Somewhat Surprisingly, (Subjective) Fuzzy Technique Can Help to Better Combine Measurement Results and Expert Estimates into a Model with Guaranteed Accuracy: Digital Twins and Beyond

Niklas Winnewisser[1], Michael Beer[1],
Olga Kosheleva[2], and Vladik Kreinovich[2]

[1]Institute for Risk and Reliability, Callinstraße 34
Leibniz University Hannover, 30167 Hannover, Germany
winnewisser@irz.uni-hannover.de, beer@irz.uni-hannover.de

[2]Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA
olgak@utep.edu, vladik@utep.edu

1. Need for digital twins

- It is desirable:
  - to predict how a building, an airplane, a ship, or any other structure will behave under different circumstances, and
  - thus, e.g., to make sure that they will remain functional under expected extreme conditions.
- It is desirable:
  - to predict what will happen if we close some roads for repairs, etc., and
  - thus, to select the schedule of needed road maintenance that minimally disturbs the traffic.
- In all these cases, we need to have a digital model of the corresponding system.

2. Need for digital twins (cont-d)

- This model should be as accurate as possible.
- Ideally, the model and the system should be as close to each other as twins.
- Such models are therefore called *digital twins*.

3. Need to take into account expert knowledge

- We want to make the model of a system as accurate as possible.
- So, we need to incorporate all available information about this system into this model.
- Usually, a significant part of this information comes in precise terms:
  - in terms of the values of some quantities characterising the system, values that can be obtained by measuring these quantities, and
  - in terms of equations relating these quantities and/or describing how the values of these quantities charge with time, etc.
- In some cases, this precise knowledge is sufficient to build an accurate model of a system.
- For example, Newton's equations accurately predict the positions of all celestial bodies hundreds of years ahead.
- We know on what days there will be a solstice in the 30th century.
- However, in many other cases, the precise knowledge is not sufficient.

4. Need to take into account expert knowledge (cont-d)

- In such cases:
  - we do not have a fully automated controllers for these systems,
  - we still need expert controllers.
- The experience of these controllers and other experts provide additional knowledge.
- The most extreme case is probably medicine, where human doctors are still irreplaceable.
- It is therefore necessary to incorporate expert knowledge into our models.

5. How can we avoid subjectivity when we incorporate expert knowledge into a model

- Experts have biases.
- Also, expert estimates come with uncertainty that varies from one expert to another.
- In other words, values produced by experts are somewhat subjective.
- At first glance, it may seem that:
  - if we add expert knowledge into a model,
  - the model becomes somewhat subjective – and thus, less reliable.
- However, the situation is not so bad if we take into account that:
  - the ultimate sources of precise information – measuring instruments
  - also have individual biases and varying degrees of uncertainty.

6. How can we avoid subjectivity when we incorporate expert knowledge into a model (cont-d)

- To take that into account, we *calibrate* each measuring instrument – by comparing:
  - the values measured by this instrument and
  - the values measured in the same situation by a much more accurate instrument.
- In measurement theory, such a much more accurate instrument is called *standard* measuring instrument.
- In each such measurement, the measured quantity has some actual value x (that we do not know).

7. How can we avoid subjectivity when we incorporate expert knowledge into a model (cont-d)

- By applying our measuring instrument to this quantity, we get a value $\widetilde{x}$ which is, in general, different from x.
- There is a *measurement error* $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$.
- By applying the standard measuring instrument to this same physical quantity, we get a different measurement result $\widetilde{X}$.
- The fact that the standard measuring instrument is much more accurate means that its measurement error is much smaller:
$$\Delta X \stackrel{\text{def}}{=} \widetilde{X} - x \ll \Delta x.$$
- The difference $\widetilde{x} - \widetilde{X}$ between the two measurement results can be represented in the following equivalent way:
$$\widetilde{x} - \widetilde{X} = \widetilde{x} - \widetilde{X} + x - x = (\widetilde{x} - x) - (\widetilde{X} - x) = \Delta x - \Delta X.$$
- Since $\Delta X \ll \Delta x$, we can safely ignore the term ΔX in the above formula and conclude that, with good accuracy, $\Delta x \approx \widetilde{x} - \widetilde{X}$.

8. How can we avoid subjectivity when we incorporate expert knowledge into a model (cont-d)

- In this approximation:
  - by comparing the measurement results of two instruments,
  - we get one of the possible values of the measurement error Δx of the original measuring instrument.
- If we repeat this procedure many times, with different objects, we get a sample of different values Δx.
- Based on this sample, we can find the probability distribution of the measurement errors.
- A natural idea is to apply the same approach to a human expert.
- We ask the human expert to provide an estimate $\widetilde{x}$ for the value of some quantity.
- We compare this estimate withe the result $\widetilde{X}$ of accurately measuring this quantity.

9. How can we avoid subjectivity when we incorporate expert knowledge into a model (cont-d)

- We use the sample formed by the differences $\widetilde{x}-\widetilde{X} \approx \Delta x$ to determine the probability distribution of the expert's estimation errors

$$\Delta x = \widetilde{x} - x.$$

- Once we know this distribution, we can use each expert as an additional measuring instrument.
- And we add no more subjectivity than when we use sensors.

10. Can we do better?

- The main idea of dealing with expert knowledge is to treat an expert as a kind of a measuring instrument.
- However, from the viewpoint of information, there is an important different between experts and measuring instruments.
- A measuring instrument only produces a single number $\widetilde{x}$.
- In contrast, in many cases, an expert:
  - not only produces an estimate $\widetilde{x}$,
  - he/she can also describe how possible are different deviation of this estimate from the actual value of the estimated quantity.
- For example, if the expert's estimate is 1.0, the expert may add that:
  - the value 0.9 is very possible,
  - the value 0.8 is somewhat possible, and
  - the value 0.7 is hardly possible.

11. Can we do better (cont-d)

- The expert may also say that the difference between his/her estimate and the actual value is small, etc.
- How can we incorporate this subjective additional information into a model with guaranteed accuracy?
- This is what we discuss in this talk.

12. How is imprecise expert information described now: a brief reminder about fuzzy techniques

- The problem with the expert information about uncertainty – of the type described above – is that it is described:
  - not in precise terms, but
  - by using imprecise ("fuzzy") words from natural language, like "very possible", "somewhat possible", "small", etc.
- Zadeh emphasized the need to take this knowledge into account when designing computer-based algorithms – e.g., in control.
- Zadeh came up with a special technique for translating this knowledge into precise terms, a technique that he called *fuzzy*.
- Zadeh's main idea was to as the expert to assign:
  - to each possible number,
  - a degree – from the interval $[0, 1]$ – to which the expert believes it be possible that this number is the actual value of the quantity.

13. How is imprecise expert information described now: a brief reminder about fuzzy techniques (cont-d)

- Degree 1 means that the expert absolutely believes that this number is possible.
- Degree 0 means that the expert absolutely believes that this number is *not* possible.
- Degrees between 0 and 1 means that the expert is not 100% sure.
- Based on the expert's answers to these questions, we get a function $\mu(x)$ that assigns, to each value x, the corresponding degree.
- This function is known as a *membership function*, or, alternatively, as a *fuzzy set*.
- In applications, we not only deal with basic statements A, B, ...
- Expert knowledge is often formulated in terms of logical combinations of these statements such as $A \,\&\, B$, $A \vee B$, $\neg A$, etc.

14. How is imprecise expert information described now: a brief reminder about fuzzy techniques (cont-d)

- In many practical situations:
  - all we know about these statements is their degree of confidence a, b, ..., and
  - we need to estimate the expert's degree of confidence in the corresponding logical combinations based on these known degrees.
- The corresponding algorithms $f_\&(a,b)$, $f_\vee(a,b)$, $f_\neg(a)$ – known as *logical operations* – form what is known as *fuzzy logic.*
- There are many such operations.
- For example, the most widely used "or"-operation is

$$f_\vee(a,b) = \max(a,b).$$

- This is the operation that we will use in this talk – but there are others, e.g., $f_\vee(a,b) = a + b - a \cdot b$.
- In the following, we will explain why we use this specific operation.

15. But fuzzy sets are subjective, so what can we do?

- Fuzzy sets are very useful when we need, e.g., to transform natural-language expert rules into a precise control strategy.
- However, these estimates are subjective.
- So we cannot directly incorporate them into a model with guaranteed accuracy.
- So what can we do?

16. A natural idea: using calibration

- We have already encountered the situation when we need to incorporate subjective information into a model with guaranteed accuracy.
- This is what we described earlier.
- There, the solution was to calibrate the measuring instrument – and, similarly, to calibrate the expert.
- It is therefore natural to try a similar idea here.
- Let us describe how this can be done.
- For this purpose, let us recall an alternative way of representing fuzzy sets – a way that is used for computations related to fuzzy sets.

17. α-cuts: a computational-friendly representation of a fuzzy set

- For computations with fuzzy sets, it is convenient to generate α-*cuts* corresponding to different values $\alpha \in [0, 1]$:
- For $\alpha > 0$, the α-cut is defined as $\mathbf{x}(\alpha) \stackrel{\text{def}}{=} \{x : \mu(x) \geq \alpha\}$.
- For $\alpha = 0$, the corresponding α-cut is defined as
$$\mathbf{x}(0) \stackrel{\text{def}}{=} \overline{\{x : \mu(x) > 0\}}.$$
- Here $\overline{S}$ means the closure of the set S, i.e., the result of adding, to the set S, all its limit points.
- Once we know all the α-cuts, we can reconstruct the original membership function as $\mu(x) = \sup\{\alpha : x \in \mathbf{x}(\alpha)\}$.
- For most types of fuzzy information, the corresponding degree $\mu(x)$ first increases, then decreases.
- In this case, each α-cut is an interval $\mathbf{x}(\alpha) = [\underline{x}(\alpha), \overline{x}(\alpha)]$.

18. What is the meaning of an α-cut: let us brainstorm

- By definition of the α-cut:
  - for each α, and for each number x outside the α-cut interval $\mathbf{x}(\alpha)$,
  - the expert's degree of confidence $\mu(x)$ that this number is a possible value of the estimated quantity is smaller than α.
- What is the expert's degree of confidence that "*some* number x outside this interval is a possible value of the estimated quantity"?
- The statement S whose degree we want to estimate means that:
  - either one of these numbers x_1 is a possible value of the estimated quantity,
  - *or* another of these numbers x_2 is a possible value of the estimated quantity, etc.
- In other words, the statement S is, in effect, an "or"-combination of statements corresponding to individual values x.

19. What is the meaning of an α-cut (cont-d)

- There are infinitely many numbers outside the α-cut interval, so we have an "or"-combination of infinitely many such statements.
- If we use the max-operation $f_\vee(a, b)$, then we simply take the supremum of all the numbers which are smaller than α.
- This is, of course, exactly α.
- This, by the way, explains why we use maximum and not any other "or"-operation like $f_\vee(a, b) = a + b - a \cdot b$.
- One can easily check that if we apply this operation to many similar numbers, the results start tending to 1.
- In the limit, when we take all infinite numbers outside the α-cut interval into account, we get a meaningless degree 1.
- Same thing happens with many other known "or"-operations.
- So, the expert's degree of confidence that one of the numbers x outside the α-cut interval is a possible value of the estimated quantity is α.

20. What is the meaning of an α-cut (cont-d)

- Thus, the expert's degree of confidence in the opposite statement:
  - that none of the numbers outside the α-cut interval is a possible value of the estimated quantity,
  - i.e., that the actual value is inside the α-cut interval,
  - is equal to $f_\neg(\alpha)$.
- So, we arrive at the following conclusion.
- For each α, the expert's degree of confidence that the actual value x is in the α-cut interval $[\underline{x}(\alpha), \overline{x}(\alpha)]$ is equal to $f_\neg(\alpha)$.

21. What we need to know to make the expert statement more objective

- Fuzzy techniques use "degree of confidence", which is a very subjective value.
- A natural objective counter-part of this subjective notion is probability.
- So, what we need is to come up with *probabilities* that x is in each α-cut.

22. How can we transform subjective degrees into objective probabilities: a natural idea

- As we have mentioned, a natural idea of transforming subjective values into objective probabilities is by using calibration.
- Specifically, we ask the expert to provide the membership function for many different cases.
- Based on these membership functions, we come up with the corresponding α-cut intervals.
- For all these cases, we also use a standard measuring instrument to measure the actual value of the corresponding quantity.
- Then, for each α, we can count in what proportion of these cases the actual value x was in the corresponding α-cut interval.
- This proportion $p(\alpha)$ is, thus, an estimate for the (objective) probability that the actual value x is in this expert's α-cut interval.

23. We are almost done

- After the above calibration, we get the function $p(\alpha)$ corresponding to this particular expert.
- Then, for each new expert's estimate, we can conclude that:
  - based on this expert's opinion,
  - the probability that the actual value likes in the corresponding α-cut interval is equal to $p(\alpha)$.
- This *almost* determines the probability distribution, but not yet fully.
- For example, for two close α-cut intervals

$$[\underline{x}(\alpha), \overline{x}(\alpha)] \supseteq [\underline{x}(\alpha + \Delta\alpha), \overline{x}(\alpha + \Delta\alpha)] :$$

  - we know the probability $p(\alpha)$ of the larger interval and
  - we know the probability $p(\alpha + \Delta\alpha)$ of the smaller interval.

24. We are almost done (cont-d)

- Thus, we know that:
  - the probability that the actual value x is in the difference between these two intervals – which consists of two small intervals

$$[\underline{x}(\alpha), \underline{x}(\alpha + \Delta\alpha)] \text{ and } [\overline{x}(\alpha + \Delta\alpha), \overline{x}(\alpha)]$$

  - is equal to the difference $p(\alpha) - p(\alpha + \Delta\alpha)$ between these two probabilities.
- However, this information does not tell us what is the probability of each of the two small intervals.
- It only provides the sum of these two small-interval probabilities.
- To have this additional information, let us use another natural idea.

25. Final idea

- To explain this idea, let us recall that the most frequently experiences probability distribution is normal (Gaussian).
- A specific feature of this distribution is that its probability density $f(x)$ is always positive.
- In principle, all numbers are possible – but, of course, very large numbers are highly improbable.
- In practice, we ignore such practically improbable events and only consider values x that are reasonably probable.
- The degree to which each value x is probable is described by the value $f(x)$ of the corresponding probability density function.
- From this viewpoint:
  - if $f(x) \le f(y)$ and we consider x to be reasonably probable,
  - then y is even more probable – and thus, also has to be considered reasonably probable.

26. Final idea (cont-d)

- So, when we want to select a set to which x belongs to some degree of certainty, a natural idea is:
  - to select some threshold f_0 and
  - to consider all the numbers x for which $f(x) \geq f_0$.
- It is reasonable to assume that the expert-based intervals $[\underline{x}(\alpha), \overline{x}(\alpha)]$ are obtained in exactly this way.
- So, the probability density is exactly the same on both ends of this interval.
- Thus:
  - between the two small intervals $[\underline{x}(\alpha), \underline{x}(\alpha + \Delta\alpha)]$ and $[\overline{x}(\alpha + \Delta\alpha), \overline{x}(\alpha)]$
  - the probability $p(\alpha) - p(\alpha + \Delta\alpha)$ is distributed proportionally to their lengths.

27. Now we are ready

- Now, we are finally ready to formulate the desired calibration-based transformation:
  - of the experts fuzzy opinion
  - into an objective probability distribution.
- Thus, we will be able to incorporate this distribution into a model with guaranteed accuracy.

28. First stage: testing an expert – ideal case

- First, for several (N) cases, we ask the expert not only to provide a single estimate $\widetilde{x}$ for the given quantity, but to also describe:
  - for every real number x,
  - the degree $\mu(x)$ to which this number x can be the actual value of the estimated quantity.
- Based on the expert-produced degrees, for each value $\alpha \in [0,1]$, we form the α-cut $\mathbf{x}(\alpha) = [\underline{x}(\alpha), \overline{x}(\alpha)]$:
  - when $\alpha > 0$, we compute $\mathbf{x}(\alpha) = \{x : \mu(x) \geq \alpha\}$;
  - when $\alpha = 0$, we compute $\mathbf{x}(0) = \overline{\{x : \mu(x) > 0\}}$.

29. Testing an expert – practical solution

- Of course, there are infinitely many real numbers, we can only ask the expert about finitely many of them $x_1 < x_2 < \ldots < x_n$.
- To estimate the values $\mu(x)$ corresponding to intermediate values $x \in (x_i, x_{i+1})$, we can, for example, use linear interpolation

$$\mu(x) = \mu(x_i) + \frac{x - x_i}{x_{i+1} - x_i} \cdot (\mu(x_{i+1}) - \mu(x_i)).$$

- Similarly, there are infinitely many real numbers α on the interval $[0, 1]$.
- We cannot perform computations infinitely many times.
- So, in practice, we can select several possible values

$$0 = \alpha_1 < \alpha_2 < \ldots < \alpha_n = 1.$$

- For example, we can select values $0 < 0.1 < 0.2 < 0.3 < \ldots < 0.9 < 1$.

30. Testing an expert – practical solution (cont-d)

- Then, we only compute the α-cuts for these selected value α:

$$[\underline{x}(\alpha_1), \overline{x}(\alpha_1)] \supseteq [\underline{x}(\alpha_2), \overline{x}(\alpha_2)] \supseteq \ldots \supseteq [\underline{x}(\alpha_n), \overline{x}(\alpha_n)].$$

- For the endpoints of these α-cuts, we have a natural order:

$$\underline{x}(\alpha_1) \le \underline{x}(\alpha_2) \le \ldots \le \underline{x}(\alpha_n) \le \overline{x}(\alpha_n) \le \overline{x}(\alpha_{n-1}) \le \ldots \le \overline{x}(\alpha_2) \le \overline{x}(\alpha_1).$$

31. Second stage: comparing expert's estimates with actual values

- For all these cases, we compare the expert estimates:
  - with the actual value,
  - to be more precise, with the value x provided by the standard measuring instrument.
- For each of the selected values α_i, we count the number N_i of cases in which the actual value x was inside the corresponding α-cut interval

$$[\underline{x}(\alpha_i), \overline{x}(\alpha_i)].$$

- Based on these counts, we compute the ratios $p(\alpha_i) \stackrel{\text{def}}{=} n_i/N$.

32. Final stage: generating the resulting probability distribution

- As a result, we get the following probability density function $f(x)$.
- For values $x \in [\underline{x}(\alpha_i), \underline{x}(\alpha_{i+1})]$, we have

$$f(x) = \frac{p(\alpha_i) - p(\alpha_{i+1})}{(\underline{x}(\alpha_{i+1}) - \underline{x}(\alpha_i)) + (\overline{x}(\alpha_i) - \overline{x}(\alpha_{i+1}))}.$$

- For values $x \in [\underline{x}(\alpha_n), \overline{x}(\alpha_n)]$, we take

$$f(x) = \frac{p(\alpha_n)}{\overline{x}(\alpha_n) - \underline{x}(\alpha_n)}.$$

- For values $x \in [\overline{x}(\alpha_{i+1}), \overline{x}(\alpha_i)]$, we have

$$f(x) = \frac{p(\alpha_i) - p(\alpha_{i+1})}{(\underline{x}(\alpha_{i+1}) - \underline{x}(\alpha_i)) + (\overline{x}(\alpha_i) - \overline{x}(\alpha_{i+1}))}.$$

33. Mathematical comment

- In the limit, when the differences $\alpha_{i+1} - \alpha_i$ tend to 0, the above formulas take the form:

$$f(x) = \frac{\dfrac{dp}{d\alpha}}{\dfrac{d\underline{x}}{d\alpha} - \dfrac{d\overline{x}}{d\alpha}}.$$

- Here, α is the value for which either $\underline{x}(\alpha) = x$ or $\overline{x}(\alpha) = x$, i.e., the value $\alpha = \mu(x)$.

34. A brief conclusion

- We have shown that, somewhat surprisingly:
  - fuzzy techniques – which usually reflect a somewhat subjective expert opinion,
  - can help to transform these estimates into an objective probabilistic form and
  - thus, help to better combine measurement results and expert estimates into a model with guaranteed accuracy.

35. Acknowledgments

This work was supported in part by:

- National Science Foundation grants 1623190, HRD-1834620, HRD-2034030, and EAR-2225395;
- AT&T Fellowship in Information Technology;
- program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- a grant from the Hungarian National Research, Development and Innovation Office (NRDI).