STATISTICAL SCIENCE

Volume 40, Number 1 February 2025

Special Issue on Contemporary Bayesian Prediction

***Statistical Science* [ISSN 0883-4237 (print); ISSN 2168-8745 (online)]**, Volume 40, Number 1, February 2025. Published quarterly by the Institute of Mathematical Statistics, 9760 Smith Road, Waite Hill, Ohio 44094, USA. Periodicals postage paid at Cleveland, Ohio and at additional mailing offices.
POSTMASTER: Send address changes to *Statistical Science*, Institute of Mathematical Statistics, Dues and Subscriptions Office, PO Box 729, Middletown, Maryland 21769, USA.

Statistical Science Volume 40, Number 1 (1–166) February 2025

Volume 40 **Number 1** **February 2025**

Special Issue on Contemporary Bayesian Prediction

Statistical Science
2025, Vol. 40, No. 1, 1–2
https://doi.org/10.1214/24-STS963

Introduction to the Special Issue on Contemporary Bayesian Prediction

Bertrand Clarke and Pietro Rigo

REFERENCES

[1] AITCHISON, J. and DUNSMORE, I. R. (1975). *Statistical Prediction Analysis*. Cambridge Univ. Press, Cambridge-Melbourne. MR0408097

[2] BROWN, L. D. (1981). A complete class theorem for statistical problems with finite sample spaces. *Ann. Statist.* **9** 1289–1300. MR0630111

[3] DE FINETTI, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036

[4] DEGROOT, M. and FIENBERG, S. (1983). The comparison and evaluation of forecasters. *Statistician* **32** 12–22.

[5] GEISSER, S. (1993). *Predictive Inference: An Introduction*. *Monographs on Statistics and Applied Probability* **55**. CRC Press, New York. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2

[6] LEHMANN, E. L. (1959). *Testing Statistical Hypotheses*. Wiley, New York. MR0107933

[7] LEHMANN, E. L. (1983). *Theory of Point Estimation*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, New York. MR0702834 https://doi.org/10.1007/978-1-4757-2769-2

[8] POPPER, K. (1934). *Logik der Forschung*: *Zur Erkenntnistheorie der Modernen Naturwissenschaft*. Julius Springer, Hutchinson & Co. English. English translation: *Logic of Research: on the Epistemology of Modern Natural Science* (1992). Routledge, London.

[9] ROBERT, C. P. (2007). *The Bayesian Choice*, 2nd ed. *Springer Texts in Statistics*. Springer, New York. MR2723361

Bertrand Clarke is Professor, Department of Statistics, University of Nebraska–Lincoln, Lincoln, Nebraska 68583-0963, USA (e-mail: bclarke3@unl.edu). Pietro Rigo is Professor, Department of Statistical Science, University of Bologna, Italy (e-mail: pietro.rigo@unibo.it).

Statistical Science
2025, Vol. 40, No. 1, 3–24
https://doi.org/10.1214/24-STS951

A Cheat Sheet for Bayesian Prediction

Bertrand Clarke and Yuling Yao

Abstract. This paper reviews the growing field of Bayesian prediction. Bayesian point and interval prediction are defined and situated in statistical prediction more generally. Then, four general approaches to Bayesian prediction are described and we turn to predictor selection. This can be done predictively or nonpredictively; predictors can be based on single models or multiple models. We call these unitary predictors and model average predictors, respectively. Predictors can also be explanatory meaning they are based on physical modeling or algorithmic meaning they are regarded simply as input-output relations. Then we turn to aspects of prediction in the context of large observational data sets, the most recent topic to emerge in prediction. We conclude with a summary and statement of several open problems.

Key words and phrases: Predictive density, calibration, validation, predictor selection, predictor evaluation, Bayes rule, Ionescu Tulcea theorem, decision theory, prequential, model average, median model, streaming data.

REFERENCES

[1] ABDELLAOUI, M. and WAKKER, P. P. (2020). Savage for dummies and experts. *J. Econom. Theory* **186** 104991, 20. MR4057518 https://doi.org/10.1016/j.jet.2020.104991

[2] AITCHISON, J. (1975). Goodness of prediction fit. *Biometrika* **62** 547–554. MR0391353 https://doi.org/10.1093/biomet/62.3.547

[3] AITCHISON, J. and DUNSMORE, I. R. (1975). *Statistical Prediction Analysis*. Cambridge Univ. Press, Cambridge-Melbourne. MR0408097

[4] BAHRI, M., MANIU, S. and BIFET, A. (2018). Sketch-based naive Bayes algorithms for evolving data streams. In *IEEE International Conference on Big Data* 604–613, Seattle, WA.

[5] BARBER, D., CEMGIL, A. T. and CHIAPPA, S., eds. (2011). *Bayesian Time Series Models*. Cambridge Univ. Press, Cambridge. MR2894230 https://doi.org/10.1017/CBO9780511984679

[6] BARBER, R. F., CANDÈS, E. J., RAMDAS, A. and TIBSHIRANI, R. J. (2024). De Finetti's theorem and related results for infinite weighted exchangeable sequences. *Bernoulli* **30** 3004–3028. MR4779857 https://doi.org/10.3150/23-bej1704

[7] BARBIERI, M. M. and BERGER, J. O. (2004). Optimal predictive model selection. *Ann. Statist.* **32** 870–897. MR2065192 https://doi.org/10.1214/009053604000000238

[8] BARBIERI, M. M., BERGER, J. O., GEORGE, E. I. and ROČKOVÁ, V. (2021). The median probability model and correlated variables. *Bayesian Anal.* **16** 1085–1112. MR4381128 https://doi.org/10.1214/20-BA1249

[9] BATES, J. M. and GRANGER, C. W. (1969). The combination of forecasts. *Oper. Res. Quart.* **20** 451–468.

[10] BAYARRI, M. J. and BERGER, J. O. (2000). *p* values for composite null models. *J. Amer. Statist. Assoc.* **95** 1127–1142. With comments and a rejoinder by the authors. MR1804239 https://doi.org/10.2307/2669749

[11] BERGER, J. O. and PERICCHI, L. R. (1996). The intrinsic Bayes factor for model selection and prediction. *J. Amer. Statist. Assoc.* **91** 109–122. MR1394065 https://doi.org/10.2307/2291387

[12] BERGER, J. O. and PERICCHI, L. R. (2001). Objective Bayesian methods for model selection: Introduction and comparison. In *Model Selection* (P. Lahiri, ed.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **38** 135–207. IMS, Beachwood, OH. With discussion by J. K. Ghosh, Tapas Samanta, and Fulvio De Santis, and a rejoinder by the authors. MR2000753 https://doi.org/10.1214/lnms/1215540968

[13] BERGER, J. O. and WOLPERT, R. L. (1984). *The Likelihood Principle*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **6**. IMS, Hayward, CA. MR0773665

[14] BERK, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Ann. Math. Stat.* **37** 51–58. MR0189176 https://doi.org/10.1214/aoms/1177699477

[15] BERNARDO, J.-M. and SMITH, A. F. M. (1994). *Bayesian Theory*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

[16] BERSSON, E. and HOFF, P. (2022). Optimal conformal prediction for small areas. Available at arXiv:2204.08122. https://doi.org/10.48550/arXiv.2204.08122

[17] BERTI, P., DREASSI, E., LEISEN, F., PRATELLI, L. and RIGO, P. (2024). A predictive approach to Bayesian forecasting. *Statist. Sci.* **40** 25–39.

[18] BERTI, P., DREASSI, E., PRATELLI, L. and RIGO, P. (2021). A class of models for Bayesian predictive inference. *Bernoulli* **27** 702–726. MR4177386 https://doi.org/10.3150/20-BEJ1255

[19] BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *J. Roy. Statist. Soc. Ser. A*

Bertrand Clarke is a Professor in the Department of Statistics at the University of Nebraska–Lincoln, Lincoln, Nebraska 68583-0963, USA (e-mail: bclarke3@unl.edu). Yuling Yao is an Assistant Professor in the Department of Statistics and Data Sciences at the University of Texas at Austin, Austin, Texas 78705, USA (e-mail: yulingyao6@gmail.com).

143 383–430. With discussion. MR0603745 https://doi.org/10.2307/2982063

[20] BREHMER, J. R. and GNEITING, T. (2020). Properization: Constructing proper scoring rules via Bayes acts. *Ann. Inst. Statist. Math.* **72** 659–673. MR4091430 https://doi.org/10.1007/s10463-019-00705-7

[21] BREHMER, J. R. and STROKORB, K. (2019). Why scoring functions cannot assess tail properties. *Electron. J. Stat.* **13** 4015–4034. MR4015787 https://doi.org/10.1214/19-EJS1622

[22] BROEMELING, L. D. (2019). *Bayesian Analysis of Time Series*. CRC Press, Boca Raton, FL. MR4599114

[23] CALSTER, V. B., MCLERNON, D. J., VAN SMEDEN, M., WYNANTS, L. and STEYERBERG, E. (2019). Calibration: The Achilles heel of predictive analytics. *BMC Med.* **17** 230.

[24] CANDELA, J. (2004). Learning with uncertainty—Gaussian processes and relevance vector machines Ph.D. thesis, Department of Informatics and Mathematical Modelling, Technical Univ. Denmark.

[25] CHAKRABARTI, A. (2020). Data stream algorithms. Lecture notes. Available at https://www.cs.dartmouth.edu/~ac/Teach/CS35-Spring20/.

[26] CHAKRABARTI, A. (2024). Data stream algorithms. Lecture notes. Available at https://www.cs.dartmouth.edu/~ac/Teach/data-streams-lecnotes.pdf.

[27] CHANDA, A. and CLARKE, B. (2024). Point prediction for streaming data. In revision.

[28] CIFARELLI, D. M. and REGAZZINI, E. (1996). De Finetti's contribution to probability and statistics. *Statist. Sci.* **11** 253–282. MR1445983 https://doi.org/10.1214/ss/1032280303

[29] CLARKE, B. (2007). Information optimality and Bayesian modelling. *J. Econometrics* **138** 405–429. MR2391317 https://doi.org/10.1016/j.jeconom.2006.05.003

[30] CLARKE, B. and GUSTAFSON, P. (1998). On the overall sensitivity of the posterior distribution to its inputs. *J. Statist. Plann. Inference* **71** 137–150. MR1651796 https://doi.org/10.1016/S0378-3758(98)00014-7

[31] CLARKE, B. and SUN, D. (1999). Asymptotics of the expected posterior. *Ann. Inst. Statist. Math.* **51** 163–185. MR1704652 https://doi.org/10.1023/A:1003891404142

[32] CLARKE, B. S. and CLARKE, J. L. (2018). *Predictive Statistics: Analysis and Inference Beyond Models. Cambridge Series in Statistical and Probabilistic Mathematics* **46**. Cambridge Univ. Press, Cambridge. MR3791464 https://doi.org/10.1017/9781139236003

[33] CLARKE, J. and CLARKE, B. (2009). Prequential analysis of complex data with adaptive model reselection. *Stat. Anal. Data Min.* **2** 274–290. MR2559266 https://doi.org/10.1002/sam.10052

[34] CLARKE, J. L., CLARKE, B. and YU, C.-W. (2013). Prediction in M-complete problems with limited sample size. *Bayesian Anal.* **8** 647–690. MR3102229 https://doi.org/10.1214/13-BA826

[35] CLEMEN, R. (1989). Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.* **5** 559–583.

[36] CORMODE, G. and MUTHUKRISHNAN, S. (2005). An improved data stream summary: The count-min sketch and its applications. *J. Algorithms* **55** 58–75. MR2132028 https://doi.org/10.1016/j.jalgor.2003.12.001

[37] DAWID, A. P. (1982). The well-calibrated Bayesian. *J. Amer. Statist. Assoc.* **77** 605–613. With comments by Joseph B. Kadane and D. V. Lindley and a rejoinder by the author. MR0675887

[38] DAWID, A. P. (1984). Statistical theory. The prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–292. MR0763811 https://doi.org/10.2307/2981683

[39] DAWID, A. P. (1992). Prequential data analysis. In *Current Issues in Statistical Inference: Essays in Honor of D. Basu. Institute of Mathematical Statistics Lecture Notes—Monograph Series* **17** 113–126. IMS, Hayward, CA. MR1194413 https://doi.org/10.1214/lnms/1215458842

[40] DAWID, A. P. and MUSIO, M. (2015). Bayesian model selection based on proper scoring rules. *Bayesian Anal.* **10** 479–499. MR3420890 https://doi.org/10.1214/15-BA942

[41] DAWID, A. P. and VOVK, V. G. (1999). Prequential probability: Principles and properties. *Bernoulli* **5** 125–162. MR1673572 https://doi.org/10.2307/3318616

[42] DAWID, A. P. and WANG, J. L. (1993). Fiducial prediction and semi-Bayesian inference. *Ann. Statist.* **21** 1119–1138. MR1241260 https://doi.org/10.1214/aos/1176349253

[43] DE FINETTI, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036

[44] DEGROOT, M. and FIENBERG, S. (1983). The comparison and evaluation of forecasters. *J. R. Stat. Soc., Ser. D, Stat.* **32** 12–22.

[45] DEMPSTER, A. P. (1973). Alternatives to least squares in multiple regression. In *Multivariate Statistical Inference* (*Proc. Res. Sem., Dalhousie Univ., Halifax, N.S.*, 1972) (D. G. Kabe and R. P. Gupta, eds.). 25–40. Elsevier, New York. MR0331661

[46] DRAPER, D. (1995). Assessment and propagation of model uncertainty. *J. Roy. Statist. Soc. Ser. B* **57** 45–97. With discussion and a reply by the author. MR1325378

[47] DUPUIS, J. A. and ROBERT, C. P. (2003). Variable selection in qualitative models via an entropic explanatory power. *J. Statist. Plann. Inference* **111** 77–94. MR1955873 https://doi.org/10.1016/S0378-3758(02)00286-0

[48] EBRAHIMI, N., SOOFI, E. S. and SOYER, R. (2010). On the sample information about parameter and prediction. *Statist. Sci.* **25** 348–367. MR2791672 https://doi.org/10.1214/10-STS329

[49] FISSLER, T. and PESENTI, S. M. (2023). Sensitivity measures based on scoring functions. *European J. Oper. Res.* **307** 1408–1423. MR4543545 https://doi.org/10.1016/j.ejor.2022.10.002

[50] FONG, E. and HOLMES, C. (2021). Conformal Bayesian computation. In *Adv. in Neural Processing Systems* (M. Ranzato, A. Beygelzimer, Y. Dauphin, P. S. Liang and J. Wortman Vaughan, eds.). **34** 18268–18279. Curran Associates, Red Hook, NY.

[51] FONG, E., HOLMES, C. and WALKER, S. G. (2023). Martingale posterior distributions. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **85** 1357–1391. MR4726953 https://doi.org/10.1093/jrsssb/qkad005

[52] FORTINI, S., LADELLI, L. and REGAZZINI, E. (2000). Exchangeability, predictive distributions and parametric models. *Sankhyā Ser. A* **62** 86–109. MR1769738

[53] FORTINI, S. and PETRONE, S. (2012). Predictive construction of priors in Bayesian nonparametrics. *Braz. J. Probab. Stat.* **26** 423–449. MR2949087 https://doi.org/10.1214/11-BJPS176

[54] FREEDMAN, D. A. and PURVES, R. A. (1969). Bayes' method for bookies. *Ann. Math. Stat.* **40** 1177–1186. MR0240914 https://doi.org/10.1214/aoms/1177697494

[55] GALTON, F. (1907). Vox populi. *Nature* **75** 450–451.

[56] GEISSER, S. (1993). *Predictive Inference: An Introduction. Monographs on Statistics and Applied Probability* **55**. CRC Press, New York. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2

[57] GELFAND, A. E., DEY, D. K. and CHANG, H. (1992). Model determination using predictive distributions with implementation via sampling-based methods. In *Bayesian Statistics*, 4 (*Peñíscola*, 1991) 147–167. Oxford Univ. Press, New York. MR1380275

[58] GELMAN, A. (2003). A Bayesian formulation of exploratory data analysis and goodness-of-fit testing. *Int. Stat. Rev.* **71** 369–382.

[59] GELMAN, A. (2013). Two simple examples for understanding posterior p-values whose distributions are far from unform. *Electron. J. Stat.* **7** 2595–2602. MR3121624 https://doi.org/10.1214/13-EJS854

[60] GELMAN, A., HWANG, J. and VEHTARI, A. (2014). Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24** 997–1016. MR3253850 https://doi.org/10.1007/s11222-013-9416-2

[61] GELMAN, A., MENG, X.-L. and STERN, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statist. Sinica* **6** 733–807. With comments and a rejoinder by the authors. MR1422404

[62] GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge Univ. Press, Cambridge. MR3587782 https://doi.org/10.1017/9781139029834

[63] GNEITING, T. and KATZFUSS, M. (2014). Probabilistic forecasting. *Annu. Rev. Stat. Appl.* **1** 125–151.

[64] GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548 https://doi.org/10.1198/016214506000001437

[65] GOUTIS, C. and ROBERT, C. P. (1998). Model choice in generalised linear models: A Bayesian approach via Kullback-Leibler projections. *Biometrika* **85** 29–37. MR1627250 https://doi.org/10.1093/biomet/85.1.29

[66] GURAJALA, T., CHOPPALA, P., MEKA, J. and TEAL, P. (2021). Derivation of the Kalman filter in a Bayesian filtering perspective. In *Proceedings of the 2nd International Conference on Range Technology*. 1–5. Chandipur, Balasore, India.

[67] HOFF, P. (2023). Bayes-optimal prediction with frequentist coverage control. *Bernoulli* **29** 901–928. MR4550209 https://doi.org/10.3150/22-bej1484

[68] HUGGINS, J. H. and MILLER, J. W. (2023). Reproducible model selection using bagged posteriors. *Bayesian Anal.* **18** 79–104. MR4515726 https://doi.org/10.1214/21-ba1301

[69] JOHNSON, V. E. (2004). A Bayesian χ^2 test for goodness-of-fit. *Ann. Statist.* **32** 2361–2384. MR2153988 https://doi.org/10.1214/009053604000000616

[70] JORDAN, A., KRÜGER, F. and LERCH, S. (2018). Evaluating probablistic forecasts with scoringRules. Available at arxiv.org/pdf/1709.04743.pdf.

[71] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[72] KONTKANEN, P., MYLLYM AKI, P., SILANDER, T., TIRRI, H. and GRÜNWALD, P. (1997). Comparing predictive inference methods for discrete domains. In *Proceedings of the Sixth International Workshop on Artificial Intelligence and Statistics*. R1:311–R1:318. PMLR.

[73] KONTKANEN, P. and MYLLYMÄKI, P. (2007). A linear-time algorithm for computing the multinomial stochastic complexity. *Inform. Process. Lett.* **103** 227–233. MR2339339 https://doi.org/10.1016/j.ipl.2007.04.003

[74] LE, T. and CLARKE, B. (2016). Using the Bayesian Shtarkov solution for predictions. *Comput. Statist. Data Anal.* **104** 183–196. MR3540994 https://doi.org/10.1016/j.csda.2016.06.018

[75] LE, T. and CLARKE, B. (2017). A Bayes interpretation of stacking for $\mathcal{M}$-complete and $\mathcal{M}$-open settings. *Bayesian Anal.* **12** 807–829. MR3655877 https://doi.org/10.1214/16-BA1023

[76] LE, T. and CLARKE, B. (2020). In praise of partially interpretable predictors. *Stat. Anal. Data Min.* **13** 113–133. MR4086137 https://doi.org/10.1002/sam.11450

[77] LE, T. M. and CLARKE, B. (2022). Model averaging is asymptotically better than model selection for prediction. *J. Mach. Learn. Res.* **23** Paper No. [33], 53. MR4420758

[78] LE, T. M. and CLARKE, B. (2022). Interpreting uninterpretable predictors: Kernel methods, Shtarkov solutions, and random forests. *Stat. Theory Relat. Fields* **6** 10–28. MR4390850 https://doi.org/10.1080/24754269.2021.1974157

[79] MADIGAN, D., RAFTERY, A., VOLINSKY, C. and HOETING, J. (1996). Bayesian model averaging. In *Proceedings of the AAAI Workshop on Integrating Multiple Learned Models*. 77–83.

[80] MARSHALL, E. C. and SPIEGELHALTER, D. J. (2003). Approximate cross-validatory predictive checks in disease mapping models. *Stat. Med.* **22** 1649–1660.

[81] MARTIN, G. M., FRAZIER, D. T., LOAIZA-MAYA, R., HUBER, F., KOOP, G., MAHEU, J., NIBBERING, D. and PANAGIOTELIS, A. (2024). Bayesian Forecasting in the 21st Century: a Modern Review. Working Paper 1/23. Department of econometrics and business statistics, Monash University.

[82] MCLATCHIE, Y., RÖGNVALDSSON, S., WEBER, F. and VEHTARI, A. (2024). Advances in projection predictive inference. *Statist. Sci.* **40** 128–147.

[83] MEEKER, W., HAHN, G. and ESCOBAR, L. (2017). *Statistical Intervals*, 2nd ed. Wiley, Hoboken NJ.

[84] MERHAV, N. and FEDER, M. (1998). Universal prediction. *IEEE Trans. Inf. Theory* **44** 2124–2147. MR1658815 https://doi.org/10.1109/18.720534

[85] MÜLLER, P., QUINTANA, F. A., JARA, A. and HANSON, T. (2015). *Bayesian Nonparametric Data Analysis*. *Springer Series in Statistics*. Springer, Cham. MR3309338 https://doi.org/10.1007/978-3-319-18968-0

[86] MUTHUKRISHNAN, S. (2005). Data streams: Algorithms and applications. *Found. Trends Theor. Comput. Sci.* **1** 117–236. MR2379507 https://doi.org/10.1561/0400000002

[87] MUTHUKRISHNAN, S. (2009). Data stream algorithms. Lecture notes. Available at https://www.cs.mcgill.ca/~denis/notes09.pdf.

[88] O'HAGAN, A. (1995). Fractional Bayes factors for model comparison. *J. Roy. Statist. Soc. Ser. B* **57** 99–138. MR1325379

[89] PARMIGIANI, G. (2023). Defining replicability of prediction rules. *Statist. Sci.* **38** 543–556. MR4665025 https://doi.org/10.1214/23-sts891

[90] PAVONE, F., PIIRONEN, J., BÜRKNER, P.-C. and VEHTARI, A. (2023). Using reference models in variable selection. *Comput. Statist.* **38** 349–371. MR4556024 https://doi.org/10.1007/s00180-022-01231-6

[91] PETRIS, G., PETRONE, S. and CAMPAGNOLI, P. (2009). *Dynamic Linear Models with R*. *Use R!* Springer, New York. MR2730074 https://doi.org/10.1007/b135794

[92] PIIRONEN, J. and VEHTARI, A. (2017). Comparison of Bayesian predictive methods for model selection. *Stat. Comput.* **27** 711–735. MR3613594 https://doi.org/10.1007/s11222-016-9649-y

[93] POLE, A., WEST, M. and HARRISON, J. (2019). *Applied Bayesian Forecasting and Time Series Analysis*. CRC Press, Boca Raton, FL.

[94] RAFTERY, A. and ZHENG, Y. (2003). Discussion: Performance of Bayesian model averaging. *J. Amer. Statist. Assoc.* **98** 931–938.

[95] RISSANEN, J. (1984). Universal coding, information, prediction, and estimation. *IEEE Trans. Inf. Theory* **30** 629–636. MR0755791 https://doi.org/10.1109/TIT.1984.1056936

[96] ROSTEK, M. (2010). Quantile maximization in decision theory. *Rev. Econ. Stud.* **77** 339–371. MR2599290 https://doi.org/10.1111/j.1467-937X.2009.00564.x

[97] RUBIN, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Statist.* **12** 1151–1172. MR0760681 https://doi.org/10.1214/aos/1176346785

[98] SATULURI, V. and PARTHASARATHY, S. (2012). Bayesian locality sensitive hashing for fast similarity search. *Proc. VLDB Endow.* **5** 430–44.

[99] SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York. MR0063582

[100] SHAFER, G. and VOVK, V. (2008). A tutorial on conformal prediction. *J. Mach. Learn. Res.* **9** 371–421. MR2417240

[101] SHALIZI, C. (2007). Class notes. Available at https://www.stat.cmu.edu/~cshalizi/754/notes/lecture-03.pdf.

[102] SHAO, J. (1997). An asymptotic theory for linear model selection. *Statist. Sinica* **7** 221–264. With comments and a rejoinder by the author. MR1466682

[103] SHMUELI, G. (2010). To explain or to predict? *Statist. Sci.* **25** 289–310. MR2791669 https://doi.org/10.1214/10-STS330

[104] SHTARKOV, YU. M. (1987). Universal sequential coding of individual messages. *Problemy Peredachi Informatsii* **23** 3–17. MR0914346

[105] SIEGEL, E. (2016). *Predictive Analytics*. Wiley, Hoboken, NJ.

[106] SILVER, N. (2012). *The Signal and the Noise*. Penguin, NY, NY.

[107] SKOURAS, K. and DAWID, A. P. (1998). On efficient point prediction systems. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 765–780. MR1649472 https://doi.org/10.1111/1467-9868.00153

[108] SKOURAS, K. and DAWID, A. P. (1999). On efficient probability forecasting systems. *Biometrika* **86** 765–784. MR1741976 https://doi.org/10.1093/biomet/86.4.765

[109] SMYTH, P. and WOLPERT, D. (1998). An evaluation of linearly combining density estimators via stacking Tech. Rep. 98-25, Information and Computer Science Department, Univ. Irvine.

[110] TIAN, Q., NORDMAN, D. J. and MEEKER, W. Q. (2022). Methods to compute prediction intervals: A review and new results. *Statist. Sci.* **37** 580–597. MR4497233 https://doi.org/10.1214/21-sts842

[111] VAN ERVEN, T., GRÜNWALD, P., DE ROOIJ, S. and JANDHYALA, V. K. (2012). Catching up faster by switching sooner: A predictive approach to adaptive estimation with an application to the AIC-BIC dilemma. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **74** 361–417. MR2925369 https://doi.org/10.1111/j.1467-9868.2011.01025.x

[112] VEHTARI, A., GELMAN, A. and GABRY, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27** 1413–1432. MR3647105 https://doi.org/10.1007/s11222-016-9696-4

[113] VEHTARI, A. and OJANEN, J. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Stat. Surv.* **6** 142–228. MR3011074 https://doi.org/10.1214/12-SS102

[114] VOVK, H., SHAFER, G. and DAWID, A. P. D A conversation with A. Philip Dawid. *Statist. Sci.* **40** 148–166.

[115] VOVK, V., GAMMERMAN, A. and SHAFER, G. (2005). *Algorithmic Learning in a Random World*. Springer, New York. MR2161220

[116] WALLIS, K. F. (2014). Revisiting Francis Galton's forecasting competition. *Statist. Sci.* **29** 420–424. MR3264553 https://doi.org/10.1214/14-STS468

[117] WANG, C. M., HANNIG, J. and IYER, H. K. (2012). Fiducial prediction intervals. *J. Statist. Plann. Inference* **142** 1980–1990. MR2903406 https://doi.org/10.1016/j.jspi.2012.02.021

[118] WATANABE, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J. Mach. Learn. Res.* **11** 3571–3594. MR2756194

[119] WOLPERT, D. (1992). Stacked generalization. *Neural Netw.* **5** 241–259.

[120] WONG, H. and CLARKE, B. (2004). Improvement over Bayes prediction in small samples in the presence of model uncertainty. *Canad. J. Statist.* **32** 269–283. MR2101756 https://doi.org/10.2307/3315929

[121] WU, X., LIU, M. and WU, Y. (2012). In-situ soil moisture sensing: Optimal sensor placement and field estimation. *ACM Trans. Sens. Netw.* **8** 33.

[122] YANG, Z. and ZHU, T. (2018). Bayesian selection of misspecified models is overconfident and may cause spurious posterior probabilities for phylogenetic trees. *Proc. Natl. Acad. Sci. USA* **115** 1854–1859. MR3779786 https://doi.org/10.1073/pnas.1712673115

[123] YAO, Y., PIRŠ, G., VEHTARI, A. and GELMAN, A. (2022). Bayesian hierarchical stacking: Some models are (somewhere) useful. *Bayesian Anal.* **17** 1043–1071. MR4506021 https://doi.org/10.1214/21-ba1287

[124] YAO, Y., RÉGALDO-SAINT BLANCARD, B. and DOMKE, J. (2024). Simulation-based stacking. In *International Conference on Artificial Intelligence and Statistics*.

[125] YAO, Y., VEHTARI, A., SIMPSON, D. and GELMAN, A. (2018). Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Anal.* **13** 917–1007. Including a rejoinder by the authors. MR3853125 https://doi.org/10.1214/17-BA1091

[126] YU, C. W. and CLARKE, B. (2011). Median loss decision theory. *J. Statist. Plann. Inference* **141** 611–623. MR2732931 https://doi.org/10.1016/j.jspi.2010.08.013

[127] YU, Q., MACEACHERN, S. N. and PERUGGIA, M. (2013). Clustered Bayesian model averaging. *Bayesian Anal.* **8** 883–908. MR3150472 https://doi.org/10.1214/13-BA859

[128] YUAN, Y. and JOHNSON, V. E. (2012). Goodness-of-fit diagnostics for Bayesian hierarchical models. *Biometrics* **68** 156–164. MR2909864 https://doi.org/10.1111/j.1541-0420.2011.01668.x

[129] ZELLNER, A. (1988). Optimal information processing and Bayes's theorem. *Amer. Statist.* **42** 278–280. With comments and a reply by the author. MR0971095 https://doi.org/10.2307/2685143

Statistical Science
2025, Vol. 40, No. 1, 25–39
https://doi.org/10.1214/23-STS884

A Probabilistic View on Predictive Constructions for Bayesian Learning

Patrizia Berti, Emanuela Dreassi, Fabrizio Leisen, Luca Pratelli and Pietro Rigo

Abstract. Given a sequence $X = (X_1, X_2, \dots)$ of random observations, a Bayesian forecaster aims to predict X_{n+1} based on $(X_1, \dots, X_n)$ for each $n \geq 0$. To this end, in principle, she only needs to select a collection $\sigma = (\sigma_0, \sigma_1, \dots)$, called "strategy" in what follows, where $\sigma_0(\cdot) = P(X_1 \in \cdot)$ is the marginal distribution of X_1 and $\sigma_n(\cdot) = P(X_{n+1} \in \cdot \mid X_1, \dots, X_n)$ the nth predictive distribution. Because of the Ionescu–Tulcea theorem, σ can be assigned directly, without passing through the usual prior/posterior scheme. One main advantage is that no prior probability is to be selected. In a nutshell, this is the predictive approach to Bayesian learning. A concise review of the latter is provided in this paper. We try to put such an approach in the right framework, to make clear a few misunderstandings, and to provide a unifying view. Some recent results are discussed as well. In addition, some new strategies are introduced and the corresponding distribution of the data sequence X is determined. The strategies concern generalized Pólya urns, random change points, covariates and stationary sequences.

Key words and phrases: Bayesian inference, conditional identity in distribution, exchangeability, predictive distribution, sequential predictions, stationarity.

REFERENCES

[1] Airoldi, E. M., Costa, T., Bassetti, F., Leisen, F. and Guindani, M. (2014). Generalized species sampling priors with latent beta reinforcements. *J. Amer. Statist. Assoc.* **109** 1466–1480. MR3293604 https://doi.org/10.1080/01621459.2014.950735

[2] Bassetti, F., Crimaldi, I. and Leisen, F. (2010). Conditionally identically distributed species sampling sequences. *Adv. in Appl. Probab.* **42** 433–459. MR2675111 https://doi.org/10.1239/aap/1275055237

[3] Bassetti, F. and Ladelli, L. (2020). Asymptotic number of clusters for species sampling sequences with non-diffuse base measure. *Statist. Probab. Lett.* **162** 108749, 7. MR4079614 https://doi.org/10.1016/j.spl.2020.108749

[4] Berti, P., Dreassi, E., Leisen, F., Pratelli, L. and Rigo, P. (2023). Kernel based Dirichlet sequences. *Bernoulli* **29** 1321–1342. MR4550225 https://doi.org/10.3150/22-BEJ1500

[5] Berti, P., Dreassi, E., Leisen, F., Pratelli, L. and Rigo, P. (2023). Bayesian predictive inference without a prior. *Statist. Sinica* **33**. https://doi.org/10.5705/ss.202021.0238

[6] Berti, P., Dreassi, E., Pratelli, L. and Rigo, P. (2021). A class of models for Bayesian predictive inference. *Bernoulli* **27** 702–726. MR4177386 https://doi.org/10.3150/20-BEJ1255

[7] Berti, P., Pratelli, L. and Rigo, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32** 2029–2052. MR2073184 https://doi.org/10.1214/009117904000000676

[8] Berti, P., Pratelli, L. and Rigo, P. (2012). Limit theorems for empirical processes based on dependent data. *Electron. J. Probab.* **17** no. 9, 18. MR2878788 https://doi.org/10.1214/EJP.v17-1765

[9] Berti, P., Pratelli, L. and Rigo, P. (2013). Exchangeable sequences driven by an absolutely continuous random measure. *Ann. Probab.* **41** 2090–2102. MR3098068 https://doi.org/10.1214/12-AOP786

[10] Berti, P., Regazzini, E. and Rigo, P. (1997). Well-calibrated, coherent forecasting systems. *Theory Probab. Appl.* **42** 82–102.

[11] Blackwell, D. and MacQueen, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

Patrizia Berti is Professor of Probability, Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy (e-mail: patrizia.berti@unimore.it). Emanuela Dreassi is Professor of Statistics, Dipartimento di Statistica, Informatica, Applicazioni, Università di Firenze, viale Morgagni 59, 50134 Firenze, Italy (e-mail: emanuela.dreassi@unifi.it). Fabrizio Leisen is Professor of Statistics, School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK (e-mail: fabrizio.leisen@gmail.com). Luca Pratelli is Professor of Probability, Accademia Navale, viale Italia 72, 57100 Livorno, Italy (e-mail: pratel@mail.dm.unipi.it). Pietro Rigo is Professor of Probability, Dipartimento di Scienze Statistiche "P. Fortunati," Università di Bologna, via delle Belle Arti 41, 40126 Bologna, Italy (e-mail: pietro.rigo@unibo.it).

[12] BLADT, M. and MCNEIL, A. J. (2022). Time series with infinite-order partial copula dependence. *Depend. Model.* **10** 87–107. MR4426880 https://doi.org/10.1515/demo-2022-0105
[13] CANALE, A., LIJOI, A., NIPOTI, B. and PRÜNSTER, I. (2017). On the Pitman–Yor process with spike and slab base measure. *Biometrika* **104** 681–697. MR3694590 https://doi.org/10.1093/biomet/asx041
[14] CASSESE, A., ZHU, W., GUINDANI, M. and VANNUCCI, M. (2019). A Bayesian nonparametric spiked process prior for dynamic model selection. *Bayesian Anal.* **14** 553–572. MR3934097 https://doi.org/10.1214/18-BA1116
[15] CHEN, K., SHEN, W. and ZHU, W. (2023). Covariate dependent Beta-GOS process. *Comput. Statist. Data Anal.* **180** Paper No. 107662, 13. MR4515770 https://doi.org/10.1016/j.csda.2022.107662
[16] CIFARELLI, D. M. and REGAZZINI, E. (1996). De Finetti's contribution to probability and statistics. *Statist. Sci.* **11** 253–282. MR1445983 https://doi.org/10.1214/ss/1032280303
[17] CLARKE, B., FOKOUE, E. and ZHANG, H. H. (2009). *Principles and Theory for Data Mining and Machine Learning*. Springer, New York.
[18] CLARKE, B. S. and CLARKE, J. L. (2018). *Predictive Statistics: Analysis and Inference Beyond Models*. *Cambridge Series in Statistical and Probabilistic Mathematics* **46**. Cambridge Univ. Press, Cambridge. MR3791464 https://doi.org/10.1017/9781139236003
[19] DAWID, A. P. (1984). Statistical theory. The prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–292. MR0763811 https://doi.org/10.2307/2981683
[20] DAWID, A. P. (1992). Prequential data analysis. In *Current Issues in Statistical Inference: Essays in Honor of D. Basu* (M. Ghosh and P. K. Pathak, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **17** 113–126. IMS, Hayward. MR1194413 https://doi.org/10.1214/lnms/1215458842
[21] DAWID, A. P. and VOVK, V. G. (1999). Prequential probability: Principles and properties. *Bernoulli* **5** 125–162. MR1673572 https://doi.org/10.2307/3318616
[22] DE FINETTI, B. (1931). Sul significato soggettivo della probabilità. *Fund. Math.* **17** 298–329.
[23] DE FINETTI, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann. Inst. Henri Poincaré* **7** 1–68. MR1508036
[24] DIACONIS, P. and FREEDMAN, D. A. (1990). Cauchy's equation and de Finetti's theorem. *Scand. J. Stat.* **17** 235–249. MR1092946
[25] DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281. MR0520238
[26] DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble If You Must. Inequalities for Stochastic Processes*. McGraw-Hill, New York. MR0236983
[27] EFRON, B. (2020). Prediction, estimation, and attribution. *J. Amer. Statist. Assoc.* **115** 636–655. MR4107663 https://doi.org/10.1080/01621459.2020.1762613
[28] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949
[29] FONG, E., HOLMES, C. and WALKER, S. G. (2023). Martingale posterior distributions (with discussion). *J. Roy. Statist. Soc. Ser. B*. To appear.
[30] FONG, E. and LEHMANN, B. (2022). A predictive approach to Bayesian nonparametric survival analysis. Available at arXiv:2202.10361v1 [stat.ME].
[31] FORTINI, S., LADELLI, L. and REGAZZINI, E. (2000). Exchangeability, predictive distributions and parametric models. *Sankhyā Ser. A* **62** 86–109. MR1769738
[32] FORTINI, S. and PETRONE, S. (2012). Predictive construction of priors in Bayesian nonparametrics. *Braz. J. Probab. Stat.* **26** 423–449. MR2949087 https://doi.org/10.1214/11-BJPS176
[33] FORTINI, S. and PETRONE, S. (2017). Predictive characterization of mixtures of Markov chains. *Bernoulli* **23** 1538–1565. MR3624870 https://doi.org/10.3150/15-BEJ787
[34] FORTINI, S. and PETRONE, S. (2020). Quasi-Bayes properties of a procedure for sequential learning in mixture models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 1087–1114. MR4136504
[35] FORTINI, S., PETRONE, S. and SPORYSHEVA, P. (2018). On a notion of partially conditionally identically distributed sequences. *Stochastic Process. Appl.* **128** 819–846. MR3758339 https://doi.org/10.1016/j.spa.2017.06.008
[36] GEISSER, S. (1993). *Predictive Inference: An Introduction*. *Monographs on Statistics and Applied Probability* **55**. CRC Press, New York. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2
[37] GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge Univ. Press, Cambridge. MR3587782 https://doi.org/10.1017/9781139029834
[38] GNEDIN, A. (2010). A species sampling model with finitely many types. *Electron. Commun. Probab.* **15** 79–88. MR2606505 https://doi.org/10.1214/ECP.v15-1532
[39] GNEDIN, A. and PITMAN, J. (2006). Exchangeable Gibbs partitions and Stirling triangles. *J. Math. Sci.* **138** 5674–5685.
[40] HAHN, P. R. (2017). Predictivist Bayes density estimation. Unpublished technical report. Available at https://math.la.asu.edu/prhahn/pred-bayes.pdf.
[41] HAHN, P. R., MARTIN, R. and WALKER, S. G. (2018). On recursive Bayesian predictive distributions. *J. Amer. Statist. Assoc.* **113** 1085–1093. MR3862341 https://doi.org/10.1080/01621459.2017.1304219
[42] HANSEN, B. and PITMAN, J. (2000). Prediction rules for exchangeable sequences related to species sampling. *Statist. Probab. Lett.* **46** 251–256. MR1745692 https://doi.org/10.1016/S0167-7152(99)00109-1
[43] HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR2722294 https://doi.org/10.1007/978-0-387-84858-7
[44] HILL, B. M. (1993). Parametric models for A_n: Splitting processes and mixtures. *J. Roy. Statist. Soc. Ser. B* **55** 423–433. MR1224406
[45] HJORT, N. L., HOLMES, C., MÜLLER, P. and WALKER, S. G., eds. (2010). *Bayesian Nonparametrics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **28**. Cambridge Univ. Press, Cambridge. MR2722987 https://doi.org/10.1017/CBO9780511802478
[46] HOFFMANN-JØRGENSEN, J. (1994). *Probability with a View Toward Statistics, Vol. II*. Chapman & Hall, New York.
[47] KALLENBERG, O. (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16** 508–534. MR0929061
[48] LEE, J., QUINTANA, F. A., MÜLLER, P. and TRIPPA, L. (2013). Defining predictive probability functions for species sampling models. *Statist. Sci.* **28** 209–222. MR3112406 https://doi.org/10.1214/12-sts407
[49] LIJOI, A., PRÜNSTER, I. and WALKER, S. G. (2008). Bayesian nonparametric estimators derived from conditional Gibbs structures. *Ann. Appl. Probab.* **18** 1519–1547. MR2434179 https://doi.org/10.1214/07-AAP495

[50] MORVAI, G. and WEISS, B. (2021). On universal algorithms for classifying and predicting stationary processes. *Probab. Surv.* **18** 77–131. MR4255241 https://doi.org/10.1214/20-ps345

[51] NEWTON, M. A. (2002). On a nonparametric recursive estimator of the mixing distribution **64** 306–322. Sankhyā Ser. A, 2, Selected articles from San Antonio Conference in honour of C. R. Rao (San Antonio, TX, 2000). MR1981761

[52] NEWTON, M. A. and ZHANG, Y. (1999). A recursive algorithm for nonparametric analysis with missing data. *Biometrika* **86** 15–26. MR1688068 https://doi.org/10.1093/biomet/86.1.15

[53] PITMAN, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249 https://doi.org/10.1007/BF01213386

[54] PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics*, *Probability and Game Theory*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 245–267. IMS, Hayward. MR1481784 https://doi.org/10.1214/lnms/1215453576

[55] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math*. **1875**. Springer, Berlin. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002, With a foreword by Jean Picard. MR2245368

[56] PITMAN, J. and YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann*. *Probab*. **25** 855–900. MR1434129 https://doi.org/10.1214/aop/1024404422

[57] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statist*. *Sinica* **4** 639–650. MR1309433

[58] SHMUELI, G. (2010). To explain or to predict? *Statist*. *Sci*. **25** 289–310. MR2791669 https://doi.org/10.1214/10-STS330

[59] SMITH, A. F. M. and MAKOV, U. E. (1978). A quasi-Bayes sequential procedure for mixtures. *J. Roy. Statist. Soc. Ser. B* **40** 106–112. MR0512148

Statistical Science
2025, Vol. 40, No. 1, 40–67
https://doi.org/10.1214/24-STS965

Exchangeability, Prediction and Predictive Modeling in Bayesian Statistics

Sandra Fortini and Sonia Petrone

Abstract. There is currently a renewed interest in the Bayesian *predictive* approach to statistics. This paper offers a review on foundational concepts and focuses on "predictive modeling," which by directly reasoning on prediction bypasses inferential models or may characterize them. We detail predictive characterizations in exchangeable and partially exchangeable settings, for a large variety of data structures, and hint at new directions. The underlying concept is that Bayesian predictive rules are probabilistic *learning* rules, formalizing through conditional probability how we learn on future events given the available information. This concept has implications in any statistical problem and in inference, from classic contexts to less explored challenges, such as providing Bayesian uncertainty quantification to predictive algorithms in data science, as we show in the last part of the paper. The paper not only gives a historical overview but also includes a few new results, presents some recent developments and poses some open questions.

Key words and phrases: Bayesian foundations, predictive characterizations, Bayesian nonparametrics, predictive sufficiency, partial exchangeability, recursive algorithms.

REFERENCES

[1] ALDOUS, D. J. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11** 581–598. MR0637937 https://doi.org/10.1016/0047-259X(81)90099-3

[2] ALDOUS, D. J. (1985). Exchangeability and related topics. In *École D'été de Probabilités de Saint-Flour, XIII*—1983. *Lecture Notes in Math*. **1117** 1–198. Springer, Berlin. MR0883646 https://doi.org/10.1007/BFb0099421

[3] ALETTI, G., CRIMALDI, I. and GHIGLIETTI, A. (2023). Interacting innovation processes. *Nature*, *Sci*. *Rep*. **13**.

[4] ANTONIAK, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2** 1152–1174. MR0365969

[5] AUGENBLICK, N. and RABIN, M. (2021). Belief movement, uncertainty reduction, and rational updating. *Q. J. Econ.* **136** 933–985.

[6] BACALLADO, S. (2011). Bayesian analysis of variable-order, reversible Markov chains. *Ann. Statist.* **39** 838–864. MR2816340 https://doi.org/10.1214/10-AOS857

[7] BACALLADO, S., BATTISTON, M., FAVARO, S. and TRIPPA, L. (2017). Sufficientness postulates for Gibbs-type priors and hierarchical generalizations. *Statist. Sci.* **32** 487–500. MR3730518 https://doi.org/10.1214/17-STS619

[8] BACALLADO, S., FAVARO, S., POWER, S. and TRIPPA, L. (2022). Perfect sampling of the posterior in the hierarchical Pitman–Yor process. *Bayesian Anal.* **17** 685–709. MR4483235 https://doi.org/10.1214/21-ba1269

[9] BEAL, M. J., GHAHRAMANI, Z. and RASMUSSEN, C. E. (2002). The infinite hidden Markov model. In *Mach. Learn* 239–245. MIT Press, Cambridge.

[10] BERNARDO, J.-M. and SMITH, A. F. M. (1994). *Bayesian Theory. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

[11] BERTI, P., CRIMALDI, I., PRATELLI, L. and RIGO, P. (2015). Central limit theorems for an Indian Buffet model with random weights. *Ann. Appl. Probab.* **25** 523–547. MR3313747 https://doi.org/10.1214/14-AAP1002

[12] BERTI, P., DREASSI, E., LEISEN, F., PRATELLI, L. and RIGO, P. (2023). A probabilistic view on predictive constructions for Bayesian learning. *Statist. Sci.* 1–15.

[13] BERTI, P., DREASSI, E., LEISEN, F., PRATELLI, L. and RIGO, P. (2023). Kernel based Dirichlet sequences. *Bernoulli* **29** 1321–1342. MR4550225 https://doi.org/10.3150/22-bej1500

[14] BERTI, P., DREASSI, E., LEISEN, F., PRATELLI, L. and RIGO, P. (2023). Bayesian predictive inference without a prior. *Statist. Sinica* **33** 2405–2429. MR4647040

[15] BERTI, P., PRATELLI, L. and RIGO, P. (2004). Limit theorems for a class of identically distributed random variables. *Ann. Probab.* **32** 2029–2052. MR2073184 https://doi.org/10.1214/009117904000000676

Sandra Fortini is Associate Professor of Statistics at the Department of Decision Sciences, Bocconi University, Milano, Italy (e-mail: sandra.fortini@unibocconi.it). Sonia Petrone is Professor of Statistics, Department of Decision Sciences and Bocconi Institute of Data Science and Analytics, Bocconi University, Milano, Italy (e-mail: sonia.petrone@unibocconi.it).

[16] BERTI, P., PRATELLI, L. and RIGO, P. (2013). Exchangeable sequences driven by an absolutely continuous random measure. *Ann*. *Probab*. **41** 2090–2102. MR3098068 https://doi.org/10.1214/12-AOP786

[17] BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann*. *Statist*. **1** 353–355. MR0362614

[18] BREIMAN, L. (2001). Statistical modeling: The two cultures. *Statist*. *Sci*. **16** 199–231. With comments and a rejoinder by the author. MR1874152 https://doi.org/10.1214/ss/1009213726

[19] CAMERLENGHI, F., FAVARO, S., MASOERO, L. and BRODERICK, T. (2024). Scaled process priors for Bayesian nonparametric estimation of the unseen genetic variation. *J*. *Amer*. *Statist*. *Assoc*. **119** 320–331. MR4713895 https://doi.org/10.1080/01621459.2022.2115918

[20] CAMERLENGHI, F., LIJOI, A., ORBANZ, P. and PRÜNSTER, I. (2019). Distribution theory for hierarchical processes. *Ann*. *Statist*. **47** 67–92. MR3909927 https://doi.org/10.1214/17-AOS1678

[21] CAMERLENGHI, F., LIJOI, A. and PRÜNSTER, I. (2017). Bayesian prediction with multiple-samples information. *J*. *Multivariate Anal*. **156** 18–28. MR3624682 https://doi.org/10.1016/j.jmva.2017.01.010

[22] CATALANO, M., DEL SOLE, C., LIJOI, A. and PRÜNSTER, I. (2024). A Unified Approach to Hierarchical Random Measures. *Sankhya A* **86** 255–287. MR4827151 https://doi.org/10.1007/s13171-023-00330-w

[23] CIFARELLI, D. M. and REGAZZINI, E. (1996). De Finetti's contribution to probability and statistics. *Statist*. *Sci*. **11** 253–282. MR1445983 https://doi.org/10.1214/ss/1032280303

[24] CLARKE, B. S. and CLARKE, J. L. (2018). *Predictive Statistics*: *Analysis and Inference Beyond Models*. *Cambridge Series in Statistical and Probabilistic Mathematics* **46**. Cambridge Univ. Press, Cambridge. MR3791464 https://doi.org/10.1017/9781139236003

[25] COPPERSMITH, D. and DIACONIS, P. (1986). Random walk with reinforcement. Unpublished manuscript.

[26] CRIMALDI, I. (2009). An almost sure conditional convergence result and an application to a generalized Pólya urn. *Int*. *Math*. *Forum* **4** 1139–1156. MR2524635

[27] DAWID, A. P. (1984). Present position and potential developments: Some personal views. Statistical theory. The prequential approach (with discussion). *J*. *Roy*. *Statist*. *Soc*. *Ser*. *A* **147** 278–292. MR0763811 https://doi.org/10.2307/2981683

[28] DE BLASI, P., FAVARO, S., LIJOI, A., MENA, R. H., PRÜNSTER, I. and RUGGIERO, M. (2015). Are Gibbs-type priors the most natural generalization of the Dirichlet process? *IEEE Trans*. *Pattern Anal*. *Mach*. *Intell*. **37** 212–229. https://doi.org/10.1109/TPAMI.2013.217

[29] DE FINETTI, B. (1931). Sul significato soggettivo della probabilità. *Fund*. *Math*. **17** 298–329.

[30] DE FINETTI, B. (1937). La prévision : Ses lois logiques, ses sources subjectives. *Ann*. *Inst*. *Henri Poincaré* **7** 1–68. MR1508036

[31] DE FINETTI, B. (1937). Sur la condition de "equivalence partielle". In *Colloque consacré á la théorie des probabilités*, *Université de Genéve*, 12–16 *octobre*, 1937 5–18. Hermann et C.ie, Paris. 1938-39, Actual. Sci. Industr. 730, vol. VI. [English translation: (1980) on the Condition of Partial Exchangeability. In *Studies in Inductive Logic and Probability, Vol. 2* 193–205].

[32] DE FINETTI, B. (1959). La probabilità e la statistica nei rapporti con l'induzione secondo i diversi punti di vista. In *Centro Internazionale Matematico estivo* (*CIME*), *Induzione e Statistica* 1–115. Cremonese. [English translation: (1972) *Probability, Induction and Statistics* 147–227. Wiley, New York].

[33] DE FINETTI, B. (1975). *Theory of Probability*: *A Critical Introductory Treatment*, *Vol*. 2. *Wiley Series in Probability and Mathematical Statistics*. Wiley, London. Translated from the Italian by Antonio Machí and Adrian Smith. MR0440641

[34] DEGROOT, M. H. and FIENBERG, S. E. (1982). The comparison and evaluation of forecasters. *Statistician* **32** 12–22.

[35] DIACONIS, P. (1988). Recent progress on de Finetti's notions of exchangeability. In *Bayesian Statistics* (*Valencia*, 1987), *Vol*. 3 (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.). *Oxford Sci*. *Publ*. 111–125. Oxford Univ. Press, New York. MR1008047

[36] DIACONIS, P. and FREEDMAN, D. (1980). de Finetti's theorem for Markov chains. *Ann*. *Probab*. **8** 115–130. MR0556418

[37] DIACONIS, P. and FREEDMAN, D. (1984). Partial exchangeability and sufficiency. In *Statistics*: *Applications and New Directions* (J. K. Ghosh and J. Roy, eds.). *Proceedings of the Indian Statistical Institute Golden Jubilee International Conference*, *Calcutta*, 1981 205–236. MR0786142

[38] DIACONIS, P. and FREEDMAN, D. (1990). On the uniform consistency of Bayes estimates for multinomial probabilities. *Ann*. *Statist*. **18** 1317–1327. MR1062710 https://doi.org/10.1214/aos/1176347751

[39] DIACONIS, P. and JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rend*. *Mat*. *Appl*. (7) **28** 33–61. MR2463439

[40] DIACONIS, P. and ROLLES, S. W. W. (2006). Bayesian analysis for reversible Markov chains. *Ann*. *Statist*. **34** 1270–1292. MR2278358 https://doi.org/10.1214/009053606000000290

[41] DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann*. *Statist*. **7** 269–281. MR0520238

[42] DOKSUM, K. (1974). Tailfree and neutral random probabilities and their posterior distributions. *Ann*. *Probab*. **2** 183–201. MR0373081 https://doi.org/10.1214/aop/1176996703

[43] DOLERA, E. and REGAZZINI, E. (2019). Uniform rates of the Glivenko–Cantelli convergence and their use in approximating Bayesian inferences. *Bernoulli* **25** 2982–3015. MR4003571 https://doi.org/10.3150/18-BEJ1077

[44] DOOB, J. L. (1949). Application of the theory of martingales. In *Le Calcul des Probabilités et Ses Applications*. *Colloq*. *Internat*. *CNRS*, *No*. 13 23–27. CNRS, Paris. MR0033460

[45] DUBINS, L. E. and SAVAGE, L. J. (1965). *How to Gamble If You Must*. *Inequalities for Stochastic Processes*. McGraw-Hill, New York. MR0236983

[46] EFRON, B. and HASTIE, T. (2016). *Computer Age Statistical Inference*. *Algorithms*, *Evidence*, *and Data Science*. *Institute of Mathematical Statistics* (*IMS*) *Monographs* **5**. Cambridge Univ. Press, New York. MR3523956 https://doi.org/10.1017/CBO9781316576533

[47] EGGENBERGER, F. and PÓLYA, G. (1923). Über die Statistik verketteter Vorgänge. *ZAMM Z*. *Angew*. *Math*. *Mech*. **3** 279–289.

[48] EWENS, W. J. (1972). The sampling theory of selectively neutral alleles. *Theor*. *Popul*. *Biol*. **3**. MR0325177 https://doi.org/10.1016/0040-5809(72)90035-4

[49] FASANO, A. and PETRONE, S. (2022). A predictive approach to Aldous–Hoover representation theorem for exchangeable random graphs. Manuscript.

[50] FONG, E., HOLMES, C. and WALKER, S. G. (2023). Martingale posterior distributions. *J*. *R*. *Stat*. *Soc*. *Ser*. *B*. *Stat*. *Methodol*. **85** 1357–1391. With discussion. MR4726953 https://doi.org/10.1093/jrsssb/qkad005

[51] FORTINI, S., LADELLI, L., PETRIS, G. and REGAZZINI, E. (2002). On mixtures of distributions of Markov chains. *Stochastic Process*. *Appl*. **100** 147–165. MR1919611 https://doi.org/10.1016/S0304-4149(02)00093-5

[52] FORTINI, S., LADELLI, L. and REGAZZINI, E. (2000). Exchangeability, predictive distributions and parametric models. *Sankhyā Ser. A* **62** 86–109. MR1769738

[53] FORTINI, S. and PETRONE, S. (2012). Hierarchical reinforced urn processes. *Statist. Probab. Lett.* **82** 1521–1529. MR2930656 https://doi.org/10.1016/j.spl.2012.04.012

[54] FORTINI, S. and PETRONE, S. (2016). Predictive distribution (de Finetti's view). In *Wiley StatsRef: Statistics Reference Online* 1–9. Wiley, New York.

[55] FORTINI, S. and PETRONE, S. (2017). Predictive characterization of mixtures of Markov chains. *Bernoulli* **23** 1538–1565. MR3624870 https://doi.org/10.3150/15-BEJ787

[56] FORTINI, S. and PETRONE, S. (2020). Quasi-Bayes properties of a procedure for sequential learning in mixture models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **82** 1087–1114. MR4136504

[57] FORTINI, S. and PETRONE, S. (2023). Prediction-based uncertainty quantification for exchangeable sequences. *Philos. Trans. R. Soc. A* **381** Paper No. 20220142, 17. MR4590505

[58] FORTINI, S. and PETRONE, S. (2025). Supplement to "Exchangeability, prediction and predictive modeling in Bayesian statistics." https://doi.org/10.1214/24-STS965SUPP

[59] FORTINI, S., PETRONE, S. and SARIEV, H. (2021). Predictive constructions based on measure-valued Pólya urn processes. *Mathematics* **9** 2845.

[60] FORTINI, S., PETRONE, S. and SPORYSHEVA, P. (2018). On a notion of partially conditionally identically distributed sequences. *Stochastic Process. Appl.* **128** 819–846. MR3758339 https://doi.org/10.1016/j.spa.2017.06.008

[61] GAEL, J.V. and GHAHRAMANI, Z. (2011). Nonparametric hidden Markov models. In *Bayesian Time Series Models* 317–340. Cambridge Univ. Press, Cambridge. MR2894245

[62] GEISSER, S. (1993). *Predictive Inference: An Introduction*. *Monographs on Statistics and Applied Probability* **55**. CRC Press, New York. MR1252174 https://doi.org/10.1007/978-1-4899-4467-2

[63] GOLDSTEIN, M. (1999). Bayes linear analysis. In *Encyclopedia of Statistical Sciences, Vol.* 3 (S. Kotz, C. B. Read and D. L. Banks, eds.) 29–34. Wiley, New York.

[64] GOOD, I. J. (1967). A Bayesian significance test for multinomial distributions. *J. Roy. Statist. Soc. Ser. B* **29** 399–431. With discussion. MR0225428

[65] GRIFFITHS, T. L. and GHAHRAMANI, Z. (2005). Infinite latent feature models and the Indian Buffet process. In *Advances in Neural Information Processing Systems* (*NIPS*) 475–482.

[66] HAHN, P. R., MARTIN, R. and WALKER, S. G. (2018). On recursive Bayesian predictive distributions. *J. Amer. Statist. Assoc.* **113** 1085–1093. MR3862341 https://doi.org/10.1080/01621459.2017.1304219

[67] HOLMES, C. C. and WALKER, S. G. (2023). Statistical inference with exchangeability and martingales. *Philos. Trans. R. Soc. A* **381** Paper No. 20220143, 17. MR4590508

[68] HOOVER, D. N. (1979). Relations on probability spaces and arrays of random variables. Technical Report, Institute of Advanced Study, Princeton.

[69] HOPPE, F. M. (1984). Pólya-like urns and the Ewens' sampling formula. *J. Math. Biol.* **20** 91–94. MR0758915 https://doi.org/10.1007/BF00275863

[70] HOPPE, F. M. (1987). The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.* **25** 123–159. MR0896430 https://doi.org/10.1007/BF00276386

[71] HOROWITZ, J. (1985). Measure-valued random processes. *Z. Wahrsch. Verw. Gebiete* **70** 213–236. MR0799147 https://doi.org/10.1007/BF02451429

[72] ISHWARAN, H. and JAMES, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96** 161–173. MR1952729 https://doi.org/10.1198/016214501750332758

[73] JAMES, L. F. (2017). Bayesian Poisson calculus for latent feature modeling via generalized Indian Buffet process priors. *Ann. Statist.* **45** 2016–2045. MR3718160 https://doi.org/10.1214/16-AOS1517

[74] JOHNSON, W. E. (1932). Probability: The deductive and inductive problems. *Mind* **41** 409–423.

[75] KALLENBERG, O. (1988). Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* **16** 508–534. MR0929061

[76] KALLENBERG, O. (1999). Multivariate sampling and the estimation problem for exchangeable arrays. *J. Theoret. Probab.* **12** 859–883. MR1702867 https://doi.org/10.1023/A:1021692202530

[77] KALLENBERG, O. (2002). *Foundations of Modern Probability*, 2nd ed. *Probability and Its Applications* (*New York*). Springer, New York. MR1876169 https://doi.org/10.1007/978-1-4757-4015-8

[78] KALLENBERG, O. (2005). *Probabilistic Symmetries and Invariance Principles*. *Probability and Its Applications* (*New York*). Springer, New York. MR2161313

[79] KALMAN, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.* **82** 35–45. MR3931993

[80] KINGMAN, J. F. C. (1978). Uses of exchangeability. *Ann. Probab.* **6** 183–197. MR0494344 https://doi.org/10.1214/aop/1176995566

[81] KINGMAN, J. F. C. (1978). The representation of partition structures. *J. Lond. Math. Soc.* (2) **18** 374–380. MR0509954 https://doi.org/10.1112/jlms/s2-18.2.374

[82] KINGMAN, J. F. C. (1980). *Mathematics of Genetic Diversity*. *CBMS-NSF Regional Conference Series in Applied Mathematics* **34**. SIAM, Philadelphia, PA. MR0591166

[83] LAURITZEN, S. L. (1984). Extreme point models in statistics. *Scand. J. Stat.* **11** 65–91. With discussion. MR0754971

[84] LAURITZEN, S. L. (1988). *Extremal Families and Systems of Sufficient Statistics*. *Lecture Notes in Statistics* **49**. Springer, New York. MR0971253 https://doi.org/10.1007/978-1-4612-1023-8

[85] LIJOI, A., NIPOTI, B. and PRÜNSTER, I. (2014). Bayesian inference with dependent normalized completely random measures. *Bernoulli* **20** 1260–1291. MR3217444 https://doi.org/10.3150/13-BEJ521

[86] LO, A. Y. (1988). A Bayesian bootstrap for a finite population. *Ann. Statist.* **16** 1684–1695. MR0964946 https://doi.org/10.1214/aos/1176351061

[87] LO, A. Y. (1991). A characterization of the Dirichlet process. *Statist. Probab. Lett.* **12** 185–187. MR1130354 https://doi.org/10.1016/0167-7152(91)90075-3

[88] LOVÁSZ, L. (2014). Large networks and graph limits. *Bull. Amer. Math. Soc.* (*N.S.*). MR3363148 https://doi.org/10.1090/S0273-0979-2014-01455-7

[89] LOVÁSZ, L. and SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combin. Theory Ser. B* **96** 933–957. MR2274085 https://doi.org/10.1016/j.jctb.2006.05.002

[90] MOYA, B. and WALKER, S. G. (2023). Martingale posterior distributions for time series models. *Statist. Sci.* 1–15.

[91] MULIERE, P., SECCHI, P. and WALKER, S. G. (2000). Urn schemes and reinforced random walks. *Stochastic Process. Appl.* **88** 59–78. MR1761692 https://doi.org/10.1016/S0304-4149(99)00119-2

[92] MULIERE, P. and WALKER, S. (1998). Extending the family of Bayesian bootstraps and exchangeable urn schemes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 175–182. MR1625636 https://doi.org/10.1111/1467-9868.00116
[93] NEWTON, M. A. (2002). On a nonparametric recursive estimator of the mixing distribution. *Sankhyā* **64** 306–322. MR1981761
[94] NEWTON, M. A., QUINTANA, F. A. and ZHANG, Y. (1998). Nonparametric Bayes methods using predictive updating. In *Practical Nonparametric and Semiparametric Bayesian Statistics*. *Lect. Notes Stat.* **133** 45–61. Springer, New York. MR1630075 https://doi.org/10.1007/978-1-4612-1732-9_3
[95] NEWTON, M. A. and ZHANG, Y. (1999). A recursive algorithm for nonparametric analysis with missing data. *Biometrika* **86** 15–26. MR1688068 https://doi.org/10.1093/biomet/86.1.15
[96] ORBANZ, P. and ROY, D. M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 437–461.
[97] ORBANZ, P. and SZEGEDY, B. (2016). Borel liftings of graph limits. *Electron. Commun. Probab.* **21** Paper No. 65, 4. MR3548777 https://doi.org/10.1214/16-ECP14
[98] PARMIGIANI, G. and INOUE, L. Y. T. (2009). *Decision Theory: Principles and Approaches*. *Wiley Series in Probability and Statistics*. Wiley, Chichester. MR2604978 https://doi.org/10.1002/9780470746684
[99] PEMANTLE, R. (2007). A survey of random processes with reinforcement. *Probab. Surv.* **4** 1–79. MR2282181 https://doi.org/10.1214/07-PS094
[100] PERMAN, M., PITMAN, J. and YOR, M. (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Related Fields* **92** 21–39. MR1156448 https://doi.org/10.1007/BF01205234
[101] PETRONE, S. (1999). Random Bernstein polynomials. *Scand. J. Stat.* **26** 373–393. MR1712051 https://doi.org/10.1111/1467-9469.00155
[102] PETRONE, S. and VERONESE, P. (2010). Feller operators and mixture priors in Bayesian nonparametrics. *Statist. Sinica* **20** 379–404. MR2640700
[103] PITMAN, J. (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Related Fields* **102** 145–158. MR1337249 https://doi.org/10.1007/BF01213386
[104] PITMAN, J. (1996). Some developments of the Blackwell–MacQueen urn scheme. In *Statistics, Probability and Game Theory* (T. S. Ferguson, L. S. Shapley and J. B. MacQueen, eds.). *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **30** 245–267. IMS, Hayward, CA. MR1481784 https://doi.org/10.1214/lnms/1215453576
[105] PITMAN, J. (2006). *Combinatorial Stochastic Processes*. *Lecture Notes in Math.* **1875**. Springer, Berlin. Lectures from the 32nd Summer School on Probability Theory held in Saint-Flour, July 7–24, 2002. With a foreword by Jean Picard. MR2245368
[106] PITMAN, J. and YOR, M. (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25** 855–900. MR1434129 https://doi.org/10.1214/aop/1024404422
[107] PÓLYA, G. (1930). Sur quelques points de la théorie des probabilités. *Ann. Inst. Henri Poincaré* **1** 117–161. MR1507985
[108] RAMSEY, F. P. (1926). Truth and probability. In *Philosophy of Probability: Contemporary Readings* (A. Eagle, ed.) 52–94. Routledge, London.
[109] REGAZZINI, E. (1999). Old and recent results on the relationship between predictive inference and statistical modelling either in nonparametric or parametric form. In *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting June* 6–10, 1998. Oxford Univ. Press, New York. MR1724874
[110] ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400–407. MR0042668 https://doi.org/10.1214/aoms/1177729586
[111] SARIEV, H., FORTINI, S. and PETRONE, S. (2023). Infinite-color randomly reinforced urns with dominant colors. *Bernoulli* **29** 132–152. MR4497242 https://doi.org/10.3150/21-bej1452
[112] SARIEV, H. and SAVOV, M. (2024). Characterization of exchangeable measure-valued Pólya urn sequences. *Electron. J. Probab.* **29** Paper No. 73, 23. MR4743435 https://doi.org/10.1214/24-ejp1132
[113] SARIEV, H. and SAVOV, M. (2025). Sufficientness postulates for measure-valued Pólya urn sequences. *J. Appl. Probab.* In press.
[114] SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York. MR0063582
[115] SCHERVISH, M. J. (1995). *Theory of Statistics*. *Springer Series in Statistics*. Springer, New York. MR1354146 https://doi.org/10.1007/978-1-4612-4250-5
[116] SHMUELI, G. (2010). To explain or to predict? *Statist. Sci.* **25** 289–310. MR2791669 https://doi.org/10.1214/10-STS330
[117] SMALE, S. and YAO, Y. (2006). Online learning algorithms. *Found. Comput. Math.* **6** 145–170. MR2228737 https://doi.org/10.1007/s10208-004-0160-z
[118] SMITH, A. F. M. and MAKOV, U. E. (1978). A quasi-Bayes sequential procedure for mixtures. *J. Roy. Statist. Soc. Ser. B* **40** 106–112. MR0512148
[119] TEH, Y. W. (2006). A hierarchical Bayesian language model based on Pitman–Yor processes. In *Proceedings of the International Conference on Computational Linguistics* 985–992.
[120] TEH, Y. W. and JORDAN, M. I. (2010). Hierarchical Bayesian nonparametric models with applications. In *Bayesian Nonparametrics* (N. L. Hjort, C. C. Holmes, P. Muller and S. G. Walker, eds.). *Camb. Ser. Stat. Probab. Math.* **28** 158–207. Cambridge Univ. Press, Cambridge. MR2730663
[121] TEH, Y. W., JORDAN, M. I., BEAL, M. J. and BLEI, D. M. (2006). Hierarchical Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 1566–1581. MR2279480 https://doi.org/10.1198/016214506000000302
[122] THIBAUX, R. and JORDAN, M. I. (2007). Hierarchical beta processes and the Indian Buffet process. In *AISTATS Proceedings* 564–571.
[123] VON PLATO, J. (1982). The generalization of de Finetti's representation theorem to stationary probabilities. In *Philosophy of Science Association: Proceedings of the Biennial Meeting of the Philosophy of Science* 137–144 Univ. Chicago Press, Chicago.
[124] WADE, S., MONGELLUZZO, S. and PETRONE, S. (2011). An enriched conjugate prior for Bayesian nonparametric inference. *Bayesian Anal.* **6** 359–385. MR2843536 https://doi.org/10.1214/ba/1339616468
[125] WADE, S., WALKER, S. G. and PETRONE, S. (2014). A predictive study of Dirichlet process mixture models for curve fitting. *Scand. J. Stat.* **41** 580–605. MR3249418 https://doi.org/10.1111/sjos.12047
[126] WALKER, S. and MULIERE, P. (1997). Beta-Stacy processes and a generalization of the Pólya-urn scheme. *Ann. Statist.* **25** 1762–1780. MR1463574 https://doi.org/10.1214/aos/1031594741
[127] WALKER, S. and MULIERE, P. (1999). A characterization of a neutral to the right prior via an extension of Johnson's sufficientness postulate. *Ann. Statist.* **27** 589–599. MR1714716 https://doi.org/10.1214/aos/1018031208
[128] WILLIAMS, P. M. (1976). Indeterminate probabilities. In *Formal Methods in the Methodology of Empirical Sciences* (*Proc.*

Conf., *Warsaw*, 1974). *Synthese Lib.* **103** 229–246. Reidel, Dordrecht. MR0535905

[129] YIU, A., FONG, E., WALKER, S. and HOLMES, C. (2022). Causal predictive inference and target trial emulation. arXiv preprint. Available at arXiv:2207.12479.

[130] ZABELL, S. L. (1982). W. E. Johnson's "sufficientness" postulate. *Ann. Statist.* **10** 1090–1099. MR0673645

[131] ZABELL, S. L. (1995). Characterizing Markov exchangeable sequences. *J. Theoret. Probab.* **8** 175–178. MR1308676 https://doi.org/10.1007/BF02213460

[132] ZABELL, S. L. (1997). The continuum of inductive methods revisited. In *The Cosmos of Science: Essays in Exploration* (J. Earman and J. D. Norton, eds.). Universty of Pittsburgh Press, Pittsburgh.

[133] ZABELL, S. L. (2005). *Symmetry and Its Discontents: Essays on the History of Inductive Probability. Cambridge Studies in Probability, Induction, and Decision Theory*. Cambridge Univ. Press, New York. With a preface by Brian Skyrms. MR2199124 https://doi.org/10.1017/CBO9780511614293

[134] ZAMAN, A. (1984). Urn models for Markov exchangeability. *Ann. Probab.* **12** 223–229. MR0723741

Statistical Science
2025, Vol. 40, No. 1, 68–80
https://doi.org/10.1214/23-STS914

Martingale Posterior Distributions for Time-Series Models

Blake Moya and Stephen G. Walker

Abstract. The paper considers the analysis of time-series data from a future missing data perspective. Using predictive models with random multistep, updating of the parameters is equivalent to an iterative ensemble approach where initial conditions are randomized. The output is a sample from a posterior distribution equivalent in structure to the martingale posterior distribution. Specifically, the paper provides a posterior distribution of the parameters of a time-series model by forward sampling from the data using a predictive model based on iteratively sampling and updating the model until the sequence of parameter estimators has converged. Repetition of this produces independent samples from the posterior. Illustrative and real data examples are provided.

Key words and phrases: Martingale, missing future data, multistep forecasting, score function.

REFERENCES

[1] BERGER, J. (2006). The case for objective Bayesian analysis. *Bayesian Anal*. **1** 385–402. MR2221271 https://doi.org/10.1214/06-BA115

[2] BERNARDO, J.-M. and SMITH, A. F. M. (1994). *Bayesian Theory*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

[3] BERTI, P., PRATELLI, L. and RIGO, P. (2006). Almost sure weak convergence of random probability measures. *Stochastics* **78** 91–97. MR2236634 https://doi.org/10.1080/17442500600745359

[4] BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann*. *Statist*. **1** 353–355. MR0362614

[5] BORNER, K., ROUSE, W. B., TRUNIFO, P. and STANLEY, H. E. (2018). Forecasting innovations in science, technology, and education. *Proc*. *Natl*. *Acad*. *Sci*. **115** 12573–12581.

[6] CHATFIELD, C. (2004). *The Analysis of Time Series*: *An Introduction*, 6th ed. *Chapman & Hall/CRC Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR2001158

[7] CHENG, H., TAN, P. N., GAO, J. and SCRIPPS, J. (2006). Multistep–ahead time series prediction. In *PAKDD* 2006: *Advances in Knowledge Discovery and Data Mining* (W. K. Ng, M. Kitsuregawa, J. Li and K. Chang, eds.). *Lecture Notes in Computer Science* **3918**. Springer, Berlin.

[8] COX, D. R. (2006). *Principles of Statistical Inference*. Cambridge Univ. Press, Cambridge. MR2278763 https://doi.org/10.1017/CBO9780511813559

[9] DE FINETTI, B. (1937). Foresight: Its logical laws, its subjective sources. *Ann*. *Inst*. *Henri Poincaré* **7** 1–68. (English translation by H. E. Kyburg, Jr. and H. E. Smokier, eds. in Studies in Subjective Probability. (2nd ed. 1980), Robert E. Krieger, Huntington, New York.)

[10] DINOV, I. D. (2018). *Data Science and Predictive Analytics*: *Biomedical and Health Applications Using R*. Springer, Cham. MR3889981

[11] DOOB, J. L. (1949). Application of the theory of martingales. In *Le Calcul des Probabilités et Ses Applications*. *Colloq*. *Internat*. **13** 23–27. CNRS, Paris. MR0033460

[12] EFRON, B. (1979). Bootstrap methods: Another look at the jackknife. *Ann*. *Statist*. **7** 1–26. MR0515681

[13] EFRON, B. (2012). Bayesian inference and the parametric bootstrap. *Ann*. *Appl*. *Stat*. **6** 1971–1997. MR3058690 https://doi.org/10.1214/12-AOAS571

[14] ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007. MR0666121 https://doi.org/10.2307/1912773

[15] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann*. *Statist*. **1** 209–230. MR0350949

[16] FONG, E., HOLMES, C. and WALKER, S. G. (2024). Margingale posterior distributions. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B*. To appear.

[17] HEWITT, E. and SAVAGE, L. J. (1955). Symmetric measures on Cartesian products. *Trans*. *Amer*. *Math*. *Soc*. **80** 470–501. MR0076206 https://doi.org/10.2307/1992999

[18] HOLMES, E. E. (2013). Derivation of an EM algorithm for constrained and unconstrained multivariate autoregressive state-space (MARSS) models. arXiv preprint. Available at arXiv:1302.3919.

[19] HOLMES, E. E., SCHEUERELL, M. D. and WARD, E. J. (2021). *Applied Time Series Analysis for Fisheries and Environmental Data*. Contacts eeholmes@uw.edu, warde@uw.edu, and scheuerl@uw.edu.

Blake Moya is Ph.D. Student, Department of Statistics and Data Sciences, University of Texas at Austin, Austin, Texas 78712, USA (e-mail: BlakeMoya@utexas.edu). Stephen G. Walker is Professor, Department of Statistics and Data Sciences, University of Texas at Austin, Austin, Texas 78712, USA (e-mail: S.G.Walker@math.utexas.edu).

[20] HOLMES, E. E., WARD, E. J. and KELLIE, W. (2012). MARSS: Multivariate autoregressive state-space models for analyzing time-series data. *R J.* **4** 11–19. https://doi.org/10.32614/RJ-2012-002

[21] LATZ, J. (2021). Analysis of stochastic gradient descent in continuous time. *Stat. Comput.* **31** Paper No. 39, 25. MR4257846 https://doi.org/10.1007/s11222-021-10016-8

[22] LAWLER, G. F. (2006). *Introduction to Stochastic Processes*, 2nd ed. CRC Press, Boca Raton, FL. MR2255511

[23] LIJOI, A., PRÜNSTER, I. and WALKER, S. G. (2007). Bayesian consistency for stationary models. *Econometric Theory* **23** 749–759. MR2364385 https://doi.org/10.1017/S0266466607070314

[24] MAGNUSSON, L., HUEY, J.-H., LIN, S.-J., ZHOU, L. and CHIN, X. (2019). Dependence on initial conditions versus model formulations for medium range forecast error variations. *Q. J. R. Meteorol. Soc.* **145** 2085–2100.

[25] MAITRA, A. (1977). Integral representations of invariant measures. *Trans. Amer. Math. Soc.* **229** 209–225. MR0442197 https://doi.org/10.2307/1998506

[26] MOYA, B. and Stephen G. Walker (2023). Supplement to "Martingale posterior distributions for time-series models." https://doi.org/10.1214/23-STS914SUPP

[27] PITT, M. K., CHATFIELD, C. and WALKER, S. G. (2002). Constructing first order stationary autoregressive models via latent processes. *Scand. J. Stat.* **29** 657–663. MR1988417 https://doi.org/10.1111/1467-9469.00311

[28] RUBIN, D. B. (1981). The Bayesian bootstrap. *Ann. Statist.* **9** 130–134. MR0600538

[29] SARMA, A. and KAY, M. (2020). Prior Setting in Practice: Strategies and Rationales Used in Choosing Prior Distributions for Bayesian Analysis. OSF Preprints.

[30] SLINGO, J. and PALMER, T. (2011). Uncertainty in weather and climate prediction. *Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **369** 4751–4767.

[31] STAN DEVELOPMENT TEAM (2022). Stan modeling language users guide and reference manual, 2.30, https://mc-stan.org.

[32] WU, H., MENG, K., FAN, D., ZHANG, Z. and LIU, Q. (2022). Multistep short–term wind speed forecasting using transformer. *Energy* **261** 125231.

Statistical Science
2025, Vol. 40, No. 1, 81–108
https://doi.org/10.1214/24-STS966

Bayesian Dependent Mixture Models: A Predictive Comparison and Survey

Sara Wade and Vanda Inácio

Abstract. For exchangeable data, mixture models are an extremely useful tool for density estimation due to their attractive balance between smoothness and flexibility. When additional covariate information is present, mixture models can be extended for flexible regression by modeling the mixture parameters, namely, the weights and atoms, as functions of the covariates. These types of models are interpretable and highly flexible, allowing not only the mean but the whole density of the response to change with the covariates. This article reviews Bayesian covariate-dependent mixture models, focusing on nonparametric constructions which we broadly organize into into three categories: (1) joint models of the responses and covariates, (2) conditional models with single-weights and covariate-dependent atoms, and (3) conditional models with covariate-dependent weights. The diversity and variety of the available models in the literature raises the question of how to choose among them for the application at hand. We attempt to shed light on this question through a careful analysis of the predictive equations for the conditional mean and density function as well as predictive comparisons in a variety of simulated data examples.

Key words and phrases: Density regression, dependent Dirichlet process, mixture of experts, nonparametric regression, stick-breaking representation.

REFERENCES

[1] ABDULSAMAD, H., NICKL, P., KLINK, P. and PETERS, J. (2021). A variational infinite mixture for probabilistic inverse dynamics learning. In 2021 *IEEE International Conference on Robotics and Automation* (*ICRA*) 4216–4222. IEEE Press, New York.

[2] ALBERT, J. H. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. MR1224394

[3] AMBROGIONI, L., GÜÇLÜ, U., VAN GERVEN, M. A. and MARIS, E. (2017). The kernel mixture network: a nonparametric method for conditional density estimation of continuous random variables. ArXiv preprint. Available at arXiv:1705.07111.

[4] AN, Q., WANG, C., SHTEREV, I., WANG, E., CARIN, L. and DUNSON, D. B. (2008). Hierarchical kernel stick-breaking process for multi-task image analysis. In *Proceedings of the* 25*th International Conference on Machine Learning* 17–24.

[5] ANTONIAK, C. E. (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *Ann. Statist.* **2** 1152–1174. MR0365969

[6] ANTONIANO-VILLALOBOS, I., WADE, S. and WALKER, S. G. (2014). A Bayesian nonparametric regression model with normalized weights: A study of hippocampal atrophy in Alzheimer's disease. *J. Amer. Statist. Assoc.* **109** 477–490. MR3223726 https://doi.org/10.1080/01621459.2013.879061

[7] BANDYOPADHYAY, D. and CANALE, A. (2016). Nonparametric spatial models for clustered ordered periodontal data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **65** 619–640. MR3522956 https://doi.org/10.1111/rssc.12150

[8] BAO, J. and HANSON, T. E. (2015). Bayesian nonparametric multivariate ordinal regression. *Canad. J. Statist.* **43** 337–357. MR3388321 https://doi.org/10.1002/cjs.11253

[9] BARRIENTOS, A. F., JARA, A. and QUINTANA, F. A. (2012). On the support of MacEachern's dependent Dirichlet processes and extensions. *Bayesian Anal.* **7** 277–309. MR2934952 https://doi.org/10.1214/12-BA709

[10] BISHOP, C. M. (1994). Mixture density networks. Technical Report, Aston University.

[11] BLACKWELL, D. and MACQUEEN, J. B. (1973). Ferguson distributions via Pólya urn schemes. *Ann. Statist.* **1** 353–355. MR0362614

[12] BLEI, D. M. and JORDAN, M. I. (2006). Variational inference for Dirichlet process mixtures. *Bayesian Anal.* **1** 121–143. MR2227367 https://doi.org/10.1214/06-BA104

[13] BLEI, D. M., NG, A. Y. and JORDAN, M. I. (2003). Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3** 993–1022.

[14] BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. and STONE, C. J. (1984). *Classification and Regression Trees*.

Sara Wade is Reader (Associate Prof.), School of Mathematics, University of Edinburgh, Edinburgh, UK (e-mail: sara.wade@ed.ac.uk). Vanda Inácio is Reader (Associate Prof.), School of Mathematics, University of Edinburgh, Edinburgh, UK (e-mail: vanda.inacio@ed.ac.uk).

Wadsworth Statistics/Probability Series. Wadsworth Advanced Books and Software, Belmont, CA. MR0726392

[15] BURGETTE, L. F. and REITER, J. P. (2012). Nonparametric Bayesian multiple imputation for missing data due to mid-study switching of measurement methods. *J. Amer. Statist. Assoc.* **107** 439–449. MR2980056 https://doi.org/10.1080/01621459.2011.643713

[16] CAI, J.-H., SONG, X.-Y., LAM, K.-H. and IP, E. H.-S. (2011). A mixture of generalized latent variable models for mixed mode and heterogeneous data. *Comput. Statist. Data Anal.* **55** 2889–2907. MR2813054 https://doi.org/10.1016/j.csda.2011.05.011

[17] CANALE, A. and DE BLASI, P. (2017). Posterior asymptotics of nonparametric location-scale mixtures for multivariate density estimation. *Bernoulli* **23** 379–404. MR3556776 https://doi.org/10.3150/15-BEJ746

[18] CANALE, A. and DUNSON, D. B. (2011). Bayesian kernel mixtures for counts. *J. Amer. Statist. Assoc.* **106** 1528–1539. MR2896854 https://doi.org/10.1198/jasa.2011.tm10552

[19] CANALE, A. and PRÜNSTER, I. (2017). Robustifying Bayesian nonparametric mixtures for count data. *Biometrics* **73** 174–184. MR3632363 https://doi.org/10.1111/biom.12538

[20] CARON, F., DAVY, M., DOUCET, A., DUFLOS, E. and VANHEEGHE, P. (2008). Bayesian inference for linear dynamic models with Dirichlet process mixtures. *IEEE Trans. Signal Process.* **56** 71–84. MR2439814 https://doi.org/10.1109/TSP.2007.900167

[21] CAROTA, C. and PARMIGIANI, G. (2002). Semiparametric regression for count data. *Biometrika* **89** 265–281. MR1913958 https://doi.org/10.1093/biomet/89.2.265

[22] CARVALHO, C. M., LOPES, H. F., POLSON, N. G. and TADDY, M. A. (2010). Particle learning for general mixtures. *Bayesian Anal.* **5** 709–740. MR2740154 https://doi.org/10.1214/10-BA525

[23] CHANDRA, N. K., CANALE, A. and DUNSON, D. B. (2023). Escaping the curse of dimensionality in Bayesian model-based clustering. *J. Mach. Learn. Res.* **24** Paper No. 144. MR4596091

[24] CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2010). BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4** 266–298. MR2758172 https://doi.org/10.1214/09-AOAS285

[25] CHUNG, Y. and DUNSON, D. B. (2009). Nonparametric Bayes conditional distribution modeling with variable selection. *J. Amer. Statist. Assoc.* **104** 1646–1660. MR2750582 https://doi.org/10.1198/jasa.2009.tm08302

[26] CHUNG, Y. and DUNSON, D. B. (2011). The local Dirichlet process. *Ann. Inst. Statist. Math.* **63** 59–80. MR2748934 https://doi.org/10.1007/s10463-008-0218-9

[27] CIFARELLI, D. M., MULIERE, P. and SCARSINI, M. (1981). Il modello lineare nell'approccio Bayesiano non parametrico. *Istituto Matematico G. Castelnuovo, Universitá degli Studi di Roma La Sapienza* **15**.

[28] CIFARELLI, D. M. and REGAZZINI, E. (1978). Problemi Statistici Nonparametrici in Condizioni di Scambiabilità Parziale e Impiego di Medie Associative. *Quaderni Istituto di Matematica Finanziaria, Università di Torino* **12** 1–36.

[29] DANIELS, M. J., LEE, M. and FENG, W. (2023). Dirichlet process mixture models for the analysis of repeated attempt designs. *Biometrics* **79** 3907–3915. MR4680769 https://doi.org/10.1111/biom.13894

[30] DE IORIO, M., JOHNSON, W. O., MÜLLER, P. and ROSNER, G. L. (2009). Bayesian nonparametric nonproportional hazards survival modeling. *Biometrics* **65** 762–771. MR2649849 https://doi.org/10.1111/j.1541-0420.2008.01166.x

[31] DE IORIO, M., MÜLLER, P., ROSNER, G. L. and MACEACHERN, S. N. (2004). An ANOVA model for dependent random measures. *J. Amer. Statist. Assoc.* **99** 205–215. MR2054299 https://doi.org/10.1198/016214504000000205

[32] DENISON, D. G. T., HOLMES, C. C., MALLICK, B. K. and SMITH, A. F. M. (2002). *Bayesian Methods for Nonlinear Classification and Regression. Wiley Series in Probability and Statistics*. Wiley, Chichester. MR1962778

[33] DEVICK, K. L., VALERI, L., CHEN, J., JARA, A., BIND, M.-A. and COULL, B. A. (2022). The role of body mass index at diagnosis of colorectal cancer on Black–White disparities in survival: A density regression mediation approach. *Biostatistics* **23** 449–466. MR4424980 https://doi.org/10.1093/biostatistics/kxaa034

[34] DEYOREO, M. and KOTTAS, A. (2015). A fully nonparametric modeling approach to binary regression. *Bayesian Anal.* **10** 821–847. MR3432241 https://doi.org/10.1214/15-BA963SI

[35] DEYOREO, M. and KOTTAS, A. (2017). A Bayesian nonparametric Markovian model for non-stationary time series. *Stat. Comput.* **27** 1525–1538. MR3687324 https://doi.org/10.1007/s11222-016-9702-x

[36] DEYOREO, M. and KOTTAS, A. (2018). Bayesian nonparametric modeling for multivariate ordinal regression. *J. Comput. Graph. Statist.* **27** 71–84. MR3788302 https://doi.org/10.1080/10618600.2017.1316280

[37] DEYOREO, M. and KOTTAS, A. (2018). Modeling for dynamic ordinal regression relationships: An application to estimating maturity of rockfish in California. *J. Amer. Statist. Assoc.* **113** 68–80. MR3803440 https://doi.org/10.1080/01621459.2017.1328357

[38] DE BLASI, P., FAVARO, S., LIJOI, A., MENA, R. H., PRÜNSTER, I. and RUGGIERO, M. (2013). Are Gibbs-type priors the most natural generalization of the Dirichlet process? *IEEE Trans. Pattern Anal. Mach. Intell.* **37** 212–229.

[39] DE LA CRUZ-MESÍA, R., QUINTANA, F. A. and MÜLLER, P. (2007). Semiparametric Bayesian classification with longitudinal markers. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **56** 119–137. MR2359237 https://doi.org/10.1111/j.1467-9876.2007.00569.x

[40] DI LUCCA, M. A., GUGLIELMI, A., MÜLLER, P. and QUINTANA, F. A. (2013). A simple class of Bayesian nonparametric autoregression models. *Bayesian Anal.* **8** 63–87. MR3036254 https://doi.org/10.1214/13-BA803

[41] DING, D. and KARABATSOS, G. (2021). Dirichlet process mixture models with shrinkage prior. *Stat* **10** Paper No. e371. MR4235616 https://doi.org/10.1002/sta4.371

[42] DUAN, J. A., GUINDANI, M. and GELFAND, A. E. (2007). Generalized spatial Dirichlet process models. *Biometrika* **94** 809–825. MR2416794 https://doi.org/10.1093/biomet/asm071

[43] DUNN, P. K. and SMYTH, G. K. (1996). Randomized quantile residuals. *J. Comput. Graph. Statist.* **5** 236–244.

[44] DUNSON, D. and BHATTACHARYA, A. (2010). Nonparametric Bayes regression and classification through mixtures of product kernels. *Bayesian Stat.* **9** 145–144.

[45] DUNSON, D. B. and HERRING, A. H. (2006). Semiparametric Bayesian latent trajectory models. Technical Report, ISDS Discussion Paper 16, Duke University.

[46] DUNSON, D. B. and PARK, J.-H. (2008). Kernel stick-breaking processes. *Biometrika* **95** 307–323. MR2521586 https://doi.org/10.1093/biomet/asn012

[47] DUNSON, D. B. and PEDDADA, S. D. (2008). Bayesian nonparametric inference on stochastic ordering. *Biometrika* **95** 859–874. MR2461216 https://doi.org/10.1093/biomet/asn043

[48] DUNSON, D. B., PILLAI, N. and PARK, J.-H. (2007). Bayesian density regression. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*

69 163–183. MR2325270 https://doi.org/10.1111/j.1467-9868.2007.00582.x

[49] ETIENAM, C., LAW, K. and WADE, S. (2020). Ultra-fast Deep Mixtures of Gaussian Process Experts. ArXiv preprint. Available at arXiv:2006.13309.

[50] FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230. MR0350949

[51] FERGUSON, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629. MR0438568

[52] FOTI, N. and WILLIAMSON, S. (2012). Slice sampling normalized kernel-weighted completely random measure mixture models. In *Advances in Neural Information Processing Systems* (F. Pereira, C. J. Burges, L. Bottou and K. Q. Weinberger, eds.) **25**. Curran Associates, Red Hook .

[53] FRONCZYK, K. and KOTTAS, A. (2014). A Bayesian nonparametric modeling framework for developmental toxicity studies. *J. Amer. Statist. Assoc.* **109** 873–888. MR3265660 https://doi.org/10.1080/01621459.2013.830445

[54] FRONCZYK, K. and KOTTAS, A. (2014). A Bayesian approach to the analysis of quantal bioassay studies using nonparametric mixture models. *Biometrics* **70** 95–102. MR3251670 https://doi.org/10.1111/biom.12120

[55] FRONCZYK, K. and KOTTAS, A. (2017). Risk assessment for toxicity experiments with discrete and continuous outcomes: A Bayesian nonparametric approach. *J. Agric. Biol. Environ. Stat.* **22** 585–601. MR3736563 https://doi.org/10.1007/s13253-017-0293-6

[56] FRUHWIRTH-SCHNATTER, S., CELEUX, G. and ROBERT, C. P. (2019). *Handbook of Mixture Analysis*. Taylor and Francis/CRC Press.

[57] FRÜHWIRTH-SCHNATTER, S. and PYNE, S. (2010). Bayesian inference for finite mixtures of univariate and multivariate skew-normal and skew-t distributions. *Biostatistics* **11** 317–336. https://doi.org/10.1093/biostatistics/kxp062

[58] GADD, C., WADE, S. and BOUKOUVALAS, A. (2020). Enriched mixtures of generalised Gaussian process experts. In *Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics* (S. Chiappa and R. Calandra, eds.). *Proceedings of Machine Learning Research* **108** 3144–3154. PMLR.

[59] GELFAND, A. E., KOTTAS, A. and MACEACHERN, S. N. (2005). Bayesian nonparametric spatial modeling with Dirichlet process mixing. *J. Amer. Statist. Assoc.* **100** 1021–1035. MR2201028 https://doi.org/10.1198/016214504000002078

[60] GEWEKE, J. and KEANE, M. (2007). Smoothly mixing regressions. *J. Econometrics* **138** 252–290. MR2380699 https://doi.org/10.1016/j.jeconom.2006.05.022

[61] GHOSAL, S., GHOSH, J. K. and RAMAMOORTHI, R. V. (1999). Posterior consistency of Dirichlet mixtures in density estimation. *Ann. Statist.* **27** 143–158. MR1701105 https://doi.org/10.1214/aos/1018031105

[62] GHOSAL, S. and VAN DER VAART, A. (2007). Posterior convergence rates of Dirichlet mixtures at smooth densities. *Ann. Statist.* **35** 697–723. MR2336864 https://doi.org/10.1214/009053606000001271

[63] GHOSAL, S. and VAN DER VAART, A. (2017). *Fundamentals of Nonparametric Bayesian Inference*. *Cambridge Series in Statistical and Probabilistic Mathematics* **44**. Cambridge Univ. Press, Cambridge. MR3587782 https://doi.org/10.1017/9781139029834

[64] GHOSAL, S. and VAN DER VAART, A. W. (2001). Entropies and rates of convergence for maximum likelihood and Bayes estimation for mixtures of normal densities. *Ann. Statist.* **29** 1233–1263. MR1873329 https://doi.org/10.1214/aos/1013203453

[65] GIUDICI, P., MEZZETTI, M. and MULIERE, P. (2003). Mixtures of products of Dirichlet processes for variable selection in survival analysis. *J. Statist. Plann. Inference* **111** 101–115.

[66] GOODFELLOW, I., BENGIO, Y. and COURVILLE, A. (2016). *Deep Learning*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. MR3617773

[67] GOODMAN, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61** 215–231. MR0370936 https://doi.org/10.1093/biomet/61.2.215

[68] GRAMACY, R. B. and LEE, H. K. H. (2008). Bayesian treed Gaussian process models with an application to computer modeling. *J. Amer. Statist. Assoc.* **103** 1119–1130. MR2528830 https://doi.org/10.1198/016214508000000689

[69] GRIFFIN, J. and LEISEN, F. (2018). Modelling and computation using NCoRM mixtures for density regression. *Bayesian Anal.* **13** 897–916. MR3807871 https://doi.org/10.1214/17-BA1072

[70] GRIFFIN, J. E. (2016). An adaptive truncation method for inference in Bayesian nonparametric models. *Stat. Comput.* **26** 423–441. MR3439383 https://doi.org/10.1007/s11222-014-9519-4

[71] GRIFFIN, J. E. and STEEL, M. F. J. (2006). Order-based dependent Dirichlet processes. *J. Amer. Statist. Assoc.* **101** 179–194. MR2268037 https://doi.org/10.1198/016214505000000727

[72] GRIFFIN, J. E. and STEEL, M. F. J. (2010). Bayesian nonparametric modelling with the Dirichlet process regression smoother. *Statist. Sinica* **20** 1507–1527. MR2777334

[73] GRIFFIN, J. E. and STEEL, M. F. J. (2011). Stick-breaking autoregressive processes. *J. Econometrics* **162** 383–396. MR2795625 https://doi.org/10.1016/j.jeconom.2011.03.001

[74] GUHA, A., HO, N. and NGUYEN, X. (2021). On posterior contraction of parameters and interpretability in Bayesian mixture modeling. *Bernoulli* **27** 2159–2188. MR4303879 https://doi.org/10.3150/20-BEJ1275

[75] GUTIÉRREZ, I., GUTIÉRREZ, L. and ALVARES, D. (2023). A new flexible Bayesian hypothesis test for multivariate data. *Stat. Comput.* **33** Paper No. 50. MR4556307 https://doi.org/10.1007/s11222-023-10214-6

[76] GUTIÉRREZ, L., BARRIENTOS, A. F., GONZÁLEZ, J. and TAYLOR-RODRÍGUEZ, D. (2019). A Bayesian nonparametric multiple testing procedure for comparing several treatments against a control. *Bayesian Anal.* **14** 649–675. MR3959876 https://doi.org/10.1214/18-BA1122

[77] GUTIÉRREZ, L., MENA, R. H. and RUGGIERO, M. (2016). A time dependent Bayesian nonparametric model for air quality analysis. *Comput. Statist. Data Anal.* **95** 161–175. MR3425946 https://doi.org/10.1016/j.csda.2015.10.002

[78] GUTIÉRREZ, L. and QUINTANA, F. A. (2011). Multivariate Bayesian semiparametric models for authentication of food and beverages. *Ann. Appl. Stat.* **5** 2385–2402. MR2907119 https://doi.org/10.1214/11-AOAS492

[79] HANNAH, L. A., BLEI, D. M. and POWELL, W. B. (2011). Dirichlet process mixtures of generalized linear models. *J. Mach. Learn. Res.* **12** 1923–1953. MR2819022

[80] HARTIGAN, J. A. (1990). Partition models. *Comm. Statist. Theory Methods* **19** 2745–2756. MR1088047 https://doi.org/10.1080/03610929008830345

[81] HEINER, M. and KOTTAS, A. (2022). Bayesian nonparametric density autoregression with lag selection. *Bayesian Anal.* **17** 1245–1273. MR4506028 https://doi.org/10.1214/21-ba1296

[82] HO, N. and NGUYEN, X. (2016). On strong identifiability and convergence rates of parameter estimation in finite mixtures. *Electron. J. Stat.* **10** 271–307. MR3466183 https://doi.org/10.1214/16-EJS1105

[83] HORIGUCHI, A., CHAN, C. and MA, L. (2022). A tree perspective on stick-breaking models in covariate-dependent mixtures. ArXiv preprint. Available at arXiv:2208.02806.

[84] HUANG, Y. and MENG, S. (2020). A Bayesian nonparametric model and its application in insurance loss prediction. *Insurance Math*. *Econom*. **93** 84–94. MR4093699 https://doi.org/10.1016/j.insmatheco.2020.04.010

[85] HWANG, B. S. and PENNELL, M. L. (2014). Semiparametric Bayesian joint modeling of a binary and continuous outcome with applications in toxicological risk assessment. *Stat*. *Med*. **33** 1162–1175. MR3247787 https://doi.org/10.1002/sim.6007

[86] IGLESIAS, P. L., ORELLANA, Y. and QUINTANA, F. A. (2009). Nonparametric Bayesian modelling using skewed Dirichlet processes. *J*. *Statist*. *Plann*. *Inference* **139** 1203–1214. MR2479861 https://doi.org/10.1016/j.jspi.2008.07.009

[87] INÁCIO, V. and RODRÍGUEZ-ÁLVAREZ, M. X. (2022). The covariate-adjusted ROC curve: The concept and its importance, review of inferential methods, and a new Bayesian estimator. *Statist*. *Sci*. **37** 541–561. MR4497231 https://doi.org/10.1214/21-sts839

[88] INÁCIO DE CARVALHO, V., JARA, A., HANSON, T. E. and DE CARVALHO, M. (2013). Bayesian nonparametric ROC regression modeling. *Bayesian Anal*. **8** 623–645. MR3102228 https://doi.org/10.1214/13-BA825

[89] ISHWARAN, H. and JAMES, L. F. (2001). Gibbs sampling methods for stick-breaking priors. *J*. *Amer*. *Statist*. *Assoc*. **96** 161–173. MR1952729 https://doi.org/10.1198/016214501750332758

[90] JACOBS, R. A. and JORDAN, M. I. (1994). Hierarchical mixtures of experts and the EM algorithm. *Neural Comput*. **6** 181–214.

[91] JACOBS, R. A., JORDAN, M. I., NOWLAN, S. and HINTON, G. E. (1991). Adaptive mixtures of local experts. *Neural Comput*. **3** 1–12.

[92] JARA, A., LESAFFRE, E., DE IORIO, M. and QUINTANA, F. (2010). Bayesian semiparametric inference for multivariate doubly-interval-censored data. *Ann*. *Appl*. *Stat*. **4** 2126–2149. MR2829950 https://doi.org/10.1214/10-AOAS368

[93] KALLI, M. and GRIFFIN, J. E. (2018). Bayesian nonparametric vector autoregressive models. *J*. *Econometrics* **203** 267–282. MR3770826 https://doi.org/10.1016/j.jeconom.2017.11.009

[94] KALLI, M., GRIFFIN, J. E. and WALKER, S. G. (2011). Slice sampling mixture models. *Stat*. *Comput*. **21** 93–105. MR2746606 https://doi.org/10.1007/s11222-009-9150-y

[95] KARLIS, D. and XEKALAKI, E. (2005). Mixed Poisson distributions. *Int*. *Stat*. *Rev*. **73** 35–58.

[96] KOTTAS, A., DUAN, J. A. and GELFAND, A. E. (2008). Modeling disease incidence data with spatial and spatio-temporal Dirichlet process mixtures. *Biom*. *J*. **50** 29–42. MR2414636 https://doi.org/10.1002/bimj.200610375

[97] KOTTAS, A. and KRNJAJIĆ, M. (2009). Bayesian semiparametric modelling in quantile regression. *Scand*. *J*. *Stat*. **36** 297–319. MR2528986 https://doi.org/10.1111/j.1467-9469.2008.00626.x

[98] KOTTAS, A., MÜLLER, P. and QUINTANA, F. (2005). Nonparametric Bayesian modeling for multivariate ordinal data. *J*. *Comput*. *Graph*. *Statist*. **14** 610–625. MR2170204 https://doi.org/10.1198/106186005X63185

[99] KOTTAS, A., WANG, Z. and RODRÍGUEZ, A. (2012). Spatial modeling for risk assessment of extreme values from environmental time series: A Bayesian nonparametric approach. *Environmetrics* **23** 649–662. MR3019057 https://doi.org/10.1002/env.2177

[100] KRNJAJIĆ, M., KOTTAS, A. and DRAPER, D. (2008). Parametric and nonparametric Bayesian model specification: A case study involving models for count data. *Comput*. *Statist*. *Data Anal*. **52** 2110–2128. MR2418492 https://doi.org/10.1016/j.csda.2007.07.010

[101] KUO, L. and MALLICK, B. (1997). Bayesian semiparametric inference for the accelerated failure-time model. *Canad*. *J*. *Statist*. **25** 457–472.

[102] LEE, C. J., ZITO, A., SANG, H. and DUNSON, D. B. (2024). Logistic-beta processes for modeling dependent random probabilities with beta marginals. ArXiv preprint. Available at arXiv:2402.07048.

[103] LEE, S. and MCLACHLAN, G. J. (2014). Finite mixtures of multivariate skew t-distributions: Some recent and new results. *Stat*. *Comput*. **24** 181–202. MR3165547 https://doi.org/10.1007/s11222-012-9362-4

[104] LI, P., BANERJEE, S., HANSON, T. A. and MCBEAN, A. M. (2015). Bayesian models for detecting difference boundaries in areal data. *Statist*. *Sinica* **25** 385–402. MR3328821

[105] LI, X., GUINDANI, M., NG, C. S. and HOBBS, B. P. (2021). A Bayesian nonparametric model for textural pattern heterogeneity. *J*. *R*. *Stat*. *Soc*. *Ser*. *C*. *Appl*. *Stat*. **70** 459–480. MR4226677 https://doi.org/10.1111/rssc.12469

[106] LIJOI, A. and PRÜNSTER, I. (2010). Models beyond the Dirichlet process. In *Bayesian Nonparametrics* (N. L. Hjort, C. C. Holmes, P. Müller and S. G. Walker, eds.). *Camb*. *Ser*. *Stat*. *Probab*. *Math*. **28** 80–136. Cambridge Univ. Press, Cambridge. MR2730661

[107] LIU, J., WADE, S. and BOCHKINA, N. (2022). Shared differential clustering across single-cell RNA sequencing datasets with the hierarchical Dirichlet process.

[108] LO, A. Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann*. *Statist*. **12** 351–357. MR0733519 https://doi.org/10.1214/aos/1176346412

[109] MACEACHERN, S. N. (1999). Dependent nonparametric processes. In *ASA Proceedings of the Section on Bayesian Statistical Science* 50–55. Amer. Statist. Assoc., Alexandria.

[110] MACEACHERN, S. N. (2000). Dependent Dirichlet processes. Technical Report, Department of Statistics, Ohio State University.

[111] MACEACHERN, S. N. (2001). Decision theoretic aspects of dependent nonparametric processes. In *Bayesian Methods with Applications to Science*, *Policy*, *and Official Statistics* (E. George, ed.) 551–560. ISBA.

[112] MEEDS, E. and OSINDERO, S. (2005). An alternative infinite mixture of Gaussian process experts. *Adv*. *Neural Inf*. *Process*. *Syst*. **18**.

[113] MENA, R. H. and RUGGIERO, M. (2016). Dynamic density estimation with diffusive Dirichlet mixtures. *Bernoulli* **22** 901–926. MR3449803 https://doi.org/10.3150/14-BEJ681

[114] MIRA, A. and PETRONE, S. (1996). Bayesian hierarchical nonparametric inference for change-point problems. In *Bayesian Statistics*, 5 (*Alicante*, 1994) (J. M. Bernardo, J. O. Berger, A. P. Dawid and A. F. M. Smith, eds.). *Oxford Sci*. *Publ*. 693–703. Oxford Univ. Press, New York. MR1425440

[115] MULIERE, P. and PETRONE, S. (1993). A Bayesian predictive approach to sequential search for an optimal dose: Parametric and nonparametric models. *J*. *Ital*. *Stat*. *Soc*. **2** 349–364.

[116] MÜLLER, P., ERKANLI, A. and WEST, M. (1996). Bayesian curve fitting using multivariate normal mixtures. *Biometrika* **83** 67–79. MR1399156 https://doi.org/10.1093/biomet/83.1.67

[117] MÜLLER, P. and QUINTANA, F. (2010). Random partition models with regression on covariates. *J*. *Statist*. *Plann*. *Inference* **140** 2801–2808. MR2651966 https://doi.org/10.1016/j.jspi.2010.03.002

[118] MÜLLER, P., ROSNER, G. L., DE IORIO, M. and MACEACHERN, S. (2005). A nonparametric Bayesian model for inference in related longitudinal studies. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **54** 611–626. MR2137257 https://doi.org/10.1111/j.1467-9876.2005.05475.x

[119] NEAL, R. M. (1996). *Bayesian Learning for Neural Networks*. *Lecture Notes in Statistics*. Springer, Berlin.

[120] NEAL, R. M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *J. Comput. Graph. Statist.* **9** 249–265. MR1823804 https://doi.org/10.2307/1390653

[121] NIETO-BARAJAS, L. E., MÜLLER, P., JI, Y., LU, Y. and MILLS, G. B. (2012). A time-series DDP for functional proteomics profiles. *Biometrics* **68** 859–868. MR3055190 https://doi.org/10.1111/j.1541-0420.2011.01724.x

[122] NORETS, A. and PELENIS, J. (2012). Bayesian modeling of joint and conditional distributions. *J. Econometrics* **168** 332–346. MR2923772 https://doi.org/10.1016/j.jeconom.2012.02.001

[123] NORETS, A. and PELENIS, J. (2014). Posterior consistency in conditional density estimation by covariate dependent mixtures. *Econometric Theory* **30** 606–646. MR3205608 https://doi.org/10.1017/S026646661300042X

[124] NORETS, A. and PELENIS, J. (2022). Adaptive Bayesian estimation of discrete-continuous distributions under smoothness and sparsity. *Econometrica* **90** 1355–1377. MR4436055

[125] PAGE, G., BHATTACHARYA, A. and DUNSON, D. (2013). Classification via Bayesian nonparametric learning of affine subspaces. *J. Amer. Statist. Assoc.* **108** 187–201. MR3174612 https://doi.org/10.1080/01621459.2013.763566

[126] PAN, T., SHEN, W., DAVIS-STOBER, C. P. and HU, G. (2024). A Bayesian nonparametric approach for handling item and examinee heterogeneity in assessment data. *Br. J. Math. Stat. Psychol.* **77** 196–211.

[127] PAPAGEORGIOU, G. (2019). Bayesian density regression for discrete outcomes. *Aust. N. Z. J. Stat.* **61** 336–359. MR4034102 https://doi.org/10.1111/anzs.12273

[128] PAPASPILIOPOULOS, O. and ROBERTS, G. O. (2008). Retrospective Markov chain Monte Carlo methods for Dirichlet process hierarchical models. *Biometrika* **95** 169–186. MR2409721 https://doi.org/10.1093/biomet/asm086

[129] PARK, J.-H. and DUNSON, D. B. (2010). Bayesian generalized product partition model. *Statist. Sinica* **20** 1203–1226. MR2730180

[130] PATI, D., DUNSON, D. B. and TOKDAR, S. T. (2013). Posterior consistency in conditional distribution estimation. *J. Multivariate Anal.* **116** 456–472. MR3049916 https://doi.org/10.1016/j.jmva.2013.01.011

[131] PETRONE, S. and RAFTERY, A. E. (1997). A note on the Dirichlet process prior in Bayesian nonparametric inference with partial exchangeability. *Statist. Probab. Lett.* **36** 69–83. MR1491076 https://doi.org/10.1016/S0167-7152(97)00050-3

[132] PETRONE, S. and VERONESE, P. (2010). Feller operators and mixture priors in Bayesian nonparametrics. *Statist. Sinica* **20** 379–404. MR2640700

[133] PETRONE, S. and WASSERMAN, L. (2002). Consistency of Bernstein polynomial posteriors. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 79–100. MR1881846 https://doi.org/10.1111/1467-9868.00326

[134] POLSON, N. G., SCOTT, J. G. and WINDLE, J. (2013). Bayesian inference for logistic models using Pólya-Gamma latent variables. *J. Amer. Statist. Assoc.* **108** 1339–1349. MR3174712 https://doi.org/10.1080/01621459.2013.829001

[135] POPE, C. A., GOSLING, J. P., BARBER, S., JOHNSON, J. S., YAMAGUCHI, T., FEINGOLD, G. and BLACKWELL, P. G. (2021). Gaussian process modeling of heterogeneity and discontinuities using Voronoi tessellations. *Technometrics* **63** 53–63. MR4205691 https://doi.org/10.1080/00401706.2019.1692696

[136] POYNOR, V. and KOTTAS, A. (2017). Bayesian nonparametric mean residual life regression. Technical Report. UCSC-SOE-17-08, Jack Baskin School of Engineering, University of California, Santa Cruz, 2017.

[137] PRATOLA, M. T., CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2020). Heteroscedastic BART via multiplicative regression trees. *J. Comput. Graph. Statist.* **29** 405–417. MR4116052 https://doi.org/10.1080/10618600.2019.1677243

[138] QUINLAN, J. J., PAGE, G. L. and QUINTANA, F. A. (2018). Density regression using repulsive distributions. *J. Stat. Comput. Simul.* **88** 2931–2947. MR3832183 https://doi.org/10.1080/00949655.2018.1491578

[139] QUINTANA, F. A. (2006). A predictive view of Bayesian clustering. *J. Statist. Plann. Inference* **136** 2407–2429. MR2279815 https://doi.org/10.1016/j.jspi.2004.09.015

[140] QUINTANA, F. A. (2010). Linear regression with a dependent skewed Dirichlet process. *Chil. J. Stat.* **1** 35–49. MR2756118

[141] QUINTANA, F. A. and IGLESIAS, P. L. (2003). Bayesian clustering and product partition models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **65** 557–574. MR1983764 https://doi.org/10.1111/1467-9868.00402

[142] QUINTANA, F. A., JOHNSON, W. O., WAETJEN, L. E. and GOLD, E. B. (2016). Bayesian nonparametric longitudinal data analysis. *J. Amer. Statist. Assoc.* **111** 1168–1181. MR3561940 https://doi.org/10.1080/01621459.2015.1076725

[143] QUINTANA, F. A., MÜLLER, P., JARA, A. and MACEACHERN, S. N. (2022). The dependent Dirichlet process and related models. *Statist. Sci.* **37** 24–41. MR4371095 https://doi.org/10.1214/20-sts819

[144] RAO, V. and TEH, Y. (2009). Spatial normalized gamma processes. *Adv. Neural Inf. Process. Syst.* **22**.

[145] RASMUSSEN, C. E. and WILLIAMS, C. K. I. (2006). *Gaussian Processes for Machine Learning*. *Adaptive Computation and Machine Learning*. MIT Press, Cambridge, MA. MR2514435

[146] REICH, B. J., BONDELL, H. D. and WANG, H. J. (2010). Flexible Bayesian quantile regression for independent and clustered data. *Biostatistics* **11** 337–352.

[147] REICH, B. J. and FUENTES, M. (2007). A multivariate semiparametric Bayesian spatial modeling framework for hurricane surface wind fields. *Ann. Appl. Stat.* **1** 249–264. MR2393850 https://doi.org/10.1214/07-AOAS108

[148] REN, L., DU, L., CARIN, L. and DUNSON, D. B. (2011). Logistic stick-breaking process. *J. Mach. Learn. Res.* **12** 203–239. MR2773552

[149] RICHARDSON, R. and HARTMAN, B. (2018). Bayesian nonparametric regression models for modeling and predicting healthcare claims. *Insurance Math. Econom.* **83** 1–8. MR3886482 https://doi.org/10.1016/j.insmatheco.2018.06.002

[150] RIGON, T. and DURANTE, D. (2021). Tractable Bayesian density regression via logit stick-breaking priors. *J. Statist. Plann. Inference* **211** 131–142. MR4117446 https://doi.org/10.1016/j.jspi.2020.05.009

[151] RODRÍGUEZ, A. and DUNSON, D. B. (2011). Nonparametric Bayesian models through probit stick-breaking processes. *Bayesian Anal.* **6** 145–177. MR2781811 https://doi.org/10.1214/11-BA605

[152] RODRÍGUEZ, A., DUNSON, D. B. and GELFAND, A. E. (2009). Bayesian nonparametric functional data analysis through density estimation. *Biometrika* **96** 149–162. MR2482141 https://doi.org/10.1093/biomet/asn054

[153] RODRÍGUEZ, A. and MARTÍNEZ, J. C. (2014). Bayesian semiparametric estimation of covariate-dependent ROC curves. *Biostatistics* **15** 353–369.

[154] RODRIGUEZ, A. and TER HORST, E. (2008). Bayesian dynamic density estimation. *Bayesian Anal*. **3** 339–365. MR2407430 https://doi.org/10.1214/08-BA313

[155] RODRÍGUEZ-ÁLVAREZ, M. X., INÁCIO, V. and KLEIN, N. (2024). Density regression via Dirichlet process mixtures of normal structured additive regression models. ArXiv preprint. Available at arXiv:2401.03881.

[156] SCOTT, D. W. (2015). *Multivariate Density Estimation*: *Theory*, *Practice*, *and Visualization*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR3329609

[157] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statist*. *Sinica* **4** 639–650. MR1309433

[158] SHAHBABA, B. and NEAL, R. (2009). Nonlinear models using Dirichlet process mixtures. *J*. *Mach*. *Learn*. *Res*. **10** 1829–1850. MR2540778

[159] SHEN, W., TOKDAR, S. T. and GHOSAL, S. (2013). Adaptive Bayesian multivariate density estimation with Dirichlet mixtures. *Biometrika* **100** 623–640. MR3094441 https://doi.org/10.1093/biomet/ast015

[160] SHI, Y., LAUD, P. and NEUNER, J. (2021). A dependent Dirichlet process model for survival data with competing risks. *Lifetime Data Anal*. **27** 156–176. MR4202724 https://doi.org/10.1007/s10985-020-09506-0

[161] SPENCER, N. A. and MURRAY, J. S. (2020). A Bayesian hierarchical model for evaluating forensic footwear evidence. *Ann*. *Appl*. *Stat*. **14** 1449–1470. MR4152141 https://doi.org/10.1214/20-AOAS1334

[162] TADDY, M. A. and KOTTAS, A. (2009). Markov switching Dirichlet process mixture regression. *Bayesian Anal*. **4** 793–816. MR2570089 https://doi.org/10.1214/09-BA430

[163] TADDY, M. A. and KOTTAS, A. (2010). A Bayesian nonparametric approach to inference for quantile regression. *J*. *Bus*. *Econom*. *Statist*. **28** 357–369. MR2723605 https://doi.org/10.1198/jbes.2009.07331

[164] TADDY, M. A. and KOTTAS, A. (2012). Mixture modeling for marked Poisson processes. *Bayesian Anal*. **7** 335–361. MR2934954 https://doi.org/10.1214/12-BA711

[165] TOKDAR, S. T. (2006). Posterior consistency of Dirichlet location-scale mixture of normals in density estimation and regression. *Sankhyā* **68** 90–110. MR2301566

[166] TOKDAR, S. T. (2011). Adaptive convergence rates of a Dirichlet process mixture of multivariate normals.

[167] VIDAKOVIC, B. (1999). *Statistical Modeling by Wavelets*. *Wiley Series in Probability and Statistics*: *Applied Probability and Statistics*. Wiley, New York. MR1681904 https://doi.org/10.1002/9780470317020

[168] WADE, S., DUNSON, D. B., PETRONE, S. and TRIPPA, L. (2014). Improving prediction from Dirichlet process mixtures via enrichment. *J*. *Mach*. *Learn*. *Res*. **15** 1041–1071. MR3195338

[169] WADE, S. and GHAHRAMANI, Z. (2018). Bayesian cluster analysis: Point estimation and credible balls (with discussion). *Bayesian Anal*. **13** 559–626. MR3807860 https://doi.org/10.1214/17-BA1073

[170] WADE, S. and INÁCIO, V. (2025). Supplement to "Bayesian dependent mixture models: a predictive comparison and survey." https://doi.org/10.1214/24-STS966SUPP

[171] WADE, S., MONGELLUZZO, S. and PETRONE, S. (2011). An enriched conjugate prior for Bayesian nonparametric inference. *Bayesian Anal*. **6** 359–385. MR2843536 https://doi.org/10.1214/ba/1339616468

[172] WADE, S., PICCARRETA, R., CREMASCHI, A. and ANTONIANO-VILLALOBOS, I. (2022). Colombian women's life patterns: A multivariate density regression approach. *Bayesian Anal*. **17** 405–433. MR4483225 https://doi.org/10.1214/20-ba1256

[173] WADE, S., WALKER, S. G. and PETRONE, S. (2014). A predictive study of Dirichlet process mixture models for curve fitting. *Scand*. *J*. *Stat*. **41** 580–605. MR3249418 https://doi.org/10.1111/sjos.12047

[174] WALDMANN, E., KNEIB, T., YUE, Y. R., LANG, S. and FLEXEDER, C. (2013). Bayesian semiparametric additive quantile regression. *Stat*. *Model*. **13** 223–252. MR3179525 https://doi.org/10.1177/1471082X13480650

[175] WALKER, S. G., LIJOI, A. and PRÜNSTER, I. (2007). On rates of convergence for posterior distributions in infinite-dimensional models. *Ann*. *Statist*. **35** 738–746. MR2336866 https://doi.org/10.1214/009053606000001361

[176] WARREN, J. L., CAI, J., JOHNSON, N. P. and DEZIEL, N. C. (2022). A discrete kernel stick-breaking model for detecting spatial boundaries in hydraulic fracturing wastewater disposal well placement across Ohio. *J*. *R*. *Stat*. *Soc*. *Ser*. *C*. *Appl*. *Stat*. **71** 175–193. MR4376850 https://doi.org/10.1111/rssc.12527

[177] WEST, M., MÜLLER, P. and ESCOBAR, M. D. (1994). Hierarchical priors and mixture models, with application in regression and density estimation. In *Aspects of Uncertainty*. *Wiley Ser*. *Probab*. *Math*. *Statist*. *Probab*. *Math*. *Statist*. 363–386. Wiley, Chichester. MR1309702

[178] WOOD, S. N. (2017). *Generalized Additive Models*: *An Introduction with R*. *Texts in Statistical Science Series*. CRC Press, Boca Raton, FL. MR3726911

[179] WU, Q. and LUO, X. (2022). Nonparametric Bayesian two-level clustering for subject-level single-cell expression data. *Statist*. *Sinica* **32** 1835–1856. MR4478180

[180] WU, Y. and GHOSAL, S. (2008). Kullback Leibler property of kernel mixture priors in Bayesian density estimation. *Electron*. *J*. *Stat*. **2** 298–331. MR2399197 https://doi.org/10.1214/07-EJS130

[181] WU, Y. and GHOSAL, S. (2010). The L_1-consistency of Dirichlet mixtures in multivariate Bayesian density estimation. *J*. *Multivariate Anal*. **101** 2411–2419. MR2719871 https://doi.org/10.1016/j.jmva.2010.06.012

[182] XIAO, S., KOTTAS, A. and SANSÓ, B. (2015). Modeling for seasonal marked point processes: An analysis of evolving hurricane occurrences. *Ann*. *Appl*. *Stat*. **9** 353–382. MR3341119 https://doi.org/10.1214/14-AOAS796

[183] XU, D., DANIELS, M. J. and WINTERSTEIN, A. G. (2018). A Bayesian nonparametric approach to causal inference on quantiles. *Biometrics* **74** 986–996. MR3860719 https://doi.org/10.1111/biom.12863

[184] XU, L., JORDAN, M. and HINTON, G. E. (1994). An alternative model for mixtures of experts. *Adv*. *Neural Inf*. *Process*. *Syst*. **7**.

[185] XU, Y., MÜLLER, P., WAHED, A. S. and THALL, P. F. (2016). Bayesian nonparametric estimation for dynamic treatment regimes with sequential transition times. *J*. *Amer*. *Statist*. *Assoc*. **111** 921–950. MR3561917 https://doi.org/10.1080/01621459.2015.1086353

[186] XU, Y., SCHARFSTEIN, D., MÜLLER, P. and DANIELS, M. (2022). A Bayesian nonparametric approach for evaluating the causal effect of treatment in randomized trials with semi-competing risks. *Biostatistics* **23** 34–49. MR4366034 https://doi.org/10.1093/biostatistics/kxaa008

[187] XU, Y., THALL, P. F., HUA, W. and ANDERSSON, B. S. (2019). Bayesian non-parametric survival regression for optimizing precision dosing of intravenous busulfan in allogeneic stem cell

transplantation. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **68** 809–828. MR3937475 https://doi.org/10.1111/rssc.12331

[188] YUAN, C. and NEUBAUER, C. (2008). Variational mixture of Gaussian process experts. *Adv. Neural Inf. Process. Syst.* **21**.

[189] ZHANG, H., WADE, S. and BOCHKINA, N. (2024). Covariate-dependent hierarchical Dirichlet process. ArXiv preprint. Available at arXiv:2407.02676.

[190] ZHOU, H., HANSON, T. and KNAPP, R. (2015). Marginal Bayesian nonparametric model for time to disease arrival of threatened amphibian populations. *Biometrics* **71** 1101–1110. MR3436735 https://doi.org/10.1111/biom.12345

Statistical Science
2025, Vol. 40, No. 1, 109–127
https://doi.org/10.1214/24-STS954

Bayesian Predictive Synthesis with Outcome-Dependent Pools

Matthew C. Johnson and Mike West

Abstract. This paper reviews background and examples of Bayesian predictive synthesis (BPS), and develops details in a subset of BPS mixture models. BPS expands on standard Bayesian model uncertainty analysis for model mixing to provide a broader foundation for calibrating and combining predictive densities from multiple models or other sources. One main focus here is BPS as a framework for justifying and understanding generalized "linear opinion pools," where multiple predictive densities are combined with flexible mixing weights that depend on the forecast variable itself– that is, the setting of outcome-dependent model mixing. BPS also defines approaches to incorporating and exploiting dependencies across models defining forecasts, and to formally addressing the problem of model set incompleteness within the subjective Bayesian framework. In addition to an overview of general mixture-based BPS, new methodological developments for dynamic BPS—involving calibration and pooling of sets of predictive distributions in a univariate time series setting—are presented. These developments are exemplified in summaries of an analysis in a univariate financial time series study.

Key words and phrases: Bayesian predictive synthesis, density forecast combination, forecaster dependence, forecasting, forecast calibration, generalized opinion pools, model combination, model set incompleteness, time series prediction.

REFERENCES

AASTVEIT, K. A., CROSS, J. L. and VAN DIJK, H. K. (2023). Quantifying time-varying forecast uncertainty and risk for the real price of oil. *J. Bus. Econom. Statist.* **41** 523–537. MR4568040 https://doi.org/10.1080/07350015.2022.2039159

AASTVEIT, K. A., GERDRUP, K. R., JORE, A. S. and THORSRUD, L. A. (2014). Nowcasting GDP in real time: A density combination approach. *J. Bus. Econom. Statist.* **32** 48–68. MR3173707 https://doi.org/10.1080/07350015.2013.844155

AASTVEIT, K. A., MITCHELL, J., RAVAZZOLO, F. and VAN DIJK, H. K. (2019). The evolution of forecast density combinations in economics. Oxford Research Encyclopedia of Economics and Finance. https://doi.org/10.1093/acrefore/9780190625979.013.381

AMISANO, G. and GIACOMINI, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *J. Bus. Econom. Statist.* **25** 177–190. MR2367773 https://doi.org/10.1198/073500106000000332

BASSETTI, F., CASARIN, R. and RAVAZZOLO, F. (2018). Bayesian nonparametric calibration and combination of predictive distributions. *J. Amer. Statist. Assoc.* **113** 675–685. MR3832218 https://doi.org/10.1080/01621459.2016.1273117

BERNARDO, J.-M. and SMITH, A. F. M. (1994). *Bayesian Theory*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

BILLIO, M., CASARIN, R., RAVAZZOLO, F. and VAN DIJK, H. K. (2012). Combination schemes for turning point predictions. *Q. Rev. Econ. Finance* **52** 402–412. https://doi.org/10.1016/j.qref.2012.08.002

BILLIO, M., CASARIN, R., RAVAZZOLO, F. and VAN DIJK, H. K. (2013). Time-varying combinations of predictive densities using nonlinear filtering. *J. Econometrics* **177** 213–232. MR3118557 https://doi.org/10.1016/j.jeconom.2013.04.009

BLUM, A. (1998). On-Line Algorithms in Machine Learning. In *Online Algorithms*: *The State of the Art* 306–325. Springer, Berlin. https://doi.org/10.1007/BFb0029575

CESA-BIANCHI, N., FREUND, Y., HAUSSLER, D., HELMBOLD, D. P., SCHAPIRE, R. E. and WARMUTH, M. K. (1997). How to use expert advice. *J. ACM* **44** 427–485. MR1470152 https://doi.org/10.1145/258128.258179

CESA-BIANCHI, N. and LUGOSI, G. (2006). *Prediction*, *Learning*, *and Games*. Cambridge Univ. Press, Cambridge. MR2409394 https://doi.org/10.1017/CBO9780511546921

CLEMEN, R. T. (1989). Combining forecasts: A review and annotated bibliography. *Int. J. Forecast.* **5** 559–583. https://doi.org/10.1016/0169-2070(89)90012-5

Matthew C. Johnson is Research Scientist, Amazon, Seattle, Washington 98109-5210, USA (e-mail: jnmmatt@amazon.com). Mike West is The Arts & Sciences Distinguished Professor Emeritus of Statistics & Decision Sciences, Duke University, Durham, North Carolina 27708-0251, USA (e-mail: Mike.West@duke.edu).

CLEMEN, R. T. and WINKLER, R. L. (1999). Combining probability distributions from experts in risk analysis. *Risk Anal*. **19** 187–203. https://doi.org/10.1111/j.1539-6924.1999.tb00399.x

CLEMEN, R. T. and WINKLER, R. . L. (2007). Aggregating probability distributions. In *Advances in Decision Analysis*: *From Foundations to Applications* (R. F. Miles, W. Edwards and D. von Winterfeldt, eds.) 154–176 9. Cambridge Univ. Press, Cambridge. https://doi.org/10.1017/CBO9780511611308.010

CLYDE, M. and GEORGE, E. I. (2004). Model uncertainty. *Statist*. *Sci*. **19** 81–94. MR2082148 https://doi.org/10.1214/088342304000000035

CLYDE, M. and IVERSEN, E. S. (2013). Bayesian model averaging in the M-open framework. In *Bayesian Theory and Applications* (P. Damien, P. Dellaportes, N. G. Polson and D. A. Stephens, eds.) 483–498. Oxford Univ. Press, Oxford. MR3221178 https://doi.org/10.1093/acprof:oso/9780199695607.003.0024

DE HEIDE, R., KIRICHENKO, A., MEHTA, N. and GRÜNWALD, P. (2019). Safe-Bayesian generalized linear regression. Available at arXiv:1910.09227.

DIACONIS, P. and ZABELL, S. L. (1982). Updating subjective probability. *J. Amer. Statist*. *Assoc*. **77** 822–830. MR0686405

DIEBOLD, F. X. and SHIN, M. (2019). Machine learning for regularized survey forecast combination: Partially-egalitarian LASSO and its derivatives. *Int*. *J*. *Forecast*. **35** 1679–1691. https://doi.org/10.1016/j.ijforecast.2018.09.006

DIEBOLD, F. X., SHIN, M. and ZHANG, B. (2020). On the aggregation of probability assessments: Regularized mixtures of predictive densities for Eurozone inflation and real interest rates. Available at arXiv:2012.11649.

GENEST, C. and SCHERVISH, M. J. (1985). Modeling expert judgments for Bayesian updating. *Ann*. *Statist*. **13** 1198–1212. MR0803766 https://doi.org/10.1214/aos/1176349664

GEWEKE, J. and AMISANO, G. (2011). Optimal prediction pools. *J. Econometrics* **164** 130–141. MR2821798 https://doi.org/10.1016/j.jeconom.2011.02.017

GEWEKE, J. and AMISANO, G. G. (2012). Prediction with misspecified models. *Amer. Econ*. *Rev*. **102** 482–486. https://doi.org/10.1257/aer.102.3.482

GIANNONE, D., LENZA, M. and PRIMICERI, G. E. (2021). Economic predictions with big data: The illusion of sparsity. *Econometrica* **89** 2409–2437. https://doi.org/10.3982/ECTA17842

GRUBER, L. and WEST, M. (2016). GPU-accelerated Bayesian learning and forecasting in simultaneous graphical dynamic linear models. *Bayesian Anal*. **11** 125–149. MR3447094 https://doi.org/10.1214/15-BA946

GRUBER, L. F. and WEST, M. (2017). Bayesian online variable selection and scalable multivariate volatility forecasting in simultaneous graphical dynamic linear models. *Econom*. *Stat*. **3** 3–22. MR3666239 https://doi.org/10.1016/j.ecosta.2017.03.003

HALL, S. G. and MITCHELL, J. (2007). Combining density forecasts. *Int*. *J*. *Forecast*. **23** 1–13. https://doi.org/10.1016/j.ijforecast.2006.08.001

HOOGERHEIDE, L., KLEIJN, R., RAVAZZOLO, F., VAN DIJK, H. K. and VERBEEK, M. (2010). Forecast accuracy and economic gains from Bayesian model averaging using time-varying weights. *J. Forecast*. **29** 251–269. MR2752013 https://doi.org/10.1002/for.1145

IRIE, K. and WEST, M. (2019). Bayesian emulation for multi-step optimization in decision problems. *Bayesian Anal*. **14** 137–160. MR3910041 https://doi.org/10.1214/18-BA1105

JEFFREY, R. C. (1983). *The Logic of Decision*, 2nd ed. Univ. Chicago Press, Chicago, IL. MR0745132

KAPETANIOS, G., MITCHELL, J., PRICE, S. and FAWCETT, N. (2015). Generalised density forecast combinations. *J. Econometrics* **188** 150–165. MR3371665 https://doi.org/10.1016/j.jeconom.2015.02.047

KASCHA, C. and RAVAZZOLO, F. (2010). Combining inflation density forecasts. *J. Forecast*. **29** 231–250. MR2752012 https://doi.org/10.1002/for.1147

KOOP, G., MCINTYRE, S. and MITCHELL, J. (2020). UK regional nowcasting using a mixed frequency vector auto-regressive model with entropic tilting. *J. Roy. Statist. Soc. Ser. A* **183** 91–119. MR4049656 https://doi.org/10.1111/rssa.12491

KOOP, G., MCINTYRE, S., MITCHELL, J. and POON, A. (2020). Regional output growth in the United Kingdom: More timely and higher frequency estimates from 1970. *J. Appl. Econometrics* **35** 176–197. MR4098531 https://doi.org/10.1002/jae.2748

LAVINE, I., LINDON, M. and WEST, M. (2021). Adaptive variable selection for sequential prediction in multivariate dynamic models. *Bayesian Anal*. **16** 1059–1083. MR4381127 https://doi.org/10.1214/20-BA1245

LINDLEY, D. V., TVERSKY, A. and BROWN, R. V. (1979). On the reconciliation of probability assessments. *J. Roy. Statist. Soc. Ser. A* **142** 146–180. With discussion. MR0547236 https://doi.org/10.2307/2345078

LITTLESTONE, N. and WARMUTH, M. K. (1994). The weighted majority algorithm. *Inform. and Comput*. **108** 212–261. MR1265851 https://doi.org/10.1006/inco.1994.1009

LOAIZA-MAYA, R., MARTIN, G. M. and FRAZIER, D. T. (2021). Focused Bayesian prediction. *J. Appl. Econometrics* **36** 517–543. MR4309597 https://doi.org/10.1002/jae.2810

MASUDA, R. and IRIE, K. (2024). On the proofs of the predictive synthesis formula. Technical Report, Department of Economics, University of Tokyo. Available at arXiv:2409.09660.

MCALINN, K. (2021). Mixed-frequency Bayesian predictive synthesis for economic nowcasting. *J. R. Stat. Soc. Ser. C. Appl. Stat*. **70** 1143–1163. MR4347707 https://doi.org/10.1111/rssc.12500

MCALINN, K., AASTVEIT, K. A., NAKAJIMA, J. and WEST, M. (2020). Multivariate Bayesian predictive synthesis in macroeconomic forecasting. *J. Amer. Statist. Assoc*. **115** 1092–1110. MR4143452 https://doi.org/10.1080/01621459.2019.1660171

MCALINN, K., AASTVEIT, K. A. and WEST, M. (2018). Bayesian predictive synthesis. Discussion of: Using stacking to average Bayesian predictive distributions, by Y. Yao *et al*. *Bayesian Anal*. **13** 971–973. https://doi.org/10.1214/17-BA1091

MCALINN, K. and WEST, M. (2019). Dynamic Bayesian predictive synthesis in time series forecasting. *J. Econometrics* **210** 155–169. MR3944768 https://doi.org/10.1016/j.jeconom.2018.11.010

PETTENUZZO, D. and RAVAZZOLO, F. (2016). Optimal portfolio choice under decision-based model combinations. *J. Appl. Econometrics* **31** 1312–1332. MR3580902 https://doi.org/10.1002/jae.2502

PRADO, R., FERREIRA, M. A. R. and WEST, M. (2021). *Time Series*: *Modeling*, *Computation & Inference*, 2nd ed. CRC Press, Boca Raton. https://doi.org/10.1201/9781351259422

RUFO, M. J., MARTÍN, J. and PÉREZ, C. J. (2012). Log-linear pool to combine prior distributions: A suggestion for a calibration-based approach. *Bayesian Anal*. **7** 411–438. MR2934957 https://doi.org/10.1214/12-BA714

TALLMAN, E. and WEST, M. (2022). On entropic tilting and predictive conditioning. Technical Report, Department of Statistical Science, Duke University. Available at arXiv:2207.10013.

TALLMAN, E. and WEST, M. (2024a). Bayesian predictive decision synthesis. *J. R. Stat. Soc. Ser. B. Stat. Methodol*. **86** 340–363. MR4754087 https://doi.org/10.1093/jrsssb/qkad109

TALLMAN, E. and WEST, M. (2024b). Predictive decision synthesis for portfolios: Betting on better models. In *Recent Developments in Bayesian Econometrics and Their Applications*. Springer. Available at arXiv:2405.01598.

TIMMERMANN, A. (2004). Forecast combinations. In *Handbook of Economic Forecasting* (C. W. J. Granger, G. Elliott and A. Timmermann, eds.) **1** 135–196 4. North Holland, Amsterdam. https://doi.org/10.1016/S1574-0706(05)01004-9

WEST, M. (1984). Bayesian aggregation. *J. Roy. Statist. Soc. Ser. A* **147** 600–607. MR0782071 https://doi.org/10.2307/2981847

WEST, M. (1988). Modelling expert opinion. In *Bayesian Statistics*, 3 (*Valencia*, 1987) (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.). *Oxford Sci. Publ.* 493–508. Oxford Univ. Press, New York. MR1008062

WEST, M. (1992). Modelling agent forecast distributions. *J. Roy. Statist. Soc. Ser. B* **54** 553–567. MR1160482

WEST, M. (2020). Bayesian forecasting of multivariate time series: Scalability, structure uncertainty and decisions (with discussion). *Ann. Inst. Statist. Math.* **72** 1–44. https://doi.org/10.1007/s10463-019-00741-3

WEST, M. (2024). Perspectives on constrained forecasting. *Bayesian Anal.* **19** 1013–1039. MR4801590 https://doi.org/10.1214/23-ba1379

WEST, M. and CROSSE, J. (1992). Modelling probabilistic agent opinion. *J. Roy. Statist. Soc. Ser. B* **54** 285–299. MR1157726

WEST, M. and HARRISON, J. (1997). *Bayesian Forecasting and Dynamic Models*, 2nd ed. *Springer Series in Statistics*. Springer, New York. MR1482232

ZHAO, Z. Y., XIE, M. and WEST, M. (2016). Dynamic dependence networks: Financial time series forecasting and portfolio decisions. *Appl. Stoch. Models Bus. Ind.* **32** 311–332. MR3518020 https://doi.org/10.1002/asmb.2161

Statistical Science
2025, Vol. 40, No. 1, 128–147
https://doi.org/10.1214/24-STS949

Advances in Projection Predictive Inference

Yann McLatchie, Sölvi Rögnvaldsson, Frank Weber and Aki Vehtari

Abstract. The concepts of Bayesian prediction, model comparison, and model selection have developed significantly over the last decade. As a result, the Bayesian community has witnessed a rapid growth in theoretical and applied contributions to building and selecting predictive models. Projection predictive inference in particular has shown promise to this end, finding application across a broad range of fields. It is less prone to over-fitting than naïve selection based purely on cross-validation or information criteria performance metrics, and has been known to out-perform other methods in terms of predictive performance. We survey the core concept and contemporary contributions to projection predictive inference, and present a safe, efficient, and modular workflow for prediction-oriented model selection therein. We also provide an interpretation of the projected posteriors achieved by projection predictive inference in terms of their limitations in causal settings.

Key words and phrases: Bayesian model selection, cross-validation, projection predictive inference.

REFERENCES

AFRABANDPEY, H., PELTOLA, T., PIIRONEN, J., VEHTARI, A. and KASKI, S. (2020). A decision-theoretic approach for model interpretability in Bayesian framework. *Mach. Learn.* **109** 1855–1876. MR4156281 https://doi.org/10.1007/s10994-020-05901-8

AGUILAR, J. E. and BÜRKNER, P.-C. (2023). Intuitive joint priors for Bayesian linear multilevel models: The R2D2M2 prior. *Electron. J. Stat.* **17** 1711–1767. MR4609453 https://doi.org/10.1214/23-ejs2136

AMIRKHIZ, R. G., DIXON, M. D., PALMER, J. S. and SWANSON, D. L. (2021). Investigating niches and distribution of a rare species in a hierarchical framework: Virginia's warbler (leiothlypis virginiae) at its northeastern range limit. *Landsc. Ecol.* **36** 1039–1054.

BARTONICEK, A., WICKHAM, S. R., PAT, N. and CONNER, T. S. (2021). The value of Bayesian predictive projection for variable selection: An example of selecting lifestyle predictors of young adult well-being. *BMC Public Health* **21** 1–13.

BATES, D., MÄCHLER, M., BOLKER, B. and WALKER, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**.

BERNARDO, J. M. and RUEDA, R. (2002). Bayesian hypothesis testing: A reference approach. *Int. Stat. Rev.* **70** 351–372.

BOHN, M., ECKERT, J., HANUS, D., LUGAUER, B., HOLTMANN, J. and HAUN, D. B. M. (2023). Great ape cognition is structured by stable cognitive abilities and predicted by developmental conditions. *Nat. Ecol. Evol.*.

BRATULIC, S., LIMETA, A., DABESTANI, S., BIRGISSON, H., ENBLAD, G., STÅLBERG, K., HESSELAGER, G., HÄGGMAN, M., HÖGLUND, M. et al. (2022). Noninvasive detection of any-stage cancer using free glycosaminoglycans. *Proc. Natl. Acad. Sci. USA* **119** e2115328119.

CARBONETTO, P. and STEPHENS, M. (2012). Scalable variational inference for Bayesian variable selection in regression, and its accuracy in genetic association studies. *Bayesian Anal.* **7** 73–107. MR2896713 https://doi.org/10.1214/12-BA703

CATALINA, A., BÜRKNER, P. and VEHTARI, A. (2021). Latent space projection predictive inference.

CATALINA, A., BÜRKNER, P.-C. and VEHTARI, A. (2020). Projection predictive inference for generalized linear and additive multilevel models.

CATALINA, A., BÜRKNER, P.-C. and VEHTARI, A. (2022). Projection predictive inference for generalized linear and additive multilevel models. In *Proceedings of the* 25th *International Conference on Artificial Intelligence and Statistics* (G. Camps-Valls, F. J. R. Ruiz and I. Valera, eds.). *Proceedings of Machine Learning Research* **151** 4446–4461. PMLR.

CHENG, L., RAMCHANDRAN, S., VATANEN, T., LIETZÉN, N., LAHESMAA, R., VEHTARI, A. and LÄHDESMÄKI, H. (2019). An additive Gaussian process regression model for interpretable nonparametric analysis of longitudinal data. *Nat. Commun.* **10** 1798. https://doi.org/10.1038/s41467-019-09785-8

CHIPMAN, H. A., GEORGE, E. I. and MCCULLOCH, R. E. (2010). BART: Bayesian additive regression trees. *Ann. Appl. Stat.* **4** 266–298. MR2758172 https://doi.org/10.1214/09-AOAS285

CINELLI, C., FORNEY, A. and PEARL, J. (2024). A crash course in good and bad controls. *Sociol. Methods Res.* **53** 1071–1104. MR4783049 https://doi.org/10.1177/00491241221099552

Yann McLatchie is Research Assistant, Department of Computer Science, Aalto University, Espoo, Finland (e-mail: yann.mclatchie.23@ucl.ac.uk). Sölvi Rögnvaldsson is Research Assistant, Department of Computer Science, Aalto University, Espoo, Finland. Frank Weber is Research Assistant, Institute for Biostatistics and Informatics in Medicine and Ageing Research, Rostock University Medical Center, Rostock, Germany. Aki Vehtari is Professor, Department of Computer Science, Aalto University, Espoo, Finland.

CLYDE, M. (2022). BAS: Bayesian variable selection and model averaging using Bayesian adaptive sampling. R package, version 1.6.4.

CORTEZ, P. and SILVA, A. M. G. (2008). Using data mining to predict secondary school student performance. In *Proceedings of 5th Annual Future Business Technology Conference*.

CSÁRDI, G. (2019). Cranlogs: Download logs from the 'RStudio' 'CRAN' mirror. R package, version 2.1.1.

DEAN, R. T. and TAYLOR, J. R. (2023). Continuous affect responses to a large diverse set of unfamiliar music: Bayesian time-series and cluster analyses. *Psychomusic*. *Music Mind Brain* **33** 7–25. https://doi.org/10.1037/pmu0000295

DIGBY, A., EASON, D., CATALINA, A., LIERZ, M., GALLA, S., URBAN, L., LE LEC, M. F., GUHLIN, J., STEEVES, T. E. et al. (2023). Hidden impacts of conservation management on fertility of the critically endangered kākāpō. *PeerJ* **11** e14675.

DUPUIS, J. and ROBERT, C. (2003). Variable selection in qualitative models via an entropic explanatory power. *J*. *Statist*. *Plann*. *Inference* **111** 77–94.

EFRON, B. (2008). Microarrays, empirical Bayes and the two-groups model. *Statist*. *Sci*. **23** 1–22. MR2431866 https://doi.org/10.1214/07-STS236

EFRON, B. (2010). *Large-Scale Inference*: *Empirical Bayes Methods for Estimation*, *Testing*, *and Prediction*. *Institute of Mathematical Statistics* (*IMS*) *Monographs* **1**. Cambridge Univ. Press, Cambridge. MR2724758 https://doi.org/10.1017/CBO9780511761362

EILERS, P. H. C. (2003). A perfect smoother. *Anal*. *Chem*. **75** 3631–3636. https://doi.org/10.1021/ac034173t

EILERS, P. H. C. (2005). Unimodal smoothing. *J*. *Chemom*. **19** 317–328.

EILERS, P. H. C. and MARX, B. D. (1996). Flexible smoothing with B-splines and penalties. *Statist*. *Sci*. **11** 89–121. MR1435485 https://doi.org/10.1214/ss/1038425655

FARAGGI, D., LEBLANC, M. and CROWLEY, J. (2001). Understanding neural networks using regression trees: An application to multiple myeloma survival data. *Stat*. *Med*. **20** 2965–2976. https://doi.org/10.1002/sim.912

GABRY, J. and MAHR, T. (2022). Bayesplot: Plotting for Bayesian models. R package version 1.9.0.

GABRY, J., SIMPSON, D., VEHTARI, A., BETANCOURT, M. and GELMAN, A. (2019). Visualization in Bayesian workflow. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *A* **182** 389–402. MR3902665 https://doi.org/10.1111/rssa.12378

GALATENKO, V. V., SHKURNIKOV, M. Y., SAMATOV, T. R., GALATENKO, A. V., MITYAKINA, I. A., KAPRIN, A. D., SCHUMACHER, U. and TONEVITSKY, A. G. (2015). Highly informative marker sets consisting of genes with low individual degree of differential expression. *Sci*. *Rep*. **5** 14967.

GARCIA-DONATO, G. and FORTE, A. (2018). Bayesian testing, variable selection and model averaging in linear models using R with BayesVarSel. *R J*. **10** 155–174.

GELMAN, A. and RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist*. *Sci*. **7**.

GELMAN, A., VEHTARI, A., SIMPSON, D., MARGOSSIAN, C. C., CARPENTER, B., YAO, Y., KENNEDY, L., GABRY, J., BÜRKNER, P.-C. et al. (2020). Bayesian workflow.

GEORGE, E. I. and MCCULLOCH, R. E. (1993). Variable selection via Gibbs sampling. *J*. *Amer*. *Statist*. *Assoc*. **88** 881–889.

GOTO, S., MORI, H., UCHIYAMA, K., ISHIZUKA, W., TANEDA, H., KONO, M., KAJIYA-KANEGAE, H. and IWATA, H. (2021). Genetic dissection of growth and eco-physiological traits associated with altitudinal adaptation in Sakhalin Fir (abies sachalinensis) based on QTL mapping. *Genes* **12** 1110.

GOUTIS, C. and ROBERT, C. P. (1998). Model choice in generalised linear models: A Bayesian approach via Kullback-Leibler projections. *Biometrika* **85** 29–37. MR1627250 https://doi.org/10.1093/biomet/85.1.29

GREENOP, A., COOK, S. M., WILBY, A., PYWELL, R. F. and WOODCOCK, B. A. (2020). Invertebrate community structure predicts natural pest control resilience to insecticide exposure. *J*. *Appl*. *Ecol*. **57** 2441–2453.

HAHN, P. R. and CARVALHO, C. M. (2015). Decoupling shrinkage and selection in Bayesian linear models: A posterior summary perspective. *J*. *Amer*. *Statist*. *Assoc*. **110** 435–448. MR3338514 https://doi.org/10.1080/01621459.2014.993077

HARRELL, F. E. (2001). *Regression Modeling Strategies*: *With Applications to Linear Models*, *Logistic Regression*, *and Survival Analysis*. *Springer Series in Statistics*. Springer, New York, NY.

JOHNSTONE, I. M. and SILVERMAN, B. W. (2004). Needles and straw in haystacks: Empirical Bayes estimates of possibly sparse sequences. *Ann*. *Statist*. **32** 1594–1649. MR2089135 https://doi.org/10.1214/009053604000000030

JONES, B., CARVALHO, C., DOBRA, A., HANS, C., CARTER, C. and WEST, M. (2005). Experiments in stochastic computation for high-dimensional graphical models. *Statist*. *Sci*. **20** 388–400. MR2210226 https://doi.org/10.1214/088342305000000304

KALLIOINEN, N., PAANANEN, T., BÜRKNER, P.-C. and VEHTARI, A. (2023). Priorsense: An R package for prior diagnostics and sensitivity.

KALLIOINEN, N., PAANANEN, T., BÜRKNER, P.-C. and VEHTARI, A. (2024). Detecting and diagnosing prior and likelihood sensitivity with power-scaling. *Stat*. *Comput*. **34** 57. MR4683177 https://doi.org/10.1007/s11222-023-10366-5

KARSDORP, F. and FONTEYN, L. (2019). Cultural entrenchment of folktales is encoded in language. *Palgrave Commun*. **5** 25.

KUCK, K., NAIK, B. I., DOMINO, K. B., POSNER, K. L., SAAGER, L., STUART, A. R., JOHNSON, K. B., ALPERT, S. B., DURIEUX, M. E. et al. (2023). Prolonged opioid use and pain outcome and associated factors after surgery under general anesthesia: A prospective cohort association multicenter study. *Anesthesiology* **138** 462–476.

KULLBACK, S. and LEIBLER, R. A. (1951). On information and sufficiency. *Ann*. *Math*. *Stat*. **22** 79–86. MR0039968 https://doi.org/10.1214/aoms/1177729694

KUMAR, R., CARROLL, C., HARTIKAINEN, A. and MARTIN, O. (2019). ArviZ a unified library for exploratory analysis of Bayesian models in Python. *J*. *Open Sour*. *Softw*. **4** 1143.

KUŚMIERCZYK, T., SAKAYA, J. and KLAMI, A. (2019). Variational Bayesian decision-making for continuous utilities. In *Proceedings of the* 33*rd International Conference on Neural Information Processing Systems* Curran Associates Inc., Red Hook, NY, USA.

LACOSTE–JULIEN, S., HUSZÁR, F. and GHAHRAMANI, Z. (2011). Approximate inference for the loss-calibrated Bayesian. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics* (G. Gordon, D. Dunson and M. Dudík, eds.). *Proceedings of Machine Learning Research* **15** 416–424. PMLR, Fort Lauderdale, FL, USA.

LEAMER, E. E. (1979). Comment on: "Information criteria for discriminating among alternative regression models" by T. Sawa. *Econometrica* **47** 507–510. MR0525791 https://doi.org/10.2307/1914197

LEUNG SOO, C., BHATNAGAR, S., BARTLETT, S. J., ESMAIL, A., DHEDA, K. and PAI, N. P. (2023). Development and evaluation of a digital HIV risk assessment tool incorporated within an app-based self-testing program. *J*. *Acquir*. *Immune Defic*. *Syndr*. **93** 387–394. https://doi.org/10.1097/QAI.0000000000003210

LINDLEY, D. V. (1968). The choice of variables in multiple regression. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **30** 31–66. MR0231492

LIQUET, B. and SUTTON, M. (2017). MBSGS: Multivariate Bayesian sparse group selection with spike and slab. R package, version 1.1.0.

MAGNUSSON, M., ANDERSEN, M. R., JONASSON, J. and VEHTARI, A. (2020). Leave-one-out cross-validation for Bayesian model comparison in large data. In *Proceedings of the 23rd International Conference on Artificial Intelligence and Statistics (AISTATS)* **108** 341–351. PMLR.

MCLATCHIE, Y., MARTIN, O. and CAPRETTO, T. (2022). Kulprit: Kullback–Leibler projections for Bayesian model selection in Python.

MCLATCHIE, Y., MATAMOROS, A. A., KOHNS, D. and VEHTARI, A. (2022). Bayesian order identification of ARMA models with projection predictive inference.

MCLATCHIE, Y., RÖGNVALDSSON, S., WEBER, F. and VEHTARI, A. (2025). Supplement to "Advances in projection predictive inference." https://doi.org/10.1214/24-STS949SUPP

MERCER, J. J., LIEFERT, D. T. and WILLIAMS, D. G. (2020). Atmospheric vapour and precipitation are not in isotopic equilibrium in a continental mountain environment. *Hydrol. Process.* **34** 3078–3101.

MIKKOLA, P., MARTIN, O. A., CHANDRAMOULI, S., HARTMANN, M., ABRIL PLA, O., THOMAS, O., PESONEN, H., CORANDER, J., VEHTARI, A. et al. (2023). Prior knowledge elicitation: The past, present, and future. *Bayesian Anal.*

MILNE, A. J., DEAN, R. and BULGER, D. (2021). The effects of rhythmic structure on tapping accuracy preprint PsyArXiv.

MITCHELL, T. J. and BEAUCHAMP, J. J. (1988). Bayesian variable selection in linear regression. *J. Amer. Statist. Assoc.* **83** 1023–1036. MR0997578

MODRÁK, M., MOON, A. H., KIM, S., BÜRKNER, P., HUURRE, N., FALTEJSKOVÁ, K., GELMAN, A. and VEHTARI, A. (2023). Simulation-based calibration checking for Bayesian computation: The choice of test quantities shapes sensitivity. *Bayesian Anal.* 1–28.

MORAIS, M. J. and PILLOW, J. W. (2022). Loss-calibrated expectation propagation for approximate Bayesian decision-making.

MORI, H., UENO, S., UJINO-IHARA, T., FUJIWARA, T., YAMASHITA, K., KANETANI, S., ENDO, R., MATSUMOTO, A., UCHIYAMA, K. et al. (2022). Genotype-by-environment interaction and genetic dissection of heartwood color in cryptomeria japonica based on multiple common gardens and quantitative trait loci mapping. *PLoS ONE* **17** e0270522. https://doi.org/10.1371/journal.pone.0270522

NG, A. M. B., PONTIUS, M. L., DE RUITER, S. L. and PROPPE, D. S. (2020). Noise, avian abundance, and productivity at banding stations across the Continental United States. *Avian Conserv. Ecol.* **15** art4.

NIKOOIENEJAD, A. and JOHNSON, V. E. (2020). BVSNLP: Bayesian variable selection in high dimensional settings using nonlocal priors. R package, version 1.1.9.

NILSSON, N. J. (1998). *Artificial Intelligence: A New Synthesis*. Elsevier, Amsterdam.

NORDSTRÖM, H. and LAUKKA, P. (2019). The time course of emotion recognition in speech and music. *J. Acoust. Soc. Am.* **145** 3058–3074.

NOTT, D. J. and LENG, C. (2010). Bayesian projection approaches to variable selection in generalized linear models. *Comput. Statist. Data Anal.* **54** 3227–3241. MR2727748 https://doi.org/10.1016/j.csda.2010.01.036

NTZOUFRAS, I., FORSTER, J. J. and DELLAPORTAS, P. (2000). Stochastic search variable selection for log-linear models. *J. Stat. Comput. Simul.* **68** 23–37. MR1820600 https://doi.org/10.1080/00949650008812054

PAASINIEMI, M. (2018). Methods and tools for interpretable Bayesian variable selection Master's thesis Aalto Univ.

PAUL, D., BAIR, E., HASTIE, T. and TIBSHIRANI, R. (2008). "Preconditioning" for feature selection and regression in high-dimensional problems. *Ann. Statist.* **36** 1595–1618. MR2435449 https://doi.org/10.1214/009053607000000578

PAVONE, F., PIIRONEN, J., BÜRKNER, P.-C. and VEHTARI, A. (2023). Using reference models in variable selection. *Comput. Statist.* **38** 349–371. MR4556024 https://doi.org/10.1007/s00180-022-01231-6

PEARL, J. (2009). Causal inference in statistics: An overview. *Stat. Surv.* **3** 96–146. MR2545291 https://doi.org/10.1214/09-SS057

PENROSE, K. W., NELSON, A. G. and FISHER, A. G. (1985). Generalized body composition prediction equation for men using simple measurement techniques. *Med. Sci. Sports Exerc.* **17** 189.

PÉREZ, P. and DE LOS CAMPOS, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* **198** 483–495. https://doi.org/10.1534/genetics.114.164442

PIIRONEN, J., PAASINIEMI, M., CATALINA, A., WEBER, F. and VEHTARI, A. (2023). Projpred: Projection predictive feature selection. R package, version 2.5.0.

PIIRONEN, J., PAASINIEMI, M. and VEHTARI, A. (2020). Projective inference in high-dimensional problems: Prediction and feature selection. *Electron. J. Stat.* **14** 2155–2197. MR4097052 https://doi.org/10.1214/20-EJS1711

PIIRONEN, J. and VEHTARI, A. (2016). Projection predictive model selection for Gaussian processes. In 2016 *IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)* 1–6. IEEE, Los Alamitos.

PIIRONEN, J. and VEHTARI, A. (2017a). Comparison of Bayesian predictive methods for model selection. *Stat. Comput.* **27** 711–735. MR3613594 https://doi.org/10.1007/s11222-016-9649-y

PIIRONEN, J. and VEHTARI, A. (2017b). Sparsity information and regularization in the horseshoe and other shrinkage priors. *Electron. J. Stat.* **11** 5018–5051. MR3738204 https://doi.org/10.1214/17-EJS1337SI

PIIRONEN, J. and VEHTARI, A. (2018). Iterative supervised principal components. In *Proceedings of the Twenty-First International Conference on Artificial Intelligence and Statistics* (A. Storkey and F. Perez-Cruz, eds.). *Proceedings of Machine Learning Research* **84** 106–114. PMLR.

PYA, N. (2022). Scam: Shape constrained additive models.

RAMSAY, J. O. (1988). Monotone regression splines in action. *Statist. Sci.* **3** 425–441.

ROSSELL, D., COOK, J. D., TELESCA, D., ROEBUCK, P., ABRIL, O. and TORRENS, M. (2023). Mombf: Model selection with Bayesian methods and information criteria. R package, version 3.3.1.

ROSSO, V., ALBRECHT, J., ROBERTS, F. and JANEY, J. M. (2019). Uniting laboratory automation, DoE data, and modeling techniques to accelerate chemical process development. *React. Chem. Eng.* **4** 1646–1657.

SÄILYNOJA, T., BÜRKNER, P.-C. and VEHTARI, A. (2022). Graphical test for discrete uniformity and its applications in goodness-of-fit evaluation and multiple sample comparison. *Stat. Comput.* **32** 32. MR4402179 https://doi.org/10.1007/s11222-022-10090-6

SAN MARTINI, A. and SPEZZAFERRI, F. (1984). A predictive model selection criterion. *J. Roy. Statist. Soc. Ser. B* **46** 296–303. MR0781890

SCHEIPL, F. (2011). SpikeSlabGAM: Bayesian variable selection, model choice and regularization for generalized additive mixed models in R. *J. Stat. Softw.* **43** 1–24.

SCHILLER, B., WIRTHGEN, E., WEBER, F., SCHILLER, S., RADKE, M., CLASSEN, M., DÄBRITZ, J., BUDERUS, S., BUFLER, P. et al. (2024). Fecal calprotectin and platelet count predict histologic disease activity in pediatric ulcerative colitis: Results from a projection-predictive feature selection. *Eur. J. Pediatr.*

SCHRITTENLOCHER, S., BACHMANN, B., TIURBE, A. M., TUAC, O., VELTEN, K., SCHMIDT, D. and CURSIEFEN, C. (2019). Impact of preoperative visual acuity on Descemet Membrane Endothelial Keratoplasty (DMEK) outcome. *Graefe Arch. Clin. Exp. Ophthalmol.* **257** 321–329.

SIVULA, T., MAGNUSSON, M., MATAMOROS, A. A. and VEHTARI, A. (2022). Uncertainty in Bayesian leave-one-out cross-validation based model comparison. arXiv preprint. Available at arXiv:2008.10296.

SOLIMAN, S. A., HAQUE, A., VANARSA, K., ZHANG, T., ISMAIL, F., LEE, K. H., PEDROZA, C., GREENBAUM, L. A., MASON, S. et al. (2022). Urine ALCAM, PF4 and VCAM-1 surpass conventional metrics in identifying nephritis disease activity in childhood-onset systemic lupus erythematosus. *Front. Immunol.* **13** 885307.

SOUTH, L. F., RIABIZ, M., TEYMUR, O. and OATES, C. J. (2022). Postprocessing of MCMC. *Annu. Rev. Stat. Appl.* **9** 529–555. MR4394919 https://doi.org/10.1146/annurev-statistics-040220-091727

STELL, L. and SABATTI, C. (2015). ptycho: Bayesian variable selection with hierarchical priors. R package, version 1.1-4.

SULLIVAN, B., BARKER, E., WILLIAMS, P., MACGREGOR, L., BHAMBER, R., THOMAS, M., GURNEY, S., HYAMS, C., WHITEWAY, A. et al. (2022). Conventional and Bayesian workflows for clinical prediction modelling of severe Covid-19 outcomes based on clinical biomarker test results: LabMarCS: Laboratory markers of COVID-19 severity—bristol cohort preprint Infectious Diseases (except HIV/AIDS).

TALTS, S., BETANCOURT, M., SIMPSON, D., VEHTARI, A. and GELMAN, A. (2020). Validating Bayesian inference algorithms with simulation-based calibration. arXiv preprint. Available at arXiv:1804.06788.

TENG, K. T.-Y., MARTINEZ AVILÉS, M., UGARTE-RUIZ, M., BARCENA, C., DE LA TORRE, A., LOPEZ, G., MORENO, M. A., DOMINGUEZ, L. and ALVAREZ, J. (2020). Spatial trends in salmonella infection in pigs in Spain. *Front. Vet. Sci.* **7** 345.

TIBSHIRANI, R. (2011). Regression shrinkage and selection via the lasso: A retrospective. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 273–282. MR2815776 https://doi.org/10.1111/j.1467-9868.2011.00771.x

TRAN, M.-N., NOTT, D. J. and LENG, C. (2012). The predictive Lasso. *Stat. Comput.* **22** 1069–1084. MR2950086 https://doi.org/10.1007/s11222-011-9279-3

TRIER, H., O'REILLY, J., SPIERING, L., MA, S., KOLLING, N., RUSHWORTH, M. and SCHOLL, J. (2023). Emotions and individual differences shape foraging under threat preprint PsyArXiv.

VANARSA, K., CASTILLO, J., WANG, L., LEE, K. H., PEDROZA, C., LOTAN, Y. and MOHAN, C. (2023). Comprehensive proteomics and platform validation of urinary biomarkers for bladder cancer diagnosis and staging. *BMC Med.* **21** 133.

VEHTARI, A. (2021). On reparameterization invariant Bayesian point estimates and credible regions. arXiv preprint. Available at arXiv:2109.10843.

VEHTARI, A., GABRY, J., MAGNUSSON, M., YAO, Y., BÜRKNER, P.-C., PAANANEN, T. and GELMAN, A. (2022). loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. R package version 2.5.0.

VEHTARI, A., GELMAN, A. and GABRY, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27** 1413–1432. MR3647105 https://doi.org/10.1007/s11222-016-9696-4

VEHTARI, A., GELMAN, A., SIMPSON, D., CARPENTER, B. and BÜRKNER, P.-C. (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Anal.* **16** 667–718. MR4298989 https://doi.org/10.1214/20-ba1221

VEHTARI, A. and OJANEN, J. (2012). A survey of Bayesian predictive methods for model assessment, selection and comparison. *Stat. Surv.* **6** 142–228. MR3011074 https://doi.org/10.1214/12-SS102

VEHTARI, A., SIMPSON, D., GELMAN, A., YAO, Y. and GABRY, J. (2024). Pareto smoothed importance sampling. *J. Mach. Learn. Res.* **25** 72. MR4749108

WANG, L. and YOON, K.-J. (2022). Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks. *IEEE Trans. Pattern Anal. Mach. Intell.* **44** 3048–3068.

WEBER, F., GLASS, Ä. and VEHTARI, A. (2024). Projection predictive variable selection for discrete response families with finite support. *Comput. Statist.*

WIRTHGEN, E., WEBER, F., KUBICKOVA-WEBER, L., SCHILLER, B., SCHILLER, S., RADKE, M. and DÄBRITZ, J. (2023). Identifying predictors of clinical outcomes using the projection-predictive feature selection—a proof of concept on the example of Crohn's disease. *Front. Pediatr.* **11** 1170563.

XU, M. and DUAN, L. L. (2020). Bayesian inference with the l_1-ball prior: Solving combinatorial problems with exact zeros. *J. Roy. Statist. Soc. Ser. B* **85** 1538–1560.

YANCHENKO, E., BONDELL, H. D. and REICH, B. J. (2023). R2D2 goes to space! A principled approach to setting prior distributions on spatial parameters.

YANCHENKO, E., BONDELL, H. D. and REICH, B. J. (2024). The R2D2 prior for generalized linear mixed models. *Amer. Statist.* 1–10.

YAO, Y., CARVALHO, L. M., MESQUITA, D. and MCLATCHIE, Y. (2023). Locking and quacking: Stacking Bayesian model predictions by log-pooling and superposition.

YAO, Y., PIRŠ, G., VEHTARI, A. and GELMAN, A. (2022). Bayesian hierarchical stacking: Some models are (somewhere) useful. *Bayesian Anal.* **17** 1043–1071. MR4506021 https://doi.org/10.1214/21-ba1287

YAO, Y., VEHTARI, A., SIMPSON, D. and GELMAN, A. (2018). Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Anal.* **13** 917–1007.

ZHANG, Y. and PELEATO, N. M. (2021). Predicting cyanobacteria abundance with Bayesian zero-inflated negative binomial models. *SSRN Electron. J.*.

ZHANG, Y. D., NAUGHTON, B. P., BONDELL, H. D. and REICH, B. J. (2022). Bayesian regression using a prior on the model fit: The R2-D2 shrinkage prior. *J. Amer. Statist. Assoc.* **117** 862–874. MR4436318 https://doi.org/10.1080/01621459.2020.1825449

ZHAO, Z., BANTERLE, M., BOTTOLO, L., RICHARDSON, S., LEWIN, A. and ZUCKNICK, M. (2021). BayesSUR: An R package for high-dimensional multivariate Bayesian variable and covariance selection in linear regression. *J. Stat. Softw.* **100** 1–32.

ZHOU, G., LEE, M.-C., WANG, X., ZHONG, D. and YAN, G. (2022). Mapping potential malaria vector larval habitats for larval source management: Introduction to multi-model ensembling approaches preprint Ecology.

ZOU, H. (2006). The adaptive lasso and its oracle properties. *J. Amer. Statist. Assoc.* **101** 1418–1429. MR2279469 https://doi.org/10.1198/016214506000000735

Statistical Science
2025, Vol. 40, No. 1, 148–166
https://doi.org/10.1214/23-STS903

A Conversation with A. Philip Dawid

Vladimir Vovk and Glenn Shafer

Abstract. Beginning in the 1970s, Alexander Philip Dawid has been a leading contributor to the foundations of statistics and especially to the development and application of Bayesian statistics. He is also known for his work on causality, especially his notation for conditional independence and his critique of the overuse of counterfactuals, and for his contributions to forensic statistics.

Dawid was born in Lancashire, England, on February 1, 1946. His family moved to London soon afterwards, and he attended the City of London School from 1956 to 1963. He studied mathematics at Cambridge, earning a BA (Bachelor of Arts) degree in 1966. After earning a Diploma in Mathematical Statistics in the academic year 1966–1967, he studied for a PhD at Imperial, then at UCL, where he became a Lecturer in Statistics in 1969. In 1978, he left UCL for a position as Professor of Statistics in the Department of Mathematics, The City University, London, where he served as Head of Statistics Section and Director of the Statistical Laboratory. He returned to the Department of Statistics at UCL in 1981, serving as Head of Department from 1983 to 1993. He moved to the University of Cambridge in 2007, becoming Professor of Statistics and Fellow of Darwin College. He has continued his work in mathematical statistics after retiring from Cambridge in 2013 and was elected Fellow of the Royal Society in 2018.

Key words and phrases: A. Philip Dawid, biography, Bayesian statistics, prequential statistics, causality, University College London.

REFERENCES

BERNARDO, J.-M. and SMITH, A. F. M. (2000). *Bayesian Theory*. *Wiley Series in Probability and Mathematical Statistics*: *Probability and Mathematical Statistics*. Wiley, Chichester. MR1274699 https://doi.org/10.1002/9780470316870

BERNARDO, J. M. (2009). The Valencia story. International Society for Bayesian Analysis (ISBA) newsletter. Available on the web in August 2023. https://www.uv.es/bernardo/ValenciaStory.pdf.

DAWID, A. P. (1970). On the limiting normality of posterior distributions. *Proc*. *Camb*. *Philos*. *Soc*. **67** 625–633. MR0258170 https://doi.org/10.1017/s0305004100045953

DAWID, A. P. (1979). Conditional independence in statistical theory (with discussion). *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **41** 1–31. MR0535541

DAWID, A. P. (1982). The well-calibrated Bayesian. *J*. *Amer*. *Statist*. *Assoc*. **77** 605–613. With comments by Joseph B. Kadane and D. V. Lindley and a rejoinder by the author. MR0675887

DAWID, A. P. (1984a). Causal inference from messy data (invited discussion of "The nature and discovery of structure" by J. W. Pratt and Robert Schlaifer). *J*. *Amer*. *Statist*. *Assoc*. **79** 22–24.

DAWID, A. P. (1984b). Statistical theory. The prequential approach. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *A* **147** 278–292. MR0763811 https://doi.org/10.2307/2981683

DAWID, A. P. (1991). Fisherian inference in likelihood and prequential frames of reference. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **53** 79–109. With discussion and a reply by the author. MR1094276

DAWID, A. P. (2000). Causal inference without counterfactuals. *J*. *Amer*. *Statist*. *Assoc*. **95** 407–448. With comments and a rejoinder by the author. MR1803167 https://doi.org/10.2307/2669377

DAWID, A. P. (2004). Probability, causality and the empirical world: A Bayes–de Finetti–Popper–Borel synthesis. *Statist*. *Sci*. **19** 44–57. MR2082146 https://doi.org/10.1214/088342304000000125

DAWID, A. P. (2007). Counterfactuals, hypotheticals and potential responses: A philosophical examination of statistical causality. In *Causality and Probability in the Sciences* (F. Russo and J. Williamson). 503–532. College Sci. Pub., London.

DAWID, A. P. and STONE, M. (1972). Expectation consistency of inverse probability distributions. *Biometrika* **59** 486–489. MR0343447 https://doi.org/10.1093/biomet/59.2.486

DAWID, A. P., STONE, M. and ZIDEK, J. V. (1973). Marginalization paradoxes in Bayesian and structural inference. *J*. *Roy*. *Statist*. *Soc*. *Ser*. *B* **35** 189–233. MR0365805

DAWID, A. P. and TEWARI, A. (2022). On learnability under general stochastic processes. *Harv*. *Data Sci*. *Rev*. **4**.

DAWID, A. P. and VOVK, V. G. (1999). Prequential probability: Principles and properties. *Bernoulli* **5** 125–162. MR1673572 https://doi.org/10.2307/3318616

Vladimir Vovk is Professor, Department of Computer Science, Royal Holloway, University of London, Egham, Surrey, UK (e-mail: v.vovk@rhul.ac.uk). Glenn Shafer is University Professor, Rutgers University, Newark, New Jersey, USA (e-mail: gshafer@business.rutgers.edu).

DAWID, P. (2022). Decision-theoretic foundations for statistical causality: Response to Pearl. *J. Causal Inference* **10** 296–299.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach. Probability and Mathematical Statistics, Vol.* 1. Academic Press, New York. MR0215390

FREEDMAN, D. A. (1963). Invariants under mixing which generalize de Finetti's theorem: Continuous time parameter. *Ann. Math. Stat.* **34** 1194–1216. MR0189111 https://doi.org/10.1214/aoms/1177703856

GALBRAITH, J. I. (2021). Obituary: Mervyn Stone, 1932–2020. *J. Roy. Statist. Soc. Ser. A* **184** 396–399.

KINGMAN, J. F. C. (1972). On random sequences with spherical symmetry. *Biometrika* **59** 492–494. MR0343420 https://doi.org/10.1093/biomet/59.2.492

KINGMAN, J. F. C. (1978). Uses of exchangeability. *Ann. Probab.* **6** 183–197. MR0494344 https://doi.org/10.1214/aop/1176995566

LINDLEY, D. V. (1953). Statistical inference. *J. Roy. Statist. Soc. Ser. B* **15** 30–65, discussion, 65–76. MR0057522

LINDLEY, D. V. (1965). *Introduction to Probability and Statistics from a Bayesian Viewpoint. Part I: Probability*. Cambridge Univ. Press, New York. MR0168083

LINDLEY, D. V. (1971). *Bayesian Statistics, a Review. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No.* 2. SIAM, Philadelphia, PA. MR0329081

OLIVER, A. (2019). "The thing about statistics is there's not necessarily a right answer": An interview with Philip Dawid. Available on the web in August 2023. https://www.statisticsviews.com/article/the-thing-about-statistics-is-theres-not-necessarily-a-right-answer-an-interview-with-philip-dawid/.

PEARL, J. (2022). Causation and decision: On Dawid's "Decision theoretic foundation of statistical causality". *J. Causal Inference* **10** 221–226.

SAVAGE, L. J. (1954). *The Foundations of Statistics*. Wiley, New York. Second edition (1972): Dover, New York. MR0063582

SMITH, A. (1995). A conversation with Dennis Lindley. *Statist. Sci.* **10** 305–319. MR1390520

SMITH, A. F. M. (1981). On random sequences with centred spherical symmetry. *J. Roy. Statist. Soc. Ser. B* **43** 208–209. MR0626767

STONE, M. and DAWID, A. P. (1972). Un-Bayesian implications of improper Bayes inference in routine statistical problems. *Biometrika* **59** 369–375. MR0431449 https://doi.org/10.1093/biomet/59.2.369

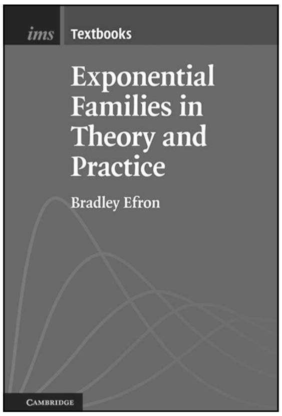
VOVK, V., GAMMERMAN, A. and SHAFER, G. (2022). *Algorithmic Learning in a Random World* 2nd ed. Springer, Cham.

WHITTLE, P. (2002). A Realised path: The Cambridge Statistical Laboratory upto 1993 (revised 2002). Available on the web in August 2023. https://www.statslab.cam.ac.uk/history-statistical-laboratory.

NEW BOOK

The Institute of Mathematical Statistics presents

IMS TEXTBOOKS



Exponential Families in Theory and Practice

Bradley Efron, Stanford University

During the past half-century, exponential families have attained a position at the center of parametric statistical inference. Theoretical advances have been matched, and more than matched, in the world of applications, where logistic regression by itself has become the go-to methodology in medical statistics, computer-based prediction algorithms, and the social sciences. This book is based on a one-semester graduate course for first year Ph.D. and advanced master's students. After presenting the basic structure of univariate and multivariate exponential families, their application to generalized linear models including logistic and Poisson regression is described in detail, emphasizing geometrical ideas, computational practice, and the analogy with ordinary linear regression. Connections are made with a variety of current statistical methodologies: missing data, survival analysis and proportional hazards, false discovery rates, bootstrapping, and empirical Bayes analysis. The book connects exponential family theory with its applications in a way that doesn't require advanced mathematical preparation.

Hardback $105.00

Paperback $39.99

IMS members are entitled to a 40% discount: email ims@imstat.org to request your code

www.imstat.org/cup/

Cambridge University Press, with the Institute of Mathematical Statistics, established the *IMS Monographs* and *IMS Textbooks* series of high-quality books. The series editors are Mark Handcock, Ramon van Handel, Arnaud Doucet, and John Aston.