

# Data-Driven Dam Safety:

## Neural Network Prediction of Embankment Dam Breach Parameters

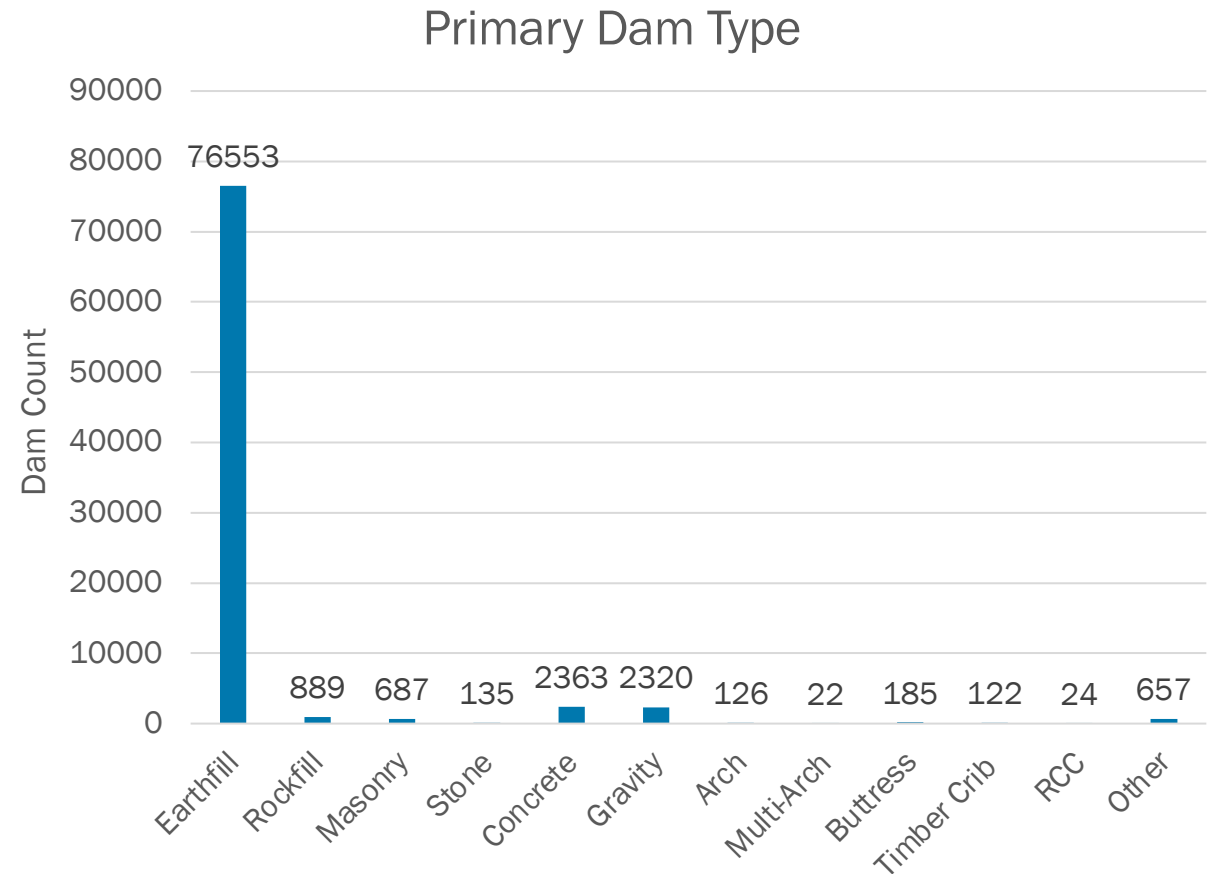
National Dam Safety Program Technical Seminar | 2024



FEMA

# Primary Dam Type (from the National Inventory of Dams)

Earthfill + Rockfill embankment dams = 92% of all known primary dam types in the USA



FEMA

# Embankment Dams (Earthfill and Rockfill)

---



**FEMA**

# Embankment Dams Occasionally Breach

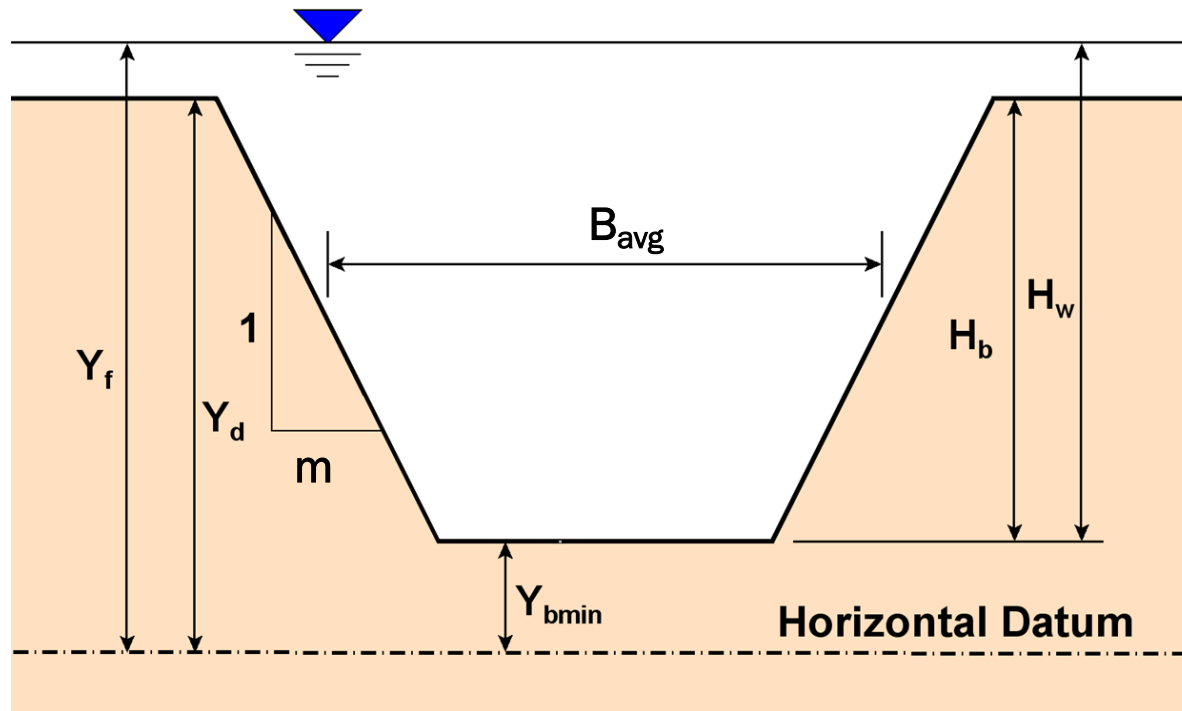
---



**FEMA**

# Empirical Trapezoidal Breach Model

# Trapezoidal Breach Model Parameters: $B_{avg}$ , $m$ , $t_f$ , $H_b$



$t_f$  = breach formation time



FEMA

# Embankment Dam Failure Mode: Mode

- About 1/3 caused by inadequate spillway capacities that result in overtopping by floodwaters (OF)
- Another 1/3 of failures are attributed internal erosion (piping) (IE)
- Remaining failures are caused by embankment slides (OS), wave action (OW), by outlet works failure (OG), intentional breaching by excavation (OX)

$$\text{Mode} = \begin{cases} 0, & \text{if internal erosion (IE) failure} \\ 1, & \text{if overtopping (OF, OS, OW, OG) failure} \end{cases}$$



FEMA

# Overtopping Failure

---



**FEMA**

## Embankment Dam Solid Corewall: Core

---



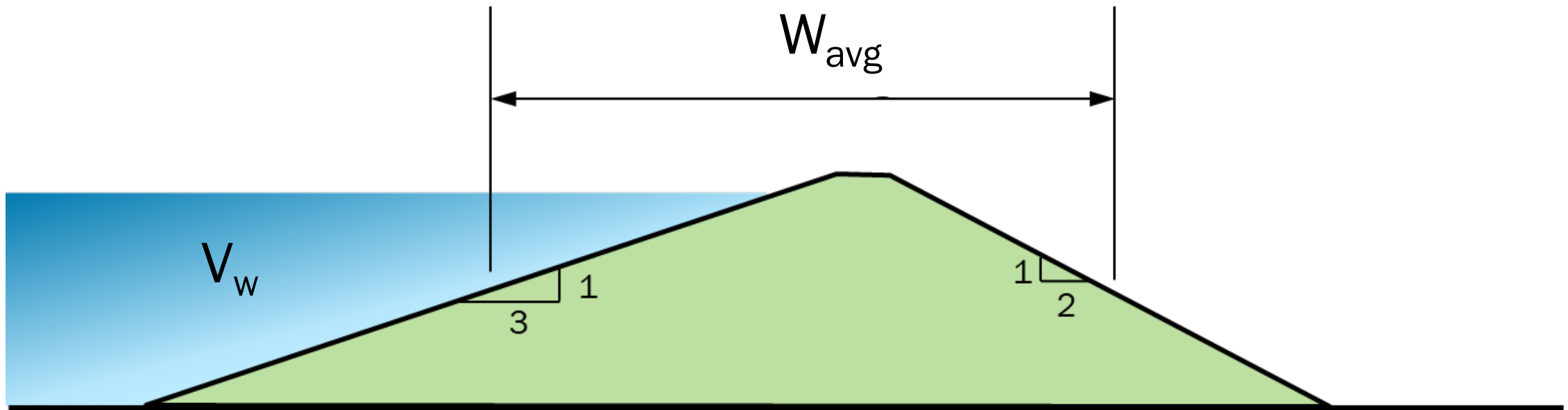
Some embankment dams use a rigid masonry, concrete, bituminous concrete, or steel corewall to create an impervious barrier within the embankment.

$$\text{Core} = \begin{cases} 0, & \text{if no corewall} \\ 1, & \text{if corewall} \end{cases}$$



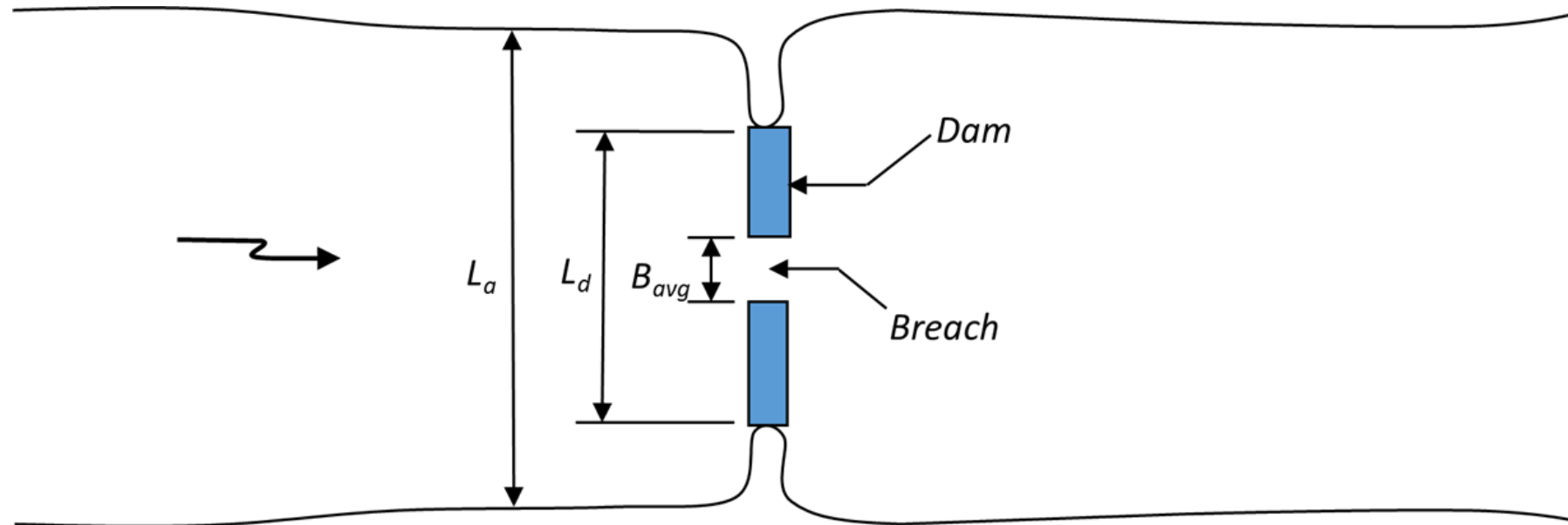
**FEMA**

# Average Embankment Width and Storage Volume: $W_{avg}$ and $V_w$



FEMA

## Approach Flow Width: $L_a$



FEMA

# Embankment Dam Failure Data

# Breach Data from 126 Dam Failures (124 for $B_{avg}$ , 123 for $m$ , 48 for $t_f$ )

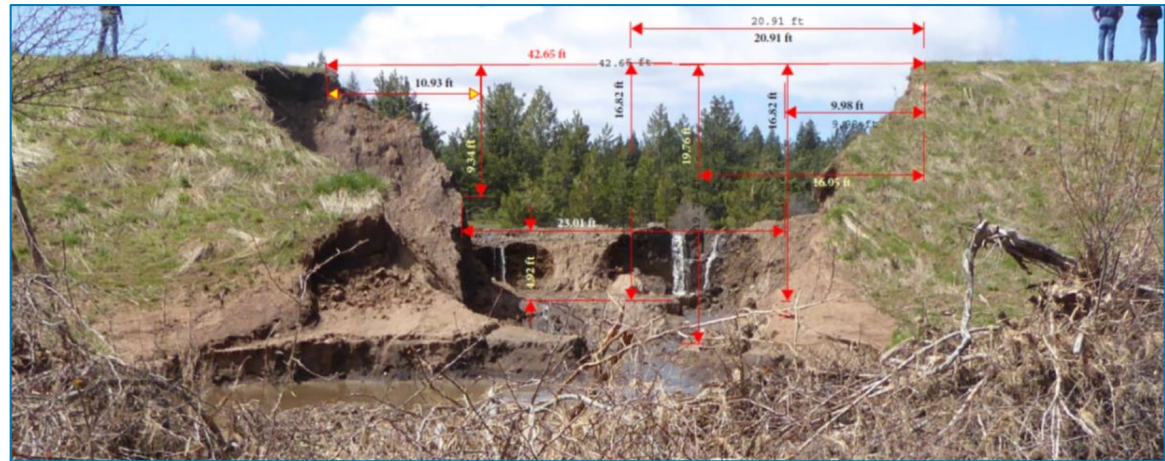
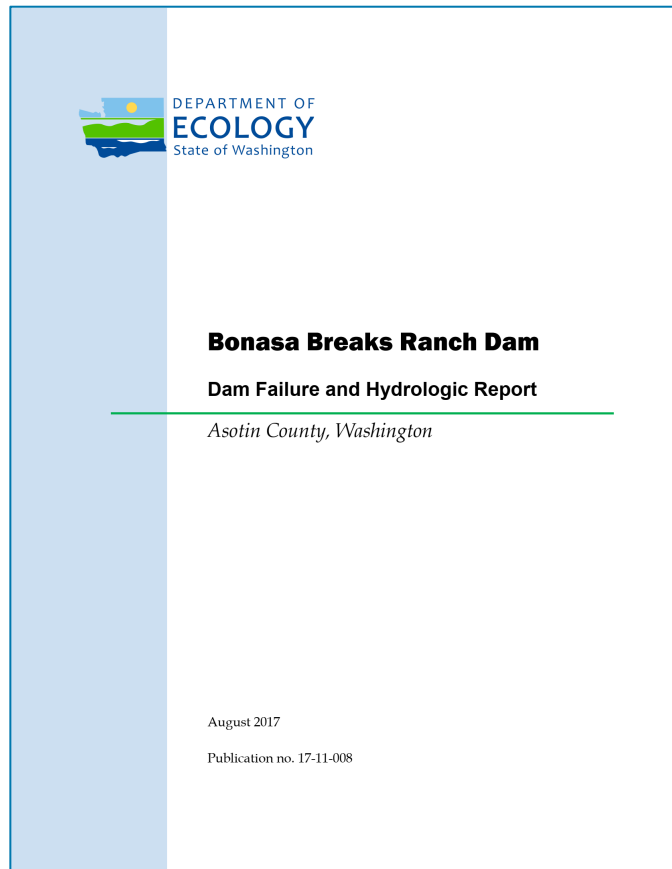
Table 1. Embankment dam breach data

No.	Dam name and location	Type <sup>a</sup>	Year built	Year failed	Failure mode <sup>b</sup>	$W_{avg}$ (m)	$V_w$ (Mm <sup>3</sup> )	$H_w$ (m)	$H_b$ (m)	$L_a$ (m)	$B_{avg}$ (m)	$m$ (h:v)	$t_f$ (h)
1	Apishapa, Colo.	E, H, C	1920	1923	IE	82.4	22.8	28.0	31.1	200	93.0	0.44	0.75
2	Baldwin Hills, Calif.	E, H	1951	1963	IE	59.6	0.950	12.2	21.3	200	25.0	0.31	0.33
3	Banqiao, Henan Province, China	E, H	1953	1975	OF	97.0	603	31.9	30.3	2100	291	2.54	5.5
4	Bass Haven Lake, Tex.	E, H	-- <sup>c</sup>	1984	OX	22.9	0.641	4.90	9.20	100	23.5	0.60	--
5	Bearwallow Lake, N.C.	E, H	1963	1976	OS	17.1	0.0493	5.79	6.40	150	12.2	1.43	--
6	Belci, Bacău County, Romania	E, Z	1963	1991	OF	37.8	12.7	15.5	15.0	400	102	0.67	1.25
7	Big Bay Lake, Miss.	E, H	1992	2004	IE	20.4	17.5	13.6	14.0	800	83.2	0.95	0.92
8	Big Lake, Tex.	E, H.	--	1996	OF	12.8 <sup>i</sup>	0.550	7.00	6.40	150	53.3	2.38	--
9	Bílá Desná, Czech Republic	E, H	1915	1916	IE	23.2	0.290	10.3	14.2	170	19.0	0.77	0.20
10	Bilberry, England	E, Z	1845	1852	OS	62.5	0.327	23.6	23.0	200	37.0	1.09	0.167
11	Bradfield (Dale Dyke), England	E, Z	1863	1864	IE	76.0	3.20	28.0	29.0	300	50.3	2.50	0.75
12	Buckhaven No. 2, Tenn.	E, H	--	1991	OF	13.4	0.0247	6.10 <sup>d</sup>	6.10	70	4.72	0.73	--
13	Bullock Draw Dike, Utah	E, H	1971	1971	IE	18.6	0.740	3.05	5.79	540	12.5	0.21	--
14	Butler Valley, Ariz.	E, H	--	1982	OF	9.63	2.38	7.16	7.16	850	62.5	0.85	--
15	Caulk Lake, Ky.	E, H	--	1973	OS	32.0	0.698	11.1	12.2	70	35.1	1.38	--
16	Chaq-Chaq, Sulaimani City, Iraq	E, Z	--	2006	OF	45.3	2.55	15.1	14.5	170	37.8 <sup>f</sup>	0.57	--
17	Clearwater Lake, Ga.	E, H	1965	1994	OF	15.0	0.466	4.05	3.78	230	22.8	1.03	--



FEMA

# Bonasa Breaks Ranch Dam Failure Report



**FEMA**

# Data Cleansing

---



**FEMA**

# Standardized Variables

$$x' = \frac{x - \bar{x}}{\sigma_x}$$

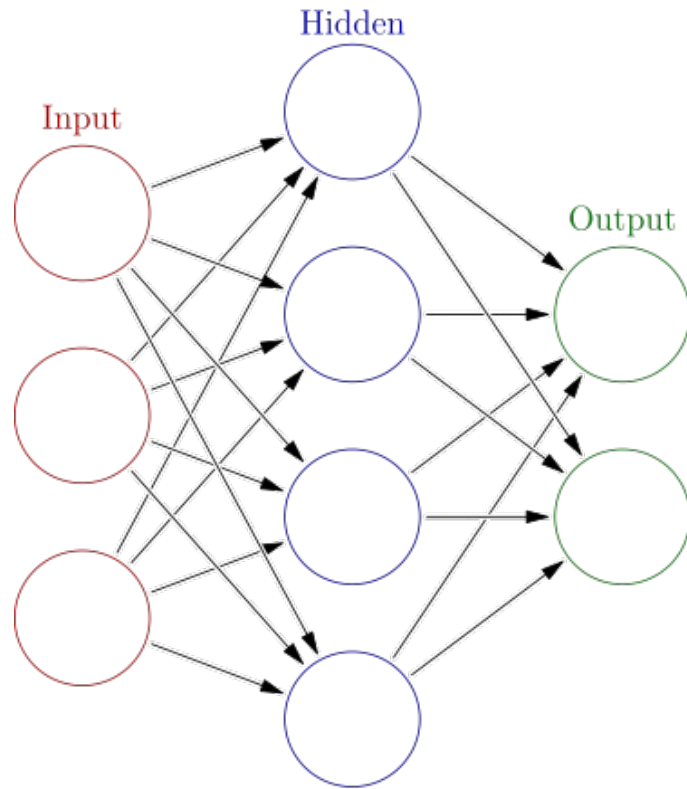
Variable	Units	Mean	Standard deviation	Minimum	Maximum
L <sub>a</sub>	m	345	469	30	3800
W <sub>avg</sub>	m	35.3	31.5	7.62	250
V <sub>w</sub>	Mm <sup>3</sup>	23.4	88.6	0.0133	660
H <sub>b</sub>	m	12.9	11.4	2.1	86.9
B <sub>avg</sub>	m	49.7	56.3	2.29	367
m	m	0.974	0.626	0.13	3.03
t <sub>f</sub>	hours	1.17	1.46	0.083	6
Q <sub>p</sub>	m <sup>3</sup> /s	6,790	14,720	30	65,120



# Neural Network Analysis

# Artificial Neural Network

---



- An artificial neural network is an interconnected group of nodes, inspired by a simplification of neurons in a brain.
- Here, each circular node represents an artificial neuron and an arrow represents a connection from the output of one artificial neuron to the input of another.
- A network is typically called a deep neural network if it has at least 2 hidden layers.



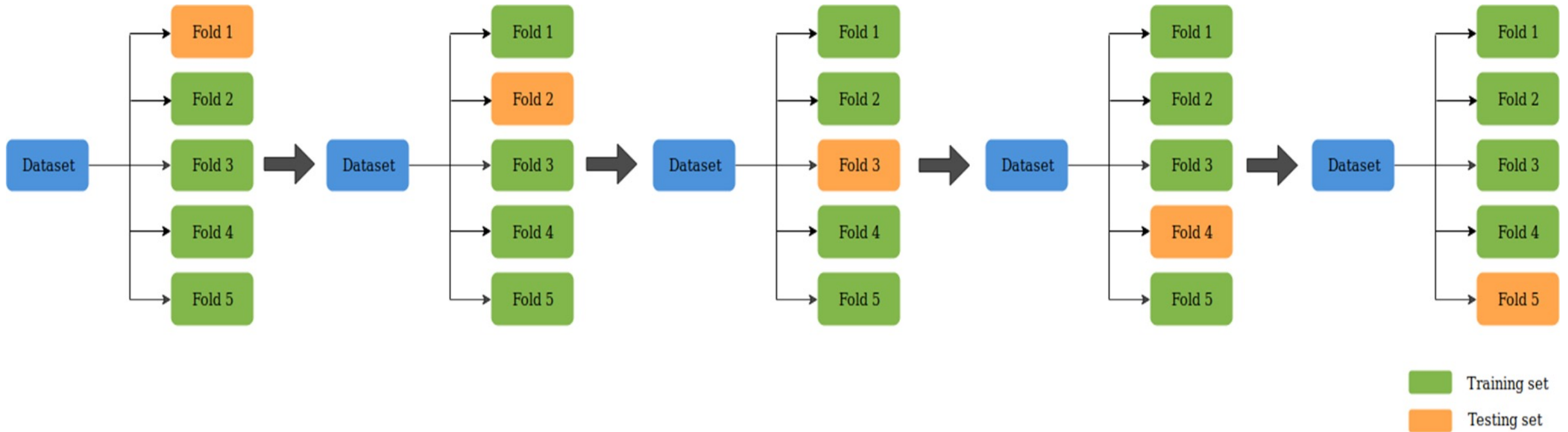
**FEMA**

# Overfitting



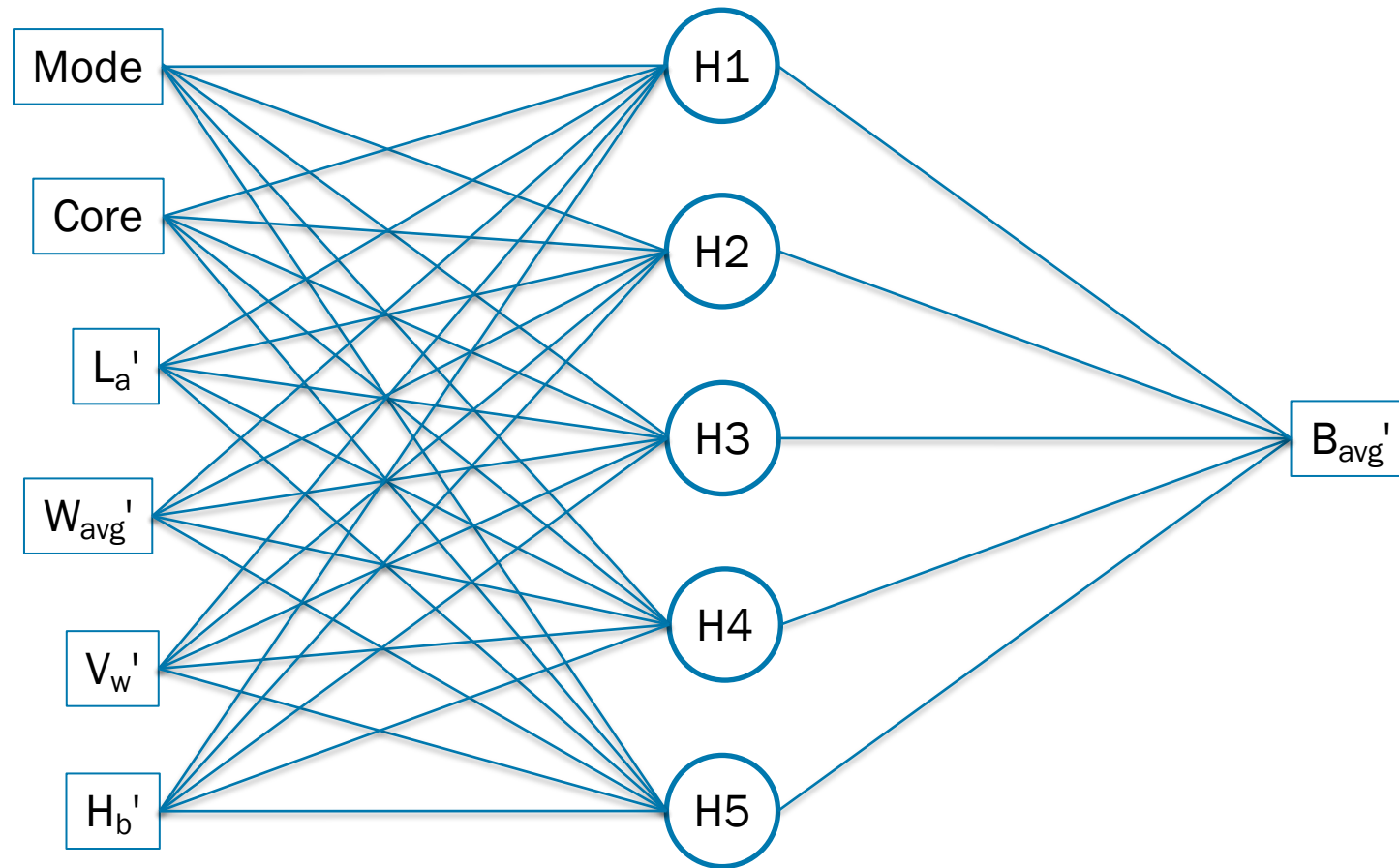
FEMA

# k-fold Cross Validation (k=5)



FEMA

# $B_{avg}'$ Neural Network Schematic



FEMA

## B<sub>avg</sub>' Neural Network

---

$$H1 = \tanh\left(-0.9847 + 0.5058 \times \text{Mode} + 2.680 \times \text{Core} + 0.5196 \times L_a' - 0.9491 \times W_{avg}' + 0.1416 \times V_w' + 0.1351 \times H_b'\right)$$

$$H2 = \tanh\left(-5.643 + 0.7383 \times \text{Mode} + 10.15 \times \text{Core} - 6.237 \times L_a' + 0.1948 \times W_{avg}' + 2.484 \times V_w' + 1.124 \times H_b'\right)$$

$$H3 = \tanh\left(-0.9922 + 0.9627 \times \text{Mode} - 0.3443 \times \text{Core} - 1.255 \times L_a' - 0.8666 \times W_{avg}' + 2.431 \times V_w' + 2.439 \times H_b'\right)$$

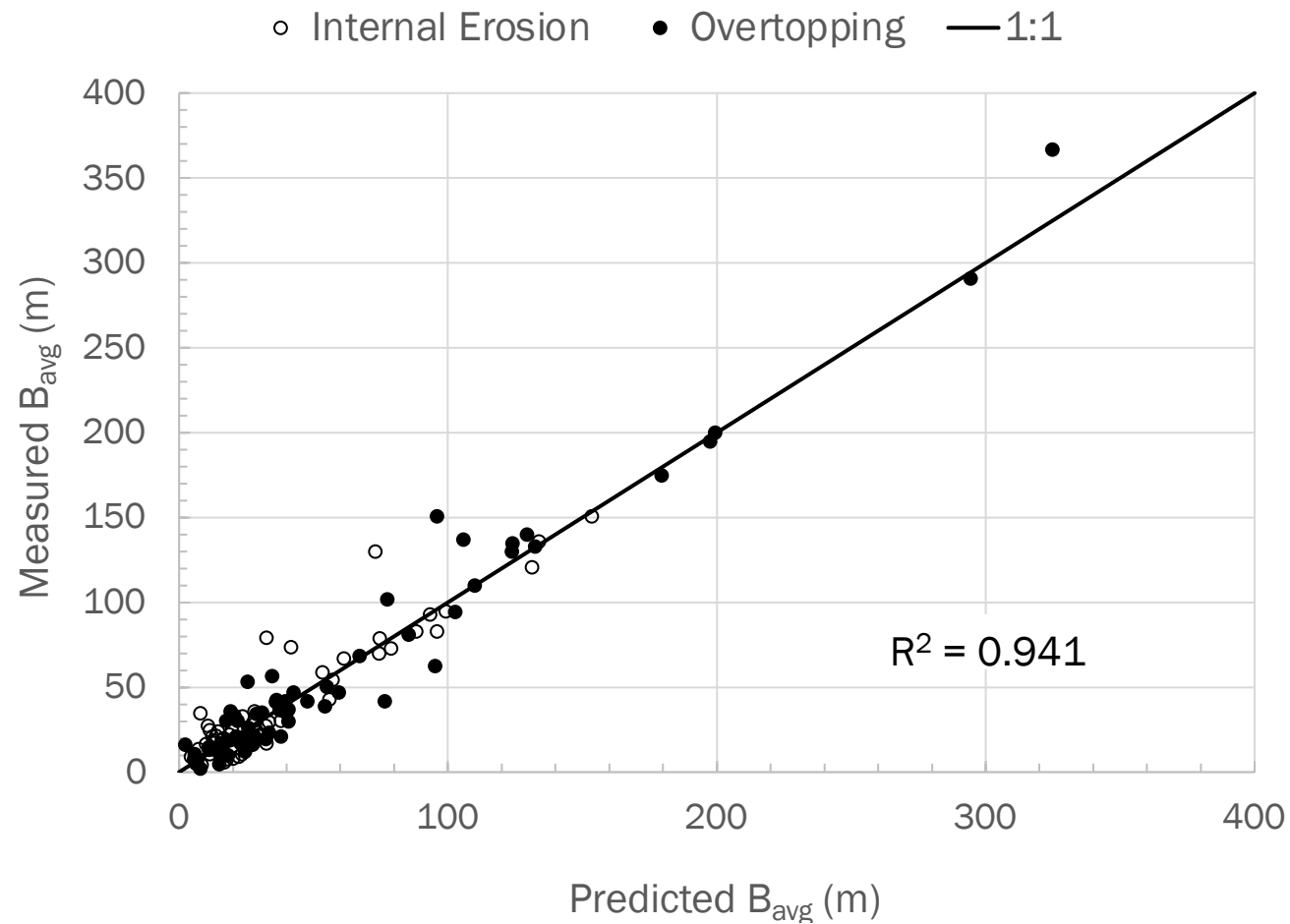
$$H4 = \tanh\left(-0.9132 + 0.5197 \times \text{Mode} + 0.4595 \times \text{Core} - 0.7917 \times L_a' - 0.2369 \times W_{avg}' + 0.7210 \times V_w' + 1.116 \times H_b'\right)$$

$$H5 = \tanh\left(-1.1169 + 0.3548 \times \text{Mode} - 3.995 \times \text{Core} - 0.6710 \times L_a' - 1.040 \times W_{avg}' + 0.4767 \times V_w' - 0.6112 \times H_b'\right)$$

$$B_{avg}' = 0.3934 + 3.457 \times H1 - 1.063 \times H2 - 1.101 \times H3 + 2.012 \times H4 - 2.450 \times H5$$

$$B_{avg} = 49.7 + 56.3 \times B_{avg}'$$

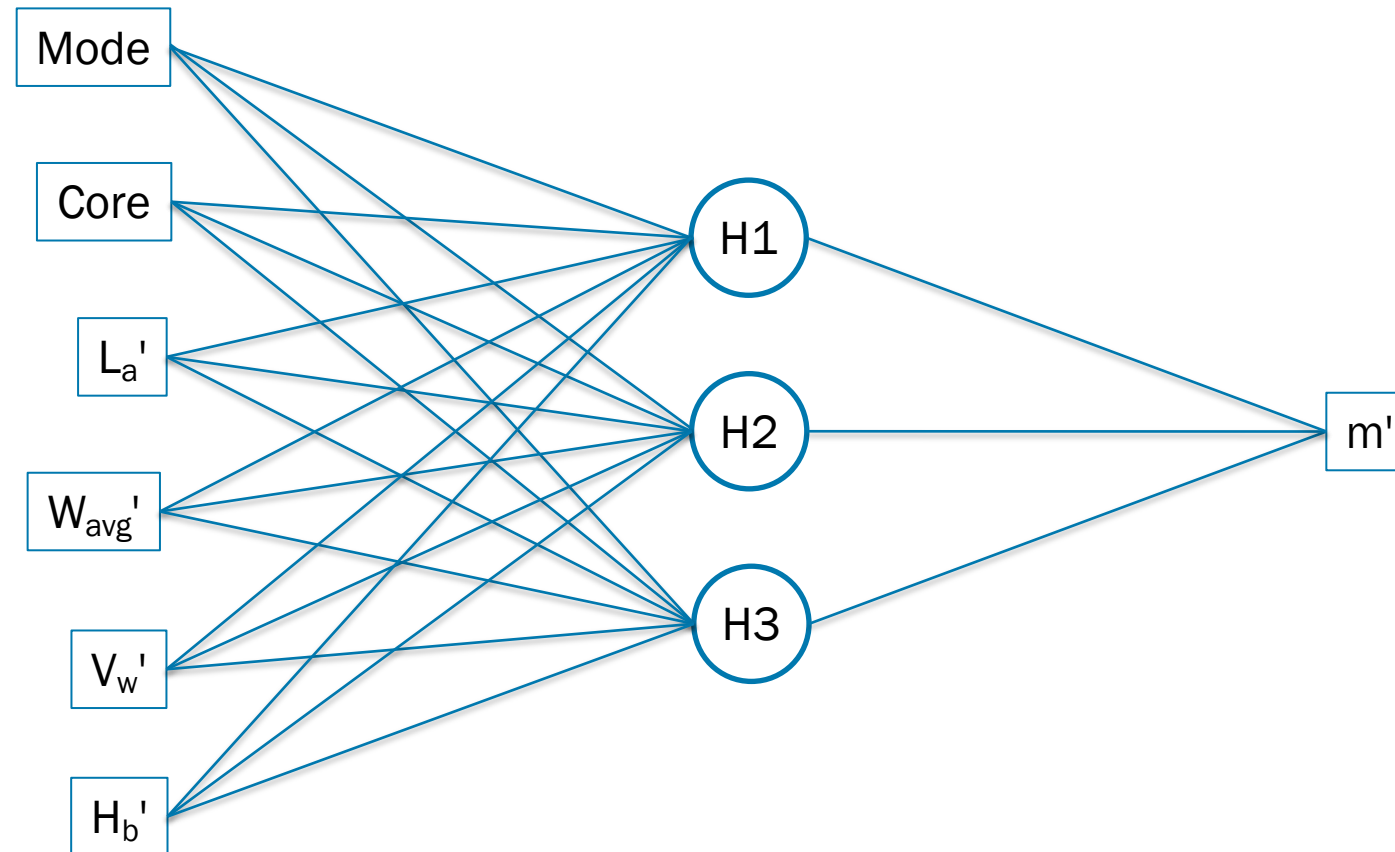
# Measured vs. Predicted $B_{avg}$



FEMA

# m' Neural Network Schematic

---



**FEMA**

## m' Neural Network

---

$$H1 = \tanh\left(-2.808 + 1.672 \times \text{Mode} + 4.144 \times \text{Core} + 0.1524 \times L_a' - 0.9152 \times W_{avg}' + 1.208 \times V_w' - 0.9376 \times H_b'\right)$$

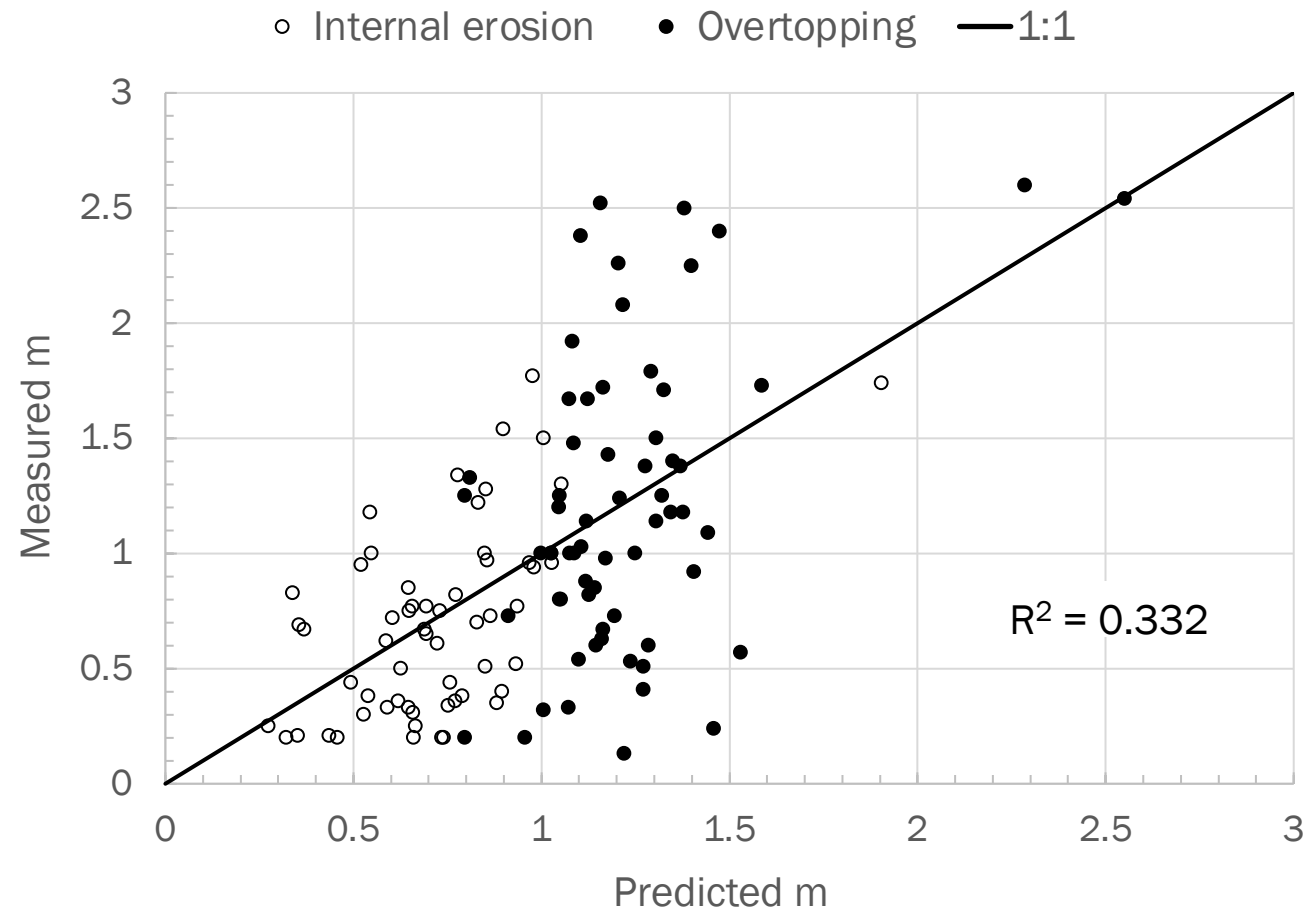
$$H2 = \tanh\left(0.5379 - 2.724 \times \text{Mode} - 1.557 \times \text{Core} + 2.383 \times L_a' + 0.03849 \times W_{avg}' + 0.4909 \times V_w' - 2.300 \times H_b'\right)$$

$$H3 = \tanh\left(0.2407 - 1.584 \times \text{Mode} + 6.083 \times \text{Core} + 1.415 \times L_a' - 0.2488 \times W_{avg}' + 0.6711 \times V_w' - 1.508 \times H_b'\right)$$

$$m' = 2.293 + 2.780 \times H1 + 2.390 \times H2 - 2.960 \times H3$$

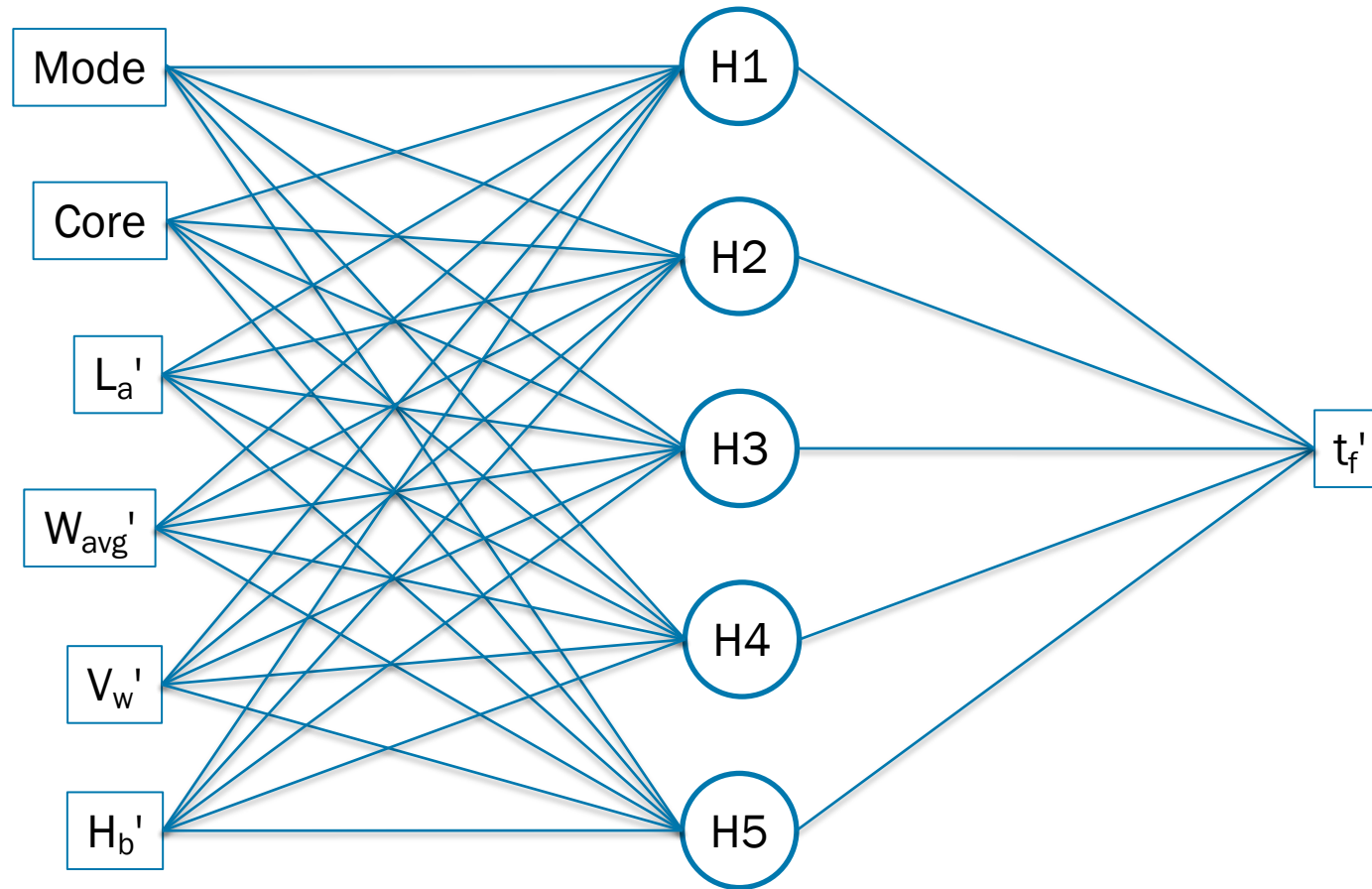
$$m = 49.7 + 56.3 \times m'$$

# Measured vs. Predicted m



FEMA

# $t_f'$ Neural Network Schematic



FEMA

## $t_f'$ Neural Network

---

$$H1 = \tanh\left(1.428 - 0.8005 \times \text{Mode} - 3.265 \times \text{Core} - 0.3579 \times L_a' + 0.1349 \times W_{avg}' - 0.4443 \times V_w' + 1.154 \times H_b'\right)$$

$$H2 = \tanh\left(0.2450 - 2.147 \times \text{Mode} + 3.819 \times \text{Core} - 0.1738 \times L_a' + 0.3629 \times W_{avg}' + 0.4377 \times V_w' - 1.073 \times H_b'\right)$$

$$H3 = \tanh\left(0.8612 + 0.1007 \times \text{Mode} - 0.8685 \times \text{Core} + 0.1701 \times L_a' + 0.1674 \times W_{avg}' + 1.466 \times V_w' + 0.2701 \times H_b'\right)$$

$$H4 = \tanh\left(0.1189 + 0.1870 \times \text{Mode} + 0.1577 \times \text{Core} - 0.1449 \times L_a' - 0.1592 \times W_{avg}' + 1.321 \times V_w' - 0.6088 \times H_b'\right)$$

$$H5 = \tanh\left(-0.2294 - 0.6111 \times \text{Mode} + 3.529 \times \text{Core} - 0.3297 \times L_a' - 0.05953 \times W_{avg}' - 0.8622 \times V_w' - 0.4747 \times H_b'\right)$$

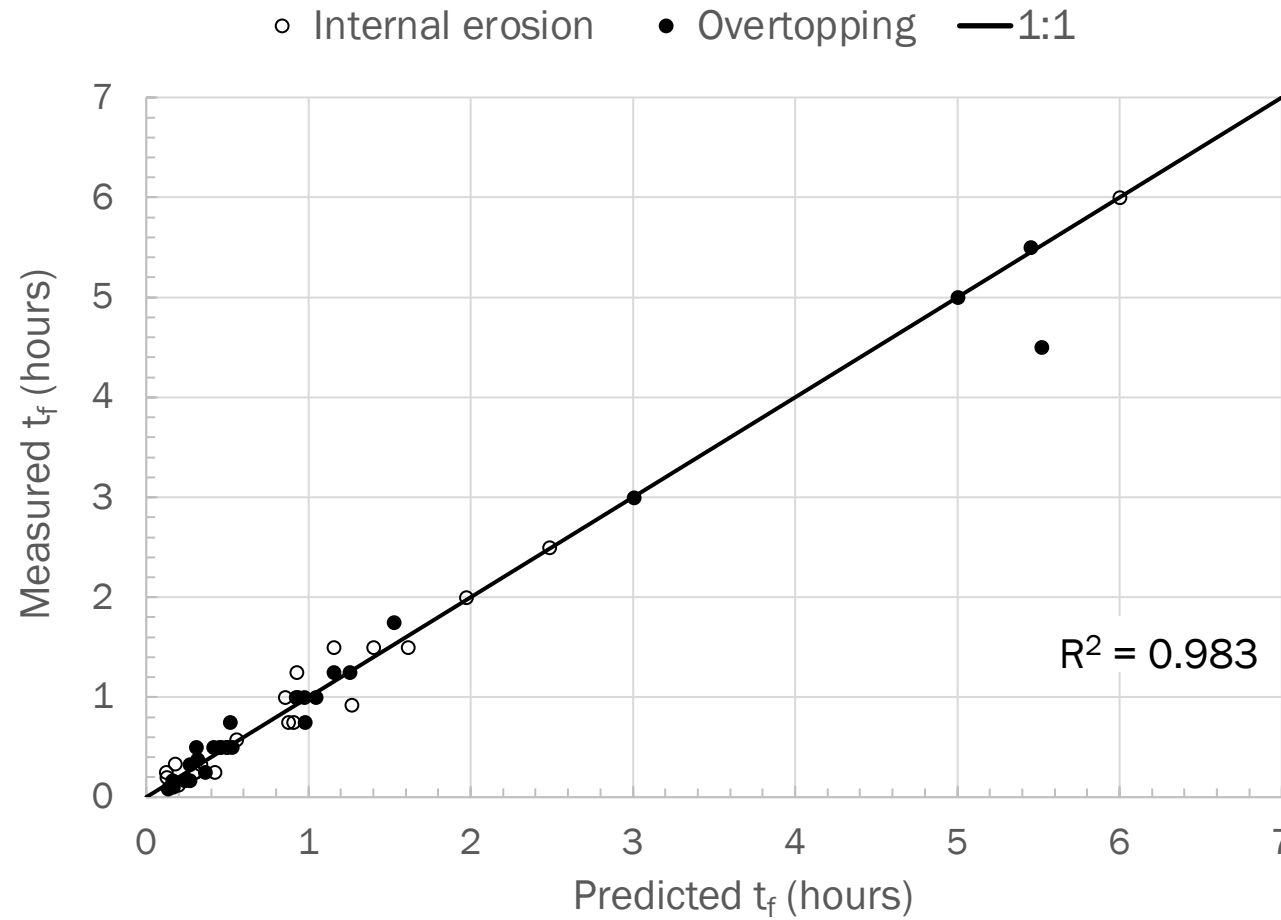
$$t_f' = -0.5275 - 0.8270 \times H1 + 0.7918 \times H2 + 1.848 \times H3 + 0.7170 \times H4 - 0.7754 \times H5$$

$$t_f = 1.17 + 1.46 \times t_f'$$



**FEMA**

# Measured vs. Predicted $t_f$



FEMA

# Breach Model Parameter Equations

## Empirical Model of Embankment Dam Breaching

D.C. Froehlich

Consulting Engineer, 303 Frenchmans Bluff Drive, Cary, North Carolina 27513, USA

**ABSTRACT.** Catastrophic flooding created by breached embankment dams needs to be evaluated when assessing potential hazards to select appropriate inflow design floods and to prepare emergency action plans. Embankment dam breaches are often considered to develop in a presupposed way, usually in the shape of a trapezoid that is defined by its final height, base width or average width, and side slopes, along with the time needed for the breach to form completely. Here data from 111 embankment dam failures are evaluated to obtain expressions for expected values of the final width, side slope, and formation time of the breach, along with expressions to calculate variances and prediction intervals of the parameters.

### 1 INTRODUCTION

The National Inventory of Dams (NID) is a database maintained by the U.S. Army Corps of Engineers (USACE) that contains information about more than 87,000 dams located in the United States and its territories (USACE 2013). About 75,000, or nearly 86%, of these dams are formed by embankments constructed from natural erodible materials (earth and rock) that rely on their weight to hold back the force of water. Because embankment dams are so numerous, potential flood hazards that would be created by uncontrolled releases of impounded water through a

breach need to be evaluated to select spillway design floods and to prepare emergency action plans.

How a breach forms in an embankment dam when it fails depends on many factors including embankment geometry, material composition, construction methods, type and degree of embankment crest and slope protective cover, reservoir dimensions, inflow to the reservoir during failure, and the manner of failure. Most dam failure models portray the process with little regard for the causal agents underlying water motion over and/or through embankments, and the resulting soil erosion. Instead, breach development is simplified greatly and is considered to proceed in a presupposed way, usually with the breach growing in the shape of a trapezoid that is defined by its final shape and the time needed to form completely as (Fig. 1). Such an empirical model requires fewer input data than more intricate models that describe the physical processes of embankment erosion in detail (Froehlich 2008).

Because all process models are abstractions of reality and cannot be considered completely accurate, they possess varying degrees of uncertainty. Consequently, variability of model parameters needs to be quantified so that bounds on their values can be established. With knowledge of parameter uncertainties, the reliabilities of predicted reservoir outflow hydrographs, peak flow rates, and water-surface elevations at downstream locations, can be estimated in a straightforward manner.

To estimate embankment dam breach model parameters and their variabilities, data from 111 dam failures are analyzed using multivariate nonlinear

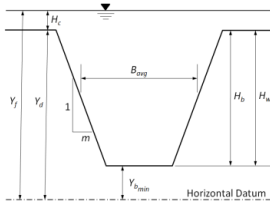


Figure 1. Final dimensions of a trapezoidal dam breach approximation, including height  $H_b$ , average width  $B_{avg}$  and side-slope ratio  $m$  (horizontal to vertical). Breaching begins when the reservoir water-surface elevation reaches the failure elevation  $Y_f$ .

$$B_{avg} = 0.23 \times k_m \times V_w^{1/3}; \quad k_m = \begin{cases} 1.0, & \text{for internal erosion failures} \\ 1.5, & \text{for overtopping failures} \end{cases}$$

$$m = \begin{cases} 0.6, & \text{for internal erosion failures} \\ 1.0, & \text{for overtopping failures} \end{cases}$$

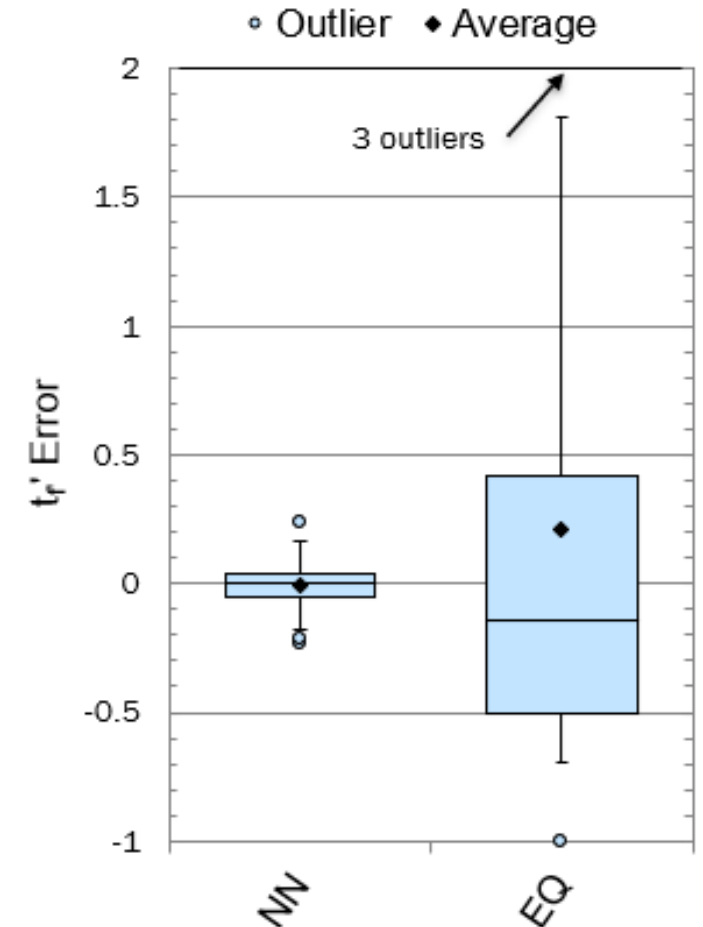
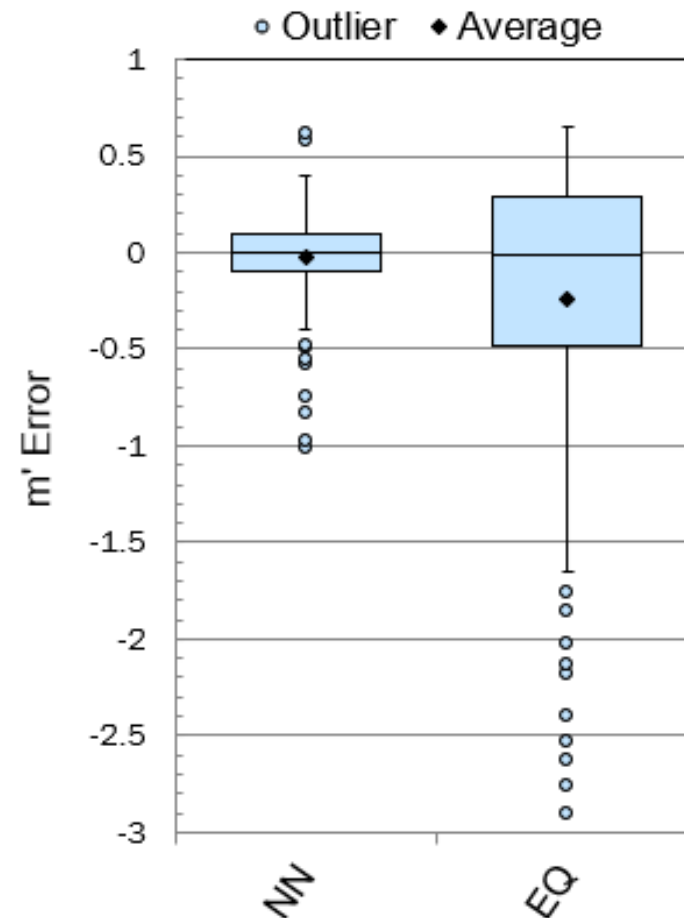
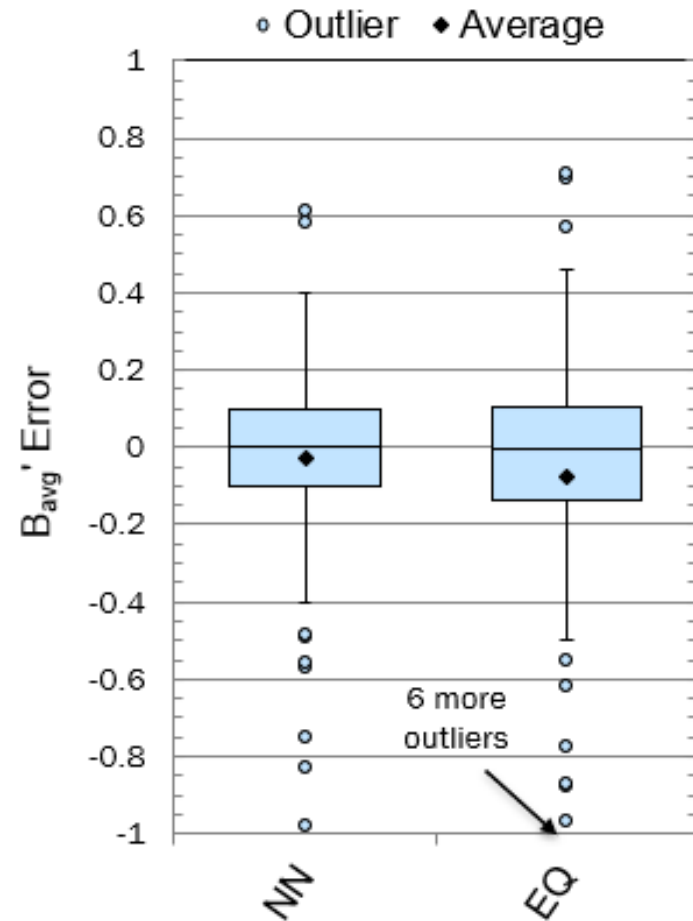
$$t_f = 60 \times \sqrt{\frac{V_w}{gH_b^2}}$$

$$V_w \text{ in } \text{Mm}^3, \quad t_f \text{ in seconds}, \quad g = 9.807 \text{ m/s}^2$$



FEMA

# Neural Network – Equation Comparison



FEMA

# Example Applications

# Gararda Dam, Rajasthan, India



Homogenous earthfill dam

Mode = 0 (Internal erosion)

Core = 0 (No corewall)

$L_a = 1150$  m

$W_{avg} = 71$  m

$V_w = 24.4$  Mm<sup>3</sup>

$H_b = 26$  m

➡  $B_{avg} = 74.8$  m

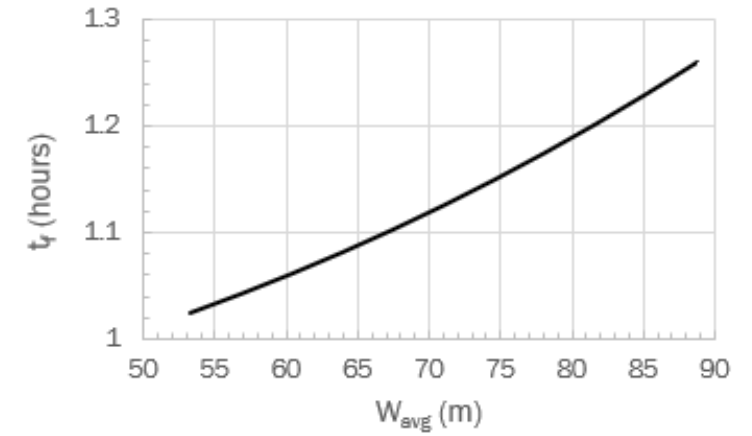
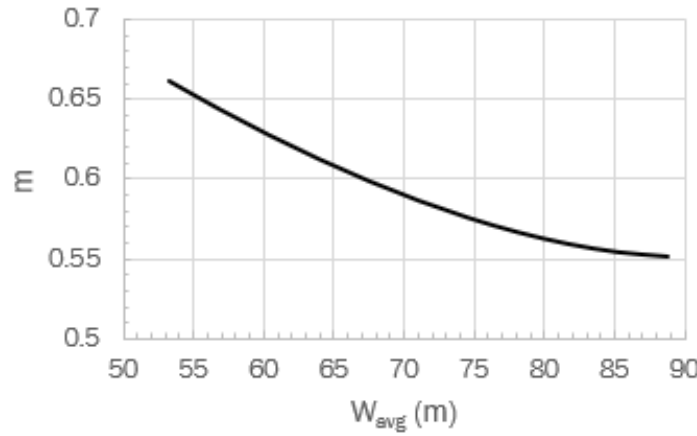
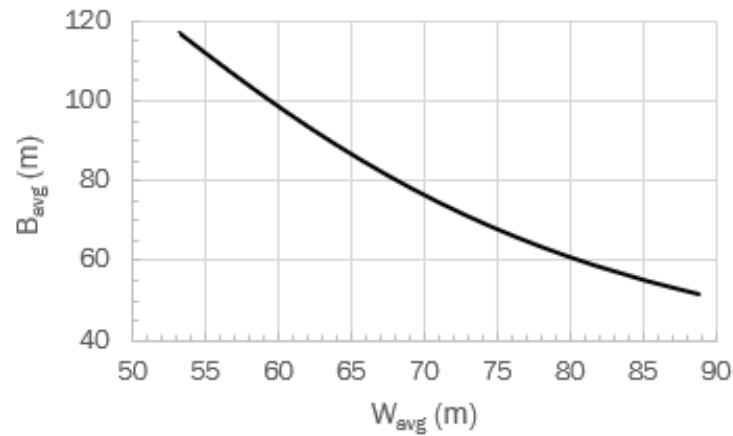
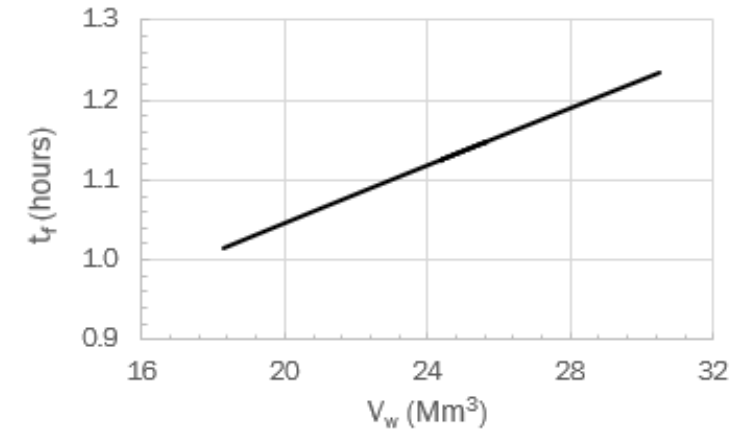
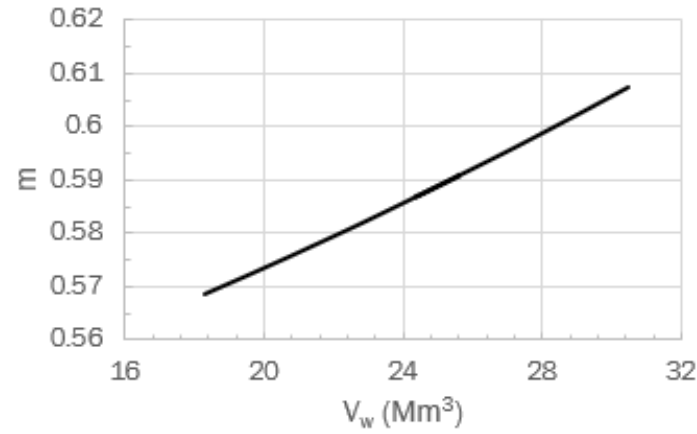
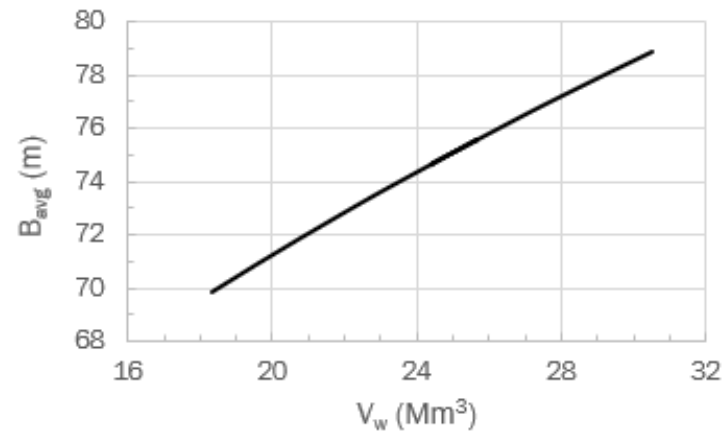
$m = 0.59$

$t_f = 1.12$  hours



FEMA

# Gararda Dam Profile Traces



FEMA

# Hirakud Dam, Odisha, India



Zoned earthfill embankment

Mode = 0 (Internal erosion)

Core = 0 (No corewall)

$L_a = 4650$  m

$W_{avg} = 108$  m

$V_w = 5700$  Mm<sup>3</sup>

$H_b = 40.2$  m

➡  $B_{avg} = 361$  m

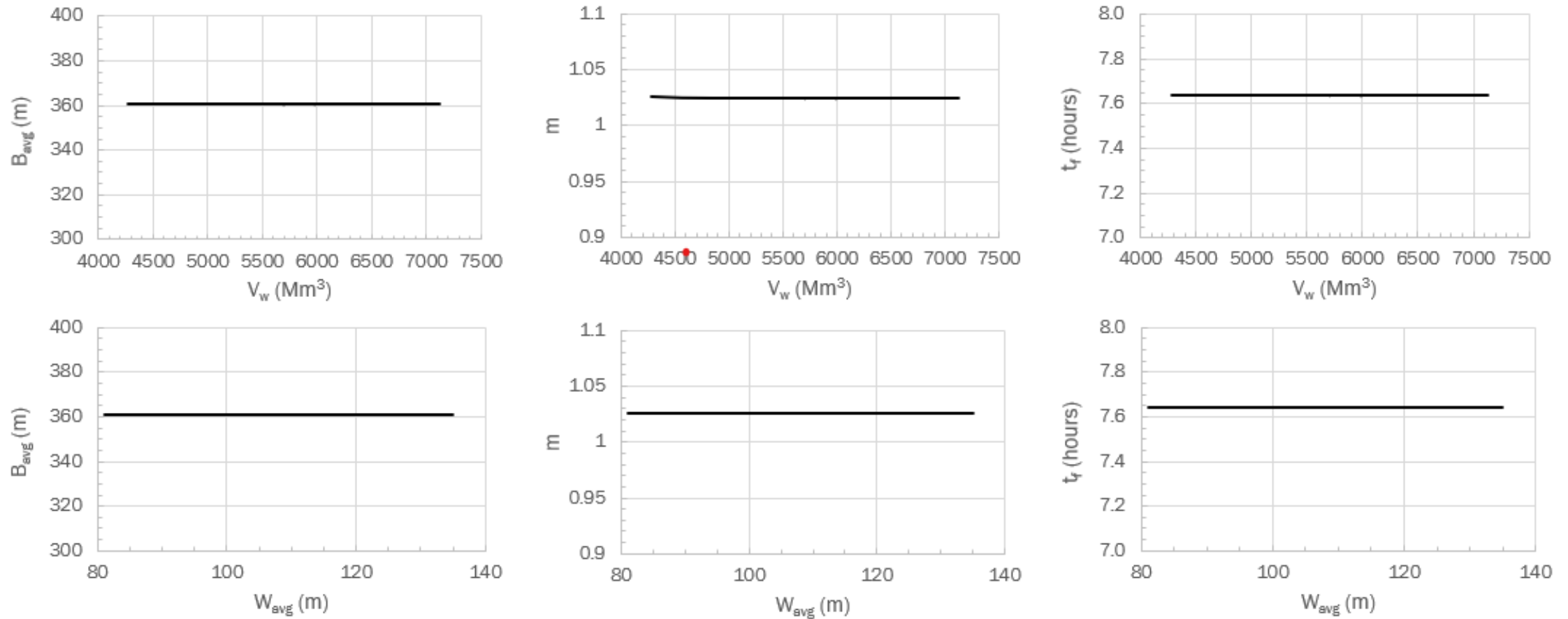
$m = 1.03$

$t_f = 7.64$  hours



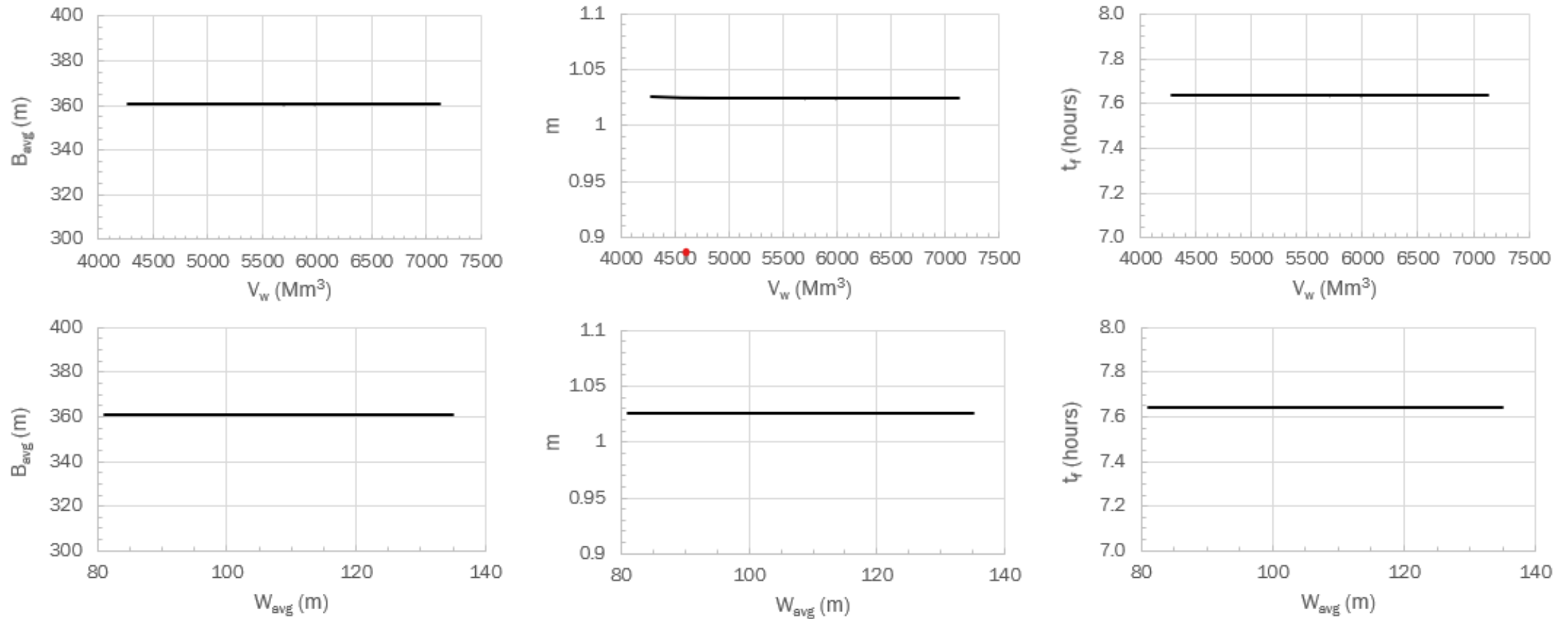
FEMA

# Hirakud Dam Profile Traces



FEMA

## Hirakud Dam Profile Traces (2)



FEMA

## Contact Information

David C. Froehlich, Ph.D., P.E., BC.WRE  
Consulting Water Resources Engineer  
Cary, North Carolina  
[dcfroehlich@aol.com](mailto:dcfroehlich@aol.com)



**FEMA**