

THE ANNALS *of* APPLIED STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles

- Elicitability and backtesting: Perspectives for banking regulation
NATALIA NOLDE AND JOHANNA F. ZIEGEL 1833
- DiscussionHAJO HOLZMANN AND BERNHARD KLAR 1875
- DiscussionPATRICK SCHMIDT 1883
- DiscussionMARK H. A. DAVIS 1886
- DiscussionCHEN ZHOU 1888
- DiscussionMARIE KRATZ 1894
- RejoinderNATALIA NOLDE AND JOHANNA F. ZIEGEL 1901
- Estimating average causal effects under general interference, with application to a social
network experimentPETER M. ARONOW AND CYRUS SAMII 1912
- Co-evolution of social networks and continuous actor attributes
NYNKE M. D. NIEZINK AND TOM A. B. SNIJDERS 1948
- Bayesian sensitivity analysis for causal effects from 2×2 tables in the presence of
unmeasured confounding with application to presidential campaign visits
LUKE KEELE AND KEVIN M. QUINN 1974
- Model-based clustering with data correction for removing artifacts in gene expression data
WILLIAM CHAD YOUNG, ADRIAN E. RAFTERY AND KA YEE YEUNG 1998
- A unified framework for variance component estimation with summary statistics in
genome-wide association studiesXIANG ZHOU 2027
- Photodegradation modeling based on laboratory accelerated test data and predictions
under outdoor weathering for polymeric materialsYUANYUAN DUAN,
YILI HONG, WILLIAM Q. MEEKER, DEBORAH L. STANLEY AND XIAOHONG GU 2052
- Variable selection for latent class analysis with application to low back pain diagnosis
MICHAEL FOP, KEITH M. SMART AND THOMAS BRENDAN MURPHY 2080
- Inference for respondent-driven sampling with misclassification
ISABELLE S. BEAUDRY, KRISTA J. GILE AND SHRUTI H. MEHTA 2111
- Learning population and subject-specific brain connectivity networks via mixed
neighborhood selectionRICARDO PIO MONTI,
CHRISTOFOROS ANAGNOSTOPOULOS AND GIOVANNI MONTANA 2142
- Multivariate spatiotemporal modeling of age-specific stroke mortality
HARRISON QUICK, LANCE A. WALLER AND MICHELE CASPER 2165
- A random-effects hurdle model for predicting bycatch of endangered marine species
E. CANTONI, J. MILLS FLEMMING AND A. H. WELSH 2178
- Empirical Bayesian analysis of simultaneous changepoints in multiple data sequences
ZHOU FAN AND LESTER MACKEY 2200
- Bayesian inference for multiple Gaussian graphical models with application to metabolic
association networksLINDA S. L. TAN, AJAY JASRA, MARIA DE IORIO
AND TIMOTHY M. D. EBBELS 2222

continued

THE ANNALS *of* APPLIED STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles—*Continued from front cover*

- Semiparametric covariate-modulated local false discovery rate for genome-wide
association studies RONG W. ZABLOCKI, RICHARD A. LEVINE,
ANDREW J. SCHORK, SHUJING XU, YUNPENG WANG, CHUN C. FAN
AND WESLEY K. THOMPSON 2252
- Point process models for spatio-temporal distance sampling data from a large-scale survey
of blue whales YUAN YUAN, FABIAN E. BACHL, FINN LINDGREN,
DAVID L. BORCHERS, JANINE B. ILLIAN, STEPHEN T. BUCKLAND,
HÅVARD RUE AND TIM GERRODETTE 2270
- Modeling node incentives in directed networks DEEPAYAN CHAKRABARTI 2298
- Automatic matching of bullet land impressions
ERIC HARE, HEIKE HOFMANN AND ALICIA CARRIQUIRY 2332
- Estimating the number of casualties in the American Indian war: A Bayesian analysis
using the power law distribution COLIN S. GILLESPIE 2357
- Statistical downscaling for future extreme wave heights in the North Sea
ROSS TOWE, EMMA EASTOE, JONATHAN TAWN AND PHILIP JONATHAN 2375
- Focusing on regions of interest in forecast evaluation
HAJO HOLZMANN AND BERNHARD KLAR 2404

THE ANNALS OF APPLIED STATISTICS

Vol. 11, No. 4, pp. 1833–2431 December 2017

INSTITUTE OF MATHEMATICAL STATISTICS

(Organized September 12, 1935)

The purpose of the Institute is to foster the development and dissemination of the theory and applications of statistics and probability.

IMS OFFICERS

- President:** Alison Etheridge, Department of Statistics, University of Oxford, Oxford, OX1 3LB, United Kingdom
- President-Elect:** Xiao-Li Meng, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA
- Past President:** Jon Wellner, Department of Statistics, University of Washington, Seattle, Washington 98195-4322, USA
- Executive Secretary:** Edsel Peña, Department of Statistics, University of South Carolina, Columbia, South Carolina 29208-001, USA
- Treasurer:** Zhengjun Zhang, Department of Statistics, University of Wisconsin, Madison, Wisconsin 53706-1510, USA
- Program Secretary:** Judith Rousseau, Université Paris Dauphine, Place du Maréchal DeLattre de Tassigny, 75016 Paris, France

IMS PUBLICATIONS

- The Annals of Statistics.** *Editors:* Edward I. George, Department of Statistics, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; Tailen Hsing, Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109-1107, USA
- The Annals of Applied Statistics.** *Editor-In-Chief:* Tilmann Gneiting, Heidelberg Institute for Theoretical Studies, HITS gGmbH, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany, and Institute for Stochastics, Karlsruhe Institute of Technology, Englerstr. 2, 76128 Karlsruhe, Germany
- The Annals of Probability.** *Editor:* Maria Eulália Vares, Instituto de Matemática, Universidade Federal do Rio de Janeiro, 21941-909 Rio de Janeiro, RJ, Brazil
- The Annals of Applied Probability.** *Editor:* Bálint Tóth, School of Mathematics, University of Bristol, University Walk, BS8 1TW, Bristol, U.K. and Institute of Mathematics, Technical University Budapest, Egry József u. 1, H-1111 Budapest, Hungary
- Statistical Science.** *Editor:* Cun-Hui Zhang, Department of Statistics, Rutgers University, Piscataway, New Jersey 08854, USA
- The IMS Bulletin.** *Editor:* Anirban DasGupta, Department of Statistics, Purdue University, West Lafayette, Indiana 47907-2068, USA

The Annals of Applied Statistics [ISSN 1932-6157 (print); ISSN 1941-7330 (online)], Volume 11, Number 4, December 2017. Published quarterly by the Institute of Mathematical Statistics, 3163 Somerset Drive, Cleveland, Ohio 44122, USA. Periodicals postage pending at Cleveland, Ohio, and at additional mailing offices.

POSTMASTER: Send address changes to *The Annals of Applied Statistics*, Institute of Mathematical Statistics, Dues and Subscriptions Office, 9650 Rockville Pike, Suite L 2310, Bethesda, Maryland 20814-3998, USA.

ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION¹

BY NATALIA NOLDE² AND JOHANNA F. ZIEGEL³

University of British Columbia and University of Bern

Conditional forecasts of risk measures play an important role in internal risk management of financial institutions as well as in regulatory capital calculations. In order to assess forecasting performance of a risk measurement procedure, risk measure forecasts are compared to the realized financial losses over a period of time and a statistical test of correctness of the procedure is conducted. This process is known as backtesting. Such traditional backtests are concerned with assessing some optimality property of a set of risk measure estimates. However, they are not suited to compare different risk estimation procedures. We investigate the proposal of comparative backtests, which are better suited for method comparisons on the basis of forecasting accuracy, but necessitate an elicitable risk measure. We argue that supplementing traditional backtests with comparative backtests will enhance the existing trading book regulatory framework for banks by providing the correct incentive for accuracy of risk measure forecasts. In addition, the comparative backtesting framework could be used by banks internally as well as by researchers to guide selection of forecasting methods. The discussion focuses on three risk measures, Value at Risk, expected shortfall and expectiles, and is supported by a simulation study and data analysis.

REFERENCES

- ACERBI, C. and SZEKELY, B. (2014). Backtesting expected shortfall. *Risk Mag.* December 76–81.
- ANDREWS, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59** 817–858. [MR1106513](#)
- BANK FOR INTERNATIONAL SETTLEMENTS (2013). Consultative document: Fundamental review of the trading book: A revised marked risk framework. Available at <http://www.bis.org/publ/bcbs265.pdf>.
- BANK FOR INTERNATIONAL SETTLEMENTS (2014). Consultative document: Fundamental review of the trading book: Outstanding issues. Available at <http://www.bis.org/bcbs/publ/d305.pdf>.
- BELLINI, F. and BIGNOZZI, V. (2015). On elicitable risk measures. *Quant. Finance* **15** 725–733. [MR3334566](#)
- BELLINI, F. and DI BERNARDINO, E. (2017). Risk management with expectiles. *Eur. J. Finance* **23** 487–506.
- BELLINI, F., KLAR, B., MÜLLER, A. and GIANIN, E. R. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.* **54** 41–48. [MR3145849](#)
- BOLLERSLEV, T. and WOOLDRIDGE, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Rev.* **11** 143–172. [MR1185178](#)

Key words and phrases. Forecasting, backtesting, elicibility, risk measurement procedure, Value at Risk, expected shortfall, expectiles.

- CHRISTOFFERSEN, P. F. (1998). Evaluating interval forecasts. *Internat. Econom. Rev.* **39** 841–862. [MR1661906](#)
- CHRISTOFFERSEN, P. (2003). *Elements of Financial Risk Management*. Academic Press.
- CONT, R., DEGUEST, R. and SCANDOLO, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quant. Finance* **10** 593–606. [MR2676786](#)
- COSTANZINO, N. and CURRAN, M. (2015). Backtesting general spectral risk measures with application to expected shortfall. *Journal of Risk Model Validation* **9** 21–33.
- DAVIS, M. H. A. (2016). Verification of internal risk measure estimates. *Stat. Risk Model.* **33** 67–93. [MR3574946](#)
- DELBAEN, F., BELLINI, F., BIGNOZZI, V. and ZIEGEL, J. F. (2016). Risk measures with the CxLS property. *Finance Stoch.* **20** 433–453. [MR3479327](#)
- DIEBOLD, F. X., GUNTHER, T. A. and TAY, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *Internat. Econom. Rev.* **39** 863–883.
- DIEBOLD, F. X. and MARIANO, R. S. (1995). Comparing predictive accuracy. *J. Bus. Econom. Statist.* **13** 253–263.
- DIEBOLD, F. X., SCHUERMANN, T. and STROUGHAIR, J. D. (2000). Pitfalls and opportunities in the use of extreme value theory in risk management. *J. Risk Finance* **1** 30–35.
- EFRON, B. (1991). Regression percentiles using asymmetric squared error loss. *Statist. Sinica* **1** 93–125. [MR1101317](#)
- EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability **57**. Chapman & Hall, New York. [MR1270903](#)
- EHM, W., GNEITING, T., JORDAN, A. and KRÜGER, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **78** 505–562. [MR3506792](#)
- EMBRECHTS, P., KLÜPPELBERG, C. and MIKOSCH, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Applications of Mathematics (New York) **33**. Springer, Berlin. [MR1458613](#)
- EMMER, S., KRATZ, M. and TASCHKE, D. (2015). What is the best risk measure in practice? A comparison of standard measures. *J. Risk* **18** 31–60.
- ENGLE, R. F. and MANGANELLI, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *J. Bus. Econom. Statist.* **22** 367–381. [MR2091566](#)
- FERNÁNDEZ, C. and STEEL, M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *J. Amer. Statist. Assoc.* **93** 359–371. [MR1614601](#)
- FISSLER, T. and ZIEGEL, J. F. (2016). Higher order elicibility and Osband’s principle. *Ann. Statist.* **44** 1680–1707. [MR3519937](#)
- FISSLER, T., ZIEGEL, J. F. and GNEITING, T. (2016). Expected shortfall is jointly elicitable with value at risk—implications for backtesting. *Risk Mag.* January 58–61.
- FÖLLMER, H. and SCHIED, A. (2002). Convex measures of risk and trading constraints. *Finance Stoch.* **6** 429–447. [MR1932379](#)
- FRONGILLO, R. and KASH, I. (2015). Vector-valued property elicitation. In *Proceedings of the 28th Conference on Learning Theory* (S. Kale, P. Grünwald and E. Hazan, eds.). *JMLR Workshop and Conference Proceedings* **40**.
- GIACOMINI, R. and WHITE, H. (2006). Tests of conditional predictive ability. *Econometrica* **74** 1545–1578. [MR2268409](#)
- GNEITING, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. [MR2847988](#)
- GNEITING, T. and RANJAN, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *J. Bus. Econom. Statist.* **29** 411–422. [MR2848512](#)
- HOLZMANN, H. and EULERT, M. (2014). The role of the information set for forecasting—with applications to risk management. *Ann. Appl. Stat.* **8** 595–621. [MR3192004](#)
- HOMMEL, G. (1983). Tests of the overall hypothesis for arbitrary dependence structures. *Biom. J.* **25** 423–430. [MR0735888](#)

- KOENKER, R. (2005). *Quantile Regression. Econometric Society Monographs* **38**. Cambridge Univ. Press, Cambridge. [MR2268657](#)
- KOU, S. and PENG, X. (2016). On the measurement of economic tail risk. *Oper. Res.* **64** 1056–1072. [MR3558435](#)
- KUAN, C.-M., YEH, J.-H. and HSU, Y.-C. (2009). Assessing value at risk with CARE, the conditional autoregressive expectile models. *J. Econometrics* **150** 261–270. [MR2535521](#)
- KUESTER, K., MITTNIK, S. and PAOLELLA, M. S. (2006). Value-at-risk prediction: A comparison of alternative strategies. *J. Financ. Econom.* **4** 53–89.
- LAMBERT, N. (2013). Elicitation and evaluation of statistical functionals. Preprint. Available at https://web.stanford.edu/~nlambert/papers/elicitatio_n_july2013.pdf.
- LAMBERT, N., PENNOCK, D. M. and SHOHAM, Y. (2008). Eliciting properties of probability distributions. In *Proceedings of the 9th ACM Conference on Electronic Commerce* 129–138. Chicago, IL. Extended abstract.
- MCNEIL, A. J. and FREY, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Finance* **7** 271–300.
- MCNEIL, A. J., FREY, R. and EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools. Princeton Series in Finance*. Princeton Univ. Press, Princeton, NJ. [MR2175089](#)
- NAU, R. F. (1985). Should scoring rules be “effective”? *Manage. Sci.* **31** 527–535.
- NEWBY, W. K. and POWELL, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* **55** 819–847. [MR0906565](#)
- NOLDE, N. and ZIEGEL, J. F. (2017). Supplement to “Elicitability and backtesting: Perspectives for banking regulation”. DOI:[10.1214/17-AOAS1041SUPP](https://doi.org/10.1214/17-AOAS1041SUPP).
- OSBAND, K. H. (1985). Providing incentives for better cost forecasting. Ph.D. thesis, Univ. California, Berkeley.
- PATTON, A. J. (2006). Volatility forecast comparison using imperfect volatility proxies. Research Paper 175, Quantitative Finance Research Centre, Univ. Technology, Sydney.
- PATTON, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *J. Econometrics* **160** 246–256. [MR2745881](#)
- PATTON, A. J. (2014). Evaluating and comparing possibly misspecified forecasts. Working paper.
- PATTON, A. J. and SHEPPARD, K. (2009). Evaluating volatility and correlation forecasts. In *Handbook of Financial Time Series* (T. Mikosch, J.-P. Kreiss, R. A. Davis and T. G. Andersen, eds.) 801–838. Springer, Berlin.
- R CORE TEAM (2015). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at <http://www.R-project.org/>.
- ROSENBLATT, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.* **23** 470–472. [MR0049525](#)
- SAERENS, M. (2000). Building cost functions minimizing to some summary statistics. *IEEE Trans. Neural Netw.* **11** 1263–1271.
- STEINWART, I., PASIN, C., WILLIAMSON, R. and ZHANG, S. (2014). Elicitation and identification of properties. *J. Mach. Learn. Res. Workshop Conf. Proc.* **35** 1–45.
- STRÄHL, C. and ZIEGEL, J. (2017). Cross-calibration of probabilistic forecasts. *Electron. J. Stat.* **11** 608–639. [MR3619318](#)
- THOMSON, W. (1979). Eliciting production possibilities from a well-informed manager. *J. Econom. Theory* **20** 360–380. [MR0540822](#)
- TSYPLAKOV, A. (2014). Theoretical guidelines for a partially informed forecast examiner. MPRA Paper, 55017. Available at <http://mpra.ub.uni-muenchen.de/55017>.
- WANG, R. and ZIEGEL, J. F. (2015). Elicitable distortion risk measures: A concise proof. *Statist. Probab. Lett.* **100** 172–175. [MR3324090](#)
- WEBER, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Math. Finance* **16** 419–441. [MR2212272](#)

ZIEGEL, J. F. (2016). Coherence and elicibility. *Math. Finance* **26** 901–918. [MR3551510](#)

DISCUSSION OF “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY HAJO HOLZMANN AND BERNHARD KLAR

Philipps-Universität Marburg and Karlsruher Institut für Technologie (KIT)

In our discussion of the insightful paper by Nolde and Ziegel, we further investigate comparative backtests based on consistent scoring rules. We use Diebold–Mariano tests in pairwise comparisons instead of mere rankings in terms of average scores, and illustrate the use of weighted proper scoring rules, which address the quality of forecasts of the full loss distribution in its upper tail rather than some specific risk measure such as the Value at Risk. Overall, at lower levels up to 95%, these allow for better discrimination between competing forecasting methods.

REFERENCES

- DAVIS, M. H. A. (2016). Verification of internal risk measure estimates. *Stat. Risk Model.* **33** 67–93. [MR3574946](#)
- DIEBOLD, F. X. and MARIANO, R. S. (1995). Comparing predictive accuracy. *J. Bus. Econom. Statist.* **13** 253–263.
- FISSLER, T. and ZIEGEL, J. F. (2016). Higher order elicibility and Osband’s principle. *Ann. Statist.* **44** 1680–1707. [MR3519937](#)
- FISSLER, T., ZIEGEL, J. F. and GNEITING, T. (2016). Expected shortfall is jointly elicitable with value at risk—implications for backtesting. *Risk Mag.* January 58–61.
- GHALANOS, A. (2014). rugarch: Univariate GARCH models. R package version 1.3-5.
- GNEITING, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. [MR2847988](#)
- GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. [MR2345548](#)
- GNEITING, T. and RANJAN, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *J. Bus. Econom. Statist.* **29** 411–422. [MR2848512](#)
- HOLZMANN, H. and KLAR, B. (2017). Focusing on regions of interest in forecast evaluation. *Ann. Appl. Stat.* **11** 2404–2431.
- MCNEIL, A. J. and FREY, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Finance* **7** 271–300.
- MCNEIL, A. J., FREY, R. and EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools. Princeton Series in Finance.* Princeton Univ. Press, Princeton, NJ. [MR2175089](#)
- R CORE TEAM (2016). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria.
- STEPHENSON, A. G. (2002). evd: Extreme value distributions. *R News* **2** 31–32.
- STRAUMANN, D. (2005). *Estimation in Conditionally Heteroscedastic Time Series Models. Lecture Notes in Statistics* **181**. Springer, Berlin. [MR2142271](#)

Key words and phrases. Backtesting, forecasting, risk management, scoring rule.

DISCUSSION OF “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY PATRICK SCHMIDT

Heidelberg Institute for Theoretical Studies and Goethe University Frankfurt

I discuss the incentive compatibility of comparative and calibration backtesting for banking regulation. In stylized models of risk reporting, calibration backtesting leads to uninformed risk reports that adapt insufficiently to volatility changes. In contrast, comparative backtesting incentivizes information for richer and more accurate models.

REFERENCES

- BANK FOR INTERNATIONAL SETTLEMENTS (2013). Fundamental review of the trading book: A revised market risk framework. Available at <http://www.bis.org/publ/bcbs265.pdf>.
- BANK FOR INTERNATIONAL SETTLEMENTS (2016). Minimum capital requirements for market risk. Available at <http://www.bis.org/bcbs/publ/d352.pdf>.
- BERKOWITZ, J., CHRISTOFFERSEN, P. and PELLETIER, D. (2011). Evaluating value-at-risk models with desk-level data. *Manage. Sci.* **57** 2213–2227.
- BERKOWITZ, J. and O'BRIEN, J. (2002). How accurate are Value-at-Risk models at commercial banks? *J. Finance* **57** 1093–1111.
- NOLDE, N. and ZIEGEL, J. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Ann. Appl. Stat.* **11** 1833–1874.
- PÉRIGNON, C. and SMITH, D. R. (2010). The level and quality of Value-at-Risk disclosure by commercial banks. *J. Bank. Financ.* **34** 362–377.
- PÉRIGNON, C., DENG, YIN, Z. and WANG, Z. (2008). Do banks overstate their Value-at-Risk? *J. Bank. Financ.* **32** 783–794.

DISCUSSION OF “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY MARK H. A. DAVIS

Imperial College London

REFERENCES

- DAVIS, M. H. A. (2016). Verification of internal risk measure estimates. *Stat. Risk Model.* **33** 67–93. [MR3574946](#)
- GIACOMINI, R. and WHITE, H. (2006). Tests of conditional predictive ability. *Econometrica* **74** 1545–1578. [MR2268409](#)
- MCLEISH, D. L. (1977). On the invariance principle for nonstationary mixingales. *Ann. Probab.* **5** 616–621. [MR0445583](#)
- WILLIAMS, D. (1991). *Probability with Martingales*. Cambridge Univ. Press, Cambridge. [MR1155402](#)

DISCUSSION ON “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY CHEN ZHOU

De Nederlandsche Bank and Erasmus University Rotterdam

REFERENCES

- ACERBI, C. and SZEKELY, B. (2014). Backtesting expected shortfall. *Risk Mag.* **December** 76–81.
- ANGELIDIS, T., BENOS, A. and DEGIANNAKIS, S. (2004). The use of GARCH models in VaR estimation. *Stat. Methodol.* **1** 105–128.
- ENGLE, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *J. Econ. Perspect.* **15** 157–168.
- MCNEIL, A. J. and FREY, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Finance* **7** 271–300.
- NOLDE, N. and ZIEGEL, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Ann. Appl. Stat.* **11** 1833–1874.
- ZWINGMANN, T. and HOLZMANN, H. (2016). Asymptotics for the expected shortfall. arXiv preprint, [arXiv:1611.07222](https://arxiv.org/abs/1611.07222).

DISCUSSION OF “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY MARIE KRATZ

ESSEC Business School, CREAR Risk Research Center

The discussion focuses on four points in the context of Basel 3. The first concerns the choice of test functions in the calibration tests. Then we discuss the interpretation of the “internal model,” as well as the choice of risk measures. Last, we consider the score difference stationarity, an important issue in practice.

REFERENCES

- ACERBI, C. and SZÉKELY, B. (2014). Backtesting expected shortfall. *Risk Mag.* **December** 76–81.
- ACERBI, C. and TASCHE, D. (2002). On the coherence of expected shortfall. *J. Bank. Financ.* **26** 1487–1503.
- ARTZNER, P., DELBAEN, F., EBER, J.-M. and HEATH, D. (1999). Coherent measures of risk. *Math. Finance* **9** 203–228. [MR1850791](#)
- BASEL COMMITTEE ON BANKING SUPERVISION (2016). Minimum capital requirements for market risk. Bank of International Settlements.
- BELLINI, F., KLAR, B., MÜLLER, A. and GIANIN, E. R. (2014). Generalized quantiles as risk measures. *Insurance Math. Econom.* **54** 41–48. [MR3145849](#)
- CAMPBELL, S. D. (2006). A review of backtesting and backtesting procedures. *J. Risk* **9** 1–17.
- CHEN, J. M. (2013). Measuring market risk under Basel II, 2.5, and III: VAR, stressed VAR, and expected shortfall. Available at SSRN: <http://ssrn.com/abstract=2252463>.
- CHRISTOFFERSEN, P. (2003). *Elements of Financial Risk Management*. Academic Press, New York.
- CHRISTOFFERSEN, P. and PELLETIER, D. (2004). Backtesting Value-at-Risk: A duration-based approach. *J. Financ. Econom.* **2** 84–108.
- COSTANZINO, N. and CURRAN, M. (2015). Backtesting general spectral risk measures with application to expected shortfall. *J. Risk Model Validation* **9** 21–33.
- COSTANZINO, N. and CURRAN, M. (2016). A simple traffic light approach to backtesting expected shortfall. Working paper.
- DIEBOLD, F. X. and MARIANO, R. S. (1995). Comparing predictive accuracy. *J. Bus. Econom. Statist.* **13** 253–263.
- EMMER, S., KRATZ, M. and TASCHE, D. (2015). What is the best risk measure in practice? A comparison of standard risk measures. *J. Risk* **18** 31–60.
- FISSLER, T. and ZIEGEL, J. F. (2016). Higher order elicibility and Osband’s principle. *Ann. Statist.* **44** 1680–1707. [MR3519937](#)
- GNENING, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. [MR2847988](#)
- KRATZ, M., LOK, Y. and MCNEIL, A. (2016a). A multinomial test to discriminate between models. In *ASTIN 2016 Proceedings*. Available online.
- KRATZ, M., LOK, Y. and MCNEIL, A. (2016b). Multinomial VaR Backtests: A simple implicit approach to backtesting expected shortfall. ESSEC Working Paper 1617, [arXiv1611.04851v1](#).

Key words and phrases. Regulation, risk measure, scoring function.

- NOLDE, N. and ZIEGEL, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Ann. Appl. Stat.* **11** 1833–1874.
- TASCHE, D. (2002). Expected shortfall and beyond. In *Statistical Data Analysis Based on the L_1 -Norm and Related Methods* (Neuchâtel, 2002) 109–123. Birkhäuser, Basel. [MR2001308](#)
- ZIEGEL, J. F. (2016). Coherence and elicibility. *Math. Finance* **26** 901–918. [MR3551510](#)

REJOINDER: “ELICITABILITY AND BACKTESTING: PERSPECTIVES FOR BANKING REGULATION”

BY NATALIA NOLDE AND JOHANNA F. ZIEGEL

University of British Columbia and University of Bern

REFERENCES

- ACERBI, C. and SZEKELY, B. (2014). Backtesting expected shortfall. *Risk Mag.* **December** 76–81.
- AGYEMAN, J. (2017). On the choice of scoring functions for forecast comparisons. Master’s thesis, Univ. British Columbia. Available at <https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0344014>.
- BEE, M., DUPUIS, D. J. and TRAPIN, L. (2016). Realizing the extremes: Estimation of tail-risk measures from a high-frequency perspective. *J. Empir. Finance* **36** 86–99.
- DAVIS, M. H. A. (2016). Verification of internal risk measure estimates. *Stat. Risk Model.* **33** 67–93. [MR3574946](#)
- DIEBOLD, F. X., GUNTHER, T. A. and TAY, A. S. (1998). Evaluating density forecasts with applications to financial risk management. *Internat. Econom. Rev.* **39** 863–883.
- DIEBOLD, F. X. and MARIANO, R. S. (1995). Comparing predictive accuracy. *J. Bus. Econom. Statist.* **13** 253–263.
- DIMITRIADIS, T. and BAYER, S. (2017). A joint quantile and expected shortfall regression framework. Preprint, [arXiv:1704.02213](#).
- ENGLE, R. F. and MANGANELLI, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *J. Bus. Econom. Statist.* **22** 367–381. [MR2091566](#)
- GNEITING, T., BALABDAOUI, F. and RAFTERY, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 243–268. [MR2325275](#)
- GNEITING, T. and KATZFUSS, M. (2014). Probabilistic forecasting. *Ann. Rev. Stat. Appl.* **1** 125–151.
- GNEITING, T. and RANJAN, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *J. Bus. Econom. Statist.* **29** 411–422. [MR2848512](#)
- GNEITING, T. and RANJAN, R. (2013). Combining predictive distributions. *Electron. J. Stat.* **7** 1747–1782. [MR3080409](#)
- GORDY, M. B., LOK, H. Y. and MCNEIL, A. J. (2017). Spectral backtests of forecast distributions with applications to risk management. Preprint, [arXiv:1708.01489](#).
- HOLZMANN, H. and EULERT, M. (2014). The role of the information set for forecasting—with applications to risk management. *Ann. Appl. Stat.* **8** 595–621. [MR3192004](#)
- HOLZMANN, H. and KLAR, B. (2016). Weighted scoring rules and hypothesis testing. Preprint, [arXiv:1611.07345](#).
- KRATZ, M., LOK, Y. and MCNEIL, A. (2016). Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. Preprint, [arXiv:1611.04851](#).
- MCNEIL, A. J. and FREY, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Finance* **7** 271–300.
- PATTON, A., ZIEGEL, J. F. and CHEN, R. (2017). Dynamic semiparametric models for expected shortfall (and value-at-risk). Preprint, [arXiv:1707.05108](#).
- SHEPHARD, N. and SHEPPARD, K. (2010). Realising the future: Forecasting with high-frequency-based volatility (heavy) models. *J. Appl. Econometrics* **25** 197–231. [MR2758633](#)
- STRÄHL, C. and ZIEGEL, J. (2017). Cross-calibration of probabilistic forecasts. *Electron. J. Stat.* **11** 608–639. [MR3619318](#)

- TAYLOR, J. W. (2017). Forecasting value at risk and expected shortfall using a semiparametric approach based on the asymmetric Laplace distribution. *J. Bus. Econom. Statist.* To appear.
- ZIEGEL, J. F., KRÜGER, F., JORDAN, A. and FASCIATI, F. (2017). Murphy diagrams for evaluating forecasts of expected shortfall. Preprint, [arXiv:1705.04537](https://arxiv.org/abs/1705.04537).

ESTIMATING AVERAGE CAUSAL EFFECTS UNDER GENERAL INTERFERENCE, WITH APPLICATION TO A SOCIAL NETWORK EXPERIMENT

BY PETER M. ARONOW AND CYRUS SAMII

Yale University and New York University

This paper presents a randomization-based framework for estimating causal effects under interference between units motivated by challenges that arise in analyzing experiments on social networks. The framework integrates three components: (i) an experimental design that defines the probability distribution of treatment assignments, (ii) a mapping that relates experimental treatment assignments to exposures received by units in the experiment, and (iii) estimands that make use of the experiment to answer questions of substantive interest. We develop the case of estimating average unit-level causal effects from a randomized experiment with interference of arbitrary but known form. The resulting estimators are based on inverse probability weighting. We provide randomization-based variance estimators that account for the complex clustering that can occur when interference is present. We also establish consistency and asymptotic normality under local dependence assumptions. We discuss refinements including covariate-adjusted effect estimators and ratio estimation. We evaluate empirical performance in realistic settings with a naturalistic simulation using social network data from American schools. We then present results from a field experiment on the spread of anti-conflict norms and behavior among school students.

REFERENCES

- AIROLDI, E. M. (2016). Estimating Causal Effects in the Presence of Interfering Units. Paper Presented at YINS Distinguished Lecture Series, Yale University.
- ANGRIST, J. D. and KRUEGER, A. B. (1999). Empirical strategies in labor economics. In *Handbook of Labor Economics* (O. C. Ahsenfelder and D. Card, eds.) **3**. North-Holland, Amsterdam.
- ARAL, S. and WALKER, D. (2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Manage. Sci.* **57** 1623–1639.
- ARONOW, P. M. (2013). Model Assisted Causal Inference. Ph.D. thesis, Dept. Political Science, Yale Univ., New Haven, CT.
- ARONOW, P. M. and MIDDLETON, J. A. (2013). A class of unbiased estimators of the average treatment effect in randomized experiments. *Journal of Causal Inference* **1** 135–144.
- ARONOW, P. M. and SAMII, C. (2013). Conservative variance estimation for sampling designs with zero pairwise inclusion probabilities. *Surv. Methodol.* **39** 231–241.
- ARONOW, P. M., SAMII, C. and ASSENOVA, V. A. (2015). Cluster robust variance estimation for dyadic data. *Polit. Anal.* **23** 546–577.

Key words and phrases. Interference, potential outcomes, causal inference, randomization inference, SUTVA, networks.

- BAIRD, S., BOHREN, J. A., MCINTOSH, C. and OZLER, B. (2016). Designing Experiments to Measure Spillover Effects. Unpublished manuscript, George Washington Univ., Univ. Pennsylvania, Univ. California, San Diego, and World Bank.
- BASU, D. (1971). An essay on the logical foundations of survey sampling. I. In *Foundations of Statistical Inference* (V. Godambe and D. A. Sprott, eds.) 203–242. Holt, Rinehart and Winston, Toronto.
- BOWERS, J., FREDRICKSON, M. M. and PANAGOPOLOUS, C. (2013). Reasoning about interference between units: A general framework. *Polit. Anal.* **21** 97–124.
- BRAMOULLÉ, Y., DJEBBARI, H. and FORTIN, B. (2009). Identification of peer effects through social networks. *J. Econometrics* **150** 41–55. [MR2525993](#)
- BREWER, K. R. W. (1979). A class of robust sampling designs for large-scale surveys. *J. Amer. Statist. Assoc.* **74** 911–915. [MR0556487](#)
- CHEN, J., HUMPHREYS, M. and MODI, V. (2010). Technology Diffusion and Social Networks: Evidence from a Field Experiment in Uganda. Manuscript, Columbia Univ., New York.
- CHEN, L. H. Y. and SHAO, Q.-M. (2004). Normal approximation under local dependence. *Ann. Probab.* **32** 1985–2028. [MR2073183](#)
- CHUNG, H., LANZA, S. T. and LOKEN, E. (2008). Latent transition analysis: Inference and estimation. *Stat. Med.* **27** 1834–1854. [MR2420348](#)
- COLE, S. R. and FRANGAKIS, C. E. (2009). The consistency statement in causal inference: A definition or an assumption? *Epidemiology* **20** 3–5.
- COX, D. R. (1958). *Planning of Experiments*. Wiley, New York. [MR0095561](#)
- ECKLES, D., KARRER, B. and UGANDER, J. (2014). Design and analysis of experiments in networks: Reducing bias from interference. Preprint. Available at [arXiv:1404.7530](#) [stat.ME].
- FATTORINI, L. (2006). Applying the Horvitz–Thompson criterion in complex designs: A computer-intensive perspective for estimating inclusion probabilities. *Biometrika* **93** 269–278. [MR2278082](#)
- FREEDMAN, D., PISANI, R. and PURVES, R. (1998). *Statistics*, 2nd ed. W.W. Norton, New York.
- FULLER, W. A. (2009). *Sampling Statistics*. Wiley, Hoboken.
- GOEL, S. and SALGANIK, M. J. (2010). Assessing respondent-driven sampling. *Proc. Nat. Acad. Sci.* **107** 6743–6747.
- GOODREAU, S. M. (2007). Advances in exponential random graph (p-star) models applied to a large social network. *Soc. Netw.* **29** 231–248.
- GOODREAU, S. M., KITTS, J. A. and MORRIS, M. (2009). Birds of a feather, or friend of a friend? Using exponential random graph models to investigate adolescents in social networks. *Demography* **46** 103–125.
- HÁJEK, J. (1971). Comment on “An essay on the logical foundations of survey sampling, part one”. In *Foundations of Statistical Inference* (V. P. Godambe and D. A. Sprott, eds.) Holt, Rinehart and Winston, Toronto.
- HANDCOCK, M. S., RAFTERY, A. E. and TANTRUM, J. M. (2007). Model-based clustering for social networks. *J. Roy. Statist. Soc. Ser. A* **170** 301–354. [MR2364300](#)
- HARTLEY, H. O. and ROSS, A. (1954). Unbiased ratio estimators. *Nature* **174** 270–271.
- HOLLAND, P. W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc.* **81** 945–970. [MR0867618](#)
- HORVITZ, D. G. and THOMPSON, D. J. (1952). A generalization of sampling without replacement from a finite universe. *J. Amer. Statist. Assoc.* **47** 663–685. [MR0053460](#)
- HUDGENS, M. G. and HALLORAN, M. E. (2008). Toward causal inference with interference. *J. Amer. Statist. Assoc.* **103** 832–842. [MR2435472](#)
- HUNTER, D. R., GOODREAU, S. M. and HANDCOCK, M. S. (2008). Goodness of fit of social network models. *J. Amer. Statist. Assoc.* **103** 248–258. [MR2394635](#)
- IMBENS, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika* **87** 706–710. [MR1789821](#)

- IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge Univ. Press, New York. [MR3309951](#)
- IMBENS, G. W. and WOOLDRIDGE, J. M. (2009). Recent developments in the econometrics of program evaluation. *J. Econ. Lit.* **47** 5–86.
- ISAKI, C. T. and FULLER, W. A. (1982). Survey design under the regression superpopulation model. *J. Amer. Statist. Assoc.* **77** 89–96. [MR0648029](#)
- LIN, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s critique. *Ann. Appl. Stat.* **7** 295–318. [MR3086420](#)
- LIU, L. and HUDGENS, M. G. (2014). Large sample randomization inference of causal effects in the presence of interference. *J. Amer. Statist. Assoc.* **109** 288–301. [MR3180564](#)
- MACKINNON, J. G. and WHITE, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *J. Econometrics* **29** 305–325.
- MANSKI, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard Univ. Press, Cambridge, MA.
- MANSKI, C. F. (2013). Identification of treatment response with social interactions. *Econom. J.* **16** S1–S23. [MR3030060](#)
- MIDDLETON, J. A. and ARONOW, P. M. (2011). Unbiased estimation of the average treatment effect in cluster-randomized experiments. Paper Presented at the Annual Meeting of the Midwest Political Science Association, Chicago, IL.
- OGBURN, E. L. and VANDERWEELE, T. J. (2014). Causal diagrams for interference. *Statist. Sci.* **29** 559–578. [MR3300359](#)
- PALUCK, E. L. (2011). Peer pressure against prejudice: A high school field experiment examining social network change. *J. Exp. Psychol.* **47** 350–358.
- PALUCK, E. L., SHEPHERD, H. and ARONOW, P. M. (2016). Changing climates of conflict: A social network experiment in 56 schools. *Proc. Nat. Acad. Sci.* **113** 566–571.
- RAJ, D. (1965). On a method of using multi-auxiliary information in sample surveys. *J. Amer. Statist. Assoc.* **60** 270–277. [MR0179875](#)
- ROBINSON, P. M. (1982). On the convergence of the Horvitz–Thompson estimator. *Austral. J. Statist.* **24** 234–238. [MR0678263](#)
- ROSENBAUM, P. R. (1999). Reduced sensitivity to hidden bias at upper quantiles in observational studies with dilated treatment effects. *Biometrics* **55** 560–564.
- ROSENBAUM, P. R. (2007). Interference between units in randomized experiments. *J. Amer. Statist. Assoc.* **102** 191–200. [MR2345537](#)
- RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. [MR0472152](#)
- RUBIN, D. B. (1990). Formal Models of Statistical Inference for Causal Effects. *J. Statist. Plann. Inference* **25** 279–292.
- SÄRNDAL, C.-E., SWENSSON, B. and WRETMAN, J. (1992). *Model Assisted Survey Sampling*. Springer, New York. [MR1140409](#)
- SOBEL, M. E. (2006). What do randomized studies of housing mobility demonstrate?: Causal inference in the face of interference. *J. Amer. Statist. Assoc.* **101** 1398–1407. [MR2307573](#)
- NEYMAN, J. S. (1990). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist. Sci.* **5** 465–472. [MR1092986](#)
- TCHETGEN TCHETGEN, E. J. and VANDERWEELE, T. J. (2012). On causal inference in the presence of interference. *Stat. Methods Med. Res.* **21** 55–75. [MR2867538](#)
- TOULIS, P. and KAO, E. (2013). Estimation of causal peer influence effects. *J. Mach. Learn. Res.* **28** 1489–1497.
- UGANDER, J., KARRER, B., BACKSTROM, L. and KLEINBERG, J. (2013). Graph cluster randomization: Network exposure to multiple universes. Preprint. Available at [arXiv:1305.6979](#) [cs.SI].
- VANDERWEELE, T. J. (2009). Concerning the consistency assumption in causal inference. *Epidemiology* **20** 880–883.

WOOD, J. (2008). On the covariance between related Horvitz–Thompson estimators. *J. Off. Stat.* **24** 53–78.

CO-EVOLUTION OF SOCIAL NETWORKS AND CONTINUOUS ACTOR ATTRIBUTES

BY NYNKE M. D. NIEZINK¹ AND TOM A. B. SNIJDERS

University of Groningen

Social networks and the attributes of the actors in these networks are not static; they may develop interdependently over time. The stochastic actor-oriented model allows for statistical inference on the mechanisms driving this co-evolution process. In earlier versions of this model, dynamic actor attributes are assumed to be measured on an ordinal categorical scale. We present an extension of the stochastic actor-oriented model that does away with this restriction using a stochastic differential equation to model the evolution of continuous actor attributes. We estimate the parameters by a procedure based on the method of moments. The proposed method is applied to study the dynamics of a friendship network among the students at an Australian high school. In particular, we model the relationship between friendship and obesity, focusing on body mass index as a continuous co-evolving attribute.

REFERENCES

- AGNEESSENS, F. and WITTEK, R. (2008). Social capital and employee well-being: Disentangling intrapersonal and interpersonal selection and influence mechanisms. *Revue Française de Sociologie* **49** 613–637.
- AMATI, V., SCHÖNENBERGER, F. and SNIJDERS, T. A. B. (2015). Estimation of stochastic actor-oriented models for the evolution of networks by generalized method of moments. *J. SFdS* **156** 140–165. [MR3432607](#)
- ARNOLD, L. (1974). *Stochastic Differential Equations: Theory and Applications*. Wiley-Interscience, New York. Translated from the German. [MR0443083](#)
- BERGSTROM, A. R. (1984). Continuous time stochastic models and issues of aggregation over time. In *Handbook of Econometrics, Vol. 2* (Z. Griliches and M. D. Intriligator, eds.). *Handbooks in Economics* **2** 1146–1212. North-Holland, Amsterdam. [MR0772383](#)
- BERGSTROM, A. R. (1988). The history of continuous-time econometric models. *Econometric Theory* **4** 365–383. [MR0985152](#)
- BLACK, F. and SCHOLES, M. (1973). The pricing of options and corporate liabilities. *J. Polit. Econ.* **81** 637–654. [MR3363443](#)
- CHRISTAKIS, N. A. and FOWLER, J. H. (2007). The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357** 370–379.
- COHEN-COLE, E. and FLETCHER, J. M. (2008). Is obesity contagious? Social networks vs. environmental factors in obesity epidemic. *J. Health Econ.* **27** 1382–1387.
- DE LA HAYE, K., ROBINS, G., MOHR, P. and WILSON, C. (2011). Homophily and contagion as explanations for weight similarities among adolescent friends. *J. Adolesc. Health* **49** 421–427.

Key words and phrases. Social networks, longitudinal data, Markov model, stochastic differential equations, method of moments.

- DIJKSTRA, J. K., LINDENBERG, S., VEENSTRA, R., STEGLICH, C., ISAACS, J., CARD, N. A. and HODGES, E. V. E. (2010). Influence and selection processes in weapon carrying during adolescence: The roles of status, aggression, and vulnerability. *Criminology* **48** 187–220.
- DIJKSTRA, J. K., GEST, S. D., LINDENBERG, S., VEENSTRA, R. and CILLESSEN, A. H. N. (2012). The emergence of weapon carrying in peer context. Testing three explanations: The role of aggression, victimization, and friends. *J. Adolesc. Health* **50** 371–376.
- HAMERLE, A., SINGER, H. and NAGL, W. (1993). Identification and estimation of continuous time dynamic systems with exogenous variables using panel data. *Econometric Theory* **9** 283–295. [MR1229417](#)
- HOLLAND, P. W. and LEINHARDT, S. (1977/1978). A dynamic model for social networks. *J. Math. Sociol.* **5** 5–20. [MR0446596](#)
- HUNTER, D. R. (2007). Curved exponential family models for social networks. *Soc. Netw.* **29** 216–230.
- KOSKINEN, J. H. and SNIJDERS, T. A. B. (2007). Bayesian inference for dynamic social network data. *J. Statist. Plann. Inference* **137** 3930–3938. [MR2368538](#)
- KUSHNER, H. J. and YIN, G. G. (2003). *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. *Applications of Mathematics (New York): Stochastic Modelling and Applied Probability* **35**. Springer, New York. [MR1993642](#)
- LEHMANN, E. L. (1999). *Elements of Large-Sample Theory*. Springer, New York. [MR1663158](#)
- MCFADDEN, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Economics* (P. Zarembka, ed.) 105–142. Academic Press, New York.
- MERTON, R. C. (1990). *Continuous-Time Finance*. Basil Blackwell, Oxford.
- ODD, J. H. L. and JANSEN, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika* **65** 199–215. [MR1763519](#)
- ODD, J. H. L., FOLMER, H., PATUELLI, R. and NIJKAMP, P. (2012). Continuous-time modeling with spatial dependence. *Geogr. Anal.* **44** 29–46.
- RIPLEY, R. M., SNIJDERS, T. A. B., BODA, Z., VÖRÖS, A. and PRECIADO, P. (2017). Manual for RSiena (version March 2017). Dept. Statistics, and Nuffield College, Univ. Oxford, Oxford.
- ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400–407. [MR0042668](#)
- SCHWEINBERGER, M. and SNIJDERS, T. A. B. (2007). Markov models for digraph panel data: Monte Carlo-based derivative estimation. *Comput. Statist. Data Anal.* **51** 4465–4483. [MR2364459](#)
- SHALIZI, C. R. and THOMAS, A. C. (2011). Homophily and contagion are generically confounded in observational social network studies. *Sociol. Methods Res.* **40** 211–239. [MR2767833](#)
- SINGER, H. (1996). Continuous-time dynamic models for panel data. In *Analysis of Change: Advanced Techniques in Panel Data Analysis* (U. Engel and J. Reinecke, eds.) 113–133. de Gruyter, Berlin.
- SNIJDERS, T. A. B. (2001). The statistical evaluation of social network dynamics. In *Sociological Methodology* (M. Sobel and M. Becker, eds.) 361–395. Basil Blackwell, Boston.
- SNIJDERS, T. A. B. (2005). Models for longitudinal network data. In *Models and Methods in Social Network Analysis* (P. J. Carrington, J. Scott and S. S. Wasserman, eds.) 215–247. Cambridge Univ. Press, New York.
- SNIJDERS, T. A. B., KOSKINEN, J. and SCHWEINBERGER, M. (2010). Maximum likelihood estimation for social network dynamics. *Ann. Appl. Stat.* **4** 567–588. [MR2758640](#)
- SNIJDERS, T. A. B., STEGLICH, C. E. G. and SCHWEINBERGER, M. (2007). Modeling the co-evolution of networks and behavior. In *Longitudinal Models in the Behavioral and Related Sciences* (K. van Montfort, H. Oud and A. Satorra, eds.) 41–71. Cambridge Univ. Press, New York.
- STEELE, J. M. (2001). *Stochastic Calculus and Financial Applications. Applications of Mathematics (New York): Stochastic Modelling and Applied Probability* **45**. Springer, New York. [MR1783083](#)

- STEGELICH, C. E. G., SNIJDERS, T. A. B. and PEARSON, M. (2010). Dynamic networks and behavior: Separating selection from influence. *Sociol. Method.* **40** 329–392.
- VOELKLE, M. C., OUD, J. H. L., DAVIDOV, E. and SCHMIDT, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychol. Methods* **17** 176–192.
- WASSERMAN, S. S. (1980). Analyzing social networks as stochastic processes. *J. Amer. Statist. Assoc.* **75** 280–294.
- WASSERMAN, S. S. and FAUST, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge Univ. Press, Cambridge.
- WEERMAN, F. (2011). Delinquent peers in context: A longitudinal network analysis of selection and influence effects. *Criminology* **49** 253–286.

BAYESIAN SENSITIVITY ANALYSIS FOR CAUSAL EFFECTS FROM 2×2 TABLES IN THE PRESENCE OF UNMEASURED CONFOUNDING WITH APPLICATION TO PRESIDENTIAL CAMPAIGN VISITS

BY LUKE KEELE AND KEVIN M. QUINN¹

Georgetown University and University of Michigan

Presidents often campaign on behalf of candidates during elections. Do these campaign visits increase the probability that the candidate will win? While one might attempt to answer this question by adjusting for observed covariates, such an approach is plagued by serious data limitations. In this paper we pursue a different approach. Namely, we ask: what, if anything, should one infer about the causal effect of a presidential campaign visit using a simple cross-tabulation of the data? We take a Bayesian approach to this problem and show that if one is willing to use substantive information to make some (possibly weak) assumptions about the nature of the unmeasured confounding, sharp posterior estimates of causal effects are easy to calculate. Using data from the 2002 midterm elections, we find that, under a reasonable set of assumptions, a presidential campaign visit on the behalf of congressional candidates helped those candidates win elections.

REFERENCES

- ANGRIST, J. D., IMBENS, G. W. and RUBIN, D. B. (1996). Identification of causal effects using instrumental variables. *J. Amer. Statist. Assoc.* **91** 444–455.
- BALKE, A. and PEARL, J. (1997). Bounds on treatment effects from studies with imperfect compliance. *J. Amer. Statist. Assoc.* **92** 1171–1176.
- CHICKERING, D. M. and PEARL, J. (1997). A clinician’s tool for analyzing non-compliance. *Computing Science and Statistics* **29** 424–431.
- COHEN, J. E., KRASSA, M. A. and HAMMAN, J. A. (1991). The impact of presidential campaigning on midterm U.S. senate elections. *Am. Polit. Sci. Rev.* **85** 165–178.
- COLE, S. R. and HERNÁN, M. A. (2008). Constructing inverse probability weights for marginal structural models. *Am. J. Epidemiol.* **168** 656–664.
- CORNFIELD, J., HAENSZEL, W., HAMMOND, E., LILIENTHAL, A., SHIMKIN, M. and WYNDER, E. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *J. Natl. Cancer Inst.* **22** 173–203.
- COX, D. R. (1958). *Planning of Experiments*. Wiley, New York. [MR0095561](#)
- DING, P. and DASGUPTA, T. (2016). A potential tale of two-by-two tables from completely randomized experiments. *J. Amer. Statist. Assoc.* **111** 157–168. [MR3494650](#)
- DING, P. and VANDERWEELE, T. J. (2016a). Sensitivity analysis without assumptions. *Epidemiology* **27** 368.
- DING, P. and VANDERWEELE, T. J. (2016b). Sharp sensitivity bounds for mediation under unmeasured mediator-outcome confounding. *Biometrika* **103** 483–490. [MR3509901](#)

Key words and phrases. Causal inference, Bayesian statistics, partial identification, sensitivity analysis.

- GELMAN, A., CARLIN, J. B., STERN, H. S. and RUBIN, D. B. (2004). *Bayesian Data Analysis*, 2nd ed. Chapman & Hall, Boca Raton, FL. [MR2027492](#)
- GILL, J. (2007). *Bayesian Methods: A Social and Behavioral Science Approach*, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.
- GILL, J. and WALKER, L. D. (2005). Elicited priors for Bayesian model specifications in political science research. *The Journal of Politics* **67** 841–872.
- GLYNN, A. N. and QUINN, K. M. (2011). Why process matters for causal inference. *Polit. Anal.* **19** 273–286.
- GRILLI, L. and MEALLI, F. (2008). Nonparametric bounds on the causal effect of university studies on job opportunities using principal stratification. *J. Educ. Behav. Stat.* **33** 111–130.
- GUSTAFSON, P. (2010). Bayesian inference for partially identified models. *Int. J. Biostat.* **6** Art. 17, 20. [MR2602560](#)
- GUSTAFSON, P., MCCANDLESS, L. C., LEVY, A. R. and RICHARDSON, S. (2010). Simplified Bayesian sensitivity analysis for mismeasured and unobserved confounders. *Biometrics* **66** 1129–1137. [MR2758500](#)
- HERRNSEN, P. S. and MORRIS, I. L. (2007). Presidential campaigning in the 2002 congressional elections. *Legis. Stud. Q.* **32** 629–648.
- HOLLAND, P. W. (1986). Statistics and causal inference. *J. Amer. Statist. Assoc.* **81** 945–970. [MR0867618](#)
- ICHINO, A., MEALLI, F. and NANNICINI, T. (2008). From temporary help jobs to permanent employment: What can we learn from matching estimators and their sensitivity? *J. Appl. Econometrics* **23** 305–327. [MR2420362](#)
- IMAI, K. and YAMAMOTO, T. (2008). Causal inference with measurement error: Nonparametric identification and sensitivity analyses of a field experiment on democratic deliberations. Paper presented at the 2008 Summer Political Methodology Meeting.
- IMBENS, G. W. (2003). Sensitivity to exogeneity assumptions in program evaluation. *Am. Econ. Rev. Pap. Proc.* **93** 126–132.
- JACOBSON, G. C. (2003). *The Politics of Congressional Elections*, 6th ed. Longman, New York.
- JIN, H. and RUBIN, D. B. (2008). Principal stratification for causal inference with extended partial compliance. *J. Amer. Statist. Assoc.* **103** 101–111. [MR2463484](#)
- KADANE, J. B. and WOLFSON, L. J. (1998). Experiences in elicitation. *Statistician* **47** 3–19.
- KEELE, L. J., FOGARTY, B. and STIMSON, J. A. (2004). The impact of presidential visits in the 2002 congressional elections. *PS Polit. Sci. Polit.* **34** 971–986.
- MANSKI, C. F. (1990). Nonparametric bounds on treatment effects. *Am. Econ. Rev. Pap. Proc.* **80** 319–323.
- MANSKI, C. F. (1993). Identification problems in the social sciences. *Sociol. Method.* **23** 1–56.
- MANSKI, C. F. (1995). *Identification Problems in the Social Sciences*. Harvard Univ. Press, Cambridge, MA.
- MANSKI, C. F. (1997). Monotone treatment response. *Econometrica* **65** 1311–1334. [MR1604297](#)
- MANSKI, C. F. (2003). *Partial Identification of Probability Distributions*. Springer, New York. [MR2151380](#)
- MCCANDLESS, L. C., GUSTAFSON, P. and LEVY, A. (2007). Bayesian sensitivity analysis for unmeasured confounding in observational studies. *Stat. Med.* **26** 2331–2347. [MR2368419](#)
- MEALLI, F. and PACINI, B. (2013). Using secondary outcomes to sharpen inference in randomized experiments with noncompliance. *J. Amer. Statist. Assoc.* **108** 1120–1131. [MR3174688](#)
- MOON, H. R. and SCHORFHEIDE, F. (2012). Bayesian and frequentist inference in partially identified models. *Econometrica* **80** 755–782. [MR2951948](#)
- NEUSTADT, R. E. (1960). *Presidential Power*. Wiley, New York.
- NUTTING, B. and STERN, A. H. (2001). *CQ's Politics in America: 2002 the 107th Congress*. Congressional Quarterly Press, Washington.

- PEARL, J. (1995). Causal diagrams for empirical research. *Biometrika* **82** 669–710. [MR1380809](#)
- PEARL, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge Univ. Press, Cambridge. [MR1744773](#)
- RICHARDSON, T. S., EVANS, R. J. and ROBINS, J. M. (2011). Transparent parametrizations of models for potential outcomes. In *Bayesian Statistics* **9** 569–610. Oxford Univ. Press, Oxford. [MR3204019](#)
- ROBINS, J. (1986). A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math. Modelling* **7** 1393–1512. [MR0877758](#)
- ROBINS, J. M. (1989). The analysis of randomized and nonrandomized AIDS treatment trials using a new approach to causal inference in longitudinal studies. In *Health Service Research Methodology: A Focus on AIDS* (L. Sechreset, H. Freeman and A. Mulley, eds.) U.S. Public Health Service, Washington, DC.
- ROBINS, J. M., HERNAN, M. A. and BRUMBACK, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* **11** 550–560.
- ROSENBAUM, P. R. (1987). Sensitivity analysis for certain permutation inferences in matched observational studies. *Biometrika* **74** 13–26. [MR0885915](#)
- ROSENBAUM, P. R. (2002). *Observational Studies*, 2nd ed. Springer, New York. [MR1899138](#)
- ROSENBAUM, P. R. and RUBIN, D. B. (1983). Assessing sensitivity to an unobserved covariate in an observational study with binary outcome. *J. R. Stat. Soc., B* **45** 212–218.
- RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. [MR0472152](#)
- RUBIN, D. B. (1990). Formal modes of statistical inference for causal effects. *J. Statist. Plann. Inference* **25** 279–292.
- RUBIN, D. B. (2008). For objective causal inference, design trumps analysis. *Ann. Appl. Stat.* **2** 808–804. [MR2516795](#)
- SELLERS, P. J. and DENTON, L. M. (2006). Presidential visits and midterm senate elections. *Pres. Stud. Q.* **36** 410–432.
- WESTERN, B. and JACKMAN, S. (1994). Bayesian inference for comparative research. *Am. Polit. Sci. Rev.* **88** 412–423.

MODEL-BASED CLUSTERING WITH DATA CORRECTION FOR REMOVING ARTIFACTS IN GENE EXPRESSION DATA¹

BY WILLIAM CHAD YOUNG, ADRIAN E. RAFTERY AND KA YEE YEUNG

University of Washington

The NIH Library of Integrated Network-based Cellular Signatures (LINCS) contains gene expression data from over a million experiments, using Luminex Bead technology. Only 500 colors are used to measure the expression levels of the 1000 landmark genes measured, and the data for the resulting pairs of genes are deconvolved. The raw data are sometimes inadequate for reliable deconvolution, leading to artifacts in the final processed data. These include the expression levels of paired genes being flipped or given the same value and clusters of values that are not at the true expression level. We propose a new method called model-based clustering with data correction (MCDC) that is able to identify and correct these three kinds of artifacts simultaneously. We show that MCDC improves the resulting gene expression data in terms of agreement with external baselines, as well as improving results from subsequent analysis.

REFERENCES

- BALL, C. A., SHERLOCK, G., PARKINSON, H., ROCCA-SERA, P., BROOKSBANK, C., CAUSTON, H. C., CAVALIERI, D., GAASTERLAND, T., HINGAMP, P., HOLSTEGE, F., RINGWALD, M., SPELLMAN, P., STOECKERT, C. J., STEWART, J. E., TAYLOR, R., BRAZMA, A., QUACKENBUSH, J. and MICROARRAY GENE EXPRESSION DATA (MGED) SOCIETY (2002). Standards for microarray data. *Science* Article ID **298** 539.
- BANFIELD, J. D. and RAFTERY, A. E. (1993). Model-based Gaussian and non-Gaussian clustering. *Biometrics* **49** 803–821.
- BANSAL, M., DELLA GATTA, G. and DI BERNARDO, D. (2006). Inference of gene regulatory networks and compound mode of action from time course gene expression profiles. *Bioinformatics* **22** 815–822.
- BASSO, K., MARGOLIN, A. A., STOLOVITZKY, G., KLEIN, U., DALLA-FAVERA, R. and CALIFANO, A. (2005). Reverse engineering of regulatory networks in human B cells. *Nat. Genet.* **37** 382–390.
- BIERNACKI, C., CELEUX, G. and GOVAERT, G. (2003). Choosing starting values for the EM algorithm for getting the highest likelihood in multivariate Gaussian mixture models. *Comput. Statist. Data Anal.* **41** 561–575.
- BINDER, H. and PREIBISCH, S. (2008). “hook”-calibration of GeneChip-microarrays: Theory and algorithm. *Algorithms Mol. Biol.* **3** 1–25.
- BLOCKER, A. W. and MENG, X.-L. (2013). The potential and perils of preprocessing: Building new foundations. *Bernoulli* **19** 1176–1211. [MR102548](#)
- BOLSTAD, B. M., IRIZARRY, R. A., ÅSTRAND, M. and SPEED, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19** 185–193.

Key words and phrases. Model-based clustering, MCDC, gene regulatory network, LINCS.

- CELEUX, G. and GOVAERT, G. (1995). Gaussian parsimonious clustering models. *Pattern Recognit.* **28** 781–793.
- CHEN, C., GRENNAN, K., BADNER, J., ZHANG, D., GERSHON, E., JIN, L. and LIU, C. (2011). Removing batch effects in analysis of expression microarray data: An evaluation of six batch adjustment methods. *PLoS ONE* **6** Article ID e17238.
- CHEN, E. Y., TAN, C. M., KOU, Y., DUAN, Q., WANG, Z., MEIRELLES, G. V., CLARK, N. R. and MA'AYAN, A. (2013a). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinform.* **14** Article ID 128. DOI:10.1186/1471-2105-14-128.
- CHEN, B., GREENSIDE, P., PAIK, H., SIROTA, M., HADLEY, D. and BUTTE, A. (2015). Relating chemical structure to cellular response: An integrative analysis of gene expression, bioactivity, and structural data across 11,000 compounds. *CPT: Pharmacom. Syst. Pharmacol.* **4** 576–584.
- COOKE, E. J., SAVAGE, R. S., KIRK, P. D. W., DARKINS, R. and WILD, D. L. (2011). Bayesian hierarchical clustering for microarray time series data with replicates and outlier measurements. *BMC Bioinform.* **12** Article ID 399. DOI:10.1186/1471-2105-12-399.
- D'HAESELEER, P., WEN, X., FUHRMAN, S. and SOMOGYI, R. (1999). Linear modeling of mRNA expression levels during CNS development and injury In *Pacific Symposium on Biocomputing* **4** 41–52. World Scientific, Singapore.
- DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood for incomplete data via the EM algorithm (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **39** 1–38.
- DUAN, Q., FLYNN, C., NIEPIL, M., HAFNER, M., MUHLICH, J. L., FERNANDEZ, N. F., ROUIL-LARD, A. D., TAN, C. M., CHEN, E. Y., GOLUB, T. R., SORGER, P. K., SUBRAMANIAN, A. and MA'AYAN A. (2014). LINC Canvas Browser: Interactive web app to query, browse and interrogate LINC L1000 gene expression signatures. *Nucleic Acids Res.* **42** W449–W460. DOI:10.1093/nar/gku476.
- DUNBAR, S. A. (2006). Applications of Luminex® xMAP™ technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin. Chim. Acta* **363** 71–82.
- FAITH, J. J., HAYETE, B., THADEN, J. T., MOGNO, I., WIERZBOWSKI, J., COTTAREL, G., KASIF, S., COLLINS, J. J. and GARDNER, T. S. (2007). Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. *PLoS Biol.* **5** Article ID e8. DOI:10.1371/journal.pbio.0050008.
- FLYNT, A. and DAEPP, M. I. (2015). Diet-related chronic disease in the northeastern United States: A model-based clustering approach. *Int. J. Health Geogr.* **14** 1–14.
- FRALEY, C. and RAFTERY, A. E. (1998). How many clusters? Which clustering method? Answers via model-based cluster analysis. *Comput. J.* **41** 578–588.
- FRALEY, C. and RAFTERY, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *J. Amer. Statist. Assoc.* **97** 611–631. MR1951635
- FRALEY, C. and RAFTERY, A. E. (2007). Model-based methods of classification: Using the mclust software in chemometrics. *J. Stat. Softw.* **18** 1–13.
- GOMEZ-ALVAREZ, V., TEAL, T. K. and SCHMIDT, T. M. (2009). Systematic artifacts in metagenomes from complex microbial communities. *ISME J.* **3** 1314–1317.
- GUSTAFSSON, M., HÖRNQUIST, M., LUNDSTRÖM, J., BJÖRKEGREN, J. and TEGNÉR, J. (2009). Reverse engineering of gene networks with LASSO and nonlinear basis functions. *Ann. N.Y. Acad. Sci.* **1158** 265–275.
- JIANG, D., TANG, C. and ZHANG, A. (2004). Cluster analysis for gene expression data: A survey. *IEEE Trans. Knowl. Data Eng.* **16** 1370–1386.
- KIM, S. Y., IMOTO, S. and MIYANO, S. (2003). Inferring gene networks from time series microarray data using dynamic Bayesian networks. *Brief. Bioinform.* **4** 228–235.
- LEEK, J. T., SCHARPF, R. B., BRAVO, H. C., SIMCHA, D., LANGMEAD, B., JOHNSON, W. E., GEMAN, D., BAGGERLY, K. and IRIZARRY, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* **11** 733–739.

- LEHMANN, R., MACHNE, R., GEORG, J., BENARY, M., AXMANN, I. M. and STEUER, R. (2013). How cyanobacteria pose new problems to old methods: Challenges in microarray time series analysis. *BMC Bioinform.* **14** Article ID 133.
- LI, Q., FRALEY, C., BUMGARNER, R. E., YEUNG, K. Y. and RAFTERY, A. E. (2005). Donuts, scratches and blanks: Robust model-based segmentation of microarray images. *Bioinformatics* **21** 2875–2882.
- LIU, X. and RATTRAY, M. (2010). Including probe-level measurement error in robust mixture clustering of replicated microarray gene expression. *Stat. Appl. Genet. Mol. Biol.* **9** Article ID 42.
- LIU, C., SU, J., YANG, F., WEI, K., MA, J. and ZHOU, X. (2015). Compound signature detection on LINCS L1000 big data. *Mol. BioSyst.* **11** 714–722.
- LO, K., RAFTERY, A. E., DOMBEK, K. M., ZHU, J., SCHADT, E. E., BUMGARNER, R. E. and YEUNG, K. Y. (2012). Integrating external biological knowledge in the construction of regulatory networks from time-series expression data. *BMC Syst. Biol.* **6** Article ID 101. DOI:[10.1186/1752-0509-6-101](#).
- LOCKHART, D. J., DONG, H., BYRNE, M. C., FOLLETTIE, M. T., GALLO, M. V., CHEE, M. S., MITTMANN, M., WANG, C., KOBAYASHI, M., HORTON, H. and BROWN, E. L. (1996). Expression monitoring by hybridization to high-density oligonucleotide arrays *Nat. Biotechnol.* **14** 1675–1680.
- MARGOLIN, A. A., NEMENMAN, I., BASSO, K., WIGGINS, C., STOLOVITZKY, G., FAVERA, R. D. and CALIFANO, A. (2006). ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. *BMC Bioinform.* **7** Article ID S7.
- MCLACHLAN, G. J. and KRISHNAN, T. (1997). *The EM Algorithm and Extensions*. Wiley, New York.
- MCLACHLAN, G. J. and PEEL, D. (2000). *Finite Mixture Models*. Wiley, New York.
- MEDVEDOVIC, M. and SIVAGANESAN, S. (2002). Bayesian infinite mixture model based clustering of gene expression profiles. *Bioinformatics* **18** 1194–1206.
- MENÉNDEZ, P., KOURMPETIS, Y. A. I., TER BRAAK, C. J. F. and VAN EEUWIJK, F. A. (2010). Gene regulatory networks from multifactorial perturbations using graphical Lasso: Application to the DREAM4 challenge. *PLoS ONE* **5** Article ID e14147. DOI:[10.1371/journal.pone.0014147](#).
- MEYER, P. E., KONTOS, K., LAFITTE, F. and BONTEMPI, G. (2007). Information-theoretic inference of large transcriptional regulatory networks. *EURASIP J. Bioinform. Syst. Biol.* **2007** Article ID 79879. DOI:[10.1155/2007/79879](#).
- MURPHY, K. and MIAN, S. (1999). Modelling gene expression data using dynamic Bayesian networks. Technical report, Computer Science Division, Univ. California, Berkeley.
- PECK, D., CRAWFORD, E. D., ROSS, K. N., STEGMAIER, K., GOLUB, T. R. and LAMB, J. (2006). A method for high-throughput gene expression signature analysis. *Genome Biol.* **7** Article ID R61. DOI:[10.1186/gb-2006-7-7-r61](#).
- PRICE, A. L., PATTERSON, N. J., PLENGE, R. M., WEINBLATT, M. E., SHADICK, N. A. and REICH, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38** 904–909.
- RAGHAVAN, V., BOLLMAN, P. and JUNG, G. S. (1989). A critical investigation of recall and precision as measures of retrieval system performance. *ACM Trans. Inf. Syst.* **7** 205–220.
- SANDELIN, A., ALKEMA, W., ENGSTRÖM, P., WASSERMAN, W. W. and LENHARD, B. (2004). JASPAR: An open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* **32** D91–D94.
- SCHENA, M., SHALON, D., DAVIS, R. W. and BROWN, P. O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270** 467–470.
- SEBASTIANI, P., GUSSONI, E., KOHANE, I. S. and RAMONI, M. F. (2003). Statistical challenges in functional genomics. *Statist. Sci.* **18** 33–70. MR1997065

- SHAO, H., PENG, T., JI, Z., SU, J. and ZHOU, X. (2013). Systematically studying kinase inhibitor induced signaling network signatures by integrating both therapeutic and side effects. *PLoS ONE* **8** Article ID e80832. DOI:[10.1371/journal.pone.0080832](https://doi.org/10.1371/journal.pone.0080832).
- SIEGMUND, K. D., LAIRD, P. W. and LAIRD-OFFRINGA, I. A. (2004). A comparison of cluster analysis methods using DNA methylation data. *Bioinformatics* **20** 1896–1904.
- STOKES, T. H., MOFFITT, R. A., PHAN, J. H. and WANG, M. D. (2007). Chip artifact CORRECTION (caCORRECT): A bioinformatics system for quality assurance of genomics and proteomics array data. *Ann. Biomed. Eng.* **35** 1068–1080.
- SUN, Z., CHAI, H. S., WU, Y., WHITE, W. M., DONKENA, K. V., KLEIN, C. J., GAROVIC, V. D., THERNEAU, T. M. and KOCHER, J.-P. A. (2011). Batch effect correction for genome-wide methylation data with illumina infinium platform. *BMC Med. Genomics* **4** Article ID 84. DOI:[10.1186/1755-8794-4-84](https://doi.org/10.1186/1755-8794-4-84).
- TEMPL, M., FILZMOSER, P. and REIMANN, C. (2008). Cluster analysis applied to regional geochemical data: Problems and possibilities. *Appl. Geochem.* **23** 2198–2213.
- VEMPATI, U. D., CHUNG, C., MADER, C., KOLETI, A., DATAR, N., VIDOVIĆ, D., WROBEL, D., ERICKSON, S., MUHLICH, J. L., BERRIZ, G., BENES, C. H., SUBRAMANIAN, A., PILLAI, A., SHAMU, C. E. and SCHÜRER, S. C. (2014). Metadata standard and data exchange specifications to describe, model, and integrate complex and diverse high-throughput screening data from the Library of Integrated Network-based Cellular Signatures (LINCS). *J. Biomol. Screen.* **19** 803–816.
- VERBIST, B., CLEMENT, L., REUMERS, J., THYS, K., VAPIREV, A., TALLOEN, W., WETZELS, Y., MEYS, J., AERSSSENS, J., BIJNENS, L. and THAS, O. (2015). ViVaMBC: Estimating viral sequence variation in complex populations from illumina deep-sequencing data using model-based clustering. *BMC Bioinform.* **16** Article ID 59. DOI:[10.1186/s12859-015-0458-7](https://doi.org/10.1186/s12859-015-0458-7).
- VIDOVIĆ, D., KOLETI, A. and SCHÜRER, S. C. (2013). Large-scale integration of small molecule-induced genome-wide transcriptional responses, Kinome-wide binding affinities and cell-growth inhibition profiles reveal global trends characterizing systems-level drug action *Front. Genetics* **5** 342–342.
- WANG, Z., CLARK, N. R. and MA'AYAN, A. (2016). Drug-induced adverse events prediction with the LINCS L1000 data. *Bioinformatics* **32** 2338–2345. DOI:[10.1093/bioinformatics/btw168](https://doi.org/10.1093/bioinformatics/btw168).
- WANG, Z., GERSTEIN, M. and SNYDER, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10** 57–63. DOI:[10.1038/nrg2484](https://doi.org/10.1038/nrg2484).
- WINGENDER, E., CHEN, X., HEHL, R., KARAS, H., LIEBICH, I., MATYS, V., MEINHARDT, T., PRÜSS, M., REUTER, I. and SCHACHERER, F. (2000). TRANSFAC: An integrated system for gene expression regulation. *Nucleic Acids Res.* **28** 316–319.
- WOLFE, J. H. (1970). Pattern clustering by multivariate mixture analysis. *Multivar. Behav. Res.* **5** 329–350.
- WU, C. F. J. (1983). On convergence properties of the EM algorithm. *Ann. Statist.* **11** 95–103. [MR0684867](https://doi.org/10.1214/aos/1176344867)
- YEUNG, K. Y., FRALEY, C., MURUA, A., RAFTERY, A. E. and RUZZO, W. L. (2001). Model-based clustering and data transformations for gene expression data. *Bioinformatics* **17** 977–987.
- YOUNG, W. C., RAFTERY, A. E. and YEUNG, K. Y. (2014). Fast Bayesian inference for gene regulatory networks using ScanBMA. *BMC Syst. Biol.* **8** Article ID 47. DOI:[10.1186/1752-0509-8-47](https://doi.org/10.1186/1752-0509-8-47).
- YOUNG, W. C., RAFTERY, A. E. and YEUNG, K. Y. (2016). A posterior probability approach for gene regulatory network inference in genetic perturbation data. *Math. Biosci. Eng.* **13** 1241–1251. DOI:[10.3934/mbe.2016041](https://doi.org/10.3934/mbe.2016041). [MR3547570](https://pubmed.ncbi.nlm.nih.gov/2616041/)
- ZELLNER, A. (1986). On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In *Bayesian Inference and Decision Techniques. Stud. Bayesian Econometrics Statist.* **6** 233–243. North-Holland, Amsterdam. [MR0881437](https://pubmed.ncbi.nlm.nih.gov/881437/)

ZOU, M. and CONZEN, S. D. (2005). A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. *Bioinformatics* **21** 71–79.

A UNIFIED FRAMEWORK FOR VARIANCE COMPONENT ESTIMATION WITH SUMMARY STATISTICS IN GENOME-WIDE ASSOCIATION STUDIES¹

BY XIANG ZHOU

University of Michigan

Linear mixed models (LMMs) are among the most commonly used tools for genetic association studies. However, the standard method for estimating variance components in LMMs—the restricted maximum likelihood estimation method (REML)—suffers from several important drawbacks: REML requires individual-level genotypes and phenotypes from all samples in the study, is computationally slow, and produces downward-biased estimates in case control studies. To remedy these drawbacks, we present an alternative framework for variance component estimation, which we refer to as MQS. MQS is based on the method of moments (MoM) and the minimal norm quadratic unbiased estimation (MINQUE) criterion, and brings two seemingly unrelated methods—the renowned Haseman–Elston (HE) regression and the recent LD score regression (LDSC)—into the same unified statistical framework. With this new framework, we provide an alternative but mathematically equivalent form of HE that allows for the use of summary statistics. We provide an exact estimation form of LDSC to yield unbiased and statistically more efficient estimates. A key feature of our method is its ability to pair marginal z -scores computed using all samples with SNP correlation information computed using a small random subset of individuals (or individuals from a proper reference panel), while capable of producing estimates that can be almost as accurate as if both quantities are computed using the full data. As a result, our method produces unbiased and statistically efficient estimates, and makes use of summary statistics, while it is computationally efficient for large data sets. Using simulations and applications to 37 phenotypes from 8 real data sets, we illustrate the benefits of our method for estimating and partitioning SNP heritability in population studies as well as for heritability estimation in family studies. Our method is implemented in the GEMMA software package, freely available at www.xzlab.org/software.html.

REFERENCES

- ABECASIS, G. R., CARDON, L. R. and COOKSON, W. O. (2000). A general test of association for quantitative traits in nuclear families. *Am. J. Hum. Genet.* **66** 279–292.
- ALLEN, H. L., ESTRADA, K., LETTRE, G. BERNDT, S. I., WEEDON, M. N., RIVADENEIRA, F., WILLER, C. J., JACKSON, A. U., VEDANTAM, S. et al. (2010). Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467** 832–838.

Key words and phrases. Genome-wide association studies, summary statistics, variance component, linear mixed model, MINQUE, method of moments.

- ALMASY, L. and BLANGERO, J. (1998). Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62** 1198–1211.
- AMOS, C. I. (1994). Robust variance-components approach for assessing genetic linkage in pedigrees. *Am. J. Hum. Genet.* **54** 535–543.
- BROWNING, S. R. (2006). Multilocus association mapping using variable-length Markov chains. *Am. J. Hum. Genet.* **78** 903–913.
- BROWNING, S. R. and BROWNING, B. L. (2013). Identity-by-descent-based heritability analysis in the Northern Finland Birth Cohort. *Hum. Genet.* **132** 129–138.
- BULIK-SULLIVAN, B. (2015). Relationship between LD score and Haseman–Elston regression. *BioRxiv* **0** 018283.
- BULIK-SULLIVAN, B. K., LOH, P.-R., FINUCANE, H. K., RIPKE, S., YANG, J., SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM, PATTERSON, N., DALY, M. J., PRICE, A. L. and NEALE, B. M. (2015a). LD score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47** 291–295.
- BULIK-SULLIVAN, B., FINUCANE, H. K., ANTTILA, V., GUSEV, A., DAY, F. R., LOH, P.-R., REPROGEN CONSORTIUM, PSYCHIATRIC GENOMICS CONSORTIUM, GENETIC CONSORTIUM FOR ANOREXIA NERVOSA OF THE WELLCOME TRUST CASE CONTROL CONSORTIUM 3 et al. (2015b). An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47** 1236–1241.
- CHEN, G.-B. (2014). Estimating heritability of complex traits from genome-wide association studies using IBS-based Haseman–Elston regression. *Front. Genet.* **5** 107.
- CHEN, W.-M., BROMAN, K. W. and LIANG, K.-Y. (2004). Quantitative trait linkage analysis by generalized estimating equations: Unification of variance components and Haseman–Elston regression. *Genet. Epidemiol.* **26** 265–272.
- CRAWFORD, L., ZENG, P., MUKHERJEE, S. and ZHOU, X. (2017). Detecting epistasis with the marginal epistasis test in genetic mapping studies of quantitative traits. *BioRxiv*.
- DE LOS CAMPOS, G., SORENSEN, D. and GIANOLA, D. (2015). Genomic heritability: What is it? *PLoS Genet.* **11** e1005048.
- DIAO, G. and LIN, D. Y. (2005). A powerful and robust method for mapping quantitative trait loci in general pedigrees. *Am. J. Hum. Genet.* **77** 97–111.
- DRIGALENKO, E. (1998). How sib-pairs reveal linkage. *Am. J. Hum. Genet.* **63** 1242–1245.
- ELSTON, R. C., BUXBAUM, S., JACOBS, K. B. and OLSON, J. M. (2000). Haseman and Elston revisited. *Genet. Epidemiol.* **19** 1–17.
- FINUCANE, H. K., BULIK-SULLIVAN, B., GUSEV, A., TRYNKA, G., RESHEF, Y., LOH, P.-R., ANTTILLA, V., XU, H., ZANG, C. et al. (2015). Partitioning heritability by functional category using GWAS summary statistics. *Nat. Genet.* **47** 1228–1235.
- FURLOTTE, N. A., HECKERMAN, D. and LIPPERT, C. (2014). Quantifying the uncertainty in heritability. *J. Hum. Genet.* **59** 269–275.
- GARCÍA-CORTÉS, L. A., MORENO, C., VARONA, L. and ALTARRIBA, J. (1992). Variance component estimation by resampling. *J. Anim. Breed. Genet.* **109** 358–363.
- GILMOUR, A. R., THOMPSON, R. and CULLIS, B. R. (1995). Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models. *Biometrics* **51** 1440–1450.
- GOLAN, D., LANDER, E. S. and ROSSETA, S. (2014). Measuring missing heritability: Inferring the contribution of common variants. *Proc. Natl. Acad. Sci. USA* **111** E5272–E5281.
- GUAN, Y. and STEPHENS, M. (2008). Practical issues in imputation-based association mapping. *PLoS Genet.* **4** e1000279.
- GUSEV, A., BHATIA, G., ZAITLEN, N., VILHJALMSSON, B. J., DIOGO, D., STAHL, E. A., GREGERSEN, P. K., WORTHINGTON, J., KLARESKOG, L. et al. (2013). Quantifying missing heritability at known GWAS loci. *PLoS Genet.* **9** e1003993.

- GUSEV, A., LEE, S. H., TRYNKA, G., FINUCANE, H., VILHJÁLMSSON, B. J., XU, H., ZANG, C., RIPKE, S., BULIK-SULLIVAN, B. et al. (2014). Partitioning heritability of regulatory and cell-type-specific variants across 11 common diseases. *Am. J. Hum. Genet.* **5** 535–552.
- HASEMAN, J. K. and ELSTON, R. C. (1972). The investigation of linkage between a quantitative trait and a marker locus. *Behav. Genet.* **2** 3–19.
- HAYES, B. J., VISSCHER, P. M. and GODDARD, M. E. (2009). Increased accuracy of artificial selection by using the realized relationship matrix. *Genet. Res. (Camb.)* **91** 47–60.
- HAYES, M. G., DEL BOSQUE-PLATA, L., TSUCHIYA, T., HANIS, C. L., BELL, G. I. and COX, N. J. (2005). Patterns of linkage disequilibrium in the type 2 diabetes gene calpain-10. *Diabetes* **54** 3573–3576.
- HOFER, A. (1998). Variance component estimation in animal breeding: A review. *J. Anim. Breed. Genet.* **115** 247–265.
- HOWIE, B., FUCHSBERGER, C., STEPHENS, M., MARCHINI, J. and ABECASIS, G. R. (2012). Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44** 955–959.
- JOSTINS, L., RIPKE, S., WEERSMA, R. K., DUERR, R. H., MCGOVERN, D. P., HUI, K. Y., LEE, J. C., SCHUMM, L. P., SHARMA, Y. et al. (2012). Host–microbe interactions have shaped the genetic architecture of inflammatory bowel disease. *Nature* **491** 119–124.
- KANG, H. M., ZAITLEN, N. A., WADE, C. M., KIRBY, A., HECKERMAN, D., DALY, M. J. and ESKIN, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics* **178** 1709–1723.
- KANG, H. M., SUL, J. H., SERVICE, S. K., ZAITLEN, N. A., KONG, S.-Y., FREIMER, N. B., SABATTI, C. and ESKIN, E. (2010). Variance component model to account for sample structure in genome-wide association studies. *Nat. Genet.* **42** 348–354.
- KOSTEM, E. and ESKIN, E. (2013). Improving the accuracy and efficiency of partitioning heritability into the contributions of genomic regions. *Am. J. Hum. Genet.* **92** 558–564.
- LEE, S. H., WRAY, N. R., GODDARD, M. E. and VISSCHER, P. M. (2011). Estimating missing heritability for disease from genome-wide association studies. *Am. J. Hum. Genet.* **88** 294–305.
- LIPPERT, C., LISTGARTEN, J., LIU, Y., KADIE, C. M., DAVIDSON, R. I. and HECKERMAN, D. (2011). FaST linear mixed models for genome-wide association studies. *Nat. Methods* **8** 833–835.
- LOH, P.-R., TUCKER, G., BULIK-SULLIVAN, B. K., VILHJÁLMSSON, B. J., FINUCANE, H. K., CHASMAN, D. I., RIDKER, P. M., NEALE, B. M., BERGER, B. et al. (2015a). Efficient Bayesian mixed model analysis increases association power in large cohorts. *Nat. Genet.* **47** 284–290.
- LOH, P.-R., BHATIA, G., GUSEV, A., FINUCANE, H. K., BULIK-SULLIVAN, B. K., POLLACK, S. J., SCHIZOPHRENIA WORKING GROUP OF THE PSYCHIATRIC GENOMICS CONSORTIUM, DE CANDIA, T. R., LEE, S. H. et al. (2015b). Contrasting genetic architectures of schizophrenia and other complex diseases using fast variance-components analysis. *Nat. Genet.* **47** 1385–1392.
- MAKOWSKY, R., PAJEWSKI, N. M., KLIMENTIDIS, Y. C., VAZQUEZ, A. I., DUARTE, C. W., ALLISON, D. B. and DE LOS CAMPOS, G. (2011). Beyond missing heritability: Prediction of complex traits. *PLoS Genet.* **7** e1002051.
- MANNING, A. K., HIVERT, M.-F., SCOTT, R. A., GRIMSBY, J. L., BOUATIA-NAJI, N., CHEN, H., RYBIN, D., LIU, C.-T., BIELAK, L. F. et al. (2012). A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat. Genet.* **44** 659–669.
- MATILAINEN, K., MÄNTYSAARI, E. A., LIDAUER, M. H., STRANDÉN, I. and THOMPSON, R. (2012). Employing a Monte Carlo algorithm in expectation maximization restricted maximum likelihood estimation of the linear mixed model. *J. Anim. Breed. Genet.* **129** 457–468.
- PICKRELL, J. K. (2014). Joint analysis of functional genomic data and genome-wide association studies of 18 human traits. *Am. J. Hum. Genet.* **94** 559–573.

- PIRINEN, M., DONNELLY, P. and SPENCER, C. C. A. (2013). Efficient computation with a linear mixed model on large-scale data sets with applications to genetic studies. *Ann. Appl. Stat.* **7** 369–390. [MR3086423](#)
- PRICE, A. L., WEALE, M. E., PATTERSON, N., MYERS, S. R., NEED, A. C., SHIANN, K. V., GE, D., ROTTER, J. I., TORRES, E. et al. (2008). Long-range LD can confound genome scans in admixed populations. *Am. J. Hum. Genet.* **1** 132–135.
- PRICE, A. L., HELGASON, A., THORLEIFSSON, G., MCCARROLL, S. A., KONG, A. and STEFANSSON, K. (2011). Single-tissue and cross-tissue heritability of gene expression via identity-by-descent in related or unrelated individuals. *PLoS Genet.* **7** e1001317.
- RAO, C. R. (1970). Estimation of heteroscedastic variances in linear models. *J. Amer. Statist. Assoc.* **65** 161–172. [MR0286221](#)
- RAO, C. R. (1971). Estimation of variance and covariance components—MINQUE theory. *J. Multivariate Anal.* **1** 257–275. [MR0301869](#)
- ROBBINS, H. and MONRO, S. (1951). A stochastic approximation method. *Ann. Math. Stat.* **22** 400–407. [MR0042668](#)
- ROBINSON, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statist. Sci.* **6** 15–51. [MR1108815](#)
- SABATTI, C., SERVICE, S. K., HARTIKAINEN, A.-L., POUTA, A., RIPATTI, S., BRODSKY, J., JONES, C. G., ZAITLEN, N. A., VARILO, T. et al. (2008). Genome-wide association analysis of metabolic traits in a birth cohort from a founder population. *Nat. Genet.* **41** 35–46.
- SHAM, P. C. and PURCELL, S. (2001). Equivalence between Haseman–Elston and variance-components linkage analyses for sib pairs. *Am. J. Hum. Genet.* **68** 1527–1532.
- SHAM, P. C., PURCELL, S., CHERNY, S. S. and ABECASIS, G. R. (2002). Powerful regression-based quantitative-trait linkage analysis of general pedigrees. *Am. J. Hum. Genet.* **71** 238–253.
- SPEED, D. and BALDING, D. J. (2015). Relatedness in the post-genomic era: Is it still useful? *Nat. Rev. Genet.* **16** 33–33.
- SPEED, D., HEMANI, G., JOHNSON, M. R. and BALDING, D. J. (2012). Improved heritability estimation from genome-wide SNPs. *Am. J. Hum. Genet.* **91** 1011–1021.
- SPELIOTES, E. K., WILLER, C. J., BERNDT, S. I., MONDA, K. L., THORLEIFSSON, G., JACKSON, A. U., ALLEN, H. L., LINDGREN, C. M., LUAN, J. et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42** 937–948.
- SPLANSKY, G. L., COREY, D., YANG, Q., ATWOOD, L. D., CUPPLES, L. A., BENJAMIN, E. J., D’AGOSTINO, R. B., FOX, C. S., LARSON, M. G. et al. (2007). The third generation cohort of the National Heart, Lung, and Blood Institute’s Framingham Heart Study: Design, recruitment, and initial examination. *Am. J. Epidemiol.* **165** 1328–1335.
- TESLOVICH, T. M., MUSUNURU, K., SMITH, A. V., EDMONDSON, A. C., STYLIANOU, I. M., KOSEKI, M., PIRRUCCELLO, J. P., RIPATTI, S., CHASMAN, D. I. et al. (2010). Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466** 707–713.
- THE 1000 GENOMES PROJECT CONSORTIUM (2012). An integrated map of genetic variation from 1092 human genomes. *Nature* **491** 56–65.
- THE WELLCOME TRUST CASE CONTROL CONSORTIUM (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3000 shared controls. *Nature* **447** 661–678.
- THOMPSON, E. A. and SHAW, R. G. (1990). Pedigree analysis for quantitative traits: Variance components without matrix inversion. *Biometrics* **46** 399–413.
- VISSCHER, P. M., HILL, W. G. and WRAY, N. R. (2008). Heritability in the genomics era—concepts and misconceptions. *Nat. Rev. Genet.* **9** 255–266.
- WEN, X. and STEPHENS, M. (2010). Using linear predictors to impute allele frequencies from summary or pooled genotype data. *Ann. Appl. Stat.* **4** 1158–1182. [MR2751337](#)
- WHITTAKER, J. C., THOMPSON, R. and DENHAM, M. (2000). Marker-assisted selection using ridge regression. *Genet. Res.* **75** 249–252.

- WRAY, N. R., YANG, J., HAYES, B. J., PRICE, A. L., GODDARD, M. E. and VISSCHER, P. M. (2013). Pitfalls of predicting complex traits from SNPs. *Nat. Rev. Genet.* **14** 507–515.
- WU, T. T., CHEN, Y. F., HASTIE, T., SOBEL, E. and LANGE, K. (2009). Genome-wide association analysis by lasso penalized logistic regression. *Bioinformatics* **25** 714–721.
- YANG, J., BENYAMIN, B., MCEVOY, B. P., GORDON, S., HENDERS, A. K., NYHOLT, D. R., MADDEN, P. A., HEATH, A. C., MARTIN, N. G. et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42** 565–569.
- YANG, J., MANOLIO, T. A., PASQUALE, L. R., BOERWINKLE, E., CAPORASO, N., CUNNINGHAM, J. M., DE ANDRADE, M., FEENSTRA, B., FEINGOLD, E. et al. (2011a). Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* **43** 519–525.
- YANG, J., WEEDON, M. N., PURCELL, S., LETTRE, G., ESTRADA, K., WILLER, C. J., SMITH, A. V., INGELSSON, E., O’CONNELL, J. R. et al. (2011b). Genomic inflation factors under polygenic inheritance. *Eur. J. Hum. Genet.* **19** 807–812.
- YANG, J., LEE, S. H., GODDARD, M. E. and VISSCHER, P. M. (2011c). GCTA: A tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88** 76–82.
- YANG, J., FERREIRA, T., MORRIS, A. P., MEDLAND, S. E., GENETIC INVESTIGATION OF ANTHROPOMETRIC TRAITS (GIANT) CONSORTIUM, DIABETES GENETICS REPLICATION AND META-ANALYSIS (DIAGRAM) CONSORTIUM, MADDEN, P. A. F., HEATH, A. C., MARTIN, N. G. et al. (2012). Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44** 369–375.
- YANG, J., ZAITLEN, N. A., GODDARD, M. E., VISSCHER, P. M. and PRICE, A. L. (2014). Advantages and pitfalls in the application of mixed-model association methods. *Nat. Genet.* **46** 100–106.
- YANG, J., BAKSHI, A., ZHU, Z., HEMANI, G., VINKHUYZEN, A. A. E., LEE, S. H., ROBINSON, M. R., PERRY, J. R. B., NOLTE, I. M. et al. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* **47** 1114–1120.
- YU, J., PRESSOIR, G., BRIGGS, W. H., BI, I. V., YAMASAKI, M., DOEBLEY, J. F., McMULLEN, M. D., GAUT, B. S., NIELSEN, D. M. et al. (2006). A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38** 203–208.
- ZAYKIN, D. V., MENG, Z. and EHM, M. G. (2006). Contrasting linkage-disequilibrium patterns between cases and controls as a novel association-mapping method. *Am. J. Hum. Genet.* **78** 737–746.
- ZHANG, Z., ERSOZ, E., LAI, C.-Q., TODHUNTER, R. J., TIWARI, H. K., GORE, M. A., BRADBURY, P. J., YU, J., ARNETT, D. K. et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* **42** 355–360.
- ZHOU, X. (2017). Supplement to “A unified framework for variance component estimation with summary statistics in genome-wide association studies.” DOI:10.1214/17-AOAS1052SUPP.
- ZHOU, X., CARBONETTO, P. and STEPHENS, M. (2013). Polygenic modelling with Bayesian sparse linear mixed models. *PLoS Genet.* **9** e1003264.
- ZHOU, X. and STEPHENS, M. (2012). Genome-wide efficient mixed-model analysis for association studies. *Nat. Genet.* **44** 821–824.
- ZHOU, X. and STEPHENS, M. (2014). Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat. Methods* **11** 407–409.
- ZHU, J. and WEIR, B. (1996). Mixed model approaches for diallele analysis based on a bio-model. *Genet. Res.* **68** 233–240.

PHOTODEGRADATION MODELING BASED ON LABORATORY ACCELERATED TEST DATA AND PREDICTIONS UNDER OUTDOOR WEATHERING FOR POLYMERIC MATERIALS

BY YUANYUAN DUAN^{*}, YILI HONG^{*,1}, WILLIAM Q. MEEKER[†],
 DEBORAH L. STANLEY[‡] AND XIAOHONG GU[‡]

Virginia Tech^{}, Iowa State University[†] and National Institute of Standards and Technology[‡]*

Photodegradation, driven primarily by ultraviolet (UV) radiation, is the primary cause of failure for organic paints and coatings, as well as many other products made from polymeric materials exposed to sunlight. Traditional methods of service life prediction involve the use of outdoor exposure in harsh UV environments (e.g., Florida and Arizona). Such tests, however, require too much time (generally many years) to do an evaluation. To overcome the shortcomings of traditional methods, scientists at the U.S. National Institute of Standards and Technology (NIST) conducted a multiyear research program to collect necessary data via scientifically-based laboratory accelerated tests. This paper presents the statistical modeling and analysis of the photodegradation data collected at NIST, and predictions of degradation for outdoor specimens that are subjected to weathering. The analysis involves identifying a physics/chemistry-motivated model that will adequately describe photodegradation paths. The model incorporates the effects of explanatory variables which are UV spectrum, UV intensity, temperature, and relative humidity. We use a nonlinear mixed-effects model to describe the sample paths. We extend the model to allow for dynamic covariates and compare predictions with specimens that were exposed in an outdoor environment where the explanatory variables are uncontrolled but recorded. We also discuss the findings from the analysis of the NIST data and some areas for future research.

REFERENCES

- BAGDONAVIČIUS, V. and NIKULIN, M. S. (2001). Estimation in degradation models with explanatory variables. *Lifetime Data Anal.* **7** 85–103. [MR1819926](#)
- BELLINGER, V. and VERDU, J. (1984). Structure-photooxidative stability relationship of amine-crosslinked epoxies. *Polym. Photochem.* **5** 295–311.
- BELLINGER, V. and VERDU, J. (1985). Oxidative skeleton breaking in epoxy-amine networks. *J. Appl. Polym. Sci.* **30** 363–374.
- CHEN, C. and FULLER, T. F. (2009). The effect of humidity on the degradation of Nafion membrane. *Polym. Degrad. Stab.* **94** 1436–1447.
- DAVIDIAN, M. and GILTINAN, D. M. (2003). Nonlinear models for repeated measurement data: An overview and update. *J. Agric. Biol. Environ. Stat.* **8** 387–419.

Key words and phrases. Degradation, nonlinear model, random effects, reliability, service life prediction, UV exposure.

- GU, X., DICKENS, B., STANLEY, D., BYRD, W., NGUYEN, T., VACA-TRIGO, I., MEEKER, W. Q., CHIN, J. and MARTIN, J. (2009). Linking accelerating laboratory test with outdoor performance results for a model epoxy coating system. In *Service Life Prediction of Polymeric Materials, Global Perspectives* (J. W. Martin, R. A. Ryntz, J. Chin and R. A. Dickie, eds.) 3–28. Springer, New York.
- HONG, Y. and MEEKER, W. Q. (2010). Field-failure and warranty prediction based on auxiliary use-rate information. *Technometrics* **52** 148–159. [MR2757212](#)
- HONG, Y., MEEKER, W. Q. and MCCALLEY, J. D. (2009). Prediction of remaining life of power transformers based on left truncated and right censored lifetime data. *Ann. Appl. Stat.* **3** 857–879. [MR2750685](#)
- HONG, Y., DUAN, Y., MEEKER, W. Q., STANLEY, D. L. and GU, X. (2015). Statistical methods for degradation data with dynamic covariates information and an application to outdoor weathering data. *Technometrics* **57** 180–193. [MR3369675](#)
- JAMES, T. (1997). *The Theory of the Photographic Process*, 4th ed. Macmillan, New York.
- KELLEHER, P. and GESNER, B. (1969). Photo-oxidation of phenoxy resin. *J. Appl. Polym. Sci.* **13** 9–15.
- KIIL, S. (2009). Model-based analysis of photoinitiated coating degradation under artificial exposure conditions. *J. Coat. Technol. Res.* **9** 375–398.
- LAWLESS, J. F. and FREDETTE, M. (2005). Frequentist prediction intervals and predictive distributions. *Biometrika* **92** 529–542. [MR2202644](#)
- LIAO, H. and ELSAYED, E. A. (2006). Reliability inference for field conditions from accelerated degradation testing. *Naval Res. Logist.* **53** 576–587. [MR2252304](#)
- LU, C. J. and MEEKER, W. Q. (1993). Using degradation measures to estimate a time-to-failure distribution. *Technometrics* **35** 161–174. [MR1225093](#)
- MARTIN, J. W., CHIN, J. W. and NGUYEN, T. (2003). Reciprocity law experiments in polymeric photodegradation: A critical review. *Prog. Org. Coat.* **47** 292–311.
- MARTIN, J. W., LECHNER, J. A. and VARNER, R. N. (1994). Quantitative characterization of photodegradation effects of polymeric materials exposed in weathering environments. In *Accelerated and Outdoor Durability Testing of Organic Materials, ASTM-STP-1202* (W. D. Ketola and D. Grossman, eds.) American Society for Testing and Materials, Philadelphia, PA.
- MARTIN, J. W., SAUNDERS, S. C., FLOYD, F. L. and WINEBURG, J. P. (1996). Methodologies for predicting the service lives of coating systems. In *Federation Series on Coating Technology* (D. Brezinski and T. Miranda, eds.) 1–32. Federation Series on Coating Technology, Blue Hill, PA.
- MEEKER, W. Q., ESCOBAR, L. A. and LU, C. J. (1998). Accelerated degradation tests: Modeling and analysis. *Technometrics* **40** 89–99.
- NELSON, W. (1990). *Accelerated Testing: Statistical Models, Test Plans, and Data Analyses*. John Wiley & Sons, New York.
- PAN, R. and CRISPIN, T. (2011). A hierarchical modeling approach to accelerated degradation testing data analysis: A case study. *Qual. Reliab. Eng. Int.* **27** 229–237.
- PENG, C.-Y. (2015). Inverse Gaussian processes with random effects and explanatory variables for degradation data. *Technometrics* **57** 100–111. [MR3318353](#)
- PINHEIRO, J. and BATES, D. (2006). *Mixed-Effects Models in S and S-PLUS*. Springer Science & Business Media, New York, NY.
- RABEK, J. (1995). *Polymer Photodegradation—Mechanisms and Experimental Methods*. Chapman & Hall, London.
- VACA-TRIGO, I. and MEEKER, W. Q. (2009). A statistical model for linking field and laboratory exposure results for a model coating. In *Service Life Prediction of Polymeric Materials* (J. Martin, R. A. Ryntz, J. Chin and R. A. Dickie, eds.) 29–43. Springer, New York, NY.
- WANG, X. and XU, D. (2010). An inverse Gaussian process model for degradation data. *Technometrics* **52** 188–197. [MR2676425](#)

- WANG, L., PAN, R., LI, X. and JIANG, T. (2013). A Bayesian reliability evaluation method with integrated accelerated degradation testing and field information. *Reliab. Eng. Syst. Saf.* **112** 38–47.
- XU, Z., HONG, Y. and JIN, R. (2016). Nonlinear general path models for degradation data with dynamic covariates. *Appl. Stoch. Models Bus. Ind.* **32** 153–167. [MR3488012](#)
- YE, Z.-S. and CHEN, N. (2014). The inverse Gaussian process as a degradation model. *Technometrics* **56** 302–311. [MR3238068](#)
- ZHANG, Y. and LIAO, H. (2015). Analysis of destructive degradation tests for a product with random degradation initiation time. *IEEE Trans. Reliab.* **64** 516–527.
- ZHOU, R., GEBRAEEL, N. and SERBAN, N. (2012). Degradation modeling and monitoring of truncated degradation signals. *IEEE Trans.* **44** 793–803.
- ZHOU, R. R., SERBAN, N. and GEBRAEEL, N. (2011). Degradation modeling applied to residual lifetime prediction using functional data analysis. *Ann. Appl. Stat.* **5** 1586–1610. [MR2849787](#)
- ZHOU, R., SERBAN, N. and GEBRAEEL, N. (2014). Degradation-based residual life prediction under different environments. *Ann. Appl. Stat.* **8** 1671–1689. [MR3271348](#)
- ZHOU, R. R., SERBAN, N., GEBRAEEL, N. and MÜLLER, H.-G. (2014). A functional time warping approach to modeling and monitoring truncated degradation signals. *Technometrics* **56** 67–77. [MR3176573](#)

VARIABLE SELECTION FOR LATENT CLASS ANALYSIS WITH APPLICATION TO LOW BACK PAIN DIAGNOSIS

BY MICHAEL FOP^{*,1}, KEITH M. SMART[†] AND
 THOMAS BRENDAN MURPHY^{*,1}

University College Dublin^{} and St. Vincent's University Hospital[†]*

The identification of most relevant clinical criteria related to low back pain disorders may aid the evaluation of the nature of pain suffered in a way that usefully informs patient assessment and treatment. Data concerning low back pain can be of categorical nature, in the form of a check-list in which each item denotes presence or absence of a clinical condition. Latent class analysis is a model-based clustering method for multivariate categorical responses, which can be applied to such data for a preliminary diagnosis of the type of pain. In this work, we propose a variable selection method for latent class analysis applied to the selection of the most useful variables in detecting the group structure in the data. The method is based on the comparison of two different models and allows the discarding of those variables with no group information and those variables carrying the same information as the already selected ones. We consider a swap-stepwise algorithm where at each step the models are compared through an approximation to their Bayes factor. The method is applied to the selection of the clinical criteria most useful for the clustering of patients in different classes. It is shown to perform a parsimonious variable selection and to give a clustering performance comparable to the expert-based classification of patients into three classes of pain.

REFERENCES

- AGRESTI, A. (2002). *Categorical Data Analysis*, 2nd ed. Wiley, New York. [MR1914507](#)
- AGRESTI, A. (2015). *Foundations of Linear and Generalized Linear Models*. Wiley, Hoboken, NJ. [MR3308143](#)
- ALBERT, A. and ANDERSON, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika* **71** 1–10. [MR0738319](#)
- ARIMA, S. (2015). Item selection via Bayesian IRT models. *Stat. Med.* **34** 487–503. [MR3301590](#)
- BADSBERG, J. H. (1992). Model search in contingency tables by CoCo. In *Computational Statistics, Vol. 1* (Y. Dodge and J. Whittaker, eds.) 251–256. Physica-Verlag, Heidelberg.
- BARTHOLOMEW, D., KNOTT, M. and MOUSTAKI, I. (2011). *Latent Variable Models and Factor Analysis: A Unified Approach*, 3rd ed. Wiley, Chichester. [MR2849614](#)
- BARTOLUCCI, F., MONTANARI, G. E. and PANDOLFI, S. (2016). Item selection by latent class-based methods: An application to nursing home evaluation. *Adv. Data Anal. Classif.* **10** 245–262. [MR3505059](#)
- BONTEMPS, D. and TOUSSILE, W. (2013). Clustering and variable selection for categorical multivariate data. *Electron. J. Stat.* **7** 2344–2371. [MR3108816](#)

Key words and phrases. Clinical criteria selection, clustering, latent class analysis, low back pain, mixture models, model-based clustering, variable selection.

- CELEUX, G., MARTIN-MAGNIETTE, M.-L., MAUGIS-RABUSSEAU, C. and RAFTERY, A. E. (2014). Comparing model selection and regularization approaches to variable selection in model-based clustering. *J. SFdS* **155** 57–71. [MR3211754](#)
- CLOGG, C. C. (1988). Latent class models for measuring. In *Latent Trait and Latent Class Models* (R. Langeheine and J. Rost, eds.) 173–205. Plenum, New York.
- CLOGG, C. C. (1995). Latent class models. In *Handbook of Statistical Modeling for the Social and Behavioral Sciences* (G. Arminger, C. C. Clogg and M. E. Sobel, eds.) 311–360. Plenum, New York.
- DEAN, N. and RAFTERY, A. E. (2010). Latent class analysis variable selection. *Ann. Inst. Statist. Math.* **62** 11–35. [MR2577437](#)
- DY, J. G. and BRODLEY, C. E. (2003/04). Feature selection for unsupervised learning. *J. Mach. Learn. Res.* **5** 845–889. [MR2248002](#)
- FOP, M., SMART, K. M. and MURPHY, T. B. (2017). Supplement to “Variable selection for latent class analysis with application to low back pain diagnosis.” DOI:[10.1214/17-AOAS1061SUPP](#).
- FRALEY, C. and RAFTERY, A. E. (2002). Model-based clustering, discriminant analysis, and density estimation. *J. Amer. Statist. Assoc.* **97** 611–631. [MR1951635](#)
- FROUD, R., PATTERSON, S., ELDRIDGE, S., SEALE, C., PINCUS, T., RAJENDRAN, D., FOS-SUM, C. and UNDERWOOD, M. (2014). A systematic review and meta-synthesis of the impact of low back pain on people’s lives. *BMC Musculoskeletal Disorders* **15** 1–14.
- GELMAN, A., JAKULIN, A., PITTAU, M. G. and SU, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2** 1360–1383. [MR2655663](#)
- GOLDBERG, D. (1989). *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, Reading.
- GOLLINI, I. and MURPHY, T. B. (2014). Mixture of latent trait analyzers for model-based clustering of categorical data. *Stat. Comput.* **24** 569–588. [MR3223542](#)
- GOODMAN, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61** 215–231. [MR0370936](#)
- GRAVEN-NIELSEN, T. and ARENDT-NIELSEN, L. (2010). Assessment of mechanisms in localized and widespread musculoskeletal pain. *Nat Rev Rheumatol* **6** 599–606.
- HABERMAN, S. J. (1979). *Analysis of Qualitative Data, Vol. 2: New Developments*. Academic Press, New York.
- HEINZE, G. and SCHEMPER, M. (2002). A solution to the problem of separation in logistic regression. *Stat. Med.* **21** 2409–2419.
- HOY, D., BAIN, C., WILLIAMS, G., MARCH, L., BROOKS, P., BLYTH, F., WOOLF, A., VOS, T. and BUCHBINDER, R. (2012). A systematic review of the global prevalence of low back pain. *Arthritis & Rheumatism* **64** 2028–2037.
- HUBERT, L. and ARABIE, P. (1985). Comparing partitions. *J. Classification* **2** 193–218.
- KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795. [MR3363402](#)
- KATZ, J. N., STOCK, S. R., EVANOFF, B. A., REMPEL, D., MOORE, J. S., FRANZBLAU, A. and GRAY, R. H. (2000). Classification criteria and severity assessment in work-associated upper extremity disorders: Methods matter. *Am. J. Ind. Med.* **38** 369–372.
- KIM, S., TADESSE, M. G. and VANNUCCI, M. (2006). Variable selection in clustering via Dirichlet process mixture models. *Biometrika* **93** 877–893. [MR2285077](#)
- LAW, M. H. C., FIGUEIREDO, M. A. T. and JAIN, A. K. (2004). Simultaneous feature selection and clustering using mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* **26** 1154–1166.
- LAZARSFELD, P. F. and HENRY, N. W. (1968). *Latent Structure Analysis*. Houghton Mifflin, Boston, MA.
- LESAFFRE, E. and ALBERT, A. (1989). Partial separation in logistic discrimination. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **51** 109–116. [MR0984997](#)

- LINZER, D. A. and LEWIS, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *J. Stat. Softw.* **42** 1–29.
- LIU, H. H. and ONG, C. S. (2008). Variable selection in clustering for marketing segmentation using genetic algorithms. *Expert Systems with Applications* **34** 502–510.
- MALSINER-WALLI, G., FRÜHWIRTH-SCHNATTER, S. and GRÜN, B. (2016). Model-based clustering based on sparse finite Gaussian mixtures. *Stat. Comput.* **26** 303–324. [MR3439375](#)
- MARBAC, M., BIERNACKI, C. and VANDEWALLE, V. (2015). Model-based clustering for conditionally correlated categorical data. *J. Classification* **32** 145–175. [MR3369412](#)
- MARBAC, M. and SEDKI, M. (2017). Variable selection for model-based clustering using the integrated complete-data likelihood. *Stat. Comput.* **27** 1049–1063. [MR3627562](#)
- MAUGIS, C., CELEUX, G. and MARTIN-MAGNIETTE, M.-L. (2009a). Variable selection for clustering with Gaussian mixture models. *Biometrics* **65** 701–709. [MR2649842](#)
- MAUGIS, C., CELEUX, G. and MARTIN-MAGNIETTE, M.-L. (2009b). Variable selection in model-based clustering: A general variable role modeling. *Comput. Statist. Data Anal.* **53** 3872–3882. [MR2749931](#)
- MCLACHLAN, G. J. and KRISHNAN, T. (2008). *The EM Algorithm and Extensions*, 2nd ed. Wiley, Hoboken, NJ. [MR2392878](#)
- MCLACHLAN, G. and PEEL, D. (2000). *Finite Mixture Models*. Wiley, New York. [MR1789474](#)
- MCLACHLAN, G. J. and RATHNAYAKE, S. (2014). On the number of components in a Gaussian mixture model. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* **4** 341–355.
- MCNICHOLAS, P. D. (2016). Model-based clustering. *J. Classification* **33** 331–373. [MR3575621](#)
- MERSKEY, H. and BOGDUK, N., eds. (2002). *Classification of Chronic Pain*. IASP Press, Washington, DC.
- MEYNET, C. and MAUGIS-RABUSSEAU, C. (2012). A sparse variable selection procedure in model-based clustering. *INRIA Report*.
- MILLER, A. (2002). *Subset Selection in Regression*, 2nd ed. *Monographs on Statistics and Applied Probability* **95**. Chapman & Hall/CRC, Boca Raton, FL. [MR2001193](#)
- MILLIGAN, G. W. and COOPER, M. C. (1986). A study of the comparability of external criteria for hierarchical cluster analysis. *Multivariate Behavioral Research* **21** 441–458.
- MURPHY, T. B., DEAN, N. and RAFTERY, A. E. (2010). Variable selection and updating in model-based discriminant analysis for high dimensional data with food authenticity applications. *Ann. Appl. Stat.* **4** 396–421. [MR2758177](#)
- NIJS, J., APELDOORN, A., HALLEGRAEFF, H., CLARK, J., SMEETS, R., MALFLIET, A., GIBBES, E. L., KOONING, M. D. and ICKMANS, K. (2015) Low back pain: Guidelines for the clinical classification of predominant neuropathic, nociceptive, or central sensitization pain. *Pain Physician* **18** E333–E346.
- PAN, W. and SHEN, X. (2007). Penalized model-based clustering with application to variable selection. *J. Mach. Learn. Res.* **8** 1145–1164.
- R CORE TEAM (2017). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- RAFTERY, A. E. and DEAN, N. (2006). Variable selection for model-based clustering. *J. Amer. Statist. Assoc.* **101** 168–178.
- RIPLEY, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge Univ. Press, Cambridge. [MR1438788](#)
- SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6** 461–464. [MR0468014](#)
- SCRUCCA, L. (2016). Genetic algorithms for subset selection in model-based clustering. In *Unsupervised Learning Algorithms* (M. E. Celebi and K. Aydin, eds.) 55–70. Springer, Berlin.
- SCRUCCA, L. and RAFTERY, A. E. (2017). Clustvarsel: A package implementing variable selection for model-based clustering in R. *J. Stat. Softw.* To appear.

- SILVESTRE, C., CARDOSO, M. G. M. S. and FIGUEIREDO, M. (2015). Feature selection for clustering categorical data with an embedded modelling approach. *Expert Systems* **32** 444–453.
- SMART, K. M., O’CONNELL, N. E. and DOODY, C. (2008). Towards a mechanisms-based classification of pain in musculoskeletal physiotherapy? *Physical Therapy Reviews* **13** 1–10.
- SMART, K. M., BLAKE, C., STAINES, A. and DOODY, C. (2010). Clinical indicators of “nociceptive”, “peripheral neuropathic” and “central” mechanisms of musculoskeletal pain. A Delphi survey of expert clinicians. *Manual Therapy* **15** 80–87.
- SMART, K. M., BLAKE, C., STAINES, A. and DOODY, C. (2011). The discriminative validity of “nociceptive”, “peripheral neuropathic”, and “central sensitization” as mechanisms-based classifications of musculoskeletal pain. *The Clinical Journal of Pain* **27** 655–663.
- STYNES, S., KONSTANTINOU, K. and DUNN, K. M. (2016). Classification of patients with low back-related leg pain: A systematic review. *BMC Musculoskelet Disord* **17** 226.
- TADESSE, M. G., SHA, N. and VANNUCCI, M. (2005). Bayesian variable selection in clustering high-dimensional data. *J. Amer. Statist. Assoc.* **100** 602–617. [MR2160563](#)
- WALKER, B. F. (2000). The prevalence of low back pain: A systematic review of the literature from 1966 to 1998. *Journal of Spinal Disorders & Techniques* **13** 205–217.
- WANG, S. and ZHU, J. (2008). Variable selection for model-based high-dimensional clustering and its application to microarray data. *Biometrics* **64** 440–448, 666. [MR2432414](#)
- WHITE, A., WYSE, J. and MURPHY, T. B. (2016). Bayesian variable selection for latent class analysis using a collapsed Gibbs sampler. *Stat. Comput.* **26** 511–527. [MR3439388](#)
- WITTEN, D. M. and TIBSHIRANI, R. (2010). A framework for feature selection in clustering. *J. Amer. Statist. Assoc.* **105** 713–726. [MR2724855](#)
- WOOLF, C. J. (2004). Pain: Moving from symptom control toward mechanism-specific pharmacologic management. *Ann. Intern. Med.* **140** 441–451.
- WOOLF, C. J., BENNETT, G. J., DOHERTY, M., DUBNER, R., KIDD, B., KOLTZENBURG, M., LIPTON, R., LOESER, J. D., PAYNE, R. and TOREBJORK, E. (1998). Towards a mechanism-based classification of pain? *Pain* **77** 227–229.
- XIE, B., PAN, W. and SHEN, X. (2008). Penalized model-based clustering with cluster-specific diagonal covariance matrices and grouped variables. *Electron. J. Stat.* **2** 168–212. [MR2386092](#)
- ZHANG, Q. and IP, E. H. (2014). Variable assessment in latent class models. *Comput. Statist. Data Anal.* **77** 146–156. [MR3210054](#)
- ZORN, C. (2005). A solution to separation in binary response models. *Polit. Anal.* **13** 157–170.

INFERENCE FOR RESPONDENT-DRIVEN SAMPLING WITH MISCLASSIFICATION¹

BY ISABELLE S. BEAUDRY*, KRISTA J. GILE[†] AND SHRUTI H. MEHTA[‡]

*Pontificia Universidad Católica de Chile**, *University of Massachusetts Amherst[†]*
and Johns Hopkins University Bloomberg School of Public Health[‡]

Respondent-driven sampling (RDS) is a sampling method designed to study hard-to-reach human populations. Beginning with a convenience sample, each participant receives a small number of coupons, which they distribute to their contacts who become eligible. RDS participants are asked to report on their number of contacts in the target population. Also, a set of characteristics is observed for each participant. Current prevalence estimators assume that these attributes are measured accurately. However, ignoring misclassification may lead to biased estimates.

The main contribution of this paper is to discuss two approaches to correct for the bias introduced by the misclassification on nodal attributes for existing RDS estimators. The two approaches leverage misclassification rates assumed to be available from external validation studies. Most importantly, our analysis identifies circumstances for which the performance of the correction methods is impaired in the specific context of RDS. The two methods that are discussed are an analytical correction for estimators of the Hájek estimator style and the Simulation Extrapolation Misclassification (SIMEX MC) approach. Extended methodology to estimate the uncertainty of the corrected estimators is also presented. The performance of the proposed methods is assessed under varying levels of known or uncertain misclassification error across simulated social networks of varying features. Finally, the methods are used to estimate HIV prevalence among people who inject drugs (PWID) and men who have sex with men (MSM) in India.

REFERENCES

- BARRON, B. A. (1977). The effects of misclassification on the estimation of relative risk. *Biometrics* **33** 414–418.
- BEAUDRY, I. S, GILE, K. J and MEHTA, S. H (2017). Supplement to “Inference for respondent-driven sampling with misclassification.” DOI:[10.1214/17-AOAS1063SUPP](https://doi.org/10.1214/17-AOAS1063SUPP).
- BIERNACKI, P. and WALDORF, D. (1981). Snowball sampling: Problem and techniques of chain referral sampling. *Sociol. Methods Res.* **10** 141–163.
- BUONACCORSI, J. P. (2010). *Measurement Error: Models, Methods, and Applications*. CRC Press, Boca Raton, FL. [MR2682774](#)
- COOK, J. R. and STEFANSKI, L. A. (1994). Simulation extrapolation estimation in parametric measurement error models. *J. Amer. Statist. Assoc.* **89** 1314–1328.
- FRANK, O. and STRAUSS, D. (1986). Markov graphs. *J. Amer. Statist. Assoc.* **81** 832–842. [MR0860518](#)

Key words and phrases. Hard-to-reach population sampling, misclassification, SIMEX MC, network sampling, social networks.

- FROST, S. D. W., BROUWER, K. C., CRUZ, M. A. F., RAMOS, R., RAMOS, M. E., LOZADA, R. M., MAGIS-RODRIGUEZ, C. and STRATHDEE, S. A. (2006). Respondent-driven sampling of injection drug users in two U.S.–Mexico border cities: Recruitment dynamics and impact on estimates of HIV and syphilis prevalence. *J. Urban Health* **83** 83–97.
- GILE, K. J. (2011). Improved inference for respondent-driven sampling data with application to HIV prevalence estimation. *J. Amer. Statist. Assoc.* **106** 135–146. [MR2816708](#)
- GILE, K. J. and HANDCOCK, M. S. (2010). Respondent-driven sampling: An assessment of current methodology. *Sociol. Method.* **40** 285–327.
- GILE, K. J., JOHNSTON, L. G. and SALGANIK, M. J. (2015). Diagnostics for respondent-driven sampling. *J. Roy. Statist. Soc. Ser. A* **178** 241–269. [MR3291770](#)
- GOODMAN, L. A. (1961). Snowball sampling. *Ann. Math. Stat.* **32** 148–170. [MR0124140](#)
- HANDCOCK, M. S., FELLOWS, I. E. and GILE, K. J. (2015). RDS: respondent-driven sampling. R package version 0.7-2, Los Angeles, CA.
- HANDCOCK, M. S. and GILE, K. J. (2011). Comment: On the concept of snowball sampling. *Sociol. Method.* **41** 367–371.
- HANDCOCK, M. S., HUNTER, D. R., BUTTS, C. T., GOODREAU, S. M., KRIVITSKY, P. N., BENDER-DEMOLL, S. and MORRIS, M. (2015). statnet: Software tools for the statistical analysis of network data. R package version 2015.6.2, The Statnet Project (<http://www.statnet.org>).
- HECKATHORN, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Soc. Probl.* **44** 174–199.
- HUNTER, D. R., GOODREAU, S. M. and HANDCOCK, M. S. (2008). Goodness of fit of social network models. *J. Amer. Statist. Assoc.* **103** 248–258. [MR2394635](#)
- HUNTER, D. R. and HANDCOCK, M. S. (2006). Inference in curved exponential family models for networks. *J. Comput. Graph. Statist.* **15** 565–583. [MR2291264](#)
- JOHNSTON, L. G., MALEKINEJAD, M., KENDALL, C., IUPPA, I. M. and RUTHERFORD, G. W. (2008). Implementation challenges to using respondent-driven sampling methodology for HIV biological and behavioral surveillance: Field experiences in international settings. *AIDS Behav.* **12** 131–141.
- KÜCHENHOFF, H., LEDERER, W. and LESAFFRE, E. (2007). Asymptotic variance estimation for the misclassification SIMEX. *Comput. Statist. Data Anal.* **51** 6197–6211. [MR2407708](#)
- KÜCHENHOFF, H., MWALILI, S. M. and LESAFFRE, E. (2006). A general method for dealing with misclassification in regression: The misclassification SIMEX. *Biometrics* **62** 85–96. [MR2226560](#)
- LIU, H., LI, J., HA, T. and LI, J. (2012). Assessment of random recruitment assumption in respondent-driven sampling in egocentric network data. *Soc. Netw.* **1** 13–21.
- LU, X. (2013). Linked ego networks: Improving estimate reliability and validity with respondent-driven sampling. *Soc. Netw.* **35** 669–685.
- LU, X., BENGTSOON, L., BRITTON, T., CAMITZ, M., KIM, B. J., THORSON, A. and LILJEROS, F. (2012). The sensitivity of respondent-driven sampling. *J. Roy. Statist. Soc. Ser. A* **175** 191–216. [MR2873802](#)
- LU, X., MALMROS, J., LILJEROS, F. and BRITTON, T. (2013). Respondent-driven sampling on directed networks. *Electron. J. Stat.* **7** 292–322. [MR3020422](#)
- LUCAS, G. M., SOLOMON, S. S., SRIKRISHNAN, A. K., AGRAWAL, A., IQBAL, S., LAEYEN-DECKER, O., MCFALL, A. M., KUMAR, M. S., OGBURN, E. L., CELENTANO, D. D., SOLOMON, S. and MEHTA, S. H. (2015). High HIV burden among people who inject drugs in 15 Indian cities. *AIDS* **29** 619–628.
- MALEKINEJAD, M., JOHNSTON, L., KENDALL, C., KERR, L., RIFKIN, M. and RUTHERFORD, G. (2008). Using respondent-driven sampling methodology for HIV biological and behavioral surveillance in international settings: A systematic review. *AIDS Behav.* **12** 105–130.
- MARKS, G., CREPAZ, N., SENTERFITT, J. W. and JANSSEN, R. S. (2005). Meta-analysis of high-risk sexual behavior in persons aware and unaware they are infected with HIV in the United States: Implications for HIV prevention programs. *J. Acquir. Immune Defic. Syndr.* **39** 446–453.

- MCCREESH, N., FROST, S. D. W., SEELEY, J., KATONGOLE, J., TARSH, M. N., NDUNGUSE, R., JICHI, F., LUNEL, N. L., MAHER, D., JOHNSTON, L. G., SONNENBERG, P., COPAS, A. J., HAYES, R. J. and WHITE, R. G. (2012). Evaluation of respondent-driven sampling. *Epidemiology* **23** 138–147.
- MILLS, H. L., JOHNSON, S., HICKMAN, M., JONES, N. S. and COLIJN, C. (2014). Errors in reported degrees and respondent driven sampling: Implications for bias. *Drug Alcohol Depend.* **142** 120–126.
- MONTEALEGRE, J. R., JOHNSTON, L. G., MURRILL, C. and MONTERROSO, E. (2013). Respondent driven sampling for HIV biological and behavioral surveillance in Latin America and the Caribbean. *AIDS Behav.* **17** 2313–2340.
- WORLD HEALTH ORGANIZATION (2015). Consolidated guidelines on HIV testing services 2015. Technical report, World Health Organization, Geneva.
- RUDOLPH, A., FULLER, C. and LATKIN, C. (2013). The importance of measuring and accounting for potential biases in respondent-driven samples. *AIDS Behav.* **17** 2244–2252.
- SALGANIK, M. J. (2006). Variance estimation, design effects, and sample size calculations for respondent-driven sampling. *J. Urban Health* **83** i98–i112.
- SALGANIK, M. J. and HECKATHORN, D. D. (2004). Sampling and estimation in hidden populations using respondent-drive sampling. *Sociol. Method.* **34** 193–239.
- SHANKS, L., KLARKOWSKI, D. and O'BRIEN, D. P. (2013). False positive HIV diagnoses in resource limited settings: Operational lessons learned for HIV programmes. *PLoS ONE* **8** 8–13.
- SMITH, R., ROSSETTO, K. and PETERSON, B. (2008). A meta-analysis of disclosure of one's HIV-positive status, stigma and social support. *AIDS Care* **20** 1266–1275.
- SOLOMON, S. S., MEHTA, S. H., SRIKRISHNAN, A. K., VASUDEVAN, C. K., MCFALL, A. M., BALAKRISHNAN, P., ANAND, S., NANDAGOPAL, P., OGBURN, E. L., LAEYENDECKER, O., LUCAS, G. M., SOLOMON, S. and CELENTANO, D. D. (2015). High HIV prevalence and incidence among MSM across 12 cities in India. *AIDS* **29** 723–731.
- TOMAS, A. and GILE, K. J. (2011). The effect of differential recruitment, non-response and non-recruitment on estimators for respondent-driven sampling. *Electron. J. Stat.* **5** 899–934. [MR2831520](#)
- TROW, M. (1957). *Right-Wing Radicalism and Political Intolerance*. Arno Press, New York. Reprinted 1980.
- UNAIDS (2014). The gap report.
- VERDERY, A. M., MERLI, M. G., MOODY, J., SMITH, J. A. and FISHER, J. C. (2015). Respondent-driven sampling estimators under real and theoretical recruitment conditions of female sex workers in China. *Epidemiology* **26** 661–665.
- VOLZ, E. and HECKATHORN, D. D. (2008). Probability based estimation theory for respondent driven sampling. *J. Off. Stat.* **24** 79–97.
- WEJNERT, C. and HECKATHORN, D. D. (2008). Web-based network sampling: Efficiency and efficacy of respondent-driven sampling for online research. *Sociol. Methods Res.* **37** 105–134. [MR2516739](#)
- YAMANIS, T. J., MERLI, M. G., NEELY, W. W., TIAN, F. F., MOODY, J., TU, X. and GAO, E. (2013). An empirical analysis of the impact of recruitment patterns on RDS estimates among a socially ordered population of female sex workers in China. *Sociol. Methods Res.* **42** 392–425. [MR3190735](#)
- YATES, F. and GRUNDY, P. M. (1953). Selection without replacement from within strata with probability proportional to size. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **15** 253–261.

LEARNING POPULATION AND SUBJECT-SPECIFIC BRAIN CONNECTIVITY NETWORKS VIA MIXED NEIGHBORHOOD SELECTION

BY RICARDO PIO MONTI^{*}, CHRISTOFOROS ANAGNOSTOPOULOS^{*} AND GIOVANNI MONTANA^{*,†}

Imperial College London^{} and King's College London[†]*

In neuroimaging data analysis, Gaussian graphical models are often used to model statistical dependencies across spatially remote brain regions known as functional connectivity. Typically, data is collected across a cohort of subjects and the scientific objectives consist of estimating population and subject-specific connectivity networks. A third objective that is often overlooked involves quantifying inter-subject variability, and thus identifying regions or subnetworks that demonstrate heterogeneity across subjects. Such information is crucial to thoroughly understand the human connectome. We propose *Mixed Neighborhood Selection* to simultaneously address the three aforementioned objectives. By recasting covariance selection as a neighborhood selection problem, we are able to efficiently learn the topology of each node. We introduce an additional mixed effect component to neighborhood selection to simultaneously estimate a graphical model for the population of subjects as well as for each individual subject. The proposed method is validated empirically through a series of simulations and applied to resting state data for healthy subjects taken from the ABIDE consortium.

REFERENCES

- BARABÁSI, A.-L. and ALBERT, R. (1999). Emergence of scaling in random networks. *Science* **286** 509–512. [MR2091634](#)
- BELILOVSKY, E., VAROQUAUX, G. and BLASCHKO, M. (2016). Testing for differences in Gaussian graphical models: Applications to brain connectivity. In *Neural Information Processing Systems* 595–603.
- BULLMORE, E. and SPORNS, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev., Neurosci.* **10** 186–198.
- CHUNG, M., HANSON, J., YE, J., DAVIDSON, R. and POLLAK, S. (2015). Persistent homology in sparse regression and its application to brain morphometry. *IEEE Trans. Med. Imag.* **34** 1928–1939.
- DAMOISEAUX, J., ROMBOUTS, S., BARKHOF, F., SCHELTENS, P., STAM, C., SMITH, S. and BECKMANN, C. (2006). Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci. USA* **103** 13848–13853.
- DANAHER, P., WANG, P. and WITTEN, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 373–397. [MR3164871](#)

Key words and phrases. Functional connectivity, neuroimaging, graphical models, inter-subject variability.

- DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **39** 1–38. [MR0501537](#)
- DI MARTINO, A., YAN, C., LI, Q., DENIO, E., CASTELLANOS, F., ALAERTS, K., ANDERSON, J., ASSAF, M., BOOKHEIMER, S. and DAPRETTO, M. (2014). The Autism Brain Imaging Data Exchange: Towards a large-scale evaluation of the intrinsic brain architecture in Autism. *Mol. Psychiatry* **19** 659–667.
- DUBOIS, J. and ADOLPHS, R. (2016). Building a science of individual differences from fMRI. *Trends Cogn. Sci.* **20** 425–443.
- FAIR, D., DOSENBACH, N., CHURCH, J., COHEN, A., BRAHMBHATT, S., MIEZIN, F., BARCH, D., RAICHLE, M., PETERSEN, S. and SCHLAGGAR, B. (2007). Development of distinct control networks through segregation and integration. *Proc. Natl. Acad. Sci. USA* **104** 13507–13512.
- FAIR, D. A., COHEN, A. L., POWER, J. D., DOSENBACH, N. U. F., CHURCH, J. A., MIEZIN, F. M., SCHLAGGAR, B. L. and PETERSEN, S. E. (2009). Functional brain networks develop from a “local to distributed” organization. *PLoS Comput. Biol.* **5** e1000381. [MR2516098](#)
- FALLANI, F., RICHIARDI, J., CHAVEZ, M. and ACHARD, S. (2014). Graph analysis of functional brain networks: Practical issues in translational neuroscience. *Philos. Trans. R. Soc. Lond. B, Biol. Sci.* **369** 1–17.
- FOX, M. and GREICIUS, M. (2010). Clinical applications of resting state functional connectivity. *Front. Syst. Neurosci.* **4** 1–13.
- FRIEDMAN, J., HASTIE, T. and TIBSHIRANI, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9** 432–441.
- FRIEDMAN, J., HASTIE, T., HÖFLING, H. and TIBSHIRANI, R. (2007). Pathwise coordinate optimization. *Ann. Appl. Stat.* **1** 302–332. [MR2415737](#)
- FRISTON, K. (2011). Functional and effective connectivity: A review. *Brain Connect.* **1** 13–36.
- GREICIUS, M., KRASNOW, B., REISS, A. and MENON, V. (2003). Functional connectivity in the resting brain: A network analysis of the default mode hypothesis. *Proc. Natl. Acad. Sci. USA* **100** 253–258.
- GUSNARD, D. and RAICHLE, M. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nat. Rev., Neurosci.* **2** 685–694.
- KANAI, R. and REES, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev., Neurosci.* **12** 231–242.
- KELLY, C., BISWAL, B., CRADDOCK, C., CASTELLANOS, X. and MILHAM, M. (2012). Characterizing variation in the functional connectome: Promise and pitfalls. *Trends Cogn. Sci.* **16** 181–188.
- KRZANOWSKI, W. J. and HAND, D. J. (2009). *ROC Curves for Continuous Data. Monographs on Statistics and Applied Probability* **111**. CRC Press, Boca Raton, FL. [MR2522628](#)
- LEE, H., LEE, D., KANG, H., KIM, B. and CHUNG, M. (2011). Sparse brain network recovery under compressed sensing. *IEEE Trans. Med. Imag.* **30** 1154–1165.
- LINDQUIST, M. A. (2008). The statistical analysis of fMRI data. *Statist. Sci.* **23** 439–464. [MR2530545](#)
- MCLACHLAN, G. J. and KRISHNAN, T. (2007). *The EM Algorithm and Extensions. Wiley Series in Probability and Statistics* **382**. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ.
- MEINSHAUSEN, N. and BÜHLMANN, P. (2006). High-dimensional graphs and variable selection with the lasso. *Ann. Statist.* **34** 1436–1462. [MR2278363](#)
- MENG, X.-L. and VAN DYK, D. (1998). Fast EM-type implementations for mixed effects models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **60** 559–578. [MR1625942](#)
- MONTI, R. P., ANAGNOSTOPOULOS, C. and MONTANA, G. (2017). Supplement to “Learning population and subject-specific brain connectivity networks via mixed neighborhood selection.” DOI:[10.1214/17-AOAS1067SUPPA](#), DOI:[10.1214/17-AOAS1067SUPPB](#).
- MUELLER, S., WANG, D., FOX, M., YEO, T., SEPULCRE, J., SABUNCU, M., SHAFEE, R., LU, J. and LIU, H. (2013). Individual variability in functional connectivity architecture of the human brain. *Neuron* **77** 586–595.

- NARAYAN, M., ALLEN, G. and TOMSON, S. (2015). Two sample inference for populations of graphical models with applications to functional connectivity. Preprint. Available at [arXiv:1502.03853](https://arxiv.org/abs/1502.03853).
- NIELSEN, J., ZIELINSKI, B., FLETCHER, T., ALEXANDER, A., LANGE, N., BIGLER, E., LAINHART, J. and ANDERSON, J. (2013). Multisite functional connectivity MRI classification of autism: ABIDE results. *Front. Human Neurosci.* **7** 72–83.
- PINHEIRO, J. and BATES, D. (2000). *Mixed-Effects Models in S and S-PLUS*. Springer Science & Business Media, Berlin.
- POWER, J., BARNES, K., SNYDER, A., SCHLAGGAR, B. and PETERSEN, S. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* **59** 2142–2154.
- RUBINOV, M. and SPORNS, O. (2010). Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage* **52** 1059–1069.
- SCHELLDORFER, J., BÜHLMANN, P. and VAN DE GEER, S. (2011). Estimation for high-dimensional linear mixed-effects models using ℓ_1 -penalization. *Scand. J. Stat.* **38** 197–214. [MR2829596](https://arxiv.org/abs/1106.2822)
- SMITH, S. (2012). The future of fMRI connectivity. *NeuroImage* **62** 1257–1266.
- SMITH, S., MILLER, K., SALIMI-KHORSHIDI, G., WEBSTER, M., BECKMANN, C., NICHOLS, T., RAMSEY, J. and WOOLRICH, M. (2011). Network modelling methods for fMRI. *NeuroImage* **54** 875–891.
- TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **58** 267–288. [MR1379242](https://arxiv.org/abs/1379242)
- VAN DEN HEUVEL, M. and POL, H. (2010). Exploring the brain network: A review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* **20** 519–534.
- VAROQUAUX, G. and CRADDOCK, C. (2013). Learning and comparing functional connectomes across subjects. *NeuroImage* **80** 405–415.
- VAROQUAUX, G., GRAMFORT, A., POLINE, J. and THIRION, B. (2010). Brain covariance selection: Better individual functional connectivity models using population prior. In *Neural Information Processing Systems* 2334–2342.
- ZUO, X., DI MARTINO, A., KELLY, C., SHEHZAD, Z., GEE, D., KLEIN, D., CASTELLANOS, X., BISWAL, B. and MILHAM, M. (2010). The oscillating brain: Complex and reliable. *NeuroImage* **49** 1432–1445.

MULTIVARIATE SPATIOTEMPORAL MODELING OF AGE-SPECIFIC STROKE MORTALITY

BY HARRISON QUICK, LANCE A. WALLER AND MICHELE CASPER

*Drexel University, Emory University and the Centers
for Disease Control and Prevention*

Geographic patterns in stroke mortality have been studied as far back as the 1960s when a region of the southeastern United States became known as the “stroke belt” due to its unusually high rates. While stroke mortality rates are known to increase exponentially with age, an investigation of spatiotemporal trends by age group at the county level is daunting due to the preponderance of small population sizes and/or few stroke events by age group. In this paper, we implement a multivariate space–time conditional autoregressive model to investigate age-specific trends in county-level stroke mortality rates from 1973 to 2013. In addition to reinforcing existing claims in the literature, this work reveals that geographic disparities in the reduction of stroke mortality rates vary by age. More importantly, this work indicates that the geographic disparity between the “stroke belt” and the rest of the nation is not only persisting, but may in fact be worsening.

REFERENCES

- ANDERSON, R. N., MINIÑO, A. M., HOYERT, D. L. and ROSENBERG, H. M. (2001). Comparability of cause of death between ICD-9 and ICD-10: Preliminary estimates. *Natl. Vital Stat. Rep.* **49** 1–32.
- BERNARDINELLI, L., CLAYTON, D. and MONTOMOLI, C. (1995). Bayesian estimates of disease maps: How important are priors? *Stat. Med.* **14** 2411–2431.
- BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *J. Roy. Statist. Soc. Ser. B* **36** 192–236. [MR0373208](#)
- BESAG, J. and HIGDON, D. (1999). Bayesian analysis of agricultural field experiments. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **61** 691–746. [MR1722238](#)
- BESAG, J., YORK, J. and MOLLIÉ, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Ann. Inst. Statist. Math.* **43** 1–59. [MR1105822](#)
- BESAG, J., GREEN, P., HIGDON, D. and Mengersen, K. (1995). Bayesian computation and stochastic systems. *Statist. Sci.* **10** 3–66. [MR1349818](#)
- BORHANI, N. O. (1965). Changes and geographic distribution of mortality from cerebrovascular disease. *Am. J. Publ. Health* **55** 673–681.
- BOTELLA-ROCAMORA, P., MARTINEZ-BENEITO, M. A. and BANERJEE, S. (2015). A unifying modeling framework for highly multivariate disease mapping. *Stat. Med.* **34** 1548–1559. [MR3334675](#)
- CASPER, M., ANDA, R. F., KNOWLES, M., POLLARD, R. A. and WING, S. (1995). The shifting stroke belt: Changes in the geographic pattern of stroke mortality in the United States. *Stroke* **26** 755–760.

Key words and phrases. Age disparities in health, Bayesian methods, geographic disparities in health, nonseparable models, small area analysis.

- CDC (2003). CDC/ATSDR Policy on Releasing and Sharing Data. Manual; Guide CDC-02. Available at <http://www.cdc.gov/maso/Policy/ReleasingData.pdf>. Accessed June 30, 2015.
- EL-SAED, A., KULLER, L. H., NEWMAN, A. B., LOPEZ, O., COSTANTINO, J., MCTIGUE, K., CUSHMAN, M. and KRONMAL, R. (2006). Geographic variations in stroke incidence and mortality among older populations in four US communities. *Stroke* **37** 1975–1979.
- ERGIN, A., MUNTNER, P., SHERWIN, R. and HE, J. (2004). Secular trends in cardiovascular disease mortality, incidence, and case fatality rates in adults in the United States. *Am. J. Med.* **117** 219–227.
- GELFAND, A. E. and VOUNATSOU, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostat.* **4** 11–25.
- GILLUM, R. F., KWAGYAN, J. and OBISESAN, T. O. (2011). Ethnic and geographic variation in stroke mortality trends. *Stroke* **42** 3294–3296.
- HOWARD, V. J. (2013). Reasons underlying racial differences in stroke incidence and mortality. *Stroke* **44** S126–S128.
- HOWARD, G., HOWARD, V. J., KATHOLI, C., OLI, M. K., HUSTON, S. and ASPLUND, K. (2001). Decline in US stroke mortality: An analysis of temporal patterns by sex, race, and geographic region. *Stroke* **32** 2213–2218.
- KLEBBA, A. J. and SCOTT, J. H. (1980). Estimates of selected comparability ratios based on dual coding of 1976 death certificates by the eighth and ninth revisions of the International Classification of Diseases. *National Vital Statistics Reports* **28**.
- KNORR-HELD, L. (2000). Bayesian modelling of inseparable space–time variation in disease risk. *Stat. Med.* **19** 2555–2567.
- KNORR-HELD, L. and BESAG, J. (1998). Modelling risk from a disease in time and space. *Stat. Med.* **17** 2045–2060.
- KNORR-HELD, L. and RUE, H. (2002). On block updating in Markov random field models for disease mapping. *Scand. J. Stat.* **29** 597–614. [MR1988414](#)
- LISABETH, L. D., DIEZ ROUX, A. V., ESCOBAR, J. D., SMITH, M. A. and MORGENSTERN, L. B. (2006). Neighborhood environment and risk of ischemic stroke. *Am. J. Epidemiol.* **165** 279–287.
- MARTINEZ-BENEITO, M. A. (2013). A general modelling framework for multivariate disease mapping. *Biometrika* **100** 539–553. [MR3094436](#)
- MORAN, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37** 17–23. [MR0035933](#)
- NCHS (2013). Bridged-race population estimates: United States July 1st resident population by state, county, age, sex, bridged-race, and Hispanic origin. Available at http://www.cdc.gov/NCHS/nvss/bridged_race.htm. Accessed June 30, 2015.
- QUICK, H., CARLIN, B. P. and BANERJEE, S. (2015). Heteroscedastic conditional auto-regression models for areally referenced temporal processes for analysing California asthma hospitalization data. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **64** 799–813. [MR3415951](#)
- QUICK, H., WALLER, L. A. and CASPER, M. (2017a). A multivariate space–time model for analyzing county-level heart disease death rates by race and sex. *J. Roy. Statist. Soc. Ser. C.* DOI:[10.1111/rssc.12215](#).
- QUICK, H., WALLER, L. A. and CASPER, M. (2017b). Supplement to “Multivariate spatiotemporal modeling of age-specific stroke mortality.” DOI:[10.1214/17-AOAS1068SUPP](#).
- RAGHUNATHAN, T. E., REITER, J. P. and RUBIN, D. B. (2003). Multiple imputation for statistical disclosure limitation. *J. Off. Stat.* **19** 1–16.
- RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability* **104**. Chapman & Hall/CRC, Boca Raton, FL. [MR2130347](#)
- SCHIEB, L. J., MOBLEY, L. R., GEORGE, M. and CASPER, M. (2013). Tracking stroke hospitalization clusters over time and associations with county-level socioeconomic and healthcare characteristics. *Stroke* **44** 146–152.

- SPIEGELHALTER, D. J., BEST, N. G., CARLIN, B. P. and VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **64** 583–639. [MR1979380](#)
- TASSONE, E. C., WALLER, L. A. and CASPER, M. L. (2009). Small-area racial disparity in stroke mortality. *Epidemiology* **20** 234–241.
- WALLER, L. A., CARLIN, B. P., XIA, H. and GELFAND, A. E. (1997). Hierarchical spatio-temporal mapping of disease rates. *J. Amer. Statist. Assoc.* **92** 607–617.
- XU, J., MURPHY, S. L., KOCHANKE, K. D. and BASTIAN, B. A. (2016). Deaths: Final data for 2013. *Natl. Vital Stat. Rep.* **64** 1–119.

A RANDOM-EFFECTS HURDLE MODEL FOR PREDICTING BYCATCH OF ENDANGERED MARINE SPECIES¹

BY E. CANTONI, J. MILLS FLEMMING AND A. H. WELSH

University of Geneva, Dalhousie University and Australian National University

Understanding and reducing the incidence of accidental bycatch, particularly for vulnerable species such as sharks, is a major challenge for contemporary fisheries management worldwide. Bycatch data, most often collected by at-sea observers during fishing trips, are clustered by trip and/or vessel and typically involve a large number of zero counts and very few positive counts. Though hurdle models are very popular for count data with excess zeros, models for clustered forms have received far less attention. Here we present a novel random-effects hurdle model for bycatch data that makes available accurate estimates of bycatch probabilities as well as other cluster-specific targets. These are essential for informing conservation and management decisions as well as for identifying bycatch hotspots, often considered the first step in attempting to protect endangered marine species. We validate our methodology through simulation and use it to analyze bycatch data on critically endangered hammerhead sharks from the U.S. National Marine Fisheries Service Pelagic Observer Program.

REFERENCES

- ALFÒ, M. and MARUOTTI, A. (2010). Two-part regression models for longitudinal zero-inflated count data. *Canad. J. Statist.* **38** 197–216. [MR2682758](#)
- BAUM, J. (2007). Population- and community-level consequences of the exploitation of large predatory marine fishes, Ph.D. thesis, Biology Department, Dalhousie University, Halifax, Canada.
- BAUM, J. K., MYERS, R. A., KEHLER, D. G., WORM, B., HARLEY, S. J. and DOHERTY, P. A. (2003). Collapse and conservation of shark populations in the Northwest Atlantic. *Science* **299** 389–392.
- BOOTH, J. G. and HOBERT, J. P. (1998). Standard errors of prediction in generalized linear mixed models. *J. Amer. Statist. Assoc.* **93** 262–272. [MR1614632](#)
- CANTONI, E., MILLS FLEMMING, J. and WELSH, A. H. (2017). Supplement to “A random-effects hurdle model for predicting bycatch of endangered marine species.” DOI:[10.1214/17-AOAS1074SUPP](#).
- BRESLOW, N. E. and CLAYTON, D. G. (1993). Approximate inference in generalized linear mixed models. *J. Amer. Statist. Assoc.* **88** 9–25.
- DE BRUIJN, N. G. (1981). *Asymptotic Methods in Analysis*, 3rd ed. Dover Publications, New York. [MR0671583](#)
- DOBBIE, M. J. and WELSH, A. H. (2001). Modelling correlated zero-inflated count data. *Aust. N. Z. J. Stat.* **43** 431–444. [MR1872202](#)

Key words and phrases. Bycatch, clustered count data, excess of zeros, random-effects hurdle models, prediction.

- FOURNIER, D. A., SKAUG, H. J., ANCHETA, J., IANELLI, J., MAGNUSSON, A., MAUNDER, M. N., NIELSEN, A. and SIBERT, J. (2012). AD model builder: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optim. Methods Softw.* **27** 233–249. [MR2901959](#)
- FULTON, K. A., LIU, D., HAYNIE, D. L. and ALBERT, P. S. (2015). Mixed model and estimating equation approaches for zero inflation in clustered binary response data with application to a dating violence study. *Ann. Appl. Stat.* **9** 275–299. [MR3341116](#)
- FURRER, R., NYCHKA, D. and SAIN, S. (2012). fields: Tools for spatial data. R package version 6.6.3.
- HALL, M. A., ALVERSON, D. L. and METUZALS, K. I. (2000). By-catch: Problems and solutions. *Mar. Pollut. Bull.* **41** 204–219.
- HUBER, P., RONCHETTI, E. and VICTORIA-FESER, M.-P. (2004). Estimation of generalized linear latent variable models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 893–908. [MR2102471](#)
- HUR, K., HEDEKER, D., HENDERSON, W., KHURI, S. and DALEY, J. (2002). Modeling clustered count data with excess zeros in health care outcomes research. *Health Serv. Outcomes Res. Methodol.* **3** 5–20.
- JIANG, J. (2003). Empirical best prediction for small-area inference based on generalized linear mixed models. *J. Statist. Plann. Inference* **111** 117–127. [MR1955876](#)
- JIANG, J. (2007). *Linear and Generalized Linear Mixed Models and Their Applications*. Springer, New York. [MR2308058](#)
- JIANG, J., JIA, H. and CHEN, H. (2001). Maximum posterior estimation of random effects in generalized linear mixed models. *Statist. Sinica* **11** 97–120. [MR1820003](#)
- JIANG, J. and LAHIRI, P. (2001). Empirical best prediction for small area inference with binary data. *Ann. Inst. Statist. Math.* **53** 217–243. [MR1841133](#)
- JIANG, J., LAHIRI, P. and WAN, S.-M. (2002). A unified jackknife theory for empirical best prediction with M -estimation. *Ann. Statist.* **30** 1782–1810. [MR1969450](#)
- LAMBERT, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics* **34** 1–14.
- LEE, Y. and NELDER, J. A. (1996). Hierarchical generalized linear models. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **58** 619–678. [MR1410182](#)
- LELE, S. R., DENNIS, B. and LUTSCHER, F. (2007). Data cloning: Easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods. *Ecol. Lett.* **10** 551–563.
- LEWISON, R. L., CROWDER, L. B., READ, A. J. and FREEMAN, S. A. (2004). Understanding impacts of fisheries bycatch on marine megafauna. *Trends Ecol. Evol.* **19** 598–604.
- LIU, L., STRAWDERMAN, R. L., COWEN, M. E. and SHIH, Y. C. T. (2010). A flexible two-part random effects model for correlated medical costs. *J. Health Econ.* **29** 110–123.
- MCCULLOCH, C. E. and NEUHAUS, J. M. (2011). Misspecifying the shape of a random effects distribution: Why getting it wrong may not matter. *Statist. Sci.* **26** 388–402. [MR2917962](#)
- MIN, Y. and AGRESTI, A. (2002). Modeling nonnegative data with clumping at zero: A survey. *J. Iran. Stat. Soc. (JIRSS)* **1** 7–33.
- MIN, Y. and AGRESTI, A. (2005). Random effect models for repeated measures of zero-inflated count data. *Stat. Model.* **5** 1–19. [MR2133525](#)
- MOLAS, M. and LESAFFRE, E. (2010). Hurdle models for multilevel zero-inflated data via h -likelihood. *Stat. Med.* **29** 3294–3310. [MR2758721](#)
- MULLAHY, J. (1986). Specification and testing of some modified count data models. *J. Econometrics* **33** 341–365. [MR0867980](#)
- MYERS, R. A., BAUM, J. K., SHEPHERD, T. D., POWERS, S. P. and PETERSON, C. H. (2007). Cascading effects of the loss of apex predatory sharks from a coastal ocean. *Science* **315** 1846–1850.

- NEELON, B., GHOSH, P. and LOEBS, P. F. (2013). A spatial Poisson hurdle model for exploring geographic variation in emergency department visits. *J. Roy. Statist. Soc. Ser. A* **176** 389–413. [MR3045852](#)
- PIKITCH, E., SANTORA, C., BABCOCK, E., BAKUN, A., BONFIL, R., CONOVER, D., DAYTON, P., DOUKAKIS, P., FLUHARTY, D. et al. (2004). Ecosystem-based fishery management. *Science* **305** 346–347.
- R DEVELOPMENT CORE TEAM (2011). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- RABE-HESKETH, S., SKRONDAL, A. and PICKLES, A. (2002). Reliable estimation of generalized linear mixed models using adaptive quadrature. *The Stata Journal* **2** 1–21.
- SALIBIÁN-BARRERA, M. A., VAN AELST, S. and WILLEMS, G. (2008). Fast and robust bootstrap. *Stat. Methods Appl.* **17** 41–71. [MR2379121](#)
- SKAUG, H., FOURNIER, D., NIELSEN, A., MAGNUSSON, A. and BOLKER, B. (2012). Generalized Linear Mixed Models using AD Model Builder. R package version 0.7.4.
- SU, L., TOM, B. D. M. and FAREWELL, V. T. (2009). Bias in 2-part mixed models for longitudinal semicontinuous data. *Biostatistics* **10** 374–389.
- WELSH, A. H., CUNNINGHAM, R. B. and CHAMBERS, R. L. (2000). Methodology for estimating the abundance of rare animals: Seabird nesting on North East herald cay. *Biometrics* **56** 22–30.
- YAU, K. K. W. and LEE, A. H. (2001). Zero-inflated Poisson regression with random effects to evaluate an occupational injury prevention programme. *Stat. Med.* **20** 2907–2920.

EMPIRICAL BAYESIAN ANALYSIS OF SIMULTANEOUS CHANGEPOINTS IN MULTIPLE DATA SEQUENCES

BY ZHOU FAN¹ AND LESTER MACKEY²

Stanford University and Microsoft Research

Copy number variations in cancer cells and volatility fluctuations in stock prices are commonly manifested as changepoints occurring at the same positions across related data sequences. We introduce a Bayesian modeling framework, BASIC, that employs a changepoint prior to capture the co-occurrence tendency in data of this type. We design efficient algorithms to sample from and maximize over the BASIC changepoint posterior and develop a Monte Carlo expectation-maximization procedure to select prior hyperparameters in an empirical Bayes fashion. We use the resulting BASIC framework to analyze DNA copy number variations in the NCI-60 cancer cell lines and to identify important events that affected the price volatility of S&P 500 stocks from 2000 to 2009.

REFERENCES

- ADAMS, R. P. and MACKAY, D. J. (2007). Bayesian online changepoint detection. Technical report. Available at [arXiv:0710.3742](https://arxiv.org/abs/0710.3742) [stat.ML].
- AKHOONDI, S. et al. (2007). FBXW7/hCDC4 is a general tumor suppressor in human cancer. *Cancer Res.* **67** 9006–9012.
- ANDRIEU, C., DOUCET, A. and HOLENSTEIN, R. (2010). Particle Markov chain Monte Carlo methods. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **72** 269–342. [MR2758115](#)
- BARDWELL, L. and FEARNHEAD, P. (2017). Bayesian detection of abnormal segments in multiple time series. *Bayesian Anal.* **12** 193–218. [MR3597572](#)
- BARRY, D. and HARTIGAN, J. A. (1993). A Bayesian analysis for change point problems. *J. Amer. Statist. Assoc.* **88** 309–319. [MR1212493](#)
- BASSEVILLE, M. and NIKIFOROV, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, Englewood Cliffs, NJ. [MR1210954](#)
- CHEN, J. and GUPTA, A. K. (2012). *Parametric Statistical Change Point Analysis: With Applications to Genetics, Medicine, and Finance*, 2nd ed. Birkhäuser/Springer, New York. [MR3025631](#)
- CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Stat.* **35** 999–1018. [MR0179874](#)
- CHIB, S. (1998). Estimation and comparison of multiple change-point models. *J. Econometrics* **86** 221–241. [MR1649222](#)
- DANG, C. V. (2012). MYC on the path to cancer. *Cell* **149** 22–35.
- DOBIGEON, N., TOURNERET, J.-Y. and DAVY, M. (2007). Joint segmentation of piecewise constant autoregressive processes by using a hierarchical model and a Bayesian sampling approach. *IEEE Trans. Signal Process.* **55** 1251–1263. [MR2464988](#)
- FAN, Z. and MACKEY, L. (2017). Supplement to “Empirical Bayesian analysis of simultaneous changepoints in multiple data sequences.” DOI:[10.1214/17-AOAS1075SUPP](https://doi.org/10.1214/17-AOAS1075SUPP).

Key words and phrases. Changepoint detection, empirical Bayes, Markov chain Monte Carlo, copy number variation, stock price volatility.

- FAN, Z., DROR, R. O., MILDORF, T. J., PIANA, S. and SHAW, D. E. (2015). Identifying localized changes in large systems: Change-point detection for biomolecular simulations. *Proc. Natl. Acad. Sci. USA* **112** 7454–7459.
- FEARNHEAD, P. (2006). Exact and efficient Bayesian inference for multiple changepoint problems. *Stat. Comput.* **16** 203–213. [MR2227396](#)
- FEARNHEAD, P. and LIU, Z. (2007). On-line inference for multiple changepoint problems. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **69** 589–605. [MR2370070](#)
- HARLÉ, F., CHATELAIN, F., GOUY-PAILLER, C. and ACHARD, S. (2016). Bayesian model for multiple change-points detection in multivariate time series. *IEEE Trans. Signal Process.* **64** 4351–4362. [MR3528848](#)
- HEALY, J. D. (1987). A note on multivariate CUSUM procedures. *Technometrics* **29** 409–412.
- HSU, D.-A. (1977). Tests for variance shift at an unknown time point. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **26** 279–284.
- HUGHES, A. E. et al. (2006). A common CFH haplotype, with deletion of CFHR1 and CFHR3, is associated with lower risk of age-related macular degeneration. *Nat. Genet.* **38** 1173–1177.
- JACKSON, B. et al. (2005). An algorithm for optimal partitioning of data on an interval. *IEEE Signal Process. Lett.* **12** 105–108.
- JENG, X. J., CAI, T. T. and LI, H. (2013). Simultaneous discovery of rare and common segment variants. *Biometrika* **100** 157–172. [MR3034330](#)
- KAMB, A. et al. (1994). A cell cycle regulator potentially involved in genesis of many tumor types. *Science* **264** 436–439.
- KILLICK, R., FEARNHEAD, P. and ECKLEY, I. A. (2012). Optimal detection of changepoints with a linear computational cost. *J. Amer. Statist. Assoc.* **107** 1590–1598. [MR3036418](#)
- LAI, W. R., JOHNSON, M. D., KUCHERLAPATI, R. and PARK, P. J. (2005). Comparative analysis of algorithms for identifying amplifications and deletions in array CGH data. *Bioinformatics* **21** 3763–3770.
- LINDORFF-LARSEN, K., PIANA, S., DROR, R. O. and SHAW, D. E. (2011). How fast-folding proteins fold. *Science* **334** 517–520.
- LONG, J. et al. (2013). A common deletion in the APOBEC3 genes and breast cancer risk. *J. Natl. Cancer Inst.* **105** 573–579.
- LOUHIMO, R., LEPIKHOVA, T., MONNI, O. and HAUTANIEMI, S. (2012). Comparative analysis of algorithms for integration of copy number and expression data. *Nat. Methods* **9** 351–355.
- MENGES, C. W., ALTOMARE, D. A. and TESTA, J. R. (2009). FAS-associated factor 1 (FAF1): Diverse functions and implications for oncogenesis. *Cell Cycle* **8** 2528–2534.
- NOBORI, T. (1994). Deletions of the cyclin-dependent kinase-4 inhibitor gene in multiple human cancers. *Trends in Genetics* **10** 228.
- NOWAK, G., HASTIE, T., POLLACK, J. R. and TIBSHIRANI, R. (2011). A fused lasso latent feature model for analyzing multi-sample aCGH data. *Biostatistics* **12** 776–791.
- OLSHEN, A. B., VENKATRAMAN, E., LUCITO, R. and WIGLER, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* **5** 557–572.
- PICARD, F., LEBARBIER, E., HOEBEKE, M., RIGAILL, G., THIAM, B. and ROBIN, S. (2011). Joint segmentation, calling, and normalization of multiple CGH profiles. *Biostatistics* **12** 413–428.
- POLLACK, J. R. and BROWN, P. O. (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat. Genet.* **23** 41–46.
- ROBBINS, H. (1956). An empirical Bayes approach to statistics. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, Vol. I* 157–163. Univ. California Press, Berkeley. [MR0084919](#)
- SHAH, S. P., LAM, W. L., NG, R. T. and MURPHY, K. P. (2007). Modeling recurrent DNA copy number alterations in array CGH data. *Bioinformatics* **23** i450–i458.
- SIEGMUND, D., YAKIR, B. and ZHANG, N. R. (2011). Detecting simultaneous variant intervals in aligned sequences. *Ann. Appl. Stat.* **5** 645–668. [MR2840169](#)

- SRIVASTAVA, M. S. and WORSLEY, K. J. (1986). Likelihood ratio tests for a change in the multivariate normal mean. *J. Amer. Statist. Assoc.* **81** 199–204. [MR0830581](#)
- STEPHENS, D. A. (1994). Bayesian retrospective multiple-changepoint identification. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **43** 159–178.
- TADA, M. et al. (2010). Prognostic significance of genetic alterations detected by high-density single nucleotide polymorphism array in gastric cancer. *Cancer Science* **101** 1261–1269.
- THEURILLAT, J.-P. et al. (2011). URI is an oncogene amplified in ovarian cancer cells and is required for their survival. *Cancer Cell* **19** 317–332.
- TRAUTMANN, K. et al. (2006). Chromosomal instability in microsatellite-unstable and stable colon cancer. *Clin. Cancer Res.* **12** 6379–6385.
- VARMA, S., POMMIER, Y., SUNSHINE, M., WEINSTEIN, J. N. and REINHOLD, W. C. (2014). High resolution copy number variation data in the NCI-60 cancer cell lines from whole genome microarrays accessible through CellMiner. *PLoS ONE* **9** e92047.
- WEI, G. C. and TANNER, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. *J. Amer. Statist. Assoc.* **85** 699–704.
- XUAN, D. et al. (2013). APOBEC3 deletion polymorphism is associated with breast cancer risk among women of European ancestry. *Carcinogenesis* **34** 2240–2243.
- YAO, Y.-C. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *Ann. Statist.* **12** 1434–1447. [MR0760698](#)
- ZHANG, N. R. and SIEGMUND, D. O. (2012). Model selection for high-dimensional, multi-sequence change-point problems. *Statist. Sinica* **22** 1507–1538. [MR3027097](#)
- ZHANG, N. R., SIEGMUND, D. O., JI, H. and LI, J. Z. (2010). Detecting simultaneous changepoints in multiple sequences. *Biometrika* **97** 631–645. [MR2672488](#)
- ZHOU, X., YANG, C., WAN, X., ZHAO, H. and YU, W. (2013). Multisample aCGH data analysis via total variation and spectral regularization. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **10** 230–235.

BAYESIAN INFERENCE FOR MULTIPLE GAUSSIAN GRAPHICAL MODELS WITH APPLICATION TO METABOLIC ASSOCIATION NETWORKS

BY LINDA S. L. TAN^{*,1}, AJAY JASRA^{*,2} MARIA DE IORIO[†] AND
TIMOTHY M. D. EBBELS^{‡,3}

National University of Singapore, University College London[†] and
Imperial College London[‡]*

We investigate the effect of cadmium (a toxic environmental pollutant) on the correlation structure of a number of urinary metabolites using Gaussian graphical models (GGMs). The inferred metabolic associations can provide important information on the physiological state of a metabolic system and insights on complex metabolic relationships. Using the fitted GGMs, we construct differential networks, which highlight significant changes in metabolite interactions under different experimental conditions. The analysis of such metabolic association networks can reveal differences in the underlying biological reactions caused by cadmium exposure. We consider Bayesian inference and propose using the multiplicative (or Chung–Lu random graph) model as a prior on the graphical space. In the multiplicative model, each edge is chosen independently with probability equal to the product of the connectivities of the end nodes. This class of prior is parsimonious yet highly flexible; it can be used to encourage sparsity or graphs with a pre-specified degree distribution when such prior knowledge is available. We extend the multiplicative model to multiple GGMs linking the probability of edge inclusion through logistic regression and demonstrate how this leads to joint inference for multiple GGMs. A sequential Monte Carlo (SMC) algorithm is developed for estimating the posterior distribution of the graphs.

REFERENCES

- ARMSTRONG, H., CARTER, C. K., WONG, K. F. K. and KOHN, R. (2009). Bayesian covariance matrix estimation using a mixture of decomposable graphical models. *Stat. Comput.* **19** 303–316. [MR2516221](#)
- ATAY-KAYIS, A. and MASSAM, H. (2005). A Monte Carlo method for computing the marginal likelihood in nondecomposable Gaussian graphical models. *Biometrika* **92** 317–335. [MR2201362](#)
- BESKOS, A., JASRA, A., KANTAS, N. and THIERY, A. (2016). On the convergence of adaptive sequential Monte Carlo methods. *Ann. Appl. Probab.* **26** 1111–1146. [MR3476634](#)
- CARVALHO, C. M. and SCOTT, J. G. (2009). Objective Bayesian model selection in Gaussian graphical models. *Biometrika* **96** 497–512. [MR2538753](#)
- CARVALHO, C. M. and WEST, M. (2007). Dynamic matrix-variate graphical models. *Bayesian Anal.* **2** 69–97. [MR2289924](#)
- CHUN, H., ZHANG, X. and ZHAO, H. (2015). Gene regulation network inference with joint sparse Gaussian graphical models. *J. Comput. Graph. Statist.* **24** 954–974. [MR3432924](#)

Key words and phrases. Gaussian graphical models, prior specification, multiplicative model, sequential Monte Carlo.

- CHUNG, F. and LU, L. (2002). The average distances in random graphs with given expected degrees. *Proc. Natl. Acad. Sci. USA* **99** 15879–15882. [MR1944974](#)
- D’SOUZA, R. M., BORGS, C., CHAYES, J. T., BERGER, N. and KLEINBERG, R. D. (2007). Emergence of tempered preferential attachment from optimization. *Proc. Natl. Acad. Sci. USA* **104** 6112–6117.
- DANAHER, P., WANG, P. and WITTEN, D. M. (2014). The joint graphical lasso for inverse covariance estimation across multiple classes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **76** 373–397. [MR3164871](#)
- DEL MORAL, P., DOUCET, A. and JASRA, A. (2006). Sequential Monte Carlo samplers. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **68** 411–436. [MR2278333](#)
- DEL MORAL, P., DOUCET, A. and JASRA, A. (2012). An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Stat. Comput.* **22** 1009–1020. [MR2950081](#)
- DEMPSTER, A. P. (1972). Covariance selection. *Biometrics* 157–175.
- DIACONIS, P. and YLVISAKER, D. (1979). Conjugate priors for exponential families. *Ann. Statist.* **7** 269–281. [MR0520238](#)
- DOBRA, A., HANS, C., JONES, B., NEVINS, J. R., YAO, G. and WEST, M. (2004). Sparse graphical models for exploring gene expression data. *J. Multivariate Anal.* **90** 196–212. [MR2064941](#)
- ELLIS, J. K., ATHERSUCH, T. J., THOMAS, L. D., TEICHERT, F., PEREZ-TRUJILLO, M., SVENDSEN, C., SPURGEON, D. J., SINGH, R., JARUP, L., BUNDY, J. G. and KEUN, H. C. (2012). Metabolic profiling detects early effects of environmental and lifestyle exposure to cadmium in a human population. *BMC Medicine* **10** 61.
- FENNER, T., LEVENE, M. and LOIZOU, G. (2007). A model for collaboration networks giving rise to a power-law distribution with an exponential cutoff. *Soc. Netw.* **29** 70–80.
- FRIEDMAN, J., HASTIE, T. and TIBSHIRANI, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9** 432–441.
- GIOT, L., BADER, J. S., BROUWER, C. et al. (2003). A protein interaction map of *Drosophila melanogaster*. *Science* **302** 1727–1736.
- GIRAUD, C., HUET, S. and VERZELEN, N. (2012). Graph selection with GGMselect. *Stat. Appl. Genet. Mol. Biol.* **11** Art. 3, 52. [MR2934907](#)
- GUO, J., LEVINA, E., MICHAELIDIS, G. and ZHU, J. (2011). Joint estimation of multiple graphical models. *Biometrika* **98** 1–15. [MR2804206](#)
- JASRA, A., STEPHENS, D. A., DOUCET, A. and TSAGARIS, T. (2011). Inference for Lévy driven stochastic volatility models via adaptive sequential Monte Carlo. *Scand. J. Stat.* **38** 1–22.
- JEONG, H., MASON, S. P., BARABASI, A. L. and OLTVAI, Z. N. (2001). Lethality and centrality in protein networks. *Nature* **411** 41–42.
- 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](#)
- LAURITZEN, S. L. (1996). *Graphical Models. Oxford Statistical Science Series* **17**. Oxford Univ. Press, New York. [MR1419991](#)
- LENKOSKI, A. and DOBRA, A. (2011). Computational aspects related to inference in Gaussian graphical models with the G-Wishart prior. *J. Comput. Graph. Statist.* **20** 140–157. [MR2816542](#)
- LIU, Y., LI, Y., LIU, K. and SHEN, J. (2014). Exposing to cadmium stress cause profound toxic effect on microbiota of the mice intestinal tract. *PLoS ONE* **9** e85323.
- MITRA, R., MÜLLER, P. and JI, Y. (2016). Bayesian graphical models for differential pathways. *Bayesian Anal.* **11** 99–124. [MR3447093](#)
- MOHAN, K., LONDON, P., FAZEL, M., WITTEN, D. and LEE, S.-I. (2014). Node-based learning of multiple Gaussian graphical models. *J. Mach. Learn. Res.* **15** 445–488. [MR3190845](#)
- MURRAY, I., GHAHRAMANI, Z. and MACKAY, D. J. C. (2006). MCMC for doubly-intractable distributions. In *Proceedings of the 22nd Annual Conference on Uncertainty in Artificial Intelligence* (T. Dechter and T. Richardson, eds.) 359–366.

- NEWMAN, M. E. J. (2001). The structure of scientific collaboration networks. *Proc. Natl. Acad. Sci. USA* **98** 404–409. [MR1812610](#)
- NEWMAN, M. E. J., STROGATZ, S. H. and WATTS, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* (3) **64**.
- OLHEDE, S. C. and WOLFE, P. J. (2013). Degree-based network models. Available at [arXiv:1211.6537](#).
- PETERSON, C., STINGO, F. C. and VANNUCCI, M. (2015). Bayesian inference of multiple Gaussian graphical models. *J. Amer. Statist. Assoc.* **110** 159–174. [MR3338494](#)
- RASTELLI, R., FRIEL, N. and RAFTERY, A. E. (2015). Properties of latent variable network models. Available at [arXiv:1506.07806](#).
- ROVERATO, A. (2002). Hyper inverse Wishart distribution for non-decomposable graphs and its application to Bayesian inference for Gaussian graphical models. *Scand. J. Stat.* **29** 391–411. [MR1925566](#)
- SALAMANCA, B. V., EBBELS, T. M. D. and DE IORIO, M. (2014). Variance and covariance heterogeneity analysis for detection of metabolites associated with cadmium exposure. *Stat. Appl. Genet. Mol. Biol.* **13** 191–201. [MR3181222](#)
- SCHAEFER, J., OPGEN-RHEIN, R. and STRIMMER, K. (2015). R package: GeneNet version 1.2.13. Available at <https://cran.r-project.org/web/packages/GeneNet/index.html>.
- SCHÄFER, C. and CHOPIN, N. (2013). Sequential Monte Carlo on large binary sampling spaces. *Stat. Comput.* **23** 163–184. [MR3016936](#)
- STEUER, R. (2006). Review: On the analysis and interpretation of correlations in metabolomic data. *Brief. Bioinform.* **7** 151–158.
- TAN, L. S., JASRA, A., DE IORIO, M. and EBBELS, T. M. (2017). Supplement to “Bayesian inference for multiple Gaussian graphical models with application to metabolic association networks.” DOI:[10.1214/17-AOAS1076SUPP](#).
- TELESCA, D., MÜLLER, P., PARMIGIANI, G. and FREEDMAN, R. S. (2012). Modeling dependent gene expression. *Ann. Appl. Stat.* **6** 542–560. [MR2976482](#)
- VALCÁRCCEL, B., WÜRTZ, P., SEICH AL BASATENA, N.-K., TUKIAINEN, T., KANGAS, A. J., SOININEN, P., JÄRVELIN, M.-R., ALA-KORPELA, M., EBBELS, T. M. and DE IORIO, M. (2011). A differential network approach to exploring differences between biological states: An application to prediabetes. *PLoS ONE* **6** e24702.
- WANG, H. and LI, S. Z. (2012). Efficient Gaussian graphical model determination under G -Wishart prior distributions. *Electron. J. Stat.* **6** 168–198. [MR2879676](#)
- WANG, H., REESON, C. and CARVALHO, C. M. (2011). Dynamic financial index models: Modeling conditional dependencies via graphs. *Bayesian Anal.* **6** 639–663. [MR2869960](#)
- YAJIMA, M., TELESCA, D., JI, Y. and MÜLLER, P. (2015). Detecting differential patterns of interaction in molecular pathways. *Biostatistics* **16** 240–251. [MR3365426](#)

SEMIPARAMETRIC COVARIATE-MODULATED LOCAL FALSE DISCOVERY RATE FOR GENOME-WIDE ASSOCIATION STUDIES¹

BY RONG W. ZABLOCKI^{*,†}, RICHARD A. LEVINE^{*}, ANDREW J. SCHORK[‡],
SHUJING XU[‡], YUNPENG WANG[§], CHUN C. FAN[‡] AND
WESLEY K. THOMPSON^{‡,¶,||,2}

San Diego State University^{}, Claremont Graduate University[†],
University of California, San Diego[‡], University of Oslo[§],
Institute of Biological Psychiatry[¶], and The Lundbeck Foundation
Initiative for Integrative Psychiatric Research^{||}*

While genome-wide association studies (GWAS) have discovered thousands of risk loci for heritable disorders, so far even very large meta-analyses have recovered only a fraction of the heritability of most complex traits. Recent work utilizing variance components models has demonstrated that a larger fraction of the heritability of complex phenotypes is captured by the additive effects of SNPs than is evident only in loci surpassing genome-wide significance thresholds, typically set at a Bonferroni-inspired $p \leq 5 \times 10^{-8}$. Procedures that control false discovery rate can be more powerful, yet these are still under-powered to detect the majority of nonnull effects from GWAS. The current work proposes a novel Bayesian semiparametric two-group mixture model and develops a Markov Chain Monte Carlo (MCMC) algorithm for a covariate-modulated local false discovery rate (*cmfdr*). The probability of being nonnull depends on a set of covariates via a logistic function, and the nonnull distribution is approximated as a linear combination of B-spline densities, where the weight of each B-spline density depends on a multinomial function of the covariates. The proposed methods were motivated by work on a large meta-analysis of schizophrenia GWAS performed by the Psychiatric Genetics Consortium (PGC). We show that the new *cmfdr* model fits the PGC schizophrenia GWAS test statistics well, performing better than our previously proposed parametric gamma model for estimating the nonnull density and substantially improving power over usual *fd*r. Using loci declared significant at *cmfdr* ≤ 0.20 , we perform follow-up pathway analyses using the Kyoto Encyclopedia of Genes and Genomes (KEGG) *Homo sapiens* pathways database. We demonstrate that the increased yield from the *cmfdr* model results in an improved ability to test for pathways associated with schizophrenia compared to using those SNPs selected according to usual *fd*r.

REFERENCES

1000 GENOMES PROJECT CONSORTIUM, ABECASIS, G. R., AUTON, A., BROOKS, L. D., DEPRISTO, M. A., DURBIN, R. M., HANDSAKER, R. E., KANG, H. M., MARTH, G. T. and

Key words and phrases. Bayesian mixture model, B-spline densities, genome-wide association study, multiple-comparison procedures, mixture of experts.

- MCVEAN, G. A. (2012). An integrated map of genetic variation from 1,092 human genomes. *Nature* **491** 56–65.
- BATTAGLINO, R., FU, J., SPÄTE, U., ERSOY, U., JOE, M., SEDAGHAT, L. and STASHENKO, P. (2004). Serotonin regulates osteoclast differentiation through its transporter. *J. Bone Miner. Res.* **19** 1420–1431.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B* **57** 289–300. [MR1325392](#)
- BUCKLEY, P. F., PILLAI, A. and HOWELL, K. R. (2011). Brain-derived neurotrophic factor: Findings in schizophrenia. *Curr. Opin. Psychiatry* **24** 122–127.
- CHIB, S. and JELIAZKOV, I. (2006). Inference in semiparametric dynamic models for binary longitudinal data. *J. Amer. Statist. Assoc.* **101** 685–700. [MR2256181](#)
- DAVIES, G., TENESA, A., PAYTON, A., YANG, J., HARRIS, S. E., LIEWALD, D., KE, X., LE HEL-LARD, S., CHRISTOFOROU, A., LUCIANO, M. et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. *Mol. Psychiatry* **16** 996–1005.
- EFRON, B. (2004). Large-scale simultaneous hypothesis testing: The choice of a null hypothesis. *J. Amer. Statist. Assoc.* **99** 96–104. [MR2054289](#)
- EFRON, B. (2007). Size, power and false discovery rates. *Ann. Statist.* **35** 1351–1377. [MR2351089](#)
- EFRON, B. and TIBSHIRANI, R. (2002). Empirical Bayes methods and false discovery rates for microarrays. *Genet. Epidemiol.* **23** 70–86.
- EILERS, P. H. C. and MARX, B. D. (1996). Flexible smoothing with *B*-splines and penalties. *Statist. Sci.* **11** 89–121. With comments and a rejoinder by the authors. [MR1435485](#)
- FERKINGSTAD, E., FRIGESSI, A., RUE, H., THORLEIFSSON, G. and KONG, A. (2008). Unsuper-vised empirical Bayesian multiple testing with external covariates. *Ann. Appl. Stat.* **2** 714–735. [MR2524353](#)
- GARDINER, E., BEVERIDGE, N., WU, J., CARR, V., SCOTT, R., TOONEY, P. and CAIRNS, M. (2012). Imprinted DLK1-DIO3 region of 14q32 defines a schizophrenia-associated miRNA sig-nature in peripheral blood mononuclear cells. *Mol. Psychiatry* **17** 827–840.
- GELMAN, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Anal.* **1** 515–533. [MR2221284](#)
- GLAZIER, A. M., NADEAU, J. H. and AITMAN, T. J. (2002). Finding genes that underlie complex traits. *Science* **298** 2345–2349.
- GREINER, A. and NICOLSON, G. (1965). Schizophrenia-melanosis. *Lancet* **286** 1165–1167.
- HOLMANS, P., GREEN, E. K., PAHWA, J. S., FERREIRA, M. A., PURCELL, S. M., SKLAR, P., OWEN, M. J., O'DONOVAN, M. C., CRADDOCK, N., THE WELLCOME TRUST CASE-CONTROL CONSORTIUM et al. (2009). Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am. J. Hum. Genet.* **85** 13–24.
- HSU, F., KENT, W. J., CLAWSON, H., KUHN, R. M., DIEKHANS, M. and HAUSSLER, D. (2006). The UCSC known genes. *Bioinformatics* **22** 1036–1046.
- KANEHISA, M. and GOTO, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28** 27–30.
- KANEHISA, M., SATO, Y., KAWASHIMA, M., FURUMICHI, M. and TANABE, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44** D457–D462.
- LANG, S. and BREZGER, A. (2004). Bayesian P-splines. *J. Comput. Graph. Statist.* **13** 183–212. [MR2044877](#)
- LEUCHT, S., BURKARD, T., HENDERSON, J., MAJ, M. and SARTORIUS, N. (2007). Physical ill-ness and schizophrenia: A review of the literature. *Acta Psychiatr. Scand.* **116** 317–333.
- LEWINGER, J. P., CONTI, D. V., BAURLEY, J. W., TRICHE, T. J. and THOMAS, D. C. (2007). Hierarchical Bayes prioritization of marker associations from a genome-wide association scan for further investigation. *Genet. Epidemiol.* **31** 871–882.
- LIDOW, M. S. (2003). Calcium signaling dysfunction in schizophrenia: A unifying approach. *Brains Res. Rev.* **43** 70–84.

- LIU, Y., LI, Z., ZHANG, M., DENG, Y., YI, Z. and SHI, T. (2013). Exploring the pathogenetic association between schizophrenia and type 2 diabetes mellitus diseases based on pathway analysis. *BMC Med. Genomics* **6** Article ID 1.
- LOPES, H. F. and DIAS, R. (2012). Bayesian mixture of parametric and nonparametric density estimation: A misspecification problem. *Braz. Rev. Econometrics* **31** 19–44.
- MAITI, S., KUMAR, K. H. B. G., CASTELLANI, C. A., O'REILLY, R. and SINGH, S. M. (2011). Ontogenetic de novo copy number variations (CNVs) as a source of genetic individuality: Studies on two families with MZD twins for schizophrenia. *PLoS ONE* **6** Article ID e17125.
- MARTIN, R. and TOKDAR, S. T. (2012). A nonparametric empirical Bayes framework for large-scale multiple testing. *Biostatistics* **13** 427–439.
- PSYCHIATRIC-GENOMICS-CONSORTIUM (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511** 421–427.
- PSYCHIATRIC-GWAS-CONSORTIUM (2011). Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* **43** 969–976.
- PURCELL, S. M., WRAY, N. R., STONE, J. L., VISSCHER, P. M., O'DONOVAN, M. C., SULLIVAN, P. F., SKLAR, P., RUDERFER, D. M., MCQUILLIN, A., MORRIS, D. W. et al. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460** 748–752.
- PUTNAM, D. K., SUN, J. and ZHAO, Z. (2011). Exploring schizophrenia drug-gene interactions through molecular network and pathway modeling. In *AMIA Annual Symposium Proceedings* 1127–1133.
- REICH, D. E., CARGILL, M., BOLK, S., IRELAND, J., SABETI, P. C., RICHTER, D. J., LAVERY, T., KOUYOUMJIAN, R., FARHADIAN, S. F., WARD, R. et al. (2001). Linkage disequilibrium in the human genome. *Nature* **411** 199–204.
- ROSEN, O. and THOMPSON, W. K. (2015). Bayesian semiparametric copula estimation with application to psychiatric genetics. *Biom. J.* **57** 468–484. [MR3340310](#)
- RUPPERT, D. (2002). Selecting the number of knots for penalized splines. *J. Comput. Graph. Statist.* **11** 735–757. [MR1944261](#)
- SCHORK, A. J., THOMPSON, W. K., PHAM, P., TORKAMANI, A., RODDEY, J. C., SULLIVAN, P. F., KELSOE, J. R., O'DONOVAN, M. C., FURBERG, H., SCHORK, N. J. et al. (2013). All SNPs are not created equal: Genome-wide association studies reveal a consistent pattern of enrichment among functionally annotated SNPs. *PLoS Genet.* **9** Article ID e1003449.
- SCOTT, J. G., KELLY, R. C., SMITH, M. A., ZHOU, P. and KASS, R. E. (2015). False discovery rate regression: An application to neural synchrony detection in primary visual cortex. *J. Amer. Statist. Assoc.* **110** 459–471. [MR3367240](#)
- THOMPSON, W. K. and ROSEN, O. (2008). A Bayesian model for sparse functional data. *Biometrics* **64** 54–63, 321. [MR2422819](#)
- VEHTARI, A. and GELMAN, A. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat. Comput.* **27** 1413–1432. [MR3647105](#)
- WAND, M. P., ORMEROD, J. T., PADOAN, S. A. and FRÜHRWIRTH, R. (2011). Mean field variational Bayes for elaborate distributions. *Bayesian Anal.* **6** 847–900. [MR2869967](#)
- WILLER, C. J., LI, Y. and ABECASIS, G. R. (2010). METAL: Fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26** 2190–2191.
- YANG, J., BENYAMIN, B., McEVOY, B. P., GORDON, S., HENDERS, A. K., NYHOLT, D. R., MADDEN, P. A., HEATH, A. C., MARTIN, N. G., MONTGOMERY, G. W. et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42** 565–569.
- YANG, J., BAKSHI, A., ZHU, Z., HEMANI, G., VINKHUYZEN, A. A., LEE, S. H., ROBINSON, M. R., PERRY, J. R., NOLTE, I. M., VAN VLIET-OSTAPTCHOUK, J. V. et al. (2015). Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index. *Nat. Genet.* **47** 1114–1120.

- ZABLOCKI, R. W., LEVINE, R. A., SCHORK, A. J., XU, S., WANG, Y., FAN, C. C. and THOMPSON, W. K. (2017). Supplement to “Semiparametric covariate-modulated local false discovery rate for genome-wide association studies.” DOI:[10.1214/17-AOAS1077SUPP](https://doi.org/10.1214/17-AOAS1077SUPP).
- ZABLOCKI, R. W., SCHORK, A. J., LEVINE, R. A., ANDREASSEN, O. A., DALE, A. M. and THOMPSON, W. K. (2014). Covariate-modulated local false discovery rate for genome-wide association studies. *Bioinformatics* **30** 2098–2104.

POINT PROCESS MODELS FOR SPATIO-TEMPORAL DISTANCE SAMPLING DATA FROM A LARGE-SCALE SURVEY OF BLUE WHALES¹

BY YUAN YUAN^{*}, FABIAN E. BACHL[†], FINN LINDGREN[†],
DAVID L. BORCHERS^{*}, JANINE B. ILLIAN^{*,‡}, STEPHEN T. BUCKLAND^{*},
HÅVARD RUE[§] AND TIM GERRODETTE[¶]

University of St Andrews^{}, University of Edinburgh[†], Norwegian University of Science and Technology[‡], King Abdullah University of Science and Technology[§] and NOAA National Marine Fisheries Service[¶]*

Distance sampling is a widely used method for estimating wildlife population abundance. The fact that conventional distance sampling methods are partly design-based constrains the spatial resolution at which animal density can be estimated using these methods. Estimates are usually obtained at survey stratum level. For an endangered species such as the blue whale, it is desirable to estimate density and abundance at a finer spatial scale than stratum. Temporal variation in the spatial structure is also important. We formulate the process generating distance sampling data as a thinned spatial point process and propose model-based inference using a spatial log-Gaussian Cox process. The method adopts a flexible stochastic partial differential equation (SPDE) approach to model spatial structure in density that is not accounted for by explanatory variables, and integrated nested Laplace approximation (INLA) for Bayesian inference. It allows simultaneous fitting of detection and density models and permits prediction of density at an arbitrarily fine scale. We estimate blue whale density in the Eastern Tropical Pacific Ocean from thirteen shipboard surveys conducted over 22 years. We find that higher blue whale density is associated with colder sea surface temperatures in space, and although there is some positive association between density and mean annual temperature, our estimates are consistent with no trend in density across years. Our analysis also indicates that there is substantial spatially structured variation in density that is not explained by available covariates.

REFERENCES

- BAILEY, H. B., MATE, B. R., PALACIOS, D. M., IRVINE, L., BOGRAD, S. J. and COSTA, D. P. (2009). Behavioural estimation of blue whale movements in the Northeast Pacific from state-space model analysis of satellite tracks. *Endanger. Species Res.* **10** 93–106.
- BORCHERS, D. L. and EFFORD, M. G. (2008). Spatially explicit maximum likelihood methods for capture-recapture studies. *Biometrics* **64** 377–385. [MR2432407](#)
- BORCHERS, D. L., BUCKLAND, S. T., GOEDHART, P. W., CLARKE, E. D. and HEDLEY, S. L. (1998). Horvitz–Thompson estimators for double-platform line transect surveys. *Biometrics* **54** 1221–1237.

Key words and phrases. Distance sampling, spatio-temporal modeling, stochastic partial differential equations, INLA, spatial point process.

- BRENNER, S. C. and SCOTT, L. R. (2008). *The Mathematical Theory of Finite Element Methods*, 3rd ed. *Texts in Applied Mathematics* **15**. Springer, New York. [MR2373954](#)
- BUCKLAND, S. T., OEDEKOVEN, C. S. and BORCHERS, D. L. (2016). Model-based distance sampling. *J. Agric. Biol. Environ. Stat.* **21** 58–75. [MR3459294](#)
- BUCKLAND, S. T., ANDERSON, D. R., BURNHAM, K. P., LAAKE, J. L., BORCHERS, D. L. and THOMAS, L. (2001). *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*, 1st ed. Oxford Univ. Press, Oxford.
- BUCKLAND, S. T., REXSTAD, E. A., MARQUES, C. S. and OEDEKOVEN, C. S. (2015). *Distance Sampling: Methods and Applications*. Springer, Berlin.
- CAMELETTI, M., LINDGREN, F., SIMPSON, D. and RUE, H. (2013). Spatio-temporal modeling of particulate matter concentration through the SPDE approach. *ASta Adv. Stat. Anal.* **97** 109–131. [MR3045763](#)
- CONN, P. B., LAAKE, J. L. and JOHNSON, D. S. (2012). A hierarchical modeling framework for multiple observer transect surveys. *PLoS ONE* **7** e42294.
- CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New York. [MR1239641](#)
- DIGGLE, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*, 2nd ed. Hodder Arnold, London.
- DORAZIO, R. M. (2012). Predicting the geographic distribution of a species from presence-only data subject to detection errors. *Biometrics* **68** 1303–1312. [MR3040037](#)
- FARIN, G. E. (2002). *Curves and Surfaces for CAGD: A Practical Guide*, 5th ed. Academic Press, New York.
- FORNEY, K. A., FERGUSON, M. C., BECKER, E. A., FIEDLER, P. C., REDFERN, J. V., BARLOW, J., VILCHIS, I. L. and BALLANCE, L. T. (2012). Habitat-based spatial models of cetacean density in the eastern Pacific Ocean. *Endanger. Species Res.* **16** 113–133.
- GERRODETTE, T. and FORCADA, J. (2005). Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. *Mar. Ecol. Prog. Ser.* **291** 1–21.
- GERRODETTE, T., PERRYMAN, W. and BARLOW, J. (2002). Calibrating group size estimates of dolphins in the Eastern Tropical Pacific Ocean. Administrative Report LJ-02-08. 20 p.
- HANKS, E. M., SCHLIEP, E. M., HOOTEN, M. B. and HOETING, J. A. (2015). Restricted spatial regression in practice: Geostatistical models, confounding, and robustness under model misspecification. *Environmetrics* **26** 243–254. [MR3340961](#)
- HAYES, R. J. and BUCKLAND, S. T. (1983). Radial-distance models for the line-transect method. *Biometrics* **39**(1) 29–42.
- HEDLEY, S. L. and BUCKLAND, S. T. (2004). Spatial models for line transect sampling. *J. Agric. Biol. Environ. Stat.* **9** 181–199.
- HEDLEY, S. L., BUCKLAND, S. T. and BORCHERS, D. L. (2004). Spatial distance sampling models. In *Advanced Distance Sampling* (S. T. Buckland, D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers and L. Thomas, eds.) 48–70. Oxford Univ. Press, Oxford.
- HEFLEY, T. J. and HOOTEN, M. B. (2016). Hierarchical species distribution models. *Curr. Landsc. Ecol. Rep.* **1** 87–97.
- HÖGMANDER, H. (1991). A random field approach to transect counts of wildlife populations. *Biom. J.* **33** 1013–1023. [MR1201875](#)
- HÖGMANDER, H. (1995). *Methods of Spatial Statistics in Monitoring Wildlife Populations*. Univ. Jyväskylä, Jyväskylä.
- ILLIAN, J. B., SØRBYE, S. H. and RUE, H. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *Ann. Appl. Stat.* **6** 1499–1530. [MR3058673](#)
- ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester. [MR2384630](#)
- JOHNSON, D. S., HOOTEN, M. B. and KUHN, C. E. (2013). Estimating animal resource selection from telemetry data using point process models. *J. Anim. Ecol.* **83** 1155–1164.

- JOHNSON, D. S., LAAKE, J. L. and VER HOEF, J. M. (2010). A model-based approach for making ecological inference from distance sampling data. *Biometrics* **66** 310–318. [MR2756719](#)
- JOHNSON, D. S., LAAKE, J. L. and VER HOEF, J. M. (2014). DSpat: Spatial modelling for distance sampling data. R package version 0.1.6.
- KINZEY, D., OLSON, P. and GERRODETTE, T. (2000). Marine Mammal data collection procedures on research ship line-transect surveys by the Southwest Fisheries Science Center. Technical Report National Oceanic and Atmospheric Administration, National Marine Fisheries Service, Southwest Fisheries Science Center.
- KREFT, I. G. G., DE LEEUW, J. and AIKEN, L. S. (1995). The effect of different forms of centering in hierarchicallinear models. *Multivar. Behav. Res.* **30** 1–21.
- LINDGREN, F. and RUE, H. (2015). Bayesian spatial modelling with R-INLA. *J. Stat. Softw.* **63**(19).
- LINDGREN, F., RUE, H. and LINDSTRÖM, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **73** 423–498. [MR2853727](#)
- MILLER, D. L., BURT, M. L., REXSTAD, E. A. and THOMAS, L. (2013). Spatial models for distance sampling data: Recent developments and future directions. *Methods Ecol. Evol.* **4** 1001–1010.
- MILLER, D. L., REXSTAD, E. A., BURT, M. L., BRAVINGTON, M. V. and HEDLEY, S. L. (2014). dsm: Density surface modelling of distance sampling data. R package version 2.2.5.
- MØLLER, J., SYVERSVEEN, A. R. and WAAGEPETERSEN, R. P. (1998). Log Gaussian Cox processes. *Scand. J. Stat.* **25** 451–482. [MR1650019](#)
- MØLLER, J. and WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes. Monographs on Statistics and Applied Probability* **100**. Chapman & Hall/CRC, Boca Raton, FL. [MR2004226](#)
- MØLLER, J. and WAAGEPETERSEN, R. P. (2007). Modern statistics for spatial point processes. *Scand. J. Stat.* **34** 643–684. [MR2392447](#)
- MOORE, J. E. and BARLOW, J. (2011). Bayesian state–space model of fin whale abundance trends from a 1991–2008 time series of line-transect surveys in the California Current. *J. Appl. Ecol.* **48** 1195–1205.
- NIEMI, A. and FERNÁNDEZ, C. (2010). Bayesian spatial point process modeling of line transect data. *J. Agric. Biol. Environ. Stat.* **15** 327–345. [MR2787262](#)
- OEDEKOVEN, C. S., LAAKE, J. L. and SKAUG, H. J. (2015). Distance sampling with a random scale detection function. *Environ. Ecol. Stat.* **22** 725–737. [MR3424098](#)
- OEDEKOVEN, C. S., BUCKLAND, S. T., MACKENZIE, M. L., EVANS, K. O. and BURGER, L. W. (2013). Improving distance sampling: Accounting for covariates and non-independency between sampled sites. *J. Appl. Ecol.* **50** 786–793.
- OEDEKOVEN, C. S., BUCKLAND, S. T., MACKENZIE, M. L., KING, R., EVANS, K. O. and BURGER, L. W. JR. (2014). Bayesian methods for hierarchical distance sampling models. *J. Agric. Biol. Environ. Stat.* **19** 219–239. [MR3257912](#)
- PARDO, M. A., GERRODETTE, T., BEIER, E., GENDRON, D., FORNEY, K. A., CHIVERS, S. J., BARLOW, J. and PALACIOS, D. M. (2015). Inferring cetacean population densities from the absolute dynamic topography of the ocean in a hierarchical Bayesian framework. *PLoS ONE* **10** e0120727. DOI:[doi:10.1371/journal.pone.0120727](#).
- RAMSAY, J. O. (1988). Monotone regression splines in action. *Statist. Sci.* **3** 425–441.
- ROYLE, J. A., DAWSON, D. K. and BATES, S. (2004). Modeling abundance effects in distance sampling. *Ecology* **85** 1591–1597.
- ROYLE, J. A. and DORAZIO, R. M. (2008). *Hierarchical Modelling and Inference in Ecology*. Academic Press, London, UK.
- ROYLE, J. A. and YOUNG, K. V. (2008). A hierarchical model for spatial capture–recapture data. *Ecology* **89** 2281–2289.

- RUE, H. and HELD, L. (2005). *Gaussian Markov Random Fields: Theory and Applications. Monographs on Statistics and Applied Probability* **104**. Chapman & Hall/CRC, Boca Raton, FL. [MR2130347](#)
- RUE, H., MARTINO, S. and CHOPIN, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **71** 319–392. [MR2649602](#)
- SCHMIDT, J. H., RATTENBURY, K. L., LAWLER, J. P. and MACCLUSKIE, M. C. (2012). Using distance sampling and hierarchical models to improve estimates of Dall’s sheep abundance. *J. Wildl. Manag.* **76**(2) 317–327.
- SIMPSON, D., LINDGREN, F. and RUE, H. (2012). Think continuous: Markovian Gaussian models in spatial statistics. *Spat. Stat.* **1** 16–29.
- SIMPSON, D., ILLIAN, J. B., LINDGREN, F., SØRBYE, S. H. and RUE, H. (2016). Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Biometrika* **103** 49–70. [MR3465821](#)
- STOYAN, D. (1982). A remark on the line transect method. *Biom. J.* **24** 191–195. [MR0662894](#)
- STOYAN, D. and GRABARNIK, P. (1991). Second-order characteristics for stochastic structures connected with Gibbs point processes. *Math. Nachr.* **151** 95–100. [MR1121200](#)
- VAN LIESHOUT, M. N. M. (2000). *Markov Point Processes and Their Applications*. Imperial College Press, London. [MR1789230](#)
- WAAGEPETERSEN, R. and SCHWEDER, T. (2006). Likelihood-based inference for clustered line transect data. *J. Agric. Biol. Environ. Stat.* **11**(3) 264–279.
- WIEGAND, T. and MOLONEY, K. A. (2014). *Handbook of Spatial Point-Pattern Analysis in Ecology*. CRC Press, Boca Raton, FL. [MR3155590](#)
- WILLIAMS, R., HEDLEY, S. L., BRANCH, T. A., BRAVINGTON, M. V., ZERBINI, A. N. and FINDLAY, K. P. (2011). Chilean blue whales as a case study to illustrate methods to estimate abundance and evaluate conservation status of rare species. *Conserv. Biol.* **25** 526–535.
- WOLTER, K. and TIMLIN, M. S. (2011). El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext). *Int. J. Climatol.* **31** 1074–1087.
- WOOD, S. N. (2006). *Generalized Additive Models: An Introduction with R*. Chapman & Hall/CRC, Boca Raton, FL. [MR2206355](#)
- YUAN, Y., BACHL, F. E., LINDGREN, F., BORCHERS, D. L., ILLIAN, J. B., BUCKLAND, S. T., RUE, H. and GERRODETTE, T. (2017). Supplement to “Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales.” DOI:[10.1214/17-AOAS1078SUPP](#).

MODELING NODE INCENTIVES IN DIRECTED NETWORKS

BY DEEPAYAN CHAKRABARTI

University of Texas, Austin

Twitter is a popular medium for individuals to gather information and express opinions on topics of interest to them. By understanding who is interested in what topics, we can gauge the public mood, especially during periods of polarization such as elections. However, while the total volume of tweets may be huge, many people tweet rarely, and tweets are short and often noisy. Hence, directly inferring topics from tweets is both complicated and difficult to scale. Instead, the network structure of Twitter (who tweets at whom, who follows whom) can telegraph the interests of Twitter users. We propose the Producer-Consumer Model (PCM) to link latent topical interests of individuals to the directed structure of the network. A key component of PCM is the modeling of incentives of Twitter users. In particular, for a user to attract more followers and become popular, she must strive to be perceived as an expert on some topic. We use this to reduce the parameter space of PCM, greatly increasing its scalability. We apply PCM to track the evolution of Twitter topics during the Italian Elections of 2013, and also to interpret those topics using hashtags. A secondary application of PCM to a citation network of machine learning papers is also shown. Extensive simulations and experiments with large real-world datasets demonstrate the accuracy and scalability of PCM.

REFERENCES

- ADAMIC, L. and ADAR, E. (2003). Friends and neighbors on the Web. *Soc. Netw.* **25** 211–230.
- AIELLO, W., CHUNG, F. and LU, L. (2000). A random graph model for massive graphs. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing* 171–180. ACM, New York. [MR2114530](#)
- AIROLDI, E. M., BLEI, D. M., FIENBERG, S. E. and XING, E. P. (2008). Mixed membership stochastic blockmodels. *J. Mach. Learn. Res.* **9** 1981–2014.
- BLEI, D. M., NG, A. Y. and JORDAN, M. I. (2003). Latent Dirichlet allocation. *J. Mach. Learn. Res.* **3** 993–1022.
- CAIMO, A. and FRIEL, N. (2011). Bayesian inference for exponential random graph models. *Soc. Netw.* **33** 41–55. DOI:[10.1016/j.socnet.2010.09.004](#).
- CALDARELLI, G., CHESSA, A., PAMMOLLI, F., POMPA, G., PULIGA, M., RICCABONI, M. and RIOTTA, G. (2014). A multi-level geographical study of Italian political elections from Twitter data. *PLoS ONE* **9** e95809.
- CARAGEA, C., WU, J., CIOBANU, A., WILLIAMS, K., FERNÁNDEZ-RAMÍREZ, J., CHEN, H.-H., WU, Z. and GILES, L. (2014). CiteSeerX: A scholarly big dataset. In *Proceedings of the 36th European Conference on Information Retrieval (ECIR'14)* 311–322.
- CHAKRABARTI, D. (2017). Supplement to “Modeling node incentives in directed networks.” DOI:[10.1214/17-AOAS1079SUPP](#).

Key words and phrases. Citation networks, social networks, stochastic blockmodel.

- CHAKRABARTI, D. and FALOUTSOS, C. (2006). Graph mining: Laws, generators, and algorithms. *ACM Comput. Surv.* **38** Article No. 2. DOI:[10.1145/1132952.1132954](https://doi.org/10.1145/1132952.1132954).
- CHAKRABARTI, D., ZHAN, Y. and FALOUTSOS, C. (2004). R-MAT: A recursive model for graph mining. In *Proceedings of the 4th SIAM International Conference on Data Mining (SDM'04)* 442–446.
- CHANG, J. (2012). Ilda: Collapsed Gibbs sampling methods for topic models. Available at <https://cran.r-project.org/web/packages/lda/index.html>.
- CHATTERJEE, S., DIACONIS, P. and SLY, A. (2011). Random graphs with a given degree sequence. *Ann. Appl. Probab.* **21** 1400–1435. DOI:[10.1214/10-AAP728](https://doi.org/10.1214/10-AAP728).
- DUIJN, M. A., SNIJDERS, T. A. and ZIJLSTRA, B. J. (2004). P2: A random effects model with covariates for directed graphs. *Stat. Neerl.* **58** 234–254.
- ERDŐS, P. and RÉNYI, A. (1959). On random graphs. I. *Publ. Math. Debrecen* **6** 290–297. [MR0120167](https://doi.org/10.1007/BF0120167)
- FOSDICK, B. K. and HOFF, P. D. (2015). Testing and modeling dependencies between a network and nodal attributes. *J. Amer. Statist. Assoc.* **110** 1047–1056. [MR3420683](https://doi.org/10.1080/01621459.2015.1056838)
- FRANK, O. and STRAUSS, D. (1986). Markov graphs. *J. Amer. Statist. Assoc.* **81** 832–842. [MR0860518](https://doi.org/10.1080/01621459.1986.10477188)
- FU, W., SONG, L. and XING, E. P. (2009). Dynamic mixed membership blockmodel for evolving networks. In *Proceedings of the 26th Annual International Conference on Machine Learning (ICML'09)* 329–336.
- GEHRKE, J., GINSPARG, P. and KLEINBERG, J. (2003). Overview of the 2003 KDD cup. *ACM SIGKDD Explor. Newsl.* **5** 149–151. DOI:[10.1145/980972.980992](https://doi.org/10.1145/980972.980992).
- GILBERT, E. N. (1959). Random graphs. *Ann. Math. Stat.* **30** 1141–1144. [MR0108839](https://doi.org/10.1080/01621459.1959.10501083)
- GOPALAN, P. K. and BLEI, D. M. (2013). Efficient discovery of overlapping communities in massive networks. *Proc. Natl. Acad. Sci. USA* **110** 14534–14539. [MR3105375](https://doi.org/10.1073/pnas.1215375110)
- HANDCOCK, M. S. and JONES, J. H. (2004). Likelihood-based inference for stochastic models of sexual network formation. *Theor. Popul. Biol.* **65** 413–422. DOI:[10.1016/j.tpb.2003.09.006](https://doi.org/10.1016/j.tpb.2003.09.006).
- HOFF, P. D. (2005). Bilinear mixed-effects models for dyadic data. *J. Amer. Statist. Assoc.* **100** 286–295. [MR2156838](https://doi.org/10.1198/016214504000000286)
- HOFF, P. D. (2009). Multiplicative latent factor models for description and prediction of social networks. *Comput. Math. Organ. Theory* **15** 261–272.
- HOFF, P. D., RAFTERY, A. E. and HANDCOCK, M. S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97** 1090–1098. [MR1951262](https://doi.org/10.1198/016214502000000262)
- HOFMANN, T. (1999). Probabilistic latent semantic indexing. In *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR'99)* 50–57.
- HOFMANN, T. (2004). Latent semantic models for collaborative filtering. *ACM Trans. Inf. Syst.* **22** 89–115. DOI:[10.1145/963770.963774](https://doi.org/10.1145/963770.963774).
- HOLLAND, P. W., LASKEY, K. B. and LEINHARDT, S. (1983). Stochastic blockmodels: First steps. *Soc. Netw.* **5** 109–137. [MR0718088](https://doi.org/10.1016/0378-3758(83)90057-X)
- HOLLAND, P. W. and LEINHARDT, S. (1981). An exponential family of probability distributions for directed graphs. *J. Amer. Statist. Assoc.* **76** 33–65. [MR0608176](https://doi.org/10.1080/01621459.1981.10477176)
- HUNTER, D. R. and HANDCOCK, M. S. (2006). Inference in curved exponential family models for networks. *J. Comput. Graph. Statist.* **15** 565–583. [MR2291264](https://doi.org/10.1198/106186006000000264)
- KARRER, B. and NEWMAN, M. E. J. (2011). Stochastic blockmodels and community structure in networks. *Phys. Rev. E* (3) **83** 016107. [MR2788206](https://doi.org/10.1103/PhysRevE.83.016107)
- KATZ, L. (1953). A new status index derived from sociometric analysis. *Psychometrika* **18** 39–43.
- KEMP, C., TENENBAUM, J. B., GRIFFITHS, T. L., YAMADA, T. and UEDA, N. (2006). Learning systems of concepts with an infinite relational model. In *Proceedings of the 21st National Conference on Artificial Intelligence (AAAI'06)* 381–388.

- KOREN, Y. (2008). Factorization meets the neighborhood: A multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'08)* 426–434.
- KRIVITSKY, P. N., HANDCOCK, M. S., RAFTERY, A. E. and HOFF, P. D. (2009). Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models. *Soc. Netw.* **31** 204–213. DOI:[10.1016/j.socnet.2009.04.001](https://doi.org/10.1016/j.socnet.2009.04.001).
- LEI, J. and RINALDO, A. (2015). Consistency of spectral clustering in stochastic block models. *Ann. Statist.* **43** 215–237. [MR3285605](#)
- LESKOVEC, J., CHAKRABARTI, D., KLEINBERG, J., FALOUTSOS, C. and GHAHRAMANI, Z. (2010). Kronecker graphs: An approach to modeling networks. *J. Mach. Learn. Res.* **11** 985–1042. [MR2600637](#)
- MILLER, K. T., GRIFFITHS, T. L. and JORDAN, M. I. (2009). Nonparametric latent feature models for link prediction. In *Advances in Neural Information Processing Systems 22 (NIPS'09)* 1276–1284.
- PALLA, K., KNOWLES, D. A. and GHAHRAMANI, Z. (2012). An infinite latent attribute model for network data. In *Proceedings of the 29th International Conference on Machine Learning (ICML'12)* 1607–1614.
- RAFTERY, A. E., NIU, X., HOFF, P. D. and YEUNG, K. Y. (2012). Fast inference for the latent space network model using a case-control approximate likelihood. *J. Comput. Graph. Statist.* **21** 901–919. [MR3005803](#)
- RICHARDSON, M., AGRAWAL, R. and DOMINGOS, P. M. (2003). Trust management for the semantic web. In *Proceedings of the 2nd International Semantic Web Conference (ISWC'03)* 351–368.
- SALTER-TOWNSHEND, M. and MURPHY, T. B. (2013). Variational Bayesian inference for the latent position cluster model for network data. *Comput. Statist. Data Anal.* **57** 661–671. [MR2981116](#)
- SARKAR, P. and MOORE, A. W. (2010). Fast nearest-neighbor search in disk-resident graphs. In *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'10)* 513–522.
- SHALIZI, C. R. and RINALDO, A. (2013). Consistency under sampling of exponential random graph models. *Ann. Statist.* **41** 508–535. [MR3099112](#)
- SNIJDERS, T. A. B. and NOWICKI, K. (1997). Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classification* **14** 75–100. [MR1449742](#)
- VU, D. Q., HUNTER, D. R. and SCHWEINBERGER, M. (2013). Model-based clustering of large networks. *Ann. Appl. Stat.* **7** 1010–1039. [MR3113499](#)
- WANG, Y. J. and WONG, G. Y. (1987). Stochastic blockmodels for directed graphs. *J. Amer. Statist. Assoc.* **82** 8–19. [MR0883333](#)
- WASSERMAN, S. and PATTISON, P. (1996). Logit models and logistic regressions for social networks. I. An introduction to Markov graphs and p . *Psychometrika* **61** 401–425. [MR1424909](#)
- XU, Z., TRESP, V., YU, K. and KRIEGEL, H. (2006). Infinite hidden relational models. In *Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence (UAI'06)* 544–551.
- YAN, T., LENG, C. and ZHU, J. (2016). Asymptotics in directed exponential random graph models with an increasing bi-degree sequence. *Ann. Statist.* **44** 31–57. [MR3449761](#)
- ZHANG, Y., LEVINA, E. and ZHU, J. (2014). Detecting overlapping communities in networks using spectral methods. ArXiv e-print. Available at <https://arxiv.org/abs/1412.3432>.

AUTOMATIC MATCHING OF BULLET LAND IMPRESSIONS

BY ERIC HARE¹, HEIKE HOFMANN AND ALICIA CARRIQUIRY

Iowa State University

In 2009, the National Academy of Sciences published a report questioning the scientific validity of many forensic methods including firearm examination. Firearm examination is a forensic tool used to help the court determine whether two bullets were fired from the same gun barrel. During the firing process, rifling, manufacturing defects, and impurities in the barrel create striation marks on the bullet. Identifying these striation markings in an attempt to match two bullets is one of the primary goals of firearm examination. We propose an automated framework for the analysis of the 3D surface measurements of bullet land impressions, which transcribes the individual characteristics into a set of features that quantify their similarities. This makes identification of matches easier and allows for a quantification of both matches and matchability of barrels. The automatic matching routine we propose manages to (a) correctly identify land impressions (the surface between two bullet groove impressions) with too much damage to be suitable for comparison, and (b) correctly identify all 10,384 land-to-land matches of the James Hamby study (Hamby, Brundage and Thorpe [*AFTE Journal* **41** (2009) 99–110]).

REFERENCES

- AFTE CRITERIA FOR IDENTIFICATION COMMITTEE (1992). Theory of identification, range striae comparison reports and modified glossary definitions—AFTE criteria for identification committee report. *AFTE Journal* **24** 336–340.
- BACHRACH, B. (2002). Development of a 3D-based automated firearms evidence comparison system. *J. Forensic Sci.* **47** 1253–1264.
- BIASOTTI, A. A. (1959). A statistical study of the individual characteristics of fired bullets. *J. Forensic Sci.* **4** 34–50.
- BREIMAN, L. (2001). Random forests. *Mach. Learn.* **45** 5–32.
- BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A. and STONE, C. J. (1984). *Classification and Regression Trees*. Wadsworth Advanced Books and Software, Belmont, CA. [MR0726392](#)
- CHU, W., SONG, J., VORBURGER, T., YEN, J., BALLOU, S. and BACHRACH, B. (2010). Pilot study of automated bullet signature identification based on topography measurements and correlations. *J. Forensic Sci.* **55** 341–347.
- CHU, W., SONG, J., VORBURGER, T., THOMPSON, R. and SILVER, R. (2011). Selecting valid correlation areas for automated bullet identification system based on striation detection. *J. Res. Natl. Inst. Stand. Technol.* **116** 649.
- CHU, W., THOMPSON, R. M., SONG, J. and VORBURGER, T. V. (2013). Automatic identification of bullet signatures based on consecutive matching striae (CMS) criteria. *Forensic Science International* **231** 137–141.

Key words and phrases. 3D topological surface measurement, data visualization, machine learning, feature importance, cross-correlation function.

- CLEVELAND, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *J. Amer. Statist. Assoc.* **74** 829–836. [MR0556476](#)
- COMMITTEE TO ASSESS THE FEASIBILITY, ACCURACY AND TECHNICAL CAPABILITY OF A NATIONAL BALLISTICS DATABASE (2008). *Ballistic Imaging*. The National Academies Press, Washington, DC.
- DE CEUSTER, J. and DUJARDIN, S. (2015). The reference ballistic imaging database revisited. *Forensic Science International* **248** 82–87.
- DE KINDER, J., TULLENERS, F. and THIEBAUT, H. (2004). Reference ballistic imaging database performance. *Forensic Science International* **140** 207–215.
- GIANNELLI, P. C. (2011). Ballistics evidence under fire. *Criminal Justice* **25** 50–51.
- HAMBY, J. E., BRUNDAGE, D. J. and THORPE, J. W. (2009). The identification of bullets fired from 10 consecutively rifled 9 mm ruger pistol barrels: A research project involving 507 participants from 20 countries. *AFTE Journal* **41** 99–110.
- HARE, E., HOFMANN, H. and CARRIQUIRY, A. (2017). Supplement to “Automatic matching of bullet land impressions.” DOI:[10.1214/17-AOAS1080SUPP](#).
- HOFMANN, H. and HARE, E. (2016). bulletr: Algorithms for Matching Bullet Lands. R package version 0.1.
- HUMMEL, J. (1996). Linked bar charts: Analysing categorical data graphically. *Journal of Computational Statistics* **11** 23–33.
- LIAW, A. and WIENER, M. (2002). Classification and regression by randomForest. *R News* **2** 18–22.
- LOCK, A. B. and MORRIS, M. D. (2013). Significance of angle in the statistical comparison of forensic tool marks. *Technometrics* **55** 548–561. [MR3176558](#)
- MA, L., SONG, J., WHITENTON, E., ZHENG, A., VORBURGER, T. and ZHOU, J. (2004). NIST bullet signature measurement system for RM (Reference Material) 8240 standard bullets. *Journal of Forensic Sciences* **49** 649–659.
- MILBORROW, S. (2015). rpart.plot: Plot “rpart” Models: An Enhanced Version of “plot.rpart.” R package version 1.5.3.
- MILLER, J. (1998). Criteria for identification of toolmarks. *AFTE Journal* **30** 15–61.
- NATIONAL RESEARCH COUNCIL (2009). *Strengthening Forensic Science in the United States: A Path Forward*. The National Academies Press, Washington, DC.
- NICHOLS, R. G. (1997). Firearm and toolmark identification criteria: A review of the literature. *Journal of Forensic Sciences* **42** 466–474.
- NICHOLS, R. G. (2003a). Consecutive matching striations (CMS): Its definition, study and application in the discipline of firearms and tool mark identification. *AFTE Journal* **35** 298–306.
- NICHOLS, R. G. (2003b). Firearm and toolmark identification criteria: A review of the literature: Part II. *Journal of Forensic Sciences* **48** 318–327.
- OPENFMC (2014). x3pr: Read/Write functionality for X3P surface metrology format. R package version 1.0.
- PETRACO, N. and CHAN, H. (2012). *Application of Machine Learning to Toolmarks: Statistically Based Methods for Impression Pattern Comparisons*. Bibliographisches Institut AG, Mannheim, Germany.
- R CORE TEAM (2016). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.
- RIVA, F. and CHAMPOD, C. (2014). Automatic comparison and evaluation of impressions left by a firearm on fired cartridge cases. *Journal of Forensic Sciences* **59** 637–647.
- ROBERGE, D. and BEAUCHAMP, A. (2006). The use of BulletTrax-3D in a study of consecutively manufactured barrels. *AFTE Journal* **38** 166–172.
- THERNEAU, T., ATKINSON, B. and RIPLEY, B. (2015). rpart: Recursive Partitioning and Regression Trees. R package version 4.1-10.
- TUKEY, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.

- VORBURGER, T. V., SONG, J. F., CHU, W., MA, L., BUI, S. H., ZHENG, A. and RENEGAR, T. B. (2011). Applications of cross-correlation functions. *Wear* **271** 529–533.
- WICKHAM, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer, New York.
- XIE, Y. (2015). *Dynamic Documents with R and Knitr*, 2nd ed. Chapman & Hall/CRC, Boca Raton, FL.

ESTIMATING THE NUMBER OF CASUALTIES IN THE AMERICAN INDIAN WAR: A BAYESIAN ANALYSIS USING THE POWER LAW DISTRIBUTION

BY COLIN S. GILLESPIE

Newcastle University

The American Indian War lasted over one hundred years, and is a major event in the history of North America. As expected, since the war commenced in late eighteenth century, casualty records surrounding this conflict contain numerous sources of error, such as rounding and counting. Additionally, while major battles such as the Battle of the Little Bighorn were recorded, many smaller skirmishes were completely omitted from the records. Over the last few decades, it has been observed that the number of casualties in major conflicts follows a power law distribution. This paper places this observation within the Bayesian paradigm, enabling modelling of different error sources, allowing inferences to be made about the overall casualty numbers in the American Indian War.

REFERENCES

- ABRAMOWITZ, M. and STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards Applied Mathematics Series **55**. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC. [MR0167642](#)
- ALSTOTT, J., BULLMORE, E. and PLENZ, D. (2014). Powerlaw: A Python package for analysis of heavy-tailed distributions. *PLoS ONE* **9** e85777.
- AXELROD, A. (1993). *Chronicle of the Indian Wars: From Colonial Times to Wounded Knee*. Konecky & Konecky, Old Saybrook, CT.
- BELL, M. J., GILLESPIE, C. S., SWAN, D. and LORD, P. (2012). An approach to describing and analysing bulk biological annotation quality: A case study using UniProtKB. *Bioinformatics* **28** i562–i568.
- BOHORQUEZ, J. C., GOURLEY, S., DIXON, A. R., SPAGAT, M. and JOHNSON, N. F. (2009). Common ecology quantifies human insurgency. *Nature* **462** 911–914.
- BREIMAN, L., STONE, C. J. and KOOPERBERG, C. (1990). Robust confidence bounds for extreme upper quantiles. *J. Stat. Comput. Simul.* **37** 127–149. [MR1082452](#)
- CEDERMAN, L. E. (2003). Modeling the size of wars: From billiard balls to sandpiles. *Am. Polit. Sci. Rev.* **97** 135–150.
- CLAUSET, A., SHALIZI, C. R. and NEWMAN, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Rev.* **51** 661–703. [MR2563829](#)
- CLAUSET, A. and WOODARD, R. (2013). Estimating the historical and future probabilities of large terrorist events. *Ann. Appl. Stat.* **7** 1838–1865. [MR3161698](#)
- CLAUSET, A., YOUNG, M. and GLEDITSCH, K. S. (2007). On the frequency of severe terrorist events. *J. Confl. Resolut.* **51** 58–87.

Key words and phrases. Power law, Bayesian, missing data, long tailed.

- CLODFELTER, M. (2008). *Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures*, 1494–2007. McFarland, Jefferson, NC.
- CRAWFORD, F. W., WEISS, R. E. and SUCHARD, M. A. (2015). Sex, lies and self-reported counts: Bayesian mixture models for heaping in longitudinal count data via birth–death processes. *Ann. Appl. Stat.* **9** 572–596. [MR3371326](#)
- DEKKERS, A. L. M. and DE HAAN, L. (1993). Optimal choice of sample fraction in extreme-value estimation. *J. Multivariate Anal.* **47** 173–195. [MR1247373](#)
- DEWEZ, T. J. B., ROHMER, J., REGARD, V. and CNUDDÉ, C. (2013). Probabilistic coastal cliff collapse hazard from repeated terrestrial laser surveys: Case study from Mesnil Val. In *Proceedings 12th International Coastal Symposium (Plymouth, England)*, *Journal of Coastal Research, Special Issue* 65. 0749–0208.
- DREES, H. and KAUFMANN, E. (1998). Selecting the optimal sample fraction in univariate extreme value estimation. *Stochastic Process. Appl.* **75** 149–172. [MR1632189](#)
- EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. *Monographs on Statistics and Applied Probability* **57**. Chapman & Hall, New York. [MR1270903](#)
- FRIEDMAN, J. A. (2015). Using power laws to estimate conflict size. *J. Confl. Resolut.* **59** 1216–1241.
- GELLER, D. S. and SINGER, J. D. (1998). *Nations at War: A Scientific Study of International Conflict* **58**. Cambridge Univ. Press, Cambridge.
- GILLESPIE, C. S. (2015). Fitting heavy tailed distributions: The powerLaw package. *J. Stat. Softw.* **64** 1–16.
- GILLESPIE, C. S. (2017). Supplement to “Estimating the number of casualties in the American Indian war: A Bayesian analysis using the power law distribution.” DOI:[10.1214/17-AOAS1082SUPP](#).
- NEWMAN, M. E. J. (2005). Power laws, Pareto distributions and Zipf’s law. *Contemp. Phys.* **46** 323–351.
- PINKER, S. (2011). *The Better Angels of Our Nature: The Decline of Violence in History and Its Causes*. Penguin, London.
- R CORE TEAM (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- RICHARDSON, L. F. (1948). Variation of the frequency of fatal quarrels with magnitude. *J. Amer. Statist. Assoc.* **43** 523–546.
- RICHARDSON, L. F. (1960). *Statistics of Deadly Quarrels*. Boxwood Press, Pittsburgh, PA.
- STUMPF, M. P. H. and PORTER, M. A. (2012). Critical truths about power laws. *Science* **335** 665–666. [MR2932329](#)
- VIRKAR, Y. and CLAUSET, A. (2014). Power-law distributions in binned empirical data. *Ann. Appl. Stat.* **8** 89–119. [MR3191984](#)
- VUONG, Q. H. (1989). Likelihood ratio tests for model selection and nonnested hypotheses. *Econometrica* **57** 307–333. [MR0996939](#)

STATISTICAL DOWNSCALING FOR FUTURE EXTREME WAVE HEIGHTS IN THE NORTH SEA

BY ROSS TOWE^{*,1}, EMMA EASTOE^{*}, JONATHAN TAWN^{*}
 AND PHILIP JONATHAN[†]

Lancaster University^{} and Shell Research Limited, UK[†]*

For safe offshore operations, accurate knowledge of the extreme oceanographic conditions is required. We develop a multi-step statistical downscaling algorithm using data from a low resolution global climate model (GCM) and local-scale hindcast data to make predictions of the extreme wave climate in the next 50-year period at locations in the North Sea. The GCM is unable to produce wave data accurately so instead we use its 3-hourly wind speed and direction data. By exploiting the relationships between wind characteristics and wave heights, a downscaling approach is developed to relate the large and local-scale data sets, and hence future changes in wind characteristics can be translated into changes in extreme wave distributions. We assess the performance of the methods using within sample testing and apply the method to derive future design levels over the northern North Sea.

REFERENCES

- BECHLER, A., VRAC, M. and BEL, L. (2015). A spatial hybrid approach for downscaling of extreme precipitation fields. *J. Geophys. Res., Atmos.* **120** 4534–4550.
- BELLOUIN, N., COLLINS, W., CULVERWELL, I., HALLORAN, P., HARDIMAN, S., HINTON, T., JONES, C., McDONALD, R., McLAREN, A., O’CONNOR, F. et al. (2011). The Hadgem2 family of Met Office unified model climate configurations. *Geosci. Model Dev.* **4** 723–757.
- BIERBOOMS, W. (2003). Wind and wave conditions. DOWEC (Dutch Offshore Wind Energy Converter Project) 47.
- CAIRES, S., SWAIL, V. and WANG, X. (2006). Projection and analysis of extreme wave climate. *Climate* **19** 5581–5605.
- CASAS-PRAT, M., WANG, X. and SIERRA, J. (2014). A physical-based statistical method for modeling ocean wave heights. *Ocean Model.* **73** 59–75.
- COLES, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer, London. [MR1932132](#)
- COLES, S. and TAWN, J. (1994). Statistical methods for multivariate extremes: An application to structural design (with discussion). *J. R. Stat. Soc. Ser. C. Appl. Stat.* **43** 1–48.
- DAVISON, A. C., PADOAN, S. A. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186. [MR2963980](#)
- DAVISON, A. C. and SMITH, R. L. (1990). Models for exceedances over high thresholds. *J. Roy. Statist. Soc. Ser. B* **52** 393–442. [MR1086795](#)
- DE HAAN, L. and DE RONDE, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes* **1** 7–45. [MR1652944](#)

Key words and phrases. Climate change, covariate modelling, environmental modelling, extreme value theory, sea waves, statistical downscaling.

- DE HAAN, L. and FERREIRA, A. (2010). *Extreme Value Theory: An Introduction*. Springer, New York. [MR2234156](#)
- EASTERLING, D., MEEHL, G., PARMESAN, C., CHANGNON, S., KARL, T. and MEARN, L. (2000). Climate extremes: Observations, modeling, and impacts. *Science* **289** 2068–2074.
- EASTOE, E. F. and TAWN, J. A. (2009). Modelling non-stationary extremes with application to surface level ozone. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **58** 25–45. [MR2662232](#)
- EDWARDS, P. (2010). *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming*. MIT Press, Cambridge, MA.
- HEFFERNAN, J. E. and RESNICK, S. I. (2007). Limit laws for random vectors with an extreme component. *Ann. Appl. Probab.* **17** 537–571. [MR2308335](#)
- HEFFERNAN, J. E. and TAWN, J. A. (2004). A conditional approach for multivariate extreme values. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **66** 497–546. With discussions and reply by the authors. [MR2088289](#)
- JONATHAN, P. and EWANS, K. (2007). The effect of directionality on extreme wave design criteria. *Ocean Eng.* **34** 1977–1994.
- JONATHAN, P., EWANS, K. and RANDELL, D. (2014). Non-stationary conditional extremes of northern North Sea storm characteristics. *Environmetrics* **25** 172–188. [MR3200308](#)
- KALLACHE, M., VRAC, M., NAVEAU, P. and MICHELANGELI, P.-A. (2011). Nonstationary probabilistic downscaling of extreme precipitation. *J. Geophys. Res.* **116** D05113.
- KEEF, C., PAPASTATHOPOULOS, I. and TAWN, J. A. (2013). Estimation of the conditional distribution of a multivariate variable given that one of its components is large: Additional constraints for the Heffernan and Tawn model. *J. Multivariate Anal.* **115** 396–404. [MR3004566](#)
- KINSMAN, B. (2012). *Wind Waves: Their Generation and Propagation on the Ocean Surface*. Dover, New York.
- KYSELÝ, J., PICEK, J. and BERANOVÁ, R. (2010). Estimating extremes in climate change simulations using the peaks-over-threshold method with a non-stationary threshold. *Glob. Planet. Change* **72** 55–68.
- LEDFORD, A. W. and TAWN, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* **83** 169–187. [MR1399163](#)
- LIU, Y. and TAWN, J. A. (2014). Self-consistent estimation of conditional multivariate extreme value distributions. *J. Multivariate Anal.* **127** 19–35. [MR3188874](#)
- LOFSTED, R. (2014). *Cases in Climate Change Policy: Political Reality in the European Union*. Routledge, London.
- LUGRIN, T., DAVISON, A. C. and TAWN, J. A. (2016). Bayesian uncertainty management in temporal dependence of extremes. *Extremes* **19** 491–515. [MR3535964](#)
- MARAUN, D., WIDMANN, M., GUTIERREZ, J., KOTLARSKI, S., CHANDLER, R., HERTIG, E., WIBIG, J., HUTH, R. and WILCKE, R. (2015). Value: A framework to validate downscaling approaches for climate change studies. *Earth's Future* **3** 1–14.
- MICHELANGELI, P.-A., VRAC, M. and LOUKOS, H. (2009). Probabilistic downscaling approaches: Application to wind cumulative distribution functions. *Geophys. Res. Lett.* **36** 1–6.
- NELSEN, R. B. (2006). *An Introduction to Copulas*, 2nd ed. Springer, New York. [MR2197664](#)
- NORTHROP, P. J. and JONATHAN, P. (2011). Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics* **22** 799–809. [MR2861046](#)
- REISTAD, M., BREIVIK, Ø., HAAKENSTAD, H., AARNES, O., FUREVIK, B. and BIDLOT, J. (2011). A high-resolution hindcast of wind and waves for the North Sea, the Norwegian Sea, and the Barents Sea. *J. Geophys. Res.* **116**. C05019.
- ROBINSON, P. M. and TAWN, J. (1997). Statistics for extreme sea currents. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **46** 183–205.
- SILVERMAN, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London. [MR0848134](#)

- SMITH, R. L. and WEISSMAN, I. (1994). Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B* **56** 515–528. [MR1278224](#)
- SOLOMON, S., QUIN, D., MANNING, M., CHEN, Z., MARQUIS, M., AVERYT, K., TIGNOR, M. and MILLER, H., eds. (2007). *Climate Change 2007—The Physical Science Basis: Working Group I Contribution to the Fourth Assessment Report of the IPCC*. Cambridge Univ. Press, Cambridge.
- TOWE, R. (2015). Modelling the Extreme Wave Climate of the North Sea. Ph.D. thesis, Univ. Lancaster, Lancaster.
- TOWE, R., EASTOE, E., TAWN, J., WU, Y. and JONATHAN, P. (2013). The extremal dependence of storm severity, wind speed and surface level pressure in the northern North Sea. In *ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering*. American Society of Mechanical Engineers, New York.
- VANEM, E., HUSEBY, A. and NATVIG, B. (2012). A Bayesian hierarchical spatio-temporal model for significant wave height in the North Atlantic. *Stoch. Environ. Res. Risk Assess.* **26** 609–632.
- VON MISES, R. (1964). *Mathematical Theory of Probability and Statistics*. Academic Press, New York. [MR0178486](#)
- WADSWORTH, J. L. and TAWN, J. A. (2012). Dependence modelling for spatial extremes. *Biometrika* **99** 253–272. [MR2931252](#)
- WANG, J. and ZHANG, X. (2010). Downscaling and projection of winter extreme daily precipitation over North America. *J. Climate* **21** 923–937.
- WILBY, R., CHARLES, S., ZORITA, E., TIMBAL, B., WHETTON, P. and MEARN, L. (2004). Guidelines for use of climate scenarios developed from statistical downscaling methods. Supporting material of the Intergovernmental Panel on Climate Change.

FOCUSING ON REGIONS OF INTEREST IN FORECAST EVALUATION

BY HAJO HOLZMANN AND BERNHARD KLAR

Philipps-Universität Marburg and Karlsruher Institut für Technologie (KIT)

Often, interest in forecast evaluation focuses on certain regions of the whole potential range of the outcome, and forecasts should mainly be ranked according to their performance within these regions. A prime example is risk management, which relies on forecasts of risk measures such as the value-at-risk or the expected shortfall, and hence requires appropriate loss distribution forecasts in the tails. Further examples include weather forecasts with a focus on extreme conditions, or forecasts of environmental variables such as ozone with a focus on concentration levels with adverse health effects.

In this paper, we show how weighted scoring rules can be used to this end, and in particular that they allow to rank several potentially misspecified forecasts objectively with the region of interest in mind. This is demonstrated in various simulation scenarios. We introduce desirable properties of weighted scoring rules and present general construction principles based on conditional densities or distributions and on scoring rules for probability forecasts. In our empirical application to log-return time series, all forecasts seem to be slightly misspecified, as is often unavoidable in practice, and no method performs best overall. However, using weighted scoring functions the best method for predicting losses can be identified, which is hence the method of choice for the purpose of risk management.

REFERENCES

- AMISANO, G. and GIACOMINI, R. (2007). Comparing density forecasts via weighted likelihood ratio tests. *J. Bus. Econom. Statist.* **25** 177–190. [MR2367773](#)
- BANK OF ENGLAND (2017). Monetary Policy Framework. Retrieved from. <http://www.bankofengland.co.uk/monetarypolicy/Pages/framework/framework.aspx>.
- BILLI, R. M. (2017). A note on nominal GDP targeting and the zero lower bound. *Macroecon. Dyn.* **11** 2138–2157.
- CASATI, B., WILSON, L. J., STEPHENSON, D. B., NURMI, P., GHELLI, A., POCERNICH, M., DAMRATH, U., EBERT, E. E., BROWN, B. G. and MASON, S. (2008). Forecast verification: Current status and future directions. *Meteorol. Appl.* **15** 3–18.
- DAWID, A. P. (1984). Statistical theory: The prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–292. [MR0763811](#)
- DE NICOLÒ, G. and LUCCHETTA, M. (2017). Forecasting tail risks. *J. Appl. Econometrics* **32** 159–170. [MR3611065](#)
- DIEBOLD, F. X. (2015). Comparing predictive accuracy, twenty years later: A personal perspective on the use and abuse of Diebold-Mariano tests. *J. Bus. Econom. Statist.* **33** 1–24.

Key words and phrases. Financial time series, predictive performance, probabilistic forecast, locally proper weighted scoring rule, misspecified forecast, rare and extreme events, risk management.

- DIEBOLD, F. X., GUNTHER, T. A. and TAY, A. (1998). Evaluating density forecasts: With applications to financial risk management. *Internat. Econom. Rev.* **39** 863–883.
- DIEBOLD, F. X. and MARIANO, R. S. (2015). Comparing predictive accuracy. *J. Bus. Econom. Statist.* **13** 253–263.
- DIKS, C., PANCHENKO, V. and VAN DIJK, D. (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *J. Econometrics* **163** 215–230. [MR2812867](#)
- ELLIOTT, G. and TIMMERMAN, A. (2016). Forecasting in economics and finance. *Ann. Rev. Econ.* **8** 81–110.
- GARÍN, J., LESTER, R. and SIMS, E. (2016). On the desirability of nominal GDP targeting. *J. Econom. Dynam. Control* **69** 21–44. [MR3521425](#)
- GHALANOS, A. (2014). rugarch: Univariate GARCH models. R package version 1.3-5.
- GIACOMINI, R. and WHITE, H. (2006). Tests of conditional predictive ability. *Econometrica* **74** 1545–1578. [MR2268409](#)
- GIESBERGEN, B. (2017). China: How realistic is the government's growth target? Economic Report. Rabobank. Retrieved at. <https://economics.rabobank.com/publications/2017/march/china-how-realistic-is-the-governments-growth-target>.
- GNEITING, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc.* **106** 746–762. [MR2847988](#)
- GNEITING, T. and RAFTERY, A. E. (2005). Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation. *Mon. Weather Rev.* **133** 1098–1118.
- GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. [MR2345548](#)
- GNEITING, T. and RANJAN, R. (2011). Comparing density forecasts using threshold- and quantile-weighted scoring rules. *J. Bus. Econom. Statist.* **29** 411–422. [MR2848512](#)
- HAIDEN, T., MAGNUSSEN, L. and RICHARDSON, D. (2014). Statistical evaluation of ECMWF extreme wind forecasts. In *European Centre for Medium—Range Weather Forecasts Newsletter, Spring 2014*.
- HOLZMANN, H. and KLAR, B. (2016). Weighted scoring rules and hypothesis testing. Available at [arXiv:1611.07345v2](#).
- HOLZMANN, H. and KLAR, B. (2017a). Supplement to “Focusing on regions of interest in forecast evaluation.” DOI:[10.1214/17-AOAS1088SUPP](#).
- HOLZMANN, H. and KLAR, B. (2017b). Discussion of “Elicitability and backtesting: Perspectives for banking regulation.” *Ann. Appl. Stat.* **11** 1875–1882.
- HYVÄRINEN, A. (2005). Estimation of non-normalized statistical models by score matching. *J. Mach. Learn. Res.* **6** 695–709. [MR2249836](#)
- KRÜGER, F., LERCH, S., THORARINSDOTTIR, T. L. and GNEITING, T. (2017). Probabilistic forecasting and comparative model assessment based on Markov chain Monte Carlo output. Preprint. Available at [arXiv:1608.06802](#).
- LERCH, S., THORARINSDOTTIR, T. L., RAVAZZOLO, F. and GNEITING, T. (2017). Forecaster's dilemma: Extreme events and forecast evaluation. *Statist. Sci.* **32** 106–127. [MR3634309](#)
- MATHESSON, J. E. and WINKLER, R. L. (1976). Scoring rules for continuous probability distributions. *Manage. Sci.* **22** 1087–1096.
- MCNEIL, A. J., FREY, R. and EMBRECHTS, P. (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton Univ. Press, Princeton, NJ. [MR2175089](#)
- NOLDE, N. and ZIEGEL, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *Ann. Appl. Statist.* To appear.
- OPSCHOOR, A., VAN DIJK, D. and VAN DER WEL, M. (2017). Combining density forecasts using focused scoring rules. *J. Appl. Econometrics* **2017** 1–16. DOI:[10.1002/jae.2575](#).
- PATTON, A. (2017). Evaluating and comparing possibly misspecified forecasts. Working paper.
- PELENIS, J. (2014). Weighted scoring rules for comparison of density forecasts on subsets of interest. Preprint. Available at <https://sites.google.com/site/jpelenis/>.

- PISONI, E., FARINA, M., PAGANI, G. and PIRODDI, L. (2011). *Environmental over-Threshold Event Forecasting Using NARX Models*. Preprints of the 18th IFAC World Congress, Milano (Italy).
- R CORE TEAM (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- THORARINSDOTTIR, T. L. and GNEITING, T. (2010). Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. *J. Roy. Statist. Soc. Ser. A* **173** 371–388. [MR2751882](#)

The Annals of Applied Statistics

Next Issues

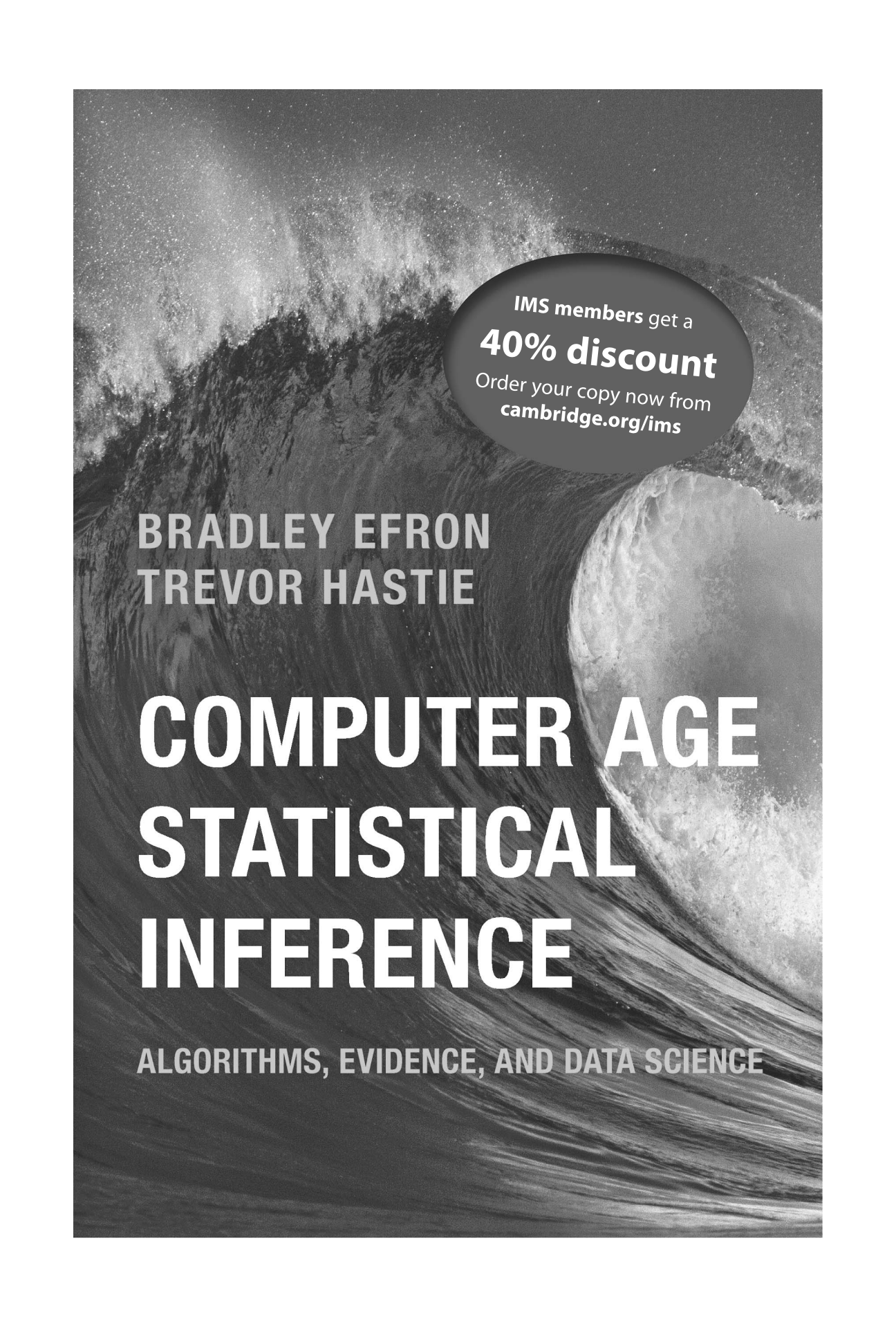
- Estimating cellular pathways from an ensemble of heterogeneous data sources
ALEXANDER M. FRANKS, FLORIAN MARKOWETZ AND EDOARDO M. AIROLDI
- Integrative exploration of large high-dimensional datasets
CHRISTOPHER PARDY, SALLY GALBRAITH AND SUSAN R. WILSON
- A spatio-temporal modeling framework for weather radar image data in tropical Southeast Asia
XIAO LIU, VIK GOPAL AND JAYANT KALAGNANAM
- Adjustment of non-confounding covariates in case-control genetic association studies
HONG ZHANG, NILANJAN CHATTERJEE, DANIEL RADER AND JINBO CHEN
- How Gaussian mixture models might miss detecting factors that impact growth patterns
BRIANNA HEGGESETH AND NICHOLAS JEWELL
- Extreme value modelling of water-related insurance claims CHRISTIAN ROHRBECK,
EMMA F. EASTOE, ARNOLDO FRIGESSI AND JONATHAN A. TAWN
- A testing based approach to the discovery of differentially correlated variable sets
KELLY BODWIN, KAI ZHANG AND ANDREW B. NOBEL
- Biomarker assessment and combination with differential covariate effects and an unknown gold
standard, with an application to Alzheimer's disease ZHEYU WANG AND
XIAO-HUA ANDREW ZHOU
- A phylogenetic scan test on Dirichlet-tree multinomial model for microbiome data
YUNFAN TANG, LI MA AND DAN LIVIU NICOLAE
- Robust dependence modeling for high-dimensional covariance matrices with financial
applications ZHE ZHU AND ROY E. WELSCH
- Time-varying extreme value dependence with applications to leading European stock markets
MIGUEL DE CARVALHO, DANIELA CASTRO AND JENNIFFER WADSWORTH
- Phenomenological forecasting of disease incidence using heteroskedastic Gaussian processes:
A dengue case study LEAH R. JOHNSON, ROBERT B. GRAMACY, JEREMY COHEN,
ERIN A. MORDECAI, COURTNEY MURDOCK, JASON ROHR, SADIE J. RYAN,
ANNA M. STEWART-IBARRA AND DANIEL WEIKEL
- Spatial capture-recapture with partial identity: An application to camera traps
BEN AUGUSTINE, J. ANDREW ROYLE, MARCELLA J. KELLY,
CHRISTOPHER B. SATTER, ROBERT S. ALONSO, ERIN E. BOYDSTON
AND KEVIN R. CROOKS
- Automated threshold selection for extreme value analysis via ordered goodness-of-fit tests with
adjustment for false discovery rate BRIAN BADER, JUN YAN
AND XUEBIN ZHANG
- An empirical study of the maximal and total information coefficients and leading measures of
dependence DAVID N. RESHEF, YAKIR A. RESHEF, PARDIS C. SABETI
AND MICHAEL MITZENMACHER
- Design of vaccine trials during outbreaks with and without a delayed vaccination comparator
NATALIE EXNER DEAN, M. ELIZABETH HALLORAN AND IRA M. LONGINI
- Two-level structural sparsity regularization for identifying lattices and defects in noisy images
XIN LI, ALEX BELIANINOV, ONDREJ DYCK, STEPHEN JESSE AND CHIWOO PARK
- Network-based feature screening with applications to genome data
MENGYUN WU, LIPING ZHU AND XINGDONG FENG
- Multivariate integer-valued time series with flexible autocovariances and their application to
major hurricane counts JAMES LIVSEY, ROBERT LUND,
STEFANOS KECHAGIAS AND VLADAS PIPIRAS
- Stochastic simulation of predictive space-time scenarios of wind speed using observations and
physical model outputs JULIE BESSAC, EMIL CONSTANTINESCU
AND MIHAI ANITESCU
- MSIQ: Joint modeling of multiple RNA-seq samples for accurate isoform quantification
WEI VIVIAN LI, ANQI ZHAO, SHIHUA ZHANG AND JINGYI JESSICA LI

Continued

The Annals of Applied Statistics

Next Issues—Continued

- Covariate balancing propensity score for a continuous treatment: Application to the efficacy of political advertisements CHRISTIAN FONG, CHAD HAZLETT AND KOSUKE IMAI
- Kernel-penalized regression for analysis of microbiome data TIMOTHY W. RANDOLPH, SEN ZHAO, WADE COPELAND, MEREDITH HULLAR AND ALI SHOJAIE
- Powerful test based on conditional effects for genome-wide screening
YAOWU LIU AND JUN XIE
- A multi-resolution model for non-Gaussian random fields on a sphere with application to ionospheric electrostatic potentials . . MINJIE FAN, DEBASHIS PAUL, THOMAS C. M. LEE AND TOMOKO MATSUO
- Reducing storage of global wind ensembles with stochastic generators
JAEHONG JEONG, STEFANO CASTRUCCIO, PAOLA CRIPPA, MARC G. GENTON
- Fast inference of individual admixture coefficients using geographic data
KEVIN CAYE, FLORA JAY, OLIVIER MICHEL AND OLIVIER FRANÇOIS
- Statistical shape analysis of simplified neuronal trees
ADAM DUNCAN, ERIC KLASSEN AND ANUJ SRIVASTAVA
- Simultaneous modelling of movement, measurement error and observer dependence in double-observer distance sampling surveys . . . PAUL B. CONN AND RAY T. ALISAUSKAS
- Covariate matching methods for testing and quantifying wind turbine updates
YEI EUN SHIN, YU DING AND JIANHUA HUANG
- A unified statistical framework for single cell and bulk RNA sequencing data
LINGXUE ZHU, JING LEI, BERNIE DEVLIN AND KATHRYN ROEDER
- Non-stationary modeling of tail dependence of two subjects' concentration
KSHITIJ SHARMA, VALÉRIE CHAVEZ-DEMOULIN AND PIERRE DILLENBOURG
- Modeling and estimation for self-exciting spatio-temporal models of terrorist activity
NICHOLAS J. CLARK AND PHILIP M. DIXON
- A spatially-varying stochastic differential equation model for animal movement
JAMES C. RUSSELL, EPHRAIM M. HANKS, MURALI HARAN AND DAVID HUGHES
- Empirical assessment of programs to promote collaboration: A network model approach
KATHERINE R. MCLAUGHLIN AND JOSHUA D. EMBREE
- Torus principal component analysis with applications to RNA structure
BENJAMIN ELTZNER, STEPHAN HUCKEMANN AND KANTI V. MARDIA
- TPRM: Tensor partition regression models with applications in imaging biomarker detection
HONGTU ZHU, MICHELLE F. MIRANDA AND JOSEPH G. IBRAHIM
- Complex-valued time series modeling for improved activation detection in fMRI studies
DANIEL WRIGHT ADRIAN, RANJAN MAITRA AND DANIEL B. ROWE
- Optimal multilevel matching using network flows: An application to a summer reading intervention SAMUEL PIMENTEL, LINDSAY PAGE, MATTHEW LENARD AND LUKE J. KEELE
- Topological data analysis of single-trial electroencephalographic signals
YUAN WANG, HERNANDO OMBAO AND MOO K. CHUNG
- Joint significance tests for mediation effects of socioeconomic adversity on adiposity via epigenetics YEN-TSUNG HUANG
- Adaptive-weight burden test for associations between quantitative traits and genotype data with complex correlations XIAOWEI WU, TING GUAN, DAJIANG J. LIU, LUIS G. LEON NOVELO AND DIPANKAR BANDYOPADHYAY
- Bayesian aggregation of average data: An application in drug development
SEBASTIAN WEBER, ANDREW GELMAN, DANIEL LEE, MICHAEL BETANCOURT, AKI VEHTARI AND AMY RACINE-POON
- BayCount: A Bayesian decomposition method for inferring tumor heterogeneity using RNA-Seq counts FANGZHENG XIE, MINGYUAN ZHOU AND YANXUN XU



IMS members get a
40% discount

Order your copy now from
cambridge.org/ims

BRADLEY EFRON
TREVOR HASTIE

COMPUTER AGE STATISTICAL INFERENCE

ALGORITHMS, EVIDENCE, AND DATA SCIENCE