A Unifying Framework for Optimal Sensor Placement Strategy and Data Analysis

Ming Jin and Professor Costas J. Spanos
Department of Electricalk Engineering and Computer Sciences
University of California, Berkeley

Supported by:
Funded by: NATIONAL RESEARCH FOUNDATION

Motivation

Sensor networks in the built environment provide critical information for building control systems to ensure residents comfort and energy efficiency. Data from sensors are used to *estimate the building's states, evaluate current power usage*, and *predict energy demands*, all of which are valuable for building managers to make decisions and understand the behavior of building energy consumption.

We are required to design the ***placement strategy*** to obtain as much information as possible with very limited number of sensors. Based on the data acquired in real time, we want to visualize the current building environment, and ***recognize the trends*** underlying the huge amount of data to predict the building states.

2012 Main Objectives

It is our objective to develop a unifying framework for sensor placement and sensing data analysis:

- Optimize the sensor locations by incorporating knowledge from physics into spatial statistical models
- Reconstruct time-varying maps of building environments, such as temperature profile
- Analyze data in real-time by dimension reduction and make predictions

Apart from its statistical nature, the spatial distribution of temperature, humidity and gas concentrations is governed by laws of physics. With relevant knowledge in the physical domain, we can improve the accuracy of statistical models even further…

Problem Definition

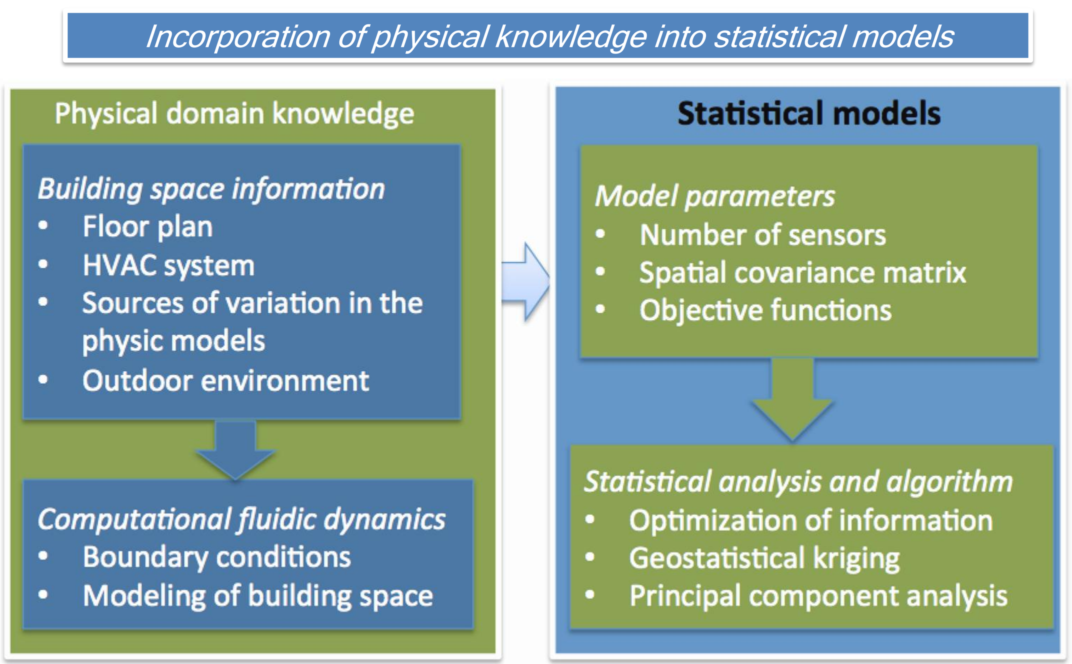
Most works either focus on the purely statistical methods by adopting questionable assumptions, or rely on physics by considerably simplifying the model. In order to improve the accuracy of our model, we need to:

Optimize the sensor placement to provide most sensing information by combining statistical methods with physical constraints

We then analyze the data collected by the sensors:

- Dimension reduction by *online PCA* methods
- Reconstruction of data by spatial kriging

Approach and Methodology



Computational Fluidic Dynamics Modeling

- Temperature distribution in the spatial domain is governed by convection, diffusion, sources and transient behaviors:

$$\underbrace{\frac{\partial \rho \phi}{\partial t}}_{\text{transient}} + \underbrace{\frac{\partial \rho U_j \phi}{\partial x_j}}_{\text{convection}} = \underbrace{\frac{\partial}{\partial x_j}\left(\Gamma_{\phi,eff}\frac{\partial \phi}{\partial x_j}\right)}_{\text{diffusion}} + \underbrace{S_\phi}_{\text{sources}}$$

- The above partial differential equation governs the system behavior. All terms are coupled and must be solved simultaneously.
- No closed form solution, yet commercial programs deliver simulation results, e.g. ANSYS FLUENT

Optimized Sensor Placement Algorithm

Input: Spatial covariance information, number of sensors k, available locations for sensors

Inputs from physical models

Output: selected sensor locations

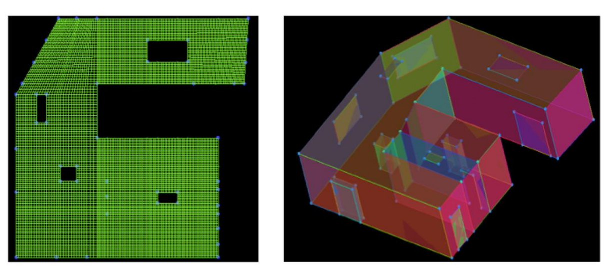
Begin:
  For i=1 **to** k
    For y **in remaining locations**,
      Calculate $I(X;y) = \sum_{x \in X} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$
    Select y which gives largest $I(X;y)$
  End

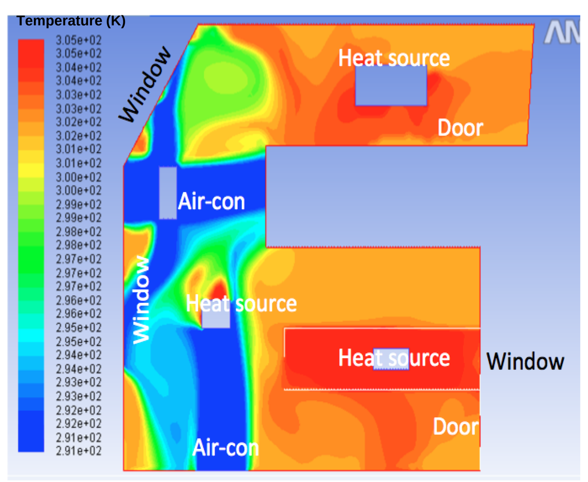
Objective: max mutual information

The sensors are chosen so that the mutual information between observed and unobserved locations are maximized

Case study: CREST Space modeling

Simplified 2D/3D model based on floor plan and possible heat sources, such as air-cons, windows, doors, *(right)* and simulated **temperature profile** *(bottom)*

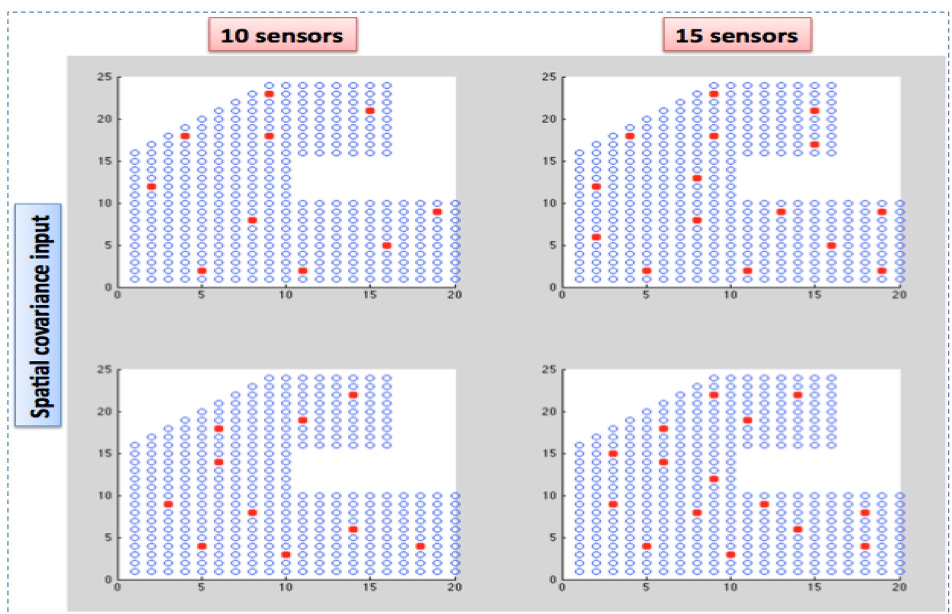




- Provide a **reference** temperature profile in a given situation at a particular time
- Estimate temperature spatial covariance
- Capable of taking real-time data as inputs to refine the model

Sensor Placement Strategy in CREST

- Sensor locations depend on spatial covariance inputs
- Locations where temperature varies most and can provide most information about others are chosen



Online Principal Component Analysis

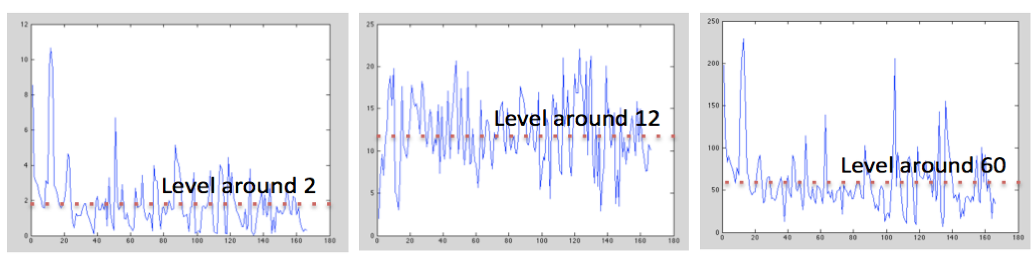
For time series analysis:

- Reduce data dimensionality
- Trend recognition and data forecasting

Online PCA as a piecewise linear approximation to the signal:

- Updates the PC dynamically
- Since the signal fluctuates a lot, it is difficult for standard PCA to capture all the information with one static PC

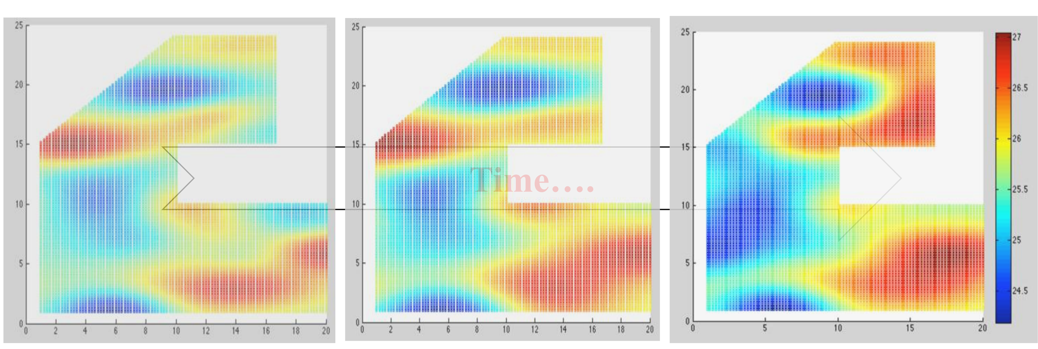
Analysis of data from dense sensor networks of strong spatial correlation. Plots of *mean square error* of PCA reconstructed signal with respect to each sensor node.



The performance of ***online PCA*** (left) is much better than ***standard PCA*** with the same and larger number of principal components (middle and right).

Continuous Map Reconstruction by Kriging

Building environment's **temperature profile** can be reconstructed in real time by spatial Kriging, which takes sensor data as inputs and estimates the temperature at **unobserved spots**, thus creating a continuous map



The evolution of temperature profile over time can help us **visualize** the occupants activity throughout the day

Conclusion and References

- Investigating optimal sensor placement strategy that incorporates physical domain knowledge based on computational fluidic dynamics into statistical models, so as to use as few sensors as possible to fulfill the sensing requirement.
- Explored data analysis methods to find trends in sensing data and reconstruct real-time building environment profiles, by online PCA and spatial statistical Kriging.
- Proposed a framework that combines sensor deployment, data analysis and data visualization

References:

1) Krause, A., A. Singh, et al. (2008). "Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies." J. Mach. Learn. Res. 9: 235-284.
2) Wang, X., X. Wang, et al. (2011). Towards Optimal Sensor Placement for Hot Server Detection in Data Centers. Proceedings of the 2011 31st International Conference on Distributed Computing Systems, IEEE Computer Society: 899-908.
3) Shumway, R. and D. Stoffer (2006). Time Series Analysis and Its Applications: With R Examples (Springer Texts in Statistics), Springer.

Future Goals

- Further explore the incorporation of physical domain knowledge in statistical models
- Develop and integrate real time building environment reconstruction programs
  - ➢ Online algorithms to detect patterns, trends, and important changes in the building atmosphere
  - ➢ Inputs to building control systems to improve comfort and energy efficiency