JOURNAL OF THE AMERICAN WATER RESOURCES ASSOCIATION

AMERICAN WATER RESOURCES ASSOCIATION

Development of an Earthen Dam Break Database

K. Hood, R.A. Perez, H.E. Cieplinski, T.V. Hromadka II, G.E. Moglen, and H.D. McInvale

Research Impact Statement: Risk assessment for earthen dams is improved by an aggregated dam break database and associated web-based regression selection tool, assisting planners regarding operational and safety decisions.

ABSTRACT: Earthen embankment dams comprise 85% of all major operational dams in the United States. Assessment of peak flow rates for these earthen dams and the impacts on dam failure are of high interest to engineers and planners. Regression analysis is a frequently used risk assessment approach for earthen dams. In this paper, we present a decision support tool for assessing the applicability of nine regression equations commonly used by practitioners. Using data from 108 case studies, six parameters were observed to be significant factors predicting for peak flow as a metric for risk analysis. We present our work on an expanded earthen dam break database that relates the regression equations and underlying data. A web application, regression selection tool, is also presented to assess the appropriateness of a given model for a given test point. This graphical display allows users to visualize how their data point compares with the data used for the regression equation. These contributions improve estimates and better inform decision makers regarding operational and safety decisions.

(KEYWORDS: earthen dam failure; regression selection tool; web application; earthen dam database; watersheds; hydraulic structures; data management; computational methods; risk assessment.)

INTRODUCTION

Earthen dams and reservoirs are frequently used for flood control and for storage of water supply. They also serve to trap sediment and debris, among other purposes. There are over 70,000 dams documented in the United States (U.S.), and approximately 85% of all major operational dams in the U.S. are earthen embankment dams (Billington and Jackson 2017). Earth filled dams are composed primarily of compacted fine-grained material and are therefore subject to erosional processes should failure occur. A topic of high interest among engineers and planners is the assessment of possible failure of these earthen dams and the possible range of inundation areas, peak flow rates, peak flow velocities, among other factors that are relevant in the assessment of flood inundation damages and risk assessment (Wahl 1998). To develop such a risk assessment, a failure scenario is typically assumed of the earth dam and then a hydrograph of the flow discharge through the earthen dam breach developed, and the released flow hydrologically or hydraulically routed downstream. Descriptions of this process can be found in the National Dam Safety Review Board Steering Committee on Dam Breach Equations (Pierce et al. 2010) along with other publications. The failure of the earthen dam is

---

Paper No. JAWRA-17-0127-P of the *Journal of the American Water Resources Association* (JAWRA). Received September 24, 2017; accepted October 24, 2018. **Discussions are open until six months from issue publication**.

Department of Mathematical Sciences (Hood, Hromadka, McInvale), United States Military Academy, West Point, New York, USA; Hromadka & Associates (Perez), Riverside, California, USA; Hromadka & Associates (Cieplinski), Glenside, Pennsylvania, USA; and Hydrology and Remote Sensing Laboratory (Moglen), USDA Agricultural Research Service, Beltsville, Maryland, USA (Correspondence to Hood: hood.karoline@gmail.com).

Citation: Hood, K., R.A. Perez, H.E. Cieplinski, T.V. Hromadka II, G.E. Moglen, and H.D. McInvale. 2018. "Development of an Earthen Dam Break Database." *Journal of the American Water Resources Association* 1–13. https://doi.org/10.1111/1752-1688.12703.

typically assessed by a coupled hydrologic and hydraulic computational model of the dam breach itself with a set of assumed erosion characteristics used to evolve the dam breach geometry over time (Singh and Snorrason 1984; Wahl 1998).

In addition, regression equations are often used to estimate some of the key outcome variables of the dam breach process (Froehlich 1995; Xu and Zhang 2009; Pierce et al. 2010; Wahl 2014). These regression equations not only may provide estimates of ultimate breach geometry dimensions but also estimates of release flow characteristics such as peak flow rate, total stored volume, and other factors. The regression equations are based upon case studies of earthen dam breach occurrences using field-measured data. Few equations exist that consider a significant proportion of the available earthen dam breach cases reported in the literature. In a subsequent section of the current paper, a review is made of some of the more frequently referenced assemblies of available earthen dam breach case studies used for development of regression equations.

These regression equations can differ in their predicted outcome variable values (Wahl 2014). An explanation for these differences is the differences in assembled datasets used to develop the equations (Pierce et al. 2010). In the earthen dam break case, the study reported data assembled in the tabulations to indicate which of the regression equations included particular data points in their reported development of their respective regression equations. That is, which reported data point was included to develop which regression equations. Some of the reported data are used in more than one regression equation effort.

At issue is the “appropriate” regression equation to be used for a particular situation. Furthermore, at issue is whether any of the regression equations are appropriate to be used as being representative of the test situation under analysis. In the typical application of these equations in an engineering assessment, the case study data that form the underpinnings of a particular regression equation are unfortunately seldom examined by the user as to whether or not the test case is well within the population of the case study data used to develop the selected regression equation. Previous papers have not addressed the issue of appropriateness of selected regression equation in relation to a test case.

The main purpose of the current paper was to assemble these various reported case study datasets for convenient reference, and to present a web application that will help to assess the test case situation within the population of the case study data that form the underpinnings of the selected regression equation.

THE ASSEMBLED DATABASE

Several sources of earthen dam break data were examined in the current study and are all included in the assembled database. These sources include reports from the U.S. Department of the Interior Bureau of Reclamation Dam Safety Office, articles published in the Journal of Geotechnical and Geoenvironmental Engineering, and the Journal of Hydraulic Engineering, among other journals and texts, and reports submitted to the National Dam Safety Review Board. Seven publications formed the basis for the assembled earthen dam break data. The key selected publications are found in Table 1.

Each of these publications contained approximately 75–125 data points. The often referenced database of Wahl (1998) assembled 108 case studies. There were variations in parameters between publications, but six common parameters were observed most frequently as factors of the response variable, peak flow. The six common parameters are: volume stored above breach invert (V_w), dam height (H_d), depth above breach (H_w), reservoir storage (S), length (L), and average width (W_{ave}). All parameters are observed in the metric system.

We identified 25 parameters in our integrated database. The different parameters in the current database can be subdivided into four subcategories: embankment dimensions (i.e., height, width, length, etc.), hydraulic characteristics (i.e., storage, surface area, and depth), breach characteristics, and time parameters (i.e., formation time, failure time, etc.). Of the 25 parameters identified, only six parameters are observed being used in the published regression

TABLE 1. Publications used to assemble the earthen dam break database.

Authors	Journal	Year
Dr. D. Froehlich	*Journal of Water Resources Planning and Management*	1995
Dr. M. Pierce et al.	*Journal of Hydrologic Engineering*	2010
Krishan P. Singh and Arni Snorrason	*Journal of Hydrology*	1984
Dr. Thornton et al.	*ASCE Journal of Hydrologic Engineering*	2011
Dr. T Wahl	*U.S. Department of the Interior Bureau of Reclamation*	1998
Dr. T. Wahl	*U.S. Department of the Interior Bureau of Reclamation*	2014
Dr. Y. Xu and Dr. L. M. Zhang	*Journal of Geotechnical and Geoenvironmental Engineering*	2009

Note: U.S., United States.

TABLE 2. Parameters and dams used in nine published regression equations.

Dam and location		Embankment dimensions				Hydraulic characteristics			Regression equation								
		Dam height H_d (m)	Average width W (m)	Length L (m)	Peak outflow Q_p (m^3/s)	Reservoir storage S (m^3)	Volume stored above breach invert V_w (m^3)	Depth above breach H_w(m)	1.	2.	3.	4.	5.	6.	7.	8.	9.
1	Apishapa, Colorado	34.14	82.4	28	6,850.00	22,500,000	22,200,000	28	*	*	*	*	*				*
2	Baimiku, China	8			—	200,000	200,000	8									
2	Baldwin Hills, California	71	59.6	12.2	1,130.00	1,100,000	910,000	12.2	*	*	*	*	*	*			*
3	Banqiao, China	24.5	—	—	78,100	492,000,000	607,500,000	31						*			
3	Bayi, China	30			5,000	30,000,000	23,000,000	28									
4	Bearwallow Lake, North Carolina		14				49,300	5.79									
4	Big Bay Dam, USA	15.6	—	—	4,160	17,500,000	17,500,000	13.5						*			
5	Bradfield, England	28.96	50	382	1,150.00	3,200,000											
5	Break Neck Run, USA	7	86	7	9.2	49,000	—	—					*				
6	Buckhaven No. 2, Tennessee		37				24,700	6.1									
6	Buffalo Creek, West Virginia	14.02	128	14.02	1,420.00	484,000	484,000	14.02		*	*	*	*				*
7	Bullock Draw Dike, Utah	5.79	18.6			1,130,000	740,000	3.05									
7	Butler, Arizona	—	9.63	7.16	810	—	2,380,000	7.16	*			*	*				
8	Canyon Lake, USA	6.1		152		985,000											
8	Castlewood, Colorado	21.34	47.4	21.6	3,570.00	4,230,000	6,170,000	21.6	*	*	*	*	*				*
9	Chenying, China	12			1,200	4,250,000	5,000,000	12									
9	Caulk Lake, Kentucky		32				698,000	11.1									
10	Cheaha Creek, USA	7.01				69,000											
10	Clearwater Lake Dam, Georgia		15				466,000	4.05									
11	Coedty, England	10.97		262		310,000	311,000	11									
11	Cougar Creek, Alberta		21.7				29,800	11.1									
12	Dalizhuang, China	12			—	600,000	600,000	12									
12	Danghe, China	46			2,500	15,600,000	10,700,000	24.5									
9	Davis Reservoir, California	11.89		—	510	58,000,000	58,000,000	11.58		*	*						
13	Dells, USA	18.3			5,440	13,000,000	13,000,000	18.3									
10	Dongchuankou, China	31			21,000	27,000,000	27,000,000	31									
14	Dushan, China	17.7			—	670,000	670,000	17.7									
11	DMAD, Utah	8.8			793	19,700,000	19,700,000										
15	East Fork Pond River, Kentucky		38.9				1,870,000	9.8									
12	Elk City, Oklahoma	9.14	50.4	564		740,000	1,180,000	9.44									
16	Emery, California		22.2				425,000	6.55									
13	Erindale, Canada	10.67		213													
17	Erlangmiao, China	12.1			—	196,000	196,000	9									

(continued)

TABLE 2. (continued)

		Embankment dimensions				Hydraulic characteristics			Regression equation								
	Dam and location	Dam height H_d (m)	Average width W (m)	Length L (m)	Peak outflow Q_p (m^3/s)	Reservoir storage S (m^3)	Volume stored above breach invert V_w (m^3)	Depth above breach H_w(m)	1.	2.	3.	4.	5.	6.	7.	8.	9.
14	Euclides de Cunha, Brazil	53.04			1,020.00	13,600,000		58.22									
18	Fogelman, Tennessee		21.3				493,000	11.1									
15	Frankfurt, Germany	9.75			79	350,000	352,000	8.23									
10	Fred Burr, Montana	10.4	30.8	10.2	654	752,000	750,000	10.2	*			*	*				*
11	French Landing, Michigan	12.19	34.3	8.53	929	—	3,870,000	8.53	*	*	*	*	*				*
12	Frenchman Creek, Montana	12.5	37.3	10.8	1,420.00	21,000,000	16,000,000	10.8	*	*	*	*	*				
12	Fengzhuang, China	10			—	625,000	625,000	8									
13	Frenchman Dam, USA	12.5			1,420	21,000,000	16,000,000	10.8									
13	Frias, Argentina			62.2													
13	Goose Creek, South Carolina	6.1	—	—	565	10,600,000	10,600,000	1.37		*	*						
14	Gouhou, China	71			2,050	3,300,000	3,180,000	44									
14	Granite Creek, Alaska	—	—	—	1,841.00	—	—	—									*
15	Haas Pond, Connecticut		16.7				23,400	2.99									
15	Hart, Michigan		31.1				6,350,000	10.7									
15	Hatchtown, Utah	19.2	44.8	16.8	3,080.00	14,800,000	14,800,000	16.8	*	*	*	*	*	*			*
16	Hatfield, USA	6.8			3,400.00	12,300,000											
16	Hebron, USA	11.58						12.19									
16	Hell Hole, California	67.06	103.2	35.1	7,360.00	30,600,000	30,600,000	35.1	*	*	*	*	*				*
17	Herrin, Illinois		28.8					10.7									
17	Horse Creek, Colorado	12.19	26.8	701		21,000,000	12,800,000	7.01									
18	Hougou, China	8			—	240,000	240,000	8									
18	Huoshishan, China	13			—	220,000	220,000	16									
19	Huqitang, China	9.9			50	734,000	424,000	5.1									
19	Hutchinson Lake Dam, Georgia		14				1,170,000	4.42									
20	Iowa Beef Processors, Washington	4.57		305		333,000	333,000	4.42									
17	Ireland No. 5, Colorado	—	18	3.81	110	—	160,000	3.81	*			*	*				
21	Jiahezi, China	18			—	80,000,000	42,000,000	12									
18	Jacobs Creek, Pennsylvania						423,000	20.1									
22	Johnston City, Illinois	4.27	21.5			575,000	575,000	3.05									
18	Johnstown (South Fork Dam, Pennsylvania)	38.1	64	24.6	8,500.00	18,900,000	18,900,000	24.6		*	*	*	*	*			

(continued)

TABLE 2. (continued)

Dam and location	Embankment dimensions				Hydraulic characteristics			Regression equation								
	Dam height H_d (m)	Average width W (m)	Length L (m)	Peak outflow Q_p (m^3/s)	Reservoir storage S (m^3)	Volume stored above breach invert V_w (m^3)	Depth above breach H_w(m)	1.	2.	3.	4.	5.	6.	7.	8.	9.
23 Kaddam, India	12.5				214,000,000											
19 Kelly Barnes, Georgia	11.58	19.4	11.3	680	505,000	777,000	11.3	*	*	*	*	*				*
24 Kodaganar, India	11.5			1,280	12,300,000	12,300,000	11.5									
20 Kendall Lake Dam, South Carolina	5.49		128		728,000											
25 Kraftsmen's Lake Dam, Georgia		8.1				177,000	3.66									
21 La Fruta, Texas		40				78,900,000	7.9									
20 Lake Avalon, New Mexico	14.5	42.7	13.7	2,320.00	7,750,000	31,500,000	13.7				*	*				
22 Lake Barcroft, USA	21.03				3,120,000											
21 Lake Frances, California	15.24	47.4			865,000	789,000	14									
23 Lake Genevieve, Kentucky		19.8				680,000	6.71									
21 Lake Latonka, Pennsylvania	13	28	6.25	290	1,590,000	4,090,000	6.25				*	*				
24 Lake Philema Dam, Georgia		28				4,780,000	9									
22 Lambert Lake, Tennessee		53.9				296,000	12.8									
22 Laurel Run, Pennsylvania	12.8	40.5	14.1	1,050.00	385,000	555,000	14.1	*	*	*	*	*				*
23 Lawn Lake, Colorado	7.9	14.2	6.71	510		798,000	6.71				*	*				
23 Lijiaju, China	25			2,950	1,140,000	1,140,000	25									
24 Lily Lake, Colorado	—	—	—	71	—	92,500	3.35	*								
25 Little Deer Creek, Utah	26.21	63.1	22.9	1,330.00	1,730,000	1,360,000	22.9	*	*	*	*	*				*
25 Liujiatai, China	35.9			28,000	40,540,000	40,540,000	35.9									
26 Longtun, China	9.5			—	30,000,000	30,000,000	9.5									
26 Long Branch Canyon, California		11.3				284,000	3.17									
26 Lower Latham, Colorado	—	25.7	5.79	340	7,080,000	7,080,000	5.79	*			*	*				
27 Lower Otay, California	41.15	53.3	172		49,300,000	49,300,000	39.6									
27 Lower Two Medicine, Montana	11.28	—	—	1,800.00	19,600,000	29,600,000	11.3	*	*	*						*
28 Lyman, Arizona	19.81				49,500,000	35,800,000	16.2									
28 Lynde Brook, Massachusetts	12.5	41.8			2,520,000	2,880,000	11.6									
29 Mahe, China	19.5			4,950	23,400,000	23,400,000	19.5									
29 Machhu II, India	60.05		4,180		110,000,000											
30 Mammoth, USA	21.3			2,520.00	13,600,000											
30 Martin Cooling Pond Dike, Florida				3,115.00	136,000,000	136,000,000	8.53									
31 Melville, Utah	10.97	25.1				24,700,000	7.92									
31 Merimac (Upper) Lake Dam, Georgia		17.5				69,600	3.44									
32 Mill River, Massachusetts	13.1			1,645.00	2,500,000	2,500,000										
32 Mossy Lake Dam, Georgia		14.3				4,130,000	4.41									
33 Niujiaoyu, China	10			—	160,000	144,000	7.2									
33 Nanaksagar, India	15.85			9,700.00	210,000,000											
34 Nahzille, New Mexico	5.49		130													
34 Noppikoski, SE	18.5				700,000	1,000,000										
28 North Branch Tributary, Pennsylvania	5.5			29.4		22,200	5.49		*	*						
35 Oakford Park, USA	6.1		107													
29 Oros, Brazil	35.36	110	35.8	9,630.00	650,000,000	660,000,000	35.8	*	*	*	*	*				*

(continued)

TABLE 2. (continued)

Dam and location	Embankment dimensions				Hydraulic characteristics			Regression equation								
	Dam height H_d (m)	Average width W (m)	Length L (m)	Peak outflow Q_p (m^3/s)	Reservoir storage S (m^3)	Volume stored above breach invert V_w (m^3)	Depth above breach H_w(m)	1.	2.	3.	4.	5.	6.	7.	8.	9.
36 Otter Lake, Tennessee		20.6				109,000	5									
30 Otto Run, USA	5.8	—	—	60	—	7,400	5.79		*	*						
37 Pierce Reservoir, Wyoming						4,070,000	8.08									
31 Potato Hill Lake, North Carolina		23.5				105,000	7.77									
31 Prospect, Colorado	—	13.1	1.68	116	—	3,540,000	1.68	*			*	*				
32 Puddingstone, California	—	—	—	480	—	617,000	15.2	*								
32 Qielinggou, China	18			2,000	700,000	700,000	18									
33 Quail Creek, Utah	—	56.6	30.5	3,110.00	—	30,800,000	16.7	*			*	*				
33 Rainbow Lake, Michigan		28.2				6,780,000	10									
34 Renegade Resort Lake, Tennessee		11				13,900	3.66									
34 Rito Manzanares, New Mexico	7.32	13.3			24,700	24,700	4.57									
35 Salles Oliveira, Brazil	35.05			7,200.00	25,900,000	71,500,000	38.4									
34 Sandy Run, Pennsylvania	8.53	—	—	435	56,800	56,700	8.53		*	*						*
35 Schaeffer, Colorado	30.5	80.8	335	4,500.00	3,920,000	4,440,000	30.5		*	*	*	*	*			*
35 Shangliuzhuang, China	14			—	110,000	110,000	14									
36 Shanhu, China	11.5			—	2,150,000	1,780,000	12.5									
36 Scott Farm Dam No. 2, Alberta		39.3				86,000	10.4									
37 Sheep Creek, USA	17.07				1,430,000	2,910,000	14.02									
37 Shilongshan, China	14			—	2,060,000	2,060,000	14									
36 Shimantan, China	25	—	—	30,000	94,400,000	117,000,000	27.4				*		*			
38 Sherburne, USA	10.36		91.4	960	42,000											
37 Sinker Creek, USA	21.34				3,330,000	3,330,000	21.34									
37 South Fork Tributary, Pennsylvania	1.8	—	—	122	—	3,700	1.83	*	*	*						*
38 Spring Lake, Rhode Island	5.49				135,000	136,000	5.49									
38 Statham Lake Dam, Georgia		12.6	—		—	564,000	5.55									
38 Swift, Montana	57.61	—	226	24,947.00	37,000,000	37,000,000	47.85		*	*	*		*			*
39 Taum Sauk, USA	—	—	—	7,743	5,390,000	5,390,000	—				*		*			
40 Teton, Idaho	92.96	250	77.4	65,120.00	356,000,000	310,000,000	77.4	*	*	*	*	*				*
40 Tongshuyuan, China	13			—	400,000	400,000	10									
41 Tiemusi, China	12			—	110,000	110,000	12									
41 Timber Lake, USA	9.3				—	1,800,000	7.33									
42 Trial Lake, Utah		7.62				1,480,000	5.18									
42 Trout Lake, North Carolina		21.6				493,000	8.53									
43 Upper Pond, Connecticut						222,000	5.18									
43 Wanshangang, China	13			—	1,500,000	1,500,000	12									
44 Wheatland No. 1, Wyoming	13.6				11,500,000	11,600,000	12.2									

(continued)

TABLE 2. (continued)

	Embankment dimensions				Hydraulic characteristics			Regression equation								
Dam and location	Dam height H_d (m)	Average width W (m)	Length L (m)	Peak outflow Q_p (m^3/s)	Reservoir storage S (m^3)	Volume stored above breach invert V_w (m^3)	Depth above breach H_w(m)	1.	2.	3.	4.	5.	6.	7.	8.	9.
44 Wilkinson Lake Dam, Georgia		13.2				533,000	3.57									
45 Winston, North Carolina	7.32	7.76	133		664,000	662,000	6.4									
45 Yuanmen, China	19.2			—	6,400,000	6,400,000	19.2									
46 Zhonghuaju, China	16			—	140,000	140,000	16									
46 Zhugou, China	23.5			11,200	15,400,000	18,430,000	23.5									
47 Zuocun, China	35			23,600	40,000,000	40,000,000	35									
41 USDA-ARS — Test #1, Okla	2.29	—	7.3	7	—	4,900	—				*		*			
42 USDA-ARS — Test #3, Okla	2.29	—	7.3	2	—	4,900	—				*		*			
43 USDA-ARS — Test #4, Okla	1.5	—	4.9	2	—	5,090	—				*		*			
44 USDA-ARS — Test #6, Okla	1.5	—	4.9	1	—	5,190	—				*		*			
45 USDA-ARS — Test #7, Okla	2.13	—	12	4	—	4,770	—				*		*			
46 Pierce Lake Dam	14	—	—	864	3,280,000	—	—							*	*	
47 Lake in the Hills Dam No. 1	12.2	—	—	238	740,000	—	—							*	*	
48 Lake in the Hills Dam No. 2	4.4	—	—	321	100,000	—	—							*	*	
49 Lake Marian Dam	15.2	—	—	90	190,000	—	—							*	*	
50 Clinton Lake Dam	19.8	—	—	4,254	91,540,000	—	—							*	*	
51 Lake Springfield Dam	14.6	—	—	3,437	66,000,000	—	—							*	*	
52 Weslake Dam	14.6	—	—	35	2,800,000	—	—							*	*	
53 Kinkaid Lake Dam	28	—	—	2,011	96,840,000	—	—							*	*	
54 Austin, Texas	21	—	—	6,683	47,070,000	—	—									*
55 St. Francis, California	56	—	—	11,327–22,653	—	—	—									*
								22	23	23	25	25	14	8	8	21

Note: Associated regression Equations (1–9) can be found in the manuscript and Table 3. Dashes indicate no information available. Asterisks indicate dam locations. Number of data points is tallied at the bottom-right.

TABLE 3. Nine published regression equations for peak flow that were cataloged in database and analyzed in the base application.

Equation number	Model name	Year	Equation
1.	Froehlich	1995	$Q_p = 0.607(V_w^{0.295} \cdot H_w^{1.24})$
2.	MacDonald and Langridge-Monopolis	1984	$Q_p = 3.85(V_w \cdot H_w)^{0.411}$
3.	MacDonald and Langridge-Monopolis	1984	$Q_p = 1.154(V_w \cdot H_w)^{0.412}$
4.	Pierce et al.	2010	$Q_p = 0.1202(L)^{1.7856}$
5.	Pierce et al.	2010	$Q_p = 0.863(V^{0.335} \cdot H^{1.833} \cdot W_{ave}^{-0.663})$
6.	Pierce et al.	2010	$Q_p = 0.012(V^{0.493} \cdot H^{1.205} \cdot L^{0.226})$
7.	Singh and Snorrason	1982	$Q_p = 13.4(H_d)^{1.89}$
8.	Singh and Snorrason	1982	$Q_p = 1.776(S)^{0.47}$
9.	U.S. Bureau of Reclamation	1982	$Q_p = 19.1(H_w)^{1.85}$

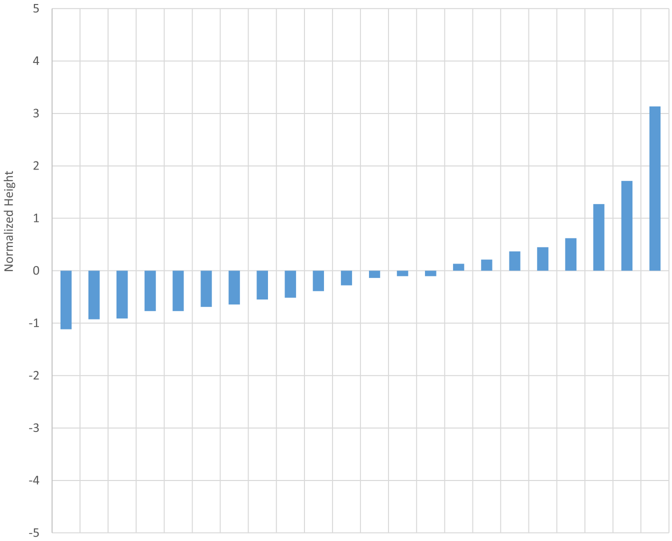


FIGURE 1. Twenty-one normalized data points of depth above breach used in the U.S. Bureau of Reclamation (1982) regression equation ($Q_p = 19.1(H_w)^{1.85}$).

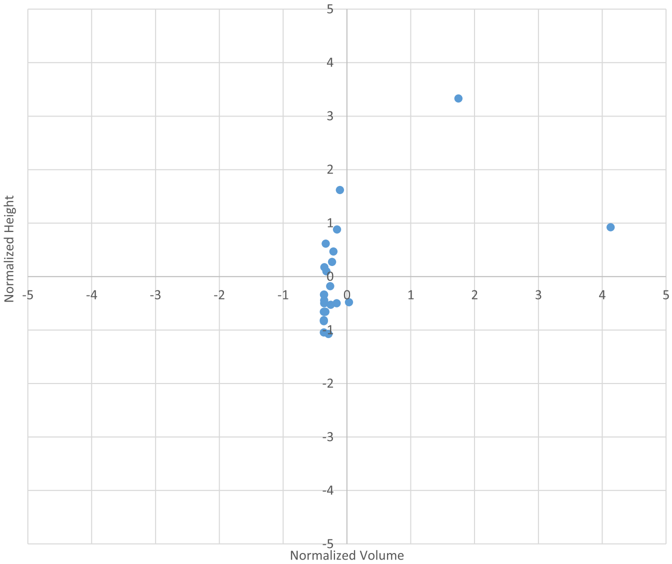


FIGURE 2. Twenty-three normalized data points of height and volume stored after breach invert used in two MacDonald and Langridge-Monopolis (1984) regression equations ($Q_p = 1.154 (V_w \cdot H_w)^{0.412}$) and ($Q_p = 3.85(V_w \cdot H_w)^{0.411}$).

equations for estimating released peak flow rates. In addition, of the 163 dams that were cataloged only 55 of the dams were used as data points to develop a regression equation. The condensed database reflecting only the parameters and dams that were used in the published regression equations can be found in Table 2 and in the web application at http://usmathematics.com/PastedGraphic-1.pdf and http://usmathematics.com/PastedGraphic-2.pdf.

Earthen dam failures are typically predicted with use of analytical equations, regression relations from numerical and analytical models, and regression relationships from laboratory tests. Most breach parameter equations stem from data developed from actual earthen dam failures. Within the assembled database, these dams range in time of construction from 1893 to 1986, and had observed modes of failure identified due to overtopping, seepage, piping, and sliding. Unfortunately, not all the reported earthen dam failure cases considered have complete information for the parameters needed for all the associated regression equations.

PUBLISHED REGRESSION EQUATIONS EXAMINED

There are many equations relating to dam failure but those for peak flow are the simplest with parameters most commonly found in the database. In addition, other categories that were commonly observed were failure time equations and breach geometry. It is noted that the equations used to estimate peak flow are often simple regression equations that relate

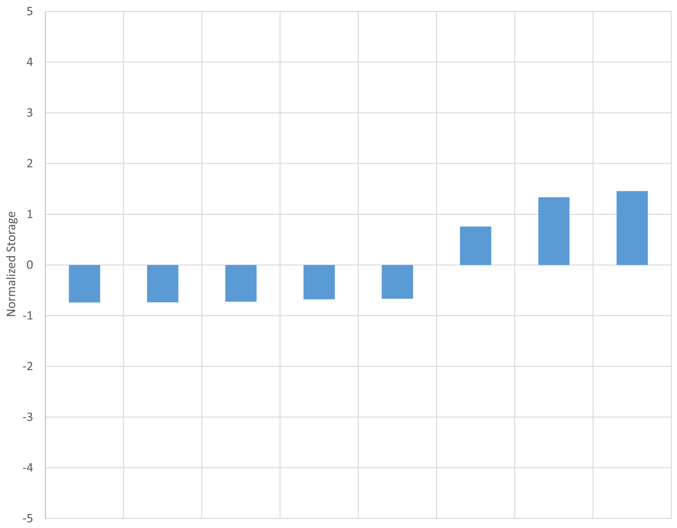


FIGURE 3. Eight normalized data points of reservoir storage used in Singh and Snorrason (1984) regression equation ($Q_p = 1.776(S)^{0.47}$).

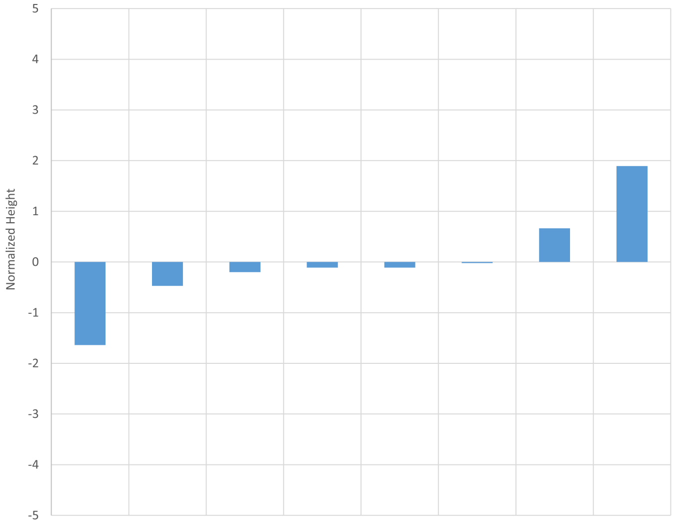


FIGURE 4. Eight normalized data points for dam height used in Singh and Snorrason (1984) regression equation ($Q_p = 13.4(H_d)^{1.89}$).

peak flow to volume of water behind the dam or the product of depth and volume (Pierce et al. 2010).

The regression equations we analyzed are shown in Table 3. All of the equations are associated with peak flow rate estimation. We were able to take the analyzed regression equations and find the data on the failed dams that were used to develop the regression equations. We included the number of original data points to develop the regression equation into our database. Future work will analyze failure time equations, and breach width equations.

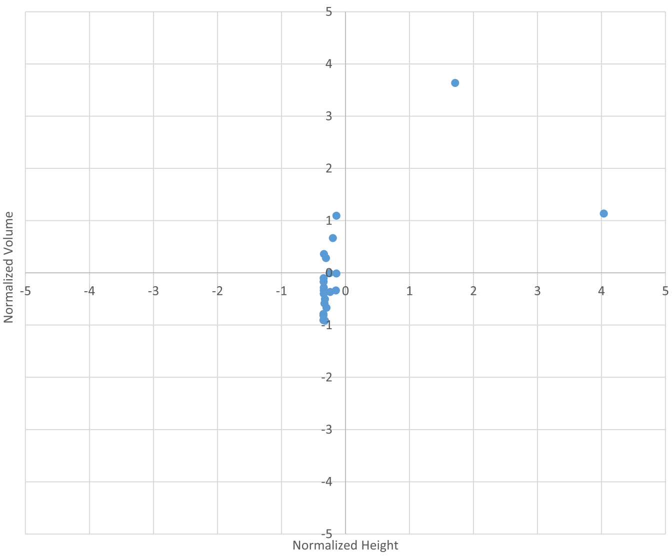


FIGURE 5. Twenty-two normalized data points for volume stored above breach invert and depth above breach for Froehlich regression equation $\left(Q_p = 0.607(V_w^{0.295} \cdot H_w^{1.24})\right)$.

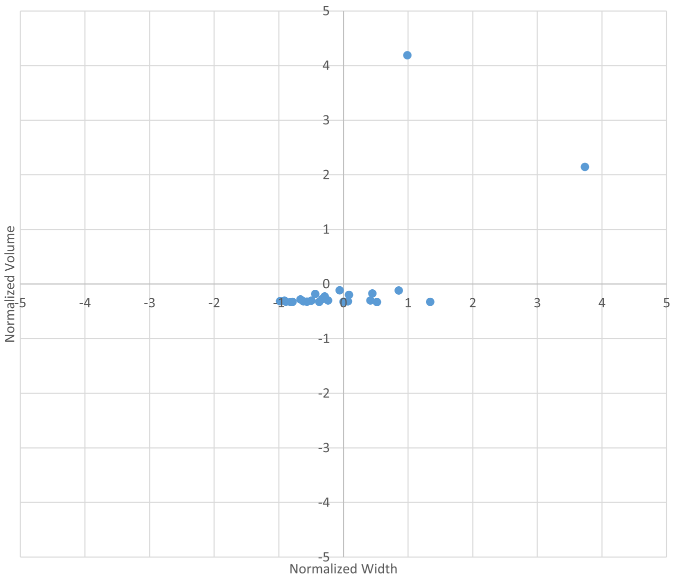


FIGURE 6. Twenty-five normalized data points for volume stored and average width for Pierce et al. (2010) regression equation $\left(Q_p = 0.863(V^{0.335} \cdot H^{1.833} \cdot W_{ave}^{-0.663})\right)$.

THE “REGRESSION SELECTION TOOL” WEB APPLICATION

For each regression equation, we standardized the data points used. The purpose of the Regression Selection Tool was to assist in determining the appropriateness of a given model for a given test point.

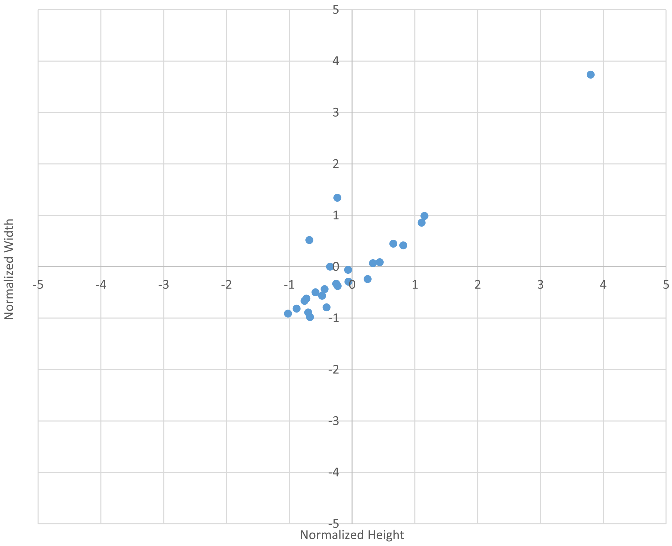


FIGURE 7. Twenty-five normalized data points for average width and height for Pierce et al. (2010) regression equation $\left(Q_{\mathrm{p}} = 0.863(V^{0.335} \cdot H^{1.833} \cdot W_{\mathrm{ave}}^{-0.663})\right)$.

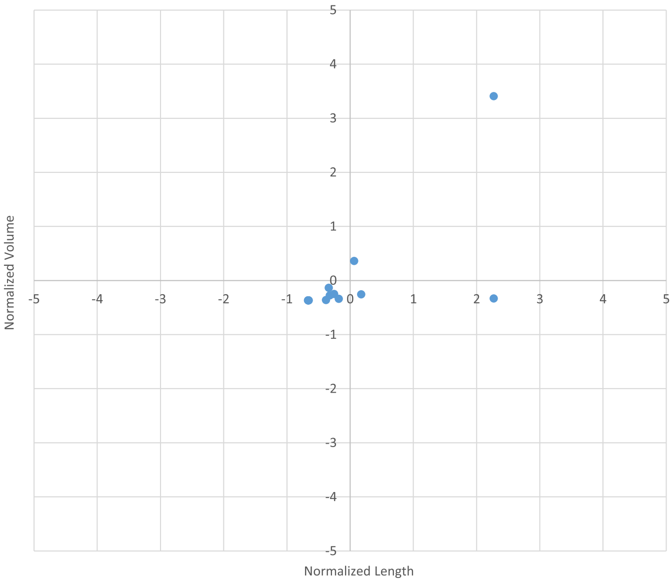


FIGURE 8. Twenty-five normalized data points for volume stored and length for Pierce et al. (2010) regression equations $\left(Q_{\mathrm{p}} = 0.012(V^{0.493} \cdot H^{1.205} \cdot L^{0.226})\right)$.

The tool also provides a visual depiction of the test point in comparison to the cluster of data points used to build the regression model. Results of the web application and fitting of variables to equations were done by normalizing all data. Normalization was performed by subtracting the data point value from the sample mean and dividing by the sample standard deviation.

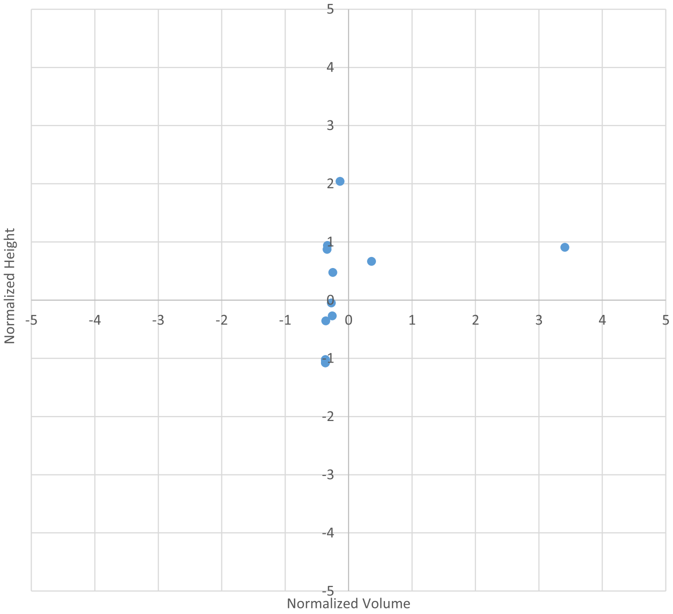


FIGURE 9. Twenty-five normalized data points for height and volume stored above breach invert for Pierce et al. (2010) regression equation $\left(Q_{\mathrm{p}} = 0.863(V^{0.335} \cdot H^{1.833} \cdot W_{\mathrm{ave}}^{-0.663})\right)$ and $\left(Q_{\mathrm{p}} = 0.012(V^{0.493} \cdot H^{1.205} \cdot L^{0.226})\right)$.

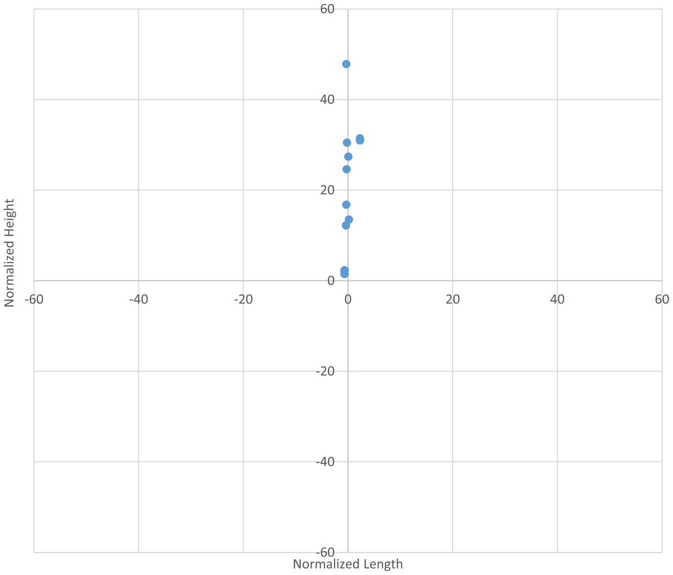


FIGURE 10. Twenty-five normalized data points for height and length for Pierce et al. (2010) regression equation $\left(Q_{\mathrm{p}} = 0.012(V^{0.493} \cdot H^{1.205} \cdot L^{0.226})\right)$.

These normalized data were then used to create scatter plots depicting the standardized associated marginal distributions (see Figures 1–11). The scatter plots can help in assessing the appropriateness of a given peak flow regression equation to an arbitrary test point. For example, in Figure 11, two of the normalized lengths used in the regression equation are relatively high, while the other lengths are very

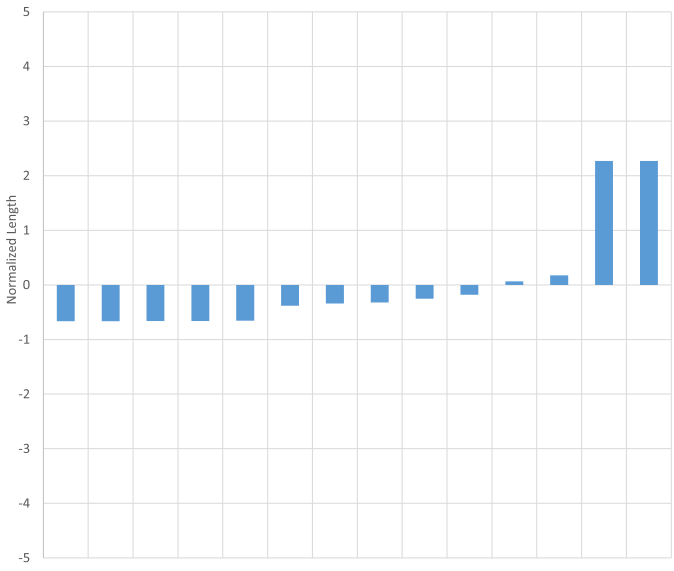
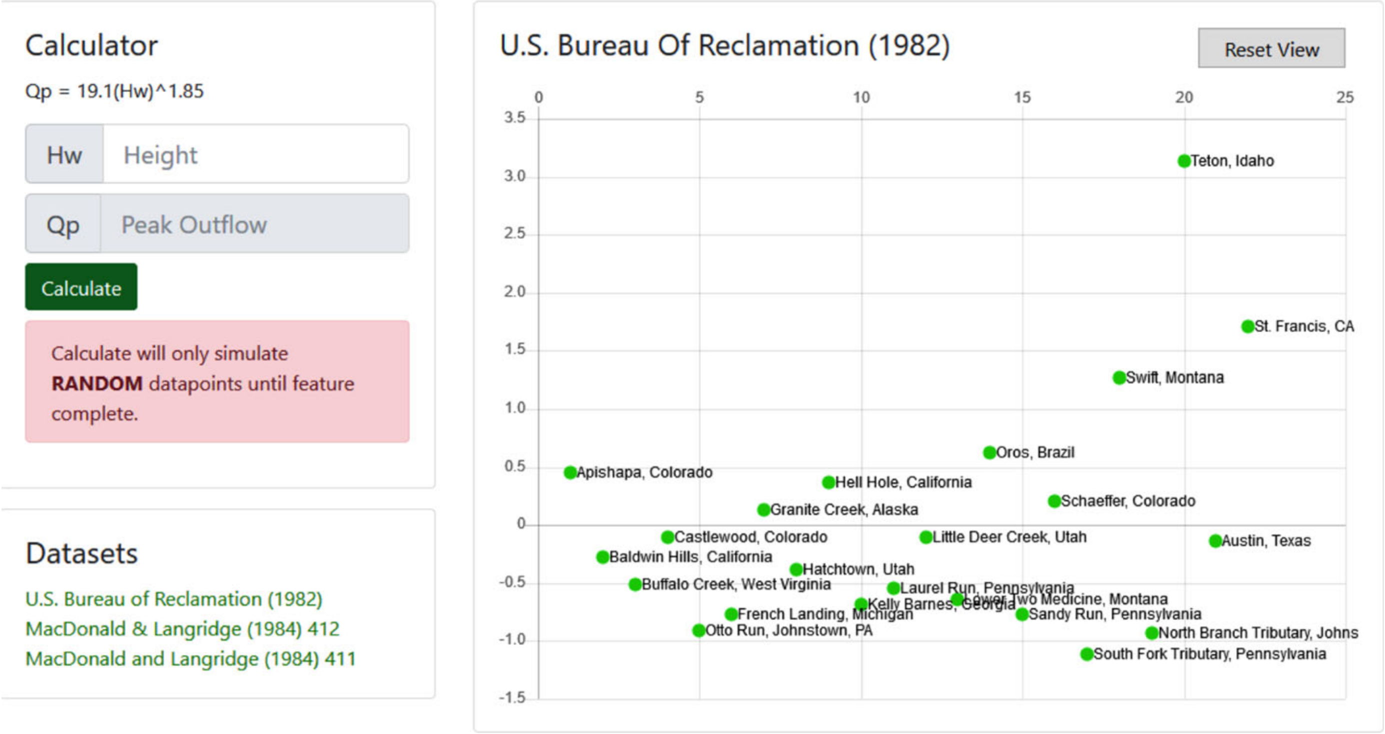


FIGURE 11. Fourteen normalized data points for length for Pierce et al. (2010) regression equation $\left(Q_{\mathrm{p}} = 0.1202(L)^{1.7856}\right)$.

small. This is partly due to a large range in the observations used to develop the regression equation. Using an arbitrary test point, a practitioner can visualize how their test point aligns with the normalized scatter plots.

Using the standardized scatter plots, the online web application provides a graphical display of the data reported in the literature that is used in the associated regression equation. Users can view all nine regression equations using a drop down menu to select which regression equation they would like to view. The database for the selected model is depicted on the web tool. The user can input test point values to add the test point to the graphical display for comparison. The graphical display allows the practitioner to visualize the consistency of the selected regression equation for the test point. At issue is whether or not the data upon which the selected model was based are appropriate for the factors of the associated test point. If the test case appears to be an “outlier” to the model points, then an alternative model may be more appropriate for peak outflow estimation. Conversely, if the test point lies within the cluster of observations, then the selected model is more likely to be suitable for estimating peak flow. See Figures 12 and 13 for a preview of the web application. In Figure 12, a practitioner can observe all available data points used in the 1982 U.S. Bureau of Reclamation equation. From there, the practitioner can observe an arbitrary test point with a specific height compared to the rest of the data. The web application will then show where the test point will nest with the other points as seen in Figure 13. Practitioners can access the website at http://usmathematics.com/dist/.



Calculator

Qp = 19.1(Hw)^1.85

Hw Height

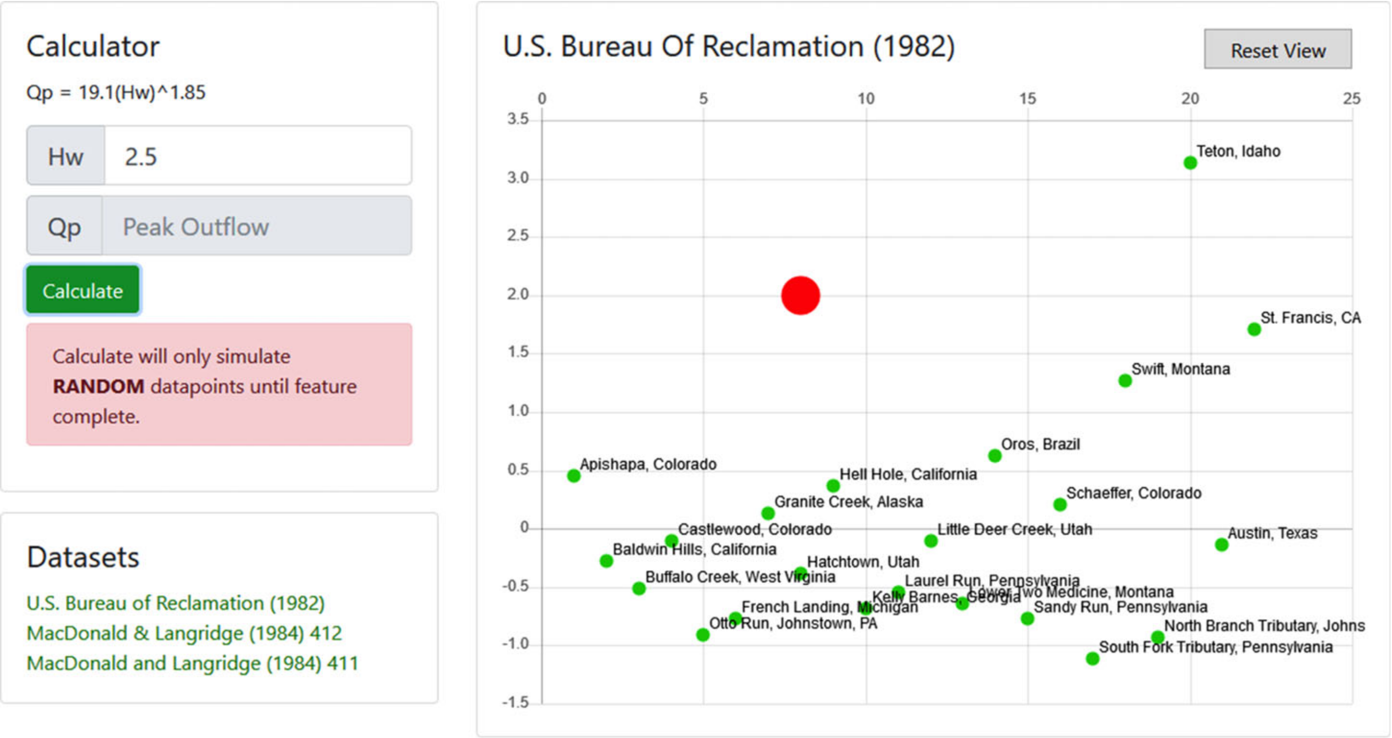
Qp Peak Outflow

Calculate will only simulate **RANDOM** datapoints until feature complete.

Datasets

U.S. Bureau of Reclamation (1982)
MacDonald & Langridge (1984) 412
MacDonald and Langridge (1984) 411

FIGURE 12. Example of the web base application using data from the U.S. Bureau of Reclamation (1982) equation. CA, California; PA, Pennsylvania.


Calculator

Qp = 19.1(Hw)^1.85

Hw 2.5

Qp Peak Outflow

Calculate

Calculate will only simulate **RANDOM** datapoints until feature complete.

Datasets

U.S. Bureau of Reclamation (1982)

MacDonald & Langridge (1984) 412

MacDonald and Langridge (1984) 411

FIGURE 13. Example test case data point under study transposed on the web base application with data from the U.S. Bureau of Reclamation (1982) equation.

Thus, the visualization provided by the Regression Selection Tool assists practitioners to better model earthen dam breaks. The information gained by improved understanding of key parameters that affect these phenomena will lead to safer and higher quality dam failure planning. The benefits of these contributions improve estimates and better inform decision support for leaders regarding operational and safety decisions.

CONCLUSIONS AND FUTURE WORK

Of all major operational dams, 85% of the dams are earthen dams. Many regression equations are used to model earthen dam failures where some regression equations are more applicable to use than other models. We developed a web application to assist practitioners using the "Regression Equation Selection Tool." The application provides a visualization of all the data points used for their respective regression equations where test case studies can be viewed. This is an evolving project with new regression equations added and documented along with additional datasets as they become available.

This database along with the web base application will help aid practitioners to properly choose the appropriate regression to use in their model based on the parameters of their earthen dam. With the use of this application, there will be an overall improvement in modeling and provide better quality, safer, and more economical earthen dams.

ACKNOWLEDGMENTS

The authors thank the editor and reviewers for their support and insight, as well as their service to the operations. In addition, the authors thank the Department of Mathematical Sciences, U.S. Military Academy for its support given throughout the research process. Contributors to the database include Dr. Tony Wahl (Senior Hydraulic Engineer for Bureau of Reclamation), Dr. Steven Abt (Professor of Civil and Environmental Engineering in the College of Engineering at Colorado State University), and Dr. Limin Zhang (Engineering Director Geotechnical Centrifuge Facility at The Hong Kong University of Science and Technology Clear Water Bay, Hong Kong) who are foundational in the literature in the development of earthen dam break assessment methodology and data compilations.

LITERATURE CITED

Billington, D.P., and D.C. Jackson. 2017. *Big Dams of the New Deal Era: A Confluence of Engineering and Politics*. Norman, OK: University of Oklahoma Press.

Froehlich, D.C. 1995. "Peak Outflow from Breached Embankment Dam." *Journal of Water Resources Planning and Management* 121 (1): 90–97.

MacDonald, T.C., and J. Langridge-Monopolis. 1984. “Breaching Characteristics of Dam Failures.” *Journal of Hydraulic Engineering* 110 (5): 567–86.

Pierce, M.W., C.I. Thornton, and S.R. Abt. 2010. “Predicting Peak Outflow from Breached Embankment Dams.” *Journal of Hydrologic Engineering* 15 (5): 338–49.

Singh, K.P., and A. Snorrason. 1984. “Sensitivity of Outflow Peaks and Flood Stages to the Selection of Dam Breach Parameters and Simulation Models.” *Journal of Hydrology* 68 (1–4): 295–310.

U.S. Bureau of Reclamation. 1982. “Guidelines for Defining Inundated Areas Downstream from Bureau of Reclamation Dams.” *Reclamation Planning Instruction No. 82-11*, Washington, D.C.

Wahl, T.L. 1998. “Prediction of Embankment Dam Breach Parameters — A Literature Review and Needs Assessment.” *Dam Safety Rep No. DSO-98-004*, U.S. Department of the Interior, Bureau of Reclamation, Denver, CO.

Wahl, T.L. 2014. “Evaluation of Erodibility-Based Embankment Dam Breach Equations.” *Interagency Agreement NRC-HQ-60-12-I-0013*, U.S. Department of the Interior, Bureau of Reclamation, Denver, CO.

Xu, Y., and L.M. Zhang. 2009. “Breaching Parameters for Earth and Rockfill Dams.” *Journal of Geotechnical and Geoenvironmental Engineering* 135 (12): 1957–70.