Blind Calibration of Sensor Networks

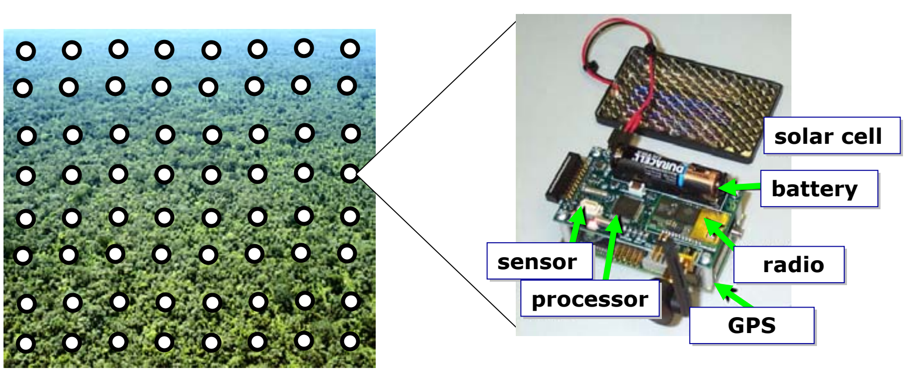
Laura Balzano
(EE, UCLA)
www.ee.ucla.edu/~sunbeam/

Mani Srivastava
(EE, UCLA)
nesl.ee.ucla.edu/people/mbs/

Steve Margulis
(Civil Engineering, UCLA)
www.seas.ucla.edu/~margulis/

Robert Nowak
(EE, UW-Madison)
www.ece.wisc.edu/~nowak

Wireless Sensing Challenges



Networking, Communications, Resource Management, Signal Processing

BUT most work assumes the data are calibrated! These inexpensive sensors are prone to drifting calibration.

Our goal: look for ways to do *in-situ calibration*.

Two Approaches to Calibration

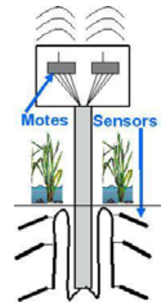
- First Approach: Non-Linear Filters
  - Description of Approach
  - Calibration Results
- Second Approach: Signal Subspace Matching
  - Description of Approach
  - Calibration Results

Outline

- First Approach: Non-Linear Filters
  - Description of Approach
  - Calibration Results
- Second Approach: Signal Subspace Matching
  - Description of Approach
  - Calibration Results

Soil Applications

- Sampling soil is an important task
  - CO_2 in our atmosphere is being absorbed by the soil
  - Ground water quality can be monitored as moisture, nutrients, toxics and pollutants percolate through the soil

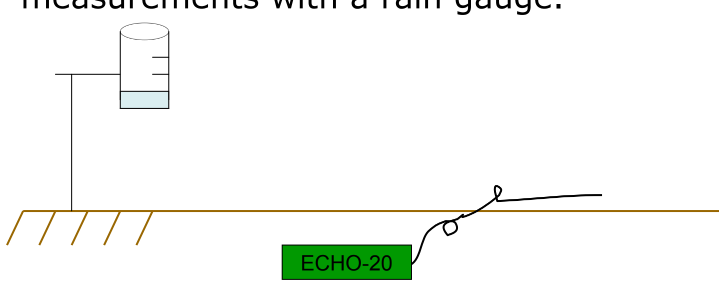


- Sampling soil has many barriers that other sampling problems do not!
  - The field is extremely heterogeneous; Oversampling is not only impossible, the disturbance caused is undesirable.
  - The behavior of the moisture and chemistry of the field depends greatly on soil types, rainfall, topography, plant types, and various other factors.

Dynamical Environment

- Fortunately, because of the importance of soil sampling, environmental engineers develop very complex but thoughtful models for soil dynamics.

- Toy example: We want to track the state of moisture at a point near the surface of the soil.
- We take measurements with an uncalibrated ECHO-20 moisture sensor and precipitation measurements with a rain gauge.



Model

Dynamics Model

$$\begin{pmatrix} y_k \\ q_k \end{pmatrix} = \begin{pmatrix} \delta y_{k-1} + \sqrt{dt}\sigma_w \rho q_k + precip_{k-1} \\ \alpha q_{k-1} + \sqrt{1-\alpha^2} w_{k-1} \end{pmatrix}$$

- Variables to note:
  - y_k is the moisture
  - δ is (time invariant) moisture decay parameter
  - q_k is model error
- Model is Non-Linear
- Model is "forced" by precipitation
- Parameter distributions are non-Gaussian (e.g., moisture is a non-negative quantity)

Model

Measurement Model

$$y = \beta_1 x + \beta_0 + v$$

- Variables to note:
  - β_1 is the calibration gain
  - β_0 is the calibration offset
  - v is the measurement noise
- Our state vector thus has 5 states:

$$Y = \begin{pmatrix} y_k \\ q_k \\ \beta_1 \\ \beta_0 \\ \delta \end{pmatrix}$$

Non-Linear Filtering

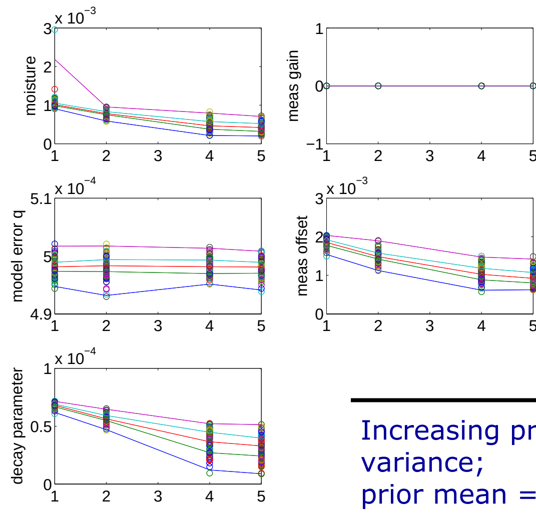
- Kalman Filter ?
  - Our system is not linear nor are the inputs Gaussian.
- Extended Kalman Filter ?
  - Introduces too much error in the linearization step.
- Ensemble Kalman Filter (our choice)
  - A Monte Carlo method which requires less computation than the particle filter, as it assumes very loose Gaussianity in the update step.
- Particle Filter or other methods…
  - Future work for comparison.

Outline

- First Approach: Non-Linear Filters
  - Description of Approach
  - Calibration Results
- Second Approach: Signal Subspace Matching
  - Description of Approach
  - Calibration Results

Results: Calibrating Offset

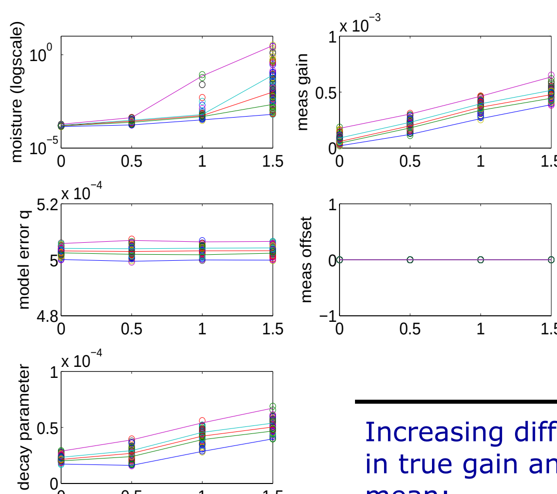
Quantiles (1,25,50,75,99) of RMS error over 100 runs



Increasing prior variance;
prior mean = 0,
true offset = 5

Results: Calibrating Gain

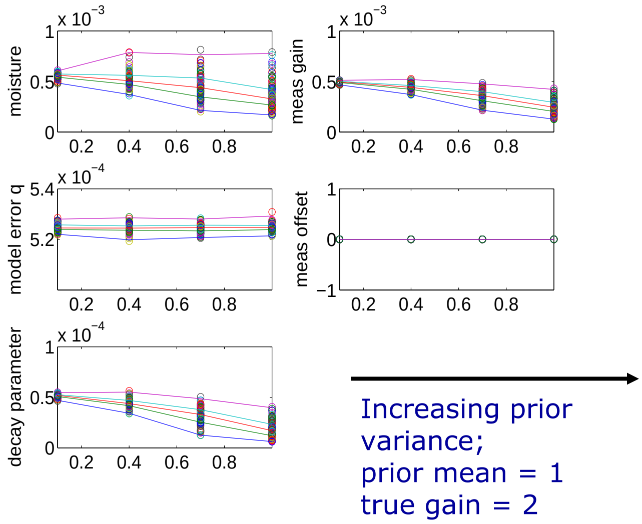
Quantiles (1,25,50,75,99) of RMS error over 100 runs



Increasing difference in true gain and prior mean;
prior variance = 0.4

Results: Calibrating Gain

Quantiles (1,25,50,75,99) of RMS error over 100 runs



Non-Linear Filters

- On this toy model, estimation of calibration parameters is sensitive…
- Next steps:
  - Incorporate other sensor modalities and side information
  - Particle filter: Importance resampling will improve robustness, but is the computational tradeoff worthwhile ?

Outline

- First Approach: Non-Linear Filters
  - Description of Approach
  - Calibration Results
- Second Approach: Signal Subspace Matching
  - Description of Approach
  - Calibration Results

A General Approach

"Uncalibrated" Sensor Measurements: (for *n* sensors)

$$\mathbf{y} = [y(1), \ldots, y(n)]^T$$

Calibrated Measurements: (a linear calibration function)

$$x(j) = \alpha(j)\, y(j) + \beta(j)$$

gain correction
for sensor j

offset correction
for sensor j

Blind Calibration Problem

$$x(j) = \alpha(j)\, y(j) + \beta(j)$$

and with the notation… $\mathrm{Y} \equiv \mathrm{diag}(\mathrm{y}) = \begin{bmatrix} y(1) & & \\ & \ddots & \\ & & y(n) \end{bmatrix}$

$$x = Y\alpha + \beta$$

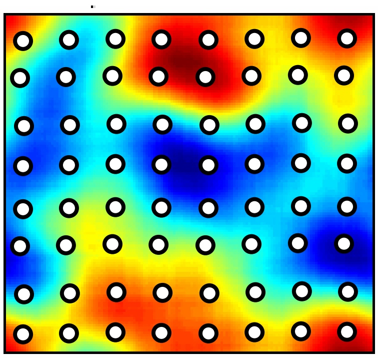
So, given k uncalibrated "snapshots" (e.g., at different times):

$$\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_k$$

Find α and β such that the equations hold for all i = 1,…,k

Without additional assumptions this is an impossible problem.

Calibration in the Field



Pseudocolor map of temperature distribution

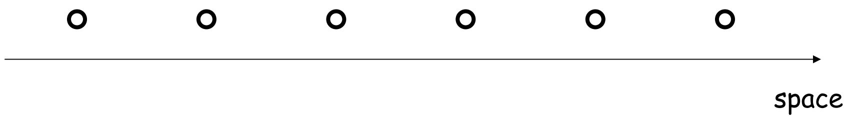
Neighboring sensors in dense deployment make very similar readings

We can automatically calibrate sensor network by forcing readings to agree locally

V. Bychkovskiy, S. Megerian, D. Estrin, and M. Potkonjak, "A collaborative approach to in-place sensor calibration," Lecture Notes in Computer Science, 2634:301–316, 2003.

Calibration by Local Agreement

Linear deployment of sensors:



Ideal (calibrated) sensor readings:

$$x_1 \quad x_2 \quad \cdots \quad x_n$$

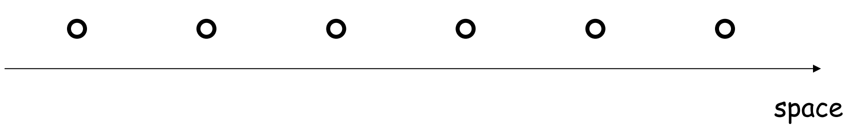
Calibration via local agreement is based on assumption:

$$x_1 - x_2 \approx 0 \quad x_2 - x_3 \approx 0 \quad \ldots \quad x_{n-1} - x_n \approx 0$$

These conditions define a signal subspace (constant functions)

Calibration by Local Agreement

Linear deployment of sensors:



Ideal (calibrated) sensor readings:

$$x_1 \quad x_2 \quad \cdots \quad x_n$$

Calibration assuming second derivatives are approx zero:

$$x_1 - 2x_2 + x_3 \approx 0 \quad \ldots \quad x_{n-2} - 2x_{n-1} + x_n \approx 0$$

These conditions define a signal subspace (linear functions)

Signal Subspaces and Calibration

Ideal (calibrated) sensor readings:

$$\mathrm{x} = [x_1,\ x_2,\ \ldots\ ,\ x_n]^T$$

Calibration via *signal subspace matching* is based on assumption:

$$\mathrm{P\,x} \approx 0$$

where **P** is an orthogonal projection matrix,
and **(I-P)** projects onto the signal subspace.

Examples: **P** could correspond to a projection onto a particular frequency band, a roughness subspace, or any other subspace where the signal should not be

This turns out to be a very exciting and useful generalization!
We ask: When do solutions exist? How do we find them ?

Identifiability

k "snapshots" give us the system of equations:

$$\mathrm{Px}_i = P(Y_i\,\alpha + \beta) = 0$$
$$i = 1, \ldots, k$$

Offsets:

Clearly, we cannot identify the component of β in the signal subspace. This component is indistinguishable from true signal.

Gains:

However, since $\mathbf{y_i}$ "modulates" α, under certain conditions on **P** it is possible exactly recover α up to a global gain factor. We cannot distinguish between α and (scalar constant) x α.

Offset Solutions

$$\mathbf{P}\left(\mathbf{Y}_i \boldsymbol{\alpha} + \boldsymbol{\beta}\right) = \mathbf{0}, \quad i = 1, \ldots, k$$

$$\boxed{\mathbf{P}\,\boldsymbol{\beta} = -\mathbf{P}\,\bar{\mathbf{Y}}\boldsymbol{\alpha}}$$

where $\bar{\mathbf{Y}} = \frac{1}{k}\sum_{i=1}^{k} \mathbf{Y}_i$

Note:

- Every offset solution is a simple function of sensor measurements and gains
- If true signals are zero mean, then offsets can be estimated!
- Otherwise, the offset component in signal subspace cannot be blindly recovered. (but we can recover them with a little bit more information)

Gain Solutions

$$\begin{aligned} \mathbf{0} &= \mathbf{P}\left(\mathbf{Y}_i \boldsymbol{\alpha} + \boldsymbol{\beta}\right) \\ &= \mathbf{P}\left(\mathbf{Y}_i - \bar{\mathbf{Y}}\right)\boldsymbol{\alpha}, \quad i = 1, \ldots, k \end{aligned}$$

A unique solution exists iff the matrix

$$\begin{bmatrix} \mathbf{P}\left(\mathbf{Y}_1 - \bar{\mathbf{Y}}\right) \\ \mathbf{P}\left(\mathbf{Y}_2 - \bar{\mathbf{Y}}\right) \\ \vdots \\ \mathbf{P}\left(\mathbf{Y}_k - \bar{\mathbf{Y}}\right) \end{bmatrix}$$

has rank n-1 (i.e., exactly n-1 columns are independent)

Exact Recovery of Calibration Gains

When does $\begin{bmatrix} \mathbf{PY}_1 \\ \mathbf{PY}_2 \\ \vdots \\ \mathbf{PY}_k \end{bmatrix} \alpha = 0$ have a unique solution?

Assumptions:

Oversampling: The ideal sensor network signals lie in a known *r-dimensional* signal subspace

Randomness: The signals are randomly distributed across snapshots according to an unknown density function with support on the signal subspace

Incoherence: The signal subspace is *incoherent* with the canonical spatial basis (i.e., δ basis)

More info: Balzano and Nowak, "Blind Calibration of Sensor Networks," IPSN 2007

Exact Recovery of Calibration Gains

Theorem 1: Under assumptions A1, A2 and A3, the gains can be perfectly recovered from any $k \geq r$ signal measurements by solving the linear system of equations

$$\begin{bmatrix} \mathbf{PY}_1 \\ \mathbf{PY}_2 \\ \vdots \\ \mathbf{PY}_k \end{bmatrix} \alpha = 0$$

Theorem 2: If the signal subspace is defined by a subset of the DFT vectors (i.e., a frequency-domain subspace), then incoherence condition is automatically satisfied, and the gains can be perfectly recovered from any $k = \lceil \frac{n-1}{n-r} \rceil + 1$ signal measurements

Robust Recovery of Calibration Gains

Gain equations may hold only approximately due to noise/errors:

$$\underbrace{\begin{bmatrix} \mathbf{P}\left(\mathbf{Y}_1 - \bar{\mathbf{Y}}\right) \\ \mathbf{P}\left(\mathbf{Y}_2 - \bar{\mathbf{Y}}\right) \\ \vdots \\ \mathbf{P}\left(\mathbf{Y}_k - \bar{\mathbf{Y}}\right) \end{bmatrix}}_{\mathbf{C}} \boldsymbol{\alpha} \approx 0$$

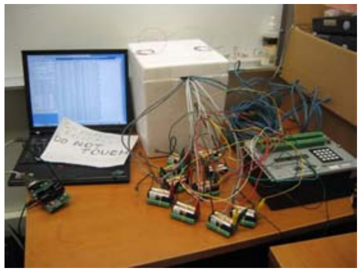
Robust solutions:

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\mathbf{C}\,\boldsymbol{\alpha}\|_2^2$$

Outline

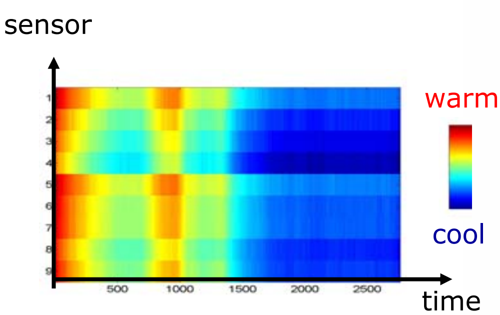
- First Approach: Non-Linear Filters
  - Description of Approach
  - Calibration Results
- Second Approach: Signal Subspace Matching
  - Description of Approach
  - Calibration Results

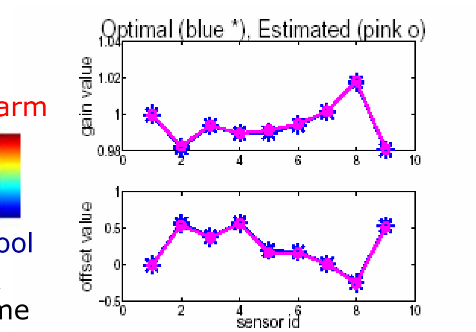
Sensing in a Box



Ideally all sensors should read the same temperature

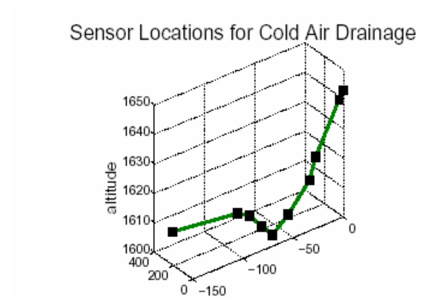
➡ 1-d signal subspace of constant functions



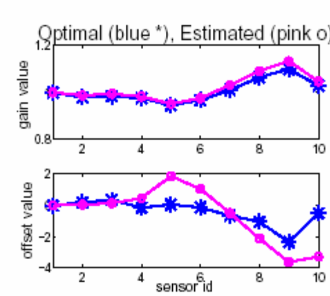


Sensing in the Wild

temperature sensing at James Reserve



4-dimensional signal subspace determined from calibrated sensor data



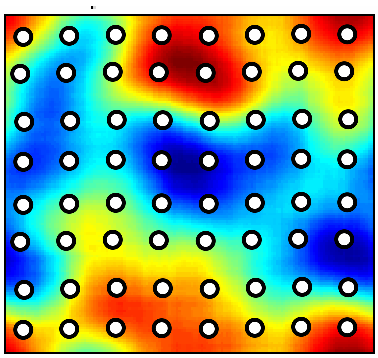
Conclusions

- This formulation is very promising and has provided some important insight into the problem of blind calibration.
- Key necessary condition is "incoherence" between signal subspace and canonical (spatial) basis; we are currently exploring this condition and its relationship to compressed sensing
- Our experience is that solutions are robust to noise and mismodeling in some cases, and sensitive in others; we do not have a good understanding of the robustness of the methodology at this time. Future work includes sensitivity analysis of the set of linear equations.

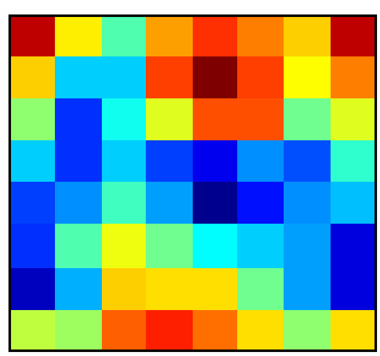
Thank you!

(extra slides)

Simulation Experiment



simulated temperature field



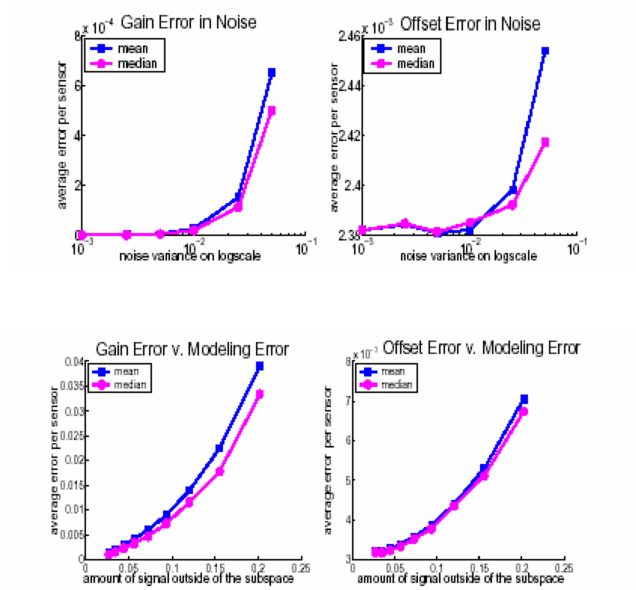
8x8 sensor readings

- field is smoothed GWN process
- approximate signal subspace = span of lowpass DFT vectors

Simulation Experiment

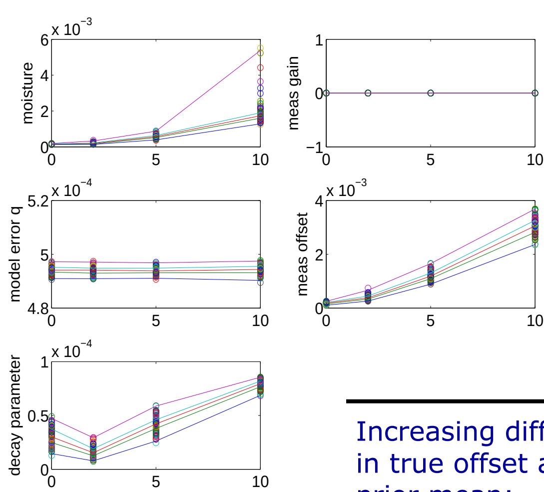
robust to noise

robust to mismodeling



Results: Calibrating Offset

Quantiles (1,25,50,75,99) of RMS error over 100 runs



Increasing difference in true offset and prior mean; prior variance = 2