

Design and Analysis of Distance Sampling Studies

Carl James Schwarz

Department of Statistics and Actuarial Science
Simon Fraser University
Burnaby, BC, Canada
cschwarz @ stat.sfu.ca

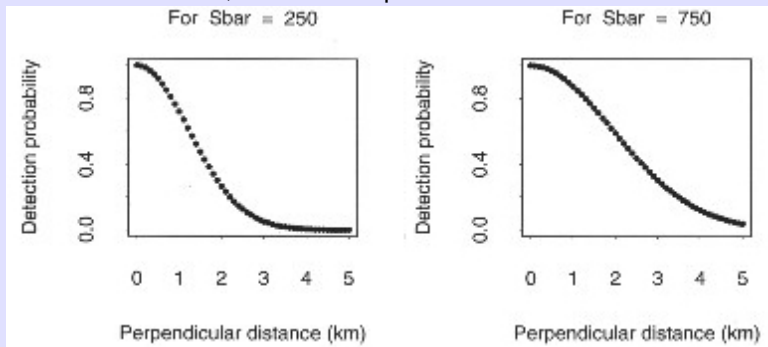
Part 2 Multiple Covariate Distance Sampling - Analysis

MCDS - Multiple Covariate Distance Sampling

MCDS - Intro

Multiple Covariate Distance Sampling:

- Similar to post-stratification where detection function can be modified by covariates (e.g. rainfall, observer, habitat type, etc.)
- Influences the *scale* of Detection function, but NOT its shape.
- Covariates affect rate at which detection function declines with distance, but not shape.



E.g. Half-normal detection function with larger scale (σ).

Covariates vs. post-stratification

- Post-stratification is very coarse; covariates are finer adjustments but need more data
- Need more contrast and more evenly distributed samples across covariate values.
- Post-stratification can have completely different detection functions; covariates only influence *scale* of detection function.

Key assumption about covariates:

- Covariates should be independent of distance
 - E.g. of failure - record habitat but only certain habitats are visible near the line (small shrubs) while other habitats are visible at large distances (tall trees).
- Avoid covariates that are correlated among themselves (e.g. height and mass of object).
- Covariates should act on *scale* parameter only.

Covariates added as more fields to data set and imported in the usual way.

Covariates can be continuous (e.g. temperature) or categorical (e.g. observer)

MCDS - Example

Amakihi on island of Hawaiian chain.

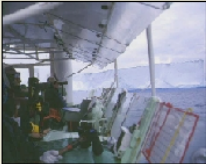
- collected as part of a larger study to assess a translocation experiment on the island of Hawaii
- Multi-species point- transect surveys at seven survey periods between July 1992 and April 1995.
- There were 41 point-count stations,
- Counted Hawaii Amakihi (*Hemignathus virens*), a generalist Hawaiian honeycreeper.
- Survey periods = strata.
- Covariates are Observer (OBS), minutes after sunrise (MAS), or hours after sunrise (HAS)
- Distances in meters.



Import the data in the usual way.

New Project Setup Wizard - Step 3: Survey Methods

In this screen, you tell Distance about your survey methods. Click 'Help' to find out more about each option.



*Mink whale line transect surveys, Antarctic Ocean
Photo: Peter Corkeron*

Type of survey

- ☐ Line transect
- ☒ Point transect
- ☐ Cue count

Observer configuration

- ☒ Single observer
- ☐ Double observer

Distance measurements

- ☒ Radial distance

Observations

- ☒ Single objects
- ☐ Clusters of objects

Sampling fraction
This option has been moved to the Multipliers page.

[Help](#) [Cancel](#) [< Back](#) [Next >](#) [Finish](#)

MCDS - Example - Amakihi

... giving

Data layers

Study area

Region

Point transect

Observation

Contents of Observation layer 'Observation' and all fields from higher layers

Study area			Region			Point transect			Observation				
ID	Label		ID	Label	Area	ID	Label	Survey effort	ID	Radial distance	OBS	MAS	HAS
ID	Label		ID	Label	Decimal	ID	Label	Decimal	ID	Decimal	Text	Decimal	Text
n/a	n/a	n/a	n/a	n/a	ha	n/a	n/a	[None]	n/a	m	n/a	[None]	n/a
Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int	Int
										1	40 TJS	50	1
										2	60 TJS	50	1
										3	45 TJS	50	1
										4	100 TJS	50	1
						1	1			5	125 TJS	50	1
										6	120 TJS	50	1
										7	140 TJS	50	1

MCDS - Example - Amakihi

Covariates affect the *scale* of the detection function, e.g. a factor covariate (categorical) with two levels:

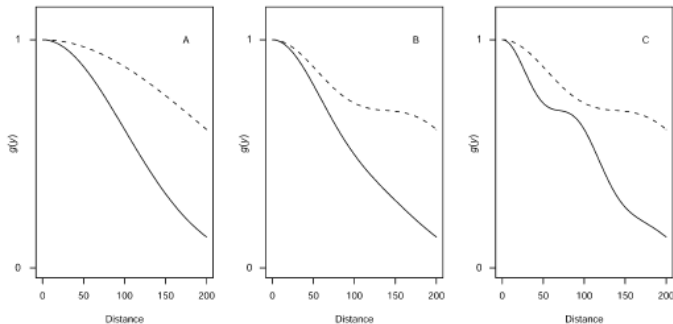


FIG. 2. Illustration of the effect of a factor covariate with two levels ($z = 1, 2$) on the key function (half-normal) with or without adjustment. (A) Key function without adjustment, with a scale term σ incorporating covariates: $\sigma = \exp[\log(50) + z \log(100/50)]$. (B) Key function with one cosine adjustment ($a_z = 0.1$) in which distances are scaled by truncation distance ($y_s = y/w$). In this case, shape of overall detection function changes depending on covariate values. (C) Key function and one adjustment term as in (B) but with distances in adjustment term scaled by covariate scale parameter ($y_s = y/\sigma$). In this case, shape of overall composite detection-function is constant at all covariate values and only the scale varies.

Model fitting advice:

- Start small, i.e. incremental add covariates.
- Start small, i.e. few factor levels (group in advance).
- Start with no series adjustments; if this converges try adding one level of adjustment
- Small probability event (i.e. extreme distances) have great effect on stability of estimates – truncate heavily, especially observations with probability of detection < 0.1 .
- Truncate at fixed width or radius value.

MCDS - Example - Amakihi

Fit a model with Observer (OBS) as a factor covariate.

Model Definition Properties: [Default Model Definition]

Analysis engine: MCDS - Multiple covariates distance sampling

Estimate | Detection function | Cluster size | Multipliers | Variance | Misc.

Stratum definition

☒ No stratification Layer type: Field name:

☐ Use layer type: Stratum

☐ Post-stratify, using: Stratum Area

Sample definition (for encounter rate)

Use layer type: Sample

Quantities to estimate and level of resolution

	Level of resolution of estimates		
	Global	Stratum	Sample
Density	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Encounter rate	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Detection function	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Cluster size (if required)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Global density estimate is Mean of stratum estimates
weighted by Stratum area ☐ Strata are replicates

Defaults Name: OBS OK Cancel

Note specification of MCDS.

MCDS - Example - Amakihi

... Fit a model with Observer (OBS) as a factor covariate ...

Model Definition Properties: [Default Model Definition]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Addjustment terms Covariates Constraints Diagnostics

Detection function models

Model	Key function	Series expansion	
1	Half-normal	Cosine	+
			-

Select among multiple models using AIC

Defaults Name: OBS-HN **OK** **Cancel**

MCDS - Example - Amakihi

... Fit a model with Observer (OBS) as a factor covariate ...

Model Definition Properties: [Default Model Definition]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate | Detection function | Cluster size | Multipliers | Variance | Misc.

Models | Adjustment terms | Covariates | Constraints | Diagnostics

Selection of adjustment terms

☐ Automated selection

Selection method: Sequential Look-ahead: 1

Selection criterion: AIC Significance level: 0.15

Maximum terms: 5

☒ Manual selection

Model	Num adj. parameters	Order of adjustment parameters (optional)
1	0	

Manually select starting values

Model	Num parameters
1	0

Scaling of distances

Scale distances by: W (right truncation distance)

Defaults Name: OBS -HN OK Cancel

MCDS - Example - Amakihi

... Fit a model with Observer (OBS) as a factor covariate ...

Model Definition Properties: [Default Model Definition]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Adjustment terms Covariates Constraints Diagnostics

Detection function covariates

Layer type containing covariate	Field name of covariate	Factor	Cluster size
Observation	OBS	<input checked="" type="checkbox"/>	<input type="checkbox"/>

Cluster size

To include cluster size as a covariate, add the cluster size field to the table of covariates and tick the 'Cluster size' box in that row.

When cluster size is a covariate, density is estimated using a different algorithm (see Help for details). Options are changed in the Estimate, Cluster Size and Variance tabs.

Defaults Name: OBS-HN OK Cancel

MCDS - Example - Amakihi

... Fit a model with Observer (OBS) as a factor covariate ...

Analysis 1: [New Analysis] Set: [Set 1]

Analysis

Name: Data(all) OBS HN Run

Created: 11/17/2012 1:40:05 PM

Run:

Survey

Set 1 [1] New Survey Details ...

Data filter

1 Default Data Filter Properties ...
New ...

Model definition

1 OBS -HN Properties ...
New ...

Comments

Inputs
Log
Results

... wait while the model is fit ...

Uses conditional MLE followed by a Horvitz-Thompson adjustment (which is slow).

Variance estimates are computed by bootstrapping (which is slow)

MCDS - Example - Amakihi

Results:

Model

Half-normal key, $k(y) = \text{Exp}(-y^{**2}/(2*s^{**2}))$

$s = A(1) * \text{Exp}(\text{fcn}(A(2)) + \text{fcn}(A(3)))$

Parameter A(1) is the intercept of the scale parameter s.

Parameter A(2) is the coefficient of level SGF of factor covariate OBS.

Parameter A(3) is the coefficient of level TJS of factor covariate OBS.

A(1) bounds = (2.5000 , 0.10000E+07)

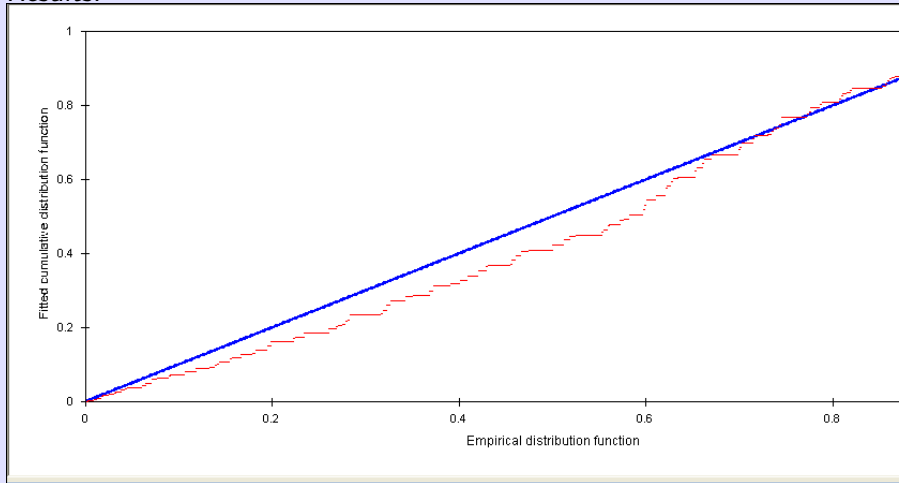
Factor covariates operate on scale and are offsets from base level (observer TKP)

Looks like strong observer effect:

Parameter	Point Estimate	Standard Error	Percent Coef. of Variation	95 Percent Confidence Interval	
A(1)	36.79	0.4172			
A(2)	-0.6990E-01	0.4936E-01			
A(3)	0.1761	0.4233E-01			
h(0)	0.58128E-03	0.15026E-04	2.58	0.55255E-03	0.61151E-03
p	0.55051E-01	0.14231E-02	2.58	0.52329E-01	0.57914E-01
EDR	58.657	0.75814	1.29	57.189	60.163

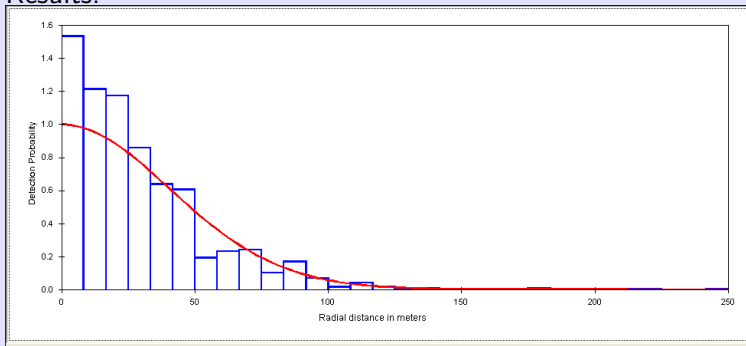
MCDS - Example - Amakihi

Results:



Hmmm ... good evidence of lack of fit. KS/CM GOF tests fail.

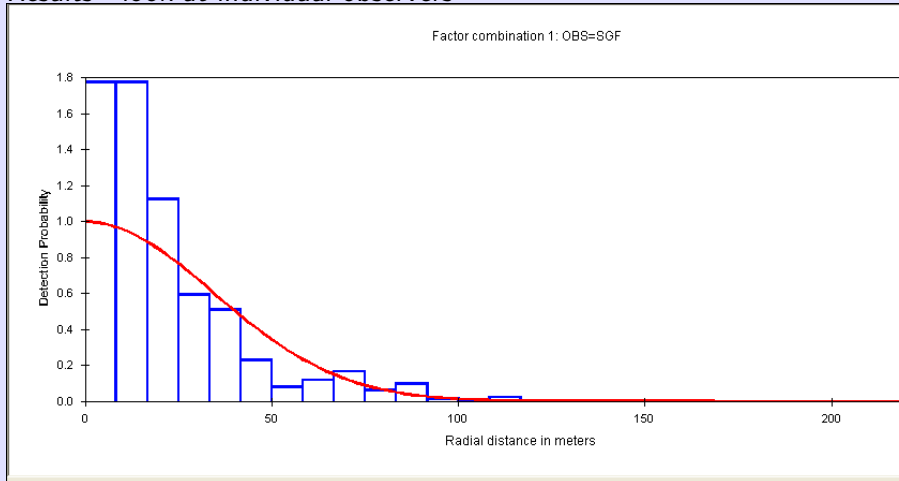
Results:



Hmmm ... good evidence of lack of fit. χ^2 GOF fail.

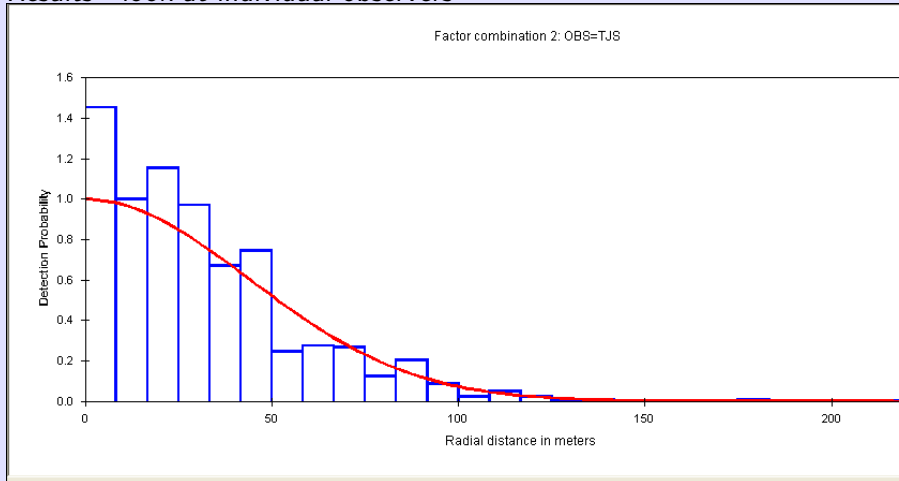
MCDS - Example - Amakihi

Results - look at individual observers



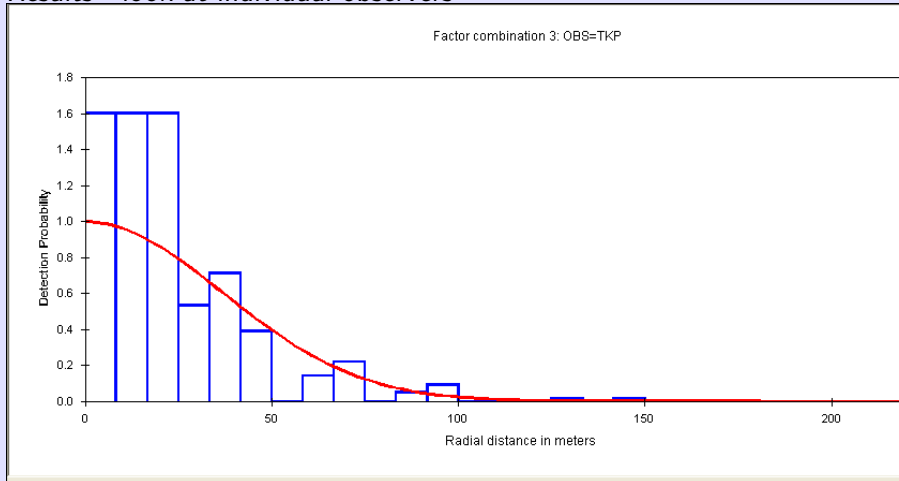
MCDS - Example - Amakihi

Results - look at individual observers



MCDS - Example - Amakihi

Results - look at individual observers



MCDS - Example - Amakihi

Results - look at final estimates

Parameter	Point Estimate	Standard Error	Percent Coef. of Variation	95% Percent Confidence Interval	
h(0)	0.58128E-03	0.15026E-04	2.58	0.55255E-03	0.61151E-03
p	0.55051E-01	0.14231E-02	2.58	0.52329E-01	0.57914E-01
EDR	58.657	0.75814	1.29	57.189	60.163
n/K	5.5618	0.19427E-01	0.35	5.5237	5.6002
D	5.1454	0.13422	2.61	4.8888	5.4155

Measurement Units

Density: Numbers/hectares

EDR: meters

Component Percentages of Var(D)

Detection probability : 98.2

Encounter rate : 1.8

MCDS - Example - Amakihi

Add *MAS* as a continuous covariate along with *OBS* and refit.

Model Definition Properties: [OBS+TIME -HN] ✕

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Addjustment terms Covariates Constraints Diagnostics

Detection function covariates

Layer type containing covariate	Field name of covariate	Factor	Cluster size	
Observation	OBS	<input checked="" type="checkbox"/>	<input type="checkbox"/>	+
Observation	MAS	<input type="checkbox"/>	<input type="checkbox"/>	

Cluster size

To include cluster size as a covariate, add the cluster size field to the table of covariates and tick the 'Cluster size' box in that row.

When cluster size is a covariate, density is estimated using a different algorithm (see Help for details). Options are changed in the Estimate, Cluster Size, and Variance tabs.

Data(all) OBS + MAS HN/cos(0) form of detection function:

Model

Half-normal key, $k(y) = \text{Exp}(-y^2/(2*s^2))$

$s = A(1) * \text{Exp}(\text{fcn}(A(2)) + \text{fcn}(A(3)) + \text{fcn}(A(4)))$

Parameter A(1) is the intercept of the scale parameter s.

Parameter A(2) is the coefficient of level SGF of factor covariate OBS.

Parameter A(3) is the coefficient of level TJS of factor covariate OBS.

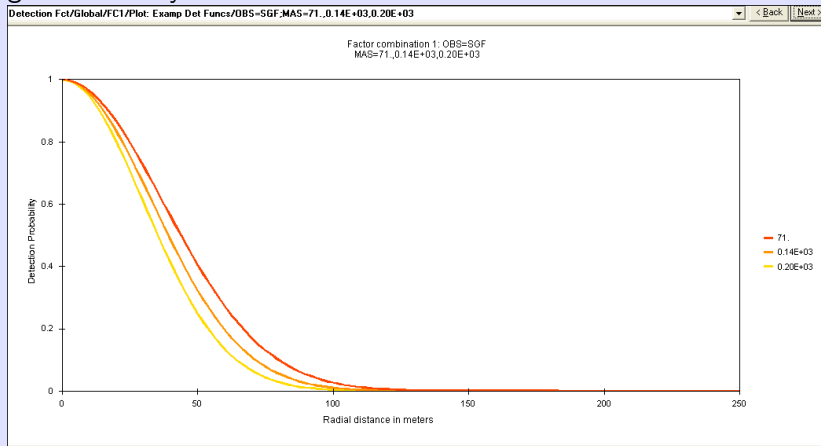
Parameter A(4) is the coefficient of covariate MAS.

Data(all) OBS + MAS HN/cos(0) - evidence of an effect of time

Parameter	Point Estimate	Standard Error	Percent Coef. of Variation	95 Percent Confidence Interval	
A(1)	48.37	0.6143			
A(2)	-0.1412	0.5086E-01			
A(3)	0.1254	0.4481E-01			
A(4)	-0.1699E-02	0.1477E-03			
h(0)	0.62202E-03	0.16661E-04	2.68	0.59019E-03	0.65557E-03
p	0.51445E-01	0.13780E-02	2.68	0.48813E-01	0.54220E-01
EDR	56.704	0.75942	1.34	55.234	58.213

MCDS - Example - Amakihi

Data(all) OBS + MAS HN/cos(0)- fit has improved but not that great and only minor effect of time



MCDS - Example - Amakihi

Try and add cosine series adjustment term to detection function:
Data(all) OBS + MAS HN/cos(1)

Model Definition Properties: [OBS+MAS -HN/Cos(1)]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Addjment terms Covariates Constraints Diagnostics

Selection of adjustment terms

☐ Automated selection

Selection method: Sequential Look-ahead: 1

Selection criterion: AIC Significance level: 0.15

Maximum terms: 5

☒ Manual selection

Model	Num adj. parameters	Order of adjustment parameters (optional)
1	1	

☐ Manually select starting values

Data(all) OBS + MAS HN/cos(1) - Results

MCDS - Example - Amakihi

Original paper results ...

Key	Adjustment terms	Covariates	Number of parameters	ΔAIC
CDS $f(0)$ pooled				
HN	Cos	–	5	21.40
Uni	Cos	–	2	25.11
HR	SP (0)	–	2	29.90
CDS $f(0)$ by strata (survey period)				
Uni	Cos	–	12	13.50
HR	SP (0)	–	14	18.59
HN	Cos	–	15	22.87
MCDS				
HR	SP (0)	OBS TIME	5	0.00
HR	SP (0)	OBS	4	1.73
HN	Cos	OBS HOUR	10	3.61
HN	Cos	OBS TIME	6	5.53
HR	SP (0)	OBS HOUR	9	5.62
HN	Cos	OBS	6	7.12
HN	Cos	TIME	4	24.56
HN	Cos	HOUR	8	24.57
HR	SP (0)	TIME	3	29.18
HR	SP (0)	HOUR	7	31.62

Fit the top MCDS model... (truncation at 82.5 m), TIME=MAS

MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0)

Model Definition Properties: [OBS -MAS HR /SP(0)]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Adjustment terms Covariates Constraints Diagnostics

Detection function models

Model	Key function	Series expansion	
1	Hazard-rate	Simple polynomial	

Select among multiple models using AIC

Defaults Name: OBS -MAS HR /SP(0) OK Cancel

MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0)

Model Definition Properties: [OBS -MAS HR /SP(0)]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate | Detection function | Cluster size | Multipliers | Variance | Misc.

Models | Adjustment terms | Covariates | Constraints | Diagnostics

Selection of adjustment terms

☐ Automated selection

Selection method: Sequential Look-ahead: 1

Selection criterion: AIC Significance level: 0.15

Maximum terms: 5

☒ Manual selection

Model	Num adj. parameters	Order of adjustment parameters (optional)
1	0	

Manually select starting values

Model	Num parameters
1	0

Scaling of distances

Scale distances by: W (right truncation distance)

Defaults Name: OBS -MAS HR /SP(0) OK Cancel

MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0)

Model Definition Properties: [OBS -MAS HR /SP(0)]

Analysis Engine: MCDS - Multiple covariates distance sampling

Estimate Detection function Cluster size Multipliers Variance Misc.

Models Adjustment terms **Covariates** Constraints Diagnostics

Detection function covariates

Layer type containing covariate	Field name of covariate	Factor	Cluster size
Observation	OBS	<input checked="" type="checkbox"/>	<input type="checkbox"/>
Observation	MAS	<input type="checkbox"/>	<input type="checkbox"/>

Cluster size

To include cluster size as a covariate, add the cluster size field to the table of covariates and tick the 'Cluster size' box in that row.

When cluster size is a covariate, density is estimated using a different algorithm (see Help for details). Options are changed in the Estimate, Cluster Size and Variance tabs.

Defaults Name: OBS -MAS HR /SP(0) OK Cancel

MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0)

Data Filter Properties: [Data(<82.5)]

Data selection | Intervals | Truncation | Units

Truncation of exact distance measurements

Right truncation

- ☐ Right truncate at largest observed distance
- ☐ Discard the largest percent of distances
- ☒ Discard all observations beyond

Left truncation

- ☒ No left truncation
- ☐ Discard all observations within

Truncation for cluster size estimation (where required)

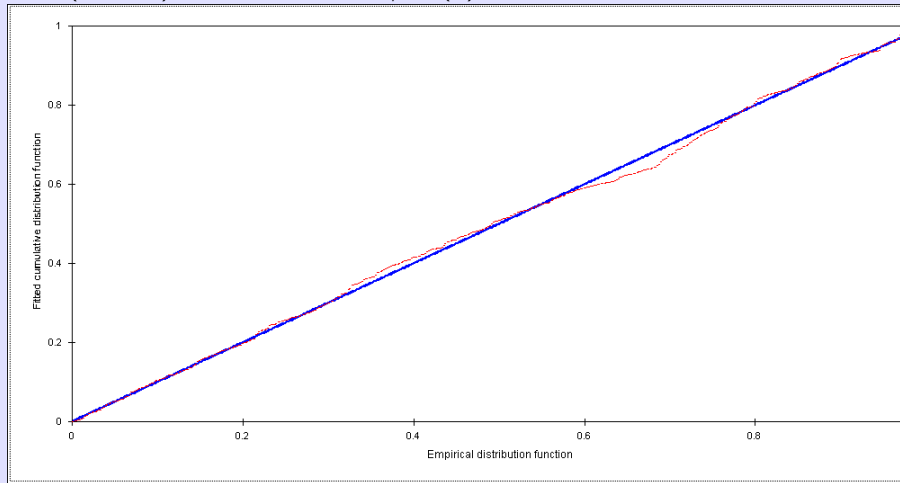
Right truncation

- ☒ Same as that specified above
- ☐ Discard all observations beyond

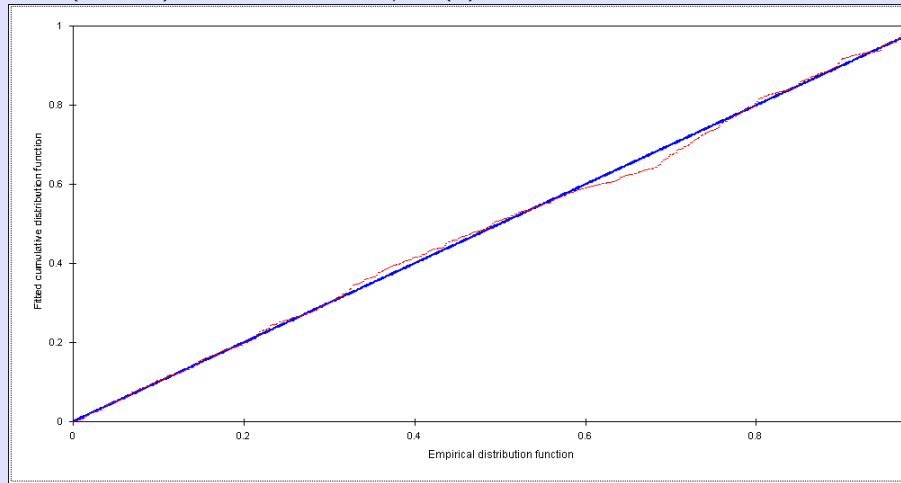
Defaults Name:

MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0): Results ...

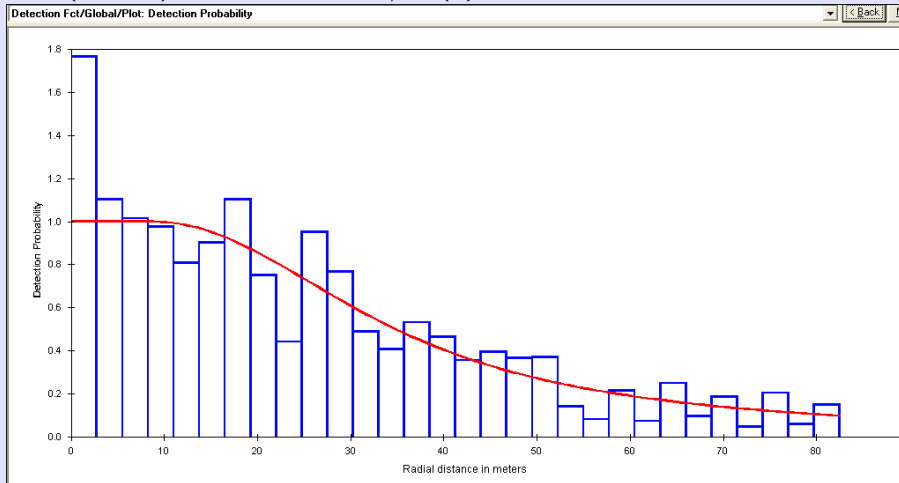


Data(< 82.5) OBS + MAS HR/SP(0): Results ...



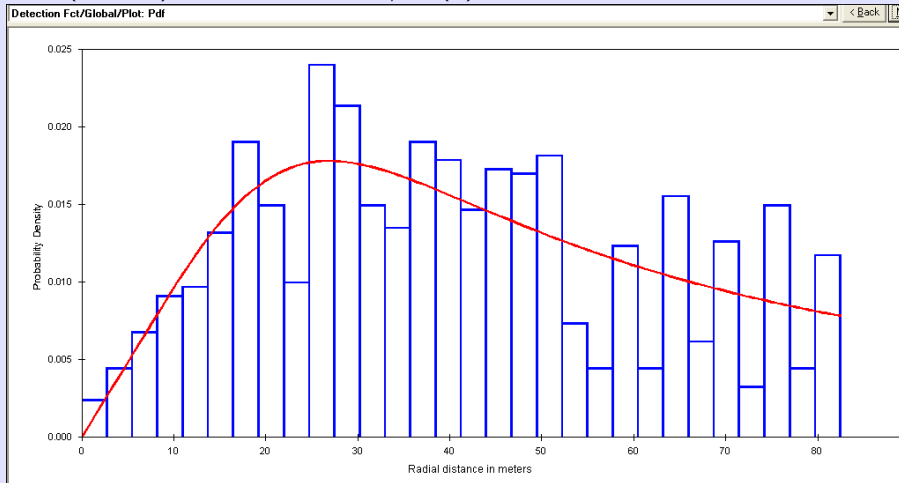
MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0): Results ...



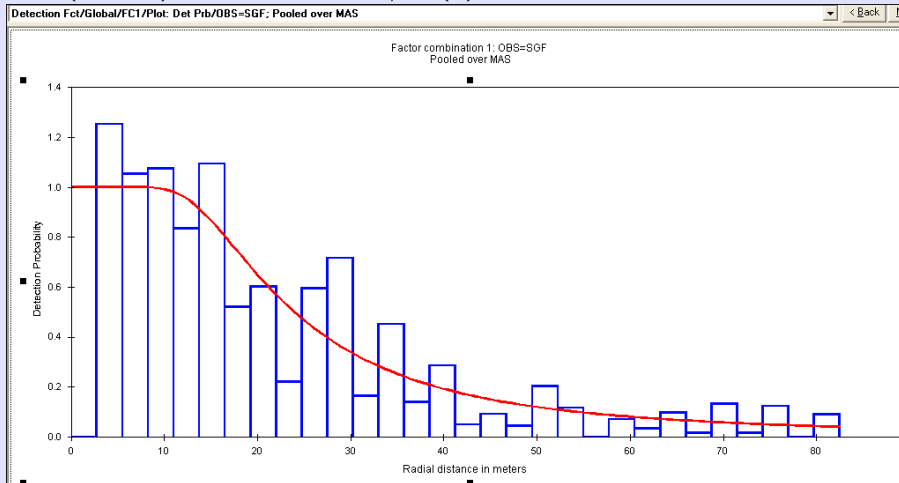
MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0): Results ...

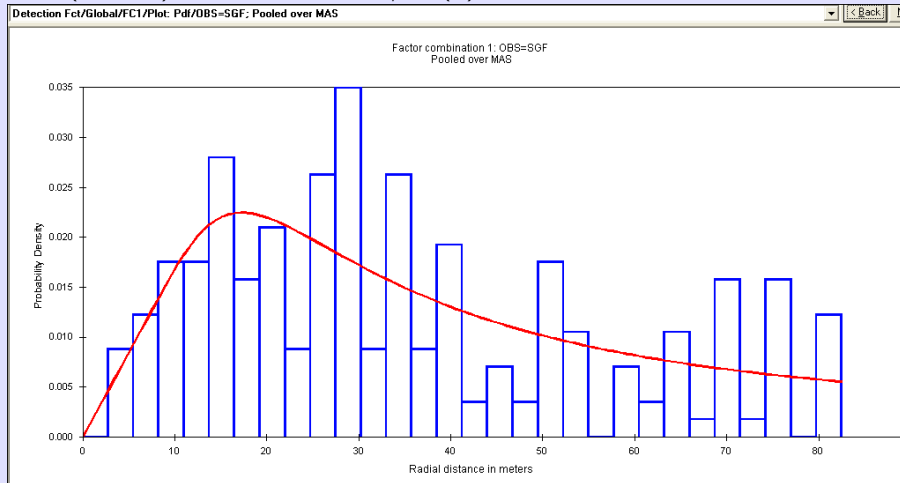


MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0): Results ...

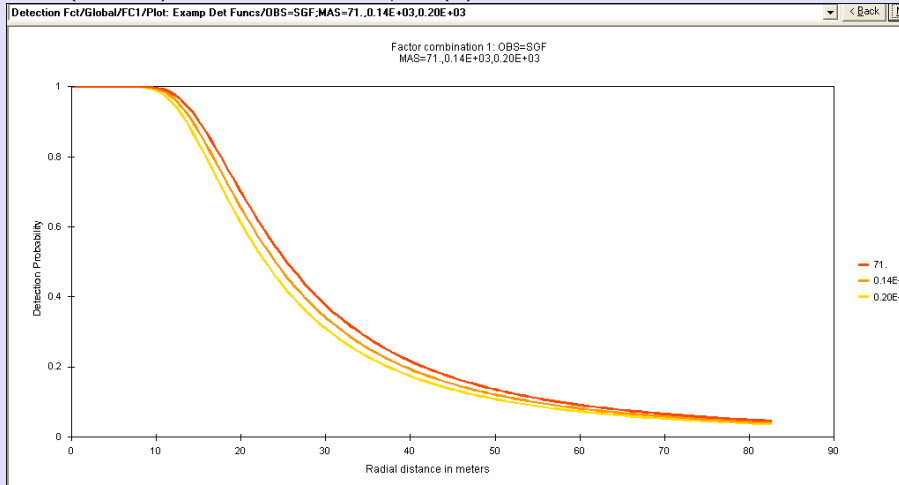


Data(< 82.5) OBS + MAS HR/SP(0): Results ...



MCDS - Example - Amakihi

Data(< 82.5) OBS + MAS HR/SP(0): Results ...



Compare the different observers detection profile.

Data(< 82.5) OBS + MAS HR/SP(0): Final Results ...

Parameter	Point Estimate	Standard Error	Percent Coef. of Variation	95% Percent Confidence Interval	
h(0)	0.96419E-03	0.25533E-04	2.65	0.91538E-03	0.10156E-02
p	0.30476	0.80706E-02	2.65	0.28934	0.32101
EDR	45.544	0.60304	1.32	44.376	46.743
n/K	4.6554	0.18534E-01	0.40	4.6191	4.6921
D	7.1440	0.19131	2.68	6.7784	7.5293

MCDS - Example - Amakihi

Multiple models – notice the difference in estimated Densities by model.

	ID				Name		# params	Delta AIC	AIC	ESW/EDF	D	D LCL	D UCL	D CV
	3	1	1	3	Data(all) OBS +MAS HN/cos(1)	11/17/	5	0.00	14031.48	54.5	5.953	5.649	6.273	0.027
	2	1	1	2	Data(all) OBS +MAS HN	11/17/	4	27.63	14059.11	56.7	5.506	5.222	5.806	0.027
	1	1	1	1	Data(all) OBS HN	11/17/	3	127.13	14158.61	58.6	5.145	4.889	5.415	0.026
	4	1	1	4	Data(all) OBS +MAS HR/SP(0)	11/17/	5	234.98	14266.46	67.0	3.944	3.753	4.145	0.025
	5	1	2	4	Data(<82.5) OBS +MAS HR/SP(0) 1	11/17/	5	0.00	10777.72	45.5	7.144	6.778	7.529	0.027

Don't compare model AICs with different data filters.

WHEW!

This was a large dataset where very small effects can be detected.

Go back to Example-Line (first example) and use ClusterSize as a covariate.

Compare to the previous results.

- When cluster size is a covariate, a different fitting procedure is used and you must obtain variance estimates using the bootstrap. Be sure to check the correct boxes.
- Not much evidence of an effect of cluster size on detection (where does it show this?)
- Look at detection function by cluster size – is this sensible?
- Small glitch in reported SE for \hat{D} but it is in the output. Look carefully for it.

Summary

- Covariates are an alternative to pre- and post-stratification.
- Covariates should be independent of distance.
- Covariate can be factor (categorical) or continuous.
- Covariates work on Detection function only and typically modify the *scale* parameters.
- Start simple and work to more complex models, esp. with series adjustments for key functions.