

Camera trap distance sampling workshop 21-25 March 2022

Source: https://unsplash.com/@satyadeep_d

Solution

Fundamental to sound inference regarding animal abundance from distance sampling analysis of camera trap data is selection of appropriate detection function models. That process is made more difficult when analysing camera trap data because of the problem of overdispersion. This practical demonstrates the process of model selection that copes with overdispersion.

1 Model selection adjustments from overdispersion

Overdispersion causes AIC to select overly-complex models, so analysts should specify the number/order of adjustment terms manually when fitting distance sampling models to data from camera traps, rather than allowing automated selection using AIC. Howe et al. (2019) describes two methods for performing model selection of distance sampling models in the face of overdispersion. The first method of Howe et al. (2019) employs a two-step process and is implemented in the most recent version of Distance for Windows.

We will walk through the process of selecting the most appropriate model from the candidate model set you fitted in Practical 2.

2 Step 1: compute (\hat{c}_1) for most complex key function models

- Open project used in Practical 2 to the analysis set you created in that practical.
- Create three additional columns using *Column browser*
 - GOF Chi2/df, c-hat and QAIC
 - * The difference between **GOF Chi2/df** and **c-hat**: **c-hat** is a user-supplied value, while **GOF Chi2/df** is a statistic computed for a fitted model. You will copy the **GOF Chi2/df** from a fitted model and input it (Fig. 3) as **c-hat** into a model about to be fitted.
- Duplicate (using the *New analysis* button, Fig. 1) these analyses and run them again
 - unicos 3, hncos 2, hrsimp 1
 - these are the most complex models for each key function family
- At this point, you will have 11 models in the analysis set; 8 from Practical 2 and the 3 models you duplicated and re-ran.



Figure 1: With the analysis of interest highlighted, press the *New analysis* button to duplicate that analysis.

ID	#	1	2	Name	# params	# adj params	# key params	Delta AIC	AIC	GOF Chi-p	P	P CV	c-hat	GOF Chi2/df	QAIC
8	1	2	16	complex uniform cosine to get chat	3	3	0	0.00	44012.20	0.000	0.33	0.04	1.00	15.63	44014.20
9	1	2	13	complex halfnorm cosine to get chat	3	2	1	19.36	44031.56	0.000	0.33	0.06	1.00	19.26	44033.56
10	1	2	15	complex hazard simple to get chat	3	1	2	9.55	44021.75	0.000	0.37	0.01	1.00	17.90	44023.75
11	1	2	10	uniform cosine 3adj	3	3	0	0.00	44012.20	0.000	0.33	0.04	15.63	15.63	2823.50
12	1	2	9	uniform cosine 2adj	2	2	0	74.62	44086.82	0.000	0.28	0.02	15.63	24.91	2826.40
13	1	2	8	uniform cosine 1adj	1	1	0	80.34	44092.54	0.000	0.30	0.01	15.63	22.34	2824.89
14	1	2	12	halfnorm cosine 1adj	2	1	1	20.36	44032.56	0.000	0.33	0.04	19.26	16.54	2292.01
15	1	2	11	halfnorm cosine 0adj	1	0	1	132.98	44145.18	0.000	0.26	0.01	19.26	29.93	2295.96
16	1	2	17	halfnorm cosine 2adj	3	2	1	19.36	44031.56	0.000	0.33	0.06	19.26	19.26	2293.85
17	1	2	18	hazard simple 1adj	3	1	2	9.55	44021.75	0.000	0.37	0.01	17.90	17.90	2466.98
18	1	2	14	hazard simple 0adj	2	0	2	20.53	44032.73	0.000	0.38	0.01	17.90	17.21	2465.70

Figure 2: Project browser after supplying \hat{c}_1 values to all candidate models and re-running them. Coloured lines show application of \hat{c}_1 values from complex model to all other models in key function family.

3 Step 2: compute QAIC for each of the 8 models

Now that the values of (\hat{c}_1) are available after refitting the complex models, we will re-fit all 8 models of the candidate model set to produce QAIC values:

$$QAIC = -2 \left\{ \frac{\log(\mathcal{L}(\hat{\theta}))}{\hat{c}} \right\} + 2K$$

- Using the *GOF Chi2/df* values, edit each of the 8 models previously fitted, inserting the *GOF Chi2/df* for a key function family into the *c-hat* field in the *Model properties* page (Fig. 3) for all models in that key function family and refit. The refit will over-write previous results; you will see the values in the QAIC and c-hat columns change to reflect the changes you made.
- For each key function family (uniform, half normal, hazard rate), examine the **QAIC** column for models in that family. The model with the smallest QAIC is the preferred model from its respective family.

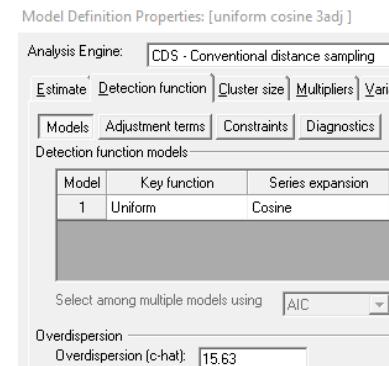


Figure 3: Supplying \hat{c}_1 value for all candidate models.

ID	#	1	2	Name	# params	# adj params	# key params	Delta AIC	AIC	GOF Chi-p	P	P CV	c-hat	GOF Chi2/df	QAIC
8	1	2	16	complex uniform cosine to get chat	3	3	0	0.00	44012.20	0.000	0.33	0.04	1.00	15.63	44014.20
9	1	2	13	complex halfnorm cosine to get chat	3	2	1	19.36	44031.56	0.000	0.33	0.06	1.00	19.26	44033.56
10	1	2	15	complex hazard simple to get chat	3	1	2	9.55	44021.75	0.000	0.37	0.01	1.00	17.90	44023.75
11	1	2	10	uniform cosine 3adj	3	3	0	0.00	44012.20	0.000	0.33	0.04	15.63	15.63	2823.50
12	1	2	9	uniform cosine 2adj	2	2	0	74.62	44086.82	0.000	0.28	0.02	15.63	24.91	2826.40
13	1	2	8	uniform cosine 1adj	1	1	0	80.34	44092.54	0.000	0.30	0.01	15.63	22.34	2824.89
14	1	2	12	halfnorm cosine 1adj	2	1	1	20.36	44032.56	0.000	0.33	0.04	19.26	16.54	2292.01
15	1	2	11	halfnorm cosine 0adj	1	0	1	132.98	44145.18	0.000	0.26	0.01	19.26	29.93	2295.96
16	1	2	17	halfnorm cosine 2adj	3	2	1	19.36	44031.56	0.000	0.33	0.06	19.26	19.26	2293.85
17	1	2	18	hazard simple 1adj	3	1	2	9.55	44021.75	0.000	0.37	0.01	17.90	17.90	2466.98
18	1	2	14	hazard simple 0adj	2	0	2	20.53	44032.73	0.000	0.38	0.01	17.90	17.21	2465.70

Figure 4: Project browser following Step 2. Emphasis is upon QAIC values within each key function family. Smallest QAIC value within each key function family are highlighted in the QAIC column. These models go forward to Step 3.

4 Step 3: select best model from Step 2 champions

The suite of candidate models has now been narrowed from 8 to 3 using the QAIC process from the previous step.

The final stage in the model selection process is to examine the column of the *Results Browser* containing *GOF Chi2/df* values for the three models selected in the previous step. The model to be used for inference for the duiker data set is the model with the smallest *GOF Chi2/df* value.

With model selection complete, in the next practical, we will examine the density estimates produced from this model following adjustments that account for partial fields of camera view and temporal animal activity.

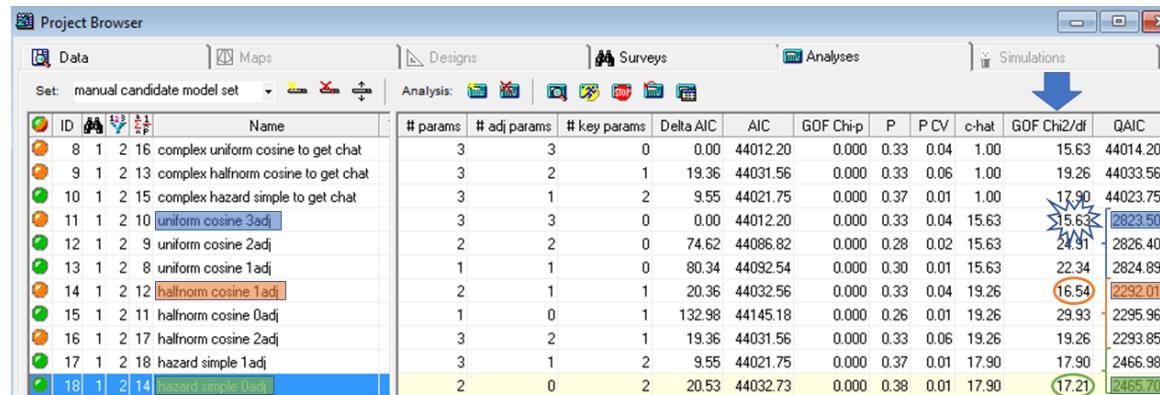


Figure 5: Project browser following Step 3. Emphasis is contrast of best models chosen from each key function family. Compare *GOF Chi2/df* values for three highlighted models. Smallest *GOF Chi2/df* value is 15.63 belonging to the uniform key function with three adjustment terms. This will be used to make inference about density in Practical 5.

5 Questions about this analysis

- Did the overdispersion approach lead to a different model choice than AIC selection algorithm?

Model selection choices

No, same model (uniform cosine with 3 adjustment terms) chosen from both approaches.

- How large is the difference in \hat{P} estimates between the preferred model and the second preferred model?

Differences in \hat{P}

The estimates are identical to two decimal places ($\hat{P}=0.33$)

References

Howe, Eric J., Stephen T. Buckland, Marie-Lyne Després-Einspenner, and Hjalmar S. Kühl (2019). "Model selection with overdispersed distance sampling data". en. In: *Methods in Ecology and Evolution* 10.1, pp. 38–47. ISSN: 2041-210X. DOI: 10.1111/2041-210X.13082.