Analysis of Populations in Clusters

- What changes when animals occur in clusters
- Size bias
- Methods to deal with size bias
- How to implement these methods in Distance

Clustered populations

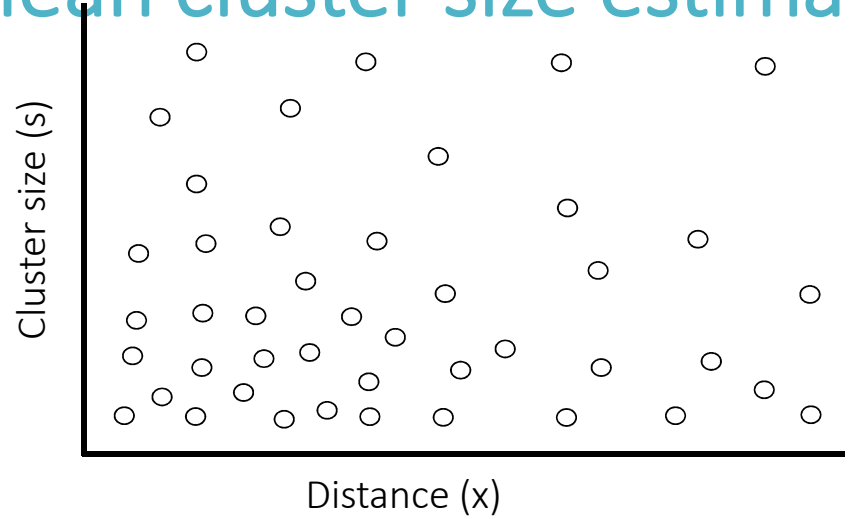
Density of clusters

$$\hat{D} = \hat{D}_s \times \hat{E}(s)$$

Mean cluster size

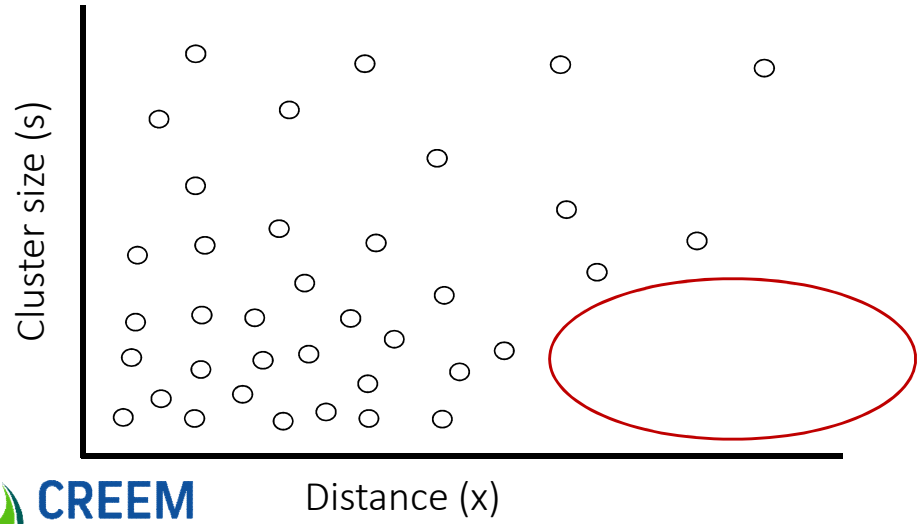
$$\left[cv(\hat{D})\right]^2 = \frac{\hat{V}(\hat{D})}{\hat{D}^2} \approx \left[cv(\text{encounter rate})\right]^2 + \left[cv\{\text{detection function}\}\right]^2 + \left[cv\{\text{cluster size}\}\right]^2$$

Mean cluster size estimation



No Size Bias

- Mean of observed sizes does not change with distance

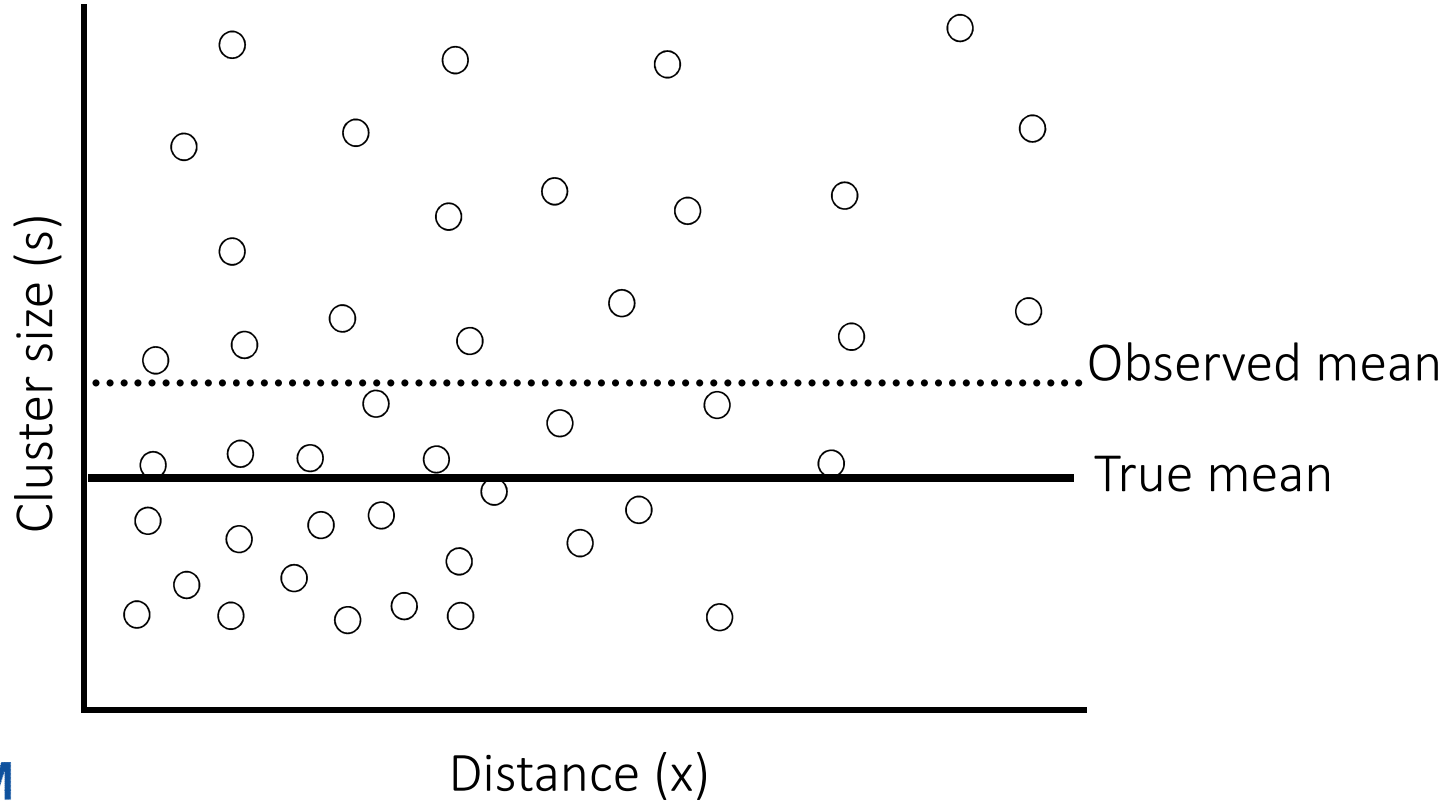


Size Bias

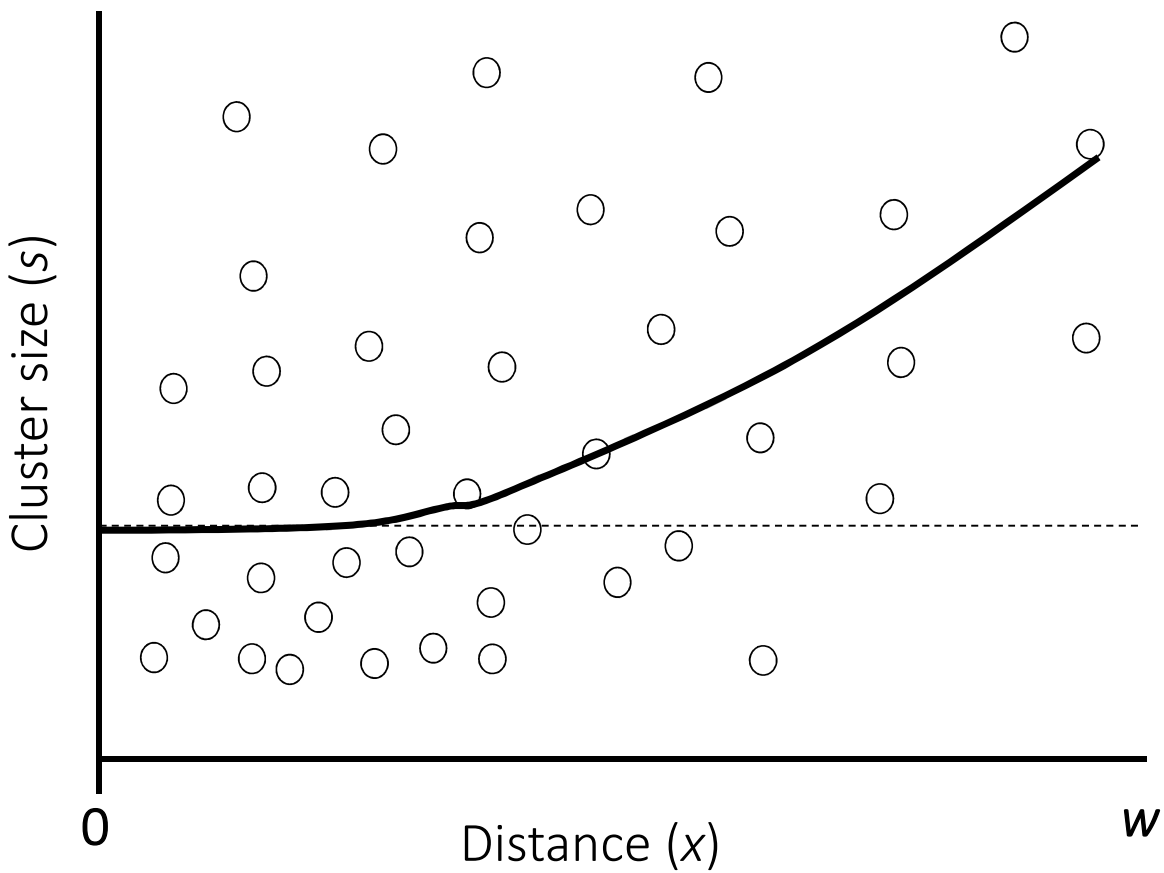
- Smaller clusters less detectable at larger distances
- Mean observed cluster size **increases** with distance

Effect of size bias on sample mean

If size bias is present, $\hat{E}(s) = \bar{s}$ will be positively biased:



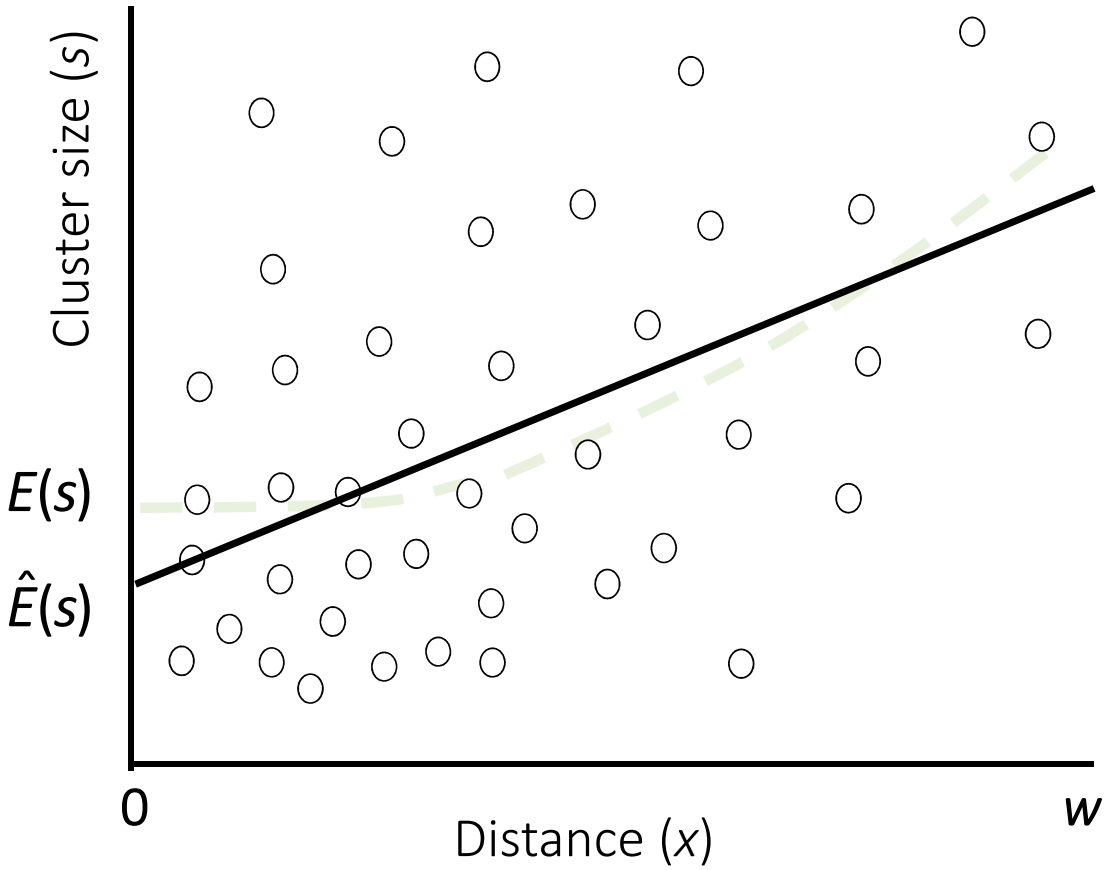
Regression methods



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Linear regression of *s* on *x*



CREEM
Centre for Research into Ecological and Environmental Modelling

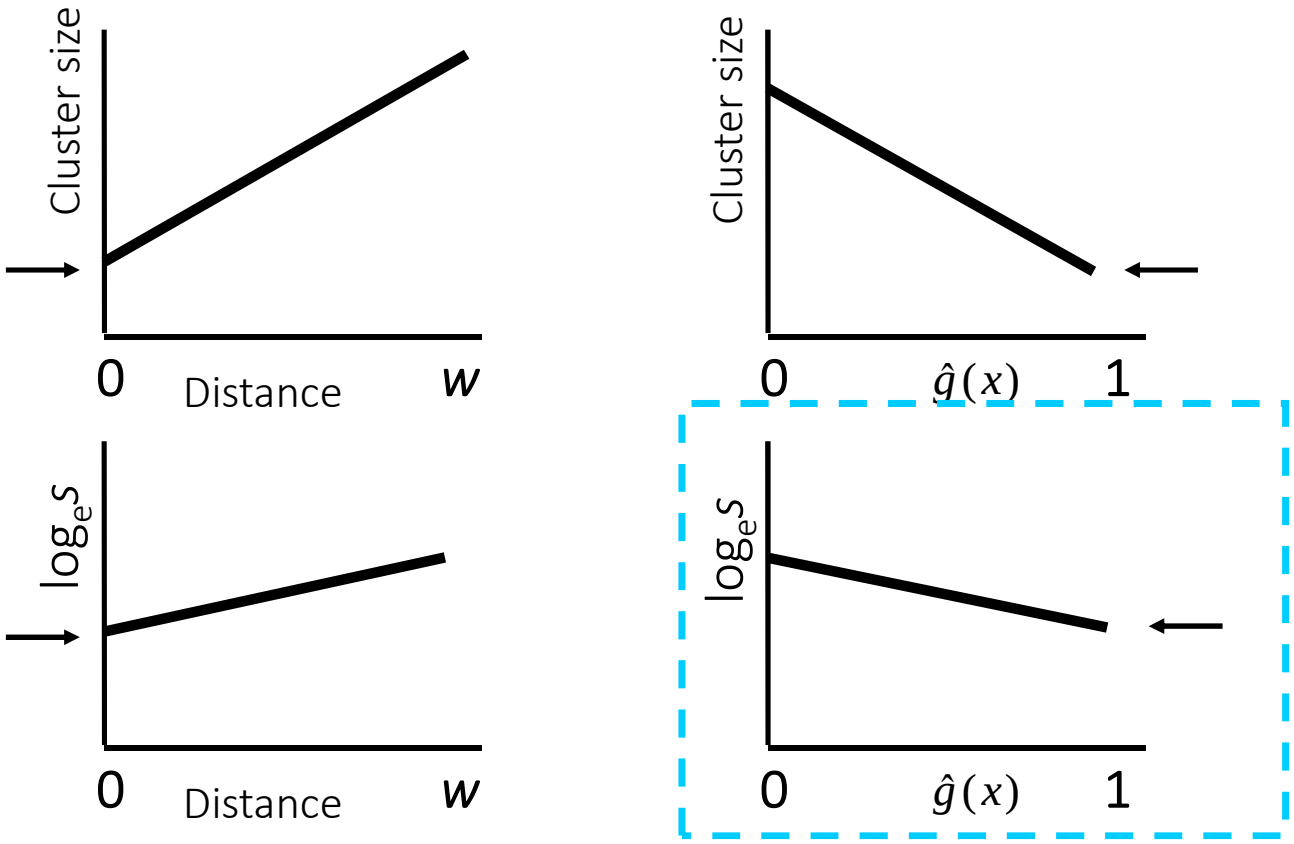
University of St Andrews

Problems with the linear regression method

- **Problem**: Relationship between *s* and *x* is not linear – no relationship when detection is certain (i.e. in the shoulder of the detection function).
    - **Solution:** Linearize by regressing s on $\hat{g}(x)$

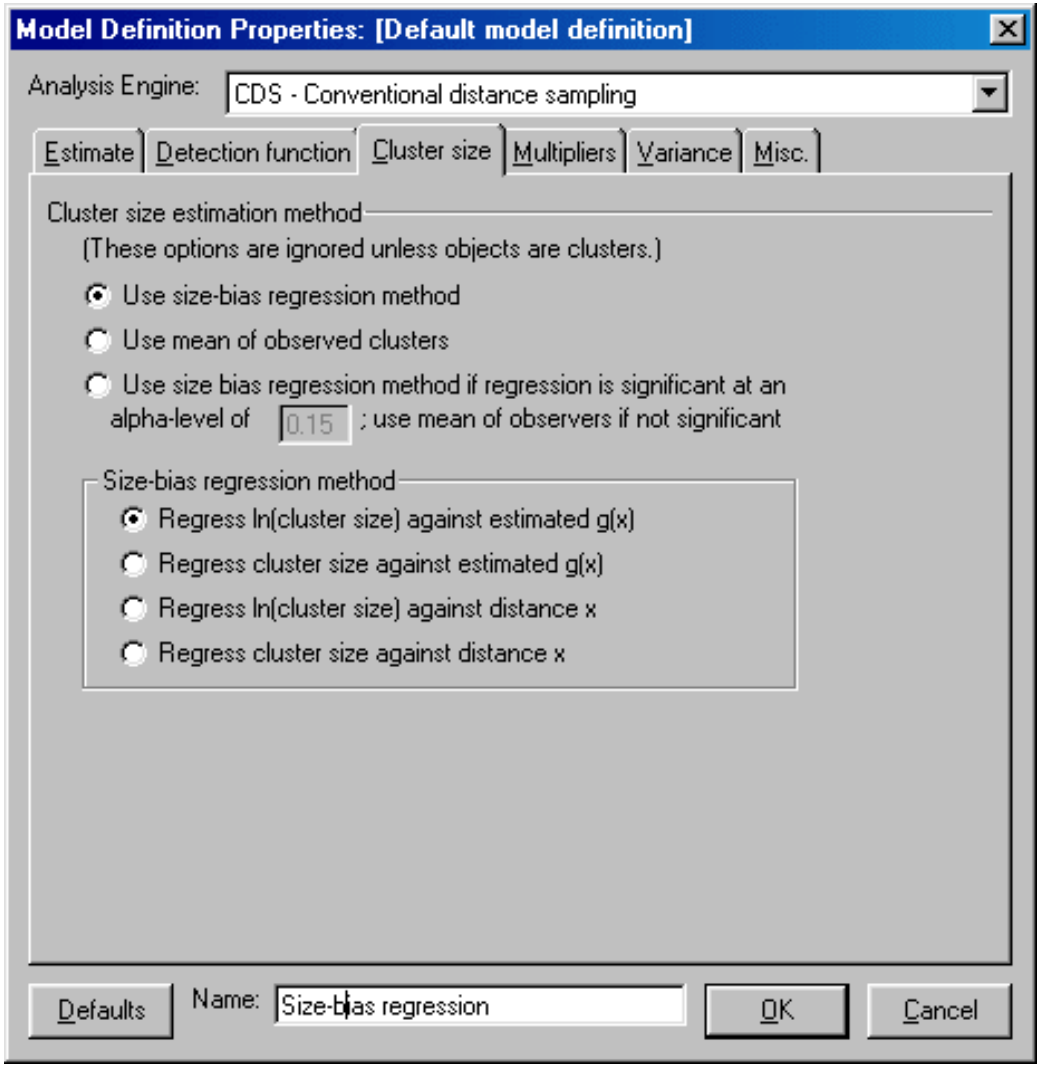
- **Problem**: Variance in *s* increases with $E(s)$ – large cluster sizes distort the fit.
    - **Solution**: Regress **log of cluster size** on $\hat{g}(x)$

Regression of log cluster size on $\hat{g}(x)$



This is the default method in Distance

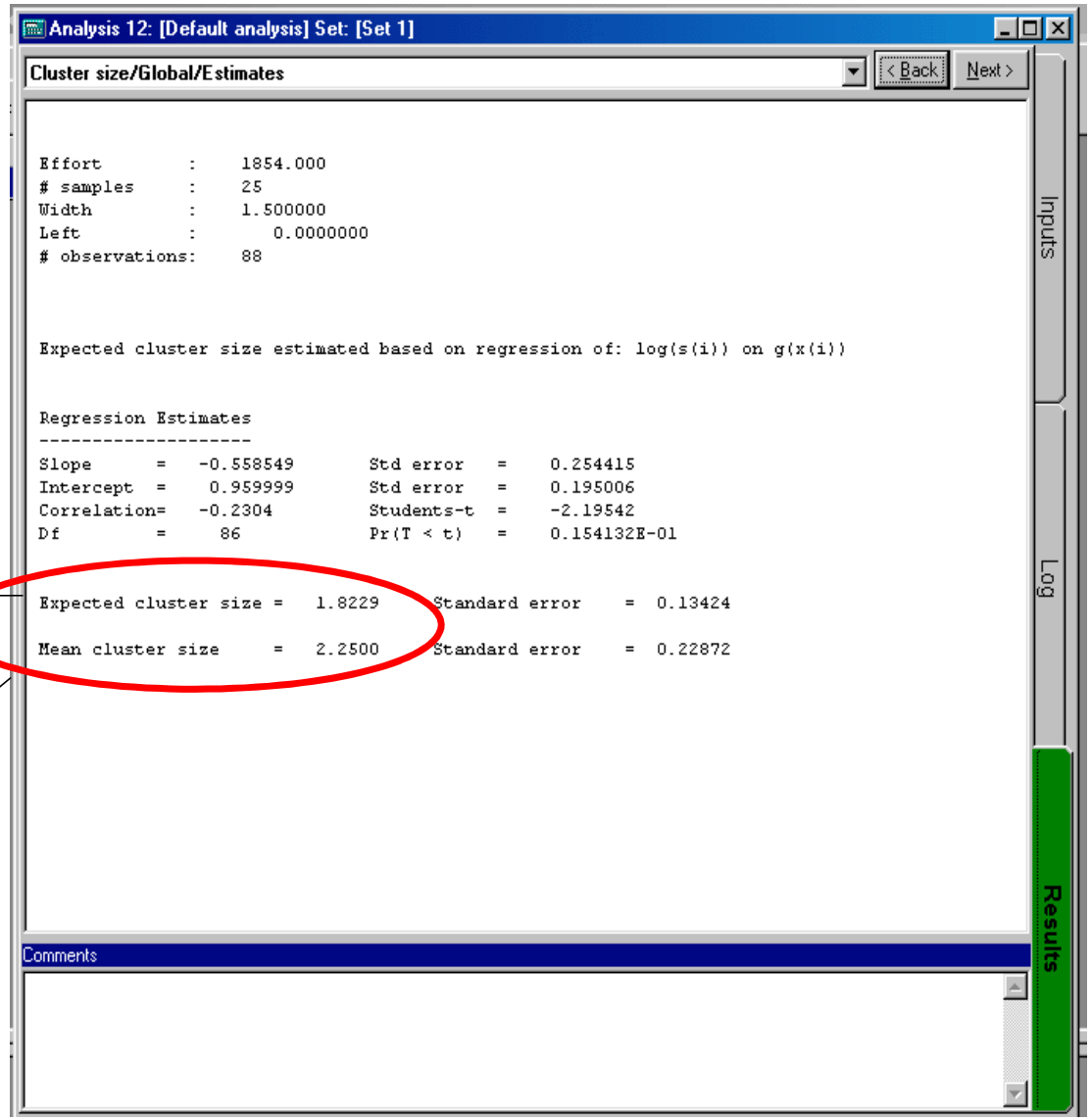
Estimating *E*(*s*) in Distance using regression methods



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Output of regression estimates



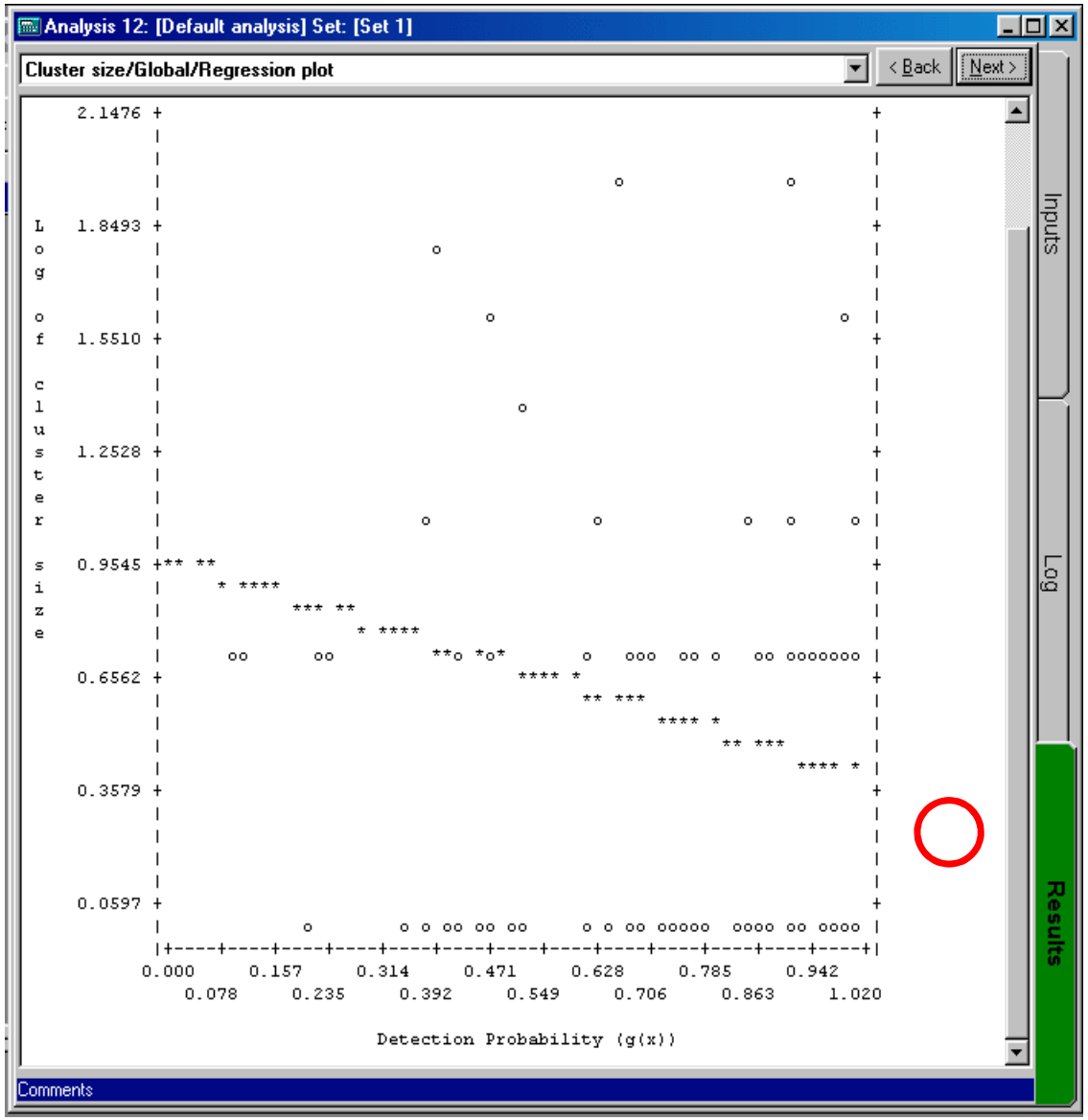
Regression estimate of cluster size

Mean cluster size from data

CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

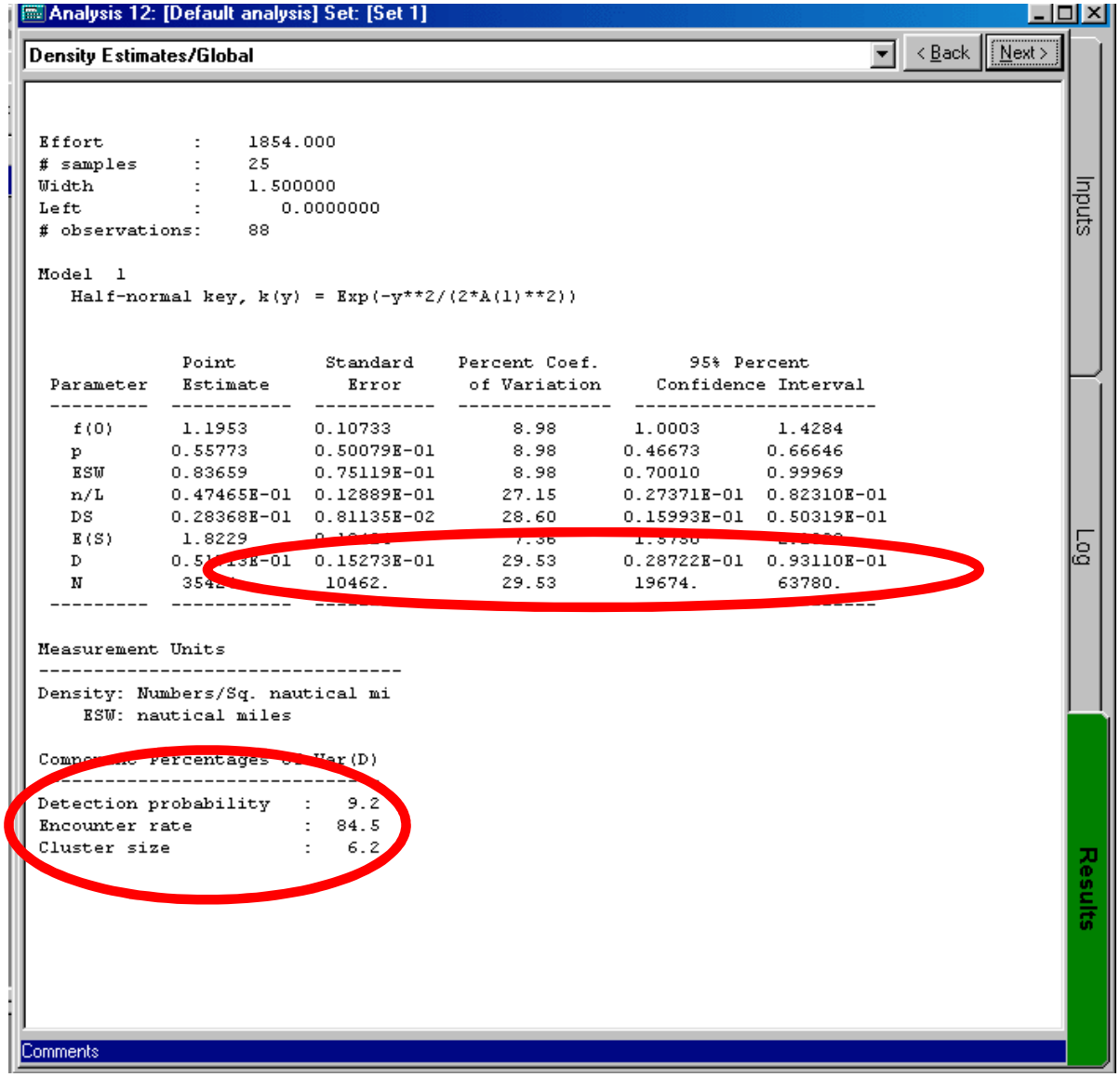
Line printer regression plot



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Precision of estimate



CREEM
Centre for Research into Ecological and Environmental Modelling

University of St Andrews

Estimating $E(s)$ in the presence of size bias

1. Regression methods — Default method in Distance
2. Include size in model for detection function — Multiple Covariate Distance Sampling
3. Stratify by cluster size — Rarely used now
4. Truncation of size-biased data — Rarely used now
5. Replace clusters by individuals — Rarely used now