Use of SCANS model-predicted density surfaces

Spatial scale of data and predictions

Density surface models (DSM) are fitted to data processed at a given spatial scale. Typically, the data may be processed into segments of continuous searching effort or the total searching effort in a grid cell, with associated sightings of animals in each case. In the SCANS models, the data are segments of target length 10km. The actual length of each segment may vary, depending on ensuring continuous searching effort in the same sighting conditions (e.g. Beaufort scale, swell), but the large majority are around 10km. The selected final models fitted to these data are predicted onto a spatial grid, also at a given scale. In the SCANS modelling, the prediction grid cells are 10x10km. Thus, there is no information in the predicted density surfaces at a spatial scale of less than 10km.

In SCANS, the density surface models are typically fitted to data collected along transect lines that cover the entire study area. The resulting modelled relationships between observed density and the explanatory environmental covariates are thus an "average" over the whole study area. How accurate the prediction is at any particular location thus depends on spatial variability in the true relationship between density and the environmental covariates. In SCANS, the entire study area encompasses a range of habitats and thus predicted density at any particular location should not be over-interpreted. Based on an analysis of residuals from modelling JCP data, Paxton et al. (2016) suggested a minimum area of 500-1,000 km^2 over which such predictions may be useful but cautioned that inferences will be less reliable in areas that are not well informed by data. In reality, the area over which it is justified to draw inferences about predicted density may be much larger. The bottom line is that predictions from density surface models are intended to be considered at a relatively large spatial scale, not at a fine scale.

Model uncertainty and prediction uncertainty

The process of fitting DSMs to data involves multiple decisions. Aspects relevant to modelling SCANS data include: the form of the model (e.g. GAM); whether to model individuals or groups; how to account for overdispersion in the data; how to deal with any auto-correlation in the data; which environmental variables to include as candidate explanatory covariates (e.g. cartesian, fixed such as seabed depth or slope, dynamic such as sea surface temperature or chlorophyll a); how to deal with correlated variables; how to determine the best model (e.g. REML, AIC) for the purposes of prediction. Decisions made about these aspects affect predicted density surfaces. Changing any decision will result in a different density surface.

Given a selected model, uncertainty in the predicted density surface can be estimated, for example using bootstrap resampling of estimated model coefficients, as a spatial surface of, for example, SE, CV or 95% confidence limits. It may be possible to "model average" predictions from two or more models to incorporate some model uncertainty but this may not be feasible or appropriate, depending on decisions made in fitting models. In any case, estimates of uncertainty in predicted density are highly unlikely to incorporate all the uncertainty in model fitting.

Reference

Paxton *et al.* (2016). Revised phase III data analysis of Joint Cetacean Protocol data resources. JNCC Report No. 517. http://data.jncc.gov.uk/data/01adfabd-e75f-48ba-9643-2d594983201e/JNCC-Report-517-FINAL-WEB.pdf