12/31/92.

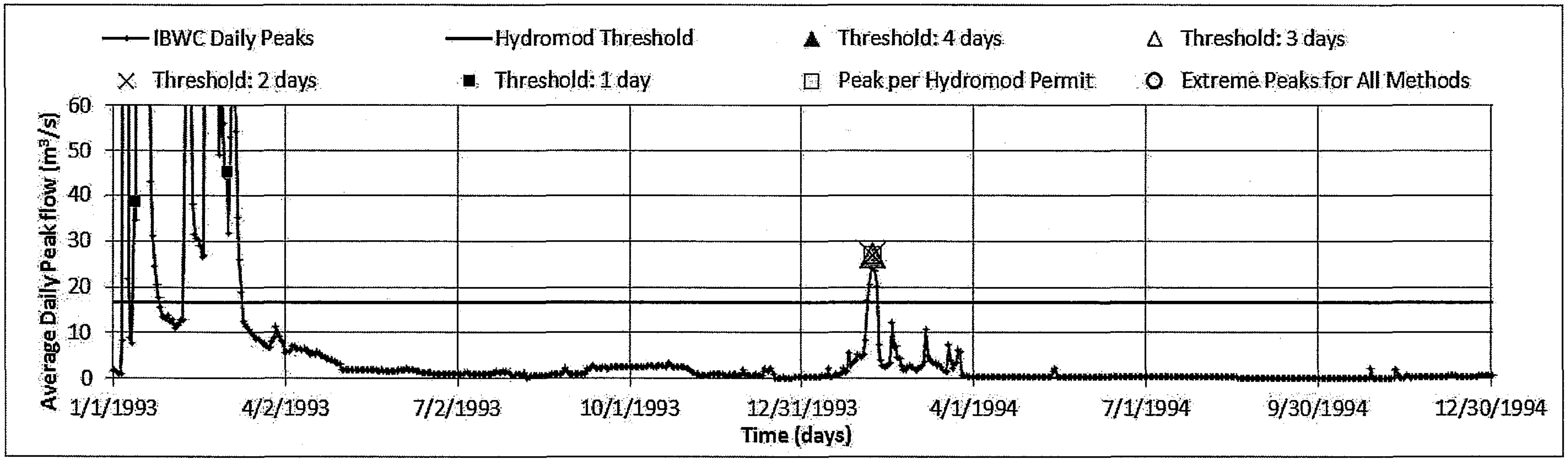


Figure A.1-29a. Peak Flow Time Series. 01/01/93 to 12/31/94.

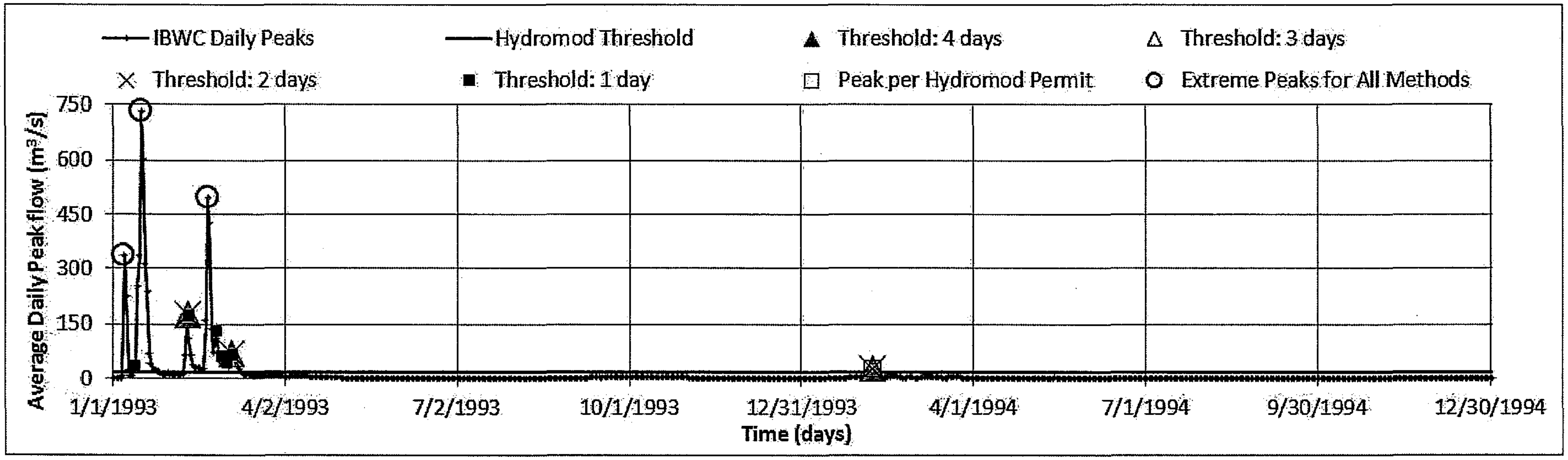


Figure A.1-29b. Peak Flow Time Series. 01/01/93 to 12/31/94. (Notice the change on vertical scale to capture extreme events).

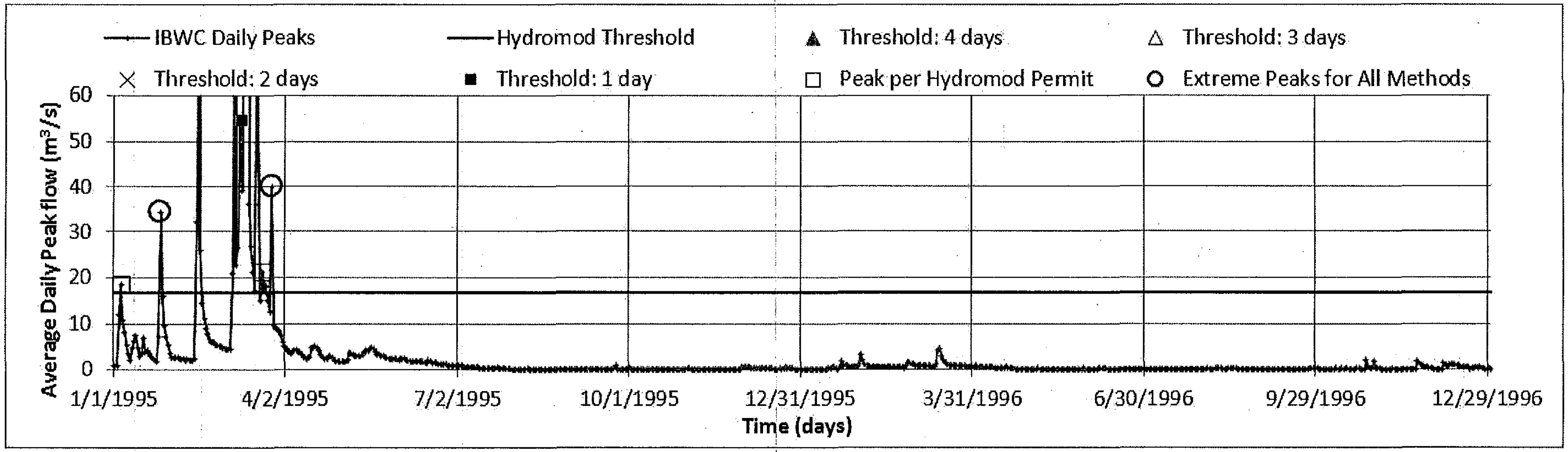


Figure A.1-30a. Peak Flow Time Series. 01/01/95 to 12/31/96.

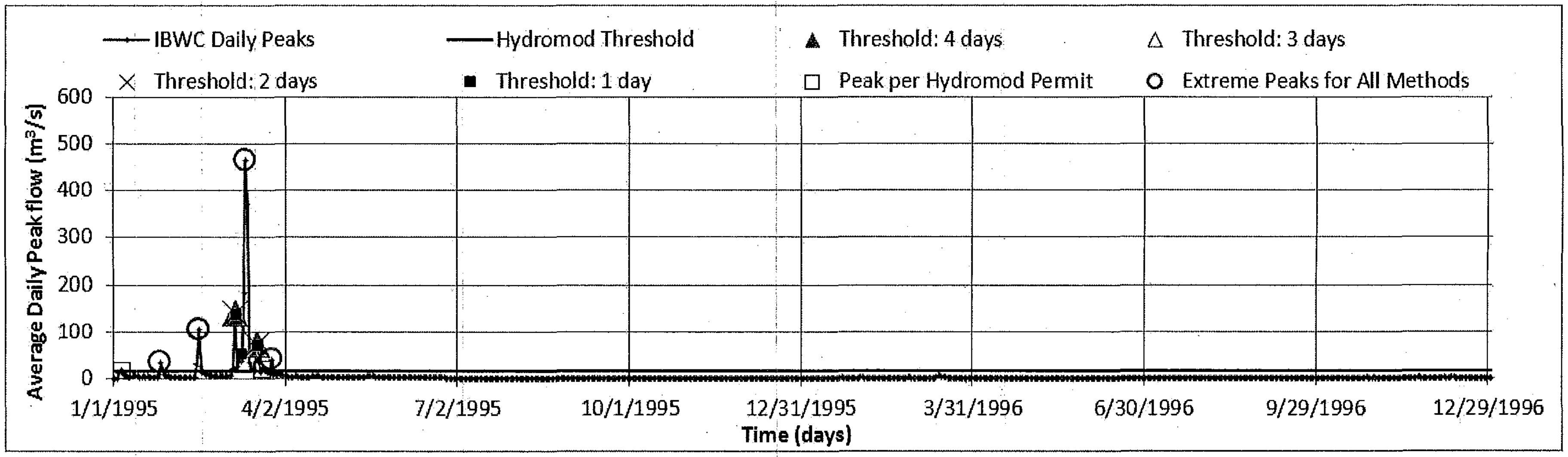


Figure A.1-30b. Peak Flow Time Series. 01/01/95 to 12/31/96. (Notice the change on vertical scale to capture extreme events).

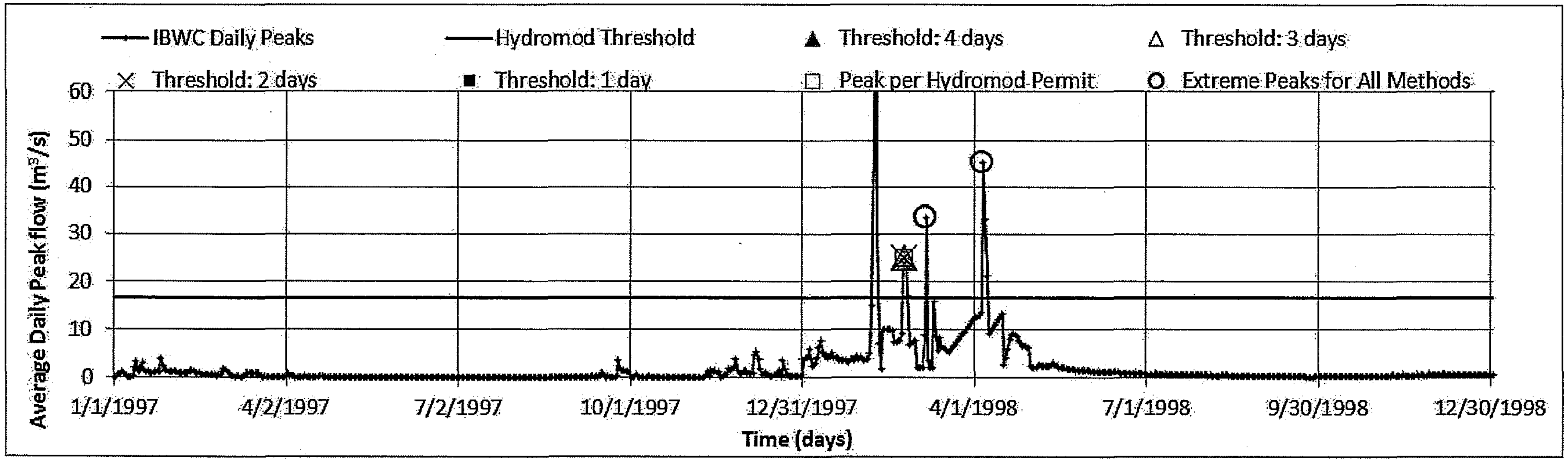


Figure A.1-31a. Peak Flow Time Series. 01/01/97 to 12/31/98.

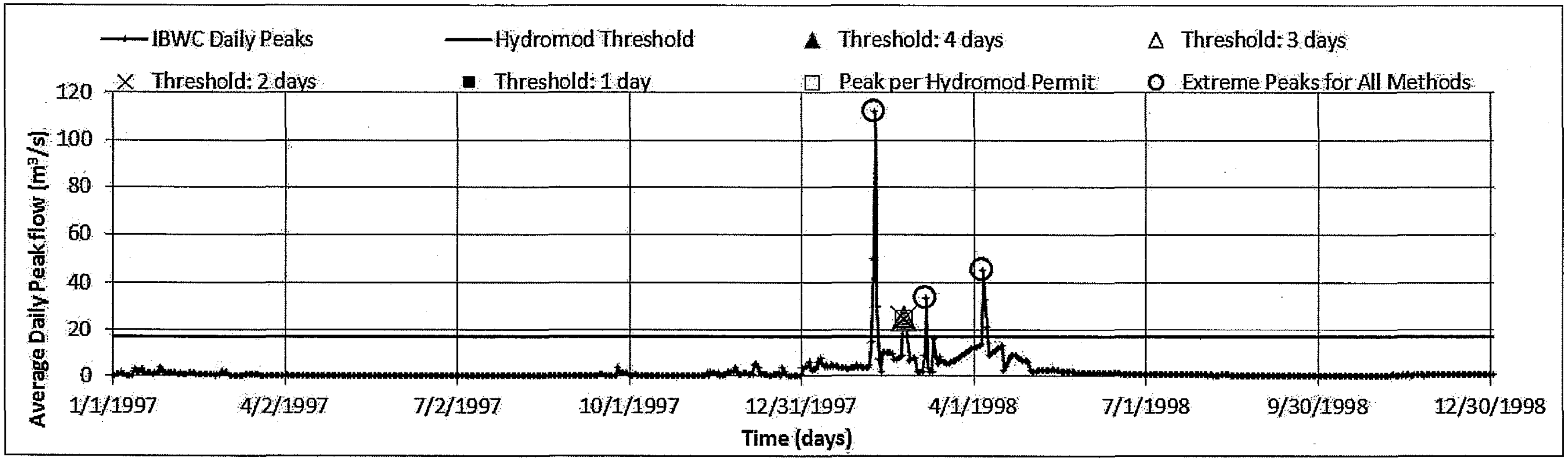


Figure A.1-31b. Peak Flow Time Series. 01/01/97 to 12/31/98. (Notice the change on vertical scale to capture extreme events).

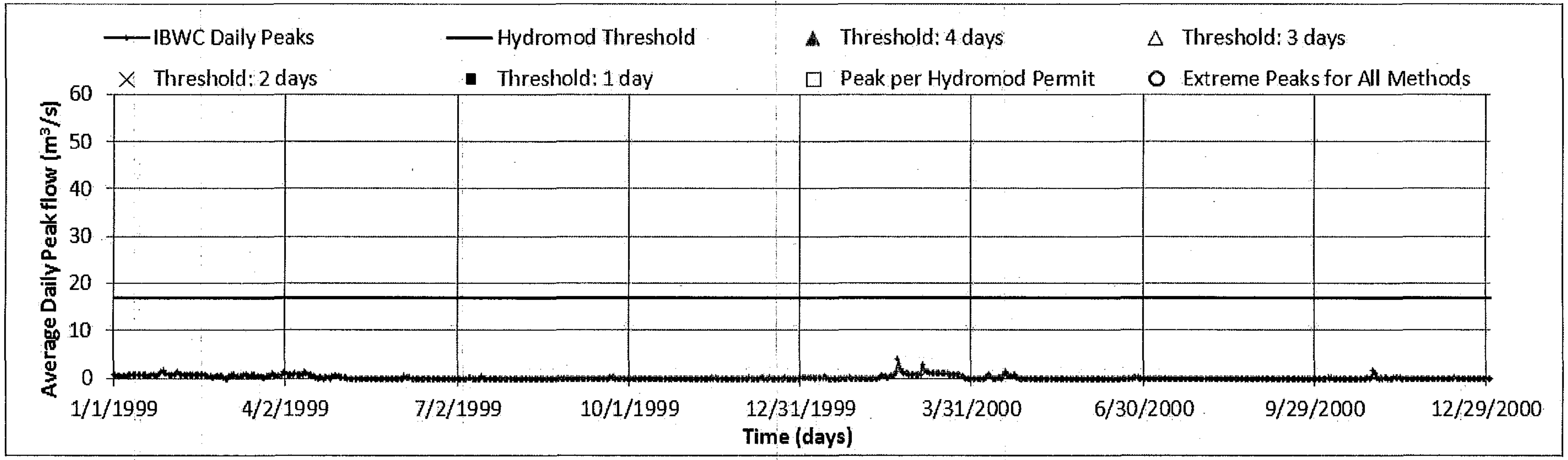


Figure A.1-32. Peak Flow Time Series. 01/01/99 to 12/31/00.

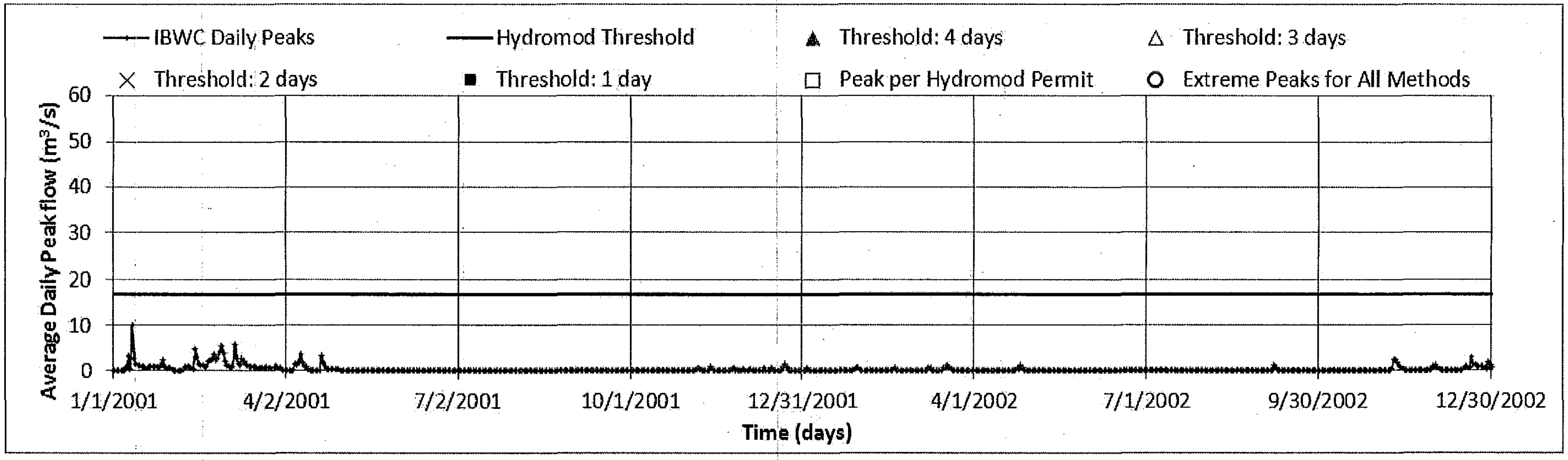


Figure A.1-33. Peak Flow Time Series. 01/01/01 to 12/31/02.

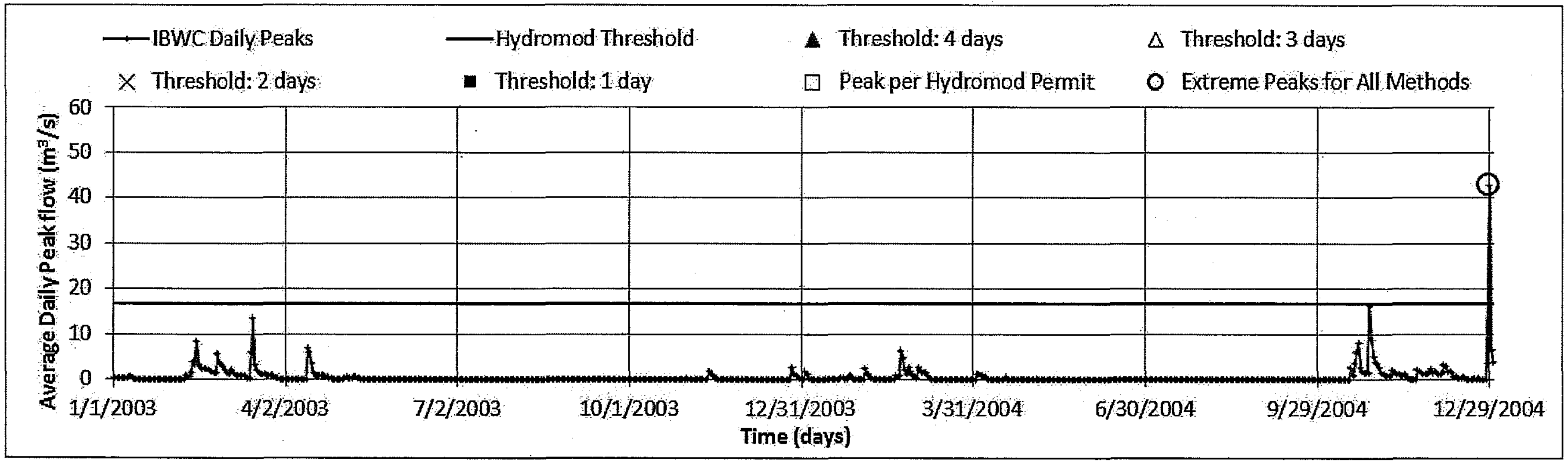


Figure A.1-34. Peak Flow Time Series. 01/01/03 to 12/31/04.

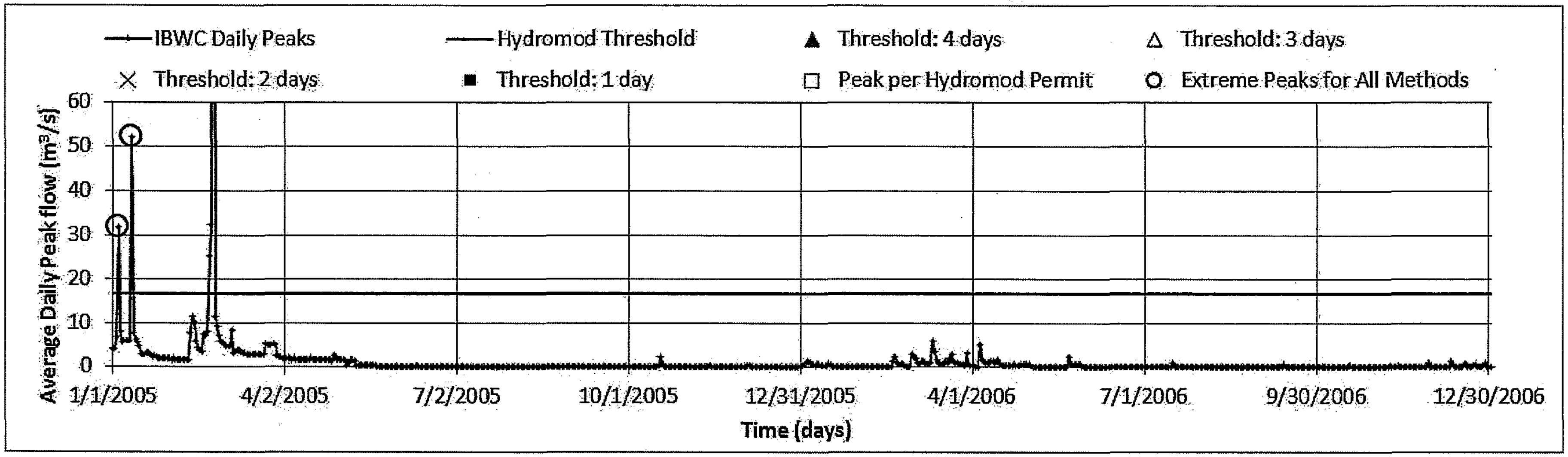


Figure A.1-35a. Peak Flow Time Series. 01/01/05 to 12/31/06.

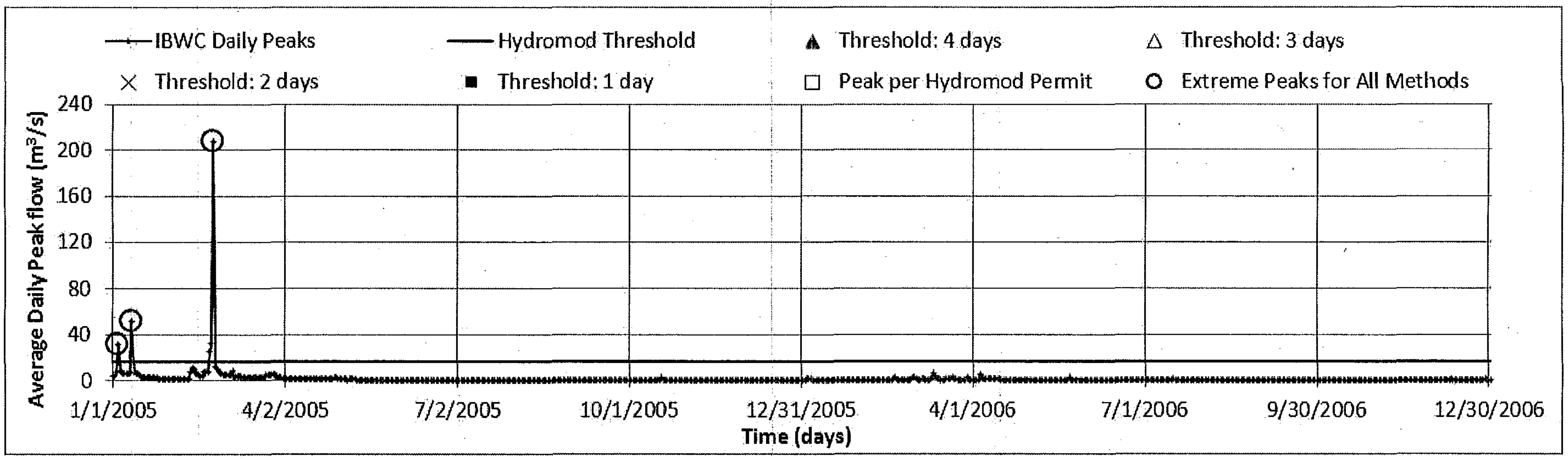


Figure A.1-35b. Peak Flow Time Series. 01/01/05 to 12/31/06. (Notice the change on vertical scale to capture extreme events).

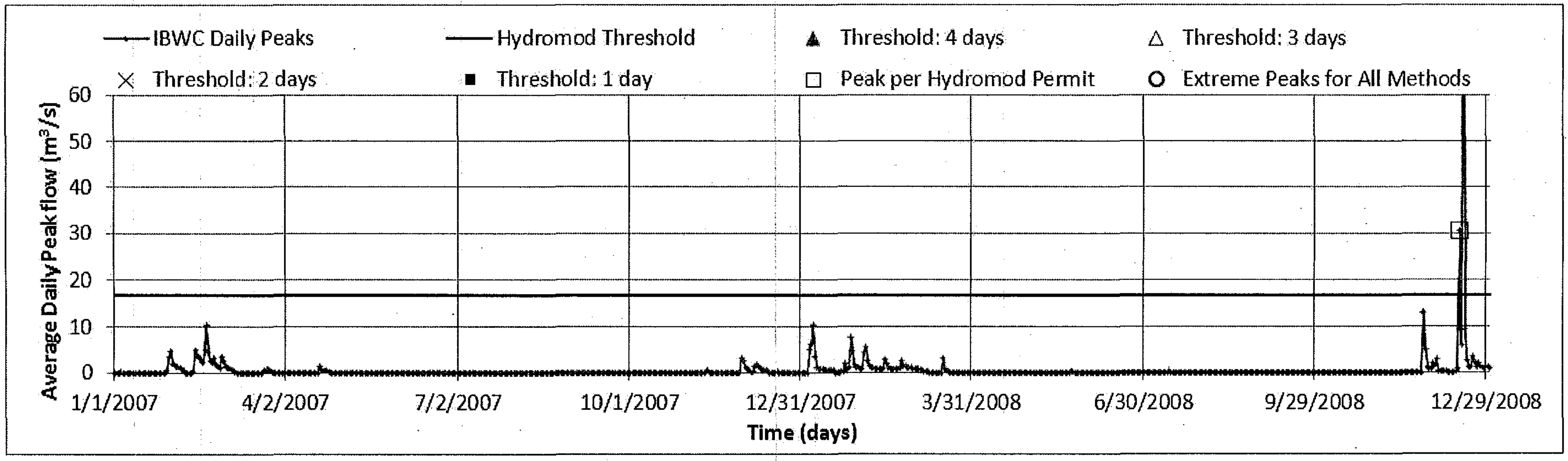


Figure A.1-36a. Peak Flow Time Series. 01/01/07 to 12/31/08.

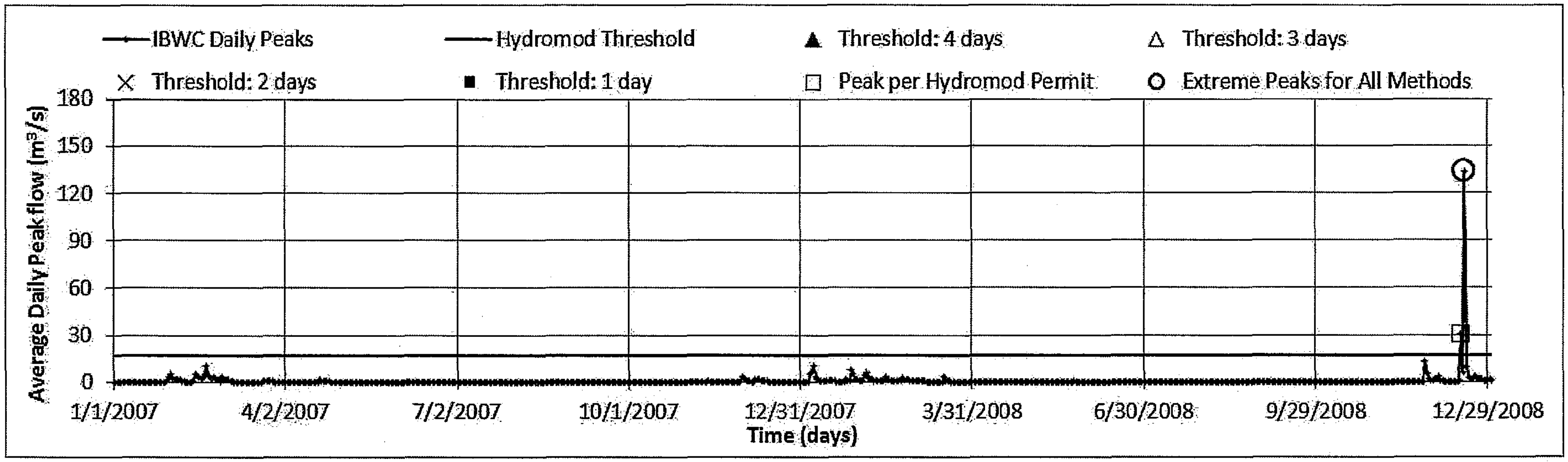


Figure A.1-36b. Peak Flow Time Series. 01/01/07 to 12/31/08. (Notice the change on vertical scale to capture extreme events).

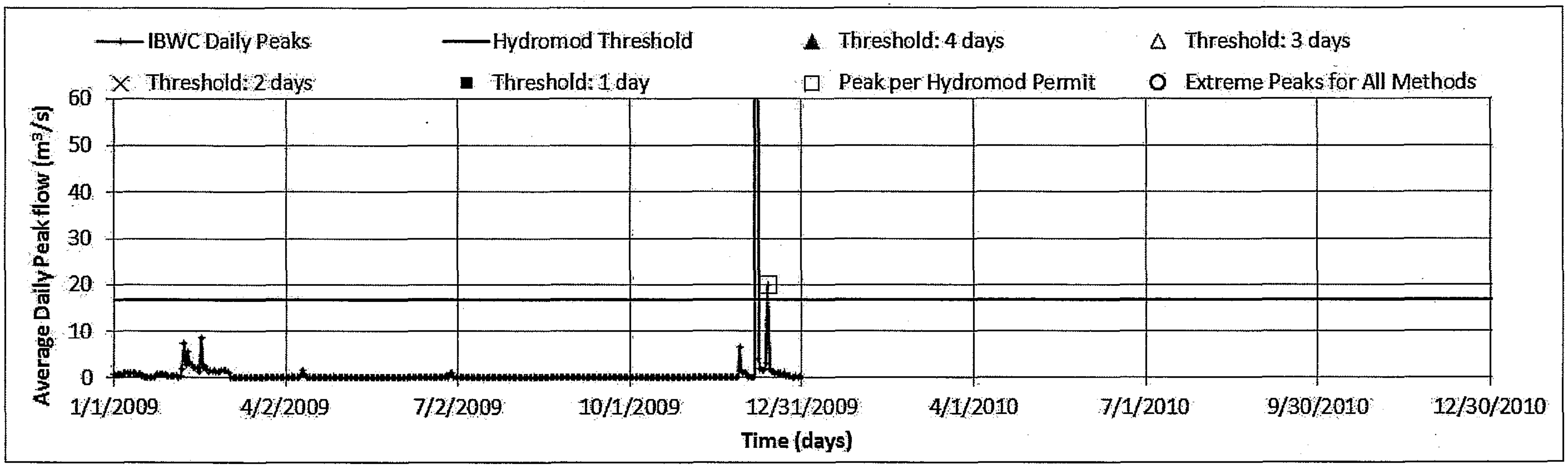


Figure A.1-37a. Peak Flow Time Series. 01/01/09 to 12/31/09.

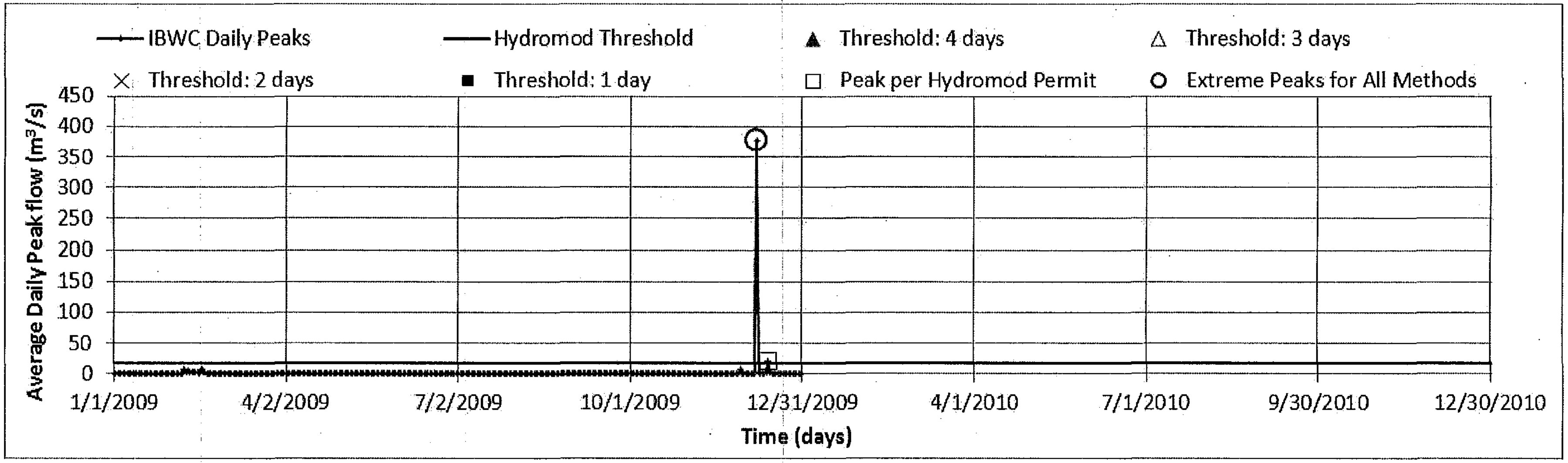


Figure A.1-37b. Peak Flow Time Series. 01/01/09 to 12/31/09. (Notice the change on vertical scale to capture extreme events).

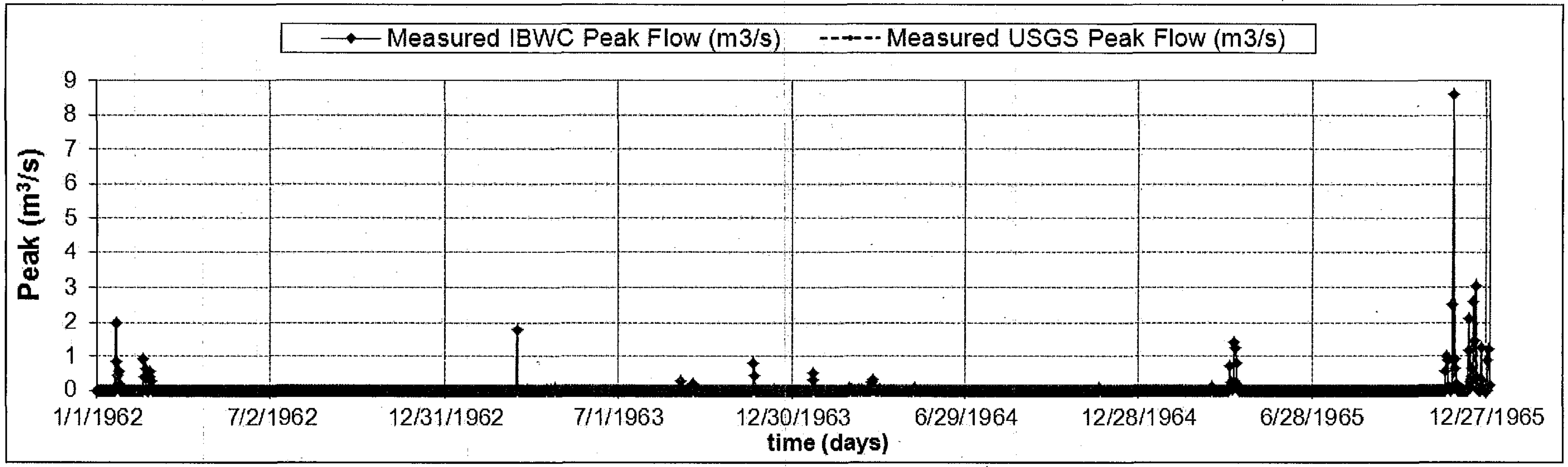


Figure A.1-38. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1962-12/31/1965

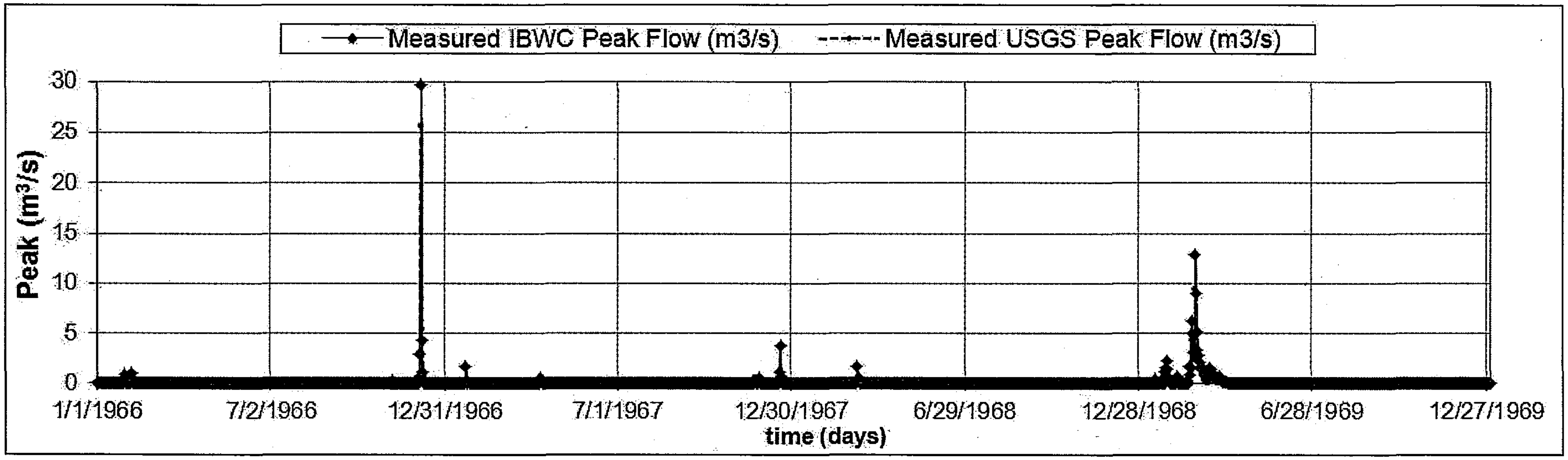


Figure A.1-39. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1966-12/31/1969

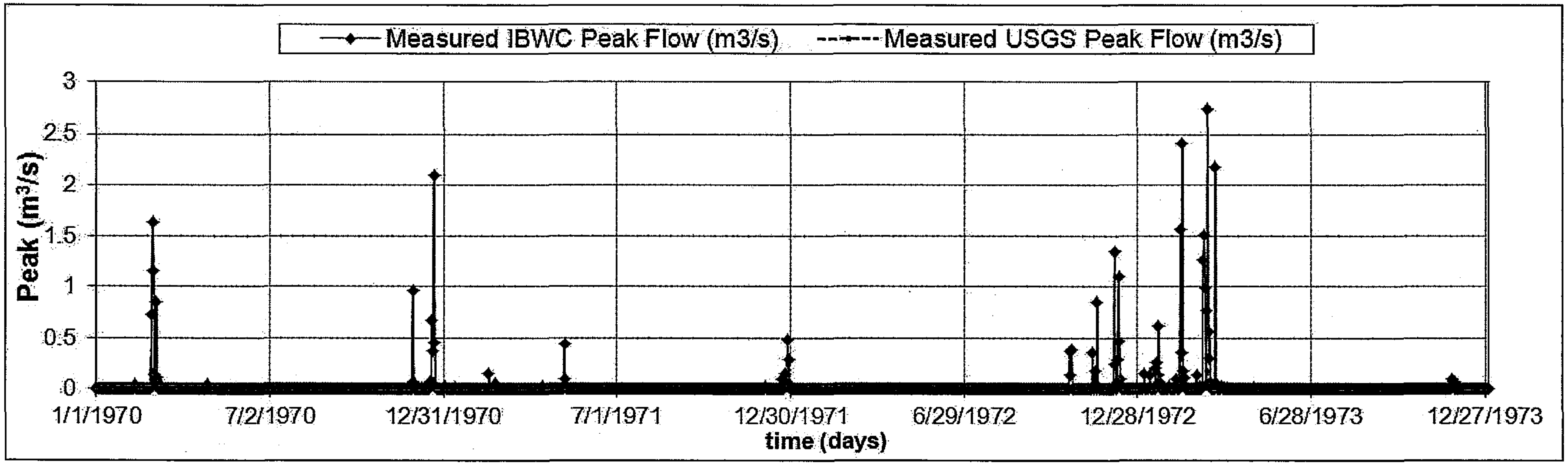


Figure A.1-40. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1970-12/31/1973

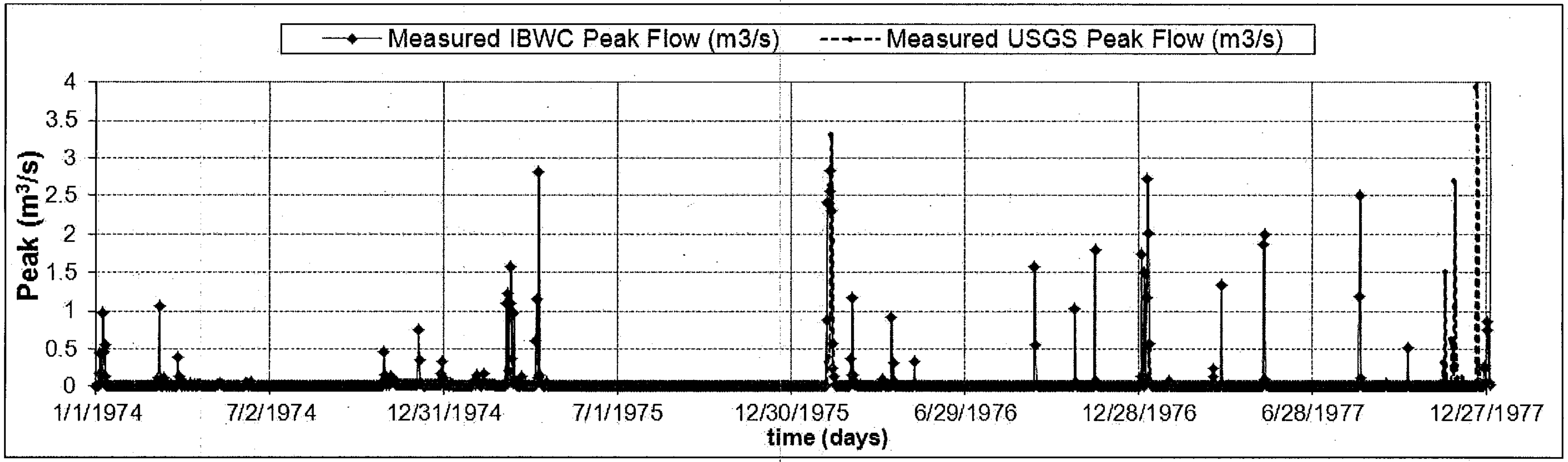


Figure A.1-41. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1974-12/31/1977

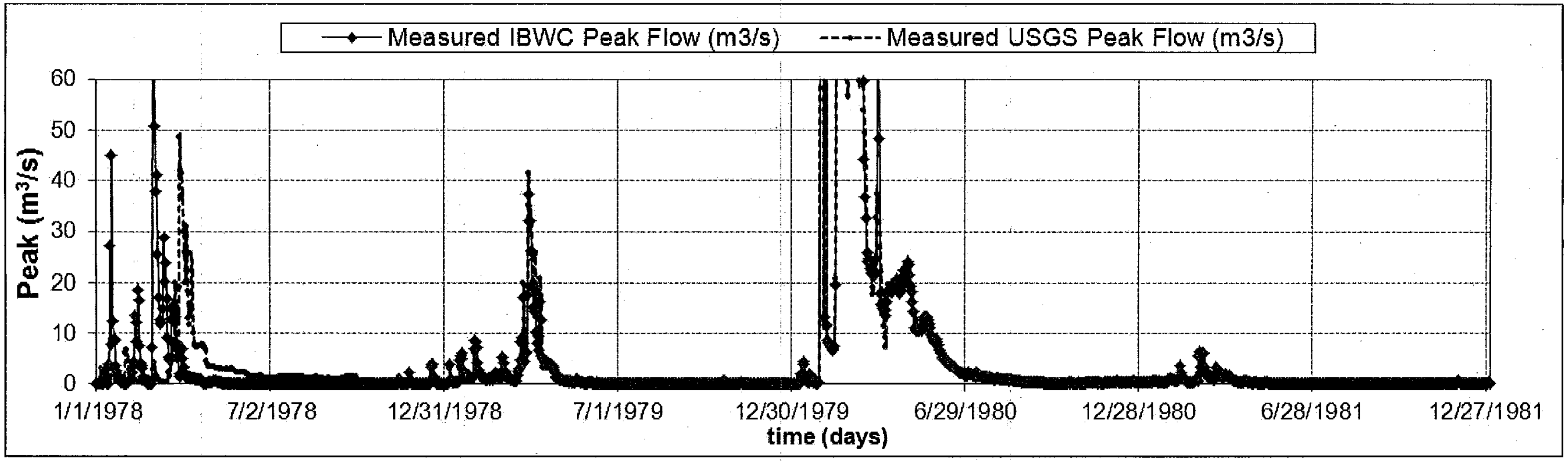


Figure A.1-42a. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1978-12/31/1981

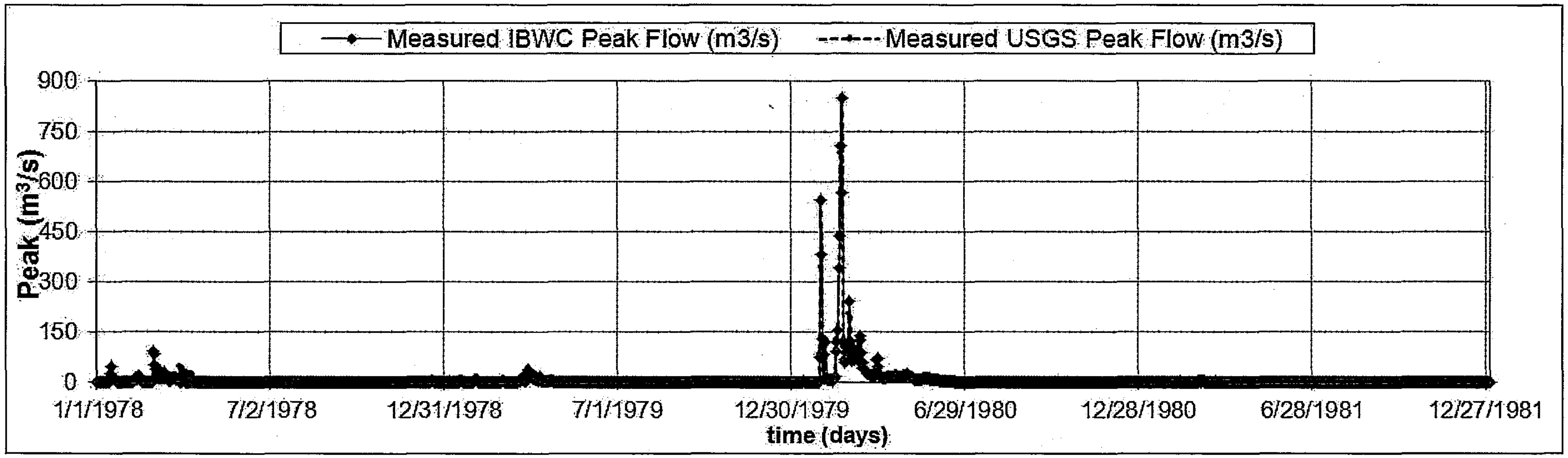


Figure A.1-42b. Peak Flow Time Series. Comparison of IBWC and USGS peak flows. 1/1/1978-12/31/1981. Notice Scale.

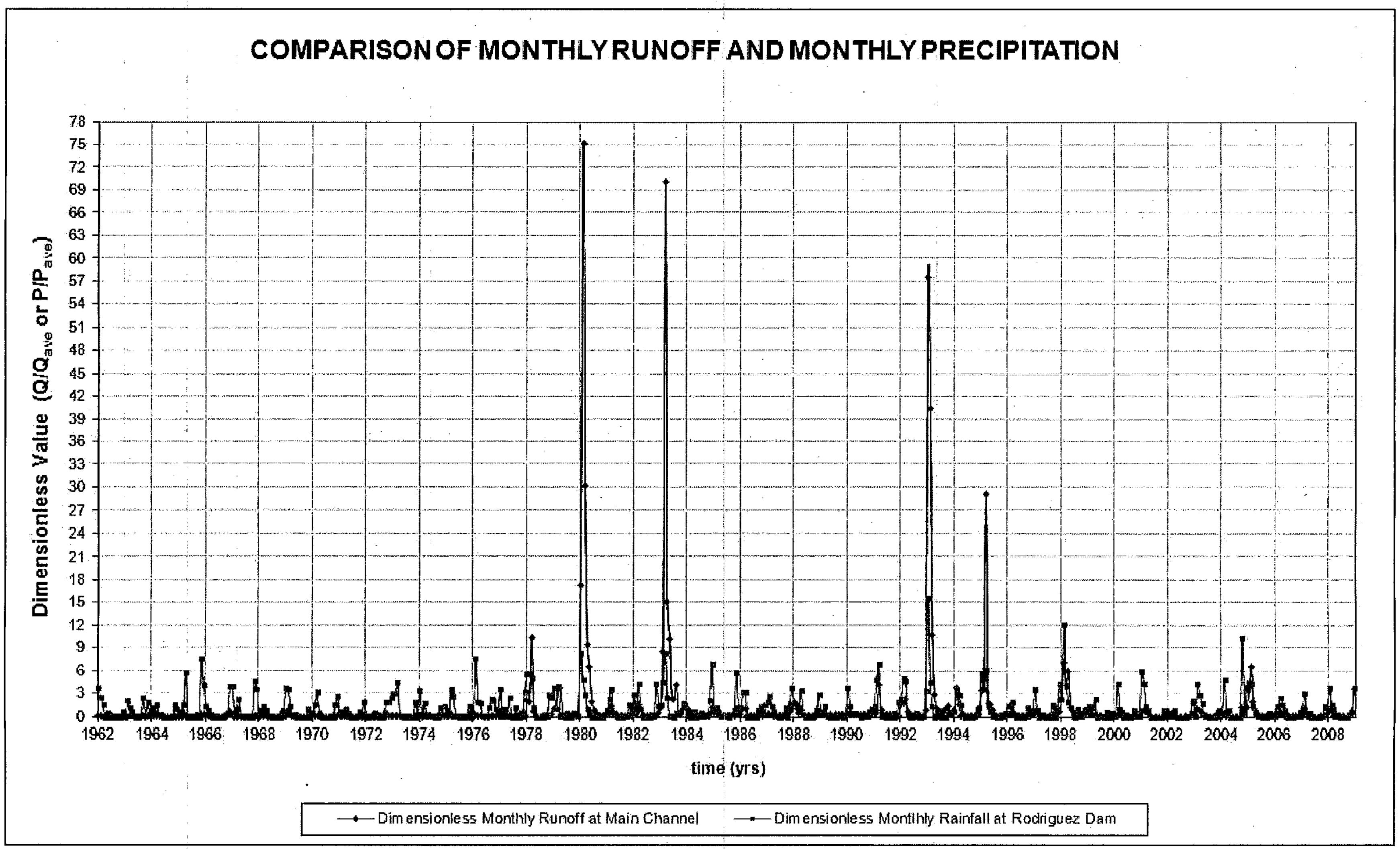


Figure A.2-1. Comparison of the Dimensionless Monthly Runoff and the Dimensionless Monthly Precipitation Time Series.

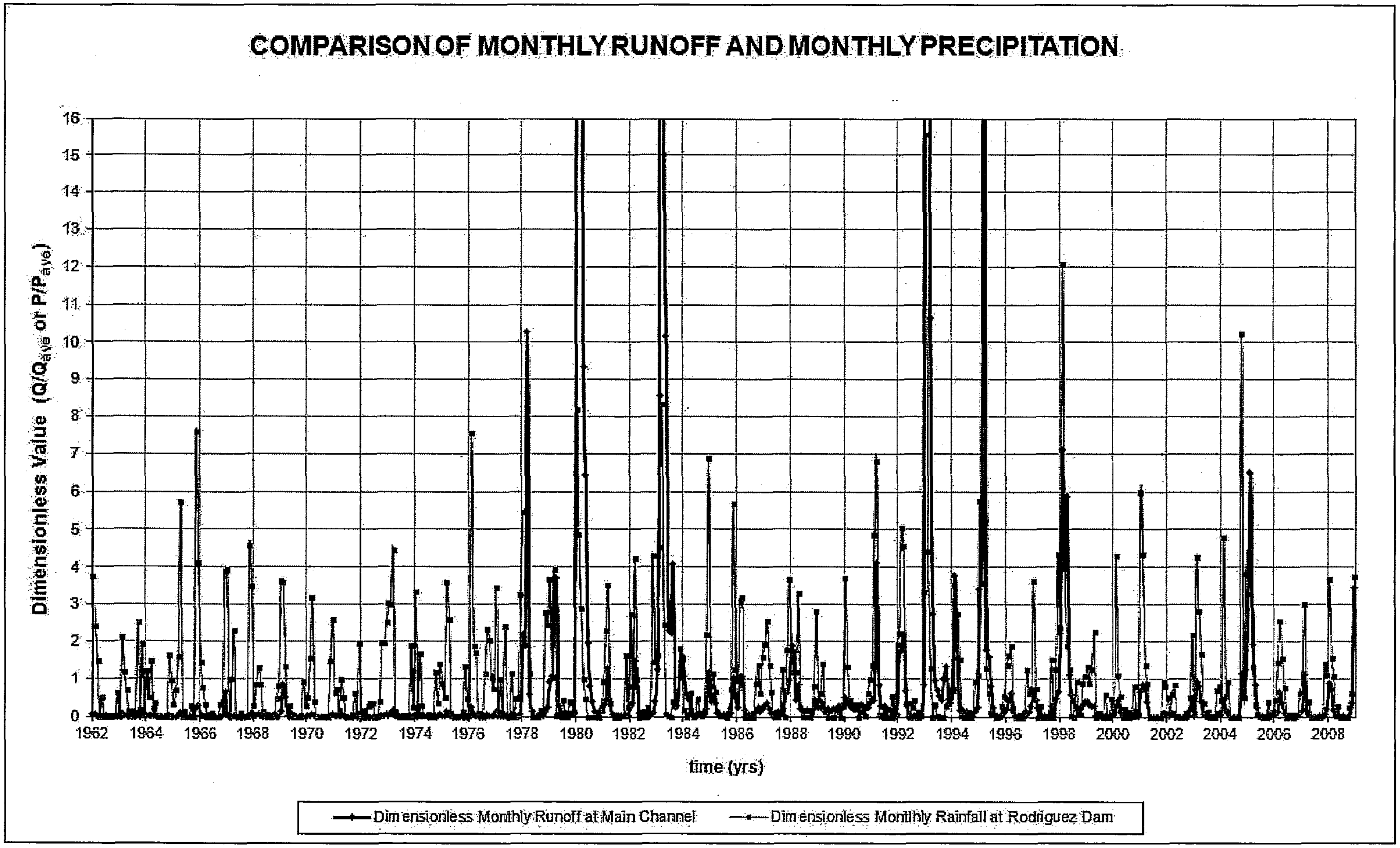
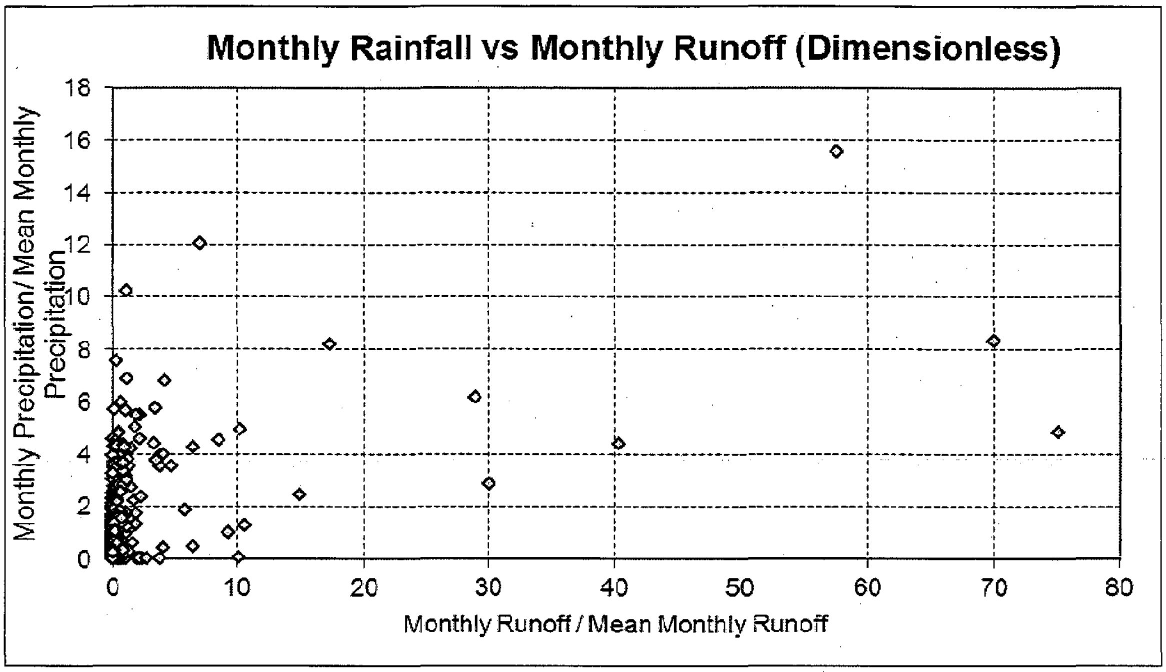
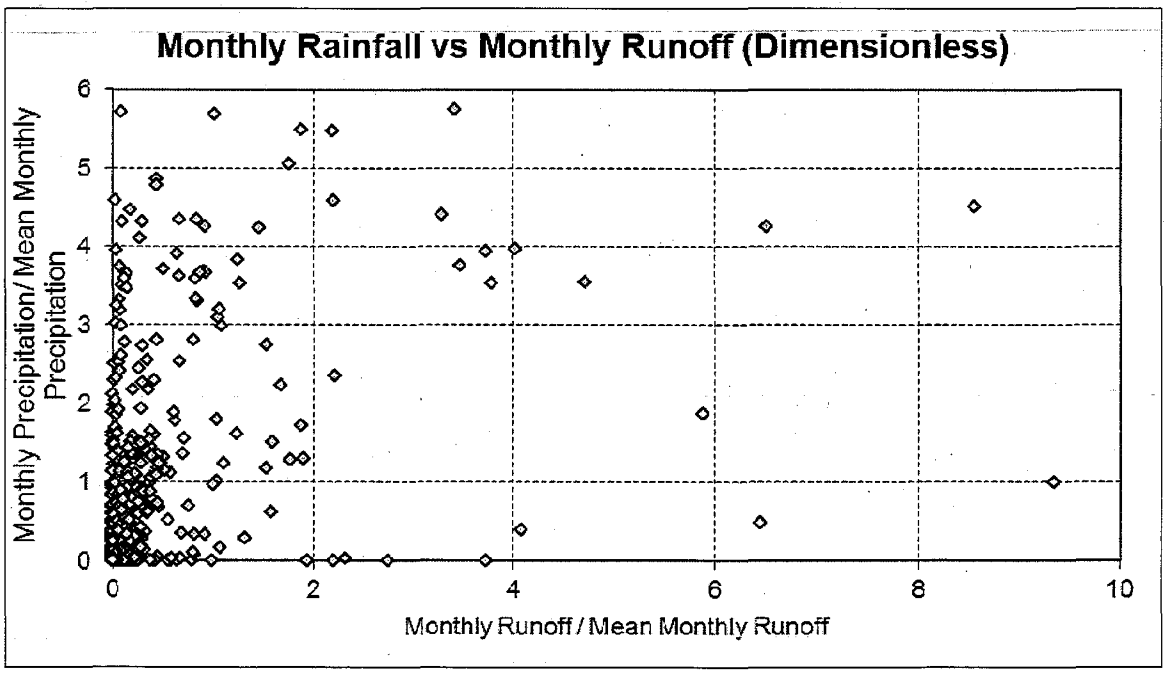
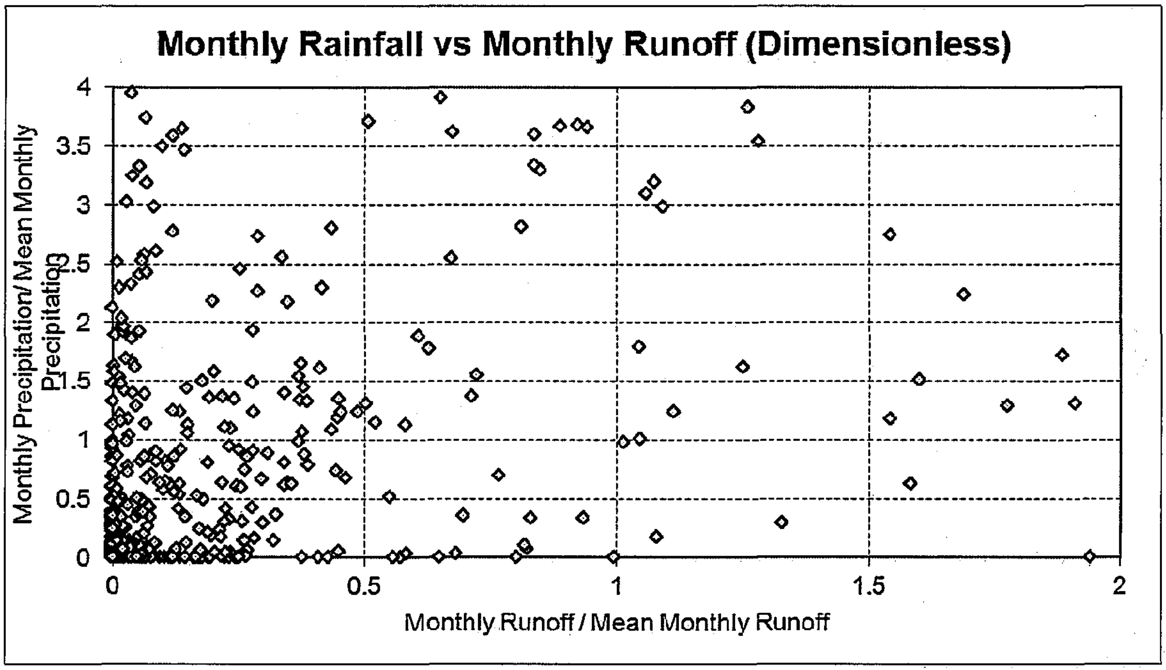


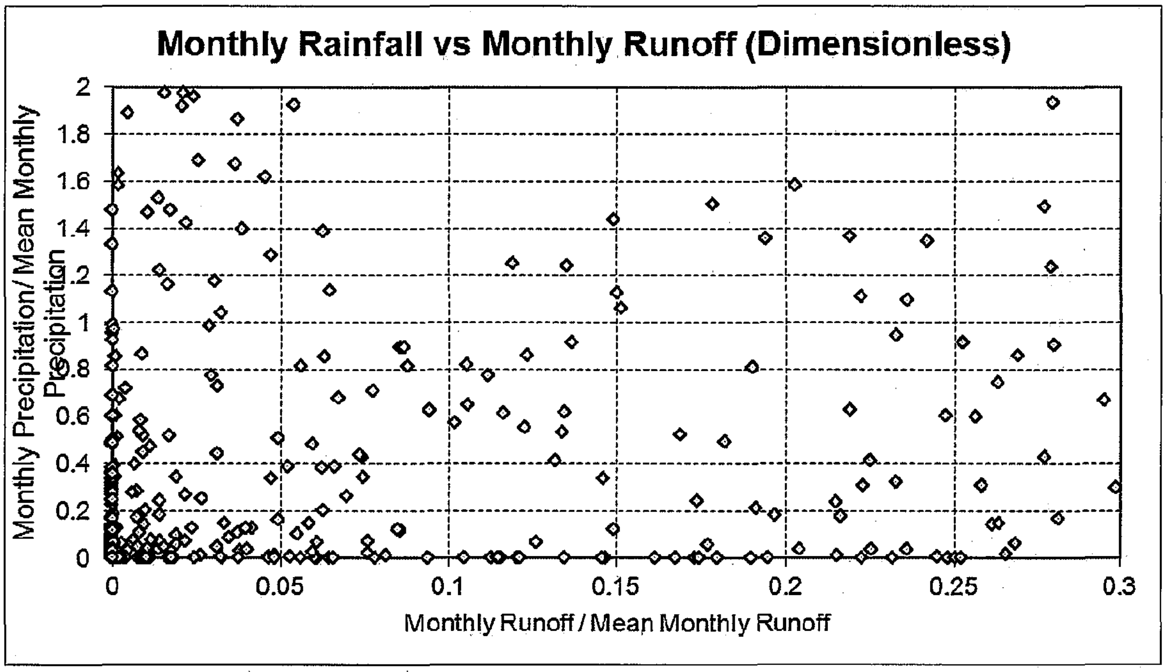
Figure A.2-2. Comparison of the Dimensionless Monthly Runoff and Monthly Precipitation Time Series (Note change in vertical scale).





Figures A.3-1 to A.3-2. Plot of Monthly Runoff and Monthly Precipitation (Different Scales).





Figures A.3-3 to A.3-4. Plot of Monthly Runoff and Monthly Precipitation (Different Scales).