

Interaction of Deterministic and Stochastic Models

Murali Haran

Department of Statistics
Penn State University

SAMSI Spatial Program Transition Workshop.
October 2010.

Working Group III

List of core members who are still actively involved:

- ▶ Susie Bayarri
- ▶ Jim Berger
- ▶ Murali Haran
- ▶ John Harlim (stochastic dynamics)
- ▶ Emily Kang (postdoc)
- ▶ Pete Kramer+ Sorin Mitran (stochastic dynamics)
- ▶ Hans Künsch

Several others participated in the group including: Sham Bhat, Anabel Forte, Cindy Greenwood, Michael Lavine, Danilo Lopes, Robert Wolpert.

Statistics and Applied Mathematical Sciences

Our working group is unique:

- ▶ This has been a genuine collaboration between the stochastic dynamics program and the spatial program: two applied mathematicians participated regularly, often driving the research problems, with discussions involving five statisticians.

Potential Research (ca. Sep 2009)

We started with a rather large list of potential areas:

- ▶ Approximations, emulation and calibration for complex computer models, typically Gaussian process-based models, inference for SDEs.
- ▶ Bridging gap between “usual” statistics idea of spatial modeling (accounting for dependence) and capturing spatial dynamic processes that are of scientific interest: one potential avenue for interacting between spatial program and stochastic dynamics program.
- ▶ Using deterministic models/process to construct flexible space-time models.

Potential Research [cont'd]

- ▶ Deterministic models as surrogate for probability/process model (“computer likelihoods”).
- ▶ Embedding deterministic models within a hierarchical framework to learn about *scientific* parameters of interest.
- ▶ Developing efficient computational approaches for simulating from multiscale models.

Note that computational challenges are a common thread through most of these problems.

Potential Scientific Applications

- ▶ Climate science. e.g. global climate, glaciology.
- ▶ Simulation of physical systems. e.g. fluid dynamics.
- ▶ Epidemiology. e.g. models for infectious disease dynamics.
- ▶ Biology. e.g. molecular motors.
- ▶ Spread of invasive species, diffusion models (biology, hydrology etc.)
- ▶ Many others...

Specific Project Discussions

Lots of research talks and project discussions in the fall.

- (i) Modeling/inference for a component model for alcohol usage.
 - (ii) Dengue fever modeling based on a data set from Peru.
 - (iii) Exploring mathematics underlying a class of spatial dynamics models.
 - (iv) Studying multiscale dynamical systems.
 - (v) Emulation/calibration for complex computer models for multivariate spatial processes.
 - (vi) Bayesian model averaging of climate models for temperature projections.
- (iv)-(vi): current/completed projects. (i)-(iii): future possibilities.

Multiscale Simulation

- ▶ Heterogeneous multiscale methods (HMMs) for simulation: uncharted waters for statisticians.
- ▶ HMMs: active research area in applied mathematics and scientific computing in the past decade (work by Vanden-Eijnden, Majda, others.)
- ▶ Lots of scientific applications, e.g. fluid dynamics, climate models.
- ▶ New ideas and terminology. E.g. HMM \neq Hidden Markov Models and “parameterization” = representation of unresolved processes by some approximation.
- ▶ There appear to be a lot of open problems, avenues for research. An exciting new area for statisticians.

Multiscale Simulation [cont'd]

- ▶ Want to simulate a physical process with both “macroscale” and “microscale” dynamics.
 - ▶ E.g. of macroscale: global climate model with spatial resolution \approx 25km to 100km.
 - ▶ E.g. of microscale: clouds. Important to try to get this right since precipitation and condensation affect macroscale.

The simulation challenge:

- ▶ Equations at macroscale need information from microscale.
 - ▶ Microscale system needs information from macroscale to provide physical constraints and sensible starting values.
 - ▶ *Microscale simulation is computationally very expensive.*
- ▶ Statistical methods: kernel density estimation, Gaussian process (emulation), filtering, sequential Monte Carlo etc.

Pete Kramer provided a more detailed discussion on Monday.

Classes of Complex Deterministic/Stochastic Models

- (I) (Relatively) simple deterministic/probability models that can be written down/approximated in closed form.
- (II) Intermediate: cannot write models down but can simulate/do forward runs fairly quickly. Maybe use 'likelihood-free' (Marjoram et al., 2002) approaches.
- (III) Very complex: can neither write them down, nor simulate fast. e.g. GCMs, EMICs. Typically build Gaussian process models as surrogates for process models.

Our group has been working primarily on (III) and on multiscale models, a combination of (II) and (III).

Working Group III: 2009–2010 and onwards

Multiscale modeling and simulation, spatio-temporal filtering:

- ▶ Harlim, J. (2010) Numerical strategies for filtering partially observed stiff stochastic differential equations, *to appear in J. Comput. Phys.*
- ▶ Kang, E.L. and Harlim, J. (2010) Filtering partially observed spatially extended chaotic dynamical systems with multiscale features, *in preparation*.
- ▶ The core group is meeting biweekly to discuss multiscale simulation/computation (HMM)-related research. This has been continuing since the spring and we anticipate several publications to emerge from this line of work.

Most importantly, we believe we have opened up possibilities for a new and fruitful line of research.

Working Group III: 2009–2010 and onwards [cont'd]

Complex computer models with multivariate spatial output:

- ▶ Bhat, K. S., Haran, M., Goes, M. (2010) Computer model calibration with multivariate spatial output: a case study in climate parameter learning, *Frontiers of Statistical Decision Making and Bayesian Analysis*.
- ▶ Bhat, K. S., Haran, M., Keller, K., Tonkonojenkov, R. (2010) Inferring likelihoods and climate system characteristics from climate models and multiple tracers, *submitted*.

Bayesian model averaging for climate projections:

- ▶ Bhat, K. S., Haran, M., Keller, K. (2010) Combining multiple climate models using Bayesian model averaging to project surface temperature, *in preparation*.

More papers in preparation related to above topics.

Working Group III: 2009–2010 and onwards [cont'd]

- ▶ Several of our research project proposals will continue through our recently formed collaborations.
- ▶ Others may find their home in future SAMSI programs. In particular, many of the complex computer models problems fit nicely into the “Uncertainty Quantification” SAMSI program in 2011.

Acknowledgments

On behalf of the working group, many thanks to SAMSI for providing the environment and opportunity to explore these new research problems and learn about a host of exciting research problems and methods in spatiotemporal modeling.

Special thanks to:

- ▶ SAMSI staff.
- ▶ SAMSI space-time program committee, including Noel Cressie, Montse Fuentes, Alan Gelfand, Richard Smith, Dongchu Sun, Michael Stein, and Jim Zidek.
- ▶ The National Science Foundation.