Towards a Fast, Practical Alternative to Joint Inversion of Multiple Datasets: Model Fusion

Omar Ochoa[1], Aaron A. Velasco[2], and Christian Servin[1]
[1]Department of Computer Science
[2]Department of Geological Sciences
University of Texas, El Paso, Texas 79968, USA
omar@miners.utep.edu, velasco@geo.utep.edu,
christians@miners.utep.edu

Abstract

In many areas of science and engineering, we have different sources of data. For example, in geophysics, there are many sources of data for Earth models: first-arrival passive seismic data (from the actual earthquakes), first-arrival active seismic data (from the seismic experiments), gravity data, etc.

Datasets coming from different sources provide complimentary information. In general, some of the datasets provide better accuracy and/or spatial resolution in some spatial areas and in some depths, while other datasets provide a better accuracy and/or spatial resolution in other areas or depths. For example: each gravity data points describes the result of measuring the gravity field at some spatial location; this field is generated by the joint effects of many locations; as a result, gravity measures the average density over a reasonably large spatial region, so estimates based on gravity measurements have (relatively) low spatial resolution. In contrast, each seismic data point comes from a narrow trajectory of a seismic signal, so the spatial resolution corresponding to this data is much higher.

At present, each of these datasets is often processed separately, resulting in several different models reflecting different aspects of the studied phenomena. It is therefore desirable to combine data from different datasets.

An ideal approach would be to use all the datasets to produce a single model. However, in many research areas – including geophysics – there are no efficient algorithms for simultaneously processing all the different datasets. While such joint inversion methods are being developed, as a first step, we propose a practical solution: to fuse the *models* coming from different datasets.

1 Need to Combine Data from Different Sources

In many areas of science and engineering, we have different sources of data.

For example, in geophysics, there are many sources of data for Earth models:

- first-arrival passive seismic data (from the actual earthquakes); see, e.g., [5];
- first-arrival active seismic data (from the seismic experiments); see, e.g., [1, 4];
- gravity data; and
- surface waves; see, e.g., [6].

Datasets coming from different sources provide complimentary information. For example, different geophysical datasets contain different information on earth structure. In general:

- some of the datasets provide better accuracy and/or spatial resolution in some spatial areas and in some depths, while
- other datasets provide a better accuracy and/or spatial resolution in other areas or depths.

For example:

- each gravity data points describes the result of measuring the gravity field at some spatial location; this field is generated by the joint effects of many locations; as a result, gravity measures the average density over a reasonably large spatial region, so estimates based on gravity measurements have (relatively) low spatial resolution;
- in contrast, each seismic data point comes from a narrow trajectory of a seismic signal, so the spatial resolution corresponding to this data is much higher.

Usually, there are several different datasets. At present, each of these datasets is often processed separately, resulting in several different models reflecting different aspects of the studied phenomena. It is therefore desirable to combine data from different datasets.

Comment. In most applications to geosciences, the corresponding quantities change very slowly with time. In these applications, we are interested in the values (such as density) at different depth and at different spatial locations; we know that these values do not change from one measurement to another. In such applications, since usually we cannot directly measure the value at a single spatial location, we measure, in effect, the average over a spatial area. The smaller the size of this area, the higher the spatial resolution.

In other applications areas, the values of the corresponding quantities change not only when we move from one location to another, but they also change with time. In such application areas, due to natural inertia of measuring instruments,

the measured values do not correspond not only to the average over a spatial area, but also to the average over a certain time interval. In such situations, in addition to *spatial* resolution, we also have *temporal* resolution: the smaller the corresponding time interval, the higher the temporal resolution. So, in general, we have *spatio-temporal* resolution.

In the following text, for simplicity, we will talk about spatial resolution, but all our discussions and formulas are applicable to the more general situation of spatio-temporal resolution as well.

2 Joint Inversion: An Ideal Future Approach

The ideal approach would be to use all the datasets to produce a single model. At present, however, in many research areas – including geophysics – there are no efficient algorithms for simultaneously processing all the different datasets.

Designing such joint inversion techniques presents an important theoretical and practical challenge.

3 Data Fusion: Brief Reminder

Our main idea. While such joint inversion methods are being developed, as a first step, we propose a practical solution: to fuse all the *models* coming from different datasets.

Comment. Some of our results have been announced in [7, 8, 11].

Simplest case: data fusion. In many real-life situations, we have several measurements and/or expert estimates $\widetilde{x}^{(1)}, \ldots, \widetilde{x}^{(n)}$ of the same quantity x.

- These values may come from the actual (direct) measurements of the quantity x.
- Alternatively, these values may come from *indirect* measurements of x, i.e., from different models, in which, based on the corresponding measurement results, the i-th model leads to an estimate $\widetilde{x}^{(i)}$ for x.

In such situations, it is desirable to fuse these estimates into a single more accurate estimate for x; see, e.g., [10].

Data fusion: case of probabilistic uncertainty (reminder). Let us start with the case when each estimate $\widetilde{x}^{(i)}$ is known with the (traditionally described) probabilistic uncertainty, e.g., when

- each estimation error $\Delta x^{(i)} \stackrel{\text{def}}{=} \widetilde{x}^{(i)} - x$ is normally distributed with 0 mean and known standard deviation $\sigma^{(i)}$, and
- estimation errors $\Delta x^{(i)}$ corresponding to different models are independent.

Comment. In practice, the estimation errors are indeed often normally distributed. This empirical fact can be justified by the Central Limit Theorem, according to which, under certain reasonable conditions, the joint effect of many relatively small errors is (approximately) normally distributed; see, e.g., [12]. For each model based on measurements of a certain type (e.g., gravity or seismic), not only the resulting error of each measurement comes from many different error sources, but also each estimate comes from several different measurements – thus further increasing the number of different error components contributing to the estimation error.

In this case, the probability density for each estimation error $\Delta x^{(i)}$ has the form

$$\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma^{(i)}}\cdot\exp\left(-\frac{(\Delta x^{(i)})^2}{2\cdot(\sigma^{(i)})^2}\right)=\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma^{(i)}}\cdot\exp\left(-\frac{(\widetilde{x}^{(i)}-x)^2}{2\cdot(\sigma^{(i)})^2}\right),$$

and the probability density $\rho(u)$ corresponding to all n estimates is (due to independence) the product of these densities:

$$\rho(x)=\prod_{i=1}^{n}\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma^{(i)}}\cdot\exp\left(-\frac{(\widetilde{x}^{(i)}-x)^2}{2\cdot(\sigma^{(i)})^2}\right)=$$

$$\left(\prod_{i=1}^{n}\frac{1}{\sqrt{2\cdot\pi}\cdot\sigma^{(i)}}\right)\cdot\exp\left(-\sum_{i=1}^{n}\frac{(\widetilde{x}^{(i)}-x)^2}{2\cdot(\sigma^{(i)})^2}\right).$$

As a single estimate x for the desired quantity, it is reasonable to select the value for which this probability (density) $\rho(x)$ is the largest (i.e., to use the *Maximum Likelihood* method). Since $\exp(z)$ is an increasing function, maximizing a function $A\cdot\exp(-B(x))$ is equivalent to minimizing $B(x)$, so we arrive at the following *Least Squares* approach: find x for which the sum

$$\sum_{i=1}^{n}\frac{(\widetilde{x}^{(i)}-x)^2}{2\cdot(\sigma^{(i)})^2}$$

is the smallest possible.

Differentiating this expression with respect to x and equating the derivative to 0, we conclude that

$$x=\frac{\sum\limits_{i=1}^{n}\widetilde{x}^{(i)}\cdot(\sigma^{(i)})^{-2}}{\sum\limits_{i=1}^{n}(\sigma^{(i)})^{-2}}.$$

The accuracy of this fused estimate can be described by the standard deviation σ for which

$$\sigma^{-2}=\sum_{i=1}^{n}(\sigma^{(i)})^{-2}.$$

Data fusion: case of interval uncertainty. In some practical situations, the value x is known with interval uncertainty, i.e., we know the interval $\mathbf{x}^{(i)} = [\widetilde{x}^{(i)} - \Delta^{(i)}, \widetilde{x}^{(i)} + \Delta^{(i)}]$ containing the actual (unknown) value of x. This happens, e.g., when we only know the upper bound $\Delta^{(i)}$ on each estimation error $\Delta x^{(i)}$: $|\Delta x^{(i)}| \le \Delta^{(i)}$. In this case, from the fact that the estimate is $\widetilde{x}^{(i)}$, we can conclude that $|x - \widetilde{x}^{(i)}| \le \Delta^{(i)}$, i.e., that $\widetilde{x}^{(i)} - \Delta^{(i)} \le x \le \widetilde{x}^{(i)} + \Delta^{(i)}$.

For interval uncertainty, it is easy to fuse several estimates. Based on each estimate $\widetilde{x}^{(i)}$, we know that the actual value x belongs to the interval $\mathbf{x}^{(i)}$. Thus, we know that the (unknown) actual value x belongs to the intersection

$$\mathbf{x} \stackrel{\text{def}}{=} \bigcap_{i=1}^{n} \mathbf{x}^{(i)} = [\max(\widetilde{x}^{(i)} - \Delta^{(i)}), \min(\widetilde{x}^{(i)} + \Delta^{(i)})]$$

of these intervals.

4 Proposed Solution – Model Fusion: Main Idea

Additional problem: we also have different spatial resolution. In many practical situations, estimates coming from different models have not only different accuracy, but also different spatial resolution.

Example. For example, in the geosciences,

- seismic data leads to estimates of the density at different locations and depths which have higher spatial resolution, while
- gravity data leads to estimates of the same densities which have lower spatial resolution.

Towards precise formulation of the problem. Estimates with higher spatial (spatio-temporal) resolution mean that we estimate the values corresponding to small spatial (spatio-temporal) cells. An estimate with a lower spatial resolution means that its results are affected by several neighboring spatial cells, i.e., that we are estimating, in effect, a weighted average of the values in several neighboring cells.

Comment. In this paper, we only consider the case when the corresponding weights are known exactly, because this case is typical for geophysical applications. For example, for gravity measurements, we know exactly Newton's formulas that describe how the gravity is determined by the densities at different locations. According to these formulas, the gravity force field $\vec{g}(\vec{x})$ at a point of measurement $\vec{x}$ is equal to the weighted combination of densities $\rho\left(\vec{x'}\right)$ at different locations $\vec{x'}$, with the weights determined by the distance between

$\vec{x}$ and $\vec{x'}$ and by the relative orientation of the line from $\vec{x}$ to $\vec{x'}$:

$$\vec{g}\,(\vec{x}) = G \cdot \int \rho\left(\vec{x'}\right) \cdot \frac{\vec{x'} - \vec{x}}{\left\|\vec{x'} - \vec{x}\right\|^3},$$

where $\|.\|$ denotes the length of a vector.

Similarly, we know the equations that describe how the propagation of the seismic waves, of electromagnetic waves, etc., depends on the parameters of the media.

In some other applications, we only have an approximate knowledge of the dependence of the measured quantity on the quantities x_i in which we are interested. In such applications, one must take into account that the corresponding weights are also only known with uncertainty.

What is given. In precise terms:

- we have resolution estimates $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of the values $x_1, \ldots, x_n$ within several small spatial cells; these estimates correspond to models with a higher spatial resolution
- we also have estimates $\widetilde{X}_j$ for the weighted averages

$$X_j = \sum_{i=1}^{n} w_{j,i} \cdot x_i;$$

these estimates correspond to models with a lower spatial resolution.

Comment. In this paper, we assume that we know the values of the weights $w_{j,i}$. This assumption makes perfect sense for geophysical problems, because in these problems, these weights are indeed known. For example:

- We know how exactly the gravity at a given point depends on the densities at different spatial locations.
- We know how exactly the travel time of a seismic signal depends on the density distribution.

In some applications, however, the corresponding weights are only approximately known. In such situations, when fusing the models, we must also take into account the uncertainty with which we know these weights. For these applications, it is desirable to extend our techniques – to accommodate such more complex situations.

What our objective is. We are interested in the values x_i. So, based on the estimates $\widetilde{x}_i$ and $\widetilde{x}$, we must provide more accurate estimates for x_i.

Example. In the geophysical example, we are interested in the values of the densities x_i.

What we do in this paper. In this paper, we describe how to fuse estimates with different accuracy and spatial resolution:

- In the case of probabilistic uncertainty, we use the Least Squares Method to derive explicit formulas for combining the estimates $\widetilde{x}_i$ and $\widetilde{X}_j$.
- In the case of interval uncertainty, we provide an efficient algorithm for estimating the ranges of x_i.

5 Model Fusion: Case of Probabilistic Uncertainty

5.1 General Case

Main idea. Our solution to the model fusion problem is to take into account several different types of approximate equalities:

- Each estimate $\widetilde{x}_i$ from a model with a high spatial resolution is approximately equal to the actual value x_i in the corresponding (smaller size) cell i, with the known accuracy $\sigma_{h,i}$:

$$\widetilde{x}_i \approx x_i.$$

- Each estimate $\widetilde{X}_j$ from (one of the) models with a lower spatial resolution is approximately equal to the weighted average of values of all the smaller cells $x_{i(1,j)}, \ldots, x_{i(k_j,j)}$ within the corresponding larger size cell, with a known accuracy $\sigma_{l,j}$:

$$\widetilde{X}_j \approx \sum_i w_{j,i} \cdot x_i,$$

for known weights $w_{j,i} \geq 0$ for which $\sum\limits_{i=1}^{n} w_{j,i} = 1$. In the simple case when these weights are equal, we get

$$\widetilde{X}_j \approx \frac{x_{i(1,j)} + \ldots + x_{i(k_j,j)}}{k_j}.$$

- We usually have a prior knowledge of the values x_i. It is reasonable to assume that this knowledge can also be described by a normal distribution, with the mean $x_{pr,i}$ and the standard deviation $\sigma_{pr,i}$:

$$x_i \approx x_{pr,i}.$$

(The case when for some i, we have no prior information at all is equivalent to setting $\sigma_{pr,i} = \infty$.)

- Finally, each estimate $\widetilde{X}_j$ from a model with a lower spatial resolution is approximately equal to the value within each of the constituent smaller size cells $x_{i(l,j)}$, with the accuracy corresponding to the (empirical) standard deviation $\sigma_{e,j}$ of the smaller-cell values within the larger cell:

$$\widetilde{X}_j \approx x_{i(l,j)},$$

where

$$\sigma_{e,j}^2 \stackrel{\text{def}}{=} \frac{1}{k_j} \cdot \sum_{l=1}^{k_j} \left(\widetilde{x}_{i(l,j)} - E_j\right)^2,$$

and

$$E_j \stackrel{\text{def}}{=} \frac{1}{k_j} \cdot \sum_{l=1}^{k_j} \widetilde{x}_{i(l,j)}.$$

We then use the Least Squares technique to combine these approximate equalities, and find the desired combined values x_i by minimizing the resulting sum of weighted squared differences.

Relation between different standard deviations. As we have mentioned earlier, there is usually a trade-off between accuracy and spatial resolution:

- if we want to estimate the value of the desired quantity with a higher spatial resolution, i.e., the value corresponding to a small spatial location, then we get lower accuracy, i.e., higher values of the standard deviation $\sigma_{h,i}$;
- on the other hand, if we are satisfied with a lower spatial resolution, i.e., with the fact that the estimated value corresponds to a larger spatial area, then we can get higher accuracy, i.e., lower values of the standard deviation $\sigma_{l,j} \ll \sigma_{h,i}$.

From the mathematical viewpoint, this trade-off makes sense. In principle, as an estimate for a model with a low spatial resolution, we can take the average of the values corresponding to high spatial resolution, and averaging usually decreases the approximation error:

$$\sigma_{l,j} \ll \sigma_{h,i} \ll \sigma_{e,j}.$$

Comment. It should be mentioned that while usually, higher spatial resolution estimates have lower accuracy, sometimes, a higher-resolution model has more accuracy in some places. For example, in the geosciences,

- the measurements from a borehole provide the most accurate estimates of the corresponding quantities,
- and for these measurements, the spatial location is also known with a very high accuracy.

Resulting formulas: general case. According to the Least Squares approach, in the general case, we minimize the following expression:

$$\sum_{i=1}^{n} \frac{(x_i - \widetilde{x}_i)^2}{\sigma_{h,i}^2} + \sum_{j=1}^{m} \frac{1}{\sigma_{l,j}^2} \cdot \left(\widetilde{X}_j - \sum_{i=1}^{n} w_{j,i} \cdot x_i \right)^2 +$$

$$\sum_{i=1}^{n} \frac{(x_i - x_{pr,i})^2}{\sigma_{pr,i}^2} + \sum_{j=1}^{m} \sum_{l=1}^{k_j} \frac{(\widetilde{X}_j - x_{i(l,j)})^2}{\sigma_{e,j}^2}.$$

In this general case, differentiation with respect to x_i leads to the following system of linear equations:

$$\frac{1}{\sigma_{h,i}^2} \cdot (x_i - \widetilde{x}_i) + \sum_{j:j \ni i} \frac{1}{\sigma_{l,j}^2} \cdot w_{j,i} \cdot \left(\sum_{i'=1}^{n} w_{j,i'} \cdot x_{i'} - \widetilde{X}_j \right) +$$

$$\frac{1}{\sigma_{pr,i}^2} \cdot (x_i - x_{pr,i}) + \sum_{j:j \ni i} \frac{1}{\sigma_{e,j}^2} \cdot (x_i - \widetilde{X}_j) = 0,$$

where $j \ni i$ means that the j-th estimate corresponding to a model with a low spatial resolution covers the i-th cell.

Towards simplification: fusing prior estimates with estimates from a model with a high spatial resolution. For each cell i for which we have both a prior estimate $x_{pr,i}$ and an estimate $\widetilde{x}_i$ from a model with a higher spatial resolution, we can fuse these two estimates by using the above-described standard data fusion technique. As a result, instead of the two terms

$$\frac{1}{\sigma_{h,i}^2} \cdot (x_i - \widetilde{x}_i) + \frac{1}{\sigma_{pr,i}^2} \cdot (x_i - x_{pr,i}),$$

we have a single term

$$\sigma_{f,i}^{-2} \cdot (x_i - x_{f,i}),$$

where

$$x_{f,i} \stackrel{\text{def}}{=} \frac{\widetilde{x}_i \cdot \sigma_{h,i}^{-2} + x_{pr,i} \cdot \sigma_{pr,i}^{-2}}{\sigma_{h,i}^{-2} + \sigma_{pr,i}^{-2}}$$

and

$$\sigma_{f,i}^{-2} \stackrel{\text{def}}{=} \sigma_{h,i}^{-2} + \sigma_{pr,i}^{-2}.$$

We can use the same formula if we only have a high spatial resolution estimate or if we only have a prior estimate:

- If we only have a high spatial resolution estimate but no prior estimate, then we should take $\sigma_{pr,i}^{-2} = 0$ (i.e., $\sigma_{pr,i} = \infty$).
- If we only have a prior estimate but no high spatial resolution estimate, then we should take $\sigma_{h,i}^{-2} = 0$ (i.e., $\sigma_{h,i} = \infty$).

As a result of this fusion, we get the following simplified formulas.

Resulting formulas: simplified equations.

$$\sigma_{f,i}^{-2} \cdot (x_i - x_{f,i}) + \sum_{j:j\ni i} \frac{1}{\sigma_{l,j}^2} \cdot w_{j,i} \cdot \left(\sum_{i'=1}^{n} w_{j,i'} \cdot x_{i'} - \widetilde{X}_j \right) +$$

$$\sum_{j:j\ni i} \frac{1}{\sigma_{e,j}^2} \cdot (x_i - \widetilde{X}_j) = 0.$$

How to solve this system of linear equations. We can use known algorithms for solving this system of linear equations.

It is worth mentioning that usually, these algorithms require that we represent the system in the standard form $Ax = b$. To represent our system of equations in this form, we need to move all the terms that do not contain unknowns to the right-hand side.

Comment. In this section, we assumed that the values x_i are independent – in the sense that in principle, we can have arbitrary combination of values $x_1, \ldots, x_n$. In some practical situations, not all combinations of x_i are physically possible, there are additional *constraints* that these values x_i must satisfy. For example, we may have prior restrictions on the individual values of x_i: for example, the density x_i at a certain location must be within known bounds (say, between 6 and 8 g/cm^3). We can also have constraints relating values in the neighboring cells: e.g., the difference between the densities in a neighboring cell cannot exceed a certain value.

In this section, we described how to find the absolute (unconstrained) minimum of the corresponding quadratic form. In situations in which there are prior constraints, we need to minimize the corresponding quadratic form under these constraints.

What is the accuracy of the fused model? In the probabilistic case, the inaccuracy $\delta a = \hat{a} - a$ of each estimate $\hat{a}$ of a quantity a can be described by its *variance* $\sigma^2 = E[(\delta a)^2]$, i.e., by the mean value of the square $(\delta a)^2$.

When we estimate several quantities $x_1, \ldots, x_n$, it is also desirable to find out how correlated are the corresponding inaccuracies $\delta x_i = \hat{x}_i - x_i$. These correlations can be described by listing all possible covariances $c_{ii'} \stackrel{\text{def}}{=} E[\delta x_i \cdot \delta x_{i'}]$. For $i = i'$, we get the variances; thus, the *covariance matrix* $c_{ii'}$ contains all the information about the inaccuracies.

In the general Least Squares approach, when the minimized quadratic expression has the form

$$a_0 + \sum_{i=1}^{n} a_i \cdot x_i + \sum_{i=1}^{n} \sum_{i'=1}^{n} a_{ii'} \cdot x_i \cdot x_{i'},$$

then the covariance matrix $c_{ii'}$ is equal to the inverse of the matrix $a_{ii'}$; see, e.g., [12]. For the above quadratic form,

$$a_{ii'} = \sigma_{h,i}^{-2} \cdot \delta_{ii'} + \sum_{j: i \in j \,\&\, i' \in j} \sigma_{l,j}^{-2} \cdot w_{j,i} \cdot w_{j,i'} + \sigma_{pr,i}^{-2} \cdot \delta_{ii'} + \sum_{j: i \in j} \sigma_{e,j}^{-2} \cdot \delta_{ii'},$$

where $\delta_{ii'} = 1$ when $i = i'$ and $\delta_{ii'} = 0$ when $i \neq i'$. Thus, we can find the values $c_{ii'}$ of the covariance matrix by inverting the above matrix $a_{ii'}$.

5.2 Case of a Single Estimate with Low Spatial Resolution

Description. Let us now consider the simplest case, when when we have exactly one estimate $\widetilde{X}_1$ from a model with a low spatial resolution. In general, we only have prior estimates and the estimates with high spatial resolution for *some* of the cells.

This situation is typical in geosciences: e.g.,

- we have an estimate originated from the gravity measurements (with a lower spatial resolution) which covers a huge area in depth, and
- we have estimates originated from seismic measurements (corresponding to higher spatial resolution) which only cover depths above the Moho surface.

For convenience, let us number the cells in such a way that the cells for which we have either prior estimates or estimates from a high spatial resolution model come first. Let h denote the total number of such cells.

This means that as the result of combining prior estimates and estimates corresponding to high spatial resolution model(s), we have h values $x_{f,1}$, $x_{f,2}$, $\ldots$, $x_{f,h}$.

Derivation. In this case, the above system of linear equations takes the following form: for $i = 1, \ldots, h$, we have

$$\sigma_{f,i}^{-2} \cdot (x_i - x_{f,i}) + \frac{1}{\sigma_{l,1}^2} \cdot w_{1,i} \cdot \left(\sum_{i'} w_{1,i'} \cdot x_{i'} - \widetilde{X}_1 \right) + \frac{1}{\sigma_{e,1}^2} (x_i - \widetilde{X}_1) = 0;$$

and for $i > h$, we have

$$\frac{1}{\sigma_{l,1}^2} \cdot w_{1,i} \cdot \left(\sum_{i'} w_{1,i'} \cdot x_{i'} - \widetilde{X}_1 \right) + \frac{1}{\sigma_{e,1}^2} (x_i - \widetilde{X}_1) = 0.$$

For $i \leq h$, multiplying both sides by $\sigma_{f,i}^2$, we conclude that

$$x_i - x_{f,i} + \frac{\sigma_{f,i}^2}{\sigma_{l,1}^2} \cdot w_{1,i} \cdot \left(\sum_{i'} w_{1,i'} \cdot x_{i'} - \widetilde{X}_1 \right) + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2} \cdot (x_i - \widetilde{X}_1) = 0.$$

If we introduce an auxiliary variable

$$\mu \stackrel{\text{def}}{=} \frac{1}{\sigma_{l,1}^2} \cdot \left(\sum_{i'} w_{1,i'} \cdot x_{i'} - \widetilde{X}_1 \right),$$

we get the equation

$$x_i - x_{f,i} + w_{1,i} \cdot \sigma_{f,i}^2 \cdot \mu + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2} \cdot (x_i - \widetilde{X}_1) = 0.$$

By keeping terms proportional to x_i in the left-hand side and by moving all the other terms to the right-hand side, we get

$$\left(1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2} \right) \cdot x_i = x_{f,i} - w_{1,i} \cdot \sigma_{f,i}^2 \cdot \mu + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2} \cdot \widetilde{X}_1,$$

hence

$$x_i = \frac{x_{f,i}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \frac{w_{1,i} \cdot \sigma_{f,i}^2}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} \cdot \mu + \widetilde{X}_1 \cdot \frac{\dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}.$$

For $i > h$, we similarly get

$$x_i - \widetilde{X}_1 + w_{1,i} \cdot \sigma_{e,1}^2 \cdot \mu = 0,$$

hence

$$x_i = \widetilde{X}_1 - w_{1,i} \cdot \sigma_{e,1}^2 \cdot \mu.$$

To make this expression practically useful, we must describe μ in terms of the given values $\widetilde{x}_i$ and $\widetilde{X}_1$. Since μ is defined in terms of the weighted average of the values x_i, let us compute the weighted average of the above expressions for x_i:

$$\sum_{i=1}^{n} w_{1,i} \cdot x_i = \sum_{i=1}^{h} w_{1,i} \cdot x_i + \sum_{i=h+1}^{n} w_{1,i} \cdot x_i,$$

where

$$\sum_{i=1}^{h} w_{1,i} \cdot x_i =$$

$$\sum_{i=1}^{h} \frac{w_{1,i} \cdot x_{f,i}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \mu \cdot \sum_{i=1}^{h} \frac{w_{1,i}^2 \cdot \sigma_{f,i}^2}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} + \widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i} \cdot \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}.$$

Similarly,

$$\sum_{i=h+1}^{n} w_{1,i} \cdot x_i = \left(\sum_{i=h+1}^{n} w_{1,i} \right) \cdot \widetilde{X}_1 - \left(\sum_{i=h+1}^{n} w_{1,i}^2 \right) \cdot \frac{\sigma_{e,1}^2}{\sigma_{l,1}^2} \cdot \mu.$$

By adding these two sums and subtracting $\widetilde{X}_1$, we conclude that

$$\sigma_{l,1}^2 \cdot \mu = \sum_{i=1}^{n} w_{1,i} \cdot x_i - \widetilde{X}_1 =$$

$$\sum_{i=1}^{h} w_{1,i} \cdot x_i + \sum_{i=h+1}^{n} w_{1,i} \cdot x_i - \widetilde{X}_1 =$$

$$\sum_{i=1}^{h} \frac{w_{1,i} \cdot x_{f,i}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \mu \cdot \sum_{i=1}^{h} \frac{w_{1,i}^2 \cdot \sigma_{f,i}^2}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} + \widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i} \cdot \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} +$$

$$\left(\sum_{i=h+1}^{n} w_{1,i} \right) \cdot \widetilde{X}_1 - \left(\sum_{i=h+1}^{n} w_{1,i}^2 \right) \cdot \sigma_{e,1}^2 \cdot \mu - \widetilde{X}_1.$$

Since

$$\sum_{i=1}^{n} w_{1,i} = \sum_{i=1}^{h} w_{1,i} + \sum_{i=h+1}^{n} w_{1,i} = 1,$$

we conclude that

$$\left(\sum_{i=h+1}^{n} w_{1,i} \right) \cdot \widetilde{X}_1 - \widetilde{X}_1 = - \left(\sum_{i=1}^{h} w_{1,i} \right) \cdot \widetilde{X}_1$$

thus,

$$\widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i} \cdot \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} + \left(\sum_{i=h+1}^{n} w_{1,i} \right) \cdot \widetilde{X}_1 - \widetilde{X}_1 =$$

$$\widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i} \cdot \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \left(\sum_{i=1}^{h} w_{1,i} \right) \cdot \widetilde{X}_1 = -\widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}.$$

So, the equation for μ takes the following simplified form:

$$\sigma_{l,1}^2 \cdot \mu = \sum_{i=1}^{h} \frac{w_{1,i} \cdot x_{f,i}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \mu \cdot \sum_{i=1}^{h} \frac{w_{1,i}^2 \cdot \sigma_{f,i}^2}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i}}{1 + \frac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} -$$

$$\left(\sum_{i=h+1}^{n} w_{1,i}^2 \right) \cdot \sigma_{e,1}^2 \cdot \mu.$$

By moving all terms containing μ to the left-hand side and all other terms to the right-hand side, we get

$$\mu \cdot \left(\sigma_{l,1}^2 + \sum_{i=1}^{h} \frac{w_{1,i}^2 \cdot \sigma_{f,i}^2}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} + \left(\sum_{i=h+1}^{n} w_{1,i}^2 \right) \cdot \sigma_{e,1}^2 \right) =$$

$$\sum_{i=1}^{h} \frac{w_{1,i} \cdot x_{f,i}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \widetilde{X}_1 \cdot \sum_{i=1}^{h} \frac{w_{1,i}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} =$$

$$\sum_{i=1}^{h} \frac{w_{1,i} \cdot (x_{f,i} - \widetilde{X}_1)}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}.$$

Thus, we can compute μ. So, we arrive at the following formulas.

Resulting formulas. First, we compute the auxiliary value μ as

$$\mu = \frac{N}{D},$$

where

$$N = \sum_{i=1}^{h} \frac{w_{1,i} \cdot (x_{f,i} - \widetilde{X}_1)}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}$$

and

$$D = \sigma_{l,1}^2 + \sum_{i=1}^{h} \frac{w_{1,i}^2 \cdot \sigma_{f,i}^2}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} + \left(\sum_{i=h+1}^{n} w_{1,i}^2 \right) \cdot \sigma_{e,1}^2.$$

Then, we compute the desired estimates for x_i, $i = 1, \ldots, h$, as

$$x_i = \frac{x_{f,i}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} - \frac{w_{1,i} \cdot \sigma_{f,i}^2}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}} \cdot \mu + \widetilde{X}_1 \cdot \frac{\dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}}{1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2}},$$

and the estimates x_i for $i = h+1, \ldots, n$ as

$$x_i = \widetilde{X}_1 - w_{1,i} \cdot \sigma_{e,1}^2 \cdot \mu.$$

5.3 Numerical Example

Simplified case: description. To illustrate the above formulas, let us consider the simplest possible case, when we have exactly one estimate $\widetilde{X}_1$ from a lower spatial resolution model, and when:

- this estimate covers all n cells;
- all the weights are all equal $w_{1,i} = 1/n$;
- for each of n cells, there is an estimate corresponding to this cell that comes from a high spatial resolution model (i.e., $h = n$);
- all estimates coming from a high spatial resolution model have the same accuracy $\sigma_{h,i} = \sigma_h$;
- the estimate corresponding to a low spatial resolution model is much more accurate than the estimates corresponding to higher spatial resolution models $\sigma_{l,1} \ll \sigma_h$, so we can safely assume that $\sigma_l = 0$; and
- there is no prior information, so $\sigma_{pr,i} = \infty$ and thus, $x_{f,i} = \widetilde{x}_i$ and $\sigma_{f,i} = \sigma_h$.

To cover the cells for which there are no estimates from a high spatial resolution model, we added a heuristic rule that the estimate from a lower spatial resolution model is approximately equal to the value within each of the constituent smaller size cells, with the accuracy corresponding to the (empirical) standard deviation $\sigma_{e,j}$. In our simplified example, we have individual estimates for each cell, so there is no need for this heuristic rule. The corresponding heuristic terms in the general least squares approach are proportional to $\dfrac{1}{\sigma_{e,1}^2}$, so ignoring these terms is equivalent to taking $\sigma_{e,1}^2 = \infty$. Thus, we have $\dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2} = 0$ and $1 + \dfrac{\sigma_{f,i}^2}{\sigma_{e,1}^2} = 1$.

Because of this and because of the fact that $w_{1,i} = \dfrac{1}{n}$ and $x_{f,i} = \widetilde{x}_i$, the formula for N takes the form

$$N = \sum_{i=1}^{n} \frac{1}{n} \cdot (\widetilde{x}_i - \widetilde{X}_1).$$

Opening parentheses and taking into account that the sum of n terms equal to $\dfrac{1}{n} \cdot \widetilde{X}_1$ is simply $\widetilde{X}_1$, we get

$$N = \frac{1}{n} \cdot \sum_{i=1}^{n} \widetilde{x}_i - \widetilde{X}_1.$$

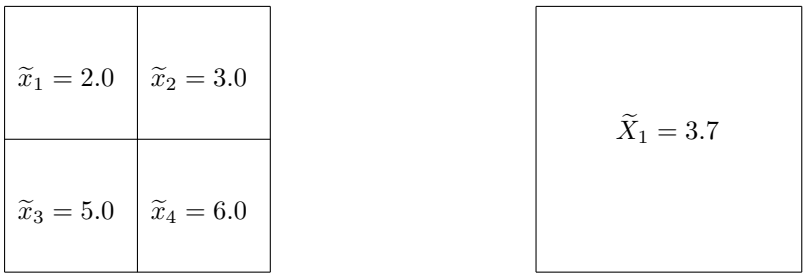

Figure 1: Higher and lower spatial resolution estimates

Similarly, due to our simplifying assumptions $\sigma_{l,1} = 0$, $w_{1,i} = \dfrac{1}{n}$, $\sigma_{f,i} = \sigma_h$, $\sigma_{e,1} = 0$, and $h = n$, we have

$$D = \sum_{i=1}^{n} \left(\frac{1}{n}\right)^2 \cdot \sigma_h^2 = \frac{1}{n} \cdot \sigma_h^2.$$

Thus,

$$\mu = \frac{N}{D} = \frac{\dfrac{1}{n} \cdot \sum\limits_{i=1}^{n} \widetilde{x}_i - \widetilde{X}_1}{\dfrac{1}{n} \cdot \sigma_h^2}.$$

The formula for x_i now turns into

$$x_i = \widetilde{x}_i - \frac{1}{n} \cdot \sigma_h^2 \cdot \mu.$$

Substituting the above expression for μ, we conclude that

$$x_i = \widetilde{x}_i - \lambda,$$

where

$$\lambda \stackrel{\text{def}}{=} \frac{1}{n} \cdot \sum_{i=1}^{n} \widetilde{x}_i - \widetilde{X}_1.$$

Numerical example: simplified case. Let us assume that we have $n = 4$ cells, and that the high spatial resolution estimates for these cells are $\widetilde{x}_1 = 2.0$, $\widetilde{x}_2 = 3.0$, $\widetilde{x}_3 = 5.0$ and $\widetilde{x}_4 = 6.0$. We also assume that each of these estimates has the same accuracy $\sigma_h = 0.5$. Let us also assume that we have an estimate $\widetilde{X}_1 = 3.7$ for the average X_1 of these four values. We assume that this estimate has a much higher accuracy $\sigma_l \ll \sigma_h$ so that we can, in effect, take $\sigma_l \approx 0$.

Since we assume that the low spatial resolution estimates are accurate ($\sigma_l \approx 0$), we therefore assume that the estimated quantity, i.e., the arithmetic average

$\widetilde{x}_1 = 1.7$	$\widetilde{x}_2 = 2.7$
$\widetilde{x}_3 = 4.7$	$\widetilde{x}_4 = 5.7$

Figure 2: The result of model fusion: simplified setting

of the four cell values, is practically exactly equal to this estimate $\widetilde{X}_1 = 3.7$:

$$\frac{x_1 + x_2 + x_3 + x_4}{4} \approx 3.7.$$

For the high spatial resolution estimates $\widetilde{x}_i$, the average is slightly different:

$$\frac{\widetilde{x}_1 + \widetilde{x}_2 + \widetilde{x}_3 + \widetilde{x}_4}{4} = \frac{2.0 + 3.0 + 5.0 + 6.0}{4} = 4.0 \neq 3.7.$$

This difference is caused by the fact that, in contrast to accurate low spatial resolution estimates, higher spatial resolution measurements are much less accurate: the corresponding estimation error has a standard deviation $\sigma_h = 0.5$. We can therefore, as we described above, use the information from the low spatial resolution estimates to "correct" the high spatial resolution estimates.

In this particular example, since $\sigma_l \approx 0$, the correcting term takes the form

$$\lambda = \frac{\widetilde{x}_1 + \ldots + \widetilde{x}_n}{n} - \widetilde{X}_1 =$$

$$\frac{2.0 + 3.0 + 5.0 + 6.0}{4} - 3.7 = 4.0 - 3.7 = 0.3,$$

so the corrected ("fused") values x_i take the form:

$$x_1 = \widetilde{x}_1 - \lambda = 2.0 - 0.3 = 1.7; \quad x_2 = \widetilde{x}_2 - \lambda = 3.0 - 0.3 = 2.7;$$

$$x_3 = \widetilde{x}_3 - \lambda = 5.0 - 0.3 = 4.7; \quad x_4 = \widetilde{x}_4 - \lambda = 6.0 - 0.3 = 5.7;$$

For these corrected values, the arithmetic average is equal to

$$\frac{x_1 + x_2 + x_3 + x_4}{4} = \frac{1.7 + 2.7 + 4.7 + 5.7}{4} = 3.7,$$

i.e., exactly to the low spatial resolution estimate.

Taking $\sigma_{e,j}$ into account. What if, in the above numerical example, we take into account the requirement that the actual values in each cell are approximately equal to $\widetilde{X}_1$, with the accuracy $\sigma_{e,1}$ equal to the empirical standard deviation?

In this case, the above formulas take the form

$$N = \frac{1}{1 + \dfrac{\sigma_h^2}{\sigma_{e,1}^2}} \cdot \left(\frac{\widetilde{x}_1 + \ldots + \widetilde{x}_n}{n} - \widetilde{X}_1 \right)$$

and

$$D = \frac{1}{1 + \dfrac{\sigma_h^2}{\sigma_{e,1}^2}} \cdot \frac{1}{n} \cdot \sigma_h^2,$$

so we get the exact same expression for μ:

$$\mu = \frac{N}{D} = \frac{\dfrac{1}{n} \cdot \sum\limits_{i=1}^{n} \widetilde{x}_i - \widetilde{X}_1}{\dfrac{1}{n} \cdot \sigma_h^2}.$$

The formulas for the fused values x_i are now somewhat more complex:

$$x_i = \frac{\widetilde{x}_i - \lambda}{1 + \dfrac{\sigma_h^2}{\sigma_{e,1}^2}} + \widetilde{X}_1 \cdot \frac{\dfrac{\sigma_h^2}{\sigma_{e,1}^2}}{1 + \dfrac{\sigma_h^2}{\sigma_{e,1}^2}}.$$

Taking $\sigma_{e,j}$ into account: numerical example. We want to take into account the requirement that the actual values in each cell are approximately equal to $\widetilde{X}_1$, with the accuracy $\sigma_{e,j}$ equal to the empirical standard deviation. In our example, the lower spatial resolution estimate $\widetilde{X}_1$ covers all four cells. In this example, the above condition takes the form $x_i \approx \widetilde{X}_1$, with the accuracy

$$\sigma_{e,1}^2 = \frac{1}{4} \cdot \sum_{i=1}^{4} (\widetilde{x}_i - E_1)^2,$$

where

$$E_1 = \frac{1}{4} \cdot \sum_{i=1}^{4} \widetilde{x}_i.$$

For our numerical example, as we have seen,

$$E_1 = \frac{1}{4} \cdot \sum_{i=1}^{4} \widetilde{x}_i = \frac{\widetilde{x}_1 + \widetilde{x}_2 + \widetilde{x}_3 + \widetilde{x}_4}{4} = 4.0$$

and thus,

$$\sigma_{e,1}^2 = \frac{(2.0-4.0)^2+(3.0-4.0)^2+(5.0-4.0)^2+(6.0-4.0)^2}{4} =$$

$$\frac{4+1+1+4}{4} = \frac{10}{4} = 2.5,$$

hence $\sigma_{e,1} \approx 1.58$.

Now, we can use the formula

$$x_i = \frac{1}{1+\dfrac{\sigma_h^2}{\sigma_{e,1}^2}} \cdot (\widetilde{x}_i - \lambda) + \frac{\dfrac{\sigma_h^2}{\sigma_{e,1}^2}}{1+\dfrac{\sigma_h^2}{\sigma_{e,1}^2}} \cdot \widetilde{X}_1$$

to find the corrected ("fused") values x_i. Here, $\sigma_h = 0.5$, $\sigma_{e,1}^2 = 2.5$, so

$$\frac{\sigma_h^2}{\sigma_{e,1}^2} = \frac{0.25}{2.5} = 0.1$$

and therefore, with two digit accuracy,

$$\frac{1}{1+\dfrac{\sigma_h^2}{\sigma_{e,1}^2}} = \frac{1}{1.1} \approx 0.91$$

and

$$\frac{\dfrac{\sigma_h^2}{\sigma_{e,1}^2}}{1+\dfrac{\sigma_h^2}{\sigma_{e,1}^2}} \cdot \widetilde{X}_1 = \frac{0.1}{1.1} \cdot 3.7 \approx 0.34.$$

Therefore, we get

$$x_1 \approx 0.91 \cdot (2.0-0.3) + 0.34 \approx 1.89;$$

$$x_2 \approx 0.91 \cdot (3.0-0.3) + 0.34 \approx 2.79;$$

$$x_3 \approx 0.91 \cdot (5.0-0.3) + 0.34 \approx 4.62;$$

$$x_4 \approx 0.91 \cdot (6.0-0.3) + 0.34 \approx 5.53.$$

The arithmetic average of these four values is equal to

$$\frac{x_1+x_2+x_3+x_4}{4} \approx \frac{1.89+2.79+4.62+5.53}{4} \approx 3.71,$$

i.e., within our computation accuracy (since we performed all the computations with two digits after the decimal point) coincides with the lower spatial resolution estimate $\widetilde{X}_1 = 3.7$.

$\widetilde{x}_1 \approx 1.89$	$\widetilde{x}_2 \approx 2.79$
$\widetilde{x}_3 \approx 4.62$	$\widetilde{x}_4 \approx 5.53$

Figure 3: The result of model fusion: general setting

6 Model Fusion: Case of Interval Uncertainty

Main idea. Our solution to the model fusion problem is to take into account three different types of approximate equalities:

- Each higher spatial resolution estimate $\widetilde{x}_i$ is approximately equal to the actual value x_i in the corresponding (smaller size) cell i, with the approximation error $x_i - \widetilde{x}_i$ bounded by the known value $\Delta_{h,i}$:
$$\widetilde{x}_i - \Delta_{h,i} \le x_i \le \widetilde{x}_i + \Delta_{h,i}.$$
- Each lower spatial resolution estimate $\widetilde{X}_j$ is approximately equal to the average of values of all the smaller cells $x_{i(1,j)}, \ldots, x_{i(k_j,j)}$ within the corresponding larger size cell, with the estimation error bounded by the known value $\Delta_{l,j}$:
$$\widetilde{X}_j - \Delta_{l,j} \le \sum_i w_{j,i} \cdot x_i \le \widetilde{X}_j + \Delta_{l,j}.$$
- Finally, we have prior bounds $\underline{x}_{pr,i}$ and $\overline{x}_{pr,i}$ on the values x_i, i.e., bounds for which
$$\underline{x}_{pr,i} \le x_i \le \overline{x}_{pr,i}.$$

Our objective is to find, for each $k = 1, \ldots, n$, the range $[\underline{x}_k, \overline{x}_k]$ of possible values of x_k.

The estimates lead to a system of linear inequalities for the unknown values $x_1, \ldots, x_n$. Thus, for each k, finding the corresponding endpoints $\underline{x}_k$ and $\overline{x}_k$ means optimizing the values x_k under linear constraints. This is a particular case of a general linear programming problem; see, e.g., [3]. So, we can use Linear Programming to find these bounds:

- the lower bound $\underline{x}_k$ can be obtained if we minimize x_k under the constraints
$$\widetilde{x}_i - \Delta_h \le x_i \le \widetilde{x}_i + \Delta_h, \quad i = 1, \ldots, n;$$
$$\widetilde{X}_j - \Delta_l \le \sum_i w_{j,i} \cdot x_i \le \widetilde{X}_j + \Delta_l; \quad \underline{x}_{pr,i} \le x_i \le \overline{x}_{pr,i}.$$

- the upper bound $\overline{x}_k$ can be obtained if we maximize x_k under the same constraints.

Comment about representing the answer. As a result of applying the linear programming techniques, for each quantity x_k, we get an interval $[\underline{x}_k, \overline{x}_k]$ of possible values.

From the user viewpoint, it is often more convenient to instead present a (numerical) estimate $\hat{x}_k$ and an upper bound Δ_k on the inaccuracy $\delta x_k = \hat{x}_k - x_k$ of this estimate. In precise terms, this means that $|\delta x_k| \le \Delta_k$, i.e., the estimation error can only take values from $-\Delta_k$ to Δ_k: $-\Delta_k \le \delta x_k \le \Delta_k$.

Since $x_k = \hat{x}_k - \delta x_k$, the value x_k is the smallest when δx_k is the largest and the value x_k is the largest when δx_k is the smallest. The estimation error δx_k can take values from $-\Delta_k$ to Δ_k, so its smallest possible value is $-\Delta_k$ and its largest possible value is Δ_k. Thus, the smallest possible values of x_k is $\hat{x}_k - \Delta_k$, and the largest possible values of x_k is $\hat{x}_k - (-\Delta_k) = \hat{x}_k + \Delta_k$. In other words, under this representation, the set of all possible values of x_k is the interval $[\hat{x}_k - \Delta_k, \hat{x}_k + \Delta_k]$.

To represent the interval $[\underline{x}_k, \overline{x}_k]$ in this form, we must find the values $\hat{x}_k$ and Δ_k for which $\hat{x}_k - \Delta_k = \underline{x}_k$ and $\hat{x}_k + \Delta_k = \overline{x}_k$. By adding and subtracting these two equalities, we conclude that

$$\hat{x}_k = \frac{\underline{x}_k + \overline{x}_k}{2} \text{ and } \Delta_k = \frac{\overline{x}_k - \underline{x}_k}{2}.$$

Mathematical comment. For each i, the two constraints $\widetilde{x}_i - \Delta_h \le x_i \le \widetilde{x}_i + \Delta_h$ and $\underline{x}_{pr,i} \le x_i \le \overline{x}_{pr,i}$ can be combined into a single set of constraints:

$$x_i^- \le x_i \le x_i^+,$$

where

$$x_i^- \stackrel{\text{def}}{=} \max(\widetilde{x}_i - \Delta_h, \underline{x}_{pr,i}); \quad x_i^+ \stackrel{\text{def}}{=} \min(\widetilde{x}_i + \Delta_h, \overline{x}_{pr,i}).$$

Simplest case: description. Let us consider the simplest case when we have a single lower spatial resolution estimate $\widetilde{X}_1$. In this case, the linear constraints take the form $x_i^- \le x_i \le x_i^+$ and

$$\widetilde{X}_1 - \Delta_l \le \sum_{i=1}^{n} w_{1,i} \cdot x_i \le \widetilde{X}_1 + \Delta_l.$$

Comment. This general expression also includes the case when some cells are not covered by the estimate $\widetilde{X}_1$: for the values corresponding to these cells, we simply have $w_{1,i} = 0$.

Simplest case: derivation. Let us select a variable x_k, $k = 1, \ldots, n$, and let us check which values of x_k are possible.

If the k-th cell is not affected by the estimate $\widetilde{X}_1$, i.e., if $w_{1,k} = 0$, then the only restrictions on x_k come from the prior bounds on x_k and from the higher spatial resolution estimates. Thus, for such a cell, the set of possible values is the interval $[x_k^-, x_k^+]$.

Let us now consider the case when the k-th cell is affected by the estimate $\widetilde{X}_1$, i.e., when $w_{1,k} > 0$. In this case, a possible value x_k must be within the interval $[x_k^-, x_k^+]$, and for the remaining variables x_i, $i = 1, \ldots, k-1, k+1, \ldots, n$, the resulting system of inequalities $x_i^- \le x_i \le x_i^+$ and

$$\widetilde{X}_1 - \Delta_l - w_{1,k} \cdot x_k \le \sum_{i \ne k} w_{1,i} \cdot x_i \le \widetilde{X}_1 + \Delta_l - w_{1,k} \cdot x_k$$

must be consistent.

All the weights $w_{1,i}$ are non-negative. Thus, when $x_i \in [x_i^-, x_i^+]$, the smallest possible value $\underline{s}$ of the sum

$$s \stackrel{\text{def}}{=} \sum_{i \ne k} w_{1,i} \cdot x_i$$

is attained when all x_i attain their smallest possible values $x_i = x_i^-$, and the largest possible value $\overline{s}$ of the sum s is attained when all x_i attain their largest possible values $x_i = x_i^+$:

$$\underline{s} = \sum_{i \ne k} w_{1,i} \cdot x_i^-; \quad \overline{s} = \sum_{i \ne k} w_{1,i} \cdot x_i^+.$$

Thus, we have

$$\sum_{i \ne k} w_{1,i} \cdot x_i^- \le \sum_{i \ne k} w_{1,i} \le \sum_{i \ne k} w_{1,i} \cdot x_i^+.$$

Now, we have two intervals

$$[\widetilde{X}_1 - \Delta_l - w_{1,k} \cdot x_k, \widetilde{X}_1 + \Delta_l - w_{1,k} \cdot x_k]$$

and

$$\left[\sum_{i \ne k} w_{1,i} \cdot x_i^-, \sum_{i \ne k} w_{1,i} \cdot x_i^+\right]$$

that contain the same sum $\sum\limits_{i \ne k} w_{1,i}$. Thus, their intersection must be non-empty, i.e., the lower endpoint of the first interval cannot exceed the upper endpoint of the second interval, and vice versa (one can easily check that if these conditions are satisfied, then the above inequalities are indeed consistent):

$$\widetilde{X}_1 - \Delta_l - w_{1,k} \cdot x_k \le \sum_{i \ne k} w_{1,i} \cdot x_i^+;$$

$$\sum_{i\neq k} w_{1,i}\cdot x_i^- \le \widetilde{X}_1 + \Delta_l - w_{1,k}\cdot x_k.$$

By moving the term $w_{1,k}\cdot x_k$ to the other side of each of the inequalities and dividing both sides of each resulting inequality by a positive number $w_{1,k}$, we conclude that

$$\frac{1}{w_{1,k}}\cdot\left(\widetilde{X}_1 - \Delta_l - \sum_{i\neq k} w_{1,i}\cdot x_i^+\right) \le x_k \le \frac{1}{w_{1,k}}\cdot\left(\widetilde{X}_1 + \Delta_l - \sum_{i\neq k} w_{1,i}\cdot x_i^-\right).$$

Simplest case: resulting formulas. For the cells k which are not affected by the estimate $\widetilde{X}_1$, the resulting bounds on x_k are $[\underline{x}_k, \overline{x}_k]$ with $\underline{x}_k = x_k^-$ and $\overline{x}_k = x_k^+$.

For the cells k which are affected by the estimate $\widetilde{X}_1$ (i.e., for which $w_{1,k} > 0$), the resulting range $[\underline{x}_k, \overline{x}_k]$ has the form

$$\underline{x}_k = \frac{1}{w_{1,k}}\cdot\left(\widetilde{X}_1 - \Delta_l - \sum_{i\neq k} w_{1,i}\cdot x_i^+\right);\ \ \overline{x}_k = \frac{1}{w_{1,k}}\cdot\left(\widetilde{X}_1 + \Delta_l - \sum_{i\neq k} w_{1,i}\cdot x_i^-\right).$$

7 Conclusions and Future Work

We propose a new approach to combining data from different sources, an approach which is a fast practical alternative to joint inversion of multiple datasets. Specifically, in this paper, we consider models that not only have different accuracy and coverage, but also different spatial resolution. To fuse such models, we must account for three different types of approximate equalities:

- each higher spatial resolution estimate is approximately equal to the actual value in the corresponding (smaller size) cell;
- each lower spatial resolution estimate is approximately equal to the average of values of all the smaller cells within the corresponding larger size cell;
- each lower spatial resolution estimate is also approximately equal to the value within each of the constituent smaller size cells, with the accuracy corresponding to the (empirical) standard deviation of the smaller-cell values within the larger cell.

Depending on whether we have probabilistic or interval uncertainty, the approach then uses the least squares or interval technique to combine these approximate equalities. For example, in the least squares approach, we find the desired combined values by minimizing the resulting sum of weighted squared differences.

On the example of simulated (synthetic) geophysical data, we show that model fusion indeed improves the accuracy and spatial resolution of individual models.

In the future, we plan to apply the model fusion techniques to more realistic simulated data and to real geophysical data (and, if necessary, use the results of these applications to further adjust the techniques).

Acknowledgments

This work was supported in part by the National Science Foundation grants:

- Cyber-ShARE Center of Excellence: A Center for Sharing Cyberresources to Advance Science and Education (HRD-0734825), and
- Computing Alliance of Hispanic-Serving Institutions CAHSI (CNS-0540592).

It was also partly supported by Grant 1 T36 GM078000-01 from the National Institutes of Health.

The authors are thankful to all the participants of the 13th International Symposium on Scientific Computing, Computer Arithmetic, and Verified Numerical Computations SCAN'2008 (El Paso, Texas, September 29 – October 3, 2008), Annual Fall Meeting of the American Geophysical Union AGU'08 (San Francisco, California, December 15–19, 2008), and the 2009 Annual Conference of the Computing Alliance of Hispanic-Serving Institutions CAHSI (Mountain View, California, January 15–18, 2009) for valuable suggestions.

References

[1] M. G. Averill, *A Lithospheric Investigation of the Southern Rio Grande Rift*, University of Texas at El Paso, Department of Geological Sciences, PhD Dissertation, 2007.

[2] M. G. Averill, K. C. Miller, G. R. Keller, V. Kreinovich, R. Araiza, and S. A. Starks, "Using Expert Knowledge in Solving the Seismic Inverse Problem", *International Journal of Approximate Reasoning*, 2007, Vol. 45, No. 3, pp. 564–587.

[3] C. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, MIT Press, Boston, Massachusetts, 2001.

[4] J. A. Hole, "Nonlinear High-Resolution Three-Dimensional Seismic Travel Time Tomography", *Journal of Geophysical Research*, 1992, Vol. 97, pp. 6553–6562.

[5] J. M. Lees and R. S. Crosson, "Tomographic inversion for three-dimensional velocity structure at Mount St. Helens using earthquake data", *Journal of Geophysical Research*, 1989, Vol. 94, pp. 5716–5728.

[6] M. Maceira, S. R. Taylor, C. J. Ammon, X. Yang, and A. A. Velasco, "High-resolution Rayleigh wave slowness tomography of Central Asia", *Journal of Geophysical Research*, 2005, Vol. 110, paper B06304.

[7] O. Ochoa, "Towards a fast practical alternative to joint inversion of multiple datasets: model fusion", *Abstracts of the 2009 Annual Conference of the Computing Alliance of Hispanic-Serving Institutions CAHSI*, Mountain View, California, January 15–18, 2009.

[8] O. Ochoa, A. A. Velasco, V. Kreinovich, and C. Servin, "Model fusion: a fast, practical alternative towards joint inversion of multiple datasets", *Abstracts of the Annual Fall Meeting of the American Geophysical Union AGU'08*, San Francisco, California, December 15–19, 2008.

[9] P. Pinheiro da Silva, A. Velasco, M. Ceberio, C. Servin, M. Averill, N. Del Rio, L. Longpré, and V. Kreinovich, "Propagation and Provenance of Probabilistic and Interval Uncertainty in Cyberinfrastructure-Related Data Processing and Data Fusion", In: R. L. Muhanna and R. L. Mullen (eds.), *Proceedings of the International Workshop on Reliable Engineering Computing REC'08*, Savannah, Georgia, February 20–22, 2008, pp. 199–234.

[10] S. Rabinovich, *Measurement Errors and Uncertainties: Theory and Practice*, Springer Verlag, New York, 2005.

[11] C. Servin, O. Ochoa, and A. A. Velasco, "Probabilistic and Interval Uncertainty of the Results of Data Fusion, With Application to Geosciences", *Abstracts of 13th International Symposium on Scientific Computing, Computer Arithmetic, and Verified Numerical Computations SCAN'2008*, El Paso, Texas, September 29 – October 3, 2008, p. 128.

[12] D.J. Sheskin, *Handbook of Parametric and Nonparametric Statistical Procedures*, Chapman & Hall/CRC Press, Boca Raton, Florida, 2007.