

Bibliography

1. H. Akaike. Information theory and an extension of the maximum likelihood principle. In B.N. Petrov and F. Csaki, editors, *Proceedings of the 2nd International Symposium on Information Theory*, pages 267–281, 1973.
2. A. R. Barron and T. M. Cover. Minimum complexity density estimation. *IEEE Transactions on Information Theory*, 37:1034–1054, 1991.
3. R. A. Baxter and D. L. Dowe. Model selection in linear regression using the MML criterion. Technical report 96/276, Department of Computer Science, Monash University, Clayton, Victoria 3800, Australia, 1996.
4. J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*. Wiley, New York, 1994.
5. R.J. Bhansali and D.Y. Downham. Some properties of the order of an autoregressive model selected by a generalization of Akaike's FPE criterion. *Biometrika*, 64:547–551, 1977.
6. D. M. Boulton and C. S. Wallace. The information content of a multistate distribution. *Journal of Theoretical Biology*, 23:269–278, 1969.
7. R. Carnap. *The Logical Structure of the World & Pseudoproblems in Philosophy*. Berkeley and Los Angeles: University of California Press, 1967. Translated by Rolf A. George.
8. G. J. Chaitin. On the length of programs for computing finite sequences. *Journal of the Association for Computing Machinery*, 13:547–569, 1966.
9. G. J. Chaitin. A theory of program size formally identical to information theory. *Journal of the Association for Computing Machinery*, 22:329–340, 1975.
10. D. M. Chickering. A transformational characterization of equivalent Bayesian network structures. In P. Besnard and S. Hanks, editors, *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI95)*, pages 87–98, Montreal, Quebec, 1995.
11. J. H. Conway and N. J. A. Sloane. On the Voronoi regions of certain lattices. *SIAM Journal on Algebraic and Discrete Methods*, 5:294–305, 1984.
12. D. L. Dowe, J. J. Oliver, and C. S. Wallace. MML estimation of the parameters of the spherical Fisher distribution. In *Proceedings of the 7th International Workshop on Algorithmic Learning Theory (ALT'96)*, *Lecture Notes in Artificial Intelligence (LNAI) 1160*, pages 213–227, Sydney, Australia, 1996. Springer Verlag.
13. D. L. Dowe and C. S. Wallace. Resolving the Neyman-Scott problem by Minimum Message Length. In *Computing Science and Statistics – Proceedings of the 28th Symposium on the Interface*, volume 28, pages 614–618, Sydney, Australia, 1997.
14. R.M. Fano. The transmission of information. Technical report 65, Research Laboratory of Electronics, MIT, Cambridge, MA, 1949.
15. G. E. Farr and C. S. Wallace. The complexity of strict minimum message length inference. *The Computer Journal*, 45(3):285–292, 2002.

18. P. D. Grünwald. *The Minimum Description Length Principle and Reasoning under Uncertainty*. Institute for Logic, Language and Computation, Amsterdam, 1998.
19. D.A. Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the Institute of Electrical and Radio Engineers*, 40(9):1098–1101, 1952.
20. M. Hutter. New error bounds for Solomonoff prediction. *Journal of Computer and System Sciences*, 62:653–667, 2001.
21. K. G. Joreskog. Some contributions to maximum likelihood factor analysis. *Psychometrika*, 32:443–482, 1967.
22. M. Kearns, Y. Mansour, A. Y. Ng, and D. Ron. An experimental and theoretical comparison of model selection methods. *Machine Learning Journal*, 27:7–50, 1997.
23. A. N. Kolmogorov. Three approaches to the quantitative definition of information. *Problems of Information Transmission*, 1:4–7, 1965.
24. T.C.M. Lee. Tree-based wavelet regression for correlated data using the Minimum Description Length principle. *Australian and New Zealand Journal of Statistics*, 44(1):23–39, 2002.
25. L. A. Levin. Universal search problems. *Problems of Information Transmission*, 9:265–266, 1973.
26. Ming Li and P.M.B. Vitanyi. *An Introduction to Kolmogorov Complexity and its Applications*. Springer-Verlag, New York, 2nd edition, 1997.
27. K. V. Mardia, J. T. Kent, and J. M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.
28. J. J. Oliver. Decision graphs – an extension of decision trees. In *Proceedings of the Fourth International Workshop on Artificial Intelligence and Statistics*, pages 343–350, 1993. Extended version available as TR 173, Department of Computer Science, Monash University, Clayton, Victoria 3800, Australia.
29. J.D. Patrick. *An Information Measure Comparative Analysis of Megalithic Geometries*. PhD thesis, Department of Computer Science, Monash University, 1978.
30. H. Price. *Time's Arrow and Archimedes' Point*. Oxford University Press, New York, 1996.
31. J. R. Quinlan and R. L. Rivest. Inferring decision trees using the Minimum Description Length principle. *Information and Computation*, 80(3):227–248, 1989.
32. J.R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Mateo, CA, 1993.
33. A. Raman and J. Patrick. Inference of stochastic automata using the Sk-strings method. Tr IS/2/95, Information Systems Department, Massey University, Palmerston North, New Zealand, 1995.
34. J. J. Rissanen. A universal prior for integers and estimation by Minimum Description Length. *Annals of Statistics*, 11(2):416–431, 1983.
35. J. J. Rissanen. Stochastic complexity. *Journal of the Royal Statistical Society (Series B)*, 49:260–269, 1987.
36. J. J. Rissanen. Hypothesis selection and testing by the MDL principle. *The Computer Journal*, 42:223–239, 1999.

37. J. J. Rissanen and G. G. Langdon. Arithmetic coding. *IBM Journal of Research and Development*, 23(2):149–162, 1979.
38. G. W. Rumantir and C. S. Wallace. Minimum Message Length criterion for second-order polynomial model selection applied to tropical cyclone intensity forecasting. In M. R. Berthold, H. J. Lenz, E. Bradley, R. Kruse, and C. Borgelt, editors, *Advances in Intelligent Data Analysis V, LNCS2810*, pages 486–496. Springer-Verlag, Berlin, 2003.
39. G. Schou. Estimation of the concentration parameter in von Mises-Fisher distributions. *Biometrika*, 65:369–377, 1978.
40. G. Schwarz. Estimating dimension of a model. *Annals of Statistics*, 6:461–464, 1978.
41. C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 623–656, 1948.
42. R. J. Solomonoff. A formal theory of inductive inference. *Information and Control*, 7:1–22, 224–254, 1964.
43. R. J. Solomonoff. Two kinds of probabilistic induction. *The Computer Journal*, 42(4):256–259, 1999.
44. A. Thom. A statistical examination of the megalithic sites in Britain. *Journal of the Royal Statistical Society (Series A)*, 118(3):275–295, 1955.
45. A. Thom. The egg shaped standing stones rings in Britain. *Archives Internationales D'Histoire Des Science*, 14:291–303, 1961.
46. A. Thom. *Megalithic Sites in Britain*. London: Oxford University Press, 1967.
47. V. N. Vapnik. *Estimation of Dependencies Based on Empirical Data*. Nauka, Moscow (in Russian), 1979. (English translation: V. Vapnik, *Estimation of Dependencies Based on Empirical Data*, Springer, New York, 1982.)
48. M. Viswanathan and C. S. Wallace. A note on the comparison of polynomial selection methods. In *Proceedings of Uncertainty 99: the 7th International Workshop on Artificial Intelligence and Statistics*, pages 169–177, Fort Lauderdale, FL, 1999. Morgan Kaufmann.
49. M. Viswanathan, C. S. Wallace, D. L. Dowe, and K. B. Korb. Finding cutpoints in noisy binary sequences – a revised empirical evaluation. In *Proceedings of the 12th Australian Joint Conference on Artificial Intelligence, Lecture Notes in Artificial Intelligence (LNAI) 1747*, pages 405–416, Sydney, Australia, 1999.
50. V. Vovk and A. Gammerman. Complexity approximation principle. *The Computer Journal*, 42(4):318–322, 1999. Special issue on Kolmogorov Complexity.
51. V. V. V'yugin. Algorithmic complexity and properties of finite binary sequences. *The Computer Journal*, 42(4):294–317, 1999.
52. C. S. Wallace and D. M. Boulton. An information measure for classification. *The Computer Journal*, 11:185–194, 1968.
53. C. S. Wallace and D. L. Dowe. MML estimation of the von Mises concentration parameter. Technical report 93/193, Department of Computer Science, Monash University, Clayton, Victoria 3800, Australia, 1993.
54. C. S. Wallace and D. L. Dowe. MML clustering of multi-state, Poisson, von Mises circular and Gaussian distributions. *Statistics and Computing*, 10(1):73–83, 2000.
55. C. S. Wallace and P. R. Freeman. Estimation and inference by compact coding. *Journal of the Royal Statistical Society (Series B)*, 49:240–252, 1987.
56. C. S. Wallace and P. R. Freeman. Single factor analysis by MML estimation. *Journal of the Royal Statistical Society (Series B)*, 54(1):195–209, 1992.
57. C. S. Wallace and K. B. Korb. Learning linear causal models by MML sampling. In A. Gammerman, editor, *Causal Models and Intelligent Data Management*, pages 89–111. Springer Verlag, 1999.

10:1102-1194, 1982.

61. P. Zador. Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Transactions on Information Theory*, 28:139-149, 1982.