Coming Soon in **Distance**...

Laura Marshall, Dave L Miller, Len Thomas

Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, Fife KY16 9LZ, Scotland.

Abstract

Distance sampling[1] is the most widely used method for obtaining estimates of animal density and population size worldwide. **Distance** is purpose-built software for the design and analysis of distance sampling surveys. Currently we are implementing a number of enhancements to deal with more complex real-world situations. We highlight the following in this poster:

- Improved spatial density surface estimation, including complex survey areas
- Inclusion of double observer analyses for point transects
- Implementation of a multi-analysis engine for line transect analyses. It will be able to include uncertainty in the following: covariate values, detection function form and species identification
- Development of reliable monotonic detection functions via a mixture model formulation.
- Implementation of a simulation engine to allow realistic survey planning and optimization

Distance

Distance provides an easy to use Windows graphical user interface to the underlying Fortran and R code which carries out the analyses. There is extensive documentation associated with **Distance** and workshops (both introductory and advanced) are held regularly.

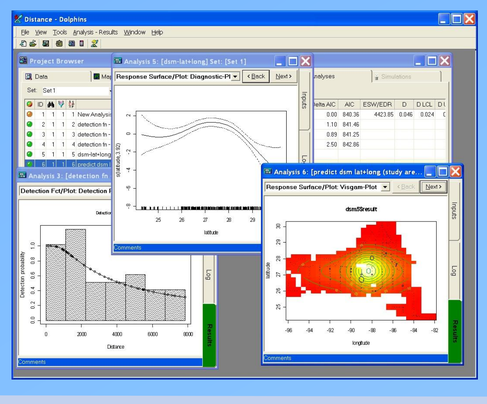
Availability

The latest public release of Distance is version 6 which to date has been downloaded over 8500 times by users in 120 different countries. We expect version 7 to be released 2013, look out for the 7.0 below which indicates the new additions. Other method developments are already available in R and may also be included in version 7, see Available Now in R!. Distance is available free of charge from http://www.ruwpa.st-and.ac.uk/distance/

Basic Distance Sampling Principles

A set of related methods used to estimate the abundance and/or density of marine mammals. The basics steps involve:

- Survey at a set of transects laid out at random within the study area.
- For each sighting, record the distance from the transect to the animal.
- Use the distribution of distances to estimate the probability of detection given distance ("detection function").
- Using the detection function, estimate the average detection probability, and use this to scale up the number of animals detected to the number of animals present.



Acknowledgements

Distance has been under development for over 30 years . Major contributors Include: D.R. Anderson, K.P. Burnham, J.R.B. Bishop, D.L. Borchers, S.T. Buckland, S.L. Hedley, J.L. Laake, F.F.C. Marques, T.A. Marques, J.H. Pollard, E. Rexstad, S. Strinberg, L. Thomas,

Funding for software and methods development has been provided by:

Colorado Cooperative Fish & Wildlife Research Unit
National Marine Mammal Laboratory (U.S.)
Wildlife Conservation Society
Office of Naval Research (U.S.)
National Park Service (U.S.)
Engineering and physical sciences research council (U.K.)
Biotechnology and biological sciences research council (U.K.)
GeoMarine Inc.
Fisheries and Oceans Canada

Distance Developments in More Detail

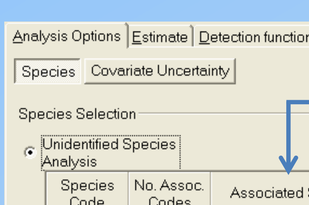


Figure: Soap film smoothing.

Spatial Density Surface Estimation 7.0

Recent developments in spatial smoothing will be incorporated into **Distance**. These include: the soap film smoother[2], which offers more reliable inference in complex topography; and improved uncertainty estimation using both bootstrap and variance propagation methods. For more information on this and other spatial density advances see our other poster: "Spatial density surface estimation from distance sampling surveys".

Double Observer Point Transect Analyses

Available Now in R!

Methods to analyse continuous double observer point transect data are now available in the `mrds R` library (http://cran.r-project.org/web/packages/mrds). Methods to analyse binned double observer point transect data[3] are available in the `mrpt R` library (https://github.com/jlaake/mrpt).

Multi-Analysis Engine 7.0

*We are currently implementing methods for dealing with complications in real-world surveys. These improvements will make the analysis of major surveys, such as the Eastern Tropical Pacific[4] NOAA surveys, possible in **Distance**.*

The new multi-analysis engine in **Distance** will deal with the following:

- Multi-species detection functions with species-specific density estimation.
- Uncertainty in cluster size estimation and other covariates, including multipliers
- Uncertain species identification in multi-species surveys
- Uncertainty in detection function

Observers may not be able to accurately determine species for some sightings. Instead they may record something like small cetacean or large dolphin, for example. The multi-analysis engine will estimate the abundance of these species groups (unidentified sightings) and prorate these to the associated species codes specified in this table.

Uncertainty may be added to covariates in any data layer using any one of a selection of sampling distributions. These include Normal, bias corrected Lognormal, bias corrected Poisson and Poisson.

Correction factors may also be applied to covariates. These may either be numeric values or variables in a specified data layer.

Model uncertainty is incorporated by selecting multiple models for each species code. Following each data resample in the bootstrap, the model with the minimum selection criterion (AIC, AICc or BIC) will be selected.

Monotonic Detection Functions using Mixture Models

Available Now in R!

*We aim to reduce the need for constrained optimisation in **Distance** through the use of mixture models to estimate the detections function.*

Detection functions which include adjustment terms can lead to fitting of non-monotonic detection functions - this means we have to constrain the optimisation to avoid unrealistic models. Miller and Thomas[5] present a new class of models for the detection function based on finite mixtures of simple parametric key functions such as the half-normal. Similar to the detection functions currently implemented in Distance, they are flexible, produce plausible shapes with a small number of parameters and allow incorporation of covariates in addition to distance. However, mixture model detection functions are by definition monotonic non-increasing and non-negative so no constraints are needed. Download the `mmds R` library from http://cran.r-project.org/web/packages/mmds.

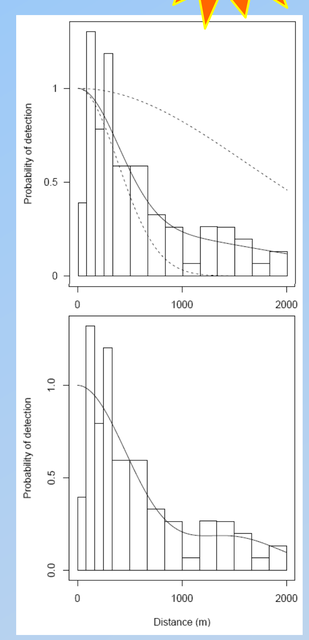
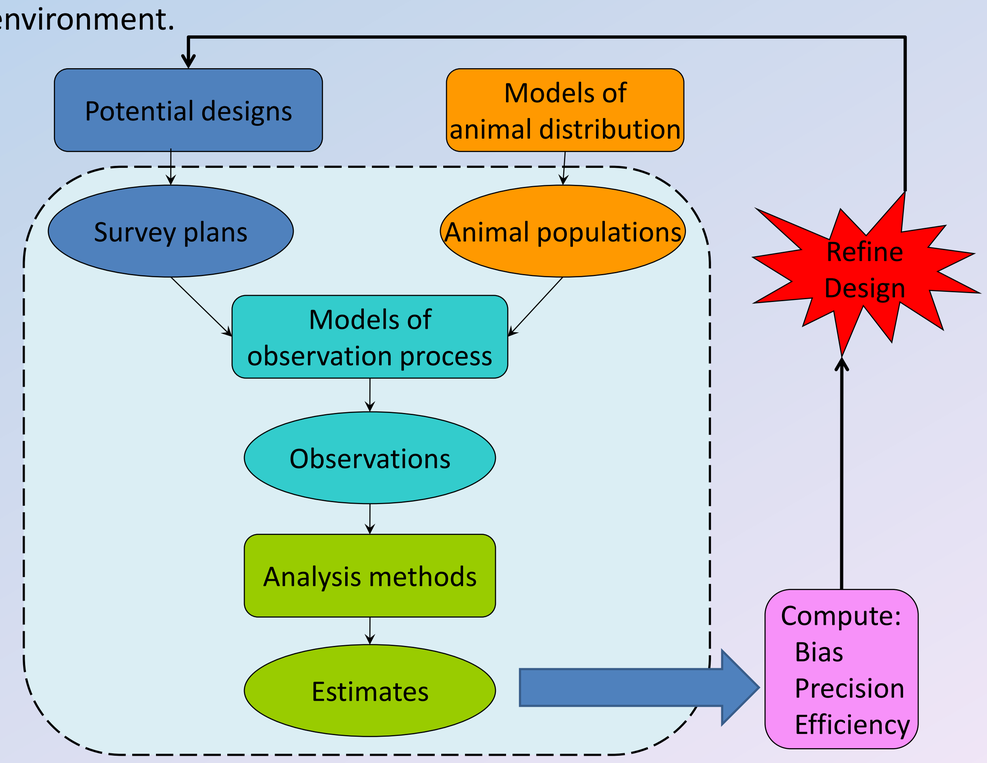


Figure: Humpback whale line transect data. Top panel: Monotonic mixture model detection function. Lower panel: non-monotonic conventional distance sampling detection function.

Simulation Engine 7.0

We aim to develop a simulation engine to allow realistic survey planning and survey design optimization.

This engine will allow different designs to be tested, allowing the user to specify the number and distribution of animals, as well as the type of survey to be performed and the amount of survey effort available. The engine will simulate many repeat surveys, and then compute bias and precision of the resulting estimates. Many different designs and levels of effort can be trialled, and an optimal design found that is also robust to different assumptions about animal distribution. Simulation helps improve efficiency of surveys by allowing the options to be trialled in a low cost environment.



References:

[1] Thomas, L., Buckland, S.T., Rexstad, E.A., Laake, J.L., Strindberg, S., Hedley, S.L., Bishop, J.R.B., Marques, T.A. and Burnham, K.P. (2010) Distance software: design and analysis of distance sampling surveys for estimating population size. *Journal of Applied Ecology* ,47: 5-14.

[2] Wood, S. N., Bravington, M.V. and Hedley, S.L. (2008) Soap film smoothing. *Journal of the Royal Statistical Society B*, 291:1-21.

[3] Laake, J. L., Collier, B. A., Morrison, M. L., & Wilkins, R. N. (2011).Point-based Mark-Recapture Distance Sampling. *Journal of Agricultural Biological and Environmental Statistics*, *16*:389-408.

[4] Gerrodette, T. and Forcada, J. (2005) Non-recovery of two spotted and spinner dolphin populations in the eastern tropical Pacific Ocean. Marine Ecology Progress Series, 291:1-21.

[5] Miller, D.L. and Thomas, L. (*In Prep*) Mixture models for distance sampling detection functions