Statistical Modeling with Spline Functions
Methodology and Theory

Mark H. Hansen
University of California at Los Angeles

Jianhua Z. Huang
University of Pennsylvania

Charles Kooperberg
Fred Hutchinson Cancer Research Center

Charles J. Stone
University of California at Berkeley

Young K. Truong
University of North Carolina at Chapel Hill

January 5, 2006

Bibliography

Abrahamowicz, M., A. Ciampi, and J. O. Ramsay (1992). Nonparametric density estimation for censored survival data: regression-spline approach. *The Canadian Journal of Statistics 20*, 171–185.

Abrahamowicz, M., T. MacKenzie, and J. M. Esdaile (1996). Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *Journal of the American Statistical Association 19*, 1432–1439.

Abramowitz, M. and I. A. Stegun (1965). *Handbook of Mathematical Functions*. New York: Dover.

Andersen, P. K., O. Borgan, R. D. Gill, and N. Keiding (1993). *Statistical Models Based on Counting Processes*. New York: Springer-Verlag.

Anderson, J. A. and V. Blair (1982). Penalized maximum likelihood estimation in logistic regression and discrimination. *Biometrika 69*, 123–136.

Anderson, J. A. and A. Senthilselvan (1990). Smooth estimates for the hazard function. *Journal of the Royal Statistical Society B 42*, 322–327.

Barron, A., M. J. Shervish, and L. Wasserman (1999). The consistency of posterior distributions in nonparametric problem. *The Annals of Statistics 27*, 536–561.

Beamish, N. and M. B. Priestley (1981). A study of autoregressive and window spectral estimation. *Applied Statistics 30*, 41–58.

Besag, J. and D. Higdon (1999). Bayesian inference for agricultural field experiments (with discussion). *Journal of the Royal Statistical Society B 61*, 691–746.

Bloxom, B. (1985). A constrained spline estimator of a hazard function. *Psychometrika 50*, 301–321.

Boneva, L. I., D. Kendall, and I. Stefanov (1971). Spline transformations: three new diagnostic aids for the statistical data-analyst (with discussion). *Journal of the Royal Statistical Society B 33*, 1–70.

Bose, S. (1996). Classification using splines. *Computational Statistics and Data Analysis 22*, 505–525.

Bourlard, H. A. and N. Morgan (1994). *Connectionist Speech Recognition*. Boston: Kluwer.

Bowie, J. U., R. Luthy, and D. Eisenberg (1991). A method to identify protein sequences that fold into a known 3-dimensional structure. *Science 253*, 164–170.

Breiman, L. (1991). The π-method for estimating multivariate functions from noisy data. *Technometrics 33*, 125–143.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone (1984). *Classification and Regression Trees*. Pacific Grove, California: Wadsworth.

Breslow, N. E. (1972). Contribution to the discussion on the paper by D. R. Cox, regression and life tables. *Journal of the Royal Statistical Society B 34*, 216–217.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory*. San Francisco: Holden Day.

Broyden, C. G. (1967). Quasi-newton methods and their application to function minimization. *Mathematics of Optimization 21*, 368–381.

Broyden, C. G. (1970). The convergence of a class of double-rank minimization algorithms. Parts I and II. *Journal of Institute of Mathematical and Its Applicaitons 6*, 76–90, 222–236.

Cheng, B. and D. M. Titterington (1994). Neural networks: a review from a statistical perspective (with discussion). *Statistical Science 9*, 2–54.

Chui, C. K. (1988). *Multivariate Splines*. Philadelphia: SIAM.

Chui, C. K. and T. X. He (1990). Bivariate C^1 quadratic finite elements and vertex splines. *Mathematics of Computation 54*, 169–187.

Chui, C. K. and M. J. Lai (1990). Multivariate vertex splines and finite elements. *Journal of Approximation Theory 60*, 245–343.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American Statistical Association 74*, 829–636.

Cleveland, W. S. and M. Fuentes (1996). Multipanel conditioning: modeling data from designed experiments, Technical Memorandum 96-1. Technical report, AT&T Bell Laboratories, Murray Hill, NJ.

Clyde, M. (2000). Model uncertainty and health effect studies for particulate matter. *Environmetrics 11*, 745–763.

Cogburn, R. and H. T. Davis (1974). Periodic splines and spectral estimation. *The Annals of Statistics 2*, 1108–1126.

Courant, R. (1943). Variational methods for the solution of problems of equilibrium and vibrations. *Bulletin of the American Mathematical Society 49*, 1–23.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society B 34*, 187–220.

Dahmen, W. (1980). On multivariate B-splines. *SIAM Journal of Numerical Analysis 17*, 179–191.

Davidon, W. C. (1959). Variable metric methods for minimization. Technical report, AEC Research and Development Report ANL 5990 (Rev.).

Davis, R. B. and J. R. Anderson (1989). Exponential survival trees. *Statistics in Medicine 8*, 947–961.

de Boor, C. (1976). A bound on the L_∞-norm of L_2-approximation by splines in terms of a global mesh ratio. *Mathematics of Computation 30*, 765–771.

de Boor, C. (1978). *A Practical Guide to Splines*. New York: Springer-Verlag.

de Boor, C. (1987). B-form basics. In G. E. Farin (Ed.), *Geometric modeling: algorithms and new trends*, Philadelphia, pp. 131–148. SIAM.

de Boor, C. and K. Höllig (1982). B-splines from parallelepipeds. *Journal d'Analyse Mathematique 42*, 99–115.

de Boor, C. and K. Höllig (1988). Approximation power of smooth bivariate pp functions. *Mathematische Zeitschrift 197*, 343–363.

de Boor, C., K. Höllig, and S. Riemenschneider (1993). *Box Splines*. New York: Springer-Verlag.

Denison, D. G. T., B. K. Mallick, and A. F. M. Smith (1998). Automatic Bayesian curve fitting. *Journal of the Royal Statistical Society B 60*, 333–350.

Dennis, Jr., J. E. and R. E. Schnabel (1983). *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Englewood Cliffs, New Jersey: Prentice Hall.

DeVore, R. A. and G. G. Lorentz (1993). *Constructive Approximation*. Berlin: Springer-Verlag.

Dixon, L. C. W. (1975). Conjugate gradient algorithms: quadratic termination without linear searches. *Journal of the Institute of Mathematics and its Applications 15*, 9–18.

Donoho, D. L. and I. M. Johnstone (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika 81*, 425–455.

Dyn, N., D. Levin, and S. Rippa (1990a). Algorithms for the construction of data dependent triangulations. In *Algorithms for Approximation*, Volume II, New York, pp. 185–192. Chapman and Hall.

Dyn, N., D. Levin, and S. Rippa (1990b). Data dependent triangulations for piecewise linear interpolation. *Journal of Numierical Analysis 10*, 137==154.

Efron, B. and R. J. Tibshirani (1993). *An Introduction to the Bootstrap.* London: Chapman and Hall.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties (with discussion). *Statistical Science 11*, 89–121.

Etezadi-Amoli, J. and A. Ciampi (1987). Extended hazard regression for censored survival data with covariates: a spline approximation for the baseline hazard function. *Biometrics 43*, 181–192.

Fahrmeier, L. and A. Klinger (1998). A nonparametric multiplicative hazard model for event history analysis. *Biometrika 85*, 581–592.

Fahrmeier, L. and S. Wagenpfeil (1996). Smoothing hazard functions and time-varying effects in discrete duration and competing risks models. *Journal of the American Statistical Association 91*, 1584–1594.

Family Expenditure Survey (1968–1983). London: Department of Employment, Statistics Division – Her Majesty's Stationary Office.

Fan, J. and I. Gijbels (1996). *Local Polynomial Modeling and its Applications.* London: Chapman and Hall.

Farin, G. (1986). Triangular Bernstein–Bézier patches. *CAGD 3*, 83–127.

Fleming, T. R. and D. P. Harrington (1991). *Counting Processes and Survival Analysis.* New York: John Wiley.

Fletcher, R. (1970). A class of methods for nonlinear programming with termination and convergence properties. In J. Abadie (Ed.), *Integer and Nonlinear Programming*, Amsterdam. North Holland.

Fletcher, R. and M. J. D. Powell (1963). A rapidly convergent descent method for minimization. *Computer Journal 8*, 163–168.

Foster and George (1996). Unknown. *UNKNOWN UNKNOWN*, UNKNOWN.

Friedman, J. H. (1991). Multivariate adaptive regression splines (with discussion). *The Annals of Statistics 19*, 1–141.

Furnival, G. M. and J. Wilson, Robert W. (1974). Regression by leaps and bounds. *Technometrics 16*, 499–511.

Gao, H.-Y. (1997). Choice of thresholds for wavelet shrinkage estimate of the spectrum. *Journal of Time Series Analysis 18*, 231–251.

Gay, D. M. (1983). Algorithm 611. Subroutines for unconstrained minimization using a model/trust-region approach. *ACM Transactions on Mathematical Software 9*, 503–524.

George, E. and R. E. McCulloch (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association 88*, 881–889.

George, E. I. and D. P. Foster (2000). Calibration and empirical Bayes variable selection. *Biometrika 87*, 731–749.

Gilbet, J. C. and J. Nocedal (1992). Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on Optimization 2*, 21–42.

Goldfarb, D. (1970). A family of variable metric methods derived by variational means. *Mathematics of Optimization 24*, 23–26.

Golub, G. H. and C. F. V. Loan (1989). *Matrix Computations* (2nd ed.). Baltimore: Johns Hopkins University Press.

Good, I. J. and R. A. Gaskins (1971). Nonparametric roughness penalties for probability densities. *Biometrika 58*, 255–277.

Gordon, L. and R. A. Olshen (1985). Tree-structured survival analysis. *Cancer Treatment Reports 69*, 1065–1069.

Gray, R. J. (1992). Flexible methods for analyzing survival data using splines, with applications to breast cancer prognosis. *Journal of the American Statistical Association 87*, 942–951.

Gray, R. J. (1994). Spline based tests in survival analysis. *Biometrics 50*, 640–652.

Green, A., M. Hauge, N. V. Holm, and L. L. Rasch (1981). Epidemiological studies of diabetes mellitus in Denmark. II. a prevalence study based on insulin prescriptions. *Diabetologia 28*, 339–342.

Green, P. J. (1995). Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika 82*, 711–732.

Green, P. J. and B. W. Silverman (1994). *Nonparametric Regression and Generalized Linear Models: a Roughness Penalty Approach.* London: Chapman and Hall.

Gu, C. (1993). Smoothing spline density estimation: a dimensionless automatic algorithm. *Journal of the American Statistical Association 88*, 495–504.

Gu, C. (1994). Penalized likelihood hazard estimation: algorithm and examples. In *Statistical Decision Theory and Related Topics V*, pp. 61–72.

Gu, C. (1996). Penalized likelihood hazard estimation: a general procedure. *Statistica Sinica 6*, 861–867.

Gu, C., D. M. Bates, Z. Chen, and G. Wahba (1990). The computation of gcv function through householder tridiagonlization with application to the fitting of interaction spline models. *Siam Journal of Matrix Analysis 10*, 457–480.

Halpern, E. (1973). Bayesian spline regression when the number of knots is unknown. *Journal of the Royal Statistical Society B 35*, 347–360.

Hansen, M. H. (1994). *Extended Linear Models, Multivariate Splines, and ANOVA*. Ph. D. thesis, University of California at Berkeley.

Hansen, M. H., C. Kooperberg, and S. Sardy (1998). Triogram models. *Journal of the American Statistical Association 93*, 101–119.

Hastie, T. J. and R. J. Tibshirani (1990). *Generalized Additive Models*. London: Chapman and Hall.

Hastie, T. J. and R. J. Tibshirani (1993). Varying-coefficient models (with discussion). *Journal of the Royal Statistical Society B 55*, 757–800.

Hastie, T. J., R. J. Tibshirani, and A. Buja (1994). Flexible discriminant analysis by optimal scoring. *Journal of the American Statistical Association 89*, 1255–1270.

Heinzl, H., A. Kaider, and G. Zlabinger (1996). Assessing interactions of binary time-dependent covariates with time in Cox proportional hazards regression models using cubic spline functions. *Statistics in Medicine 15*, 2589–2601.

Herndon, II, J. E. and F. E. Harrell, Jr. (1990). The restricted cubic spline hazard model. *Communications in Statistics, Theory and Methods 19*, 639–694.

Herndon, II, J. E. and F. E. Harrell, Jr. (1995). The restricted cubic spline as baseline hazard in the proportional hazards model with step function time–dependent covariables. *Statistics in Medicine 14*, 2119–2129.

Hess, K. R. (1994). Assessing time-by-covariate interactions in proportional hazards regression models using cubic spline functions. *Statistics in Medicine 13*, 1045–1062.

Hobohm, U., M. Scharf, R. Schneider, and C. Sander (1992). Selection of representative set of structures from the brookhaven protein data bank. *Protein Science 1*, 409–417.

Holmes, C. and M. B. K. (1998). Bayesian radial basis functions of variable dimension. *Neural Computation 10*, 1217–1235.

Hornung, R. W., J. A. Deddens, and R. J. Roscoe (1998). Modifiers of lung cancer risk in uranium miners from the Colorado Plateau. *Health Physics 74*, 12–21.

Hornung, R. W. and T. J. Meinhardt (1987). Quantitative risk assessment of lung-cancer in United-States uranium miners. *Health Physics 52*, 417–430.

Hosmer, D. and S. Lemeshow (1989). *Applied Logistic Regression*. New York: John Wiley.

Huang, J. (2001). Concave extended linear modeling: a theoretical synthesis. *Statistica Sinica 11*, 173–197.

Huang, J. Z., C. Kooperberg, C. J. Stone, and Y. K. Truong (2000). Functional ANOVA modeling for proportional hazards regression. *The Annals of Statistics 28*, 960–999.

Huang, J. Z. and C. J. Stone (1998). The L_2 rate of convergence for event history regression with time-dependent covariates. *Scandinavian Journal of Statistics 25*, 603–620.

Intrator, O. and C. Kooperberg (1995). Trees and splines in survival analysis. *Statistical Methods in Medical Research 4*, 237–261.

Izenman, A. J. and C. J. Sommer (1988). Philatelic mixtures and multimodal densities. *Journal of the American Statistical Association 83*, 941–953.

Jarjoura, D. (1988). Smoothing hazard rates with cubic splines. *Communications in Statistics, Simulation and Computation 17*, 377–392.

Jupp, D. L. P. (1978). Approximation to data by splines with free knots. *SIAM Journal of Numerical Analysis 15*, 328–343.

Kalbfleisch, J. D. and R. L. Prentice (1980). *The Statistical Analysis of Failure Time Data*. New York: John Wiley.

Kaplan, E. L. and P. Meier (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association 53*, 457–481.

Karlin, S., C. Micchelli, and Y. Rinott (1986). Multivariate splines: A probabilistic perspective. *Journal of Multivariate Analysis 20*, 69–90.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*, 773–795.

Kelly, F. P. and B. D. Ripley (1976). Note on Strauss's model for clustering. *Biometrika 63*, 357–360.

Kennedy, Jr., W. J. and J. E. Gentle (1980). *Statistical Computing*. New York: Marcel Dekker.

Kincaid, D. and W. Cheney (1996). *Numerical Analysis: Mathematics of Scientific Computing, 2nd edition.* Pacific Grove, CA: Brooks/Cole.

Klotz, J. (1982). Spline smooth estimates of survival. In *Survival Analysis.* Hayward, California: Institute of Mathematical Statistics.

Koo, J.-Y. (1996). B-splines for tensor logspline density estimation. *Computational Statistics and Data Analysis 21,* 31–42.

Koo, J.-Y., C. Kooperberg, and J. Park (1999). Logspline density estimation under censoring and truncation. *Scandinavian Journal of Statistics 26,* 87–105.

Kooperberg, C. (1998). Bivariate density estimation with an application to survival data. *Journal of Computational and Graphical Statistics 7,* 322–341.

Kooperberg, C., S. Bose, and C. J. Stone (1997). Polychotomous regression. *Journal of the American Statistical Association 92,* 117–127.

Kooperberg, C. and D. B. Clarkson (1997). Hazard regression with interval-censored data. *Biometrics 53,* 1485–1494.

Kooperberg, C. and C. J. Stone (1991). Logspline density estimation. *Computational Statistics and Data Analysis 12,* 327–347.

Kooperberg, C. and C. J. Stone (1992). Logspline density estimation for censored data. *Journal of Computational and Graphical Statistics 1,* 301–328.

Kooperberg, C. and C. J. Stone (1999). Using the stochastic gradient method to fit polychotomous regression models. *Journal of Computational and Graphical Statistics 9,* 169–189.

Kooperberg, C. and C. J. Stone (2001a). Confidence intervals for logspline density estimation. Technical report, Fred Hutchinson Cancer Research Center, Seattle, Washington.

Kooperberg, C. and C. J. Stone (2001b). Logspline density estimation with free knots. Technical report, Fred Hutchinson Cancer Research Center, Seattle, Washington.

Kooperberg, C., C. J. Stone, and Y. K. Truong (1995a). Hazard regression. *Journal of the American Statistical Association 90,* 78–94.

Kooperberg, C., C. J. Stone, and Y. K. Truong (1995b). The L_2 rate of convergence for hazard regression. *Scandinavian Journal of Statistics 22,* 143–157.

Kooperberg, C., C. J. Stone, and Y. K. Truong (1995c). Logspline estimation of a possibly mixed spectral distribution. *Journal of Time Series Analysis 16,* 359–388.

Kooperberg, C., C. J. Stone, and Y. K. Truong (1995d). Rate of convergence for logspline spectral density estimation. *Journal of Time Series Analysis 16,* 389–401.

LeBlanc, M. and J. Crowley (1992). Relative risk trees for censored data. *Biometrics 48*, 411–425.

LeBlanc, M. and J. Crowley (1999). Adaptive regression splines in the Cox model. *Biometrics 55*, 204–213.

Lindstrom, M. J. (1999). Penalized estimation of free-knot splines. *Journal of Computational and Graphical Statistics 8*, 333–352.

Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control 22*, 551–575.

Loader, C. (1999). *Local Regression and Likelihood*. New York: Springer-Verlag.

Luebeck, E. G., W. F. Heidenreich, W. D. Hazelton, H. G. Paretzke, and S. H. Moolgavkar (1999). Biological-based analysis of the Colorado uranium miners cohort data: age, dose and dose-rate effects. Technical report, Fred Hutchinson Cancer Research Center, Seattle, Washington.

Mackisack, M. S. and D. S. Poskitt (1990). Some properties of autoregressive estimates for processes with mixed spectra. *Journal of Time Series Analysis 11*, 325–37.

Madigan, D. and A. Raftery (1994). Model selection and accounting for model uncertainty in graphical models using occam's window. *Journal of the American Statistical Association 89*, 1335–1346.

Mardia, K. V., J. T. Kent, and J. M. Bibby (1979). *Multivariate Analysis*. London: Academic Press.

Marron, J. S. and P. Chaudhuri (1999). Sizer for exploration of structures in curves. *Journal of the American Statistical Association 94*, 807–823.

Miller, A. J. (1990). *Subset Selection in Regression*. London: Chapman and Hall.

Minnotte, M. C. (1998). Achieving higher-order convergence rates for density estimates with binned data. *Journal of the American Statistical Association 93*, 663–672.

Minnotte, M. C. and D. W. Scott (1993). The mode tree: a tool for visualization of nonparametric density features. *Journal of Computational and Graphical Statistics 2*, 51–68.

Møller, M. (1993). Supervised learning on large redundant training sets. *International Journal of Neural Systems 4*, 15–25.

Morandi, R. and P. Costantini (1989). Piecewise monotone quadratic histosplines. *SIAM Journal on Scientific and Statistical Computing 10*, 397–406.

Morgan, J. and R. Scott (1975). A nodal basis for C^1 piecewise polynomials of degree $n \geq 5$. *Mathematics of Computation 29*, 736–740.

Moulin, P. (1994). Wavelet thresholding techniques for power spectrum estimation. *IEEE Transactions on Signal Processing 42*, 3126–3136.

Nicholls, G. K. (1998). Bayesian image analysis with Markov chain Monte Carlo and colored continuum triangulation models. *Journal of the Royal Statistical Society B 60*, 643–659.

Nocedal, J. and S. J. Wright (1999). *Numerical Optimization.* New York: Springer-Verlag.

O'Hagan, A. (1994). *Kendall's Advanced Theory of Statistics. Volume 2B: Bayesian Inference.* Cambridge: Cambridge University Press.

O'Sullivan, F. (1988a). Fast computation of fully automated log-density and log-hazard estimators. *SIAM Journal on Scientific and Statistical Computing 9*, 363–379.

O'Sullivan, F. (1988b). Nonparametric estimation of relative risk using splines and cross-validation. *SIAM Journal on Scientific and Statistical Computing 9*, 531–542.

Pawitan, Y. and F. O'Sullivan (1994). Nonparametric spectral density estimation using penalized whittle likelihood. *Journal of the American Statistical Association 89*, 600–610.

Polak, E. (1997). *Optimization.* New York: Springer-Verlag.

Polak, E. and G. Ribiere (1969). Note sur la convergence de methodes de directions conjugees. *Revue Francaise d'Informatique et de Recherche Operationnelle 3*, 35–43.

Polak, E. and G. Ribière (1969). Note sur la convergence de methodes de directions conjugueés. *Revue Francaise d'Informatique et de Recherche Operationelle, Serie Rouge 16*, 35–43.

Priestley, M. B. (1962a). The analysis of stationary processes with mixed spectra—I. *Journal of the Royal Statistical Society B 24*, 214–233.

Priestley, M. B. (1962b). The analysis of stationary processes with mixed spectra—II. *Journal of the Royal Statistical Society B 24*, 511–529.

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* London: Academic Press.

Ripley, B. D. (1996). *Pattern recognition and neural networks.* Cambridge UK: Cambridge University Press.

Rippa, S. (1992). Long and thin triangles can be good for linear interpolation. *SIAM Journal on Numerical Analysis 29*, 257–270.

Rivlin, T. J. (1969). *Introduction to the Approximation of Functions.* Waltham, MA: Blaisdell.

Robbins, H. and S. Monro (1951). A stochastic approximation method. *The Annals of Mathematical Statistics 22*, 400–407.

Robinson, A. (1989). *Dynamic Error Propagation Networks*. Ph. D. thesis, Cambridge University.

Rosenberg, P. S. (1995). Hazard function estimation using B-splines. *Biometrics 51*, 874–887.

Schoenberg, I. J. (1946). Contributions to the problem of approximation of equidistant data by analytic functions. *Quarterly of Applied Mathematics 4*, 45–99.

Schumaker, L. L. (1981). *Spline Functions: Basic Theory*. New York: John Wiley.

Schwartz (1981). Unknown. *UNKNOWN UNKNOWN*, UNKNOWN.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics 6*, 461–464.

Schwarz, H. R. (1988). *Finite Element Methods*. New York: Academic Press.

Segal, M. R. (1988). Regression trees for censored data. *Biometrics 44*, 35–47.

Senthilselvan, A. (1987). Penalized likelihood estimation of hazard and intensity functions. *Journal of the Royal Statistical Society B 49*, 170–174.

Shanno, D. F. (1970). Conditioning of quasi-newton methods for function minimization. *Mathematics of Optimization 24*, 847–856.

Silverman, B. W. (1985). Some aspects of the spline smoothing approach to non-parametric regression curve fitting (with discussion). *Journal of the Royal Statistical Society B 47*, 1–50.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. London: Chapman and Hall.

Sleeper, L. A. and D. P. Harrington (1990). Regression splines in the Cox model with application to covariate effects in liver disease. *Journal of the American Statistical Association 85*, 941–949.

Smith, M. (1996). *Nonparametric Regression: A Markov chain Monte Carlo Approach*. Ph. D. thesis, University of New South Wales.

Smith, M. and R. Kohn (1996). Nonparametric regression using Bayesian variable selection. *Journal of Econometrics 75*, 317–344.

Stoica, P., R. L. Moses, T. Söderström, and J. Li (1991). Optimal high-order Yule–Walker estimation of sinusoidal frequencies. *IEEE Transactions on Signal Processing 39*, 1360–1368.

Stone, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *The Annals of Statistics 8*, 1348–1360.

Stone, C. J. (1985). Additive regression and other nonparametric models. *The Annals of Statistics 13*, 689–705.

Stone, C. J. (1986). The dimensionality reduction principle for generalized additive models. *The Annals of Statistics 14*, 590–606.

Stone, C. J. (1989). Uniform error bounds involving logspline models. In K. B. A. T. W. Anderson and D. L. Iglehart (Eds.), *Probability, Statistics and Mathematics: Papers in Honor of Samuel Karlin*, New York, pp. 335–355. Academic Press.

Stone, C. J. (1990). Large-sample inference for log-spline models. *The Annals of Statistics 18*, 717–741.

Stone, C. J. (1991). Asymptotics for doubly flexible logspline response models. *The Annals of Statistics 19*, 1832–1854.

Stone, C. J. (1994). The use of polynomial splines and their tensor products in multivariate function estimation (with discussion). *The Annals of Statistics 22*, 118–184.

Stone, C. J., M. H. Hansen, C. Kooperberg, and Y. K. Truong (1997). Polynomial splines and their tensor products in extended linear modeling (with discussion). *The Annals of Statistics 25*, 1371–1470.

Stone, C. J. and C.-Y. Koo (1986). Additive splines in statistics. In *Proceedings of the Statistical Computing Section*, pp. 45–48. American Statistical Association, Washington, DC.

Tapia, T. A. and J. R. Thompson (1978). *Nonparametric Probability Density Estimation*. Baltimore: Johns Hopkins University Press.

Tibshirani, R. J. and M. LeBlanc (1992). A strategy for binary description and classification. *Journal of Computational and Graphical Statistics 1*, 3–20.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York: Springer-Verlag.

Vapnik, V. (1998). *The Statistical Learning Theory*. New York: Springer-Verlag.

Villalobos, M. A. and G. Wahba (1983). Multivariate thin plate spline estimation for the posterior probabilities in the classification problem. *Communications in Statistics, Theory and Methods 12*, 1449–1479.

Wahba, G. (1976). Histosplines with knots which are order statistics. *Journal of the Royal Statistical Society B 38*, 140–151.

Wahba, G. (1980). Automatic smoothing of the log periodogram. *Journal of the American Statistical Association 75*, 122–132.

Wahba, G. (1990). *Spline Models for Observational Data*. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia: SIAM.

Walden, A. T., D. B. Percival, and E. J. McCoy (1998). Spectrum estimation by wavelet thresholding of multitaper estimators. *IEEE Transactions on Signal Processing 46*, 3153–3165.

Wand, M. P., S. J. Marron, and D. Ruppert (1991). Transformations in density estimation, (with discussion). *Journal of the American Statistical Association 86*, 343–361.

White, H. (1989). Some asymptotic results for learning in single hidden-layer feedforward network models. *Journal of the American Statistical Association 84*, 1003–1013.

Whittemore, A. S. and J. B. Keller (1986). Survival estimation using splines. *Biometrics 42*, 495–506.

Whittle, P. (1961). Gaussian estimation in stationary time series. *Bulletin of the International Statistical Institute 39*, 105–130.

Wilson, I. G. (1983). Add a new dimension to your philately. *The American Philatelist 97*, 342–349.

Wong, F., M. H. Hansen, R. Kohn, and M. Smith (1998). Focused sampling and its application to non-parametric and robust regression. Technical report, University of New South Wales.

X., S. and W. L. (1998). Unknown. *UNKNOWN UNKNOWN*, UNKNOWN.

Zeger, S. L. and M. R. Karim (1991). Generalized linear models with random effects: a Gibbs sampling approach. *Journal of the American Statistical Association 86*, 79–86.

Zellner, A. (1986). On assessing prior distributions and Bayesian regression analysis with g-prior distributions. In P. Goel and A. Zellner (Eds.), *In Bayesian Inference and Decision Techniques: Essays in Honor of Bruno de Finetti*, Amsterdam, pp. 233–243. North-Holland.

Zhang, K. and D. Eisenberg (1994). The three-dimensional profile method using residue preference as a continuous function of residue environment. *Protein Science 3*, 687–695.