WaLTER
wadden sea
LONG-TERM ECOSYSTEM RESEARCH

Sound, innovative and connective monitoring
for the Wadden Sea area

SAMPLING METHODS

SAMPLING METHODS

A COMPARISON BETWEEN THE SIBES AND THE BEUKEMA/DEKKER BENTHOS SAMPLING PROGRAM AT BALGZAND

Jaap van der Meer

Royal Netherlands Institute for Sea Research

VERSION 20150924

SEP. 2015

TABLE OF CONTENTS

ABSTRACT

Two different annual sampling programs of the marine intertidal macrozoobenthos of the Balgzand area in the western Dutch Wadden Sea are compared for the years 2008-2012 in terms of estimates of mean numerical and biomass density of eight abundant species. The SIBES program that uses a regular grid reveals lower SE/mean ratios than the BD program that uses a transect approach, despite a lower sampling effort in terms of total surface area sampled. Yet, particularly biomass density estimates are much lower in the SIBES program, which only partly can explained by inadequate sampling depth. It is recommended that the strenght of each program, efficiency in one and accuracy in the other, are combined in the future.

1 INTRODUCTION

In 1969 Jan Beukema from the Royal Netherlands Institute for Sea Research (NIOZ) started a biannual sampling program of the marine intertidal macrozoobenthos at Balgzand. This program, which was continued by Rob Dekker in the early 1990s and hereafter called the BD program, is still ongoing. Sampling occurs each year in February/March and in August/September. The same twelve transects of about one kilometer length and three permanent square stations, scattered over the area, are sampled on each occasion.

In 2008, NIOZ started another annual sampling program of the same species group over the entire Dutch Wadden Sea. This program, which is also maintained until now, uses a regularly spaced grid plus some additional points that are randomly selected on the transects between grid points [3]. It includes the entire Balgzand area. Sampling occurs in summer, mainly in June and July. The program has become known as SIBES.

In this short note, I will compare these two programs with respect to the estimates and their variances of the annual mean numerical and biomass densities of eight common species. Ideally, estimates are unbiased, which means that on average they resemble the true mean, and have low variance.

2 FIELD SAMPLING METHODS

The 12 transects and 3 stations of the BD program are said to be randomly sampled at the start of the program, but details on the precise way are lacking (Beukema pers.comm.). Neither the sampling universe (area), nor the sampling support (details on the sampling units) are well defined. The sampling universe is rougly defined as the Balgzand proper, that is the intertidal area west of the Amsteldiep. Exact geographical borders have not been defined. Generally speaking, different ways of selecting transects exists, and many methods result in uneven selection probability over the sampling region. For example, one way of selecting transects leading to uneven selection probability is to randomly select a starting point (longitude-latitude coordinate) and a direction of the transect (0-360 degrees), but without proper weighing of, for example, transects that cross the border of the area. Each transect of about 1 km contains 50 cores, taken in 10 groups of 5 cores. The succeeding groups within a transect are about 100 m apart, cores within a group are taken within a few meters of each other. But it is unclear what the precise area of a group is. In winter, the sampled surface area is 0.019 m^2 for each core. In summer a smaller core with a surface area of 0.0095 m^2 is used. The permanent stations have a size of 30 m by 30 m, and at each station 9 (previously 16) spatial points are chosen haphazardly. Around each point one box of 0.096 m2 and one core of 0.0095 m^2 is taken, both in summer and winter. Further details of the program can be found in [2].

The size of the regular SIBES grid at Balgzand varied between 250 and 308 points, where at each point a single core with a surface area of 0.0173 m^2 is taken. Here too the sampling universe is not well defined and it is also unclear whether an element of randomness was included in the selection of the grid. I deleted all points east of the Amsteldiep, which are classified as Balgzand samples in the SIBES database, but do not really belong to the Balgzand area s.s. The number of remaining points varied among years between 186 and 239. The reason for the difference in sample size among years is partly due to a different number of extra random points, but mainly to problems with reaching deeper points. Further details of the sampling program are given in [4].

3 ANALYTICAL METHODS

The variables of interest are abundances in terms of number and biomass per m^2 for five common bivalve species (the cockle *Cerastoderma edule*, the American razor clam *Ensis directus*, the Baltic tellin *Macoma balthica*, the soft-shelled clam *Mya arenaria*, and the blue mussel *Mytilus edulis*), and three polychaetes (the lugworm *Arenicola marina*, the ragworm *Hediste diversicolor*, and a small reddish worm *Heteromastus filiformis*).

For the BD program, I ignore the differences between transects and sta- tions and simply assume that 15 similar sampling units of the same surface area are taken according to a simple random sampling design. I only consider the late summer sampling data. For the SIBES program, I ignore possible bias as a result of selected sampling, for example due to non-random missing values. I treat all cores as if they were taken according to a simple random sampling design, which generally speaking yields accurate, but slightly conservative estimates of the variance [1]. Missing biomass density values are imputed on the basis of the numerical density by using a fixed mean individual weight. If a length measure is available imputing is based on a length-weight relation, see the appendix for details.

Only simple descriptive statistics, such as the mean, SE, SE/mean ratio, and Pearson correlation coefficient are used.

4 RESULTS

Trend plots show that for one species, the lugworm *Arenicola marina*, bands defined by numerical density estimates plus or minus SE rarely overlap be- tween the two programs for the period 2008-2012 (Figs. 2 and 3). Estimates for this species are much higher in the BD program. Numerical density estimates for the other two polychaete species, *Heteromastus filiformis* and *Hediste diversicolor*, are also higher in the BD program. The other (bivalve) species do not consistently differ and the square regions defined by the means plus or minus twice the SE mostly overlap with the y = x line in the bivariate plot, where estimates from the two programs are plotted against each other (Figs. 4 and 5). Biomass density estimates show much larger differences between the two programs. For all polychaetes and for the soft-shelled clam *Mya arenaria* estimates from the BD program are consistently higher (Figs 6 and 7).

.

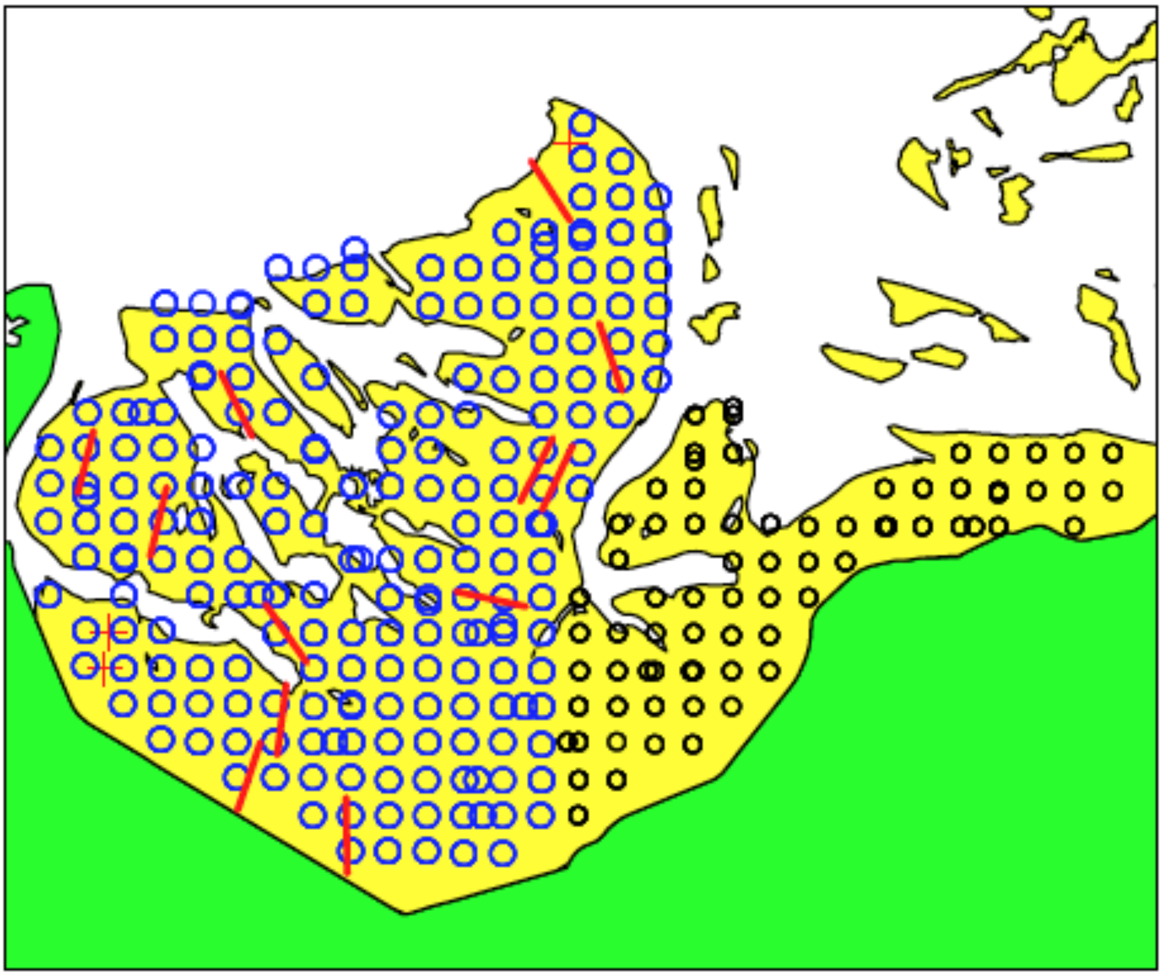

Figure 1: Map of transects (red lines), permanent quadrats (red crosses) of the BD sampling program, and SIBES points included (large blue dots) or excluded (small black dots) from the analysis.

For six species trends over time in numerical density are strongly correlated, but the correlation is weak or even absent for the other two species (Table 1). Trends over time in biomass density show a strong correlation for only three species (Table 2).

Table 1: Pearson correlation coefficients between the BD and the SIBES numerical density estimates. Period 2008-2012, n = 5.

Number	Species	Original scale	Log scale
1	*Arenicola marina*	0.71	0.82
2	*Cerastoderma edule*	0.95	0.77
3	*Ensis directus*	0.97	0.97
4	*Hediste diversicolor*	0.68	0.80
5	*Heteromastus filiformis*	-0.03	-0.04
6	*Macoma balthica*	0.84	0.82
7	*Mya arenaria*	0.85	0.93
8	*Mytilus edulis*	0.53	0.42

Table 2: As Table 1, but for biomass density estimates.

Number	Species	Original scale	Log scale
1	*Arenicola marina*	0.40	0.28
2	*Cerastoderma edule*	0.96	0.90
3	*Ensis directus*	0.95	0.98
4	*Hediste diversicolor*	0.66	0.74
5	*Heteromastus filiformis*	0.83	0.79
6	*Macoma balthica*	0.29	0.38
7	*Mya arenaria*	-0.17	-0.21
8	*Mytilus edulis*	-0.08	0.53

Standard errors and ratio between the standard error and the mean of the numerical density estimates were systematically lower in the SIBES program (Figs. 8, 9, 11, and 10). The standard error/mean ratio was about 1.5 times higher in the BD program, which is reflected in a back-transformed average log-ratio of SE/mean (SIBES/BD) that equals 0.74.

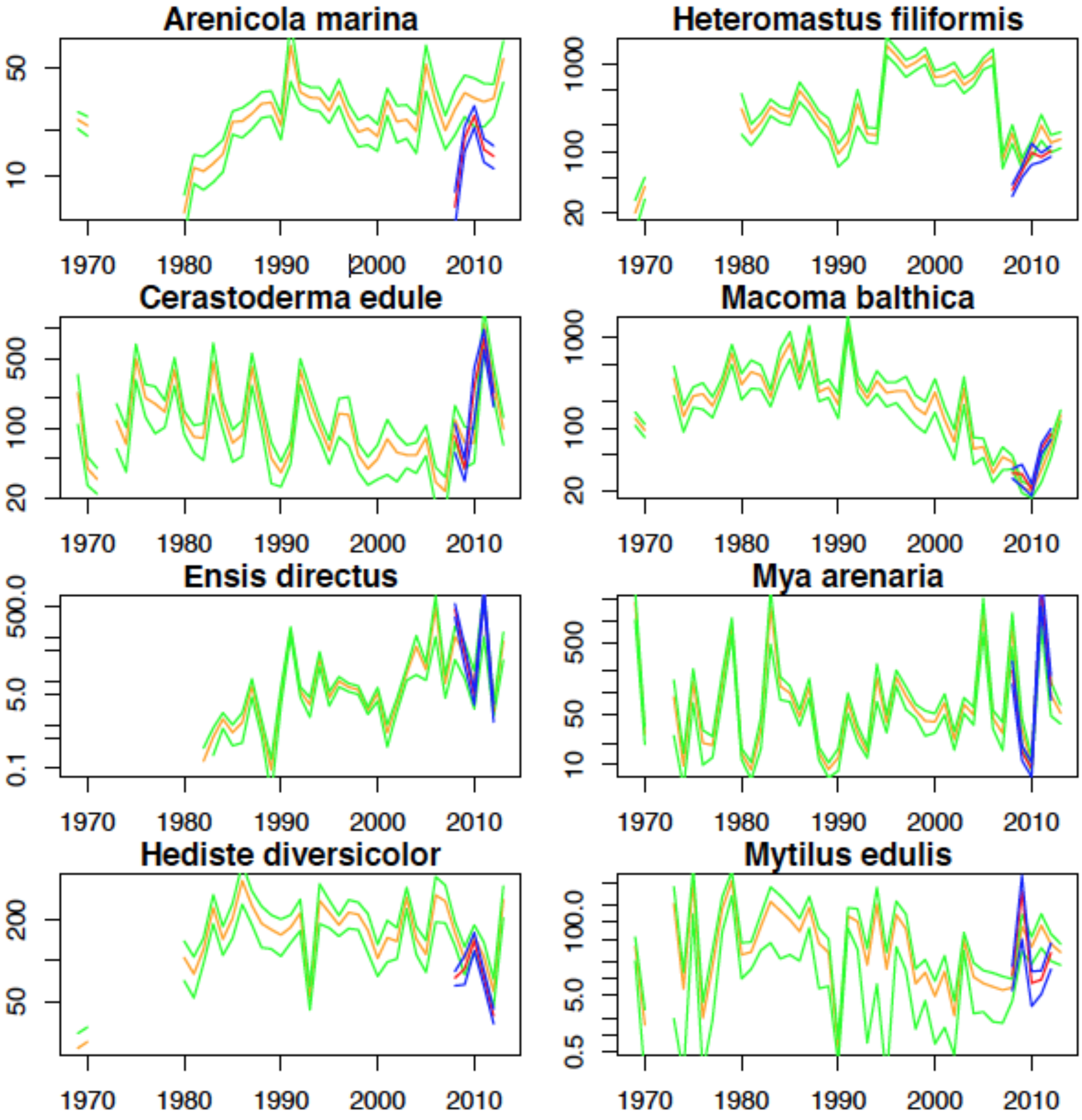

Figure 2: Trends over time in numerical density (number per m2) plus or minus SE for the BD (orange-green lines) and the SIBES (red-blue) program. Densities on log scale.

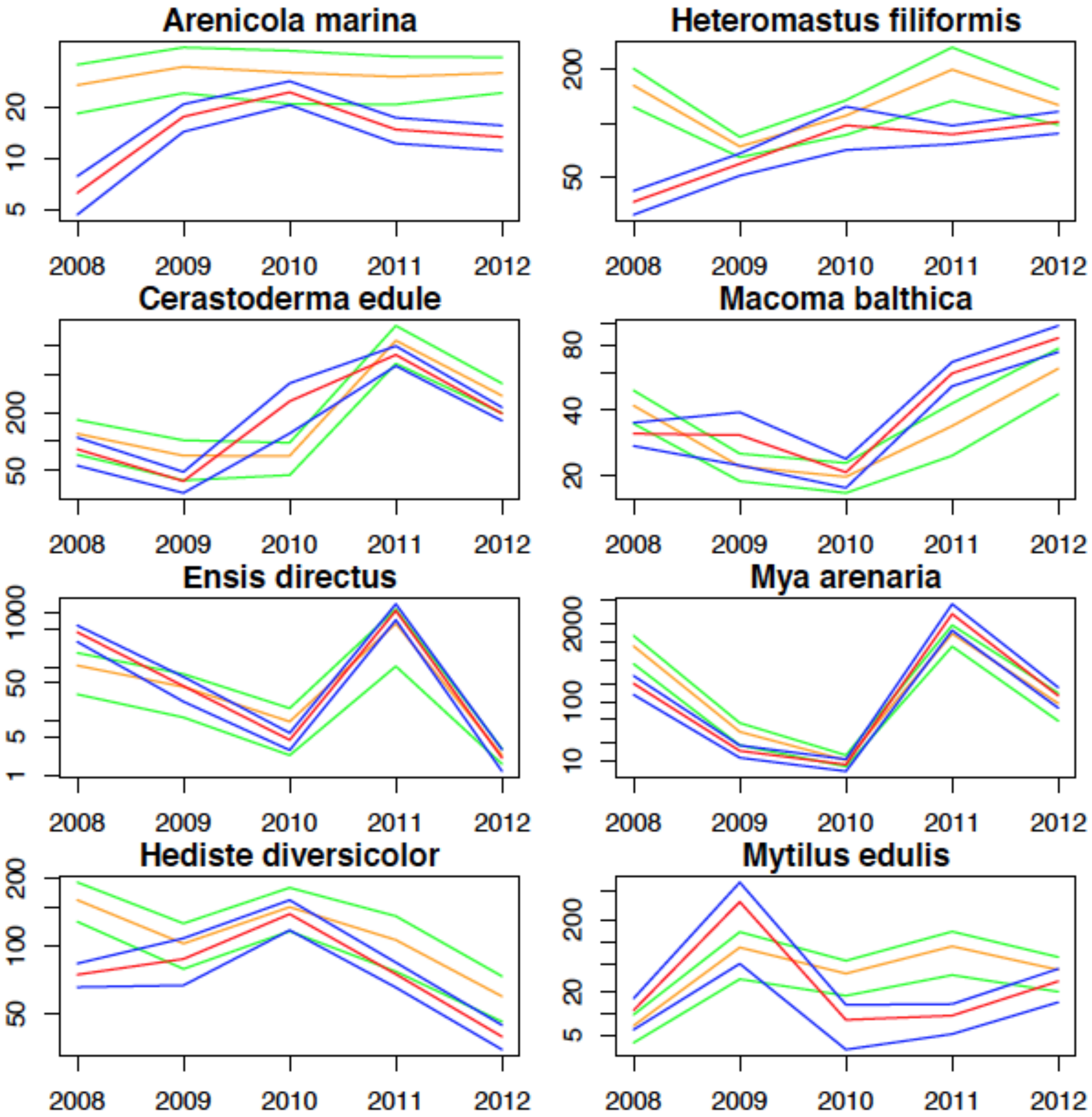

Figure 3: As Figure 2, but for period 2008-2012 only.

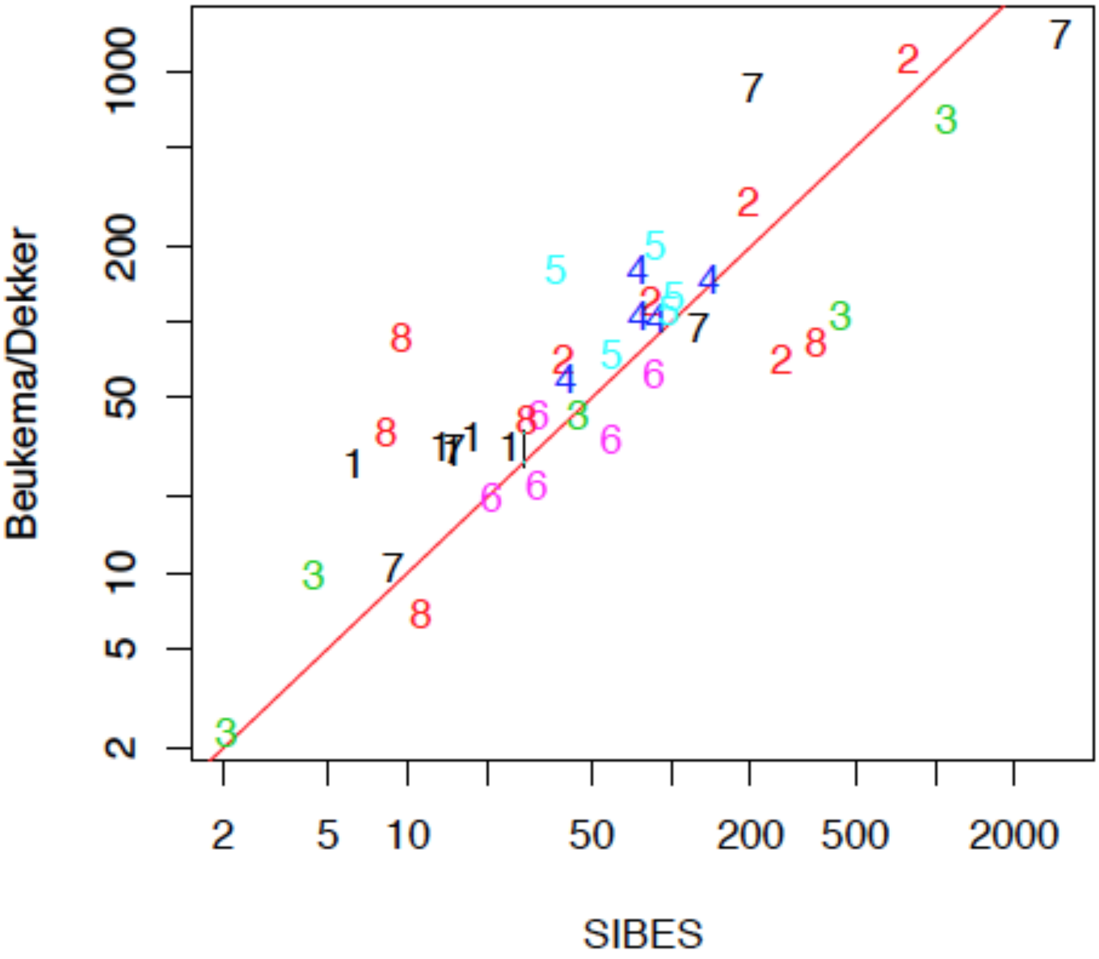

Figure 4: Estimates for the BD program plotted versus the same estimate from the SIBES program. Period 2008-2012. Log scale. Numbers refer to the species, see Table 1.

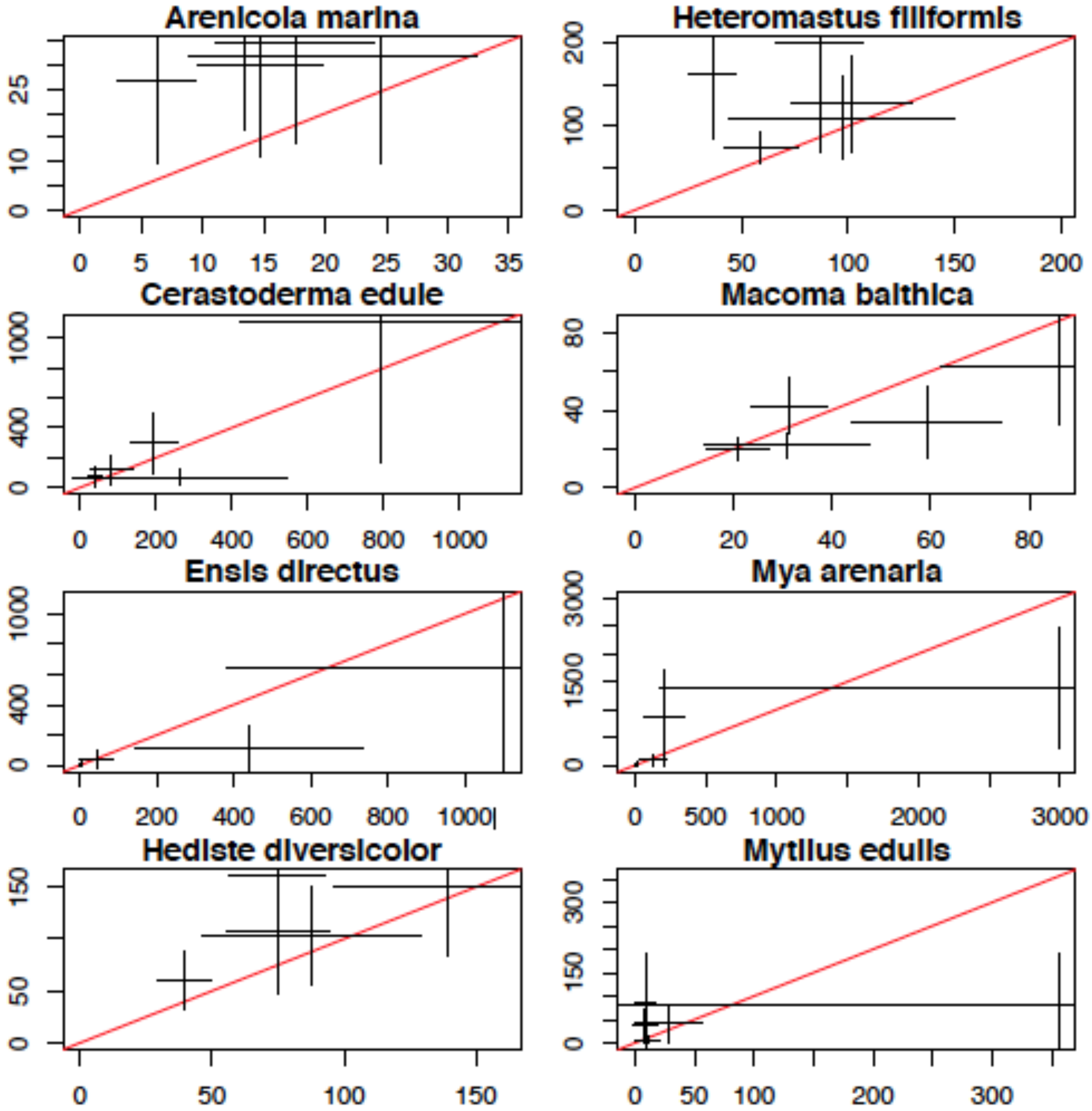

Figure 5: As Fig. 4, but now plus or minus twice the SE. Separate plot (using original scale) for each species.

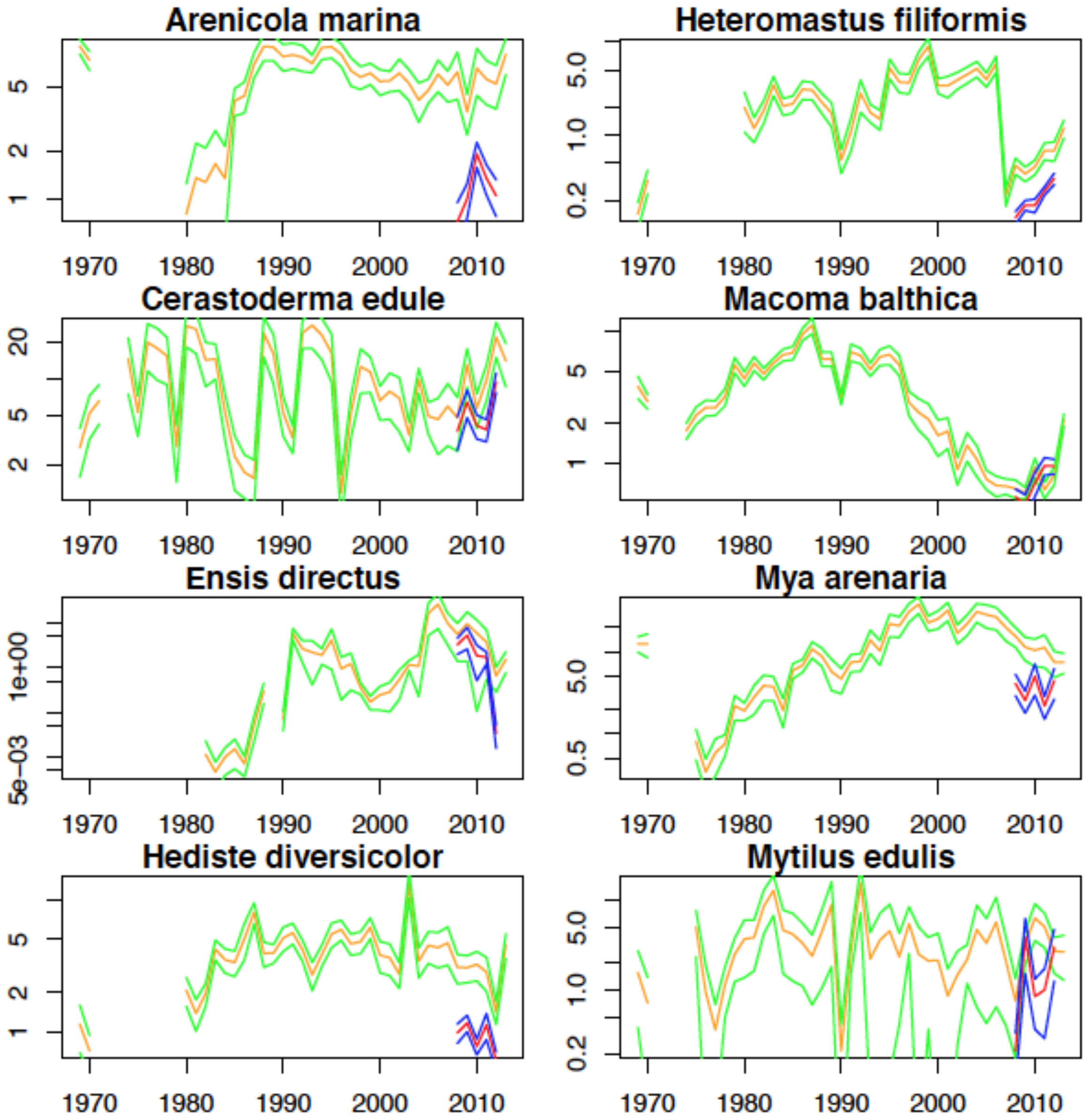

Figure 6: As Fig. 2, but with biomass density (gram per m^2).

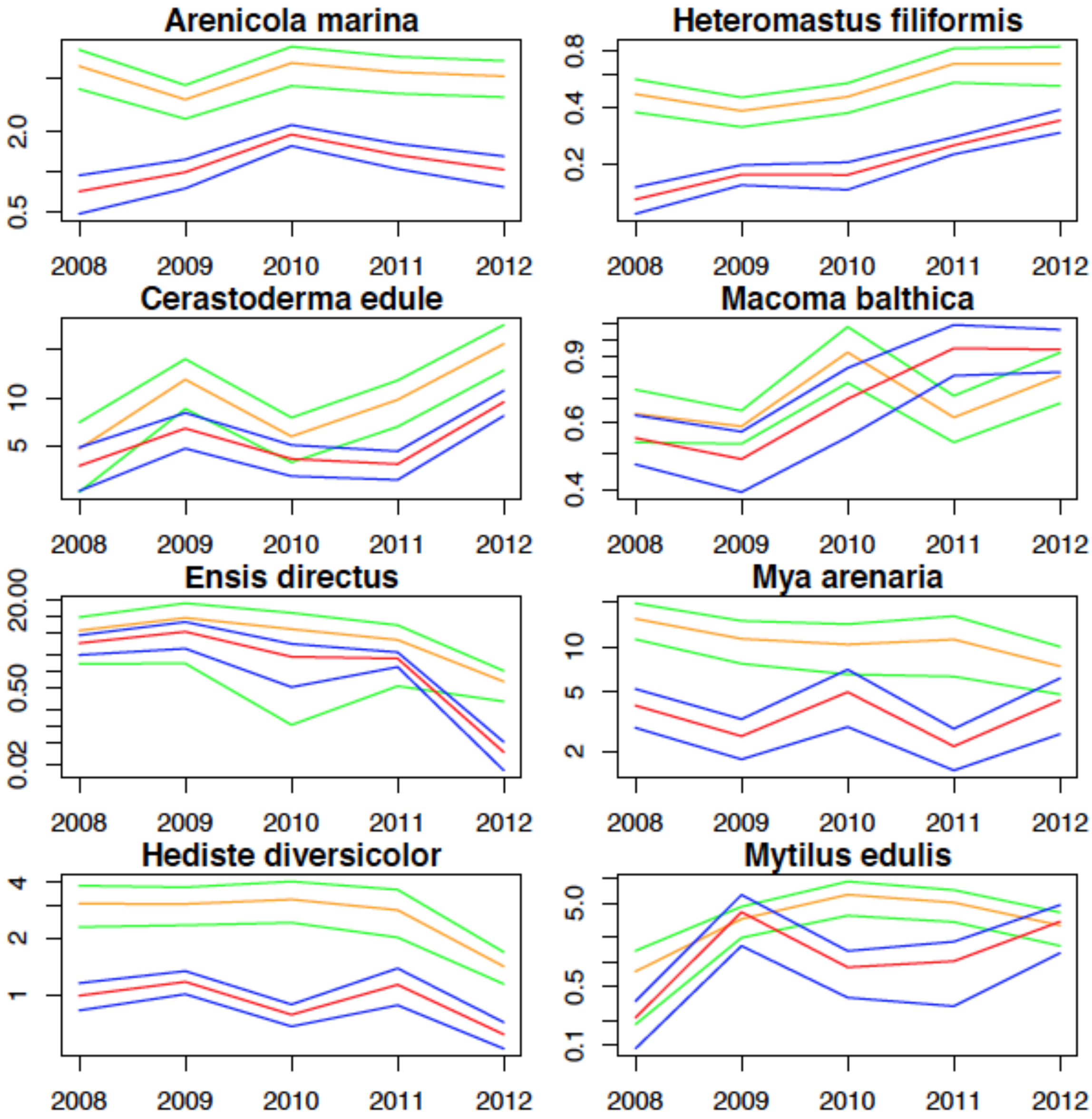

Figure 7: As Fig. 3, but with biomass density (gram per m^2).

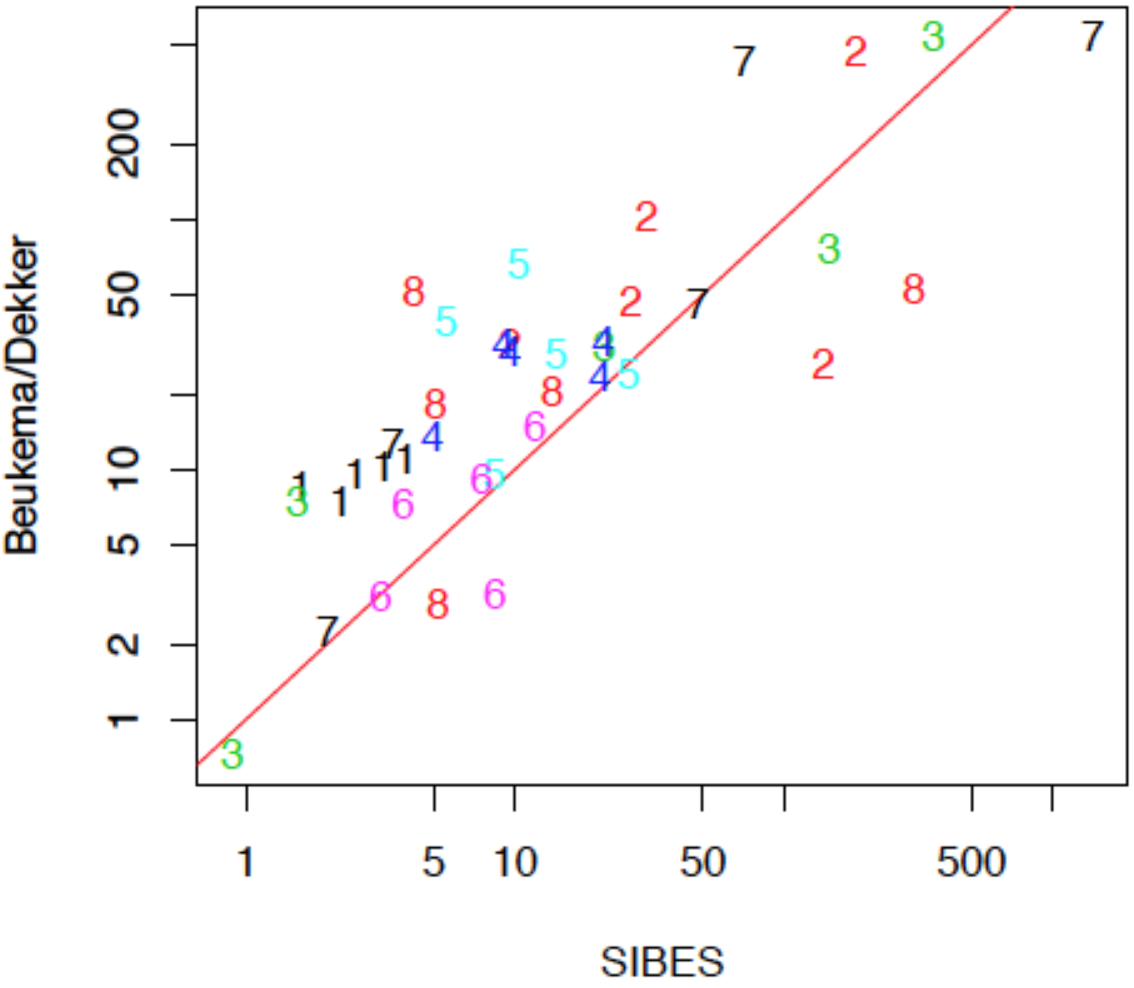

Figure 8: Estimated SEs of the mean numerical density for the BD program plotted versus those from the SIBES program. Period 2008-2012. Log scale. Numbers refer to the species, see Table 1. The red line gives the *y* = *x* line; more points are above the red line, pointing to a higher SE in the BD program.

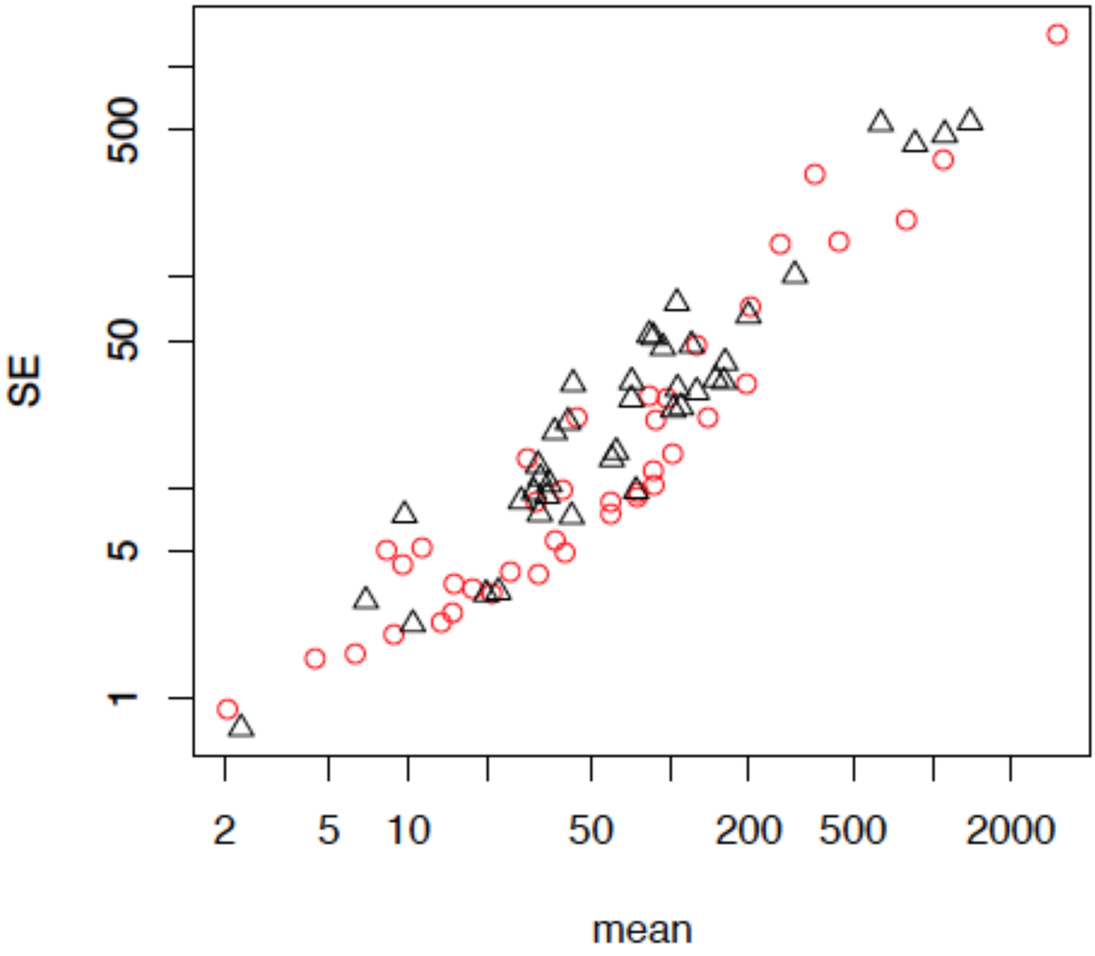

Figure 9: Estimated SEs plotted versus estimated means. Black triangles refer to the BD program, red dots to the SIBES program. Period 2008-2012. Log scale.

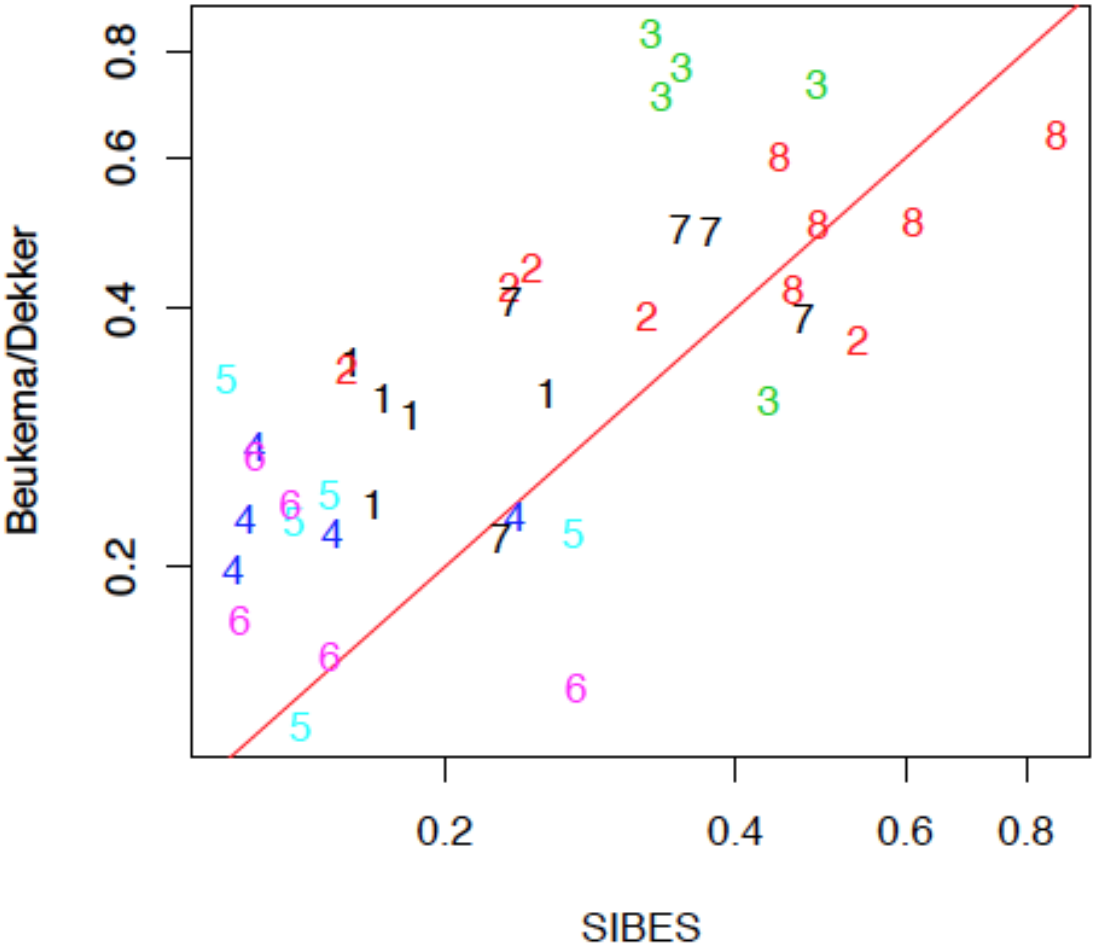

Figure 10: Estimated SE/mean of the mean numerical density for the BD program plotted versus those from the SIBES program. Period 2008-2012. Log scale. Numbers refer to the species, see Table 1. The red line gives the *y* = *x* line; more points are above the red line, pointing to a higher SE/mean in the BD program.

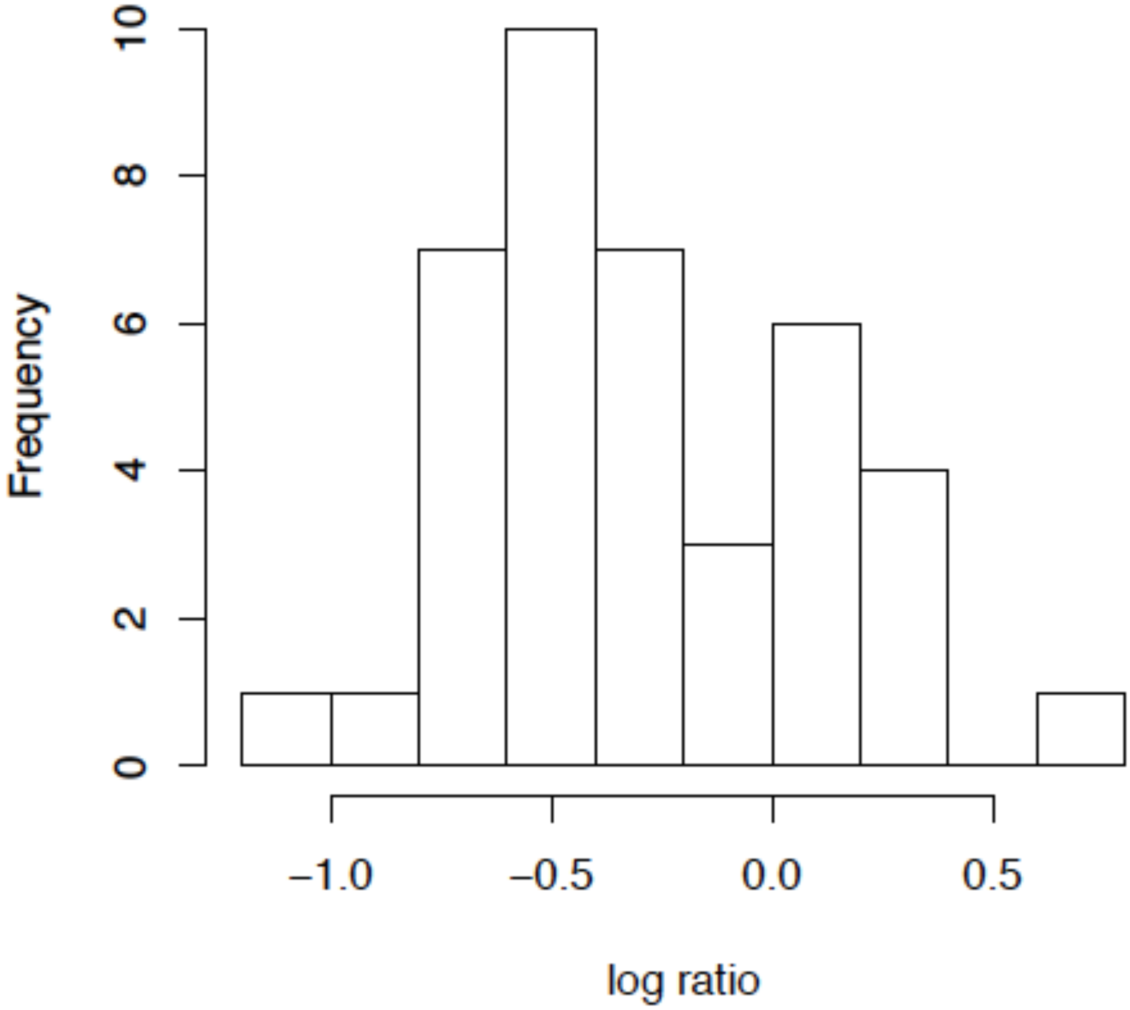

Figure 11: Frequency distribution of the log-ratio between the SE/mean for the SIBES and the SE/mean for the BD program.

5 DISCUSSION AND CONCLUSION

Numerical and biomass density estimates are on average rather similar between the two sampling programs for four bivalve species, that is all bivalve species examined apart from the soft-shelled clam *Mya arenaria*. But only two of these species, the cockle *Cerastoderma edule* and the razor clam *Ensis directus* show strongly correlated trends in both numerical and biomass density estimates. Estimated trends over time for the mussel *Mytilus edulis* and the Baltic tellin *Macoma balthica* do not correlate strongly. Apparently sampling error is larger than trend for these species.

The SIBES program provides much lower estimates for the biomass of the soft-shelled clam than the BD program, whereas estimates of numerical density are rather similar. Numerical density is mainly determined by very young animals who are extremely abundant in some years. Biomass density on the other hand is mainly determined by relatively few very large and deep living individuals. One obvious reason for the lower clam biomass estimates in the SIBES program is that sampling depth is not adequate in the SIBES program for this deep burrowing species. This explanation is in agreement with the results for the polychaete *Arenicola marina*, which is also a deep burrowing species. For this species both the estimates for the numerical density as well as for the biomass density are much lower in the SIBES program. However, it does not explain why the SIBES biomass estimates are also much lower than the BD biomass estimates for the other two polychaere species, *Hediste diversicolor* and *Heteromastus filiformis*. As we do not know the true biomass densities, it is impossible to tell what the absolute bias is for the two programs. I am tempting to believe that SIBES severely underestimates biomass density, but a more detailed look can be worthwhile.

The number of cores taken was a factor three higher in the BD program, around 650 versus around 200. The ratio for the area covered is around 2.3, since the BD program samples around 8.5 m^2 in summer and the SIBES program around 3.5 m^2. The SE/mean ratios are on average lower in the SIBES program, despite the lower effort. The SE/mean ratio in the SIBES program is on average 0.75 times the ratio in the BD program. Apparently, a geographically spreading out of sampling effort pays, as has been observed before [5].

If this would have been a game, one might say that BD wins in terms of bias, whereas SIBES wins in terms of variance. It would be nice if the best of these two programs, accuracy and efficiency, can be combined and maintained in the future. Finally, it should be stressed that average numerical and biomass density is only one aspect that can be estimated from these sampling programs. Many more issues can and have been studied, such as seasonal mortality and production, just to name something.

6 REFERENCES

Aune-Lundberg, L., Stran, G-H. (2014). Comparison of variance estimation methods for use with two-dimensional systematic sampling of land use/land cover data. Environmental Modelling & Software 61:8797.

Beukema, J.J. (1974). Seasonal changes in the biomass of the macrobenthos of a tidal flat area in the Dutch Wadden Sea. Netherlands Journal of Sea Research 8:94107.

Bijleveld, A.I., van Gils, J.A., van der Meer, J., Dekinga, A., Kraan, C., van der Veer, H.W. and Piersma, T. (2012). Designing a benthic monitoring programme with multiple conflicting objectives. Methods in Ecology and Evolution 3:526-536.

Compton, T.J. et al. (2013). Distinctly variable mudscapes: Distribution gradients of intertidal macrofauna across the Dutch Wadden Sea. Journal of Sea Research 82:103116.

Van der Meer, J. (1997). Sampling design of monitoring programmes for marine benthos: a comparison between the use of fixed versus randomly selected stations. Journal of Sea Research 37:167-179.

ACKNOWLEDGEMENTS

I thank Jan Beukema, Tanya Compton, Anne Dekinga, Rob Dekker, Jan Drent, and Anita Koolhaas for helpful comments on an earlier version.

7 APPENDIX

Table 3 gives the mean individual ash-free weight, and the parameter *a* from the length-weight relation $w = al^3$, where the median condition w/l^3 was used as the estimator of *a*. Fig. 12 shows the length-weight relationship for the five bivalve species.

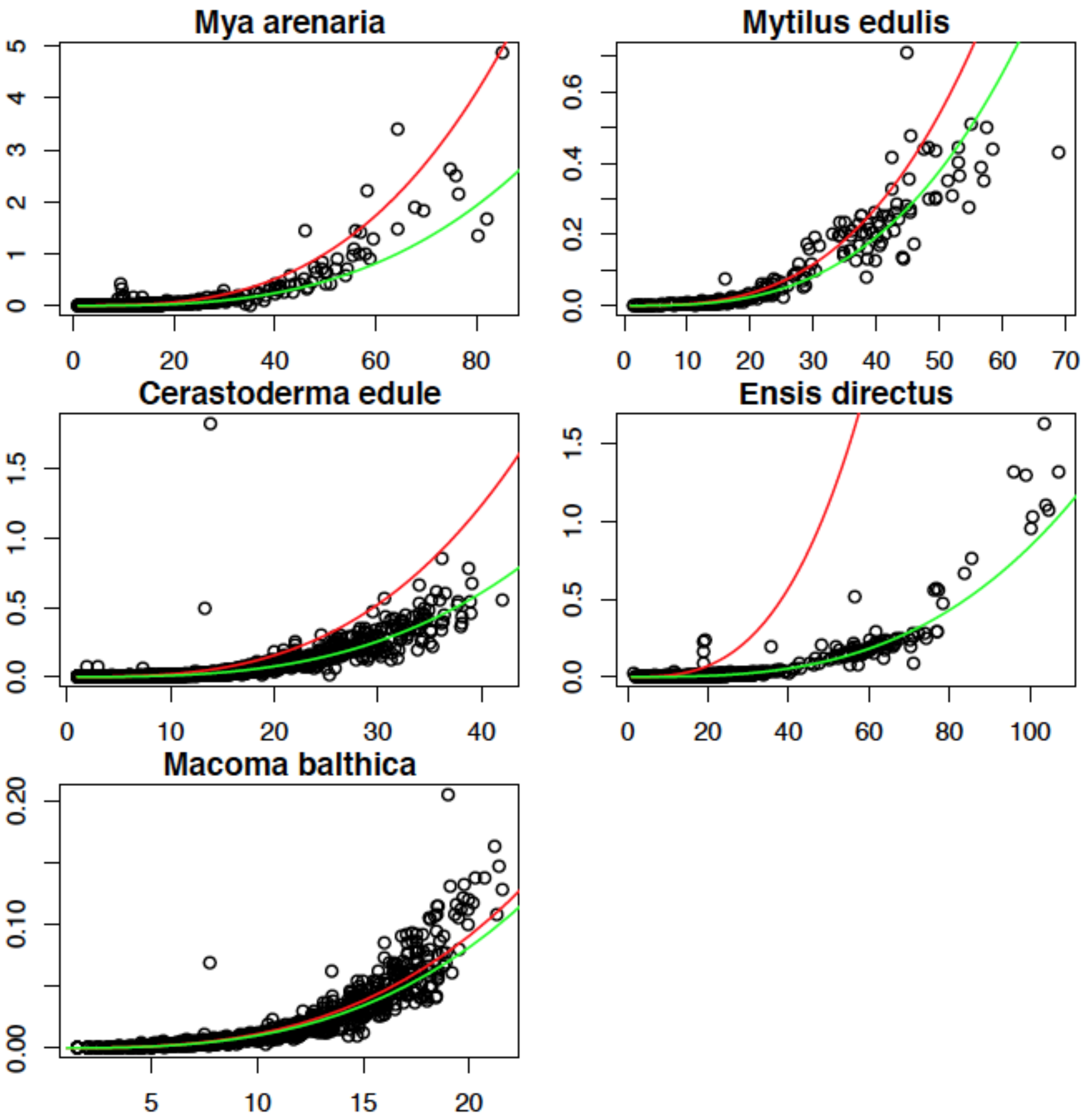

Figure 12: Length-weight relationship. Fitted lines are of the form $w = al^3$, where the mean (red line) or median (green line) condition w/l^3 was used as the estimator of *a*. Horizontal axis gives the length in mm, vertical axis the weight in g.

Table 3: Mean ash-free dry weight in g, and the parameter a from the length-weight relation $w = al^3$, where length is measured in mm and weight in g.

Species	Mean afdm	*a*
Arenicola marina	0.1073	
Cerastoderma edule	0.0527	9.483e-06
Ensis directus	0.0340	8.411e-07
Hediste diversicolor	0.0180	
Heteromastus filiformis	0.0035	
Macoma balthica	0.0180	1.018e-05
Mya arenaria	0.0651	3.766e-06
Mytilus edulis	0.0829	3.026e-06