Uncertainty Analysis for Large Dam Failure Frequencies Based on Historical Data

Fernando Ferrante*, Michelle Bensi, Jeffrey Mitman
U.S. Nuclear Regulatory Commission, Washington DC, USA [1]

Abstract: External flooding hazard assessments typically include considerations for multiple water sources, including catastrophic dam failure accidents, if applicable. For large dams in the United States, a recent study identified a number of significant failures, of which a subset is classified as events involving catastrophic large dam failure. This analysis indicates that the dominant causes of failure are about the same as those for the entire population of dams: overtopping due to exceedance of the reservoir level (usually the result of severe weather phenomena), foundation effects and internal erosion, and miscellaneous other causes including poor design and maintenance, as well as seismic events. While there are well-documented cases of significant events involving such failures around the world, the specific likelihood of such an event is challenging to predict. In attempting to quantify the frequency of large dam failure events, it was recognized in the aforementioned study that reliance on anecdotal historical events must take into account the significant ambiguity and lack of information completeness involved in using such data. At the same time, the use of available databases provides a framework to evaluate in more detail the extent to which this uncertainty may impact the understanding of bounding dam failure rate estimates. In this work, sensitivity studies were performed in order to evaluate the changes due to a number of categorization bins for large dams including dam type, construction completion date, and dam incident information. Bayesian analysis tools were also used for the derivation of posterior uncertainty distributions that include subjective information such as data quality and expert judgment considerations. The extent of the variation in the commonly derived point estimate is documented and discussed for a number of the categories and assumptions usually relied upon in available literature when estimating failure rates for large dams.

Keywords: External Flooding, Dam Failures, Initiating Event Frequency, Uncertainty Analysis.

1. INTRODUCTION

Industrial sites are vulnerable to a wide range of natural hazards including earthquakes, high winds, tornados, hurricanes and floods. In particular, external flooding can be caused by initiators such as extreme meteorological events (e.g. severe storms, tides, and waves), seiche/tsunami, and dam failures. Understanding the risks posed to industrial facilities by these hazards is important for a variety of reasons, such as resource allocation (e.g. prioritization of funds for mitigation/remediation) and emergency planning. Despite the importance of understanding the risks posed by extreme natural events, the maturity of available methodologies and data for assessing the frequency of occurrence of these hazards varies significantly from hazard to hazard. In this paper, we focus specifically on understanding the frequency of dam failure events, which have the potential to affect facilities located both upstream and downstream of the dam. While failures of dams upstream of facilities pose a potential for flooding, failure of dams downstream of facilities can cause unavailability of water to the site.

Dam failures can be caused by a variety of mechanisms including overtopping, seismic events, internal erosion and piping, operational/mechanical failures, and combinations of these initiators. While severe earthquakes and extreme flood events can cause dam failures, these events have a relatively low likelihood of occurrence. Conversely, failures from internal erosion, piping, and operational/mechanical failures can occur without a specific initiator (e.g. earthquake, large rain event).

Dam-regulating entities in the United States (US), such as the US Army Corps of Engineers (USACE) (USACE, 2006), the US Bureau of Reclamation (USBR) (USBR, 2010), and the Federal Energy Regulatory Commission (FERC) (FERC, 2005) have developed frameworks for the purposes of understanding risks, assessing public safety, and allocating resources across portfolios of dams. Thresholds for taking action (e.g. to mitigate risks through retrofits) are often based on estimates of relative risk and subjective metrics. The risk frameworks typically are based on processes involving expert elicitation and, if appropriate or necessary, dam-specific engineering assessments. However, dam-specific assessments have not been performed for all

[1] This paper was prepared, in part, by employees of the U.S. Nuclear Regulatory Commission on his or her own time apart from his or her regular duties. NRC has neither approved nor disapproved its technical content.

dams, may not be cost-justified, or may not be readily available to all stakeholders. For this reason, it is useful to have generic dam failure rate estimates than can be used as screening criteria or when more detailed assessments are not available.

The goal of this paper is to use classical and Bayesian statistical methods to develop generic estimates of dam failure frequency based on information contained in two major US databases. Because these databases were not originally developed to support the quantification of dam failure rates, there are several challenges associated with their use for the current application. The work by Ferrante et al (2011) provides a literature review of existing dam failure frequency studies and includes a detailed discussion of the challenges associated with using these databases for deriving dam failure frequencies. Due to the various limitations and caveats associated with use of the aforementioned databases, the sensitivity of statistically-based estimates to a variety of factors and assumptions is explored. These sensitivity studies are designed to help analysts understand the potential variability of the estimated generic values.

2. METHODOLOGY AND RESULTS

Two main sources of information are used in this work to perform sensitivity and uncertainty analysis on large dam failure frequency estimates for the US based on historical data: (1) the National Inventory of Dams (NID) database maintained by USACE (USACE, 2011), and (2) the National Performance of Dams Program (NPDP) database, maintained by Stanford University (Stanford, 2007). Currently, these sources facilitate the collection of (1) the total number of years of operation for US large dams of particular characteristics, such as dam type, and (2) the number of historical failures of US large dams. These databases contain the best available collection of US dams information and dam failure catalogues, and are more complete and accurate than previous efforts to compile such information. However, they also contain significant sources of uncertainty and missing information, which needs to be carefully considered (as in any effort involving data analysis on such a scale). For example, it should be noted that the databases were not created for the specific purpose of performing dam failure frequency calculations and were not designed to be fully consistent with one other. Nevertheless, these databases are still the primary source of information on existing dams and events. This paper does not intend to express judgment on the quality of the efforts made to develop these databases; instead, it highlights the challenges in the input and categorization of data for such a wide population, which potential users also need to take into account when deriving estimates for low-probability events. The sources of uncertainty from the information gathered will be discussed in more detail in the subsections below.

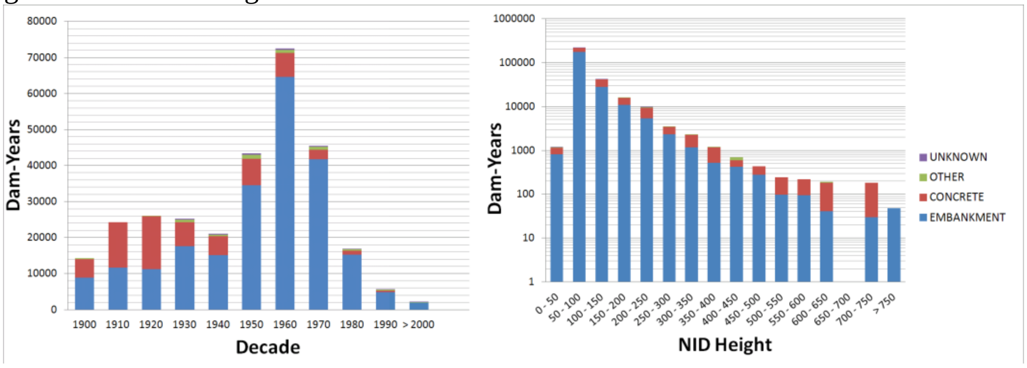
Following the framework originally developed in Ferrante et al (2011), this work is restricted to the US dam population, to which a subjective (but necessary) “large” dam definition is imposed. There are a wide range of categorization criteria used by various US and international organizations to define classes of dam size, which are based primarily on height and volume of reservoir impounded. These criteria for categorizing a dam as “large” or “small” can be highly subjective. However, for the current application, there is a clear need to establish criteria for distinguishing between large and small dams because an individual dam may be more or less susceptible to certain failure modes based on its size or reservoir volume (e.g. a dam with a large reservoir volume and substantial population downstream may have less vulnerabilities due to augmented inspection and maintenance programs). The International Commission on Large Dams (ICOLD) establishes that a dam can be defined as “large” if its height from the foundation exceeds 15 meters [49.2 feet]. For dams between 5 meters [16.4 feet] and 15 meters [49.2 feet] in height, ICOLD will apply the large dam definition if its reservoir volume exceeds 3 million cubic meters [3,923,852 cubic yards] (WCD, 2000). In USACE (1979), dams are defined according to height and reservoir requirements as well, where “small” dams are those between 7.6 meters [25 feet] and 12.2 meters [40 feet] in height, “intermediate” dams comprise those between 12.2 meters [40 feet] and 30.5 meters [100 feet] in height, and “large” dams exceed 30.5 meters [100 feet]. In the current work, we set the criteria for a “large” dam as those exceeding 12.2 meters [40 feet] in height (no reservoir volume definition is used). A sensitivity analysis for the impact of increased height thresholds is performed.

2.1. NID Database

The National Inventory of Dams (NID) database contains the most extensive listing of dams in the US. It is periodically updated and maintained by USACE with support from a number of state and federal agencies, which submit individual dam information through cooperative participation. A description of the inclusion criteria and required submittal information are described in USACE (2008), which includes 60 fields such as dam height, dam type, storage, and location. The NID database also includes a number of fields restricted from public release, which were not used in this analysis. As in Ferrante et al (2011), the only source of dam-year operational data for this work continues to be the NID database, for which the 2010 update is used. The 2010 version of the NID database includes a listing of over 84,000 dams. Applying the large dam criteria used in this work yields 11,964 dams (approximately 14% of the dams in the database). For dam height, it should be noted that the database field "NID height" is used, which corresponds to the maximum value of dam height, structural height, and hydraulic height, as submitted by NID participants and established by USACE ("NID height" is accepted as the general height of the dam). Specific fields, aside from dam height, that were explicitly considered in this analysis include: dam type, purpose, and year completed. Only dams built since 1900 (i.e., 20th and 21st century dams) are considered in this study. A small percentage of dams (less than 10%) do not have entries for year of construction completion, which is defined by NID as the year in which the original main dam structure was completed. In these cases, the authors assumed an average completion year based on the available information is used, which corresponds to 1963 (i.e., 43 dam-years per dam with a cut-off date of 2006).

For dam type, NID specifies abbreviations to be used for commonly defined dam attributes: earthfill (RE), rockfill (ER), gravity (PG), buttress (CB), arch (VA), multi-arch (MV), concrete (CN), masonry (MS), stone (ST), timber crib (TC), and other (OT). Submittals often include a combination of attributes to define impoundments with distinctly designed sections. For example, a specific site may include a buttress or an arch gravity section supported by embankments, preventing a single classification in NID. According to the classification scheme, dam type combinations are expected to be provided in order of importance such that a dam type combination initiating with RE or ER (e.g., REPGCN) will indicate an impoundment consisting mostly of embankment sections.

The vast majority of entries are comprised of single attributes (90%), of which earthfill dams account for approximately 89%. It should be noted that a small percentage of dams (2%) have not been categorized with respect to dam type. In order to develop a feasible categorization scheme, four major overarching dam types are used to bin the various single and combination entries in NID: embankment dams (including earthfill, earthfill-rockfill, rockfill), concrete dams (arch, gravity, multi-arch, buttress, and concrete), other type dams (masonry, stone, timber crib, and other), and unknown type dams (empty entries). Entries with multiple dam attributes are categorized with respect to their order of importance, unless additional sources are available that suggest a different dam type category. Additionally, in this work, impoundment structures used to retain waste material resulting from activities such as mining (commonly known as "tailings dams" and usually categorized in NID as embankment dams) are also included in the "other" category because these types of dams are not usually designed and maintained to equivalent standards as other embankment dams. To segregate tailings dams from the overall embankment dam population, a NID category that lists the purpose of individual dams is used (i.e., tailings are identified with "T" in the field "purpose"). Furthermore, dams with names that contain key words such as "tailings," "slurry impoundment," and "mining refuse" are also segregated. Figure 1a shows the range of US large dams built per decade since 1900 with respect to the major dam types considered in this study; indicating a significant period of dam construction between 1950 and 1980. The distribution of dam height for large US dams (using the 12.2 meters [40 feet] criteria) is shown in logarithmic scale in Figure 1b.



(a) (b)

Figure 1. Dam-years (a) per decade and (b) per NID height for all large dam types in NID

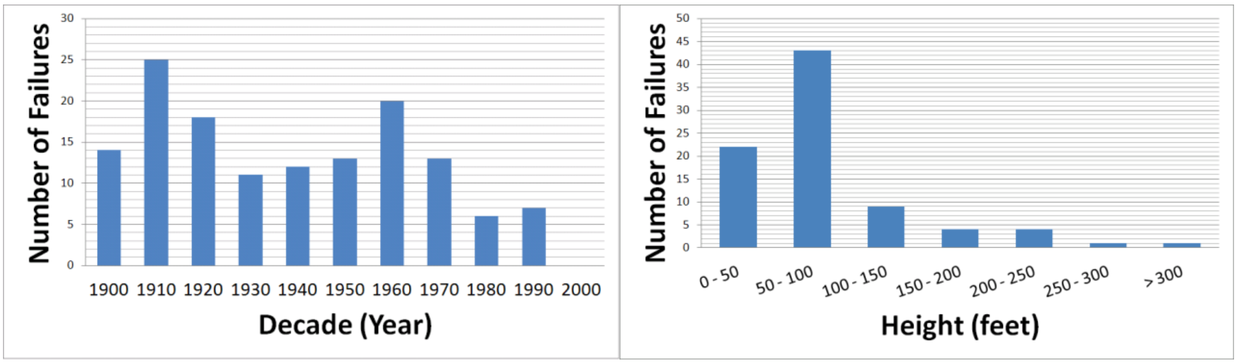
The main source of uncertainty from the NID database is the dam type characterization, as various entries could be misclassified with respect to their dominant attribute. It is unclear, for example, whether the distinction between earthfill versus rockfill composition in different entries is sufficiently robust to justify further parsing of embankment dams. The overarching major dam types used in this work are intended to address this issue to some extent, by avoiding a more granular categorization.

2.2. NPDP Database

The NPDP database was established in 1994 as an information resource for sharing dam incidents and failures within the engineering and dam safety professional community in the US, and is maintained by Stanford University (Stanford, 2007). It is the main source of data on dam failures used in this study. Similar to the NID database, it contains a large number of entries that include information on incident date, dam type, dam height, and other attributes for individual dam failures. There are 1109 dam failures identified, as well as 1776 dam incidents that are searchable by various attributes. NPDP defines a dam failure as a "breach and uncontrolled release of the reservoir." As noted in Ferrante et al. (2010), due to the difficulty in establishing accurate information for a large number of historical dam failure events, a complete description of each individual dam failure is not available for all entries. In particular, a significant number of failure events contained in the NPDP database do not have information regarding dam type, dam height H, and/or construction completion year T_{CY}.

A set of criteria similar to the one presented in the previous subsections was used to define applicable dam failure events: only dam failure events for dams with H equal to or above 12.2 meters [40 feet], built after 1900 were considered. Events with missing dam height information were excluded. However, in order to achieve as much information completeness as possible, additional sources of information were researched and reviewed to identify (1) dam failure events not included in NPDP, and (2) information missing from existing dam failure entries in the NPDP database. This was achieved by identifying individual documentation on specific dam failure events (Kocahan & Taylor, 2002) and cross-checking information with dam failure listings (e.g., VP Singh, 2010). It should be noted that several of the dams with failure events were also later rebuilt, and these dams are identified in NID. Limited cross-checking with the NID information is possible since there is a possibility that rebuilt dams do not exhibit the same attributes as the dams that failed.

Application of the height and vintage criteria utilized in this paper results in a set of 139 dam failure incidents. It is noted that a subset of these incidents are associated with NPDP database entries that are missing construction completion year information. In this report, analysts varied the construction completion year for dams that do not have this information available via NPDP (or other sources). For example, if such entries are assumed to have construction completion year equal to the year in which the failure incident occurred (i.e., infantile or early failure), the distribution of the number of failures with respect to decade and dam height, H, are shown in Figure 2. As demonstrated in this figure, failures of large dams in the US have historically clustered around dams built in the early and mid-20th century (i.e., 1910 – 1920 and around 1960) which are also associated with years of increased dam construction in the US (see Figure 2a). It is also observed that most failures impact dams with heights less than 30.5 meters [100 feet].



(a) (b)

Figure 2. Dam-years per (a) decade and (b) NID height (feet) for all large dam types in NID

The failure database constitutes a significant source of uncertainty in estimating dam failure rates, due to the incompleteness of information as well as the classification of failure modes and dam type attributes of individual events. It is unclear whether certain events identified as ‘failures’ would be best categorized as ‘incidents’ based on the event descriptions provided (e.g., certain events are categorized as ‘partial failures’). Failure mechanisms are also identified with a number of key descriptors, such as “flood,” “seepage,” “piping,” “spillway failure,” “erosion,” etc. Considering the description of specific events (including those with more detailed accounts), it is clear that developing a categorization of dam failure mechanisms poses a significant challenge because dams can fail due to a wide range of causes (including in combination with flooding events). While events such as overtopping of a dam due to extreme flooding can be identified, it was deemed in Ferrante et al (2010) that parsing selected failure modes without sufficient technical basis can produce artificially low dam failure frequencies; therefore, this is also not pursued here. Additionally, various sources indicate that certain failures were not included in NPDP or had information otherwise missing. While an attempt was made to compile a more complete list of failures for this work, an exhaustive and thorough investigation was not performed and it is possible that additional failures or more detailed information could be used to further refine dam failure events applicable to large dams. Finally, since there have been no major updates to NPDP since 2006, the estimation of dam failure frequencies will be limited to this date, as including dam-years accrued between 2006 and 2010 would not have an equivalent dam failure events contribution. In other words, although additional failures may have occurred since 2006, no effort has been made to collect such information in this work.

2.3. Point Estimate Calculation

Based on the assumptions discussed for the dam-year and dam failure events obtain primarily from NID and NPDP, a point-estimate of the annualized failure frequency, f, can be derived for various ranges of dam types, height H, and construction completion year T_{CY}.

For failure events with missing construction completion year, a value needs to be assumed for the time interval between known incident date and unknown completion year, ΔT_F. For example, assuming all failure events associated with unknown construction completion year correspond to “early mortality” such that $\Delta T_F = 0$ cases (i.e., the failures took place during or immediately after construction completion prior to operational status) and also assuming an average construction completion year of $T_{CY} = 1963$ for operating dams with missing construction completion year results in the point estimates shown in Table 1. An overall value of f = 2.71E-4/year is obtained, with a generally decreasing trend between early 20th century dams (1910 – 1920) and later periods. Table 1 also presents the results per dam height H, where the concentration of dam-years occurs at values of less than 61 meters [200 feet]. Due to the limited amount of available data, accurate estimates for dam heights above 61 meters [200 feet] are not possible given the lack of dam-years and dam failure events. However, it is noted that it is to be expected that larger dams have better maintenance and inspection programs and, therefore, lower failure frequencies.

Table 1. Dam failure frequencies for all dam types per construction year and height range

	CONSTRUCTION YEAR RANGE					**DAM HEIGHT RANGE (feet)**					
	1920	**1940**	**1960**	**1980**	**2006**	**100**	**200**	**300**	**400**	**800**	
TOTAL	**1900**	**1920**	**1940**	**1960**	**1980**	**40**	**100**	**200**	**300**	**400**	
512,745	61,194	73,366	105,060	240,332	32,793	431,276	59,851	13,721	4,708	3,189	**DAM-YEARS**
139	39	29	25	33	13	104	25	8	1	1	**FAILURES**
2.71E-4	6.37E-4	3.95E-4	2.38E-4	1.37E-4	3.96E-4	2.41E-4	4.18E-4	5.83E-4	2.12E-4	3.14E-4	***f (/year)***

Given that it may not be realistic to assume all failures with missing construction completion year have $\Delta T_F = 0$, a variation in ΔT_F was performed. Based on the existing information, the mean value for ΔT_F is 19.5 years. Using $\Delta T_F = 19.5$ the failure frequency for all dam types, between 1900 and 2006 results in a failure frequency of f = 2.40E-4/year. While the sensitivity is small for the entire period considered, a reduction is achieved for later construction ranges since an increasing ΔT_F parameter eventually results in a reduction in the number of failures considered in later decades (e.g., for dams of all types built between 1980 and 2006, there is a reduction to f = 1.22E-4/year with ΔT_F = 19.5 years). If all failure events for which ΔT_F is unknown are excluded, the failure frequency is f = 1.64E-4/year for the period 1900 – 2006.

The results discussed so far include a number of failures that may be considered representative of an early mortality period (in addition to those included by assuming $\Delta T_F = 0$). Some events are clearly indicative of failure during construction or initial filling of the reservoir, while others took place immediately after construction was completed. For a significant portion of the failures considered, it is not possible to ascertain when or how the failure took place to discern early mortality attributes. As discussed in Ferrante et al (2010), it would be expected that dams that survived through the first few years of operation would have reduced values for failure frequencies. However, the estimates are sensitive to the assumed range considered to represent an early mortality period and any assumptions need to be considered carefully. In order to assess this effect with the data developed in this work, an early mortality threshold ΔT_{EM} is used to represent the number of years for which an individual failure event should be excluded in order to assess a failure frequency for dams that survived the early mortality period. In other words, failure events with $\Delta T_F \leq \Delta T_{EM}$ are excluded from the point estimate calculation. Table 2 shows the sensitivity of f with respect to ΔT_{EM}, where ΔT_F is assigned values of either 0 years (i.e., all failures with missing information are excluded) or 19.5 years. Limited variation is observed due to changes in ΔT_F and ΔT_{EM} with all values within the 1E-4/year range. While subjectivity may be involved in choosing a specific value for ΔT_{EM}, it is clear that very high values for ΔT_{EM} will skew the estimates to potentially misleading results.

Table 2. Dam failure frequencies with varying ΔT_{EM} and ΔT_F = 0, 19.5 years

ΔT_{EM}	0	2	4	6	8	10
ΔT_F = 19.5 years	2.40E-4	2.13E-4	1.91E-4	1.79E-4	1.72E-4	1.66E-4
ΔT_F = 0 years	1.64E-4	1.37E-4	1.15E-4	1.03E-4	9.56E-5	8.97E-5

With a value of ΔT_{EM} = 8 years and ΔT_F = 19.5 years (with T_{CY} = 1963), a comparison between the major dam types considered in this analysis can be made. For all dams, a value of f_{ALL} = 1.72E-4/year is obtained, with corresponding results for embankment and concrete dams yielding, f_E = 1.69E-4/year and f_C = 1.48E-4/year, respectively. Therefore, small differences between dam types are observed for the results during the 1900 – 2006 period. The convergence of the values of f_{ALL}, f_E, and f_C, are shown in Figure 3, where f is calculated using the cumulative number of dam-years and failures in time for each major dam type. The value of f increases until approximately 1920 – 1930 as the number of dam-years and failures accumulates, when f begins to decrease, converging to the results indicated above.

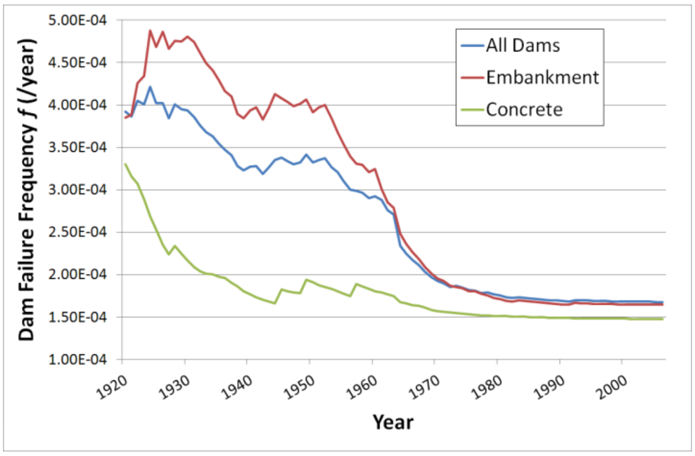


Figure 3. Convergence of dam failure frequency f (/year) for embankment, concrete, and all dams

Finally, the sensitivity to the selection of the large dam height criteria, H_{LARGE}, for all dam types is assessed by calculating f for a subset of increasing values as shown in Table 3. The value of f_{ALL} remains fairly constant as H_{LARGE} is increased from 12.2 meters [40 feet] to 61 meters [200 feet]. For values exceeding 76.2 meters [250 feet] (beyond which there are no reported failures), the value decreases to 8.31E-5/year. While there are no specific thresholds at which a distinction can be made in terms of susceptibility to failure modes for "large" versus "small" dams, care should be exercised in the choice of H_{LARGE} as selecting a high value will result in a significantly sparse subset of dam-years and failures as shown in Table 3. In fact, this would also apply to any attempt to parse the estimation of f with respect to a large number of attributes simultaneously, as this can lead to artificially low estimates. Finally, it should be recognized that all estimates calculated in this work are generic in nature and are, therefore, an approximation of the results that may be obtained by performing a more detailed probabilistic analysis for a specific dam, given that dams are very unique with respect to design and site characteristics.

Table 3. Sensitivity of *f* for all dam types with respect to large dam height criteria

LARGE DAM HEIGHT CRITERIA, H_{LARGE} (feet)					
≥ 40	≥ 50	≥ 100	≥ 200	≥ 250	
512,745	293,835	81,469	21,618	12,036	**DAM-YEARS**
96	74	22	4	1	**FAILURES**
1.87E-4	2.52E-4	2.70E-4	1.85E-4	8.31E-5	***f* (/year)**

2.4. Uncertainty Analysis

To address the limitations associated with the datasets and the uncertainty associated with classically derived statistical failure rates, an approach using a Bayesian framework is implemented (Kelly & Smith, 2011). A model based on the assumption that the occurrences of dam failure events follow a homogenous Poisson process with rate parameter λ, which is equal to the mean rate of events, is considered first. In this model, we address missing data using the same data assumptions used to derive the point estimates above. Next, we consider an exponential model that assumes the failure rate for a dam is constant over the life of the dam. In conjunction with the exponential model, we do not make any assumptions about the values of missing data. Instead, we treat observations with missing data as censored observations. We further describe these models below.

Utilizing the Poisson model, the number of dam failures events for a specified period of cumulative operating experience follows a Poisson distribution. In this paper, the conjugate Gamma prior distribution as well as a non-informative prior were considered for the parameter λ of the Poisson distribution. It is noted that the parameters of the prior distribution will be denoted with a subscript "1" and posterior parameters will use a subscript "2." Data derived from the NPDP and NID databases were used to obtain a posterior distribution based on the number of dam failure events observed and the cumulative number of observed dam-years using well-established analytical relationships for the conjugate pair. In this work, the cumulative years of operating experience is calculated only using dam-years for the dams that have not failed due to problems with repeated observations in both datasets (e.g., some failed dams appear in the NID database).

Table 4 provides a comparison of posterior mean failure frequencies (as well as 5^{th} and 95^{th} percentiles) obtained using the Gamma prior distribution with parameters $\nu_1 = 2589$ and $k_1 = 0.833$ for embankment and concrete dams, when varying the values of ΔT_{EM} and ΔT_F. The prior distribution parameters for the Gamma distribution were subjectively chosen because they yield a prior distribution with 5^{th} percentile corresponding to 1E-5/dam-year, a 95^{th} percentile corresponding to 1E-3/dam-year, and a mean consistent with the values obtained from the point estimate calculations. This is consistent with the statements in the addenda to the ASME/ANS RA-S-2008 Standard (2009) on the mean failure rate for all US dams with respect to external flooding hazard evaluations for nuclear power plant applications (ASME/ANS, 2007). As can be seen, the posterior mean values are consistent with the point estimates presented in Table 2. Furthermore, the 5^{th} and 95^{th} percentiles correspond to a relatively narrow spread around the mean, particularly for the cases in which data pertaining to embankment dams are used (which provides a larger dataset than the case utilizing data information on concrete dams).

For embankment dams, Figure 4 compares the prior and posterior distributions for a range of prior parameter values ν_1 and k_1 when $\Delta T_{EM} = 8$ was selected as the early mortality cut-off point, $\Delta T_F = 19.5$ was assigned to dams that have failed but for which the construction completion year is unknown, and the construction completion year 1963 was assigned to non-failed dams missing this information. In general, it was found that when considering the larger datasets (i.e. for the datasets containing data on all dams or embankment dams); the posterior distributions are relatively insensitive to the parameters of the prior distributions. For more finely parsed data (e.g. when considering data for concrete dams), it was found that the values of the prior distribution are relatively more influential.

Table 4. Posterior mean dam failure frequencies with varying ΔT_{EM} and ΔT_F = 0, 19.5 years for the Poisson-Gamma model with $\nu_1 = 2589$ and $k_1 = 0.833$ for embankment and concrete dams

		Embankment Dams						Concrete Dams					
ΔT_{EM}		0	2	4	6	8	10	0	2	4	6	8	10
ΔT_F = 19.5 years	μ_2	2.4E-4	2.1E-4	2.0E-4	1.8E-4	1.7E-4	1.7E-4	2.1E-4	1.9E-4	1.5E-4	1.5E-4	1.5E-4	1.4E-4
	5th	2.0E-4	1.8E-4	1.6E-4	1.5E-4	1.4E-4	1.3E-4	1.4E-4	1.2E-4	9.0E-5	9.0E-5	9.0E-5	8.1E-5
	95th	2.9E-4	2.5E-4	2.3E-4	2.2E-4	2.1E-4	2.0E-4	3.0E-4	2.7E-4	2.3E-4	2.3E-4	2.3E-4	2.1E-4
ΔT_F = 0 years	μ_2	1.7E-4	1.4E-4	1.3E-4	1.1E-4	1.0E-4	9.5E-5	1.7E-4	1.4E-4	1.1E-4	1.1E-4	1.1E-4	9.3E-5
	5th	1.4E-4	1.1E-4	9.7E-5	8.4E-5	7.5E-5	7.1E-5	9.9E-5	8.1E-5	5.5E-5	5.5E-5	5.5E-5	4.6E-5
	95th	2.1E-4	1.8E-4	1.6E-4	1.4E-4	1.3E-4	1.2E-4	2.4E-4	2.1E-4	1.7E-4	1.7E-4	1.7E-4	1.5E-4

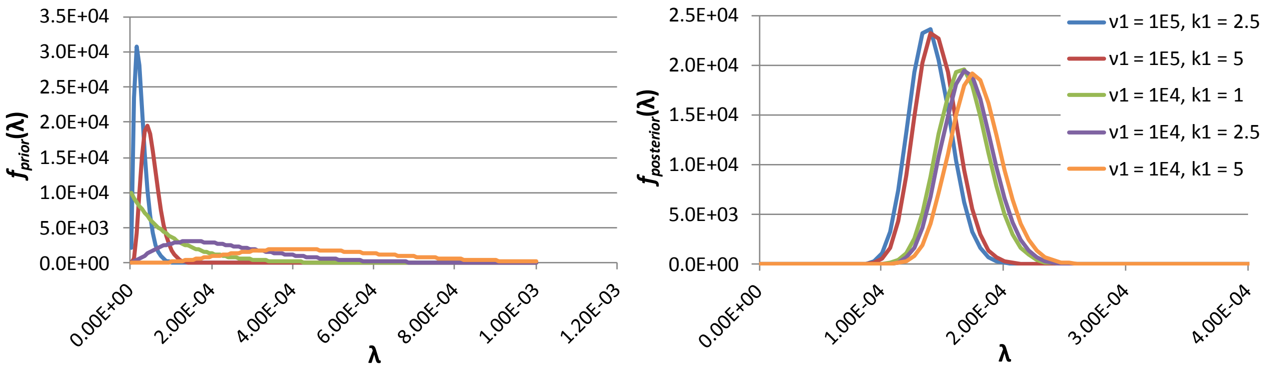


Figure 4. Prior distribution with parameters ν_1 and k_1 (left) and posterior distributions (right) when $\Delta T_F = 8$ and $\Delta T_{EM} = 19.5$ using data for embankment dams

Previously it was observed that parsing data by dam height could lead to erroneous results for point estimates when considering dams in excess of 61 meters [200 feet] due to the limited number of observations available. To address the potential uncertainty in this estimate, the Gamma prior distribution (with $\nu_1 = 2589$ and $k_1 = 0.833$) and non-informative prior were utilized to compute the posterior distribution of λ when considering all failure events (i.e. excluding no observations on the basis of early mortality) and assigning $\Delta T_F = 0$ for all failed dams missing the construction completion year (comparable to the point estimates presented in Table 2). The results are shown in Table 5. Results obtained using the informative prior are fairly consistent with the point estimate provided in Table 2. When using the non-informative prior, results are less consistent with the point estimates for the larger dam heights.

Table 5. Posterior mean dam failure frequencies for all dam types with varying dam heights

	Gamma Prior ($\nu_1 = 2589, k_1 = 0.833$)					Non-Informative Prior				
Height Range (ft)	**40-100**	**100-200**	**200-300**	**300-400**	**400-800**	**40-100**	**100-200**	**200-300**	**300-400**	**400-800**
μ_2	2.4E-4	4.1E-4	5.4E-4	2.5E-4	3.2E-4	2.4E-4	4.3E-4	6.2E-4	3.2E-4	4.7E-4
5th %	2.0E-4	2.9E-4	2.8E-4	4.0E-5	5.0E-5	2.0E-4	3.0E-4	3.2E-4	3.7E-5	5.5E-5
95th %	2.8E-4	5.6E-4	8.7E-4	6.1E-4	7.7E-4	2.8E-4	5.7E-4	1.0E-3	8.3E-4	1.2E-3

To understand the effect of assumptions made about missing observations in the context of a Bayesian assessment, we utilize an exponential model and consider missing observations as censored. Of course, the exponential model is directly related to the Poisson model used above. The exponential model is updated based on observations of individual component life-spans. The Gamma distribution is employed in this paper as the prior distribution on the parameter λ (equal to the mean rate of events) of the exponential model. There are multiple types of "life-span observations" available based on the NPDP and NID datasets. Dams that have failed and have known construction completion and failure dates provide direct information about their known lifespan. Dams that have not failed and have known construction completion dates provide information that the lifespan of the dam is at least equal to the difference between the year for which the most recent information is available (i.e. 2006 in this paper) and the construction completion year (i.e. they provide lowerbound observations). However, as described above, the NPDP and NID databases are missing construction completion dates for some dams. For failed dams missing this information, it is known that the lifespan of the dam is no more than x years, where x is equal to the year in which the dam failure event occurred minus a reference year that bounds the potential year of construction (assumed to be 1900 in this paper). We refer to these as upperbound observations to indicate that the lifespan of the dam is less than or equal to x years. For dams that have not failed and for which we do not have the construction year, the observations are assumed to be bounded at the lower-end by zero.

To compute posterior distributions using this model in conjunction with the censored observations, we utilize the WinBUGS software (Lunn et al, 2000), which uses Markov chain Monte Carlo (MCMC) methods to compute posterior distributions. Table 6 provides the posterior means (and 5th and 95th percentiles) for embankment and concrete dams when considering all dam failure events. The parameters of the Gamma prior distribution are once again $\nu_1 = 2589$ and $k_1 = 0.8333$. For comparison, Table 6 also provides the values obtained using the Poisson-Gamma model as well as the point estimate (with assumed values for missing data). In general, it is seen that the results obtained when considering observations as censored are fairly consistent with the results obtained by assuming values for missing data.

Table 6. Posterior mean dam failure frequencies using exponential-Gamma model with censoring and Poisson-Gamma model as well as point estimate

	Embankment Dams			**Concrete Dams**		
	Exponential	**Poisson**	**Point Estimate**	**Exponential**	**Poisson**	**Point Estimate**
μ_2 (or pt est.)	2.87E-4	2.76E-4	2.76E-4	2.40E-4	2.49E-4	2.46E-4
5th	2.46E-4	2.34E-4	−	1.54E-4	1.66E-4	−
95th	3.33E-4	3.20E-4	−	3.37E-4	3.44E-4	−
MC error	4.60E-7	−	−	9.01E-7	−	−

In general, the results of the Bayesian assessments are consistent with the point estimates. When working with the larger datasets, the effects of the prior distribution parameter assumptions are minimal. However, when the data is parsed into smaller subsets, the influence of the prior becomes more significant. Overall, all estimates are in the range of 1E-4/dam-year regardless of the method used to derive the estimate of dam failure frequency.

4. CONCLUSION

Sensitivity studies on the dam failure frequency for US large dams were performed in this study to evaluate the impact of various attributes and sources of uncertainty when using historical dam information. Bayesian analysis tools were also used for the derivation of posterior uncertainty distributions that include subjective information such as data quality and expert judgment considerations. The extent of the variation in the commonly derived point estimate was documented and discussed for a number of categories and assumptions. It is stressed that the goal of this work is solely to develop generic dam failure frequencies based on information contained in databases and readily available historical records. As such, it is not a replacement for more detailed probabilistic assessments and/or dam-specific studies (which could yield higher or lower failure frequency estimates). Although historical dam failure information can provide useful qualitative insights on the general performance and failure modes for certain categories, its applicability to specific dams has to be assessed to establish sufficient technical bases for decision-making. This is due to the variability in site-specific characteristics (e.g., hydrologic, geologic, and operational) and the potential contributions of site-specific failure modes. Despite the limitations of working with data-driven estimates of dam failure frequencies, this work provides insights into the variability and subjectivity of the estimates and their sensitivity to input information (particularly historical dam failure accounts). The series of assessments performed in this paper generally support dam failure frequencies in the range of 1E-4/dam-year, though it is shown that variability exists based on the assumptions utilized relative to parsing data and addressing missing observations.

Acknowledgements

The authors would like to thank Sunil Weerakkody, Branch Chief, PRA Operational Support Branch, Division of Risk Assessment, Office of Nuclear Reactor Regulation; and Benjamin Beasley, Branch Chief, Operating Experience and Generic Issues Branch, Division of Risk Analysis, Office of Nuclear Regulatory Research, for their support and feedback during the development of this paper.

References

Ferrante, F., Sancaktar, S., Mitman, J., and Wood, J. (2011) *An Assessment Of Large Dam Failure Frequencies Based On US Historical Data.* ANS PSA 2011 International Topical Meeting on Probabilistic Safety Assessment and Analysis, March 13-17, 2011, Wilmington, NC, USA.

FERC (2005) *Dam Safety Performance Monitoring Program*, Chapter 14, Engineering Guidelines, Federal Energy Regulatory Commission, Washington DC, USA.

Kelly, D., and Smith, C. (2011) *Bayesian Inference for Probabilistic Risk Assessment: A Practitioner's Guidebook.* 1st ed. Springer, USA.

Lunn, D.J., Thomas, A., Best, N., and Spiegelhalter, D. (2000) WinBUGS − a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing, 10:325−337

Stanford University (2007) *National Performance of Dams Program.* Department of Civil and Environmental Engineering, http://npdp.stanford.edu.

USACE (1979) *Recommended Guidelines for Safety Inspection of Dams*. Engineer Technical Letter No. ER 1110-2-106. US Army Corps of Engineers, Washington DC, USA.

USACE (2006) *Reliability Analysis and Risk Assessment For Seepage And Slope Stability Failure Modes For Embankment Dams*. Engineer Technical Letter No. 1110-2-561, US Army Corps of Engineers, Washington DC, USA.

USACE (2008) *National Inventory of Dams, Methodology*. Version 4.0. US Army Corps of Engineers, Washington DC, USA.

USACE (2011) *National Inventory of Dams.* US Army Corps of Engineers, http://nid.usace.army.mil.

USBR (2010) *Dam Safety Risk Analysis Best Practices Training Manual.* US Department of Interior, Bureau of Reclamation, Denver, CO, USA.

WCD (2000) *Dams and Development: A New Framework for Decision-Making*. The Report of the World Commission on Dams, World Commission on Dams, London, UK.

Kocahan, H., and Taylor, G. (2002) *Rehabilitation of Black Rock Dam "Seepage & Inadequate Spillway"*. Association of State Dam Safety Officials Conference, Arlington, VA, USA.

Singh, V.P. (2010) *Dam Breach Modeling Technology*. 2nd ed. Kluwer Academic Publishers, Netherlands.

ASME/ANS (2009) *Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications.* American Society of Mechanical Engineers/American Nuclear Society, New York, NY, USA.