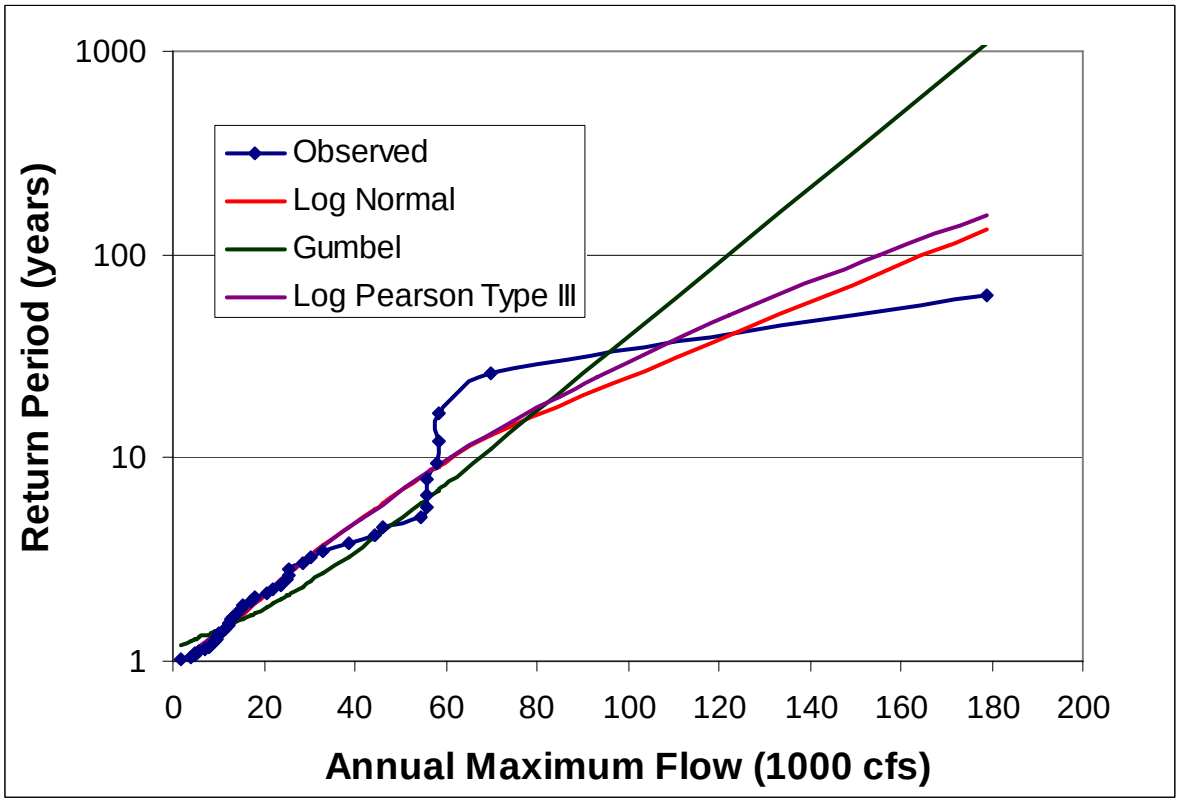
General Technical Note 1.

Concepts in Flood Return Frequency Analysis



Civilizations have gradually learned the lessons of catastrophic flood sporadic recurrence, beginning with the recognition that such floods do recur and the importance of passing on warnings to descendants.[1] Over many millennia progress has been marked by several milestones:

- Detailed flood records spanning generations,
- Increasingly sophisticated measures to mitigate the consequences of infrequent floods,
- Systematic investigations of flood causes,
- Improved forecasting, and
- Finally a more scientific understanding of extreme flood recurrence.

While the timing and magnitude of individual extreme floods are random in nature, extreme flooding at a location over the long-term exhibits a general trend of increasing magnitude (either as stage or discharge) with an increase in the recurrence interval. This trend has been the subject of a sound application of basic concepts in statistics and probability. The scientific understanding of extreme flood recurrence is rooted firmly in fundamental approaches developed by hydrologists over the course of the 20th century. Hydrologists have successfully completed countless studies of riverine flooding applying these concepts and methodologies.

This General Technical Note reviews fifteen fundamental concepts in flood return frequency analysis (RFA), including important implications and limitations.

A. Probabilistic Analysis

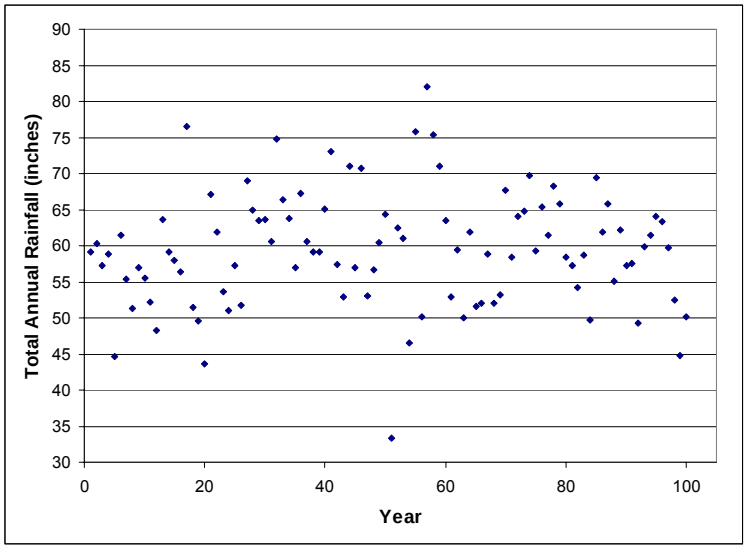
In hydrology (as well as many other physical realms) repeated observations of phenomena can yield varying results. Many variations correlate with *deterministic* (i.e., knowable, controllable) factors and hydrologists seek to define mathematical descriptions and physical confirmation of these correlations (see Deterministic Analysis below). A basic example of a deterministic process is the variation of a river stage versus flow. Apart from deterministic variation the remaining variability in historic or experimental data is termed the *random variability*. Phenomena with appreciable randomness are termed *probabilistic processes*.[2]

When a large number of hydrologic observations for probabilistic processes are collected patterns often emerge. Figure 1a, b, and c illustrate one such pattern. Figure 1a presents a histogram of a 100-year record of highly variable observations of annual total rainfall (at a hypothetical location). Figure 1b depicts this record arranged by five-inch rainfall bins (vertical bars), with the x-axis indicating annual rainfall totals

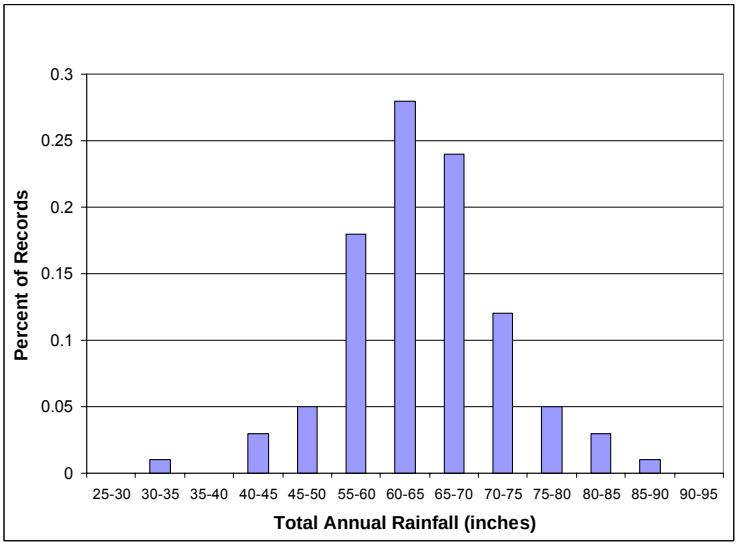
[1] The same could be said for other directly weather-related natural disasters—e.g., droughts, blizzards, and wildfires—as well as earthquakes, plagues, and pestilence.

[2] Random variability can be associated with repeated measurement of a highly controlled experiment as well as with a series of observations of naturally occurring phenomena. In the former, deterministic variations are often term *bias* and influence the *accuracy* of the measurement—as in variations in measuring the weight of an object due to improper zeroing of the scale. Random variation describes the dispersion or spread in measurements associated with uncontrolled factors. Random variation influences the *precision* of the measurement and the number of significant digits (decimal places, order of magnitude) reported.

a) Data Record



b) Data Ranges versus Percent of Records



c) Data Ranges versus Cumulative Percent of Records

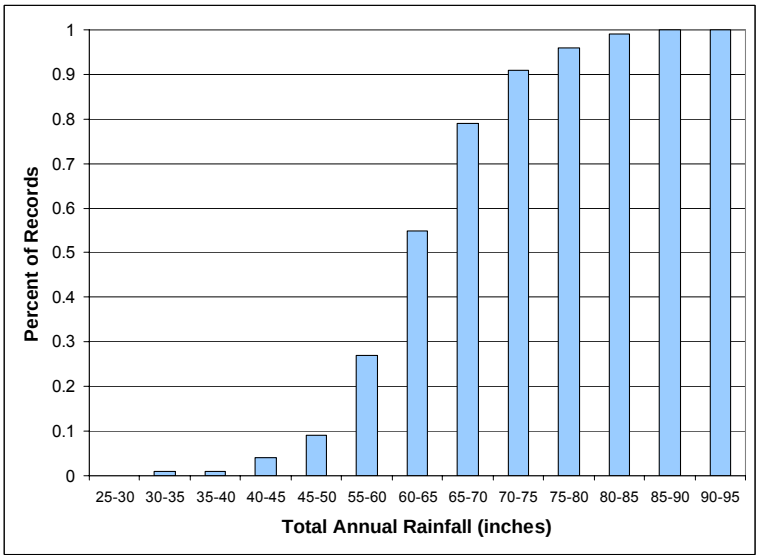


Figure 1. Variability in Total Annual Rainfall

and the y-axis indicating the fraction of total observations, or *relative frequency*, for each range. Here the x-axis spans less than one order of magnitude. Figure 1c shows the cumulative fraction of total observations. This variability pattern approaches a central peak with symmetrical tails.

Another case, shown in Figure 2a, b, and c, is the variability in a set of soil permeability (to infiltration) data. In this case, the x-axis spans nearly 3.5 orders of magnitude. These data display a very asymmetric pattern.

Mathematicians have developed idealized quantitative relationships—i.e., functions, equations—to characterize these and other patterns of relative frequency. These equations describe perfect curves and allow for precisely calculable *probabilities* for any discrete value—i.e., for each individual potential value of annual rainfall or soil permeability.

B. Normal Distribution

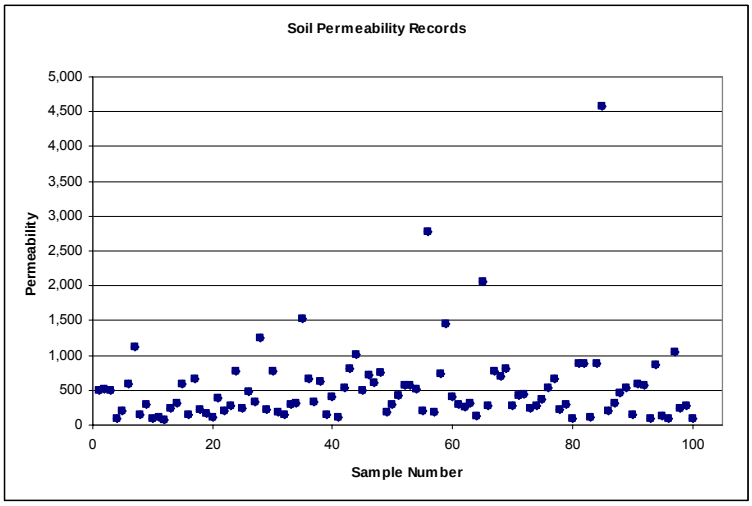
The most well-known symmetrical probability function is the *normal (or Gaussian) distribution Function.* Figures 3a and b illustrate the "bell curve shaped" probability density function (PDF) and the "S- shaped" cumulative distribution function (CDF) forms of the Gaussian distribution, which correspond to the relative and cumulative frequency plots in Figure 1b and c. A normal distribution function is defined by two parameters that describe two moments, i.e., points about which various considerations of the distribution "weight" are balanced. These are the *mean* (around which the two sides of the distribution balance, abbreviated as μ) and the *variance* (which describes the distribution on either side of the mean; the variance is an averaged square difference between each record and the mean). The square root of the variance is the *standard deviation* (abbreviated as σ) which is often referred to instead of the variance. The value of σ/μ is a normalized standard deviation and termed the *coefficient of variance* (CV). Figure 4 illustrates the changing shape of the normal distribution with changes in CV.

The formula for the probability density form of the normal distribution is :

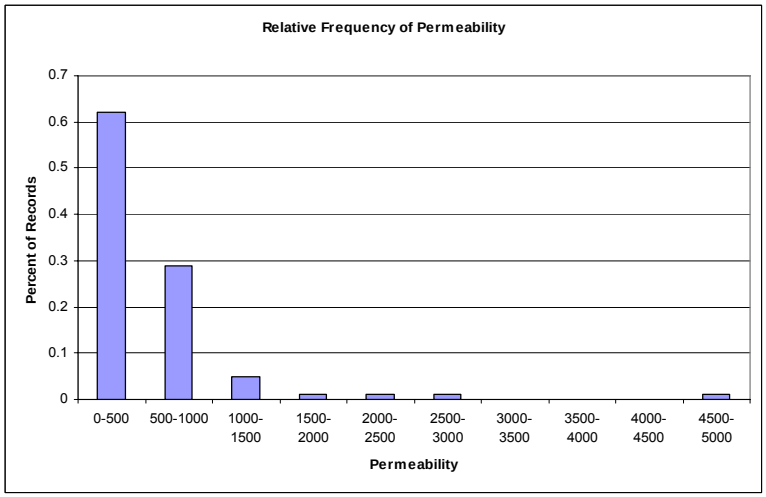
$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Figures 5a and b illustrate an important aspect of the normal distribution. For any normally distributed data, fixed percentage of data lie between μ and plus one, two, three (etc.) σ: 34.1%, 47.7%, and 49.8%, respectively. Given the symmetry of the distribution the same percentages lie below μ within multiples of minus one, two, three (etc) σ. Similarly, any percentage of values surrounding μ corresponds to a multiple of the σ. For example, 80, 90, and 95 percent of all records lie within ± 1.28σ, 1.65σ, and 1.96σ, respectively. In these cases 10%, 5%, and 2.5% of all records lie outside on the upper tail, and 10%, 5%, and 2.5% also lie on the lower tail.

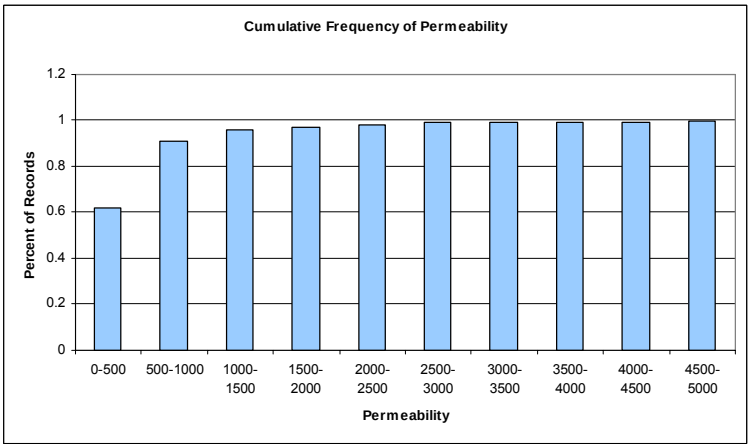
In order to describe variability with a normal distribution μ and σ. are calculated for the data—in the case of Figure 1a μ and σ are 60 and 8 inches, respectively. As depicted in Figure 6, the data plot can be compared to the ideal curve (heavy blue line). Note that most of the observed relative frequencies do not line up directly on the curve and are offset by varying amounts and there is some variability in fitting the normal distribution to the data. In other words there is some *uncertainty* regarding whether this particular normal distribution curve—which is based on the estimates of μ and σ—truly reflects the random variation in annual rainfall.



a) Data Record

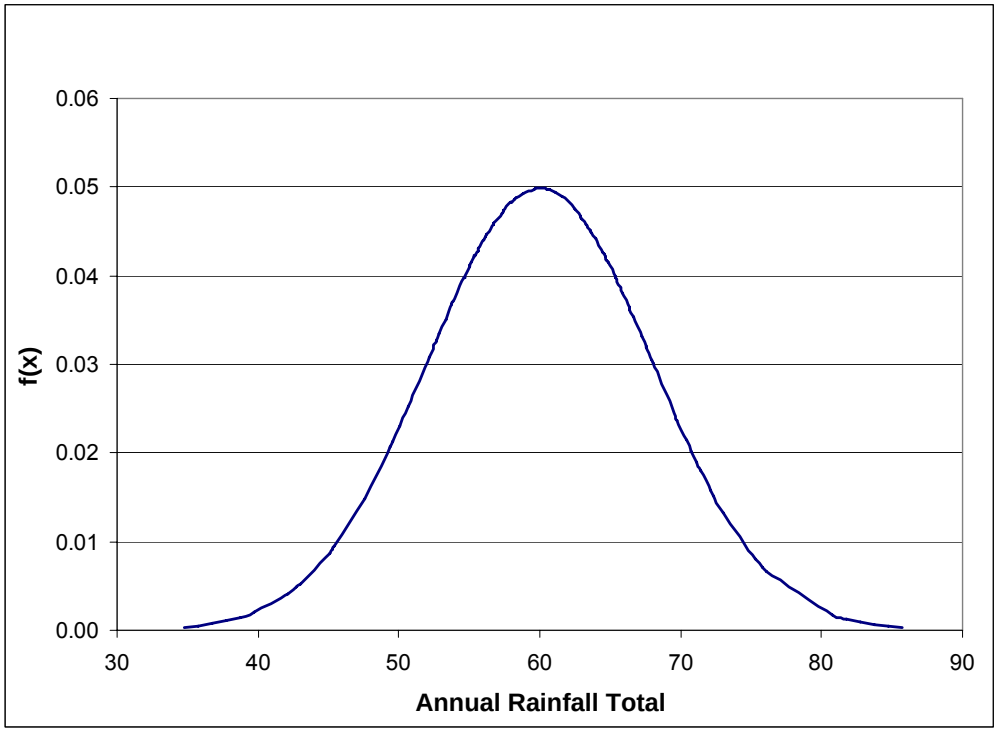


b) Data Ranges versus Percent of Records

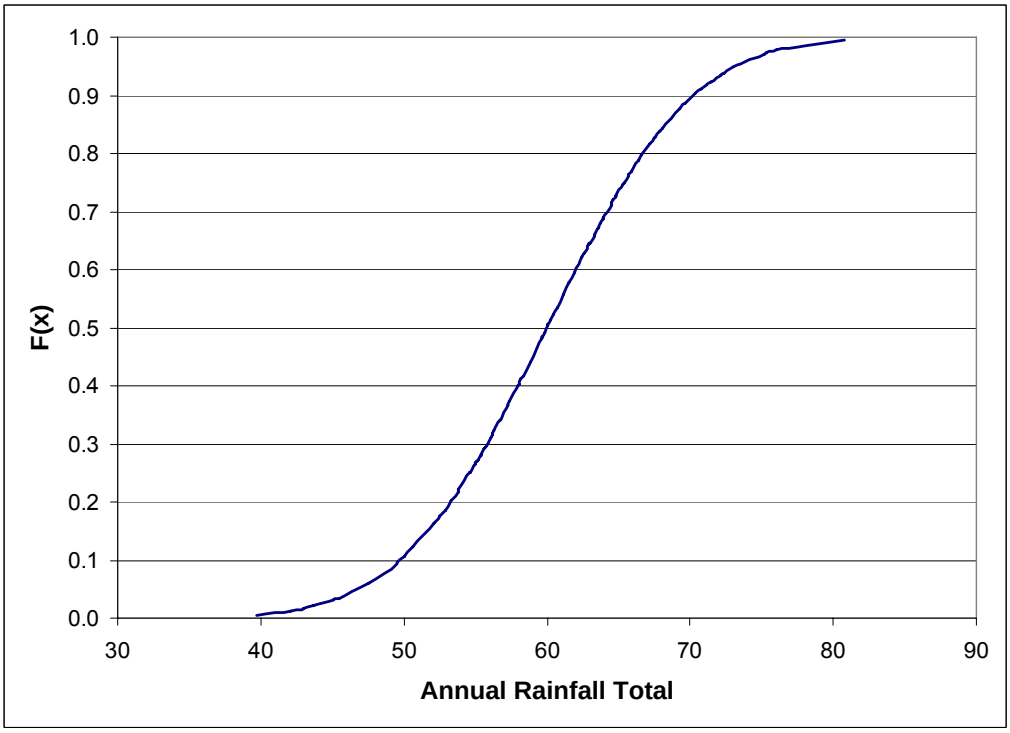


c) Data Ranges versus Cumulative Percent of Records

Figure 2. Variability in Soil Permeability
(Hypothetical Location)



a) Probability Density Form



b) Cumulative Distribution Form

Figure 3. Normal Distribution Function

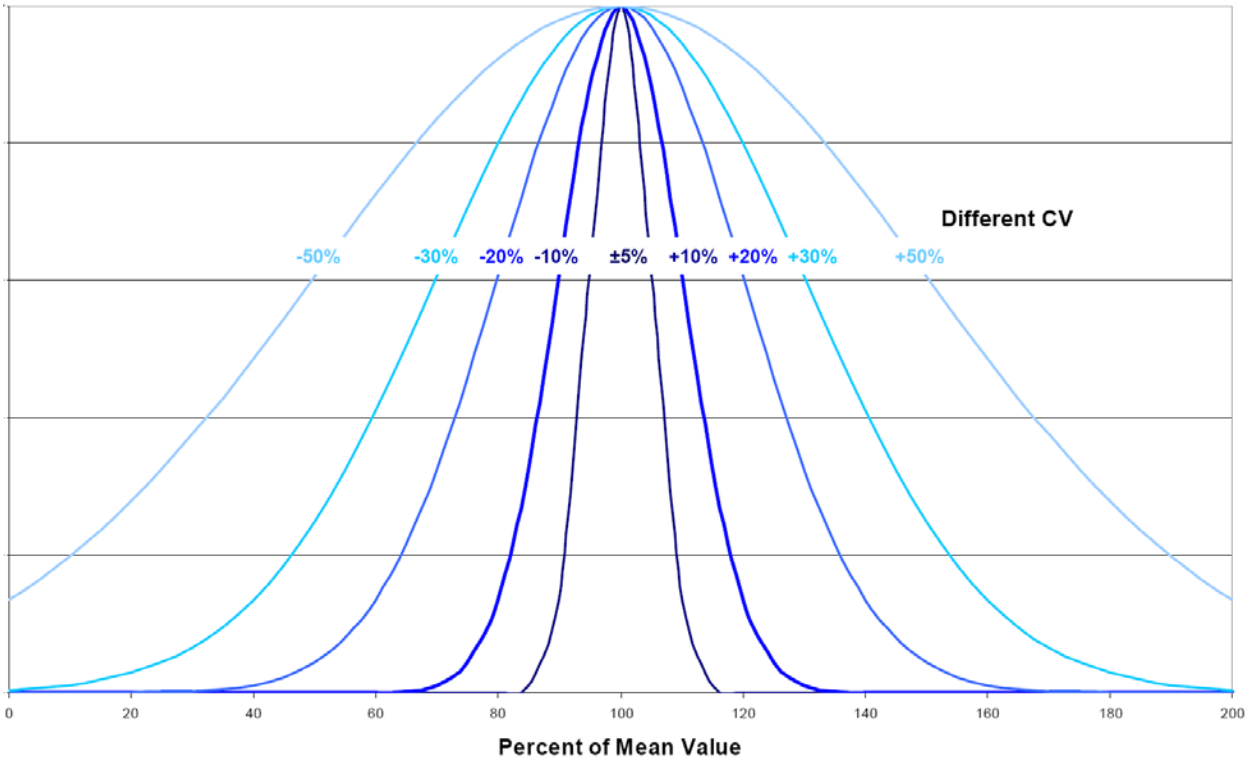
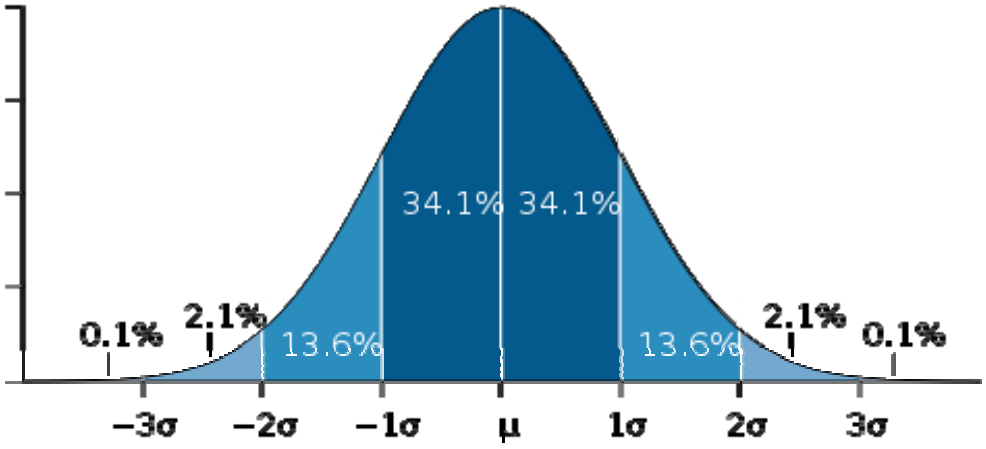
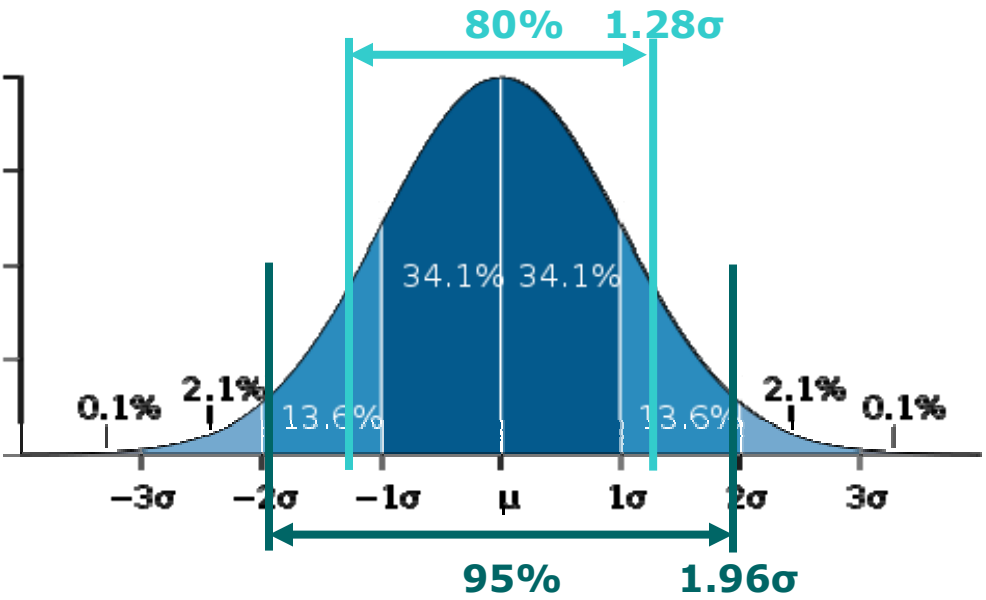


Figure 4. Changing Shape of Normal Distribution for Different CV



a) Percent of Data Lying Within +/- 1σ, 2σ, and 3σ From μ



b) Values for σ Corresponding to 80% and 95% of Data

Figure 5. Percent of Data and Standard Deviation for Normal Distribution

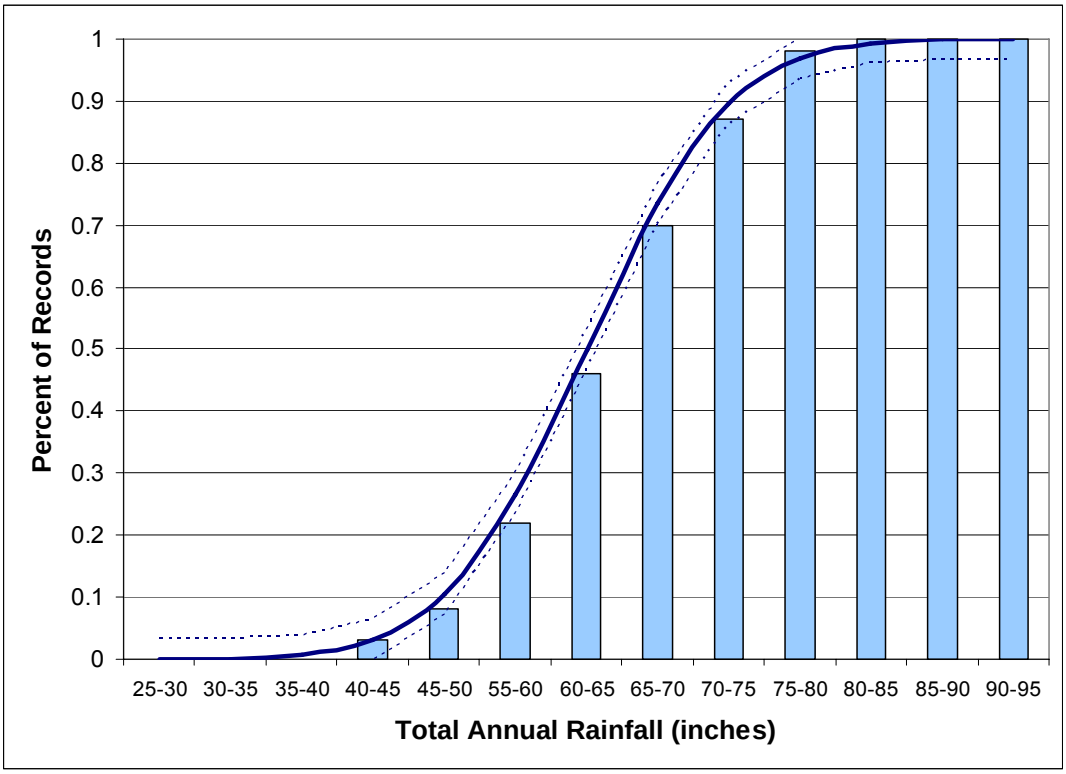


Figure 6. Cumulative Frequency Data and Cumulative Distribution Function

C. Uncertainty in Estimates

Hydrologists not only use probability distribution functions to characterize the variability of observations—such as by plugging annual rainfall record values of μ and σ into the normal distribution equations—they also use distribution equations to characterize uncertainty regarding the dataset μ and σ and the distribution. The values for μ and σ used to construct the solid blue curve in Figure 6 are the *most likely estimate of μ and σ based on the dataset*, but the actual mean and standard deviation could be expressed as a range—e.g., μ as 60 inches, ± some amount. Thus, the whole normal distribution curve could also be expressed as a range with uncertainty bands above and below the solid blue line (see Figure 6). This can seem confusing as it is basically estimating a probability for a probability estimate.

In the case of normally distributed data, the uncertainty for the estimates of *μ and 1σ* (and any point) can be calculated using well established statistical techniques. For example the ± band around the estimate of the mean uses a symmetrical distribution known as the Student's t distribution. The ± band of uncertainty for an estimate is referred to as an *Uncertainty or Confidence Interval* and the bounds are referred to as *Upper and Lower Confidence Limits, UCLs/LCLs*. UCLs/LCLs are typically set such that the Confidence Interval captures a high percentage of the possible values. For example a 95% confidence interval for the value of μ would be 60 inches ± 1.6 inches.

D. Deterministic Analysis

Hydrologists first examine data with a view toward identifying deterministic relationships, or correlations. *Correlation analysis* evaluates the variability in one dataset characteristic versus one or more physical factors, e.g., the flood stage at a point in a river versus the discharge (flow). Particular correlations between factors do not prove a specific cause-and-effect relationship—two factors may be products of a third factor—but are evidence of deterministic aspects of observed variability. For example, while flood stage and discharge at a river location are correlated, the actual physical cause-effect relationship is more complex, involving flow routing considerations over the entire river segment. A common application of correlation is *trend analysis,* which examines a characteristic versus time, termed a time-series (e.g., a stage hydrograph).

Correlation and trend analyses are usually done in conjunction with data graphs—such as Figure 7a. Using the general shape of the graph, hydrologists identify evidence for the presence of linear or other mathematical relationships—polynomial, exponential, power function, periodic oscillation, etc. In some cases the equation for a line or curve has been established—rooted in physical or mathematical theory or a large body of previous *empirical* work (experiments or observations)—and the hydrologist seeks values for the coefficients to optimize the fit. In some cases the full equation with coefficients has been established and the hydrologist will corroborate the coefficient values. In cases where a line/curve type has not been predetermined, the hydrologist relies on fundamental physical understanding to propose certain equations and then tests the equations with coefficients to fit the data. The process of fitting a line or curve to the data is known as *regression analysis*.

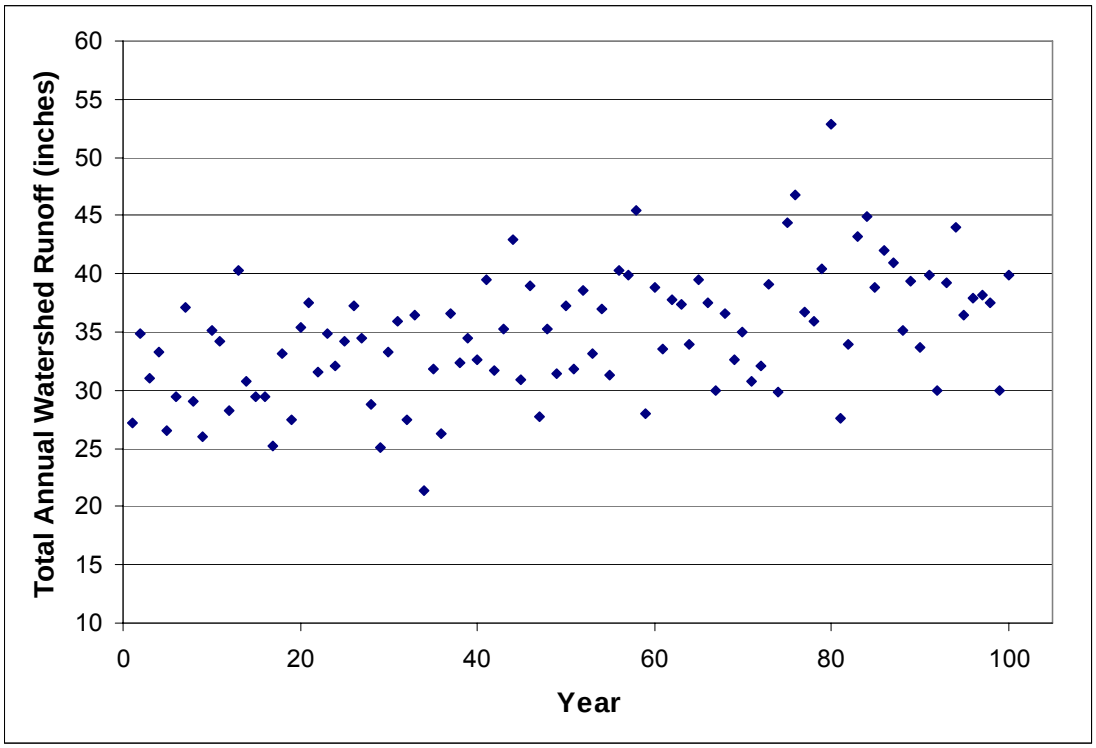
The fit of any proposed mathematical relationship to the data is analyzed in several ways:

- The error (difference) in the y-value for each point versus the y-value associated with the proposed line is calculated and squared (square difference or error) and then summed—referred to as the sum of the square error (SSE). The square root of the mean square error (RMSE) is analogous to a σ between the data and the proposed line/curve. RMSE for different proposed lines/curves and coefficient values can be compared. The normalized RMSE, dividing by μ, is termed the CV(RMSE) or the Scatter Index (SI).]
- The difference between each data point and the data mean, squared, and then summed is the sum of the square total, SST (analogous to the overall data variance). The difference between SST and SSE (SST – SSE) is the amount of variance in the data accounted for by the proposed line or curve. The quantity

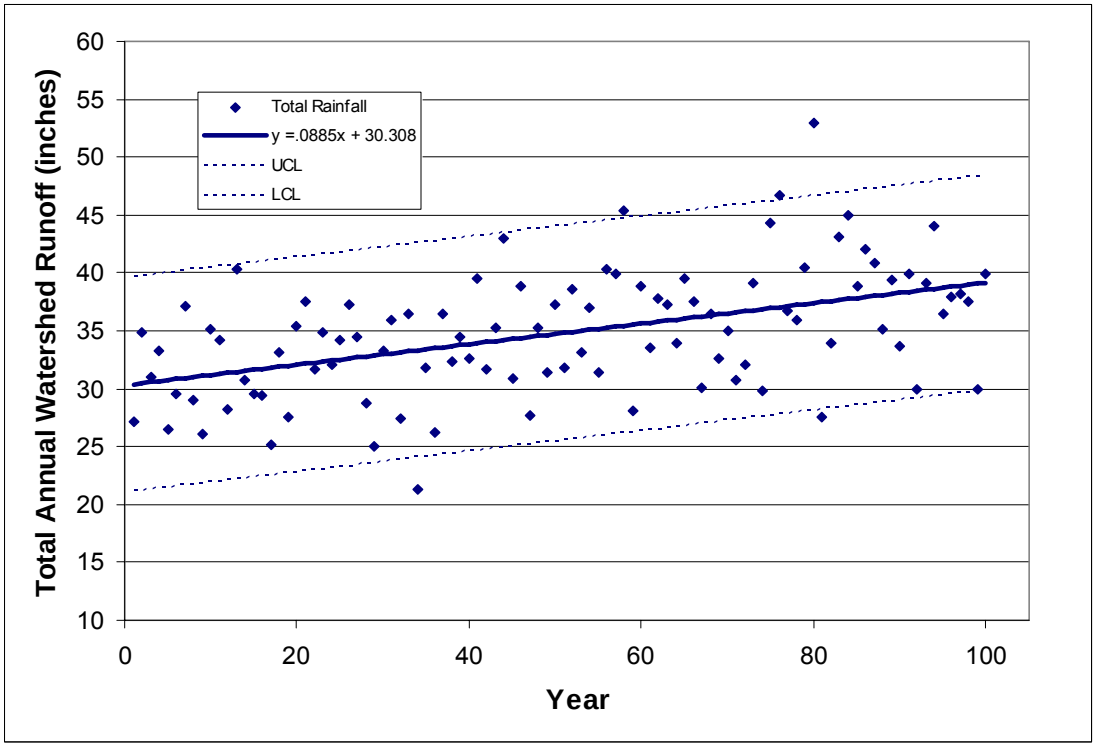
$$R^2 = \frac{(SST - SSE)}{SST} = 1 - \frac{SSE}{SST}$$

  is the fraction of total variance explained by the proposed line/curve and is referred to as the coefficient of determination (sometimes the correlation coefficient). A high R^2 (e.g., above 0.9) is indicates that a large percentage in the data variation can be explained by the proposed line/curve. A hypothesis test is also used.[3]

[3] The probability that we can reject a hypothesis that the relationship is not valid (that we can reject the null hypothesis) is called a chi squared, χ^2, test.



a) Data Record



b) Data Record with Linear Trend

Figure 7. Long-Term Trend in Watershed Runoff

- The distribution of the residual error (variance) in the data apart from the variance explained by the line/curve. Often the residual error shows some evidence of a normal distribution, and RMSE (or σ for the residual error) is then used to define the UCL and LCL above and below the correlation/ trend line (or curve). The spread of UCL/LCL bands is also an indication of fitness.

Figure 7b shows the linear trend analysis for the data in Figure 7a.

Data μ	34.8	SST	2,887
Data σ	5.4	SSE	2,234
Linear Trend:	y = 0.09x +30.1	RMSE	4.73
		R^2	0.226
95% UCL/LCL	y = 0.0885x +30.308 ± 9.27		

Importantly the R^2 of 0.226 is not that high due to considerable residual variance.

The correlation of many physical factors with flood conditions—e.g., climatological, topographical, land cover, river modifications factors, etc.—is an important subject in flood hydrology, with some correlations having a direct bearing on flood forecasting. Short- or long time changes in climate e.g., El Nino, sea level rise, etc.) and floodplain and river morphodynamics (changing river shapes, depths, gradients, etc), are major subjects of flood trend analysis.

E. Conventions for Confidence Intervals

The normal distribution is used ubiquitously in scientific and engineering to evaluate random variability and estimate uncertainty. Different scientific and engineering endeavors adopt different conventions for describing variance or uncertainty. In experimental scientific literature a 95% Confidence Interval is often used to express the quality of data statistics and trend analysis. A 95% Confidence Interval has LCLs and UCLS spanning from 2.5% to 97.5% of the distribution, i.e., with 2.5% variability on each tail outside the interval.

In hydrologic design the selection of a Confidence Interval is subject to interpretations regarding risk—the consequences associated with events outside the selected interval. Goldman in 2003 (Goldman 2003) cites the use of a 90% Confidence Interval for the certification of levees along the Upper Mississippi River.

Manufacturing quality control and assurance practices often use the term Six Sigma, which refers to six standard deviations—three for each tail. Six Sigma encompasses 99.6% of the normal distribution (see Figure 5b). Manufacturing items to a tolerance of six sigma can be a rigorous quality control standard depending on the value of σ.

When uncertainty is a concern on only the upper (or lower) tail of the distribution a convention may describe the probability of values below (or above) a reference point on that tail. Figure 8 illustrates a 90% probability of values being below a point on the upper tail of a normal distribution—referred to as a 90% probability of non-exceedance. This is the same as having 10% of the values remaining on the upper tail. This 90% probability of non-exceedance is equivalent to the 80% UCL. This can be confusing as the probability of non-exceedance and the UCL are **NOT** the same number.

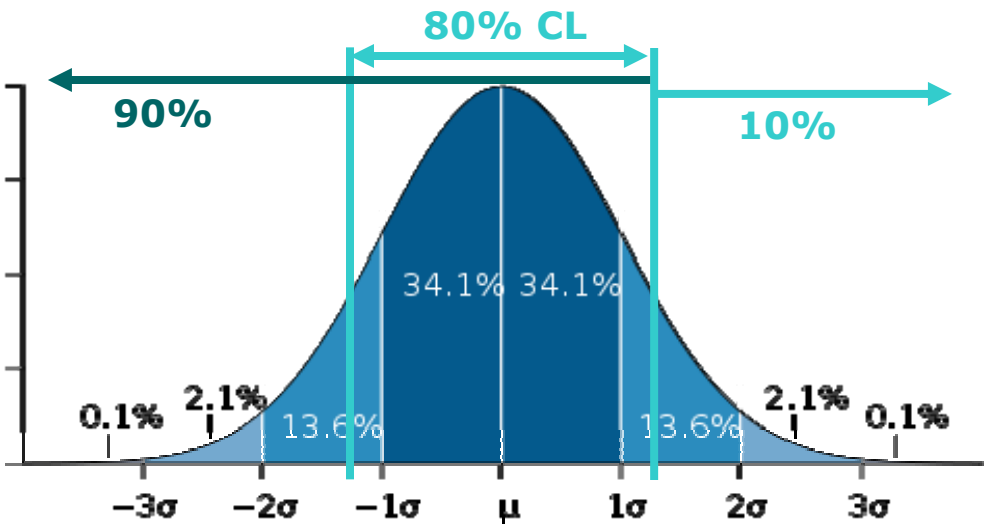
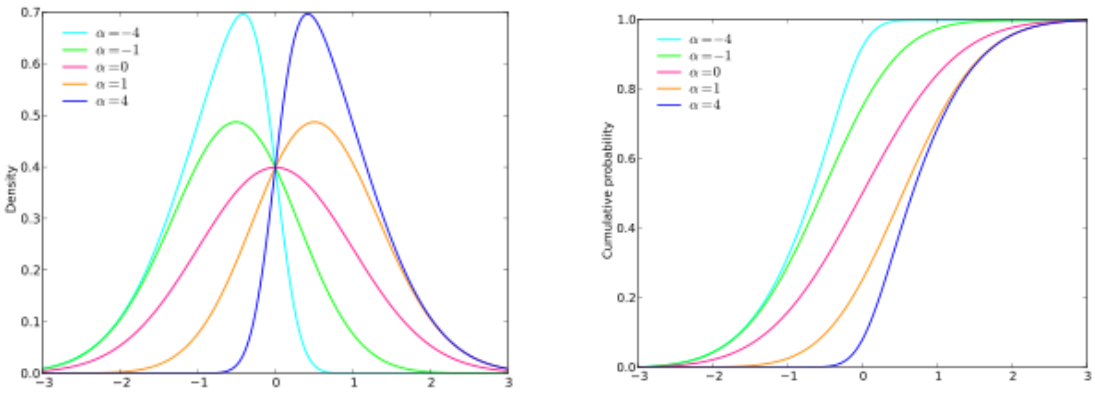


Figure 8. Probability of Non-Exceedance and Confidence Limit

F. Other Probability Distributions

The normal distribution does not adequately describe the variability of many probabilistic processes. Some common examples are:

1. Asymmetric or skewed distributions. In some data, tail observations are more prevalent and extended on one side of the distribution, resulting in differences between the mean (average), median (50th percentile) and mode (highest observed frequency). One type of asymmetric distribution is the Skewed normal distribution, which modifies the normal distribution with the addition of a third moment called the coefficient of skewness (α). Figures 9a and b show the PDFs and CDFs for the Skewed normal distribution.



a) Probability Density b) Cumulative Distribution

Figure 9. Skewed Normal Distribution

http://en.wikipedia.org/wiki/Skew_normal_distribution

2. Log normal distribution. When data are highly skewed, with x values stretched over several orders of magnitude, similar to Figure 2, the distribution of log (x) can be evaluated. (i.e., log(10) =1, log(100) =2, log (500) = 2.699, etc.) Figures 10a and b present the untransformed and the transformed probability density functions for the data from Figure 2.
3. Pinched/rounded distributions. Figure 11 shows several symmetric distributions affected by squeezing the peak and fattening the tails (e.g., red),[4] or vice versa (e.g. blue). In these example distributions the multiples of the standard deviation do **NOT** encompass the normal distribution percentages of variability. This property can be described by adding consideration of a fourth moment of a distribution called kurtosis. In addition to kurtosis it is possible to use other shape parameters comprised of higher moments. Kurtosis and higher moments can also apply to asymmetric distributions.

Mathematicians and scientists have defined numerous distribution formulas taking into account combinations of data transformation, skewness, and shape parameters. When any probability distribution is used to describe variability in an observation, the same distribution is appropriate for characterizing the uncertainty in estimates (Chow et al.1988). Thus, if random variability is fitted to a skewed distribution, the uncertainty should also be considered to be skewed.

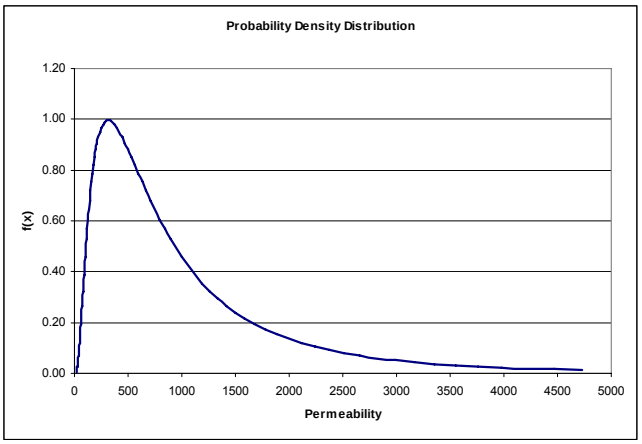
G. Return Frequency

As noted above, for a record of observations collected over a defined period of time[5]—such as the 100-year record of annual total rainfall—hydrologists can calculate relative frequencies and define idealized mathematical probabilities, expressed as percentage of observations. In the above case the calculated 5% relative frequency is 71.7 inches using the 95th ranked value and the estimated 5% probability is 73.0 inches using the normal distribution and dataset μ and σ. In addition, hydrologists can describe relative frequencies in terms of *return or recurrence interval* and estimate the probability of how often a particular annual rainfall magnitude can be expected to occur, or what magnitude is associated with a particular recurrence interval. To understand return frequencies and return probabilities it is important to first review four key terms.

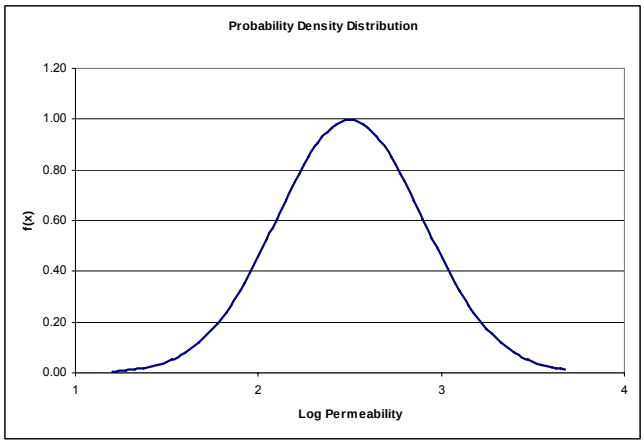
1. The *event* is a measured or calculated data point (can be an instantaneous or cumulative measurement) with a magnitude, such as an hourly, daily, monthly, and annual rainfall accumulation; a peak flood stage or discharge (flow); a daily high tide; etc.
2. The *time-frame*, τ, is a duration under consideration over which many events of varying magnitude have or may occur (regularly or irregularly). The time-frame can be an historical record (such as 30 years, 10,950 days, of daily rainfall data) or a hypothetical (or future) time length for which recurrence is being analyzed (e.g., 100 years).

[4] The term “fat tail” has been made famous in the post-2008 discussion of the massive losses on supposedly low default probability mortgage credit instruments and the ensuing financial meltdown. Credit ratings and instrument pricing are heavily dependent on statistical analyses of default probabilities. Underestimated low probabilities have been called “black swan” events. (see Taleb, 2010)

[5] It is not necessary for the data to be collected at regular intervals, e.g., peak river flood data.



a) Probability Density for Untransformed Data



b) Probability Density for Transformed Data, log(x)

Figure 10. Log Normal Distribution (data in Figure 2)

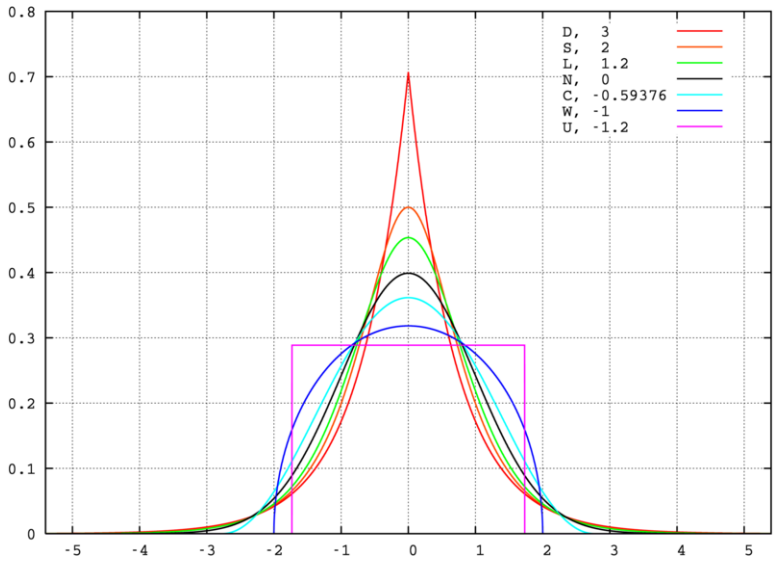


Figure 11. Distributions with Kurtosis

http://en.wikipedia.org/wiki/Kurtosis

3. The *return frequency*, F_R, for an event of *a particular magnitude or greater*. For data of duration τ the *nth* record has an *observed* F_R of n/τ . For example, if the daily rainfall has exceeded 6 inches 10 times (n = 10) during a 30-year record (τ = 10,950), F_R is 10/10,950 = 0.000913 = 0.0913% (expressed *per day*). If expressed *per year* F_R is 33.3%. The *observed* F_R is often modified as $n/(\tau +1)$ (in accordance with binomial distribution theory). In this case the F_R for a 6-inch rainfall is still 0.0913% per day. The highest ranked event for 100 years of data is 0.99% (instead of 0.1%) per year.

   For the higher ranking events, hydrologists may not consider τ to be sufficiently long enough to calculate a reasonable F_R. using $n/(\tau +1)$. To compensate hydrologists often modify *observed* F_R using $(n - a) / (\tau + 1 - 2a)$, where a is typically between 0 and 0.5. If a is taken as 0.3, the F_R of the top ranked event from a list of 100 annual records is 0.7% per year. As with the calculation of relative frequency, F_R requires a complete, accurate, representative record.

4. The *return period, T,* (or recurrence interval), is the <u>average</u> interval between the events of a particular magnitude, and equals $1/F_R$. In the above 6-inch rainfall example the *observed T* is 1095 days or 3 years. The unmodified and modified observed return periods for the top ranked flood from 100 years of data (F_R of 0.99% and 0.7%) are 101 and 143 years.

For period in years F_R is also called the annual exceedance probability, which is equal to 1- F, where F is the cumulative probability for annual non-exceedance.

Return period is a widely used term in the scientific description of flood hazards and in flood planning and engineering. The Federal Emergency Management Agency (FEMA) National Flood Insurance Program (NFIP) uses the 1% and 0.2% return frequency floods—the 100 and 500-year return period (T_{100} and T_{500}) floods—to define properties subject to various program requirements. In turn, other federal, state, and local governments use these hazard levels to manage flood mitigation and control programs (including levee design). For urban street drainage studies hydrologists often use a return frequency of 10 years. For flood safety programs involving threats to human life and critical infrastructure and population centers (e.g., dam safety analysis) hydrologists often consider a 10,000-year return period.

Conceptually, it is important to note that F_R and T are calculated averages from the record τ and not fixed values. The 30-year record noted above could contain two consecutive days of 6-inch rainfall. Following the first day, the probability of an event occurring on the succeeding day is still 0.091% based on random independent occurrence. (A rainfall *forecast*, unlike an RFA, takes into account factors other than random occurrence that might make it more likely that two events could transpire so close together). The fact that two of the ten events take place on consecutive days DOES NOT alter the overall long-term frequency and return period. Not only can hydrologic events occur at much shorter intervals than the return period, they can also occur at much longer intervals. Two of the ten rainfall events in a 30-year record could be much more than three years apart.

For any specified return period (T) flood the likely number of recurrences over any time-frame τ shorter or longer than T is provided by a probability formula—known as the *Poisson distribution*. This means, for example, that a probability can be calculated for any number of recurrences—e.g., only once, at least once, exactly three, etc.—of a 100-year (1%) flood over any given time-frame. This probability is true for any 100-year event as it is not a function of the type event, the location, or the particular 100-year event magnitude.

Figure 12a illustrates the cumulative probability for the Poisson distribution. Curves are presented for four estimated average number of recurrences (1, 2, 5, and 10) irrespective of the time-frame (τ) and show the probability of at least that number of events actually recurring in that τ. For example, the

probabilities for floods expected to occur 1, 2, 5, and 10 times over a record actually occurring that many times or more are 63, 59, 56, and 54%, respectively.

The probability *P* of *at least one T (1/F_R)* flood of occurring in a given τ can be calculated using the following version of the Poisson distribution: [6]

$$P = 1 - (1 - F_R)^{\tau}$$

For example, the probability of a 1% flood occurring during a 30-year period, the length of a typical mortgage, is 26%. Again, this is true of 100-year floods everywhere. Figures 12a and b illustrate the increasing probability of occurrence with increasing τ for T_{10} , T_{20} , T_{50} , T_{100} , T_{500} , and T_{1000}. The cumulative distributions for at least one recurrence are illustrated in Figure 12b.

Importantly, as revealed in Figure 12b, the probability of a *T* flood occurring at least once in τ equal to T is **NOT** 100%. The probability of a T_{100} flood (or greater) occurring during a given 100 year time-frame is 63%. The probabilities do not exceed 95% and 99% for at least one *T* occurrence until τ approaches 3*T* and 5T. Not illustrated on the figure but also of note, the probability for at least three *T* occurrences does not exceed 95% until τ approaches 7*T,* and 9*T* for at least five occurrences. *Thus, records for evaluating return frequency floods should be many times longer than the return periods of interest.*

H. Return Frequencies as a Function of Magnitude

In many cases sufficiently long records are not available for evaluating extreme return period floods. Furthermore, there is often reason to suspect that highly ranked floods are of much lower return frequency than would be estimated (e.g., a flood ranked 1 out of 100 years appears to be well beyond a 143-year return period based on flood records at other nearby watersheds). In these cases hydrologists need to estimate extreme *T* (or 1/F_R) as a function of flood magnitude, and vice versa. Hydrologists therefore employ probability distributions which describe (i.e., model) a correlation between the increasing flood magnitude and increasing return period (declining return frequency).

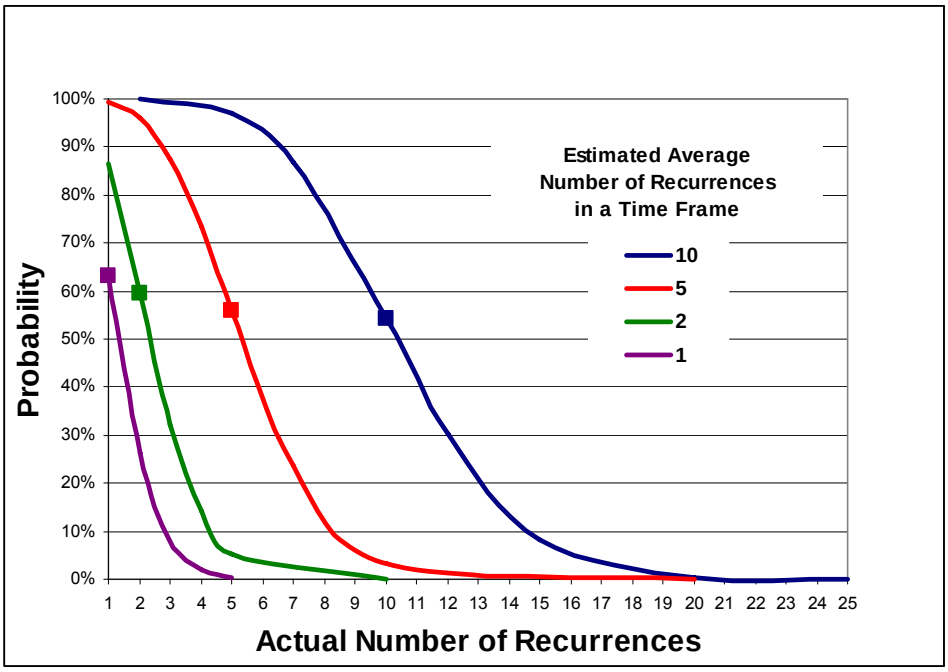
Figure 13a presents a 44-year record of annual peak river flow (Chow et al 1988). Figure 13b is a graph of the modified return period versus magnitude—thus the return period for the 1/44 ranked event is 44.4/0.7 or 63 years. Figure 13b shows that the normal distribution (using data μ and σ) is a poor fit, even below the 10-year return period. Modeling the relationship between flood magnitude and return period can usually be improved with probability distributions which accommodate combinations of a) log, exponential, or other power function transformations of the return period, b) skewness, and c) kurtosis and additional shape parameters.

---

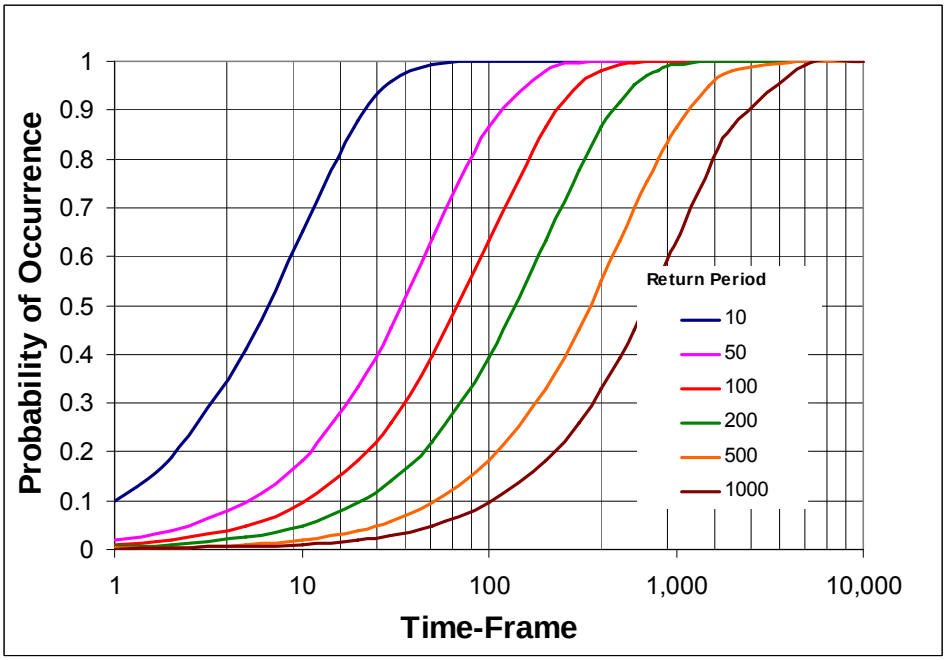
[6] This equation is a cumulative distribution for the Poisson Distribution. Note this is the case of *at least one* recurrence and the probability is greater than for the case of *exactly one* recurrence. For a Poisson Distribution the probability of *exactly one* recurrence in T is

$$P = T * F_R * \exp(-T*F_R) \text{ OR approximately } P = T * F_R * (1 - F_R)^{T-1}$$

For a 100-year flood or greater, the probability of one and only one recurrence in 100 years is exp(-1) or 37%. Over large T, the Binomial Distribution converges to the Poisson Distribution

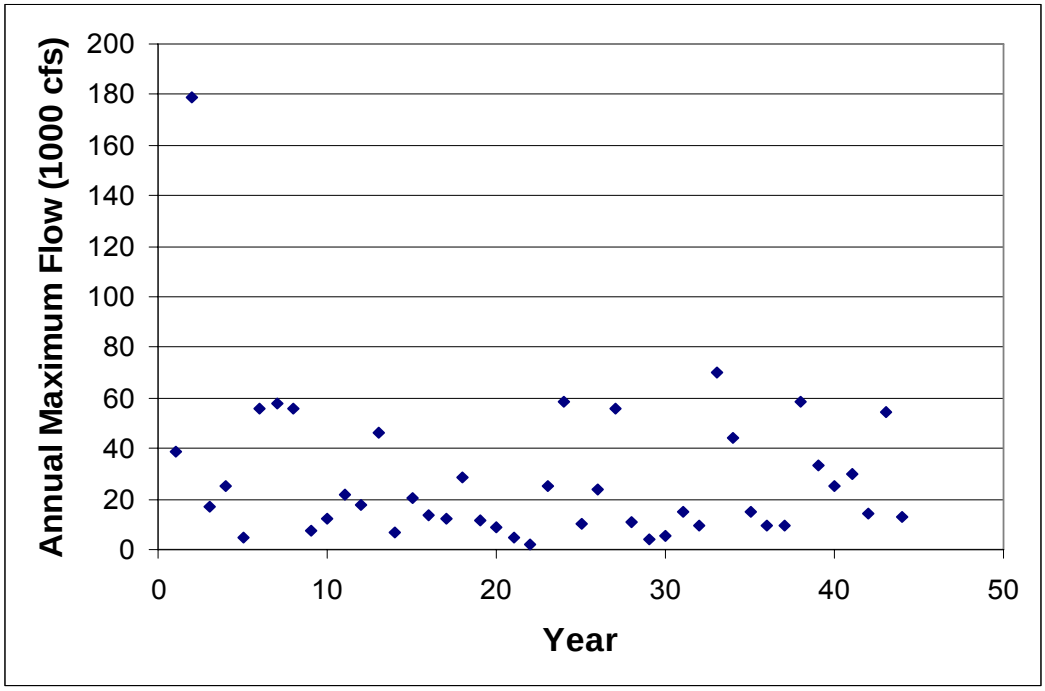


a) Cumulative Probability of Recurrences for Any Time-Frame Greater Than or Equal to Estimated Average

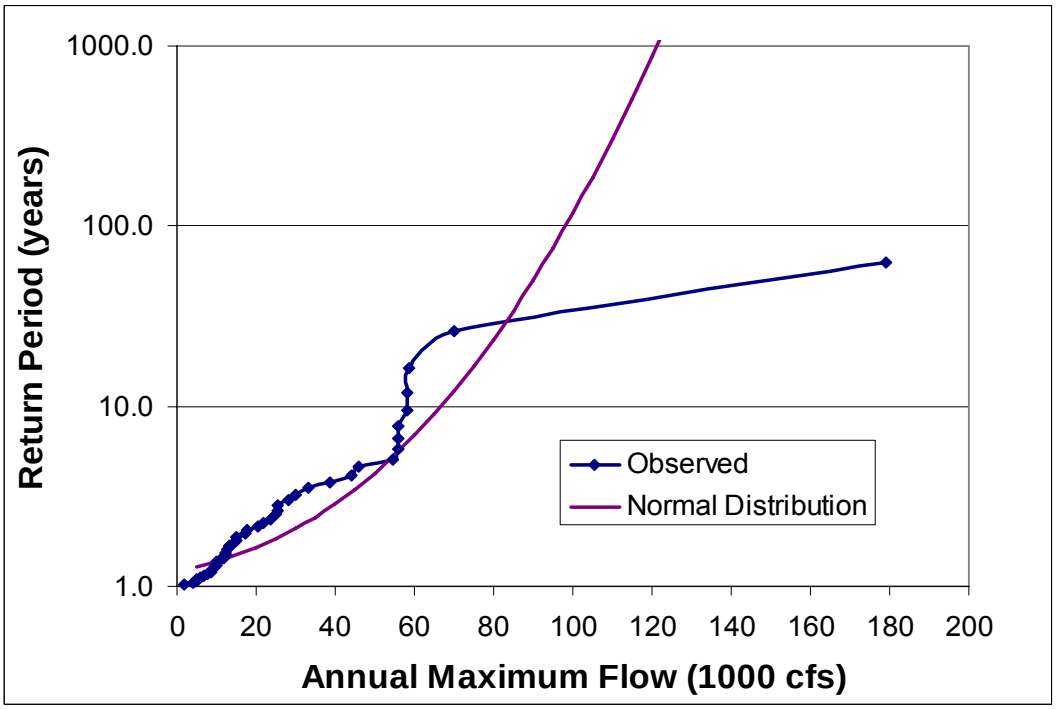


b) Cumulative Probability For At Least One Recurrence of a Return Period Event Over Different Time-Frames (Log Scale)

Figure 12. Poisson Cumulative Distribution



a) Data Record



a) Observed versus Normal Distribution Return Period

Figure 13. Example Annual River Flood Data

Several common probability distributions used in describing extreme return frequencies are:

- Log-normal distribution, takes the logarithm of the return period (log (100) = 2) and uses a normal distribution of the magnitude versus log transformed return period (or return frequency); defined by two parameters, mean and the variance.
- Log Pearson Type III distribution, adds the coefficient of skewness to the log-normal distribution and reduces to the log-normal distribution when coefficient of skewness equals zero.
- Generalized Extreme Value, (GEV) distribution includes three parameters: the mean, variance, and a shape parameter in an exponential form. The shape parameter allows for a higher probability of extreme values, i.e., a"fat tail" (Gordon 2004; Zervas 2005).

  The GEV Type I, or Gumbel, distribution, has a shape parameter equal to 0, resolving to a fixed coefficient of skewness of 1.14. The GEV Type 1 is defined by two parameters, μ and σ, with a double exponential equation.

  The GEV Type II, or Frechet, distribution has an adjustable skewness and shape parameter, allowing for a fatter tail.

  The GEV Type III, or Weibull, distribution has fatter peak and the tail goes to zero at an extreme value.

  Figure 14 illustrates the three GEV type distributions.
- Generalized Pareto distribution, a three parameter power law function.

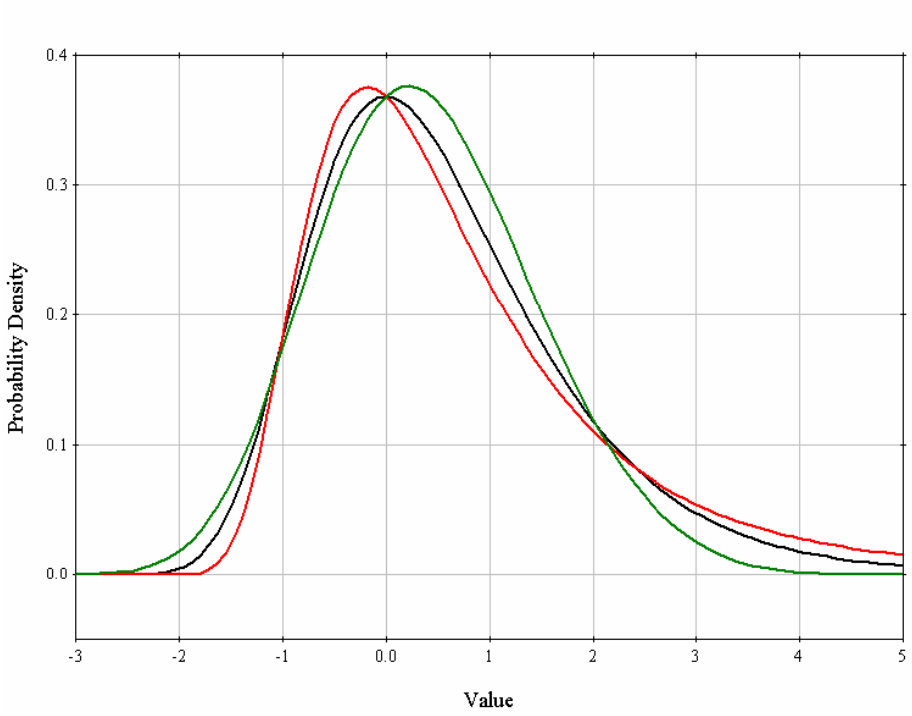


Figure 14. Example of GEV Distributions

GEV Type I (black), GEV Type II (red), GEV Type III (green). (Zervas 2005)

To apply these probability distributions to a long-term flood record hydrologists first determine dataset usability based on length, continuity, vertical accuracy (see GTN 2), and other sources of error. Next they determine dataset μ, σ, skewness, kurtosis, and other shape moments. The hydrologist then calculates values for the equation parameters using the dataset statistics. For example, the CDF, F(x), for the GEV Type I (Gumbel) distribution is:

$$F(x)=\exp\left(-\left(\exp\left(-\frac{x-u}{\alpha}\right)\right)\right)$$

$$x = flood\,magnitude$$

$$u=\mu-0.5772\alpha$$

$$\alpha=\frac{\sigma\sqrt{6}}{\pi}$$

The quality of the fit can be assessed for the lower ranked events. The purpose of using one of these distributions is that the observed return periods for the extreme events appear to be underestimated so evaluation of fit with respect to these points is not appropriate.

Figure 15 shows the observed data and log normal, Gumbel, and Log Pearson Type III distributions for the 44-year flood record given in Figure 11. As shown in Figure 14, all three distributions are better fits than the normal distribution for the 10-year return period and below. However, above the 10-year return period these equations can deviate markedly from the observations—and from each other. This is consistent with a record needing to be many times longer than the return period of interest in order to provide an accurate estimate. All three distributions indicate that the most extreme observed event (179,000 cfs) has a return period above 100 years—compared with the *observed T* of 63 years.

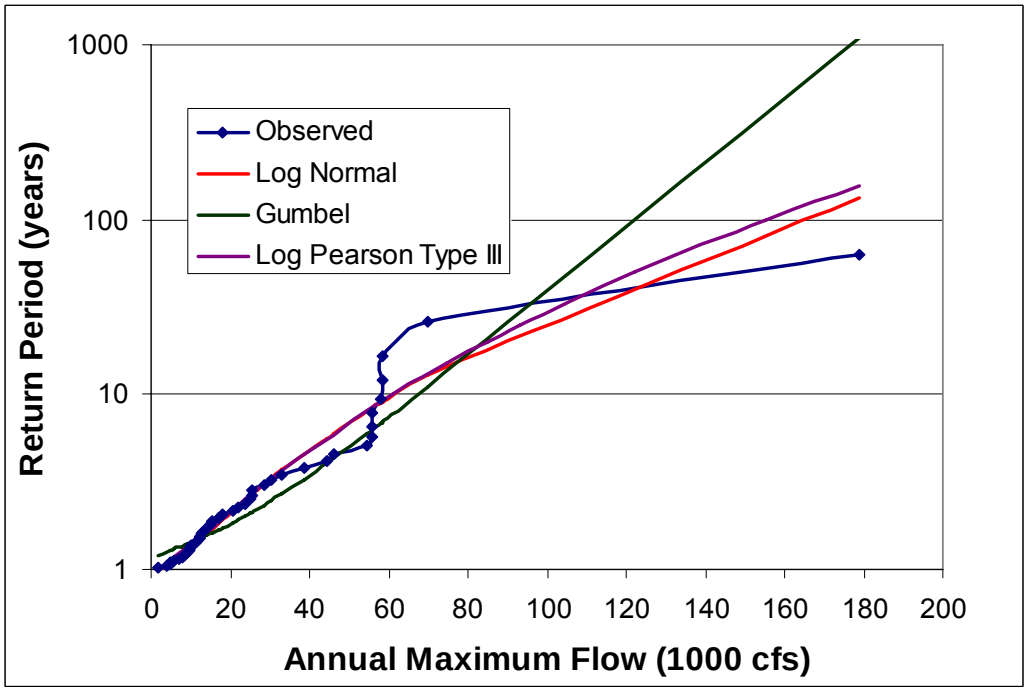


Figure 15. Frequency Distributions for Annual River Flood Data

Extreme return period flood magnitudes (e.g., 100-year or 500-year) can be estimated using a selected probability distribution. Table 1 presents the estimated 5-, 10-, 25-, 50-, 100-, 200-, and 500-year floods using the three distributions (log normal, Gumbel, and Log Pearson Type III) shown in Figure 14. These same return period floods are illustrated in Figure 16. The difference between the Log Pearson Type III and the log normal curves is a result of the skew in the log transformed record.

The estimates for the 5-, 10-, and 25-year return period floods are fairly consistent with each other and the observed return period—not surprising given the data record of 44 years. The estimates for 50-, 100-, 200-, and 500-years show increasing divergence.

Table 1. Estimated Return Period Floods for Three Distributions (1000 cfs)

Return Period (years)	Log Normal	Gumbel	Log Pearson Type III
5	41.0	49.9	41.1
10	61.7	67.5	61.3
25	95.4	89.8	93.3
50	126.3	106.2	122.1
100	162.6	122.6	155.3
200	205.0	138.9	193.3
500	271.3	160.5	251.6

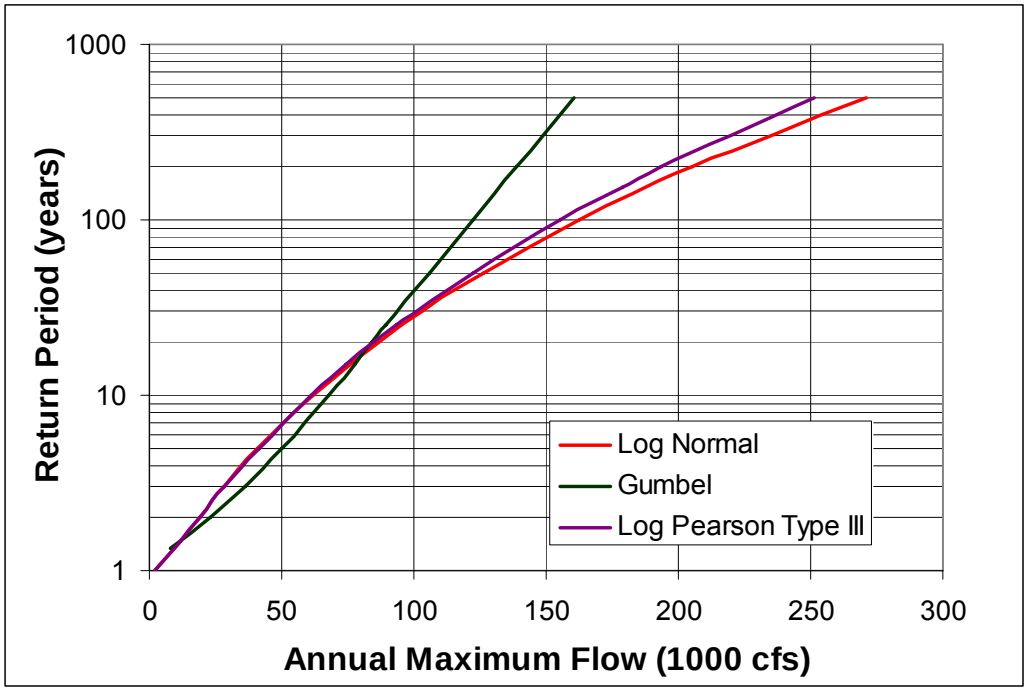


Figure 16. Estimated Return Period Floods for Three Distributions

In estimating an extreme return period flood magnitude it is important to also estimate the uncertainty of the estimate. Table 2 and Figure 17 illustrates the 90% Confidence Limits for the Log Pearson Type III distribution. Importantly, since the Log Pearson Type III distribution is skewed the confidence limits are also skewed. The interval between the mean and the UCL estimate is higher than the interval between the mean and LCL estimate.

Table 2. Confidence Limits for Estimated Return Period Floods for Log Pearson Type III Distribution (1000 cfs)

Return Period (years)	90% LCL	Mean	90% UCL
5	34.8	41.1	49.8
10	50.6	61.3	76.7
25	74.7	93.3	121.8
50	95.7	122.1	164.2
100	119.3	155.3	214.6
200	145.6	193.3	273.9
500	184.9	251.6	367.7

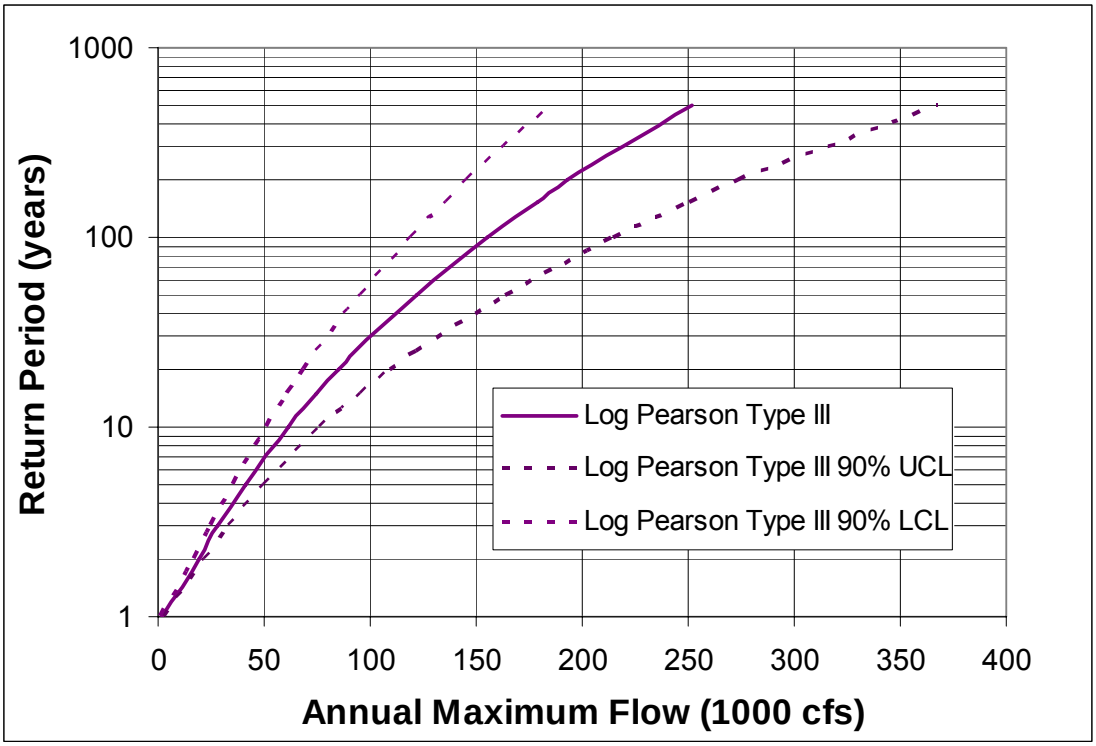


Figure 17. Confidence Limits for Estimated Return Period Floods for Log Pearson Type III Distribution

Over recent decades hydrologists around the world have used probability distributions to describe river flood return frequency, often using multi-decadal gauge records (McMahon and Srikanthan 1981; Nguyen and In-Na 1992, Karim 1995; Wang et al 2001, McCuen and Galloway 2010;), and occasionally using data on historic and paleo-floods (Kochel and Baker 1982, Benito et al 2004). Importantly, no one equation provides the best fit for all flood data, and the specific application of a particular curve type to the correlation of flood magnitude and return period could be deemed an example of "over-fitting," and is somewhat arbitrary—as the addition of more data over time could easily lead to selection of a different distribution. Any extrapolation for extreme floods also assumes that key conditions governing flood physics and meteorology during the record will remain constant. The published literature provides a basis for selecting a few probability distributions for evaluation. The Log Pearson Type III is the primary equation used by the US Geological Survey to evaluate river flood frequency (Chow et al 1988).

I. Joint and Combined Probability

When events are totally unrelated—i.e., the occurrence of one has no influence on the probability of the other—the probabilities of the two events are deemed *independent* (e.g., each successive coin toss in a series has an independent probability of 50% of turning up heads). Events can also be *dependent*, where the probability of one event depends on previous or concurrent events (e.g., drawing cards from a deck). Events are mutually exclusive if one event precludes the other from occurring.

When there is a conjunction of two or more independent events, each with its own probability distribution, joint probability is equal to the product of the individual event probabilities. When the number of independent events increases, the probability of conjunction rapidly diminishes. The joint probability for N successive coin tosses to all turn up heads is equal to 0.5^N; thus the probabilities of two through six successive tosses resulting in heads are 25%, 12.5%, 6.25%, 3.12%, and 1.56%.

Combined probabilities for multiple outcomes or multiple chances can also be computed. If the events are mutually exclusive, the combined probability is simply the sum of the event probabilities (e.g., the probability of rolling a 1 or 2 on a single dice roll is 33.3%, also the probability of rolling a 1 at least once on two rolls of a dice). When the events are not mutually exclusive the various joint probabilities must be discounted (e.g., the probability of rolling a 1 only once on two rolls of a dice). For low joint probabilities the combined probability approaches the sum of the individual probabilities. While joint probability declines with increasing number of events, the combined probability rises dramatically as the number of outcomes/chances increases. For six coin tosses, the probability of at least one heads occurring increases from 50% to 75%, 87.5%, 93.75%, 96.88%, and finally 98.4%.

The affect of uncertainty is different on joint and combined probabilities. In the former it is multiplicative whereas in the latter case it tends to be additive. If the probability of an event has been determined as a joint probability—e.g. an event has been determined to be a 100-year event as a result of the joint probability of three 21.3% events ($0.213^3 = 0.01$)—then increasing each of the three individual probabilities to 27.2% doubles the joint probability to a 50-year event.

J. Joint Probability Analysis

Joint probability analysis (JPA) becomes more complex with multiple independent events each described by a probability distribution. For example, consider a hypothetical case of three channels, with separate independent flow frequency distributions, which combine (assume steady flows) in a fourth channel. The flow frequency for the fourth channel is a joint probability of the three tributaries.

Suppose the daily flow averages range from 2,000 to 50,000 cfs for the first tributary, and from 1,000 to 20,000 and 1,000 to 18,000 cfs in the other two. If each tributary's flow is considered in 1,000 cfs increments, there are 17,640 probability combinations for the fourth channel. Any particular flow in the fourth channel—e.g., 40,000 cfs—could occur as a result of many different combinations—e.g., 20,000 plus 10,000 plus 10,000 cfs respectively; 16,000 plus 13,000 plus 11,000; etc.. If the CDFs are known for each tributary flow then their probability density forms can be used to calculated the resulting joint probability for any individual combination of tributary flows.[7] The joint probability (probability density form) for any combinations of tributary flows is the product of the corresponding probability densities.[8] To determine the flow frequency (CDF) for the fourth channel this product is integrated over the range of tributary flows.[9] Standard math software and a computer can easily calculate values for this integral using the 17,640 input combinations and thus estimate flow frequencies for the fourth channel. This approach is termed a full Joint Probability Method or Full JPM.

The flow frequency estimate (the estimate of the integral) for the fourth channel will improve with smaller input value intervals. Input values at a 100 cfs increment (17 million combinations) versus a 2,000 cfs increment (2,250 combinations) will provide a more versus less precise estimate compared to the 1,000 cfs increment. Thus, a full JPM would start with coarser inputs and then perform successive re-calculations until the resulting flow frequency estimates converge.

Another approach is the *Monte Carlo* JPM technique. This technique utilizes a large enough number of possible random combinations to satisfactorily determine the joint probability. In the Monte Carlo analysis, a frequency for each independent event is randomly drawn—for example a separate random number between 0.000 and 1.000 is selected for each tributary (three separate draws). After drawing a frequency for each tributary, the corresponding flow magnitudes for each tributary are found from the respective tributary probability distributions. These flow magnitudes are then added to give the flow in the combined (fourth) channel. The process is repeated, creating a *synthetic record* for the fourth channel, which is then evaluated using suitable probability distributions. More randomly derived records can be added until the probability distribution stabilizes. In cases where an optimal set of input combinations is difficult to determine in advance, the Monte Carlo technique can often be used to efficiently arrive at a stable estimate of the joint probability distribution.

Monte Carlo analysis can also be used to address uncertainty in the estimated magnitude associated with each individual input. For example after a frequency value is drawn for a particular tributary, an additional draw is used to determine the flow magnitude from the uncertainty distribution for possible values at that frequency—with the most likely flow value at that frequency having the highest probability of being drawn. This doubles the total draws for each calculation from three to six.

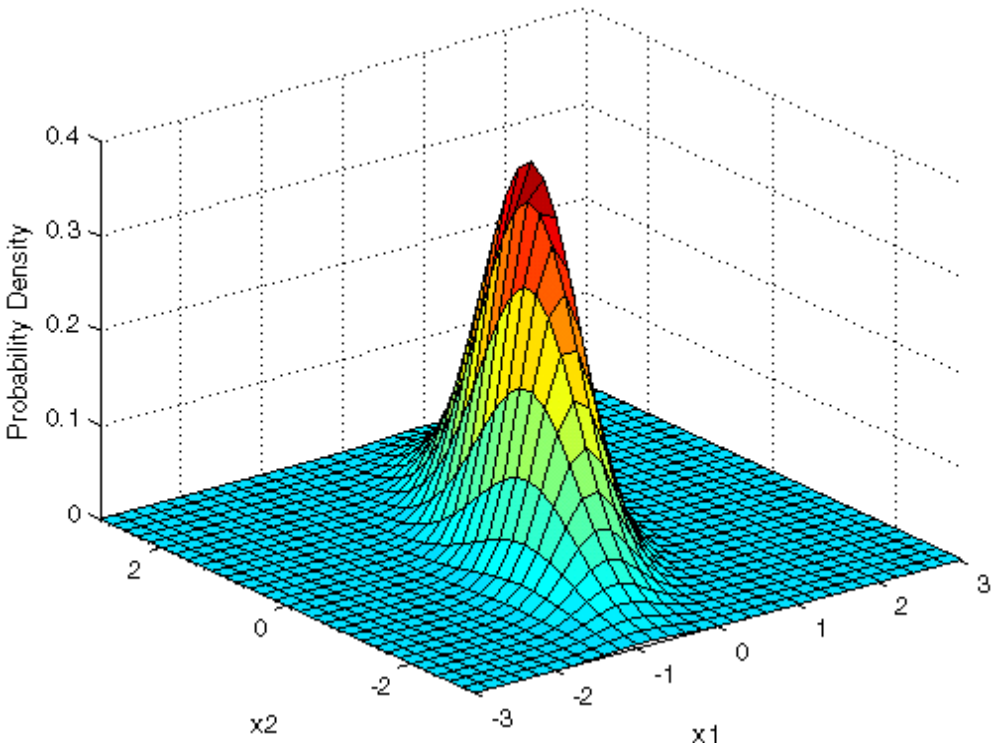
One difficulty in communicating joint probability is visualization. The joint probability "space" can be readily depicted for two independent events as a three-dimensional contour, as in Figure 18. However, for three or more events (as in the above case) such illustrations are not available.

---

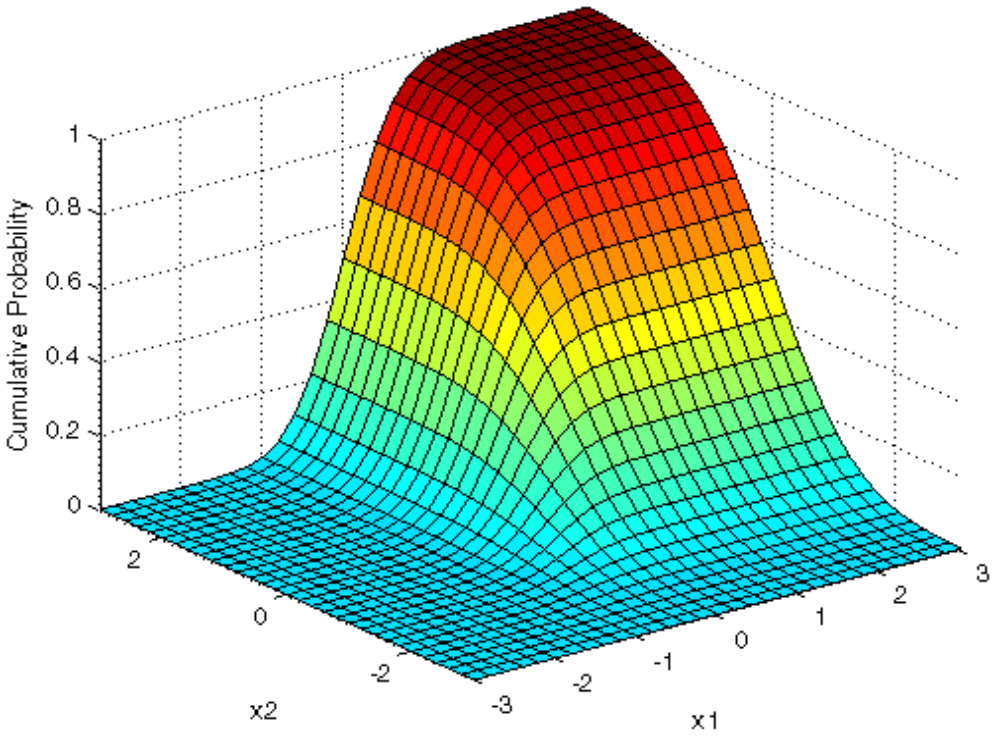
[7] The cumulative frequency distribution form is referred to as F(x), where x is a particular flow and F(x) is the cumulative probability for flow equal or exceeding x with a value between 0 and 1.0, see Figure 3b for a normal Distribution. The probability density form is referred to as f(x), where f(x) is a discrete probability density value for a specific flow, x, see Figure 3a for a normal Distribution).

[8] Or $f(x_1 , x_2 , x_3) = f(x_1) * f(x_2) * f(x_3)$

[9] In calculus this is described by a triple integral, one for each tributary. $F(x) = \iiint f(x)\, dx$



a) Probability Density



b) Cumulative Probability

Figure 18. Illustration of Joint Probability for Two Events

http://www.mathworks.com/help/toolbox/stats/brn2ivz-89.html

When the overall return frequency (probability) of a certain joint flood event is known (or reasonably estimated)—e.g., for the fourth channel say F_R of 0.02, or T_{500} , is 60,000 cfs—this return frequency can be considered the sum of all the joint probabilities for each combination of events that produces it—e.g., the three tributaries flows that can produce total flow exceeding 60,000 cfs. The joint probabilities for the various individual input combinations can then each be considered a fraction of 1 (and together they sum to 1) and the return frequency for any individual combination is thus the overall return frequency—0.02—multiplied by the fractional joint probability.

Advanced statistical techniques can be employed to select an optimal sample, or subset, of input combinations to represent the universe of combinations which sum to 0.02. For example a universe of thousands of combinations can be represented by a subset of 100 combinations. Each individual input combination is assigned a weight that is multiplied by its fractional joint probability. In this case one combination might be used to represent more than 1% of the contributing probabilities and so it could be weighted higher, e.g., at 0.015, while another representing less that 1% could be weighted lower, e.g. at 0.005. The weighted joint fractional probabilities of each combination in the subset again sum to 1, or an overall return frequency of 0.02. In this case the subset size and weightings are targeted to reduce the likelihood of error (uncertainty) in representing the joint probability "space" over a particular region—i.e., T_{500}.. This technique is known as JPM-Optimized Sample, or JPM-OS.

Depending on the number and distribution of contributing events, and the handling of uncertainty, joint probability analysis can be very computationally demanding. If conditions in one event are highly correlated to another—e.g., flow in two tributaries—these correlations can be used to reduce the number of independent events. With advances in computational resources, joint probability analysis has become a powerful tool in estimating a wide variety of recurrence events in river flood hydrology—such as the conjunction of a dam failure with extreme flood conditions.

K. Combining Deterministic and Probabilistic and Analysis

In *stochastic analysis* Hydrologists often combine deterministic and probabilistic methods. Processes that are described as *deterministic*—such as weir flow, which is a function of upstream stage, downstream stage (in the case where the weir is "drowned"), weir length, and the empirical weir crest friction coefficient—may have one or more input values and factors with appreciable random variability. Conducting a series of deterministic analyses (e.g., solutions to the weir equation) by varying an input value in accordance with its probability distribution (e.g., upstream stage) can provide the probability distribution of the output condition (weir flow). This combining of deterministic and probabilistic analysis is called *stochastic analysis*.

When more than one input value is characterized by a probability distribution, a joint probability analysis is also included. An example of stochastic analysis which also employs a joint probability analysis would be to examine the diversion flow rate at the Bonnet Carré Spillway using the defined weir length, a general value for the empirical crest friction coefficient (obtained from physical experiments on a scale model or from publications on similar weirs), and frequency distributions for the upstream and downstream stage. In this case the Monte Carlo technique could be employed to draw values for the two stages, which are then input into the weir equation to calculate the associated diversion flow. Repeating this draw hundreds of times and inputting the values into the weir equation yields a simulated record of diversion flows, which can be fitted with a probability distribution. The stochastic analysis can be expanded to also address uncertainty in the weir coefficient.

L. Flood Return Frequency Analysis

In locations with very sparse river gauge data the local flood hazard is evaluated in a flood RFA employing stochastic methods and three components:

- A rainfall return frequency probability distribution is developed using data from the nearest meteorological site, including the extrapolation of extreme rainfall return events.
- Overland flow and runoff within the basin (broken down into various sub-basins or watersheds) are analyzed for extreme rainfall events using empirical hydrologic models. In addition to the rainfall event, these models have inputs/conditions/coefficients/etc. for watershed characteristics (e.g., antecedent soil moisture, size, topography, land cover, runoff storage and losses, and distance to channel). This step translates the rainfall into input flow hydrographs at appropriate points along the overall river tributary system network.
- Flow routing within the channel network takes the set of input flow hydrographs and computes flood profiles throughout the network floodplain, using a one- or two-dimensional numerical model of the flow physics (i.e., a hydraulic model). These models require conditions/coefficients on conveyance (e.g., channel and floodplain bathymetric/ topographic data and friction—such as Manning's *n*), upstream and downstream boundary conditions, (e.g., pumping conditions for forced drainage systems), and selection of numerical settings.[10]

Local extreme (e.g., 100-year return) floods for a small basin are estimated using a single basin rainfall probability distribution input. Flood RFA for very large basins requires evaluating joint probabilities. Deterministic hydrologic and hydraulic models are employed to route a set of basin-wide rainfall events, creating a set of downstream synthetic flood events. The set of events can be developed using the full or optimized joint probability method or the Monte Carlo technique. The synthetic downstream results are then analyzed to determine the return frequency floods. The Monte Carlo technique can be used to address uncertainties for the input hydrographs, boundary conditions, conveyance properties and coefficients, and numeric parameters.

Stochastic modeling for flood RFA has been widely adopted by hydrologists throughout the world. In the U.S. most flood RFAs are performed for the federal NFIP with numerous applications in southeast Louisiana, including studies of drainage systems in the metropolitan New Orleans area. The NFIP has published guidelines for flood RFA (FEMA, 2009). The combination of rainfall RFA with deterministic hydrologic/hydraulic models can provide a very reasonable flood RFA. Three major limitations are:

a. High quality representative rainfall data from a nearby location;
b. Accurate available channel and floodplain data for setting conditions/coefficients for the hydrologic/hydraulic model; and
c. Adequate hindcast data for calibration and validation of the overall stochastic method.

[10] A discussion of river flood model numerics is beyond the scope of this General Technical Note. A discussion of numerical issues associated with two-dimensional surge modeling is presented in General Technical Note 5.

M. Sensitivity, Calibration, and Validation Tests

When hydrologists employ stochastic models—such as in a river flood RFA—the models must be tested to assess their accuracy (potential bias) and precision (uncertainty). Three types of tests are performed:

1. *Sensitivity tests* evaluate the influence of uncertainty in key input values (inflow hydrographs, boundary conditions, conveyance conditions/coefficients, numeric settings), which are not otherwise assessed (e.g., with the Monte Carlo technique). By varying an individual setting or parameter in a highly controlled way (i.e., by keeping all other inputs constant)—typically within a range of interest—a comprehensive set of sensitivity tests can rigorously evaluate the variability of model results. In practice, modelers often conduct limited sensitivity testing to assess selected variations in settings/parameters.

2. *Calibration tests* determine values for those conditions/coefficients/settings/etc. for which results are highly sensitive. The hydrologist simulate an actual historic flood—termed a *hindcast*—and compares the simulation results for flood stages and flow rates (velocities) to actual observed values. Comparing simulated versus observed peak values and/or hydrographs the hydrologist adjusts the conditions/coefficients/settings/etc. in the hydrologic and/or hydraulic model within accepted ranges to achieve a best fit. Calibration often requires judgment on whether or not to "tune" the input value to improve the simulation of a particular event. For example, it may not be appropriate to select a friction coefficient that calibrates a relatively shallow flood when many of the floods to be simulated involve deeply submerged flow. When there are multiple conditions/parameters, calibration tests must optimize the selection of values.

3. *Validation tests* assess the selected conditions/coefficients/settings/etc. for an additional separate hindcast. The results of this additional simulation are also compared to actual data. In addition to assessing the quality of hindcast values along the river network, in flood RFA an important validation test is to compare the computed flood frequency distribution with observed distributions for one or more long-term gauge stations.

The evaluation of sensitivity, calibration, and validation tests is similar to correlation and trend analysis. Graphs are usually made to compare model results versus observed data, or results from one simulation versus another in the case of sensitivity tests. These can include scatter plots of peak values across the river network and time-series hydrographs for selected locations. In the first case (see Figure 19) the test is judged by proximity of points to the line corresponding perfect agreement (slope of the line equals 1). In the second case (see Figure 20) the model hydrograph is plotted alongside the observed hydrograph.

The comparisons are then assessed for *accuracy*—bias, or deterministic variability or error—*and precision*—degree of random variation. Bias in comparing sensitivity test results (e.g., consistent offset) indicates that the model is sensitive to the input/condition/coefficient/setting/etc. being tested. Bias in a calibration or validation test—a consistent pattern of over- or under-prediction of the hindcast—indicates that the model has not captured (and may not be able to a capture) an important aspect of the physical process. Bias can be associated with local portions of the river network (e.g., along one particular reach) or they can by system-wide. Bias can arise from a limitation in the model input/conditions/coefficients/settings/etc or in the accuracy of the hindcast observations.

Bias in the results is not always easy to spot. The hydrologists must carefully review the results of calibration and validation tests to look for bias. Once UCL/LCLs are established for the random variation between model and observed results (see below), these can be used to assess the relative importance of biased results. If a sound theory on the source of bias can be identified and supported, the hydrologist can correct the model (e.g., adjusting channel depth or the friction coefficient) or the observed data (e.g., adjusting a stage observation to correct for a benchmark error). If corrections to the model and data are not supported, the data should be given priority and a method for post-simulation correction is required.

Following bias correction the random variation between two sensitivity tests, or between a calibration/validation test and the observation, is evaluated by computing RMSE and R^2 , and by examining the probability distribution for the remaining variation. The random variation distribution for a calibration/validation test is an indicator of uncertainty in the model. A normal distribution will usually be sufficient for many comparisons—such as the comparison of modeled and observed peaks and hydrographs in Figures 19 and 20. For some highly skewed comparisons—e.g., a predicted versus an observed flood return frequency distribution for a gauge station—a skewed probability distribution will be required. The distribution of uncertainty is used to determine UCL/LCLs above and below the model results.

The hydrologist may use the UCL/LCLs to note groups of points that reflect significant bias[11] or individual points that are markedly beyond the UCL/LCL—termed *outliers*. If there are reasons to believe that outliers are not valid, the hydrologist may exclude them. The uncertainty distribution and the UCL/LCLs are then recomputed following removal of bias and outliers.

The hydrologist will typically discuss the quality of test results in terms of the RMSE, R^2, and the spread of the UCL/LCL bands, and compare these values to other flood RFAs for regions with similar characteristics.

N. Application of Flood Return Frequency Analysis

Flood RFAs have five major applications:

1. *Flood Insurance Studies (FISs) under the NFIP.* Congress established the NFIP in 1968 and as part of the program set special 100- and 500-year flood hazard trigger levels for assessing insurance premiums. The NFIP establishes requirements for participating local communities to regulate floodplain development to reduce the likelihood of claims. Thus, the NFIP requires that these particular return period floods be determined and delineated on Flood Insurance Rate Maps (FIRMs).
2. *Flood Control Studies*. Flood control studies combine flood RFA—i.e., evaluation of the flood hazard—with estimates of potential consequences—e.g., property damages at the peak flood stage and lost economic production, loss of non-evacuee lives in cases of dam failure—in a *risk assessment*. Studies compare the benefits (risk reduction) associated with proposed flood control projects with the project costs. Selected projects then use the risk reduction criteria as the basis for design.

[11] Actual statistical significance of bias can be evaluated with a χ^2 Test.

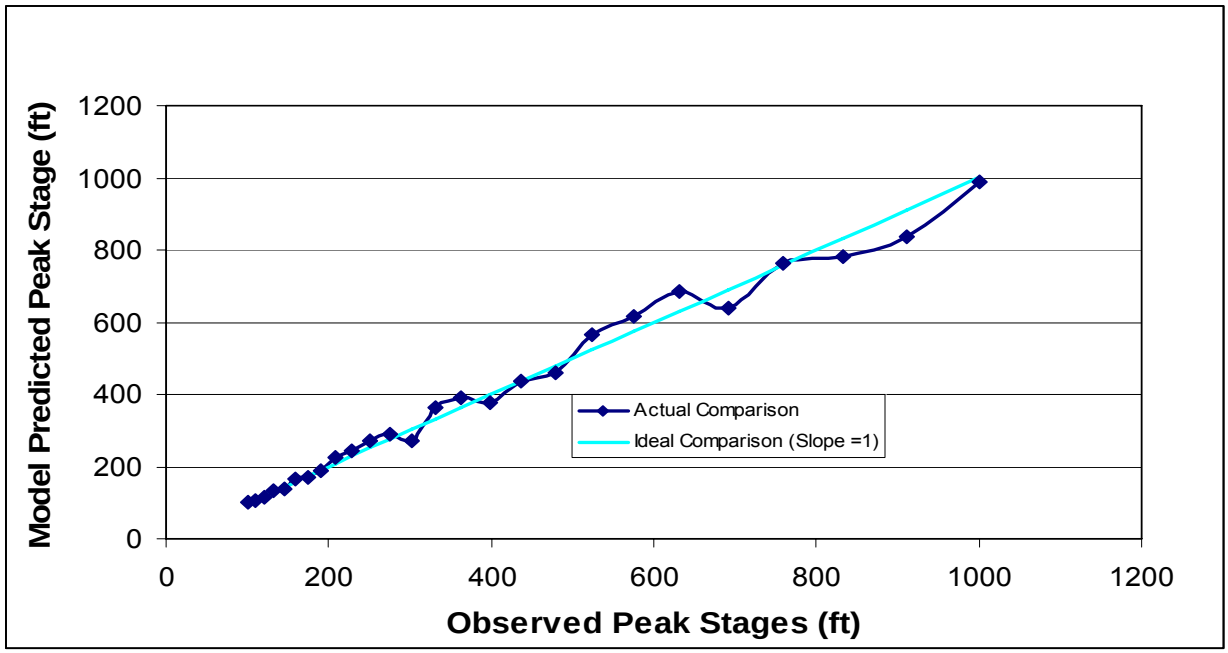


Figure 19. Example Comparison of Observed Versus Modeled Peak Stages

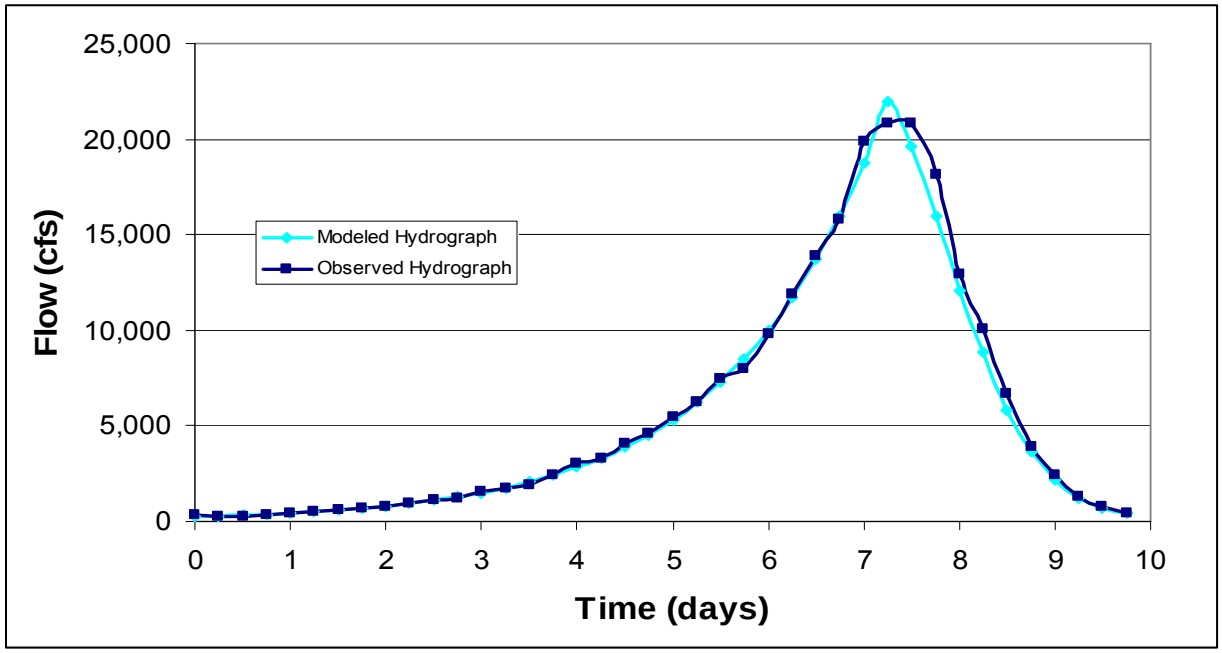


Figure 20. Example Comparison of Observed Versus Modeled Hydrograph

3. *Flood Impact Studies*. Urban and regional planners use flood RFAs to assess the potential for changing conditions to increase extreme flood returns and magnitudes. Common examples

include urbanization and regional land use changes, floodplain development, channel modifications, and flood control projects (which while reducing flooding in the target area may adversely affect flooding elsewhere). "What if" scenarios can be evaluated by changing appropriate inputs/conditions/coefficients in the flood stochastic model. The change in outcomes versus changes to the model can be examined using correlation analysis. Benefit-cost analysis is used in alternative evaluation.

4. *Flood Mitigation Studies*. Similar to flood impact studies, urban and regional planners use RFAs to assess ways of reducing extreme flood returns and magnitudes through urban runoff reduction, floodplain restoration, and channel modifications. These studies also model "what if" scenarios by changing model inputs/conditions/coefficients and examining correlations with model results. Benefit-cost analysis is used in alternative evaluation.

   Flood control, impact, and mitigation studies are strongly affected by Congress' design of the NFIP—as the benefit and costs often have abrupt discontinuities at the 100-year flood threshold. Items contributing to this threshold effect include insurance premium costs, floodplain development restrictions, property value decreases, and community growth impairment. If the NFIP addressed a comprehensive range of incremental flood hazards, then project costs and benefits would adjust more smoothly in response to a range of return period floods. Smaller communities might then opt for projects reducing 50-year floods, while larger urban areas might choose to construct 1,000-year protection (or higher).[12]

5. *Long-Term Flood Trend Studies*. Expected flood frequency distribution can be compared with actual flood recurrence at a gauge station to isolate long-term trends associated with underlying climatological, topographical, or other physical conditions. Similarly, long-term trends can be combined with the normal stochastic analysis to evaluate potential increases in return period floods. For example, the U.S. Army Corps of Engineers currently mandates use of long-term sea level rise forecasts in coastal river RFA (USACE 2009).

The treatment of bias is similar for all flood RFA applications. As previously noted, evidence of bias encountered during a flood RFA should be investigated and addressed. Recommendations for handling any significant residual bias should be explicitly discussed.

The treatment of uncertainty, however, often varies by applications of RFA. For example, FEMA NFIP guidance does not address accounting for uncertainty in the estimates of the 100-year flood elevations. This may be due, in part, to the particular actuarial approaches of the NFIP and the other tools available for management of financial risks to the program. The NFIP approach to flood thresholds and the RFA uncertainties, however, can color the treatment of uncertainty in other applications of flood RFA.

[12] The arbitrary NFIP flood insurance premium thresholds, combined with a general lack of understanding that a return period is an average recurrence interval over a much longer time-frame, contribute to distorted public perceptions of flood hazards and excessive skepticism about the science of flood RFA.

In flood control studies hydrologists may make allowances for uncertainty (e.g., use of confidence bands), and to a greater degree (e.g., wider confidence bands) consistent with considerations of

- Important gaps in model calibration/validation—e.g., lack of long-term gauge data to construct station RFAs for validating the stochastic model; and/or
- The risks associated with flooding and ramifications of uncertainty—e.g., problematic evacuation, vulnerable critical infrastructure.

Flood impact and mitigation studies may make different allowances for estimate uncertainties depending on study objectives. Thus, it is not uncommon for hydrologists to employ different estimates of the same flood hazard—e.g., the 100-year flood—depending upon the application.

O. Combined Probability of Return Floods Over Large Areas

Extreme flood events in distant river basins are basically independent events. Government agencies with large geographic jurisdictional responsibilities for flood hazards must consider the combination of return floods hazards when assessing their resources. For an agency that manages ten independent flood hazard locations (assume location floods are not mutually exclusive), the probabilities for a 100- and 500-year event occurring somewhere in a given year are 9.56% and 1.98%, respectively—not quite ten times the individual probability.

Combined probabilities can be further evaluated for longer time-frames using the Poisson distribution. Figures 21 illustrates the increasing probability of an occurrence for a 500-year return period floods with increasing time frame and increasing number of independent, not mutually exclusive, locations. For example, among 10 river basins over a 10-year time-frame there is an 18% probability for the occurrence of a 500-year event.

Uncertainty can significantly affect these combined probabilities. In the above example, if the return periods of the 500-yr events in the 10 independent river basins were really 250-yrs, the probability of one flood occurring in 10 years is 33% and not 18%.

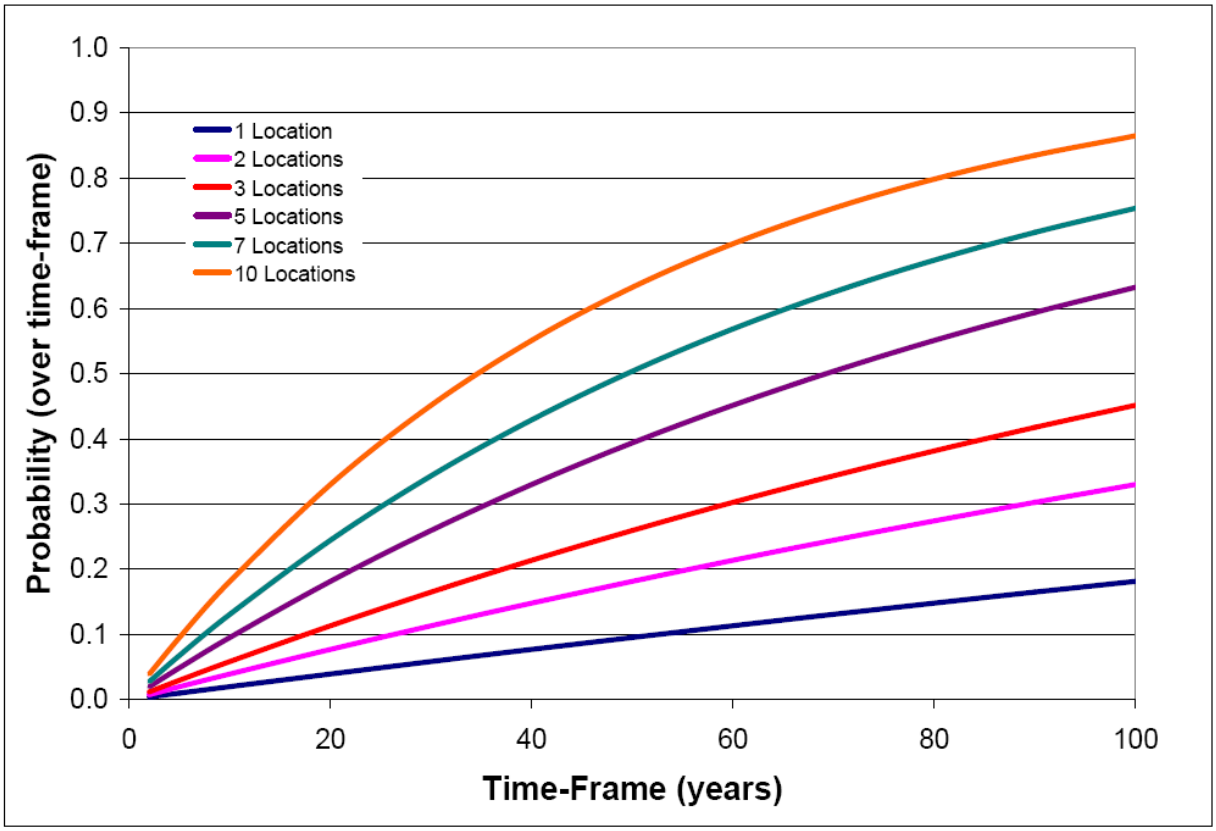


Figure 21. Return Frequency for Multiple Locations versus Time-Frame for 500-yr Flood

References

Chow, V. T., D. R. Maidment, and L. W. Mays, Applied Hydrology, McGraw-Hill, 1988.

Gerardo, B., M. Lang, M. Barriendos, M. C. Llasat, F. Francés, T. Ouarda, V. Yehouda Enzel, A. Bardossy, and A. Coeur, Use of Systematic, Palaeoflood and Historical Data for the Improvement of Flood Risk Estimation. Review of Scientific Methods, Natural Hazards Volume 31, Number 3, 623-643, 2004.

Goldman, D., Quantifying Uncertainty in Estimates of Regulated Flood Frequency Curves, Proceedings of World Water and Environmental Resources Congress 2001, Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges.

Goldman, D., Upper Mississippi Flood Frequency Study, Uncertainty in Estimates of the 1% Profile, Prepared for Rock Island District, U. S. Army Corps of Engineers, Cold Regions Research and Engineering Laboratory, U.S. Army Corps of Engineers, Hanover, N.H., August 2003

Gordon, N. D., T. A. Mahon, B. L. Finlayson, C. J. Gippel, R. J. Nathan, Stream hydrology: An Introduction for Ecologists, John Wiley and Sons, 2004.

Karim, M. A., A comparison of four distributions used in flood frequency analysis in Bangladesh, Hydrological Sciences Journal, Volume 40, Issue No. 1, 1995.

Kochel, R. C., and V. R. Baker, Paleoflood Hydrology, Science, Vol. 215 no. 4531 pp. 353-361 January 22, 1982.

Mathworks, http://www.mathworks.com/help/toolbox/stats/brn2ivz-89.html

McCuen, R. H., and K.E. Galloway, Record Length Requirements for Annual Maximum Flood Series, Journal of Hydrologic Engineering Volume 15, Issue 9, Technical Notes, p. 704, 2010.

McMahon, T. A., and R. Srikanthan, Log Pearson III distribution — Is it applicable to flood frequency analysis of Australian streams? Journal of Hydrology. Volume 52, Issus 1-2, June 1981.

Nguyen, T., and N. In-Na, Plotting formula for pearson type III distribution considering historical information, Environmental Monitoring and Assessment, Volume 23, Numbers 1-3, 137-152, 1992.

Taleb, N. N., The Black Swan, The Impact of the Highly Improbable. Second Edition, Random House, 2010

U.S. Army Corps of Engineers (USACE), Incorporating Sea-Level Change Considerations in Civil Works Programs, Engineering Circular 1165-2-211, 2009.

U.S. Federal Highway Administration, HEC-19 Workshop Manual, http://www.fhwa.dot.gov/engineering/hydraulics/pubs/hec/hec19.pdf.

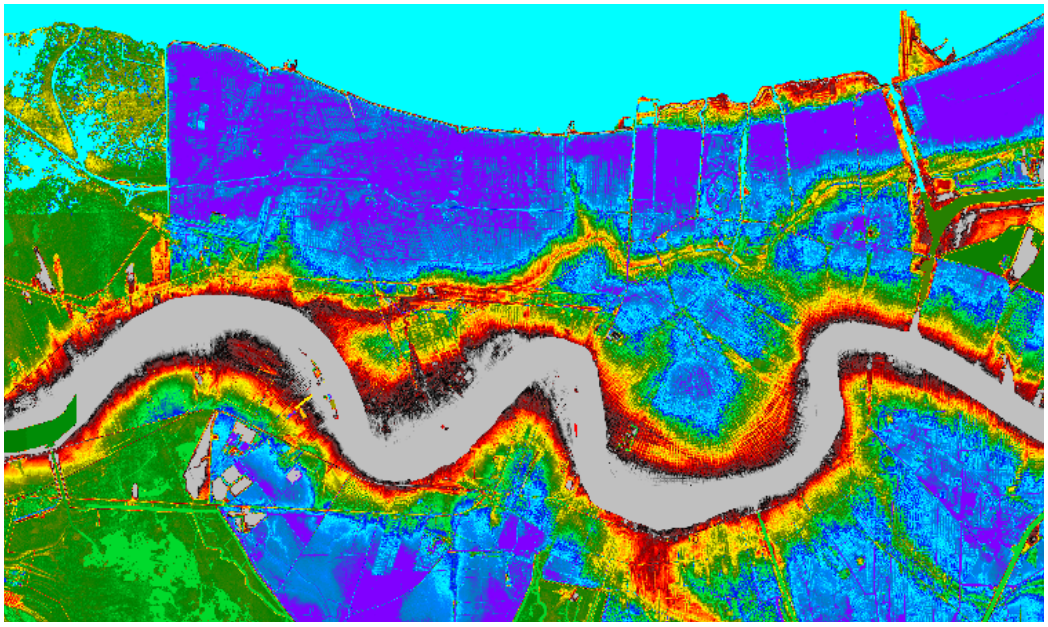
U.S. Federal Emergency Management Agency (FEMA), Guidelines and Specifications for Flood Hazard Mapping Partners, Appendix C: Guidance for Riverine Flooding Analyses and Mapping, November, 2009; http://www.fema.gov/library/viewRecord.do?id=2206.

U.S. Geological Survey, Guidelines for Determining Flood Flow Frequency, Bulletin 17B, 1982, http://water.usgs.gov/osw/bulletin17b/bulletin_17B.html.

Wang, H., D. Qin, J. Sun, and J. Wang, Study on the general model of hydrological frequency analysis, Science in China Series E: Technological Sciences, Volume 44, Supplement 1, 52-61, August 2001.

Zervas, C. E., Response of Extreme Storm Tide Levels to Long-term Sea Level Change, NOAA/National Ocean Service, Center for Operational Oceanographic Products and Services, 2005, http://tidesandcurrents.noaa.gov/est/050415-53.pdf.

General Note 2.
Elevation Data for
Hurricane Surge Hazard Analysis



A major concern with any flood hazard analysis—including inundation from hurricane surge and associated waves—is accurate vertical data. This concern applies both to surge data (e.g., high water marks and gauge readings from past events) and to topographic and bathymetric data used in surge models and flood zone mapping. Issues include proper control, or referencing, to a uniform datum; treatment of regional and local subsidence; and development of digital elevation models (DEMs).

A. Elevation Datum

National Oceanic and Atmospheric Administration (NOAA) data are typically reported relative to a station-specific *tidal datum*—e.g., Local Mean Sea Level (LMSL), Mean Lower Low Water (MLLW), or Mean Higher High Water (MHHW) determined from the analysis of multi-decadal data at that station. U.S. Geological Survey (USGS) and U. S. Army Corps of Engineers (USACE) elevations are more commonly given relative to national references—e.g., the National Geodetic Vertical Datum of 1929 (NGVD29) or the North American Vertical Datum of 1988 (NAVD88), which until recently could only obtained by leveling (surveying) from the nearest recognized benchmarks.

NGVD29 and NAVD88 are both *geoid*-based references. The geoid is a three dimensional geometric surface where the earth's gravity has an equal given magnitude. (Gravity decreases with altitude, i.e., distance from the earth's center of mass.) The geoid represents something similar to a global MSL at which all the oceans would stabilize without winds, tides, Coriolis, and other forces. The geoid incorporates irregularities in the earth's general shape (nearly ellipsoidal but not perfectly so) and regional variations in gravity with changes in the thickness and density of the earth's crust. The most consistent and reliable geoid-based reference for hydrologic data is the NAVD88. Geodesists continue to obtain new gravitational and other information and periodically improve their model of the geoid, e.g., GEOID03, GEOID09, GEOID12, and GEOID12a.

Satellite-based global positioning system (GPS) survey methods facilitate *differential* vertical measurements (see Figure 1). The satellite locations are referenced to a perfect, idealized ellipsoid (not geoid) surrounding the earth. Surveyors use GPS to measure the ellipsoid heights of a point of interest (H_A) and a nearby location of known elevation—i.e., a benchmark (H_B). The benchmark has a previously established height with respect to the geoid—termed orthometric height (h_B). The difference at each location between the geoid and ellipsoid, termed the geoid height (N_A and N_B) is developed using the latest geoid model. With values for H_A , H_B , h_B , N_A ,and N_B , a value can be estimated for h_A .

Benchmarks in portions of the Gulf Coast are notoriously problematic due to the quality of initial leveling and subsequent aging from natural subsidence. In recent years the National Geodetic Survey (NGS) has sponsored installation of a network of Continuously Operating Reference Stations (CORS)—so-called "smart benchmarks"—to facilitate more accurate and precise differential GPS elevation measurements. The CORS network also supports *real-time kinetic* (RTK, or "on the go") GPS-based leveling with or without differential surveying of a physical benchmarks. GPS-based orthometric heights relative to the CORS network can generally be measured to within ± 1 inch with respect to the applied geoid model, with proper observation (Zilkoski 2011).

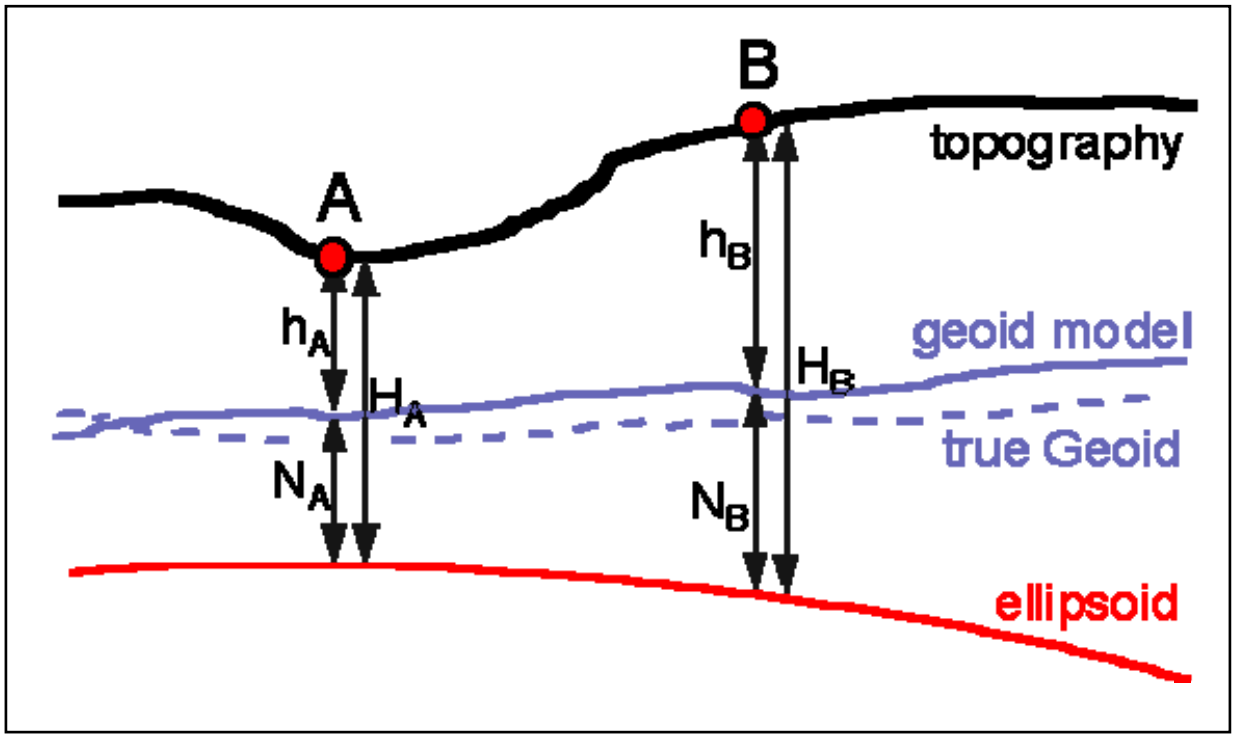


Figure 1. Illustration of Geoid and Ellipsoid
Wormley 2011

The CORS have facilitated refinements by NGS of the geoid model, significant enough to effect elevation values. Thus, references to NAVD88 should include the "epoch." "NAVD88" denotes use of older, pre-CORS control; such references are suspect as they are likely to be based on outdated benchmarks. "NAVD88-2004.65" and "NAVD88-2006.81" incorporate CORS-based control and the GEOID03 and GEOID09 models, respectively.

The LMSL reference at Grand Isle changed from 1.1 ft NAVD88, to 0.2 ft NAVD88-2004.65 with the implementation of the CORS and GEOID03, a change of 0.9 ft. Some areas were lowered even more, and a few were raised. With GEOID09, references to NAVD88-2006.81 versus NAVD88-2004.65 were generally lowered by another 0.08 to 0.16 ft (see Figure 2). Note that a change in the GEOID model is **NOT** a change in the LMSL itself. However, due to differences in the LMSL under different GEOID references it is critical that all elevation data employ the current epoch.[1]

For the purposes of hurricane surge hazard analysis, references to NAVD88-2005.65 can be readily adjusted to NAVD88-2006.81. However, data that referenced simply to NAVD88 or to NGVD29, MSL, MLLW, etc. must have their vertical control re-established. In many cases this cannot be done by a paper correction and requires resurvey of benchmarks or other approved means of re-referencing. The use of inadequately referenced vertical data, including the mixing of data of different reference quality, can contribute considerable uncertainty.

[1] The NGS and their local partner are continuing research on the GEOID model and methods of vertical control. The global error associated with the GEOID09 model is estimated to be about ± 2 inches (Zilkoski 2011).

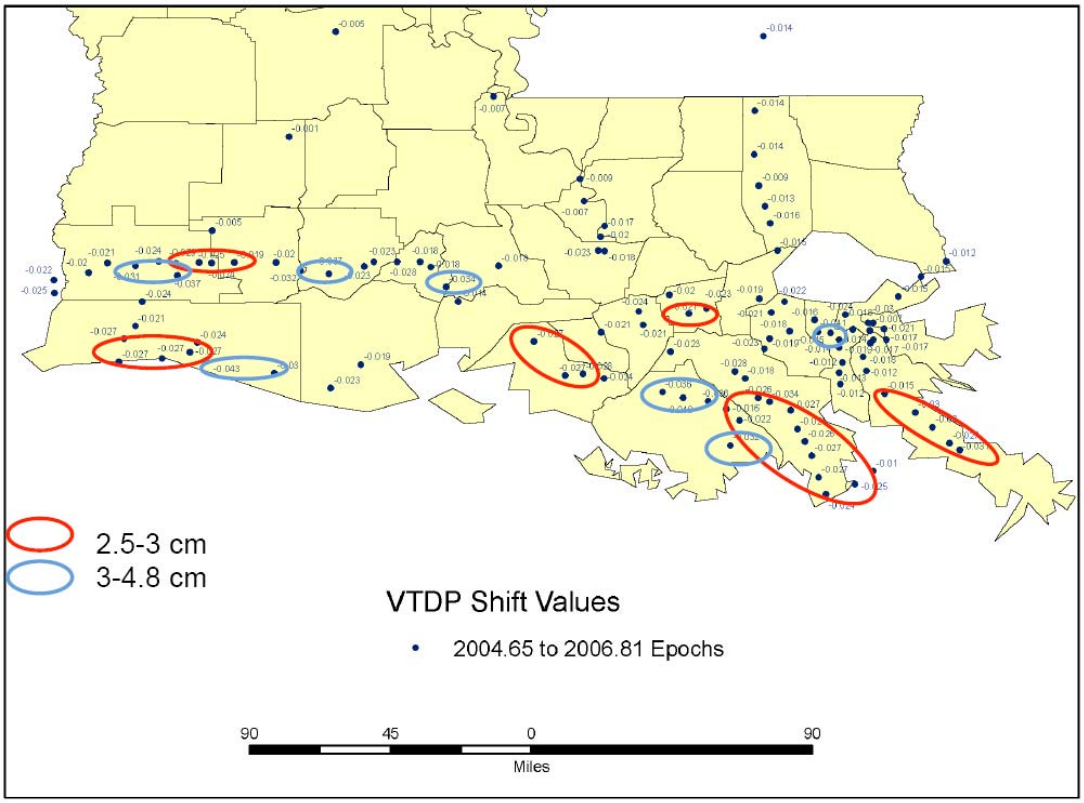


Figure 2. Differences between NAVD88-2006.81 versus NAVD88-2004.65
Shields 2011

B. Subsidence

Subsidence affects elevation measurements throughout coastal Louisiana—for land surfaces, water bottoms, roads, pile-supported structures (e.g., tide gauge structures), etc. Subsidence is influenced by numerous processes that impact different coastal regions in different ways (Reed and Yuill, 2009). These processes include:

- Tectonic subsidence
- Holocene sediment compaction
- Sediment loading
- Glacial isostatic adjustment
- Fluid withdrawal (oil and gas and groundwater)
- Surface water drainage and management

Local rates of subsidence must be considered in determining the potential for elevation data to become outdated. For example, at Grand Isle the local subsidence at the pile-supported tide station (apart from sea level rise) is on the order of 8 mm per year (Reed and Yuill 2009), or about 0.25 ft per decade. Some portions of southeast Louisiana experience higher subsidence rates. Thus, elevation data in many areas should be updated periodically to meet data quality requirements.

C. Digital Elevation Models

Hurricane surge hazard analysis makes extensive use of a regional digital elevation model (DEM), both in surge modeling and in mapping hazard areas. A DEM is a high resolution (dense pixel) representation of surface terrain and water bottoms. Figure 3 depicts a 5 meter grid (horizontal spacing) DEM for the topography of New Orleans. DEMs for surge modeling must often incorporate multiple sources of very large (terabyte) topographic and bathymetric datasets (see General Technical Note 5).

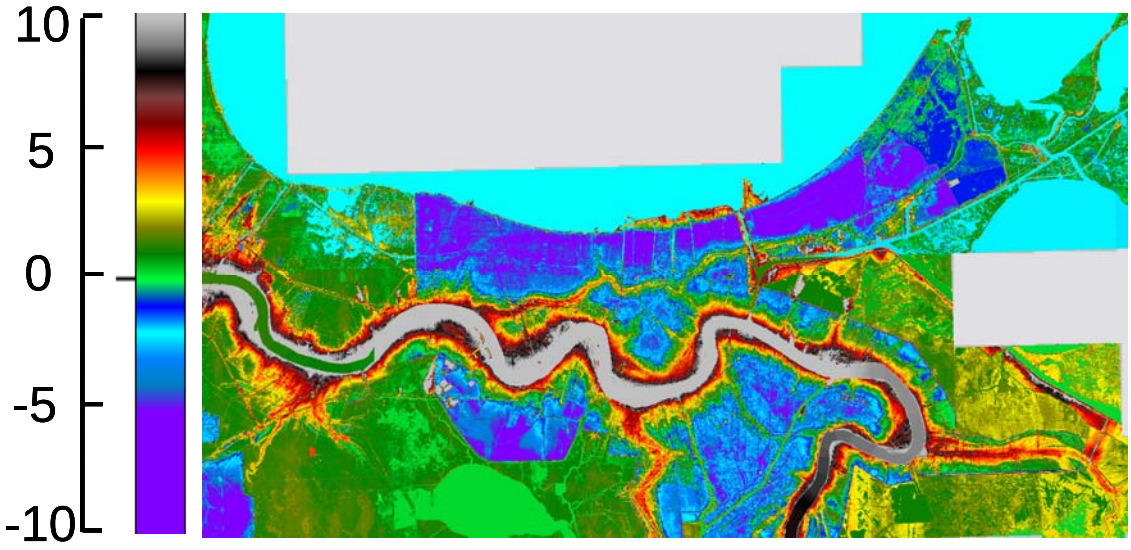


Figure 3. LIDAR DEM (ft) for New Orleans Area

Many sources of topographic data in coastal Louisiana, including numerous Light Detection and Ranging (LIDAR) data, were obtained in the 1990s and early 2000s, prior to the use of CORS corrected benchmarks and updated geoid models. Particular local and regional benchmark problems (e.g., sources of subsidence) can thus produce some significant general patterns of error in pre-CORS data. Figure 4 illustrates the discrepancies found between a 2006 CORS-based survey in Plaquemines Parish and previous pre-CORS LIDAR.

Louisiana bathymetric data sources for coastal bays, lakes, and sounds are often much older, frequently pre-dating World War II (e.g., Lake Pontchartrain). Many of these sources reference an approximate local tidal datum (e.g., local MLLW). Bathymetric data on numerous channels dredged for oil and gas activity are not even available. Some navigation channels may have initial or maintenance dredging plans but no accurate actual bank-to-bank bathymetric data.

Potential sources of data for a DEM can thus reflect widely varying accuracy (bias) and precision (uncertainty). Ultimately, all data used for a DEM should be referenced to NAVD88-2006.81 and meet objectives for accuracy and precision.[2]

[2] These objectives are set forth in project management plans consistent with particular goals of the program undertaking the hazard analysis and DEM development—e.g., a coastal Flood Insurance Study, a regional coastal planning and feasibility study, a flood control project design. DEM data quality objectives can be different under different programs.

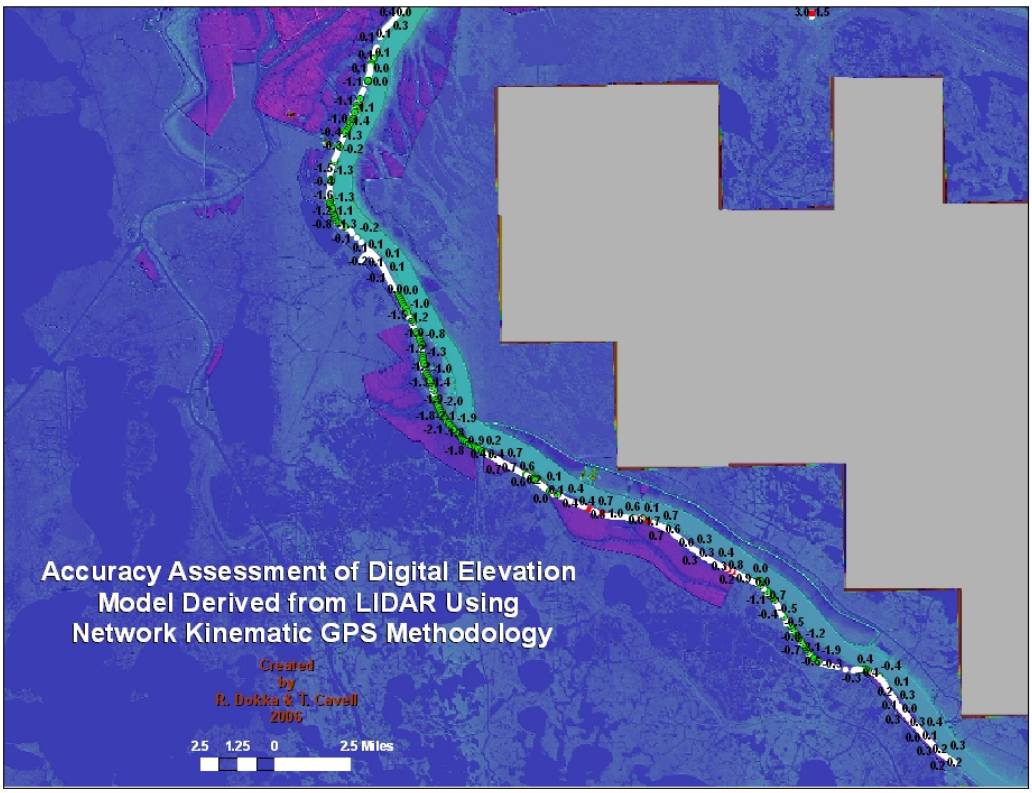


Figure 4. LIDAR DEM versus Updated Survey in Plaquemines Parish

Dokka and Cavell 2006

Green points denote (LIDAR – Current Survey) > 0.8 ft

References

Dokka, R. and T. Cavell, Accuracy Assessment of a portion of a Lidar-based Digital Elevation Model (DEM) of the Louisiana Coast—Southeast Louisiana, prepared for URS Corporation, September 21, 2006, under contract for the U.S. Army Corps of Engineers Louisiana Coastal Protection and Restoration Study.

National Oceanic and Atmospheric Administration, Tidal and Geodetic Datums: Tidal and Geodetic Relationships, http://www.laseagrant.org/pdfs/NOAA_TidalGeodeticRelationships.pdf

Reed, D. J., and B. Yuill, Understanding Subsidence in Coastal Louisiana, For the Louisiana Coastal Area Science and Technology Program, February 26, 2009, http://www.mvd.usace.army.mil/lcast/pdfs/UNO_SubsidenceinLA_09.pdf

Shields R., Vertical Geodetic Control in Southern Louisiana, in R. Dokka, T. Osborne, R. Shields, C. Mugnier, D. Zilkoski, G. Mader, and M. Schenewerk, National Static GPS/RTN Best Practices Seminar, June 20, 1011, Center for GeoInformatics, Louisiana State University, http://c4g.lsu.edu/c4g15/

Wormley, Samuel J., Orthometric Height, 2011, http://edu-observatory.org/gps/height.html

Zilkoski, D., Guidelines for Establishing GPS-Derived Ellipsoid and Orthometric Heights, in R. Dokka, T. Osborne, R. Shields, C. Mugnier, D. Zilkoski, G. Mader, and M. Schenewerk, National Static GPS/RTN Best Practices Seminar, June 20, 1011, Center for GeoInformatics, Louisiana State University, http://c4g.lsu.edu/c4g15/