Multiple covariate distance sampling (MCDS) - Complications

- Aim:
  - Discuss some complications that arise with MCDS analyses
  - Give some analysis guidelines

CREEM
Centre for Research into Ecological and Environmental Modelling
University of St Andrews

Complications 1. Clustered populations

There are two approaches to estimating number of individuals when objects are in clusters:

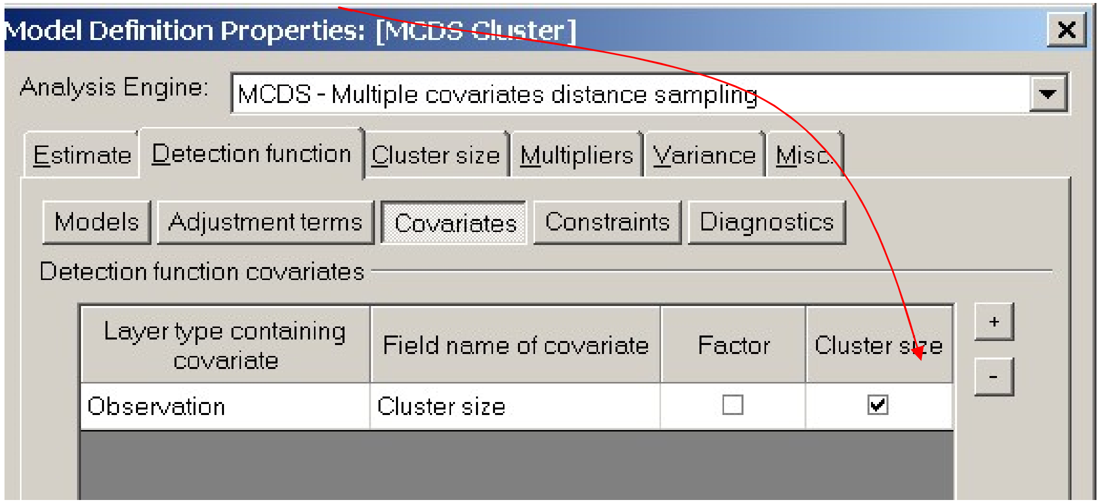
(1)
$$\hat{N} = \sum_{i=1}^{n} \frac{1}{\Pr[\textit{group included}]} \hat{E}[s]$$
$$= \frac{A}{2L} \hat{E}(s) \sum_{i=1}^{n} \hat{f}(0 \mid z_i)$$

(2)
$$\hat{N} = \sum_{i=1}^{n} \frac{\textit{group size}}{\Pr[\textit{group included}]}$$
$$= \frac{A}{2L} \sum_{i=1}^{n} s_i \hat{f}(0 \mid z_i)$$

When cluster size is *not* a covariate, we use (1); when it is a covariate, we use (2)

Clustered populations (contd.)

To tell Distance that a covariate represents cluster size, tick the box:



When cluster size is a covariate:

- Distance does not estimate variance using analytic methods: the bootstrap must be used (Reflected in the Variance tab)
- There is no need for size bias regression methods (Cluster size tab changes)
- No stratification allowed (Estimate tab)

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Complications 2. Adjustment terms

With adjustments: $$g(x,\mathbf{z}) = k(x,\mathbf{z})\left[1+\sum_{j=1}^{m} a_j p_j(x_s)\right] \Big/ c$$

Adjustment terms use *scaled* distances, x_s

- cosine adjustment of order 2: $\cos(2\pi x_s)$
- simple polynomial of order 4: x_s^4

Why scale?

- Avoid numerical problems
- Limits cosine adjustment to a small number of 'wiggles'

How to scale?

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

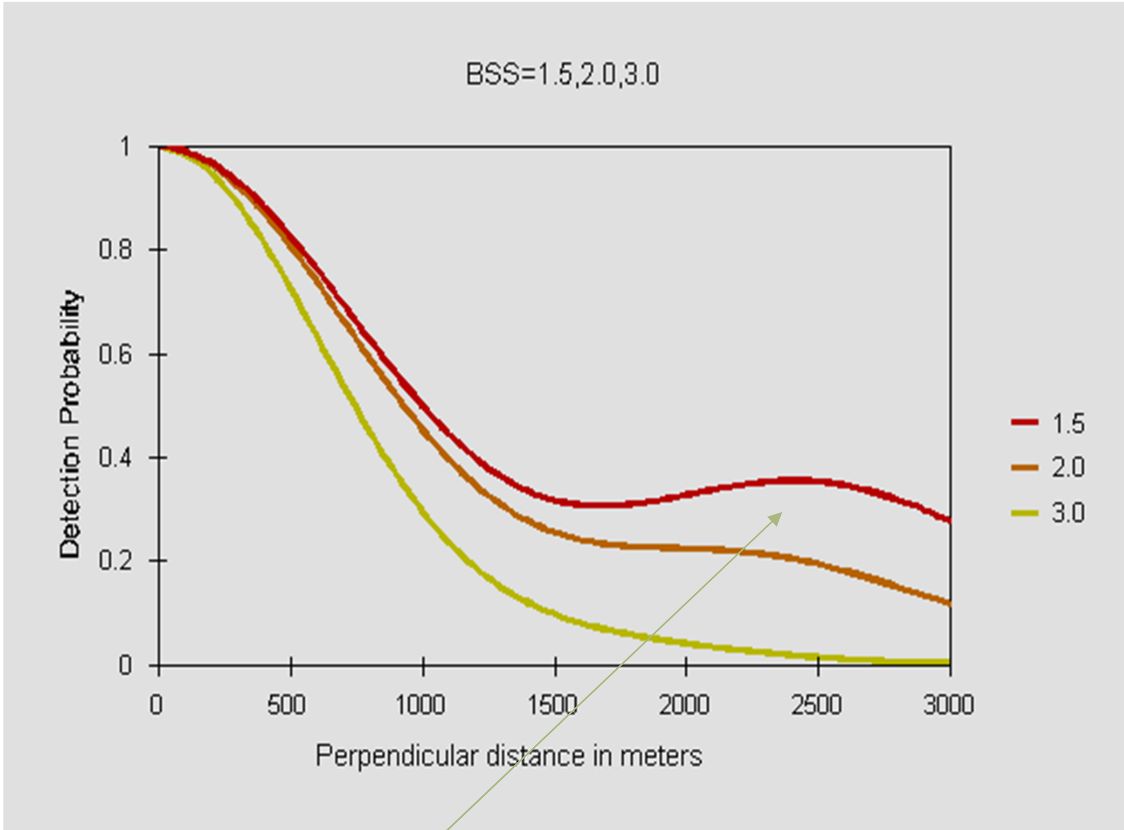
Adjustment terms (contd.)

Scenario 1: Scale distances by *w*, the right truncation distance $x_s = x/w$

Then covariates affect the scale of the key function, but adjustment terms are unaffected by covariates, so the overall shape varies with covariate value:

e.g., half-normal with 1 cosine adjustment of order 2

$$g(x \mid \mathbf{z}) \propto \exp\left(\frac{-x^2}{2\sigma(\mathbf{z})^2}\right)\left[1 + a_2 \cos\left(\frac{2\pi x}{w}\right)\right]$$



Note: no monotonicity constraint

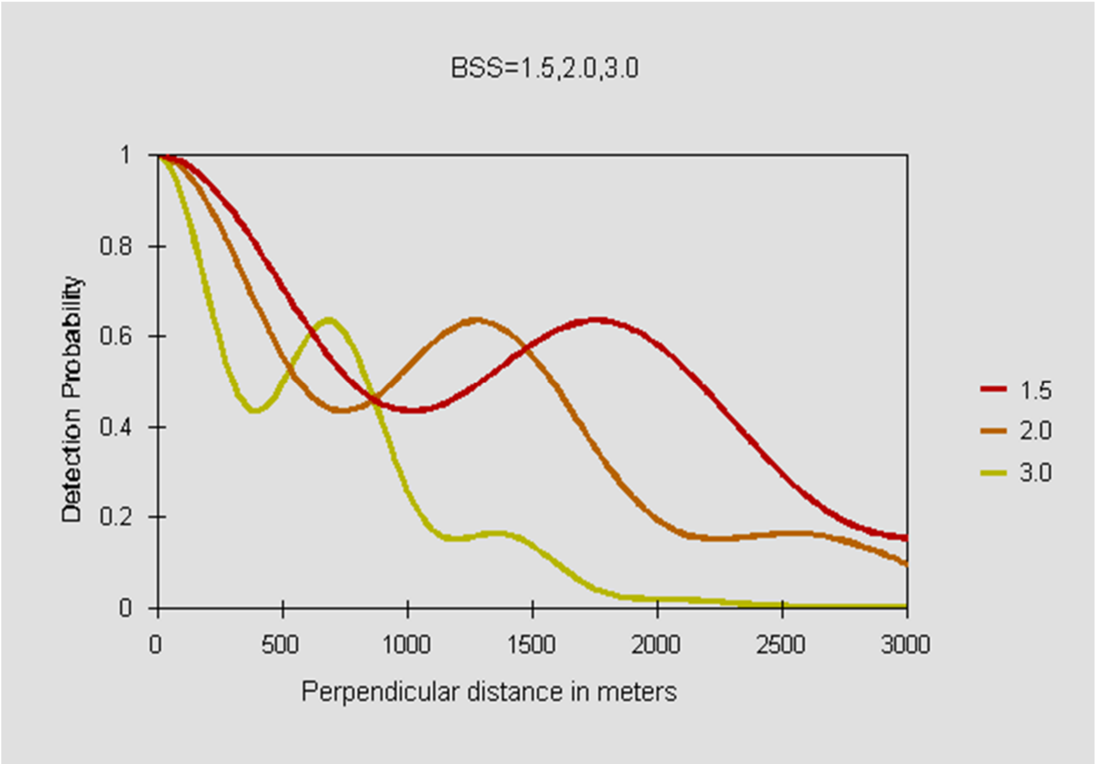
Adjustment terms (contd.)

Scenario 2: Scale distances by σ(**z**), the estimated scale parameter $x_s = x/\sigma(\mathbf{z})$

Then covariates affect the scale of the key function, and the scale of the adjustment terms, so only the scale and not the shape of the overall function is affected:

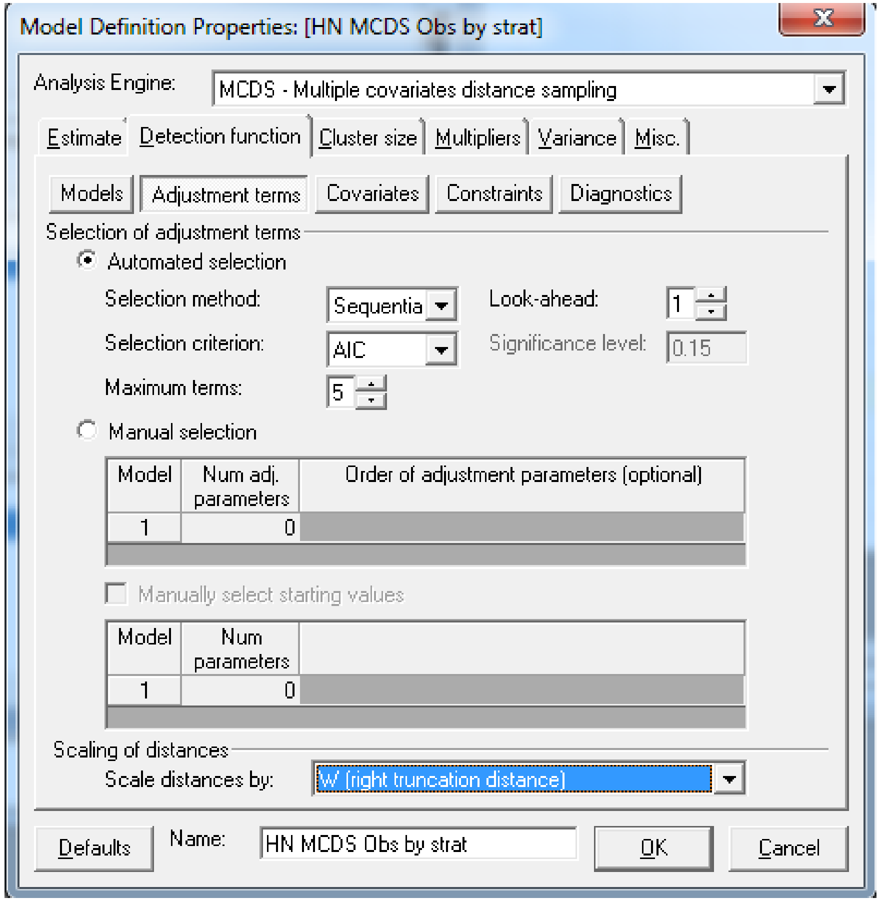
e.g., half-normal with 1 cosine adjustment of order 2

$$g(x \mid \mathrm{z}) \propto \exp\left(\frac{-x^2}{2\sigma(\mathrm{z})^2}\right)\left[1 + a_2 \cos\left(\frac{2\pi x}{\sigma(\mathrm{z})}\right)\right]$$



Adjustment terms (contd.)

- The previous was an extreme example, to illustrate the difference between scaling factors.
- Generally:
  - start with no adjustment terms and introduce covariate terms one by one
  - check the fit with adjustments looks reasonable
  - consider whether to scale by *w* or *σ*
- you may need fewer adjustment terms with MCDS than CDS analyses

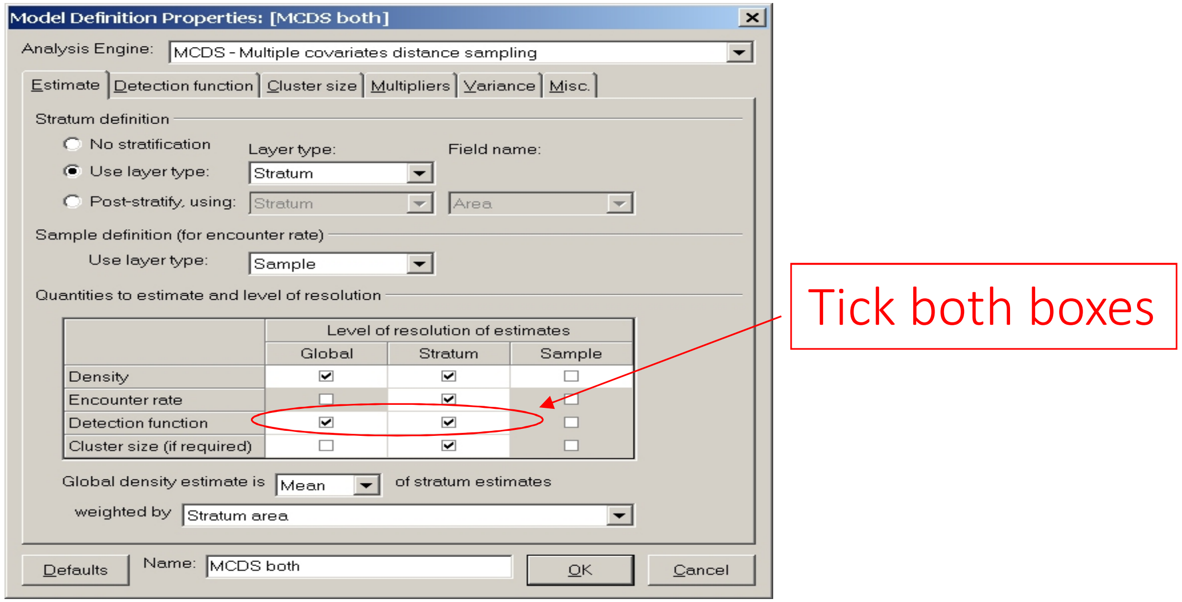


CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Complications 3. Stratification

If we want stratum-level estimates of density/abundance we can fit the detection function with covariates globally, and estimate $f(0|\mathbf{z})$ by stratum:



- Note: Global variance estimate for density/abundance must be calculated via the bootstrap

MCDS analysis guidelines

Choose covariates that are:

- independent of distance
- not strongly correlated with each other

Specifying the model:

- factor covariates generally harder to fit
- avoid or limit automatic selection of adjustment terms
- if using adjustments, consider whether to scale by *w* or σ
- check convergence and monotonicity
- add only one covariate at a time
- where necessary, use starting values and bounds for parameters
- consider reducing the truncation distance, *w*, if more than 5% of the $P_a(z_i)$ are <0.2, or if any are less than 0.1