Bayesian Analysis Users Guide
Release 4.00, Manual Version 1

G. Larry Bretthorst
Biomedical MR Laboratory
Washington University School Of Medicine,
Campus Box 8227
Room 2313, East Bldg.,
4525 Scott Ave.
St. Louis MO 63110
http://bayes.wustl.edu
Email: larry@bayes.wustl.edu

October 21, 2016

Contents

List of Figures

List of Tables

Chapter 29

The Image Model Selection Package

The Image Model Selection Package allows you to load one or more Ascii model and then use Bayesian probability theory to compute the posterior probability for the loaded models, thus allowing you to determine which model best accounts for the data.[1] To use this package you do not have to have Fortran or C installed on your server. However, if you do not have Fortran or C installed, you must use the system models. Consequently, installing both Fortran and C is strongly recommended. The interface to this package is shown in Fig. 29.1. To use this package, you must do the following:

Select the "Analyze Image Pixel" package from the Package menu.

Load one or more Fortran or C model using the "System" or "User" buttons in the "Load And Build Model" widget group.

Load an Arrayed image using Files/Load Image menu. viewer. used by the Bayesian Analysis software.

Load an Abscissa using Files/Load Abscissa menu.

Enter the Standard Deviation of the noise.

Check the Use Gaussian widget, if the number of array elements is small. Here small means if the number of parameters being estimated is within a factor of 2 of the number of array elements.

Select the server that is to process the analysis.

Check the status of the selected server to determine if the server is busy, change to another server if the selected server is busy.

Run the the analysis on the selected server by activating the Run button.

[1] I would like to build a system library of predefined models. If you have models that you think would be of general use, I would like to hear from you. To have one of your models included, I would need the source code, the parameter file, a brief description of the model equations and data requirements.

Figure 29.1: The Interface To The Image Model Selection Package

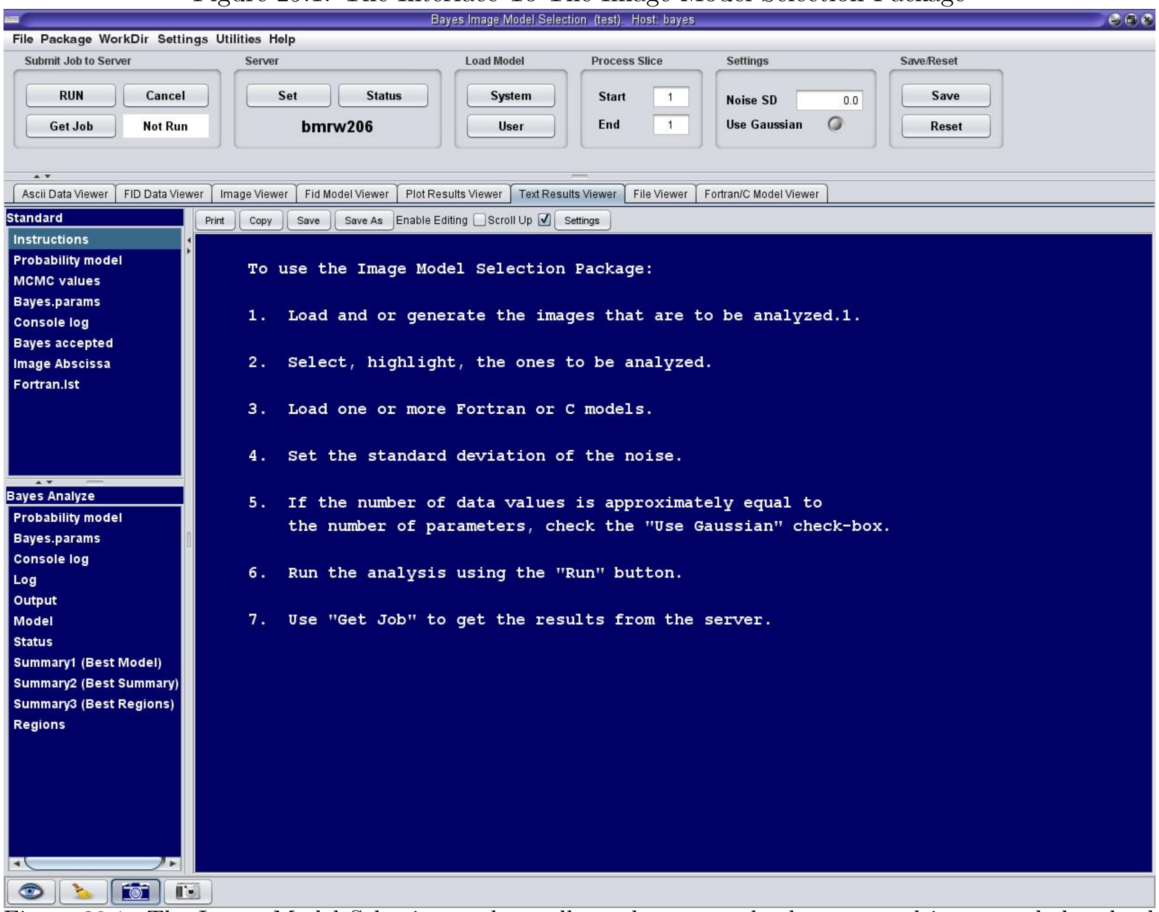

Figure 29.1: The Image Model Selection package allows the user to load an arrayed image and then load several models, either system defined or user defined, and analyze the input image on a pixel by pixel basis using the loaded models. The program that implements this package computes the posterior probability for the modes. Prior to running an analysis, you must also load an appropriate Abscissa file.

Get the the results of the analysis by activating the Get Job button. If the analysis is running, this button will return the Accepted report containing the status of the current run. Otherwise, it will fetch and display the results from the current analysis.

29.1 The Bayesian Calculations

The Bayesian calculation done in this package is identical to that done in Chapter 21 and we will not repeat that calculation here. However, we will describe how the Image Model Selection package differs from the Enter Ascii Model Selection package. First, the Image Model Selection package process each pixel in an image. That is to say the same calculation is performed on each pixel in the image. It is assumed that the image is arrayed and the models describe the signal in the arrayed pixels. The program does not process all pixels at one time using a common model; rather it process each pixel in parallel.

In the model selection calculation done by the Image Model Selection package, it assumes one has a set of models, $U_j \equiv \{U_1, \ldots, U_m\}$, and one wishes to compute the posterior probability for each of the U_j models in every pixel in the image. These models can be loaded from the System directory or they can be loaded from the user directory. They don't need to have common parameters, but they do need to have common data requirements because they all use the same data. Each of these models will have a set of parameters associated with it. Unlike the Analyze Image Pixel routine which output's parameter maps, the Image Model Selection routine cannot output a parameter map because the different models will in general have different parameterizations and consequently no map can be output.

To circumvent this problem, we require the Fortran/C models to have the same derived parameters so that an output parameter map can be generated by the user. How the user constructs these maps, is up to the individual who is writing the Fortran/C codes. For example the exponential models supplied in the system directories, define three output decay rate constants: a low, middle and high decay rate constant. The one exponential model sets the low, middle, and high derived decay rate constants to be equal. The two exponential model, sets the low decay rate to the lowest estimated rate constant, and the middle and high derived rate constant are set to the highest estimated rate constant. Finally, the three exponential model sets the low, middle and high as the three estimated low, middle and high rate constants. By setting the output derived images in this fashion, the low rate constant will always be the lowest decay rate. The middle decay rate will always be second largest decay rate and the high decay rate constant will always be the largest decay rate constant found.

The equation that relates model U_j to the data is given by:

$$d_{kl}(t_i) = U_j(t_i, \Omega_j) + n_{kl}(t_i) \tag{29.1}$$

where $d_{kl}(t_i)$ is the data at lth readout of the kth phase encode pixel sampled at abscissa values t_i. The t_i may be vector valued or it may be a single column of numbers. However, it must have the same number of columns in all loaded models. And it must be the same abscissa for all pixels. The noise is represented symbolically by $n_{kl}(t_i)$. The models U_j can be any system or user defined model and they can use marginalization or not depending on the model.

The program that implements this calculation loops over all of the pixels in the image and performs the model selection calculation described in Enter Ascii Model Selection, Chapter 21. Equation 21.60 is the joint posterior probability for the nonlinear parameters when the amplitudes are marginalized out of the posterior probability, and Eq. (21.18) is used when there are no parameters marginalized out of the posterior probability. It is these two equations that are targeted by the Markov chain Monte Carlo simulations used to implement the Image Pixel Model Selection Calculation.

29.2 Outputs Form The Image Model Selection Package

The test data set contained in the bayes.test.data/ImagePixelsAll directory is shown in Fig. 29.2. The test image shown in Fig. 29.2 is a $5 \times 5 \times 51$, having a dwell time of 0.1 seconds. So the first data value is a $t_1 = 0$ and the last data value is at $t = 5$ seconds. The decay rate constants are 1, 2, 3, 4, and 5 in each row. And the amplitudes are 0, 20, 30, 40 and 50 for each row. The noise has standard deviation of one. When this test data was run, 6 models were loaded: ExpOneNoConst_Marg.f, ExpOneConst_Marg.f, ExpTwoNoConst_Marg.f, ExpTwoConst_Marg, ExpThreeNoConst_Marg.f and ExpThreeConst_Marg.f. The no signal model, model zero, is always loaded and is used to designate which pixels do not have a signal. The posterior probability for the ExpOneNoConst.f model is shown Fig. 29.4. Note the black band on the left-hand side of this image. These are the no-signal pixels and the program correctly identified them as containing no signal and did not process them. The nonzero area contains the expected model number in the data. In this case the models are numbers are 0, 1, 2, 3, 4, and 5. Model number zero is the no-signal model, model number 1 is the ExpOneNoConst_Marg.f model, etc. I computed the mean and standard deviation in this region and its 0.986 ± 0.003 meaning that single exponential with no constant model was identified with a probability averaging 0.986.

The image model selection package outputs parameter maps of all of the derived parameters. So for example, when testing one, two and three exponential models you get a low, middle and high decay rate constant. Each parameter map comes as a set of three images: the mean parameter, the peak parameter, and the standard deviation of the parameter. The mean parameter map, is the mean or average value of the parameter in all of the Markov chain Monte Carlo simulations. The "peak" value of the parameter is the value of the parameter in the simulation that had maximum posterior probability. Finally, the standard deviation of the parameter is the standard deviation of the parameters computed from all of the samples gathered in the Markov chain Monte Carlo simulation. Additionally, you get maps of the amplitudes that occurs in the model. For example in the exponential model selection example we have been discussing you get both the amplitude and constant offset maps.

There are three additional outputs, one is the average logarithm of the likelihood. And can be used to spot outliers. The posterior probability for the model is output. In this case the expected mean value of the model number is output. For example, in the one, two and three exponential model example. I loaded the one, two and three exponential models in that order. So model number one was the one exponential model, two was the two exponential model, etc. The output model number map, is the average value of the model number for each pixel in the image. Model number zero, is the no signal model and implies that none of the models fit the data than the no signal model.

You also get parameter maps for each of the high probability models. For example if you load the "ExpOneNoConst.img" file and run the one, two and three exponential model selection, you will get parameter maps for the one exponential model, but not the two and three. That's because in this data set the two and three exponential models are never selected, and so there are no parameters to

Figure 29.2: Single Exponential Example Image

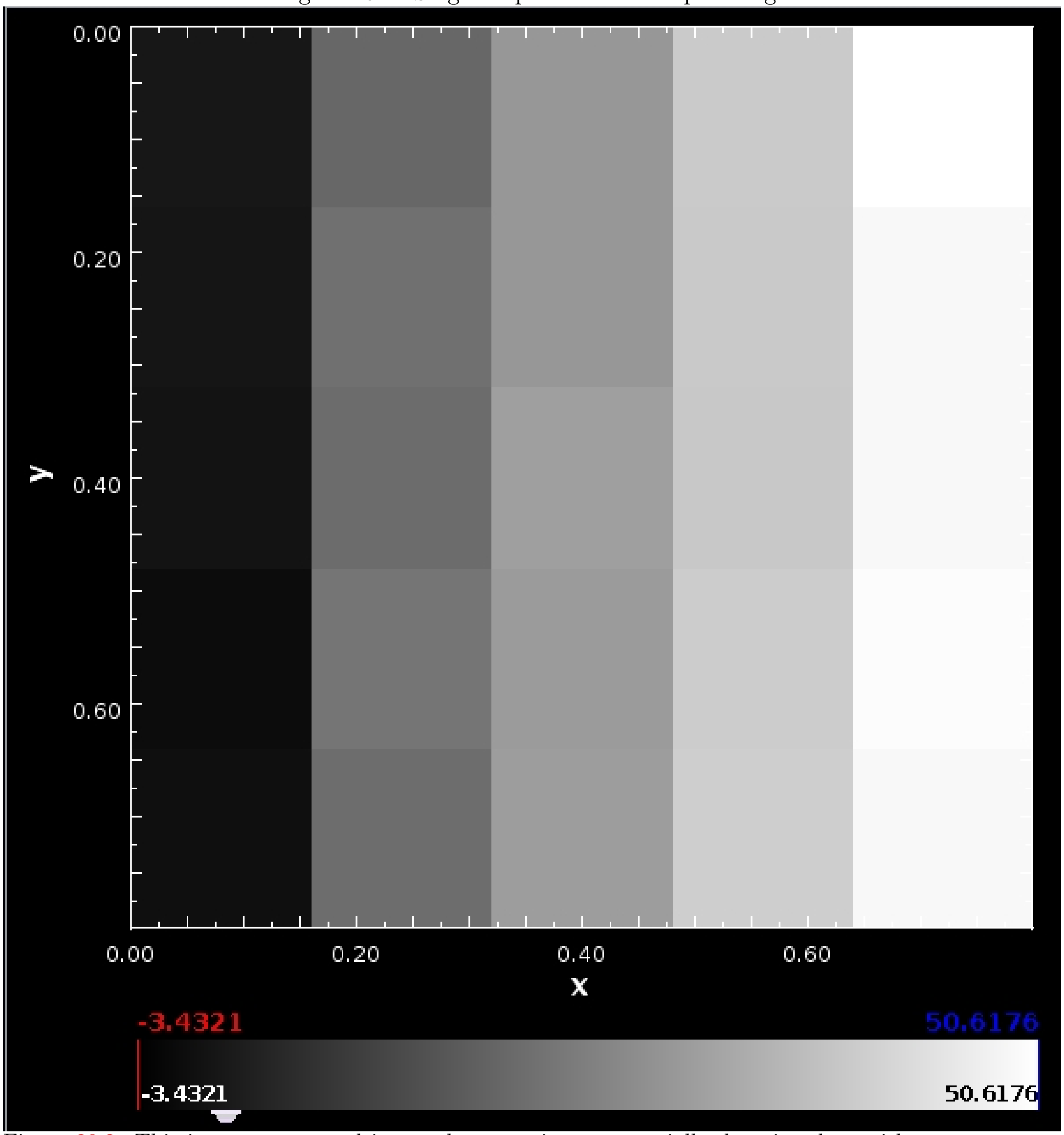

Figure 29.2: This is an test arrayed image that contains exponentially decaying data without a constant. The image was generated so that the decay rate constant in the horizontal rows have decay rate constants of 1, 2, 3, . . ., 5. The amplitude of each row is 0, 20, 30, 40, and 50.

Figure 29.3: Single Exponential Example Data

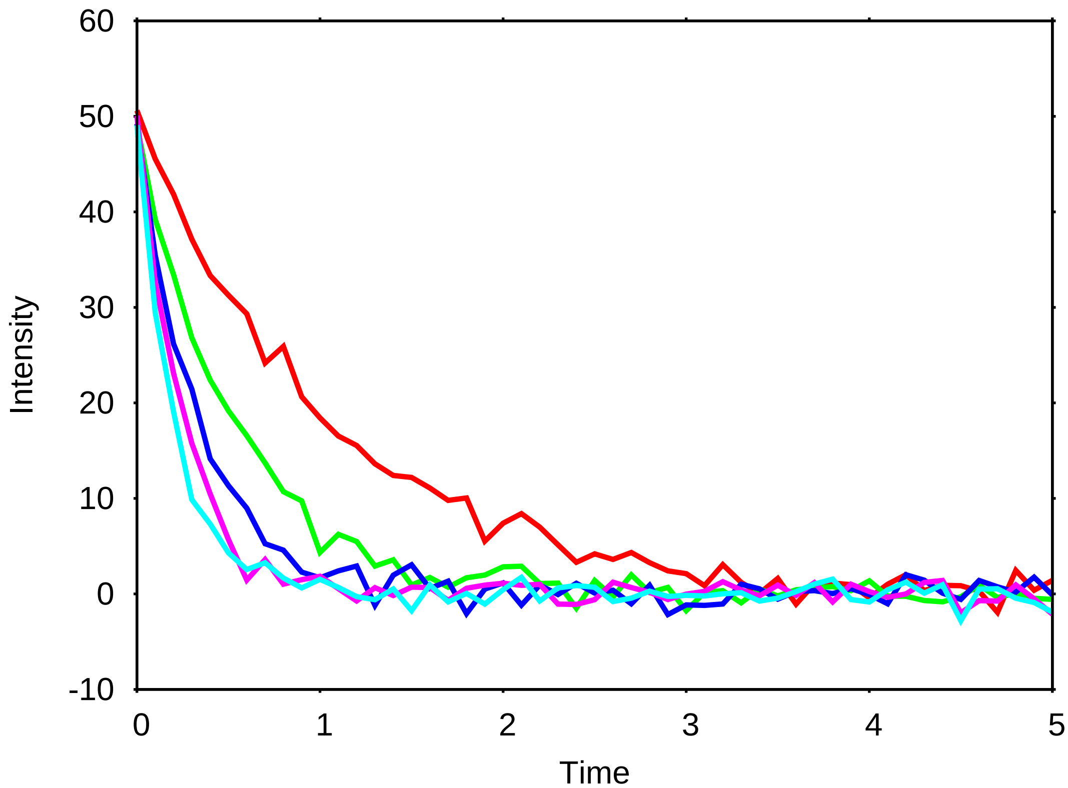

Figure 29.3: If one extracts the arrayed data in the 5th column, you will obtain the 5 exponentially decaying Ascii data sets shown here. Note that the intensity is the same in each data set, but the decay rate constant are rapidly increasing.

Figure 29.4: Posterior Probability For The ExpOneNoConst Model

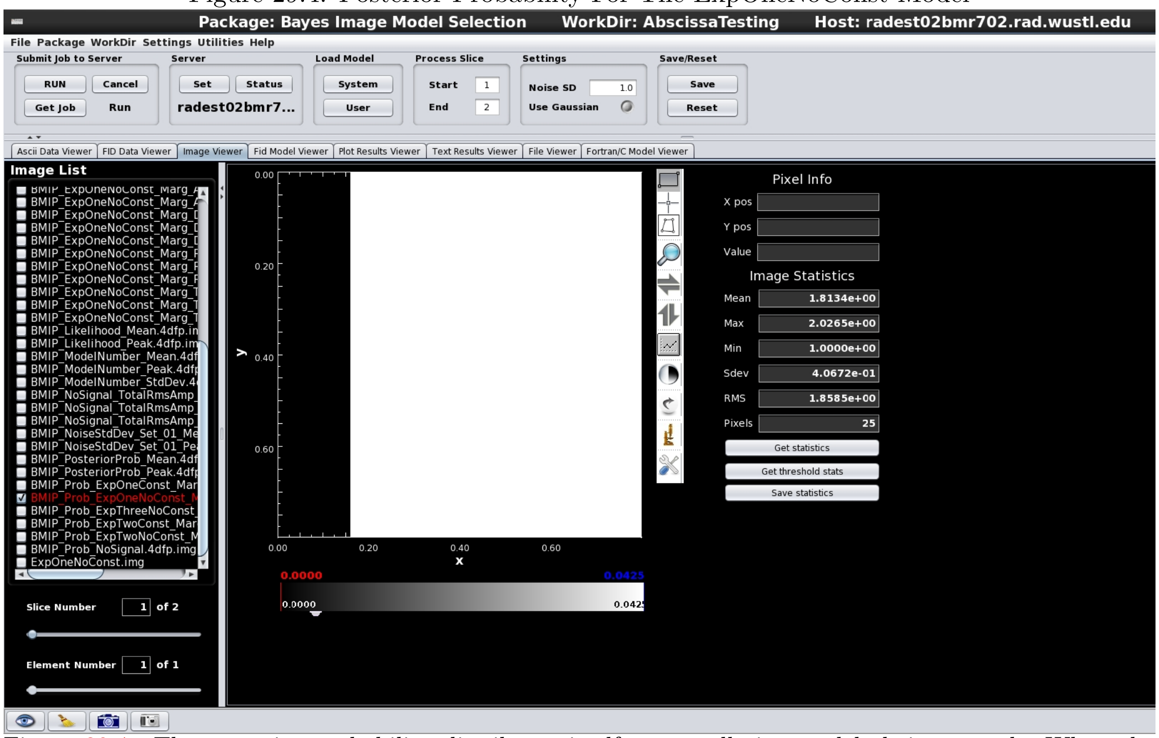

Figure 29.4: The posterior probability distributes itself across all size models being tested. When the posterior probability is near one, the model is almost certain. In this figure, we show the posterior probability for the ExpOneNoConst.f model. The posterior pro ability averages about 0.983 ± 0.003, so probability theory is almost positive all of the pixels are single exponential without a constant. And given that this is simulated test data, that is indeed the correct answer.

output.

Bibliography

[1] Rev. Thomas Bayes (1763), "An Essay Toward Solving a Problem in the Doctrine of Chances," *Philos. Trans. R. Soc. London*, **53,** pp. 370-418; reprinted in *Biometrika*, **45**, pp. 293-315 (1958), and *Facsimiles of Two Papers by Bayes,* with commentary by W. Edwards Deming, New York, Hafner, 1963.

[2] G. Larry Bretthorst (1988), "Bayesian Spectrum Analysis and Parameter Estimation," in *Lecture Notes in Statistics,* **48**, J. Berger, S. Fienberg, J. Gani, K. Krickenberg, and B. Singer (eds), Springer-Verlag, New York, New York.

[3] G. Larry Bretthorst (1990), "An Introduction to Parameter Estimation Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods,* Dartmouth College 1989, P. Fougère *ed.,* pp. 53-79, Kluwer Academic Publishers, Dordrecht the Netherlands.

[4] G. Larry Bretthorst (1990), "Bayesian Analysis I. Parameter Estimation Using Quadrature NMR Models" *J. Magn. Reson.*, **88**, pp. 533-551.

[5] G. Larry Bretthorst (1990), "Bayesian Analysis II. Signal Detection And Model Selection" *J. Magn. Reson.*, **88**, pp. 552-570.

[6] G. Larry Bretthorst (1990), "Bayesian Analysis III. Examples Relevant to NMR" *J. Magn. Reson.*, **88**, pp. 571-595.

[7] G. Larry Bretthorst (1991), "Bayesian Analysis. IV. Noise and Computing Time Considerations," *J. Magn. Reson.*, **93**, pp. 369-394.

[8] G. Larry Bretthorst (1992), "Bayesian Analysis. V. Amplitude Estimation for Multiple Well-Separated Sinusoids," *J. Magn. Reson.*, **98**, pp. 501-523.

[9] G. Larry Bretthorst (1992), "Estimating The Ratio Of Two Amplitudes In Nuclear Magnetic Resonance Data," in *Maximum Entropy and Bayesian Methods,* C. R. Smith *et al.* (eds.), pp. 67-77, Kluwer Academic Publishers, the Netherlands.

[10] G. Larry Bretthorst (1993), "On The Difference In Means," in Physics & Probability Essays in honor of Edwin T. Jaynes, W. T. Grandy and P. W. Milonni (eds.), pp. 177-194, Cambridge University Press, England.

[11] G. Larry Bretthorst (1996), "An Introduction To Model Selection Using Bayesian Probability Theory," in *Maximum Entropy and Bayesian Methods*, G. R. Heidbreder, ed., pp. 1-42, Kluwer Academic Publishers, Printed in the Netherlands.

[12] G. Larry Bretthorst (1999), "The Near-Irrelevance of Sampling Frequency Distributions," in *Maximum Entropy and Bayesian Methods,* W. von der Linden *et al.* (eds.), pp. 21-46, Kluwer Academic Publishers, the Netherlands.

[13] G. Larry Bretthorst (2001), "Nonuniform Sampling: Bandwidth and Aliasing," in *Maximum Entropy and Bayesian Methods in Science and Engineering,* Joshua Rychert, Gary Erickson and C. Ray Smith *eds.*, pp. 1-28, American Institute of Physics, USA.

[14] G. Larry Bretthorst, Christopher D. Kroenke, and Jeffrey J. Neil (2004), "Characterizing Water Diffusion In Fixed Baboon Brain," in Bayesian Inference And Maximum Entropy Methods In Science And Engineering, Rainer Fischer, Roland Preuss and Udo von Toussaint *eds.*, AIP conference Proceedings, **735**, pp. 3-15.

[15] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "Exponential parameter estimation (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 55-63.

[16] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J. H. Ackerman (2005), "Exponential model selection (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 64-72.

[17] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.H. Ackerman (2005), "How accurately can parameters from exponential models be estimated? A Bayesian view," *Concepts in Magnetic Resonance*, 27A, Issue 2, pp. 73-83.

[18] G. Larry Bretthorst, W. C. Hutton, J. R. Garbow, and Joseph J. H. Ackerman (2008), "High Dynamic Range MRS Time-Domain Signal Analysis," *Magn. Reson. in Med.*, **62**, pp. 1026-1035.

[19] V. Chandramouli, K. Ekberg, W. C. Schumann, S. C. Kalhan, J. Wahren, and B. R. Landau (1997), "Quantifying gluconeogenesis during fasting," *American Journal of Physiology*, **273**, pp. H1209-H1215.

[20] R. T. Cox (1961), "The Algebra of Probable Inference," Johns Hopkins Univ. Press, Baltimore.

[21] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1998), "Site-Specific Thermodynamics and Kinetics of a Coiled-Coil Transition by Spin Inversion Transfer NMR," *Biophysical Journal*, **74**, pp. 3190-3197.

[22] André d'Avignon, G. Larry Bretthorst, Marlyn Emerson Holtzer, and Alfred Holtzer (1999), "Thermodynamics and Kinetics of a Folded-Folded Transition at Valine-9 of a GCN4-Like Leucine Zipper," *Biophysical Journal*, **76**, pp. 2752-2759.

[23] David Freedman, and Persi Diaconis (1981), "On the histogram as a density estimator: L_2 theory," *Zeitschrift f¨r Wahrscheinlichkeitstheorie und verwandte Gebiete*, **57**, 4, pp. 453-476.

[24] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (1996), "Markov Chain Monte Carlo in Practice," Chapman & Hall, London.

[25] Paul M. Goggans, and Ying Chi (2004), "Using Thermodynamic Integration to Calculate the Posterior Probability in Bayesian Model Selection Problems," in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 23rd International Workshop*, **707**, pp. 59-66.

[26] Marlyn Emerson Holtzer, G. Larry Bretthorst, D. André d'Avignon, Ruth Hogue Angelette, Lisa Mints, and Alfred Holtzer (2001), "Temperature Dependence of the Folding and Unfolding Kinetics of the GCN4 Leucine Lipper via 13C alpha-NMR," *Biophysical Journal*, **80**, pp. 939-951.

[27] E. T. Jaynes (1968), "Prior Probabilities," *IEEE Transactions on Systems Science and Cybernetics*, SSC-4, pp. 227-241; reprinted in [30].

[28] E. T. Jaynes (1978), "Where Do We Stand On Maximum Entropy?" in *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus *Eds.*, pp. 15-118, Cambridge: MIT Press, Reprinted in [30].

[29] E. T. Jaynes (1980), "Marginalization and Prior Probabilities," in *Bayesian Analysis in Econometrics and Statistics*, A. Zellner *ed.,* North-Holland Publishing Company, Amsterdam; reprinted in [30].

[30] E. T. Jaynes (1983), "Papers on Probability, Statistics and Statistical Physics," a reprint collection, D. Reidel, Dordrecht the Netherlands; second edition Kluwer Academic Publishers, Dordrecht the Netherlands, 1989.

[31] E. T. Jaynes (1957), "How Does the Brain do Plausible Reasoning?" unpublished Stanford University Microwave Laboratory Report No. 421; reprinted in *Maximum-Entropy and Bayesian Methods in Science and Engineering* **1,** pp. 1-24, G. J. Erickson and C. R. Smith *Eds.,* 1988.

[32] E. T. Jaynes (2003), "Probability Theory—The Logic of Science," edited by G. Larry Bretthorst, Cambridge University Press, Cambridge UK.

[33] Sir Harold Jeffreys (1939), "Theory of Probability," Oxford Univ. Press, London; Later editions, 1948, 1961.

[34] John G. Jones, Michael A. Solomon, Suzanne M. Cole, A. Dean Sherry, and Craig R. Malloy (2001) "An integrated ^{2}H and ^{13}C NMR study of gluconeogenesis and TCA cycle flux in humans," *American Journal of Physiology, Endocrinology, and Metabolism,* **281,** pp. H848-H856.

[35] John Kotyk, N. G. Hoffman, W. C. Hutton, G. Larry Bretthorst, and J. J. H. Ackerman (1992), "Comparison of Fourier and Bayesian Analysis of NMR Signals. I. Well-Separated Resonances (The Single-Frequency Case)," *J. Magn. Reson.*, **98**, pp. 483–500.

[36] Pierre Simon Laplace (1814), "A Philosophical Essay on Probabilities," John Wiley & Sons, London, Chapman & Hall, Limited 1902. Translated from the 6th edition by F. W. Truscott and F. L. Emory.

[37] N. Lartillot, and H. Philippe (2006), "Computing Bayes Factors Using Thermodynamic Integration," *Systematic Biology*, **55** (2), pp. 195-207.

[38] D. Le Bihan, and E. Breton (1985), "Imagerie de diffusion in-vivo par rsonance," Comptes rendus de l'Acadmie des Sciences (Paris), **301** (15), pp. 1109-1112.

[39] N. R. Lomb (1976), "Least-Squares Frequency Analysis of Unevenly Spaced Data," *Astrophysical and Space Science*, **39**, pp. 447-462.

[40] T. J. Loredo (1990), "From Laplace To SN 1987A: Bayesian Inference In Astrophysics," in *Maximum Entropy and Bayesian Methods,* P. F. Fougere (ed), Kluwer Academic Publishers, Dordrecht, The Netherlands.

[41] Craig R. Malloy, A. Dean Sherry, and Mark Jeffrey (1988), "Evaluation of Carbon Flux and Substrate Selection through Alternate Pathways Involving the Citric Acid Cycle of the Heart by 13C NMR Spectroscopy," *Journal of Biological Chemistry*, **263** (15), pp. 6964-6971.

[42] Craig R. Malloy, Dean Sherry, and Mark Jeffrey (1990), "Analysis of tricarboxylic acid cycle of the heart using ^{13}C isotope isomers," *American Journal of Physiology*, **259**, pp. H987-H995.

[43] Lawrence R. Mead and Nikos Papanicolaou, "Maximum entropy in the problem of moments," *J. Math. Phys.* **25**, 2404–2417 (1984).

[44] K. Merboldt, Wolfgang Hanicke, and Jens Frahm (1969), "Self-diffusion NMR imaging using stimulated echoes," Journal of Magnetic Resonance, **64** (3), pp. 479-486.

[45] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller (1953), "Equation of State Calculations by Fast Computing Machines," Journal of Chemical Physics. The previous link is to the Americain Institute of Physics and if you do not have access to Science Sitations you many not be able to retrieve this paper.

[46] Radford M. Neal (1993), "Probabilistic Inference Using Markov Chain Monte Carlo Methods," technical report CRG-TR-93-1, Dept. of Computer Science, University of Toronto.

[47] Jeffrey J. Neil, and G. Larry Bretthorst (1993), "On the Use of Bayesian Probability Theory for Analysis of Exponential Decay Data: An Example Taken from Intravoxel Incoherent Motion Experiments," *Magn. Reson. in Med.*, **29**, pp. 642–647.

[48] H. Nyquist (1924), "Certain Factors Affecting Telegraph Speed," *Bell System Technical Journal*, **3**, pp. 324-346.

[49] H. Nyquist (1928), "Certain Topics in Telegraph Transmission Theory," *Transactions AIEE*, **3**, pp. 617-644.

[50] William H. Press, Saul A. Teukolsky, William T. Vetterling and Brian P. Flannery (1992), "Numerical Recipes The Art of Scientific Computing Second Edition," Cambridge University Press, Cambridge UK.

[51] Emanuel Parzen (1962), "On Estimation of a Probability Density Function and Mode," *Annals of Mathematical Statistics* **33**, 1065–1076

[52] Karl Pearson (1895), "Contributions to the Mathematical Theory of Evolution. II. Skew Variation in Homogeneous Material," *Phil. Trans. R. Soc. A* **186**, 343–326.

[53] Murray Rosenblatt, "Remarks on Some Nonparametric Estimates of a Density Function," *Annals of Mathematical Statistics* **27**, 832–837 (1956).

[54] Jeffery D. Scargle (1981), "Studies in Astronomical Time Series Analysis I. Random Process In The Time Domain," *Astrophysical Journal Supplement Series*, **45**, pp. 1-71.

[55] Jeffery D. Scargle (1982), "Studies in Astronomical Time Series Analysis II. Statistical Aspects of Spectral Analysis of Unevenly Sampled Data," *Astrophysical Journal*, **263**, pp. 835-853.

[56] Jeffery D. Scargle (1989), "Studies in Astronomical Time Series Analysis. III. Fourier Transforms, Autocorrelation Functions, and Cross-correlation Functions of Unevenly Spaced Data," *Astrophysical Journal*, **343**, pp. 874-887.

[57] Arthur Schuster (1905), "The Periodogram and its Optical Analogy," *Proceedings of the Royal Society of London*, **77**, p. 136-140.

[58] Claude E. Shannon (1948), "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, **27**, pp. 379-423.

[59] John E. Shore, and Rodney W. Johnson (1981), "Properties of cross-entropy minimization," *IEEE Trans. on Information Theory*, **IT-27,** No. 4, pp. 472-482.

[60] John E. Shore and Rodney W. Johnson (1980), "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Trans. on Information Theory*, **IT-26** (1), pp. 26-37.

[61] Devinderjit Sivia, and John Skilling (2006), "Data Analysis: A Bayesian Tutorial," Oxford University Press, USA.

[62] Edward O. Stejskal and Tanner, J. E. (1965), "Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient." Journal of Chemical Physics, **42** (1), pp. 288-292.

[63] D. G. Taylor and Bushell, M. C. (1985), "The spatial mapping of translational diffusion coefficients by the NMR imaging technique," Physics in Medicine and Biology, **30** (4), pp. 345-349.

[64] Myron Tribus (1969), "Rational Descriptions, Decisions and Designs," Pergamon Press, Oxford.

[65] P. M. Woodward (1953), "Probability and Information Theory, with Applications to Radar," McGraw-Hill, N. Y. Second edition (1987); R. E. Krieger Pub. Co., Malabar, Florida.

[66] Arnold Zellner (1971), "An Introduction to Bayesian Inference in Econometrics," John Wiley and Sons, New York.