Air Quality Monitoring Network Design Using Pareto Optimality Methods for Multiple Objective Criteria

Paul D. Sampson [1]

with

Peter Guttorp [1] & **David M. Holland** [2]

[1] NRCSE, University of Washington

[2] U.S. Environmental Protection Agency

supported by a grant from the U.S. EPA

http://www.nrcse.washington.edu/

Outline

1. Spatial Monitoring Network Design Objectives in the Statistics Literature
2. Air Quality Monitoring Objectives
3. Multiple Objective Monitoring Network Design
   A. References to selected applications
   B. Pareto Optimality approach
4. Summary

Three components of an optimal spatial design problem (D.L. Zimmerman, Optimal spatial design.

In: *Encyclopedia of Environmetrics*)

1. Specification of a <u>design space</u> of candidate sites (finite, or in principle, continuous spatial domain).
2. Specification of a <u>model</u> for the existing observations (if any) and the potential observations at candidate sites.
3. Specification of **an** <u>optimality criterion</u>.

National and Regional Surveys

Program names link to pages with maps, measurements, and other program information.

Table of Contents

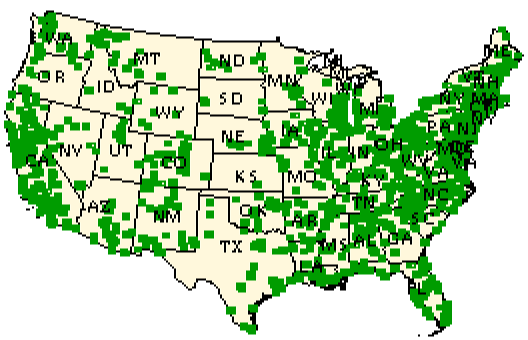
http://www.epa.gov/cludygxb/programs/index2.html

National

And

Regional

Surveys



http://www.epa.gov/cludygxb/programs/namslam.html

NAMS/SLAMS- National Air Monitoring Stations/ State and Local Air Monitoring Stations

Program Name (Acronym)	NAMS/SLAMS
Agency	EPA,State &loc agen ow
Year Initiated	1979
Measures	AIR-criteria pollutnts, visibility/fine particulates, toxics
Collection Source	
Point	Yes
Source	No
Transect	No
Other area	No
Locations for Data Collection	5000 samplrs,3150 sites
Temporal Interval	Hourly,Pb&PM10 variable
Sampling Design	**Selected**
Data Available	Yes
Accessible	EPA reg offices,AIRS
Extent for Reporting	Primarily urban,some rura
Annual Funding	FY 96-$36 M in fed fnds
Cost per Site for Installation	$5 K-$100 K per site
Cost per Site for Op. & Mgmt	Avrg $1 K per site/yrly
Partners	
International	No
Agency	EPA Regions
State	State agencies
Local	Local agencies,contrctrs
Authorities/Reason for Running Prg.	40CFR58
Users of Data per Year	450
Program Meets Metadata Standards	No
Expansion of Prog (Needed/Not)	Not needed
Contact Person	David Lutz
Phone #	919-541-5476

1. Spatial Monitoring Network Design Objectives:

Frameworks for optimal spatial design:

1. **Exploratory, random sampling, and/or space-filling designs.**
   Nychka, D. and Saltzman, N. (1998), Design of air quality networks. In *Case Studies in Environmental Statistics*
2. **Designs for estimating a regression function in a spatially correlated field (with known spatial covariance):**
   W.G. Müller (2000, *Collecting Spatial Data: Optimum Design of Experiments for Random Fields*)
   R.L. Smith (NSF-CBMS Lecture Notes, in prep).

3. **Designs for estimation of the spatial covariance or variogram:**
   W.G. Müller and D.L. Zimmerman (1999, *Environmetrics*)
   R.L. Smith (NSF-CBMS Lecture Notes, in prep).

4. **Designs for optimal spatial prediction**, including designs specifically concerned with **assessment of regulatory thresholds**: identification of sites or regions exceeding thresholds (for one or more pollutant measures) and/or the risk or expected cost of misclassifying sites according to a threshold (classifying a "contaminated" site as safe or vice versa).

2. Air Quality Monitoring Objectives

Multiple scientific objectives are explicit in current guidelines for air quality monitoring networks. Three perspectives:

1. **Four general purposes for the ambient air monitoring program are** (http://www.epa.gov/oar/oaqps/qa/monprog.html):

- **to judge compliance with and/or progress made toward meeting ambient air quality standards;**
- **to activate emergency control procedures that prevent or alleviate air pollution episodes;**
- **to observe pollution trends throughout the region, including non-urban areas; and**
- **to provide data base for research evaluation of effects: urban, land-use, and transportation planning; development and evaluation of: abatement strategies and diffusion models.**

2. Specific objectives of monitoring sites in the SLAMS network according to U.S. Code of Federal Regulations, Part 58, Appendix D are:

- **to determine representative concentrations in areas of high population density;**
- **to determine highest concentrations expected to occur in the area covered by the network;**
- **to observe pollution trends throughout the region, including non-urban areas; and**
- **to determine general background concentration levels.**

See also EPA guidelines (U.S.EPA, 1998, EPA-454/R-98-002):

- **to determine the extent of air pollution transport into and out of an area.**

So, what does this suggest for numerical objective design criteria?

1. **Criterion associated with maps of <u>probability of exceeding standards</u> (**computed how?**)**
2. **Criteria assessing <u>spatial prediction accuracy</u>** (kriging error, entropy)
3. **Utility functions for other criteria:**
   - **network representation of <u>population</u>**
   - **network representation of <u>sources</u>**
   - **<u>Cost</u>**

- **Remark:** 1 & 2 above are achievable using recent methods of spatial analysis

- **Many of these objectives assume some form of <u>prior information</u> regarding where:**
  - **people live,**
  - **pollutant sources are, and**
  - **high and background levels of pollutant concentrations are expected.**
- **Information about where high and background concentrations may be expected requires, probably, a <u>combination of available monitoring data and</u> <u>air quality model predictions</u> in network design calculations.**
- **Statistical network design methodology has apparently never been recommended to attempt to meet these objectives.**

From **"Guidance for Network Design and Optimum Site Exposure for $PM_{2.5}$ and PM_{10}"** (EPA-454/R-99-022):

Network Design Philosophies: statistical methods accounting for correlation, model-based methods, random sampling, systematic sampling, judgmental sampling, heterogeneous strategies.

"Monitoring networks for criteria pollutants always use judgmental sampling strategies that consider where source emissions are in relation to populations and which way the wind blows. … Most of this guidance is based on judgmental network design, though it is expected that networks will involve more of the hybrid approach as they are evaluated as future $PM_{2.5}$ measurements and improved aerosol modeling techniques are developed."

3. Multiple Objective Monitoring Network Design:

A. References to selected applications in publications

B. Pareto optimality approach

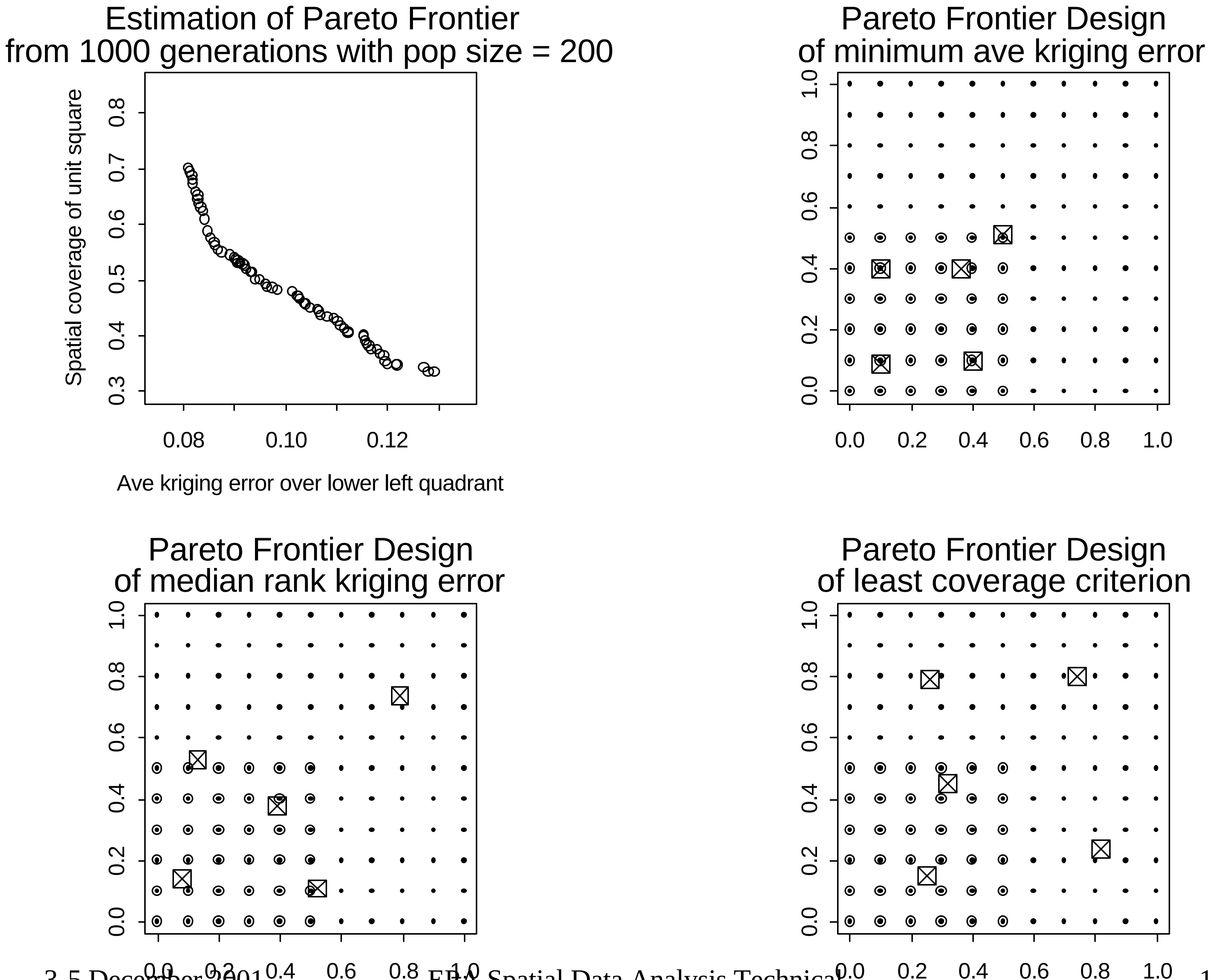
B. Pareto optimal designs:

Alternative to optimization of a composite design criterion is to investigate the space of "Pareto optimal" designs.

Given a vector of n design criteria, $X_1,\ldots,X_n$, a design having attained numerical criteria values $a_1,\ldots,a_n$, is said to dominate another design attaining values $b_1,\ldots,b_n$, if $a_i \leq b_i$ for all i, and for at least one j, $a_j < b_j$.

A design that is not dominated by any other is said to be Pareto optimal, and the Pareto optimal set or Pareto frontier is the set of all Pareto optimal designs.

Consideration of the Pareto optimal set will allow better understanding (compared with optimization of a single criterion) of the trade-offs necessary to obtain greater relative efficiency on given criteria.


Estimation of Pareto Frontier
from 1000 generations with pop size = 200
Spatial coverage of unit square
Ave kriging error over lower left quadrant
0.3
0.4
0.5
0.6
0.7
0.8
0.08
0.10
0.12
Pareto Frontier Design
of minimum ave kriging error
Pareto Frontier Design
of median rank kriging error
Pareto Frontier Design
of least coverage criterion
0.0
0.2
0.4
0.6
0.8
1.0

4. Summary

- Fact: Multiple air quality monitoring objectives
- Some objectives require incorporation of prior information
  - In models and model-based estimates of errors for spatial estimation
  - In particular objectives of air monitoring that concern effects of pollutant sources & transport, and effects on human health
- Pareto optimal design calculations provide an effective way to make decisions in the context of multiple objectives.
- Evolutionary computation algorithms provide feasible tools for Pareto optimization.