**Supplementary Information**

**Text S1: Treatment of secondary parameters in power analysis.**

The first consideration when designing an experiment, survey or monitoring based on power is quantifying the magnitude and/or rate of change to detect. But secondary parameters used in power calculations also need to be realistic. For example, if measurements have a Normal distribution, will the SD remain the same if the mean increases or will it also increase? Selection of secondary parameters is based on statistical experience and biological knowledge, and valid power calculations depend on appropriate choices.

In *emon*, secondary parameters are treated as follows. For the Normal distribution, generally, our approach has been to assume that the standard deviation remains the same for any future change in mean. But, *emon* does give users the choice to specify any value for the future standard deviation. For the Poisson distribution, the mean equals the variance. Thus, changing the mean automatically changes the standard deviation. For the negative binomial distribution, the mean *μ*, variance *σ2* and size *k* are linked by

 (S1)

Thus, an increase in the mean increases the variance, when size remains constant. However, it is possible to keep the variance constant by changing the size. Thus, options when choosing secondary parameters with a change in mean are (i) keep the variance constant but change the size; (ii) keep the size constant but change the variance. Option (ii) is adopted in *emon* if the size for time 2 is the same as the size for time 1. However, to achieve a constant variance between two successive time points or two areas, the second size is chosen such that

 (S2)

Because the size parameter must be positive, a constant variance can be achieved only when the denominator of (S2) is positive.

The Lognormal distribution is defined such that the random variable X has a Lognormal distribution if Y=ln(X) has a Normal distribution. The means and variances of the Normal and Lognormal distributions are linked by the standard formulae (e.g. Johnson, Kotz & Balakrishnan 1994):

 (S3)

 (S4)

where *μ* and *σ* are the mean and standard deviation of *Y,* and *m* and *V* are the corresponding mean and standard deviation of *X*.

To explain how *emon* interprets these mean and variance parameters, we will use the *power.groups* function (Table 1, main paper) as an example. The values for the *emon* parameter pars1 are μ1 and σ1, the mean and standard deviation of Y at time 1. Within *emon*, equation (S3) is used to generate the mean *m1* on the original scale from pars1. The value of *m2* is then generated from *m1* with the additive or multiplicative change specified in the function. The entry in pars2 specifies the standard deviation *σ2*. *Emon* calculates *μ2* using a rearrangement of equation (S3) to give *µ2 = ln(m2) - σ22/2*.

The values of *μ1*, *σ1*, *μ2* and *σ2* are then used to simulate the data for the power analysis. This process makes an implicit assumption about the value of *V2*. Using equation (S4), if we assume that *σ1=σ2*:

 (S5)

Thus, by increasing the mean from *μ1* to *μ2*, we also increase the variance by a factor of

*exp[2(µ2-μ1)]*.

**Example S1: Power study for Dogger Bank survey**

This example describes how *emon* was used to design a survey with a specified probability of detecting chosen changes in the species richness of the benthic community on the Dogger Bank prior to the introduction of management measures.

In 2014 the Joint Nature Conservation Committee (JNCC) and the Centre for Environment, Fisheries and Aquaculture Science (Cefas) conducted a joint monitoring survey at the Dogger Bank Site of Community Importance (SCI; a conservation designation) to test alternate sampling designs for detecting meaningful change following the implementation of management measures to stop some types of demersal (bottom) fishing on some parts of the Dogger Bank (Ware and McIlwaine 2015). One of the tested concepts was Investigative Monitoring, where an experimental approach is used to determine the effectiveness of a management measure. BACI experiments were designed at four (A, B, C and D) proposed demersal fishery closed areas within the Dogger Bank SCI (Figure S1) and biological samples (infaunal and sediment samples, seabed imagery and 2m beam trawl samples) were collected inside and outside of the boundaries of the proposed closed areas.

Separate power analyses were carried out for each of the areas – so, effectively, the study was stratified into relatively homogenous areas. For the purposes of this example, Area D was selected to illustrate the way *emon* is used to perform the power study. One reason for this was that 41 sandy sediment samples, which were collected in 2012 by Forewind (in support of their Environmental Impact Assessment for the proposed Dogger Bank R3 windfarm development), were available to estimate the parameters of the appropriate probability distribution to calculate the power. The example illustrates a situation in which the choice of probability distribution can have a large effect on the power.

As noted earlier, the power calculations in e*mon* assume that points are independent of one another. Semi-variograms calculated on the Forewind data didn’t show any obvious signs of spatial dependence as a function of separation distance between the points.

For species richness, the Negative Binomial distribution was the most like the estimated density function generated from the Forewind data (Fig. S2a) and so this was chosen for the power calculations. However, to illustrate the effects of using the ‘wrong’ distribution, power was also calculated for the Poisson distribution. This calculation was intended to highlight the risk of selecting an inappropriate distribution when data used to help plan the study, such as those collected by Forewind in this example, are not available.

There are two options when postulating the form of the negative binomial distribution at time 2 when the mean *μ* has been subject to change from time 1 to time 2. As described above, the first option is to keep the size of the distribution the same. The second is to keep the variance the same by changing the size (implemented in the *emon* function *size2.samevar*). For the 50% change, both options are shown in Figure S2(b). The ‘same variance’ (variance=171) density has a similar shape at times 1 and 2, apart from the shift in the mean. However, the ‘same size’ density is much wider, reflecting its increased variance of 366.3. For the power analysis, we chose to use the ‘same variance’ option.

Figures S2(c) and (d) show power, based on 1000 simulations, of detecting 10% to 50% changes in the mean species richness in the closed area. If the poorly-fitting Poisson distribution were selected (c), it would lead to higher powers than for the Negative Binomial distribution (d), which provided a better fit to the Forewind data.

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

This example describes how *emon* was used to design a sampling scheme with a specified probability of detecting a relatively rare non-indigenous species.

Some non-indigenous species (NIS) pose a serious risk to indigenous species and may lead to direct economic costs by impacting ecosystem services and indirect costs by generating remediation requirements. In the marine environment, where NIS can be very difficult to detect and therefore eradicate, the impact of NIS is often irreversible. A paper by Tidbury et al. (2016) showed that the area surrounding the Port of Dover was at the highest risk of introductions of NIS in England.

Current legislation, (including the Water Framework Directive and the Marine Strategy Framework Directive) and future legislation (including the Ballast Water Convention and invasive species regulations) for controlling NIS focuses on the detection of high impact species. In a recent report (Cefas 2015), NIS of importance to UK waters were prioritised for development of new, molecular-based detection methods. Species specific PCR assays were developed to allow detection of the high priority species from mixed samples. The top ten species were: *Didemnum vexillum, Styesela clava, Crepidula fornicata, Crassostrea gigas, Ficopomatus enigmaticus, Bugula neritina, Botrylloides violaceus, Tricellaria ipinata, Corella eumyesota, Grateloupia turuturu*. These all have both a pelagic and benthic lifestage. Thus, to sample these species, we need to consider survey designs for sampling in the water column and the benthic environment (in their benthic phases these species are sessile and so likely to form patches). We consider sampling in the water column first, where we assume that individuals have a random distribution.

In the case of detecting the pelagic phases, we note that the Port of Dover covers an area of approximately 2,265,000 m² and has a mean depth of 6.2 m at mean low tide and so it has an approximate volume of V=1.4043x1010 l at mean low tide. The volume of the sample is v=10 l.

Using *detect.prop*, Figure S3 (Top) shows the number of individuals (N) per litre needed for the probability that an individual is detected to be equal to 0.9. This is done by redefining the parameter *θ* in equation (4) as *θ* = *nv/V*, where *n* is the number of individuals, *v* is the volume of the sampling unit and *V* is the volume of the survey area. Note that this approach assumes that no more than one individual is captured in a sampling unit. The function *detect.prop* calculates from which  for the values of *N*, *pprop*, *v* and *V* specified above. For example, if sample size is *N*=30 we would need to have at least 0.007 individuals of a given NIS per litre to have a 90% chance of detecting the species.

In the case of detecting the benthic phase we want to estimate the patch size (radius) of a given NIS which we can be 90% sure of detecting as a function of sample size in a triangular design. This can be calculated using the *detect* function in *emon* by specifying the *statistic=”R”* argument. However, in this instance, because the sampling units have radius 0.25 m, we need to reduce the answer given by *detect* by 0.25 m. This is because the theory on which *detect* is based assumes that the sampling units are points of negligible size. If the sampling units have some finite size, they are more likely to detect patches. Figure S3 (Bottom) shows the required patch radius as a function of sample size. For example, the dashed lines illustrate that with a sample size of 100, a patch with radius 76.6 m will be detected with probability 0.9 for the triangular design. Of course, if habitat preferences of the benthic phase of a given NIS are known and habitat distributions in the search area can be described, then the calculations can be conducted for a smaller search area.



Figure S1: Location of the proposed closure areas and position of biological sampling stations in area D within the Dogger Bank SCI site.



Figure S2: Power study for the Dogger Bank Example S1: (a) Fitted probability density functions for the richness data collected by Forewind; (b) potential Negative Binomial distributions for the power simulations: Time 2 (i) = same size as for Time 1; and Time 2 (ii) = same variance as for Time 1; (c) power plots for Poisson and (d) Negative Binomial, with the same variance at Times 1 and 2; The lines on each plot in (c) and (d) (from bottom to top) represent the power to detect increases in species richness of 10%, 20%, 30%, 40% and 50% between time 1 and 2 in the closed area.



Figure S3: Plots for Port of Dover Examples S2. Top: Number of individuals per litre needed for the probability that an individual is detected = 0.9. Bottom: Patch radius needed to be 90% sure of detecting a patch as a function of sample size.