Final Report on Workshop: 17w5120 - New perspectives on State Space Models

Marie Auger-Méthé (The University of British Columbia),
Dave Campbell (Simon Fraser University),
Len Thomas (University of St. Andrews)

August 28, 2017–Sept 1, 2017

1   Introduction

This workshop brought together 29 participants from 4 continents for 5 days of research and discussion. On Monday, all participants gave 5 minute flash talks where they briefly discussed their work, interests, research, and concerns. The workshop schedule provided the flexibility for break out group discussion as topics arose. Those discussion groups enabled deeper exploration of ideas than a standard conference format. The rest of the document is a description of the topics presented and discussed throughout the workshop.

2   Monday

2.1   Opening Talk

Dave Campbell gave an overview of the area of State Space Models including recursive and distributional properties and implications of the basic linear model with observations Y_t of the process X_t with parameters α, β:

$$Y_t = \alpha X_t + \epsilon, \quad \epsilon \sim N(0, \sigma_\epsilon^2) \tag{1}$$

$$X_t = \beta X_{t-1} + \delta, \quad \delta \sim N(0, \sigma_\delta^2) \tag{2}$$

These models are used in diverse applications including animal movement models where animal locations Y are observed with error, and their behaviour X changes according to a process with noise. SSMs are also used for stock assessment where the number of animals Y are observed with noise and the life evolution is the process.

Typical methods for parameter estimation include maximum likelihood but the complete data likelihood:

$$P(Y, X \mid \theta) = P(X_0 \mid \theta) \prod_{t=1}^{T} P(Y_t \mid X_t, \theta) P(X_t \mid X_{t-1}, \theta)$$

is intractable. Maximum likelihood is greatly simplified by first integrating out the $X(t = 0 : T)$ values giving the observed data likelihood:

$$P(Y \mid \theta) = \int_\chi \ldots \int_\chi P(X_0 \mid \theta) \prod_{t=1}^T P(Y_t \mid X_t, \theta) P(X_t \mid X_{t-1}, \theta) dX_0 \ldots dX_T,$$

however this requires high dimensional integration.

The main strategies for achieving this integral include approximating it via Laplace Approximation or Monte Carlo integration either by MCMC or particle Filter.

2.2 Flash Talks

Much of Monday afternoon was spent in flash talks where attendees described their work, interests and background as it relates to State Space Models.

Vianey Leos-Barajas discussed Bayesian model monitoring in dynamic modelling comparing the predictive ability of the model to the new observations to identify structural changes in the process model

Nicholas Michaud discussed epidemiological Susceptible Infectious Recovery models for forecasting disease outbreaks. He discussed a hierarchical model for dealing with annual disease outbreaks with similar but different parameters in successive years. He is interested in knowing how to use up-to-date but biased data along with unbiased but lagged data.

Israel Martinez Hernandez, discussed long memory time series and functional data analysis. Dynamic factor models such as functional auto-regressive models and Poisson processes based on functional dynamic factor models.

Diana Cole discussed identifiability. Her example state space model highlighted parameter redundancy and methods for detecting lack of parameter identifiability through symbolic computation and differential algebra. She specifically mentioned a paper which wrongly say that they can estimate a parameter based on prior vs posterior, but they actually can't. She also advertised her upcoming talk and tutorial sessions on Thursday.

William Aeberhard works on robustness. In New Delhi they have particulate matter observations every 15 minutes. The intent is to place monitors across the city to obtain better spatial data. He is working on optimal experimental design for deployment of monitoring devices. Where to place the observations to obtain a good estimate will depend on what you would like to do with it.

Carolina Euan works on clustering time series which involves defining similarity metrics that are relevant to time series structures. She uses the spectral density for stationary methods and coherence (correlation in spectral domain) She is using dynamic (windowed) clustering to look for time varying clustering model.

Juan Morales discussed landscape heterogeneity, animal movement models, and seed dispersal. Movement is affected by heterogeneity of the landscape. Juan is interested in improvements in animal tracking models. He is using wearables, accelerometers and location monitors to examine behaviour and population dynamics. Accelerometers are at 10Hz on 3 axes and GPS location data every 5 minutes. Devices are strapped onto sheep who are re-captured monthly to examine their condition. The issues of multi-resolution approaches came up several times throughout the week.

Christoffer Albertsen discussed fisheries stock assessment where the normal models do not necessarily make sense. Model validation is challenging, there are limits to Laplace approximations. He discussed the complexity of big data, which in his context involves having a space-time grid, a species grid, and an age structure grid. If the state is a radial movement process then a Laplace approximation will fail unless the process is reparameterized into angle and distance rather than euclidean distance. Multimodal states also cause problems.

Daniel Dinsdale discussed his work on preferential sampling. The observation locations may be due to the factor you wish to measure. Ignoring that relationship will introduce bias. For example, estimating the pollution level over an area will be biased in some locations if we only measure near the pollution source.

Ruth Joy works on high frequency time series models where the observed state is the location of the animal but interest is in the actual animal behaviour. She discussed how to incorporate the fine time scale behaviour into the larger scale time-location behaviour. She has worked on identifying foraging hotspots and their overlap with commercial fisheries. Her interests include parameter estimation rather than state,

limitations in types of data compared to the actual interest. She is also interested in animal behaviour and matching it with complex observation types and scales. She also deals with wind speed distributions over the sea and opportunistic fishing data.

Ethan Lawler discussed his work on age structured populations. Fish are modelled by age, but observed by length. They use Age-Length keys to map fish length into fish age. More generally, his models involve indirect proxy observation of variables. He has data from a wide variety of sources to incorporate into models.

Kim Whoriskey is working on animal movement models. She is using a hierarchical model and is interested in population level parameters and individual level parameters. She is interested in switching the implementation from state space models to a Hidden Markov Models or hybrid schemes thereof. The main current challenge is combining approaches that will allow the use of Switching Hidden Markov Models. She uses multiple data sources in her models.

Dave Campbell discussed diagnostics for Laplace approximation via probabilistic integration. His goal is to assess the validity of the Laplace approximation via a method that will numerically integrate the high dimensional surface over a coarse grid. The method is slower than Laplace Approximation, but faster than alternatives yet carries forward a measure of uncertainty in the numerical integral.

Sofiá Ruiz Suárez is working on demographics of movement and internal variables of animal movement models. She has been fitting models using an Approximate Bayesian Computing approach. She has a variety of variables. One of the ABC issues is the scale of the summary statistic.

Aaron King talked about how we don't observe the same effectiveness of math in biology compared to other sciences. He discussed the trade off between fidelity to the mechanism and tractability of the model. Aaron is scheduled to discuss his R software package *pomp* later in the week.

Ken Newman discussed his work in state space models in the context of fisheries management decisions. He has worked on age based life cycle models, abundance observation models, and Bayesian inference. Model issues include overfitting the data. He discussed using the state space process noise as a model fitting diagnostic. In some cases his interests are in probabilities of extinction or population viability. His group have also used RJMCMC as a tool for model selections.

Mike Dowd gave an overview on his work on large scale nonlinear dynamic models. Large scale models require approximation methods. He has done work on state space models for fisheries in Bayesian models, oceanographic and hydrodynamic models, and animal movement models. In particular he has considerable work on particle filtering.

Len Thomas works on the interface between statistics, ecology, and computing. He discussed his work on modelling population dynamics of grey seals from airplane images. He has incorporated age structure, fecundity, and seal mobility in the process model along with an observation model on population counts. He has done model selection via integrated likelihood and posterior model selection. He has also worked on animal movement models and developed a quick and dirty reconstruction of animal movement and tracking.

Fanny Empacher is starting to work on improving efficiency of methods by combining model fitting methods. She has worked on comparing different Kalman filter variants for estimating state space models.

Gonçalo Ferraz discussed space time models in ecology and moments of biological change. He discussed a study of birds repopulating a region where observations were taken through a marked recapture process. He has also been working on tracking mosquitos to monitor occupancy dynamics as part of a pest management system. There is a huge variety of data quality and different data sources. He wants to estimate abundance and uncertainty.

Marc Genton discussed his work on functional box plots and Gaussian Random fields.

Edward Ionides discussed his experiences in teaching state space models. He has a dozen graduate student groups from a course that he taught that completed state space model projects with code and examples on github. He discussed some research work on the trade off between computing time and precision for some profile likelihood work.

Marie Auger-Méthé works on how environmental changes affect the behaviour of animals. She discussed work on animal movement models and noisy observations. The marine animal tracking data is highly error prone. She uses state space models to reconstruct the location path of the animal. Some of her recent work has related to moving models from a Bayesian MCMC framework into a Maximum likelihood methodology. She is also working on models involving discrete states and constraints such as avoiding fish location estimates that are out of the water.

Giovanni Petris discussed Functional State Space Models for stochastically evolving elements of a function space. The state process is a continuous smooth evolving process. The smoothing and filtering recursion carry over to the functional case. These functional state space models allow a straightforward way to incorporate continuous time properties and discrete time observations.

Matías Salibián-Berrera discussed his background in robustness and recent work on animal movement models.

Ying Sun described her work on environmental statistics methodology. Her goal is to provide fast and reliable estimates of spatial temporal patterns of expensive computer models. She discussed rainfall models as inputs into agricultural models. Rainfall is seen and felt but difficult to accurately measure. Precipitation has amount and occupancy models. Her rainfall data is very high frequency. She discussed a spatial random field model with a truncation that splits the dynamics into a binary process. She has an observation process of rainfall and an underlying autoregressive spatial-temporal process model. Her current interests are in model diagnostics.

Andrew Edwards works on fisheries stock assessment. He will be discussing more about Empirical Dynamic Models on Thursday. He discussed how model output is used in practice by model managers.

2.3 Open Discussion

This session focused on challenges and interests from within the group. This was used to devise some break out sessions that will run in parallel to other break out sessions.

3 Tuesday

3.1 Christoffer Albertsen: Introduction to the Template Model Builder (TMB) package

Christoffer Albertsen gave an introduction to the Template Model Builder (TMB), a software library for R that calculates parameters and states from SSMs. He walked the group through examples with code. TMB uses C++ code to compile the model. The model is differentiated using operation derivatives (automatic differentiation) within the C++ functions combined via the chain rule. Part of the key to speed in TMB is the way that it turns the model into a graph of operations and acts on them in different ways. The operation graph and derivatives thereof are determined only once. TMB figures out sparsity in matrices and uses appropriate matrix shortcuts. TMB uses Laplace approximations to marginalize out the random effects or nuisance parameters.

Christoffer showed several examples beginning with a state space model based on estimating the temperature of a beaver through various (latent) activity regimes. He then walked the group through an animal movement model from ringed seal GPS data, several more introductory examples and another real data example for the spatial distribution of deer and coyotes in Banff National park. In the advanced section Christoffer showed us how to parallelize code, simulate from the model, include our own features, define our likelihood and observation model, and fit Gaussian Markov Random Fields. Additional experimental features like checking the Laplace approximation via average gradient method were also demonstrated.

3.2 William Aeberhard: Robust state-space models

William discussed the different meanings behind the term robust. He discussed robustness to model misspecification, robustness to unusual data points, and robustness to assumptions. He discussed a bias correction method for robust estimation of parameters and dynamics from state space models. He later moved into current challenges including problems of asymptotic normality, inference, and standard errors. He showed how to down weigh the impact of highly influential data points which do not otherwise fit the data and correct for the induced bias in the likelihood.

3.3 Lunch discussion of prior specification

Discussion was led by Gonçalo Ferraz. Some of the issues addressed include how one determines if a prior is uninformative, and whether or not such a prior exists or should be used at all. Discussion addressed mainly philosophical issues between practitioners, modellers, and statisticians. Discussion continued about how informative the priors on parameters are when acting on the model.

3.4 Michael Dowd: High dimensional applications of state-space models

Mike Dowd described combining data and oceanographic models. The State Space Models he described involve the ocean chemistry and are closely related to problems in weather prediction. His goals include model selection, estimating the state and its parameters, and predicting future states. The specific process equations he described come from partial differential equations which are discretized over a grid of 180 by 82 cells over 36 depth levels. This example has 7 geochemical variables to predict as state variables. This model must meld together several different observation types including satellite images and information from ships and buoys. More recently new data comes from ocean gliders. With simpler data and simpler models he would be able to compute the full hierarchical Bayesian framework. To make estimation feasible in the current context he needs to treat the numerical model as a deterministic system and typically apply a Data Assimilation strategy based on the ensemble Kalman Filter. He ended with discussion about Approximate Bayesian Computation strategies and additional Model Approximation strategies using computer experiment emulators.

3.5 Break Out Discussion: High dimensional SSMs

This breakout group discussed Approximate Bayesian Computation. This seems a promising approach but is necessarily application specific and subjective. The role of emulators for SSMs were discussed as a promising strategy for sample generation but we need to quantify the approximation errors. Particle smoothing is usually the real target for most SSMs (i.e. retrospective analysis). But note inference for parameters is accomplished with filtering, and filtering is the pre-cursor for smoothing.

Discussion about Particle filters began with focus on SIR, but other strategies are often better for look-ahead filters. The group discussed how to resample in a particle filter. Additional complexity may involve integrating animal movement models within ocean models, using the Ensemble Kalman filter in SSMs, and multiple model inference via particle filters and otherwise.

3.6 Break Out Discussion: Evaluating the accuracy of the Laplace approximation

This group began by discussing where the Laplace approximations is used. There was some question about whether or not the condition of $E_X \left(\frac{dP(Y, X | \theta)}{d\theta} \right) = 0$ is sufficient or necessary for evaluating the efficacy of the Laplace approximation.

There was some discussion about parameter types and interpretations. In the context of SSMs, the dimension of integral increases with sample size. Laplace Approximation is asymptotically exact under the condition that the data is growing faster than the dimension of the system, whereas in SSM models that condition is violated. The lack of asymptotic consistency implies a bias that in many cases will in not vanish with sample size. For the Laplace approximation to be reasonable, the joint distribution needs to be checked. It must be unimodal, symmetric, and have 'non-heavy' tails. The group wandered if there is a cheap way of checking assumptions.

Alternative methods of evaluating the marginal likelihood perhaps involve a full scale particle filter or a probabilistic numerics based approach.

There was some further discussion about software and confounding issues between software breaking and the Laplace approximation breaking.

3.7 Edward Ionides: Iterated filtering and overdispersion for discrete SSMs

Ed described a new iterative filtering algorithm that incorporates a particle filter to marginalize out the states X while slowly decrementing a time varying perturbation on the model parameters θ. This method performs well in a variety of contexts including epidemiological scenarios. He gave as an example a Cholera epidemiology model operating in continuous time. It was noted that if you take a continuous time model and discretize it you could be destabilizing the model considerably. The observations were monthly reported cholera deaths, whereas the infections and death occur at much faster time scales. The presentation continued by highlighting how over-dispersion can change parameter estimates as well as their uncertainties. This is in contrast to generalized linear models where over-dispersion typically only affects uncertainty.

4 Wednesday

4.1 Aaron King: Introduction to the POMP package

The pomp library for data analysis using partially observed Markov process (POMP) models. Algorithmically the library uses calls to some of:

- *rprocess* draws from (2)

- *dprocess* evaluates (2) for a given $X_t \mid X_{t-1}, \theta$ or set thereof

- *rmeasure* draws from (1)

- *dmeasure* evaluates (1) for a given $Y_t \mid X_t, \theta$ or set thereof

- *initialize* draws a sample from $P(X_0 \mid \theta)$

The data is passed into the pomp platform, and then a model can be defined (optionally using C) within R. The user defines some or all of *rprocess, dprocess, rmeasure, dmeasure*, and *initialize*. Depending on the components defined, different methods are available for implementation. As methods are applied, the pomp object is modified allowing different types of output and actions.

Aaron showcased how to simulate from the state process for the Ricker model. He went into detail with a Susceptible, Infectious, Recovered epidemiological model, showing how to simulate the discrete process from the continuous model, how to add in covariates, as well as some general tips and tricks. The pomp package can handle a variety of methods including the multiple iterative filtering (MIF) algorithm that Ed Ionides described the day before.

For the MIF algorithm, in order to evaluate the log likelihood of the model at the end of the algorithm we can not just extract the likelihood from the results. Instead we need to run an additional particle filter because the MIF model is slightly different from the target; MIF has dynamic parameters rather than static ones as the specified model implies.

The *probe* function allows you to simulate data and produce summaries thereof to allow you to make comparisons. A consequence thereof is that probe matching could be used for a synthetic likelihood approach. Furthermore this enables one to construct Approximate Bayesian Computing algorithms.

4.2 Break Out Discussion: Advanced POMP

This section examined pomp source code. Several examples were shown along with implemention in R and C while highlighting design considerations.

4.3 Break Out Discussion: Quality of Laplace Approximation Part 2

This group discussed ways of incorporating models written in TMB (or pomp) and applying them in the opposite software. This became a recurring theme in later discussions.

There was discussion about adding higher order terms to the Laplace approximation but that would miss structure that occurs far from the single point around which the Taylor expansion occurs.

The group discussed the idea that during the incremental optimization process, you should be checking the quality of the Laplace approximation for every θ. The problem that was mentioned is that there may exist a point in the likelihood surface for which the Laplace approximation is bad and therefore the approximated likelihood is under estimated. However, at another point on the parameter space, the Laplace approximation may be exact and therefore not underestimated. At this second point, the likelihood may evaluate at a higher value and therefore be selected as the optima despite being sub-optimal in the non-approximated likelihood.

Discussion took a brief tangent into Bayesian vs Frequentist methods.

4.4 Discussion session: Diagnostics for state-space models a posteriori model selection

Diagnostics depend on the goal of the analysis and should be tailored to the inferential needs.

The group discussed the use of WAIC for model selection as well as informal model selection via minimization of the process noise term σ_δ^2. In the case of nested models, one could split parameters into the product of a binary indicator variable and a continuous valued parameter. This is equivalent to having a spike and slab prior in the Bayesian context and using Elastic Nets in the frequentist LASSO literature.

The pomp software package probe tool allows you to simulate data and see if summaries of the simulated data match summaries of the observed data. This is akin to using Approximate Bayesian Computing or Synthetic Likelihood methods for model checking based on using only the fixed $\hat{\theta}$ values returned as the maximum likelihood estimator.

Model checking can also be performed using the step ahead prediction cross validation or applying cross validation blockwise to removed sections of the time series.

5 Thursday

5.1 Diana Cole: Parameter identifiability

Diana discussed the different ways in which parameter identifiability can arise. If parameters are confounded: $Y = \alpha\beta X + \epsilon$, then α and β can not be uniquely identified. This type of parameter redundancy can be assessed using symbolic methods and differential algebra. Those methods involve taking derivatives of the likelihood and making substitutions so as to do an expansion exhaustive summary.

Practical identifiability can be assessed using profile likelihoods. In the above example that implies plotting $P(Y \mid \alpha, (\beta, \sigma_\epsilon^2) = argmax(\beta, \sigma_\epsilon^2))$ with respect to α and seeing if the resulting profile likelihood is flat. Flat likelihoods imply ridges in the parameter space.

To assess parameter estimability one must estimate parameters and then examine the curvature of the likelihood, i.e. the Fisher Information matrix. The Fisher Information matrix is the asymptotic estimator of the variance. If the Fisher Information matrix is not full rank then parameters are not estimable.

Diana guided groups through 3 parallel identifiability tool sets: Data Cloning based approaches, differential algebra, and profile likelihood. These well designed hands on tutorials included software and led to discussion about strengths and weaknesses of the different strategies. Groups performed analyses and experimentally examined the cases in which the different strategies apply.

5.2 Break Out Discussion: Parameter identifiability

Within this group a specific SSM example was examined where the differential algebra approach said that all parameters were identifiable, but in fact the likelihood is multi-modal and therefore only locally identifiable. This type of identifiability is difficult to determine in practice.

Since there are many reasons for lack of identifiability, there is a need for multiple methods and strategies.

5.3 Break Out Discussion: Multi-temporal scale models; continuous vs discrete models

This group discussed mismatches in time scale of observations, such as accelerometer measurements every second and other gps location measurements occurring every 5 minutes. There was some discussion about

the use of Hidden Markov Models to model switching between behaviour types. In many cases fine scale behaviours inform large scale behaviours. There may be some summary statistic about the behaviour such as collapsing the signal via summaries or perhaps functional data approaches that could be used in order to deal with multi-resolution information.

5.4   Lunch Discussion: Data to alleviate problems of parameter redundancy

Collect or include different kinds of data via additional publications and /or other experiments. Reformulate the SSM to make parameters estimable via reparameterization. Enforce constraints in the parameter space.

In the special case of distinguishing between σ_y^2 and σ_x^2, obtain repeated measures. Calculate estimates of observation noise externally and possibly increase that estimate when plugging into the observation process because what you estimate in the lab may be too optimistic for real world data. Alternatively, place probability distributions on the variance terms (or use fully Bayesian methods.)

Produce a simpler model.

There were questions about whether or not the presence of a method of moments estimator is a sufficient condition for parameter identifiability.

Examination of the adjoint sensitivity equations is another potential avenue of research in this area.

5.5   Andrew Edwards: SSMs in management & alternatives to SSMs (Empirical Dynamic Modeling [EDM])

Andrew discussed statistical catch-age models for rock fish and hake. He described the process of modelling and sharing results with managers and stake holders. He showed the state space models that are used for hake which include age structure dynamics, constant mortality, and random recruitment. Hake has excellent data along the entire west coast of the US and Canada. This allows age distribution information.

Currently they use ADMB for estimation but have recently recoded the problem into TMB. While TMB allows for maximum likelihood estimation, in the long term, it also permits incorporation into Hamiltonian Monte Carlo because of the way that the derivatives are calculated in that software.

The current stock status, recent trends, and biomass projections, feed into determining the allowable catch. State space models are needed to account for observation error and process error. Giving explicit uncertainties is appreciated by fisheries managers.

Andrew then shifted to a talk about Empirical Dynamic Modelling (EDM). The EDM approach uses a non-parametric, 'equation free' approach to exploit information from the previous times that the data was in a similar state. Data are then projected forward in time by considering where the data moved in similar situations in the past.

EDM is based on Takens embedding theorem which states that lags can substitute for unobserved variables. Takens embedding theorem provides a system in which a $X(t)$, $X(t-1)$, and $X(t-2)$ can be used to produce a manifold with a one to one correspondence to the full systems comprised of $X(t), Y(t), Z(t)$. Since X, Y, Z are not all observed we can use the 'shadow manifold' of X based on $X(t)$, $X(t-1)$, and $X(t-2)$ as a surrogate for the dynamics. EDM is a predictive estimation strategy which takes the nearest neighbours within the symplex of the system in lagged space and projects the point forward based on the average location of the projected nearest neighbours. Sugihara et al (Science 2012) provide a video supplement explaining EDM.

This approach was recently applied to salmon on the west coast in a 2015 PNAS paper. Data include spawners, recruits, sea surface temperatures, the Pacific decadal oscillation, and Fraser River Discharge. The data span 1948-2010. The sockeye salmon come back every four or five years, making them nearly distinct, parallel time series.

With long time series, Andrew showed some results about detecting direction of influence between multiple time series using an example of sardines and sea surface temperature. The methodology seems easy to implement via rEDM.

5.6 Breakout Group Discussion: Fitting space-time dynamical models

Several papers were discussed. Many people have worked in this area, but it remains a difficult problem. The consensus seems that one needs to approximate. Particle filtering methods were discussed in this context. It's not clear theoretically how to do space time filtering. Smart proposals were discussed as was the curse of dimensionality including a review of a paper that shows how to circumvent dimensionality challenges.

5.7 Breakout Group Discussions: Fitting methods and model formulation decision trees

Discussion about which methods to use depend heavily on the goal of the analysis and the available data types. There was considerable discussion about the structure of the advice that we could provide. The cost is an important factor in making methodological decisions, whether it be computation cost, overhead in learning new methods, and modelling expertise. Model formulation is an important element in the selection of appropriate methodological tools. The two concepts are therefore difficult to separate. The group began gathering ideas with the goal of formalizing their advice. Discussion continued throughout the week.

5.8 Breakout Group Discussion: Empirical Dynamic Modelling

The group worked through code and data examples using empirical dynamic models. They discussed how it would work on real data examples including stochastic systems. When there is process noise, the dynamics will not be embedded in a finite dimension and Takens theorem does not apply. The group tried some disease data and models and were unable to estimate extreme events. The inability to estimate extremes such as extinctions and extreme high values is intrinsic to the EMD approach because it depends on the average of a simplex of points that will not allow predictions outside the scope of previous observations.

6 Friday

6.1 Nicholas Michaud: Introduction to the NIMBLE Package

Nimble is a R library for building hierarchical models. Nimble allows you to use BUGS language as programmable objects in R, allowing you to turn BUGS code into objects that can be used in a wider variety of algorithms in R. As a result, Nimble increases the library of BUGS algorithms and models allowing fine tuning of MCMC details, Sequential Monte Carlo methods, and more.

You can compile a nimble model into C++ code without making any other changes. This will let you run your modeling in R to debug errors and then compile it into a much faster running model.

In the SSM example code, the following samplers were automatically selected random walk sampler (Metropolis Hastings), conjugate distributions, and posterior predictive samplers (this is only used for terminal stochastic nodes that can be simulated directly). Nimble by default samples parameters one at a time but offers the flexibility to sample parameters through block samplers and Metropolis Hastings could be performed on the log scale for a variance parameter so as to avoid proposing values which are negative. In nimble you use the *configureMCMC* function to select specific sampler strategies. Nimble also has slice sampling, and automated factor slice sampling which performs slice sampling in the eigen space of the parameters. This amounts to being automatic reparameterization after learning the correlation structure of the data.

The R code that nimble generates is quite slow and is mainly used for debugging. The C++ code it generates is very efficient. MCMC diagnostics can be performed using *library(coda)*. Once you have completed your MCMC, you can extract the WAIC for model comparison.

Several examples with code were shown. The nimble development team is building in automatic differentiation which opens the doors to a wider class of methods and samplers.

Nimble is easy for BUGS and JAGS users to learn and exposes them to a wider variety of methods. There was some discussion about how a nimble model could be used as the input to pomp.

7 Discussion: Next Steps and Lesson Learned

Discussion began with combining packages so you can recycle code sets to access more methodologies. This seemed like it would not take a lot of additional effort.

There was some discussion about modernizing older code to more efficient packages. Multi resolution approaches were mentioned as an ongoing direction of interest.

Incorporating functional responses as inputs or outputs of a system would be insightful.

In some cases discrete models are being used for continuous processes whereas continuous time processes may be more appropriate. When incorporated into a multi-resolution approach some parameters may be heavily influenced by having the higher frequency data than the other data type.

There was some discussion about the diversity of model fitting options available.

In some cases SSMs are called hierarchical models but those fields would be better served by the methodological literature if they knew there is another name for the process.

Empirical Dynamic Modelling was discussed as an interesting and potentially fast way to produce predictions for special goals.

The applied code sessions were well appreciated. It would be useful to marginalized some parameters via Laplace and others via particle filter.

More attention should be placed on marginalizing parameters analytically, perhaps as part of the process of assessing identifiability.

The software package participants were all offering help with getting methods to work and asking for information about which features would be useful to include.

There was interest in writing a review paper outlining methodological choices and diagnostics with attention to the cost function both in goal and utility.