Nat Hazards (2017) 88:559–574
DOI 10.1007/s11069-017-2880-9

ORIGINAL PAPER

Base flood estimates compared and linked to engineering modifications of the Missouri River

Mingming Luo[1,2] · **Robert E. Criss**[2]

Received: 4 January 2017 / Accepted: 11 April 2017 / Published online: 18 April 2017

Abstract A novel stage projection method is used to estimate present-day flood levels at 12 sites on the Missouri River, using present-day rating curves and historical discharge estimates. These results are compared to several other flood estimates, including methods based only on historical stage data, and official "100-year" flood (base flood) levels determined from the statistics of discharge. Differences among these estimates vary with location, and their utility depends on river management style. At sites in the upper basin, channel configuration has changed little, but peak discharges have decreased slightly, due to tributary reservoirs and withdrawals. Little difference is seen between the various estimates of flood levels, and historical changes appear to be minimal. In contrast, flow behavior and channel character have been drastically modified along the middle Missouri River by a system of dams and huge reservoirs that were constructed and filled between 1933 and 1964; estimates of flood levels depend on location relative to these facilities. Further downstream, the lower Missouri River is lined with levees and has been transformed into a narrow navigational channel. Results are complex at and above Kansas City, because the channel at many sites has become incised due to decreased sediment loads. Below Kansas City, the water levels of significant floods are now higher than historical values, and official base flood levels are underestimated. Profound changes to the Missouri River have destabilized it in many complex ways, causing it to be less predictable, and many decades or centuries will be required for a new state of equilibrium to develop.

Keywords Flood level · Stage change · Flood control · Missouri River

---

✉ Mingming Luo
luomingming@cug.edu.cn

✉ Robert E. Criss
criss@wustl.edu

[1] School of Environmental Studies, China University of Geosciences, Wuhan 430074, Hubei, China

[2] Department of Earth and Planetary Sciences, Washington University in St. Louis, St. Louis, MO 63130, USA

1 Introduction

The Missouri River is the longest in North America, flowing 3767 km before entering the Mississippi River at St. Louis, Missouri (Fig. 1). Headwaters originate near the continental divide and combine to form the "upper Missouri River" which retains much of its natural character. Further downstream, the "middle Missouri River" is controlled by six main stem dams and reservoirs, as well as by hundreds of smaller reservoirs on tributaries, that together have drastically changed the flow regime. Finally, the "lower Missouri River," representing the reach from Sioux City to St. Louis, has been converted to a navigational channel that is narrowed twofold by thousands of wing dikes. Management of the system to accommodate diverse demands for power generation, navigation, water supply, wildlife protection, and flood control has become increasingly difficult.

A succession of recent record floods demands scientific explanation and has brought both river management and flood risk calculations into national focus. Belt (1975) showed that flood stages for a given discharge have become higher on the middle Mississippi River because of river constriction by levees and wing dikes. Similar effects were documented on the lower Missouri River, which is managed similarly (Criss and Shock 2001). Other reports (e.g., GAO 1995; USACE 2012; Collenteur et al. 2015) show that levees have significant effects and likewise show that river stages have become higher for given high discharges at many sites. Ehlmann and Criss (2006) showed that the day-to-day variability

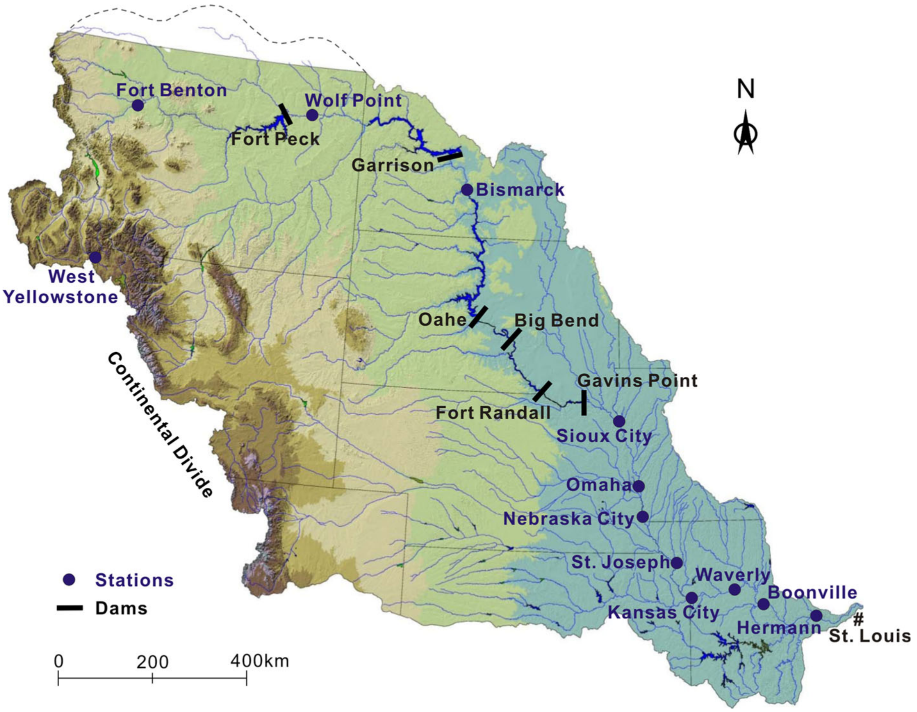

Fig. 1 Locations of selected gauging stations (*dots*), major dams (*bars*), and main stem reservoirs in the Missouri River basin. The river extends from the continental divide to ultimately join the Mississippi River near St. Louis. The *shading* on the base map shows the generalized topography of the basin, after USACE (2012); note that the basin extends into southern Canada (*dashed line*, *no shading*)

of river stages has increased twofold along the lower Missouri River and attributed this twofold increase to the twofold decrease in channel width. Numerous studies, most recently Criss and Luo (2017), have argued that flood levels are greatly underestimated in many areas.

This paper uses long historical records to examine the evolving nature of peak annual floods along different major reaches of the Missouri River that have been modified in different ways. We present and compare several different estimates of flood risk for different recurrence intervals, including a novel approach, and discuss the limitations and advantages of these different models.

2 Data and methods

2.1 Data sources

Data analyzed in this study represent 11 stations on the main stem of the Missouri River, plus one on a headwaters tributary near the continental divide (Fig. 1). These sites were selected to maximize geographic and temporal coverage and have data on both annual peak stage and discharge back to the 1920s (USGS 2016). Moreover, eight of these sites, notably Hermann, Boonville, Kansas City, St. Joseph, Omaha, Sioux City, Bismarck, and Fort Benton, have data on peak annual stages that extend back into the 1870s or 1880s (Jarvis 1936; NOAA 2016). Finally, we also examine long, >100 year records for daily stage that are available for Bismarck and Hermann, which represent the middle and lower Missouri River.

2.2 Flood level estimation

Statistical methods are essential to the evaluation of flood risk. These methods necessarily presume that the quantity of interest is a random variable, and that the data conform to some particular distribution curve. The normal distribution is widely used, but skewed, log normal, or other distributions have also been used.

In traditional statistical analysis, the likelihood that the value of an appropriate quantity "S", for example river stage, will exceed some particular magnitude can be calculated using:

$$S = \mu + K_L \sigma \tag{1}$$

where μ is the mean and σ is the standard deviation of the population of interest (Chow 1964). Values of K_L are tabulated for different distributions by USGS (1981); for a normal distribution, K_L is given by:

$$\mathrm{Erfc}\left(K_L/\sqrt{2}\right) = 2/L \tag{2}$$

where Erfc is the complimentary error function, and L is the expected recurrence interval. For example, for a 100-year flood (base flood), which is actually a flood with a 1% probability of occurrence in any given year, L is 100, so K_{100} can be calculated as 2.32635….

Estimates of flood risk are made using tables of historical data. Traditional analyses use tables of peak annual water levels at a site of interest. The means and standard deviations of the flood population are easily found, so it would appear that Eq. 1 can be immediately used to estimate the water levels expected for any recurrence interval of interest. However, different assumptions are possible, giving rise to different methods that yield different estimates.

2.2.1 Historical stage method

In the simplest case, where environmental conditions are presumed to be static, the flood population is homogeneous and ordinarily presumed to fit a normal distribution. Such calculations are provided below that utilize peak annual river stage as the statistical variable and assume that the normal distribution applies so that the K_L values are given by Eq. 2. This is referred to as the historical stage method (Table 1). This elementary method is independent of discharge estimates and is also amenable to analysis using a standard probability plot.

2.2.2 Criss (2016) method

A different method proposed by Criss (2016) also uses stage only, but addresses the possibility that linear changes may have affected the flood population. Basically, this method uses historical tables of peak annual river stages to estimate the "present-day" means and standard deviations to be used in Eq. 1; these values may be higher, lower, or the same as those provided by ordinary statistical analysis of the data. Estimates of the "present-day" values are provided by dividing the historical population into an early and a late half, and determining the mean and standard deviation of each; if the environmental conditions have evolved in a manner that has systematically affected the flood population, this will be reflected in a difference between those respective means and standard deviations of the two subpopulations. Presuming that any changes are gradual and linear, Eqs. 17a and 18a of Criss (2016) can be used to estimate the "present-day" means and standard deviations for the flood population, and then, these values can be inserted into Eq. 1 (above) to estimate the present-day flood risk.

2.2.3 Stage projection method

In contrast to the above, most recent studies of flood risk use river discharge as the primary statistical variable (e.g., USACE 2004, 2016; Salas and Obeysekera 2014; Benameur et al. 2017). These methodologies are considered here to be less than optimal. In contrast to stage, discharge is not a simple measurement, and values are much less accurate, especially for times long past. In addition, historical discharge records are much shorter and available for far fewer sites than stage records; additional problems are discussed by Criss (2016) and Criss and Luo (2017). Nevertheless, for comparison to discharge-based methods, we provide a new "stage projection" method that combines historical discharge "data" with modern rating curves.

In particular, the "stage projection" method accepts the discharge estimates made for prior "peak annual" floods as given and then uses the present-day rating curves available from USGS (2016) to estimate the stages that those historical floods would have attained, had they occurred today. The means and standard deviations of this projected stage population can readily be determined; then, Eq. 1 uses these quantities to estimate modern flood risk.

Table 1 Estimated base flood levels and associated statistical data for sites on the Missouri River, in meters relative to the local gauge datum, for the interval ~1929 to present

Stations	Data since[a]	River distance[b]	Historical stage			Projected stage			Criss method[c]			USACE (2004)		
		km	μ	σ	100-year flood	μ	σ	100-year flood	μ_{pd}	σ_{pd}	100-year flood	μ^d	σ^d	100-year flood[e]
West Yellowstone[f]	1929		0.87	0.11	1.12	0.85	0.10	1.08	0.92	0.13	1.23			
Fort Benton	1924	3336.4	1.99	0.68	3.57	2.01	0.68	3.59	1.73	0.53	2.96			
Wolf Point	1929	2738.1	5.34	0.95	7.55	4.73	0.77	6.52	4.18	0.73	5.89			5.01
Bismarck	1929	2115.4	3.43	1.27	6.38	3.28	1.55	6.88	2.97	0.47	4.05			5.71
Sioux City	1929	1178.3	7.82	1.80	12.01	6.40	2.20	11.52	5.50	0.94	7.68	5.81	1.45	9.45
Omaha	1929	991.3	7.31	1.24	10.20	8.12	1.49	11.59	7.02	0.99	9.31	7.00	1.54	10.47
Nebraska City	1930	905.6	5.28	1.15	7.96	6.07	1.15	8.74	5.94	1.21	8.74	5.51	1.39	8.82
St. Joseph	1922	721.3	5.79	1.21	8.60	6.58	1.51	10.09	6.79	1.41	10.07	5.83	1.54	9.51
Kansas City	1929	589.2	8.78	1.80	12.98	8.05	2.19	13.16	8.38	2.07	13.19	7.48	2.45	13.14
Waverly	1929	472.2	6.72	1.42	10.02	7.15	1.30	10.18	7.95	1.28	10.92	7.24	1.02	9.60
Boonville	1926	317.2	7.23	1.70	11.20	7.28	1.69	11.22	8.54	1.43	11.87	6.59	1.81	11.12
Hermann	1929	157.6	7.68	1.73	11.71	7.97	1.87	12.32	8.80	1.47	12.22	7.31	1.86	11.29

[a] Beginning of interval when both discharge and stage data are available from USGS (2016)

[b] River distance indicates kilometers above the mouth of the Missouri River

[c] Method of Criss (2016), where μ_{pd} and σ_{pd} are the estimated, present-day means and standard deviations

[d] Estimated from Fig. 6, with the μ values also including the level of the local NWS flood stage for each site

[e] Base flood levels reported by USACE (2004), except at Wolf Point (FEMA map 3001580175A, 2007) and Bismarck (FEMA map 38059C0515D, 2005)

[f] Site on the Madison River, a tributary of the upper Missouri River

Note that all the calculations are in meters

Alternatively, a standard probability plot can be made of the set of projected stages, and flood levels for different recurrence intervals directly read off. Linearity of the historical stage data or of the stages projected to present-day conditions shows conformity to a simple normal distribution. This computational method has certain affinities to the "stage indexing" procedure (e.g., Pinter 2001), but is more accurate and much more direct, because there is no need to gather and process extensive sets of stage–discharge pairs to generate "specific gauge" curves for each of numerous discharges. Moreover, too few data are normally available to calculate "specific gauge" curves for the large, but uncommon, floods that are of greatest interest. Most importantly, the stage projection and the stage indexing method only reveal systematic changes, such as modifications to the river channel, that have modified the rating curve. Thus, flat "specific gauge" trend lines do not prove that the flood population and the associated risk have not changed, as shown below.

2.2.4 *Official estimates*

Official estimates of flood risk are based on complex methodologies that are fundamentally based on historical estimates of peak annual discharge and presume that those values conform to a "log Pearson Type 3" statistical distribution (USGS 1981). Then, different methodologies including complex computer simulations of the longitudinal profiles of flowing water surfaces for different flows are used to estimate flood levels at different sites (e.g., USACE 2004). Such calculations are fundamentally not reproducible using the information provided, but values for comparison are readily available from USACE (2016), and/or from flood insurance rate maps for different localities (FEMA 2016).

2.3 Daily change calculations

The daily change in water level provides an important measure of river behavior and historical change that can be very accurately defined (Ehlmann and Criss 2006). The daily change of river stage is defined as the absolute value of the difference between the river stage for any given day and the previous day. These short-term changes are inherently quite variable. Accordingly, averages over a running, specified time interval, here taken as 10 years, were conducted to define temporal changes in these daily changes. The standard deviation (σ) of the daily stage changes was also calculated over running 10-year intervals.

We also evaluated the annual amplitude of discharge, simply defined as the difference between the annual maximum and annual minimum discharge. Simple regressions define any temporal trends. Alternatively, the historical data set of river stage can be divided into an early half and a late (more recent) half; then, any systematic changes can be recognized by comparing the probability distribution diagrams.

2.4 Data presentation

Summary data for the 12 gauging stations examined here are provided in Table 1, and shown in Fig. 2. Table 1 provides the length of the continuous, or nearly continuous, interval where records for peak annual stage and peak annual discharge are both available for these sites ($\sim$1929 to present). Also reported are the means and standard deviations of the relevant flood populations as determined with the various methods described above. While longer records for stage alone are available for some of these sites and are utilized in Figs. 4 and 5, these are not used elsewhere. Thus, the various statistical estimates reported

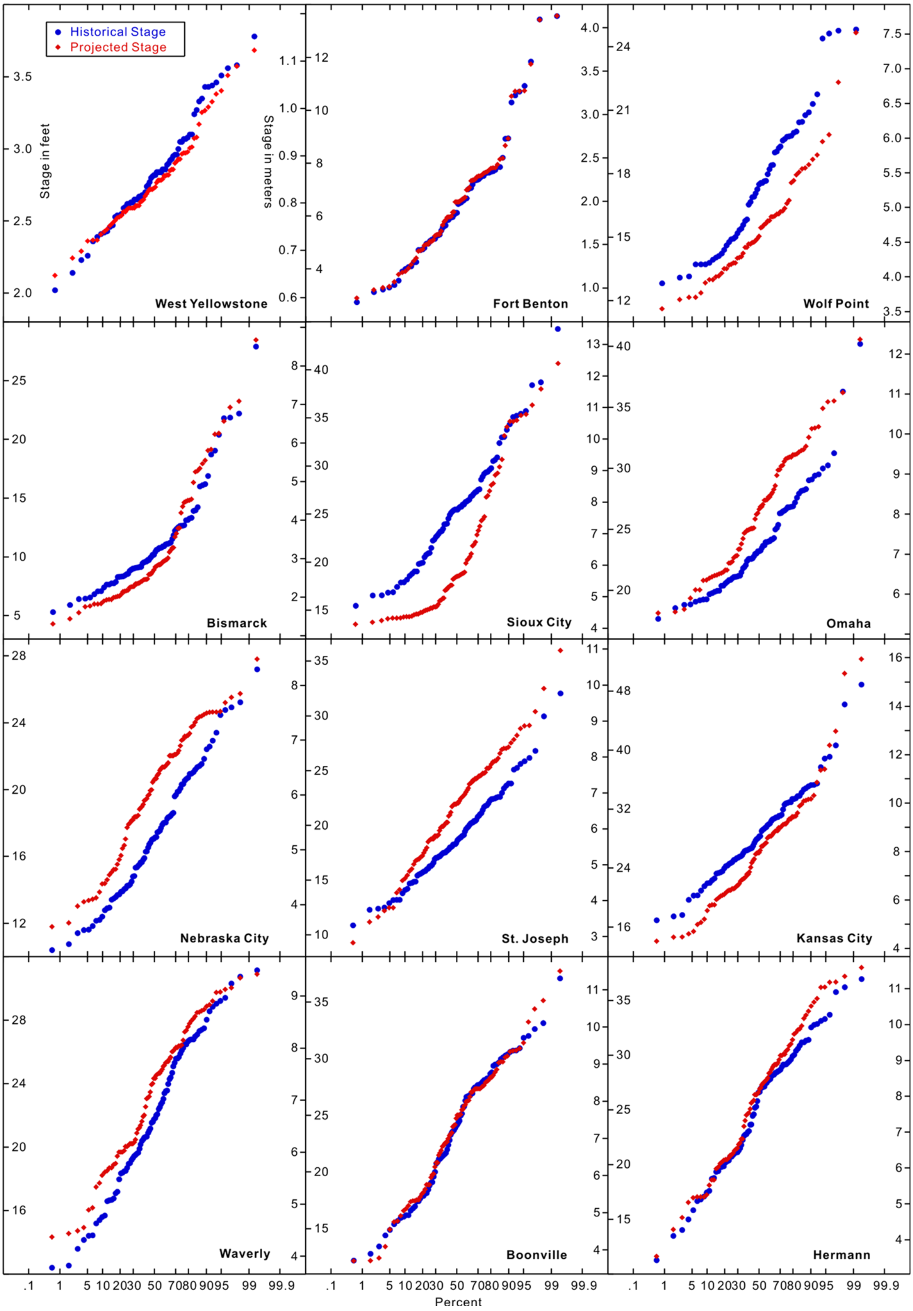

Fig. 2 Standard probability plots comparing historical stages (*blue*) to the projected stages (*red*) calculated from the annual peak discharge. For each plot, the left vertical scale is in feet, while the right scale is in meters

in Table 1 for each site were all calculated using the same data set and can be directly compared. Results are first presented by considering the three major geographic segments along the Missouri River, from upstream to downstream.

3 Results and discussion

3.1 Upper Missouri River and headwaters

The upstream part of the Missouri River basin is represented by the sites at West Yellowstone, on the Madison River tributary, and at Fort Benton on the main stem of the Missouri River. The river channels at both sites are little modified, and human modifications of river flow are small, although some small reservoirs and diversions exist upstream of Fort Benton. Note that the projected, “modernized” stages are almost the same as the stages measured in the past (Fig. 2), so the means and standard deviations of the historical and projected populations are similar. This similarity necessarily extends to the estimated levels of the base flood (Table 1). This result basically confirms that the rating curve has changed little at these sites, over the period of record.

The Criss method gives a similar result to the above methods at West Yellowstone, but provides lower values for the mean, standard deviation, and base flood level at Fort Benton (Table 1). This difference probably originates in a systematic loss of discharge at this site, due to a few impoundments and diversions. Average annual flows have decreased by 17% from 1891 to 2015 at this site (USGS 2016), and peak flows appear to have also decreased by about a factor of two. Unlike the other methods, including the stage projection method, the Criss method automatically accommodates any type of temporal trend that would systematically affect peak annual water levels, including systematic changes of the discharge population, whether discharge has been recorded or not.

The USACE (2004) flow frequency study did not extend into Montana, nor has FEMA produced maps that depict the base flood levels at Fort Benton or West Yellowstone. Thus, no official estimates are available for comparison.

3.2 Middle Missouri River

The middle Missouri River basin has been greatly modified by the construction of six main stem dams and reservoirs, variously constructed between 1933 and 1964 under the “Pick-Sloan Plan.” Three of these number among the largest impoundments in the world, with Fort Peck, Garrison, and Oahe dams impounding reservoirs with gross capacities of 18.5, 23.8 and 23.1 million acre feet of water. The collective residence time of water in this system is nearly 3 years, and it follows that the flow regime of the river has been hugely altered.

USACE (2012) presents useful specific gauge trends for several sites in this region, and these show that strong temporal trends are ubiquitous, but that their character varies with location. In several reaches directly below dams, where reservoir discharge enters a flowing channel, the river channel has become incised by 3 m (10 feet) or more, primarily because the dams have cut off the supply of sediment. The rate of incision has generally decreased over time, but it sharply increased during the large 2011 flood. In contrast, stages have generally increased in reservoir “headwater” regions where flowing rivers first encounter static pools, due to sedimentation and delta formation. Stage increases can be several centimeters per year, in places amounting to 1.2–1.8 m (4–6 feet) since dam construction.

Long-term data were secured for Wolf Point, a “tailwater” site about 113 km (70 miles) downstream of Fort Peck dam, but well upstream of Lake Sakakawea (Garrison Dam), and for Bismarck, close to the head of Lake Oahe. First, the historical stages at Wolf Point are systematically higher than the projected stages, indicating 0.6–1.2 m (2–4 feet) of channel incision. This difference is clearly reflected in the lower mean of the projected stage data, as well as in the lower calculated level of the projected base flood (Table 1); this same difference is seen in Fig. 2 for the more frequent floods. The standard deviation of the projected stage data is also lower than for the raw historical data, probably reflecting the desired, regulatory attenuation of high discharges. However, Fig. 2 shows a significant departure from linearity at this site and does not show a large difference between the raw and projected base flood levels, indicating that the modest incision has had little influence on high flows. In contrast, the Criss method shows that the estimated base flood levels have become significantly reduced below Fort Peck dam (Table 1). This clearly shows that the frequency of high peak flows has been diminished due to flood control efforts, so the present-day flood population is different than the historical average. Finally, the base flood level estimated by FEMA (2016) at Wolf Point is even lower (Table 1), yet the peak stage of 7.55 m seen at Wolf Point in 2011 is much higher than the official base flood of 5.01 m. Of course, the official estimate predated the 2011 flood, which overwhelmed the flood control system in a manner that was not anticipated.

Results at Bismarck, a site located at and slightly above the “headwaters” of Oahe reservoir, differ from those at Wolf Point. The specific gauge data (USACE 2012) suggest that minor sedimentation has occurred, elevating water levels associated with low and modest annual peak flows; however, incision during the 2011 flood appears to have reversed this effect, at least temporarily. However, water levels for flows of 2830 m^3/s (100,000 cfs), and probably above, appear to have increased by 0.9–1.5 m (3–5 feet) over pre-reservoir conditions.

The above changes are not clearly defined, and this uncertainty is reflected in the probability plots and base flood estimates at Bismarck. For example, a comparison of the probability plots for the historical and projected stages (Fig. 2) suggests that the recent incision that began during the 2011 flood has continued, because the stages for the small, more frequent annual floods are lower than the historical stages. This incision is contrary to the main, post-construction trend at Bismarck, but the recent rating curve probably reflects this incised condition. Explanation might be found in the post 2011 average and peak annual flows being above the long-term averages. For large annual floods, the projected stages are slightly higher than the historical stages (Fig. 2), so the estimated stage of the base flood is somewhat higher than the estimate based on the raw historical record, or the FEMA estimate. The base flood level estimated by the Criss method is much too low, because the abrupt discontinuity in the flood record caused by the construction of major structures is poorly represented by a linear trend (see below).

3.3 Lower Missouri River

Below the lowest of the main stem dams, the lower 1207 km (750 miles) of the Missouri River has been converted to a navigation channel, and most areas are now isolated from the floodplain by levees. Thousands of wing dams have narrowed the channel twofold over its natural state, while man-made cutoffs have reduced the river length by about 8%, and practically all of the islands and sand bars have been eliminated (Funk and Robinson 1974). However, the flood response of the river is different above and below Kansas City.

At and above Kansas City, stages have generally decreased for given low flows, including small annual peak flows (USACE 2012); this probably reflects channel incision caused by the greatly reduced sediment supply below the large, main stem reservoirs. In contrast, stages have generally increased for high flows. The overall effect on the annual flood population at any site depends on the water level where these stage decreases transition to stage increases, and to confounding effects such as the natural cutoff at St. Joseph that occurred during the 1952 flood.

Several of these effects are apparent in comparing the probability plots for the raw historical stages and projected stages. The channel incision effect dominates for annual floods of any magnitude at Sioux City, and for all <10 year floods at Kansas City, but applies only to the smallest annual floods at Omaha and St. Joseph. For the highest, most damaging floods, and for all annual floods at Nebraska City, the stages of the projected flood population have increased, indicating that severe floods have worsened (Fig. 2). This effect is likewise seen in the higher level of the projected base flood at all of these sites except Sioux City, as compared to the base flood levels estimated from the raw historical data (Table 1). The base flood stages estimated by the Criss method are similar to the projected stage estimates at and below Nebraska City. However, the Criss method estimates are too low at Sioux City and Omaha, probably reflecting a lack of linearity as was found at Bismarck, as these sites are closest to the major dams. The official estimates of flood risk may also be too low at most sites above Kansas City (Table 1).

Below Kansas City, stages appear to have increased for annual floods of practically any recurrence interval (Fig. 2). The projected base flood stages are higher than the raw historical estimates, and the Criss method provides similar or higher stages (Table 1). The official stage estimates for the base flood levels are generally similar to the estimates determined from the raw historical data, which are clearly too low.

3.4 Annual discharge amplitude

The annual discharge amplitude, representing the difference between the highest and lowest flows in any given year, provides insight into the complex relationships discussed

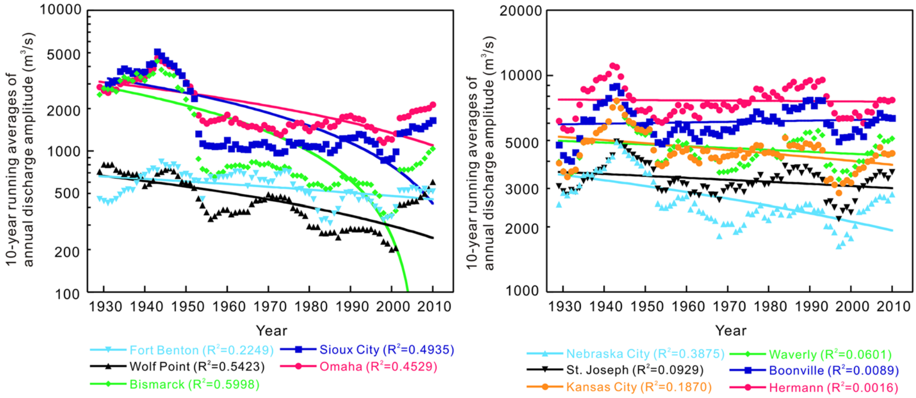

Fig. 3 10-year running averages of the annual discharge amplitude at various sites on the Missouri River. *Solid lines* are linear least-squares fits, which can appear to be curved on this logarithmic plot

above. Figure 3 shows that the behavior of this quantity differs greatly among the various sites.

The discharge amplitude appears to have slowly and steadily decreased at Fort Benton, on the upper Missouri River. As discussed above, this decrease probably reflects lower peak annual flood flows, likely due to small impoundments and modest diversions in this area. A flood population now constituted of systematically lower flows would reduce the risk of high stages. This reduction is reflected in the lower base flood level estimated by the Criss method (Table 1), which collectively accommodates all types of steady temporal trends. However, this real reduction is overlooked by risk estimates that are based on the raw and projected annual peak stages, which are nearly identical.

The discharge amplitude has significantly decreased in the dam and reservoir area, as represented by data at Wolf Point and Bismarck (Fig. 3). This decrease in amplitude also mostly reflects lowered peak flows, which in this reach reflects flood control effected by the reservoirs. A direct consequence is that the base flood estimated by the Criss method is lower than that determined from the raw and projected stages. A difference in the realism of the Criss estimate arises from the much earlier impoundment of Fork Peck reservoir, above Wolf Point, than of Lake Sakakawea above Bismarck. That is, Fort Peck Lake was filled at the beginning of the data interval that was analyzed, so that interval is dominated by relatively steady, post-construction changes, for which the Criss method is well suited. In contrast, Lake Sakakawea was filled in the middle of the data interval analyzed, and that interval spans the abrupt change in behavior that is not well represented by linearity; this nonlinear character causes the Criss method to greatly underestimate the present-day, base flood level at Bismarck.

Interestingly, the magnitude of the discharge amplitude reduction below the reservoirs decreases with downstream distance. Thus, the amplitude reduction is the greatest at Sioux City, strong at Omaha, smaller at Nebraska City, and slight at St. Joseph, Kansas City and Waverly. As at Bismarck, the sharp overall reduction in annual flood flows in the middle of the analyzed historical interval causes the Criss method to significantly underestimate the base flood levels at Sioux City and Omaha.

Finally, furthest downstream of the dams, discharge magnitude has increased slightly at Boonville and Hermann. This increase, combined with channelization and levee effects, has elevated the stages of severe floods. Estimates made using the Criss method are probably the best in this region, as all types of temporal change are accommodated. In contrast, the official estimates are too low in this area, as they are too similar to estimates based on the raw historical stages.

3.5 Daily stage change

Long-term records for daily river stage are available at Bismarck (since 1897; USGS 2016) and Hermann (since 1874; USACE 2016; Ehlmann and Criss 2006). These data can be readily used to determine the daily stage change. The result is rather chaotic, but smoothing is easily achieved by conducting running averages of the absolute values of daily change. Flow regulation along the middle Missouri River has significantly reduced the magnitude of those daily changes, as well as of the standard deviation of those values, as exemplified by Bismarck (Fig. 4). In contrast, channelization of the lower Missouri River has greatly amplified the magnitude of daily change, as pointed out by Ehlmann and Criss (2006), and also amplified the standard deviation of that change. Although daily data and annual peak data are different, it is clear that engineering modifications have made the middle Missouri River less chaotic, and the lower Missouri River more chaotic, and these

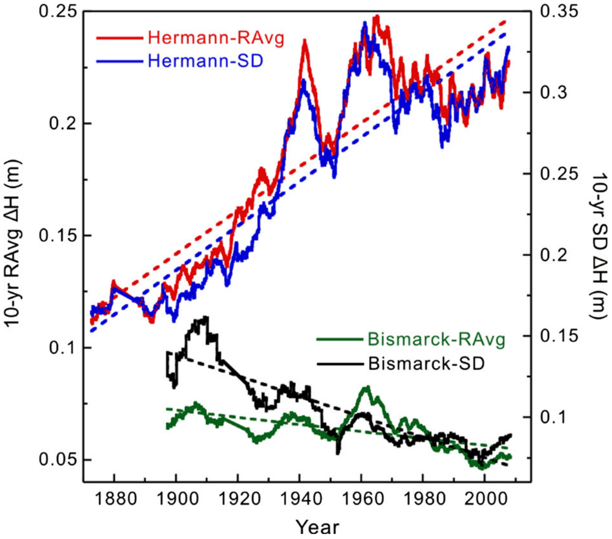

Fig. 4 10-year running averages (RAvg) and standard deviations (SD) of daily stage change at Hermann and Bismarck. *Dashed lines* are least-squares fits

changes clearly influence the probability of flooding. This effect seems to be reflected in the estimated water levels for the base floods at these sites, particularly when the raw historical estimates are compared to those made by the Criss method (Table 1).

3.6 Comparison of different base flood estimators

Flood risk is difficult to estimate, particularly for the severe, extreme events that have a low probability of occurrence (e.g., Klemes 2000). Historical records are too short to accurately define the recurrence intervals for such events, and engineering structures along rivers have modified the flood population in ways that render invalid the techniques that treat the data as a static, historical population. Any long-term changes to precipitation patterns or to watershed character that would affect the amount and delivery of discharge will add additional confounding effects.

Sites that have the longest historical records best illustrate the above differences. It is particularly useful to examine standard probability plots of the entire set of historical stages and then to equally divide those data into an “early” half and the “late”, i.e., most recent half, and compare the results. When this is done, the three distributions are nearly identical on the upper Missouri River at Fort Benton (Fig. 5). However, the distributions differ greatly at Sioux City, just downstream of the huge flood control projects, where the recent data show the lowering of flood risk because of channel incision and flow regulation efforts. Finally, on the lower Missouri River far downstream of the dams, flood levels for practically any return interval have significantly increased at Boonville and Hermann (Fig. 5).

The different methodologies for flood risk estimation treat these matters in different ways. Simple statistical analysis of the raw historical stage data, which involves treating the data as if they represented a homogeneous population, is approximately valid at Fort Benton, but overestimates modern flood risk at Sioux City, and underestimates this risk

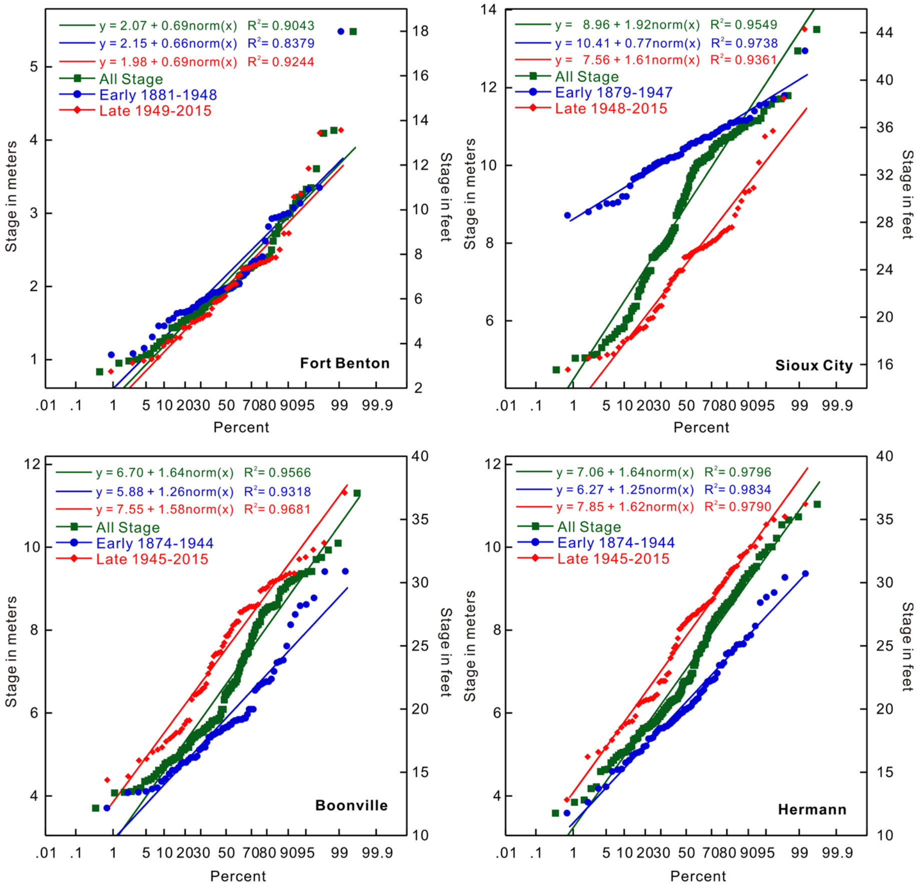

Fig. 5 Comparisons of stage probability distribution between the early and late half periods at the selected sites on the Missouri River; *lines* are least-squares fits. Note that Jarvis (1936) and NOAA (2016) provide maximum annual stages for these sites over the long intervals that are indicated

much further downstream. As for the official calculations, as demonstrated by Criss and Luo (2017), the effective values of μ and σ can be extracted from the stages tabulated by USACE (2016) for floods of various recurrence intervals. The tabulated water levels for floods of different recurrence intervals can be directly plotted against the associated values of K_L for a symmetrical distribution (Fig. 6). The slope of that regression line provides a good estimate for σ, and the y intercept provides μ. High linearity of the points indicates that the underlying distribution closely approaches a symmetrical, un-skewed Gaussian curve of river stage. Thus, even though they were calculated with a very complicated protocol, the official base flood estimates (USGS 1981; USACE 2004) are most similar to the simple statistical analysis of historical stages, and they commonly underestimate flood risk.

The new stage projection method "modernizes" the historical stage data, but relies on the accuracy of historical discharge estimates as well as on the accuracy of the modern rating curve. This method should greatly improve the accuracy of risk calculations, but only accounts for changes that can be seen in rating curves, and will not account for

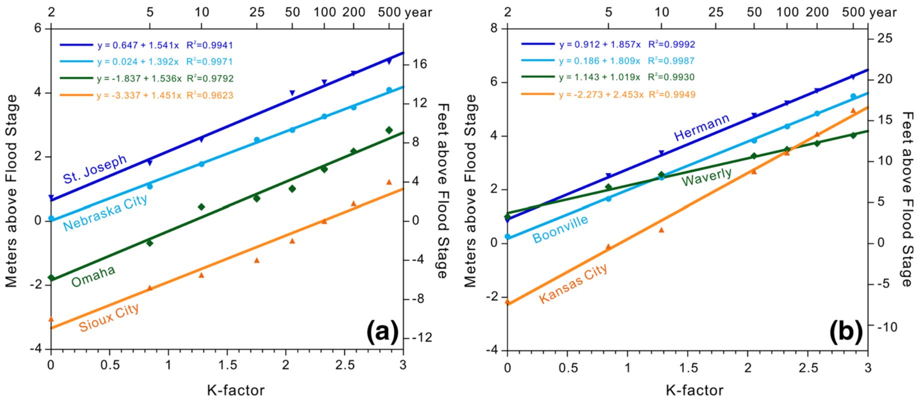

Fig. 6 Water levels for 2–500-year floods have been estimated by USACE (2016) for several sites on the lower Missouri River. When these levels are plotted against the K value for the indicated recurrence interval (*top*), the slopes of the linear trends define the effective standard deviation of the estimated flood population. Additionally, the y intercepts define the effective mean of that population (cf. Eq. 1; Table 1); these intercepts can differ from zero because local flood stage is arbitrary defined. Despite the complex, discharge-based, computational protocols used to define these official flood levels, the strong linearity ($R^2 > 0.98$) seen for these sites indicates that the estimated population closely resembles a simple Gaussian distribution

temporal differences in the "early" and "late" populations that result from changes in the amount or delivery of flow. Finally, the Criss method attempts to use differences between the "early" and "late" populations to estimate the present-day means and standard deviations that are then used to define present-day risk. This method accommodates any type or combination of temporal changes and works well if those changes are either negligible or occurred steadily, because linearity of the combined effects on the flood population is assumed. However, this method can overcorrect for the effect of any temporal change, be it positive or negative, if the changes occurred in a nonlinear manner, and particularly if the change is large and occurred abruptly.

4 Conclusions

Profound but different changes to major reaches of the Missouri River make it a useful laboratory to examine the effects of engineering works, and to test the effectiveness of different methodologies used to estimate flood levels. The upper Missouri River and its headwaters remain most similar to natural, historical conditions, although both average and peak annual discharges have become somewhat diminished by diversions and small impoundments. The population of peak annual water levels has been basically static, and several different methods provide similar estimates for base flood stage.

The middle Missouri River has been isolated by levees, and profoundly and abruptly changed from 1933 to 1964 by the construction of several huge, main stem dams and reservoirs. Flow and stage variations have become attenuated, on both daily and annual timeframes. Stages have been lowered downstream of dams by channel incision and restricted sediment supply, but have increased due to aggradation in the heads of reservoirs. Simple statistical analysis of the flood population, which presumes stationarity of conditions, can either overestimate or underestimate flood risk, but due to the abruptness of dam construction, the Criss method can overcorrect for such effects. The stage projection method presumes the correctness of historical discharge estimates and rating curves, and does not accommodate effects due to flow regulation.

Further downstream, the free flowing lower Missouri River has been isolated by levees and greatly narrowed and channelized for navigation. The upstream reservoirs have cut off much of the sediment supply, causing significant channel incision at and above Kansas City. Also, peak discharges have been lowered, but the reduction in annual discharge amplitude is the largest just below the reservoirs; this effect becomes progressively smaller downstream, being small at Kansas City and Waverly, and apparently negligible at Boonville and Hermann. Estimating flood risk from a simple statistical analysis of raw historical stage data is clearly not appropriate, yet at many sites the official methodology provides estimates that closely resemble those static calculations. The stage projection method better accommodates changes to channel morphology, but does not properly account for flow regulation effects that are quite significant at Sioux City and Omaha, unless only a very short, post-construction record is analyzed, which introduces other problems. The Criss methodology might be the best at and downstream of Nebraska City, but overcorrects further upstream where abrupt, nonlinear changes to river behavior have occurred. Available data suggest that the middle and lower Missouri River are continuing to adjust to the profound effects of flow regulation, sediment trapping, and channelization that vary with location, and will require many decades if not centuries to stabilize. Flood risk estimation will be particularly challenging until a new state of equilibrium is attained along the river, and a long record of subsequent flood behavior is assembled.

References

Belt CB (1975) The 1973 flood and man's constriction of the Mississippi River. Science 189:681–684

Benameur S, Benkhaled A, Meraghni D, Chebana F, Necir A (2017) Complete flood frequency analysis in Abiod watershed, Biskra (Algeria). Nat Hazards 86:519–534

Chow VT (1964) Handbook of applied hydrology, vol 8. McGraw-Hill, New York, p 23

Collenteur RA, Moel H, Jongman B, Baldassarre GD (2015) The failed-levee effect: do societies learn from flood disasters? Nat Hazards 76:373–388

Criss RE (2016) Statistics of evolving populations and their relevance to flood risk. J Earth Sci 27(1):2–8

Criss RE, Luo M (2017) Increasing risk and uncertainty of flooding in the Mississippi River basin. Hydrol Process 31(6):1283–1292

Criss RE, Shock EL (2001) Flood enhancement through flood control. Geology 29:875–878

Ehlmann BL, Criss RE (2006) Enhanced stage and stage variability on the lower Missouri River benchmarked by Lewis and Clark. Geology 34:977–980

FEMA (2016) Flood maps. http://msc.fema.gov/portal

Funk JL, Robinson JW (1974) Changes in the channel of the lower Missouri River and effects on fish and wildlife. Missouri Department of Conservation, Aquatic Series, vol 11, p 52

GAO (1995) Midwest flood, information of the performance, effects, and control of levees. GAO/RCED-95-125

Jarvis CS (1936) Floods in the United States: magnitude and frequency. U.S. Geological Survey Water Supply Paper, vol 771, pp 1–497

Klemes V (2000) Tall tales about the tails of hydrological distribution. J Hydrol Eng 5(3):227–239

NOAA (2016) Record of climatological observations. http://www.ncdc.noaa.gov/IPS/coop/coop.html

Pinter N (2001) Assessing flood hazards on dynamic rivers. EOS 82(31):3p

Salas JD, Obeysekera J (2014) Revisiting the concepts of return period and risk for nonstationary hydrologic extreme events. J Hydrol Eng 19:554–568

USACE (2004) Upper Mississippi River system flow frequency study: final report. January 2004. http://www.mvr.usace.army.mil/pdw/pdf/FlowFrequency/flowfreq.html

USACE (2012) Missouri River stage trends, Technical report. August 2012, 46p

USACE (2016) Flow frequency query: Missouri River. http://rivergages.mvr.usace.army.mil/flow_freq/flow_freq.cfm

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 (2016) Peak streamflow for the nation. http://nwis.waterdata.usgs.gov/usa/nwis/peak