**SUPPORTING INFORMATION**

This document provides supporting information on the design and implementation of the simulation study.

**Appendix S2: Simulation study**

To evaluate our modelling framework and quantify differences in precision and coverage probability between the CSDM method and the BHSDM method we used a simple simulation study. We simulated spatial variation in abundance over 300 hypothetical grid cells that were all 100 km2 in area. All grid cells were surveyed where the amount of trackline effort varied between 1 and 10 km. The search half-width was set to 5 km so the maximum amount of area covered in a grid cell was 100 km2 (i.e. the entire grid cell). Each grid cell was randomly assigned a habitat covariate value that was simulated from a standard normal distribution. To simulate variation in abundance we assumed a quadratic relationship between the covariate and the mean number of animals in each grid cell. Given this mean value, the actual number of whales in each grid cell was simulated from a Poisson distribution (the limiting form of the Tweedie distribution with scale parameter 1.0). Within each grid cell we assumed whales were homogeneously distributed, therefore, the probability of being within range of the search vessel was directly proportional to the observation effort in that grid cell.

We simulated both surface detectability and surface availability. For surface detectability, we simulated two team observer data using a hazard rate function with separate probabilities of detection on the trackline (i.e. g(0)) for each team. Each whale was randomly assigned a distance between 0 and 5 km. Group size was set to 1 such that each group consisted of a single animal. For surface availability, we set the true probability of being at the surface to 0.5. However, availability is never known perfectly, and we emulated scientific uncertainty about its true value by simulating an estimate of surface availability from a beta distribution with a coefficient of variation (CV) of 0.05. For the CSDM method we used this estimate as a fixed measure of surface availability. For the BHSDM method we used an informed beta prior with the mean set to the simulated estimate of surface availability and the CV set to 0.05.

We applied both the CSDM method and the BHSDM method to each of 1000 independently simulated datasets. For each simulation, we estimated population size by summing up the estimated number of animals in each grid cell. For the BHSDM method this estimate was calculated by calculating the mean of the posterior distribution for population size. We estimated coverage by determining whether or not the true mean population size fell within the 95% confidence intervals of the CSDM method or the 95% credible region of the posterior estimate of the BHSDM method. For the CSDM we calculated lognormal confidence intervals and for the BHSDM 95% credible intervals were calculated directly from the posterior distribution using the upper 0.975 and lower 0.025 quantiles. In addition to coverage, we calculated percent relative bias for each simulation and method as (-)/i\*100 where is the true mean population size for simulation i and is the estimate of population size for simulation *i* and method *j*.