Journal of Earth Science, February 2016
Printed in China
DOI: 10.1007/s12583-015-0641-9

ISSN 1674-487X

Statistics of Evolving Populations and Their Relevance to Flood Risk

Robert E. Criss*
Department of Earth and Planetary Sciences, Washington University, St. Louis MO 63130, *USA*

ABSTRACT: Statistical methods are commonly used to evaluate natural populations and environmental variables, yet these must recognize temporal trends in population character to be appropriate in an evolving world. New equations presented here define the statistical measures of aggregate historical populations affected by linear changes in population means and standard deviations. These can be used to extract the statistical character of present-day populations, needed to define modern variability and risk, from tables of historical data that are dominated by measurements made when conditions were different. As an example, many factors such as climate change and in-channel structures are causing flood levels to rise, so realistic estimation of future flood levels must take such secular changes into account. The new equations provide estimates of water levels for "100-year" floods in the USA Midwest that are 0.5 to 2 m higher than official calculations that routinely assume population stationarity. These equations also show that flood levels will continue to rise by several centimeters per year. This rate is nearly ten times faster than the rise of sea level, and thus represents one of the fastest and most damaging rates of change that is documented by robust data.
KEY WORDS: flood risk, statistical theory, Mississippi River.

0 INTRODUCTION

Considerable evidence shows that environmental conditions are changing and their variability increasing, with great and mostly negative impacts on human welfare. Common arguments are that storms are becoming more intense, droughts and floods more frequent, and rainfall less predictable, all as consequences of global warming (Karl et al., 2009). Such claims are oft disputed, as trend lines for many environmental variables are quite flat compared to annual variations. The resultant correlation coefficients are low and unconvincing to some, while good historical records are short and confused by complex feedbacks and oscillations that can span many decades. These difficulties underscore the need for robust theoretical tools that quantitatively incorporate temporal change into statistical analysis. Equations provided here facilitate estimation of present-day environmental variability from historical data, and are applicable to diverse scientific problems, risk assessment, resource management, and the protection of life and property.

Floods provide an important example, as they have afflicted humans since the dawn of civilization and remain a leading cause of weather-related death. Despite our best efforts to control them, the severity and frequency of floods are both increasing. Consequently, annual economic flood damages in the USA continue to rise and fatalities remain high (NWS, 2013). Factors contributing to these trends include climate change, progressive changes to watersheds that increase or accelerate runoff, flow impedance by in-channel navigation structures, and the isolation of rivers from their floodplains by levees (e.g., Jha et al., 2004; GAO, 1995; Belt, 1975; Funk and Robinson, 1974). Economic losses are rising because floods are becoming higher and more frequent, even as more costly infrastructure is constructed in floodplains (Pinter, 2005).

A key index of flood severity is the "100-year" flood. Levees are commonly designed to exceed "100-year" levels and "100-year" flood zones are delineated on detailed maps produced by FEMA (2013). These maps are then used to define insurance rates, construction standards, etc. Calculation of theoretical flood levels is a complex process that involves a statistical analysis of historical flood records that incorporates several assumptions (see Klemes, 2000). For example, river discharges are presumed to conform to a "Log Pearson Type III" distribution, and even more importantly, population stationarity is assumed (USGS, 1981). Note that official calculations clearly assume that the character of the flood population has not changed over time (USGS, 1981, p. 6), when available evidence suggests otherwise.

This article examines the statistical nature of populations constituted of time-series data, and applies the results to historical series of peak annual floods. Arguments are presented why river stages (water levels), not discharges, should be used in statistical calculations. The new equations are used to analyze flood populations whose means and standard deviations are changing over time. In particular, several methods are provided to extract the mean and standard deviation of present-day flood stages from the aggregate historical population, which

*Corresponding author: criss@wustl.edu

Manuscript received May 2, 2014.
Manuscript accepted August 27, 2014.

are needed to define realistic levels for the "10-year", "100-year", etc., floods that will occur in the future.

1 WHY A DIFFERENT APPROACH IS NEEDED

1.1 Unrealistic Predictions

It is easy to demonstrate that the official levels predicted for regulatory "100-year", etc., flood events are typically incorrect. For example, Criss and Winston (2008) used a standard chi-square test to show that recent floods experienced at Hannibal, Missouri occupied the far extreme (>99.5%) part of the "critical region", allowing one to easily reject the presumed correctness of the official flood risk calculations. Subsequent to that publication, floods at Hannibal exceeded the "10-year" flood levels in 2009, 2010 and 2011, and in 2013 experienced what is officially a "50-year" flood. Such outcomes are far too unlikely to be attributed to a nearly continuous succession of statistical flukes, and instead must be attributed to faulty calculation of flood risk. Criss and Kusky (2008) demonstrated similar problems with official flood probabilities at numerous other sites.

1.2 The Case for Stages

Official calculations emphasize discharges (flows) in flood-frequency analysis, yet many compelling reasons show that water levels (stages) should be used instead. (1) Stages are easily and accurately measured, while discharges are calculations. (2) Stages are easily understood and are, in fact, the most relevant quantity. If floodwaters are encroaching a home, the owner is concerned about the water level, not about the discharge of the river. (3) Discharge calculations depend on the measured stage as well as on channel configuration, water velocity that varies with position in the channel, changes in channel configuration due to scour or deposition, water temperature, etc. The resultant calculations have significant uncertainties. Even worse, few accurate determinations are available for the rare, high flows that are crucial to flood risk assessment, and so extrapolations of rating curves are commonly used to estimate discharges of great floods. (4) Because they are not simple measurements, there have been unfortunate attempts to "correct" historical discharge data reported in published catalogues (see Pinter, 2010). (5) Pairs of "100-year" discharges and stages reported for many sites are grossly incompatible with current rating curves for the site.

1.3 Evidence for Temporal Change

Historical records of peak annual floods are available for thousands of sites, and these provide robust evidence that flood populations are changing. The most rapid and straightforward way to demonstrate temporal change at a given site is to divide the time series data into an early half and a late (most recent) half, and compare the statistical properties of each, particularly the mean, standard deviation and skewness. If the record is relatively long and substantial differences appear between these well-constrained parameter pairs, the nature of the flood population is clearly changing with time. A subsequent section of this paper shows how such differences can be used to extract present-day flood probabilities from the historical record.

Another method of defining temporal dependence is to make a simple plot of peak annual stage vs. year, and perform a least squares analysis. The flood data will be highly scattered and the correlation coefficients unimpressively low, but positive slopes typically appear. This method cannot be used to extract changes in standard deviation or skewness as these parameters cannot be defined for single points. Making running averages of the historical data, over a 10-year interval, for example, can circumvent this problem. Such analysis typically reveals progressive increases in the mean annual peak stage as well as in the standard deviation, defining better-constrained trends whose upper ends provide another estimate for the statistical nature of the most recent flood population. Deviations of the running averages from the overall trend provide information on lengthy periods of heavy precipitation or drought.

In what follows, temporal changes in flood populations or other environmental conditions are assumed to be linear. This is the simplest assumption that recognizes that temporal changes exist. Rough linear trends for peak flood data have been reported previously for many sites (e.g., Criss and Shock, 2001).

1.4 Statistical Character of Flood Populations

No compelling theory explains why annual flood peaks should constitute a normal population, a log-normal population, a skewed population, or any other population. However, given a historical record of annual flood peaks for any site, standard probability plots are easily made of the stages, discharge, or the logs of discharge. This exercise typically reveals curvilinear trends for discharge, more linear trends for log discharge, and slightly stronger linear trends for stage. The following sections reveal how systematic temporal changes can be embodied within such historical records that can nevertheless retain the approximate character of a normal population.

2 THEORERETICAL RESULTS

2.1 Population Mean Changes with Time

For a normal population with mean (μ) and standard deviation (σ), the probability density function defines the well known "bell curve"

$$P_x = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-(x-\mu)^2/(2\sigma^2)} \tag{1}$$

where P_x is the probability of observing the value x, taken here to represent the peak annual stage. For such a function, the mean, median and mode (most probable value) are all coincident. In what follows, the population of interest is envisioned to comprise a historical table constituted of peak stage measurements reported once per calendar year over a period of observation, but the parameter of interest could represent practically any environmental observable.

Now suppose that the observable of interest, here flood levels, has a systematic tendency to increase linearly with time. In such a case, the mean will increase from value μ_i for the first year of the record at a rate at, where t is time in years, so

$$\mu_t = \mu_i + at \tag{2}$$

Note that μ_i is not equal to the earliest measurement, but rather represents the mean value that would have been recorded during the initial year if peak flood levels could have been measured in several hundred morphologically identical watersheds, that responded to conditions such as precipitation varying in a manner that was statistically appropriate for that period. In the next year (t=1), the mean value will differ slightly not only because of systematic climate trends, but because of systematic changes to the river channel and watershed; the combination of all such changes is incorporated in factor a. Now, what does the aggregate histogram of annual peak stages at a given site look like, after a large number of years has elapsed? The result can be found by integrating the evolving histograms over the time interval, giving

$$P_x = \frac{1}{2aT}\left\{ Erf\,\frac{x-\mu_i}{\sqrt{2}\sigma} - Erf\,\frac{x-\mu_i-aT}{\sqrt{2}\sigma} \right\} \tag{3}$$

where T is the total number of years of observation, and factor a is $\partial\mu/\partial t$. This is a symmetrical distribution, with an identical mean, median and mode at value

$$\mu=\mu_i+aT/2 \tag{4}$$

The effective standard deviation of this integrated distribution is no longer σ, which is the standard deviation for any given year, but for the aggregate population has a wider value

$$\sigma_{\text{all}} = \sqrt{\sigma^2 + (aT)^2/12} \tag{5}$$

For large values of aT, this distribution approximates a square wave, but for small values of aT it closely resembles a standard bell curve about mean μ with a standard deviation of σ_{all}. Note that the values for μ and σ_{all} represent the mean and standard deviation of the collective population of historical data, not the values for the most recent year (present day, abbrev. pd) in the distribution. The latter values are simply

$$\mu_{\text{pd}}=\mu_i+aT \text{ and } \sigma_{\text{pd}}=\sigma \tag{6}$$

2.2 Standard Deviation Changes with Time

Suppose that flood stages at any time have a normal probability of occurrence, and that the average value remains invariant but that the standard deviation of the distribution changes with time, such that $\partial\sigma/\partial t=b$. Suitable integration of Eq. 1 yields the following distribution for a historical series of data

$$P_x = \frac{1}{bT\sqrt{8\pi}}\left\{ E_i\,\frac{-(x-\mu)^2}{2\sigma_i^2} - E_i\,\frac{-(x-\mu)^2}{2(\sigma_i+bt)^2} \right\} \tag{7}$$

where Ei is the exponential integral. This is a symmetrical distribution with identical mean, median and mode all sharing the invariant value μ. The standard deviation of the overall population can be shown to be

$$\sigma_{\text{all}} = \sqrt{\left(\sigma_i + \frac{bT}{2}\right)^2 + \frac{b^2T^2}{12}} \tag{8}$$

which, for small values of b, is close to

$$\sigma_{\text{all}} \sim \sigma_i+bT/2 \tag{9}$$

Of course, the present-day value for the standard deviation is different than that of the aggregate historical population

$$\sigma_{\text{pd}}=\sigma_i+bT \tag{10}$$

while the mean, median and mode are identical and remain invariant.

2.3 Both Mean and Standard Deviation Change with Time

The author was unable to develop an analytical solution for the probability density function where flood peaks have a normal probability of occurrence in any given year, but both the mean and the standard deviation of the levels change linearly with time, such that $\partial\mu/\partial t=a$ and $\partial\sigma/\partial t=b$. However, key properties of the resultant distribution were secured. For the mean

$$\mu=\mu_i+aT/2 \tag{11}$$

and for the mode

$$\text{mode} = \mu_i + \frac{a\sigma T}{bT+2\sigma} \tag{12}$$

Clearly, Eq. 12 shows that the mean and mode are identical if either b, or a, or both are zero, corresponding to the symmetrical distributions in Eqs. 3, 7 or 1, respectively. However, if both b and a are finite, the mean and mode will differ so the aggregate distribution is skewed. Finally, the standard deviation for the aggregate historical population was found to be a combination of the separate effects identified above

$$\sigma_{\text{all}} = \sqrt{\left(\sigma_i + \frac{bT}{2}\right)^2 + \frac{T^2(a^2+b^2)}{12}} \tag{13}$$

This equation was successfully tested against extensive numerical simulations where the parameters were varied over large ranges. The above results suggest an exact, practical method of dealing with historical data sets where the means and standard deviations are both changing, see below.

3 PRACTICAL METHODS FOR REAL DATA SETS

3.1 General Remarks

The flood level expected to be exceeded within a given time interval L can be computed from (e.g., Chow, 1964)

$$\text{Stage}=\mu+K_L\sigma \tag{14}$$

where μ and σ are the mean and standard deviation of peak flood stages at the particular site, and the K factor depends on the interval length L and the nature of the governing probability distribution. For a normal distribution, the K factors can be found in tables, or directly computed using

$$L = 2/Erfc(K/\sqrt{2}) \tag{15}$$

where $Erfc$ is the complimentary error function. For example, for a "100-year" flood and assuming a normal distribution, L is 100 and K can be computed from Eq. 15 as 2.326 348 ... Extensive tables of factors for skewed distributions that conform to Pearson type III characteristics are provided by USGS (1981).

If environmental conditions are static, as assumed in available calculations (USGS, 1981), evaluating flood risk from a set of historical data is straightforward. The mean, standard deviation and skewness of the historical series of annual peak floods at any site are readily calculated. Appropriate K

factors can be calculated for the probabilistic distribution that is assumed to describe flooding at that site. Finally, given μ, σ and K, the levels expected for a "100-year" flood, or for a flood expected for any other recurrence interval, can be easily computed using Eq. 14.

The problem is less straightforward if environmental conditions are changing, so that μ and σ are functions of time. This is because μ and σ defined by the aggregate, historical flood population do not represent the values today. The problem reduces to extracting the present-day values μ and σ from historical records that are dominated by measurements made in prior times when conditions were different. Three practical methods for accomplishing this are provided below.

3.2 Simple Method

If the mean and/or standard deviation of the annual flood peaks are changing over time, the values μ and σ calculated from the early half of the historical data series will differ significantly from the values μ_2 and σ_2 calculated for the most recent half. Those values are easily extracted from a table of historical data, and assuming linear change, the values for $\partial\mu/\partial t$ and $\partial\sigma/\partial t$ are

$$\frac{\partial\mu}{\partial t}=\frac{2(\mu_2-\mu_1)}{L} \text{ and } \frac{\partial\sigma}{\partial t}=\frac{2(\sigma_2-\sigma_1)}{L} \tag{16}$$

where L is an integer representing the length of the historical record in years. The present-day values are

$$\mu_{\text{pd}}=\mu_2\left(\frac{3L-2}{2L}\right)-\mu_1\left(\frac{L-2}{2L}\right) \text{ and}$$

$$\sigma_{\text{pd}}=\sigma_2\left(\frac{3L-2}{2L}\right)-\sigma_1\left(\frac{L-2}{2L}\right) \tag{17}$$

Analogous relations obtain for skewness, if needed. Note that, for long historical records, the values in parentheses are very close to the simple fractions 3/2 and 1/2. Once the present-day values are determined, they can be used in equation 14 to calculate realistic levels for "100-year", etc., floods at the particular site at the present time.

A minor problem arises for data sets that encompass an odd number of years. Results presented below utilize an algorithm that assigns one-half the statistical weight of the central datum to the early half of the data, and one-half to the late set. Several simpler methods, such as averaging the results where the central datum is first assigned to the early set and then to the late, are evident, and while these are less accurate, they provide similar estimates for present-day values if the period of observation is long.

3.3 Running Average Method

Given a historical record, a linear regression on a plot of annual flood peaks vs. year of observation can be used to estimate $\partial\mu/\partial t$ and μ_{pd}. While useful, this approach has several disadvantages in flood-frequency analysis. (1) Because the data are highly scattered, correlation coefficients are low. (2) Least squares analysis places great emphasis on anomalous data points, as it minimizes the sum of the squares of the deviations about the regression line. (3) No analogous procedure can provide the requisite estimates of $\partial\sigma/\partial t$ and σ_{pd}, because the unbiased standard deviation is not defined for single points.

The above defects can be partly overcome by use of n-way running averages of the mean and standard deviation of flood data. To effect this, the table of flood data should be inverted so that the most recent data are on top. For any choice of n, the n-way running averages can be computed down the table. Finally, ordinary least squares regressions of the smoothed data provide estimates of $\partial\mu/\partial t$, $\partial\sigma/\partial t$, μ_{pd}, and σ_{pd}. In detail, these regressions should be made by plotting the running average against the central year of the running interval. If this is not done, the slopes will be correct, but the present day values calculated from the regression will be systematically off.

In practice, results returned by the running average method depend somewhat on the arbitrary interval n. The correlation coefficients of the regression lines increase sharply as n increases, but that does not prove that the requisite estimates are better. Present-day values for the mean (μ_{pd}) tend to be similar over a wide range of n, but the calculated slopes $\partial\mu/\partial t$ vary by up to ~3% percent of their value as n ranges from 2 to 30 years. Present-day values for the standard deviation (σ_{pd}) vary from about -15% to +33% of the value for the entire data set, depending on the interval chosen for n. For a data set of appreciable length, a choice for n of 10-years is suggested as providing reasonable smoothing and intermediate estimates for slopes and present-day values.

3.4 Theoretical Method

The above methods provide similar estimates for μ_{pd}, but a wider range of values for factor σ_{pd}. Fortunately, Eq. 13 can be used to extract a more precise estimate for σ_{pd} from a historical data set, which is first used to compute the standard deviations of the entire population (σ_{all}), the first half of the data (σ_1), and the second (most recent) half of the data (σ_2). The variance of the present day population can then be calculated from those three values using the following equation, derived from Eqs. 13 and 16

$$\sigma_{\text{pd}}^2=\frac{A-\sqrt{A^2-B}}{6} \tag{18a}$$

where

$$A=14\sigma_2^2-4\sigma_{\text{all}}^2+2\sigma_1^2 \text{ and } B=12\left(4\sigma_2^2-\sigma_{\text{all}}^2\right)^2 \tag{18b}$$

3.5 Annual Change of Flood Stages

Chow's equation (Eq. 14) can be differentiated to define the annual increase in flood stages expected at any site. Thus

$$\partial\text{Stage}/\partial t=\partial\mu/\partial t+K_L\,\partial\sigma/\partial t+\sigma\partial K_L/\partial t \tag{19}$$

The first term on the right is factor a, while the second (middle) term is bK_L. The term on the far right is small, and is zero if the skewness is not changing.

4 EXAMPLES

4.1 Mississippi River at St. Louis

Measurements of the stage of the Mississippi River at St. Louis have been made almost every day since 1861, and are available at USACE (2013a). The highest stage recorded every year can be easily extracted from these data, and the results processed by the above methods to define expected flood levels.

Examination of this historical data set (1861 to 2013) provides a simple mean of 122.43 m, a standard deviation of 2.023 m, and a skewness of 0.198 9 (Table 1). If all these data are assumed to be part of a static normal distribution, and their aggregate mean and standard deviation are viewed as being appropriate for the "present-day", then Chow's equation indicates that the "100-year" flood level is only 129.1 m. If the skewness is taken into account using the USGS (1981) tables, the "100-year" level would be 129.4 m. These theoretical stages are somewhat lower than the official "100-year" levels calculated by USACE and FEMA (see below); reasons are complex but include the use of discharge determinations, rather than stages, in making the official estimates.

However, all of the above calculations grossly underestimate the current "100-year" flood level, because the historical population is not homogeneous, but instead its statistical character has changed over time. A simple plot of the annual peak stages vs. the year (Yr) of observation provides the following regression line for St. Louis

$$\text{Stage}=100.44+0.012\,38\cdot(\text{Yr}),\quad R=0.271 \tag{20}$$

The correlation coefficient is low, but the trend is distinct and shows that annual peak stages have increased by 1.9 m over the period of observation. Use of a 10-year running average for stage provides a similar slope, but the same method additionally permits regression lines to be constructed for the standard deviation and skewness, from which their present-day (2013) values can be estimated (Fig. 1, Table 1). Correlation coefficients are much improved because of the smoothing that running averages provide. These various methods indicate that the present-day mean of the annual peak stage is about 1 m higher than the value suggested by the simple mean of the entire population; this level corresponds to that expected for a "2-year" flood because K_2=0. Indicated values for the "100-

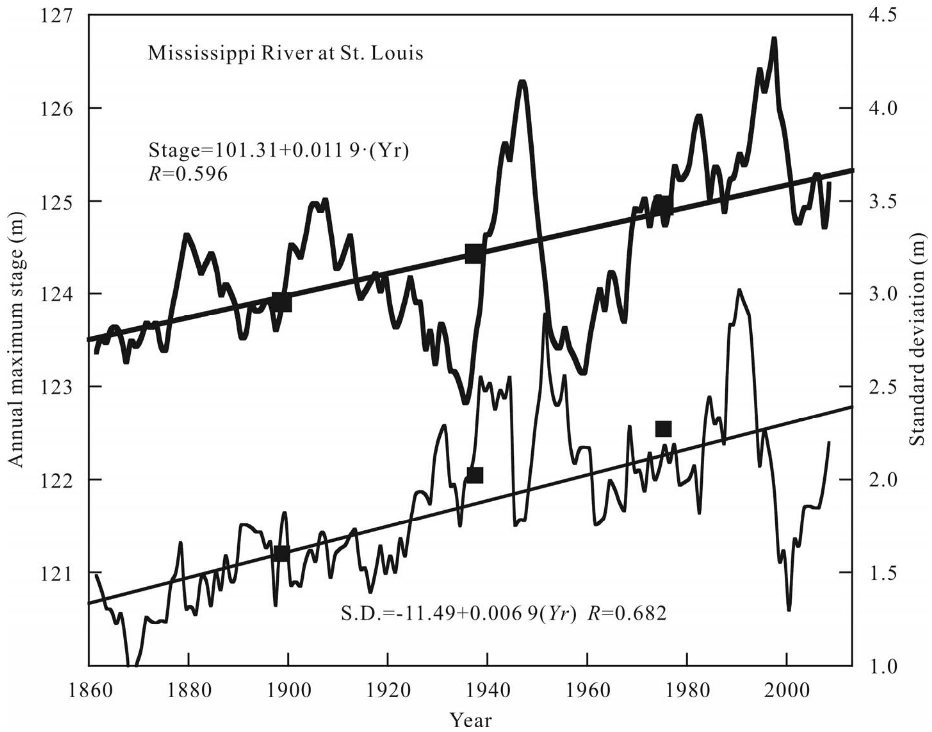

Figure 1. Ten-year running averages of the annual maximum stage (m rel. MSL) and the associated standard deviation, plotted vs. the mid-year of the running interval. Linear regressions are shown. Squares plotted at 1937 indicate the mean and standard deviation of the entire (1861–2013) population, while the flanking squares indicate the respective values for the first and second halves of the data.

Table 1 Mississippi River at St. Louis, 1861–2013

Parameter	Normal distribution	Linear regression	10-Yr run	Simple method	Eqs. 18a, 18b
a (m/y)	0	0.012 4	0.011 9	0.013 4	
b (m/y)	0		0.006 9	0.008 8	
μ_{pd} (ft)	124.43	125.4	125.3	125.5	
σ_{pd} (ft)	2.02		2.39	2.60	2.56
∂Skewness/∂t	0		0.000 9	-0.001 78	
Skewness$_{pd}$	0.199		0.056	-0.146 7	
"100-yr" stage*	129.1	>130.1	130.9	131.5	131.4

*Elevation in m rel. MSL; skewness neglected.

year” flood levels are about 5 m higher than this (Table 1).

Calculated values are even higher if the simple method, or the theoretical equation, are used, mostly because the present-day estimates for the standard deviation are also greater than that defined by the entire population. Thus, for a “100 year’ flood, the present day (2013) stage would be about 131.5 m; the values would be ~0.1 m higher if the small, present-day skewness is taken into account. For comparison, the official “100-year” flood levels at St. Louis are estimated to be 129.8 m by USACE (2004) and 129.9 m by FEMA (2011). In other words, the official calculations for the “100-year” flood level at St. Louis are ~1.7 m too low, primarily because they neglect both the tendency for the flood levels to increase over time and the concomitant increase in the standard deviation of the flood population.

4.2 Additional Examples

Long historical data sets are available for river stages at numerous sites in the United States (USACE 2013a, b; Jarvis, 1936). In practically all cases examined, these data show significant positive slopes for parameters *a* and *b*, indicating that annual peak stages have increased by several meters over the last century, while the associated standard deviations have also significantly increased (Fig. 2; Table 2). These ubiquitous trends require reevaluation of the stages expected for the “10-year”, “50-year”, “100-year” etc., floods generated under

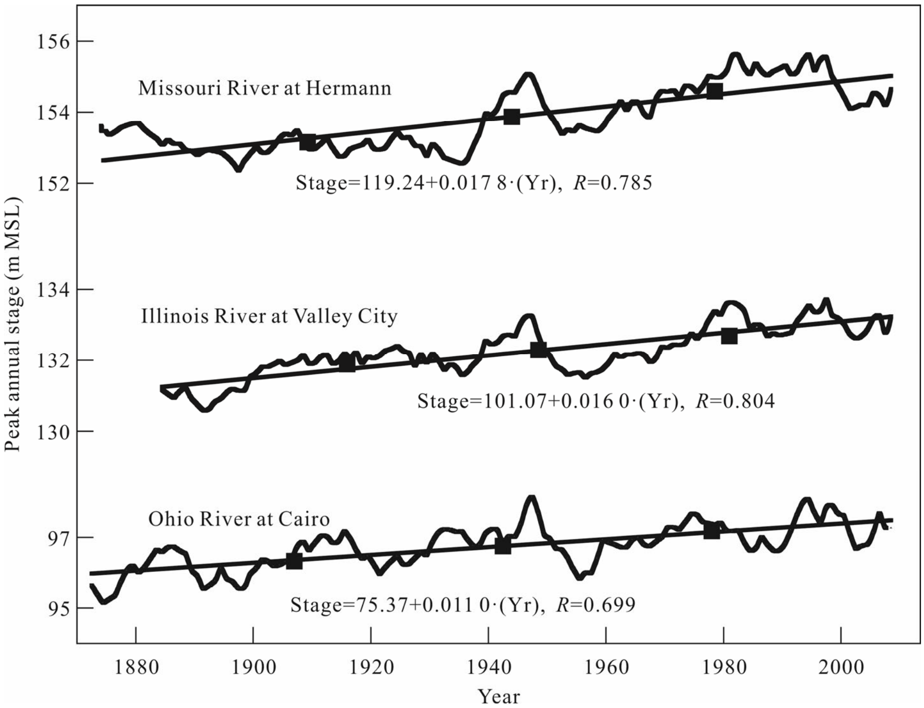

Figure 2. Ten-year running averages and linear regressions of the annual maximum stage (m rel. MSL) for sites along the lower Missouri, Illinois and Ohio rivers. Center squares indicate the means of each entire population, while the flanking squares indicate the respective means for the first and second halves of the data; the squares are collinear and were used to define factor *a* in Table 1. Mean annual flood stages increased ~2 m at all three sites over the period of record.

Table 2 Parameters and estimates of 100-yr flood stages for various sites

Site	Data since	*a* (m/y)	*b* (m/y)	μ_{pd} (m)	σ_{pd} (m)	100-yr stage (m)*	Official 100-yr stage (m)	Δ (m)
Miss. R. at Burlington	1917[2]	0.019 9	0.003 9	161.33	1.059	163.8	162.8[5]	1.0
Miss. R. at Hannibal	1879[2]	0.017 6	0.003 1	143.44	1.446	146.8	145.4[5]	1.4
Miss. R. at St. Louis	1861[1]	0.013 4	0.008 8	125.45	2.600	131.5	129.8[5]	1.7
Miss. R. at Chester	1892[1]	0.020 5	0.006 5	113.99	2.449	119.7	118.6[5]	1.1
Missouri R. at Omaha	1872[3, 4]	0.022 3	0.013 5	296.26	2.194	301.4	299.5[5]	1.9
Missouri R. at Hermann	1874[1,4]	0.020 8	0.005 3	155.31	1.795	159.5	158.1[5]	1.4
Illinois R. at Henry	1869[2]	0.012 6	0.000 6	137.85	1.159	140.5	140.6[5]	-0.1
Illinois R. at Havana	1896[2]	0.011 0	0.003 4	135.58	1.267	138.5	138.1[5]	0.4
Illinois R. at Valley City	1884[1]	0.012 4	0.001 4	133.09	1.495	136.6	135.8[5]	0.8
Meramec R. at Eureka	1922[3]	0.019 4	0.002 4	130.90	2.457	136.6	135.9[6]	0.7
Ohio R. at Cairo	1872[1]	0.012 0	-0.000 7	97.60	1.690	101.5	100.9[7]	0.6

*Water elevation in m rel. MSL from Eqs. 14 and 17, skewness neglected. Data sources: [1]. USACE, 2013a; [2]. USACE, 2013b; [3]. USGS, 2013; [4]. Jarvis, 1936; [5]. USACE, 2013c, 2004; [6]. FEMA, 2006; [7]. FEMA, 2009.

present conditions. Table 2 provides "100-year" stage estimates made by Eqs. 14 and 17; skewness was neglected but is positive for 7 of the 11 sites examined, so considering it would mostly increase the tabulated values. Comparison of these values with official estimates made by USACE or FEMA, which were found to be within 0.2 m of each other, suggests that the official levels reported along the upper Mississippi and lower Missouri rivers are 1.0 to 1.9 m too low. This difference explains the current overuse of terms such as "200-year" floods, which arises from the incorrect assumption that flood populations have remained static.

Finally, Eq. 19 can be used to define the annual increase in L-year flood stages expected at any site. Because both a and b are generally positive (Table 1), the available data suggest that flood levels will continue to increase by several centimeters per year (Eq. 13; Table 2), a damaging rise that is ~10× times faster than that of sea level.

5 CONCLUSIONS

Historical data sets on major midwestern rivers in the USA show that flood levels are increasing, and that standard deviations of flood levels have concomitantly increased. These effects must be incorporated in calculations of expected flood levels to accurately define present day risk. Mathematical integration of histograms with temporally-varying means and standard deviations can be used to deduce the characteristics of composite histograms constituted of historical data. New equations presented here allow the present day means and standard deviations of dynamic populations to be extracted from their aggregate historical data. When applied to historical flood records, these equations show that the stages officially estimated for "100-year" flood levels at many important USA sites are too low, generally by 1 to 2 m. Moreover, these levels will continue to rise by several cm/y, a damaging trend that dwarfs the rate of rise of sealevel.

REFERENCES CITED

Belt, C. B., 1975. The 1973 Flood and Man's Constriction of the Mississippi River. *Science*, 189: 681–684

Chow, V. T., 1964. Handbook of Applied Hydrology. McGraw-Hill, New York. 8: 23

Criss, R. E., Shock, E. L., 2001. Flood Enhancement through Flood Control. *Geology*, 29: 875–878

Criss, R. E., Winston, W. E., 2008. Public Safety and Faulty Flood Statistics. *Environmental Health Perspectives*, 116(12): A516. doi:10.1289/ehp.12042

Criss, R. E., Kusky, T. M., 2008. Finding the Balance between Floods, Flood Protection, and River Navigation. http://ces.slu.edu/annualreport/FloodForum_Book_final.pdf(2008)

FEMA, 2011. Flood Insurance Rate Map, City of St. Louis, Missouri. Map Number 2903850068C, May 24, 2011

FEMA, 2006. Flood Insurance Rate Map, Jefferson County, Missouri. Map Number 29099C0025E, April 5, 2006

FEMA, 2009. Flood Insurance Rate Map, Alexander County, Illinois. Map Number 17003C0220E, May 4, 2009

FEMA, 2013. Flood Insurance Rate Maps, Various Sites and Years. https://msc.fema.gov/webapp/wcs/stores/CategoryDisplay?catalogId=10001&storeId=servlet/10001&categoryId=12001&langId=-1&userType=G&type=1

Funk, J. L., Robinson, J. W., 1974. Changes in the Channel of the Lower Missouri River and Effects on Fish and Wildlife. *Missouri Department of Conservation*, *Aquatic Series*, 11: 52

GAO, 1995. Midwest Flood, Information of the Performance, Effects, and Control of Levees, GAO/RCED-95-125

Jarvis, C. S., 1936. Floods in the United States: Magnitude and Frequency. *U.S. Geological Survey Water Supply Paper*, 771: 1–497

Jha, M., Pan, Z., Takle, E. S., et al., 2004. Impacts of Climate Change on Streamflow in the Upper Mississippi River Basin: A Regional Climate Model Perspective. *Journal of Geophysical Research*, 109: D09105. doi:10.1029/2003JD003686

Karl, T. R., Melillo, J. M., Peterson, T. C., 2009. Global Climate Change Impacts in the United States. Cambridge University Press, New York. 188

Klemes, V., 2000. Tall Tales about Tails of Hydrological Distributions. *Journal of Hydrologic Engineering*, 5(3): 227–231

NWS, 2013. Flood Loss Data. http://www.nws.noaa.gov/hic/ [accessed Oct. 2013]

Pinter, N., 2005. One Step Forward, Two Steps Back on U.S. Floodplains. *Science*, 308: 207–208

Pinter, N., 2010. Historical Discharge Measurements on the Middle Mississippi River, USA: No Basis for "Changing History". *Hydrological Processes*, 24: 1088–1093

USACE, 2004. Upper Mississippi River System Flow Frequency Study: Final Report. http://www.mvr.usace.army.mil/pdw/pdf/FlowFrequency/flowfreq.htm [accessed Oct. 2013]

USACE, 2013a. Mississippi River Basin-Historic Data. http://mvs-wc.mvs.usace.army.mil/archive/mi.html [accessed Oct. 2013]

USACE, 2013b. Water Levels of Rivers and Lakes. http://rivergages.mvr.usace.army.mil/WaterControl/new/layout.cfm [accessed Oct. 2013]

USACE, 2013c. Flow Frequency Query: Upper Mississippi River. Available: http://rivergages.mvr.usace.army.mil/flow_freq/flow_freq.cfm [accessed Oct. 2013]

USGS, 1981. Guidelines for Determining Flood Flow Frequency. Interagency Advisory Committee on Water Data, Bulletin #17B of the Hydrology Subcommittee. U.S. Department of Geological Survey, Office of Water Data Coordination Reston, Virginia

USGS, 2013. Peak Streamflow for the Nation. http://nwis.waterdata.usgs.gov/usa/nwis/peak [accessed Oct. 2013]