LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

An Information Value Function For Nonparametric Gaussian Processes

H. Wei, W. Lu, and S. Ferrari
Laboratory for Intelligent Systems and Controls (LISC),
Department of Mechanical Engineering and Materials Science,
Duke University

MURI 2 Workshop 2012, MIT (Boston, MA)
April 18, 2012

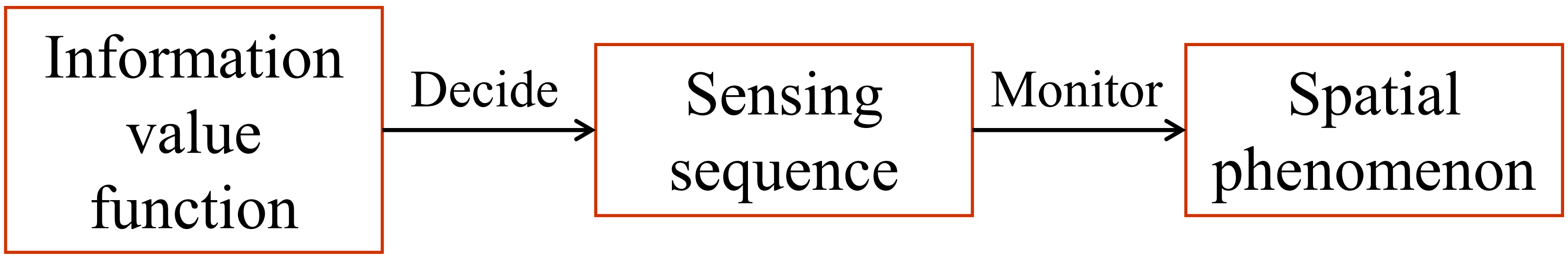
LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Outline

- Introduction and Motivation
- Problem Formulation
  - Sensor planning
- Methodology
  - Gaussian process: data representation
  - Information value function: greedy algorithm
- Simulation
  - Criteria: estimating error and estimating variance
  - Comparison with random algorithm
- Summary and Conclusions

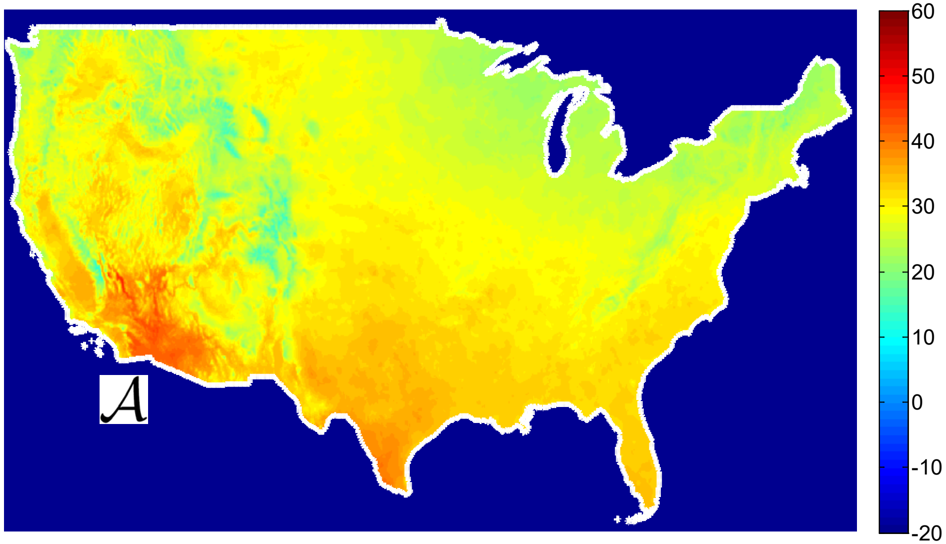
Introduction and Motivation

Information Value Functions for Sensor Planning:



- State of the art
  - "An Information Potential Approach to Integrated Sensor Path Planning and Control", G. Zhang, et al.
  - "A Comparison of Information Functions and Search Strategies for Sensor Planning", 2012. S. Ferrari, et al.
- Main contribution:
  - Method for continuous spatial phenomenon

Spatial Phenomenon



Nomenclature

$\mathcal{A}$: Region of Interest

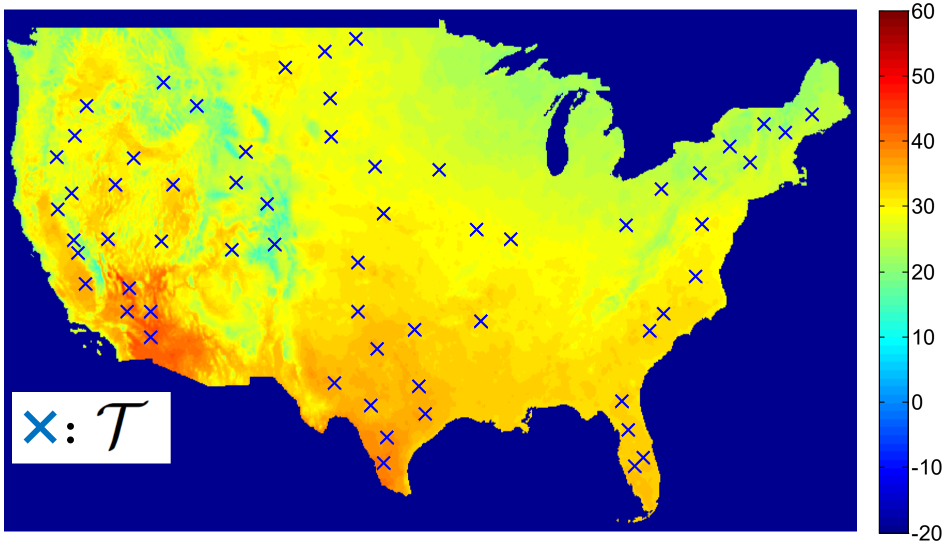
g: Spatial phenomenon

Spatial phenomenon,

$$g(\mathbf{x}), \mathbf{x} \in \mathcal{A}$$

- Defined over two-dimensional region of interest, $\mathcal{A} \subset \mathbb{R}^2$
- Time invariant
- Max temperature of the continental United States in August

Targets



Nomenclature

$\mathcal{A}$: Region of Interest

g: Spatial phenomenon

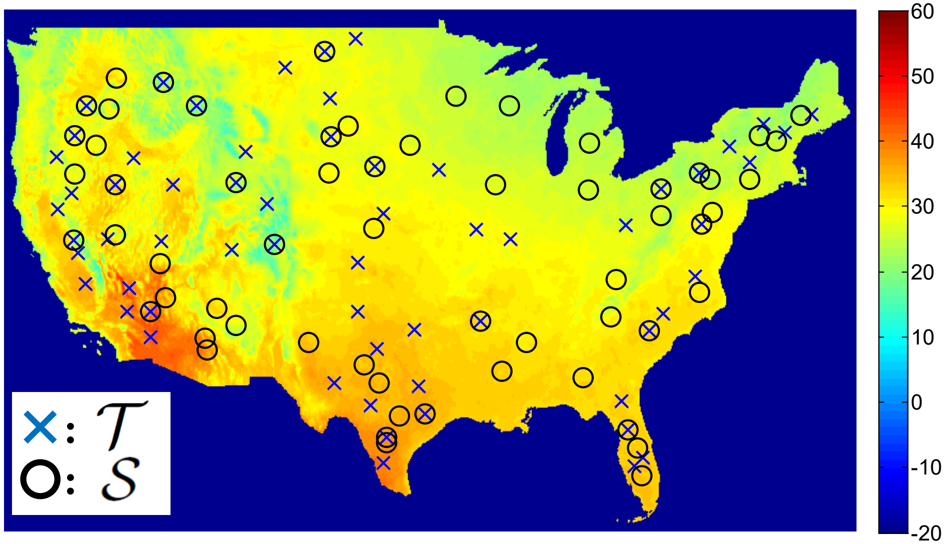
$\mathcal{T}$: Targets

Set of targets,

$$\mathcal{T} = \{\mathbf{t}_i | i = 1, \cdots, r\}, \mathbf{t}_i \in \mathcal{A}$$

- Represent points of highest interest in $\mathcal{A}$
- $\mathcal{T}$ and r can change over time

Accessible Sensing Locations



Nomenclature

$\mathcal{A}$: Region of Interest

g: Spatial phenomenon

$\mathcal{T}$: Targets

$\mathcal{S}$: Accessible sensing locations

Set of accessible sensing locations, $\mathcal{S}$

$$\mathcal{S} = \{\mathbf{s}_i | i = 1, \cdots, l\} \subset \mathcal{A}$$

- Known *a priori*
- Size of $\mathcal{S}$ is limited

LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Sensor Measurement Model

At the kth time step t_k,

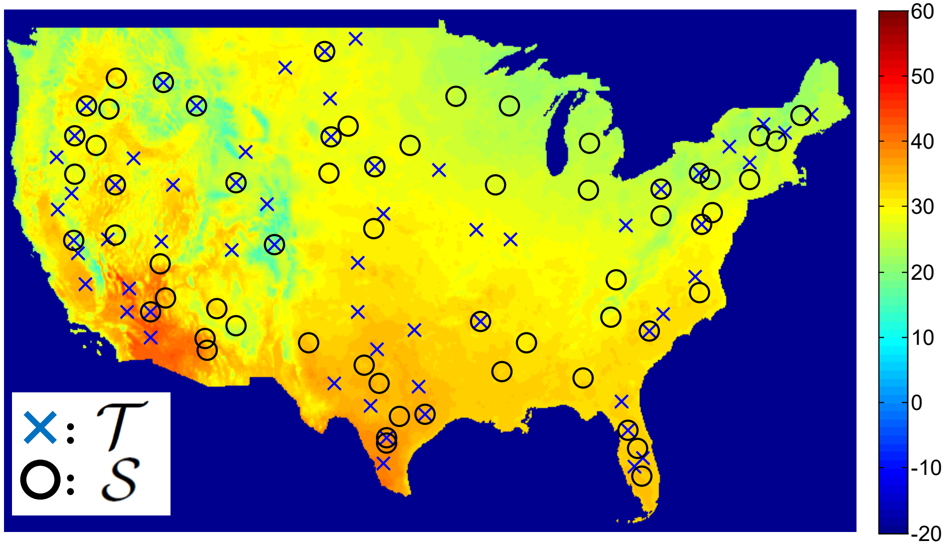
- Sensor takes one measurement from,
  - $\mathbf{y}_k \in \mathcal{S}$
- sensor model:
  - $z_k = g(\mathbf{y}_k) + \varepsilon$
  - ε : additive Gaussian noise, $\mathcal{N}(0, \sigma^2)$
- History of sensing sequence

$$\mathbf{Y}_k = [\mathbf{y}_1 \mid \cdots \mid \mathbf{y}_k]$$

- History of observations

$$\mathbf{Z}_k = [z_1 \mid \cdots \mid z_k]$$

Sensor Planning Problem



Nomenclature

$\mathcal{A}$: Region of Interest

g: Spatial phenomenon

$\mathcal{T}$: Targets

$\mathcal{S}$: Accessible sensing locations

$\mathbf{Y}_k = [\mathbf{y}_1 | \cdots | \mathbf{y}_k]$

$\mathbf{Z}_k = [z_1 | \cdots | z_k]$

Sensor planning:

Decide $\mathbf{Y}_k^* = [\mathbf{y}_1^* | \cdots | \mathbf{y}_k^*]$ that minimizes

$$err = \frac{1}{r}\sqrt{\sum_{\mathbf{x}_i \in \mathcal{T}} (g(\mathbf{x}_i) - \mathrm{E}[f(\mathbf{x}_i)|\mathbf{Y}_k, \mathbf{Z}_k])^2}$$

for $\mathbf{y}_i^* \in \mathcal{S}$, $i = 1, \ldots, k$

LISC
LABORATORY FOR INTELLIGENT SYSTEMS AND CONTROLS

Gaussian Process: Model the Spatial Phenomenon

Estimation of spatial phenomenon:

- $g(\mathbf{x}) \xleftarrow{\text{estimation}} f(\mathbf{x}), \mathbf{x} \in \mathcal{A}$
- $f(\mathbf{x}) \sim$ Gaussian process

Gaussian process:

$$f(\mathbf{x}) \sim \mathrm{GP}(m(\mathbf{x}), k(\mathbf{x}_1, \mathbf{x}_2))$$

$$m(\mathbf{x}) = \mathrm{E}[f(\mathbf{x})]$$

$$k(\mathbf{x}_1, \mathbf{x}_2) = \mathrm{E}[(f(\mathbf{x}_1)\text{-}m(\mathbf{x}_1))(f(\mathbf{x}_2)\text{-}m(\mathbf{x}_2))]$$

Notation:

- $\mathbf{X}_{\mathcal{T}} = [\mathbf{x}_1 \mid \cdots \mid \mathbf{x}_r]$, $\mathbf{x}_i \in \mathcal{T}$, $i = 1,\ldots,r$
- $\mathbf{f}(\mathbf{X}_{\mathcal{T}}) = [\, f(\mathbf{x}_1) \;\; \cdots \;\; f(\mathbf{x}_r) \,]^{\mathrm{T}}$
- $\mathbf{m}(\mathbf{X}_{\mathcal{T}}) = [\, m(\mathbf{x}_1) \;\; \cdots \;\; m(\mathbf{x}_r) \,]^{\mathrm{T}}$
- $\mathbf{K}(\mathbf{X},\mathbf{Y})[i,\, j] = k\,(\mathbf{x}_i, \mathbf{y}_j)$, $\mathbf{x}_i = \mathbf{X}[:, i]$, $\mathbf{y}_j = \mathbf{Y}[:, j]$

Gaussian Process: Model the Spatial Phenomenon

Prior distribution on targets:

$$\mathbf{f}(\mathbf{X}_T) \sim \mathcal{N}\left(\mathbf{m}(\mathbf{X}_T), \mathbf{K}[\mathbf{X}_T, \mathbf{X}_T]\right)$$

Prediction:

$$\begin{bmatrix} \mathbf{Z}_k \\ \mathbf{f}(\mathbf{X}_T) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{m}(\mathbf{Y}_k) \\ \mathbf{m}(\mathbf{X}_T) \end{bmatrix}, \begin{bmatrix} \mathbf{K}(\mathbf{Y}_k, \mathbf{Y}_k) + \sigma^2 I & \mathbf{K}(\mathbf{Y}_k, \mathbf{X}_T) \\ \mathbf{K}(\mathbf{X}_T, \mathbf{Y}_k) & \mathbf{K}(\mathbf{X}_T, \mathbf{X}_T) \end{bmatrix}\right)$$

Posterior distribution on targets:

$$\mathbf{f}(\mathbf{X}_T)|\mathbf{Y}_k, \mathbf{Z}_k \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)\ ,$$

where,

$$\boldsymbol{\mu}_k = \mathbf{m}(\mathbf{X}_T) + \mathbf{K}(\mathbf{X}_T, \mathbf{Y}_k)[\mathbf{K}(\mathbf{Y}_k, \mathbf{Y}_k) + \sigma^2 I]^{-1}(\mathbf{Z}_k - \mathbf{m}(\mathbf{Y}_k))$$

$$\boldsymbol{\Sigma}_k = \mathbf{K}(\mathbf{X}_T, \mathbf{X}_T) - \mathbf{K}(\mathbf{X}_T, \mathbf{Y}_k)[\mathbf{K}(\mathbf{Y}_k, \mathbf{Y}_k) + \sigma^2 I]^{-1}\mathbf{K}(\mathbf{Y}_k, \mathbf{X}_T)$$

EDG: Measure Expected Difference by Sensing Action

Kullback-Leibler (KL) divergence: difference between *P*(x) and *Q*(x)

$$D(P||Q) = -\int_{-\infty}^{\infty} \ln \frac{P(x)}{Q(x)} P(x)dx$$

Choose $\{\mathbf{y}_k, z_k\}$ to maximize

$$D(p(\mathbf{f}(\mathbf{X}_T)|\mathbf{Y}_k,\mathbf{Z}_k)||p(\mathbf{f}(\mathbf{X}_T)|\mathbf{Y}_{k-1},\mathbf{Z}_{k-1}))$$

But z_k is unknown ⇨ expected discrimination gain (EDG)

$$\hat{\varphi}_D(\mathbf{f}(\mathbf{X}_T); z_k|\mathbf{y}_k,\mathbf{Y}_{k-1},\mathbf{Z}_{k-1}) = \int D(p(\mathbf{f}(\mathbf{X}_T)|\mathbf{Y}_{k-1},\mathbf{Z}_{k-1},\mathbf{y}_k,z_k)||p(\mathbf{f}(\mathbf{X}_T)|\mathbf{Y}_{k-1},\mathbf{Z}_{k-1})) \times p(z_k|\mathbf{Y}_{k-1},\mathbf{Z}_{k-1},\mathbf{x}_k)dz_k.$$

EDG for Multivariate Gaussian Distribution

For multivariate Gaussian distributions, the EDG is

$$\hat{\varphi}_D(\mathbf{f}(\mathbf{X}_T); z_k|\mathbf{y}_k, \mathbf{Y}_{k-1}, \mathbf{Z}_{k-1}) =$$
$$\int_{-\infty}^{\infty} \frac{1}{2}(\mathrm{tr}(\mathbf{\Sigma}_{k-1}^{-1}\mathbf{\Sigma}_k) - \ln(\frac{\det(\mathbf{\Sigma}_k)}{\det(\mathbf{\Sigma}_{k-1})}) - r$$
$$+ (\boldsymbol{\mu}_k - \boldsymbol{\mu}_{k-1})^T \mathbf{\Sigma}_{k-1}^{-1}(\boldsymbol{\mu}_k - \boldsymbol{\mu}_{k-1}))\mathcal{N}(\mu_{z_k}, \sigma_{z_k})dz_k$$

where

$$\mu_{z_k} = \mathbf{K}(\mathbf{y}_k, \mathbf{Y}_{k-1})[\mathbf{K}(\mathbf{Y}_{k-1}, \mathbf{Y}_{k-1}) + \sigma^2 I]^{-1}(\mathbf{Z}_{k-1} - \mathbf{m}(\mathbf{Y}_{k-1}))$$
$$+ m(\mathbf{y}_k)$$

and

$$\sigma_{z_k} = \mathbf{K}(\mathbf{y}_k, \mathbf{Y}_{k-1})[\mathbf{K}(\mathbf{Y}_{k-1}, \mathbf{Y}_{k-1}) + \sigma^2 I]^{-1}\mathbf{K}(\mathbf{Y}_{k-1}, \mathbf{y}_k)$$
$$- k(\mathbf{y}_k, \mathbf{y}_k)$$

Integrate EDG Analytically

Integrating EDG analytically over z_k to reduce computation

$$\hat{\varphi}_D(\mathbf{f}(\mathbf{X}_T); z_k|\mathbf{y}_k, \mathbf{Y}_{k-1}, \mathbf{Z}_{k-1}) = \frac{1}{4}c\sigma^3_{z_k}\sqrt{\pi} + \frac{1}{2}\sigma_{z_k}\sqrt{\pi}\Big(\mathrm{tr}(\boldsymbol{\Sigma}^{-1}_{k-1}\boldsymbol{\Sigma}_k) - \ln(\frac{\det(\boldsymbol{\Sigma}_k)}{\det(\boldsymbol{\Sigma}_{k-1})}) - r + \mathbf{V}_1^T\mathbf{M}_1^T\boldsymbol{\Sigma}^{-1}_{k-1}(\mathbf{M}_1\mathbf{V}_1 - 2\mathbf{M}_2\mathbf{V}_2) + \mathbf{V}_2^T\mathbf{M}_2^T\boldsymbol{\Sigma}^{-1}_{k-1}\mathbf{M}_2\mathbf{V}_2\Big)$$

where

$$\mathbf{M}_1 = \mathbf{K}(\mathbf{X}_T, \mathbf{Y}_{k-1})(\mathbf{K}(\mathbf{Y}_{k-1}, \mathbf{Y}_{k-1}) + \sigma^2 I)^{-1}$$

$$\mathbf{M}_2 = \mathbf{K}(\mathbf{X}_T, \mathbf{Y}_k)(\mathbf{K}(\mathbf{Y}_k, \mathbf{Y}_k) + \sigma^2 I)^{-1}$$

$$\mathbf{V}_1 = \mathbf{Z}_{k-1} - \mathbf{m}(\mathbf{Y}_{k-1})$$

$$\mathbf{V}_2 = [(\mathbf{Z}_{k-1} - \mathbf{m}(\mathbf{Y}_{k-1}))^T \quad \mu_{z_k} - m(\mathbf{y}_k)]^T$$

$$c = \mathrm{diag}(\mathbf{M}_2^T \Sigma^{-1}_{k-1}\mathbf{M}_2)[k]$$

Information Value function: Greedy Algorithm

Greedy algorithm for sensor planning:

Input: functions: $\mathbf{m}, \mathbf{K}(\cdot,\cdot)$;
sets: $\mathcal{S}, \mathcal{T}$;
scalars: maximum number of observations, N_f
Output: sensing location sequence $\mathbf{Y}_{N_f}$
begin
$\mathbf{Y}_{N_f} \leftarrow \emptyset$;
for $k = 1 : N_f$
$\mathbf{y}_k = \underset{\mathbf{y}_k \in \mathcal{S}}{\text{argmax}}\, \hat{\varphi}_D(\mathbf{f}(\mathbf{X}_T); z_k | \mathbf{y}_k, \mathbf{Y}_{k-1}, \mathbf{Z}_{k-1})$
$\mathbf{Y}[k] = \mathbf{y}_k$
$z_k = g(\mathbf{y}_k) + \varepsilon$
$\mathbf{Z}[k] = z_k$
endfor
return Y
end

Data and Assumptions

Data:

Maximum temperature distribution of the continental United States territory in August, 1997

Prior distribution:

- $m(\mathbf{x}) = 0$
- $k(\mathbf{x}_1, \mathbf{x}_2) = e^{-\|\mathbf{x}_1 - \mathbf{x}_2\|^2}$

Other parameters:

parameter	value	memo
r	41	number of targets
l	100	number of accessible sensing locations
σ	1.0 [°C]	standard variance of sensing noise

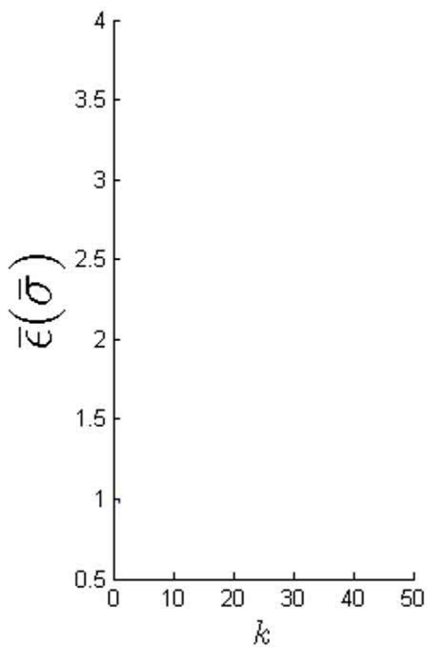
Performance Criteria and Result

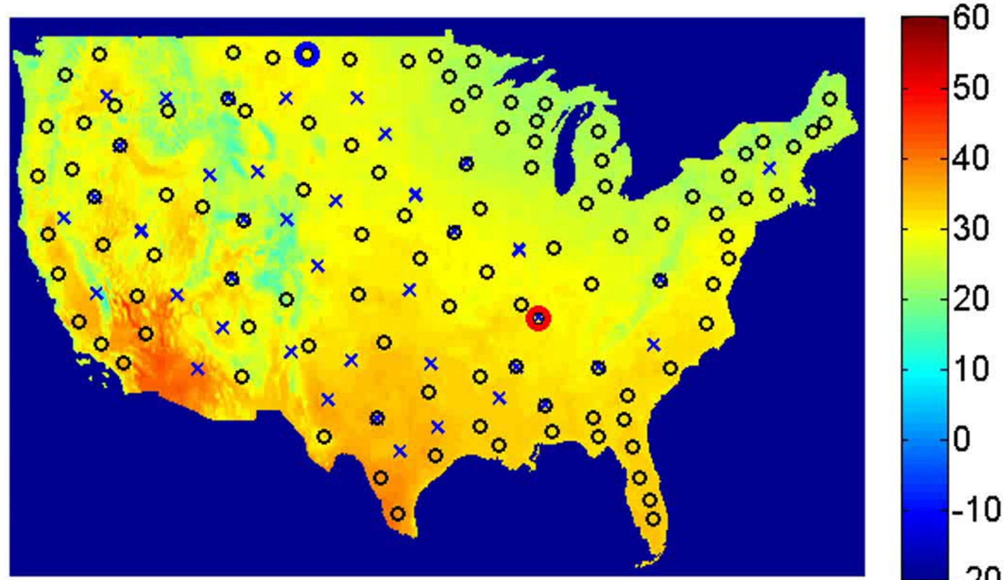
Estimation error: $\bar{\epsilon}(\hat{\phi}_D) = \frac{1}{r}\|\boldsymbol{\mu}_k - \mathbf{g}(\mathcal{T})\|$

Estimation variance: $\bar{\sigma}(\hat{\phi}_D) = \frac{1}{r}\mathrm{tr}(\boldsymbol{\Sigma}_k)$

Random algorithm: $\mathbf{y}_k \in \mathcal{S} \sim 1/l$

- greedy error
- random error
- greedy variance
- random variance





Summary

- Problem formulation
  - Sensor planning
- Methodology
  - Gaussian process: data representation
  - Information value function: greedy algorithm
- Contribution
  - Greedy algorithm for continuous spatial phenomenon
- Future work
  - More covariance function
  - More nonparametric Bayesian models for various phenomena

Thanks
Welcome Questions

References:

- G. Zhang, W. Lu, and S. Ferrari, "An Information Potential Approach to Integrated Sensor Path Planning and Control," IEEE Transactions on Robotics, submitted.
- W. Lu, G. Zhang, S. Ferrari, M. Anderson, and R. Fierro, "An Information Potential Approach for Tracking and Surveilling Multiple Moving Targets using Mobile Sensor Agents, " Journal of Defense Modeling and Simulation, accepted.
- S. Ferrari, G. Zhang, and C. Cai, "A Comparison of Information Functions and Search Strategies for Sensor Planning," IEEE Transactions on Systems, Man, and Cybernetics - Part B, Vol. 42, No. 1, 2012.