SAMSI Spring 2008 Program on Environmental Sensor Networks

Final Report, April 1, 2009

Paul G. Flikkema
Chair, Program Leaders

1 Introduction

The core purpose of the SAMSI Program on Environmental Sensor Networks is to identify research challenges and opportunities in the use of wireless environmental sensor networks to address critical contemporary problems. They include understanding the effects of global climate change, human activity, and invasive species on ecosystem function, and drive the our need to understand the dynamics of diverse environmental phenomena and their causes. This problem domain offers unique interdisciplinary research challenges. First, the labor cost of deploying and maintaining these networks is very high, which is limiting adoption of the technology. Secondly, uncertainty is dominant, with noise, numerous failure modes, and over/under-sampling issues driven by conflicts between the needs of network connectivity and spatial design for processes of interest. These problems are compounded by inherent issues of dimensionality and scale: datasets for the biological and physical problems of interest consist of sampled multivariate spatio-temporal process with natural scales ranging from minutes to decades and meters to hundreds or thousands of kilometers.

2 Program Organization

A remarkable characteristic of this program is the diversity of disciplines represented in the participants of the Opening Workshop and both Working Groups. Researchers in ecology, computer science, mathematics, electrical and computer engineering, and environmental engineering are working with statisticians specializing in, among other fields, experiment design, sampling techniques, linear models, spatial statistics and hierarchical Bayesian methods.

The program was led by Paul Flikkema (Northern Arizona University) who was in residence at SAMSI during January - May 2008. Two *Working Groups* were formed, whose principal functions were to identify, organize, and nurture collaborative research initiatives. The majority of participants were from outside the Triangle area.

The **Sensor Networks Datasets** working group led by Paul Flikkema (Northern Arizona University) included: Ankit Agarwal (University of Kansas), David Bell (Duke University), Michela Cameletti (SAMSI/Bergamo University), Zoe Cardon (Ecosystems Center, Marine Biological Laboratory), Jim Clark (Duke University), Alan Gelfand (Duke University), Scott Holan (University of Missouri), Rosaria Ignaccolo (SAMSI/Universita' di Torino), Natallia Katenka (University of Michigan), Yongku Kim (SAMSI/Duke), Ernst Linder (University of New Hampshire), Kristian Lum (Duke University), John McGee (UNC Chapel Hill, Renaissance Computing Institute), Yajun Mei (Georgia Institute of Technology), George Michailidis (University of Michigan), Long Nguyen (SAMSI/Duke), Michael Porter (SAMSI/NCSU), Ilka Reis (National Institute for Space Research, Brazil), Karl Rohe (University of California at Berkeley), Sande Satoskar (RENCI), Lance Waller (Emory University), Kim Weems (NCSU), and Bin Yu (UC Berkeley).

The **Sensor Design** Working Group, led by James S. Clark (Duke University) and Jun Yang (Duke University), included: Ankit Agarwal (University of Kansas), David Bell (Duke University), Michael Breen (EPA), Michela Cameletti (SAMSI/Bergamo University), Zoe Cardon (Ecosystems Center Marine Biological Laboratory), Jorge Cortes (UC San Diego), Jessica Croft (University of Utah), Todd Dawson (UC Berkeley), Carla Ellis (Duke University), Marco Ferreira (University of Missouri), Paul Flikkema (Northern Arizona University), Jeff Frolik (University of Vermont), Alan Gelfand (Duke University), Joe Fred Gonzalez, Jr. (Center for Disease Control), Scott Holan (University of Missouri), Sheryl Howard (Nothern Arizona University), Rosaria Ignaccolo (SAMSI/Universita' di Torino), Chris Jones (University of North Carolina) Yongku Kim (SAMSI/Duke University), Hamid Krim (North Carolina State University), Soumen Lahiri (Texas A&M University), David Leslie (Bristol University), Kristian Lum (Duke University), Yajun Mei (Georga Institute of Technology), George Michailidis (University of Michigan), Long Nguyen (SAMSI/Duke University), Neal Patwari (University of Utah), Michael Porter (North Carolina State University), Ilka Reis (National Institute for Space Research, Brazil), Christine Shoemaker (Cornell University), Bin Yu (UC Berkeley), Yi Zhang (Duke University), and Zhengyuan Zhu (University of North Carolina).

3 Achieving Diversity

The program has had strong participation by female faculty, post-doctoral researchers, and students. Zoe Cardon (Marine Biological Laboratory) and Deborah Estrin (UCLA) serve on the Program Leaders Committee. Estrin, Jennifer Hoeting (Colorado State University), and Kiona Ogle (University of Wyoming) contributed invited presentations at the Opening Workshop. Carla Ellis (Duke University) organized the Fall 2007 SAMSI graduate course on Environmental Sensor Networks.

Participating faculty and researchers include Michela Camelleti, Sheryl Howard, Rosaria Ignaccolo, Cari Kaufman, Christine Shoemaker, Kimberly Weems, and G. Beate Zimmer.

Zoe Cardon (Ecosystems Center, Marine Biological Laboratory) was was a leader of the Sensor Networks Datasets Working Group, and provided a critical dataset that will be used in papers now in preparation as well as crucial support in both the interpretation of metadata and development of models.

Participant **Rosaria Ignaccolo**, a SAMSI New Researcher and an Assistant Professor in the Department of Statistics and Applied Mathematics at the Universita' degli Studi di Torino, was an active participant in the Sensor Data working group, and was a key contributor to cleaning and exploratory data analysis of the group's datasets. She has been interested in exploring functional data analysis methods for ecological data and also presented a talk "Functional Analysis and Clustering with Spline Libraries" at the Transition Workshop.

Participant **Sheryl Howard**, an assistant professor in electrical engineering at Northern Arizona University, attended both the Opening and Transition Workshops, and presented an invited talk at the Transition Workshop entitled "Coded Compressive Estimation in Environmental Sensor Networks." Her work led to an NSF grant award for her proposal "BRIGE: Energy-Efficient Communication with Combined Decoding/ Inference". It has also resulted in a Science Foundation Arizona graduate fellowship award for her student Rui Chen, and she is now supporting two undergraduate students (Forrest Schwynn and Hristo Taralov) and one female graduate student (Fauzia Ahmed). She also presented a talk "Combined Source-Channel Decoding and Transmission Censoring for Power Reduction in a Wireless Sensor Network" at the 2008 International Analog Decoding Workshop (Logan, UT, July 12, 2008). Her student Rui Chen collaborated with Flikkema's student Saiyi Wang (also awarded a SFAz graduate fellowship) on a poster "Energy Efficiency in Environmental Sensing Networks: Cross-Layer Approaches for Transmission Censoring" at the Science Foundation Arizona Graduate Research Fellows Grand Challenge Summit, March 27-29, 2009.

Graduate students include Christina Bentrup (Northern Arizona University), Jessica Croft (University of Utah) and Natallia Katenka (University of Michigan), Kristian Lum (Duke University), and Ilka Reis (National Institute for Space Research, Brazil).

Participant **Kristian Lum** is a Ph.D. student in statistics at Duke University. She participated in the Fall 2008 Sensor Networks for Environmental Modeling Course as well as the Opening and Transition workshops. For her preliminary exam in April 2008, she studied how inference is degraded with decreased transmission rates for various models of the sensed data and inference schemes. More recently, she has been analyzing transmission suppression schemes based on approximating the dynamics of the sensed data with linear temporal models, as well as stochastic differential equation models using Ornstein-Uhlenbeck processes.

4 Research Progress

Both Working Groups conducted weekly distributed meetings throughout the program period. Meeting schedule information, notes, presentation slides, reading lists, and participant directories are all available on-line on the SAMSI website. Both groups have focused on crossdisciplinary challenges involving statistics, applied mathematics, engineering, and computer science that are driven by important ecological questions and the characteristics of wireless sensor networks. Each Working Group developed and pursued a detailed research agenda, as outlined in the following.

4.1 Sensor Networks Datasets

Because this group brought together researchers from a very broad set of disciplinary perspectives, our weekly telemeetings were initially dedicated to two types of discussions: (i) short, informal talks and discussions by all working group members about their backgrounds and research interests, and (ii) discussions on research perspectives of the fields represented by the working group members.

Since the group felt strongly that research should be grounded in knowledge of issues that occur in actual experiments and real datasets, these discussions ran concurrently with the process of acquisition and evaluation of example datasets. Unlike in other disciplines, such bioinformatics, datasets from sensor arrays or wireless networks are very new and extremely rare. They are also very unwieldy, with diverse variables, different sampling intervals, and numerous faults of varying types and severity. Furthermore, necessary physical conversions require interaction with other sensed variables, coupling and propagating uncertainty.

Early on, the group identified three prerequisites for a successful research agenda that is grounded in real data: the dataset should be sufficiently rich in terms of its statistical properties and association with relevant research in ecology or environmental sciences; we should have sufficient knowledge of the data collection process used; and we should be able to closely interact with a scientist familiar with the experiment and dataset. With these in mind, we studied two datasets:

- Zoe Cardon (Ecosystems Center, Marine Biological Laboratory) presented a dataset for an experiment based on measurement of *water potential* at multiple sites around sagebrush plants in Utah. This experiment was designed to further understanding of soil microbial activity as a function of water in the soil. This dataset is of interest in part because it is from a wired sensor array, and thus is an excellent vehicle for exploring the effects of posited wireless networks. One of the important lessons of this dataset is that, in wireless sensor networking, a rich spectrum of errors and faults will occur regardless of, and in addition to, whatever effects wireless networking may have on the data.

- The group has also assessed a *root structure/soil respiration* dataset from the UCLA Center for Embedded Networked Sensing that used a wireless sensor network with the goal of characterizing the spatio-temporal properties and regulation of soil moisture. The network monitored dynamics of soil respiration, soil moisture, and fine root and rhizomorph (fungi) structure using mini-rhizotrons, with the objectives of understanding ecological processes related to the coupling between soil moisture and fine root and rhizomorph dynamics.

In broad terms, the goal of the working group was to answer the question: *How, and how well, can we answer important and inherently statistical questions in the ecological sciences with data from wireless sensor networks, and how do networks affect our ability to answer those questions?* The specific research questions that fall under this umbrella are challenging; they typically involve multiple, coupled dynamical processes with latent variables, as well as issues of scaling and dimensionality. With respect to wireless sensor networks, the group is working to develop approaches to the crucial open question of modeling energy consumption and efficiency in wireless data gathering networks.

The working group's research plan coalesced around two thematic areas. The first is *Data Analysis*, motivated by the following questions that are relevant to the water potential dataset, but represent challenges found in a wide spectrum of research questions in ecology and environmental science:

- Can we assume that the rate of water loss from deeper soil layers (via transpiration) is proportional to conductance across the soil-root interface?
- Is there a change in the relationship between root-soil conductance for water and soil water potential, e.g., as the season progresses?
- Is the amplitude of daily oscillations in soil water content a driver of soil biogeochemical processes that affect plant root function and growth?

A second theme that we developed is *System-Data-Network Interaction* to explore statistical tools that can be used to address questions such as:

- How do faults and errors interact with various network algorithms to affect our ability to answer ecological questions across time scales?
- How does the energy cost of computation at sensor nodes factor into decisions about network signal processing, inference, coding, and transmission algorithms, given the panoply of errors and faults that may occur?
- How can the model-mediated gathering of data be tuned, based on its explanatory power, given the energy requirements of sampling, computation, and data transmission processes?

The Sensor Datasets Working Group self-organized into two subgroups along the Data Analysis and System-Data-Network Interaction themes, and, in telecon meetings, zeroed in on two problem areas.

The Data Analysis subgroup tackled the above questions in the context of aridland soil-plant-air systems. A unique aspect of this work is inference of hydraulic redistribution and its drivers, the latter including variations of plant-air conductivity in the summer monsoon season. A paper in preparation will compare Bayesian and classical inference techniques for capturing the relative effects of water potential gradients among plants, the atmosphere, and soil at different depths.

The Data Analysis subgroup also addressed automated detection of anomalous data from sensor networks, where Ernst Linder has led the development of algorithms based on the median polish. As part of this work, he advised Jared Murray (Undergraduate Student, Department of Mathematics & Statistics, U. of New Hampshire), on an undergraduate research project (Fall 2008-Spring 2009) to develop an interactive graphical tool called MP-TUNER for automated anomaly detection for multiple time series data from environmental sensor network. Software is currently under development and almost finished, and can be accessed at http://pubpages.unh.edu/ jsb28/.

There are two specific thrusts in the In System-Data-Network Interaction area. In the first (Howard and Flikkema), we are studying how to couple network-aware source coding, channel coding, transmission control, and Bayesian source-channel inference (a generalization of MAP decoding) at the destination node. This work focuses on the trade-off of uncertainty reduction and energy consumption rather than focusing on information rate, since channel capacity is not limiting in this environmental sensing application. Published results address both global inference at the information sink and a form of in-network cooperative communication wherein nodes use local information to make communication decisions based on prediction of the consequences of candidate decisions using a Bayesian framework. The second thrust (Flikkema with Undergraduate Student Kenji Yamamoto, EE Dept., Northern Arizona U.) is addressing the gap in understanding between theoretical results and practical implementation of in-network algorithms on energy-limited sensor nodes. For example, if the energy cost (stemming from computational complexity) of inference algorithms is too high, it may exceed the reduction in communication energy cost enabled by those algorithms. A real-time power/energy measurement system has been designed, developed, and is under test that will provide accurate and precise estimates of the energy cost of algorithms running on sensor nodes, where electrical current demands vary over five orders of magnitude in both magnitude and time scales.

4.2 Sensor Design

During the first two meetings (Jan. 24 and 31), the working group focused on studying specific ecological applications of wireless sensor networks, in order to better understand the needs of the ecology researchers. The two applications studied were the redwood tree monitoring project of Todd E. Dawson (Integrative Biology, UC Berkeley), and the Duke Forest monitoring project of James S. Clark. Neal Patwari (Electrical Engineering, U. Utah), David Bell (Environmental Science, Duke U.), and Yongku Kim (Statistics, SAMSI) led the discussions.

In the third meeting (Feb. 7), XuanLong Nguyen (Statistics and Computer Science, SAMSI) gave an overview and survey of *suppression* and related techniques in distributed systems and sensor networks. By exploiting redundancy that naturally arise in sensor data, these techniques reduce the amount of data that needs to communicated to a gateway or base station for collection, thereby conserving energy and prolonging the lifetime of the sensor network deployment.

The next three meetings (Feb. 14, 21, and 28) were devoted to a series of roundtable discussions, wherein each participant prepared a couple of ideas of potential interest to the working group and led the group discussion on these ideas. Paul Flikkema (Electrical Engineering, Northern Arizona U.) talked about joint coding, estimation, and transmission censoring. Marco A. R. Ferreira (Statistics, U. Missouri, Columbia) presented a Bayesian decision-theoretic setup for tackling sensor design problems. Scott Holan (Statistics, U. Missouri, Columbia) proposed looking at adaptive sampling and design problem, and studying models of network failure. Jun Yang (Computer Science, Duke U.) argued for reducing the total maintenance cost of the network instead of total energy consumption, and applying model-driven techniques to the system health monitoring of the network itself. Zhengyuan Zhu (Statistics, UNC Chapel Hill) discussed opportunities of improving data collection efficiency using spatio-temporal sampling design. Neal Patwari and Jessica Croft (Electrical Engineering, U. Utah) proposed considering adaptive deployment and survival strategies for the network. Yongku Kim talked about challenges in statistical analysis, suppression scheme design, and dynamic models. Michela Cameletti and Rosaria Ignaccolo (Statistics, SAMSI) discussed their experience with the Piedmonte PM10 monitoring network and problems in adaptive sampling and model- and entropy-based network design. XuanLong Nguyen presented two specific problems: study of data reduction vs. statistical efficiency of suppression schemes, and sensor selection driven by a spatial model. Ilka A. Reis (Statistics, SAMSI) talked about design of better temporal suppression schemes. Christine Shoemaker (Environmental Engineering, Cornell U.) presented her project on monitoring Cannonsville Reservoir Basin, and challenges in efficient simulation and assessing uncertainty in simulation models. James S. Clark elaborated on the idea of model-based data suppression using soil moisture data as an example.

In the meeting on Mar. 6, Jun Yang summarized the main threads among the problems of interest presented during the roundtable discussions. The group decided to focus on two specific design problems:

- *Design and analysis of data collection schemes.* Given a time series of raw readings, the system can employ a variety of techniques to save communication (and hence energy): a) randomly transmit a reading with some probability; b) transmit only readings that differ by more than ϵ from their predicted values; c) quantize each reading and only transmit the quantized value if it is different from the last transmitted reading; and d) compress the readings and then transmit the compressed data. Although there has been a lot of work based on these ideas, more rigorous analysis is needed in order to quantify the cost/benefit tradeoff among them and to better understand their relationships and differences. Once formal definitions of cost and benefit are chosen, the design problem involves choosing the best data collection scheme and the optimal parameter setting for it (e.g., probability of transmission, value of ϵ, level of quantization, or compression method).

- *Spatial sensor deployment design.* Given a spatio-temporal model that one wishes to learn using a collection of sensors, where should the sensors be placed to achieve the desired cost/benefit target? While the general experimental design problem has been studied extensively in statistics, the traditional cost models and design constraints are probably inappropriate in the sensor network setting. With cost models and constraints unique to sensor networks, the problem will be a novel and interesting one.

In preparation for tackling the above two problems, the working group reviewed the necessary background knowledge. On Mar. 13, Neal Patwari, representing the electrical engineering perspective, presented models for path loss, interference, and battery power in wireless sensor networks. On Mar. 20, Zhengyuan Zhu, representing the statistical perspective, gave an overview of known results on spatial sampling design.

After these meetings and presentations, subgroups of the Sensor Design Working Group were formed to focus on specific research problems. These subgroups include *sampling/routing design subgroup*, *suppression design subgroup*, and *review article subgroup.*

The *sampling/routing design subgroup* consists of XuanLong Nguyen (statistics and computer science, SAMSI), Jun Yang (computer science, Duke U.), Yi Zhang (computer science, Duke U.), and Zhengyuan Zhu (statistics, U. North Carolina at Chapel Hill). This group is working on jointly designing data sampling and network routing strategies for environmental wireless sensor networks. Traditionally, these two aspects of the design problem have been tackled separately: sampling design (to achieve specific modeling goals) mostly has been the concern of the statistics community, while energy-efficient routing is the focus of the computer science community. Obviously, a truly optimal design must tackle both these aspects, because both sampling and routing have large impact on the consumption of energy, often the most precious resource on battery-powered sensor nodes. The sampling/routing design subgroup is now tackling the design problem by jointly considering these two aspects.

The sampling/routing design subgroup has been meeting regularly since April 2008, and has made considerable progress. Zhengyuan Zhu is now leading the effort in writing an article to summarize the subgroup's findings.

A sizable subset of the Sensor Design Working Group participated in the transitional workshop in October 2008. On behalf of the sampling/routing design subgroup, Zhengyuan Zhu (statistics, U. North Carolina) summarized the findings on the optimal joint design of data sampling and message routing and in wireless networks. Members of the Sensor Design Working Group present at the workshop also discussed possible next steps for the working group.

The sampling/routing design subgroup also sought feedback from and collaboration with the DDDAS team at Duke U. and Northern Arizona, who has been working on deploying sensors in the Duke Forest to study forest growth. The DDDAS team already includes many active members of the SAMSI working groups: James S. Clark, Paul Flikkema, Alan Gelfand, Kristian Lum, XuanLong Nguyen, Jun Yang, and Yi Zhang. Zhengyuan Zhu joined one of the DDDAS project meeting in October 2008 and discussed the possibility of applying the results of the sampling/routing design subgroup in the practical context of the DDDAS project.

The *suppression design subgroup* consists of Kristian Lum (statistics, Duke U.), Jun Yang, and Yi Zhang. This subgroup is interested in the design of *suppression* schemes, which is a way to reduce communication (and therefore save energy) in sensor networks by using predictive models to suppress reporting of predictable data. However, in the presence of communication failures, missing data is difficult to interpret because it could have been either suppressed or lost in transmission. To date, there has been no solution for handling failures for general, spatio-temporal suppression that uses *cascading*, where a node can use suppression in reporting its readings to another node, which can then use suppression again in further reporting this reading together with other readings to a third node, etc. While cascading further reduces communication, it makes failure handling very difficult, because nodes can act on incomplete and incorrect information and in turn affect other nodes. The subgroup has developed a cascaded suppression framework that fully exploits both temporal and spatial data correlation to reduce communication, and applies coding theory and Bayesian inference to recover missing data resulted from suppression and communication failures. A paper on this subject is currently under submission.

The *review article subgroup* of the Sensor Design Working Group is led by Soumen Lahiri (statistics, Texas A&M U.), and includes XuanLong Nguyen, Jun Yang, and Zhengyuan Zhu. This subgroup is working on a survey article to be submitted to the journal *Statistical Science*. This article will provide the background on wireless sensor networks, and highlight the probabilistic and statistical challenges.

5 Graduate Student Involvement

Three local graduate students actively participated in this program:

David Bell (Duke University) is the SAMSI Graduate Fellow associated with the Sensor Network Dataset working group (Spring 2008) and also the Environmental Risk working group (Fall 2007). He has been active in attending meetings as well as acquisition and exploration of data. He has also been involved in modeling soil moisture data from a local wireless sensor network in Duke Forest with James Clark, Paul Flikkema, Alan Gelfand, Yongku Kim, and XuanLong Nguyen. With his advisor, James Clark, he is developing his dissertation research plan which will involve the use of environmental sensor network data in examining plant-insect interactions in a mixed-hardwood forest. His experience in the SAMSI program prepared him for dealing with data analysis of often faulty sensor data. During his graduate fellowship, he presented a poster at SAMSI's Environmental Sensor Network

Workshop (January 2008) concerning modeling of battery data from a wireless sensor network to identify the effects of transmission and data collection on sensor node longevity. He has also given a presentation as an introduction to ecological modeling with sap flux data during SAMSI's PostDoc/Graduate Student Seminar (November 2007), a presentation regarding the use of mathematics and statistics in ecology and environmental sciences at the SAMSI Undergraduate Workshop (March 2008), and gave another presentation during SAMSI's PostDoc/Graduate Student Seminar (April 2008).

Kristian Lum (Duke University) – please see entry under Section 3.

Ilka A. Reis (National Institute for Space Research [INPE], Brazil) has a background in statistics and, currently is pursuing a doctorate in Remote Sensing at INPE. She is interested in developing methods for data collection in sensor networks, especially using data suppression. She attended the kickoff workshop of the SAMSI Environmental Sensor Networks program (January, 13-16, 2008), where she presented the work "Temporal suppression by outlier detection for data collection in sensor networks". She spent the following 8 weeks visiting SAMSI, where she attended the initial meetings of the Working Groups formed as a result of the workshop. While she has since returned to Brazil, she is continuing her involvement in the telemeetings. During her visit to SAMSI, she interacted with researchers to explore the statistical issues involved in the environmental data collection using sensor networks. As a result, she started extending her previous work on temporal data suppression to a more general spatial-temporal suppression scheme. This work, in addition to previous efforts, is expected to form her dissertation.

Another graduate student, Jessica Croft (University of Utah), participated remotely in the telemeetings.

6 Publications and Presentations

Lance Waller organized a session "Monitoring Sensor Networks in Ecology" at the American Statistical Association Environmental Statistics Section's Workshop on Environmetrics (NCAR, Boulder, CO, October 22-24, 2008). Program participants contributed talks as follows:

- Paul Flikkema (collaborative work with Sheryl Howard) - The Roles of Compression and Coding in Inference on Wireless Sensor Networks
- Ernst Linder (collaborative work with Yongku Kim, Zoe Cardon, Scott Holan, Ernst Linder, and Paul Flikkema) - Median Polish Algorithms for Automated Anomaly Detection in Environmental Sensor Networks
- Yongku Kim (collaborative work with Long Nguyen and Scott Holan) - A Correlation Process Prior For Anomaly Detection of Functional Data

- Zhengyuan Zhu (collaborative work with Long Nguyen and Jun Yang) - Optimal Design of Sensor Networks under Energy Budget Constraints

List of Presentations

(Presentations at SAMSI workshops not included.)

Flikkema, P. The Roles of Compression and Coding in Inference on Wireless Sensor Networks. American Statistical Association, Environmental Statistics Section, Workshop on Environmetrics, NCAR, Boulder, CO, October 22 - 24, 2008.

Kim, Y. A Correlation Process Prior For Anomaly Detection of Functional Data American Statistical Association, Environmental Statistics Section, Workshop on Environmetrics, NCAR, Boulder, CO, October 22 - 24, 2008.

Linder, E. Median polish algorithms for automated anomaly detection in Environmental Sensor Networks. American Statistical Association, Environmental Statistics Section, Workshop on Environmetrics, NCAR, Boulder, CO, October 22 - 24, 2008.

Linder, E. Median polish algorithms for automated anomaly detection in environmental sensor networks. ENAR: Eastern North American Regional Meetings of the International Biometrics Society, San Antonio, TX, March 15 - 18, 2009.

Zhu, Z. Optimal Design of the Sensor Network under Energy Budget Constraints American Statistical Association, Environmental Statistics Section, Workshop on Environmetrics, NCAR, Boulder, CO, October 22 - 24, 2008.

List of Publications

Cardon, Z.G., Flikkema, P., Herron, P.M., Holan, S., Kim, Y., Linder, E., and Stark, J.M. A new view of hydraulic redistribution of soil water during rainstorms. To be submitted to Ecology.

Cardon, Z.G., Stark, J. M., Herron, P.M. (2009) Hydraulic redistribution and the fate of root-derived carbon in soil. Abstract submitted for Ecological Society of America meetings, August 2009, Albuquerque, NM.

Gelfand, A.E. and Puggioni, G. Analyzing Space-time Sensor Network Data under Suppression and Failure in Transmission, Statistics and Computing (forthcoming).

He, Y. and Flikkema, P. System-Level Characterization of Single-Chip Radios for Wireless Sensor Network Applications. IEEE WAMICON 2009, April 20-21, 2009, Clearwater, FL USA.

Howard, S. and Flikkema, P. Integrated Source-Channel Decoding for Correlated Data-Gathering Sensor Networks. IEEE Wireless Communications and Networking Conference (WCNC 2008), March-April 2008

Howard, S. and Flikkema, P. Progressive Joint Coding, Estimation and Transmission Censoring in Energy-Centric Wireless Data Gathering Networks. Fifth IEEE International Conference on Mobile Ad-hoc and Sensor Systems (MASS 2008), Sept-Oct 2008.

Linder, E., Cardon, Z., Murray, J., Holan, S., Flikkema, P., Ignaccolo R., Kim, Y. A Sequential Median Polish for Automated Data Cleaning and Anomaly Detection in Environmental Sensor Networks. Paper in preparation.

Murray, J. Median polish algorithm for automated anomaly detection in sensor networks (MP-Tuner). Entry to 2009 Student Computing Competition by the American Statistical Association (Section on Computing and Graphical Statistics).

Murray, J. Median polish algorithm for automated anomaly detection in sensor networks (MP-Tuner). Entry for the 2009 U. of New Hampshire Undergraduate Research Conference. Interactive presentations to be given April 22 and April 24, 2009 (U. of New Hampshire).

Nguyen, X., Bell, D., Clark, J., Gelfand, A. and Kim, Y. Modeling and computation of wireless sensor network data for environmental monitoring. In preparation.

Nguyen, X., Yang, J., Yang, Y., and Zhu, Z. Optimal sensor network design under budget constraints. In preparation.

Nguyen, X., Holand, S., and Kim, Y. A correlation process prior for anomaly detection with functional data. In preparation.

Nguyen, X., Huang, L. and Joseph, A. (2008). Support vector machines, data reduction and approximate kernel matrices. Proceedings of the 19th European Conference on Machine Learning (ECML), September, Antwerp, Belgium.

Rajagopal, R., Nguyen, X., Ergen, S. and Varaiya, P. Theory of multiple sequential change-point detection. To be submitted to IEEE Trans. on Signal Processing.

Rajagopal, R., Nguyen, X., Ergen, S. and Varaiya, P. (2008). Distributed online simultaneous fault detection for multiple sensors. International Conference on Information Processing in Sensor Networks (IPSN), St. Louis, MO.

Rajagopal, R., Nguyen, X., Coleri-Ergen, S., and Varaiya, P. (2009). Theory of simultaneous fault detection for multiple sensors. Second International Workshop on Sequential Methodologies (IWSM), Troyes, France (invited extended abstract).

Silberstein, A., Puggioni, G., Gelfand, A., Munagala, K., and Yang, J. Suppression and Failures in Sensor Networks: A Bayesian Approach, Proceedings of the 2007 International Conference on Very Large Data Bases (VLDB '07), Vienna, Austria 2007; 842–853.

Silberstein, A., Braynard, R., Filpus, G., Puggioni, G., Gelfand, A., Munagala, K., and Yang, J. Data-Driven Processing in Sensor Networks. Proceedings of the 3rd Biennial Conference on Innovative Data Systems Research (CIDR), 2007, Asilomar, California; 10–21.

Yamamoto, K. and Flikkema, P. Prospector: Multiscale Energy Measurement of Embedded Systems with Wideband Power Supply Signals. In preparation.

7 Workshops

7.1 Planning meeting

The Program Leaders Committee was able to define the program and organize the Opening Workshop via email and teleconference, so a formal planning meeting was not required.

7.2 Opening workshop

The opening workshop for the program was held on January 13-16, 2008, attracted 78 attendees from diverse fields, and met the goal of establishing the composition and activities of the Working Groups. Details of the workshop program are at `http://www.samsi.info/workshops/2007sensor-opening200801.shtml`, and all the presentations at the opening workshop are available at the SAMSI web site.

7.3 Transition workshop

The transition workshop was held October 20-21, 2008, and featured talks by eight program participants (including three female researchers and one graduate student) as well as time for extensive discussions within and between the two Working Groups. Details of the workshop program, including all presentations, are at `http://www.samsi.info/workshops/2008sensor-transition200810.shtml`.

8 Education and Outreach

The Opening Workshop for the Program was preceded by a day of Tutorial Overviews with the following speakers:

- Paul Flikkema, Northern Arizona University: *Ecosystem Inferential Models to Control Data Acquisition and Assimilation*
- Bill Kaiser, Univ. of California-Los Angeles: *Sensor Network Platforms for Rapidly Deployable, Configurable, and Sustainable Observatories*
- Jennifer Hoeting, Colorado State University: *Hierarchical Modeling*
- Kiona Ogle, University of Wyoming: *Data-Model Integration: Examples from Below-ground Ecosystem Ecology*

All the tutorial presentations are available on-line on the SAMSI website.

Paul Flikkema organized a session on Environmental Sensor Networks at a SAMSI Undergraduate Workshop, Feb. 29 - Mar. 1. Speakers were Kenji Yamamoto (undergraduate student, Northern Arizona University), Dave Bell (graduate student, SAMSI/Duke University), XuanLong Nguyen (SAMSI postdoctoral fellow), Yongku Kim (SAMSI postdoctoral fellow), and Michael Porter (SAMSI postdoctoral associate). Paul Flikkema moderated the session.

9 Industrial and Governmental Participation

Yuliy Baryshnikov (Bell Laboratories) and Mike Godin (Monterey Bay Aquarium Research Institute) were invited speakers for the Opening Workshop. There was also participation in the opening workshop and working groups from government agencies, laboratories, and industry, including EPA, Centers for Disease Control and Prevention (CDC), Marine Biological Laboratory, the IBM T. J. Watson Research Center, the Center for Wireless Communications, University of Oulu (Finland), and the National Institute for Space Research (Brazil).

10 External Support

This program did not have external support.

11 Affiliates Participation

There were working group participants from each of the following university affiliates: University of California - Berkeley, Duke University, University of Michigan, North Carolina State University, and University of North Carolina at Chapel Hill.