Received Date : 06-Oct-2016

Revised Date : 22-Dec-2016

Accepted Date : 17-Jan-2017

Article type : Application

Handling editor: Prof. Robert Freckleton

Emon: an R-package to support the design of marine ecological and environmental studies, surveys and monitoring programmes

Jon Barry[1], David Maxwell[1], Simon Jennings[1,3], David Walker[2] and Joanna Murray[1]

[1]Centre for Environment, Fisheries and Aquaculture Science, Lowestoft Laboratory, Pakefield Road, Lowestoft, Suffolk NR33 OHT, UK.

[2]Centre for Environment, Fisheries and Aquaculture Science, Weymouth Laboratory, The Nothe, Barrack Road, Weymouth DT4 8UB, UK.

[3]School of Environmental Sciences, University of East Anglia, Norwich Research Park, NR4 7TJ, UK.

Corresponding Author. Tel +44 1502 652244 email jon.barry@cefas.co.uk

Running title: Tools to support design of surveys and monitoring

Keywords: marine, survey, monitoring, design, power, distribution, abundance, feature detection.

SUMMARY

1. Marine systems and their biota are always changing, in response to environmental and human pressures such as climate variation and change, eutrophication, fisheries exploitation, litter, noise and accidental releases or regular discharges of contaminants, radioactivity and hazardous substances. Studies, surveys and monitoring help to describe and understand system responses to these pressures, and provide evidence to assess the needs for, and effects of, management interventions. Studies, surveys and monitoring are often costly, especially offshore, so small investments in preliminary data collection and systematic planning of these activities can help to make best use of resources and inform trade-offs between budgets and expectations.
2. To meet recognised needs for accessible tools to plan some aspects of studies, surveys and monitoring, we developed the R package *emon*. *Emon* includes routines for study design through power analysis (assuming independence of observations) and feature detection; which are the focus of this paper.
3. We hope that access to the functions in *emon*, many drawing on work which was previously published but without code for implementation, will raise awareness of what marine studies, surveys and monitoring can achieve, thus encouraging cost-effective, needs- and evidence-based designs.

Introduction

Studies, surveys and monitoring are fundamental methods for describing (i) the state of marine systems, (ii) human and environmental impacts on those systems and (iii) the needs for, and effects of, management measures. Surveys are used to characterise the abundance, distribution and occurrence of biota or concentrations of chemicals. Surveys may be synoptic, to provide a one-off characterisation of the distribution of habitats or contaminants or to search for a rare or invasive species, for example. But surveys may also be repeated – when this is done regularly, we describe the process as monitoring. Monitoring can fulfil single or multiple functions, depending on frequency, scale and range of measurements made. These range from describing the timing of single or cyclical events (phenology), to describing changes in biological, chemical and physical states and processes, which may provide feedback on the effects of human and environmental pressure and the response of a system to management intervention.

Sampling at sea can be particularly challenging and expensive when the sea is deep, opaque, inaccessible, rough and/or seasonally and spatially variable. While sampling and monitoring methods are increasingly automated, many methods are still ship-based or require ship support, and costs can range from £500 per day for a smaller inshore vessel to >£10K for larger vessels working offshore, with additional costs for staff time, consumables, sample and data processing. Thus, as well as ensuring studies, surveys and monitoring are likely to meet their intended objectives, small investments in preliminary data collection and systematic planning can save significant costs by contributing to designs that use resources as efficiently as possible and informing the choice of options based on quantified trade-offs between budgets and expectations. There is a reasonably large literature focused on the design of studies, surveys and monitoring the sea (e.g. Underwood & Chapman 2013; Kimura & Somerton 2006; Underwood 2000; Nicholson, Fryer & Ross 1997), that contributes to a wider set of questions about the rationale and objectives of surveys and monitoring (e.g. Lindenmayer & Likens 2010; Nichol & Williams 2006; Hellawell 1991).

Several recurring questions emerge when designing studies, surveys and monitoring, but the tools to address many of these questions are not readily available to practitioners or rely on assumptions that are violated in many circumstances. These include questions about the probabilities that studies and surveys detect specified effects and differences (statistical power), the probabilities of detecting different types of environmental change or trend when monitoring, and the probabilities of detecting individuals or patches with a given sampling density and design. Several methods specifically developed to design and improve marine monitoring programmes are described in the literature. For instance, Nicholson, Fryer & Ross (1997) outlined approaches to detect year-to-year differences and linear trends in contaminant levels in fish and shellfish and Fryer & Nicholson (1999) extended this to show how smoothers could be used to assess the power of detecting more complex trends. Fryer & Nicholson (1993) considered power to detect linear trends and incidents where between-year variation was modelled by a random effect. Other research has assessed the power of a large-scale annual monitoring survey to detect changes in community metrics (Nicholson & Jennings 2004) or increases and decreases in fish abundance (Maxwell & Jennings 2005), and Blanchard, Maxwell and Jennings (2008) extended these approaches to compare the power of different survey designs to detect specified trends in abundance. Often, the survey objective is to determine whether a given region contains an important feature – for example, a plant or animal of conservation interest or a non-indigenous species that one wishes to eradicate. Barry and Nicholson (1993) calculated this probability on the assumption that the feature was a single circle and that the sampling pattern is either random, square lattice, triangular lattice or randomly placed on transects. This work was developed by Nicholson and Barry (1996), with the design objective that if one or more patches exceeded some specified size, then the probability of patch detection should be high.

Emon **R package**

To better meet needs for accessible statistical tools to plan studies, surveys and monitoring, we developed the R package *emon* (Barry & Maxwell, 2015). *Emon* (short for "environmental monitoring") makes accessible, generalises and extends methods of power analysis and feature detection which were formerly published without code for application. The functions in *emon* also address the limitations of many parametric methods for assessing power. To fully understand the uses of *emon* we recommend that readers download the *emon* library and use the help facilities as a supplement to the information provided here.

Emon includes five main groups of functions for running calculations to support study, survey and monitoring design (Table 1). Function *power.groups* calculates the power for detecting difference in the mean between two ‘groups’ (e.g. two areas or time points). Function *power.BACI* calculates power for BACI designs. Function *power.trend* is used with a function *generate.trend* to determine the power to detect linear, incident, step and “updown” trends. Functions *detect* and *detect.prop* calculate sample sizes and probabilities for detection of patches and features with different spatial sampling patterns. Other functions in the package support the main functions and deal with topics not covered here.

Power analysis

Statistical power analysis can be used when planning studies, surveys or monitoring to assess the feasibility of detecting a change and to achieve efficient designs given resources available (Gerrodette 1987; Holt, Gerrodette & Cologne 1987; Peterman 1990; Peterman & M’Gonigle 1992; Steidl, Hayes & Schauber 1997; Ellis *et al*. 2015). There are several algebraic expressions which define power for common methods of analysis (e.g. Cohen 1988). The advantage of using these expressions is that they are quick to compute. The disadvantages are that they apply to only a subset of potential survey designs and they make specific assumptions about the statistical distribution of the data, usually that it is Normal with constant variance. Consequently, the expressions lack flexibility to deal with many observed distributions and survey designs. Power analysis by simulation provides greater flexibility, because any appropriate distribution can be assumed. *Emon* uses the simulation approach and data can take any of four distributions: Normal, Poisson, Negative Binomial or Lognormal. The power functions assume that data are statistically independent. Valid power calculations depend on appropriate choice of secondary parameters (Supporting Information, Text S1). For example, if measurements have a Normal distribution, will the standard deviation remain the same if the mean increases or will it also increase? Selection of secondary parameters in power calculations should be based on statistical experience and biological knowledge.

When pairs of surveys are planned, investigators often seek to assess the power to detect given differences between 'groups'. The function *power.groups* calculates the power to detect a difference between means for two groups. These groups could be different areas or the same area at two points in time. For *power.groups*, either parametric or nonparametric randomisation tests can be used to test for differences between means. For parametric tests, if the data are assumed to be Normally distributed, a t-test is used; for the other distributions, likelihood ratio tests are performed based on the relevant Generalized Linear Model.

If a study is planned to detect the impact of changes in pressures or management measures then a Before-After Control-Impact (BACI) design is widely used (Smith 2002). Observations are taken in control and impacted areas before and after a change. The function *power.BACI* calculates power for simple BACI designs. If we define the factor *treatment* which indicates whether an observation is from the control or impact area and a factor *time*, which indicates whether sampling occurred before or after the change then the effect is assessed by the interaction between *treatment* and *time*. The p-value for this interaction is determined using the appropriate Generalized Linear Model for the response distribution. The interaction can also be calculated non-parametrically using a randomisation procedure - *emon* uses unrestricted permutation of the raw data (Anderson and Ter Braak, 2003). An example of the application of *power.BACI* is provided in the Supporting Information (Example S1).

The design, improvement and evaluation of monitoring can often be informed by knowledge of the power to detect changes in a quantity through time. The function *power.trend* is used with *generate.trend* to determine the power to detect trends. Four types of trend for the mean function can be created by *generate.trend*: linear, incident, step and updown with parameters in *generate.trend* defining slopes and change-points. Alternatively, a user may create their own mean function. Within *power.trend*, realisations of the data can be created assuming Normal, Poisson or Negative Binomial probability distributions. Options for analysis of the trend include linear regression for monotonically increasing or decreasing trends, the nonparametric Mann-Kendall method (Mann, 1945; Kendall 1975), with significance determined through a randomisation test, and Generalised Additive Models (Wood, 2006) with time as the dependent variable. These latter models are fitted using the function *gam* in the R-library *mgcv*.

Detecting features

Often, point sample surveys are conducted to find features of interest. These may be, for example, habitats or non-indigenous species. *Emon* considers situations where the feature is (i) in the form of a circular patch and (ii) where it occupies some proportion of the survey area.

For patches, the factors governing the probability of detection are (i) the size of the patch (ii) the number and size of sampling points and (iii) the pattern of sampling points. The function *detect* calculates the probability of detection of a circular patch of specified radius for a specified density of sampling points, the sampling density needed to achieve a specified probability of detection or the radius of the feature that will be detected with specified probability and sampling density. There are options for random, square lattice, and triangular lattice spatial sampling designs. When the feature is not a patch but occupies a proportion θ of the survey area, the function *detect.prop* is used. However, the calculations done by this function are appropriate only when the sampling points are located randomly on the survey area.

The function *detect* is mainly based on the approach of Barry and Nicholson (1993), who give the probabilities that at least one sampling point detects a patch for random, square lattice and triangular lattice designs. These probabilities are a function of the standardised patch radius $R=r(A/N)^{0.5}$, where r is the actual patch radius, N is the sample size and A is the survey area. The only difference is that *detect* uses a more exact formula for the random design of

$$p_r = 1-(1-a/A)^N \tag{1}$$

where a is the patch area. Note that these formulae assume that the sampling unit radius is of negligible size and that sampling units can overlap. However, if the sampling unit radius has some finite value r_s, r is replaced by $r' = r + r_s$.

The approach can be turned around to ask "*which value of N is needed to obtain a specified probability of patch detection?*" For the random design, we rearrange equation (1) to give

$$\hat{N} = \frac{\log(1-p_r)}{\log(1-a/A)} \tag{2}$$

For the other designs, numerical methods are used to calculate $\hat{N}$.

A third approach is to consider the value of r that would be detected with probability p given sample size N. As before, numerical methods must be used for the two lattice designs. For the random design, we get the expression

$$r = \sqrt{A(1-(1-p_r)^{(1/N)})/\pi} \tag{3}$$

For a random sampling pattern, if the object of interest is not a patch, but comprises individuals or a feature which covers some proportion θ of the survey area A, we use the function *detect.prop*. The probability of at least one sampling point detecting the object is:

$$p_{prop} = 1-(1-\theta)^N \tag{4}$$

A similar expression was given by McArdle (1990) and extended into a Bayesian context by Nicholson and Barry (1995). Function *detect.prop* can calculate values for θ and N given those of the other two parameters.

Conclusions

Emon provides easy access to a range of flexible methods to support the design of studies, monitoring or surveys. Our approaches have focused on questions and challenges relating to the marine environment, but some of the methods for assessing power and probabilities of detection may also have relevance to questions encountered when monitoring terrestrial biota. We hope that access to the functions in *emon* will raise awareness of what marine monitoring and surveys can be expected to achieve, thus encouraging cost-effective, needs- and evidence-based designs.

An obvious future extension to *emon* is to build a suite of functions to design surveys based on the desired precision of estimates. Currently this is available only for the simple case of estimating a mean abundance. Other extensions include power for models that allow for temporal and spatial correlations (Elston *et al*. 2011) and/or incorporate the precision of estimates of variance components used in the power simulations (Sims *et al*. 2007). In terms of detecting features, if the sampling points are of some finite area then the probability of detecting the feature will depend on the spatial aggregation of the feature. Theory or simulations to quantify this would be a useful addition to *emon*. Similar work on comparing sampling grab sizes for sampling is reported in Boyd, Barry & Nicholson (2006).

Acknowledgments

This work was supported by Cefas grant 67104DG and the Cefas/JNCC partnership.

Data Accessibility Statement

Data used for the power analysis reported in Example S1 can be found in the Supporting Information file "*R code for examples 1 and 2.R*".

Author contribution statement

JB led the coding and first draft of the manuscript. DM contributed to coding and led on turning functions into a package. JB, DM and SJ developed the project and JM and DW developed examples. All authors contributed to revisions of the manuscript and gave final approval for submission.

References

Anderson, M.J. and Ter Braak, C.J.F. (2003). Permutation tests for multi-factorial analysis of variance. Journal of Computation and Simulation, 73, 85-113.

Barry, J. & Nicholson, M.D. (1993) Measuring the probabilities of patch detection for four spatial sampling schemes. *Journal of Applied Statistics*, 20, 353-362.

Barry, J. & Maxwell, D. (2015) *emon: tools for environmental and ecological survey design and analysis*. https://cran.r-project.org/.

Blanchard, J.L., Maxwell, D.L. & Jennings, S. (2007) Power of monitoring surveys to detect abundance trends in depleted populations: the effects of density-dependent habitat use, patchiness, and climate change. *ICES Journal of Marine Science*, 65, 111-120.

Boyd, S., Barry, J. and Nicholson, M.D. (2006) A comparative study of a 0.1m^2 and 0.25m^2 Hamon grab for sampling macrobenthic fauna from offshore marine gravels. *Journal of the Marine Biological Association of the UK,* 86, 1315-1328.

Cefas (2015) *Developing molecular tools for the identification of non-native species in environmental samples*. Unpublished report on behalf of Defra.

Cohen, J. (1988) *Statistical power analysis for the behavioural sciences*, 2nd edn. Lawrence Erlbaum Associates, London.

Ellis, M. M. ,Ivan, J.S., Tucker, J.M. & Schwartz, M. K. (2015) rSPACE: spatially based power analysis for conservation and ecology. *Methods in Ecology and Evolution*, 6, 621–625.

Elston, D.A., Nevison, I.M., Scott, W.A., Sier, A.R.J. & Morecroft, M.D. (2011) Power calculations for monitoring studies: a case study with alternative models for random variation. *Environmetrics*, 22, 618-625.

Fryer, R.J. & Nicholson. M.D. (1993) The power of a contaminant monitoring programme to detect linear trends and incidents. *ICES Journal of Marine Science*, 50, 161-168.

Fryer, R.J. & Nicholson, M.D. (1999) Using smoothers for comprehensive assessments of contaminant time series in marine biota. *ICES Journal of Marine Science*, 56, 779-790.

Gerrodette, T. (1987) A power analysis for detecting trends. *Ecology*, 68, 1364–1372.

Hellawell, J.M. (1991) Development of a rationale for monitoring *Monitoring for Conservation and Ecology* (ed. F.B. Goldsmith), pp 1-14. Chapman and Hall, London..

Holt, R.S., Gerrodette, T. & Cologne, J.B. (1987) Research vessel survey design for monitoring dolphin abundance in the eastern tropical Pacific. *Fishery Bulletin U.S.*, 85, 435–446.

Johnson, N.L., Kotz, S. and Balakrishnan, N. (1994) *Continuous Univariate Distributions, Volume 1*, 2nd edn. Wiley, New York.

Kendall, M.G. (1975) *Rank correlation methods*, 4th edn. Charles Griffith: London.

Kimura, D.K., & Somerton, D.A. (2006). Review of statistical aspects of survey sampling for marine fisheries. *Reviews in Fisheries Sciences*, 14, 245–283.

Lindenmayer, D.B. & Likens, G.E. (2010) The science and application of ecological monitoring. *Biological Conservation*, 143, 1317-1328.

McArdle, B.H. (1990) When are rare species not there? *Oikos*, 57, 276–277.

Manly, B.F.J. (2006) *Randomization, bootstrap and Monte Carlo methods in biology*, 3rd edn. Chapman and Hall, London.

Mann, H. B. (1945) Non-parametric tests against trend. *Econometrica*, 13, 245-259.

Maxwell, D. & Jennings, S. (2005) Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. *Journal of Applied Ecology*, 42, 25-37.

Nichols, J.D. & Williams, B.K. (2006) *Monitoring for conservation*. Trends in Ecology and Evolution. 21, 668–673.

Nicholson M.D., Fryer, R.J. & Ross, C.A. (1997) Designing monitoring programmes for detecting temporal trends in contaminants in fish and shellfish. *Marine Pollution Bulletin*, 34, 821-826.

Nicholson, M.D. & Jennings, S. (2004) Testing candidate indicators to support ecosystem-based management: the power of monitoring surveys to detect temporal trends in fish community metrics. *ICES Journal of Marine Science*, 61, 35–42.

Nicholson, M.D. & Barry, J. (1995) Inferences from spatial surveys about the presence of an unobserved species, *Oikos*, 72, 74-78.

Nicholson, M.D. & Barry, J. (1996) Survey design for detecting patches. *Journal of Applied Statistics*, 23, 361-368.

Peterman, R.M. (1990) Statistical power analysis can improve fisheries research management. *Canadian Journal of Fisheries and Aquatic Sciences*, 47, 2–15.

Peterman, R.M. & M'Gonigle, M. (1992) Statistical power analysis and the precautionary principle. *Marine Pollution Bulletin*, 24, 231–234.

Shen, H., Brown, L.D. & Zhi, H. (2006) Efficient estimation of log-normal means with application to pharmacokinetic data. *Statistics in Medicine*, 25, 3023-3038.

Sims, M., Elston, D.A., Harris, M.P. & Wanless, S. (2007) Incorporating variance uncertainty into a power analysis of monitoring designs. *Journal of Agricultural, Biological and Environmental Statistics,* 12, 236-249.

Smith, E. P. (2002) BACI design. *Encyclopedia of Environmetrics* (eds A.H. El-Shaarawi & W.W. Piegorsch), pp. 141 –148, Wiley, Chichester.

Steidl, R.J., Hayes, J.P. & Schauber, E. (1997) Statistical power analysis in wildlife research. *Journal of Wildlife Management*, 61, 270–279.

Tidbury, H.J., Taylor, N.G.H., Copp, G.H., Garnacho, E. & Stebbing, P.D. (2016) Predicting and mapping the risk of introduction of marine non-indigenous species into Great Britain and Ireland. *Biological Invasions*, 18, 1–16.

Underwood, A.J. (2000) Importance of experimental design in detecting and measuring stresses in marine populations. *Journal of Aquatic Ecosystem Stress and Recovery*, 7, 3-24.

Underwood, A.J. & Chapman, M.G. (2013). Design and analysis in benthic surveys in environmental sampling. *Methods for the study of marine benthos,* 4th edn. (ed. A. Eleftheriou), pp 1-45, Wiley, Chichester.

Ware, S. & Mcilwaine, P. (2015) *Dogger Bank SCI Monitoring Survey – CEND 10/14 Cruise Report*. Marine Evidence Survey Data, No. 3. Available from: http://jncc.defra.gov.uk/page-7054.

Wood, S.N. (2006) *Generalized Additive Models: An introduction with R*. Chapman and Hall/CRC, London.

Supporting information

File: *Secondary parameters and examples* contains

Text S1: Treatment of secondary parameters in power analysis;

Example S1: Power study for Dogger Bank survey;

Example S2: Detecting non-indigenous species at the Port of Dover.

File: *R code for examples 1 and 2.R*

Gives code to run the examples above.

Table 1. Summary of user-manipulated functions in *emon* described in this paper.

Function	**Purpose**	**References**
power.BACI	Tests power of BACI designs to detect differences between a control and a treatment.	Manly (2006); Shen, Brown & Zhi (2006)
permute.BACI	Non-parametric randomisation test of treatment effect in BACI design.	Manly (2006)
power.groups	Calculates power by simulation for comparing the mean of two groups of independent observations.	Fryer & Nicholson (1993,1999)
permute.groups	Non-parametric randomisation test for	Manly (2006)

	comparing the means of two groups.	
power.trend	Calculates power by simulation for a specified trend. The time and mean value of the trend can use input from *generate.trend*. The distribution of the data is created by *addnoise*. The user can specify the statistical method to detect the trend to be either linear regression, the Mann-Kendall statistic or a Generalised Additive Model.	Fryer & Nicholson (1993, 1999); Wood (2006).
generate.trend	Generates one of four scenarios for the change in mean value as function of time. Results can be used in *power.trend*.	
addnoise	Used in *power.trend* to add random variation to the mean values. The distribution, with mean given by the mean values, can be either Normal, Poisson or Negative Binomial.	
mannkendall	Calculates the Mann-Kendall statistic for a monotonic trend and the p-value by simulation against the null hypothesis of no trend.	Mann (1945); Kendall (1975)
detect	Calculates either the: (i) probability of detection of a circular patch of specified radius for a specified density of points; (ii) density needed to achieve a specified probability of detection; or (iii) radius of the patch that will be detected with specified probability and sampling density. The user can choose either random, square lattice, and triangular lattice spatial sampling designs.	Barry & Nicholson (1993)
detect.prop	Calculate the probability of detecting: (i) a	McArdle (1990)

	feature that occupies a proportion θ of the sampling area and where the number of sampling points is specified; (ii) the number of sampling points needed to achieve a specified probability of detection, where θ is also specified; or (iii) the value of θ that will be detected with specified probability and sampling density.	