Uncertainty Nested in Uncertainty: Modeling in Sensor Networks

Gregory J Pottie
UCLA Electrical Engineering Department
Email: pottie@icsl.ucla.edu

***Abstract*--Many models are required for undertaking research into physical phenomena using sensor networks: the source characteristics, the propagation medium, the noise and interference, the sensor transducer and its coupling to the media and the communications among the sensor nodes. Moreover, models are also used for various levels of signal processing and networking abstractions. The entire chain from observation to conclusion is subject to uncertainties, not least whether the question the network is supposed to answer is meaningful. We enumerate the sources of uncertainty in sensor networks and suggest a framework for dealing with the different forms.**

I. INTRODUCTION

In undertaking a scientific experiment using sensor networks, the objective is often to determine which of a set of candidate models best describes the physical phenomenon of interest. For example, light patterns on a forest floor play a critical role in determining which plant species will be present and their rate of growth. Photosynthesis takes place when the intensity over a range light frequencies exceeds some threshold and increases until some saturation limit, in a species-dependent fashion. There are also temporal dynamics effects. All of this can be directly measured in individual locations using photosynthetically active region (PAR) light sensors. However, as these can only cover one point at a time, statistical models of incident sun intensity patterns are often used to extend the data to larger regions. One might derive such models for example using camera images of the tree canopy and/or the ground cover, with validation of the resultant model at selected locations using PAR sensors on the ground. Note that a down-looking camera does not actually directly measure the relevant scientific quantity since (1) it measures reflected, rather than incident intensity (2) there are geometric issues and (3) the intensity response of the camera will not match that of a plant. There are many other examples of applications where point measurements must be supplemented by some combination of wider-area means of gathering data which is related to but not exactly what is of interest [1].

More generally, models are required in multiple domains:

- Source phenomena: discrete set vs. continuous, coupling to medium, propagation medium, noise and interference processes
- Sensor transduction: coupling to medium, conversion to electrical signal, drift and error sources
- Processing abstractions: transformation of observations to reduced representations (for processing convenience, dealing with communications and other systems constraints relevance to particular scientific questions, fusion among diverse sensor types)
- System performance: reliability of components, time of storage of data at various levels of representation

In composing a sensing system there are thus many layers of uncertainty [2]:

- Observations (noisy, subject to imperfections of conversion to electrical signals, interference, etc.)
- Model Parameters (weighting of statistical and deterministic components)
- Model (particular probability density function family or in general combination of components)
- Interactions with other system components (procedure for setting objectives, how to report uncertainties)

In most scientific experiments there is a considerable body of prior knowledge that may be drawn upon in reducing these uncertainties. The scientific objective is usually to decide among some finite number of models that will explain the facts to some desired fidelity, with testing of alternative models if the pre-selected ones turn out to be

inadequate and likely generation of new scientific questions based upon the results. The tools available include a body of prior knowledge concerning the phenomenon of interest, a set of instruments of differing capabilities and costs that generate data, a set of procedures of varying complexity and accuracy for carrying out the experiment and analyzing the data, and a community of researchers who can provide critical input on the questions, methods, and analysis. Here we are often assisted by some *reference experiment* which produces results for certain selected cases with high fidelity, but which cannot be extended to the entire hypothesis domain due to reasons of cost or inflexibility of components. We similarly may have components (sensors, models at particular fidelity levels, in general prior information) that are trusted, and others that are not. Thus, we must approach the design of the overall system with components in which we have varying levels of trust (prior information on their reliability or correspondence with reality). As resources are finite, it is of interest to devise systematic procedures that can minimize the use of the more expensive of these resources while still producing convincing results.

A typical experimental cycle is thus as follows. A reference experiment is described in a peer-reviewed journal, which establishes some data points concerning some phenomenon but leaves many questions unanswered. The reference procedure is resource-intensive, but reliable. If it could be deployed on a large scale it might answer many more questions, but the cost precludes this. Thus a new procedure is devised with less expensive equipment and/or use of human resources. It is tested in conjunction with the reference procedure in some limited domain to validate its operation, and then applied on a larger scale. Certain data points are spot-checked with the reference procedure, particularly those that are critical in distinguishing between alternate hypotheses (models). New questions arise, resulting in the desire for additional experiments, and so forth.

Very similar procedures are employed in the design cycle of most new systems. For example, simulations are done because the theory is known only for certain limiting cases, prototyping is performed because the models used in simulation carry many uncertainties, and several stages of testing are necessary due to interactions among hardware and software that are far too complicated to exhaustively model and test. Product goals may change under testing as being either too ambitious or not ambitious enough. Results of the tests are carried forward into theory and models for the next product. In this cycle, relatively little time is spent on the design concept with the vast majority of effort devoted to avoiding known types of error through careful construction of modular interfaces (software and hardware) that also allow re-use of validated subsystems, and to verification of the operations of modules and the complete system.

Given the near-universal use of this design approach for experiments and systems, it is remarkable how little literature there is on quantifying its most basic tradeoff: the deliberate sacrifice of efficiency and design aesthetics in favor of reducing the time/effort to validate for a *sequence* of designs or experiments. What appears to be extremely difficult to characterize is how much the validation effort is reduced in going to a layered design with well-structured interfaces as compared to an all-in-one design for each experimental instance, and at what efficiency or performance cost. This paper will present a framework for approaching this problem, but at this early stage of research no quantitative results.

The remainder of the paper is organized as follows. In section II, we use the example of calibration to illustrate a number of the validation problems in experiment design. In section III, we describe how a modular approach reduces the validation effort when a sequence of experiments is considered. Section IV presents our conclusions.

II. CALIBRATION IN SENSOR NETWORKS

Consider the problem of calibration in sensor networks, in which there are two classes of instruments:

- Static nodes subject to drift, according to some parametric model with stochastic parameters; the nodes are relatively low cost so that we can deploy more than minimum required to characterize physical phenomenon (e.g. a field) if only they were calibrated.
- A mobile trusted instrument (either with robotics or carried by personnel), which can be calibrated with high precision.

Given a cost of deployments, a requirement for reconstruction fidelity of the physical phenomenon under study, and a cost associated with auditing using the mobile instrument, one problem is to find the lowest cost system.

The central component of calibration is comparison of the measurements of a sensor with some reference standard. This reference standard may be a trusted model, a trusted instrument, a data fusion result from a set of instruments that are collectively trusted whatever their individual uncertainties, or some combination of the above. Careful design or modeling can extend the time between calibration events. For example, pressure sensors are subject to thermal effects, but

sensors can be designed that are insensitive to temperature variations over some limited range, with still greater accuracy if temperature is measured and the results compensated. In another example, time standards are typically derived from networks of clocks each subject to random drifts; if the drifts are iid and Gaussian, then it is readily apparent that the average time reported will have a lower expected drift than individual clocks. Where some clocks are subject to malfunctioning, also required is a mechanism to determine outliers and then exclude them from the collective decision. Suppose for example sensor *i* returns measurement M_i, which differs from the true value M_0 due to a linear drift term $D_i(t)$, and some bias B_i, that may be due to it being some physical distance from the source. Then if the drift and bias terms are zero mean iid random variables, the normalized sum of the M_i tends to M_0 as the number of sensors goes to infinity. But of course with a finite number of sensors the drift will eventually dominate, and strategies to exclude the sensors with large bias (e.g. those that are far from the source) can reduce the bias of the fused result.

There are many statistical approaches to measuring how well a set of measurements from similar or very different sensors support or fit a model, or in determining what models may be constructed that are consistent with the data. For the latter, the principal division is between parametric and non-parametric models. In a parametric model, a model is proposed with some parameters that are unknown (e.g. a polynomial or a set of Gaussians with unknown weights). Training data is then used to fit a model according to one's favored statistical measure. Typically, one chooses among a small set of competing models (e.g. polynomials of degree 2 and 3), with relatively modest data sets being required to converge to an answer. In non-parametric methods, no assumption is made as to the type of model. While attractive in requiring less a priori knowledge of the model, this requires in practice orders of magnitude more observations. Since engineered systems and science experiments are usually characterized by a significant degree of prior experience, parametric approaches are generally favored.

Bayesian techniques are common in engineering problems. In this context, a model is selected if it is the most probable given the observations and prior knowledge, among those considered. Bayesian approaches are natural in this setting: uncertainty in the conclusions is directly modeled and a clear rule is given for fusing data among diverse types of sensors. Moreover, there are intimate connections to criteria such as maximization of mutual information between observations and the conclusions [3,4], and least squares techniques. Nonetheless a Bayes formulation of the model fitting problem is almost always subject to some modification to account for (a) imperfect prior information (b) the perception that less complicated models are better than more complicated ones (Occam's razor) (c) intractability in many robust optimization problems and (d) the desire to uncover the underlying causal relationships [5]. Examples of specific procedures are the minimum description length (MDL) [6], BIC, and AIC [7].

Returning now to the problem of calibration in sensor networks, there are three basic approaches that can be considered. The first is to buy self-calibrating instruments for which all the careful compensation for environmental factors has been done. This is impractical for many reasons, not least that there is often no such instrument available. The second technique is to over-deploy with respect to the minimum number of reliable sensors required to characterize the physical phenomenon of interest. Then groups of nodes may cooperate to report results, using some combination of fusion and exclusion of reports, based upon the record of behavior with respect to the consensus of the other members of the group, and some model of how the phenomenon is expected to behave. Reputation systems have been applied to this end [8]. This does not actually result in calibration, but can extend the time period that data might be trusted to be within some range of the true value. The third and essential method is to periodically audit the deployed nodes using a trusted instrument. Given an objective of establishing some statistical measure of reliability, the auditor instrument will not necessarily need to visit the locations of all the sensors in the network on each pass; how often it inspects the network and what trajectory is followed will depend on both the model of physical phenomenon and the drift model of the sensors (as compensated by the redundant deployment).

III. VALIDATION IN TWO-LEVEL MODELS

Now supposing that a successful calibration procedure is in place, an important question in sensor networks is how much testing with reference techniques is required to verify that a new model/procedure produces sufficiently reliable results. To make the discussion concrete, consider again the problem of modeling light patterns in forest. There are two classes of instruments:

- Cameras, which measure reflected light intensity over a large area in a series of rapid snapshots

- Mobile PAR sensors, which directly and accurately measure the quantity of interest at discrete points.

The process varies too quickly to achieve coverage with the PAR sensors at the required density with the mobility means available. The goal is instead to calibrate the camera images to interpret the reflected intensity in terms of samples of the incident intensity. In this case, the mobile element is measuring a related but different quantity than the primary sensor. This is similar to remote sensing applications, except that the instruments that provide the interpretation are moved to regions as required to best answer the current scientific question (could be for average intensity, exploration of shadow regions to find thresholds for photosynthesis or photosynthesis saturation). A sequence of experiments will be conducted to characterize light conditions in different regions, seasons, and weather conditions.

One approach to such problems is to use a multi-level modeling procedure whereby reference models and procedures are available at each level, with higher levels directing validation activity at lower levels where required. A diagram of the procedure for the special case of a two-level model is given in Figure 1.

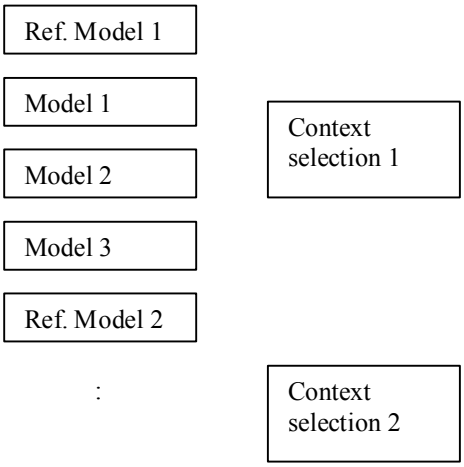

Figure 1. Two-level Model

At the low level (to the left) it is hypothesized that one of several models is the correct one for a given context. For example, different sets of models might be hypothesized for shade compared to full sun. One or more reference models or procedures are available for providing observations that can distinguish among the alternatives for each context. The second level of the model may have a similar internal structure, with the additional element that decisions are made based on some combinations of sensor data from this second level and the first to determine adaptively where additional measurements should be taken, and which context is most likely so that particular models are examined. This allows different tools and models to be employed where they are more likely to fit the problem at hand. Thus for example in full sun it is unlikely that any part of the region will be below the threshold at which photosynthesis saturates, and so little investigation is required. Full sun conditions are relatively easily distinguished from shadow conditions using a camera. On the other hand, shaded regions have considerable biologically interesting variations in intensity and so extensive testing may be done to examine alternative statistical descriptions of the incident intensity, fusing camera and PAR sensor data. Cloudy conditions constitute yet another context for which the entire scene will experience diffuse light.

The multi-level modeling approach enables simple model components to be re-used in different contexts, and can result in simpler validation tasks for each using well-known procedures, as compared to the general problem of matching a one-level model to a highly spatially heterogeneous physical phenomenon. This is because in a sequence of experiments that deal with similar sets of conditions, models and analysis techniques can be re-used, with validation effort concentrated on the regions/conditions for which the prior tool kit yields unsatisfactory results. In this way the tool kit of models, statistical tools, and algorithms can be gradually extended from simple physical situations to more complicated ones. This approach is very much in line with what are termed functional causal models in which each variable of interest is modeled as being a function of some set of parent variables which are explicitly modeled, and some noise or omitted factors. A set of such functions constitutes the overall model. These models may then be nested to produce the higher level abstractions according to the Markov property: each variable is independent of all its non-descendents, to avoid spurious correlations [5].

A number of research questions naturally arise. One way to view this scheme is that a functional model is a *compression* rule for the observation set. This compression is the essential first step to generalizing (i.e., statement of causal relations). The mutual information between observation sets and alternative models may be used to assess which model is the best, and whether a new one needs to be constructed. The two level scheme can be viewed as a problem of compression with side information; the primary sensor is at the first level, and the side information is the sensor at the second level. The question is then given

that the science goal is evaluation of some function of the data and side information what layering will result in the minimum number or cost of observations. This is related to the problem of functional compression [9].

Suppose now that the model correctly identifies contexts and the causal model that is supported by the data. Even in this case, the model can suffer from two difficulties when applied in a new setting: there may be spurious models to check, or there may be a need for new models to deal with new contexts. A causal model that is specially constructed for the new situation can in principle be more efficient in that the correct set of causes lie at its base and the decision graph (e.g. a Bayesian network) implies the minimum resource usage. The model which evolves to this new setting will have baggage of unneeded modules, which however will provide robustness as they have been needed in some prior situations. An interesting research question is how well the evolving model compares to the most efficient model, and the validation effort required for the two approaches across the full range of experimental conditions.

IV. CONCLUSION

The central problem in sensor networks is validation of data and models. By considering that any given experiment is part of a sequence of experiments which seek to build knowledge, it is clear that modular approach can greatly reduce the scope of this validation task. What remains to be characterized is the quantitative tradeoff between reduction of validation effort and model fidelity for this modular approach compared to a custom design.

REFERENCES

[1] Aman Kansal and William J Kaiser and Gregory J Pottie and Mani B Srivastava, "Actuation Techniques for Sensing Uncertainty Reduction," *CENS Technical Report #51*, March 2005.

[2] K.B. Laskey, "Model uncertainty: theory and practical implications," *IEEE Trans. Systems, Man and Cybernetics*, vol 26 No 3 May 1996, pp. 340-348.

[3] M. Ahmed and G. Pottie, "Fusion in the context of information theory," in Distributed Sensor Networks, Ed. S.R. Iyengar and R.R. Brooks, CRC Press: 2005.

[4] J.W. Fisher, T. Darrell, W.T. Freeman, and P. Viola, "Learning joint statistical models for audio-visual fusion and segregation," in *Advances in Neural Information Processing Systems* 13, edited by T. K. Leen, T. G. Dietterich, and V. Tresp, pp. 772-778, 2001.

[5] J. Pearl. Causality: Models, Reasoning and Inference. Cambridge University Press, Cambridge, 2000.

[6] A. Barron, J. Rissanen and B. Yu, "The minimum description length principle in coding and modeling," *IEEE Trans. Inform. Theory*, Vol. 44, No. 6, Oct. 1998, pp. 2743-2760.

[7] L. Wasserman, "Bayesian model selection and model averaging," *CMU Department of Statistics Technical Report 666*, Oct. 1997.

[8] P. Resnik, R. Zeckhauser, E. Friedman, and K. Kuwabara, "Reputation systems," *Comm. ACM*, vol. 43, No. 12, Dec. 2000, pp. 45-48.

[9] A. Orlitsky and J.R. Roche, "Coding for computing," IEEE Trans. Inform Theory, vol. 47, No. 3, Marchh 2001, pp. 903-917.