From Gauging Accuracy of Quantity Estimates to Gauging Accuracy and Resolution of Field Measurements: Geophysical Case Study

IrinaPerfilieva\textsuperscript{1}, Roumen Anguelov\textsuperscript{2}, Vladik Kreinovich\textsuperscript{3}, and Matt Averill\textsuperscript{4}

\textsuperscript{1}Inst. for Research and Applications of Fuzzy Modeling 
University of Ostrava, Czech Republic
\textsuperscript{2}Department of Math. and Applied Math. 
Univ. of Pretoria, South Africa
\textsuperscript{3}Department of Computer Science 
\textsuperscript{4}Department of Geological Sciences 
University of Texas at El Paso 
El Paso, Texas 79968, USA 
vladik@utep.edu
1. Traditional Applications of Interval Computations: Reminder

- **Objective:** estimate a difficult-to-measure quantity $y$.
- **Approach:** measure quantities $x_1, \ldots, x_n$ related to $x_i$ by a known dependence $y = f(x_1, \ldots, x_n)$.
- **Fact:** measurements are never absolutely accurate.
- **Conclusion:** the measurement results $\tilde{x}_i$ are, in general, different from the actual (unknown) values $x_i$.
- **Conclusion:** the result $\tilde{y} = f(\tilde{x}_1, \ldots, \tilde{x}_n)$ of data processing differs from $y = f(x_1, \ldots, x_n)$.
- **Frequent situation:** we only know the upper bound $\Delta_i$ on the measurement errors $\Delta x_i \triangleq \tilde{x}_i - x_i$: $|\Delta x_i| \leq \Delta_i$.
- **So:** we only know that $x_i \in x_i \triangleq [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$.
- **Interval computations:** find the corresponding range $y = \{ f(x_1, \ldots, x_n) : x_i \in x_i \}$ of $y$. 
2. In Practice, the Situation is Often More Complex

- **Dynamics**: we measure the values \( v(t) \) of a quantity \( v \) at a certain moment of time \( t \).
- **Spatial dependence**: we measure the value \( v(x, t) \) at a certain location \( x \).
- **Geophysical example**: we are interested in the values of the density at different locations and at different depth.
- **Traditional**: uncertainty in the measured value, \( \tilde{v} \approx v \).
- **New**: uncertainty in location \( x \), \( \tilde{x} \approx x \).
- **Additional uncertainty**: the sensor picks up the “average” value of \( v \) at locations close to \( \tilde{x} \).
- **Question**: how to describe and process the new uncertainty (resolution)?
3. Outline

- **Question** (reminder): how to describe and process uncertainty both
  - in the measured value $\tilde{v}$ and
  - in the spatial resolution $\tilde{x}$?

- **Natural idea**: the answer depends on what we know about the spatial resolution.

- **Possible situations**:
  - we know exactly how the measured values $\tilde{v}_i$ are related to $v(x)$, i.e., $\tilde{v}_i = \int w_i(x) \cdot v(x) \, dx + \Delta v_i$;
  - we only know the upper bound $\delta$ on the location error $\tilde{x} - x$ (this is similar to the interval case);
  - we do not even know $\delta$.

- **What we do**: describe how to process all these types of uncertainty.
4. Situations in Which We Have Detailed Knowledge

- **Fact:** all our information about \( v(x) \) is contained in the measured values \( \tilde{v}_i \).

- **Linearity assumption:** \( \tilde{v}_i = v_i + \Delta v_i \), where:
  - we have \( v_i \overset{\text{def}}{=} \int w_i(x) \cdot v(x) \, dx \); and
  - \( \Delta v_i \) is the measurement error; e.g., \( |\Delta v_i| \leq \Delta_i \).

- **Comment:** \( v_i \) can be viewed as the value of \( v(x) \) at a “fuzzy” point characterized by uncertainty \( w_i(x) \).

- **Description of the situation:** we know the weights \( w_i(x) \).

- **Find:** range \( [y, \bar{y}] \) for \( y \overset{\text{def}}{=} \int w(x) \cdot v(x) \, dx \).

- **LP solution:** minimize (maximize) \( \int w(x) \cdot v(x) \, dx \) under the constraints
  \[
  v_i \overset{\text{def}}{=} \tilde{v}_i - \Delta_i \leq \int w_i(x) \cdot v(x) \, dx \leq \bar{v}_i \overset{\text{def}}{=} \tilde{v}_i + \Delta_i.
  \]
5. Situations With Detailed Knowledge (cont-d)

- **Reminder**: find the range of $\int w(x) \cdot v(x) \, dx$ when $v_i \leq \int w_i(x) \cdot v(x) \, dx \leq \bar{v}_i$.
- **General case**: when no bounds on $v(x)$, bounds are infinite – unless $w(x)$ is a linear combination of $w_i(x)$.
- **In practice** (e.g., in geophysics): $v(x) \geq 0$.
- **Similar**: imprecise probabilities (Kuznetsova, Walley).
- **Solution**: dual LP problem provides guaranteed bounds

$$\underline{v} = \sup \left\{ \sum y_i \cdot v_i : \sum y_i \cdot w_i(x) \leq w(x) \right\};$$

$$\overline{v} = \inf \left\{ \sum y_i \cdot \bar{v}_i : w(x) \leq \sum y_i \cdot w_i(x) \right\}.$$

- **Easier** than in IP: $w_i(x)$ are localized, and we often have $\leq 2$ non-zero $w_i(x)$ at each $x$.
- **Piece-wise linear** $w_i(x)$ and $w(x)$ – sufficient to check inequality at endpoints.
6. Situations in Which We Only Know Upper Bounds

- **Situation:** we only know;
  - the upper bound $\Delta$ on the measurement inaccuracy $\Delta v \overset{\text{def}}{=} \tilde{v} - v$: $|\Delta v| \leq \Delta$, and
  - the upper bound $\delta$ on the location error $\Delta x \overset{\text{def}}{=} \tilde{x} - x$: $|\Delta v| \leq \delta$.
- **Consequence:** the measured value $\tilde{v}$ is $\Delta$-close to a convex combination of values $v(x)$ for $x$ s.t. $\|x - \tilde{x}\| \leq \Delta x$.
- **Conclusion:** $v_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq \overline{v}_\delta(\tilde{x}) + \Delta$, where:
  - $v_\delta(\tilde{x}) \overset{\text{def}}{=} \inf \{ v(x) : \|x - \tilde{x}\| \leq \delta \}$, and
  - $\overline{v}_\delta(\tilde{x}) \overset{\text{def}}{=} \sup \{ v(x) : \|x - \tilde{x}\| \leq \delta \}$.
- **Fact:** measurement errors are random.
- **So:** it makes sense to only consider essential ess inf and ess sup (i.e., inf and sup modulo measure 0 sets).
7. What If a Model Is Only Known With Interval Uncertainty

- **Reminder:** we can tell when an observation \((\tilde{v}, \tilde{x})\) is consistent with a model \(v(x)\):
  \[
  v_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq \bar{v}_\delta(\tilde{x}) + \Delta.
  \]
- **Fact:** in real life, we rarely have an *exact* model \(v(x)\).
- **Usually:** we have *bounds* on \(v(x)\), i.e., an interval-valued model \(v(x) = [v^-(x), v^+(x)]\).
- **Question:** when is an observation \((\tilde{v}, \tilde{x})\) consistent with an *interval-valued* model?
- **General answer:** when the observation \((\tilde{v}, \tilde{x})\) is consistent with *one* of the models \(v(x) \in v(x)\).
- **A checkable answer:** an observation \((\tilde{v}, \tilde{x})\) is consistent with an interval-valued model \([v^-(x), v^+(x)]\) when
  \[
  v^-_\delta(\tilde{x}) - \Delta \leq \tilde{v} \leq v^+_\delta(\tilde{x}) + \Delta.
  \]
8. Situations in Which We Only Know Upper Bounds (cont-d)

- **Fact:** the actual $v(x)$ is often continuous.

- **Case of continuous $v(x)$:** we can simplify the above criterion.

- **Simplification:** the set $\tilde{m}$ of all measurement results $(\tilde{x}, \tilde{x})$ is consistent with the model $v(x)$ iff

  \[ \forall (\tilde{v}, \tilde{x}) \in \tilde{m} \exists (v(x), x) \in v ((\tilde{v}, \tilde{x}) \text{ is } (\Delta, \delta)-\text{close to } (v(x), x)), \]

  i.e., $|\tilde{v} - v| \leq \Delta$ and $\|x - \tilde{x}\| \leq \delta$.

- **Hausdorff metric:** $d_H(A, B) \leq \varepsilon$ means that:

  \[ \forall a \in A \exists b \in B (d(a, b) \leq \varepsilon) \text{ and } \forall b \in B \exists a \in A (d(a, b) \leq \varepsilon). \]

- **Conclusion:** we have an asymmetric version of Hausdorff metric (““quasi-metric”).
9. Example of Asymmetry

- **Case 1:**
  - The actual field: \( v(0) = 1 \) and \( v(x) = 0 \) for \( x \neq 0 \);
  - Measurement results: all zeros, i.e., \( \tilde{v} = 0 \) for all \( \tilde{x} \).
  - Conclusion: all the measurements are consistent with the model.
  - Reason: the value \( \tilde{v} = 0 \) for \( \tilde{x} = 0 \) is consistent with \( v(x) = 0 \) for \( x = \delta \) s.t. \( |\tilde{x} - x| \leq \delta \).

- **Case 2:**
  - The actual field: all zeros, i.e., \( v(x) = 0 \) for all \( x \).
  - Measurement results: \( \tilde{v} = 1 \) for \( \tilde{x} = 0 \), and \( \tilde{v} = 0 \) for all \( \tilde{x} \neq 0 \).
  - Conclusion (for \( \Delta < 1 \)): the measurement \((1, 0)\) is inconsistent with the model.
  - Reason: for all \( x \) which are \( \delta \)-close to \( \tilde{x} = 0 \), we have \( v(x) = 0 \) hence we should have \( |\tilde{x} - v(x)| = |\tilde{x}| \leq \Delta \).
10. Situations with No Information about Location Accuracy

- **Example**: when we solve the seismic inverse problem to find the velocity distribution.

- **Natural heuristic idea**:
  - add a perturbation of size $\Delta_0$ to the reconstructed field $\tilde{v}(x)$;
  - simulate the new measurement results;
  - apply the same algorithm to the simulated results, and reconstruct the new field $\tilde{v}_\text{new}(x)$.

- **Case 1**: perturbations are *not visible* in $\tilde{v}_\text{new}(x) - \tilde{v}(x)$.
- **So**: details of size $\Delta_0$ *cannot* be reconstructed: $\delta > \Delta_0$.

- **Case 2**: perturbations are *visible* in $\tilde{v}_\text{new}(x) - \tilde{v}(x)$.
- **So**: details of size $\Delta_0$ *can* be reconstructed: $\delta \leq \Delta_0$. 
11. Towards Optimal Selection of Perturbations

- **Fact:** since perturbations are small, we can safely linearize their effects.

- **Conclusion:**
  - based on the results of perturbations $e_1(x), \ldots, e_k(x)$,
  - we can get the results of any linear combination
    \[ e(x) = c_1 \cdot e_1(x) + \ldots + c_k \cdot e_k(x). \]

- **Fact:** usually, there is no preferred spatial location.

- **Conclusion:** we can choose different locations as origins ($x = 0$) of the coordinate system.

- **Natural requirement:** the results of perturbations should not change if we change the origin $x = 0$. 
12. Towards Optimal Perturbations (cont-d)

- **Reminder:** the class of perturbations should not change when we change the origin $x = 0$.
- **Fact:** in new coordinates, $x_{\text{new}} = x + x_0$.
- **Conclusion:** the set $\{c_1 \cdot e_1(x) + \ldots + c_k \cdot e_k(x)\}$ must be shift-invariant: $e_i(x + x_0) = \sum_{j=1}^{k} c_{ij}(x_0) \cdot e_j(x)$.
- When $x_0 \to 0$, we get a system of linear differential equations with constant coefficients.
- **General solution:** linear combination of expressions $\exp(\sum a_i \cdot x_i)$ with complex $a_i$.
- **Fact:** perturbations must be uniformly small.
- **So:** the only bounded perturbations are linear combinations of sinusoids.
- **Conclusion:** use sinusoidal perturbations.
13. Acknowledgments

This work was supported in part:

- by NSF Cyber-Share grant HRD-0734825:
  – A Center for Sharing Cyberresources to Advance Science and Education;
- by Grant 1 T36 GM078000-01 from the National Institutes of Health:
  – Enhancement of Qualitative Science;
- by the Japan Advanced Institute of Science and Technology (JAIST) Int’l Joint Research Grant 2006-08.