Initial Analyses and Demonstration of a Soil Moisture Smart Sensor Web

Mahta Moghaddam(1), Dara Entekhabi(2), Leila Farhadi(2), Yuriy Goykhman(1), Mingyan Liu(1), Aditya Mahajan(1), Ashutosh Nayyar(1), David Shuman(1), and Demosthenis Teneketzis(1)

[1] University of Michigan, Electrical Engineering and Computer Science
1301 Beal Avenue, Ann Arbor, MI, 48109
Tel: 734-647-0244, email: mmoghadd@umich.edu

[2] Department of Civil and Environmental Engineering
Massachusetts Institute of Technology, Cambridge, MA

Abstract **-** We have developed a new concept for a smart sensor web technology for measurements of soil moisture that include spaceborne and in-situ assets. The objective of the technology is to enable a guided/adaptive sampling strategy for the in-situ sensor network to meet the measurement validation objectives of the spaceborne sensors, with respect to resolution and accuracy. One potential application is the Soil Moisture Active/Passive (SMAP) mission, The sensor nodes are guided to perform as a macro-instrument measuring processes at the scale of the satellite footprint, hence meeting the requirements for the difficult problem of validation of satellite measurements. The science measurements considered are the surface-to-depth profiles of soil moisture estimated from satellite radars and radiometers, with calibration and validation using in-situ sensors. Satellites allow global mapping but with coarse footprints. The total variability in soil-moisture fields comes from variability in processes on various scales. Installing an in-situ network to sample the field for all ranges of variability is impractical. However, a sparser but smarter network can provide the validation estimates by operating in a guided fashion with guidance from its own sparse measurements. The feedback and control take place in the context of a dynamic data assimilation system. The overall design of the smart sensor web - including the control architecture, assimilation framework, and actuation hardware - will be presented in this paper. We also present results of initial numerical and laboratory demonstrations of the sensor web concept, which includes a small number of soil moisture sensors and their measurement model, a dynamic soil moisture time-evolution model (SWAP), and an optimal control strategy. Based on these results, the TRL has been advanced to 3 from the initial level of 2.

I. INTRODUCTION

The long-term vision of Earth Science measurements involves sensor webs that can provide information at conforming spatial and temporal sampling scales, and at selectable times and locations, depending on the phenomena under observation. Each of the six strategic focus areas of NASA Earth Science (climate, carbon, surface, atmosphere, weather, and water) has a number of measurement needs, many of which will ultimately need to be measured via such a sensor web architecture. Here, we develop technologies that enable key components of a sensor web for an example measurement need, namely, soil moisture. Soil moisture is a measurement need in four out of the six strategic focus area roadmaps (it appears in climate, carbon, weather, and water roadmaps). It is used in all land surface models, all water and energy balance models, general circulation models, weather prediction models, and ecosystem process simulation models. Depending on the particular application area, this quantity may need to be measured with a number of different sampling characteristics. It is therefore necessary to develop sensor web capabilities to enable flexible and guided sampling scenarios, as well as calibration and validation strategies to support them.

This project seeks to develop and demonstrate, via numerical and laboratory experiments, the architecture and algorithms for a sensor web control system that interconnects the elements of the web and enables "smart sensing" through the integration of a data assimilation framework. The sensor nodes will be guided to serve as a macro-instrument compatible with the large-scale effective measurements by satellite sensors.

II. THE SENSOR WEB CONCEPT

The ground footprints of remote sensors are often coarser than the scale of variations of the variables. As a result, the remote sensing estimate is only a coarse-resolution estimate of a field mean. In-situ sensors often sample a point location in the heterogeneous field. Statistics of errors of retrieval are indicative of errors in measurements, and errors in representativeness of in-situ samples. These two errors cannot be separated using existing sampling networks.

For soil moisture profile fields, for example, the total variability is derived from variability in processes that influence it on a wide spectrum of scales ranging from meters to several kilometers. This broad spectrum of variability and multiple causes is not unique to soil moisture, but is a characteristic of many Earth system variables. A key challenge is how to calibrate and validate the satellite footprint estimate, for example from SMAP, which is an average of the field that may be 10s or 100s of km^2 for the radar and radiometer, resepectively. To install an in-situ sensor network that samples the field across all ranges of variability is impractical and cost-prohibitive. Our hypothesis is that a much smaller but smarter network can provide the needed validation estimates for satellite measurements.

The sensor web has to operate in a guided fashion. The guidance comes from the sparse measurements themselves, which, through a control system, guide the sensor web to modify the sampling rate and other parameters such that their observations yield the most representative picture of the satellite footprint conditions. The control and feedback take place in the context of a data assimilation system that merges data from forecast models, sensors, and relevant auxiliary information to produce the best estimate of the variable field and its anticipated evolution, balanced against measurement costs. This means that even if a measurement may improve the value of the soil moisture estimate, if it is too costly in term of power usage, the optimal decision could be not to take that measurement.

Here, we develop and demonstrate this control system for guided sampling by a sensor web. The guidance is towards producing representative and statistically unbiased estimates of the remote sensing footprint variable estimate based on a finite-size sensor web with dynamic operations. The duty cycle and sampling at the network nodes will be driven by a data assimilation system that can provide guidance on the worth of each measurement at different sampling intervals. Uncertainty in the model and current estimates can form the basis for the quantitative evaluation of the worth of data at each sensor web node. Dynamic commanding and data ingestion from those nodes optimize the value of the sparse ground-truth network in validating the remote sensing-based coarse-resolution retrieval. Here, we build a prototype of the semiclosed control system for the sensor web, coupled with a data assimilation system, for the case of soil moisture remote sensing as an example. The remote sensing instruments could produce observations at km-scale. The instruments operating at an in situ node could include meteorological sensors (temperature, precipitation, wind, solar radiation, etc.), soil moisture probes installed on surface and at varying depths, and multifrequency tower-mounted radars for O(100)m observations of soil moisture profile fields. Ancillary data such as topography, vegetation cover, and soil texture could also be provided at the spatial scale of in situ observations. There are specific challenges with validating remote sensing estimates using these point and/or small-footprint (O(100m)) in situ samples. These challenges will allow the demonstration of the advantages of the proposed approach.

The real-time data assimilation will track the conditions for variability in soil moisture and guide the sensor web to modulate its measurement duty-cycle and other parameters across the network. This is an adaptive sampling network guided by the data assimilation system that can feed back the value of each additional measurement. A block diagram depicting the interrelationships of the elements of the proposed system is shown in Figure 1.

The in-situ sensor web data, fed to the coupled assimilation-control system, will be used to determine the parameters required for the next set of sensor web measurements. Depending on the meteorological and other physical scene variations that are judged to influence the soil moisture profiles between the remote sensing measurement intervals, the sensors could be commanded to turn on or off, and depending on sensor type, to modify their sampling

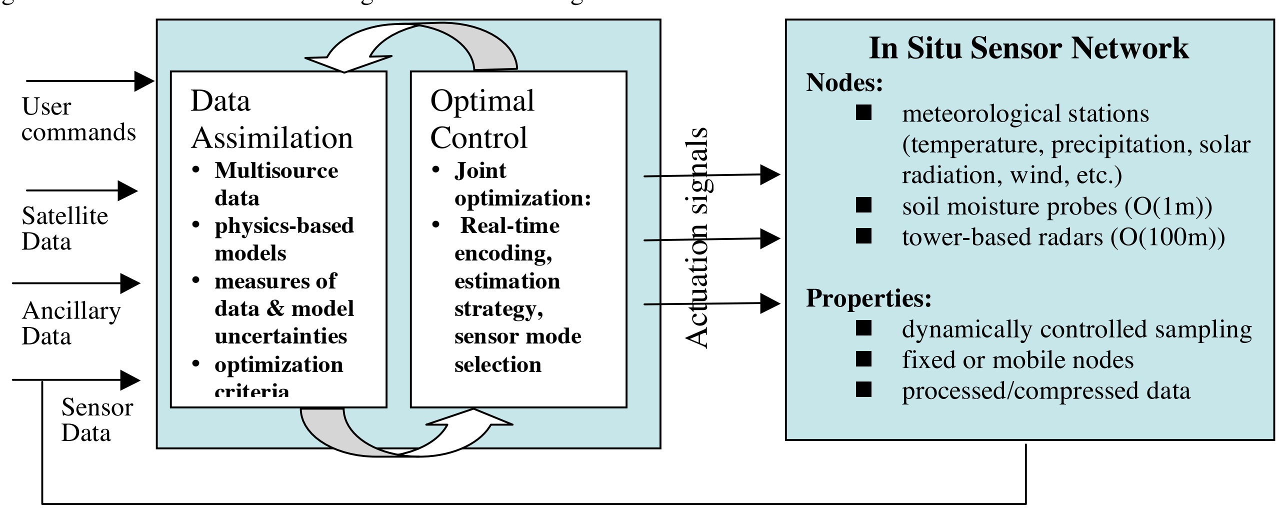

Figure 1. Elements of the sensor web technology and their interrelationships. The semi-closed system generates guidance to the sensor web, through actuators, for modifying its sampling characteristics using a coupled data assimilation and control system, antecedent sensor data, and ancillary data (e.g., topography and soil texture). User command can also be incorporated.

parameters. For a tower-mounted radar, for example, parameters to control are the frequencies for different depths of observation or vegetation conditions, bandwidth for variable spatial resolutions, and number of samples to average depending on expected measurement noise levels. An actuator will transmit a data packet containing the control signals to each relevant node. The nodes will decode the received signals and set their sampling parameters accordingly.

III. Approach

A. Data Assimilation

Data assimilation is a statistical estimation framework that combines physics-based model forecasts with observations [1]-[3]. In data assimilation it is assumed that models have uncertainty. It is also assumed that observations have errors. The relevant measures of the probability density function of the model forecasts are propagated in time until measurements are available. The probability density function or measures of it are updated based on the relative uncertainty of model forecasts and observations. The data assimilation and the sensor web will be coupled through an optimal control system. The duty cycle and weight given to each node measurement is evaluated against the value of that measurement in the data assimilation system. Reduction in resulting covariance will determine which observation will have the most value. This information is passed on to the control system to dynamically adjust the sample averaging and data collection. Since the model in the data assimilation will incorporate and integrate auxiliary information on vegetation, terrain drainage, soil texture, precipitation, etc., the possibility of fixing sensor positions representative of larger area will be explored. Data assimilation and sensor web-guided identification of station locations is beneficial since it allows long-term (multi-year) monitoring of a satellite footprint-scale estimate based on a sparse in-situ network.

Data assimilation has to take place in the context of a time-evolution model describing the physical process of soil moisture variations. The time and depth evolution of soil moisture fields can be expressed via a pair of coupled partial different equations (PDE) in space and time. This model has a number of parameters associated with terrain and meteorological conditions. The solution to the coupled differential equation is an estimate of future states of soil moisture fields with the knowledge of the current state and the model parameters.

The Soil-Water-Atmosphere-Plant (SWAP) model [4] is a community standard solver for such a model. SWAP incorporates surface energy balance by including micrometeorological data such as precipitation, winds, air temperature, and humidity. It also incorporates soil physics properties such as amplitude and phase characteristics of flow dynamics. It then solves the coupled differential equations numerically. We have used SWAP to develop a time-series of soil moisture variations using actual values of rainfall measurements for sample areas.

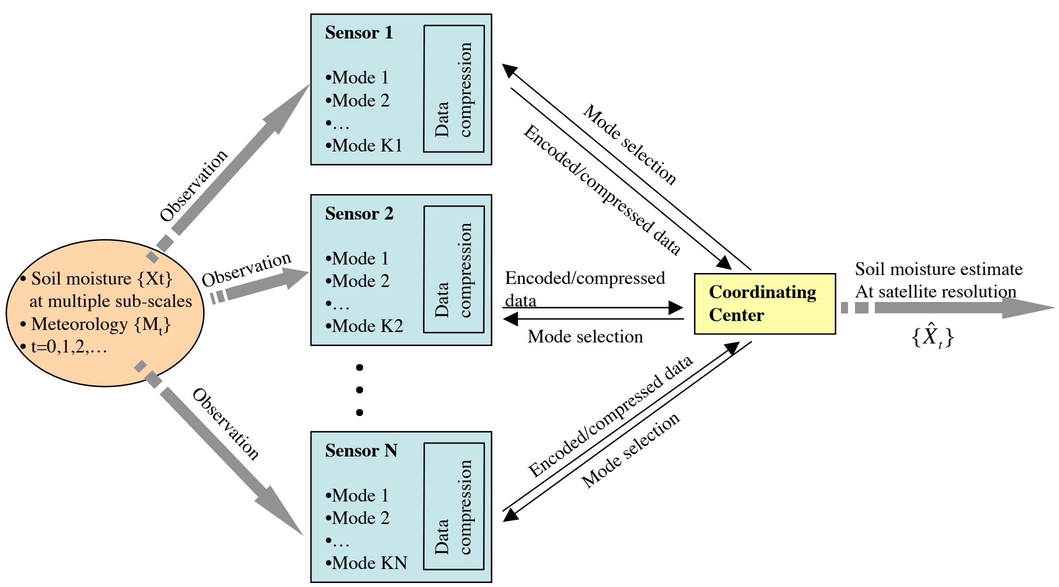

Figure 2. Control architecture. Each sensor measures variables over a finite period of time. Variables are correlated with the soil moisture field. Data are compressed at each sensor node and transmitted to the coordinating center, which derives an optimal control instruction set for the sensors, as well as an unbiased estimate of the soil moisture field at the remote sensor resolution, guided by data assimilation.

B. Control Architecture

The physics-based models have uncertainty and observations have errors. Thus, we model soil moisture at any point location in a spatial field as a discrete time stochastic process $\{X_t, t = 0,1,2,...\}$, the evolution of which is described by a stochastically forced hydrologic model. At specified times that maximize the information content of a measurement, each sensor can be activated to sense and transmit information. The data gathered by each sensor is encoded/compressed and transmitted in real-time through noisy channels to the coordinating center. At any time t the coordinating center utilizes the information it has gathered up to t to estimate field mean and to specify the mode each sensor will employ at time $t+1$, so as to gather additional data. Thus, the objective is to determine: (i) a sensor mode selection strategy for the coordinating center; (ii) an estimation strategy for the coordinating center; and (iii) real-time encoding strategies for the sensors, so as to minimize the expected value of the sum of a function of the difference of the estimate $\hat{X}_t$ and field mean from $t = 0$ up to time horizon T. The control system architecture is shown in Figure 2.

The above-described interdependence of sensor mode selection and real-time encoding strategies results in a challenging optimization/control problem [5]-[13]. We have established a common mathematical framework and terminology for the different elements of the project as shown in Figure 3.

is X_{t+1}. The vector of all soil moisture measurements X_t is also called the *state vector*. The dynamic model f_t is also called the *state transition operator*. The state transition operator describes the evolutions of soil moisture in space and time. Both variations are functions of scene parameters such as topography, soil type and texture, vegetation. They are also functions of external forcings such as rain, cloud cover, solar radiation, and temperature. The evolution of the soil moisture state vector is generally a dissipative process, but one that is forced with these exogenous discontinuities. The discontinuities, could pose barriers to information if the sensor network does not adapt its sampling strategy to capture the rapidly varying nature of the discontinuities.

We also include the sensor measurement model as part of the assimilation system, since it provides the sensor data as an input to the time-evolution model. For the i[th] sensor, the measured value Y_t^i is related to the value of the variable soil moisture X_t via a physical model h_t^i and sensor parameter configuration U_t^i. These parameters could be frequency, polarization, power level, etc. Measurement noise is added to the true signal and denoted as V_t^i. Sensors make observations that can be translated into estimates of unknown variables. Sensor models do not include any time evolution or dynamic nature. They can, however, include the probabilistic nature of the unknowns at time t. The models and unknowns could be scalar (1-D) or vector (N-D), depending on how many variables are being measured and

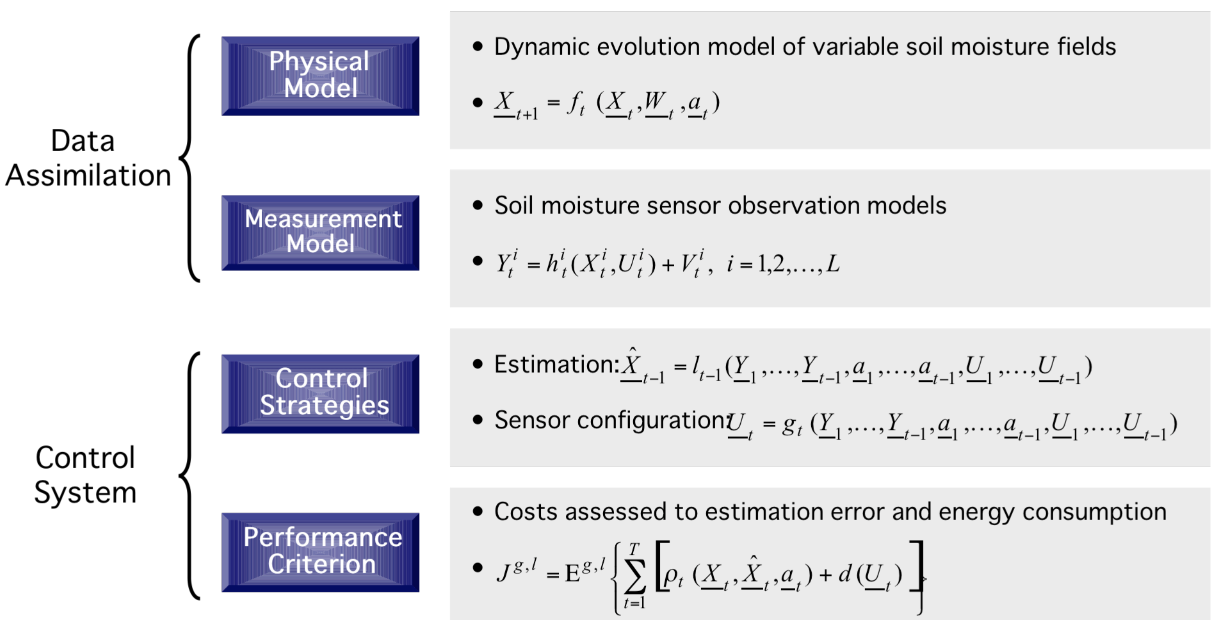

Figure 3. Overall problem formulation and mathematical notation, showing the relationship between the different project components.

The data assimilation component consists of a physical dynamic evolution model f_t, which, given the knowledge of values of the variable soil moisture X_t (up to time t), parametric uncertainties W_t (such as uncertainties in topography or temperature), and *exogenous forcings* a_t at time t (such as rain), predicts the value of variable soil moisture at the next point in time ($t+1$). This predicted value

how many sensors there are. Different sensors allow estimates of the unknowns at different spatial scales. Sensors could be in-situ (moisture probes) or remote (tower-based, airborne, or spaceborne SARs and radiometers). In general, the estimation of unknowns is a complex task, depending on the degree of model nonlinearity, measurement noise, and sensor calibration. It is assumed that each sensor is calibrated independently of the rest of the sensors in the web, but potentially in coordination with the entire web in terms of scheduling and resource usage.

The control strategy is derived for the objective of minimizing a cost measure, which is a combination of achieving the best possible variable estimate at any given time and minimizing resource usage for making the required measurements. This means that even if a measurement may improve the value of the soil moisture estimate, if it is too costly in term of power usage, the optimal decision could be not to take that measurement. This strategy holds for a centralized stochastic optimization problem with imperfect observations, as we have assumed. Fundamental issues in selecting a sensor configuration are:

- Energy consumption cost of current sensor configuration
- Effect on the quality of the current state estimate
- Effect on future decisions for sensor configurations and their effect on quality of future state estimates
- Trade-off between the first and last two items above

This problem belongs to the class of optimization problems known as Partially Observable Markov Decision Processes (POMDP). To solve such problems, backward induction is typically used to determine optimal sensor selection and estimation strategies sequentially in time, by moving backwards in time. The solution method has the following features:

- Compute conditional probability p_t of current state $\underline{X}_t$ using all previous measurements (and all previous sensor configurations)
- Choose optimal sensor configuration $\underline{U}_t$ and optimal estimate using p_t
- Sensor selection strategy g_t and estimation strategy l_t are determined by specifying the optimal sensor configuration and optimal state estimate for every possible realization of p_t

The above off-line computations are numerically very expensive. The complexity increases with increasing number of sensors and increasing number of sensor modes. The number of sensor data quantization levels is also a factor in increasing the complexity. Once the off-line computations are performed, the on-line implementation of the control strategy is rather straightforward. We have successfully applied this strategy to a 1-D problem and with varying levels of success to a 2-D problem. In each case, a control policy table has been generated. The results will be shown at the presentation.

C. Sensor Models

We envision a sensor web that will ultimately comprise of different varieties of sensors. In particular, the soil moisture sensors could be localized, such as probes, or could be remote, such as tower-mounted or aircraft-based radars or radiometers. Deriving physics-based remote sensor models to relate their measurements to estimates of soil moisture is generally rather complicated. The in-situ sensors, on the other hand, offer an opportunity for accurate measurements that are related to soil moisture values via simple empirical models. Several standard methods of in-situ sensing exist, such as time-domain reflectometer (TDR) probes, neutron probes, capacitance probes, and ring resonators. We have chosen an in-situ soil moisture probe making highly localized measurements. We selected and procured capacitance probes from Decagon, model ECH$_2$O EC-5, and developed its calibration curve in form of a third-order

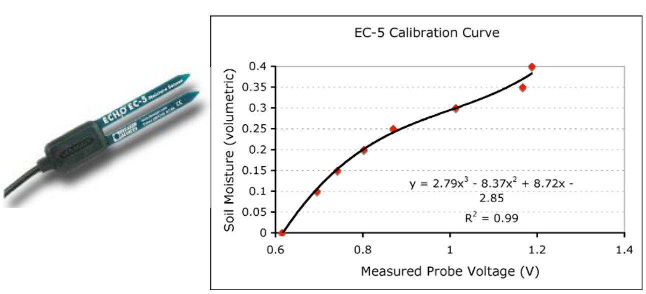

Figure 4. Left: the Decagon ECH$_2$O EC-5 soil moisture probe. Right: calibration curve (or "sensor model") derived from experimental data and used in the control algorithm.

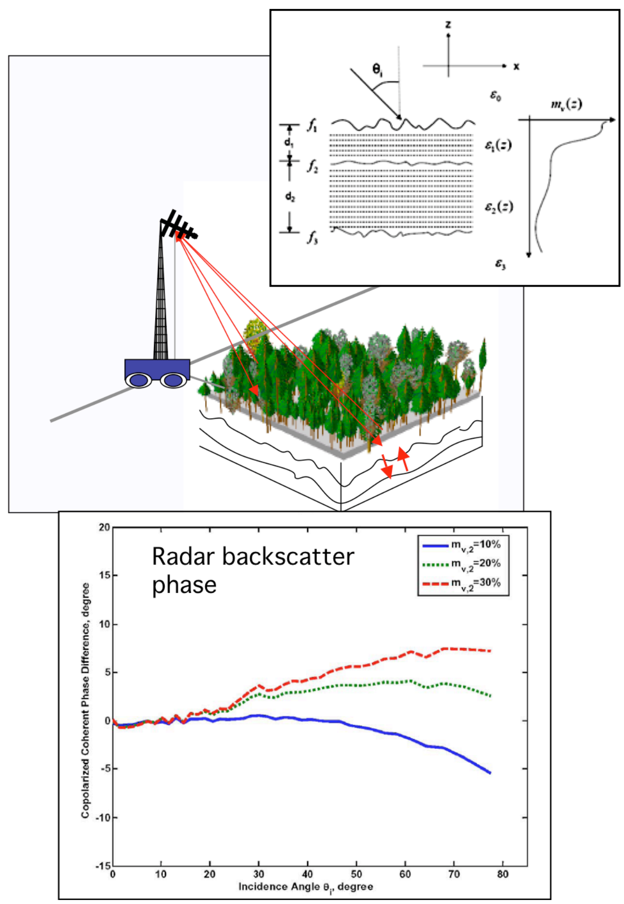

Figure 5. Top: a realistic soil moisture profile from surface to a depth of d1+d2. Middle: remote sensor could be a low-frequency radar (tower-mounted, airborne, or spaceborne) which can produce measurements related to soil moisture profiles at varying depths. Bottom: Example of backscattered co-pol phase dependence on moisture profile, to be used in developing inversion algorithms.

polynomial. This polynomial, shown in Figure 4 along with the probe and experimental data points, was used as the initial sensor model input to the control system. The model generated with the empirical data represents a calibration accuracy of about 1%.

For remote sensors, which could be tower-mounted, airborne, or spaceborne, the physics-based retrieval models of soil moisture involve solutions to nonlinear optimization problems. Considering a low-frequency radar as an example (Figure 5), its measured backscattering coefficients could be related to the profiles of soil moisture via models derived from Maxwell's equations. A number of models that relate radar backscattering coefficients to soil moisture have recently been developed (the "forward" problem). The models could be numerical or analytical. The "inverse" problem, or the retrieval problem, has also been addressed in our previous works, but needs further advancement. The basic strategy is to derive multi-dimensional polynomial expressions that are derived from the more complicated numerical models in several unknowns. The closed-form nature of the fitted model allows us to apply a number of optimization techniques, both local and global. The statistical nature of the unknowns (e.g., soil moisture and surface roughness) can be systematically included in development of the optimization algorithm.

D. Actuation hardware

A key enabling technology in the proposed system is the proper actuation of the sensors using output computed by the control algorithms. Actuation allows the measurement parameters to be dynamically adjusted according to the data collected so far, the inferred soil condition, as well as the overall objective of the sensing task. This process involves the sensor radio transceiver receiving the control message from the coordinating center, decoding the message into a set of parameter values associated with the measurement device modes, and issuing the actuation command that leads to the parameter adjustment of the measurement device. Commonly available sensor platforms for R&D purposes, e.g., the MICA2 motes [16], often lack sophisticated actuation capabilities. The typical actuation on these platforms is limited to setting data sampling rate and specifying the duty cycling rate of the sensors. The proposed system involves the dynamic tuning of a wide range of parameters representing various modes of each instrument. For example, for the tower radar, a number of parameters such as frequencies, power levels, and polarizations could be controlled. This requires the development of a customized actuator for the proposed system that will become part of the sensor board and is connected to the measurement device. For this development, we will leverage the existing Narada board already developed by a colleague at the University of Michigan.

IV. Laboratory Test-Bed

The control signals generated by the central coordinator need to be conveyed to the sensors via wireless links and actuators at the sensor locations. The objective of the laboratory test-bed is to provide experimental proof-of-concept for the actuation of sensors, given the control signal and antecedent sensor data that are available to the coordinator via wireless links.

We have planned two major phases for the actuation experiments (phases A and B). In Phase A, recently completed, COTS devices were used for actuation and wireless communication. The control feedback loop was

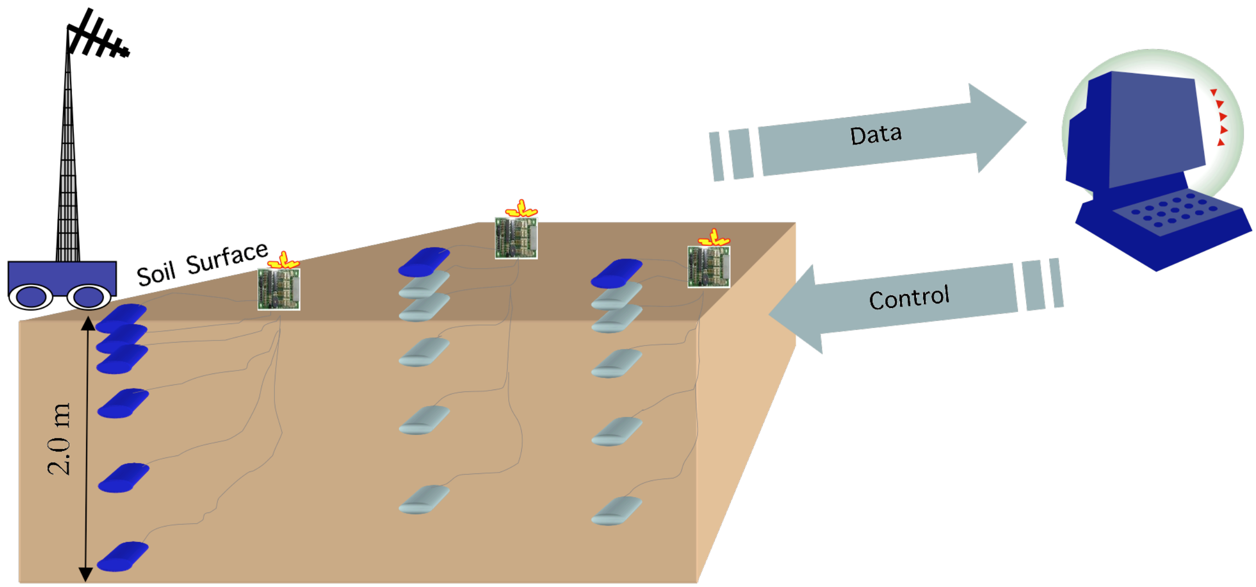

Figure 6. Field measurement conceptual setup: both remote and in-situ sensors are present, and send data to the coordinator. The coordinator issues command signals via the wireless link to actuators at sensor locations, which in turn set sensor measurement parameters.

implemented to command a single sensor at a single location via an actuation device. The control policy for the 1-D problem was successfully integrated with the lab set up and used to actuate the sensor at intervals prescribed by the control algorithm.

In Phase B, custom actuation and communication devices will be built (possibly using some COTS components). Furthermore, in parallel with the control algorithm progress, multiple sensors will be included in the demonstration, each of which can be controlled by the coordinator. Phase B will also include optimization criteria for power management. Phase A experiments were in the laboratory only. In Phase B, we plan to set up a field-analog experiment. Figure 6 shows the field experiment concept, where both in-situ and remote sensors may be used, each of which will receive commands from the coordinator and actuated accordingly. Each sensor can in turn send its data back to the coordinator.

V. Summary

The proposed technology for coupling a data assimilation framework into a sensor web control system to achieve an optimal dynamic sampling strategy is fundamentally new. Previous studies related to this topic exist, but have used an empirical approach to search for temporal stability of network nodes for capturing the mean conditions of the observed field [18]-[19]. No previous work has been done to implement such dependencies within a control system to guide the sampling of a sensor web.

The novel aspects and benefits of this technology are:

- It uses a physics-based approach to relate the variations of soil moisture to soil texture, terrain, vegetation, and meteorological conditions, and hence the decisions on weighting the node measurements are solidly tractable, regardless of geographic location.
- It enables, for the first time, a dynamically guided sampling strategy for the sensor web by integrating in situ data, real-time processing, data assimilation, and an optimal control algorithm. The new sampling strategy enables representative estimates of the time-varying field mean provided by space-based remote sensing assets.

The methodology for data collection and data processing described here is also applicable to several other technological areas including transportation systems, wireless sensor networks, and Mobile Ad hoc Networks.

Acknowledgement

This work was carried out at the University of Michigan and at the Massachusetts Institute of Technology through a grant from the National Aeronautics and Space Administration.

References

[1] Dunne, S. and D. Entekhabi, 2005: An Ensemble-based Reanalysis Approach to Land Data Assimilation, *Water Resources Research*, 41(2), W0201310.1029/2004WR003449.

[2] Boni, G., D. Entekhabi, and F. Castelli, 2001: Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation, *Water Resources Research*, 37(6), 1713-1722.

[3] Van Dam, J.C., 2000. *Field-scale water flow and solute transport. SWAP model concepts, parameter estimation, and case studies.* PhD-thesis, Wageningen University, Wageningen, The Netherlands, 167 p., English and Dutch summaries

[4] Reichle, R., D. B. McLaughlin, and D. Entekhabi, 2001: Variational data assimilation of microwave radiobrightness observations for land surface hydrologic applications, *IEEE Transactions on Geoscience and Remote Sensing*, 39(8), 1708-1718.

[5] Song, N. O., and D. Teneketzis, 2004: "Discrete search with multiple sensors," *Mathematical Methods of Operations Research*, vol. 60, no. 1, pp. 1-14.

[6] Meier, L., J. Peschon, and R. M. Dressler, 1967: "Optimal control of measurement subsystems," *IEEE Trans. Automatic Control*, vol. 12, no. 5.

[7] Sonmez, T., and G. L. Blankenship, 2001: "Modeling of target vehicle traffic and implementation of suboptimal sensor scheduling," *Proc. of the IEEE Conference on Control Applications*.

[8] Andersland, M. S., and D. Teneketzis, 1996: "Measurement scheduling for recursive team estimation," *Journal of Optimization Theory and Applications*, vol. 89, pp. 615-636.

[9] Athans, M., 1972: "On the determination of optimal costly measurement strategies for linear stochastic systems," *Automatica*, vol. 8, pp. 397-412.

[10] Athans, M., and C. Schweppe, 1967: "Optimal waveform design via control theoretic concepts," *Information and Control*, vol. 10, pp. 335-377.

[11] Baras, J. S., and A. Bensoussan, 1988: "Sensor scheduling problems," *Proc. of the 27th IEEE Conference on Decision and Control (CDC).*

[12] Chu, M., H. Haussecker, and F. Zhao, 2002: "Scalable information-driven sensor querying and routing for ad hoc heterogeneous sensor networks," *International Journal of High Performance Computing Applications*.

[13] Isler, V., and R. Bajcsy, 2005: "The sensor selection problem for bounded uncertainty sensing models," *Proc. 4th International Symposium on Information Processing in Sensor Networks (IPSN)*, pp. 151-158.

[14] Witsenhausen, H. S., 1971: "Separation of estimation and control for discrete-time stochastic systems," *Proc. of the IEEE*.

[15] Mahajan, A., and D. Teneketzis, 2005: "On jointly optimal encoding, decoding and memory update for noisy real-time communication systems," submitted to *IEEE Trans. Information Theory*.

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

[17] Pierce, L., and M. Moghaddam, (2005): The MOSS VHF/UHF Spaceborne SAR System Testbed, IGARSS'05, Seoul, Korea, July 2005.
[18] Martinez-Fernandez, J. and A. Ceballos, 2005: Mean soil moisture estimation using temporal stability analysis, *Journal of Hydrology*, 312, 28–38.
[19] Cosh, M.H., Jackson, T.J., Bindlish, R., Prueger, J.H., 2004. Watershed Scale Temporal Persistence Of Soil Moisture And Its Role In Validation Satellite Estimates, *Remote Sensing Of Environment*, 92(4), 427-435.