Joslyn, Cliff A: (2004) "GIT Analysis of the Crushable Foam Experiment and Simulations", Los Alamos Technical Report LAUR 04-6207

GIT Analysis of the Crushable Foam Experiment and Simulations

Cliff Joslyn*

August, 2004

Contents

Abstract

We provide an analysis of the uncertainty quantification of the "Crushable Foam" experiment and simulation from the perspective of random intervals. We represent the measured data as fuzzy numbers, random intervals, and minmaxmean p-boxes. We compare with simulated data from two sources.

We've been working with the "crushable foam" model setup for a while now. See [4] for full technical details, including the data sets used here. We will assume from this point forward that the reader is familiar with these sources, as well as the fundamentals of the mathematics of Generalized Information Theory (GIT) from Joslyn's perspective [6, 7, 8].

1 Data

In this analysis, we concern ourselves only with the PAC feature for configuration one with drop height $D_I = 13$ in and foam thickness $h = 0.25$ in. Table 1 shows measured sensors readings from the drop table for all three sensors [4]. Table 2 shows the results at the three sensors simulated by the finite element (FE) model under three different preload settings [4]. Finally, table 3 shows three single-point values calculated by the single degree-of-freedom (SDOF) analytical model for three different material models [4].

*Knowledge Systems and Computational Biology Team, Computer and Computational Sciences, Mail Stop B265, Los Alamos National Laboratory, Los Alamos, NM 87545, USA, joslyn@lanl.gov, http://www.c3.lanl.gov/~joslyn, (505) 667-9096.

Replicate	Sensor 1	Sensor 2	Sensor 3
1	1.457	1.262	1.039
2	1.525	1.218	1.064
3	1.536	1.339	1.237
4	1.604	1.332	1.208
5	1.467	1.221	1.124
6	1.594	1.347	1.217
7	1.574	1.338	1.164
8	1.598	1.284	1.081
9	1.525	1.348	0.972
10	1.407	1.394	0.900

Table 1: Measured PAC results (units 10^{-3} g) for configuration 1 [4].

Preload Setting	Sensor 1	Sensor 2	Sensor 3
Nominal, 187.5 psi	1.6009	1.7590	1.5707
Minimum, 62.5 psi	1.1625	1.2828	1.1376
Maximum, 437.5 psi	2.4101	2.6606	2.3607

Table 2: PAC features (units 10^{-3} g) for finite element simulation [4].

It's evident from inspection that there isn't a lot of similarity between, well, much of anything: while the measured data appear to be within a region, there a lot of variability between replicates and among sensors; the FE data show a wider variation, sometimes intersecting the measured data; and finally the SDOF data are significantly different again.

The questions become: how do we analyze these data according to the principles of Generalized Information Theory (GIT), in particular that portion of GIT rooted in random interval theory [7]?

2 Reminder on Random Intervals and P-Boxes

A finite random interval (all random intervals considered here will be finite) $\mathcal{A} := \{\langle I_j, m(I_j)\rangle\}, 1 \leq j \leq M$ is a finite Dempster-Shafer structure on the Borel field of $\mathbb{R}$, so that each focal element $I_j \subseteq \mathbb{R}$ is a half-open interval, $m(I_j)$ is its evidence function (basic probability assignment), and $\sum_{j=1}^{M} m(I_j) = 1$.

Material Model	PAC
Cubic stiffness whose coefficients are best fitted to the measured PAC and TOA data	0.248
Cubic stiffness best fitted to material data	0.285
Bi-linear stiffness whose coefficients are best fitted to the measured PAC and TOA data	0.857

Table 3: Values of PAC (units 10^{-3} g) calculated by SDOF model for parameters $D_I = 13$ in, $h = 0.25$ in [4].

Given a random interval $\mathcal{A}$, denote its trace as $\rho^{\mathcal{A}}\colon \mathbb{R} \to [0,1]$ with $\rho^{\mathcal{A}}(x) := \mathrm{Pl}(\{x\}) = \sum_{I_j \ni x} m(I_j)$, and its core and support respectively as

$$\mathbf{C}(\mathcal{A}) := \bigcap_{i=1}^{M} I_j, \qquad \mathbf{U}(\mathcal{A}) := \bigcup_{i=1}^{N} I_j.$$

When $\mathcal{A}$ is consistent with $\mathbf{C}(\mathcal{A}) \neq \emptyset$, then $\rho^{\mathcal{A}}(x)$ is a possibility distribution $\pi(x)$ called a possibilistic histogram [5].

Definition 1 (P-Box) [2, 3, 8]

A **p-box** is a structure $\mathcal{B} := \left\langle \underline{B}, \overline{B} \right\rangle$, where $\underline{B}, \overline{B}\colon \mathbb{R} \to [0,1]$ with:

1. $\lim_{x \longrightarrow -\infty} \underline{B}(x) \longrightarrow 0, \quad \lim_{x \longrightarrow \infty} \underline{B}(x) \longrightarrow 1$
2. $\lim_{x \longrightarrow -\infty} \overline{B}(x) \longrightarrow 0, \quad \lim_{x \longrightarrow \infty} \overline{B}(x) \longrightarrow 1$
3. $\underline{B}(x), \overline{B}(x)$ are non-decreasing in x, and
4. $\underline{B} \leq \overline{B}$.

Given a p-box $\mathcal{B}$, define its trace $\rho^{\mathcal{B}}\colon \mathbb{R} \to [0,1]$ as $\rho^{\mathcal{B}} := \overline{B} - \underline{B}$.

$\underline{B}$ and $\overline{B}$ are interpreted as bounds on cumulative distribution functions (CDFs). In other words, $\mathcal{B} = \left\langle \underline{B}, \overline{B} \right\rangle$ can be identified with the set of all functions $\{F : \underline{B} \leq F \leq \overline{B}\}$ such that F is the CDF of some probability measures Pr on $\mathbb{R}$. In this way, each p-box defines such a class of probability measures.

Each random interval $\mathcal{A}$ determines a unique p-box $\mathcal{B}(\mathcal{A}) := \langle \mathrm{BEL}, \mathrm{PL} \rangle$ [8], where

$$\mathrm{BEL}(x) := \mathrm{Bel}([-\infty, x)), \quad \mathrm{PL}(x) := \mathrm{Pl}([-\infty, x)),$$

are the cumulative belief and plausibility defined by Yager [9]. Conversely, any given p-box $\mathcal{B}$ determines only an equivalence class of corresponding random intervals [8]. Joslyn and Ferson have shown [8] that $\rho^{\mathcal{A}} = \rho^{\mathcal{B}(\mathcal{A})}$, that is, the random interval-trace of a random interval is the same as the p-box-trace of its p-box.

P-boxes are frequently determined based on information available about a quantity which partially constrain a distribution F. For example, Fig. 1 shows six p-boxes with increasing constraint. For `x.1`, shown in black, we know that $\mathbf{U}(F) = [.5, 1.5]$, where $\mathbf{U}$ is the support of F, and thus $\mathcal{B}$ is effectively just that interval. Denote this as $\mathcal{B}([.5, 1.5])$. Any CDF F which falls inside $\mathcal{B}([.5, 1.5])$ is consistent with this information, which we denote as $F \in \mathcal{B}([.5, 1.5])$.

For `x.2`, shown in cyan, we additionally know that the first moment of $F = 1.0$, which then cuts off the "corners" from that box. For `x.3`, shown in teal, we instead know that the mode $= 1.0$, which provides a stronger constraint. For `x.4`, shown in blue, in addition to min, max, and mean, we also have a variance of .1. `x.5`, shown in red, is a single distribution F, the uniform on $[.5, 1.5]$, which as the single maximum entropy distribution member of $\mathcal{B}$, denoted $U([.5, 1.5]) \in \mathcal{B}([.5, 1.5])$, satisfies the demmands of a Bayesian analysis. Finally, `x.6`, shown in green, is the single *number* 1, which can be interpreted in multiple ways, for example the Dirac distribution with infinite mass at $x = 1$, or the degnerate interval $[1, 1]$. Note that also, in this sense, $1 \in \mathcal{B}$.

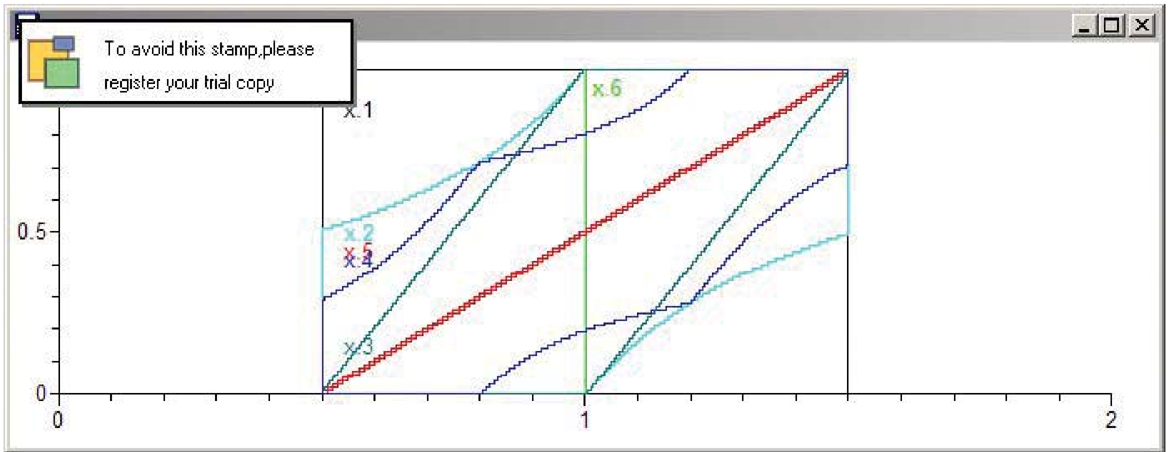

Figure 1: P-boxes for `x.1 = minmax( .5, 1.5 )`, `x.2 = minmaxmean( .5, 1.5, 1.0 )`, `x.3 = minmaxmode( .5, 1.5, 1.0 )`, `x.4 = minmaxmeanvar( .5, 1.5, 1.0, .1 )`, `x.5 = uniform( .5, 1.5 )`, and `x.6 = 1`.

3 Fuzzy Number Representation

Our first observation is that each replicate $r^i := \langle r^i_1, r^i_2, r^i_3 \rangle, 1 \leq i \leq N$, where for us $N = 10$, of the measured data is specified by three points, which represent the PAC readings for the three sensors for a single drop. Since each triple has a minium, maximum, and some kind of "intermediate" value (not *necessarily* a mean or mode), it is natural to create a permutation $r^i_{j1} \leq r^i_{j2} \leq r^i_{j3}$ by sorting, and then let $r'^i := \left\langle r^i_{j_1}, r^i_{j_2}, r^i_{j_3} \right\rangle$, which are shown in Table 4[1]. It will also be useful to construct the mean of the sorted replicates

$$r'^{\text{avg}} = \langle r'_1, r'_2, r'_3 \rangle := \frac{\sum_{i=1}^N r'^i}{N} = \frac{\left\langle \sum_{i=1}^N r^i_{j_1}, \sum_{i=1}^N r^i_{j_2}, \sum_{i=1}^N r^i_{j_3} \right\rangle}{N} = \langle 1.1006, 1.3083, 1.5287 \rangle .$$

The most parsiminious representation of a sorted replicate r'^i is to consider it as a triangular fuzzy number. To be precise, using a common notation [1], let $\widetilde{r}^i := \left[r^i_{j_1}, r^i_{j_2}, r^i_{j_3} \right] \mathrel{\widetilde{\subseteq}} \mathbb{R}$ by linearly interpolating up from membership zero at the min $r^i_{j_1}$ and max $r^i_{j_3}$ to membership 1 at the intermediate value $r^i_{j_2}$. We thereby create the sorted matrix of fuzzy numbers $\widetilde{R} = [\widetilde{r}^i] := \left[r^i_{j_1}, r^i_{j_2}, r^i_{j_3} \right]$, which is effectively equivalent to Table 4.

The triangles for each replicate are show in Fig. 2. Note that we while we're using the same display mode as for the p-boxes in Fig. 1, these structures are *not* p-boxes of the form in Def. 1, but are rather the membership functions of fuzzy quantities $\mu\colon \mathbb{R} \to [0, 1]$.

Note the fairly wide range of variation, including replicate $\widetilde{r}^{10}$, which is quite skewed. It's natural

[1] We are somewhat chagrined to note that Table 4 is actually just a reordering of the whole columns of Table 1. In other words, there is a systematic bias in the PAC data such that Sensor 1 always reads higher than Sensor 2 which itself always reads higher than Sensor 3. While this does not show our method to its best advantage, there is nothing significant in this for the demonstration here.

Replicate	Sensor 1	Sensor 2	Sensor 3
1	1.039	1.262	1.457
2	1.064	1.218	1.525
3	1.237	1.339	1.536
4	1.208	1.332	1.604
5	1.124	1.221	1.467
6	1.217	1.347	1.594
7	1.164	1.338	1.574
8	1.081	1.284	1.598
9	0.972	1.348	1.525
10	0.900	1.394	1.407

Table 4: Measured PAC results (units 10^{-3} g) for configuration 1, sorted for construction of triangular fuzzy numbers.

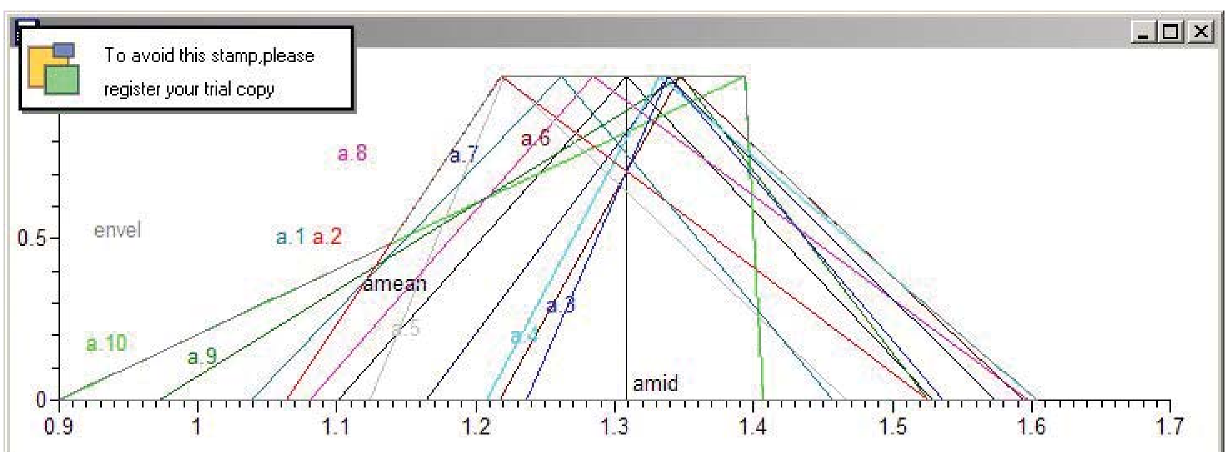

Figure 2: Fuzzy triangular representation of measured data.

to compute the literal mean of these fuzzy numbers, which in this case is equivalent to constructing

$$\widetilde{r}^{\text{avg}} := \frac{\sum_{i=1}^{N} \widetilde{r}^i}{N} = [r'_1, r'_2, r'_3] = [1.1006, 1.3083, 1.5287],$$

which is shown in black in Fig. 2 as `amean`. Naturally, $\widetilde{r}^{\text{avg}}$ is highly symmetrical, since between replicates i each point is iid, thus we can expect the quantity

$$\left| \left(r^i_{j2} - r^i_{j1} \right) - \left(r^i_{j3} - r^i_{j2} \right) \right|,$$

expressing the differences between the sizes of the left and rights sides of the triangles, to be iid around zero. Also shown is r'_2, the central value of $\widetilde{r}^{\text{avg}}$, as `amid`.

Finally, shown in Fig. 2 in dark gray as `envel` is what RiskCalc calls the "envelope" of the $\widetilde{r}^i$, but is rather the fuzzy union

$$\widetilde{r}^{\cup} := \bigcup_{i=1}^{N} \widetilde{r}^i,$$

using the standard maximum t-conorm $\sqcup = \vee$. In this case, $\widetilde{r}^{\cup}$ is actually a fuzzy interval [1] identified by the "convex hull" of all the $\widetilde{r}^i$.

4 Interval and Random Interval Representation

A fuzzy number interpretation of the replicates r^i treats them as expressing a vague outcome. However, it thus doesn't consider the stochastic variation present between the replicates. We can include this factor by moving to a random set and random interval approach.

For each replicate r^i, construct its enveloping interval $\bar{r}^i := [r^i_{j_1}, r^i_{j_3}]$. Then the structure

$$\mathcal{A} := \left\{ \left\langle \bar{r}^i, m\left(\bar{r}^i\right) \right\rangle \right\}, \quad 1 \leq i \leq N,$$

is the random interval constructed by weighting each interval uniformly, so that $\forall 1 \leq i \leq N, m\left(\bar{r}^i\right) = 1/N = 1/10$.

Fig. 3 shows the details of these intervals and the constructed random interval and p-box, in particular:

- For each interval, $\bar{r}^i$ is shown as a pair of vertical bars.
- The mean interval

  $$\bar{r}^{\text{avg}} := \frac{\sum_{i=1}^{N} \bar{r}^i}{N} = [\widetilde{r}^{\text{avg}}_1, \widetilde{r}^{\text{avg}}_3] = [1.1006, 1.5287],$$

  which is exactly the support of the mean fuzzy interval $\mathbf{U}(\widetilde{r}^{\text{avg}})$, and is shown as `bmean` in black.
- Note that $\mathcal{A}$ is consistent in that $\mathbf{C}(\mathcal{A}) \neq \emptyset$. Therefore the trace $\rho(x) = \sum_{\bar{r}^i \ni x} m(\bar{r}^i)$ is a possibility distribution $\pi(x) = \rho(x)$, which is shown as `cons` in purple.
- The full p-box $\mathcal{B}(\mathcal{A})$ is shown as `m` in olive. We can point it out especially as the pair of stepwise-constant traces trending from zero on the left to one on the right.

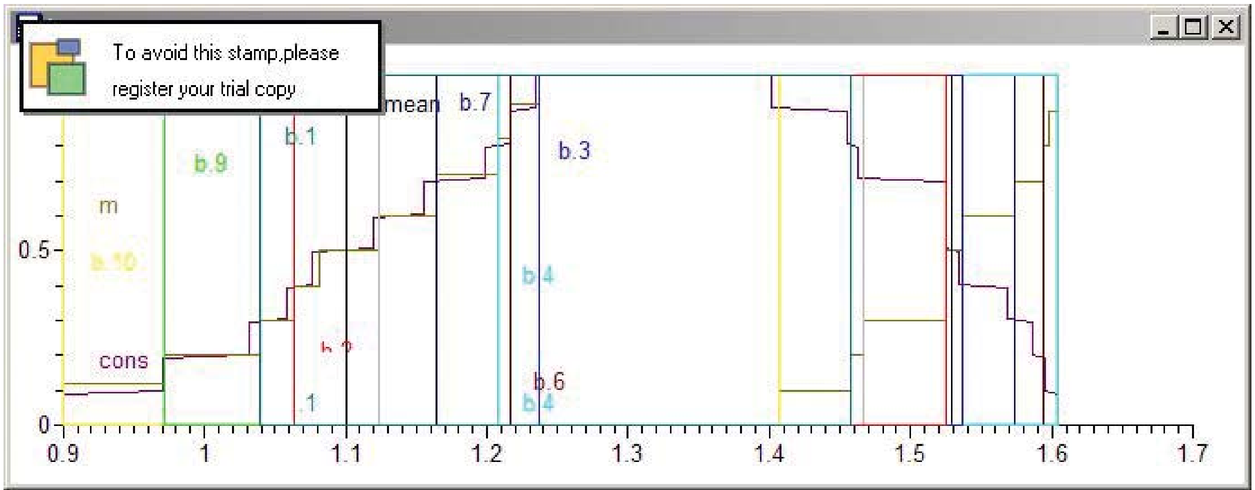

Figure 3: Intervals $\bar{r}^i$, their possibility distribution and p-box.

Note that

$$\forall x \leq \max_{y \in \mathbf{C}(\mathcal{A})} y, \quad \mathrm{BEL}(x) = 0 \text{ and } \mathrm{PL}(x) = \pi(x),$$

and $\forall x \geq \mathbf{C}(\mathcal{A}), \pi(x)$ and $\mathrm{BEL}(x)$ are inverses reflected around .5.

Fig. 4 shows some summary information about the p-box structure, including:

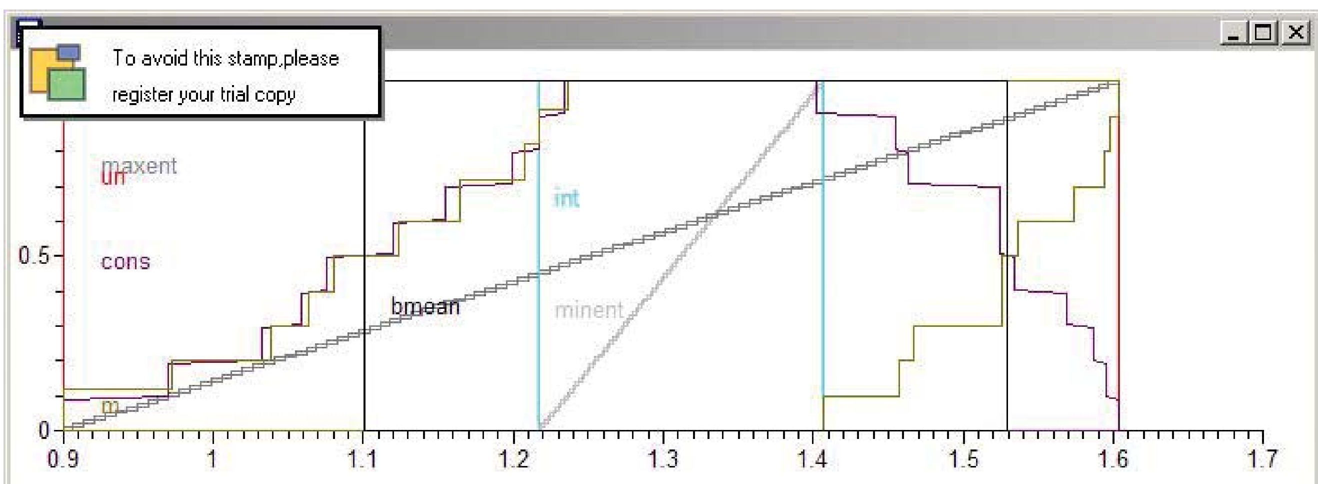

Figure 4: Summary information about p-box representation.

- The mean interval $\bar{r}^{\mathrm{avg}}$, the envelope $\tilde{r}^{\cup}$, the possibility distribution $\pi(x)$, and p-box $\mathcal{B}(\mathcal{A})$ are shown again.
- The core and support respectively $\mathbf{C}(\mathcal{A}), \mathbf{U}(\mathcal{A})$ are shown as `un` in red and `int` in cyan.
- We have selected two single cumulative distributions which are in some sense optimal for our situation. In particular, we can construct the uniform distributions $U(\mathbf{C}(\mathcal{A})) = U(1.217, 1.407)$

and $U(\mathbf{U}(\mathcal{A})) = U(0.900, 1.604)$ respectively, which are shown as `maxent` and `minent` respectively. Despite the choice of names (read nothing of significance into these), neither of these is the distribution $F \in \mathcal{B}(\mathcal{A})$ with maximal entropy, which would itself be a justified representative distribution from $\mathcal{B}(\mathcal{A})$.

5 MinMaxMean P-Box Representation

The random interval analysis above ignores the intermediate value $r^i_{j_2}$ for each replicate, whereas this could be interpreted as a mean value for each replicate, and thus we can represent each as the p-box $\texttt{minmaxmean}\left(r^i_{j_1}, r^i_{j_3}, r^i_{j_2}\right)$. These are shown in Fig. 5, along with the mean such p-box

$$\texttt{minmaxmean}\left(r^{\text{avg}}_{j_1}, r^{\text{avg}}_{j_3}, r^{\text{avg}}_{j_2}\right) = \texttt{minmaxmean}\,(1.1006, 1.5287, 1.3083)$$

shown as `mmean`.

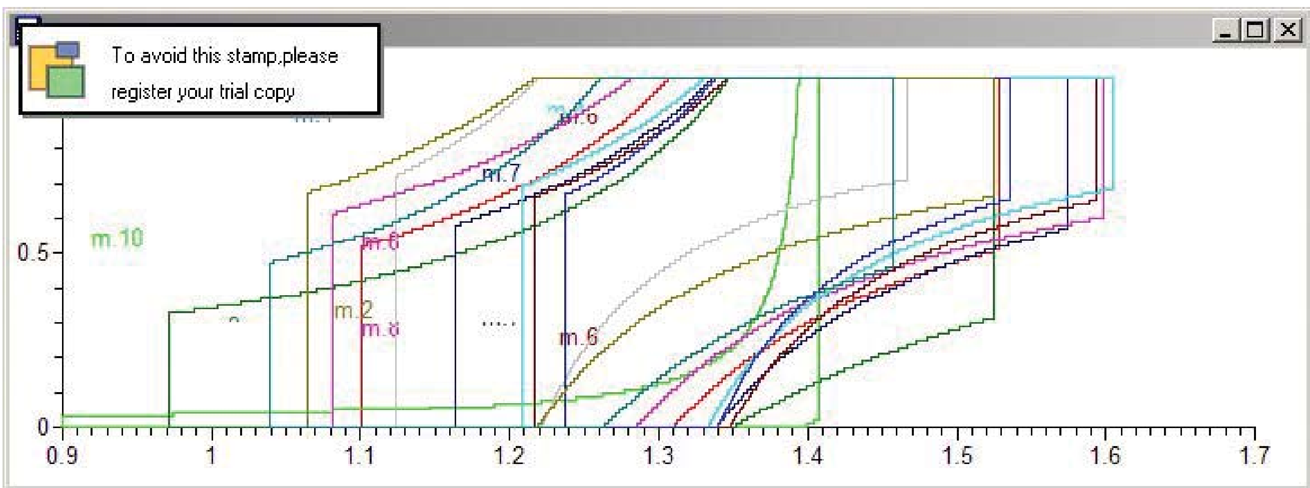

Figure 5: $\texttt{minmaxmean}(r'^i)$.

6 Summary and Comparison with Simulated Data

Fig. 6 shows summary information on all the representations provided, including:

Mean Fuzzy Number: $\tilde{r}^{\text{avg}}$ as `amean` in black.

Intermediate Value: f'_2 as `amid` in dark green.

Fuzzy Union: $\tilde{r}^{\cup}$ as `envel` in teal.

Mean Interval: $\bar{r}^{\text{avg}}$ as `bmean` in blue.

Core: The core of the random interval $\mathbf{C}(\mathcal{A})$ as `int` in cyan.

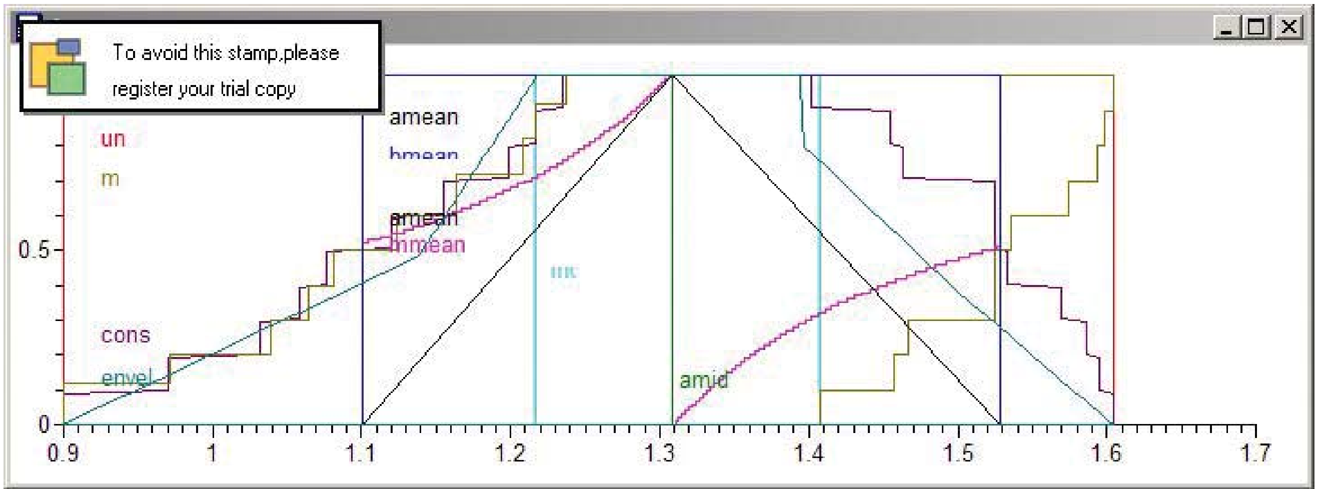

Figure 6: Summary of measured data.

Support: The support of the random interval $\mathbf{U}(\mathcal{A})$ as `un` in red.

Possibility Distribution: The possibility distribution $\pi(x)$ as `cons` in purple.

Random Interval P-Box: The full p-box $\mathcal{B}(\mathcal{A})$ as `m` in olive.

Mean MinMaxMean P-Box: The mean MinMaxMean p-box `minmaxmean` $\left(r_{j_1}^{\text{avg}}, r_{j_3}^{\text{avg}}, r_{j_2}^{\text{avg}}\right)$ as `mmean` in magenta.

Note the following:

- The average fuzzy number is completely contained by the possibility distribution trace: $\widetilde{r}^{\text{avg}} \leq \pi$, although it is not clear whether this is necessary or not.
- The intermediate point of the average fuzzy number $r_2' = 1.3083 \in \mathbf{C}(\mathcal{A})$.
- $\mathbf{C}(\mathcal{A}) \subseteq \bar{r}^{\text{avg}} \subseteq \mathbf{U}(\mathcal{A})$, which is always true.
- I believe that it is guaranteed that if $\mathbf{C}(\mathcal{A})$ exists, then $U(\mathbf{C}(\mathcal{A})) \in \mathcal{B}(\mathcal{A})$, but this seems unlikely for $U(\mathbf{U}(\mathcal{A}))$, and indeed, from inspection it appears that it might just nick the corner of $\mathcal{B}(\mathcal{A})$ precariously.
- Finally, the intervals $\bar{r}^i, \bar{r}^{\text{avg}}, \mathbf{C}(\mathcal{A}), \mathbf{U}(\mathcal{A})$ and the possibility distribution $\pi(x)$ are fuzzy quantities of the form $\mu: \mathbb{R} \to [0,1]$, while the random interval $\mathcal{A}$ and the distributions $U(\mathbf{C}(\mathcal{A}))$ and $U(\mathbf{U}(\mathcal{A}))$ are effectively p-boxes $\mathcal{B}$ of the form defined in (1).

This is all very nice, but what we're missing are the *simulated data* to compare the measured data to. These are shown together with all the summary measured data in Fig. 7.

Represent the FE data from Table 2 as replicates $f^j, 1 \leq j \leq 3$. So by similar reasoning to the above, we know that by first principles, it's appropriate to represent these as triangular fuzzy numbers denoted $\widetilde{f^j} \widetilde{\subseteq} \mathbb{R}$, shown in Fig. 7 as `fe.j`. Similarly, the SDOF data from Table 3 are appropriately represented as single numbers $s^k \in \mathbb{R}, 1 \leq k \leq 3$, shown as `sd.k`.

What we're ultimately interested in are measures of the similarity between the FE, SDOF, and the measured data, both pairwise among these data sets, and among all three simultaneously. This is an analytical question, which to my knowledge doesn't have a specific methodology yet. Issues include:

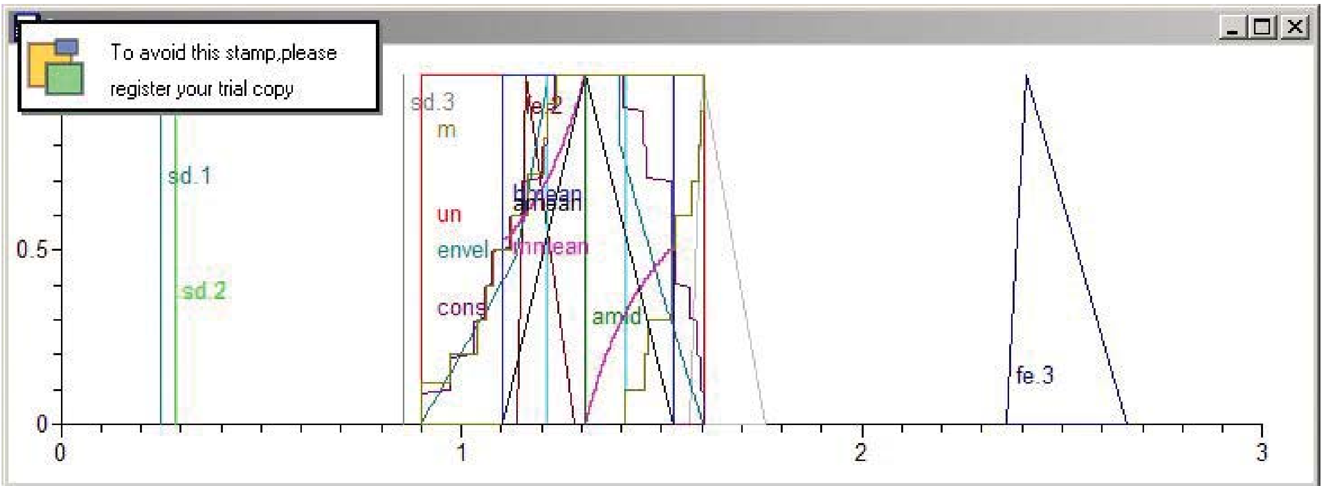

Figure 7: Summary of measured and simulated data.

Apples to Apples: Given the replicates f^j represented as fuzzy numbers $\widetilde{f^j}$, presumably they should be compared to the average fuzzy number $\widetilde{r}^{\text{avg}}$. On the other hand, would it be wise to instead represent the f^j as a random interval, albeit almost degenerate, in order to be able to compare it properly to $\mathcal{A}$?

Measures of Difference: The ultimate goal is measures of distance or overlap among the quantities. This, in turn, involves at least two components:

Overlap as Commonality of Support: We can calculate the amounts of literal overlap between the various structures, for example their cores and supports. We note, for example, that

$$\mathbf{U}\left(\widetilde{f^1}\right) \cap \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right) \neq \emptyset, \quad \mathbf{U}\left(\widetilde{f^2}\right) \subseteq \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right), \quad \mathbf{U}\left(\widetilde{f^3}\right) \cap \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right) = \emptyset,$$

$$\forall 1 \leq k \leq 3, s^k \notin \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right),$$

and thus measures like

$$\left|\mathbf{U}\left(\widetilde{f^j}\right) \cap \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right)\right|, \qquad \frac{\left|\mathbf{U}\left(\widetilde{f^j}\right) \cap \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right)\right|}{\left|\mathbf{U}\left(\widetilde{f^j}\right) \cup \mathbf{U}\left(\mathcal{B}(\mathcal{A})\right)\right|}$$

are available to indicate this.

Measures of Distance: In general, points, intervals, fuzzy intervals, p-boxes, and random intervals all have convolutions available for operators $\oplus \in \{+, -, \times, \div, \hat{}\}$ [8], and in particular for $\oplus = -$, so that, for example, we can calculate

$$\forall 1 \leq j \leq 3, \left| \widetilde{f}^j - \widetilde{r}^{\text{avg}} \right|$$

Acknowledgements

Calculations were performed in RAMAS RiskCalc 4.0[2] [2].

References

[1] Dubois, Didier and Prade, Henri: (1978) "Operations on Fuzzy Numbers", *Int. J. Systems Science*, v. **9**, pp. 613-626

[2] Ferson, Scott: (2002) *RAMAS Rick Calc 4.0 Software: Risk Assessment with Uncertain Numbers*, CRC Press

[3] Ferson, Scott; Kreinovich, V; Ginzburg, LM; D.S. Myers and K. Sentz: (2002) "Constructing Probability Boxes and Dempster-Shafer Structures", *SAND Report 2002-4015*, Sandia National Lab, Albuquerque NM, `http://www.sandia.gov/epistemic/Reports/SAND2002-4015.pdf`

[4] Hemez, Francois M and Ben-Haim, Yakov: (2004) "Info-Gap Robustness for the Correlation of Tests and Simulations of a Non-Linear Transient", *Mechanical Systems and Signal Processing*, v. **18**, pp. 1443-1467

[5] Joslyn, Cliff: (1997) "Measurement of Possibilistic Histograms from Interval Data", *Int. J. General Systems*, v. **26**:1-2, pp. 9-33

[6] Joslyn, Cliff: (2004) "A Formerly Breathless Introduction to Generalized Information Theory", LAUR 04-4078, `ftp://ftp.c3.lanl.gov/pub/users/joslyn/shortcourse.pdf`

[7] Joslyn, Cliff and Booker, Jane: (2004) "Generalized Information Theory for Engineering Modeling and Simulation", in: *Engineering Design Reliability Handbook*, ed. E Nikolaidis and D Ghiocel, CRC Press, in press

[8] Joslyn, Cliff and Ferson, Scott: (2004) "Approximate Representations of Random Intervals for Hybrid Uncertainty Quantification", in: *4th Int. Conf. on Sensitivity Analysis of Model Output (SAMO 04)*, `ftp://ftp.c3.lanl.gov/pub/users/joslyn/samo04.pdf`

[9] Yager, Ronald R: (1986) "Arithmetic and Other Operations on Dempster-Shafer Structures", *Int. J. Man-Machine Studies*, v. **25**, pp. 357-366

[2] `http://www.ramas.com/riskcalc.htm`