DenMod: Working group for the advancement of marine species density surface modelling

Report of the 3rd working group meeting and 2nd public workshop

5-7 December 2019

Third working group meeting.

Participants: Jason Roberts, Ana Canadas, Len Thomas, Dave Miller, Phil Bouchet, Catriona Harris, Karin Forney, Elizabeth Becker, Jeff Moore, Jay ver Hoef, Megan Ferguson, Debi Palka, Sam Chavez, Doug Sigourney, Anu Kumar. Working group members unable to attend were Paul Conn, Brett McClintock, Devin Johnson and Pat Halpin.

Summary

This section briefly reports on the third meeting of the US Navy-sponsored DenMod working group, held on 5-6 December 2019 prior to the World Marine Mammal Conference in Barcelona. The overall aim of the working group is to develop and implement innovative approaches for advancing spatial modelling methods to best characterize the seasonal abundance and distribution of marine species, focusing on marine mammals. In this third meeting we shared updates on the year 1 and 2 tasks and agreed the project priorities for years 3 and 4.

Background and objectives of the working group

The Navy has a pressing need for reliable estimates of density for many marine species, and how these densities vary in space and time. Such information is central to estimates of “takes” arising from testing and training activities, and for long-term population monitoring; it can also be used when planning operations, where appropriate, to minimize expected impact.

Many approaches have been proposed and applied to density surface models used by the US Navy (and more widely in the ecological research community). Four key points are: (1) different approaches can lead to very divergent estimates; (2) the relative merits of many methods are at

present largely unclear; (3) many issues that make modelling these data complex remain unresolved; and (4) there are several approaches that remain untested in this context. Given this, the reliability of current estimates is somewhat open to question.

We have created a working group of scientists involved in density surface modelling to coordinate advances on this topic. The group is supported by dedicated post-doctoral staff, tasked with solving issues identified by the group. Specific goals of the group are:
1. Bring together scientific leaders in density surface modelling.
2. Share information about best practices, without being prescriptive.
3. Develop and implement innovative approaches for advancing spatial modelling methods to best characterize seasonal abundance and distribution of marine species, focused on US Navy training and testing areas.
4. Provide user-friendly no-cost tools, where possible, implementing new approaches.
5. Provide statistical support to those tasked with undertaking density surface modelling for the US Navy's Phase IV analysis.
6. Provide accessible guidance for practitioners in the form of public reports or scientific publications.
7. Highlight priority areas for continued research.
8. Solicit input from the wider scientific community and share findings through public workshops.

The project represents a collaboration between the University of St Andrews, Duke University and the regional NOAA Fisheries Science Centers that are largely responsible for collection and analysis of line transect visual survey data used in Navy impact assessments.

The group hold annual face-to-face working group meetings to share information, establish and update priorities, discuss potential solutions, and receive feedback on solutions implemented.

<u>Overview of meeting</u>

The first day consisted of a series of updates on each of the seven topics that the group have been working on during the first two years of the project: uncertainty estimation; extrapolation; model unification; workflow (e.g., documenting the end-to-end workflow to ease analysis and enhance repeatability); visual and acoustic data integration; data integration; pinnipeds.

Key updates on each of these topics include:

Uncertainty estimation

- R tool developed for variance propagation (varprop), and available within R package dsm (density surface modelling).
- Paper submitted: Bravington, MV, DL Miller and S Hedley. Estimating variance in density surface models. Preprint: https://arxiv.org/abs/1807.07996

Extrapolation

- A report on extrapolation in density surface models is now available online (in open-access form): Bouchet, PJ, Miller, DL, Roberts, J, Mannocci, L, Harris, CM & Thomas, L (2019) From here and now to there and then: practical recommendations for extrapolating cetacean density surface models to novel conditions. CREEM Technical Report, no. 2019-1, University of St Andrews, 59 p. http://hdl.handle.net/10023/18509
- The report is accompanied by a dedicated R package (dsmextra), which provides a toolkit for quantifying and visualising extrapolation in density surface models (as implemented in package dsm) projected into novel environmental space. The package can be downloaded freely from Github at https://densitymodelling.github.io/dsmextra/

Model unification

- Completed formal (mathematical) comparison of multiple spatial modelling approaches.
- 2 publications:
    - Miller, D. L., Glennie, R., Seaton, A. E. (2019) Understanding the Stochastic Partial Differential Equation Approach to Smoothing. Journal of Agricultural, Biological, and Environmental Statistics. Online early: https://doi.org/10.1007/s13253-019-00377-z
    - Pedersen, EJ, DL Miller, GL Simpson, N Ross (2019) Hierarchical generalized additive models in ecology: an introduction with mgcv. PeerJ https://peerj.com/articles/6876/

Workflow

- A workflow wiki was made publicly available prior to the working group meeting. The wiki documents the workflow for production of Navy density surfaces along with FAQs. It is a citeable resource with a DOI (https://osf.io/5eza8/wiki/).

Visual and acoustic data integration

- A list of possible case study datasets where there are spatially and temporally overlapping visual and acoustic data have been compiled and assessed for suitability.
- No case studies investigated thus far have been found to be ideal (for various reasons). More datasets are being investigated.

Data integration

- The primary focus thus far has been on the integration of telemetry data into density surface modelling of pinnipeds.
- A technical group has been set up to consider data integration more broadly, but the main effort under this topic will occur in year 4.

Pinnipeds

- Working group members from Alaska Fisheries Science Center are working to integrate pinniped telemetry data, aerial surveys, rookery counts and opportunistic data into a density surface model.

The second day was spent discussing priority research topics for years 3 and 4 of the DenMod project. A number of the topics listed above, such as visual and acoustic data integration, will continue for another year. In addition, new topics for year 3 will include model checking and validation, and the development of tools for density surface modellers. Topics that will commence in year 4 include data integration and methods for dealing with group size.

Second public workshop

Summary

This section briefly reports on the second public workshop of the US Navy-sponsored DenMod working group, held on 7 December 2019 as a pre-conference workshop at the World Marine Mammal Conference, Barcelona. One important goal of the DenMod project is to "solicit input from the wider scientific community and share findings through public workshops." We plan to achieve this through holding three public workshops over the course of the project. This second workshop provided an opportunity for stakeholders to provide feedback on progress thus far and the proposed priorities for the remainder of the project. There were 91 registered participants, including members of the working group.

Overview of meeting

The project team presented an overview of density surface modelling in general, followed by details of the DenMod project goals and objectives. These introductory presentations were followed by five presentations given by working group members. Four presentations outlined progress over the last four years, while the last presentation presented the plan for years 3 and 4. Summaries of each of the five main presentations are given below.

Len Thomas – "*Summary of main focus areas in years 1 and 2 (workflow, uncertainty, extrapolation, model unification, acoustics, data integration, pinnipeds)*"
This presentation provided an overview of progress on each of the seven topics that were the focus of efforts in years 1 and 2 of the project. The key updates reflect those included above under the summary of the working group meeting.

Jason Roberts – "Density surface modelling workflow wiki"
The wiki and FAQ, which documents the workflow for producing Navy density surfaces, was made publicly available prior to the meeting (https://osf.io/5eza8/wiki/). This presentation provided an overview of the scope of the wiki and walked participants through some example FAQs, such as "How should I validate my model predictions?".

Phil Bouchet – "*Extrapolation*"
A technical report on extrapolation was made publicly available prior to the meeting (Bouchet et al. http://hdl.handle.net/10023/18509). This presentation provided an overview of the contents of this report which synthesizes the ecological literature on extrapolation metrics and provides recommendations for diagnosing extrapolation within the context of cetacean studies. The

presentation also included details of the dsmextra package (https://densitymodelling.github.io/dsmextra/) which provides a toolkit for quantifying and visualizing extrapolation in density surface models (as implemented in package dsm) projected into novel environmental space. The idea behind dsmextra is to aid ecologists, practitioners, and model end-users in identifying conditions (e.g. areas) under which predicted density surfaces may be prone to errors.

Dave Miller – "*Progress on uncertainty estimation*"
This presentation focused on propagating uncertainty through the stages of density surface modelling. An R tool has been developed for variance propagation (varprop), which is available within R package dsm, and is described in a submitted paper (Bravington et al. https://arxiv.org/abs/1807.07996). This presentation discussed the approaches taken by different providers of density surfaces to the US Navy and discussed how to unify these approaches and how to integrate variance propagation.

Dave Miller – "*Overview of project plan for year 3*."
This presentation outlined the plan for year 3 and also for year 4 of the project. Current tasks which will continue into year 3 include uncertainty estimation, visual and acoustic data integration, and pinnipeds. New tasks for year 3 include model checking and validation, and the development of tools for dsm package enhancement, particularly for complex detection scenarios. New tasks for year 4 include addressing group size issues and data integration.

There followed an hour long open discussion session to provide stakeholders with the opportunity to ask questions and provide feedback on progress and proposed future priorities.

The following points were raised and discussed by participants:

- What can be done with uncertainty propagation when there are multiple detection functions?
    - This is an area of active research and is relatively straightforward in principle, though operationalizing this in code (and making such an analysis simple in the dsm package) is more complicated due to several bits of internal book keeping. We hope to complete this work in year 3 with an update to the dsm package and a tutorial.

- Spatial autocorrelation, uncertainty and incorporating dynamic process into density surface models.
    - Spatial autocorrelation is a big issue to deal with. Firstly it is important to have good diagnostics to identify situations where one should even spend time exploring autocorrelated residual issues. The team noted that there are models that can include a spatial autocorrelation structure, such as CAR models. These models can be hard to fit in standard software and therefore may require a move into another computational framework, such as JAGS or TMB (therefore requiring bespoke code). It is difficult to incorporate a general solution into the current structure of the dsm package.

    - In terms of population dynamics, it is recognized that populations are not static. People already acknowledge this when they make predictions at multiple time points, or by taking appropriate averages at particular times. At another level there is interest in incorporating information available on health or population dynamics explicitly. The working group are interested in look at this, it is likely to be complex but not impossible, but will require bespoke code.

- What does one do when the best model and the next best model give very different estimates of abundance?
    - The team acknowledged that this is not an uncommon issue. They advised against model averaging. Within mgcv and dsm, there are ways to fit density surface models with all of the covariates of interest included while at the same time shrinking those that do not matter. This method incorporates model structural uncertainty, but it was noted that this may result in a model with large uncertainty.
    - It was noted that the workflow wiki contains information about model selection and will be updated as more tools become available on this topic over the coming year.
    - There should be careful consideration about the goal of the model – understanding the system, or prediction.
    - It was advised that practitioners evaluate extrapolation as it may be that the model is predicting across conditions that were not sampled, particularly when dynamic covariates are being used. This could result in large uncertainty in some places.

- Can models be used to differentiate between different states of behaviour?
    - Models haven't typically been used to describe distributions of behaviour, but rather distribution of animals. However we could consider the situation where we might model animals on a calving ground versus animals on a foraging ground. We could go further than this if behavioural covariates are available. Predicting feeding ground vs breeding ground hotspots could be a candidate for inclusion of a factor-smooth interaction in the model. It was noted that this would only indicate correlation and not causation.

- Lessons from other taxa
    - The work presented at the public meeting was taxonomically biased towards marine mammals because of project funding and therefore participants questioned whether the group were learning from developments relating to other species, e.g. seabirds or fish stocks.
    - The working group members from the University of St Andrews in particular have experience working on many different species including seabirds, turtles, and various terrestrial animals. The efforts within the DenMod project are trying to synthesize the research developments across species, not just within the marine mammal sphere.

    - There are many methodological links between studies on different species, and the group are aware of the links with methods such as GEEs and spatial point processes, however it was acknowledged that there is a gap in our knowledge about fishery stock assessment where the survey methods can be quite different.

- Uncertainty related to availability bias.
    - This topic had been discussed at the Distance Sampling workshop held earlier the same day. There are basically three kinds of methods to address this. Aerial line transect is the easiest as the plane is faster than the animals. Focal follows and tagging is another approach to take. Another method is to try to collect data simultaneously with the survey, e.g., observing in the area long enough that animals change their availability. There is often information in the distance sampling data that can be helpful, e.g. forward distance.
    - There are also cases where people have built completely different models for availability, such as GAM-based temperature models for turtle availability.
    - It is hoped that by the end of 2020 we will have methods that will propagate this uncertainty into the final model estimates.

- Mark-recapture distance sampling
    - Propagating the uncertainty from this method would be really useful. It is a high priority for the project and statistically not too hard, so should be achievable in 2020.

- Methods to disentangle effect of season/time and temperature.
    - It was advised to determine any confounding between these covariates as there is no way to tease them apart if the warmer part of the study site is sampled at one time of year, and the colder part at another time of year.
    - There may be complex combinations of space and time where space isn't important in one direction, but matters in another. This could be identified by looking at interaction plots.

All participants were thanked for their attendance and contributions to discussions. It was announced that the next public workshop will be coincident with the Society for Marine Mammalogy conference in 2021.