

THE ANNALS *of* APPLIED STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles

- Editorial SUSAN M. PADDOCK 1921
- Spatial accessibility of pediatric primary healthcare: Measurement and inference
MALLORY NOBLES, NICOLETA SERBAN AND JULIE SWANN 1922
- Discussion LAURA A. HATFIELD 1947
- Discussion AMELIA M. HAVILAND 1952
- Rejoinder LANCE A. WALLER 1956
- Rejoinder MALLORY NOBLES, NICOLETA SERBAN AND JULIE SWANN 1961
- Reconstructing past temperatures from natural proxies and estimated climate forcings
using short- and long-memory models
LUIS BARBOZA, BO LI, MARTIN P. TINGLEY AND FREDERI G. VIENS 1966
- Equivalence testing for functional data with an application to comparing pulmonary
function devices COLIN B. FOGARTY AND DYLAN S. SMALL 2002
- A multiple filter test for the detection of rate changes in renewal processes with varying
variance MICHAEL MESSER, MARIETTA KIRCHNER, JULIA SCHIEMANN,
JOCHEN ROEPER, RALPH NEININGER AND GABY SCHNEIDER 2027
- Bayesian protein structure alignment ... ABEL RODRIGUEZ AND SCOTT C. SCHMIDLER 2068
- Isolation in the construction of natural experiments
JOSÉ R. ZUBIZARRETA, DYLAN S. SMALL AND PAUL R. ROSENBAUM 2096
- Global estimation of child mortality using a Bayesian B-spline Bias-reduction model
LEONTINE ALKEMA AND JIN ROU NEW 2122
- Imputation of truncated p -values for meta-analysis methods and its genomic application
SHAOWU TANG, YING DING, ETIENNE SIBILLE, JEFFREY S. MOGIL,
WILLIAM R. LARIVIERE AND GEORGE C. TSENG 2150
- Longitudinal high-dimensional principal components analysis with application
to diffusion tensor imaging of multiple sclerosis
VADIM ZIPUNNIKOV, SONJA GREVEN, HAOCHANG SHOU, BRIAN S. CAFFO,
DANIEL S. REICH AND CIPRIAN M. CRAINICEANU 2175
- A novel spectral method for inferring general diploid selection from time series genetic
data MATTHIAS STEINRÜCKEN, ANAND BHASKAR AND YUN S. SONG 2203
- Spatio-temporal modelling of extreme storms THEODOROS ECONOMOU,
DAVID B. STEPHENSON AND CHRISTOPHER A. T. FERRO 2223
- Voronoi residual analysis of spatial point process models with applications to California
earthquake forecasts ANDREW BRAY, KA WONG,
CHRISTOPHER D. BARR AND FREDERIC PAIK SCHOENBERG 2247
- Longitudinal Mixed Membership trajectory models for disability survey data
DANIEL MANRIQUE-VALLIER 2268
- A fast algorithm for detecting gene–gene interactions in genome-wide association
studies JIAHAN LI, WEI ZHONG, RUNZE LI AND RONGLING WU 2292

continued

THE ANNALS *of* APPLIED STATISTICS

AN OFFICIAL JOURNAL OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

Articles—*Continued from front cover*

- A permutational-splitting sample procedure to quantify expert opinion on clusters of chemical compounds using high-dimensional data
ELASMA MILANZI, ARIEL ALONSO, CHRISTOPHE BUYCK,
GEERT MOLENBERGHS AND LUC BIJNENS 2319
- The use of covariates and random effects in evaluating predictive biomarkers under a potential outcome framework
ZHIWEI ZHANG, LEI NIE, GUOXING SOON AND AIYI LIU 2336
- Evaluating epoetin dosing strategies using observational longitudinal data
CECILIA A. COTTON AND PATRICK J. HEAGERTY 2356
- Synthesising evidence to estimate pandemic (2009) A/H1N1 influenza severity in 2009–2011 ANNE M. PRESANIS, RICHARD G. PEBODY, PAUL J. BIRRELL,
BRIAN D. M. TOM, HELEN K. GREEN, HAYLEY DURNALL,
DOUGLAS FLEMING AND DANIELA DE ANGELIS 2378
- Detecting duplicates in a homicide registry using a Bayesian partitioning approach
MAURICIO SADINLE 2404
- Functional response additive model estimation with online virtual stock markets
YINGYING FAN, NATASHA FOUTZ, GARETH M. JAMES AND WOLFGANG JANK 2435
- Probit models for capture–recapture data subject to imperfect detection, individual heterogeneity and misidentification BRETT T. MCCLINTOCK,
LARISSA L. BAILEY, BRIAN P. DREHER AND WILLIAM A. LINK 2461
- Do debit cards increase household spending? Evidence from a semiparametric causal analysis of a survey ANDREA MERCATANTI AND FAN LI 2485
- Reduced-rank spatio-temporal modeling of air pollution concentrations in the Multi-Ethnic Study of Atherosclerosis and Air Pollution
CASEY OLIVES, LIANNE SHEPPARD, JOHAN LINDSTRÖM, PAUL D. SAMPSON,
JOEL D. KAUFMAN AND ADAM A. SZPIRO 2509
- Nonparametric inference in a stereological model with oriented cylinders applied to dual phase steel K. Š. MCGARRITY, J. SIETSMA AND G. JONGBLOED 2538

Acknowledgment of priority

- Usage of the Lambert W function in statistics GEORG M. GOERG 2567

EDITORIAL: SPATIAL ACCESSIBILITY OF PEDIATRIC PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE

BY SUSAN M. PADDOCK

RAND Corporation

REFERENCES

- ADAY, L. A. and ANDERSEN, R. (1974). A framework for the study of access to medical care. *Health Serv. Res.* **9** 208–220.
- NOBLES, M., SERBAN, N. and SWANN, J. (2014). Spatial accessibility of pediatric primary health-care: Measurement and inference. *Ann. Appl. Stat.* **8** 1922–1946.

SPATIAL ACCESSIBILITY OF PEDIATRIC PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE^{1,2}

BY MALLORY NOBLES, NICOLETA SERBAN AND JULIE SWANN³

Georgia Institute of Technology

Although improving financial access is in the spotlight of the current U.S. health policy agenda, this alone does not address universal and comprehensive healthcare. Affordability is one barrier to healthcare, but others such as availability and accessibility, together defined as *spatial accessibility*, are equally important. In this paper, we develop a measurement and modeling framework that can be used to infer the impact of policy changes on disparities in spatial accessibility within and across different population groups. The underlying model for measuring spatial accessibility is optimization-based and accounts for constraints in the healthcare delivery system. The measurement method is complemented by statistical modeling and inference on the impact of various potential contributing factors to disparities in spatial accessibility. The emphasis of this study is on children's accessibility to primary care pediatricians, piloted for the state of Georgia. We focus on disparities in accessibility between and within two populations: children insured by Medicaid and other children. We find that disparities in spatial accessibility to pediatric primary care in Georgia are significant, and resistant to many policy interventions, suggesting the need for major changes to the structure of Georgia's pediatric healthcare provider network.

REFERENCES

- ASSUNCAO, R. M. (2003). Space varying coefficient models for small area data. *Environmetrics* **14** 453–473.
- BERMAN, S., DOLINS, J. et al. (2002). Factors that influence the willingness of private primary care pediatricians to accept more Medicaid patients. *Pediatrics* **110** 239–248.
- BRAVEMAN, P. and GRUSKIN, S. (2003). Defining equity in health. *Journal of Epidemiology and Community Health* **57** 254–258.
- BUJA, A., HASTIE, T. and TIBSHIRANI, R. (1989). Linear smoothers and additive models. *Ann. Statist.* **17** 453–555. [MR0994249](#)
- DIGGLE, P. (1985). A Kernel method for smoothing point process data. *J. R. Stat. Soc. Ser. C Appl. Stat.* **34** 138–147.
- DRAKE, B. and RANK, M. R. (2009). The racial divide among American children in poverty: Re-assessing the importance of neighborhood. *Children and Youth Services Review* **31**.
- FAN, J. and ZHANG, J.-T. (2000). Two-step estimation of functional linear models with applications to longitudinal data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62** 303–322. [MR1749541](#)
- FISCELLA, K. and WILLIAMS, D. (2004). Health disparities based on socioeconomic inequities: Implications for urban health care. *Academic Medicine* **79** 1139–1147.

Key words and phrases. Healthcare access, optimization model, pediatric healthcare, spatial accessibility, spatial-varying coefficient model.

- FLEURBAEY, M. and SCHOKKAERT, E. (2009). Unfair inequalities in health and health care. *Journal of Health Economics* **28** 73–90.
- GELFAND, A. E., KIM, H.-J., SIRMANS, C. F. and BANERJEE, S. (2003). Spatial modeling with spatially varying coefficient processes. *J. Amer. Statist. Assoc.* **98** 387–396. [MR1995715](#)
- GUAGLIARDO, M. F. (2004). Spatial accessibility of primary care: Concepts, methods and challenges. *International Journal of Health Geographics* **3** 1–13.
- GUAGLIARDO, M. F., RONZIO, C. R., CHEUNG, I., CHACKO, E. and JOSEPH, J. G., (2004). Physician accessibility: An urban case study of pediatric providers. *Health Place* **10** 273–283.
- HAMBIDGE, S., EMSERMANN, C. et al. (2007). Disparities in pediatric preventive care in the United States. *Archives of Pediatric and Adolescent Medicine* **161** 30–36.
- HASTIE, T. J. and TIBSHIRANI, R. J. (1990). *Generalized Additive Models. Monographs on Statistics and Applied Probability* **43**. Chapman & Hall, London. [MR1082147](#)
- HASTIE, T. and TIBSHIRANI, R. (1993). Varying-coefficient models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **55** 757–796. [MR1229881](#)
- HOOVER, D. R., RICE, J. A., WU, C. O. and YANG, L.-P. (1998). Nonparametric smoothing estimates of time-varying coefficient models with longitudinal data. *Biometrika* **85** 809–822. [MR1666699](#)
- JIANG, H. (2010). Statistical computation and inference for functional data analysis. Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA.
- JOSEPH, A. and PHILLIPS, D. (1984). *Accessibility and Utilization: Geographical Perspectives on Health Care Delivery*. Harper and Row, New York.
- KHAN, A. (1992). An integrated approach to measuring potential spatial access to healthcare services. *Socio-Economic Planning Sciences* **26** 275–287.
- KIDS COUNT NATIONAL INDICATORS (2010). Data Center, Baltimore MD.
- KRIVOBOKOVA, T., KNEIB, T. and CLAESKENS, G. (2010). Simultaneous confidence bands for penalized spline estimators. *J. Amer. Statist. Assoc.* **105** 852–863. [MR2724866](#)
- LI, Y. and RUPPERT, D. (2008). On the asymptotics of penalized splines. *Biometrika* **95** 415–436. [MR2521591](#)
- LUO, W. and QI, Y. (2009). An enhanced two-step floating catchment area (E2SFCA) method for measuring spatial accessibility to primary care physicians. *Health Place* **15** 1100–1107.
- MARCIN, J., ELLIS, J. et al. (2004). Using telemedicine to provide pediatric subspecialty care to children with special health care needs in an underserved rural community. *Pediatrics* **113** 1–6.
- MARMOT, M., FRIEL, S. et al. (2008). Closing the gap in a generation: Health equity through action on the social determinants of health. *Lancet* **372** 1661–1669.
- MARSH, M. T. and SCHILLING, D. A. (1994). Equity measurement in facility location analysis: A review and framework. *European J. Oper. Res.* **74** 1–17.
- MCGRAIL, M. R. and HUMPHREYS, J. S. (2009). The index of rural access: An innovative integrated approach for measuring primary care access. *BMC Health Services Research* **9**. DOI:10.1186/1472-6963-9-124.
- NOBLES, M., SERBAN, N. and SWANN, J. (2014). Supplement to “Spatial accessibility of pediatric primary healthcare: Measurement and inference.” DOI:10.1214/14-AOAS728SUPP.
- ODOKI, J., KERALI, H. and SANTORINI, F. (2001). An integrated model for quantifying accessibility-benefits in developing countries. *Transportation Research Part A: Policy and Practice* **35** 601–623.
- PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076. [MR0143282](#)
- PEARCE, L. D. (2002). Integrating survey and ethnographic models for systematic anomalous case analysis. *Sociological Methodology* **32** 103–132.
- PENCHANSKY, R. and THOMAS, J. W. (1981). The concept of access. *Med. Care* **19** 127–140.
- PERLOFF, J., KLETKE, P. and FOSSETT, J. (1995). Which physicians limit their Medicaid participation, and why. *Health Services Research* **30** 7–25.

- REARDON, S. F., MATTHEWS, S. A., O'SULLIVAN, D., LEE, B. A., FIREBAUGH, G., FARRELL, C. R. and BISCHOFF, K. (2008). The geographic scale of metropolitan racial segregation. *Demography* **45** 489–514.
- RUPPERT, D., WAND, M. P. and CARROLL, R. J. (2003). *Semiparametric Regression. Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge Univ. Press, Cambridge. MR1998720
- SERBAN, N. (2011). A space–time varying coefficient model: The equity of service accessibility. *Ann. Appl. Stat.* **5** 2024–2051. MR2884930
- TALEN, E. and ANSELIN, L. (1998). Assessing spatial equity: An evaluation of measures of accessibility to public playgrounds. *Environment and Planning B: Planning and Design* **30** 595.
- TIBSHIRANI, R. and KNIGHT, K. (1999). The covariance inflation criterion for adaptive model selection. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **61** 529–546. MR1707859
- WALKER, L. O., STERLING, B. S., HOKE, M. M. and DEARDEN, K. A. (2007). Applying the concept of positive deviance to public health data: A tool for reducing health disparities. *Public Health Nurs.* **24** 571–576.
- WALLER, L. A., ZHU, L., GOTWAY, C. A., GORMAN, D. M. and GRUENEWALD, P. J. (2007). Quantifying geographic variations in associations between alcohol distribution and violence: A comparison of geographically weighted regression and spatially varying coefficient models. *Stoch. Environ. Res. Risk Assess.* **21** 573–588. MR2380676
- WANG, F. and LUO, W. (2005). Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. *Health Place* **11** 131–146.
- WILLIAMS, D. R. and MOHAMMED, S. A. (2009). Discrimination and racial disparities in health: Evidence and needed research. *Journal of Behavioral Medicine* **32** 20–47.
- WU, H. and LIANG, H. (2004). Backfitting random varying-coefficient models with time-dependent smoothing covariates. *Scand. J. Stat.* **31** 320–330.

DISCUSSION OF “SPATIAL ACCESSIBILITY OF PEDIATRIC PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE”

BY LAURA A. HATFIELD

Harvard Medical School

REFERENCES

- CARNATHAN, M. (2012). Atlanta Metropolitan Area: High population growth, employment losses adding to income woes. Urban Institute, Washington, DC. Available at <http://metrotrends.org/spotlight/atlanta.cfm>.
- CHEUNG, P. T., WILER, J. L., LOWE, R. A. and GINDE, A. A. (2012). National study of barriers to timely primary care and emergency department utilization among Medicaid beneficiaries. *Ann. Emerg. Med.* **60** 4–10.
- KANGOVI, S., BARG, F. K., CARTER, T., LONG, J. A., SHANNON, R. and GRANDE, D. (2013). Understanding why patients of low socioeconomic status prefer hospitals over ambulatory care. *Health Aff. (Millwood)* **32** 1196–1203.
- NOBELS, M., SERBAN, N. and SWANN, J. (2014). Spatial accessibility of pediatric primary health-care: Measurement and inference. *Ann. Appl. Stat.* **8** 1922–1946.
- SMULOWITZ, P. B., O’MALLEY, J., YANG, X. and LANDON, B. E. (2014). Increased use of emergency department after health care reform in Massachusetts. *Ann. Emerg. Med.* **64** 107–115.

**DISCUSSION OF “SPATIAL ACCESSIBILITY OF PEDIATRIC
PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE”**

BY AMELIA M. HAVILAND

Carnegie Mellon University

REFERENCES

- EFRON, B. (2013). Estimation and accuracy after model selection. *StatLib*. Stanford Univ., Stanford, CA.
- NATIONAL RESEARCH COUNCIL (2012). *Deterrence and the Death Penalty*. National Academies Press, Washington, D.C.

DISCUSSION OF “SPATIAL ACCESSIBILITY OF PEDIATRIC PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE”

BY LANCE A. WALLER

Emory University

REFERENCES

- GASTWIRTH, J. L., NAYAK, T. K. and WANG, J.-L. (1989). Statistical properties of measures of between-group income differentials. *J. Econometrics* **42** 5–19. [MR1029927](#)
- HANDCOCK, M. S. and MORRIS, M. (1999). *Relative Distribution Methods in the Social Sciences*. Springer, New York. [MR1705294](#)
- INSTITUTE OF MEDICINE (2012). *Geographic Adjustment in Medicare Payment: Phase II: Implications for Access, Quality, and Efficiency*. National Academies Press, Washington, DC.
- LIU, F. (2001). *Environmental Justice Analysis: Theories, Methods, and Practice*. CRC Press, Boca Raton.
- MARSH, M. T. and SCHILLING, D. A. (1994). Equity measurement in facility location analysis: A review and framework. *European Journal of Operations Research* **74** 1–17.
- NATIONAL RESEARCH COUNCIL (2007). *Using the American Community Survey: Benefits and Challenges*. National Academies Press, Washington, DC.
- NOBELS, M., SERBAN, N. and SWANN, J. (2014). Spatial accessibility of pediatric primary healthcare: Measurement and inference. *Ann. Appl. Stat.* **8** 1922–1946.
- SPIELMAN, S. E., FOLCH, D. and NAGLE, N. (2014). Patterns and causes of uncertainty in the American Community Survey. *Applied Geography* **46** 147–157.
- TASSONE, E. C., WALLER, L. A. and CASPER, M. L. (2009). Small-area racial disparity in stroke mortality: An application of Bayesian spatial hierarchical modeling. *Epidemiology* **20** 234–241.
- WALLER, L. A., LOUIS, T. A. and CARLIN, B. P. (1999). Environmental justice and statistical summaries of differences in exposure distributions. *J. Expo. Anal. Environ. Epidemiol.* **9** 56–65.

REJOINDER: “SPATIAL ACCESSIBILITY OF PEDIATRIC PRIMARY HEALTHCARE: MEASUREMENT AND INFERENCE”

BY MALLORY NOBLES, NICOLETA SERBAN AND JULIE SWANN

Georgia Institute of Technology

REFERENCES

- BARTON, R. R., NELSON, B. L. and XIE, W. (2014). Quantifying input uncertainty via simulation confidence intervals. *INFORMS J. Comput.* **26** 74–87. [MR3171248](#)
- DASKIN, M. S. and DEAN, L. K. (2004). Location of health care facilities. In *Handbook of OR/MS in Health Care: A Handbook of Methods and Applications* (M. B. F. Sainfort and W. Pierskalla, eds.) 43–76. Kluwer Academic, Norwell, MA.
- GRABER, A., CARTER, M. and VERTER, V. (2014). Multi criteria optimization modeling for primary care facility network design the regulator’s perspective. Preprint.
- JACOBSON, S. H., HALL, S. N. and SWISHER, J. R. (2006). Discrete-event simulation of health care systems. In *Patient Flow: Reducing Delay in Healthcare Delivery* (R. W. Hall, ed.) 211–252. Springer, New York.
- ROUSE, W. B. and SERBAN, N. (2014). *Understanding and Managing the Complexity of Healthcare*. MIT Press, Cambridge, MA.
- SMITH, R. C. (2013). *Uncertainty Quantification: Theory, Implementation, and Applications*. *Computational Science & Engineering* **12**. SIAM, Philadelphia, PA. [MR3155184](#)
- SNYDER, L. V. (2006). Facility location under uncertainty: A review. *IIE Transactions* **38** 547–564.

RECONSTRUCTING PAST TEMPERATURES FROM NATURAL PROXIES AND ESTIMATED CLIMATE FORCINGS USING SHORT- AND LONG-MEMORY MODELS

BY LUIS BARBOZA^{*,1}, BO LI^{†,2}, MARTIN P. TINGLEY^{‡,§,3}
AND FREDERICK G. VIENS^{¶,4}

Universidad de Costa Rica^{*}, *University of Illinois at Urbana–Champaign*[†],
Pennsylvania State University[‡], *Harvard University*[§] and *Purdue University*[¶]

We produce new reconstructions of Northern Hemisphere annually averaged temperature anomalies back to 1000 AD, and explore the effects of including external climate forcings within the reconstruction and of accounting for short-memory and long-memory features. Our reconstructions are based on two linear models, with the first linking the latent temperature series to three main external forcings (solar irradiance, greenhouse gas concentration and volcanism), and the second linking the observed temperature proxy data (tree rings, sediment record, ice cores, etc.) to the unobserved temperature series. Uncertainty is captured with additive noise, and a rigorous statistical investigation of the correlation structure in the regression errors is conducted through systematic comparisons between reconstructions that assume no memory, short-memory autoregressive models, and long-memory fractional Gaussian noise models.

We use Bayesian estimation to fit the model parameters and to perform separate reconstructions of land-only and combined land-and-marine temperature anomalies. For model formulations that include forcings, both exploratory and Bayesian data analysis provide evidence against models with no memory. Model assessments indicate that models with no memory underestimate uncertainty. However, no single line of evidence is sufficient to favor short-memory models over long-memory ones, or to favor the opposite choice. When forcings are not included, the long-memory models appear to be necessary. While including external climate forcings substantially improves the reconstruction, accurate reconstructions that exclude these forcings are vital for testing the fidelity of climate models used for future projections.

Finally, we use posterior samples of model parameters to arrive at an estimate of the transient climate response to greenhouse gas forcings of 2.5°C (95% credible interval of [2.16, 2.92]°C), which is on the high end of, but consistent with, the expert-assessment-based uncertainties given in the recent Fifth Assessment Report of the IPCC.

REFERENCES

- AMMANN, C. M., JOOS, F., SCHIMMEL, D. S., OTTO-BLIESNER, B. L. and TOMAS, R. A. (2007). Solar influence on climate during the past millennium: Results from transient simulations with the NCAR climate system model. *Proc. Natl. Acad. Sci. USA* **104** 3713–3718.

Key words and phrases. External forcings, long-memory, proxies, temperature reconstruction.

- BARBOZA, L., LI, B., TINGLEY, M. P. and VIENS, F. G. (2014). Supplement to “Reconstructing past temperatures from natural proxies and estimated climate forcings using short- and long-memory models.” DOI:10.1214/14-AOAS785SUPP.
- BENTH, F. E. and ŠALTYTĖ-BENTH, J. (2005). Stochastic modelling of temperature variations with a view towards weather derivatives. *Appl. Math. Finance* **12** 53–85.
- BERAN, J. (1992). A goodness-of-fit test for time series with long range dependence. *J. Roy. Statist. Soc. Ser. B* **54** 749–760. MR1185220
- BERAN, J. (1994). *Statistics for Long-Memory Processes. Monographs on Statistics and Applied Probability* **61**. Chapman & Hall, New York. MR1304490
- BERLINER, L. M., WIKLE, C. K. and CRESSIE, N. (2000). Long-lead prediction of pacific SSTs via Bayesian dynamic modeling. *J. Climate* **13** 3953–3968.
- BINDOFF, N. L., STOTT, P. A., ACHUTARAO, K. M., ALLEN, M. R., GILLETT, N., GUTZLER, D., HANSINGO, K., HEGERL, G., HU, Y., JAIN, S., MOKHOV, I. I., OVERLAND, J., PERLWITZ, J., SEBBARI, R. and ZHANG, X. (2013). Detection and attribution of climate change: From global to regional. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.). Cambridge Univ. Press, Cambridge.
- BRIFFA, K. R., JONES, P. D., SCHWEINGRUBER, F. H. and OSBORN, T. J. (1998). Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years. *Nature* **393** 450–455.
- BRODY, D. C., SYROKA, J. and ZERVOS, M. (2002). Dynamical pricing of weather derivatives. *Quant. Finance* **2** 189–198. MR1913206
- BROHAN, P., KENNEDY, J. J., HARRIS, I., TETT, S. F. B. and JONES, P. D. (2006). Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.* **111** D12106+.
- BROOKS, S. P. and GELMAN, A. (1998). General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Statist.* **7** 434–455. MR1665662
- BROOKS, S. P. and ROBERTS, G. O. (1997). Assessing convergence of Markov chain Monte Carlo algorithms. *Statist. Comput.* **8** 319–335.
- CHRISTIANSEN, B. (2011). Reconstructing the NH mean temperature: Can underestimation of trends and variability be avoided? *J. Climate* **24** 674–692.
- CHRISTIANSEN, B., SCHMITH, T. and THEJLL, P. (2009). A surrogate ensemble study of climate reconstruction methods: Stochasticity and robustness. *J. Climate* **22** 951–976.
- CHRONOPOULOU, A. and VIENS, F. G. (2010). Hurst index estimation for self-similar processes with long-memory. In *Recent Development in Stochastic Dynamics and Stochastic Analysis* (J. Duan, S. Luo and C. Wang, eds.). *Interdiscip. Math. Sci.* **8** 91–117. World Sci. Publ., Hackensack, NJ. MR2807815
- CHRONOPOULOU, A. and VIENS, F. G. (2012). Estimation and pricing under long-memory stochastic volatility. *Ann. Finance* **8** 379–403. MR2922802
- CHRONOPOULOU, A., VIENS, F. G. and TUDOR, C. A. (2009). Variations and Hurst index estimation for a Rosenblatt process using longer filters. *Electron. J. Stat.* **3** 1393–1435. MR2578831
- COLLINS, M. R., KNUTTI, R., ARBLASTER, J., DUFRESNE, J.-L., FICHEFET, T., FRIEDLINGSTEIN, P., GAO, X., GUTOWSKI, W. J., JOHNS, T., KRINNER, G., SHONGWE, M., TEBALDI, C., WEAVER, A. J. and WEHNER, M. (2013). Long-term climate change: Projections, commitments and irreversibility. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.). Cambridge Univ. Press, Cambridge.
- COOK, E. R., BRIFFA, K. R. and JONES, P. D. (1994). Spatial regression methods in dendroclimatology: A review and comparison of two techniques. *Int. J. Climatol.* **14** 379–402.

- COWLES, M. K. and CARLIN, B. P. (1996). Markov chain Monte Carlo convergence diagnostics: A comparative review. *J. Amer. Statist. Assoc.* **91** 883–904. [MR1395755](#)
- CROWLEY, T. J., CRISTE, T. A. and SMITH, N. R. (1993). Reassessment of Crete (Greenland) ice core acidity/volcanism link to climate change. *Geophys. Res. Lett.* **20** 209–212.
- CROWLEY, T. and KIM, K. (1993). Towards development of a strategy for determining the origin of decadal–centennial scale climate variability. *Quat. Sci. Rev.* **12** 375–385.
- CROWLEY, T. J. and KIM, K.-Y. (1996). Comparison of proxy records of climate change and solar forcing. *Geophys. Res. Lett.* **23** 359–362.
- DAVIES, R. B. and HARTE, D. S. (1987). Tests for Hurst effect. *Biometrika* **74** 95–101. [MR0885922](#)
- DEMPSTER, A. P., LAIRD, N. M. and RUBIN, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **39** 1–38. [MR0501537](#)
- FLATO, G., MAROTZKE, J., ABIODUN, B., BRACONNOT, P., CHOU, S. C., COLLINS, W., COX, P., DRIOUÉCH, F., EMORI, S., EYRING, V., FOREST, C., GLECKLER, P., GUILYARDI, E., JAKOB, C., KATSOV, V., REASON, C. and RUMMUKAINEN, M. (2013). Evaluation of climate models. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.). Cambridge Univ. Press, Cambridge.
- GELMAN, A. and RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–472.
- GENZ, A. and BRETZ, F. (2009). *Computation of Multivariate Normal and t Probabilities. Lecture Notes in Statistics* **195**. Springer, Dordrecht. [MR2840595](#)
- GILLET, N. P., ARORA, V. K., FLATO, G. M., SCINOCCA, J. F. and VON SALZEN, K. (2012). Improved constraints on 21st-century warming derived using 160 years of temperature observations. *Geophys. Res. Lett.* **39** L01704.
- GILLET, N. P., ARORA, V. K., MATTHEWS, D. and ALLEN, M. R. (2013). Constraining the ratio of global warming to cumulative CO₂ emissions using CMIP5 simulations. *J. Climate* **26** 6844–6858.
- 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 RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. [MR2345548](#)
- GNEITING, T. and SCHLATHER, M. (2004). Stochastic models that separate fractal dimension and the Hurst effect. *SIAM Rev.* **46** 269–282 (electronic). [MR2114455](#)
- GSCHLÖSSL, S. and CZADO, C. (2007). Spatial modelling of claim frequency and claim size in non-life insurance. *Scand. Actuar. J.* **3** 202–225. [MR2361126](#)
- HARRIS, G. R., SEXTON, D. M. H., BOOTH, B. B. B., COLLINS, M. and MURPHY, J. M. (2013). Probabilistic projections of transient climate change. *Clim. Dyn.* **40** 2937–2972.
- HASLETT, J., WHILEY, M., BHATTACHARYA, S., SALTER-TOWNSHEND, M., WILSON, S. P., ALLEN, J. R. M., HUNTLEY, B. and MITCHELL, F. J. G. (2006). Bayesian palaeoclimate reconstruction. *J. Roy. Statist. Soc. Ser. A* **169** 395–438. [MR2236914](#)
- HEGERL, G. C., ZWIERS, F. W., BRACONNOT, P., GILLET, N. P., LUO, Y., MARENGO ORSINI, J. A., NICHOLLS, N., PENNER, J. E. and STOTT, P. A. (2007). Understanding and attributing climate change. In *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller, eds.). Cambridge Univ. Press, Cambridge.
- HUYBERS, P. and CURRY, W. (2006). Links between annual, Milankovitch and continuum temperature variability. *Nature* **441** 329–332.
- IMBERS, J., LOPEZ, A., HUNTINGFORD, C. and ALLEN, M. (2014). Sensitivity of climate change detection and attribution to the characterization of internal variability. *J. Climate* **27** 3477–3491.

- JACKSON, C. (2011). Multi-state models for panel data: The msm package for R. *J. Stat. Softw.* **38** 1–28.
- JONES, P. D., BRIFFA, K. R., OSBORN, T. J., LOUGH, J. M., VAN OMMEN, T. D., VINTHER, B. M., LUTERBACHER, J., WAHL, E. R., ZWIERS, F. W., MANN, M. E., SCHMIDT, G. A., AMMANN, C. M., BUCKLEY, B. M., COBB, K. M., ESPER, J., GOOSSE, H., GRAHAM, N., JANSEN, E., KIEFER, T., KULL, C., KÜTTEL, M., MOSLEY-THOMPSON, E., OVERPECK, J. T., RIEDWYL, N., SCHULZ, M., TUDHOPE, A. W., VILLALBA, R., WANNER, H., WOLFF, E. and XOPLAKI, E. (2009). High-resolution palaeoclimatology of the last millennium: A review of current status and future prospects. *Holocene* **19** 3–49.
- JUCKES, M. N., ALLEN, M. R., BRIFFA, K. R., ESPER, J., HEGERL, G. C., MOBERG, A., OSBORN, T. J., WEBER, S. L. and ZORITA, E. (2006). Millennial temperature reconstruction inter-comparison and evaluation. *Clim. Past Discuss.* **2** 1001–1049.
- KALMAN, R. E. and BUCY, R. S. (1961). New results in linear filtering and prediction theory. *Trans. ASME Ser. D, J. Basic Engrg.* **83** 95–108. [MR0234760](#)
- KAUFMAN, D. S., SCHNEIDER, D. P., MCKAY, N. P., AMMANN, C. M., BRADLEY, R. S., BRIFFA, K. R., MILLER, G. H., OTTO-BLIESNER, B. L., OVERPECK, J. T., VINTHER, B. M. and OTHERS (2009). Recent warming reverses long-term arctic cooling. *Science* **325** 1236–1239.
- KOLMOGOROFF, A. N. (1940). Wienerische Spiralen und einige andere interessante Kurven im Hilbertschen Raum. *C. R. (Doklady) Acad. Sci. USSR (N.S.)* **26** 115–118. [MR0003441](#)
- LANDRUM, L., OTTO-BLIESNER, B. L., WAHL, E. R., CONLEY, A., LAWRENCE, P. J., ROSEN-BLOOM, N. and TENG, H. (2013). Last millennium climate and its variability in CCSM4. *J. Climate* **26** 1085–1111.
- LEAN, J., BEER, J. and BRADLEY, R. (1995). Reconstruction of solar irradiance since 1610: Implications for climate change. *Geophys. Res. Lett.* **22** 3195–3198.
- LEE, T. C. K., ZWIERS, F. W. and TSAO, M. (2008). Evaluation of proxy-based millennial reconstruction methods. *Clim. Dyn.* **31** 263–281.
- LI, B., NYCHKA, D. W. and AMMANN, C. M. (2010). The value of multiproxy reconstruction of past climate. *J. Amer. Statist. Assoc.* **105** 883–895. [MR2752583](#)
- LOSO, M. G. (2009). Summer temperatures during the Medieval Warm Period and Little Ice Age inferred from varved proglacial lake sediments in southern Alaska. *J. Paleolimnol.* **41** 117–128.
- LUTERBACHER, J., DIETRICH, D., XOPLAKI, E., GROSJEAN, M. and WANNER, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303** 1499–1503.
- MANDELBROT, B. (1965). Une classe processus stochastiques homothétiques à soi; application à la loi climatologique H. E. Hurst. *C. R. Math. Acad. Sci. Paris* **260** 3274–3277. [MR0176521](#)
- MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* **10** 422–437. [MR0242239](#)
- MANN, M. E., BRADLEY, R. S. and HUGHES, M. K. (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392** 779–787.
- MANN, M. E., BRADLEY, R. S. and HUGHES, M. K. (1999). Northern Hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophys. Res. Lett.* **26** 759–762.
- MANN, M. E., RUTHERFORD, S., WAHL, E. and AMMANN, C. (2005). Testing the fidelity of methods used in proxy-based reconstructions of past climate. *J. Climate* **18** 4097–4107.
- MANN, M. E., RUTHERFORD, S., WAHL, E. and AMMANN, C. (2007). Robustness of proxy-based climate field reconstruction methods. *J. Geophys. Res.* **112** D12109+.
- MANN, M. E., ZHANG, Z., HUGHES, M. K., BRADLEY, R. S., MILLER, S. K., RUTHERFORD, S. and NI, F. (2008a). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proc. Natl. Acad. Sci. USA* **105** 13252–13257.

- MANN, M. E., ZHANG, Z., HUGHES, M. K., BRADLEY, R. S., MILLER, S. K., RUTHERFORD, S. and NI, F. (2008b). Supporting information for “Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia.” *Proc. Natl. Acad. Sci. USA* **105** 13252–13257.
- MANN, M. E., ZHANG, Z., RUTHERFORD, S., BRADLEY, R. S., HUGHES, M. K., SHINDELL, D., AMMANN, C. M., FALUVEGI, G. and NI, F. (2009). Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science* **326** 1256–1260.
- MARTIN, A. D., QUINN, K. M. and PARK, J. H. (2011). MCMCpack: Markov chain Monte Carlo in R. *J. Stat. Softw.* **42** 1–21.
- MASSON-DELMOTTE, V., SCHULZ, M., ABE-OUCHI, A., BEER, J., GANOPOLSKI, A., ROUCO, J. F. G., JANSEN, E., LAMBECK, K., LUTERBACHER, J., NAISH, T., OSBORN, T., OTTO-BLIESNER, B., QUINN, T., RAMESH, R., ROJAS, M., SHAO, X. and TIMMERMANN, A. (2013). Information from paleoclimate archives. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (T. F. Stocker, D. Qin, G. K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.). Cambridge Univ. Press, Cambridge.
- MCLEOD, A. I., YU, H. and KROUGLY, Z. L. (2007). Algorithms for linear time series analysis: With R package. *Journal of Statistical Software* **23** 1–26.
- MCSHANE, B. B. and WYNER, A. J. (2011). A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable? *Ann. Appl. Stat.* **5** 5–44.
- MOBERG, A., SONECHKIN, D. M., HOLMGREN, K., DATSENKO, N. M., KARLÉN, W. and LAURITZEN, S.-E. (2005). Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature* **433** 613–617.
- MURPHY, A. and IZZELDIN, M. (2009). Bootstrapping long memory tests: Some Monte Carlo results. *Comput. Statist. Data Anal.* **53** 2325–2334. [MR2665831](#)
- NRC (2006). *Surface Temperature Reconstructions for the Last 2000 Years*. The National Academies Press, Washington, DC.
- OTTO, A., OTTO, F. E. L., BOUCHER, O., CHURCH, J., HEGERL, G., FORSTER, P. M., GILLETT, N. P., GREGORY, J., JOHNSON, G. C., KNUTTI, R., LEWIS, N., LOHMANN, U., MAROTZKE, J., MYHRE, G., SHINDELL, D., STEVENS, B. and ALLEN, M. R. (2013). Energy budget constraints on climate response. *Nat. Geosci.* **6** 415–416.
- OVERPECK, J., HUGHEN, K., HARDY, D., BRADLEY, R., CASE, R., DOUGLAS, M., FINNEY, B., GAJEWSKI, K., JACOBY, G., JENNINGS, A., LAMOUREUX, S., LASCA, A., MACDONALD, G., MOORE, J., RETELLE, M., SMITH, S., WOLFE, A. and ZIELINSKI, G. (1997). Arctic environmental change of the last four centuries. *Science* **278** 1251–1256.
- PAGES 2K CONSORTIUM (2013). Continental-scale temperature variability during the past two millennia. *Nature Geoscience* **6** 339–346.
- PALMA, W. and BONDON, P. (2003). On the eigenstructure of generalized fractional processes. *Statist. Probab. Lett.* **65** 93–101. [MR2017253](#)
- PERCIVAL, D. B., OVERLAND, J. E. and MOFJELD, H. O. (2001). Interpretation of North Pacific variability as a short- and long-memory process. *J. Climate* **14** 7–11.
- PLUMMER, M., BEST, N., COWLES, K. and VINES, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6** 7–11.
- ROBINSON, P. M. (1995). Gaussian semiparametric estimation of long range dependence. *Ann. Statist.* **23** 1630–1661. [MR1370301](#)
- RUTHERFORD, S., MANN, M. E., DELWORTH, T. L. and STOUFFER, R. J. (2003). Climate field reconstruction under stationary and nonstationary forcing. *J. Climate* **16** 462–479.

- RUTHERFORD, S., MANN, M. E., OSBORN, T. J., BRIFFA, K. R., JONES, P. D., BRADLEY, R. S. and HUGHES, M. K. (2005). Proxy-based Northern Hemisphere surface temperature reconstructions: Sensitivity to method, predictor network, target season, and target domain. *J. Climate* **18** 2308–2329.
- SCHNEIDER, T. (2001). Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *J. Climate* **14** 853–871.
- SMERDON, J. E., KAPLAN, A., CHANG, D. and EVANS, M. N. (2010). A pseudoproxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium. *J. Climate* **23** 4856–4880.
- SMITH, R. L. (2010). Understanding sensitivities in paleoclimatic reconstructions. Preprint.
- STEIG, E. J., SCHNEIDER, D. P., RUTHERFORD, S. D., MANN, M. E., COMISO, J. C. and SHINDELL, D. T. (2009). Warming of the Antarctic ice-sheet surface since the 1957 international geophysical year. *Nature* **457** 459–462.
- STOCKER, T. F., DAHE, Q., PLATTNER, G.-K., ALEXANDER, L., ALLEN, S., BINDOFF, N., BREON, F.-M., CHURCH, J., CUBASCH, U., EMORI, S., FORSTER, P., FRIEDLINGSTEIN, P., GILLET, N., GREGORY, J., HARTMANN, D., JANSEN, E., KIRTMAN, B., KNUTTI, R., KUMAR KANIKICHARLA, K., LEMKE, P., MAROTZKE, J., MASSON-DELMOTTE, V., MEEHL, G., MOKHOV, I., PIAO, S., RAMASWAMY, V., RANDALL, D., RHEIN, M., ROJAS, M., SABINE, C., SHINDELL, D., TALLEY, L., VAUGHAN, D. and XIE, S.-P. (2013). Technical summary. In *Climate Change 2013: The Physical Science Basis* (T. F. Stocker, Q. Dahe, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P. M. Midgley, eds.). Cambridge Univ. Press, Cambridge.
- TAQQU, M. S. (2013). Benoît Mandelbrot and fractional Brownian motion. *Statist. Sci.* **28** 131–134. [MR3075342](#)
- TILJANDER, M., SAARNISTO, M., OJALA, A. E. K. and SAARINEN, T. (2003). A 3000-year palaeoenvironmental record from annually laminated sediment of lake Korttajärvi, central Finland. *Boreas* **32** 566–577.
- TINGLEY, M. P. and HUYBERS, P. (2010a). A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: Development and applications to paleoclimate reconstruction problems. *J. Climate* **23** 2759–2781.
- TINGLEY, M. P. and HUYBERS, P. (2010b). A Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: Comparison with the regularized expectation-maximization algorithm. *J. Climate* **23** 2782–2800.
- TINGLEY, M. P. and HUYBERS, P. (2013). Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature* **496** 201–205.
- TINGLEY, M. P., CRAIGMILE, P. F., HARAN, M., LI, B., MANNSHARDT, E. and RAJARATNAM, B. (2012). Piecing together the past: Statistical insights into paleoclimatic reconstructions. *Quat. Sci. Rev.* **35** 1–22.
- TUDOR, C. A. and VIENS, F. G. (2007). Statistical aspects of the fractional stochastic calculus. *Ann. Statist.* **35** 1183–1212. [MR2341703](#)
- TUNG, K. K., ZHOU, J. S. and CAMP, C. D. (2008). Constraining model transient climate response using independent observations of solar-cycle forcing and response. *Geophys. Res. Lett.* **35** L17707.
- WAHL, E. R. and SMERDON, J. E. (2012). Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noise-only predictors. *Geophysical Research Letters* **39** L06703.
- WERNER, J. P., LUTERBACHER, J. and SMERDON, J. E. (2013). A pseudoproxy evaluation of Bayesian hierarchical modeling and canonical correlation analysis for climate field reconstructions over Europe. *J. Climate* **26** 851–867.
- YANG, X., XING, K., SHI, K. and PAN, Q. (2008). Joint state and parameter estimation in particle filtering and stochastic optimization. *J. Control Theory Appl.* **6** 215–220. [MR2418951](#)

ZHANG, Z., MANN, M. E. and COOK, E. R. (2004). Alternative methods of proxy-based climate field reconstruction: Application to summer drought over the conterminous United States back to AD 1700 from tree-ring data. *Holocene* **14** 502–516.

EQUIVALENCE TESTING FOR FUNCTIONAL DATA WITH AN APPLICATION TO COMPARING PULMONARY FUNCTION DEVICES

BY COLIN B. FOGARTY AND DYLAN S. SMALL

University of Pennsylvania

Equivalence testing for scalar data has been well addressed in the literature, however, the same cannot be said for functional data. The resultant complexity from maintaining the functional structure of the data, rather than using a scalar transformation to reduce dimensionality, renders the existing literature on equivalence testing inadequate for the desired inference. We propose a framework for equivalence testing for functional data within both the frequentist and Bayesian paradigms. This framework combines extensions of scalar methodologies with new methodology for functional data. Our frequentist hypothesis test extends the Two One-Sided Testing (TOST) procedure for equivalence testing to the functional regime. We conduct this TOST procedure through the use of the nonparametric bootstrap. Our Bayesian methodology employs a functional analysis of variance model, and uses a flexible class of Gaussian Processes for both modeling our data and as prior distributions. Through our analysis, we introduce a model for heteroscedastic variances within a Gaussian Process by modeling variance curves via Log-Gaussian Process priors. We stress the importance of choosing prior distributions that are commensurate with the prior state of knowledge and evidence regarding practical equivalence. We illustrate these testing methods through data from an ongoing method comparison study between two devices for pulmonary function testing. In so doing, we provide not only concrete motivation for equivalence testing for functional data, but also a blueprint for researchers who hope to conduct similar inference.

REFERENCES

- ALBEROLA-LOPEZ, C. and MARTIN-FERNANDEZ, M. (2003). A simple test of equality of time series. *Signal Process.* **83** 1343–1348.
- ANDERSON, S. and HAUCK, W. W. (1990). Consideration of individual bioequivalence. *Journal of Pharmacokinetics and Pharmacodynamics* **18** 259–273.
- BARNARD, J., MCCULLOCH, R. and MENG, X.-L. (2000). Modeling covariance matrices in terms of standard deviations and correlations, with application to shrinkage. *Statist. Sinica* **10** 1281–1311. [MR1804544](#)
- BEHSETA, S. and KASS, R. E. (2005). Testing equality of two functions using BARS. *Stat. Med.* **24** 3523–3534. [MR2210791](#)
- BERGER, R. L. (1982). Multiparameter hypothesis testing and acceptance sampling. *Technometrics* **24** 295–300. [MR0687187](#)

Key words and phrases. Equivalence testing, functional data analysis, bootstrap, Bayesian, Gaussian processes.

- BERGER, R. L. and HSU, J. C. (1996). Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statist. Sci.* **11** 283–319. [MR1445984](#)
- BROWN, L. D., HWANG, J. and MUNK, A. (1995). An unbiased test for the bioequivalence problem. Technical report, Cornell Univ., Ithaca, NY.
- BUJA, A. and ROLKE, W. (2003). (Re)sampling methods for simultaneous inference with applications to function estimation and functional data. Unpublished manuscript.
- CHAMBERS, R. and CHANDRA, H. (2013). A random effect block bootstrap for clustered data. *J. Comput. Graph. Statist.* **22** 452–470. [MR3173724](#)
- CHEN, K. and MÜLLER, H.-G. (2012). Modeling repeated functional observations. *J. Amer. Statist. Assoc.* **107** 1599–1609. [MR3036419](#)
- CHOW, S.-C. and LIU, J.-P. (1992). On the assessment of variability in bioavailability/bioequivalence studies. *Commun. Stat.* **21** 2591–2607.
- DAVISON, A. C. and HINKLEY, D. V. (1997). *Bootstrap Methods and Their Application*. Cambridge Series in Statistical and Probabilistic Mathematics **1**. Cambridge Univ. Press, Cambridge. [MR1478673](#)
- EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*. Monographs on Statistics and Applied Probability **57**. Chapman & Hall, New York. [MR1270903](#)
- FOGARTY, C. B. and SMALL, D. S. (2014). Supplement to “Equivalence testing for functional data with an application to comparing pulmonary function devices.” DOI:10.1214/14-AOAS763SUPP.
- FUENTES, M. (2002). Spectral methods for nonstationary spatial processes. *Biometrika* **89** 197–210. [MR1888368](#)
- FUENTES, M. and SMITH, R. L. (2001). A new class of nonstationary spatial models. Technical report, North Carolina State Univ., Raleigh, NC.
- GOUDSOUZIAN, N. G. and KARAMANIAN, A. (1984). *Physiology for the Anesthesiologist*, 2nd ed. Appleton-Century-Crofts, Norwalk, CT.
- JEFFREYS, H. (1961). *Theory of Probability*, 3rd ed. Clarendon, Oxford. [MR0187257](#)
- KASS, R. E. and RAFTERY, A. E. (1995). Bayes factors. *J. Amer. Statist. Assoc.* **90** 773–795.
- KAUFMAN, C. G. and SAIN, S. R. (2010). Bayesian functional ANOVA modeling using Gaussian process prior distributions. *Bayesian Anal.* **5** 123–149. [MR2596438](#)
- MATÉRN, B. (1986). *Spatial Variation*, 2nd ed. *Lecture Notes in Statistics* **36**. Springer, Berlin. [MR0867886](#)
- MORRIS, J. S. and CARROLL, R. J. (2006). Wavelet-based functional mixed models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68** 179–199. [MR2188981](#)
- NYCHKA, D., WIKLE, C. and ROYLE, J. A. (2002). Multiresolution models for nonstationary spatial covariance functions. *Stat. Model.* **2** 315–331. [MR1951588](#)
- PACIOREK, C. J. and SCHERVISH, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17** 483–506. [MR2240939](#)
- R DEVELOPMENT CORE TEAM (2011). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer, New York. [MR2168993](#)
- SEARLE, S. R., CASELLA, G. and MCCULLOCH, C. E. (1992). *Variance Components*. Wiley, New York. [MR1190470](#)
- SHAO, J. and TU, D. S. (1995). *The Jackknife and Bootstrap*. Springer, New York. [MR1351010](#)
- TIBSHIRANI, R. (1988). Variance stabilization and the bootstrap. *Biometrika* **75** 433–444. [MR0967582](#)

A MULTIPLE FILTER TEST FOR THE DETECTION OF RATE CHANGES IN RENEWAL PROCESSES WITH VARYING VARIANCE

BY MICHAEL MESSER^{*,1}, MARIETTA KIRCHNER[†], JULIA SCHIEMANN^{*,‡},
JOCHEN ROEPER^{*,1,2}, RALPH NEININGER^{*} AND GABY SCHNEIDER^{1,2*}

Goethe University Frankfurt^{}, Heidelberg University[†]
and University of Edinburgh[‡]*

Nonstationarity of the event rate is a persistent problem in modeling time series of events, such as neuronal spike trains. Motivated by a variety of patterns in neurophysiological spike train recordings, we define a general class of renewal processes. This class is used to test the null hypothesis of stationary rate versus a wide alternative of renewal processes with finitely many rate changes (change points). Our test extends ideas from the filtered derivative approach by using multiple moving windows simultaneously. To adjust the rejection threshold of the test, we use a Gaussian process, which emerges as the limit of the filtered derivative process. We also develop a multiple filter algorithm, which can be used when the null hypothesis is rejected in order to estimate the number and location of change points. We analyze the benefits of multiple filtering and its increased detection probability as compared to a single window approach. Application to spike trains recorded from dopamine midbrain neurons in anesthetized mice illustrates the relevance of the proposed techniques as preprocessing steps for methods that assume rate stationarity. In over 70% of all analyzed spike trains classified as rate nonstationary, different change points were detected by different window sizes.

REFERENCES

- ANTOCH, J. and HUŠKOVÁ, M. (1994). Procedures for the detection of multiple changes in series of independent observations. In *Asymptotic Statistics (Prague, 1993)*. *Contrib. Statist.* 3–20. Physica, Heidelberg. [MR1311925](#)
- BASSEVILLE, M. and NIKIFOROV, I. V. (1993). *Detection of Abrupt Changes: Theory and Application*. Prentice Hall, Englewood Cliffs, NJ. [MR1210954](#)
- BENDA, J. and HERZ, A. V. M. (2003). A universal model for spike-frequency adaptation. *Neural Comput.* **15** 2523–2564.
- BERTRAND, P. (2000). A local method for estimating change points: The “hat-function.” *Statistics* **34** 215–235. [MR1802728](#)
- BERTRAND, P. R., FHIMA, M. and GUILLIN, A. (2011). Off-line detection of multiple change points by the filtered derivative with p -value method. *Sequential Anal.* **30** 172–207. [MR2801138](#)
- BILLINGSLEY, P. (1999). *Convergence of Probability Measures*, 2nd ed. Wiley, New York. [MR1700749](#)
- BRODSKY, B. E. and DARKHOVSKY, B. S. (1993). *Nonparametric Methods in Change-Point Problems*. Kluwer Academic, Dordrecht. [MR1228205](#)

Key words and phrases. Stochastic processes, renewal processes, change point detection, nonstationary rate, multiple filters, multiple time scales.

- BRODY, C. D. (1999). Correlations without synchrony. *Neural Comput.* **11** 1537–1551.
- BROWN, E. N., KASS, R. E. and MITRA, P. P. (2004). Multiple neural spike train data analysis: State-of-the-art and future challenges. *Nat. Neurosci.* **7** 456–461.
- CSÖRGŐ, M. and HORVÁTH, L. (1987). Asymptotic distributions of pontograms. *Math. Proc. Cambridge Philos. Soc.* **101** 131–139. [MR0877707](#)
- CSÖRGŐ, M. and HORVÁTH, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley, Chichester. [MR2743035](#)
- DREYER, J. K., HERRIK, K. F., BERG, R. W. and HOUNSGAARD, J. D. (2010). Influence of phasic and tonic dopamine release on receptor activation. *J. Neurosci.* **30** 14273–14283.
- FARKHOI, F., STRUBE-BLOSS, M. and NAWROT, M. P. (2009). Serial correlation in neural spike trains: Experimental evidence, stochastic modelling, and single neuron variability. *Phys. Rev. E* **79** 021905.
- GONON, F. G. (1988). Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by in vivo electrochemistry. *Neuroscience* **24** 19–28.
- GOURÉVITCH, B. and EGGERMONT, J. J. (2007). A nonparametric approach for detection of bursts in spike trains. *J. Neurosci. Methods* **160** 349–358.
- GRÜN, S., DIESMANN, M. and AERTSEN, A. (2002). “Unitary Events” in multiple single-neuron activity. II. Non-Stationary data. *Neural Comput.* **14** 81–119.
- GRÜN, S., RIEHLE, A. and DIESMANN, M. (2003). Effect of cross-trial nonstationarity on joint-spike events. *Biol. Cybernet.* **88** 335–351.
- GRÜN, S. and ROTTER, S., eds. (2010). *Analysis of Parallel Spike Trains. Springer Series in Computational Neuroscience 7*. Springer, New York.
- HOWE, M. W., TIERNEY, P. L., SANDBERG, S. G., PHILLIPS, P. E. M. and GRAYBIEL, A. M. (2013). Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. *Nature* **500** 575–579.
- HUŠKOVÁ, M. and SLABÝ, A. (2001). Permutation tests for multiple changes. *Kybernetika (Prague)* **37** 605–622. [MR1877077](#)
- JOHNSON, D. H. (1996). Point process models of single-neuron discharges. *J. Comput. Neurosci.* **3** 275–299.
- JONES, M. C., MARRON, J. S. and SHEATHER, S. J. (1996). A brief survey of bandwidth selection for density estimation. *J. Amer. Statist. Assoc.* **91** 401–407. [MR1394097](#)
- KASS, R. E., VENTURA, V. and BROWN, E. N. (2005). Statistical issues in the analysis of neuronal data. *J. Neurophysiol.* **94** 8–25.
- KENDALL, D. G. and KENDALL, W. S. (1980). Alignments in two-dimensional random sets of points. *Adv. in Appl. Probab.* **12** 380–424. [MR0569434](#)
- LEGÉNDY, C. R. and SALCMAN, M. (1985). Bursts and recurrences of bursts in the spike trains of spontaneously active striate cortex neurons. *J. Neurophysiol.* **53** 926–939.
- MESSER, M., KIRCHNER, M., SCHIEMANN, J., ROEPER, J., NEININGER, R. and SCHNEIDER, G. (2014). Supplement to “A multiple filter test for the detection of rate changes in renewal processes with varying variance.” DOI:[10.1214/14-AOAS782SUPP](#).
- MILEYKOVSKIY, B. and MORALES, M. (2011). Duration of inhibition of ventral tegmental area dopamine neurons encodes a level of conditioned fear. *J. Neurosci.* **31** 7471–7476.
- MUHSAL, B. (2013). Change-point methods for multivariate autoregressive models and multiple structural breaks in the mean. Dissertation, available at <http://nbn-resolving.org/urn:nbn:de:swb:90-355368>.
- NAWROT, M. P., AERTSEN, A. and ROTTER, S. (1999). Single-trial estimation of neuronal firing rates: From single-neuron spike trains to population activity. *J. Neurosci. Methods* **94** 81–92.
- NAWROT, M. P., BOUCSEIN, C., RODRIGUEZ MOLINA, V., RIEHLA, A., AERTSEN, A. and ROTTER, S. (2008). Measurement of variability dynamics in cortical spike trains. *Journal of Neuroscience Methods* **169** 347–390.

- PERKEL, D. H., GERSTEIN, G. L. and MOORE, G. P. (1967a). Neuronal spike trains and stochastic point processes. I. The single spike train. *Biophys. J.* **7** 391–417.
- PERKEL, D. H., GERSTEIN, G. L. and MOORE, G. P. (1967b). Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys. J.* **7** 419–440.
- PETROV, V. V. (1995). *Limit Theorems of Probability Theory: Sequences of Independent Random Variables*. Clarendon, Oxford. [MR1353441](#)
- POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer, New York. [MR0762984](#)
- REDGRAVE, P., RODRIGUEZ, M., SMITH, Y., RODRIGUEZ-OROZ, M. C., LEHERICY, S., BERGMAN, H., AGID, Y., DELONG, M. R. and OBESO, J. A. (2010). Goal-directed and habitual control in the basal ganglia: Implications for Parkinson’s disease. *Nat. Rev., Neurosci.* **11** 760–772.
- RIEKE, F., WARLAND, D., DE RUYTER VAN STEVENINCK, R. and BIALEK, W. (1999). *Spikes: Exploring the Neural Code*. MIT Press, Cambridge, MA. [MR1983010](#)
- SCHIEMANN, J., KLOSE, V., SCHLAUDRAFF, F., BINGMER, M., SEINO, S., MAGILL, P. J., SCHNEIDER, G., LISS, B. and ROEPER, J. (2012). K-ATP channels control in vivo burst firing of dopamine neurons in the medial substantia nigra and novelty-induced behavior. *Nat. Neurosci.* **15** 1272–1280.
- SCHNEIDER, G. (2008). Messages of oscillatory correlograms: A spike train model. *Neural Comput.* **20** 1211–1238. [MR2402061](#)
- SHIMAZAKI, H. and SHINOMOTO, S. (2007). A method for selecting the bin size of a time histogram. *Neural Comput.* **19** 1503–1527. [MR2316960](#)
- STAUDE, B., ROTTER, S. and GRÜN, S. (2008). Can spike coordination be differentiated from rate covariation? *Neural Comput.* **20** 1973–1999. [MR2419958](#)
- STEINEBACH, J. and EASTWOOD, V. R. (1995). On extreme value asymptotics for increments of renewal processes. *J. Statist. Plann. Inference* **45** 301–312. [MR1342102](#)
- STEINEBACH, J. and ZHANG, H. Q. (1993). On a weighted embedding for pontograms. *Stochastic Process. Appl.* **47** 183–195. [MR1239836](#)
- TOKDAR, S., XI, P., KELLY, R. C. and KASS, R. E. (2010). Detection of bursts in extracellular spike trains using hidden semi-Markov point process models. *J. Comput. Neurosci.* **29** 203–212.

BAYESIAN PROTEIN STRUCTURE ALIGNMENT¹

BY ABEL RODRIGUEZ AND SCOTT C. SCHMIDLER

University of California, Santa Cruz and Duke University

The analysis of the three-dimensional structure of proteins is an important topic in molecular biochemistry. Structure plays a critical role in defining the function of proteins and is more strongly conserved than amino acid sequence over evolutionary timescales. A key challenge is the identification and evaluation of structural similarity between proteins; such analysis can aid in understanding the role of newly discovered proteins and help elucidate evolutionary relationships between organisms. Computational biologists have developed many clever algorithmic techniques for comparing protein structures, however, all are based on heuristic optimization criteria, making statistical interpretation somewhat difficult. Here we present a fully probabilistic framework for pairwise structural alignment of proteins. Our approach has several advantages, including the ability to capture alignment uncertainty and to estimate key “gap” parameters which critically affect the quality of the alignment. We show that several existing alignment methods arise as *maximum a posteriori* estimates under specific choices of prior distributions and error models. Our probabilistic framework is also easily extended to incorporate additional information, which we demonstrate by including primary sequence information to generate simultaneous sequence–structure alignments that can resolve ambiguities obtained using structure alone. This combined model also provides a natural approach for the difficult task of estimating evolutionary distance based on structural alignments. The model is illustrated by comparison with well-established methods on several challenging protein alignment examples.

REFERENCES

- ALTSCHUL, S. F., GISH, W., MILLER, W., MYERS, E. W. and LIPMAN, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215** 403–410.
- BISHOP, M. J. and THOMPSON, E. A. (1986). Maximum likelihood alignment of DNA sequences. *J. Mol. Biol.* **190** 159–165.
- BRONNER, C. E. et al. (1994). Mutation in the DNA mismatch repair gene homologue hMLH1 is associated with hereditary non-polyposis colon cancer. *Nature* **368** 258–261.
- BROWN, N., ORENGO, C. and TAYLOR, W. (1996). A protein structure comparison methodology. *Computational Chemistry* **20** 359–380.
- CHOTHIA, C. and LESK, A. M. (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J.* **5** 823–826.
- COHEN, G. (1997). ALIGN: A program to superimpose protein coordinates, accounting for insertions and deletions. *Acta Crystallographica* **30** 1160–1161.

Key words and phrases. Protein alignment, structure alignment, affine gap, dynamic programming, Procrustes distance.

- DAYHOFF, M. and ECK, R., eds. (1968). *Atlas of Protein Sequence and Structure* **3**. National Biomedical Research Foundation, Silver Spring, MD.
- DRYDEN, I. L., HIRST, J. D. and MELVILLE, J. L. (2007). Statistical analysis of unlabeled point sets: Comparing molecules in chemoinformatics. *Biometrics* **63** 237–251, 315. [MR2345594](#)
- DRYDEN, I. L. and MARDIA, K. V. (1998). *Statistical Shape Analysis*. Wiley, Chichester. [MR1646114](#)
- DURBIN, R., EDDY, S., KROGH, A. and MITCHISON, G. (1998). *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge Univ. Press, Cambridge.
- EIDHAMMER, I., JONASSEN, I. and TAYLOR, W. R. (2000). Structure comparison and structure patterns. *J. Comput. Biol.* **7** 685–716.
- FALICOV, A. and COHEN, F. E. (1996). A surface of minimum area metric for the structural comparison of proteins. *J. Mol. Biol.* **258** 871–892.
- FISCHER, D., WOLFSON, H., LIN, S. and NUSSINOV, R. (1994). Three-dimensional, sequence order-independent structural comparison of a serine protease against the crystallographic database reveals active site similarities: Potential implications to evolution and to protein folding. *Protein Sci.* **3** 768–778.
- GELMAN, A. and RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–472.
- GERSTEIN, M. and LEVITT, M. (1998). Comprehensive assessment of automatic structural alignment against a manual standard, the scop classification of proteins. *Protein Sci.* **7** 445–456.
- GIBRAT, J.-F., MADEJ, T. and BRYANT, S. (1996). Surprising similarities in structure comparison. *Current Opinion in Structural Biology* **6** 377–385.
- GODZIK, A. (1996). The structural alignment between two proteins: Is there a unique answer? *Protein Eng.* **5** 1325–1338.
- GREEN, P. J. and MARDIA, K. V. (2006). Bayesian alignment using hierarchical models, with applications in protein bioinformatics. *Biometrika* **93** 235–254. [MR2278080](#)
- GRISHIN, N. V. (1997). Estimation of evolutionary distances from protein spatial structures. *J. Mol. Evol.* **45** 359–369.
- HENIKOFF, S. and HENIKOFF, J. (1992). Amino acid substitution matrices from protein blocks. *Proceedings of National Academy of Science* **89** 10915–10919.
- HOLM, L. and SANDER, C. (1993). Protein structure comparison by alignment of distance matrices. *J. Mol. Biol.* **233** 123–138.
- JOHNSON, M. S., SUTCLIFFE, M. J. and BLUNDELL, T. L. (1990). Molecular anatomy: Phyletic relationships derived from three-dimensional structures of proteins. *J. Mol. Evol.* **30** 43–59.
- KENDALL, D. G., BARDEN, D., CARNE, T. K. and LE, H. (1999). *Shape and Shape Theory*. Wiley, Chichester. [MR1891212](#)
- KENOBI, K. and DRYDEN, I. L. (2012). Bayesian matching of unlabeled point sets using Procrustes and configuration models. *Bayesian Anal.* **7** 547–565. [MR2981627](#)
- KOEHL, P. and LEVITT, M. (2002). Sequence variations within protein families are linearly related to structural variations. *J. Mol. Biol.* **323** 551–562.
- KOTLOVYI, V., NICHOLS, W. L. and EYCK, L. F. T. (2003). Protein structural alignment for detection of maximally conserved regions. *Biophys. Chem.* **105** 595–608.
- LACKNER, P., KOPPENSTEINER, W. A., SIPPL, M. J. and DOMINGUES, F. S. (2000). ProSup: A refined tool for protein structure alignment. *Protein Eng.* **13** 745–752.
- LEMMEN, C. and LENGAUER, T. (2000). Computational methods for the structural alignment of molecules. *J. Comput.-Aided Mol. Des.* **14** 215–232.
- LEVITT, M. and GERSTEIN, M. (1998). A unified statistical framework for sequence comparison and structure comparison. *Proceedings of the National Academy of Sciences USA* **95** 5913–5920.
- LIPMAN, D. J. and PEARSON, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227** 1435–1441.

- LIU, J. S. and LAWRENCE, C. E. (1999). Bayesian inference on biopolymer models. *Bioinformatics* **15** 38–52.
- MIZUGUCHI, K. and GO, N. (1995). Seeking significance in three-dimensional protein structure comparisons. *Curr. Opin. Struct. Biol.* **5** 377–382.
- NEEDLEMAN, S. B. and WUNSCH, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48** 443–453.
- ORTIZ, A. R., STRAUSS, C. E. M. and OLMEA, O. (2002). Mammoth (Matching molecular models obtained from theory): An automated method for model comparison. *Protein Sci.* **11** 2606–2621.
- PAPADOPOULOS, N., NICOLAIDES, N. C., WEI, Y. F., RUBEN, S. M., CARTER, K. C., ROSEN, C. A., HASELTINE, W. A., FLEISCHMANN, R. D., FRASER, C. M., ADAMS, M. D. et al. (1994). Mutation of a mutL homolog in hereditary colon cancer. *Science* **263** 1625–1629.
- RAO, S. and ROSSMANN, M. (1973). Comparison of super-secondary structures in proteins. *J. Mol. Biol.* **105** 241–256.
- ROSSMANN, M. and ARGOS, P. (1975). A comparison of heme binding pocket in globins and cytochrome b5*. *J. Biol. Chem.* **250** 7523–7532.
- ROSSMANN, M. and ARGOS, P. (1976). Exploring structural homology in proteins. *J. Mol. Biol.* **105** 75–96.
- SATOW, Y., COHEN, G., PADLAN, E. and AVIES, D. (1986). Phosphocholine binding immunoglobulin fab mcpc603: An x-ray diffraction study at 2.7 Å. *J. Mol. Biol.* **190** 593–604.
- SAUL, L. K. and JORDAN, M. I. (1995). Boltzmann chains and hidden Markov models. In *Advances in Neural Information Processing Systems (NIPS) 7* (G. Tesauro, D. S. Touretzky and T. Leen, eds.). MIT Press, Cambridge, MA.
- SCHMIDLER, S. C. (2003) Statistical shape analysis of protein structure families. In *Presentation at the LASR workshop on Stochastic Geometry, Biological Structure and Images*, Leeds, UK.
- SCHMIDLER, S. C. (2004). Bayesian shape matching and protein structure alignment. In *Presentation at the 6th World Congress of the Bernoulli Society and the 67th Annual Meeting of the Institute of Mathematical Statistics*, Barcelona, Spain.
- SCHMIDLER, S. C. (2007a). Fast Bayesian shape matching using geometric algorithms. In *Bayesian Statistics 8* (M. Bernardo, J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. Smith and M. West, eds.) 471–490. Oxford Univ. Press, Oxford. [MR2433204](#)
- SCHMIDLER, S. C. (2007b). Bayesian flexible shape matching with applications to structural bioinformatics. Technical report, Dept. Statistical Sciences, Duke Univ., Durham, NC.
- SCHMIDLER, S. C., LUCAS, J. E. and OAS, T. G. (2007). Statistical estimation of statistical mechanical models: Helix-coil theory and peptide helicity prediction. *J. Comput. Biol.* **14** 1287–1310. [MR2373277](#)
- SHINDYALOV, I. N. and BOURNE, P. E. (1998). Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **11** 739–747.
- SMALL, C. G. (1996). *The Statistical Theory of Shape*. Springer, New York. [MR1418639](#)
- TAYLOR, W. R. (2002). Protein structure comparison using bipartite graph matching and its application to protein structure classification. *Mol. Cell. Proteomics* **1** 334–339.
- TAYLOR, W. R. and ORENGO, C. A. (1989). Protein structure alignment. *J. Mol. Biol.* **208** 1–22.
- THOMPSON, J. D., PLEWNIAK, F. and POCH, O. (1999). BALiBASE: A benchmark alignment database for the evaluation of multiple alignment programs. *Bioinformatics* **15** 87–88.
- THOMPSON, J. D., KOEHL, P., RIPP, R. and POCH, O. (2005). BALiBASE 3.0: Latest developments of the multiple sequence alignment benchmark. *Proteins* **61** 127–136.
- WALLACE, A., LASKOWSKI, R. and THORNTON, J. (1996). Tess: A geometric hashing algorithm for deriving 3d coordinate templates for searching structural databases: Applications to enzyme active sites. *Protein Sci.* **6** 2308–2323.
- WANG, R. and SCHMIDLER, S. C. (2008). Bayesian multiple protein structure alignment and analysis of protein families. In preparation.

- WEBB, B.-J. M., LIU, J. S. and LAWRENCE, C. E. (2002). BALSAs: Bayesian algorithm for local sequence alignment. *Nucleic Acids Res.* **30** 1268–1277.
- WOOD, T. C. and PEARSON, W. R. (1999). Evolution of protein sequences and structures. *J. Mol. Biol.* **291** 977–995.
- WU, T. D., SCHMIDLER, S. C., HASTIE, T. and BRUTLAG, D. L. (1998). Regression analysis of multiple protein structures. *J. Comput. Biol.* **5** 597–607.
- ZHU, J., LIU, J. S. and LAWRENCE, C. E. (1998). Bayesian adaptive sequence alignment algorithms. *Bioinformatics* **14** 25–39.
- ZU-KANG, F. and SIPPL, M. (1996). Optimum superimposition of protein structures: Ambiguities and implications. *Fold. Des.* **1** 123–132.

ISOLATION IN THE CONSTRUCTION OF NATURAL EXPERIMENTS

BY JOSÉ R. ZUBIZARRETA*, DYLAN S. SMALL^{1,†}
AND PAUL R. ROSENBAUM^{1,†}

Columbia University and University of Pennsylvania[†]*

A natural experiment is a type of observational study in which treatment assignment, though not randomized by the investigator, is plausibly close to random. A process that assigns treatments in a highly nonrandom, inequitable manner may, in rare and brief moments, assign aspects of treatments at random or nearly so. Isolating those moments and aspects may extract a natural experiment from a setting in which treatment assignment is otherwise quite biased, far from random. Isolation is a tool that focuses on those rare, brief instances, extracting a small natural experiment from otherwise useless data. We discuss the theory behind isolation and illustrate its use in a reanalysis of a well-known study of the effects of fertility on workforce participation. Whether a woman becomes pregnant at a certain moment in her life and whether she brings that pregnancy to term may reflect her aspirations for family, education and career, the degree of control she exerts over her fertility, and the quality of her relationship with the father; moreover, these aspirations and relationships are unlikely to be recorded with precision in surveys and censuses, and they may confound studies of workforce participation. However, given that a woman is pregnant and will bring the pregnancy to term, whether she will have twins or a single child is, to a large extent, simply luck. Given that a woman is pregnant at a certain moment, the differential comparison of two types of pregnancies on workforce participation, twins or a single child, may be close to randomized, not biased by unmeasured aspirations. In this comparison, we find in our case study that mothers of twins had more children but only slightly reduced workforce participation, approximately 5% less time at work for an additional child.

REFERENCES

- ANGRIST, J. D. and EVANS, W. N. (1998). Children and their parent's labor supply: Evidence from exogenous variation in family size. *Amer. Econ. Rev.* **88** 450–477.
- 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.
- ANTHONY, J. C., BREITNER, J. C., ZANDI, P. P., MEYER, M. R., JURASOVA, I., NORTON, M. C. and STONE, S. V. (2000). Reduced prevalence of AD in users of NSAIDs and H2 receptor antagonists: The cache county study. *Neurology* **54** 2066–2071.
- APEL, R., BLOKLAND, A. A. J., NIEUWBEERTA, P. and VAN SCHELLEN, M. (2010). The impact of imprisonment on marriage and divorce: A risk-set matching approach. *J. Quant. Criminol.* **26** 269–300.

Key words and phrases. Differential effect, generic bias, risk-set matching, sensitivity analysis.

- ARPINO, B. and AASSVE, A. (2013). Estimating the causal effect of fertility on economic wellbeing: Data requirements, identifying assumptions and estimation methods. *Empir. Econ.* **44** 355–385.
- BAIOCCHI, M., SMALL, D. S., LORCH, S. and ROSENBAUM, P. R. (2010). Building a stronger instrument in an observational study of perinatal care for premature infants. *J. Amer. Statist. Assoc.* **105** 1285–1296. [MR2796550](#)
- BOUND, J., JAEGER, D. A. and BAKER, R. M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *J. Amer. Statist. Assoc.* **90** 443–450.
- BROOKHART, M. A., WANG, P. S., SOLOMON, D. H. and SCHNEEWEISS, S. (2006). Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* **17** 268–275.
- CAMPBELL, D. T. (1986). Relabeling internal and external validity for applied social scientists. In *Advances in Quasi-Experimental Design and Analysis* (W. M. K. Trochim, ed.) 67–77. Jossey-Bass, San Francisco, CA.
- CORNFIELD, J., HAENZEL, W., HAMMOND, E. C., LILIENTHAL, A. M., SHIMKIN, M. B. and WYNDER, E. L. (1959). Smoking and lung cancer: Recent evidence and a discussion of some questions. *J. Natl. Cancer Inst.* **22** 173–203.
- COX, D. R. (1972). Regression models and life-tables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **34** 187–220. [MR0341758](#)
- DIPRETE, T. A. and GANGL, M. (2004). Assessing bias in the estimation of causal effects. *Sociolog. Method.* **34** 271–310.
- EGLSTON, B. L., SCHARFSTEIN, D. O. and MACKENZIE, E. (2009). On estimation of the survivor average causal effect in observational studies when important confounders are missing due to death. *Biometrics* **65** 497–504. [MR2751473](#)
- GASTWIRTH, J. L. (1992). Methods for assessing the sensitivity of statistical comparisons used in title VII cases to omitted variables. *Jurimetrics* **33** 19–34.
- GIBBONS, R. D., AMATYA, A. K., BROWN, C. H., HUR, K., MARCUS, S. M., BHAUMIK, D. K. and MANN, J. (2010). Post-approval drug safety surveillance. *Ann. Rev. Pub. Health* **31** 419–437.
- HANSEN, B. B. (2007). Optmatch. *R News* **7** 18–24.
- HOLLAND, P. W. H. (1988). Causal inference, path analysis, and recursive structural equations models. *Sociolog. Method.* **18** 449–484.
- HOSMAN, C. A., HANSEN, B. B. and HOLLAND, P. W. (2010). The sensitivity of linear regression coefficients' confidence limits to the omission of a confounder. *Ann. Appl. Stat.* **4** 849–870. [MR2758424](#)
- HSU, J. Y. and SMALL, D. S. (2013). Calibrating sensitivity analyses to observed covariates in observational studies. *Biometrics* **69** 803–811. [MR3146776](#)
- IMAL, K., KEELE, L., TINGLEY, D. and YAMAMOTO, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *Amer. Polit. Sci. Rev.* **105** 765–789.
- IMBENS, G. W. and ROSENBAUM, P. R. (2005). Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *J. Roy. Statist. Soc. Ser. A* **168** 109–126. [MR2113230](#)
- KENNEDY, E. H., TAYLOR, J. M. G., SCHAUBEL, D. E. and WILLIAMS, S. (2010). The effect of salvage therapy on survival in a longitudinal study with treatment by indication. *Stat. Med.* **29** 2569–2580. [MR2756944](#)
- LI, Y. P., PROPERT, K. J. and ROSENBAUM, P. R. (2001). Balanced risk set matching. *J. Amer. Statist. Assoc.* **96** 870–882. [MR1946360](#)
- LIN, D. Y., PSATY, B. M. and KRONMAL, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounders in observational studies. *Biometrics* **54** 948–963.
- LIU, W., KURAMOTO, J. and STUART, E. (2013). Sensitivity analysis for unobserved confounding in nonexperimental prevention research. *Prev. Sci.* **14** 570–580.

- LU, B. (2005). Propensity score matching with time-dependent covariates. *Biometrics* **61** 721–728. [MR2196160](#)
- LU, B., GREEVY, R., XU, X. and BECK, C. (2011). Optimal nonbipartite matching and its statistical applications. *Amer. Statist.* **65** 21–30. [MR2899649](#)
- MARCUS, S. M. (1997). Using omitted variable bias to assess uncertainty in the estimation of an AIDS education treatment effect. *J. Ed. Behav. Statist.* **22** 193–201.
- MARCUS, S. M., SIDDIQUE, J., TEN HAVE, T. R., GIBBONS, R. D., STUART, E. A. and NORMAND, S. L. T. (2008). Balancing treatment comparisons in longitudinal studies. *Psychiatr. Ann.* **38** 805–811.
- MCCANDLESS, L. C., GUSTAFSON, P. and LEVY, A. (2007). Bayesian sensitivity analysis for unmeasured confounding in observational studies. *Stat. Med.* **26** 2331–2347. [MR2368419](#)
- MEYER, B. D. (1995). Natural and quasi-experiments in economics. *J. Bus. Econom. Statist.* **13** 151–161.
- MURRAY, J., LOEBER, R. and PARDINI, D. (2012). Parental involvement in the criminal justice system and the development of youth theft, marijuana use, depression, and poor academic performance. *Criminology* **50** 255–302.
- NAGIN, D. S. and SNODGRASS, G. M. (2013). The effect of incarceration on re-offending: Evidence from a natural experiment in Pennsylvania. *J. Quant. Criminol.* **29** 601–642.
- NEYMAN, J. (1923). On the application of probability theory to agricultural experiments. *Statist. Sci.* **5** 463–464.
- NIEUWBEERTA, P., NAGIN, D. S. and BLOKLAND, A. A. J. (2009). Assessing the impact of first-time imprisonment on offender's subsequent criminal career development: A matched samples comparison. *J. Quant. Criminol.* **25** 227–257.
- ROBINS, J. M., ROTNITZKY, A. and SCHARFSTEIN, D. O. (2000). Sensitivity analysis for selection bias and unmeasured confounding in missing data and causal inference models. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials (Minneapolis, MN, 1997)* (E. Halloran and D. Berry, eds.). *IMA Vol. Math. Appl.* **116** 1–94. Springer, New York. [MR1731681](#)
- ROSENBAUM, P. R. (1984). The consequences of adjustment for a concomitant variable that has been affected by the treatment. *J. Roy. Statist. Soc. Ser. A* **147** 656–666.
- ROSENBAUM, P. R. (1987). Sensitivity analysis for certain permutation tests in matched observational studies. *Biometrika* **74** 13–26.
- ROSENBAUM, P. R. (1996). Comment. *J. Amer. Statist. Assoc.* **91** 465–468.
- ROSENBAUM, P. R. (2006). Differential effects and generic biases in observational studies. *Biometrika* **93** 573–586. [MR2261443](#)
- ROSENBAUM, P. R. (2007). Sensitivity analysis for m -estimates, tests, and confidence intervals in matched observational studies. *Biometrics* **63** 456–464. [MR2370804](#)
- ROSENBAUM, P. R. (2010). *Design of Observational Studies*. Springer, New York. [MR2561612](#)
- ROSENBAUM, P. R. (2013a). Using differential comparisons in observational studies. *Chance* **26** 18–25.
- ROSENBAUM, P. R. (2013b). Impact of multiple matched controls on design sensitivity in observational studies. *Biometrics* **69** 118–127. [MR3058058](#)
- ROSENBAUM, P. R. and SILBER, J. H. (2009). Amplification of sensitivity analysis in matched observational studies. *J. Amer. Statist. Assoc.* **104** 1398–1405. [MR2750570](#)
- RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *J. Educ. Psych.* **66** 688–701.
- RUTTER, M. (2007). Proceeding from observed correlation to causal inference: The use of natural experiments. *Perspect. Psychol. Sci.* **2** 377–395.
- RYAN, P. B., MADIGAN, D., STANG, P. E., OVERHAGE, J. M., RACOOSIN, J. A. and HARTZEMA, A. G. (2012). Empirical assessment of methods for risk identification in health-care data: Results from the experiments of the observational medical outcomes partnership. *Stat. Med.* **31** 4401–4415. [MR3040089](#)

- SEKHON, J. S. and TITIUNIK, R. (2012). When natural experiments are neither natural nor experiments. *Amer. Polit. Sci. Rev.* **106** 35–57.
- SMALL, D. S. (2007). Sensitivity analysis for instrumental variables regression with overidentifying restrictions. *J. Amer. Statist. Assoc.* **102** 1049–1058. [MR2411664](#)
- SMALL, D. S. and ROSENBAUM, P. R. (2008). War and wages: The strength of instrumental variables and their sensitivity to unobserved biases. *J. Amer. Statist. Assoc.* **103** 924–933. [MR2528819](#)
- STUART, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statist. Sci.* **25** 1–21. [MR2741812](#)
- SUSSER, M. (1973). *Causal Thinking in the Health Sciences*. Oxford, New York.
- SUSSER, M. (1981). Prenatal nutrition, birthweight, and psychological development: An overview of experiments, quasi-experiments, and natural experiments in the past decade. *Amer. J. Clin. Nutrition* **34** 784–803.
- VANDENBROUCKE, J. P. (2004). When are observational studies as credible as randomised trials? *Lancet* **363** 1728–1731.
- VAN DER LAAN, M. J. and ROBINS, J. M. (2003). *Unified Methods for Censored Longitudinal Data and Causality*. *Springer Series in Statistics*. Springer, New York. [MR1958123](#)
- WANG, L. and KRIEGER, A. M. (2006). Causal conclusions are most sensitive to unobserved binary covariates. *Stat. Med.* **25** 2257–2271. [MR2240099](#)
- WILDEMAN, C., SCHNITTKER, J. and TURNER, K. (2012). Despair by association? *Amer. Sociol. Rev.* **77** 216–243.
- YU, B. B. and GASTWIRTH, J. L. (2005). Sensitivity analysis for trend tests: Application to the risk of radiation exposure. *Biostatistics* **6** 201–209.
- ZUBIZARRETA, J. R. (2012). Using mixed integer programming for matching in an observational study of kidney failure after surgery. *J. Amer. Statist. Assoc.* **107** 1360–1371. [MR3036400](#)
- ZUBIZARRETA, J. R., SMALL, D. S., GOYAL, N. K., LORCH, S. and ROSENBAUM, P. R. (2013). Stronger instruments via integer programming in an observational study of late preterm birth outcomes. *Ann. Appl. Stat.* **7** 25–50. [MR3086409](#)

GLOBAL ESTIMATION OF CHILD MORTALITY USING A BAYESIAN B-SPLINE BIAS-REDUCTION MODEL¹

BY LEONTINE ALKEMA AND JIN ROU NEW

National University of Singapore

Estimates of the under-five mortality rate (U5MR) are used to track progress in reducing child mortality and to evaluate countries' performance related to Millennium Development Goal 4. However, for the great majority of developing countries without well-functioning vital registration systems, estimating the U5MR is challenging due to limited data availability and data quality issues.

We describe a Bayesian penalized B-spline regression model for assessing levels and trends in the U5MR for all countries in the world, whereby biases in data series are estimated through the inclusion of a multilevel model to improve upon the limitations of current methods. B-spline smoothing parameters are also estimated through a multilevel model. Improved spline extrapolations are obtained through logarithmic pooling of the posterior predictive distribution of country-specific changes in spline coefficients with observed changes on the global level.

The proposed model is able to flexibly capture changes in U5MR over time, gives point estimates and credible intervals reflecting potential biases in data series and performs reasonably well in out-of-sample validation exercises. It has been accepted by the United Nations Inter-agency Group for Child Mortality Estimation to generate estimates for all member countries.

REFERENCES

- ALKEMA, L. and NEW, J. R. (2012). Progress toward global reduction in under-five mortality: A bootstrap analysis of uncertainty in Millennium Development Goal 4 estimates. *PLoS Med.* **9** e1001355.
- ALKEMA, L. and NEW, J. R. (2013). Global estimation of child mortality using a Bayesian B-spline bias-reduction model. Technical report. Available at <http://arxiv.org/abs/1309.1602>.
- ALKEMA, L. and NEW, J. (2014). Supplement to "Global estimation of child mortality using a Bayesian B-spline Bias-reduction model." DOI:10.1214/14-AOAS768SUPPA, DOI:10.1214/14-AOAS768SUPPB.
- ALKEMA, L., WONG, M. B. and SEAH, P. R. (2012). Monitoring progress towards Millennium Development Goal 4: A call for improved validation of under-5 mortality rate estimates. *Statistics, Politics and Policy* **3** Article ID 2.
- AMOUZOU, A. (2011). Real-time results tracking. Technical report, Institute for International Programs, Johns Hopkins Bloomberg School of Public Health, Baltimore. Available at http://www.jhsph.edu/departments/international-health/centers-and-institutes/institute-for-international-programs/_documents/RRT_Technical_Note.pdf.

Key words and phrases. Bayesian hierarchical model, Millennium Development Goal 4, logarithmic pooling, penalized B-spline regression model, under-five mortality rate, United Nations Inter-agency Group for Child Mortality Estimation.

- BRASS, W. (1964). Uses of census or survey data for the estimation of vital rates. In *African Seminar on Vital Statistics*, 14–19 December 1964, United Nations, New York.
- CENSUS OF INDIA (2011). Sample registration. Available at http://censusingia.gov.in/Vital_Statistics/SRS/Sample_Registration_System.aspx.
- CLARK, S. J., WAKEFIELD, J., MCCORMICK, T. and ROSS, M. (2012). Hyak mortality monitoring system: Innovative sampling and estimation methods. Working Paper 118. Available at <http://www.csss.washington.edu/Papers/wp118.pdf>.
- CURRIE, I. D. and DURBAN, M. (2002). Flexible smoothing with P -splines: A unified approach. *Stat. Model.* **2** 333–349. MR1951589
- EILERS, P. H. C. (1999). Discussion of “The analysis of designed experiments and longitudinal data using smoothing splines” by A. P. Verbyla, B. R. Cullis, M. G. Kenward and S. J. Welham. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **48** 300–311.
- EILERS, P. H. C. and MARX, B. D. (1996). Flexible smoothing with B -splines and penalties. *Statist. Sci.* **11** 89–121. MR1435485
- EILERS, P. H. C. and MARX, B. D. (2010). Splines, knots, and penalties. *Wiley Interdisciplinary Reviews: Computational Statistics* **2** 637–653.
- GELMAN, A. and RUBIN, D. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–511.
- GNEITING, T. and RAFTERY, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *J. Amer. Statist. Assoc.* **102** 359–378. MR2345548
- HILL, K., YOU, D., INOUE, M. and OESTERGAARD, M. Z. (2012). Child mortality estimation: Accelerated progress in reducing global child mortality *PLoS Med.* **9** e1001303.
- OESTERGAARD, M. Z., ALKEMA, L. and LAWN, J. E. (2013). Millennium Development Goals national targets are moving targets and the results will not be known until well after the deadline of 2015. *International Journal of Epidemiology* **42** 645–647.
- PEDERSEN, J. and LIU, J. (2012). Child mortality estimation: Appropriate time periods for child mortality estimates from full birth histories. *PLoS Med.* **9** e1001289.
- PLUMMER, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003)*, Vienna. Available at <http://mcmc-jags.sourceforge.net/>.
- RAFTERY, A. E. and LEWIS, S. M. (1992). How many iterations in the Gibbs sampler? In *Bayesian Statistics 4* (J. M. Bernardo et al., eds.) 763–773. Oxford Univ. Press, Oxford.
- RAFTERY, A. E. and LEWIS, S. M. (1996). Implementing MCMC. In *Markov Chain Monte Carlo in Practice* (W. R. Gilks, D. J. Spiegelhalter and S. Richardson, eds.) 115–130. Chapman & Hall, London. MR1397966
- RAJARATNAM, J. K., MARCUS, J., FLAXMAN, A., WANG, H., LEVIN-RECTOR, A., DWYER, L., COSTA, M., LOPEZ, A. and MURRAY, C. (2010). Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: A systematic analysis of progress towards Millennium Development Goal 4. *The Lancet* **375** 1988–2008.
- SULLIVAN, J. M. (2008). An assessment of the credibility of child mortality declines estimated from DHS mortality rates. Working Paper 1, United Nations Children’s Fund, New York. Available at http://www.childinfo.org/files/Overall_Results_of_Analysis.pdf.
- UN SYSTEM TASK TEAM ON THE POST-2015 UN DEVELOPMENT AGENDA (2012). Addressing inequalities: The heart of the post-2015 agenda and the future we want for all. Available at http://www.un.org/millenniumgoals/pdf/10_inequalities_20July.pdf.
- UNITED NATIONS (1983). *Manual X: Indirect Techniques for Demographic Estimation*. United Nations, New York.
- UNITED NATIONS CHILDREN’S FUND AND USAID (2012). Real-time child mortality monitoring meeting. United Nations Children’s Fund, New York. Available at <http://newsletter.childrenandaids.org/real-time-child-mortality-monitoring-meeting-december-19-2012/>.

- UNITED NATIONS CHILDREN'S FUND, DIVISION OF POLICY AND STRATEGY (2013). Committing to child survival: A promise renewed progress report 2013. United Nations Children's Fund, New York.
- UNITED NATIONS, DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS, POPULATION DIVISION (2011). World population prospects. The 2010 Revision.
- WANG, H., DWYER-LINDGREN, L., LOFGREN, K. T., RAJARATNAM, J. K., MARCUS, J. R., LEVIN-RECTOR, A., LEVITZ, C. E., LOPEZ, A. D. and MURRAY, C. J. (2012). Age-specific and sex-specific mortality in 187 countries, 1970–2010: A systematic analysis for the global burden of disease study 2010. *The Lancet* **380** 2071–2094.
- UNITED NATIONS INTER-AGENCY GROUP FOR CHILD MORTALITY ESTIMATION (2012). Levels & trends in child mortality: Report 2012. United Nations Children's Fund, New York.
- UNITED NATIONS INTER-AGENCY GROUP FOR CHILD MORTALITY ESTIMATION (2013). Levels & trends in child mortality: Report 2013. United Nations Children's Fund, New York.

IMPUTATION OF TRUNCATED p -VALUES FOR META-ANALYSIS METHODS AND ITS GENOMIC APPLICATION¹

BY SHAOWU TANG^{*}, YING DING[†], ETIENNE SIBILLE[‡], JEFFREY S. MOGIL[§],
WILLIAM R. LARIVIERE[¶] AND GEORGE C. TSENG[¶]

Roche Molecular Systems, Inc.^{*}, *Pfizer*[†], *University of Toronto*[‡],
McGill University[§] and *University of Pittsburgh*[¶]

Microarray analysis to monitor expression activities in thousands of genes simultaneously has become routine in biomedical research during the past decade. A tremendous amount of expression profiles are generated and stored in the public domain and information integration by meta-analysis to detect differentially expressed (DE) genes has become popular to obtain increased statistical power and validated findings. Methods that aggregate transformed p -value evidence have been widely used in genomic settings, among which Fisher's and Stouffer's methods are the most popular ones. In practice, raw data and p -values of DE evidence are often not available in genomic studies that are to be combined. Instead, only the detected DE gene lists under a certain p -value threshold (e.g., DE genes with p -value < 0.001) are reported in journal publications. The truncated p -value information makes the aforementioned meta-analysis methods inapplicable and researchers are forced to apply a less efficient vote counting method or naïvely drop the studies with incomplete information. The purpose of this paper is to develop effective meta-analysis methods for such situations with partially censored p -values. We developed and compared three imputation methods—mean imputation, single random imputation and multiple imputation—for a general class of evidence aggregation methods of which Fisher's and Stouffer's methods are special examples. The null distribution of each method was analytically derived and subsequent inference and genomic analysis frameworks were established. Simulations were performed to investigate the type I error, power and the control of false discovery rate (FDR) for (correlated) gene expression data. The proposed methods were applied to several genomic applications in colorectal cancer, pain and liquid association analysis of major depressive disorder (MDD). The results showed that imputation methods outperformed existing naïve approaches. Mean imputation and multiple imputation methods performed the best and are recommended for future applications.

REFERENCES

- BEGUM, F., GHOSH, D., TSENG, G. C. and FEINGOLD, E. (2012). Comprehensive literature review and statistical considerations for GWAS meta-analysis. *Nucleic Acids Research* **40** 3777–3784.
BELLOT, G. L., TAN, W. H., TAY, L. L., KOH, D. and WANG, X. (2012). Reliability of tumor primary cultures as a model for drug response prediction: Expression profiles comparison of

Key words and phrases. Microarray analysis, meta-analysis, Fisher's method, Stouffer's method, missing value imputation.

- tissues versus primary cultures from colorectal cancer patients. *J. Cancer Res. Clin. Oncol.* **138** 463–482.
- BENJAMINI, Y. and HOCHBERG, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **57** 289–300. [MR1325392](#)
- BENJAMINI, Y. and YEKUTIELI, D. (2001). The control of the false discovery rate in multiple testing under dependency. *Ann. Statist.* **29** 1165–1188. [MR1869245](#)
- BIANCHINI, M., LEVY, E., ZUCCHINI, C., PINSKI, V., MACAGNO, C., SANCTIS, P. D., VALVASORI, L., CARINCI, P. and MORDOH, J. (2006). Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa. *Int. J. Oncol.* **29** 83–94.
- BIRNBAUM, A. (1954). Combining independent tests of significance. *J. Amer. Statist. Assoc.* **49** 559–574. [MR0065101](#)
- BIRNBAUM, A. (1955). Characterizations of complete classes of tests of some multiparametric hypotheses, with applications to likelihood ratio tests. *Ann. Math. Statist.* **26** 21–36. [MR0067438](#)
- BOROVECKI, F. et al. (2005). Genome-wide expression profiling of human blood reveals biomarkers for huntingtons disease. *Proc. Natl. Acad. Sci. USA* **102** 11023–11028.
- CARDOSO, J. et al. (2007). Expression and genomic profiling of colorectal cancer. *Biochimica et Biophysica Acta-Reviews on Cancer* **1775** 103–137.
- CHAN, S. K., GRIFFITH, O. L., TAI, I. T. and JONES, S. J. M. (2008). Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers. *Cancer Epidemiol Biomarkers Prev.* **17** 543–552.
- CHANG, L.-C., LIN, H.-M., SIBILLE, E. and TSENG, G. C. (2013). Meta-analysis methods for combining multiple expression profiles: Comparisons, statistical characterization and an application guideline. *BMC Bioinformatics* **14** 368.
- CHOI, J. K. et al. (2003). Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* **19** 84–90.
- CHOI, H. et al. (2007). A latent variable approach for meta-analysis of gene expression data from multiple microarray experiments. *BMC Bioinformatics* **8** 364–383.
- DUVAL, S. and TWEEDIE, R. L. (2000a). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics* **56** 455–463.
- DUVAL, S. and TWEEDIE, R. (2000b). A nonparametric “trim and fill” method of accounting for publication bias in meta-analysis. *J. Amer. Statist. Assoc.* **95** 89–98. [MR1803144](#)
- FISHER, R. A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd, Edinburgh.
- GORLOV, I. P., BYUN, J., GORLOVA, O. Y., APARICIO, A. M., EFSTATHIOU, E. and LOGOTHETIS, C. J. (2009). Candidate pathways and genes for prostate cancer: A meta-analysis of gene expression data. *BMC Med. Genomics* **2** 48.
- GRIFFITH, O. L., JONES, S. J. M. and WISEMAN, S. M. (2006). Meta-analysis and meta-review of thyroid cancer gene expression profiling studies identifies important diagnostic biomarkers. *J. Clin. Oncol.* **24** 5043–5051.
- HEDGES, L. V. and OLKIN, I. (1980). Vote-counting methods in research synthesis. *Psychological Bulletin* **88** 359.
- HEDGES, L. V. and OLKIN, I. (1985). *Statistical Methods for Meta-Analysis*. Academic Press, Orlando, FL. [MR0798597](#)
- IOANNIDIS, J. P. A., ALLISON, D. B., BALL, C. A. et al. (2009). Repeatability of published microarray gene expression analysis. *Nature Genetics* **41** 149–155.
- JIANG, X., TAN, J., LI, J., KIVIMÄE, S., YANG, X., ZHUANG, L., LEE, P. L., CHAN, M. T. W., STANTON, L. W., LIU, E. T., CHEYETTE, B. N. R. and YU, Q. (2008). DACT3 is an epigenetic regulator of Wnt/beta-catenin signaling in colorectal cancer and is a therapeutic target of histone modifications. *Cancer Cell* **13** 529–541.
- LACROIX-FRALISH, M. L., AUSTIN, J.-S., ZHENG, F. Y., LEVITIN, D. J. and MOGIL, J. S. (2011). Patterns of pain: Meta-analysis of microarray studies of pain. *Pain* **152** 1888–1898.

- LI, K. C. (2002). Genome-wide coexpression dynamics: Theory and application. *Proc. Natl. Acad. Sci. USA* **99** 16875–16880.
- LI, J. and TSENG, G. C. (2011). An adaptively weighted statistic for detecting differential gene expression when combining multiple transcriptomic studies. *Ann. Appl. Stat.* **5** 994–1019. [MR2840184](#)
- LITTELL, R. C. and FOLKS, J. L. (1971). Asymptotic optimality of Fisher’s method of combining independent tests. *J. Amer. Statist. Assoc.* **66** 193–194. [MR0312634](#)
- LITTELL, R. C. and FOLKS, J. L. (1973). Asymptotic optimality of Fisher’s method of combining independent tests. II. *J. Amer. Statist. Assoc.* **68** 193–194. [MR0375577](#)
- LITTLE, R. J. A. and RUBIN, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd ed. Wiley, Hoboken, NJ. [MR1925014](#)
- MCCARLEY, R. W., WIBLE, C. G., FRUMIN, M., HIRAYASU, Y., LEVITT, J. J. and SHENTON, M. E. (2001). Why vote-count reviews don’t count [letter to the editor]. *Biological Psychiatry* **49** 161–163.
- MOREAU, Y. et al. (2003). Comparison and meta-analysis of microarray data: From the bench to the computer desk. *Trends in Genetics* **19** 570–577.
- OLKIN, I. and SANER, H. (2001). Approximations for trimmed Fisher procedures in research synthesis. *Stat. Methods Med. Res.* **10** 267–276.
- OWEN, A. B. (2009). Karl Pearson’s meta-analysis revisited. *Ann. Statist.* **37** 3867–3892. [MR2572446](#)
- PIROOZANIA, M., NAGARAJAN, V. and DENG, Y. (2007). Gene venn—a web application for comparing gene lists using venn diagram. *Bioinformatics* **1** 420–422.
- RHODES, D. R., BARRETTE, T. R., RUBIN, M. A., GHOSH, D. and CHINNAIYAN, A. M. (2002). Meta-analysis of microarrays: Interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.* **62** 4427–4433.
- ROY, S. N. (1953). On a heuristic method of test construction and its use in multivariate analysis. *Ann. Math. Statistics* **24** 220–238. [MR0057519](#)
- SEGAL, E. et al. (2004). A module map showing conditional activity of expression modules in cancer. *Nature Genetics* **3** 1090–1098.
- SONG, C. and TSENG, G. C. (2014). Hypothesis setting and order statistic for robust Genomic meta-analysis. *Ann. Appl. Stat.* **8** 777–800.
- STOUFFER, S. et al. (1949). *The American Soldier, Volume I: Adjustment During Army Life*. Princeton Univ. Press, Princeton, NJ.
- TANG, S., DING, Y., SIBILLE, E., MOGIL, J. S., LARIVIERE, W. R. and TSENG, G. C. (2014). Supplement to “Imputation of truncated p -values for meta-analysis methods and its genomic application.” DOI:10.1214/14-AOAS747SUPP.
- TSENG, G. C., GHOSH, D. and FEINGOLD, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.* **40** 3785–3799.
- WANG, X., LIN, Y., SONG, C., SIBILLE, E. and TSENG, G. C. (2012). Detecting disease-associated genes with confounding variable adjustment and the impact on genomic meta-analysis: With application to major depressive disorder. *BMC Bioinformatics* **13** 52.

LONGITUDINAL HIGH-DIMENSIONAL PRINCIPAL COMPONENTS ANALYSIS WITH APPLICATION TO DIFFUSION TENSOR IMAGING OF MULTIPLE SCLEROSIS

BY VADIM ZIPUNNIKOV^{*,1}, SONJA GREVEN^{†,2}, HAOCHANG SHOU[‡],
BRIAN S. CAFFO^{*,1}, DANIEL S. REICH^{§,3} AND CIPRIAN M. CRAINICEANU^{*,1}

Johns Hopkins University^{*}, *Ludwig-Maximilians-Universität München*[†],
University of Pennsylvania[‡] and *National Institutes of Health*[§]

We develop a flexible framework for modeling high-dimensional imaging data observed longitudinally. The approach decomposes the observed variability of repeatedly measured high-dimensional observations into three additive components: a subject-specific imaging random intercept that quantifies the cross-sectional variability, a subject-specific imaging slope that quantifies the dynamic irreversible deformation over multiple realizations, and a subject-visit-specific imaging deviation that quantifies exchangeable effects between visits. The proposed method is very fast, scalable to studies including ultrahigh-dimensional data, and can easily be adapted to and executed on modest computing infrastructures. The method is applied to the longitudinal analysis of diffusion tensor imaging (DTI) data of the corpus callosum of multiple sclerosis (MS) subjects. The study includes 176 subjects observed at 466 visits. For each subject and visit the study contains a registered DTI scan of the corpus callosum at roughly 30,000 voxels.

REFERENCES

- 3D-SLICER (2011). <http://www.slicer.org/>.
- ASTON, J. A. D., CHIOU, J.-M. and EVANS, J. P. (2010). Linguistic pitch analysis using functional principal component mixed effect models. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **59** 297–317. [MR2744475](#)
- BIGELOW, J. L. and DUNSON, D. B. (2009). Bayesian semiparametric joint models for functional predictors. *J. Amer. Statist. Assoc.* **104** 26–36. [MR2663031](#)
- BUDAVARI, T., WILD, V., SZALAY, A. S., DOBOS, L. and YIP, C.-W. (2009). Reliable eigenspectra for new generation surveys. *Monthly Notices of the Royal Astronomical Society* **394** 1496–1502.
- CRAINICEANU, C. M., STAIU, A.-M. and DI, C.-Z. (2009). Generalized multilevel functional regression. *J. Amer. Statist. Assoc.* **104** 1550–1561. [MR2750578](#)
- CRAINICEANU, C. M., CAFFO, B. S., LUO, S., ZIPUNNIKOV, V. M. and PUNJABI, N. M. (2011). Population value decomposition, a framework for the analysis of image populations. *J. Amer. Statist. Assoc.* **106** 775–790. [MR2894733](#)
- DEMMELE, J. W. (1997). *Applied Numerical Linear Algebra*. SIAM, Philadelphia, PA. [MR1463942](#)
- DI, C., CRAINICEANU, C. M. and JANK, W. S. (2010). Multilevel sparse functional principal component analysis. *Stat.* **3** 126–143.
- DI, C.-Z., CRAINICEANU, C. M., CAFFO, B. S. and PUNJABI, N. M. (2009). Multilevel functional principal component analysis. *Ann. Appl. Stat.* **3** 458–488. [MR2668715](#)

Key words and phrases. Principal components, linear mixed model, diffusion tensor imaging, brain imaging data, multiple sclerosis.

- EVERSON, R. and ROBERTS, S. (2000). Inferring the eigenvalues of covariance matrices from limited, noisy data. *IEEE Trans. Signal Process.* **48** 2083–2091. [MR1824643](#)
- GOLDSMITH, J., CRAINICEANU, C. M., CAFFO, B. S. and REICH, D. S. (2011). Penalized functional regression analysis of white-matter tract profiles in multiple sclerosis. *NeuroImage* **57** 431–439.
- GOLUB, G. H. and VAN LOAN, C. F. (1996). *Matrix Computations*, 3rd ed. Johns Hopkins Univ. Press, Baltimore, MD. [MR1417720](#)
- GREVEN, S., CRAINICEANU, C., CAFFO, B. and REICH, D. (2010). Longitudinal functional principal component analysis. *Electron. J. Stat.* **4** 1022–1054. [MR2727452](#)
- GUO, W. (2002). Functional mixed effects models. *Biometrics* **58** 121–128. [MR1891050](#)
- HALL, P., MÜLLER, H.-G. and YAO, F. (2008). Modelling sparse generalized longitudinal observations with latent Gaussian processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **70** 703–723. [MR2523900](#)
- HARVILLE, D. (1976). Extension of the Gauss–Markov theorem to include the estimation of random effects. *Ann. Statist.* **4** 384–395. [MR0398007](#)
- HUA, Z. W., DUNSON, D. B., GILMORE, J. H., STYNER, M. and ZHU, H. T. (2012). Semiparametric Bayesian local functional models for diffusion tensor tract statistics. *NeuroImage* **63** 460–474.
- KARHUNEN, K. (1947). Über lineare Methoden in der Wahrscheinlichkeitsrechnung. *Annales Academiæ Scientiarum Fennicæ* **37** 1–79.
- LI, Y., ZHU, H., SHEN, D., LIN, W., GILMORE, J. H. and IBRAHIM, J. G. (2011). Multiscale adaptive regression models for neuroimaging data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **73** 559–578. [MR2853730](#)
- LOËVE, M. (1978). *Probability Theory II*, 4th ed. Springer, New York. [MR0651018](#)
- MCCULLOCH, C. E. and SEARLE, S. R. (2001). *Generalized, Linear, and Mixed Models*. Wiley, New York. [MR1884506](#)
- MINKA, T. P. (2000). Automatic choice of dimensionality for PCA. *Adv. Neural Inf. Process. Syst.* **13** 598–604.
- MIPAV (2011). <http://mipav.cit.nih.gov>.
- MOHAMED, A. and DAVATZIKOS, C. (2004). *Medical Image Computing and Computer-Assisted Intervention*. Springer, Berlin.
- MORI, S. (2007). *Introduction to Diffusion Tensor Imaging*. Elsevier, Amsterdam.
- MORRIS, J. S. and CARROLL, R. J. (2006). Wavelet-based functional mixed models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68** 179–199. [MR2188981](#)
- MORRIS, J. S., BALADANDAYUTHAPANI, V., HERRICK, R. C., SANNA, P. and GUTSTEIN, H. (2011). Automated analysis of quantitative image data using isomorphic functional mixed models, with application to proteomics data. *Ann. Appl. Stat.* **5** 894–923. [MR2840180](#)
- PUJOL, S. (2010). 3D-Slicer (tutorial). National Alliance for Medical Image Computing (NA-MIC).
- RAINE, C. S., MCFARLAND, H. and HOHLFELD, R. (2008). *Multiple Sclerosis: A Comprehensive Text*. Saunders, Philadelphia, PA.
- REICH, D. S., OZTURK, A., CALABRESI, P. A. and MORI, S. (2010). Automated vs conventional tractography in multiple sclerosis: Variability and correlation with disability. *NeuroImage* **49** 3047–3056.
- REISS, P. T. and OGDEN, R. T. (2008). Functional generalized linear models with applications to neuroimaging. In *Poster presentation Workshop on Contemporary Frontiers in High-Dimensional Statistical Data Analysis*, Isaac Newton Institute, University of Cambridge, UK.
- REISS, P. T. and OGDEN, R. T. (2010). Functional generalized linear models with images as predictors. *Biometrics* **66** 61–69. [MR2756691](#)
- REISS, P. T., OGDEN, R. T., MANN, J. and PARSEY, R. V. (2005). Functional logistic regression with PET imaging data: A voxel-level clinical diagnostic tool. *Journal of Cerebral Blood Flow & Metabolism* **25** s635.

- RODRÍGUEZ, A., DUNSON, D. B. and GELFAND, A. E. (2009). Bayesian nonparametric functional data analysis through density estimation. *Biometrika* **96** 149–162. [MR2482141](#)
- ROWEIS, S. (1997). EM algorithms for PCA and SPCA. *Adv. Neural Inf. Process. Syst.* **10** 626–632.
- SHINOHARA, R., CRAINICEANU, C., CAFFO, B., GAITA, M. I. and REICH, D. S. (2011). Population wide model-free quantification of blood-brain-barrier dynamics in multiple sclerosis. *NeuroImage* **57** 1430–1446.
- SHOU, H., ZIPUNNIKOV, V., CRAINICEANU, C. and GREVEN, S. (2013). Structured functional principal component analysis. Available at [arXiv:1304.6783](#).
- STAIU, A.-M., CRAINICEANU, C. M. and CARROLL, R. J. (2010). Fast analysis of spatially correlated multilevel functional data. *Biostatistics* **11** 177–194.
- WENG, J., ZHANG, Y. and HWANG, W.-S. (2003). Candid covariance-free incremental principal component analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25** 1034–1040.
- XIAO, L., RUPPERT, D., ZIPUNNIKOV, V. and CRAINICEANU, C. (2013). Fast covariance estimation for high-dimensional functional data. Available at [arXiv:1306.5718](#).
- YUAN, Y., GILMORE, J. H., GENG, X., STYNER, M., CHEN, K., WANG, J. L. and ZHU, H. (2014). Fmem: Functional mixed effects modeling for the analysis of longitudinal white matter tract data. *NeuroImage* **84** 753–764.
- ZHAO, H., YUEN, P. C. and KWOK, J. T. (2006). A novel incremental principal component analysis and its application for face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics* **36** 873–886.
- ZHU, H., BROWN, P. J. and MORRIS, J. S. (2011). Robust, adaptive functional regression in functional mixed model framework. *J. Amer. Statist. Assoc.* **106** 1167–1179. [MR2894772](#)
- ZIPUNNIKOV, V., CAFFO, B., YOUSEM, D. M., DAVATZIKOS, C., SCHWARTZ, B. S. and CRAINICEANU, C. (2011a). Multilevel functional principal component analysis for high-dimensional data. *J. Comput. Graph. Statist.* **20** 852–873. [MR2878951](#)
- ZIPUNNIKOV, V., CAFFO, B., YOUSEM, D. M., DAVATZIKOS, C., SCHWARTZ, B. S. and CRAINICEANU, C. M. (2011b). Functional principal component models for high dimensional brain volumetrics. *NeuroImage* **58** 772–784.
- ZIPUNNIKOV, V., GREVEN, S., SHOU, H., CAFFO, B., REICH, D. S. and CRAINICEANU, C. (2014). Supplement to “Longitudinal high-dimensional principal components analysis with application to diffusion tensor imaging of multiple sclerosis.” DOI:[10.1214/14-AOAS748SUPP](#).

A NOVEL SPECTRAL METHOD FOR INFERRING GENERAL DIPLOID SELECTION FROM TIME SERIES GENETIC DATA

BY MATTHIAS STEINRÜCKEN^{1,2}, ANAND BHASKAR¹ AND YUN S. SONG¹

University of California, Berkeley

The increased availability of time series genetic variation data from experimental evolution studies and ancient DNA samples has created new opportunities to identify genomic regions under selective pressure and to estimate their associated fitness parameters. However, it is a challenging problem to compute the likelihood of nonneutral models for the population allele frequency dynamics, given the observed temporal DNA data. Here, we develop a novel spectral algorithm to analytically and efficiently integrate over all possible frequency trajectories between consecutive time points. This advance circumvents the limitations of existing methods which require fine-tuning the discretization of the population allele frequency space when numerically approximating requisite integrals. Furthermore, our method is flexible enough to handle general diploid models of selection where the heterozygote and homozygote fitness parameters can take any values, while previous methods focused on only a few restricted models of selection. We demonstrate the utility of our method on simulated data and also apply it to analyze ancient DNA data from genetic loci associated with coat coloration in horses. In contrast to previous studies, our exploration of the full fitness parameter space reveals that a heterozygote advantage form of balancing selection may have been acting on these loci.

REFERENCES

- BOLLBACK, J. P., YORK, T. L. and NIELSEN, R. (2008). Estimation of $2N_e s$ from temporal allele frequency data. *Genetics* **179** 497–502.
- BURKE, M. K., DUNHAM, J. P., SHAHRESTANI, P., THORNTON, K. R., ROSE, M. R. and LONG, A. D. (2010). Genome-wide analysis of a long-term evolution experiment with *Drosophila*. *Nature* **467** 587–590.
- EWENS, W. J. (2004). *Mathematical Population Genetics: I. Theoretical Introduction*, 2nd ed. Springer, New York. MR2026891
- FEARNHEAD, P. (2003). Ancestral processes for non-neutral models of complex diseases. *Theor. Popul. Biol.* **63** 115–130.
- FEARNHEAD, P. (2006). The stationary distribution of allele frequencies when selection acts at unlinked loci. *Theor. Popul. Biol.* **70** 376–386.
- FEDER, A. F., KRYAZHIMSKIY, S. and PLOTKIN, J. B. (2014). Identifying signatures of selection in genetic time series. *Genetics* **196** 509–522.
- GENZ, A. and JOYCE, P. (2003). Computation of the normalizing constant for exponentially weighted Dirichlet distribution integrals. *Computing Science and Statistics* **35** 181–212.

Key words and phrases. Population genetics, spectral method, transition density function, hidden Markov model.

- GREEN, R. E., KRAUSE, J., BRIGGS, A. W., MARICIC, T., STENZEL, U., KIRCHER, M., PATTERSON, N., LI, H., ZHAI, W., FRITZ, M. H.-Y. et al. (2010). A draft sequence of the Neandertal genome. *Science* **328** 710–722.
- GUTENKUNST, R. N., HERNANDEZ, R. D., WILLIAMSON, S. H. and BUSTAMANTE, C. D. (2009). Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet.* **5** e1000695.
- HUMMEL, S., SCHMIDT, D., KREMEYER, B., HERRMANN, B. and OPPERMAN, M. (2005). Detection of the CCR5-Delta32 HIV resistance gene in Bronze Age skeletons. *Genes Immun.* **6** 371–374.
- LANG, G. I., RICE, D. P., HICKMAN, M. J., SODERGREN, E., WEINSTOCK, G. M., BOTSTEIN, D. and DESAI, M. M. (2013). Pervasive genetic hitchhiking and clonal interference in forty evolving yeast populations. *Nature* **500** 571–574.
- LUDWIG, A., PRUVOST, M., REISSMANN, M., BENECKE, N., BROCKMANN, G. A., CAS-TAÑOS, P., CIESLAK, M., LIPPOLD, S., LLORENTE, L., MALASPINAS, A.-S., SLATKIN, M. and HOFREITER, M. (2009). Coat color variation at the beginning of horse domestication. *Science* **324** 485.
- LUKIĆ, S., HEY, J. and CHEN, K. (2011). Non-equilibrium allele frequency spectra via spectral methods. *Theor. Popul. Biol.* **79** 203–219.
- MALASPINAS, A. S., MALASPINAS, O., EVANS, S. N. and SLATKIN, M. (2012). Estimating allele age and selection coefficient from time-serial data. *Genetics* **192** 599–607.
- MATHAR, R. J. (2009). A Java Math.BigDecimal implementation of core mathematical functions. Available at [arXiv:0908.3030](https://arxiv.org/abs/0908.3030).
- MATHIESON, I. and MCVEAN, G. (2013). Estimating selection coefficients in spatially structured populations from time series data of allele frequencies. *Genetics* **193** 973–984.
- ORLANDO, L., GINOLHAC, A., ZHANG, G., FROESE, D., ALBRECHTSEN, A., STILLER, M., SCHUBERT, M., CAPPELLINI, E., PETERSEN, B., MOLTKE, I. et al. (2013). Recalibrating *Equus* evolution using the genome sequence of an early Middle Pleistocene horse. *Nature* **499** 74–78.
- OROZCO-TERWENGEL, P., KAPUN, M., NOLTE, V., KOFLER, R., FLATT, T. and SCHLÖT-TERER, C. (2012). Adaptation of *Drosophila* to a novel laboratory environment reveals temporally heterogeneous trajectories of selected alleles. *Mol. Ecol.* **21** 4931–4941.
- PRESS, W. H., TEUKOLSKY, S. A., VETTERLING, W. T. and FLANNERY, B. P. (2007). *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. Cambridge Univ. Press, Cambridge. [MR2371990](https://arxiv.org/abs/1203.7071)
- REICH, D., GREEN, R. E., KIRCHER, M., KRAUSE, J., PATTERSON, N., DURAND, E. Y., VIOLA, B., BRIGGS, A. W., STENZEL, U., JOHNSON, P. L. F. et al. (2010). Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* **468** 1053–1060.
- SHANKARAPPA, R., MARGOLICK, J. B., GANGE, S. J., RODRIGO, A. G., UPCHURCH, D., FARZADEGAN, H., GUPTA, P., RINALDO, C. R., LEARN, G. H., HE, X., HUANG, X. L. and MULLINS, J. I. (1999). Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J. Virol.* **73** 10489–10502.
- SONG, Y. S. and STEINRÜCKEN, M. (2012). A simple method for finding explicit analytic transition densities of diffusion processes with general diploid selection. *Genetics* **190** 1117–1129.
- STEINRÜCKEN, M., BHASKAR, A. and SONG, Y. (2014). Supplement to “A novel spectral method for inferring general diploid selection from time series genetic data.” DOI:[10.1214/14-AOAS764SUPP](https://doi.org/10.1214/14-AOAS764SUPP).
- STEINRÜCKEN, M., WANG, Y. X. R. and SONG, Y. S. (2013). An explicit transition density expansion for a multi-allelic Wright–Fisher diffusion with general diploid selection. *Theor. Popul. Biol.* **83** 1–14.
- STEPHENS, M. and DONNELLY, P. (2003). Ancestral inference in population genetics models with selection (with discussion). *Aust. N. Z. J. Stat.* **45** 395–430. [MR2018460](https://arxiv.org/abs/1203.7071)

- WILLIAMSON, E. G. and SLATKIN, M. (1999). Using maximum likelihood to estimate population size from temporal changes in allele frequencies. *Genetics* **152** 755–761.
- WISER, M. J., RIBECK, N. and LENSKI, R. E. (2013). Long-term dynamics of adaptation in asexual populations. *Science* **342** 1364–1367.

SPATIO-TEMPORAL MODELLING OF EXTREME STORMS¹

BY THEODOROS ECONOMOU, DAVID B. STEPHENSON
AND CHRISTOPHER A. T. FERRO

University of Exeter

A flexible spatio-temporal model is implemented to analyse extreme extra-tropical cyclones objectively identified over the Atlantic and Europe in 6-hourly re-analyses from 1979–2009. Spatial variation in the extremal properties of the cyclones is captured using a 150 cell spatial regularisation, latitude as a covariate, and spatial random effects. The North Atlantic Oscillation (NAO) is also used as a covariate and is found to have a significant effect on intensifying extremal storm behaviour, especially over Northern Europe and the Iberian peninsula. Estimates of lower bounds on minimum sea-level pressure are typically 10–50 hPa below the minimum values observed for historical storms with largest differences occurring when the NAO index is positive.

REFERENCES

- BAILEY, T. C. and GATRELL, A. C. (1995). *Interactive Spatial Data Analysis*. Longman, Harlow.
- BANERJEE, S., CARLIN, B. P. and GELFAND, A. E. (2004). *Hierarchical Modelling and Analysis for Spatial Data*. Chapman & Hall, London.
- BONAZZI, A., CUSACK, S., MITAS, C. and JEWSON, S. (2012). The spatial structure of European wind cyclones as characterized by bivariate extreme-value copulas. *Nat. Hazards Earth Syst. Sci.* **12** 1769–1782.
- CASSON, E. and COLES, S. (1999). Spatial regression models for extremes. *Extremes* **1** 449–468.
- COLES, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer, London. [MR1932132](#)
- COLES, S., HEFFERNAN, J. and TAWN, J. (1999). Dependence measures for extreme value analyses. *Extremes* **2** 339–365.
- COOLEY, D., NYCHKA, D. and NAVEAU, P. (2007). Bayesian spatial modeling of extreme precipitation return levels. *J. Amer. Statist. Assoc.* **102** 824–840. [MR2411647](#)
- COOLEY, D. and SAIN, S. R. (2010). Spatial hierarchical modeling of precipitation extremes from a regional climate model. *J. Agric. Biol. Environ. Stat.* **15** 381–402. [MR2787265](#)
- DAVISON, A. C., PADOAN, S. A. and RIBATET, M. (2012). Statistical modeling of spatial extremes. *Statist. Sci.* **27** 161–186. [MR2963980](#)
- DE HAAN, L. and FERREIRA, A. (2006). *Extreme Value Theory: An Introduction*. Springer, New York. [MR2234156](#)
- DELLA-MARTA, P. M. and PINTO, J. G. (2009). Statistical uncertainty of changes in winter cyclones over the North Atlantic and Europe in an ensemble of transient climate simulations. *Geophysical Research Letters* **36** 1–5.

Key words and phrases. Bayesian hierarchical model, spatial random effects, natural hazards, extra-tropical cyclones, extreme value distribution, European windstorms.

- DELLA-MARTA, P. M., MATHIS, H., FREI, C., LINIGER, M. A., KLEINN, J. and APPENZELLER, C. (2009). The return period of wind cyclones over Europe. *International Journal of Climatology* **29** 437–459.
- EINMAHL, J. H. J. and MAGNUS, J. R. (2008). Records in athletics through extreme-value theory. *J. Amer. Statist. Assoc.* **103** 1382–1391. [MR2504198](#)
- ELSNER, J. B., HODGES, R. E. and JAGGER, T. H. (2012). Spatial grids for hurricane climate research. *Climate Dynamics* **39** 21–36.
- FRIGESSI, A., HAUG, O. and RUE, H. (2002). A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes* **5** 219–235. [MR1995776](#)
- GAETAN, C. and GRIGOLETTO, M. (2007). A hierarchical model for the analysis of spatial rainfall extremes. *J. Agric. Biol. Environ. Stat.* **12** 434–449. [MR2405533](#)
- GELMAN, A., CARLIN, J. B., STERN, H. S., DUNSON, D. B., VEHTARI, A. and RUBIN, D. B. (2013). *Bayesian Data Analysis*. CRC Press, Boca Raton, FL.
- 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](#)
- HEATON, M. J., KATZFUSS, M., RAMACHANDAR, S., PEDINGS, K., GILLELAND, E., MANNSHARDT-SHAMSELDIN, E. and SMITH, R. L. (2011). Spatio-temporal models for large-scale indicators of extreme weather. *Environmetrics* **22** 294–303. [MR2843385](#)
- HODGES, K. (1994). A general method for tracking analysis and its application to meteorological data. *Monthly Weather Review* **122** 2573–2585.
- ILLIAN, J., PENTTINEN, A., STOYAN, H. and STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester. [MR2384630](#)
- KATZ, R. W. (2010). Statistics of extremes in climate change. *Climate Change* **100** 71–76.
- LIONELLO, P., BOLDRIN, U. and GIORGI, F. (2008). Future changes in cyclone climatology over Europe as inferred from a regional climate simulation. *Climate Dynamics* **30** 657–671.
- MAILIER, P. J., STEPHENSON, D. B., FERRO, C. A. T. and HODGES, K. I. (2006). The seriality of extratropical cyclones. *Bulletin of the American Meteorological Society* **87** 1317–1318.
- MITCHELL-WALLACE, K. and MITCHELL, A. (2007). Willis' detailed report for wintercyclone Kyrill. Available at <http://www.willisresearchnetwork.com>.
- NISSEN, K. M., LECKEBUSCH, G. C., PINTO, J. G., RENGGLI, D., ULBRICH, S. and ULBRICH, U. (2010). Cyclones causing wind cyclones in the Mediterranean: Characteristics, trends and links to large-scale patterns. *Natural Hazards and Earth System Science* **10** 1379–1391.
- ODELL, L., KNIPPERTZ, P., PICKERING, S., PARKES, B. and ROBERTS, A. (2013). The Braer cyclone revisited. *Weather* **68** 105–111.
- PINTO, J. G., ZACHARIAS, S., FINK, A. H., LECKENBUSCH, G. C. and ULBRICH, U. (2009). Factors contributing to the development of extreme North Atlantic cyclones and their relationship with the NAO. *Climate Dynamics* **32** 711–737.
- R DEVELOPMENT CORE TEAM (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- REICH, B. J. and SHABY, B. A. (2012). A hierarchical max-stable spatial model for extreme precipitation. *Ann. Appl. Stat.* **6** 1430–1451. [MR3058670](#)
- 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](#)
- SAHA, S. ET AL. (2010). The NCEP climate forecast system reanalyses. *Bulletin of the American Meteorological Society* **91** 1015–1057.
- SANG, H. and GELFAND, A. E. (2009). Hierarchical modeling for extreme values observed over space and time. *Environ. Ecol. Stat.* **16** 407–426. [MR2749848](#)
- SANG, H. and GELFAND, A. E. (2010). Continuous spatial process models for spatial extreme values. *J. Agric. Biol. Environ. Stat.* **15** 49–65. [MR2755384](#)

- SEIERSTAD, I. A., STEPHENSON, D. B. and KVAMSTO, N. G. (2006). How useful are teleconnection patterns for explaining variability in extratropical storminess. *Tellus A* **59** 170–181.
- SIENZ, F., SCHNEIDERREIT, A., BLENDER, R., FRAEDRICH, K. and LUNKEIT, F. (2010). Extreme value statistics for North Atlantic cyclones. *Tellus A* **62** 347–360.
- SMITH, R. L. (1989). Extreme value analysis of environmental time series: An application to trend detection in ground-level ozone. *Statist. Sci.* **4** 367–393. [MR1041763](#)
- 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](#)
- STEPHENSON, D. B., CASATI, B., FERRO, C. A. T. and WILSON, C. A. (2008). The extreme dependency score: A non-vanishing measure for forecasts of rare events. *Meteorological Applications* **15** 41–50.
- TURKMAN, K. F., TURKMAN, M. A. A. and PEREIRA, J. M. (2010). Asymptotic models and inference for extremes of spatio-temporal data. *Extremes* **13** 375–397. [MR2733939](#)
- ULBRICH, U., FINK, A. H., KLAWA, M. and PINTO, J. G. (2001). Three extreme cyclones over Europe in December 1999. *Weather* **56** 70–80.
- VITOLO, R., STEPHENSON, D. B., COOK, I. M. and MITCHELL-WALLACE, K. (2009). Serial clustering of intense European storms. *Meteorologische Zeitschrift* **18** 411–424.

VORONOI RESIDUAL ANALYSIS OF SPATIAL POINT PROCESS MODELS WITH APPLICATIONS TO CALIFORNIA EARTHQUAKE FORECASTS

BY ANDREW BRAY*, KA WONG[†], CHRISTOPHER D. BARR[‡]
AND FREDERIC PAIK SCHOENBERG[§]

*University of Massachusetts, Amherst**, *Google[†]*, *Yale University[‡]*
and University of California, Los Angeles[§]

Many point process models have been proposed for describing and forecasting earthquake occurrences in seismically active zones such as California, but the problem of how best to compare and evaluate the goodness of fit of such models remains open. Existing techniques typically suffer from low power, especially when used for models with very volatile conditional intensities such as those used to describe earthquake clusters. This paper proposes a new residual analysis method for spatial or spatial–temporal point processes involving inspecting the differences between the modeled conditional intensity and the observed number of points over the Voronoi cells generated by the observations. The resulting residuals can be used to construct diagnostic methods of greater statistical power than residuals based on rectangular grids.

Following an evaluation of performance using simulated data, the suggested method is used to compare the Epidemic-Type Aftershock Sequence (ETAS) model to the Hector Mine earthquake catalog. The proposed residuals indicate that the ETAS model with uniform background rate appears to slightly but systematically underpredict seismicity along the fault and to overpredict seismicity in along the periphery of the fault.

REFERENCES

- ADELFO, G. and SCHOENBERG, F. P. (2009). Point process diagnostics based on weighted second-order statistics and their asymptotic properties. *Ann. Inst. Statist. Math.* **61** 929–948. [MR2556772](#)
- BADDELEY, A., MØLLER, J. and PAKES, A. G. (2008). Properties of residuals for spatial point processes. *Ann. Inst. Statist. Math.* **60** 627–649. [MR2434415](#)
- BADDELEY, A., RUBAK, E. and MØLLER, J. (2011). Score, pseudo-score and residual diagnostics for spatial point process models. *Statist. Sci.* **26** 613–646. [MR2951394](#)
- BADDELEY, A. and TURNER, R. (2000). Practical maximum pseudolikelihood for spatial point patterns (with discussion). *Aust. N. Z. J. Stat.* **42** 283–322. [MR1794056](#)
- BADDELEY, A., TURNER, R., MØLLER, J. and HAZELTON, M. (2005). Residual analysis for spatial point processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** 617–666. [MR2210685](#)
- BARR, C. D. and DIEZ, D. M. (2012). Sizes of Voronoi regions in a spatial network designed by an inhomogeneous Poisson process. Unpublished manuscript.
- BARR, C. D. and SCHOENBERG, F. P. (2010). On the Voronoi estimator for the intensity of an inhomogeneous planar Poisson process. *Biometrika* **97** 977–984. [MR2746166](#)

Key words and phrases. Epidemic-Type Aftershock Sequence models, Hector Mine, residuals analysis, point patterns, seismology, Voronoi tessellations.

- BRAY, A. and SCHOENBERG, F. P. (2013). Assessment of point process models for earthquake forecasting. *Statist. Sci.* **28** 510–520. [MR3161585](#)
- CLEMENTS, R. A., SCHOENBERG, F. P. and SCHORLEMMER, D. (2011). Residual analysis methods for space-time point processes with applications to earthquake forecast models in California. *Ann. Appl. Stat.* **5** 2549–2571. [MR2907126](#)
- CLEMENTS, R. A., SCHOENBERG, F. P. and VEEN, A. (2012). Evaluation of space-time point process models using super-thinning. *Environmetrics* **23** 606–616. [MR3020078](#)
- CZADO, C., GNEITING, T. and HELD, L. (2009). Predictive model assessment for count data. *Biometrics* **65** 1254–1261. [MR2756513](#)
- DALEY, D. J. and VERE-JONES, D. (1988). *An Introduction to the Theory of Point Processes*. Springer, New York. [MR0950166](#)
- DAWID, A. P. (1984). Present position and potential developments: Some personal views: Statistical theory: The prequential approach. *J. Roy. Statist. Soc. Ser. A* **147** 278–292. [MR0763811](#)
- DU, Q., FABER, V. and GUNZBURGER, M. (1999). Centroidal Voronoi tessellations: Applications and algorithms. *SIAM Rev.* **41** 637–676. [MR1722997](#)
- FIELD, E. H. (2007). Overview of the working group for the development of regional earthquake models (RELM). *Seismol. Res. Lett.* **78** 7–16.
- GUAN, Y. (2008). A goodness-of-fit test for inhomogeneous spatial Poisson processes. *Biometrika* **95** 831–845. [MR2461214](#)
- HELMSTETTER, A., KAGAN, Y. Y. and JACKSON, D. D. (2007). High-resolution time-independent grid-based forecast $M \geq 5$ earthquakes in California. *Seismol. Res. Lett.* **78** 78–86.
- HELMSTETTER, A. and WERNER, M. (2012). Adaptive spatio-temporal smoothing of seismicity for long-term earthquake forecasts in California. *Bull. Seismol. Soc. Amer.* **102** 2518–2529.
- HINDE, A. L. and MILES, R. E. (1980). Monte Carlo estimates of the distributions of the random polygons of the Voronoi tessellation with respect to a Poisson process. *J. Stat. Comput. Simul.* **10** 205–223.
- JACKSON, D. D. and KAGAN, Y. Y. (1999). Testable earthquake forecasts for 1999. *Seismol. Res. Lett.* **70** 393–403.
- JOHNSON, E. A. and MIYANISHI, K. (2001). *Forest Fire: Behavior and Ecological Effects*. Academic Press, San Diego.
- JORDAN, T. H. (2006). Earthquake predictability, brick by brick. *Seismol. Res. Lett.* **77** 3–6.
- KEELEY, J. E., SAFFORD, H., FOTHERINGHAM, C. J. et al. (2009). The 2007 Southern California wildfires: Lessons in complexity. *J. For.* **September** 287–296.
- LAWSON, A. (1993). A deviance residual for heterogeneous spatial Poisson processes. *Biometrics* **49** 889–897.
- LAWSON, A. (2005). Comment on “Residual analysis for spatial point processes” by A. Baddeley, R. Turner, J. Møller and M. Hazelton. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** 654.
- MALAMUD, B. D., MILLINGTON, J. D. A. and PERRY, G. L. W. (2005). Characterizing wildfire regimes in the United States. *Proc. Natl. Acad. Sci. USA* **102** 4694–4699.
- MASSEY, F. (1951). The Kolmogorov–Smirnov test for goodness of fit. *J. Amer. Statist. Assoc.* **42** 68–78.
- MEIJERING, J. L. (1953). Interface area, edge length, and number of vertices in crystal aggregation with random nucleation. *Philips Res. Rep.* **8** 270–290.
- MEYER, P. A. (1971). Démonstration simplifiée d’un théorème de Knight. In *Séminaire de Probabilités V Université de Strasbourg. Lecture Notes in Mathematics* **191** 191–195. Springer, Heidelberg.
- OGATA, Y. (1998). Space-time point process models for earthquake occurrences. *Ann. Inst. Statist. Math.* **50** 379–402.
- OGATA, Y. (2011). Significant improvements of the space-time ETAS model for forecasting of accurate baseline seismicity. *Earth, Planets and Space* **63** 217–229.

- OGATA, Y., JONES, L. M. and TODA, S. (2003). When and where the aftershock activity was depressed: Contrasting decay patterns of the proximate large earthquakes in southern California. *J. Geophys. Res.* **108** 2318–2329.
- OKABE, A., BOOTS, B., SUGIHARA, K. and CHIU, S. (2000). *Spatial Tessellations*, 2nd ed. Wiley, Chichester.
- RHOADES, D. A., SCHORLEMMER, D., GERSTENBERGER, M. C. et al. (2011). Efficient testing of earthquake forecasting models. *Acta Geophys.* **59** 728–747.
- RIPLEY, B. D. (1976). The second-order analysis of stationary point processes. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **39** 172–212.
- SCHOENBERG, F. (1999). Transforming spatial point processes into Poisson processes. *Stochastic Process. Appl.* **81** 155–164. [MR1694573](#)
- SCHOENBERG, F. P. (2003). Multidimensional residual analysis of point process models for earthquake occurrences. *J. Amer. Statist. Assoc.* **98** 789–795. [MR2055487](#)
- SCHOENBERG, F. P. (2013). Facilitated estimation of ETAS. *Bull. Seismol. Soc. Amer.* **103** 1–7.
- SCHORLEMMER, D. and GERSTENBERGER, M. C. (2007). RELM testing center. *Seismol. Res. Lett.* **78** 30–35.
- SCHORLEMMER, D., GERSTENBERGER, M. C., WEIMER, S., JACKSON, D. D. and RHOADES, D. A. (2007). Earthquake likelihood model testing. *Seismol. Res. Lett.* **78** 17–27.
- SCHORLEMMER, D., ZECHAR, J. D., WERNER, M. J., FIELD, E. H., JACKSON, D. D., JORDAN, T. H. and THE RELM WORKING GROUP (2010). First results of the regional earthquake likelihood models experiment. *Pure Appl. Geophys.* **167** 859–876.
- TANEMURA, M. (2003). Statistical distributions of Poisson Voronoi cells in two and three dimensions. *Forma* **18** 221–247. [MR2040084](#)
- THORARINSDOTTIR, T. L. (2013). Calibration diagnostics for point process models via the probability integral transform. *Stat* **2** 150–158.
- VEEN, A. and SCHOENBERG, F. P. (2006). Assessing spatial point process models for California earthquakes using weighted K-functions: Analysis of California earthquakes. In *Case Studies in Spatial Point Process Modeling* (A. Baddeley, P. Gregori, J. Mateu, R. Stoica and D. Stoyan, eds.) 293–306. Springer, New York. [MR2229141](#)
- XU, H. and SCHOENBERG, F. P. (2011). Point process modeling of wildfire hazard in Los Angeles County, California. *Ann. Appl. Stat.* **5** 684–704. [MR2840171](#)
- ZECHAR, J. D., GERSTENBERGER, M. C. and RHOADES, D. A. (2010). Likelihood-based tests for evaluating space-rate-magnitude earthquake forecasts. *Bull. Seismol. Soc. Amer.* **100** 1184–1195.
- ZECHAR, J. D., SCHORLEMMER, D., WERNER, M. J., GERSTENBERGER, M. C., RHOADES, D. A. and JORDAN, T. H. (2013). Regional earthquake likelihood models I: First-order results. *Bull. Seismol. Soc. Amer.* **103** 787–798.
- ZHUANG, J., HARTE, D., WERNER, M. J., HAINZL, S. and ZHOU, S. (2012). Basic models of seismicity: Temporal models. *Community Online Resource for Statistical Seismicity Analysis*. DOI:10.5078/corssa-79905851.

LONGITUDINAL MIXED MEMBERSHIP TRAJECTORY MODELS FOR DISABILITY SURVEY DATA¹

BY DANIEL MANRIQUE-VALLIER

Indiana University

We develop methods for analyzing discrete multivariate longitudinal data and apply them to functional disability data on the U.S. elderly population from the National Long Term Care Survey (NLTCSS), 1982–2004. Our models build on a Mixed Membership framework, in which individuals are allowed multiple membership on a set of extreme profiles characterized by time-dependent trajectories of progression into disability. We also develop an extension that allows us to incorporate birth-cohort effects, in order to assess inter-generational changes. Applying these methods, we find that most individuals follow trajectories that imply a late onset of disability, and that younger cohorts tend to develop disabilities at a later stage in life compared to their elders.

REFERENCES

- AIROLDI, E. M., FIENBERG, S. E., JOUTARD, C. and LOVE, T. M. (2007). Discovering latent patterns with hierarchical Bayesian mixed-membership models. In *Data Mining Patterns: New Methods and Applications* (P. Poncelet, F. Masseglia and M. Teisseire, eds.) 240–275. Idea Group Inc., Hershey, PA.
- 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.
- AIROLDI, E. M., EROSHEVA, E. A., FIENBERG, S. E., JOUTARD, C., LOVE, T. and SHRINGARPURE, S. (2010). Reconceptualizing the classification of PNAS articles. *Proc. Natl. Acad. Sci. USA* **107** 20899–20904.
- AKAIKE, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory* 267–281. Akademiai Kiado, Budapest.
- BERTOLET, M. (2008). To weight or not to weight? Incorporating sampling designs into model-based analyses. Ph.D. thesis, Dept. Statistics, Carnegie Mellon Univ., Pittsburgh, PA.
- BHATTACHARYA, A. and DUNSON, D. B. (2012). Simplex factor models for multivariate unordered categorical data. *J. Amer. Statist. Assoc.* **107** 362–377. MR2949366
- CLARK, R. F. (1998). An introduction to the National Long-Term Care Survey. Office of Disability, Aging, and Long-Term Care Policy with the U.S. Dept. Health and Human Services. Available at <http://aspe.hhs.gov/daltcp/reports/nltcssu2.htm>.
- CONNOR, J. T. (2006). Multivariate mixture models to describe longitudinal patterns of frailty in American seniors. Ph.D. thesis, Dept. Statistics & H. John Heinz III School of Public Policy & Management, Carnegie Mellon Univ., Pittsburgh, PA.
- CONNOR, J. T., FIENBERG, S. E., EROSHEVA, E. A. and WHITE, T. (2006). Towards a restructuring of the National Long Term Care Survey: A longitudinal perspective. Prepared for presentation at an Expert Panel Meeting on the National Long Term Care Survey, Committee on National Statistics, National Research Council.

Key words and phrases. NLTCSS, Mixed Membership, trajectories, multivariate analysis, MCMC, cohort analysis.

- CORDER, L. S. and MANTON, K. G. (1991). National surveys and the health and functioning of the elderly: The effects of design and content. *J. Amer. Statist. Assoc.* **86** 513–525.
- EROSHEVA, E. A. (2002). Grade of membership and latent structures with application to disability survey data. Ph.D. thesis, Dept. Statistics, Carnegie Mellon Univ., Pittsburgh, PA.
- EROSHEVA, E. A. and FIENBERG, S. E. (2005). Bayesian mixed membership models for soft clustering and classification. In *Classification—The Ubiquitous Challenge* (C. Weihs and W. Gaul, eds.) 11–26. Springer, Berlin.
- EROSHEVA, E. A., FIENBERG, S. E. and JOUTARD, C. (2007). Describing disability through individual-level mixture models for multivariate binary data. *Ann. Appl. Stat.* **1** 502–537. [MR2415745](#)
- EROSHEVA, E. A., FIENBERG, S. E. and LAFFERTY, J. D. (2004). Mixed-membership models of scientific publications. *Proc. Natl. Acad. Sci. USA* **101** 5220–5227.
- FERRUCCI, L., GURALNIK, J. M., SIMONSICK, E., SALIVE, M. E., CORTI, C. and LANGLOIS, J. (1996). Progressive versus catastrophic disability: A longitudinal view of the disablement process. *The Journals of Gerontology: Series A* **51** M123.
- GOODMAN, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika* **61** 215–231. [MR0370936](#)
- HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed. Springer, New York. [MR2722294](#)
- JASRA, A., HOLMES, C. C. and STEPHENS, D. A. (2005). Markov chain Monte Carlo methods and the label switching problem in Bayesian mixture modeling. *Statist. Sci.* **20** 50–67. [MR2182987](#)
- KREUTER, F. and MUTHÉN, B. (2008). Analyzing criminal trajectory profiles: Bridging multilevel and group-based approaches using growth mixture modeling. *J. Quant. Criminol.* **24** 1–31.
- KURLAND, B. F. and HEAGERTY, P. J. (2005). Directly parameterized regression conditioning on being alive: Analysis of longitudinal data truncated by deaths. *Biostatistics* **6** 241–258.
- KURLAND, B. F., JOHNSON, L. L., EGGLESTON, B. L. and DIEHR, P. H. (2009). Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. *Statist. Sci.* **24** 211–222. [MR2655850](#)
- MANRIQUE-VALLIER, D. (2014). Mixed membership trajectory models. In *Handbook on Mixed Membership Models* (E. M. Airoldi, D. M. Blei, E. A. Erosheva and S. E. Fienberg, eds.). Chapman & Hall/CRC, London.
- MANRIQUE-VALLIER, D. (2014). Supplement to “Longitudinal Mixed Membership trajectory models for disability survey data.” DOI:[10.1214/14-AOAS769SUPP](#).
- MANRIQUE-VALLIER, D. and FIENBERG, S. E. (2008). Population size estimation using individual level mixture models. *Biom. J.* **50** 1051–1063. [MR2649394](#)
- MANTON, K. G. (2008). Recent declines in chronic disability in the elderly U.S. population: Risk factors and future dynamics. *Annu. Rev. Public Health* **29** 91–113.
- MANTON, K. G., CORDER, L. and STALLARD, E. (1997). Chronic disability trends in elderly United States populations: 1982–1994. *Proc. Natl. Acad. Sci. USA* **94** 2593–2598.
- MANTON, K. G., GU, X. L. and LAMB, V. L. (2006). Change in chronic disability from 1982 to 2004/2005 as measured by long-term changes in function and health in the US elderly population. *Proc. Natl. Acad. Sci. USA* **103** 18374.
- MANTON, K. G., LAMB, V. L. and GU, X. (2007). Medicare cost effects of recent US disability trends in the elderly future implications. *J. Aging Health* **19** 359–381.
- MANTON, K. G., STALLARD, E. and WOODBURY, M. A. (1991). A multivariate event history model based upon fuzzy states: Estimation from longitudinal surveys with informative nonresponse. *J. Off. Stat.* **7** 261–293.
- NAGIN, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. *Psychol. Methods* **4** 139–157.

- RAFTERY, A. E., NEWTON, M. A., SATAGOPAN, J. M. and KRIVITSKY, P. N. (2007). Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. In *Bayesian Statistics* (J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith and M. West, eds.). *Oxford Sci. Publ.* **8** 371–416. Oxford Univ. Press, Oxford. [MR2433201](#)
- SCHWARZ, G. (1978). Estimating the dimension of a model. *Ann. Statist.* **6** 461–464. [MR0468014](#)
- 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](#)
- STALLARD, E. (2005). Trajectories of morbidity, disability, and mortality among the US elderly population: Evidence from the 1984–1999 NLTCS. *N. Amer. Actuar. J.* **11** 16–53.
- TANNER, M. A. (1996). *Tools for Statistical Inference: Methods for the Exploration of Posterior Distributions and Likelihood Functions*, 3rd ed. Springer, New York. [MR1396311](#)
- WHITE, T. A. and EROSHEVA, E. A. (2013). Using group-based latent class transition models to analyze chronic disability data from the National Long-Term Care Survey 1984–2004. *Stat. Med.* **32** 3569–3589. [MR3092251](#)
- WOODBURY, M. A., CLIVE, J. and GARSON, A. JR. (1978). Mathematical typology: A grade of membership technique for obtaining disease definition. *Comput. Biomed. Res.* **11** 277–98.
- XING, E. P., FU, W. and SONG, L. (2010). A state-space mixed membership blockmodel for dynamic network tomography. *Ann. Appl. Stat.* **4** 535–566. [MR2758639](#)

A FAST ALGORITHM FOR DETECTING GENE–GENE INTERACTIONS IN GENOME-WIDE ASSOCIATION STUDIES

BY JIAHAN LI*, WEI ZHONG^{†,1}, RUNZE LI^{‡,2} AND RONGLING WU[‡]

*University of Notre Dame**, *Xiamen University[†]*
and Pennsylvania State University[‡]

With the recent advent of high-throughput genotyping techniques, genetic data for genome-wide association studies (GWAS) have become increasingly available, which entails the development of efficient and effective statistical approaches. Although many such approaches have been developed and used to identify single-nucleotide polymorphisms (SNPs) that are associated with complex traits or diseases, few are able to detect gene–gene interactions among different SNPs. Genetic interactions, also known as epistasis, have been recognized to play a pivotal role in contributing to the genetic variation of phenotypic traits. However, because of an extremely large number of SNP–SNP combinations in GWAS, the model dimensionality can quickly become so overwhelming that no prevailing variable selection methods are capable of handling this problem. In this paper, we present a statistical framework for characterizing main genetic effects and epistatic interactions in a GWAS study. Specifically, we first propose a two-stage sure independence screening (TS-SIS) procedure and generate a pool of candidate SNPs and interactions, which serve as predictors to explain and predict the phenotypes of a complex trait. We also propose a rates adjusted thresholding estimation (RATE) approach to determine the size of the reduced model selected by an independence screening. Regularization regression methods, such as LASSO or SCAD, are then applied to further identify important genetic effects. Simulation studies show that the TS-SIS procedure is computationally efficient and has an outstanding finite sample performance in selecting potential SNPs as well as gene–gene interactions. We apply the proposed framework to analyze an ultrahigh-dimensional GWAS data set from the Framingham Heart Study, and select 23 active SNPs and 24 active epistatic interactions for the body mass index variation. It shows the capability of our procedure to resolve the complexity of genetic control.

REFERENCES

- ALTSHULER, D., DALY, M. J. and LANDER, E. S. (2008). Genetic mapping in human disease. *Science* **322** 881–888.
- AYERS, K. L. and CORDELL, H. J. (2010). SNP selection in genome-wide and candidate gene studies via penalized logistic regression. *Genet. Epidemiol.* **34** 879–891.
- BREIMAN, L. (2001). Random forests. *Mach. Learn.* **45** 5–32.

Key words and phrases. Gene–gene interaction, GWAS, high-dimensional data, sure independence screening, variable selection.

- BURTON, P. R., CLAYTON, D. G., CARDON, L. R., CRADDOCK, N., DELOUKAS, P., DUNCANSON, A. et al. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447** 661–678.
- CHIPMAN, H. (1996). Bayesian variable selection with related predictors. *Canad. J. Statist.* **24** 17–36. [MR1394738](#)
- CHO, S., KIM, H., OH, S., KIM, K. and PARK, T. (2009). Elastic-net regularization approaches for genome-wide association studies of rheumatoid arthritis. *BMC Proceedings* **3** S25.
- CORDELL, H. J. (2009). Detecting gene-gene interactions that underlie human diseases. *Nat. Rev. Genet.* **10** 392–404.
- DALY, A. K. (2010). Genome-wide association studies in pharmacogenomics. *Nat. Rev. Genet.* **11** 241–246.
- DAS, K., LI, J., WANG, Z., TONG, C., FU, G., LI, Y., XU, M., AHN, K., MAUGER, D., LI, R. and WU, R. (2011). A dynamic model for genome-wide association studies. *Hum. Genet.* **8** 1–8.
- DAWBER, T. R., MEADORS, G. F. and MOORE, F. E. (1951). Epidemiological approaches to heart disease: The Framingham study. *Am. J. Publ. Health* **41** 279–286.
- DUGGAL, P., GILLANDERS, E. M., HOLMES, T. N. and BAILEY-WILSON, J. E. (2008). Establishing an adjusted p-value threshold to control the family-wide type 1 error in genome wide association studies. *BMC Genomics* **9** 516.
- FAN, J. and LI, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *J. Amer. Statist. Assoc.* **96** 1348–1360. [MR1946581](#)
- FAN, J. and LV, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **70** 849–911. [MR2530322](#)
- FRAYLING, T. M., TIMPSON, N. J., WEEDON, M. N., ZEGGINI, E., FREATHY, R. M., LINDGREN, C. M. et al. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* **316** 889–894.
- GORLOV, I. P., GORLOVA, O. Y., SUNYAEV, S. R., SPITZ, M. R. and AMOS, C. I. (2008). Shifting paradigm of association studies: Value of rare single-nucleotide polymorphisms. *Am. J. Hum. Genet.* **82** 100–112.
- HARLEY, J. B., ALARCN-RIQUELME, M. E., CRISWELL, L. A., JACOB, C. O., KIMBERLY, R. P., MOSER, K. L. et al. (2008). Genome-wide association scan in women with systemic lupus erythematosus identifies susceptibility variants in ITGAM, PXX, KIAA1542 and other loci. *Nat. Genet.* **40** 204–210.
- HE, Q. and LIN, D. (2011). A variable selection method for genome-wide association studies. *Bioinformatics* **27** 1–8.
- HIRSCHHORN, J. N. (2009). Genomewide association studies—illuminating biologic pathways. *N. Engl. J. Med.* **360** 1699–1701.
- HOLDEN, M., DENG, S., WOJNOWSKI, L. and KULLE, B. (2008). GSEA-SNP: Applying gene set enrichment analysis to SNP data from genome-wide association studies. *Bioinformatics* **24** 2784–2785.
- HUNTER, D. R. and LI, R. (2005). Variable selection using MM algorithms. *Ann. Statist.* **33** 1617–1642. [MR2166557](#)
- JACOBS, K. B., YEAGER, M., WACHOLDER, S., CRAIG, D., KRAFT, P., HUNTER, D. J. et al. (2009). A new statistic and its power to infer membership in a genome-wide association study using genotype frequencies. *Nat. Genet.* **41** 1253–1257.
- JAQUISH, C. E. (2007). The Framingham Heart Study, on its way to becoming the gold standard for cardiovascular genetic epidemiology? *BMC Med. Genet.* **8** 63.
- KIM, Y., WOJCIECHOWSKI, R., SUNG, H., MATHIAS, R., WANG, L., KLEIN, A., LENROOT, R., MALLEY, J. and BAILEY-WILSON, J. (2009). Evaluation of random forests performance for genome-wide association studies in the presence of interaction effects. *BMC Proceedings* **3** S64.

- LANGE, K., CANTOR, R., HORVATH, S., PEROLA, M., SABATTI, C., SINSHEIMER, J. and SOBEL, E. (2001). Mendel version 4.0: A complete package for the exact genetic analysis of discrete traits in pedigree and population data sets. *Am. J. Hum. Genet.* **69** (Suppl. 1) A1886.
- LANGE, K., PAPP, J. C., SINSHEIMER, J. S., SRIPRACHA, R., ZHOU, H. and SOBEL, E. M. (2013). Mendel: The Swiss army knife of genetic analysis programs. *Bioinformatics* **29** 1568–1570.
- LI, J., DAS, K., FU, G., LI, R. and WU, R. (2011). The Bayesian lasso for genome-wide association studies. *Bioinformatics* **27** 516–523.
- MANOLIO, T. A., COLLINS, F. S., COX, N. J., GOLDSTEIN, D. B., HINDORFF, L. A., HUNTER, D. J. et al. (2009). Finding the missing heritability of complex diseases. *Nature* **461** 747–753.
- PSYCHIATRIC GCCC (2009). Genomewide association studies: History, rationale, and prospects for psychiatric disorders. *Am. J. Psychiatr.* **166** 540–556.
- RITCHIE, M. D., HAHN, L. W., ROODI, N., BAILEY, L. R., DUPONT, W. D., PARL, F. F. and MOORE, J. H. (2001). Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am. J. Hum. Genet.* **69** 138.
- SCUTERI, A., SANNA, S., CHEN, W.-M., UDA, M., ALBAI, G., STRAIT, J. et al. (2007). Genome-wide association scan shows genetic variants in the FTO gene are associated with obesity-related traits. *PLoS Genet.* **3** e115.
- SPELIOTES, E. K., WILLER, C. J., BERNDT, S. I., MONDA, K. L., THORLEIFSSON, G. et al. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat. Genet.* **42** 937–948.
- SZYMCZAK, S., BIERNACKA, J. M., CORDELL, H. J., GONZALEZ-RECIO, O., KONIG, I. R., ZHANG, H. and SUN, Y. V. (2009). Machine learning in genome-wide association studies. *Genet. Epidemiol.* **33** S51–S57.
- TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. [MR1379242](#)
- UEKI, M. and TAMIYA, G. (2012). Ultrahigh-dimensional variable selection method for whole-genome gene-gene interaction analysis. *BMC Bioinformatics* **13** 72.
- WAN, X., YANG, C., YANG, Q., XUE, H., FAN, X., TANG, N. and YU, W. (2010a). BOOST: A fast approach to detecting gene-gene interactions in genome-wide case-control studies. *Am. J. Hum. Genet.* **87** 325–340.
- WAN, X., YANG, C., YANG, Q., XUE, H., TANG, N. and YU, W. (2010b). Predictive rule inference for epistatic interaction detection in genome-wide association studies. *Bioinformatics* **26** 30–37.
- WANG, G., VOLKOW, N., LOGAN, J., PAPPAS, N., WONG, C., ZHU, W., NETUSLL, N. and FOWLER, J. (2001). Brain dopamine and obesity. *The Lancet* **357** 354–357.
- WANG, H., LI, R. and TSAI, C.-L. (2007). Tuning parameter selectors for the smoothly clipped absolute deviation method. *Biometrika* **94** 553–568. [MR2410008](#)
- WANG, K., LI, M. and BUCAN, M. (2007). Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.* **81** 1278–1283.
- WANG, Y., LIU, G., FENG, M. and WONG, L. (2011). An empirical comparison of several recent epistatic interaction detection methods. *Bioinformatics* **27** 2936–2943.
- WEEDON, M. N. and FRAYLING, T. M. (2008). Reaching new heights: Insights into the genetics of human stature. *Trends Genet.* **24** 595–603.
- WU, J., DEVLIN, B., RINGQUIST, S., TRUCCO, M. and ROEDER, K. (2010). Screen and clean: A tool for identifying interactions in genome-wide association studies. *Genet. Epidemiol.* **34** 275–285.
- 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.
- WU, T. T. and LANGE, K. (2008). Coordinate descent algorithms for lasso penalized regression. *Ann. Appl. Stat.* **2** 224–244. [MR2415601](#)

- WU, Z. and ZHAO, H. (2009). Statistical power of model selection strategies for genome-wide association studies. *PLoS Genet.* **5** e1000582.
- YANG, C., HE, Z., WAN, X., YANG, Q., XUE, H. and YU, W. (2009). SNPHarvester: A filtering-based approach for detecting epistatic interactions in genome-wide association studies. *Bioinformatics* **25** 504–511.
- YANG, C., WAN, X., YANG, Q., XUE, H. and YU, W. (2010). Identifying main effects and epistatic interactions from large-scale SNP data via adaptive group lasso. *BMC Bioinformatics* **11** S18.
- YI, N., KAKLAMANI, V. G. and PASCHE, B. (2011). Bayesian analysis of genetic interactions in case-control studies, with application to adiponectin genes and colorectal cancer risk. *Ann. Hum. Genet.* **75** 90–104.
- ZHANG, X., HUANG, S., ZOU, F. and WANG, W. (2010). TEAM: Efficient two-locus epistasis tests in human genome-wide association study. *Bioinformatics* **26** 217–227.
- ZHANG, Y. and LIU, J. S. (2007). Bayesian inference of epistatic interactions in case-control studies. *Nat. Genet.* **39** 1167–1173.
- ZHOU, H., SEHL, M. E., SINSHEIMER, J. S. and LANGE, K. (2010). Association screening of common and rare genetic variants by penalized regression. *Bioinformatics* **26** 2375–2382.
- ZHU, L.-P., LI, L., LI, R. and ZHU, L.-X. (2011). Model-free feature screening for ultrahigh-dimensional data. *J. Amer. Statist. Assoc.* **106** 1464–1475. [MR2896849](#)
- ZOU, H. and HASTIE, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** 301–320. [MR2137327](#)
- ZOU, H. and LI, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Ann. Statist.* **36** 1509–1533. [MR2435443](#)

A PERMUTATIONAL-SPLITTING SAMPLE PROCEDURE TO QUANTIFY EXPERT OPINION ON CLUSTERS OF CHEMICAL COMPOUNDS USING HIGH-DIMENSIONAL DATA¹

BY ELASMA MILANZI^{*}, ARIEL ALONSO[†], CHRISTOPHE BUYCK[‡],
GEERT MOLENBERGHS^{*,§} AND LUC BIJNENS[‡]

Hasselt University^{}, Maastricht University[†], Janssen Pharmaceuticals[‡]
and University of Leuven[§]*

Expert opinion plays an important role when selecting promising clusters of chemical compounds in the drug discovery process. We propose a method to quantify these qualitative assessments using hierarchical models. However, with the most commonly available computing resources, the high dimensionality of the vectors of fixed effects and correlated responses renders maximum likelihood unfeasible in this scenario. We devise a reliable procedure to tackle this problem and show, using theoretical arguments and simulations, that the new methodology compares favorably with maximum likelihood, when the latter option is available. The approach was motivated by a case study, which we present and analyze.

REFERENCES

- AGRAFIOTIS, D. K., ALEX, S., DAI, H., DERKINDEREN, A., FARNUM, M., GATES, P., IZRAILEV, S., JAEGER, E. P., KONSTANT, P., LEUNG, A., LOBANOV, V. S., MARICHAL, P., MARTIN, D., RASSOKHIN, D. N., SHEMANAREV, M., SKALKIN, A., STONG, J., TABRUYN, T., VERMEIREN, M., WAN, J., XU, X. Y. and YAO, X. (2007). Advanced Biological and Chemical Discovery (ABCD): Centralizing discovery knowledge in an inherently decentralized world. *J. Chem. Inf. Model* **47** 1999–2014.
- BERGER, M. and WONG, W. (2009). *An Introduction to Optimal Designs for Social and Biomedical Research*. Wiley-Blackwell, Oxford.
- BIRKEL, T. (1992). Laws of large numbers under dependence assumptions. *Statist. Probab. Lett.* **14** 355–362. MR1179641
- BOND, T. G. and FOX, C. M. (2007). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 2nd ed. Lawrence Erlbaum, Mahwah, NJ.
- CHEN, X. and XIE, M. (2012). A split-and-conquer approach for analysis of extra ordinary large data. DIMACS Technical Report 2012-01 [cited 2013 June 15]. Available at <http://dimacs.rutgers.edu/TechnicalReports/TechReports/2012/2012-01.pdf>.
- DE BOECK, P. and WILSON, M., eds. (2004). *Explanatory Item Response Models: A Generalized Linear and Nonlinear Approach*. Springer, New York. MR2083193
- DONOHO, D. L. (2000). High-dimensional data analysis: The curses and blessings of dimensionality. Aide-Memoire [cited 2013 June 15]. Available at <http://www-stat.stanford.edu/~donoho/Lectures/AMS2000/Curses.pdf>.
- FAN, J., GUO, S. and HAO, N. (2012). Variance estimation using refitted cross-validation in ultra-high dimensional regression. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **74** 37–65. MR2885839

Key words and phrases. Maximum likelihood, pseudo-likelihood, rater, split samples.

- FAN, J. and LI, R. (2006). Statistical challenges with high dimensionality: Feature selection in knowledge discovery. In *International Congress of Mathematicians, Vol. III* (M. Sanz-Sole, J. Soria, J. L. Varona and J. Verdera, eds.) 595–622. Eur. Math. Soc., Zürich. [MR2275698](#)
- FIEUWS, S. and VERBEKE, G. (2006). Pairwise fitting of mixed models for the joint modeling of multivariate longitudinal profiles. *Biometrics* **62** 424–431. [MR2227490](#)
- HACK, M. D., RASSOKHIN, D. N., BUYCK, C., SEIERSTAD, M., SKALKIN, A., TEN HOLTE, P., JONES, T. K., MIRZADEGAN, T. and AGRAFIOTIS, D. K. (2011). Library enhancement through the wisdom of crowds. *J. Chem. Inf. Model* **51** 3275–3286.
- JOHNSON, R. A. and WICHERN, D. W. (2007). *Applied Multivariate Statistical Analysis*, 6th ed. Pearson Prentice Hall, Upper Saddle River, NJ. [MR2372475](#)
- LI, K.-C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86** 316–342. [MR1137117](#)
- MOLENBERGHS, G. and VERBEKE, G. (2005). *Models for Discrete Longitudinal Data*. Springer, New York. [MR2171048](#)
- MOLENBERGHS, G., VERBEKE, G. and IDDI, S. (2011). Pseudo-likelihood methodology for partitioned large and complex samples. *Statist. Probab. Lett.* **81** 892–901. [MR2793758](#)
- NELSON, K. P., LIPSITZ, S. R., FITZMAURICE, G. M., IBRAHIM, J., PARZEN, M. and STRAWDERMAN, R. (2006). Use of the probability integral transformation to fit nonlinear mixed-effects models with nonnormal random effects. *J. Comput. Graph. Statist.* **15** 39–57. [MR2269362](#)
- NEWMAN, C. M. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. In *Inequalities in Statistics and Probability (Lincoln, Neb., 1982)*. *Institute of Mathematical Statistics Lecture Notes—Monograph Series* **5** 127–140. IMS, Hayward, CA. [MR0789244](#)
- OXMAN, A. D., LAVIS, J. N. and FRETHEIM, A. (2007). Use of evidence in WHO recommendations. *Lancet* **369** 1883–1889.
- PINHEIRO, J. C. and BATES, D. M. (1995). Approximations to the log-likelihood function in the nonlinear mixed-effects model. *J. Comput. Graph. Statist.* **4** 12–35.
- TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **58** 267–288. [MR1379242](#)
- XIA, Y., TONG, H., LI, W. K. and ZHU, L.-X. (2002). An adaptive estimation of dimension reduction space. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **64** 363–410. [MR1924297](#)

THE USE OF COVARIATES AND RANDOM EFFECTS IN EVALUATING PREDICTIVE BIOMARKERS UNDER A POTENTIAL OUTCOME FRAMEWORK

BY ZHIWEI ZHANG*, LEI NIE[†], GUOXING SOON[†] AND AIYI LIU[‡]

*Center for Devices and Radiological Health, Food and Drug Administration**,
Center for Drug Evaluation and Research, Food and Drug Administration[†]
and Eunice Kennedy Shriver National Institute
of Child Health and Human Development[‡]

Predictive or treatment selection biomarkers are usually evaluated in a subgroup or regression analysis with focus on the treatment-by-marker interaction. Under a potential outcome framework (Huang, Gilbert and Janes [*Biometrics* **68** (2012) 687–696]), a predictive biomarker is considered a predictor for a desirable treatment benefit (defined by comparing potential outcomes for different treatments) and evaluated using familiar concepts in prediction and classification. However, the desired treatment benefit is unobservable because each patient can receive only one treatment in a typical study. Huang et al. overcome this problem by assuming monotonicity of potential outcomes, with one treatment dominating the other in all patients. Motivated by an HIV example that appears to violate the monotonicity assumption, we propose a different approach based on covariates and random effects for evaluating predictive biomarkers under the potential outcome framework. Under the proposed approach, the parameters of interest can be identified by assuming conditional independence of potential outcomes given observed covariates, and a sensitivity analysis can be performed by incorporating an unobserved random effect that accounts for any residual dependence. Application of this approach to the motivating example shows that baseline viral load and CD4 cell count are both useful as predictive biomarkers for choosing antiretroviral drugs for treatment-naïve patients.

REFERENCES

- COHEN, C. J., ANDRADE-VILLANUEVA, J., CLOTET, B., FOURIE, J., JOHNSON, M. A., RUXRUNGTHAM, K., WU, H., ZORRILLA, C., CRAUWELS, H., RIMSKY, L. T., VANVEGEL, S., BOVEN, K. and THRIVE STUDY GROUP (2011). Rilpivirine versus efavirenz with two background nucleoside or nucleotide reverse transcriptase inhibitors in treatment-naïve adults infected with HIV-1 (THRIVE): A phase 3, randomised, non-inferiority trial. *Lancet* **378** 229–237.
- DODD, L. E. and PEPE, M. S. (2003). Semiparametric regression for the area under the receiver operating characteristic curve. *J. Amer. Statist. Assoc.* **98** 409–417. [MR1995717](#)
- FOSTER, J. C., TAYLOR, J. M. G. and RUBERG, S. J. (2011). Subgroup identification from randomized clinical trial data. *Stat. Med.* **30** 2867–2880. [MR2844689](#)

Key words and phrases. Conditional independence, counterfactual, ROC regression, sensitivity analysis, treatment effect heterogeneity, treatment selection.

- GADBURY, G. L. and IYER, H. K. (2000). Unit-treatment interaction and its practical consequences. *Biometrics* **56** 882–885.
- GAIL, M. and SIMON, R. (1985). Testing for qualitative interactions between treatment effects and patient subsets. *Biometrics* **41** 361–372.
- HOLLAND, P. W. (1986). Statistics and causal inference (with discussion). *J. Amer. Statist. Assoc.* **81** 945–970. [MR0867618](#)
- HUANG, Y., GILBERT, P. B. and JANES, H. (2012). Assessing treatment-selection markers using a potential outcomes framework. *Biometrics* **68** 687–696. [MR3055173](#)
- KARAPETIS, C. S., KHAMBATA-FORD, S., JONKER, D. J., O'CALLAGHAN, C. J., TU, D., TEBBUTT, N. C., SIMES, R. J., CHALCHAL, H., SHAPIRO, J. D., ROBITAILLE, S., PRICE, T. J., SHEPHERD, L., AU, H.-J., LANGER, C., MOORE, M. J. and ZALCBERG, J. R. (2008). K-ras mutations and benefit from cetuximab in advanced colorectal cancer. *N. Engl. J. Med.* **359** 1757–1765.
- MANSKI, C. F. (2003). *Partial Identification of Probability Distributions*. Springer, New York. [MR2151380](#)
- PAIK, S., SHAK, S., TANG, G., KIM, C., BAKER, J., CRONIN, M., BAEHNER, F. L., WALKER, M. G., WATSON, D., PARK, T., HILLER, W., FISHER, E. R., WICKERHAM, D. L., BRYANT, J. and WOLMARK, N. (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N. Engl. J. Med.* **351** 2817–2826.
- PEPE, M. S. (2003). *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford Statistical Science Series **28**. Oxford Univ. Press, Oxford. [MR2260483](#)
- POCOCK, S. J., ASSMANN, S. E., ENOS, L. E. and KASTEN, L. E. (2002). Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. *Stat. Med.* **21** 2917–2930.
- POULSON, R. S., GADBURY, G. L. and ALLISON, D. B. (2012). Treatment heterogeneity and individual qualitative interaction. *Amer. Statist.* **66** 16–24. [MR2934737](#)
- QIAN, M. and MURPHY, S. A. (2011). Performance guarantees for individualized treatment rules. *Ann. Statist.* **39** 1180–1210. [MR2816351](#)
- RUSSEK-COHEN, E. and SIMON, R. M. (1998). Evaluating treatments when a gender by treatment interaction may exist. *Stat. Med.* **16** 455–464.
- SIMON, R. (2008). Development and validation of biomarker classifiers for treatment selection. *J. Statist. Plann. Inference* **138** 308–320. [MR2412589](#)
- SIMON, R. (2010). Clinical trials for predictive medicine: New challenges and paradigms. *Clin. Trials* **7** 516–524.
- SU, X., ZHOU, T., YAN, X., FAN, J. and YANG, S. (2008). Interaction trees with censored survival data. *Int. J. Biostat.* **4** Art. 2, 28. [MR2383729](#)
- TIAN, L., ALIZADEH, A. A., GENTLES, A. J. and TIBSHIRANI, R. (2012). A simple method for detecting interactions between a treatment and a large number of covariates. Available at [arXiv:1212.2995](#).
- VAN DER LAAN, M. J. and ROBINS, J. M. (2003). *Unified Methods for Censored Longitudinal Data and Causality*. Springer, New York. [MR1958123](#)
- ZHANG, B., TSIATIS, A. A., LABER, E. B. and DAVIDIAN, M. (2012). A robust method for estimating optimal treatment regimes. *Biometrics* **68** 1010–1018. [MR3040007](#)
- ZHANG, Z., WANG, C., NIE, L. and SOON, G. (2013). Assessing the heterogeneity of treatment effects via potential outcomes of individual patients. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **62** 687–704. [MR3118329](#)
- ZHANG, Z., NIE, L., SOON, G. and LIU, A. (2014). Supplement to “The use of covariates and random effects in evaluating predictive biomarkers under a potential outcome framework.” DOI:10.1214/14-AOAS773SUPP.
- ZHOU, X.-H., OBUCHOWSKI, N. A. and MCCLISH, D. K. (2002). *Statistical Methods in Diagnostic Medicine*. Wiley, New York. [MR1915698](#)

ZOU, K. H., LIU, A., BANDOS, A. I., OHNO-MACHADO, L. and ROCKETTE, H. E. (2011). *Statistical Evaluation of Diagnostic Performance: Topics in ROC Analysis*. CRC Press, Boca Raton, FL. MR3052764

EVALUATING EPOETIN DOSING STRATEGIES USING OBSERVATIONAL LONGITUDINAL DATA

BY CECILIA A. COTTON¹ AND PATRICK J. HEAGERTY²

University of Waterloo and University of Washington

Epoetin is commonly used to treat anemia in chronic kidney disease and End Stage Renal Disease subjects undergoing dialysis, however, there is considerable uncertainty about what level of hemoglobin or hematocrit should be targeted in these subjects. In order to address this question, we treat epoetin dosing guidelines as a type of dynamic treatment regimen. Specifically, we present a methodology for comparing the effects of alternative treatment regimens on survival using observational data. In randomized trials patients can be assigned to follow a specific management guideline, but in observational studies subjects can have treatment paths that appear to be adherent to multiple regimens at the same time. We present a cloning strategy in which each subject contributes follow-up data to each treatment regimen to which they are continuously adherent and artificially censored at first nonadherence. We detail an inverse probability weighted log-rank test with a valid asymptotic variance estimate that can be used to test survival distributions under two regimens. To compare multiple regimens, we propose several marginal structural Cox proportional hazards models with robust variance estimation to account for the creation of clones. The methods are illustrated through simulations and applied to an analysis comparing epoetin dosing regimens in a cohort of 33,873 adult hemodialysis patients from the United States Renal Data System.

REFERENCES

- BESARAB, A., BOLTON, W. K., BROWNE, J. K., EGRIE, J. C., NISSENSON, A. R., OKAMOTO, D. M., SCHWAB, S. J. and GOODKIN, D. A. (1998). The effects of normal as compared with low hematocrit values in patients with cardiac disease who are receiving hemodialysis and epoetin. *N. Engl. J. Med.* **339** 584–590. PMID: 9718377.
- BROOKHART, M. A., SCHNEEWEISS, S., AVORN, J., BRADBURY, B. D., LIU, J. and WINKELMAYER, W. C. (2010). Comparative mortality risk of anemia management practices in incident hemodialysis patients. *JAMA: The Journal of the American Medical Association* **303** 857–864.
- CAIN, L. E., ROBINS, J. M., LANOY, E., LOGAN, R., COSTAGLIOLA, D. and HERNÁN, M. A. (2010). When to start treatment? A systematic approach to the comparison of dynamic regimes using observational data. *Int. J. Biostat.* **6** Art. 18, 26. [MR2644387](#)
- CAIN, L. E., LOGAN, R., ROBINS, J. M., STERNE, J. A. C., SABIN, C., BANSI, L., JUSTICE, A., GOULET, J., VAN SIGHEM, A., DE WOLF, F., BUCHER, H. C., VON WYL, V., ESTEVE, A., CASABONA, J., DEL AMO, J., MORENO, S., SENG, R., MEYER, L., PEREZ-HOYOS, S., MUGA, R., LODI, S., LANOY, E., COSTAGLIOLA, D. and HERNAN, M. A. (2011). When to initiate combined antiretroviral therapy to reduce mortality and AIDS-defining illness in HIV-infected persons in developed countries: An observational study. *Ann. Intern. Med.* **154** 509–515.

Key words and phrases. Marginal Structural Models, observational studies, survival analysis.

- CANADIAN ERYTHROPOIETIN STUDY GROUP (1990). Association between recombinant human erythropoietin and quality of life and exercise capacity of patients receiving haemodialysis. *BMJ* **300** 573–578.
- CHAKRABORTY, B. and MOODIE, E. E. M. (2013). *Statistical Methods for Dynamic Treatment Regimes*. Springer, New York. [MR3112454](#)
- COTTON, C. A. and HEAGERTY, P. J. (2011). A data augmentation method for estimating the causal effect of adherence to treatment regimens targeting control of an intermediate measure. *Statistics in Biosciences* **3** 28–44.
- COTTON, C. A. and HEAGERTY, P. J. (2014). Supplement to “Evaluating epoetin dosing strategies using observational longitudinal data.” DOI:[10.1214/14-AOAS774SUPP](#).
- DRÜEKE, T. B., LOCATELLI, F., CLYNE, N., ECKARDT, K.-U., MACDOUGALL, I. C., TSAKIRIS, D., BURGER, H.-U. and SCHERHAG, A. (2006). Normalization of hemoglobin level in patients with chronic kidney disease and anemia. *N. Engl. J. Med.* **355** 2071–2084.
- ESCHBACH, J. W. (1994). Erythropoietin: The promise and the facts. *Kidney International Supplements* **44** S70–S76.
- HERNÁN, M. A., BRUMBACK, B. and ROBINS, J. M. (2000). Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* **11** 561–570.
- HERNÁN, M. A., BRUMBACK, B. and ROBINS, J. M. (2001). Marginal structural models to estimate the joint causal effect of nonrandomized treatments. *J. Amer. Statist. Assoc.* **96** 440–448. [MR1939347](#)
- HERNÁN, M. A., LANOY, E., COSTAGLIOLA, D. and ROBINS, J. M. (2006). Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin. Pharmacol. Toxicol.* **98** 237–242.
- JUNG, S.-H. (1999). Rank tests for matched survival data. *Lifetime Data Anal.* **5** 67–79. [MR1750342](#)
- LEE, E. W., WEI, L. J. and AMATO, D. A. (1992). Cox-type regression analysis for large numbers of small groups of correlated failure time observations. In *Survival Analysis: State of the Art (Columbus, OH, 1991)* (J. P. Klein and P. K. Goel, eds.) 237–247. Kluwer Academic, Dordrecht. [MR1175646](#)
- LIANG, K. Y. and ZEGER, S. L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika* **73** 13–22. [MR0836430](#)
- MISKULIN, D. C., WEINER, D. E., TIGHIOUART, H., LADIK, V., SERVILLA, K., ZAGER, P. G., MARTIN, A., JOHNSON, H. K. and MEYER, K. B. (2009). Computerized decision support for EPO dosing in hemodialysis patients. *Am. J. Kidney Dis.* **54** 1081–1088.
- MISKULIN, D. C., ZHOU, J., TANGRI, N., BANDEEN-ROCHE, K., COOK, C., EPHRAIM, P. L., CREWS, D. C., SCIALLA, J. J., SOZIO, S. M., SHAFI, T. et al. (2013). Trends in anemia management in US hemodialysis patients 2004–2010. *BMC Nephrology* **14** 264.
- NATIONAL KIDNEY FOUNDATION (2006). K/DOQI clinical practice guidelines and clinical practice recommendations for anemia in chronic kidney disease. *American Journal of Kidney Diseases* **47: Suppl 3** S11–S145.
- ORELLANA, L., ROTNITZKY, A. and ROBINS, J. M. (2010). Dynamic regime marginal structural mean models for estimation of optimal dynamic treatment regimes, Part I: Main content. *Int. J. Biostat.* **6** Art. 8, 49. [MR2602551](#)
- PALMER, S. C., NAVANEETHAN, S. D., CRAIG, J. C., JOHNSON, D. W., TONELLI, M., GARG, A. X., PELLEGRINI, F., RAVANI, P., JARDINE, M., PERKOVIC, V., GRAZIANO, G., MCGEE, R., NICOLUCCI, A., TOGNONI, G. and STRIPPOLI, G. F. M. (2010). Meta-analysis: Erythropoiesis-stimulating agents in patients with chronic kidney disease. *Ann. Intern. Med.* **153** 23–33.
- PEPE, M. S. and COUPER, D. (1997). Modeling partly conditional means with longitudinal data. *J. Amer. Statist. Assoc.* **92** 991–998. [MR1482129](#)

- PEPE, M. S., HEAGERTY, P. J. and WHITAKER, R. (1999). Prediction using partly conditional time-varying coefficients regression models. *Biometrics* **55** 944–950.
- PFEFFER, M. A., BURDMANN, E. A., CHEN, C.-Y., COOPER, M. E., DE ZEEUW, D., ECKARDT, K.-U., FEYZI, J. M., IVANOVICH, P., KEWALRAMANI, R., LEVEY, A. S., LEWIS, E. F., MCGILL, J. B., MCMURRAY, J. J. V., PARFREY, P., PARVING, H.-H., REMUZZI, G., SINGH, A. K., SOLOMON, S. D. and TOTO, R. (2009). A trial of darbepoetin alfa in type 2 diabetes and chronic kidney disease. *N. Engl. J. Med.* **361** 2019–2032.
- ROBINS, J. M. (1993). Information recovery and bias adjustment in proportional hazards regression analysis of randomized trials using surrogate markers. In *Proceedings of the Biopharmaceutical Section, American Statistical Association* 24–33. American Statistical Association, Alexandria, VA.
- ROBINS, J. M. (1998). Marginal structural models. In 1997 *Proceedings of the Section on Bayesian Statistical Science* 1–10. American Statistical Association, Alexandria, VA.
- ROBINS, J. M. (2000). Marginal structural models versus structural nested models as tools for causal inference. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials (Minneapolis, MN, 1997)* (M. E. Halloran and D. Berry, eds.) 95–133. Springer, New York. [MR1731682](#)
- ROBINS, J. M. and FINKELSTEIN, D. M. (2000). Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics* **56** 779–788.
- ROBINS, J. M., HERNÁN, M. A. and BRUMBACK, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* **11** 550–560.
- ROBINS, J., ORELLANA, L. and ROTNITZKY, A. (2008). Estimation and extrapolation of optimal treatment and testing strategies. *Stat. Med.* **27** 4678–4721. [MR2528576](#)
- ROBINS, J. M., ROTNITZKY, A. and ZHAO, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J. Amer. Statist. Assoc.* **90** 106–121. [MR1325118](#)
- ROBINS, J. M., BLEVINS, D., RITTER, G. and WULFSOHN, M. (1992). G-estimation of the effect of prophylaxis therapy for pneumocystis carinii pneumonia on the survival of AIDS patients. *Epidemiology* **3** 319–336.
- RUBIN, D. B. (1980). Discussion of “Randomization analysis of experimental data in the Fisher randomization test,” by D. Basu. *J. Amer. Statist. Assoc.* **75** 591–593.
- SHORTREED, S. M. and MOODIE, E. E. M. (2012). Estimating the optimal dynamic antipsychotic treatment regime: Evidence from the sequential multiple-assignment randomized clinical antipsychotic trials of intervention and effectiveness schizophrenia study. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **61** 577–599. [MR2960739](#)
- SINGH, A. K., SZCZECH, L., TANG, K. L., BARNHART, H., SAPP, S., WOLFSON, M. and REDDAN, D. (2006). Correction of anemia with epoetin alfa in chronic kidney disease. *N. Engl. J. Med.* **355** 2085–2098.
- UNGER, E. F., THOMPSON, A. M., BLANK, M. J. and TEMPLE, R. (2010). Erythropoiesis-stimulating agents, a time for a reevaluation. *N. Engl. J. Med.* **362** 189–192.
- UNITED STATES GOVERNMENT ACCOUNTABILITY OFFICE (2006). Report to the Chairman, Committee on Ways and Means, House of Representatives. End-stage renal disease: Bundling of Medicare’s payment for drugs with payment for all ESRD services would promote efficiency and clinical flexibility (GAO-07-77). Accessed May 2, 2013, at <http://www.gao.gov/assets/260/253347.pdf>.
- WANG, O., KILPATRICK, R. D., CRITCHLOW, C. W., LING, X., BRADBURY, B. D., GILBERTSON, D. T., COLLINS, A. J., ROTHMAN, K. J. and ACQUAVELLA, J. F. (2010). Relationship between epoetin alfa dose and mortality: Findings from a marginal structural model. *Clin. J. Am. Soc. Nephrol.* **5** 182–188.

- XIE, J. and LIU, C. (2005). Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data. *Stat. Med.* **24** 3089–3110. [MR2209045](#)
- YOUNG, J., CAIN, L., ROBINS, J., O’REILLY, E. and HERNÁN, M. (2011). Comparative effectiveness of dynamic treatment regimes: An application of the parametric g-formula. *Statistics in Biosciences* **3** 119–143.
- ZHANG, Y., THAMER, M., STEFANIK, K., KAUFMAN, J. and COTTER, D. J. (2004). Epoetin requirements predict mortality in hemodialysis patients. *Am. J. Kidney Dis.* **44** 866–876.
- ZHANG, Y., THAMER, M., KAUFMAN, J. S., COTTER, D. J. and HERNÁN, M. A. (2011). High doses of epoetin do not lower mortality and cardiovascular risk among elderly hemodialysis patients with diabetes. *Kidney Int.* **80** 663–669.
- ZHENG, Y. and HEAGERTY, P. J. (2005). Partly conditional survival models for longitudinal data. *Biometrics* **61** 379–391. [MR2140909](#)

SYNTHESISING EVIDENCE TO ESTIMATE PANDEMIC (2009) A/H1N1 INFLUENZA SEVERITY IN 2009–2011¹

BY ANNE M. PRESANIS*, RICHARD G. PEBODY[†], PAUL J. BIRRELL*,
BRIAN D. M. TOM*, HELEN K. GREEN[†], HAYLEY DURNALL[‡],
DOUGLAS FLEMING[‡] AND DANIELA DE ANGELIS*

*Medical Research Council Biostatistics Unit**, *Public Health England[†]*
and Royal College of General Practitioners[‡]

Knowledge of the severity of an influenza outbreak is crucial for informing and monitoring appropriate public health responses, both during and after an epidemic. However, case-fatality, case-intensive care admission and case-hospitalisation risks are difficult to measure directly. Bayesian evidence synthesis methods have previously been employed to combine fragmented, under-ascertained and biased surveillance data coherently and consistently, to estimate case-severity risks in the first two waves of the 2009 A/H1N1 influenza pandemic experienced in England. We present in detail the complex probabilistic model underlying this evidence synthesis, and extend the analysis to also estimate severity in the third wave of the pandemic strain during the 2010/2011 influenza season. We adapt the model to account for changes in the surveillance data available over the three waves. We consider two approaches: (a) a two-stage approach using posterior distributions from the model for the first two waves to inform priors for the third wave model; and (b) a one-stage approach modelling all three waves simultaneously. Both approaches result in the same key conclusions: (1) that the age-distribution of the case-severity risks is “u”-shaped, with children and older adults having the highest severity; (2) that the age-distribution of the infection attack rate changes over waves, school-age children being most affected in the first two waves and the attack rate in adults over 25 increasing from the second to third waves; and (3) that when averaged over all age groups, case-severity appears to increase over the three waves. The extent to which the final conclusion is driven by the change in age-distribution of those infected over time is subject to discussion.

REFERENCES

- ADES, A. E. and SUTTON, A. J. (2006). Multiparameter evidence synthesis in epidemiology and medical decision-making: Current approaches. *J. Roy. Statist. Soc. Ser. A* **169** 5–35. MR2222010
- ALBERT, I., ESPIÉ, E., DE VALK, H. and DENIS, J.-B. B. (2011). A Bayesian evidence synthesis for estimating campylobacteriosis prevalence. *Risk Analysis* **31** 1141–1155.
- BIRD, S. M. (2010). Like-with-like comparisons? *The Lancet* **376** 684+.
- BIRRELL, P. J., KETSETZIS, G., GAY, N. J., COOPER, B. S., PRESANIS, A. M., HARRIS, R. J., CHARLETT, A., ZHANG, X.-S., WHITE, P. J., PEBODY, R. G. and DE ANGELIS, D. (2011). Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *Proc. Natl. Acad. Sci. USA* **108** 18238–18243.

Key words and phrases. Evidence synthesis, Bayesian, influenza, severity.

- BOX, G. E. P. (1980). Sampling and Bayes' inference in scientific modelling and robustness. *J. Roy. Statist. Soc. Ser. A* **143** 383–430. [MR0603745](#)
- BROOKS, S. P. and GELMAN, A. (1998). General methods for monitoring convergence of iterative simulations. *J. Comput. Graph. Statist.* **7** 434–455. [MR1665662](#)
- CAMPBELL, C. N. J., MYTTON, O. T., MCLEAN, E. M., RUTTER, P. D., PEBODY, R. G., SACHEDINA, N., WHITE, P. J., HAWKINS, C., EVANS, B., WAIGHT, P. A., ELLIS, J., BIRMINGHAM, A., DONALDSON, L. J. and CATCHPOLE, M. (2011). Hospitalization in two waves of pandemic influenza A(H1N1) in England. *Epidemiology and Infection* **139** 1560–1569.
- DEPARTMENT OF HEALTH (2011). Department of Health Winter Watch. Accessed 25/02/2011.
- DONALDSON, L. J., RUTTER, P. D., ELLIS, B. M., GREAVES, F. E. C., MYTTON, O. T., PEBODY, R. G. and YARDLEY, I. E. (2009). Mortality from pandemic A/H1N1 2009 influenza in England: Public health surveillance study. *BMJ* **339** b5213.
- EDDY, D. M., HASSELBLAD, V. and SHACHTER, R. (1992). *Meta-Analysis by the Confidence Profile Method*. Academic Press, Boston, MA.
- FLEMING, D. M. (1999). Weekly returns service of the royal college of general practitioners. *Communicable Disease and Public Health / PHLIS* **2** 96–100.
- GARSKE, T., LEGRAND, J., DONNELLY, C. A., WARD, H., CAUCHEMEZ, S., FRASER, C., FERGUSON, N. M. and GHANI, A. C. (2009). Assessing the severity of the novel influenza A/H1N1 pandemic. *BMJ* **339** b2840.
- GOUBAR, A., ADES, A. E., DE ANGELIS, D., MCGARRIGLE, C. A., MERCER, C. H., TOOKEY, P. A., FENTON, K. and GILL, O. N. (2008). Estimates of human immunodeficiency virus prevalence and proportion diagnosed based on Bayesian multiparameter synthesis of surveillance data. *J. Roy. Statist. Soc. Ser. A* **171** 541–580. [MR2432503](#)
- HARDELID, P., ANDREWS, N. J., HOSCHLER, K., STANFORD, E., BAGUELIN, M., WAIGHT, P. A., ZAMBON, M. and MILLER, E. (2011). Assessment of baseline age-specific antibody prevalence and incidence of infection to novel influenza A/H1N1 2009. *Health Technology Assessment* **14** 115–192.
- HEALTH PROTECTION AGENCY (2010). Epidemiological report of pandemic (H1N1) 2009 in the UK. Technical report, Health Protection Agency.
- HEALTH PROTECTION AGENCY, HEALTH PROTECTION SCOTLAND, COMMUNICABLE DISEASE SURVEILLANCE CENTRE NORTHERN IRELAND AND NATIONAL PUBLIC HEALTH SERVICE FOR WALES (2009). First Few Hundred (FF100) Project: Epidemiological Protocols for Comprehensive Assessment of Early Swine Influenza Cases in the United Kingdom Technical report, Health Protection Agency.
- HOSCHLER, K., THOMPSON, C., ANDREWS, N., GALIANO, M., PEBODY, R., ELLIS, J., STANFORD, E., BAGUELIN, M., MILLER, E. and ZAMBON, M. (2012). Seroprevalence of influenza A(H1N1)pdm09 virus antibody, England, 2010 and 2011. *Emerging Infect. Dis.* **18** 1894–1897.
- JACKSON, C., VYNNYCKY, E. and MANGTANI, P. (2010). Estimates of the transmissibility of the 1968 (Hong Kong) influenza pandemic: Evidence of increased transmissibility between successive waves. *Am. J. Epidemiol.* **171** 465–478.
- LAURIE, K. L., HUSTON, P., RILEY, S., KATZ, J. M., WILLISON, D. J., TAM, J. S., MOUNTS, A. W., HOSCHLER, K., MILLER, E., VANDEMAELE, K., BROBERG, E., VAN KERKHOVE, M. D. and NICOLL, A. (2013). Influenza serological studies to inform public health action: Best practices to optimise timing, quality and reporting. *Influenza Other Respir Viruses* **7** 211–224.
- LIPSITCH, M., RILEY, S., CAUCHEMEZ, S., GHANI, A. C. and FERGUSON, N. M. (2009). Managing and reducing uncertainty in an emerging influenza pandemic. *New England Journal of Medicine* **361** 112–115.
- LIPSITCH, M., FINELLI, L., HEFFERNAN, R. T., LEUNG, G. M., REDD, S. C. and FOR THE 2009 H1N1 SURVEILLANCE GROUP (2011). Improving the evidence base for decision making during

- a pandemic: The example of 2009 influenza A/H1N1. *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science* **9** 89–114.
- LU, G. and ADES, A. E. (2006). Assessing evidence inconsistency in mixed treatment comparisons. *J. Amer. Statist. Assoc.* **101** 447–459. [MR2256166](#)
- LUNN, D., SPIEGELHALTER, D., THOMAS, A. and BEST, N. (2009). The BUGS project: Evolution, critique and future directions. *Stat. Med.* **28** 3049–3067. [MR2750401](#)
- MCDONALD, S. A., PRESANIS, A. M., DE ANGELIS, D., VAN DER HOEK, W., HOOIVELD, M., DONKER, G. and KRETZSCHMAR, M. E. (2014). An evidence synthesis approach to estimating the incidence of seasonal influenza in the Netherlands. *Influenza Other Respi. Viruses* **8** 33–41.
- MILLER, M. A., VIBOUD, C., BALINSKA, M. and SIMONSEN, L. (2009). The signature features of influenza pandemics—implications for policy. *N. Engl. J. Med.* **360** 2595–2598.
- MILLER, E., HOSCHLER, K., HARDELID, P., STANFORD, E., ANDREWS, N. and ZAMBON, M. (2010). Incidence of 2009 pandemic influenza A/H1N1 infection in England: A cross-sectional serological study. *The Lancet* **375** 1100–1108.
- O’HAGAN, A. (2003). HSSS model criticism. In *Highly Structured Stochastic Systems*. Oxford *Statist. Sci. Ser.* **27** 423–453. Oxford Univ. Press, Oxford. [MR2082418](#)
- PEBODY, R. G., MCLEAN, E., ZHAO, H., CLEARY, P., BRACEBRIDGE, S., FOSTER, K., CHARLETT, A., HARDELID, P., WAIGHT, P., ELLIS, J., BIRMINGHAM, A., ZAMBON, M., EVANS, B., SALMON, R., MCMENAMIN, J., SMYTH, B., CATCHPOLE, M. and WATSON, J. (2010). Pandemic influenza A(H1N1) 2009 and mortality in the United Kingdom: Risk factors for death, April 2009 to March 2010. *Euro Surveillance* **15**.
- PRESANIS, A. M., DE ANGELIS, D., SPIEGELHALTER, D. J., SEAMAN, S., GOUBAR, A. and ADES, A. E. (2008). Conflicting evidence in a Bayesian synthesis of surveillance data to estimate human immunodeficiency virus prevalence. *J. Roy. Statist. Soc. Ser. A* **171** 915–937. [MR2530293](#)
- PRESANIS, A. M., DE ANGELIS, D., NEW YORK CITY SWINE FLU INVESTIGATION TEAM, HAGY, A., REED, C., RILEY, S., COOPER, B. S., FINELLI, L., BIEDRZYCKI, P. and LIPSITCH, M. (2009). The severity of pandemic H1N1 influenza in the United States, from April to July 2009: A Bayesian analysis. *PLoS Med.* **6** e1000207.
- PRESANIS, A. M., DE ANGELIS, D., GOUBAR, A., GILL, O. N. and ADES, A. E. (2011a). Bayesian evidence synthesis for a transmission dynamic model for HIV among men who have sex with men. *Biostatistics* **12** 666–681.
- PRESANIS, A. M., PEBODY, R. G., PATERSON, B. J., TOM, B. D. M., BIRRELL, P. J., CHARLETT, A., LIPSITCH, M. and DE ANGELIS, D. (2011b). Changes in severity of 2009 pandemic A/H1N1 influenza in England: A Bayesian evidence synthesis. *BMJ* **343**.
- PRESANIS, A. M., OHLSEN, D., SPIEGELHALTER, D. J. and DE ANGELIS, D. (2013). Conflict diagnostics in directed acyclic graphs, with applications in Bayesian evidence synthesis. *Statist. Sci.* **28** 376–397. [MR3135538](#)
- PRESANIS, A. M., PEBODY, R. G., BIRRELL, P. J., TOM, B. D. M., GREEN, H., DURNELL, H., FLEMING, D. and DE ANGELIS, D. (2014). Supplement to “Synthesising evidence to estimate pandemic (2009) A/H1N1 influenza severity in 2009–2011.” DOI:10.1214/14-AOAS775SUPP.
- REED, C., ANGULO, F. J., SWERDLOW, D. L., LIPSITCH, M., MELTZER, M. I., JERNIGAN, D. and FINELLI, L. (2009). Estimates of the prevalence of pandemic (H1N1) 2009, United States, April–July 2009. *Emerging Infect. Dis.* **15** 2004–2007.
- SHUBIN, M., VIRTANEN, M., TOIKKANEN, S., LYYTIKÄINEN, O. and AURANEN, K. (2013). Estimating the burden of A(H1N1)pdm09 influenza in Finland during two seasons. *Epidemiology & Infection* **142** 964–974.
- SPIEGELHALTER, D. J., ABRAMS, K. R. and MYLES, J. P. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation (Statistics in Practice)*. Wiley, New York.
- SWEETING, M. J., DE ANGELIS, D., HICKMAN, M. and ADES, A. E. (2008). Estimating hepatitis C prevalence in England and Wales by synthesizing evidence from multiple data sources. Assessing data conflict and model fit. *Biostatistics* **9** 715–734.

- SYPSA, V., BONOVAS, S., TSIODRAS, S., BAKA, A., EFSTATHIOU, P., MALLIORI, M., PANAGIOTOPOULOS, T., NIKOLAKOPOULOS, I. and HATZAKIS, A. (2011). Estimating the disease burden of 2009 pandemic influenza A(H1N1) from surveillance and household surveys in Greece. *PLoS ONE* **6** e20593.
- TRUELOVE, S. A., CHITNIS, A. S., HEFFERNAN, R. T., KARON, A. E., HAUPT, T. E. and DAVIS, J. P. (2011). Comparison of patients hospitalized with pandemic 2009 influenza A(H1N1) virus infection during the first two pandemic waves in Wisconsin. *J. Infect. Dis.* **203** 828–837.
- VAN KERKHOVE, M. D., VANDEMAELE, K. A. H., SHINDE, V., JARAMILLO-GUTIERREZ, G., KOUKOUNARI, A., DONNELLY, C. A., CARLINO, L. O., OWEN, R., PATERSON, B., PELLETIER, L., VACHON, J., GONZALEZ, C., HONGJIE, Y., ZIJIAN, F., CHUANG, S. K., AU, A., BUDA, S., KRAUSE, G., HAAS, W., BONMARIN, I., TANIGUCHI, K., NAKAJIMA, K., SHOBAYASHI, T., TAKAYAMA, Y., SUNAGAWA, T., HERAUD, J. M., ORELLE, A., PALACIOS, E., VAN DER SANDE, M. A. B., WIELDERS, C. C. H. L., HUNT, D., CUTTER, J., LEE, V. J., THOMAS, J., SANTA-OLALLA, P., SIERRA-MOROS, M. J., HANSHAOWORAKUL, W., UNGCHUSAK, K., PEBODY, R., JAIN, S., MOUNTS, A. W. and ON BEHALF OF THE WHO WORKING GROUP FOR RISK FACTORS FOR SEVERE H1N1PDM INFECTION (2011). Risk factors for severe outcomes following 2009 influenza A(H1N1) infection: A global pooled analysis. *PLoS Med* **8** e1001053+.
- WELTON, N. J. and ADES, A. E. (2005). A model of toxoplasmosis incidence in the UK: Evidence synthesis and consistency of evidence. *J. R. Stat. Soc. Ser. C. Appl. Stat.* **54** 385–404. [MR2135881](#)
- WIELDERS, C. C. H., VAN LIER, E. A., VAN'T KLOOSTER, T. M., VAN GAGELDONK-LAFEBER, A. B., VAN DEN WIJNGAARD, C. C., HAAGSMA, J. A., DONKER, G. A., MEIJER, A., VAN DER HOEK, W., LUGNÉR, A. K., KRETZSCHMAR, M. E. E. and VAN DER SANDE, M. A. B. (2012). The burden of 2009 pandemic influenza A(H1N1) in the Netherlands. *The European Journal of Public Health* **22** 150–157.
- WILSON, N. and BAKER, M. G. (2009). The emerging influenza pandemic: Estimating the case fatality ratio. *Euro Surveill.* **14**.
- WONG, J. Y., KELLY, H., IP, D. K., WU, J. T., LEUNG, G. M. and COWLING, B. J. (2013). Case fatality risk of influenza A(H1N1pdm09): A systematic review. *Epidemiology (Cambridge, Mass.)* **24** 830–841.
- WU, J. T., MA, E. S. K., LEE, C. K., CHU, D. K. W., HO, P.-L., SHEN, A. L., HO, A., HUNG, I. F. N., RILEY, S., HO, L. M., LIN, C. K., TSANG, T., LO, S.-V., LAU, Y.-L., LEUNG, G. M., COWLING, B. J. and PEIRIS, J. S. M. (2010). The infection attack rate and severity of 2009 pandemic H1N1 influenza in Hong Kong. *Clinical Infectious Diseases* **51** 1184–1191.
- YANG, J.-R., HUANG, Y.-P., CHANG, F.-Y., HSU, L.-C., LIN, Y.-C., SU, C.-H., CHEN, P.-J., WU, H.-S. and LIU, M.-T. (2011). New variants and age shift to high fatality groups contribute to severe successive waves in the 2009 influenza pandemic in Taiwan. *PLoS ONE* **6** e28288+.

DETECTING DUPLICATES IN A HOMICIDE REGISTRY USING A BAYESIAN PARTITIONING APPROACH¹

BY MAURICIO SADINLE

Carnegie Mellon University

Finding duplicates in homicide registries is an important step in keeping an accurate account of lethal violence. This task is not trivial when unique identifiers of the individuals are not available, and it is especially challenging when records are subject to errors and missing values. Traditional approaches to duplicate detection output independent decisions on the coreference status of each pair of records, which often leads to nontransitive decisions that have to be reconciled in some ad-hoc fashion. The task of finding duplicate records in a data file can be alternatively posed as partitioning the data file into groups of coreferent records. We present an approach that targets this partition of the file as the parameter of interest, thereby ensuring transitive decisions. Our Bayesian implementation allows us to incorporate prior information on the reliability of the fields in the data file, which is especially useful when no training data are available, and it also provides a proper account of the uncertainty in the duplicate detection decisions. We present a study to detect killings that were reported multiple times to the United Nations Truth Commission for El Salvador.

REFERENCES

- BILENKO, M., MOONEY, R. J., COHEN, W. W., RAVIKUMAR, P. and FIENBERG, S. E. (2003). Adaptive name matching in information integration. *IEEE Intelligent Systems* **18** 16–23.
- BUERGENTHAL, T. (1994). The United Nations Truth Commission for El Salvador. *Vanderbilt Journal of Transnational Law* **27** 497–544.
- BUERGENTHAL, T. (1996). La Comisión de la Verdad para El Salvador. In *Estudios Especializados de Derechos Humanos I* 11–62. Instituto Interamericano de Derechos Humanos, San José, Costa Rica.
- CHRISTEN, P. (2005). Probabilistic data generation for deduplication and data linkage. In *Proceedings of the Sixth International Conference on Intelligent Data Engineering and Automated Learning (IDEAL'05)* 109–116. Springer, Berlin.
- CHRISTEN, P. (2012a). *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Springer, Berlin.
- CHRISTEN, P. (2012b). A survey of indexing techniques for scalable record linkage and deduplication. *IEEE Transactions on Knowledge and Data Engineering* **24** 1537–1555.
- CHRISTEN, P. and PUDJIJONO, A. (2009). Accurate synthetic generation of realistic personal information. In *Advances in Knowledge Discovery and Data Mining* (T. Theeramunkong, B. Kijssirikul, N. Cercone and T.-B. Ho, eds.). *Lecture Notes in Computer Science* **5476** 507–514. Springer, Berlin.

Key words and phrases. Deduplication, duplicate detection, distribution on partitions, United Nations Truth Commission for El Salvador, entity resolution, homicide records, Hispanic names, human rights, record linkage, string similarity.

- CHRISTEN, P. and VATSALAN, D. (2013). Flexible and extensible generation and corruption of personal data. In *Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM 2013)*. ACM, New York.
- COMMISSION ON THE TRUTH FOR EL SALVADOR (1993). From madness to hope: The 12-year war in El Salvador: Report of the Commission on the Truth for El Salvador. Available at <http://www.usip.org/files/file/ElSalvador-Report.pdf> [Accessed October 15, 2014]. UN Security Council.
- CSARDI, G. and NEPUSZ, T. (2006). The igraph software package for complex network research. *InterJournal Complex Systems* 1695.
- ELMAGARMID, A. K., IPEIROTIS, P. G. and VERYKIOS, V. S. (2007). Duplicate record detection: A survey. *IEEE Transactions on Knowledge and Data Engineering* 19 1–16.
- FAY, R. E. (2004). An analysis of person duplication in census 2000. In *Proceedings of the Section on Survey Research Methods* 3478–3485. Amer. Statist. Assoc., Alexandria, VA.
- FELLEGI, I. P. and SUNTER, A. B. (1969). A theory for record linkage. *J. Amer. Statist. Assoc.* 64 1183–1210.
- FERNÁNDEZ, E. and GARCÍA, A. M. (2003). Accuracy of referencing of Spanish names in Medline. *The Lancet* 361 351–352.
- FORTINI, M., LISEO, B., NUCCITELLI, A. and SCANU, M. (2001). On Bayesian record linkage. *Research in Official Statistics* 4 185–198.
- FORTINI, M., NUCCITELLI, A., LISEO, B. and SCANU, M. (2002). Modeling issues in record linkage: A Bayesian perspective. In *Proceedings of the Section on Survey Research Methods* 1008–1013. Amer. Statist. Assoc., Alexandria, VA.
- GUTMAN, R., AFENDULIS, C. C. and ZASLAVSKY, A. M. (2013). A Bayesian procedure for file linking to analyze end-of-life medical costs. *J. Amer. Statist. Assoc.* 108 34–47. MR3174601
- HERZOG, T. N., SCHEUREN, F. J. and WINKLER, W. E. (2007). *Data Quality and Record Linkage Techniques*. Springer, New York.
- HOOVER GREEN, A. (2011). Repertoires of violence against noncombatants: The role of armed group institutions and ideologies. Ph.D. thesis, Yale Univ.
- HSU, W., LEE, M. L., LIU, B. and LING, T. W. (2000). Exploration mining in diabetic patients databases: Findings and conclusions. In *Proceedings of the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '00)* 430–436. ACM, New York.
- JARO, M. A. (1989). Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. *J. Amer. Statist. Assoc.* 84 414–420.
- KEENER, R., ROTHMAN, E. and STARR, N. (1987). Distributions on partitions. *Ann. Statist.* 15 1466–1481. MR0913568
- LARSEN, M. D. (2002). Comments on hierarchical Bayesian record linkage. In *Proceedings of the Section on Survey Research Methods* 1995–2000. Amer. Statist. Assoc., Alexandria, VA.
- LARSEN, M. D. (2005). Advances in record linkage theory: Hierarchical Bayesian record linkage theory. In *Proceedings of the Section on Survey Research Methods* 3277–3284. Amer. Statist. Assoc., Alexandria, VA.
- LARSEN, M. D. (2012). An experiment with hierarchical Bayesian record linkage. Preprint. Available at <http://arxiv.org/abs/1212.5203>.
- LARSEN, M. D. and RUBIN, D. B. (2001). Iterative automated record linkage using mixture models. *J. Amer. Statist. Assoc.* 96 32–41. MR1973781
- LITTLE, R. J. A. and RUBIN, D. B. (2002). *Statistical Analysis with Missing Data*, 2nd ed. Wiley, Hoboken, NJ. MR1925014
- LUM, K., PRICE, M. E. and BANKS, D. (2013). Applications of multiple systems estimation in human rights research. *Amer. Statist.* 67 191–200.
- MARSHALL, L. (2008). Potential duplicates in the census: Methodology and selection of cases for followup. In *Proceedings of the Section on Survey Research Methods* 4237–4244. Amer. Statist. Assoc., Alexandria, VA.

- MATSAKIS, N. E. (2010). Active duplicate detection with Bayesian nonparametric models. Ph.D. thesis, Massachusetts Institute of Technology.
- MCCULLAGH, P. (2011). Random permutations and partition models. In *International Encyclopedia of Statistical Science* 1170–1177. Springer, Berlin.
- MILLER, P. L., FRAWLEY, S. J. and SAYWARD, F. G. (2000). IMM/Scrub: A domain-specific tool for the deduplication of vaccination history records in childhood immunization registries. *Computers and Biomedical Research* **33** 126–143.
- R CORE TEAM (2013). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- ROTA, G.-C. (1964). The number of partitions of a set. *Amer. Math. Monthly* **71** 498–504. MR0161805
- RUIZ-PÉREZ, R., LÓPEZ-CÓZAR, E. D. and JIMÉNEZ-CONTRERAS, E. (2002). Spanish personal name variations in national and international biomedical databases: Implications for information retrieval and bibliometric studies. *Journal of the Medical Library Association* **90** 411–430.
- SADINLE, M. (2014). Supplement to “Detecting duplicates in a homicide registry using a Bayesian partitioning approach.” DOI:10.1214/14-AOAS779SUPP.
- SADINLE, M. and FIENBERG, S. E. (2013). A generalized Fellegi–Sunter framework for multiple record linkage with application to homicide record systems. *J. Amer. Statist. Assoc.* **108** 385–397. MR3174628
- SARIYAR, M. and BORG, A. (2010). The RecordLinkage package: Detecting errors in data. *The R Journal* **2** 61–67.
- SARIYAR, M., BORG, A. and POMMERENING, K. (2009). Evaluation of record linkage methods for iterative insertions. *Methods Inf. Med.* **48** 429–437.
- SARIYAR, M., BORG, A. and POMMERENING, K. (2012). Missing values in deduplication of electronic patient data. *Journal of the American Medical Informatics Association* **19** e76–e82.
- STEORTS, R. C., HALL, R. and FIENBERG, S. E. (2013). A Bayesian approach to graphical record linkage and deduplication. Preprint. Available at <http://arxiv.org/abs/1312.4645>.
- TANCREDI, A. and LISEO, B. (2011). A hierarchical Bayesian approach to record linkage and population size problems. *Ann. Appl. Stat.* **5** 1553–1585. MR2849786
- WINKLER, W. E. (1988). Using the EM algorithm for weight computation in the Fellegi–Sunter model of record linkage. In *Proceedings of the Section on Survey Research Methods* 667–671. Amer. Statist. Assoc., Alexandria, VA.
- WINKLER, W. E. (1989). Frequency-based matching in the Fellegi–Sunter model of record linkage. In *Proceedings of the Section on Survey Research Methods* 778–783. Amer. Statist. Assoc., Alexandria, VA.
- WINKLER, W. E. (1990). String comparator metrics and enhanced decision rules in the Fellegi–Sunter model of record linkage. In *Proceedings of the Section on Survey Research Methods* 354–359. Amer. Statist. Assoc., Alexandria, VA.

FUNCTIONAL RESPONSE ADDITIVE MODEL ESTIMATION WITH ONLINE VIRTUAL STOCK MARKETS

BY YINGYING FAN^{*,1}, NATASHA FOUTZ[†],
GARETH M. JAMES^{*} AND WOLFGANG JANK[‡]

University of Southern California^{}, University of Virginia[†]
and University of South Florida[‡]*

While functional regression models have received increasing attention recently, most existing approaches assume both a linear relationship and a scalar response variable. We suggest a new method, “Functional Response Additive Model Estimation” (FRAME), which extends the usual linear regression model to situations involving both functional predictors, $X_j(t)$, scalar predictors, Z_k , and functional responses, $Y(s)$. Our approach uses a penalized least squares optimization criterion to automatically perform variable selection in situations involving multiple functional and scalar predictors. In addition, our method uses an efficient coordinate descent algorithm to fit general nonlinear additive relationships between the predictors and response.

We develop our model for novel forecasting challenges in the entertainment industry. In particular, we set out to model the decay rate of demand for Hollywood movies using the predictive power of online virtual stock markets (VSMs). VSMs are online communities that, in a market-like fashion, gather the crowds’ prediction about demand for a particular product. Our fully functional model captures the pattern of pre-release VSM trading prices and provides superior predictive accuracy of a movie’s post-release demand in comparison to traditional methods. In addition, we propose graphical tools which give a glimpse into the causal relationship between market behavior and box office revenue patterns, and hence provide valuable insight to movie decision makers.

REFERENCES

- BAR-JOSEPH, Z., GERBER, G. K., GIFFORD, D. K., JAAKKOLA, T. S. and SIMON, I. (2003). Continuous representations of time-series gene expression data. *J. Comput. Biol.* **10** 341–356.
- BASS, F. M., GORDON, K., FERGUSON, T. L. and GITH, M. L. (2001). DIRECTV: Forecasting diffusion of a new technology prior to product launch. *Interfaces* **31** S82–S93.
- CHEN, D., HALL, P. and MÜLLER, H.-G. (2011). Single and multiple index functional regression models with nonparametric link. *Ann. Statist.* **39** 1720–1747. [MR2850218](#)
- CRAINICEANU, C., CAFFO, B. and MORRIS, J. (2011). Multilevel functional data analysis. In *The SAGE Handbook of Multilevel Modeling*. SAGE, Thousand Oaks, CA.
- DI, C.-Z., CRAINICEANU, C. M., CAFFO, B. S. and PUNJABI, N. M. (2009). Multilevel functional principal component analysis. *Ann. Appl. Stat.* **3** 458–488. [MR2668715](#)
- EINAV, L. (2007). Seasonality in the U.S. motion picture industry. *Rand J. Econ.* **38** 127–145.

Key words and phrases. Functional data, nonlinear regression, penalty functions, forecasting, virtual markets, movies, Hollywood.

- EINAV, L. (2010). Not all rivals look alike: Estimating an equilibrium model of the release date timing game. *Econ. Inq.* **48** 369–390.
- ELIASHBERG, J., ELBERSE, A. and LEENDERS, M. (2006). The motion picture industry: Critical issues in practice, current research, and new research directions. *Mark. Sci.* **25** 638–661.
- FAN, Y., JAMES, G. and RADCHENKO, P. (2014). Functional additive regression. Under review.
- FERRATY, F. and VIEU, P. (2002). The functional nonparametric model and application to spectrometric data. *Comput. Statist.* **17** 545–564. [MR1952697](#)
- FERRATY, F. and VIEU, P. (2003). Curves discrimination: A nonparametric functional approach. *Comput. Statist. Data Anal.* **44** 161–173. [MR2020144](#)
- FOUTZ, N. and JANK, W. (2010). Pre-release demand forecasting for motion pictures using functional shape analysis of virtual stock markets. *Mark. Sci.* **29** 568–579.
- HASTIE, T., TIBSHIRANI, R. and FRIEDMAN, J. (2001). *The Elements of Statistical Learning*. Springer, New York. [MR1851606](#)
- JAMES, G. M. and HASTIE, T. J. (2001). Functional linear discriminant analysis for irregularly sampled curves. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** 533–550. [MR1858401](#)
- JAMES, G. M., HASTIE, T. J. and SUGAR, C. A. (2000). Principal component models for sparse functional data. *Biometrika* **87** 587–602. [MR1789811](#)
- JAMES, G. M. and SILVERMAN, B. W. (2005). Functional adaptive model estimation. *J. Amer. Statist. Assoc.* **100** 565–576. [MR2160560](#)
- JAMES, G. M. and SUGAR, C. A. (2003). Clustering for sparsely sampled functional data. *J. Amer. Statist. Assoc.* **98** 397–408. [MR1995716](#)
- JANK, W. and SHMUELI, G. (2006). Functional data analysis in electronic commerce research. *Statist. Sci.* **21** 155–166. [MR2324075](#)
- RAMSAY, J. O. and SILVERMAN, B. W. (2005). *Functional Data Analysis*, 2nd ed. Springer, New York. [MR2168993](#)
- RAVIKUMAR, P., LAFFERTY, J., LIU, H. and WASSERMAN, L. (2009). Sparse additive models. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **71** 1009–1030. [MR2750255](#)
- RICE, J. A. and WU, C. O. (2001). Nonparametric mixed effects models for unequally sampled noisy curves. *Biometrics* **57** 253–259. [MR1833314](#)
- SAWHNEY, M. S. and ELIASHBERG, J. (1996). A parsimonious model for forecasting gross box office revenues of motion pictures. *Mark. Sci.* **15** 113–131.
- SIMON, N., FRIEDMAN, J., HASTIE, T. and TIBSHIRANI, R. (2013). A sparse-group lasso. *J. Comput. Graph. Statist.* **22** 231–245. [MR3173712](#)
- SPANN, M. and SKIERA, B. (2003). Internet-based virtual stock markets for business forecasting. *Manage. Sci.* **49** 1310–1326.
- TIBSHIRANI, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. [MR1379242](#)
- YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005a). Functional linear regression analysis for longitudinal data. *Ann. Statist.* **33** 2873–2903. [MR2253106](#)
- YAO, F., MÜLLER, H.-G. and WANG, J.-L. (2005b). Functional data analysis for sparse longitudinal data. *J. Amer. Statist. Assoc.* **100** 577–590. [MR2160561](#)
- YUAN, M. and LIN, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **68** 49–67. [MR2212574](#)
- ZEGER, S. L. and DIGGLE, P. J. (1994). Semiparametric models for longitudinal data with applications to CD4 cell numbers in HIV seroconverters. *Biometrics* **50** 689–699.
- ZHANG, S., JANK, W. and SHMUELI, G. (2010). Real-time forecasting of online auctions via functional k-nearest neighbors. *Int. J. Forecast.* **26** 666–683.
- ZOU, H. and LI, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Ann. Statist.* **36** 1509–1533. [MR2435443](#)

PROBIT MODELS FOR CAPTURE–RECAPTURE DATA SUBJECT TO IMPERFECT DETECTION, INDIVIDUAL HETEROGENEITY AND MISIDENTIFICATION¹

BY BRETT T. MCCLINTOCK*, LARISSA L. BAILEY[†],
BRIAN P. DREHER[‡] AND WILLIAM A. LINK[§]

*NOAA National Marine Mammal Laboratory**, *Colorado State University[†]*,
Colorado Parks and Wildlife[‡], and *USGS Patuxent Wildlife Research Center[§]*

As noninvasive sampling techniques for animal populations have become more popular, there has been increasing interest in the development of capture–recapture models that can accommodate both imperfect detection and misidentification of individuals (e.g., due to genotyping error). However, current methods do not allow for individual variation in parameters, such as detection or survival probability. Here we develop misidentification models for capture–recapture data that can simultaneously account for temporal variation, behavioral effects and individual heterogeneity in parameters. To facilitate Bayesian inference using our approach, we extend standard probit regression techniques to latent multinomial models where the dimension and zeros of the response cannot be observed. We also present a novel Metropolis–Hastings within Gibbs algorithm for fitting these models using Markov chain Monte Carlo. Using closed population abundance models for illustration, we re-visit a DNA capture–recapture population study of black bears in Michigan, USA and find evidence of misidentification due to genotyping error, as well as temporal, behavioral and individual variation in detection probability. We also estimate a salamander population of known size from laboratory experiments evaluating the effectiveness of a marking technique commonly used for amphibians and fish. Our model was able to reliably estimate the size of this population and provided evidence of individual heterogeneity in misidentification probability that is attributable to variable mark quality. Our approach is more computationally demanding than previously proposed methods, but it provides the flexibility necessary for a much broader suite of models to be explored while properly accounting for uncertainty introduced by misidentification and imperfect detection. In the absence of misidentification, our probit formulation also provides a convenient and efficient Gibbs sampler for Bayesian analysis of traditional closed population capture–recapture data.

REFERENCES

- ALBERT, J. H. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. [MR1224394](#)
- BAILEY, L. L. (2004). Evaluating elastomer marking and photo-identification methods for terrestrial salamanders: Marking effects and observer bias. *Herpetological Review* **35** 38–41.

Key words and phrases. Data augmentation, individual heterogeneity, latent multinomial, mark-recapture, missing data, population size, probit regression, record linkage.

- BASU, S. and EBRAHIMI, N. (2001). Bayesian capture–recapture methods for error detection and estimation of population size: Heterogeneity and dependence. *Biometrika* **88** 269–279. [MR1841274](#)
- BONNER, S. J. and HOLMBERG, J. (2013). Mark-recapture with multiple, non-invasive marks. *Biometrics* **69** 766–775. [MR3106605](#)
- BONNER, S. J. and SCHOFIELD, M. R. (2013). MC(MC)MC: Exploring Monte Carlo integration with MCMC for mark-recapture models with individual covariates. *Methods in Ecology and Evolution*. DOI:10.1111/2041-210X.12095.
- CASTLEDINE, B. J. (1981). A Bayesian analysis of multiple-recapture sampling for a closed population. *Biometrika* **68** 197–210. [MR0614956](#)
- COULL, B. A. and AGRESTI, A. (1999). The use of mixed logit models to reflect heterogeneity in capture–recapture studies. *Biometrics* **55** 294–301.
- DARROCH, J. N. (1958). The multiple-recapture census. I. Estimation of a closed population. *Biometrika* **45** 343–359. [MR0119360](#)
- DORAZIO, R. M. and RODRIGUEZ, D. T. (2012). A Gibbs sampler for Bayesian analysis of site-occupancy data. *Methods in Ecology and Evolution* **3** 1093–1098.
- DREHER, B. P., WINTERSTEIN, S. R., SCRIBNER, K. T., LUKACS, P. M., ETTER, D. R., ROSA, G. J. M., LOPEZ, V. A., LIBANTS, S. and FILCEK, K. B. (2007). Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bears. *Journal of Wildlife Management* **71** 2684–2693.
- FIENBERG, S. E., JOHNSON, M. S. and JUNKER, B. W. (1999). Classical multilevel and Bayesian approaches to population size estimation using multiple lists. *J. Roy. Statist. Soc. Ser. A* **162** 383–406.
- FIENBERG, S. E. and MANRIQUE-VALLIER, D. (2009). Integrated methodology for multiple systems estimation and record linkage using a missing data formulation. *ASTA Adv. Stat. Anal.* **93** 49–60. [MR2476299](#)
- GEORGE, E. I. and ROBERT, C. P. (1992). Capture–recapture estimation via Gibbs sampling. *Biometrika* **79** 677–683. [MR1209469](#)
- GIMENEZ, O. and CHOQUET, R. (2010). Individual heterogeneity in studies on marked animals using numerical integration: Capture–recapture mixed models. *Ecology* **91** 951–957.
- HALL, A. J., MCCONNELL, B. J. and BARKER, R. J. (2001). Factors affecting first-year survival in grey seals and their implications for life history strategy. *Journal of Animal Ecology* **70** 138–149.
- HASTINGS, K. K., HIBY, L. A. and SMALL, R. J. (2008). Evaluation of a computer-assisted photograph-matching system to monitor naturally marked harbor seals at Tugidak Island, Alaska. *Journal of Mammalogy* **89** 1201–1211.
- JOHNSON, D. S., CONN, P. B., HOOTEN, M. B., RAY, J. C. and PONE, B. A. (2013). Spatial occupancy models for large data sets. *Ecology* **94** 801–808.
- KARANTH, K. U. and NICHOLS, J. D. (1998). Estimation of tiger densities in India using photographic captures and recaptures. *Ecology* **79** 2852–2862.
- KAUFFMAN, M. J., FRICK, W. F. and LINTHICUM, J. (2003). Estimation of habitat-specific demography and population growth for peregrine falcons in California. *Ecological Applications* **13** 1802–1816.
- KERNIGHAN, B. W. and RITCHIE, D. M. (1988). *The C Programming Language*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ.
- KING, R. and BROOKS, S. P. (2008). On the Bayesian estimation of a closed population size in the presence of heterogeneity and model uncertainty. *Biometrics* **64** 816–824. [MR2526632](#)
- KING, R., BROOKS, S. P. and COULSON, T. (2008). Analyzing complex capture–recapture data in the presence of individual and temporal covariates and model uncertainty. *Biometrics* **64** 1187–1195. [MR2522267](#)
- LANGTIMM, C. A., O’SHEA, T. J., PRADEL, R. and BECK, C. A. (1998). Estimates of annual survival probabilities for adult Florida manatees *Trichechus manatus latirostris*. *Ecology* **79** 981–997.

- LINK, W. A. (2013). A cautionary note on the discrete uniform prior for the binomial N . *Ecology* **94** 2173–2179.
- LINK, W. A., YOSHIKAZI, J., BAILEY, L. L. and POLLOCK, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* **66** 178–185. [MR2756704](#)
- LUKACS, P. M. and BURNHAM, K. P. (2005). Estimating population size from DNA-based closed capture–recapture data incorporating genotyping error. *Journal of Wildlife Management* **69** 396–403.
- MACKEY, B. L., DURBAN, J. W., MIDDLEMAS, S. J. and THOMPSON, P. M. (2008). A Bayesian estimate of harbour seal survival using sparse photo-identification data. *Journal of Zoology* **274** 18–27.
- MANRIQUE-VALLIER, D. and FIENBERG, S. E. (2008). Population size estimation using individual level mixture models. *Biom. J.* **50** 1051–1063. [MR2649394](#)
- MCCLINTOCK, B. T., CONN, P. B., ALONSO, R. S. and CROOKS, K. R. (2013a). Integrated modeling of bilateral photo-identification data in mark-recapture analyses. *Ecology* **94** 1464–1471.
- MCCLINTOCK, B. T., HILL, J. M., FRITZ, L., CHUMBLEY, K., LUXA, K. and DIEFENBACH, D. R. (2013b). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology and Evolution*. DOI:10.1111/2041-210X.12140.
- MORRISON, T. A., YOSHIKAZI, J., NICHOLS, J. D. and BOLGER, D. T. (2011). Estimating survival in photographic capture–recapture studies: Overcoming misidentification error. *Methods in Ecology and Evolution* **2** 454–463.
- OTIS, D. L., BURNHAM, K. P., WHITE, G. C. and ANDERSON, D. R. (1978). Statistical-inference from capture data on closed animal populations. *Wildlife Monographs* **62** 7–135.
- PLEDGER, S. (2000). Unified maximum likelihood estimates for closed capture–recapture models using mixtures. *Biometrics* **56** 434–442.
- PLUMMER, M., BEST, N., COWLES, K. and VINES, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6** 7–11.
- POLSON, N. G., SCOTT, J. G. and WINDLE, J. (2013). Bayesian inference for logistic models using Pólya–Gamma latent variables. *J. Amer. Statist. Assoc.* **108** 1339–1349. [MR3174712](#)
- PRADEL, R. (2005). Multievent: An extension of multistate capture–recapture models to uncertain states. *Biometrics* **61** 442–447. [MR2140915](#)
- R CORE TEAM (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
- ROYLE, J. A. (2008). Modeling individual effects in the Cormack–Jolly–Seber model: A state-space formulation. *Biometrics* **64** 364–370, 664. [MR2432405](#)
- ROYLE, J. A., DORAZIO, R. M. and LINK, W. A. (2007). Analysis of multinomial models with unknown index using data augmentation. *J. Comput. Graph. Statist.* **16** 67–85. [MR2345748](#)
- RUPELL, E. W., RILEY, S. P. D., DOUGLAS, M. R., POLLINGER, J. P. and CROOKS, K. R. (2009). Estimating bobcat population sizes and densities in a fragmented urban landscape using noninvasive capture–recapture sampling. *Journal of Mammalogy* **90** 129–135.
- TANCREDI, A. and LISEO, B. (2011). A hierarchical Bayesian approach to record linkage and population size problems. *Ann. Appl. Stat.* **5** 1553–1585. [MR2849786](#)
- TANCREDI, A., AUGER-MÉTHÉ, M., MARCOUX, M. and LISEO, B. (2013). Accounting for matching uncertainty in two stage capture–recapture experiments using photographic measurements of natural marks. *Environ. Ecol. Stat.* **20** 647–665. [MR3128764](#)
- THOMPSON, S. K. (1992). *Sampling*. Wiley, New York. [MR1193031](#)
- WHITE, G. C. and BURNHAM, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study* **46 Supplement** 120–138.
- WILLIAMS, B. K., NICHOLS, J. D. and CONROY, M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.

- WRIGHT, J. A., BARKER, R. J., SCHOFIELD, M. R., FRANTZ, A. C., BYROM, A. E. and GLEESON, D. M. (2009). Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* **65** 833–840. [MR2649856](#)
- YIP, P. S. F., BRUNO, G., TAJIMA, N., SEBER, G. A. F., BUCKLAND, S. T., CORMACK, R. M., UNWIN, N., CHANG, Y. F., FIENBERG, S. E., JUNKER, B. W., LAPORTE, R. E., LIBMAN, I. M. and MCCARTY, D. J. (1995a). Capture–recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology* **142** 1047–1058.
- YIP, P. S. F., BRUNO, G., TAJIMA, N., SEBER, G. A. F., BUCKLAND, S. T., CORMACK, R. M., UNWIN, N., CHANG, Y. F., FIENBERG, S. E., JUNKER, B. W., LAPORTE, R. E., LIBMAN, I. M. and MCCARTY, D. J. (1995b). Capture–recapture and multiple-record systems estimation II: Applications in human diseases. *American Journal of Epidemiology* **142** 1059–1068.
- YOSHIZAKI, J. (2007). Use of natural tags in closed population capture–recapture studies: Modeling misidentification. Ph.D. thesis, North Carolina State Univ., Raleigh, NC.
- YOSHIZAKI, J., POLLOCK, K. H., BROWNIE, C. and WEBSTER, R. A. (2009). Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. *Ecology* **90** 3–9.
- YOSHIZAKI, J., BROWNIE, C., POLLOCK, K. H. and LINK, W. A. (2011). Modeling misidentification errors that result from use of genetic tags in capture–recapture studies. *Environ. Ecol. Stat.* **18** 27–55. [MR2783681](#)

DO DEBIT CARDS INCREASE HOUSEHOLD SPENDING? EVIDENCE FROM A SEMIPARAMETRIC CAUSAL ANALYSIS OF A SURVEY

BY ANDREA MERCATANTI¹ AND FAN LI²

Bank of Italy and Duke University

Motivated by recent findings in the field of consumer science, this paper evaluates the causal effect of debit cards on household consumption using population-based data from the Italy Survey on Household Income and Wealth (SHIW). Within the Rubin Causal Model, we focus on the estimand of population average treatment effect for the treated (PATT). We consider three existing estimators, based on regression, mixed matching and regression, propensity score weighting, and propose a new doubly-robust estimator. Semiparametric specification based on power series for the potential outcomes and the propensity score is adopted. Cross-validation is used to select the order of the power series. We conduct a simulation study to compare the performance of the estimators. The key assumptions, overlap and unconfoundedness, are systematically assessed and validated in the application. Our empirical results suggest statistically significant positive effects of debit cards on the monthly household spending in Italy.

REFERENCES

- ABADIE, A. and IMBENS, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica* **74** 235–267. [MR2194325](#)
- ABADIE, A. and IMBENS, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *J. Bus. Econom. Statist.* **29** 1–11. [MR2789386](#)
- ABADIE, A., DRUKKER, D., HERR, J. L. and IMBENS, G. W. (2003). Implementing matching estimators for average treatment effects in Stata. *The Stata Journal* **4** 290–311.
- ATTANASIO, O., GUISO, L. and JAPPELLI, T. (2002). The demand for money, financial innovation and the welfare cost of inflation: An analysis with household data. *Journal of Political Economy* **110** 318–351.
- BURMAN, D. (1974). Do people overspend with credit cards? *The Journal of Consumer Credit Management* **5** 98–103.
- CHEN, X. (2007). Large sample sieve estimation of semi-nonparametric models. In *Handbook of Econometrics* 6 (J. Heckman and E. Leamer, eds.) 5549–5632. Elsevier Science, New York.
- COLE, C. (1998). Identifying interventions to reduce credit card misuse through consumer behavior research. In *Proc. of the Marketing and Public Policy Conference* 11–13. Georgetown Univ. Press, Washington, DC.
- COMMITTEE ON PAYMENT AND SETTLEMENT SYSTEM (2001). *A Glossary of Terms Used in Payments and Settlement Systems*. Bank for International Settlements, Basel, Switzerland.
- CRUMP, R. K., HOTZ, V. J., IMBENS, G. W. and MITNIK, O. A. (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics* **90** 389–405.

Key words and phrases. Causal inference, potential outcomes, payment instruments, power series, propensity score, overlap, unconfoundedness.

- DAS, M., NEWEY, W. K. and VELLA, F. (2003). Nonparametric estimation of sample selection models. *Rev. Econom. Stud.* **70** 33–58. [MR1952565](#)
- FEINBERG, R. A. (1986). Credits cards as spending facilitating stimuli: A conditioning interpretation. *Journal of Consumer Research* **13** 348–356.
- HAHN, J. (1998). On the role of the propensity score in efficient semiparametric estimation of average treatment effects. *Econometrica* **66** 315–331. [MR1612242](#)
- HAUSMAN, J. A. and NEWEY, W. K. (1995). Nonparametric estimation of exact consumers surplus and deadweight loss. *Econometrica* **63** 1445–1476.
- HEAT, C. and SOLL, J. B. (1996). Mental budgeting and consumer decisions. *Journal of Consumer Research* **23** 40–52.
- HECKMAN, J. J., ICHIMURA, H. and TODD, P. (1997). Matching as an econometric evaluation estimator: Evidence from a job training program. *Rev. Econom. Stud.* **64** 605–654.
- HELSON, H. (1964). *Adaptation-Level Theory*. Harper & Row, New York.
- HIRANO, K. and IMBENS, G. W. (2001). Estimation of causal effects using propensity score weighting: An application to data on right heart catheterization. *Health Services and Outcomes Research Methodology* **2** 259–278.
- HIRANO, K., IMBENS, G. W. and RIDDER, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71** 1161–1189. [MR1995826](#)
- HIRSCHMAN, E. (1979). Differences in consumer purchase behavior by credit card payment system. *Journal of Consumer Research* **6** 58–66.
- 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](#)
- HSEE, C. K., LOEWENSTEIN, G. F., BLOUNT, S. and BAZERMAN, M. H. (1999). Preference reversals between joint and separate evaluation of options: A review and theoretical analysis. *Psychological Bulletin* **125** 576–590.
- 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](#)
- IMBENS, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics* **86** 4–29.
- IMBENS, G. W., NEWEY, W. and RIDDER, G. (2005). Mean-square-error calculations for average treatment effects. Working Paper 05-34, IEPR.
- KAHNEMAN, D. and MILLER, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review* **93** 136–153.
- LI, F., MORGAN, L. K. and ZASLAVSKY, A. M. (2014). Balancing covariates via propensity score weighting. Available at [arXiv:1404.1785](#).
- MANN, R. J. (2006). *Charging Ahead: The Growth and Regulation of Payment Card Markets Around the World*. Cambridge Univ. Press, Cambridge.
- MOREWEDGE, C. K., HOTZMAN, L. and EPLEY, N. (2007). Unfixed resources: Perceived costs, consumption, and the accessible account effect. *Journal of Consumer Research* **34** 459–467.
- PARDUCCI, A. (1995). *Happiness, Pleasure and Judgment: The Contextual Theory and Its Application*. Lawrence Erlbaum Associates, Mahwah, NJ.
- PRELEC, D. and LOEWENSTEIN, G. F. (1998). The red and the black: Mental accounting of savings and debt. *Marketing Science* **17** 4–28.
- PRELEC, D. and SIMESTER, D. (2001). Always leave home without it: A further investigation of the credit-card effect on willingness to pay. *Marketing Letters* **12** 5–12.
- ROBINS, J. M., ROTNITZKY, A. and ZHAO, L. P. (1995). Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J. Amer. Statist. Assoc.* **90** 106–121. [MR1325118](#)

- ROSENBAUM, P. R. and RUBIN, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70** 41–55. [MR0742974](#)
- RUBIN, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66** 688–701.
- RUBIN, D. B. (1977). Assignment to treatment group on the basis of a covariate. *Journal of Educational Statistics* **2** 1–26.
- RUBIN, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Ann. Statist.* **6** 34–58. [MR0472152](#)
- RUBIN, D. B. (1980). Discussion of “Randomization analysis of experimental data: The Fisher randomization test” by D. Basu. *J. Amer. Statist. Assoc.* **75** 591–593.
- SOMAN, D. (2001). Effects of payment mechanism on spending behavior: The role of rehearsal and immediacy of payments. *Journal of Consumer Research* **27** 460–474.
- SOMAN, D. and CHEEMA, A. (2002). The effects of credit on spending decisions: The role of credit limit and credibility. *Marketing Science* **21** 32–53.
- THALER, R. H. (1985). Mental accounting and consumer choice. *Marketing Science* **4** 199–214.
- THALER, R. H. (1999). Mental accounting matter. *Journal of Behavioral Decision Making* **12** 183–206.
- TOKUNAGA, H. (1993). The use and abuse of consumer credit: Applications of psychological theory and research. *Journal of Economic Psychology* **14** 285–316.

REDUCED-RANK SPATIO-TEMPORAL MODELING OF AIR POLLUTION CONCENTRATIONS IN THE MULTI-ETHNIC STUDY OF ATHEROSCLEROSIS AND AIR POLLUTION¹

BY CASEY OLIVES*, LIANNE SHEPPARD*, JOHAN LINDSTRÖM[†],
PAUL D. SAMPSON*, JOEL D. KAUFMAN* AND ADAM A. SZPIRO*

University of Washington and Lund University[†]*

There is growing evidence in the epidemiologic literature of the relationship between air pollution and adverse health outcomes. Prediction of individual air pollution exposure in the Environmental Protection Agency (EPA) funded Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air) study relies on a flexible spatio-temporal prediction model that integrates land-use regression with kriging to account for spatial dependence in pollutant concentrations. Temporal variability is captured using temporal trends estimated via modified singular value decomposition and temporally varying spatial residuals. This model utilizes monitoring data from existing regulatory networks and supplementary MESA Air monitoring data to predict concentrations for individual cohort members.

In general, spatio-temporal models are limited in their efficacy for large data sets due to computational intractability. We develop reduced-rank versions of the MESA Air spatio-temporal model. To do so, we apply low-rank kriging to account for spatial variation in the mean process and discuss the limitations of this approach. As an alternative, we represent spatial variation using thin plate regression splines. We compare the performance of the outlined models using EPA and MESA Air monitoring data for predicting concentrations of oxides of nitrogen (NO_x)—a pollutant of primary interest in MESA Air—in the Los Angeles metropolitan area via cross-validated R^2 .

Our findings suggest that use of reduced-rank models can improve computational efficiency in certain cases. Low-rank kriging and thin plate regression splines were competitive across the formulations considered, although TPRS appeared to be more robust in some settings.

REFERENCES

- BANERJEE, S., GELFAND, A. E., FINLEY, A. O. and SANG, H. (2008). Gaussian predictive process models for large spatial data sets. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **70** 825–848. MR2523906
- BRAUER, M., HOEK, G., VAN VLIET, P., MELIEFSTE, K., FISCHER, P., GEHRING, U., HEINRICH, J., CYRYS, J., BELLANDER, T., LEWNE, M. and BRUNEKREEF, B. (2003). Estimating long-term average particulate air pollution concentrations: Application of traffic indicators and geographic information systems. *Epidemiology* **14** 228–239.
- BYRD, R. H., LU, P., NOCEDAL, J. and ZHU, C. Y. (1995). A limited memory algorithm for bound constrained optimization. *SIAM J. Sci. Comput.* **16** 1190–1208. MR1346301

Key words and phrases. Spatiotemporal modeling, reduced-rank, air pollution, kriging, thin plate splines.

- CARROLL, M. L., DIMICELI, C. M., SOHLBERG, R. A. and TOWNSHEND, J. R. G. (2004). 250m MODIS Normalized Difference Vegetation Index, 250ndvi28920033435, Collection 4. Univ. Maryland, College Park, Maryland, Day 289, 2003.
- COHEN, M. A., ADAR, S. D., ALLEN, R. W., AVOL, E., CURL, C. L., GOULD, T., HARDIE, D., HO, A., KINNEY, P., LARSON, T. V., SAMPSON, P., SHEPPARD, L., STUKOVSKY, K. D., SWAN, S. S., LIU, L. J. S. and KAUFMAN, J. D. (2009). Approach to estimating participant pollutant exposures in the multi-ethnic study of atherosclerosis and air pollution (MESA air). *Environmental Science & Technology* **43** 4687–4693.
- CRAINICEANU, C. M., DIGGLE, P. J. and ROWLINGSON, B. (2008). Bivariate binomial spatial modeling of *Loa loa* prevalence in tropical Africa. *J. Amer. Statist. Assoc.* **103** 21–37. [MR2420211](#)
- DOCKERY, D. W., POPE, C. A. 3RD, XU, X., SPENGLER, J. D., WARE, J. H., FAY, M. E., FERRIS, B. G. JR and SPEIZER, F. E. (1993). An association between air pollution and mortality in six U.S. cities. *N. Engl. J. Med.* **329** 1753–1759.
- ECKHOFF, P. A. and BRAVERMAN, T. N. (1995). Addendum to the user's guide to CAL3QHC version 2.0 (CAL3QHCR user's guide). Technical Support Division, Office of Air Quality Planning and Standards, Research Triangle Park, NC.
- FRY, J., XIAN, G., JIN, S., DEWITZ, J., HOMER, C., YANG, L., BARNES, C., HEROLD, N. and WICKHAM, J. (2011). Completion of the 2006 National Land Cover Database for the Conterminous United States. *Photogrammetric Engineering & Remote Sensing* **77** 858–864.
- FUENTES, M. (2007). Approximate likelihood for large irregularly spaced spatial data. *J. Amer. Statist. Assoc.* **102** 321–331. [MR2345545](#)
- FUENTES, M., GUTTORP, P. and SAMPSON, P. D. (2006). Using transforms to analyze space-time processes. *Monogr. Statist. Appl. Probab.* **107** 77.
- GELFAND, A. E., BANERJEE, S. and GAMERMAN, D. (2005). Spatial process modelling for univariate and multivariate dynamic spatial data. *Environmetrics* **16** 465–479. [MR2147537](#)
- GREEN, P. J. and SILVERMAN, B. W. (1994). *Nonparametric Regression and Generalized Linear Models: A Roughness Penalty Approach*. *Monographs on Statistics and Applied Probability* **58**. Chapman & Hall, London. [MR1270012](#)
- HODGES, J. S. (2013). *Richly Parameterized Linear Models*. Chapman & Hall, Boca Raton.
- HODGES, J. and CLAYTON, M. K. (2011). Random effects old and new. Technical report, Univ. Minnesota, Minneapolis, MN.
- HOEK, G., BELEN, R., DE HOOGH, K., VIENNEAU, D., GULLIVER, J., FISCHER, P. and BRIGGS, D. (2008). A review of land-use regression models to assess spatial variation of outdoor air pollution. *Atmospheric Environment* **42** 7561–7578.
- JERRETT, M., ARAIN, A., KANAROGLOU, P., BECKERMAN, B., POTOGLOU, D., SAHSUVAROGLU, T., MORRISON, J. and GIOVIS, C. (2005a). A review and evaluation of intraurban air pollution exposure models. *J. Expo. Anal. Environ. Epidemiol.* **15** 185–204.
- JERRETT, M., BURNETT, R. T., MA, R., POPE, C. A. 3RD, KREWSKI, D., NEWBOLD, K. B., THURSTON, G., SHI, Y., FINKELSTEIN, N., CALLE, E. E. and THUN, M. J. (2005b). Spatial analysis of air pollution and mortality in los angeles. *Epidemiology* **16** 727–736.
- KAMMANN, E. E. and WAND, M. P. (2003). Geoadditive models. *J. Roy. Statist. Soc. Ser. C* **52** 1–18. [MR1963210](#)
- KAUFMAN, J. K., ADAR, S. D., ALLEN, R. W., BARR, R. G., BUDOFF, M. J., BURKE, G. L., CASILLAS, A. M., COHEN, M. A., CURL, C. L., DAVIGLUS, M. L., DIEZ ROUX, A. V., JACOBS, D. R. JR, KRONMAL, R. A., LARSON, T. V., LIU, S. L., LUMLEY, T., NAVASACIEN, A., O'LEARY, D. H., ROTTER, J. I., SAMPSON, P. D., SHEPPARD, L., SISCOVICK, D. S., STEIN, J. H., SZPIRO, A. A. and TRACY, R. P. (2012). Prospective study of particulate air pollution exposures, subclinical atherosclerosis, and clinical cardiovascular disease the multi-ethnic study of atherosclerosis and air pollution (MESA air). *American Journal of Epidemiology* **176** 825–837.

- KELLER, J. P., OLIVES, C., KIM, S.-Y., SHEPPARD, L., SAMPSON, P. D., SZPIRO, A. A., ORON, A. P., LINDSTRÖM, J., VEDAL, S. and KAUFMAN, J. D. (2014). A unified spatiotemporal modeling approach for prediction of multiple air pollutants in the multi-ethnic study of atherosclerosis and air pollution. *Environ. Health Perspect.* To appear.
- KIMELDORF, G. S. and WAHBA, G. (1970). A correspondence between Bayesian estimation on stochastic processes and smoothing by splines. *Ann. Math. Statist.* **41** 495–502. [MR0254999](#)
- KÜNZLI, N., JERRETT, M., MACK, W. J., BECKERMAN, B., LABREE, L., GILLILAND, F., THOMAS, D., PETERS, J. and HODIS, H. N. (2005). Ambient air pollution and atherosclerosis in Los Angeles. *Environ. Health Perspect.* **113** 201–206.
- LINDSTRÖM, J., SZPIRO, A. A., SAMPSON, P. D., ORON, A., RICHARDS, M., LARSON, T. and SHEPPARD, L. (2013). A flexible spatio-temporal model for air pollution with spatial and spatio-temporal covariates. *Environ. Ecol. Stat.* 1–23.
- MILLER, K. A., SISCOVICK, D. S., SHEPPARD, L., SHEPHERD, K., SULLIVAN, J. H., ANDERSON, G. L. and KAUFMAN, J. D. (2007). Long-term exposure to air pollution and incidence of cardiovascular events in women. *N. Engl. J. Med.* **356** 447–458.
- NYCHKA, D. W. (2000). Spatial-process estimates as smoothers. In *Smoothing and Regression: Approaches, Computation, and Application* 393–424. Wiley, New York.
- NYCHKA, D. and SALTZMAN, N. (1998). Design of air quality networks. In *Case Studies in Environmental Statistics* (D. Nychka, L. Cox and W. Piegorsch, eds.). *Lecture Notes in Statistics* **132** 51–76. Springer, New York.
- OLIVES, C., SHEPPARD, L., LINDSTRÖM, J., SAMPSON, P. D., KAUFMAN, J. D. and SZPIRO, A. A. (2014). Supplement to “Reduced-rank spatio-temporal modeling of air pollution concentrations in the Multi-Ethnic Study of Atherosclerosis and Air Pollution.” DOI:[10.1214/14-AOAS786SUPP](#).
- PACE, R. and LESAGE, J. (2009). A sampling approach to estimate the log determinant used in spatial likelihood problems. *Journal of Geographical Systems* **11** 209–225.
- PACIOREK, C. J., YANOSKY, J. D., PUETT, R. C., LADEN, F. and SUH, H. H. (2009). Practical large-scale spatio-temporal modeling of particulate matter concentrations. *Ann. Appl. Stat.* **3** 370–397. [MR2668712](#)
- POPE, C. A. 3RD, THUN, M. J., NAMBOODIRI, M. M., DOCKERY, D. W., EVANS, J. S., SPEIZER, F. E. and HEATH, C. W. JR (1995). Particulate air pollution as a predictor of mortality in a prospective study of U.S. adults. *Am. J. Respir. Crit. Care Med.* **151** 669–674.
- POPE, C. A. 3RD, BURNETT, R. T., THUN, M. J., CALLE, E. E., KREWSKI, D., ITO, K. and THURSTON, G. D. (2002). Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *Journal of the American Medical Association* **287** 1132–1141.
- RITZ, B., WILHELM, M. and ZHAO, Y. (2006). Air pollution and infant death in southern California, 1989–2000. *Pediatrics* **118** 493–502.
- RUPPERT, D., WAND, M. P. and CARROLL, R. J. (2003). *Semiparametric Regression. Cambridge Series in Statistical and Probabilistic Mathematics* **12**. Cambridge Univ. Press, Cambridge. [MR1998720](#)
- SAMET, J. M., DOMINICI, F., CURRIERO, F. C., COURSAK, I. and ZEGER, S. L. (2000). Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. *N. Engl. J. Med.* **343** 1742–1749.
- SAMPSON, P. D., SZPIRO, A. A., SHEPPARD, L., LINDSTRÖM, J. and KAUFMAN, J. D. (2011). Pragmatic estimation of spatio-temporal air quality model with irregular monitoring data. *Atmospheric Environment* **45** 6593–6606.
- STEIN, M. L. (2007). Spatial variation of total column ozone on a global scale. *Ann. Appl. Stat.* **1** 191–210. [MR2393847](#)
- STEIN, M. L. (2008). A modeling approach for large spatial datasets. *J. Korean Statist. Soc.* **37** 3–10. [MR2420389](#)

- STROUD, J. R., MÜLLER, P. and SANSÓ, B. (2001). Dynamic models for spatiotemporal data. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **63** 673–689. [MR1872059](#)
- SZPIRO, A. A. and PACIOREK, C. J. (2013). Measurement error in two-stage analyses, with application to air pollution epidemiology. *Environmetrics* **24** 501–517. [MR3161971](#)
- SZPIRO, A. A., SHEPPARD, L. and LUMLEY, T. (2011). Efficient measurement error correction with spatially misaligned data. *Biostatistics* **12** 610–623.
- SZPIRO, A. A., SAMPSON, P. D., SHEPPARD, L., LUMLEY, T., ADAR, S. D. and KAUFMAN, J. D. (2010). Predicting intra-urban variation in air pollution concentrations with complex spatio-temporal dependencies. *Environmetrics* **21** 606–631. [MR2842271](#)
- TELEATLAS (2000). TeleAtlas Dynamap 2000. [CD_ROM], TeleAtlas, Lebanon, NH.
- WAHBA, G. (1981). Spline interpolation and smoothing on the sphere. *SIAM J. Sci. Statist. Comput.* **2** 5–16. [MR0618629](#)
- WOOD, S. N. (2003). Thin plate regression splines. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65** 95–114. [MR1959095](#)
- ZHANG, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based geostatistics. *J. Amer. Statist. Assoc.* **99** 250–261. [MR2054303](#)

NONPARAMETRIC INFERENCE IN A STEREOLOGICAL MODEL WITH ORIENTED CYLINDERS APPLIED TO DUAL PHASE STEEL¹

BY K. S. MCGARRITY^{*,†}, J. SIETSMA[†] AND G. JONGBLOED[†]

Materials innovation institute (M2i)^{} and Delft University of Technology[†]*

Oriented circular cylinders in an opaque medium are used to represent certain microstructural objects in steel. The opaque medium is sliced parallel to the cylinder axes of symmetry and the cut-plane contains the observable rectangular profiles of the cylinders. A one-to-one relation between the joint density of the squared radius and height of the 3D cylinders and the joint density of the squared half-width and height of the observable 2D rectangles is established. We propose a nonparametric estimation procedure to estimate the distributions and expectations of various quantities of interest, such as the cylinder radius, height, aspect ratio, surface area and volume from the observed 2D rectangle widths and heights. Also, the covariance between the radius and height of a cylinder is estimated. The asymptotic behavior of these estimators is established to yield point-wise confidence intervals for the expectations and point-wise confidence sets for the distributions of the quantities of interest. Many of these quantities can be linked to the mechanical properties of the material, and are, therefore, useful for industry. We illustrate the mathematical model and estimation procedures using a banded microstructure for which nearly 90 μm of depth have been observed via serial sectioning.

REFERENCES

- ANDERSEN, S. J., HOLME, B. and MARIOARA, C. D. (2008). Quantification of small, convex particles by TEM. *Ultramicroscopy* **108** 750–762.
- ANEVSKI, D. and SOULIER, P. (2011). Monotone spectral density estimation. *Ann. Statist.* **39** 418–438. [MR2797852](#)
- BOLTE, S. and CORDELIÈRES, F. P. (2006). A guided tour into subcellular colocalization analysis in light microscopy. *J. Microsc.* **224** 213–232. [MR2312809](#)
- CHOW, Y. S. and TEICHER, H. (1988). *Probability Theory: Independence, Interchangeability, Martingales*, 2nd ed. Springer, New York. [MR0953964](#)
- CRUZ-ORIVE, L. M. and WEIBEL, E. R. (1990). Recent stereological methods for cell biology: A brief survey. *American Journal of Physiology* **258** L148–L156.
- CRUZ-ORIVE, L.-M., HOPPELER, H., MATHIEU, O. and WEIBEL, E. R. (1985). Stereological analysis of anisotropic structures using directional statistics. *J. Roy. Statist. Soc. Ser. C* **34** 14–32. [MR0793336](#)
- FULLMAN, R. L. (1953). Measurement of particle sizes in opaque bodies. *Transactions AIME (Journal of Metals)* **197** 447–452.

Key words and phrases. Banded microstructures, grain size distribution, isotonic estimation, stochastic modeling, Wicksell's problem.

- GIUMELLI, A. K., MILITZER, M. and HAWBOLT, E. B. (1999). Analysis of the austenite grain size distribution in plain carbon steels. *ISIJ International* **39** 271–280.
- GROENEBOOM, P. and JONGBLOED, G. (1995). Isotonic estimation and rates of convergence in Wicksell's problem. *Ann. Statist.* **23** 1518–1542. [MR1370294](#)
- GROENEBOOM, P. and JONGBLOED, G. (2010). Generalized continuous isotonic regression. *Statist. Probab. Lett.* **80** 248–253. [MR2575453](#)
- HALL, P. and SMITH, R. L. (1988). The kernel method for unfolding sphere size distributions. *J. Comput. Phys.* **74** 409–421. [MR0930210](#)
- HIGGINS, M. D. (2000). Measurement of crystal size distributions. *American Mineralogist* **85** 1105–1116.
- JENSEN, D. G. (1995). Estimation of the size distribution of spherical, disc-like or ellipsoidal particles in thin foil. *Journal of Physics D: Applied Physics* **28** 549–558.
- JEPPSSON, J., MANNESSON, K., BORGSTAM, A. and ÅGREN, J. (2011). Inverse Saltykov analysis for particle-size distributions and their time evolution. *Acta Materialia* **59** 874–882.
- LI, M., GHOSH, S., RICHMOND, O., WEILAND, H. and ROUNS, T. N. (1999). Three dimensional characterization and modeling of particle reinforced metal matrix composites: Part I quantitative description of microstructural morphology. *Materials Science and Engineering A* **A265** 153–173.
- MASE, S. (1995). Stereological estimation of particle size distributions. *Adv. in Appl. Probab.* **27** 350–366. [MR1334818](#)
- MAYHEW, T. M. (1991). The new stereological methods for interpreting functional morphology from slices of cells and organs. *Exp. Physiol.* **76** 639–665.
- MCGARRITY, K. S. (2013). Stochastic modeling of anisotropic microstructural features in metals. Ph.D. thesis, Technical Univ. Delft.
- MCGARRITY, K. S., SIETSMA, J. and JONGBLOED, G. (2012a). Characterisation and quantification of microstructural banding in dual phase steels part I: A general 2D study. *Materials Science and Technology* **28** 895–902. DOI:[10.1179/1743284712Y.0000000004](https://doi.org/10.1179/1743284712Y.0000000004).
- MCGARRITY, K. S., SIETSMA, J. and JONGBLOED, G. (2012b). Characterization and quantification of microstructural banding in dual phase steels part II: A case study extending to 3D. *Materials Science and Technology* **28** 903–910. DOI:[10.1179/1743284712Y.0000000028](https://doi.org/10.1179/1743284712Y.0000000028).
- MCGARRITY, K. S., SIETSMA, J. and JONGBLOED, G. (2014). Supplement to “Nonparametric inference in a stereological model with oriented cylinders applied to dual phase steel.” DOI:[10.1214/14-AOAS787SUPP](https://doi.org/10.1214/14-AOAS787SUPP).
- MEHNERT, K., OHSER, J. and KLIMANEK, P. (1998). Testing stereological methods for the estimation of spatial size distributions by means of computer-simulated grain structures. *Materials Science and Engineering A* **246** 207–212.
- MIYAMOTO, K. (1994). Particle number and sizes estimated from sections—A history of stereology. In *Research of Pattern Formation* 507–516. KTK Scientific Publishers, Tokyo.
- OAKESHOTT, R. B. S. and EDWARDS, S. F. (1992). On the stereology of ellipsoids and cylinders. *Phys. A* **189** 208–233. [MR1187504](#)
- OHSER, J. and MÜCKLICH, F. (2000). *Statistical Analysis of Microstructures in Materials Science*. Wiley, New York.
- RASBAND, W. S. (1997–2009). ImageJ. Available at <http://rsb.info.nih.gov/ij/>. U.S. National Institutes of Health, Bethesda, MD.
- RUSS, J. C. and DEHOFF, R. T. (2000). *Practical Stereology*, 2nd ed. Plenum Press, New York, NY.
- SAHAGIAN, DORIS L. and PROUSSEVITCH, A. A. (1998). 3D particle size distributions from 2D observations: Stereology for natural applications. *Journal of Volcanology and Geothermal Research* **84** 173–196.
- SEN, B. and WOODROOFE, M. (2012). Bootstrap confidence intervals for isotonic estimators in a stereological problem. *Bernoulli* **18** 1249–1266. [MR2995794](#)

- SILVERMAN, B. W., JONES, M. C., NYCHKA, D. W. and WILSON, J. D. (1990). A smoothed EM approach to indirect estimation problems, with particular reference to stereology and emission tomography. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **52** 271–324. [MR1064419](#)
- SPIESS, M. and SPODAREV, E. (2011). Anisotropic Poisson processes of cylinders. *Methodol. Comput. Appl. Probab.* **13** 801–819. [MR2851838](#)
- TEWARI, A. and GOKHALE, A. M. (2001). Estimation of three-dimensional grain size distribution from microstructural serial sections. *Mater. Charact.* **46** 329–335.
- THOULESS, M. D., DALGLEISH, B. J. and EVANS, A. G. (1988). Determining the shape of cylindrical second phases by two-dimensional sectioning. *Materials Science and Engineering A* **102** 57–68.
- VAN ES, B. and HOOGENDOORN, A. (1990). Kernel estimation in Wicksell's corpuscle problem. *Biometrika* **77** 139–145. [MR1049415](#)
- WICKSELL, S. D. (1925). A mathematical study of a biometric problem. *Biometrika* **17** 84–99.

ACKNOWLEDGMENT OF PRIORITY USAGE OF THE LAMBERT W FUNCTION IN STATISTICS

BY GEORG M. GOERG

Google, Inc.

REFERENCES

- GOERG, G. M. (2011). Lambert W random variables—a new family of generalized skewed distributions with applications to risk estimation. *Ann. Appl. Stat.* **5** 2197–2230. [MR2884937](#)
- STEHLÍK, M. (2003). Distributions of exact tests in the exponential family. *Metrika* **57** 145–164. [MR1969249](#)
- STEHLÍK, M. (2006). Exact likelihood ratio scale and homogeneity testing of some loss processes. *Statist. Probab. Lett.* **76** 19–26. [MR2213239](#)
- STEHLÍK, M. (2014). Exact testing of the scale parameter with the missing time-to-failure information. *Commun. Dependability Qual. Manag.* **10** 124–129.
- STEHLÍK, M., POTOCKÝ, R., WALDL, H. and FABIÁN, Z. (2010). On the favorable estimation for fitting heavy tailed data. *Comput. Statist.* **25** 485–503. [MR2720398](#)