Annals of the Institute of Statistical Mathematics (2020) 72:41–44
https://doi.org/10.1007/s10463-019-00744-0

INVITED ARTICLE: SECOND AKAIKE MEMORIAL LECTURE

Reply to Discussion of "Bayesian forecasting of multivariate time series: scalability, structure uncertainty and decisions"

Mike West[1]

Received: 16 October 2019 / Published online: 9 December 2019

I am most grateful to the invited discussants, Professor Chris Glynn and Dr. Jouchi Nakajima, for their thoughtful and constructive comments and questions. Their discussion contributions speak clearly to some of the key areas of advance in Bayesian forecasting and time series modeling reviewed in the paper, and critically address important areas of "Challenges and Opportunities" with some new suggestions and connections. My responses here speak directly to their specific comments and questions. I hope and expect that this conversation will additionally contribute to promoting new research developments in dynamic models for increasingly complex and challenging problems in multivariate time series analysis and forecasting—and the broader fields of statistical modeling and decision analysis—in the Akaike tradition.

The discussants focus primarily on issues of model structure specification and learning in dynamic graphical models. These issues raise hard questions in multivariate models generally, as discussed in Section 2 of the paper. More specifically, they represent key current challenges in parental set specifications and modeling choices in DDNMs (Section 3) and the more general class of SGDLMs (Section 4). In two recent and current applied projects of my own and with collaborators, the exploration of multiple models based on ranges of parental sets has been—and is—the main effort in the research enterprise. Some of the examples in the paper highlight these kinds of endeavors, using both traditional Bayesian model uncertainty approaches and shotgun stochastic search methods, while comparing models on ranges of forecast and decision criteria as well as standard model probabilities. The model classes are now well understood, with immense flexibility to adapt to complex but inherently structured inter-dependencies among time series, and their changes in time. However, model choice and specification is challenging.

The Related Articles are https://doi.org/10.1007/s10463-019-00741-3; https://doi.org/10.1007/s10463-019-00742-2; https://doi.org/10.1007/s10463-019-00743-1.

✉ Mike West
Mike.West@duke.edu

1 Department of Statistical Science, Duke University, Durham, NC 27708-0251, USA

Dr. Nakajima highlights the general problem based on his expertise and detailed experience with DDNMs linked, primarily but not exclusively, to macroeconomic time series modeling and forecasting. With a cogent discussion some of the seminal background of decouple/recouple thinking in VAR and TV-VAR models, he wisely suggests that a blend of informed prior structuring coupled with empirical statistical model assessments is likely needed in applications in any other than a few dimensions. I very much agree. From the viewpoints of applied macroeconomics in areas such as monetary policy, bringing clear thinking about context and theoretically justified or required constraints on models is vital. Then coupling that with (also clearly thought out) statistical assessments and comparisons is critical. Traditionally, those applied areas have been dominated by the "theory first" view, but are increasingly integrating with the "let the data speak" view. Of course, the increasing impact of Bayesian methodology, pioneered by influential time series econometricians (e.g. Sims 2012), is central to this evolution, and TV-VAR models are routinely adopted. My hope and expectation is that this will continue and that the approaches reviewed in my paper—that extend traditional models with graphical/sparse structures more aggressively—will be increasingly adopted in macroeconomics and in other fields. That said, it remains the case that detailed evaluation of potentially many model choices—testing partial constraints inspired by theory and context, and balanced by empirical testing with specific sets of defined forecast and/or decision goals in the use of the models—will remain central to application.

These challenges of model comparison with respect to parental set selection and structure are echoed in Professor Glynn's comments and questions. Professor Glynn appropriately connects with more traditional Bayesian sparsity prior modeling approaches, whether "point-mass mixtures" or "spike-and-slab" structures. I do agree that there are benefits of the latter over the former in technical senses, and some of the recent literature on bringing these ideas more aggressively into the sequential forward/filtering analysis of dynamic models is indeed interesting and exciting. The initial motivations for dynamic latent threshold models (LTMs, as discussed and noted in the several Nakajima et al. references in the paper, and others) were in fact based on that traditional Bayesian thinking. LTMs very naturally represent not only the interest in learning about changes over time in relevant variables—here, relevant members of parental sets—but also, in fact, imply "smooth" thresholding that corresponds to a class of dynamic spike-and-slab structures. The applied relevance and impact of this thinking is, I believe, very clear from the several publications referenced in the paper, and other substantive applications such as Kimura and Nakajima (2016). However, LTMs—like other dynamic sparsity modeling approaches—are inherently challenging to fit in a forward/sequential format, and some of the recent innovations in "dynamic sparsity" research that might be more conducive to efficient, and effective, sequential analysis are clearly of interest. Bayesian optimization-based analysis and opportunities to exploit importance sampling in new ways are certainly promising directions, in my view. In addition to the new directions by Rockova and McAlinn referenced by Professor Glynn, I note related developments of Bitto and Frühwirth-Schnatter (2019), and quite novel dynamic sparsity structures—and their Bayesian analyses—of Irie (2019) that open up new ground entirely with impressive examples.

To both discussants and readers, however, I will summarize main concerns raised in the paper that are directly relevant to this core issue of model structure assessment.

First, and critically: as dimensions scale the issues of predictor inter-dependencies generate messy problems of multiplicities leading—inevitably—to model uncertainties spread over increasing numbers of models that are exchangeable in any practical sense. Typically, many parental set choices will generate similar "fit" to the data measured in the usual ways and in terms of other specified forecast and decision outcomes. Standard statistical thinking fails as many similar models are aggregated or selected, and the basic premise of sparsity modeling is violated (e.g. Giannone et al. 2018). I encourage a more decision analytic view, i.e., selecting one or a small number of models, rather than the usual model averaging view. This, of course, requires articulation of forecasting and decision goals, and of relevant utility functions.

Second, to emphasize: we model for reasons. Purely statistical assessments via posterior distributions over models—whether combined with insightful theoretical constraints or not—are valid only if we choose "purely statistical" to define utility functions for model uses. Examples in the paper highlight this, and I hope that this paper and discussion will aid the broader community in considering modeling usage goals as part of the broader enterprise in model comparison and selection.

Third, but not at all least: dynamics and sequential settings. Much traditional Bayesian machinery—dominated by MCMC in the last three decades—simply does not translate to the sequential setting. Currently fashionable methods of sequential Monte Carlo face tremendous challenges in any but small problems, and have yet to properly impact in large-scale applications. New ideas and methodology for finding, evaluating, comparing and combining models—generally as well as in connection with parental sets in DDNMs and SGDLMs—are critically needed in the sequential context. Some of the perspective mooted in Section 4 of the paper—of integrating more formal Bayesian decision theoretic thinking into the model uncertainty context—seem very worth embracing and developing. The conceptual advances in Lavine et al. (2019) represent some of my own recent thinking and collaborative development in this direction. Adopting such perspectives will, I predict, open up opportunities for core research and advance methodology in the Akaike spirit: challenging statistical modeling and decision analysis issues motivated by hard, important applications, that engage existing and new researchers in conceptual and theoretical innovation to bring back to address those real-world problems.

References

Bitto, A., Frühwirth-Schnatter, S. (2019). Achieving shrinkage in a time-varying parameter model framework. *Journal of Econometrics*, *210*, 75–97.

Giannone, D., Lenza, M., Primiceri, G. E. (2018). *Economic predictions with big data*: *The illusion of sparsity*. Federal Reserve Board of New York, Staff Report No. 847.

Irie, K. (2019). *Bayesian dynamic fused LASSO*. Technical Report, Department of Economics, The University of Tokyo, arXiv:1905.12275.

Kimura, T., Nakajima, J. (2016). Identifying conventional and unconventional monetary policy shocks: A latent threshold approach. *The BE Journal of Macroeconomics*, *16*, 277–300.

Lavine, I., Lindon, M., West, M. (2019). *Adaptive variable selection for sequential prediction in multivariate dynamic models*, arXiv:1906.06580.

Sims, C. A. (2012). Statistical modeling of monetary policy and its effects. *American Economic Review, 102*, 1187–1205, also: Nobel Prize in Economics documents, Nobel Prize Committee, 2011.