

ON THE ON THE MEASUREMENT ERROR EFFECT ON ESTIMATES OF THE EFFECT OF FISHING PARAMETER λ

Doug S. Butterworth and N. Moosa
Marine Resource Assessment and Management Group (MARAM)
Department of Mathematics and Applied Mathematics
University of Cape Town, Rondebosch 7701, South Africa

In MARAM/IWS/DEC14/Peng/A10, Bergh presents arguments that measurement error associated with survey biomass estimates ($MB = x$) leads to positive bias in GLM estimates of the fishing effect parameter λ .

It is first important to understand how this bias arises. Since what is under consideration here is estimates of biomass, the relationship between measured and true biomass ($B = y$) will likely be (as near as makes no odds) a straight line through the origin: say $y = mx$.

Ignoring measurement error in x : $\hat{m} = \sum xy / \sum x^2$

If at the other extreme there was no error in y , but error in x , then the slope estimates becomes:

$$\hat{m}^* = \sum y^2 / \sum xy$$

(Of course there is no error in the “true biomass” in actuality, but this explanation is in the context of B being a covariate in a GLM for which there is an overall error term on the right hand side.)

When $r=1$, these two estimates coincide. Without such perfect correlation it is readily shown that:

$$\hat{m}^* > \hat{m}$$

Taking measurement error into account in the regression will (depending on relative variances) hence lead to an unbiased estimate $\hat{m}^\#$ where $\hat{m}^* > \hat{m}^\# > \hat{m}$.

It is the fact that $\hat{m}^\# > \hat{m}$ which leads to the regression parameter multiplying biomass in the GLMs being biased low, so that the λ parameter is biased high to compensate.

However, this argument relies on the assumption the error term in the original $y=mx$ regression has a constant standard deviation. If instead we consider the standard assumption associated with surveys of a log normal distribution with constant CV, i.e.: $\ln y = \ln m + \ln x$ with a constant variance error term, then:

$$\ln \hat{m} = \sum \ln y / \sum \ln x$$

but also:

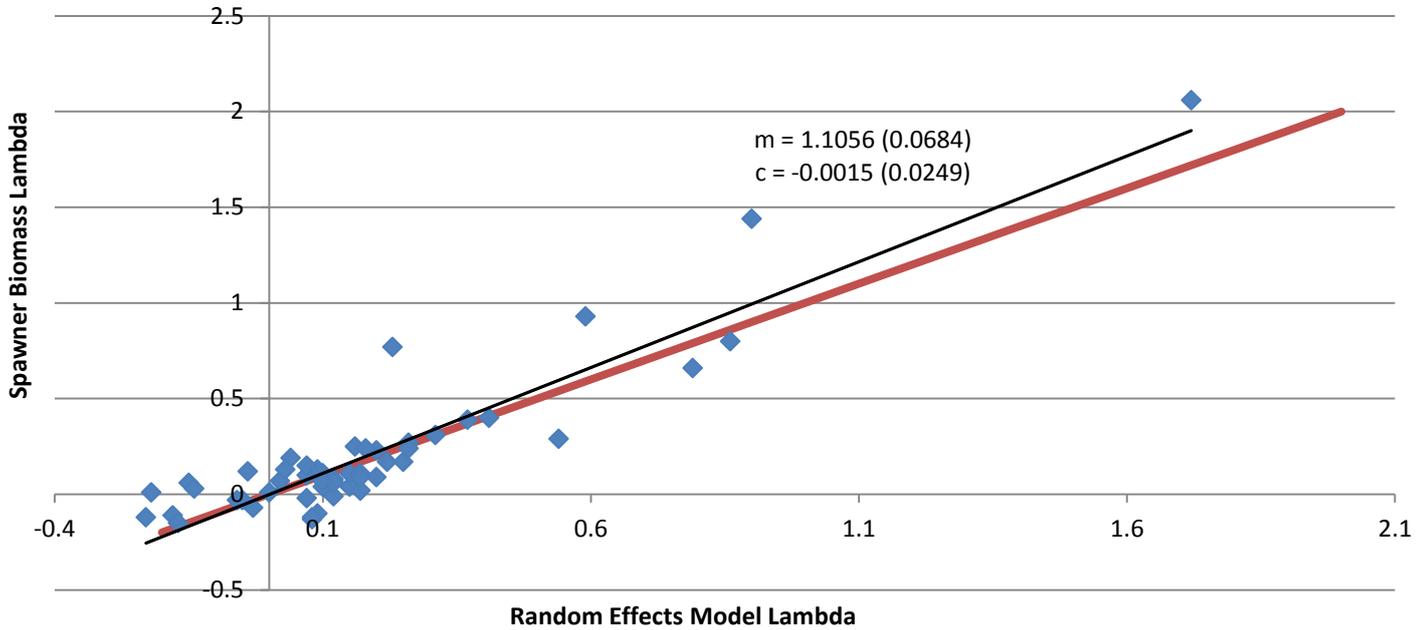
$$\ln \hat{m}^* = \sum \ln y / \sum \ln x$$

i.e. the two estimates are identical, so that the measurement error does not bias the estimate.

For SA pelagic surveys, previous attempts to find relationships of survey CVs to survey effort or to biomass have enjoyed little success. Effectively multiplicative process error seems to dominate, so that a constant CV is the most appropriate form to assume for measurement error effect evaluations, the effects of which would consequently be expected to be minimal.

The Figures attached provide an empirical test of this argument, based on the estimates of λ from the GLM results reported in MARAM/IWS/DEC14/Peng/B17 for Dassen and for Robben Islands. The random effects models considered there do not incorporate explicit survey estimates. Instead they effectively estimate the true survey biomass internally, and in a manner where the error, if any, would be expected to be much less than for the models which include the actual survey biomass estimates explicitly. Hence, if the measurement error bias suggested in MARAM/IWS/DEC14/Peng/A10 was non-trivial, when the λ estimates from based on either recruit or spawner biomass survey estimates are plotted against the corresponding random year effects estimates, one would expect to see a regression line through these data show a positive intercept of the vertical axis to reflect the bias suggested in Peng/A10. Instead in all four cases, the regression line passes through the origin (as near as makes no odds) (indeed the slope would also clearly be close to 1 in all cases, were it not for the impact of a few influential points of higher values). These Figures thus suggest that the measurement error bias for these analyses is insubstantial.

a) Scatterplot of the Lambdas for the Spawner Biomass and Random Effects Models for Dassen Island



b) Scatterplot of the Lambdas for the Recruit Biomass and Random Effects Models for Dassen Island

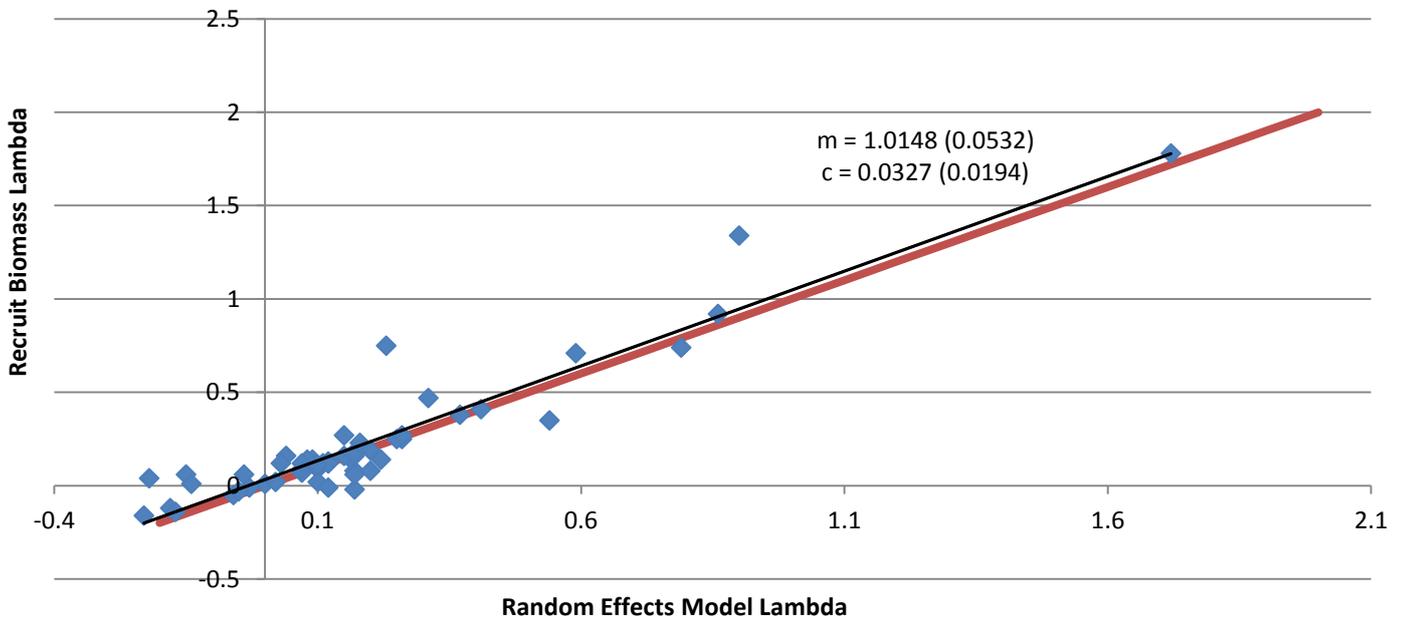


Figure 1: Scatterplot of the lambdas for a) the spawner biomass model and b) the recruit biomass model against the random effects model for Dassen Island. The red line is an equality relation, $y = x$. The black line is an ordinary linear regression, $y = mx + c$ with estimates and standard errors in parenthesis shown for m and c .

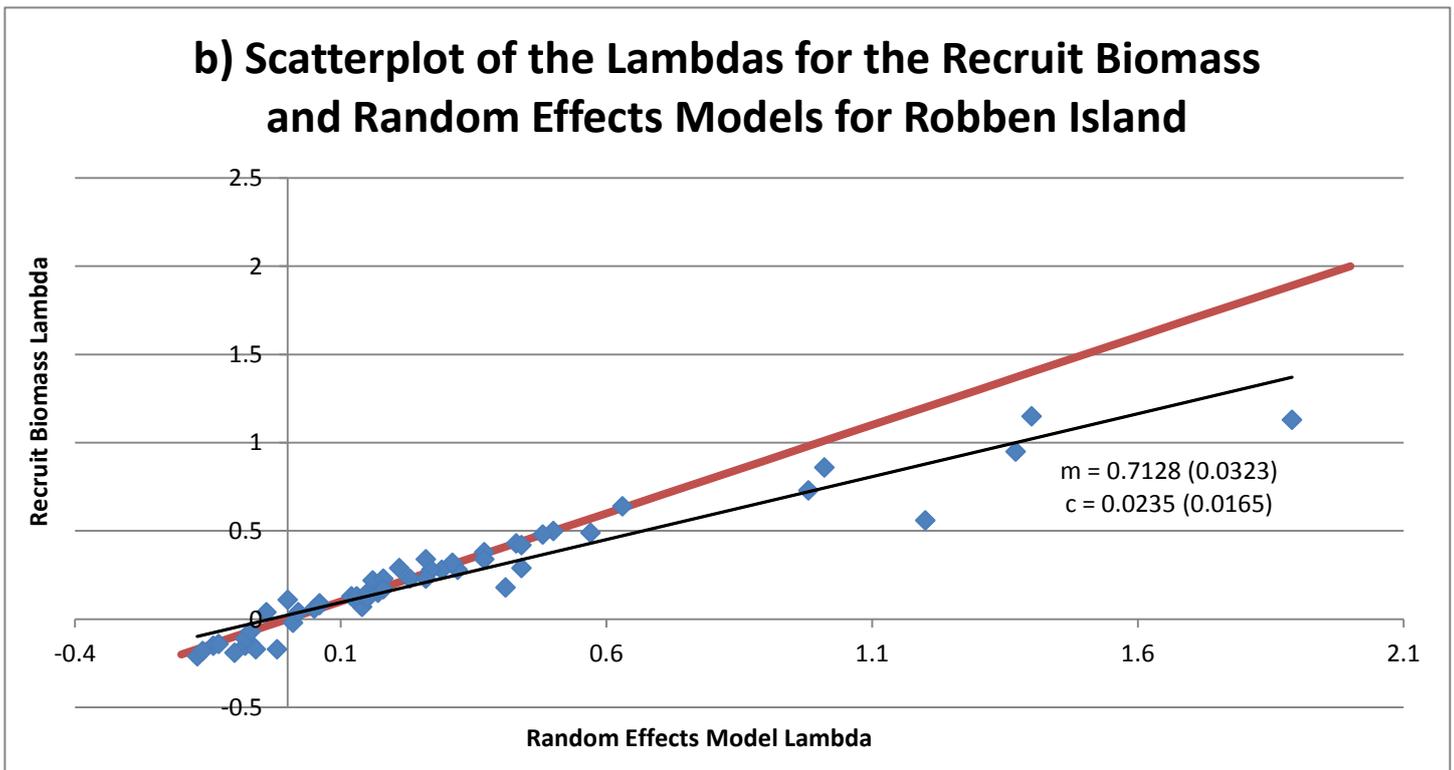
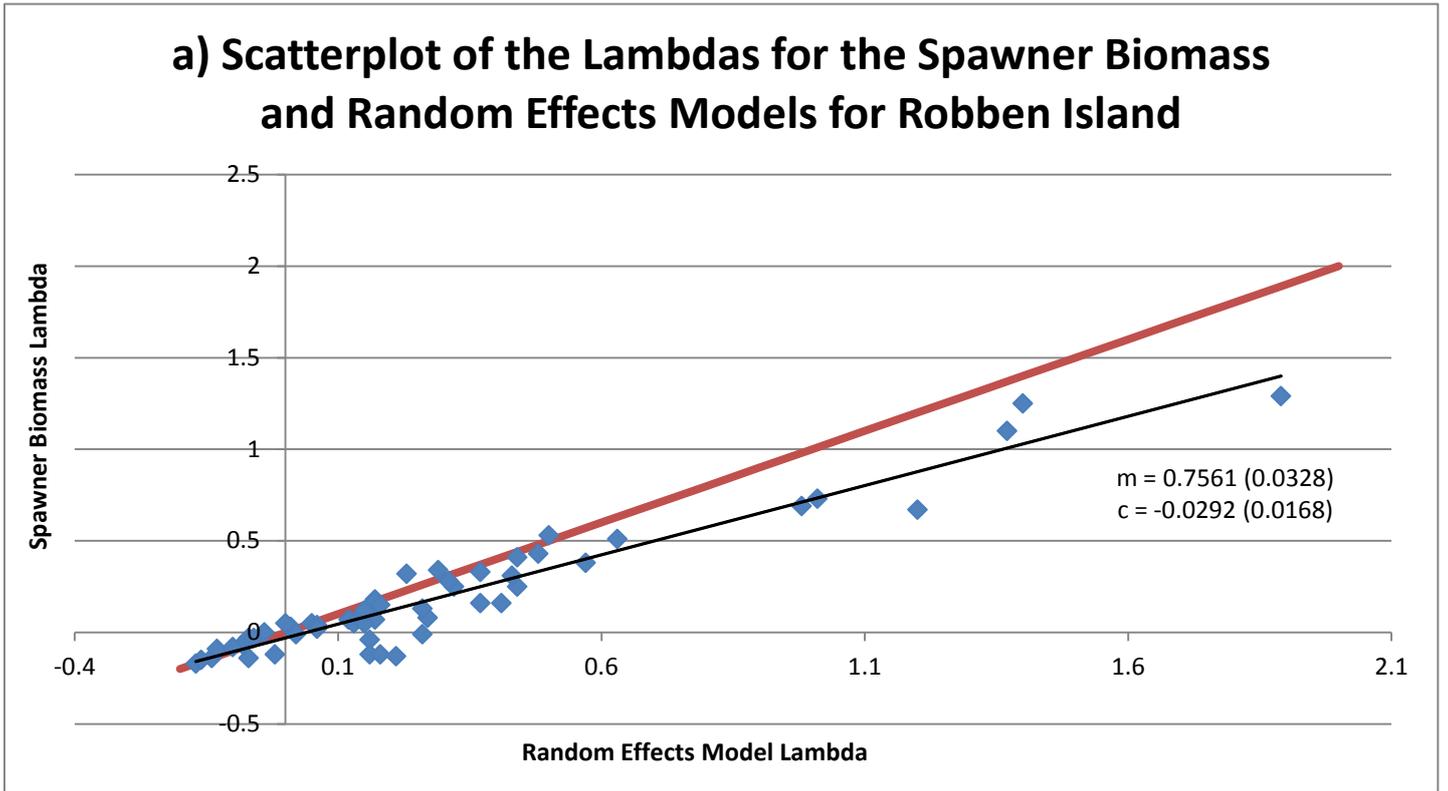


Figure 2: Scatterplot of the lambdas for a) the spawner biomass model and b) the recruit biomass model against the random effects model for Robben Island. The red line is an equality relation, $y = x$. The black line is an ordinary linear regression, $y = mx + c$ with estimates and standard errors in parenthesis shown for m and c .