Cyberinfrastructure for Risk Forecasting and Communication: Application to the Great Lakes Observatory System

Peter Adriaens*,**, Anna Michalak*, Nancy Love*, Mingyan Liu**, Edith Parker ***, Thomas Finholt****Joseph DePinto***** , Tad Slawecki*****, Steven Ruberg******, David Schwab******

*Department of Civil and Environmental Engineering, The University of Michigan, 1351 Beal Ave., Ann Arbor, MI 48109-2125 (E-mail: adriaens@umich.edu; amichala@umich.edu), **Department of Electrical Engineering and Computer Science, The University of Michigan, 1301 Beal Ave., Ann Arbor, MI 48109-2122 (E-mail: Mingyan@umich.edu), ***School for Public Health – Health Behavior and Health Education, The University of Michigan, Observatory Drive, Ann Arbor MI 48109 (Email: eparker@umich.edu), ****School of Information, University of Michigan, (Email: finholt@umich.edu), ****Limno-Tech, Inc., 501 Avis Drive, Ann Arbor, MI 48108 (E-mail: jdepinto@limno.com), *****Great Lakes Environmental Research Laboratory (GLERL)

ABSTRACT

Environmental sensing encompasses a broad range of systems (surface waters, groundwater, atmospheric, etc…) and spatially distributed attributes (microbial, chemical, physical) that describe the system and its perturbations in response to human interference (industrial discharge, agricultural runoff, etc…). Microbiological pollution represents one of the most widespread impairments of potable and recreational waters, yet the sampling design for public health protection exhibits significant uncertainties due to sample selection constraints, and the delay in obtaining analytical results for effective decision-making. With the advances in wireless sensor technology, networks have the promise to provide useful spatio-temporal representation of environmental signals that cover large geographical areas. As State and Federal agencies have started to invest in these systems, there is a need to design network optimization strategies that explicitly incorporate sensor characteristics, microbiological criteria and economic constraints in the decisions. To enable adoption of these technologies, we need to ensure that the value proposition of environmental sensing cyberinfrastructure is defined not only from a technical perspective (i.e. What technical problems are we trying to solve?), but also from a market-based perspective (i.e. What are the business fundamentals that will allow these technologies to come to the market?).

These concepts capitalize on on-going early stage data collection efforts by the Great Lakes Environmental Research Laboratory (GLERL) in Lake Erie (wireless sensor hubs to analyze causes and forecasting of hypoxia), the Macomb County Health Department in Lake St. Clair (wireless chemical, and off-line microbial analysis at water intakes), the Great Lakes Commission (GLC) managed Great Lakes Observatory System (GLOS), the NSF-supported WATer and Environmental Research Systems Network (WATERS Network; http://www.watersnet.org/), and the Macomb County Public Health Department which manages the Lake St. Clair Regional Monitoring Project (http://www.lakestclairdata.net/). Leveraging each of these, currently separate, efforts towards deploying a data-driven sensing network for risk forecasting and communication in the Lake Huron-Lake Erie Corridor will allow the University of Michigan and its private and government partners to position itself to establish a cyberinfrastructure testbed in the Great Lakes region under the NSF Major Research Equipment and Facilities Construction (MREFC) program. Funded by congressional appropriation in 2011, this program releases funding to provide unique capabilities at the frontiers of science and engineering. Several authors of this white paper (Finholt, DePinto), and the Association of Environmental Engineering and Science Professors (AEESP; Adriaens, Vice President; Love, Board Member) have been involved over for the last few years in defining the science, education, and cyberinfrastructure needs of the WATERS Network. This paper illustrates technical advances and market-based strategies in cyberinfrastructure (hardware, software, data integration and visualization) as applied to address microbial sensing needs.

The overarching goal of the white paper is to develop a framework to establish a microbial and chemical risk forecasting cyberinfrastructure network along the Lake Huron-Lake Erie corridor.

1. THE WATERS NETWORK

The WATERS Network is a distributed network for research on complex environmental systems. It emphasizes research on the nation's water resources related to human-dominated natural and built environments. The network will be comprised of: interacting field sites with an integrated CI; a centralized technical resource staff and management infrastructure to support interdisciplinary research through data collection from advanced sensor systems, data mining and aggregation from multiple sources and databases; cyber-tools for analysis, visualization, and predictive multi-scale modeling that is dynamically driven. As such, the network will transform workforce development in the water-related intersection of environmental science and engineering, as well as enable educational engagement opportunities for all age levels. The scientific goal and strategic intent of the Network is to transform our understanding of the earth's water cycle and associated biogeochemical cycles across spatial and temporal scales-enabling quantitative forecasts of critical water-related processes, especially those that affect and are affected by human activities. This strategy will develop scientific and engineering tools that will enable more effective adaptive approaches for resource management.

The need for the network is based on three critical deficiencies in current abilities to understand large-scale environmental systems and thereby develop more effective management strategies. First we lack basic data and the infrastructure to collect them at the needed resolution. Second, we lack the means to integrate data across scales from different media (paper records, electronic worksheets, web-based) and sources (observations, experiments, simulations). Third, we lack sufficiently accurate modeling and decision-support tools to predict the underlying processes or subsequently forecast the effects of different management strategies. The network will foster cutting-edge science and engineering research that addresses major national needs (public and governmental) related to water and include, for example: (i) water resource problems, such as impaired surface waters, contaminated ground water, water availability for human use and ecosystem needs, floods and floodplain management, urban storm water, agricultural runoff, and coastal hypoxia; (ii) understanding environmental impacts on public health; (iii) achieving a balance of economic and environmental sustainability; (iv) reversing environmental degradation; and (v) protecting against chemical and biological threats. WATERS has been supported by $ 1M of NSF funding for 3 years has a Program Office, and has supported pilot projects across the US. The Program Office and committee have been established to prepare an application for MREFC funding in 2011.

Figure 1. Lake Huron –Erie Corridor

Why the Great Lakes? There is a strong rationale for choosing the Great Lakes in general, and the lake Huron-Lake Erie corridor as a test bed. First, they are representative of the challenges facing the management of large water bodies in the United States and elsewhere in the world. Second, the lakes are a critical freshwater resource for the United States and Canada, impacting national economic sectors such as agriculture, power generation, steel production, shipping, drinking water and bottling industries, tourism, and natural resources management. Also, given the potential impacts of climate change on water resources in more arid areas of the western (and now Southeastern) U.S., there will be a growing pressure to divert Great Lakes water to those areas, thus creating a significant management issue. In addition, it will be possible to take advantage of existing programs and partnerships in collecting data and planning research activities.

The CI will take into account the properties of collected data, and the representation and utility of the collected data. To maximize the utility of the network information to public stakeholders and for research purposes, the proposed project will focus on translating the modeled data using (near) real time communication modes, including interactive axis grid displays, and teleconference uplinks. Advancing the theoretical understanding of spatio-temporal risks through an application-driven sensor network design, and implementation of an effective risk communication strategy through environmental CI, offers unique advantages to integrate sensor network implementation in policy designs for multi-stakeholder source water protection.

2. MICROBIAL SENSING: PROBLEM DEFINITION AND MOTIVATION

Microbiological contamination represents one of the most widespread impairments of potable source water and recreational waters. To protect public health and ensure the microbiological and chemical safety of source water, the USEPA has required the development of source water assessment and protection programs (SWAP) to delineate the source water protection area, to inventory contaminant sources and to determine the susceptibility of the source area to contamination. Implementation of these plans requires extensive monitoring programs for biological and physical water quality indicators, in support of the quantitative guidelines developed for human health protection. Specifically for microbiological pollution, the guidelines usually recommend collecting a minimum number of water samples, estimating the density of a target (indicator) organism in the samples, computing a number of summary statistics and comparing them with predefined limits [EPA-86, NRC-04]. Information on the presence and levels of pathogenic organisms is required routinely and quickly in order to monitor the risk of waterborne disease to human health, often using a combination of microbial analysis and predictive (pathogen loading) modeling tools [EPA-99]. In most cases, there continues to be a lack of scientific data to support monitoring schemes that would provide the most meaningful information about whether the guidelines are met. For example, the EMPACT (Environmental Monitoring for Public Access and Community Tracking) Beaches project [EPA-05] recently reported on how the compounded uncertainty of the collected summary statistics and the delay in obtaining analytical results led to misdiagnoses of problems and therefore to incorrect recreational water management recommendations.

To reduce the uncertainty associated with delayed decision-making, a wide range of model types, and complexities, are used [summarized in EPA-99]. Considering the quick turnaround time needed for decisions of a recreational and consumptive nature, simple models are most often used. Computer models that predict pathogen concentrations by simulating the dominant mixing and transport processes in the receiving water include modules to characterize point and non-point sources of pathogens to establish the loading rates, estimate the dominant fate and transport processes to predict pathogen distribution, and interpret the model output to find the pathogen concentration at a point of interest to determine the need for an advisory. The predictive capability of available methods and models exhibits considerable uncertainty in space and time, and requires either measured pathogen concentrations (diagnostic) or surrogate triggers to validate the likely occurrence of a microbial event.

Advances in wireless technology and the design of intelligent sensors have the promise to improve environmental monitoring of water bodies, by providing useful spatio-temporal representations of environmental signals over large geographical areas [EGPS-01, SOPHME-04]. A wireless sensor network needs to be co-optimized to sample physical heterogeneity while maximizing network longevity and robustness under harsh environmental conditions, characterized by internal (e.g. calibration) and external (e.g. weather) failure [SBWMAMRTRG-03]. Network designs such as multi-hop wireless mesh, biologically-inspired decentralized management, and tiered embedded networks have been proposed to address the robustness and longevity concerns, but have to date not been significantly constrained by specific environmental applications [e.g. PK-05].

The current deployment of wireless networks is largely limited to fixed-buoy systems focused on water quantity and physical characteristics of oceans and watersheds. Yet, public health officials are increasingly taking an active interest in wireless sensing technology to improve modeling and forecasting of public health and ecological risks, as well as for real-time monitoring of existing exposure [ASEGE-03]. For example, NOAA's Great Lakes Environmental Research Laboratory (GLERL) has deployed the Real Time Environmental Coastal Observation Network (RECON) pilot in the Lake Huron-Erie corridor. RECON (Figure 1) consists of six wireless observation systems that make detailed real-time scientific data - including various combinations of wind and wave information, air temperature, current and temperature profiles, dissolved oxygen, pH, turbidity, and water temperature - available through the Internet. Collected information is housed on servers located at GLERL (Ann Arbor) for both real-time and archival access. GLERL has agreed to make this data freely accessible to share experiences in design and implementation -- such as difficulties in establishing reliable data transfer protocols -- of the RECON infrastructure in order to inform the development of tools for the optimization of sensor networks.

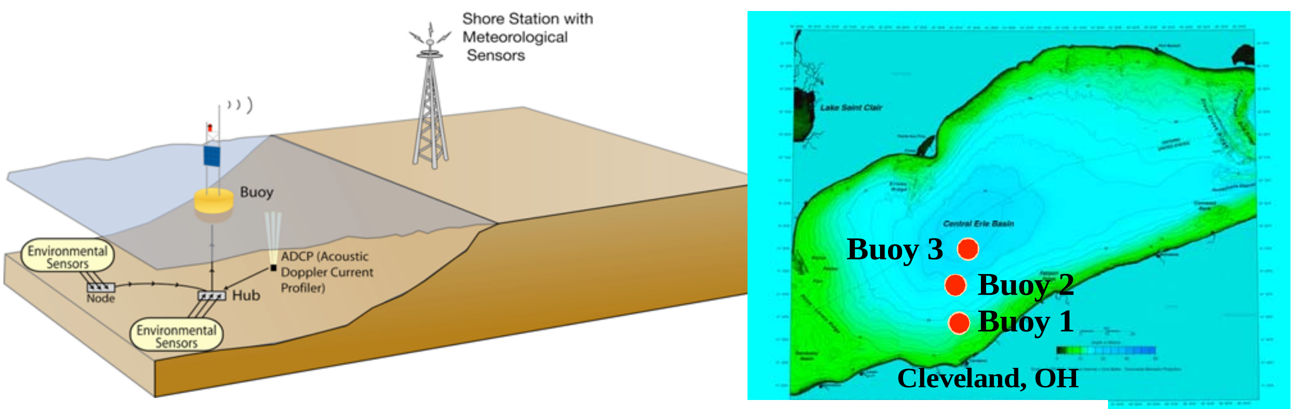

Figure 2. *Configuration and deployment of RECON in Lake Erie (NOAA-GLERL)*

These properties would be applicable to microbial systems, whose behavior is adaptive and responds to variations of various parameters in their environment, such as temperature, redox potential, pH and organic carbon, as well as extreme weather events such as precipitation and wind- or current-induced mixing [RDG-88, CPRL-01, JGRPCW-05]. The challenge presented to a wireless network design with application to microbial contamination is to align microbial action criteria with sensor characteristics (surrogate measurements vs. diagnostic analysis) as constrained by the economics of network deployment within the physical boundaries and a priori knowledge of the natural system (Figure 3). The opportunity to advance the theoretical understanding through an application-driven network optimization strategy will

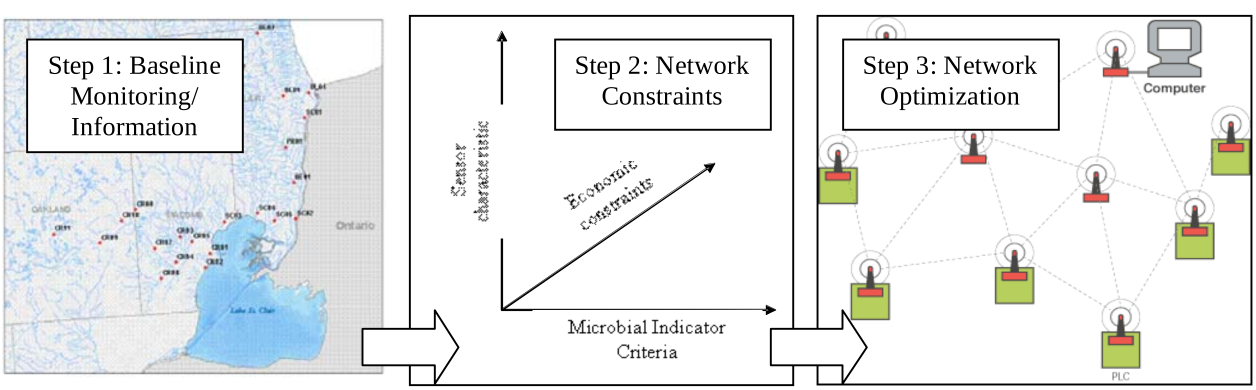

Figure 3. *Conceptualization of the proposed wireless network optimization strategy for microbial attributes constrained by environmental monitoring and economic considerations*

aid in the integration of wireless sensor networks for source water protection.

3. DYNAMICS AND MEASUREMENT OF MICROBIAL INDICATORS AND ANCILLARY DATA IN SOURCE WATERS

The main causes of microbial contamination in surface waters include municipal discharges such as non-disinfected or poorly disinfected sewage treatment plant (STP) effluents [HEHHSLAC-93], combined sewage overflows (CSO) [MDMT-96], and urban stormwater [e.g. DM-93]. While the fecal pollution impacts of STPs have been recognized for a long time (and considered in locating STP outfalls either offshore in deep water or downstream from urban areas) CSOs and stormwater are discharged throughout urban areas and impact the receiving water and its uses [e.g. NRC-98]. Extreme precipitation and high water tables decrease the efficiency of onsite sewage disposal and may increase the likelihood of microorganisms in water systems. Increased urbanization has and will continue to alter watersheds and freshwater flows, resulting in contamination from both point sources and nonpoint sources.

Waterborne bacterial pathogens include *Salmonella*, *Shigella*, *Vibrio*, *Campylobacter*, *Yersinia*, and pathogenic *E. coli* strains. Indicator bacterial analysis includes total coliforms, fecal coliforms, E. coli, fecal streptococci, enterococci, and *Clostridium perfringens* [FMH-00, NRC-98], though problems remain in their use. This is in part due to the discrepancy in microbiological screening between finished waters, and source or recreational waters [AGGLSY-99], considering the range of point and nonpoint sources of fecal contamination. To fulfill various regulations, early warning of potential microbial contamination requires a rapid, simple, broadly applicable technique. For applications of health risk confirmation and to identify the source of a microbial contamination problem, the time frame and investment in indicators, indicator approaches, and methods must be greater [NRC-04]. ***The necessity for both descriptive and predictive analysis of these highly complex microbial ecosystems requires accurate detection and quantification of microbial parameters as a function of matrix variables describing the relevant physico-chemical environments on both a spatial and temporal basis.*** Databases such as EPAs Storage and Retrieval (STORET) system indicate that the distribution of water quality indicators, and the relationship between microbial and surrogate parameters for microbial events is highly site specific. Surrogate indicators for microbial pathogens, such as particles [BHBRLMA – 05], bacteriophages [LMMCCGCGSSJ-03], pH, temperature, redox potential, turbidity, and organic carbon [BROV-02; A-05], and wet/dry weather events [e.g. JGRPCW-05], have been observed at statistically significant levels, and are used in predictive models.

The prediction of pathogen concentrations in source waters is largely based on simulating the dominant mixing and transport processes, and by incorporating multiple input parameters [EPA-99]. Considering the non-conservative nature of microbial contamination, its irregular distribution in space and time [e.g. EPA-05, JGRPCW-05], and variable concentration due to growth and interaction between microbial communities, the statistical output afforded by these models often results in misdiagnosis of water quality problems and incorrect management decisions. New modeling approaches such as artificial neural networks (ANN) have shown promising applications for microbial quality prediction in surface waters [e.g. NLB-02, BL-03], due to the capacity of these models to be trained, and interpret complex, inter-related, and often non-linear, relationships between multiple parameters. However, ***these models do not explicitly address the spatial and temporal microbial distribution and its associated uncertainty, and therefore do not inform monitoring network placement.***

3.1. Water Quality Metrics

Quantitative tests for indicator bacteria are used in monitoring surface drinking water intakes because these waters often show evidence of fecal contamination and are usually treated with filtration and disinfection. Interpretation of indicator data in recreational water applications is different again because the exposure can be more irregular and involves a more limited population at risk. Hence, wireless network design can only be optimized once data quality criteria are defined and quantified. ***Data quality is defined as the ability of the sensor network to represent the parameter distribution within the entire system, where this ability is defined based on specific criteria.*** Depending on the application or water quality management goal, these criteria may include: (i) Minimization of average uncertainty in spatial distribution; (ii) Minimization of single maximum point uncertainty in the system; (iii) Maximized ability to determine whether the parameter is above/below a given threshold at a receptor node. The choice of criteria and its associated requirements for network design will be aligned with NRC [NRC-04] recommendations (Figure 4) for phased monitoring. In all cases, however, the data quality assessment goes well beyond issues of measurement precision and sensor calibration. The implications for this task are that speed, cost, specificity, and sensitivity will drive the achievable data quality metrics and thus the associated network design objectives.

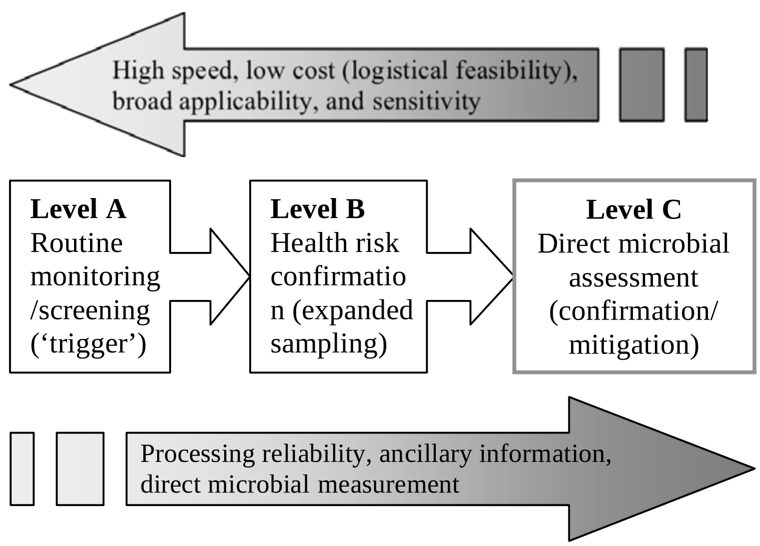

***Figure 4**. Phased monitoring framework for selection of data quality criteria for waterborne pathogens, indicating needs and criteria of network attributes*

Level A. Data quality for screening and routine monitoring is characterized by an emphasis on good spatial and/or temporal coverage and frequency of monitoring. Hence, the cost per node is a major driver for comprehensive sampling for screening purposes with broad applicability to a number of geographic locations, various types of watersheds, and different water matrices are preferred. Considering this driver, the sensor nodes will focus on collecting ancillary data, and on using spatial/temporal correlations between physical/chemical measurements and microbial events to inform the spatial characterization of possible microbial events, rather than using direct microbial analysis. Work in this task could include setting up univariate and multivariate statistical correlations using ANOVA- and ANN-based analysis of Lake St. Clair Data (http://www.lakestclairdata.net/), as well as review of published literature to inform the goal of minimizing the average spatial uncertainty of a potential emergent trigger of an emergent microbial problem in Step 2. Once screening has identified a potential problem ('trigger' event), the second phase involves more detailed information to confirm public health risks.

Level B. Data quality for health risk assessment is characterized by more extensive sampling to "maximize the ability to determine whether a microbial threshold criterion is exceeded at selected receptor nodes." The aim of Level 2 actions is to assess the need for further management actions (e.g., beach closures, boil-water orders). Hence, the data quality is driven by risk classification with quantified uncertainty (e.g. 95% chance that a target threshold is exceeded). A typical approach involves expanded sampling with supplemental information (ancillary and microbial) to determine whether the response is repeatable over space and time and to determine whether the signal persists. From a sensor network perspective, ancillary data predictions will be informed by increasing microbial indicator attributes (i.e. increase in sensor heterogeneity), and more reliable processing methods are used to confirm that the result is not an artifact. Exceedance criteria of source water for drinking and recreational uses, as well as sampling frequency criteria (e.g. days to weeks for recreational waters) may be used as recommended by EPA [e.g. EPA-86. EPA-98, EPA-03], using either a maximum concentration or summary statistics (geometric mean of at least five samples collected over 30 days). Monte-Carlo type simulations could be applied to the correlations developed or used for Level A, to evaluate the impact of measurement and correlation uncertainty on the predictive outcome of the threshold or the relevant summary statistics used for guidance criteria.

Level C. Data quality for source identification and mitigation is a diagnostic assessment of confirmed microbial contamination and identification (source attribution), and represents the highest level of sensing in terms of specificity and sensitivity (impacting cost, response time, and power requirements). The assumption is that source identification will be based as a first approximation on error-prone, but easy-to-use, metrics such as the fecal coliform:fecal streptococci (FC:FS) ratios (>4: predominantly human, <0.7 non-human) to imply human and non-human sources of contamination [e.g. F-74, ECDVMM-97]. Rather than using mean values, a frequency distribution within indicative values will be used to inform the network design. Molecular deterministic methods used for identifying microorganisms are expensive and time-consuming and will be incorporated as a cost-function for network design considerations, rather than as a water quality measure.

3.2. 'Trigger' vs. Diagnostic Sensing Technology

The highly heterogeneous nature of microbial occurrence and activity or viability in natural systems indicates a need for distributed microbial sensing capabilities. The current approaches for microbial (indicator) detection and quantification in source water emphasize off-line membrane-culturing (24 hour), molecular tools (hours), or flow-cytometry (FCM)-based (minutes) analyses, techniques which to date are sub-optimal because of time delay, cost, or ease-of-use reasons. Hence, the integration of distributed sensing with economic constraints of the cost of the network will require a technology implementation commensurate with the phased criteria described in Figure 4. 'Trigger' technology provides early warning of impending hazard, whereas diagnostic technology quantitatively and specifically describes the type of hazard (Figure 5).

State-of-the-art microbial sensors compatible with wireless network systems emphasize optical detection methods for chlorophyll content, and flow cytometers, which are capable of quantifying total numbers, and some target organisms. Both the RECON system and the Networked Aquatic Microbial System (NAMOS), which consists of ten stationary buoys and one mobile robotic boat for real-time, in-situ measurements and analysis of chemical and physical factors governing the abundances and dynamics of microorganisms at biologically-relevant spatiotemporal scales, focus on chlorophyll measurements. Woods Hole and UT-Galveston have deployed field FCM systems for microbial monitoring of plankton; bacterial sensors are either off-line or near real-time but have not been deployed in open water systems. The array of different technology platforms currently available for microbial sensing is provided in Table 1.

Near-term technological capabilities for achieving automated, low-cost ($500 per sensor node), and robust detection of microbial parameters of interest in field settings have focused on micro-electromechanical systems (MEMS) and improvements in the environmental application of molecular tools [IAAW-99, SSR-02, GSA-04, NRC-04]. To achieve acceptable specificity and applicability in the environment, the performance characteristics (e.g., detection limit, setup time, adaptability, matrix interferences) of these approaches vary [reviewed in GSA-04] Once relevant spatio-temporal correlations between surrogate (non-microbial) and diagnostic variables for microbial events are incorporated in a spatial modeling and uncertainty analysis framework, sensor network designs can be informed based on the types and characteristics of sensors that would be required to capture the parameter field [ASEGE-03]. The development of a sensor that can function in a real environment is a far greater challenge than one that can operate in a lab; to achieve the sensitivity, selectivity and reliability that is necessary, we use a multi-modal approach where one "sensor" can sense many different analytes, or can sense one analyte with many different sensitivities.

Table 1. *Comparison of Selected Technology Platform Attributes for Microbial Sensing*

Methodology	Detection Limit[a] (cfu/mL or cfu/g)	Setup Time	Adaptability	Matrix Interference
Plating Techniques	1	1-3 days	Excellent	Low
Bioluminescence	10^3-10^4	½ h	Low	Medium
Piezoelectric	10^6	5h	Good	High
Impedance	1-10^5	6-24 h	Moderate/good	Medium
Flow Cytometry	10^2-10^3	½ h	Good	Medium
Acoustic	5x10^4-10^6	3h	Moderate	High
Electrochemical	10^3	½-2 h	Low	Low

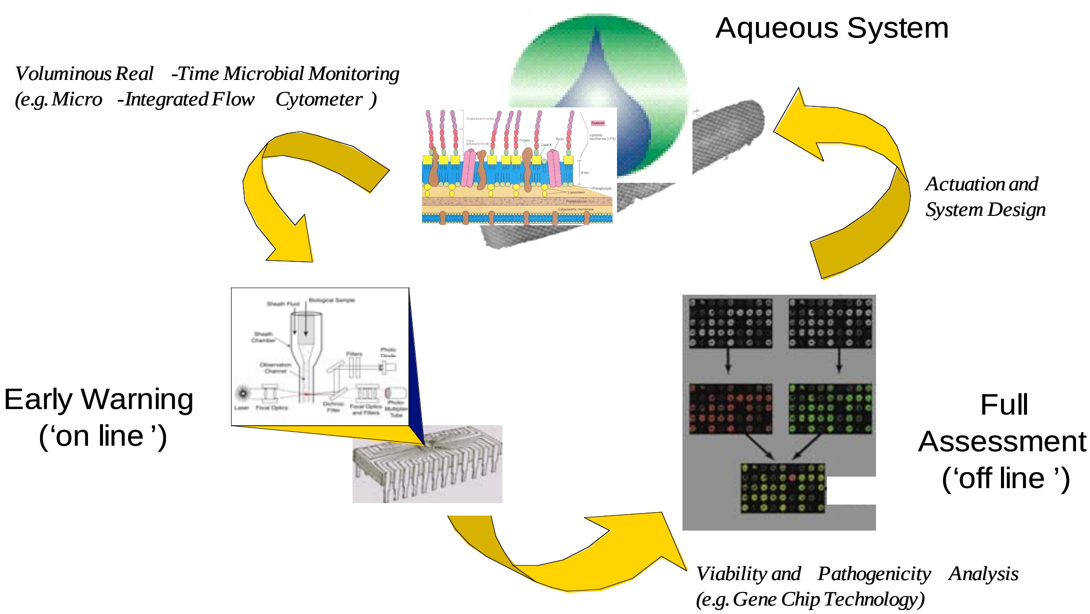

Figure 5. *Microbial Sensing for System Management: Trigger vs. Diagnostics*

4. SPATIAL MODELING AND UNCERTAINTY ANALYSIS

Microbial parameters exhibit significant variability both spatially and temporally. A recently released report put together by the US EPA [WBMSSD-05], for example, notes that microbial parameters at beaches exhibit "some form of systematic spatial variation" that was not adequately accounted for by the depth at which samples were taken. Although this has been recognized for some time, few methods are available to identify and take into account this heterogeneity. In order to effectively propagate the information provided by measurements in time and space, statistical methods need to be able to quantify the degree of spatial variability and use this information to estimate microbial parameters at unsampled locations in a probabilistic framework. The methods should be able to quantify and ultimately reduce the uncertainty associated with the spatial distribution of microbial parameters.

Geostatistical and multiscale statistical modeling tools provide a framework for achieving the goal outlined above, and have begun to be incorporated in a limited number of studies. The field of geostatistics, or the theory of regionalized variables, was introduced by Matheron [M-63, M-71] and is an adaptation of least squares methods to quantities that are correlated in space. Geostatistical interpolation methods were first developed in mining engineering. Much of the early applications were in subsurface environments, describing the spatial distributions of geological structural parameters [JH-78], hydrogeological parameters [K-97] and soil properties [G-97]. More recently, the generality and adaptability of the geostatistical approach has led to a wide range of applications in the earth and environmental sciences, including a limited number of applications to the estimation of surface water quality parameters and contaminant loads [LEP-97, JCC-97, BML-00, AGSEE-03, GS-04, MK-05].

Geostatistical approaches model the spatial and/or temporal distribution of a parameter as a random field, described by a covariance function that captures the degree to which information is lost as one moves away from sampled locations. Kriging-based methods not only provide a Best Linear Unbiased Estimate (BLUE) of the distribution of a parameter of interest, but also characterize the uncertainty associated with that estimate. This makes them ideally suited to analyzing limited environmental data and to designing monitoring network that can effectively capture the spatial and temporal heterogeneity of physical, chemical and biological parameters.

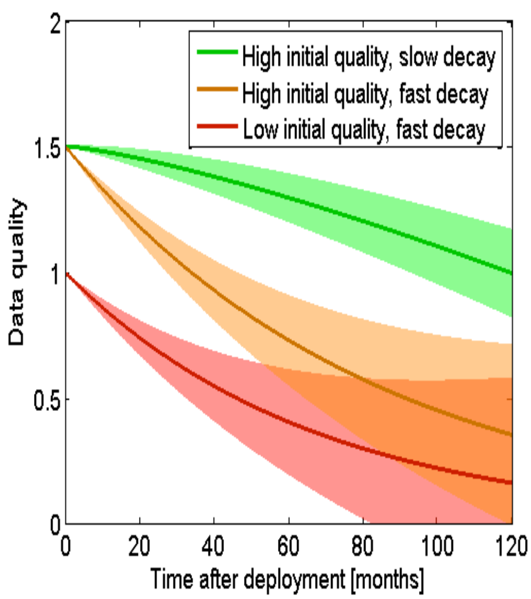

Figure 6*. The objective of sensor networks is to develop a methodology that is able to deliver high initial data quality and a slow decrease in data quality as sensors begin to fail. An additional desirable feature is a relatively low level of uncertainty about a network's ability to continue delivering high data quality over time. The green represents the optimal scenario of a high initial data quality, a slow anticipated decrease in data quality over time, and a relatively low level of uncertainty regarding future performance. The orange and red lines represent sub-optimal design, with either a lower initial data quality level, a faster decrease in quality over time, and/or a higher level of uncertainty with regard to future performance.*

In surface water quality monitoring, basic geostatistical methods have been applied in a few instances. As part of its EMPACT Beaches Project [WBMSSD-05], the EPA constructed variograms (although did not call them by that name) designating the expected variance between pairs of measurements of indicator densities taken at a given separation distance. This work clearly indicated that spatial correlation was present. Posa and Rossi [PR-91] applied a geostatistical approach to estimate dissolved oxygen concentrations in the Mar Piccolo of Tarato, Italy. Little et al. [LEP-97] used ordinary kriging (a geostatistical interpolation approach where the underlying process is assumed to have a constant but unknown mean) to interpolate measurements in Murrells Inlet, South Carolina, including coliform and other data. No cross-correlations between variables were considered. Bellehumeur et al. [BML-00] used geostatistical simulation methods to estimate the spatial distribution of lake acidity (pH) on the Canadian Shield. Gardner et al. [GSL-03] and Gardner and Sullivan [GS-04] used geostatistical methods to analyze stream temperature data. They considered different spatial and temporal scales of variability by implementing a complex nested variogram model to explain the observed distributions. These recent works demonstrate the potential of geostatistical approaches for interpreting spatial patterns of water quality data. However, these studies all focused on analyzing either a single type of data, or considered different data types independently. Novel approaches are needed to jointly assimilate the variety of data that can be collected using a wireless sensor network and that can take into account the time-dependent decay of the data quality as a function of power depletion, accidental loss of a sensor, and other decays. Three data decay functions are illustrated in Figure 6; the goal for an optimized sensing network is to have high data quality and slow decay rates of sensor quality.

A variety of works have also explored the design of monitoring networks based on a criterion of minimizing the uncertainty of the interpolated product [e.g. C-91]. For surface water systems, the possibility of using geostatistics as a monitoring network design tool was described by Jassby et al. [JCC-97] for San Francisco Bay. The design of a wireless sensing network for microbial attributes involves several factors and constraints not considered by existing methods, typically developed for groundwater applications. First, the hydrodynamics of surface water systems make it unlikely that the covariance structure of the parameter fields will be stationary in space. As such, the field cannot be described by a single variogram or covariance structure. Recent and ongoing work at the University of Michigan and Michigan Technological University has focused on developing tools for identifying and quantifying spatially-variable covariance structures that can represent the variability in the degree of physical heterogeneity of a natural system [AM-06]. This work is also developing novel methods for sampling network design based on such heterogeneity. Second, a variety of auxiliary (or trigger) variables can be measured in surface water systems that may provide additional information about microbial parameter distributions. Geostatistical methods can incorporate information on such auxiliary information in a statistically rigorous manner. Third, existing methods optimize monitoring arrays based primarily on the criterion of reducing uncertainty in areas of interest, but without considering constraints and criteria imposed by other elements of the network**. *In the case of a wireless sensor network deployed in a surface water environment, technological issues of network longevity, robustness and communication must be considered in conjunction with scientific monitoring criteria.***

5. WIRELESS SENSOR NETWORKS AND TESTBEDS

Over the past decades, antennas, radio transceivers and processors have been greatly improved in terms of form, size, and power efficiency. Such progress in wireless communication technologies together with marked advances in micro-electromechanical systems (MEMS) has enabled the integration of sensing, actuation, processing, and wireless communication capabilities into tiny sensor devices, which are envisioned to be made increasingly inexpensive, energy-efficient, and reliable. At the same time, the advances in mobile ad hoc networks have enabled the development of *wireless sensor network*, a subject of extensive study within the networking research community in recent years, whereby sensors may be deployed to self-organize into networks that serve a variety of applications.

These applications range from scientific data gathering, environmental monitoring and pollution detection [AC-00, IG-99, BGS-00, BEGH-01], to building smart homes and laboratories [AS-00, E-00, HK-00, PGPMG-00]. Specific examples include the ZebraNet project [LSZM-04], which used GPS-enabled sensors to monitor the zebra migrations in Kenya, the monitoring of nesting habitats of birds and environmental conditions such as temperature and humidity in Maine [MPSCA-02, SPMC-04], which used the Berkeley Motes [MICA2], habitat monitoring [CEHZ-01], health [KKP-99], and the monitoring and detection of car theft [PK-00]. Wireless sensor networks are well suited for these applications due to their rapid and inexpensive deployment (e.g., compared to wired solutions). They can be deployed (e.g., airborne) to areas otherwise inaccessible by land. The low cost and low energy nature of these sensors also makes them easily disposable (if made bio-degradable).

Along with this wide range of emerging applications, there have been extensive studies on building protocols, software, as well as simulation and emulation tools that help bring these applications to reality, for example, different sensor platforms [WINS, MICA2, MBCISSWC-02, SS-02] as well as operating systems [TinyOS, LC-02, LSZM-04]. Applications most relevant to the project outlined here are waterborne sensor networks for the monitoring of marine ecosystems, water quality, and contaminants. Few examples exist, and those that do focus on multivariate ocean monitoring (e.g. ARGO system) for plankton, algal blooms and other ecosystem degradation parameters [EGPS-01, PK-05].

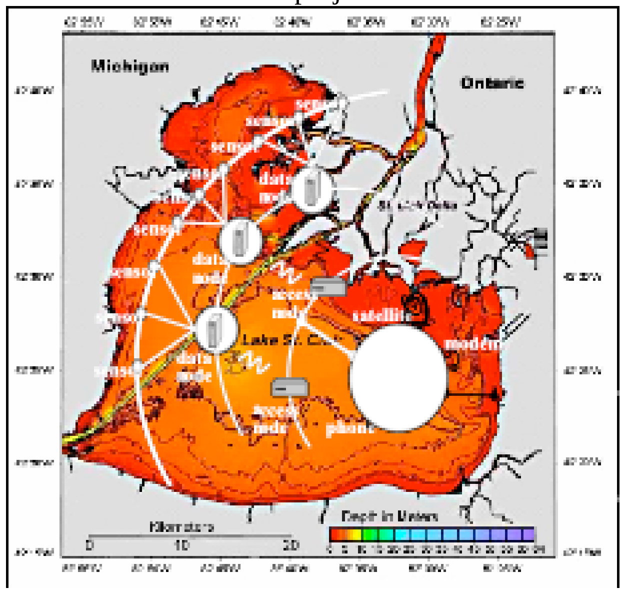

***Figure 7**. Conceptual testbed in Lake St. Clair*

The network optimization approach could be based on a tiered communication structure such as shown in Figure 7. Wireless waterborne sensors are then deployed similar to RECON. The ones closest to shore can form the first tier of the network and potentially serve as relays for other sensors that located further away. Each tier is located further away from shore. The number of tiers formed will explicitly incorporate the area of the sensing region and the transceiver capability of the sensors. Failure functions will be incorporated in the network design to evaluate the impact of sensor degradation and data loss on the spatio-temporal interpretation of microbial data. Generalization of the network design methodology for monitoring programs in other lakes, with emphasis on the Great Lakes basin, will use a scenario-testing approach in the iteration model. Since the data quality metrics impact the design criteria for the network, both in terms of environmental (spatio-temporal) constraints and network robustness (in terms of communication and data heterogeneity), this Step will provide a data-driven design space for wireless network implementation in environmental systems.

Various co-authors have been working to apply these techniques with various plug-in (or off line) sensors to the specifics of an environmental system (scale, available data, and data quality objectives). Real-world constraints will prove invaluable to answer the question: “How does a priori information influence the design of networks co-optimized for environmental and network failure constraints?” The application testbed for this project may capitalize on the Lake St. Clair system, which is part of the Lake Erie basin and is located on the border between southeastern Michigan and southern Ontario. The lake is connected to Lake Huron to the North by the St. Clair River and to Lake Erie to the south by the Detroit River, and exhibits the largest delta system (and thus contributing contaminant sources) of all the Great Lakes. This testbed may be proposed because: (i) a database with parameters collected over eight years is available online (http://www.lakestclairdata.net/); (ii) a wireless data collection system is being implemented at the water treatment plants that can serve as a benchmark for the models, (iii) the availability of a high resolution numerical hydrodynamic model, developed by Prof. Guy Meadows at the University of Michigan, designed to model the transport and dispersion of microbial and other constituents in this basin; and (iv) results generated from network optimization based on a well-characterized system will help with generalizing the methodological approach to the Lake Huron-Erie corridor, currently instrumented by GLERL. Given the monitoring that has taken place in the Lake St. Clair basin, and the impending data from the wireless systems at the water intakes, this would enable us to apply our methods at a 200m resolution within the lake area.

6. DATA COMPILATION AND COMMUNICATION

Ecosystem management at the scale of the Great Lakes will require significant advances in data management and computational capabilities. Existing programs designed to support environmental observatories have emphasized the compilation and federation of existing data from multiple sources. To complement and leverage these efforts, the ERC will implement a new grid-based computational infrastructure with a research focus on the deployment of high performance simulation models. The new *Great Lakes Grid* (GLGrid) will be based on the New York State Grid (NYS Grid), developed by project co-leader Miller with support from NSF. The NYS Grid, which currently includes computational and data systems throughout New York State (including UB and Cornell), will be extended to include UM and other systems involved with Great Lakes Research, and will provide access to government researchers and regulatories (USEPA and NOAA Great Lakes and Ecosystems Research labs), as well as private-sector partners. The administration of GLGrid will be coordinated by UB's Center for Computational Research under Miller's direction.

In addition to deploying a dedicated grid infrastructure, current efforts will develop and evaluate the necessary interfaces and "middleware" needed to deploy and optimize high-performance ERC simulation models. Research needs will emphasize (1) development of secure and high-performance grid technologies that allow for the integration of high-end computers, data, networking, and visualization, as well as sensors, imaging devices, and databases; (2) implementation of grid technologies, dynamic resource classification for fast processing on homogeneous parallel platforms, and the distributed computation for individual computational tasks on heterogeneous platforms; (3) development of technology for building a common core database platform on the grid, the development of distributed search technology utilizing heterogeneous databases, large-scale distributed text searching, and intelligent storage controller development; and (4) portal development to promote access to a wide range of users at all levels, pre-college, college and the general public.

Dr. Finholt and colleagues at UM's Collaboratory for Research on Electronic Work (CREW) have conducted extensive research on how new technologies enable new ways of organizing work, both for providers and users of the technology, as well as in both academic and business environments. Similar to a recent NSF-sponsored evaluation of the Teragrid, the GLGrid evaluation will: (a) provide specific information to GLGrid managers that will increase the likelihood of GLGrid success; and (b) give ERC and NSF leaders and policy makers general data that will assist them in making strategic decisions about future directions for cyberinfrastructure. In particular, GLGrid will be evaluated in terms of progress in meeting user requirements, impact on research practice and outcomes, quality and content of GLGrid education, outreach, and training efforts, and satisfaction among GLGrid partners.

7. CONCLUSIONS AND RESEARCH QUESTIONS

Technology development for environmental sensing infrastructure is governed by addressing technological and societal/market needs. Substantial investment in materials, devices and modeling software for implementation of sensor networks has advanced our technological understanding of the critical elements required to accomplish this objective. Yet, the deployment of sensor nodes (multiple sensors in a spatial network) remains costly and fragile, and the data feedback from these nodes is often too slow to impact decision-making or to forecast hazardous events, either using trigger or diagnostic data. Leveraging the strengths at our Michigan institutions has the potential to position us to capture a cyberinfrastructure node as part of the WATERS program.

To better understand the needs and integration of the various aspects of building an appropriate cyberinfrastructure for the Great lakes system, the breakout session will seek to resolve outstanding questions and approaches in the following areas, not limited to microbial sensing needs :

1. *Data Specification and Collection: What do you want to Forecast About the Ecosystem?*

Wireless network design can only be optimized once data quality criteria are specified. We define data quality as the ability of the sensor network to represent the parameter distribution within the entire system, where this ability is defined based on specific criteria. Depending on the application or water quality management goal, these criteria may include: (i) Minimization of average uncertainty in spatial distribution; (ii) Minimization of single maximum point uncertainty in the system; (iii) Maximized ability to determine whether the parameter is above/below a given threshold at a receptor node. The choice of criteria and its associated requirements for network design include speed, cost, specificity, and sensitivity that will drive achievable data quality metrics and thus the associated network design objectives. For example, the Lake St. Clair pilot collects general water quality data (e.g. oxygen, pH, conductivity) every 15 minutes, chemical data (e.g. halogenated compounds, petroleum hydrocarbons) every few hours, and microbial data on a weekly basis). The RECON system provides general water quality data in similar timeframes.

2. *Modeling and Interpretation: What is the Quality of Information Provided by the Measured Attributes?*

The design of a wireless monitoring network that maximizes a specific measure of data quality must be based on the spatial and temporal covariance structure of microbial and related environmental parameters, available baseline measurements, and information about the hydrodynamics of the system. Spatial statistical methods (geostatistical and multiscale modeling tools) provide an integrated framework for defining the optimal design criteria for minimizing the uncertainty associated with the characterization of microbial parameters in recreational and potable waters. This topic will utilize this framework to (i) discuss modeling tools to quantify the uncertainty in the spatial and temporal distribution of water quality parameters related to microbial attributes, (ii) translate the spatial distributions and their uncertainty into an overall measure of data quality, and (iii) optimize the network to maximize the overall data quality at the time of deployment and in the presence/absence of sensor failure.

3. *Visualization and Communication: Conceptualize the Utility of the System and its Interactions*

This section will address how models inform our physical understanding of the ecosystem. In addition, we will discuss (1) ontology and model integration across the biotic, abiotic, and socio-economic domains, and (2) integration of computationally-, data-, instrument-, and sensor-based grids along with the design, development, and deployment of fundamental cyberinfrastructure middleware and tools to make the use of such systems transparent to the end user. These projects will complement and extend similar ongoing national efforts within the environmental science and engineering communities, which emphasize cyberinfrastructure projects for disciplinary-focused programs (e.g., hydroinformatics and ecoinformatics). The outcome of this topical area is to provide tools for a variety of stakeholders that will enable them to assess and manage the impacts of natural and anthropogenic stressors on regional ecosystem resources.

4. *Sensor Network Deployment: Data- and Model-Driven Infrastructure for Forecasting*

This topic will address the development of sensor network configurations that satisfy (i) low power requirements for network placement, and (ii) data utility. The configurations will include nodes that can be as simple as a temperature sensor, or as complex as a complete environmental sensor system that measures and correlates temperature, pH, dissolved oxygen levels, and specific microbia biomarkers. While some of these sensors can be “off the shelf” components, the geographical extent and diversity of the data collection region dictates that specialized sensor systems be developed. These sensor systems will be designed to simplify addition to an existing sensor network; the interface to the network will be standardized, while the sensing components will have whatever functionality is desired for a particular location. Experiences from currently deployed systems and utility-driven networks (e.g. tiered communication systems) will be discussed.

8. REFERENCES

[A-05] M.E. Aull “Water Quality Indicators in Watershed Subbasins With Multiple Land Uses.” *Thesis, Worcester Polytechnic Institute.* 113 p., 2005.

[A-05] P. Adriaens. “Scaling Contaminant Distributions and Contaminant Processes in Sediments”, *Proc. Assoc. Environ. Health Sci., Amherst, MA*, October 2005.

[AC-00] J. Agre and L. Clare, “An integral architecture for cooperative sensing networks,” *IEEE Computer Magazine*, May 2000.

[AGLW-05] P. Adriaens, C. Gruden, M.-Y. Li, and J. Wolfe. 2005. “Scaling of Microbial Competence for Sediment Remediation.” *3rd Int. Conf. on Remediation of Contaminated Sediments, New Orleans, LA.* January 2005.

[AGSEE-03] P. Adriaens, P. Goovaerts, S. Skerlos, E. Edwards, and T. Egli, "Intelligent infrastructure for sustainable potable water: a roundtable for emerging transnational research and technology development needs," *Biotechnology Advances*, vol. 22, pp. 119-134, 2003.

[AM-05] E. Altman and D. Miorandi, “Coverage and connectivity of ad-hoc network in presence of channel randomness,” in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, April 2005. Miami, FL.

[AM-06] A. Alkhaled and A. M. Michalak, "Spatial Covariance Structure of Modeled Column Integrated CO2 Distributions and its Impact on Representativeness of Satellite Measurements," to be presented at American Geophysical Union Fall Meeting, San Francisco, CA, 2006.

[AS-00] G. D. Abowd and J. P. G. Sterbenz, “Final report on the interagency workshop issues for smart environments,” *IEEE Personal Communications*, October 2000.

[ASEGE-03] P. Adriaens, S. Skerlos, E.A. Edwards, P. Goovaerts, and T. Egli. “Intelligent Infrastructure for Sustainable Potable Water Supplies: A roundtable for emerging transnational research and technology development needs.” *Biotechnol. Adv*. 22, 119-134, 2003.

[ATKGMW-04] P. Adriaens, T. Towey, A. Khijniak, C. Gruden, R McCulloch, and J. Wolfe. “Scaling of Sediment Bioremediation Efficacy Assessment Using Microbial Sensing and Geostatistical Analysis” *SETAC World Congress, Portland, OR.*, 2004

[BAG-04] N. Barabas, P Adriaens, and P. Goovaerts. 2004. “Modified polytopic vector analysis to identify and quantify dioxin dechlorination signatures in sediments. 1. Theory”. *Environ. Sci. Technol.*, 38: 1813-1820, 2004.

[BC-02] M. Bhardwaj and A. P. Chandrakasan, “Bounding the lifetime of sensor networks via optimal role assignments,” in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, June 2002, New York, NY.

[BEGH-01] N. Bulusu, D. Estrin, L. Girod, and J. Heidemann, “Scalable coordination for wireless sensor networks: self-configuring localization systems,” in *Proc. International Symposium on Communication Theory and Applications (ISCTA)*, July 2001.

[BGA-01] N. Barabas, P. Goovaerts, and P. Adriaens. “Geostatistical Assessment and Validation of Uncertainty for Three-Dimensional Dioxin Data from Sediments in an Estuarine River”, *Environ. Sci. Technol*. 35: 3294-3301, 2001.

[BGA-04] N. Barabas, P. Goovaerts, and P. Adriaens. “Modified Polytopic Vector Analysis to Identify and Quantify Dioxin Dechlorination Signatures in Sediments. 2. Application to the Passaic River Superfund Site”, *Environ. Sci. Technol*., 38: 1821-1827, 2004

[BGC-01] M. Bhardwaj, T. Garnett, and A. P. Chandrakasan, “Upper bounds on the lifetime of sensor networks,” in *Proc. IEEE International Conference on Communications (ICC)*, June 2001, Helsinki, Finland.

[BGS-00] P. Bonnet, J. Gehrke, and P. Seshadri, “Querying the physical world,” *IEEE Personal Communications*, October 2000.

[BHBRLMA-05] J. Brookes, M.R. Hipsey, M.D. Burch, R.H. Regel, L.G. Linden, C.M. Ferguson, and J.P. Antenucci “Relative Value of Surrogate Indicators for Detecting Pathogens in Lakes and Reservoirs”. *Environ. Sci. Technol*. ASAP, 2005.

[BL-03] G.M. Brion, and S. Lingireddy. “Artifical Neural Network Modelling: A Summary of Successful Applications Relative to Microbial Water Quality” *Wat. Sci. Technol*. 47: 235-240, 2003.

[BML-00] C. Bellehumeur, D. Marcotte, and P. Legendre, "Estimation of regionalized phenomena by geostatistical methods: lake acidity on the Canadian Shield," *Environmental Geology*, vol. 39, pp. 211-220, 2000.

[BROV-02] L. Bonadonna, R. Briancesco, M. Ottaviani and E. Verschetti. “ Occurrence of Cryptosporidium Oocysts in Sewage Effluents and Correlation with Microbial, Chemical, and Physical Water Variables”. *Environ. Monit. Assess*. 75: 241-252, 2002.

[C-91] N. A. C. Cressie, *Statistics for Spatial Data*. New York: Wiley-Interscience, 1991.

[CBV-04] R. Cristescu, B. Beferull-Lozano, and M. Vetterli, “On network correlated data gathering,” in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, March 2004, Hong Kong.

[CEHZ-01] A. Cerpa, J. Elson, M. Hamilton, and J. Zhao, “Habitat monitoring: application driver for wireless communications technology,” in *Proc. ACM SIGCOMM*, April 2001, Costa Rica.

[CJBM-01] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris, “SPAN: an energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks,” in *Proc. ACM International Conference on Mobile Computing and Networking (MOBICOM)*, 2001.

[CL-04] N. Chang and M. Liu, “Revisiting TTL-based controlled flooding search: optimality and randomization,” in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, pp. 85-99, September 2004, Philadelphia, PA.

[CL-05] N. Chang and M. Liu, “Optimal controlled flooding search in large wireless networks,” in *Proc. International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, pp. 229-237, March 2005, Trentino, Italy.

[CL-06] N. Chang and M. Liu, “Controlled flooding search with delay constraints,” *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, April 2006, Barcelona, Spain.

[CPRL-01] F.C. Curriero, J.A. Patz, J.B. Rose, and B. Lele. “The Association Between Extreme Precipitation and Waterborne Disease Outbreaks in the United States, 1948-1994.” *Am. J. Pub. Health* 91: 1194-1199.

[CT-00] J. Chang and L. Tassiulas, “Energy Conserving Routing in Wireless Ad-hoc Networks,” in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, 2002.

[DJ-98] C.V. Deutsh. and Journel, A.G. GSLIB: Geostatistical Software Library and User’s Guide. *Oxford University Press, New York*, 1998.

[DL-02-1] E. J. Duarte-Melo and M. Liu, “Energy efficiency in many-to-one communications in wireless networks”, in *Proc. IEEE Midwest Symposium on Circuits and Systems (MWSCAS)*, August 2002, Tulsa, OK.

[DL-02-2] E. J. Duarte-Melo and M. Liu, “Analysis of energy consumption and lifetime of heterogeneous wireless sensor networks,” in *Proc. IEEE Global Communications Conference (GLOBECOM)*, vol. 1, pp. 21-25, November 2002, Taipei, Taiwan.

[DL-03] E. J. Duarte-Melo and M. Liu, “Data-Gathering Wireless Sensor Networks: Organization and Capacity,” *Elsevier Journal of Computer Networks (COMNET)*, Special Issue on Wireless Sensor Networks, vol. 43, issue 4, pp. 519-537, November 2003.

[DLM-03] E. J. Duarte-Melo, M. Liu and A. Misra, “A computational approach to the joint design of distributed data compression and data dissemination in a field-gathering wireless sensor network,” in *Proc. Annual Allerton Conference on Communication, Control and Computation (Allerton)*, October 2003, Allerton, IL.

[DLM-04] E. J. Duarte-Melo, M. Liu and A. Misra, “A modeling framework for computing lifetime and information capacity in wireless sensor networks,” in *Proc. International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)*, March 2004, Cambridge, UK.

[DLM-05] E. J. Duarte-Melo, M. Liu and A. Misra, “An Efficient and Robust Computational Framework for Studying Lifetime and Information Capacity in Sensor Networks,” to appear in *ACM Mobile Networks and Applications (MONET)*, Special Issue on Energy Constraints and Lifetime Performance in Wireless Sensor Networks, 2005.

[DM-93] J.B. Dutka and J. Marsalek. „Urban impacts on river shoreline microbiologica pollution“ *J. Great Lakes Res*. 19: 665-674, 1993.

[E-00] I. A. Essa, “Ubiquitous sensing for smart and aware environment,” *IEEE Personal* Communications, October 2000.

[ECDVMM-97] E. Edwards, M. Coyne, T. Daniel, P. Vendrell, J. Murdoch, and P. Moore. “Indicator bacteria concentrations of two Northwest Arkansas streams in relation to flow and season.” Am Soc. Agric. Eng. 40: 103-109, 1997.

[EE-00] J. Elson and D. Estrin, “An address-free architecture for dynamic sensor networks,” Technical Report 00-724, Computer Science Department, USC, January 2000.

[EGHK-99] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar, “Next century challenges: scalable coordination in sensor networks,” in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, 1999, Seattle, WA.

[EGHS-99] D. Estrin, R. Govindan, J. Heidemann, and S. Kumar. “Next Century Challenges: Scalable Coordination in Sensor Networks.” *Mobile Computing and Networking*, pp. 263-270, 1999.

[EGPS-01] D. Estrin, L Girod, G Pottie, M Srivastava. “Instrumenting the world with wireless sensor networks” *ICASSP IEEE Int. Conf. Acoust. Speach Signal Process. Proceed*.. 2001.

[EM-99] A.H. El-Shaarawi and J. Marsalek. “Guidelines for Indicator Bacteria in Waters: Uncertainties in Applications.” *Environmetrics* 10: 521-529, 1999.

[EPA-03] USEPA “Bacterial Water Quality Standards for Recreational Waters. Status Report.” *EPA-823-R-03-008. U.S. Environmental Protection Agency, Office of Water, Washington DC.*, 2003.

[EPA-05] USEPA. “The EMPACT Beaches Project: Results from a Study on Microbiological Monitoring in Recreational Waters”. *EPA 600/R-04/023. Office of Research and Development, Cincinnati OH*. 2005.

[EPA-86] USEPA. “Ambient Water Quality Criteria for Bacteria” *U.S. Environmental Protection Agency, Office of Water, Criteria and Standards Division, Washington, DC*, 1996

[EPA-97] USEPA. “Compendium of Tools for Watershed Assessment and TMDL Development” *U.S. Environmental Protection Agency, Office of Wetlands, Oceans, and Watersheds, Washington, DC.* 1997.

[EPA-98] USEPA. “Bacteria Water Quality Standard Status Report”. *U.S. Environmental Protection Agency, Office of Water, Washington, DC.*, 1998.

[EPA-99] USEPA. “Review of Potential Modeling Tools and Approaches to Support the BEACH Program”. EPA *823-R-99-002. Office of Science and Technology, Washington DC*., 1999.

[ERD-97] K.W. Easter, M.W. Rosegrant, and A. Dinar. Markets for Water: Potential and Performance. Kluwer Publishers, Norwell, MA. 1997.

[EWW-05] T. A. Erickson, M. W. Williams, and A. Winstral, "Persistence of topographic controls on the spatial distribution of snow in rugged mountain terrain, Colorado, United States," *Water Resources Research*, vol. 41, 2005.

[F-74] R. Feachem. “An improved role for faecal coliform to faecal streptococci ratios in the differentiation between human and non-human pollution sources.” *Wat. Res*. 9: 689-690, 1974.

[FMH-00] D.S. Francy, D.N. Myers, and D.R. Helsel. “Microbiological Monitoring for the USGS-National Water Quality Assessment Program”. *Water Resources Investigations Report 00-4018, 31 pp.*, 2000.

[G-97] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*. New York: Oxford University Press, 1997.

[GBHH-94] A. Gutjahr, B. Bullard, S. Hatch, and L. Hughson, "Joint Conditional Simulations and the Spectral Approach for Flow Modeling," *Stochastic Hydrology And Hydraulics*, pp. 79-108, 1994.

[GK-00] P. Gupta and P. R. Kumar, “The capacity of wireless networks,” *IEEE Trans. on Information Theory*, vol. 46, no. 2, March 2000.

[GMHM-05] S. M. Gourdji, K. Mueller, C. Humphriss, and A. M. Michalak, "Fine Spatial Resolution Global CO2 Flux Estimates for 1997 to 2001 Obtained Using Remote-Sensing Derived Environmental Data Within a Geostatistical Inverse Model," presented at American Geophysical Union Fall Meeting, San Francisco, CA, 2005.

[GS-04] B. Gardner and P. J. Sullivan, "Spatial and temporal stream temperature prediction: Modeling nonstationary temporal covariance structures," *Water Resources Research*, vol. 40, 2004.

[GSA-04] C. Gruden, Skerlos, S.J., and P. Adriaens. “Flow Cytometry for Microbial Sensing in Environmental Sustainability Applications: Current Status and Future Prospects”. *FEMS Microbiol. Ecol*. 49: 37-49, 2004.

[GSL-03] B. Gardner, P. J. Sullivan, and A. J. Lembo, "Predicting stream temperatures: geostatistical model comparison using alternative distance metrics," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 60, pp. 344-351, 2003.

[GT-01] M. Grossglauser and D. Tse, “Mobility increases the capacity of ad-hoc wireless networks, in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, 2001.

[HCB-00] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-efficient communication protocol for wireless microsensor networks,” in *Proc. Hawaii International Conference on System Sciences*, January 2000.

[HEHHSLAC-93] M.A. House, J.B. Ellis, E.E.Herricks, T. Hvitved-Jacobson, J. Seager, L. Lijklema, H. Aalderinck, and I.T. Clifforde. “Urban drainage impacts on receiving water quality” *Wat. Sci. Technol*. 27: 117-158, 1993.

[HK-00] C. Herring and S. Kaplan, “Component-based software systems for smart environments,” *IEEE Personal Communications*, October 2000.

[HL-02] C. Hsin and M. Liu, “A distributed monitoring mechanism for wireless sensor networks,” in *Proc. ACM Workshop on Wireless Security (WiSe)*, pp. 57-66, September 2002, Atlanta, GA.

[HL-04] C. Hsin and M. Liu, “Network coverage using low duty-cycled sensors: random and coordinated algorithms,” in *Proc. International Workshop on Information Processing in Sensor Networks (IPSN)*, vol. 1, pp. 433-442, April 2004, Berkeley, CA.

[HL-05] C. Hsin and M. Liu, “Partial clustering: maintaining connectivity in a low duty-cycled dense wireless sensor network,” *IEEE Workshop on Algorithms for Ad Hoc and Sensor Networks (WMAN)*, April 2005, Denver, CO.

[HL-05-1] C. Hsin and M. Liu, “A Two-Phase Self-Monitoring Mechanism for Wireless Sensor Networks,” to appear in *Elsevier Journal of Computer Communications*, Special Issue on Sensor Networks, 2005.

[HL-05-2] C. Hsin and M. Liu, “Randomly Duty-cycled Wireless Sensors Networks: the Dynamics of Coverage,” under revision for *IEEE Transactions on Wireless Communications*, 2005.

[IAAW-99] D. Ivnitski, Abdel-Hamid, I., Atanasov, P., and E. Wilkins, ”Biosensors for Detection of Pathogenic Bacteria” *Biosens. Bioelectron*. 14, 599-624, 1999.

[IG-99] T. Imielinski and S. Goel, “DataSpace: querying and monitoring deeply networked collections in physical space,” in *Proc. ACM International Workshop on Data Engineering for Wireless and Mobile Access (MobiDE)*, 1999, Seattle, WA.

[IGE-00] C. Intanagonwiwat, R. Govindan, and D. Estrin, “Directed diffusion: a scalable and robust communication paradigm for sensor networks,” in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, 2000.

[JCC-97] A. D. Jassby, B. E. Cole, and J. E. Cloern, "The design of sampling transects for characterizing water quality in estuaries," *Estuarine Coastal and Shelf Science*, vol. 45, pp. 285-302, 1997.

[JGRPCW-05] Y. Jeong, S.B. Grant, S. Ritter, A. Pednekar, L. Candelaria, and C. Winant. “Identifying Pollutant Sources in Tidally Mixed Systems: Case Study of Fecal Indicator Bacteria from Marinas in Newport Bay, Southern California.” *Environ. Sci Technol. ASAP*, 2005.

[JH-78] A. G. Journel and C. J. Huijbregts, *Mining Geostatistics*. London: Academic Press, 1978.

[K-97] P. K. Kitanidis, *Introduction to Geostatistics Applications in Hydrogeology*. New York: Cambridge University Press, 1997.

[KHL-05] D. Kim, C. Hsin and M. Liu, “Asymptotic connectivity of low duty-cycled wireless sensor networks,” in *Proc. IEEE Military Communication Conference (MILCOM)*, October 2005, Atlantic City, NJ.

[KKP-99] J. M. Kahn, R. H. Katz, and K. S. J. Pister, “Next century challenges: mobile networking for smart dust,” in *Proc. ACM International Conference on Mobile Computing and Networking (MobiCom)*, 1999, Seattle, WA.

[L-03] M. Liu, “Sequential use of wireless sensors for target estimate and tracking,” in *Proc. IEEE Military Communication Conference (MILCOM)*, vol. 1, pp. 664-669, October 2003, Boston, MA.

[L-04] T. Little, T. Value Creation and Capture: A Model of the Software Development Process," IEEE Software, vol. 21, no. 3, pp. 48-53, May/Jun, 2004.

[LBA-05] M.-Y. Li, N. Barabas, and P. Adriaens. “M-Scale model for multi-scale estimation of spatially-distributed datasets: 2. Application to Passaic River dioxin data” *Environ. Sci. Technol*. In review, 2005.

[LBA-05] M.Y. Li, N. Barabas, and P. Adriaens. “M-Scale model for multi-scale estimation of dioxin data” *Organohalogen Compounds* 2005.

[LC-02] P. Levis and D. Culler, “Mate: a tiny virtual machine for sensor networks,” in *Proc. of ASPLOS*, October 2002.

[LEP-97] L. S. Little, D. Edwards, and D. E. Porter, "Kriging in estuaries: As the crow flies, or as the fish swims?," *Journal of Experimental Marine Biology and Ecology*, vol. 213, pp. 1-11, 1997.

[LMMCCGCGSSJ-03] F.X. Lucena, Mendez, A. Moron, E. Calderon, C. Campos, A. Guerrero, M. Cardenas, C. Gantzer, L. Shwartzbrood, S. Skraber, and J. Jofre. "Occurrence and Densities of Bacteriophages Proposed as Indicators and Bacterial Indicators in River Waters from Europe and South America" *J. Appl. Microbiol*. 94: 808-815, 2003.

[LSZM-04] T. Liu, C. Sadler, P. Zhang, and M. Martonosi, "Implementing software on resource-constrained mobile sensors: experiences with IMPALA", in *Proc. of MOBISYS*, June 2004.

[M-63] G. Matheron, "Principles of geostatistics," *Econ. Geol.*, vol. 58, pp. 1246-1266, 1963.

[M-71] G. Matheron, "The theory of regionalized variables and its applications," Paris School of Mines publication 1971.

[MBCISSWC-02] R. Min, M. Bhardwaj, S. Cho, N. Ickes, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan, "Energy-centric enabling technologies for wireless sensor networks," *IEEE Wireless Communications*, vol. 9, pp. 28-39, 2002.

[MBT-04] A. M. Michalak, L. Bruhwiler, and P. P. Tans, "A geostatistical approach to surface flux estimation of atmospheric trace gases," *Journal of Geophysical Research-Atmospheres*, vol. 109, 2004.

[MDLN-03] D. Marco, E. J. Duarte-Melo, M. Liu and D. L. Neuhoff, "On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data," in *Proc. International Workshop on Information Processing in Sensor Networks (IPSN)*, vol. 1, pp. 1-16, April 2003, Palo Alto, CA.

[MDMT-96] J. Marsalek, B.J. Dutka, A.J. McCorquodale, aand I.K. Tsanis. "Microbiological pollution in the Canadian upper Great Lakes connecting channels. Wat. Sci. Technol. 33: 349-356, 1996.

[MGKHM-05] K. Mueller, S. Gourdji, K. Schaefer, C. Humphriss, and A. M. Michalak, "Using Remote Sensing Data to Help Constrain Fluxes of CO2 in a Geostatistical Inverse Modeling Framework," presented at American Geophysical Union Fall Meeting, San Francisco, CA, 2005.

[MICA2] Crossbow Technology Inc: Mica2 sensor node platform, http://www.xbow.com/Products/Wireless_Sensor_Networks.htm.

[MK-02] A. M. Michalak and P. K. Kitanidis, "Application of Bayesian Inference Methods to Inverse Modeling for Contaminant Source Identification at Gloucester Landfill, Canada," in *Computational Methods in Water Resources XIV, Volume 2*, S. M. Hassanizadeh, R. J. Schotting, W. G. Gray, and G. F. Pinders, Eds. Amsterdam, The Netherlands: Elsevier, 2002, pp. 1259-1266.

[MK-03] A. M. Michalak and P. K. Kitanidis, "A method for enforcing parameter nonnegativity in Bayesian inverse problems with an application to contaminant source identification," *Water Resources Research*, vol. 39, 2003.

[MK-04-01] A. M. Michalak and P. K. Kitanidis, "Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling," *Water Resources Research*, vol. 40, 2004.

[MK-04-02] A. M. Michalak and P. K. Kitanidis, "Application of geostatistical inverse modeling to contaminant source identification at Dover AFB, Delaware," *Journal of Hydraulic Research*, vol. 42, pp. 9-18, 2004.

[MK-05] A. M. Michalak and P. K. Kitanidis, "A method for the interpolation of nonnegative functions with an application to contaminant load estimation," *Stochastic Environmental Research and Risk Assessment*, vol. 19, pp. 8-23, 2005.

[MMGHBST-05] A. M. Michalak, K. Mueller, S. Gourdji, C. Humphriss, L. Bruhwiler, K. Schaefer, and P. P. Tans, "Application of Geostatistical Kalman Smoother to the Estimation of Monthly Gridscale Fluxes of Carbon Dioxide," presented at Seventh International Carbon Dioxide Conference, Boulder, CO, 2005.

[MMWLL-94] M. Munawar, Munawar, IF, Weisse, T, Leppard, GG and M. Legner "The significance and future potential of using microbes for assessing ecosystem health: The Great Lakes example." *J. Aquat. Ecosyst. Health*. Vol. 3, no. 4, pp. 295-310, 1994.

[MPSCA-02] A. Mainwaring, J. Polastre, R. Szewczyk, D. Culler, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *Proc. ACM International Workshop on Wireless Sensor Networks and Applications*, September 2002.

[MS-06] A.M. Michalak and S. Shlomi "A geostatistical data assimilation approach for estimating groundwater plume distributions from multiple monitoring events," invited paper submitted for review to AGU monograph on Data Integration in Subsurface Hydrology, July 2006.

[NLB-02] T.S. Neelakantan, Lingireddy, and G.M. Brion. "Relative Performance of Different ANN Training Algorithms in Predicting Protozoa Concentration in Surface Waters" *ASCE J. Environ. Engng*. 128: 533-542, 2002.

[NRC-04] National Research Council. "Indicators for Waterborne Pathogens". *National Academies Press*. 316 p., 2004.

[NRC-97] National Research Council. "Valuing Ground Water: Economic Concepts and Approaches". NRC Press, Washington, DC. 204 p. 1997.

[NRC-98] National Research Council. "Issues in Potable Reuse. The Viability of Augmenting Drinking Water Supplies With Reclaimed Water". *NRC Press, Washington, DC.*, 1998.

[PGPMG-00] E. M. Petriu, N. D. Georganas, D. C. Petriu, D. Makrakis, and V. Z. Groza, "Sensor-based information appliances," *IEEE Instrumentation and Measurement Magazine*, December 2000.

[PK-00] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors, *Communications of the ACM*, vol. 43, no. 5, 2000.

[PK-05] G. Pottie and W. Kaiser. "Principles of Embedded Networked Systems Design" *Cambridge University Press, Cambridge UK*., 2005.

[PR-91] D. Posa and M. E. Rossi, "Geostatistical Modeling of Dissolved-Oxygen Distribution in Estuarine Systems," *Environmental Science & Technology*, vol. 25, pp. 474-481, 1991.

[RDG-88] J.B. Rose, H. Darbin, and C.P. Gerba. "Correlations of the protozoa Cryptosporidium and Giardia with Water Quality Variables in a Watershed" *Wat. Sci. Technol*. 20: 271-276, 1998.

[RGGLSY-99] J.B. Rose, R.M. Atlas, C.P. Gerba, M.J.R. Gilchrist, M.W. Le Chevalier, M.D. Sobsey, and M.V. Yates. "Microbial Pollutants in our Nation's Water-

Environmental and Public Health Issues." *American Society for Microbiology, Washington, DC.*, 1999.

[S-00] R.N. Stavins "Market-based Environmental Policies," *Public Policies for Environmental Protection 2nd Edition*. Resources For the Future: Washington, DC. 2000

[SBWMAMRTRG-03] L. Sacks, M. Britton, I. Wokoma, M. Marbini, T. Adebutu, I. Marshall, C. Roadknight, J. Tateson, D. Robinson, and A. Gonzalez-Velasquez. "The development of a robust, autonomous sensor network platform for environmental monitoring", *Sensors and Applications XXII, Limerick, Ireleand*, September 2003.

[SHS-03] Y. Shi, Y. T. Hou and H. D. Sherali, "On Lexicographic Max-Min Node Lifetime Problem for Energy-Constrained Wireless Sensor Networks", Technical Report, Bradley Department of ECE, Virginia Tech, September 2003.

[SK-00] L. Subramanian and R. H. Katz, "An architecture for building self-configurable systems," in *IEEE/ACM Workshop on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2000.

[SM-01] B. Swift, and J. Mazurek. "Getting More for Four: Principles for Comprehensive Emissions Trading," *Policy Report*. Progressive Policy Institute. 2001.

[SM-05] S. Shlomi and A. M. Michalak, "A Geostatistical Framework for Incorporating Transport Information in Estimating the Distribution of a Groundwater Contaminant Plume," presented at American Geophysical Union Fall Meeting, San Francisco, CA, 2005.

[SM-06] S. Shlomi and A.M. Michalak "A geostatistical framework for incorporating transport information in estimating the distribution of a groundwater contaminant plume," submitted for review in Water Resources Research, April 2006.

[SOPHME-04] R. Szewczyk, E. Osterweil, J. Polastre, M Hamilton, A Mainwaring, D. Estrin. "Habitat Monitoring with Sensor Networks" *Communications of the ACM*, Vol. 47, No. 6, pp. 34-40, 2002.

[SPMC-04] R Szewczyk, J. Polastre, A. Mainwaring, and D. Culler, "Lessons from a sensor network expedition," in *Proc. First European Workshop on Wireless Sensor Networks*, January 2004.

[SS-02] A. Savvides and M. Srivastava, "A distributed computation platform for wireless embedded sensing," in *Proc. International Conference on Computer Design*, 2002.

[SSR-02] J.M. Simpson, Santo Domingo, J.W. and Reasoner, D.J. "Microbial Source Tracking: State of the Science" *Environ. Sci. Tech*nol. 36(24): 5279-5288, 2002.

[SWAMSB-05] L.L. Shum, I. Wokoma, T. Adebutu, A.D. Marbini, L. Sacks, and M. Britton. "Distributed algorithm implementation and interaction in wireless sensor networks" *Proceed. European Workshop on Wireless Sensing Networks, Istanbul, Turkey*, 2005.

[T-97] M. Thobanl. Formal Water Markets: Why, When, and How to Introduce Tradable Water Rights. World Bank Research Observer 12, 161-179. 1997

[TG-02] D. Tian and N. D. Georganas, "A coverage-preserving node scheduling scheme for large wireless sensor networks," in *Proc. First ACM International Workshop on Wireless Sensor Networks and Applications (WSNA)*, 2002.

[TinyOS] J. Hill, R. Szewczyk, et al., “System architecture directions for networked sensors,” in *Proc. of ASPLOS*, April 200.

[W-03] H. Wackernagel, *Multivariate Geostatistics*. Berlin: Springer, 2003.

[WBMSSD-05] L. J. Wymer, K. P. Brenner, J. W. Martinson, W. R. Stutts, S. A. Schaub, and A. P. Dufour, "The EMPACT Beaches Project: Results from a Study on Microbiological Monitoring in Recreational Waters," U.S. Environmental Protection Agency, Cincinnati, OH August, 2005 2005.

[WINS] Rockwell Science Center, “Wireless integrated network sensors (WINS),” http://wins.rsc.rockwell.com.

[WSSM-05] I. Wokoma, L.L. Shum, L. Sacks, and I. Marshall. “A Biologically-Inspired Clustering Algorithm Dependent on Spatial Data in Sensor Networks” *Proceed. European Workshop on Wireless Sensing Networks, Istanbul, Turkey,* pp 1-10, 2005.

[XK-03] F. Xue and P. R. Kumar, “The number of neighbors needed for connectivity of wireless networks,” *Wireless Networks*, 2003.

[YHE-02] W. Ye, J. Heidemann, and D. Estrin, “An energy-efficient MAC protocol for wireless sensor networks,” in *Proc. IEEE Annual Conference on Computer Communications (INFOCOM)*, June 2002, New York, NY.