

**ICMS workshop on  
Smoothing-Based and Gaussian-Process-Based Methods for Non-Parametric  
Regression in Environmental Problems**

**Date:** March 26-30 , 2007

**Venue:** 14 India St of Edinburgh

The workshop ran largely as proposed. There were some departures from the originally-proposed list of participants, since a) it was still the US term time and b) a research centre in the States was running a 6-month programme which attracted a number of our originally identified US speakers; nonetheless the meeting brought together leading experts in (a) modern smoothing techniques in data analysis; (b) Gaussian process based Bayesian statistical methodology; and (c) environmental statistics.

The objectives were to:

- explore the relationship between smoothing-based and Bayesian-Gaussian-process-based approaches to non-parametric regression and model uncertainty, and assess the potential for transfer of computational tools and theoretical constructs between the two;
- confront non-parametric regression methodology with large scale environmental problems of trend detection and assessment, and of model uncertainty, and to formulate proposals for practical solutions;
- explore the power of non-parametric regression techniques in spatial and spatio-temporal contexts and consider advantages and disadvantages of the techniques with any computational issues.

Discussion themes included:

Theme 1: Challenges in modern smoothing techniques in regression and time series: model-based opportunities?

Theme 2: Gaussian processes and Bayesian methodology: strategies for implementation in high-dimensional environmental contexts.

Theme 3: Environmental science problems and questions needing a statistical solution. Too much data, too little information?

Theme 4: Spatial and spatio-temporal non-parametric modelling

Theme 5: Computational challenges

The recent flourishing of publications and research interest in smoothing techniques and Gaussian processes is a reminder that these methods are evolving into mainstream statistical techniques for the analysis of data from today's vast range of application areas. In particular, the workshop focused on environmental applications.

A total of thirty-five participants from the UK and other European countries attended the four-day workshop. These included a number of PhD students whose research activities are in the relevant areas. The PhD students were all asked to give a short 10 minute presentation to their work, and this was generally considered as an extremely important

aspect of the workshop. Most of the participants are experts in the thematic areas of the workshop and most people spoke for 30 minutes, followed by a series of scheduled 2-hours breakout discussion sessions with a chairman. Coffee breaks were widely used for informal extensions of discussions.

The workshop aimed to provide an overview of the current state of research in smoothing and Gaussian process based approaches to non-parametric regression with applications in environmental sciences. A brief summary of the content is as follows: six speakers gave a series of presentations that reflected new developments in Bayesian and Gaussian process modeling, four speakers outlined the challenges in environmental science; three talks discussed connections between the three main areas of the workshop; and six speakers addressed new developments and applications. For example, Professor Marian Scott and Vic Barnett opened the workshop with a brief introduction to SPRUCE (co scientific sponsor) and to the themes of the workshops, these being then followed by more technical introductions to smoothing, Gaussian process and the nature of environmental problems. Mark Hallard (SEPA), Gavin Simpson (UCL- Environmental change) and Jan Dick and Peter Levy (CEH) all spoke about their needs as environmental scientists.

. The feedback from questionnaires was uniformly positive, as is illustrated by the selection of responses below:

• **What, for you, was the highlight of the workshop?**

\* listening to the environmental scientists presenting the many problems and challenges which face them. It was clear to me that the inferential methodology which I work on had (and already has) an important role to play.

\*The discussion sessions after the talks.

\*Interaction with statisticians hearing about the problems from their perspective;  
\*Meeting scientist from other fields related to environment, who use statistics and probability for their problems.

\*The successful integration of two/three different communities of statisticians/environmental scientists and computer scientists. Interesting discussions (often over coffee).

\*The break-out discussion groups were very interesting, as were the larger group discussions on the first and last days.

• **What was your impression of the overall academic value of the workshop?**

I found that it was very interesting to meet with statisticians, Bayesian GP modellers and environmental scientists and there was a very good grouping of these three areas – open discussion between these three groupings was extremely valuable from an academic perspective.

\*I think it was particularly valuable to bring together the two different technical approaches of standard nonparametric regression and Gaussian processes.

\*Very high - integrating statisticians and users was a brave move and very worth while

\*The scientific level was very high and focused. The programme was very well constructed.

\*Very much the exchange of ideas/ identifying similarities and differences, even accounting for the different terminology.

\*It was indeed of great academic value. The interaction between scientist from different areas (ecologist, environmental scientist, mathematicians and statisticians) is really helpful for understanding the real problems. I think that these kind of meetings is the only way of realizing which kind of techniques need further development.

\*The presentation and the discussions were of a good academic level and, which is more remarkable, near from the real problems, with applications and taking account the numerical constraint.

\* First class; the interaction of individuals from so many different specialities and professional bases (universities, research institutes etc) lead to stimulating interchange especially in the discussion sessions and prompted many to re-examine their standpoints in relation to some of the key problems.

**• Did the workshop help you to develop/sustain contacts likely to result in new research?**

\*There are no immediate plans for new research as a result of the workshop. However, it is possible that new research avenues will be pursued in the future.

\*Very definitely - well worth while

\* some very useful contacts made

\*I met a lot of people who I had heard of from within my field, and there is potential for collaboration in the future.

\*Yes; I met some really interesting people. Some I knew already, some I had never met. At present we are planning some collaboration using our data assimilation methods to new areas in environmental monitoring of large water bodies.

\*Yes. I was able to sustain a number of existing contacts in the field of environmental statistics, and felt that I was able to develop my contacts with the statistics group at Sheffield University.

**What in your view are the key future research areas / directions in your field?**

Solve the trade-off between accurate models and reliable/fast inference.

\*The workshop identified that there are possibly useful lines of research in trying to combine the different approaches of nonparametric regression and Gaussian processes.

\*Ability to compare landscapes with a element of statistical uncertainty in our estimates.

\*Numerical methods for dealing with high dimension which was amply mentioned during the workshop

\*Extremal modelling of temporal-spatial data.

\*In the environmental modelling area, it seems necessary to add stochasticity to process-based models. SDE type models (e.g. as discussed by Manfred Opper) seem hold promise.

\*It seemed to me that apart from the climate change arena (and meteorology), the Gaussian process techniques have not been taken up widely. Some new areas (such as carbon cycle modelling) were identified, but here the techniques have not quite caught up with the complexity of the system being modelled. There is still some scope for work on dealing with more complex environmental situations.

\*As a statistician, I've collected lots of great ideas for further research. I would say that a key topic from the statistical point of view is the construction of complete models for spatial (or spatial-temporal) varying processes, that describe the large and the small scale variations.

\* The key direction is to make GP methods available to all scientists by teaching them what it means to use this type of model.

\*Critical issues for me are to address the correct probabilistic treatment of physically based models, and putting real effort into relating these models to reality.

\*So many! Forthcoming involvements for me personally and for SPRUCE may be likely to be directed to problems in environmental sampling and monitoring and in environmental standards.

\* I think that the "machine learning" practitioners should consider more the data they are dealing with - similarly to the methodology presented by some of the speakers.

- Development of statistical techniques for the analysis of output from complex stochastic (rather than deterministic) environmental models;

- Development of methods to deal with rare, unexpected and extreme environmental events;

- Closer linkage between areas of statistical research, conceptual and technical developments in the environmental sciences, and the use of this science to provide an evidence base for environmental policy.

- Environmental problems from different views, different methods and techniques and from different fields, but with similar main goals

- Comparison of techniques in terms of complexity, flexibility, prediction or exploratory power.

**Did the programme result in new ideas or the acquisition of new techniques or methods?**

\*I feel that I gained a much improved understanding of the Gaussian process approach to smoothing and that I now understand rather more clearly the machine-learning approach to problems. This is very helpful.

\*It led to future collaboration which will hopefully lead to the acquisition of new techniques

\*Yes. Above all, the great importance of Gaussian processes techniques, which were quite new for me.

\*I am not sure that it did as the group was so diverse.

\*For me it helped to remind me of the breadth of work going on (one often gets too focussed on ones own area) and I think helped link together various aspects. I did learn a few new things too.

\*New directions of thought.

\*Methodologically the talks helped a lot to get a clearer picture of the problems / solutions / research directions.

\*I gained a lot of background knowledge about Gaussian process methods, and about how these methods can be used to analyse outputs from complex models. I have encountered quite a few of these ideas before, but the intensive nature of the workshop gave me a much clearer understanding of the advantages and limitations of these methods, and of the practical issues involved in applying them to real data.

\*I received a number of interesting suggestions regarding a project that I am currently working on - analyzing the impact of climate uncertainty upon outputs from a complex ecosystem model - many of which will feed directly into follow-up research work that we do in this area.

### **Facilities and staff**

These were excellent.

\*The facilities were fine and the staff were very helpful. The smallish size of the venue, and its interesting history, helps to make the atmosphere of the workshop friendly and intimate.

\*All excellent

\*Excellent facilities. Ideal place to organize a workshop.

\*For a meeting of this size, India St worked very well. All admin seemed to be handled very smoothly. Staff were very efficient and extremely helpful.

\*I just can say that everything was perfect.

\* Everything was very well organized and we get all the information on the website of the workshop to find accommodation.

\*I loved the ICMS; it has such a great history; I wonder whether the break-out rooms could have been a bit more equipped - I am a whiteboard fan .... but they were not in all rooms? Hard to keep the character I guess and have all mod-cons.

\*The facilities at ICMS were excellent, and the administrative arrangements all seemed to run very smoothly. The vegan lunch that was provided to me on the Monday was fantastic, and probably the most substantial lunch that I have ever had an academic meeting. I did not stay in accommodation, so I cannot comment on that.

\*A well run meeting integrating a range of people with diverse requirements - not an easy task but on this occasion well executed

\*I would like to congratulate the organizers and the administrative staff for their great job.

\*Mainly thanks for organising/hosting it. I enjoyed the opportunity to mix with people from a wide range of backgrounds

Finally the proposers wish to extend their sincere thanks to ICMS for its guidance on the workshop proposal and for making such a valuable event possible. We would also particularly like to thank Morag Burton for her administration of the workshop.