Sparse sampling: Spatial design for monitoring stream networks

Melissa J. Dobbie, *CSIRO*
Brent L. Henderson, *CSIRO*
Don L. Stevens, Jr, *Oregon State University*

Abstract

Spatial designs for monitoring stream networks, especially ephemeral systems, are typically non-standard, `sparse' and can be very complex, reflecting the complexity of the ecosystem being monitored, the scale of the population, and the competing multiple monitoring objectives. The main purpose of this paper is to present a review of approaches to spatial design to enable informed decisions to be made about developing practical and optimal spatial designs for future monitoring of streams.

Full Text: PDF

Dobbie, Melissa J., Henderson, Brent L., Stevens, Don L., Jr, , Sparse sampling: Spatial design for monitoring stream networks, Statistics Surveys, , (), 113-153 (electronic). DOI: 10.1214/07-SS032.

References

[1] Angulo, J.M., Ruiz-Medina, M.D., Alonso, F.J. and Bueso, M.C. (2005). Generalized approaches to spatial sampling design. Environmetrics 16, 523–534. MR2147541

[2] Baddeley M.C, Curtis A. and Wood, R. (2004). An introduction to prior information derived from probabilistic judgments: elicitation of knowledge, cognitive bias and herding. In Geological Prior Information: Informing Science and Engineering, A. Curtis and R. Wood, Eds. Geological Society of London, 15–27.

[3] Bergen, M. (1996). The Southern California Bight pilot project: sampling design. In Southern California Coastal Water Research Project Annual Report 1994-95. Westminster, CA, 109–113.

[4] Berry, B.J.L. and Baker, A.M. (1968). Geographic sampling. In Spatial Analysis, B.J.L. Berry and D.F. Marble, Eds. Prentice Hall, Englewood Cliffs, NJ.

[5] Bogaert, P. and Russo, D. (1999). Optimal spatial sampling design for the estimation of the variogram based on a least squares approach. Water Resources Research 35, 1275–1289.

[6] Brus, D.J. and de Gruijter, J.J. (1993). Design-based versus model-based estimates of spatial means: Theory and application in environmental soil science. Environmetrics 4, 123–152.

[7] Brus, D.J. and de Gruijter, J.J. (1997). Random sampling or geostatistical modeling? Choosing between design-based and model-based sampling strategies for soil (with Discussion). Geoderma 60, 1–44.

[8] Bueso, J.M., Angulo, J.M. and Alonso, F.J. (1998). A state-space model approach to optimum spatial sampling design based on entropy. Environmental and Ecological Statistics 5, 29–44.

[9] Butler, D. (2006). Everything, Everywhere. Nature 440, 402–405.

[10] Caeiro, S., Painho, M., Goovaerts, P., Costa, H. and Sousa, S. (2003). Spatial sampling design for sediment quality assessment in estuaries. Environmental Modelling and Software 18, 853–859.

[11] Casselton, W.F. and Zidek, J.V. (1984). Optimal monitoring network designs. Statistics and Probability Letters 2, 223–227.

[12] Casselton, W.F., Kan, L. and Zidek, J.V. (1992). Quality data networks that minimize entropy. In Statistics in the Environmental and Earth Sciences, A.T. Walden and P. Guttorp, Eds. Edward Arnold, London, 10–38.

[13] Chao, C.-T. and Thompson, S.K. (2001). Optimal adaptive selection of sampling sites. Environmetrics 12, 517–538.

[14] Cochran, W.G. (1977). Sampling techniques. 3rd ed., Wiley, New York. MR0474575

[15] Cotter, J. and Nealon, J. (1987). Area Frame Design for Agricultural Surveys. Area Frame Section, Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture.

[16] Cressie, N. (1993). Statistics for spatial data. 2nd ed., Wiley, New York. MR1239641

[17] Cressie, N., Calder, C.A., Clark, J.S., Ver Hoef, J.M. and Wikle, C.K. (2007). Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. Department of Statistics Preprint No. 798, The Ohio State University.

[18] Cressie, N., Gotway, C.A., and Grondona, M.O. (1990). Spatial prediction from networks. Chemometrics and Intelligent Laboratory Systems I, 251–271.

[19] de Gruijter, J.J., and Ter Braak, C.J.F. (1990). Model free estimation from survey samples: A reappraisal of classical sampling theory. Mathematical Geology 22, 407–415. MR1047605

[20] Diaz-Ramos, S., Stevens, D.L., Jr and Olsen, A.R. (1996). EMAP Statistical Methods Manual. Rep. EPA/620/R-96/002, U.S. Environmental Protection Agency, Office of Research and Development, NHEERL-WED, Corvallis, Oregon.

[21] Diggle, P.J. and Lophaven, S. (2006). Bayesian Geostatistical Design. Scandanavian Journal of Statistics 33, 53–64. MR2255109

[22] Dixon, W. and Chiswell, B. (1996). Review of aquatic monitoring design. Water Resources 30(9), 1935–1948.

[23] Dixon, W., Smyth, G., and Chiswell, B. (1999). Optimized selection of river sampling sites. Water Resources 33(4), 971–978.

[24] Fairweather, P. (1991). Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research 42, 555–567.

[25] Fedorov, V.V. (1996). Design of spatial experiments: model fitting and prediction. In Handbook of Statistics, Volume 13, C.R. Rao and S. Gosh, Eds. North Holland. MR1492578

[26] Fuentes M., Chaudhuri A., and Holland D.M. (2005). Bayesian entropy for spatial sampling design of environmental data. Institute of Statistics Mimeo Series No. 2571

[27] Fuller, W.A. (1999). Environmental surveys over time. Journal of Agricultural, Biological and Environmental Statistics 4(4), 331–345. MR1816371

[28] Gilbert, R.0. (1987). Statistical Methods for Environmental Pollution Monitoring. Van Nostrand Reinhold, New York.

[29] Hansen, M.H., Madow, W.G., and Tepping, B.J. (1983). An evaluation of model dependent and probability sampling inferences in sample surveys. Journal of the American Statistical Association 78, 776–760.

[30] Hazard, J.W. and Law, B.E. (1989). Forest Survey Methods Used in the USDA Forest Service. EPA/600/3 89/065. NTIS PB89 220 594/AS. U.S. EPA Environmental Research Laboratory, Corvallis, OR.

[31] Herlihy, A.T., Larsen, D.P., Paulsen, S.G., Urquhart, N.S., and Rosenbaum, B.J. (2000). Designing a spatially balanced, randomized site selection process for regional stream surveys: the EMAP mid-Atlantic pilot study. Environmental Monitoring and Assessment 63, 95–113.

[32] Horvitz, D.G. and Thompson, D.J. (1952). A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47, 663–685. MR0053460

[33] Kaufmann, P.R., Herlihy, A.T., Elwood, J.W., Mitch, M.E., Overton, W.S., Sale, M.J., Cougan, K.A., Peck, D.V., Reckhow, K.H., Kinney, A.J., Christie, S.J., Brown, D.D., Hagley, C.A., and Jager, H.I. (1988). Chemical Characteristics of Streams in the Mid-Atlantic and Southeastern United States Volume I: Population Descriptions and Physico-Chemical Relationships. EPA/600/3-88/021a U.S. Environmental Protection Agency, Washington, D.C.

[34] Kott, P.S. (1986). Some asymptotic results for the systematic and stratified sampling of a finite population. Biometrika 73, 485–491. MR0855909

[35] Kutz, F.W. and Linthurst, R.A. (1990). A systems-level approach to environmental assessment. Toxicological and Environmental Chemistry 28, 105–114.

[36] Kynn, M. (2005). Eliciting expert knowledge for Bayesian logistic regression in species habitat modeling. Unpublished PhD thesis, Queensland University of Technology, Brisbane, Australia.

[37] Landers, D.H., Eilers, J.M., Brakke, D.F., Overton, W.S., Kellar, P.E., Silverstein, M.E., Schonbrod, R.D., Crowe, R.E., Linthurst, R.A., Omernik, J.M., Teague, S.A. and Meier, E.P. (1987). Characteristics of Lakes in the Western United States Volume I: Population Descriptions and Physico Chemical Relationships. EPA 600/3 86/054a. U.S. Environmental Protection Agency, Washington, D.C.

[38] Lark, R.M. (2002). Modelling complex soil properties as contaminated regionalised variables. Geoderma 106, 173–190.

[39] Lark, R.M. (2003). Two robust estimators of the cross-variogram for multivariategeostatistical analysis of soil properties. European Journal of Soil Science 54, 187–201.

[40] Le, N.D. and Zidek, J.V. (1994). Network designs for monitoring multivariate random spatial fields. In Recent Advances in Statistics and Probability, J.P. Vilaplana and M.L. Puri, Eds. 191–206.

[41] Le, N.D., Sun, L. and Zidek, J.V. (2003). Designing networks for monitoring multivariate environmental fields using data with monotone pattern. TR 2003-5, Statistical and Applied Mathematical Sciences Institute, RTP, NC.

[42] Lele, S. and Allen, K. (2006). On using expert opinion in ecological analyses: a frequentist approach. Environmetrics 17(7), 683–704. MR2252325

[43] Lele, S. and Das, A. (2000). Elicited data and incorporation of expert opinion for statistical inference in spatial studies. Mathematical Geology 32(4), 465–487.

[44] Lesser, V.M. (2001). Applying survey research methods to account for denied access to research sites on private property. Wetlands 21(4), 639–647.

[45] Lettenmaier, D.P. (1978). Design considerations for ambient stream quality monitoring. Water Resources Bulletin 14, 884–902.

[46] Liebetrau, A.M. (1979). Water quality sampling: some statistical considerations. Water Resources Research 15, 1717–1725.

[47] Linthurst, R.A., Landers, D.H., Eilers, J.M., Brakke, D.F., Overton, W.S., Meier, E.P. and Crowe, R.E. (1986). Characteristics of Lakes in the Eastern United States Volume I: Population Descriptions and Physico Chemical Relationships. EPA 600/4 86/007a. U.S. Environmental Protection Agency, Washington, D.C.

[48] Little, R.J.A. (2002). Statistical Analysis With Missing Data. 2nd ed., John Wiley and Sons, Inc., New Jersey. MR1925014

[49] Lozano, S.J., Scharold, J.V. and Nalepa, T.F. (2001). Recent declines in benthic macroinvertebrate densities in Lake Ontario. Canadian Journal of Fisheries and Aquatic Sciences 58, 518–529.

[50] Maher, W.A., Cullen, P.W. and Norris, R.H. (1994). Framework for designing sampling programs. Environmental Monitoring and Assessment 30, 134–162.

[51] Marchant, B.P. and Lark, R.M. (2007). Optimized sample schemes for geostatistical surveys. Mathematical Geology 39, 113–134. MR2322251

[52] Martin, T.G., Kuhnert, P.K., Mengersen, K. and Possingham, H. (2004). The power of expert opinion in ecological models: A Bayesian approach examining the impact of livestock grazing on birds. Ecological Applications 15, 266–280.

[53] McBratney, A.B., Webster, R. and Burgess, T.M. (1981). The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I. Theory and method. Computers and Geosciences 7, 331–334.

[54] McDonald, T. (2003). Review of Environmental monitoring methods: Survey Design. Environmental Monitoring and Assessment 85, 277–292.

[55] Messer, J.J., Ariss, C.W., Baker, J.R., Drousé, S.K., Eshelman, K.N., Kaufmann, P.R., Linthurst, R.A., Omernik, J.M., Overton, W.S., Sale, M.J., Schonbrod, R.D., Stambaugh, S.M. and Tuschall, J.R., Jr. (1986). National Stream Survey Phase I - Pilot Survey. EPA 600/4 86/026. U.S. Environmental Protection Agency, Washington, D.C.

[56] Messer, J.J., Linthurst, R.A. and Overton, W.S. (1991). An EPA Program for monitoring ecological status and trends. Environmental Management 17, 67–78.

[57] Muller, W.G. (2000). Collecting Spatial Data: Optimum Design of Experiments for Random Fields. 2nd ed., Physica Verlag, Heidelberg. MR1795688

[58] Müller, W. (2005). A comparison of spatial design methods for correlated observations. Environmetrics 16, 495–505. MR2147539

[59] Muller, W.G. and Zimmerman, D.L. (1999). Optimal designs for variogram estimation. Environmetrics 10, 23–37.

[60] Munoz-Hernandez, B., Lesser, V.M. and Smith, R. (2006). Adjustment procedures to account for non-ignorable missing data in environmental surveys. Environmetrics 17, 653–662. MR2247175

[61] Nychka, D. and Saltzman, N. (1998). Design of air-quality monitoring networks. In Case Studies in Environmental Statistics, D. Nychka, W. Piegorsch, and L. Cox, Eds. Springer, New York, 51–76.

[62] Olea, R.A. (1984). Sampling design optimization for spatial functions. Mathematical Geology 16, 369–392.

[63] Olsen, A.R., Stevens, D.L., Jr, and White, D. (1998). Application of global grids in environmental sampling. Computing Science and Statistics 30, 279–284.

[64] Olsen, A.R., Sedransk, J., Edwards, D., Gotway, C.A., Liggett, W., Rathburn, S., Reckhow, K.H. and Young, L.J. (1999). Statistical issues for monitoring ecological and natural resources in the United States. Environmental Monitoring and Assessment 54, 1–45

[65] Overton, W.S. (1993). Probability sampling and population inference in monitoring programs. In Environmental Monitoring with GIS, M.R. Goodchild, S.O. Parks, and C.T. Steyaert, Eds. Oxford University Press, NY, 470–480.

[66] Overton, W.S. and Stehman, S.V. (1993). Properties of designs for sampling continuous spatial resources from a triangular grid. Communications in Statistics, Part A - Theory and Methods 22, 2641–2660. MR1237705

[67] Overton, W.S. and Stehman, S.V. (1996). Desirable design considerations for long-term monitoring of ecological variables. Environmental and Ecological Statistics 3, 349–361.

[68] Overton, W.S., White, D., and Stevens, D.L., Jr. (1990). Environmental Monitoring and Assessment Program: design report. EPA/600/3-91/053, Washington (DC): US Environmental Protection Agency, 52pp.

[69] Royall, R. M. and Herson, J. (1973). Robust estimation in finite populations. II: Stratification on a size variable. Journal of the American Statistical Association 68, 890–893. MR0386088

[70] Royle, J.A. and Nychka, D. (1998). An Algorithm for the Construction of Spatial Coverage Designs with Implementation in Splus. Computers and Geosciences 24, 479–488.

[71] Sanders, T.G. (Editor) (1987). Design of Networks for Monitoring Water Quality. 2nd ed., WRP Publication, Littleton, Colorado.

[72] Sarndal, C. (1978). Design-based and model-based inference for survey sampling. Scandinavian Journal of Statistics 5, 27–52. MR0471132

[73] Sarndal, C.E., Swensson, B. and Wretman, J.H. (1992). Model assisted survey sampling. Springer, New York. MR1140409

[74] Sarno, E. (2005). Testing information redundancy in environmental monitoring networks, Environmetrics 16, 71–79. MR2146899

[75] Sharp, W.E. (1971). A topologically optimum water sampling plan for rivers and streams. Water Resources Research 7, 1641–1646.

[76] Shreve, R.L. (1966). Statistical law of stream numbers. Journal of Geology 74, 17–37.

[77] Skalski, J.R. (1990). A design for long term status and trends monitoring. Journal of Environmental Management 30, 139–144.

[78] Smith, M.J. and Storey, A.W. (2001). Design and Implementation of Baseline Monitoring (DIBM3): Developing an Ecosystem Health Monitoring Program for Rivers and Streams in Southeast Queensland. Report to the South East Queensland Regional Water Quality Management Strategy, Brisbane, 416pp.

[79] Smith, R.L. (2001). Design of a monitoring network. In Environmental Statistics course notes. Department of Statistics, University of North Carolina, 219–268. Available on web at http://www.stat.unc.edu/postscript/rs/envnotes.pdf

[80] SRA (2004). Sustainable Rivers Audit Program. MDBC Publication No. 38/04

[81] Stehman, S.V. (1999). Basic probability sampling designs for thematic map accuracy assessments. International Journal of Remote Sensing 20(12), 2423–2441.

[82] Stein, M. L. (1988). Asymptotically efficient prediction of a random field with a

misspecified covariance function. Annals of Statistics 16, 55–63. MR0924856

[83] Stein, A. and Ettema, C. (2003). An overview of spatial sampling procedures and experimental design of spatial studies for ecosystem comparison. Agriculture, Ecosystems and Environment 94, 31–47.

[84] Stevens, D.L., Jr (1997). Variable density grid-based sampling designs for continuous spatial populations. Environmetrics 8, 167–95.

[85] Stevens, D.L., Jr (2002). Sampling Design and Statistical Analysis Methods for the Integrated Biological and Physical Monitoring of Oregon Streams, Oregon Department of Fish and Wildlife Report Number OPSW-ODFW-2002-07, 14 pages + appendices.

[86] Stevens, D.L., Jr, and Jensen, S.F. (2007). Sample Design, Execution and Analysis for Wetland Assessment. Wetlands 27(3), 515–523.

[87] Stevens, D.L., Jr and Olsen, A.R. (1999). Spatially restricted surveys over time for aquatic resources. Journal of Agricultural, Biological, and Environmental Statistics 4, 415–28. MR1816377

[88] Stevens, D.L., Jr. and Olsen, A.R. (2004). Spatially-balanced sampling of natural resources. Journal of American Statistical Association 99(465): 262–278. MR2061889

[89] Strahler, A.N. (1952). Dynamic basis of geomorphology. Geological Society of America Bulletin 63, 923–938.

[90] Theobald, D.M., Stevens, D.L. Jr, White, D., Urquhart, N.S., Olsen, A.R. and Norman, J.B. (2007). Using GIS to generate spatially-balanced random survey designs for natural resource applications. Environmental Management 40, 134–146.

[91] Thompson, S.K. (1990). Adaptive cluster sampling. Journal of American Statistical Association 85, 1050–1059. MR1134501

[92] Thompson, S.K. (1991). Stratified adaptive cluster sampling. Biometrika 78, 389–397. MR1131173

[93] Thompson, S.K. (2002). Sampling. 2nd ed., John Wiley and Sons, New York. MR1891249

[94] Thompson, S.K. (2006). Adaptive Web Sampling. Biometrics 62, 1224–1234. MR2307448

[95] Thompson, S.K. and Seber G.A.F. (1996). Adaptive sampling. Wiley, New York. MR1390995

[96] Urquhart, N.S. and Kincaid, T.M. (1999). Designs for detecting trend from repeated surveys of ecological resources. Journal of Agricultural, Biological and Environmental Statistics 4, 404–414. MR1816376

[97] Urquhart, N.S., Overton, W.S., and Birkes, D.S. (1993). Comparing sampling designs for monitoring ecological status and trends: impact of temporal patterns. Chapter 3 in Statistics for the Environment, V. Barnett and K.F. Turkman, Eds. John Wiley and Sons, 71–85.

[98] Urquhart, N.S., Paulsen, S.G., and Larsen, D.P. (1998). Monitoring for policy-relevant regional trends over time. Ecological Applications 8, 246–257.

[99] Van Groenigen, J.W. and Stein, A. (1998). Constrained optimization of spatial sampling using continuous simulated annealing. Journal of Environmental Quality 43, 684–691.

[100] Ver Hoef, J.M. (2002). Model-assisted sampling. In Encyclopaedia of Environmetrics Vol. 3, A. El-Shaarawi and W. Piegorsch, Eds. John Wiley and Sons, 1283–1284.

[101] Waller, L.A. (2002). Optimal spatial design. In Encyclopaedia of Environmetrics Vol. 3, A. El-Shaarawi and W. Piegorsch, Eds. John Wiley and Sons, 2067–2073.

[102] Wardrop, D.H., Kentula, M.E., Stevens, D.L.,Jr, Jensen, S.F. and Brooks, R.P. (2007). Assessment of Wetland Condition: An Example from the Upper Juniata Watershed in Pennsylvania, USA. Wetlands 27, 416–431.

[103] Wardrop, D.H., Kentula, M.E., Stevens, D.L., Jr, Jensen, S.F. and Hychka, K.C. (2007). Assessment of Wetlands in the Upper Juniata Watershed in Pennsylvania, USA, using the Hydrogeomorphic Approach. Wetlands 27, 432–445.

[104] Warwick, A.W. and Myers, D.E. (1987). Optimization of sampling locations for variogram calculations. Water Resources Research 23(3), 496–500.

[105] Whigham, D.F., Jacobs, A.D., Weller, D.E., Jordan, T.E., Kentula, M.E., Jensen, S.F., and Stevens, D.L., Jr (2007). Combining HGM and EMAP Procedures to Assess Wetlands at the Watershed Scale - Status of Flats and Non-Tidal Riverine Wetlands in the Nanticoke River Watershed, Delaware And Maryland (USA). Wetlands 27, 462–478.

[106] White, D., Kimmerling, A.J., and Overton, W.S. (1992). Cartographic and geometric components of a global sampling design for environmental monitoring. Cartography and Geographic Information Systems 19, 5–22.

[107] Wiens, D.P. (2005). Robustness in Spatial Studies II: Minimax Design. Environmetrics 16, 205–217. MR2146908

[108] Wikle, C.K. and Royle, J.A. (1999). Space-time dynamic design of environmental monitoring networks. Journal of Agricultural, Biological, and Environmental Statistics 4, 489–507. MR1816382

[109] Wikle, C.K. and Royle, J.A. (2005). Dynamic design of ecological monitoring networks for non-Gaussian spatio-temporal data. Environmetrics 16, 507–522. MR2147540

[110] Winkels, H.J. and Stein, A. (1997). Optimal cost-effective sampling for monitoring and dredging of contaminated sediments. Journal of Environmental Quality 26, 933–946.

[111] Xia, G., Miranda, M. L., and Gelfand A.E. (2006). Approximately optimal spatial design approaches for environmental health data. Environmetrics 17, 363–385 MR2239678

[112] Yates, F. (1946). A review of recent statistical developments in sampling and sample surveys. Journal of the Royal Statistical Society 109, 12–43.

[113] Zhu, Z. and Stein, M.L. (2005). Spatial Sampling Design for Parameter Estimation of the Covariance Function. Journal of Statistical Planning and Inference 134, 583–603. MR2200074

[114] Zhu, Z. and Zhang, H. (2006). Spatial Sampling Design under In-fill Asymptotic Framework. Environmetrics 17, 323–337. MR2239675

[115] Zidek, J.V., Sun, W., and Le, N.D. (2000). Designing and integrating composite networks for monitoring multivariate (Gaussian) pollution fields. Journal of Applied Statistics 49, 63–79. MR1817875

[116] Zimmerman, D L. (2006). Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction. Environmetrics 17, 635–652. MR2247174