Developing informative prior distributions for the bias of hydro-acoustic survey estimates of the biomass of the South African round herring population

C.L. de Moor[*], **J.C. Coetzee**[≠] **and D.S. Butterworth**[*]

Correspondence email: *carryn.demoor@uct.ac.za*

Prior distributions for the bias of the hydro-acoustic survey estimates of the biomass of the South African round herring are developed by considering the bias associated with the individual sources of error.

Keywords: acoustic bias, informative prior distribution, round herring, South Africa

Introduction

An assessment of the South African round herring (*Etrumeus whiteheadi)* resource is currently underway. That assessment is conditioned on (among other data) the hydro-acoustic survey estimates of biomass from the November and May/June surveys. The biases associated with those survey estimates of biomass are estimated in the assessment, informed by the use of informative prior distributions. This document details the development of each of these prior distributions[1].

A BENEFIT workshop was held in December 2000 with the aim of improving estimates of the accuracy and/or precision of survey estimates of fish stock abundances in the Benguela region (Anon. 2000). This workshop produced a summary of the main sources of error (bias) inherent in acoustic surveys for each of the main commercial species, including South African sardine (*Sardinops sagax*) and anchovy (*Engraulis encrasicolus*), and their likely ranges. The summary was based on a combination of published information and expert opinion, with "experts" including senior scientists from South Africa, Namibia, Angola and Norway. Although these error factors pertained mainly to surveys of anchovy and sardine, they are also pertinent to acoustic surveys of round herring although the sizes of the individual errors may differ. An attempt has thus been made to define the central value, and minimum and maximum possibilities for these errors for round herring. Where necessary, new error terms have been added and their ranges specified based on available knowledge. It is likely that these values may be improved as more data become available. Rationale for the derivation of parameter values describing each error factor is provided below, and should be read in conjunction with those published in Anon. (2000).

Sources of bias in the November hydroacoustic survey estimate of total biomass

For each source of error (bias), five values of limits are provided corresponding to the break-points of a trapezium-type probability density function which decreases quadratically on either side of the likely limits to a minimum and a maximum value (Table 1, Figure 1). The "likely lower" and "likely upper" values were symmetrically placed about the "likely" midpoint. The sources of bias were considered as those by which the survey estimate should be multiplied to provide an estimate of true biomass. Positive errors therefore correct for negative biases and negative errors for positive biases.

---

[*] MARAM (Marine Resource Assessment and Management Group), Department of Mathematics and Applied Mathematics, University of Cape Town, Rondebosch, 7701, South Africa.

[≠] Department of Forestry, Fisheries and the Environment – Branch Fisheries, Private Bag X2, Vlaeberg, 8018, South Africa.

[1] Parts of this text were previously recorded in the Annex of Appendix C of de Moor and Butterworth (2010); they are repeated here for completeness.

Target Strength (TS)

The TS of round herring is unknown and no published data on round herring TS exists. Barange *et al*. (1996) produced estimates of the TS of anchovy and sardine off the coast of South Africa. Preliminary unpublished data (D. Merkle pers comm) suggests that the TS of round herring should be higher than that of anchovy, but lower than that of sardine. This is supported by observations of anchovy, sardine and round herring schools where maximum volume backscattering coefficients (s_v, dB re m^2/m^3) within dense round herring schools are typically higher than those of anchovy, but lower than those of sardine. Estimates of round herring biomass are currently obtained using the Barange *et al*. (1996) TS derived for similarly sized sardine. Assuming that the likely TS value of round herring (b_{20} = -73.0) is between the TS of anchovy (b_{20} = -76) and sardine (b_{20}=-71.5) as estimated by Barange *et al*. (1996), with a 2/3 weighting towards sardine, the likely bias correction is set at 1.41 (the difference in biomass that would be obtained when using the sardine TS rather than the assumed round herring TS).

Some recent published TS data for anchovy (Sawada *et al*. 2009, Madirolas *et al*. 2017, Sobradillo *et al.* 2021), however, suggest that the TS (b_{20}) of anchovy is much higher (by up to 10 dB) than that determined for anchovy by Barange *et al.* (1996). Use of such a high TS value for anchovy is unrealistic given that it would result in anchovy biomass estimates that are about only one tenth of the current annual harvest levels. Additionally, none of these higher TS estimates appear to be routinely applied in acoustic surveys of anchovy resources. According to Doray *et al*. (2021), acoustic surveys coordinated under the ICES Working Group on Acoustic and Egg Surveys for Small Pelagic Fish continue to use older TS estimates, which were established some 40 years ago for herring and sprat, for the estimation of sardine and anchovy biomass.

Given this uncertainty, and local observations described above, the TS of round herring should be higher than that derived for anchovy (Barange *et al*. (1996). Given the similarity of sardine to round herring, compared to anchovy, the maximum bias is set considering a TS that is 1dB higher than that of anchovy, giving a maximum bias of 2.24 (the difference in biomass that would be estimated had a TS of 1dB higher than anchovy rather than the TS of sardine been used). Similarly, and given the generally similar size and morphology of round herring and sardine and their similar acoustic signature at high density, it is unlikely that the TS of round herring is higher than that derived for sardine by Barange *et al*. (1996) and hence the minimum bias is set at 1. A symmetrical distribution around the central value of 1.41 has been assumed for the likely lower and upper range (1.18 – 1.64), informed by the biomass that would result from increasing or decreasing the assumed round herring TS by 1dB. Further work to estimate the TS of round herring from *in-situ* data should be prioritised so that the effect of this bias on estimates of round herring biomass can be determined more accurately.

TS dependence on depth

This error was not considered important for anchovy and sardine at the time of Anon. (2000), but is an additional error that has been considered important in the context of round herring biomass estimation. Round herring are close to the surface at night, migrating to deeper water before dawn and staying close to the bottom during the day. At dusk they again migrate upwards in the water column. The surveys off South Africa are conducted by day and night, so round herring should be deep for approximately half of the time. However, during migration (both up and down in the water column) the tilt angle will be substantially increased and this will also lead to a reduction in TS and lower biomass estimates than those obtained assuming a TS regression derived from fish close to the surface.

Published findings for herring (Ona 2003) suggest a strong depth dependence of TS, with halving of TS between the surface and a depth of 200m, and with the steepest decrease in TS in the first few (upper 50) meters of the water column. For anchovy, Zhao *et al*. (2008) estimated that night-time estimates of biomass would be 58% higher than day-time ones for fish migrating from a depth of 50m during the day to 20m at night. If their regression is applied to fish migrating from 200m during the day to 20m at night (as is typical for adult round herring), night-time estimates would be some 3.66 times higher than day-time estimates. A value of 3.66 was therefore used to set the maximum error value and because it is absolutely certain that TS decreases with depth and during vertical migration, the minimum value was set at 1.

Round herring are found distributed across the entire continental shelf, from close inshore to beyond the 200m depth contour but the maximum correction will likely only apply to a small portion of the biomass. Applying the Zhao *et al*. (2008) regression instead to a depth of 50m and 150m results in a likely error range of between 1.59 and 3.05 and the likely value was taken as the midpoint of this range. However, this error will only apply during the day (approximately 12 hours) and during periods of migration (approximately 4 hours) and therefore the likely lower, likely, likely upper and maximum values were multiplied by 0.67 to gives values of 1.06, 1.55, 2.04 and 2.44 respectively. Further work that compares day-time and night-time estimates to those derived using both day and night data may provide useful insights into the scale of this negative bias.

Calibration

Calibration errors are likely to be similar to those of anchovy and sardine, and have therefore not been changed from those agreed by Anon. (2000).

Attenuation

The effect of attenuation on round herring estimates has not been determined. This error factor is, however, most likely to be substantially less than that of denser schooling sardine (which averages at around 1.15 for November surveys). A maximum error of 1.15 has thus been selected, with 1 for the minimum error (this cannot be less), with a symmetrical distribution around the likely value of 1.075. It would be possible to estimate the effect of attenuation on round herring biomass estimates using a similar method to that used for quantifying attenuation effects in dense schools of sardine.

Target Identification

The same parameter values specified for the minimum and maximum target identification error for anchovy have been applied (de Moor *et al*. 2020), but the likely range has been increased. This is to account for larger overestimation of round herring (when the assumption is made that deep targets during the day are most likely to be round herring, but could possibly include horse mackerel). Conversely, diving behaviour at dawn may result in (larger relative to other pelagic species) under-sampling of round herring in some trawls, with consequent underestimation of their biomass.

Weather effects

Weather effects are likely to play a larger role when fish are deeper (vessel pitch and roll effects are amplified at depth), and as such a slightly wider likely range compared to that for anchovy has been selected. Nevertheless, the maximum effect is assumed to be the same as for anchovy (de Moor *et al*. 2020). It is likely that the size of this error may be more accurately estimated in the future.

Sources of bias in the May/June hydroacoustic survey estimate of recruitment

The error ranges for some sources of bias are assumed to differ between the November survey which targets the total biomass and the May/June survey which targets the recruits of the year (Table 2). The rationale for the derivation of values of the parameters describing the distributions of the error factors that differ in range from that applicable to November surveys is provided below. Again this should be read in conjunction with the information published in Anon. (2000).

TS dependence on depth

Juvenile round herring tend to be distributed closer inshore during the recruit survey than adults are during the November survey, so that the maximum error is likely to be lower. However, given that the largest reduction in TS occurs within the first 50m of the water column, this is still considered to be an important source of negative bias. To calculate the maximum error and likely error range during the day, we used a maximum depth of 150m, instead of 200m and a likely depth range of 30 to 100m, instead of 50 to 150m. This results in a maximum error for day-time of 3.05 and likely error range of 1.59 to 2.38. The duration of daytime was reduced to 9 hours and the time for vertical migration was also reduced from 4 hours to 3 hours to reflect the shorter vertical migration range and crepuscular period. The values determined for the day were therefore halved (12/24 hours) and result in a maximum value of 1.53. However, halving the likely lower value results in a value that is <1, so this value is set equal to the minimum range = 1 and the likely value (1.09) is set at the midpoint of the likely lower and likely upper (1.19) range. Further analyses using the mean depth of round herring recruits during the May survey may improve estimates of the likely effect.

Attenuation

As for sardine recruits, attenuation biases are likely to be smaller than those applicable to denser schooling adults. No information on the likely reduction in this effect for round herring recruits is available, but the maximum error is assumed to be 10% and the distribution of the error is assumed to be symmetrical around a midpoint of 5%.

Target Identification

The range of this error should be smaller than that for November surveys, given the closer inshore distribution and smaller overlap between round herring recruits and adult horse mackerel. Similarly, the under-sampling of round herring during trawling is likely to be less than that for adults because the slower swimming juvenile round herring are less likely to avoid capture than adults. The maximum and minimum values have therefore been reduced, although the distribution is still centred on 1 (corresponding to an equal chance of under- or over estimation).

Weather effects

The maximum range for this error is assumed to be similar for May and November surveys, although the likely range has been halved to account for the greater inshore distribution of recruits relative to adults, and consequent reduction in the depth range during the day.

A probability density function for all sources of survey bias

The acoustic errors above were separated into two different types: constant or variable (Tables 1 and 2). Constant error relates to a factor whose value is not known exactly, but whatever it is, it remains the same for each year. In contrast, variable error relates to a factor whose true value will change from one year to the next.

The probability density functions (pdf) for the bias in the hydro-acoustic surveys that relates directly to the acoustic survey (rather than, for example, the coverage of the stock), k_{ac} and k_{rec} for total and recruitment biomass, respectively, are based on all these errors. The CV associated with variable error factors only, ϕ_{ac} and ϕ_{rec}, are then additionally calculated to be included in the likelihood when conditioning the assessment to the hydro-acoustic survey estimates of abundance.

Ten thousand samples were drawn from the individual trapezium-type pdfs for each source of error. These sampled numbers were then all inverted so as to correspond directly to the model parameters k_{ac} and k_{rec} which account for bias in the model biomass rather than in the survey estimated biomass (as considered in Tables 1 and 2). The inverted sample of constant errors is denoted as $C_j^k, j = 1, \cdots, 10000$ for the constant error factor k (target strength and calibration – beam factor), and the inverted sample of variable errors – or errors that vary inter-annually – is denoted as $V_j^k, j = 1, \cdots, 10000$ for error factor k. Histograms of the samples are given in Figure 2.

As the survey biomass estimates are considered in log-space in the likelihood (de Moor 2022), the distributions of k_{ac} and k_{rec} and ϕ_{ac} and $\phi_{rec}$are similarly displayed in log-space. Histograms of the $ln(C_j^k)$ and $ln(V_j^k)$ samples are given in Figure 3. The median of $ln(k_{ac})$ and $ln(k_{rec})$ is subsequently calculated as the median of the sample: $\sum_k ln(C_j^k) + \sum_k ln(V_j^k), j = 1, \cdots, 10000$, which is -1.092 and -0.715, respectively. The standard deviation of $ln(k_{ac})$ is based only on the log of the constant factors, and thus it is calculated as the standard deviation of the sample: $\sum_k ln(C_j^k), j = 1, \cdots, 10000$, which is 0.182 for November and for May/June. The prior distributions for $ln(k_{ac})$ and $ln(k_{rec})$ are accordingly taken to be normally distributed, i.e. $ln(k_{ac}) \sim N(-1.092, 0.182^2)$ and $ln(k_{rec}) \sim N(-0.715, 0.182^2)$ (Figure 4). The standard deviation of the log of the variable factors is considered similar to additional standard error in the likelihood calculation, and is calculated as the standard deviation of the sample: $\sum_k ln(V_j^k)$, $j = 1, \cdots, 10000$, giving $\phi_{ac} = 0.304$ for November and $\phi_{rec} = 0.196$ for May/June.

References

Anon. 2000. Survey Errors Workshop. Benguela Environment and Fisheries Interaction and Training programme report. 4-7 December, Breakwater Lodge, Cape Town.

Barange M, Hampton I, Soule M. 1996. Empirical determination of *in situ* target strengths of three loosely aggregated pelagic fish species. *ICES Journal of Marine Science* 53:225-232

de Moor 2022. Assessment of South African round herring: Initial results. Department of Forestry, Fisheries and the Environment: Branch Fisheries Document FISHERIES/2022/JUL/SWG-PEL/21.

de Moor CL, Coetzee JC, Butterworth DS. 2020. Calculating an informative prior distribution for the bias of hydro-acoustic survey estimates of the biomass of the South African anchovy. Department of Forestry, Fisheries and the Environment: Branch Fisheries Document FISHERIES/2020/MAY/SWG-PEL/39.

de Moor CL, Butterworth DS. 2010. Further work towards the assessment of the South African round herring (Etrumeus whiteheadi) resource using data from 1987 to 2010. Marine and Coastal Management: Branch Fisheries Document MCM/2010/SWG-PEL/46.

Doray M, Boyra G, van der Kooij J (Eds.). 2021. ICES Survey Protocols – Manual for acoustic surveys coordinated under ICES Working Group on Acoustic and Egg Surveys. ICES Techniques in Small Pelagic Fish (WGACEGG). 1st Edition. Marine Environmental Sciences Vol. 64. 100 pp. https://doi. org/10.17895/ices.pub.7462

Madirolas A, Membiela FA, Gonzalez JD, Cabreira AG, dell'Erba M, Prario IS, Blanc S. 2017. Acoustic target strength (TS) of argentine anchovy (*Engraulis anchoita*): the nighttime scattering layer. *ICES Journal of Marine Science*. 74(5) 1408–1420. https://doi.org/10.1093/icesjms/fsw185

Ona E. 2003. An expanded target-strength relationship for herring. *ICES Journal of Marine Science* 60:493-499.

Sawada K, Takahashi H, Abe K, Ichii T, Watanabe K, Takao Y. 2009. Target-strength, length, and tilt-angle measurements of Pacific saury (*Cololabis saira*) and Japanese anchovy (*Engraulis japonicus*) using an acoustic-optical system. *ICES Journal of Marine Science*. 66:1212-1218.

Sobradillo B, Boyra G, Pe´rez-Arjona I, Martinez U, Espinosa V. 2021. Ex situ and in situ target strength measurements of European anchovy in the Bay of Biscay. *ICES Journal of Marine Science.* 78(3): 782-796.

Zhao X, Wang Y and Dai F. 2008. Depth-dependent target strength of anchovy (*Engraulis japonicus*) measured in situ. *ICES Journal of Marine Science*. 65: 882–888.

Table 1. Individual error factors for November hydro-acoustic surveys of round herring biomass, where the values define trapezium-type pdfs. Note that these error factors apply to the observed biomass, i.e. they reflect the inverse of the multiplicative bias (applied to predicted biomass) in de Moor (2022). Constant error relates to a factor whose value is not known exactly, but whatever it is, it remains the same for each year, while variable error relates to a factor whose true value will change from one year to the next.

Error	Minimum	Likely (lower)	Likely (midpoint)	Likely (upper)	Maximum	Nature
Target Strength	1.00	1.18	1.41	1.64	2.24	Constant
Target strength dependence on depth	1.00	1.06	1.55	2.04	2.44	Variable
Calibration						
(On-axis sensitivity)	0.90	0.95	1.00	1.05	1.10	Variable[2]
(Beam factor)	0.75	0.90	1.00	1.10	1.25	Constant
Attenuation	1.00	1.05	1.075	1.10	1.15	Variable
Target Identification	0.50	0.80	1.00	1.20	1.50	Variable[2]
Weather Effects	1.00	1.10	1.20	1.30	1.50	Variable

Table 2. Individual error factors for May/June hydro-acoustic surveys of round herring recruitment, where the values define trapezium-type pdfs. Note that these error factors apply to the observed recruitment, i.e. they reflect the inverse of the multiplicative bias (applied to predicted recruitment) in de Moor (2022).

Error	Minimum	Likely (lower)	Likely (midpoint)	Likely (upper)	Maximum	Nature
Target Strength	1.00	1.18	1.41	1.64	2.24	Constant
Target strength dependence on depth	1.00	1.00	1.09	1.19	1.53	Variable
Calibration						
(On-axis sensitivity)	0.90	0.95	1.00	1.05	1.10	Variable[2]
(Beam factor)	0.75	0.90	1.00	1.10	1.25	Constant
Attenuation	1.00	1.025	1.05	1.075	1.10	Variable
Target Identification	0.60	0.85	1.00	1.15	1.40	Variable[2]
Weather Effects	1.00	1.05	1.10	1.15	1.50	Variable

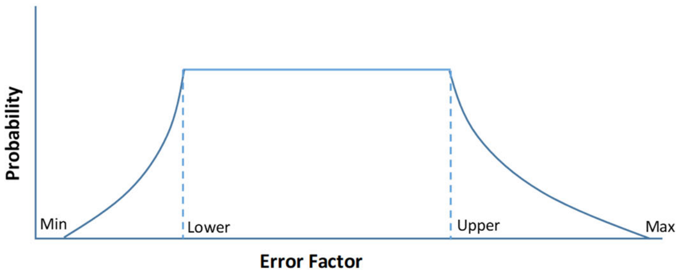


Figure 1. Trapezium-type distribution function used in error model, with parabolas used to describe the decrease from the maximum probability, but not necessarily equally on both sides.

---

[2] This was recorded in Anon. (2000) as random error, denoting that it could be positive or negative rather than purely positive or negative.

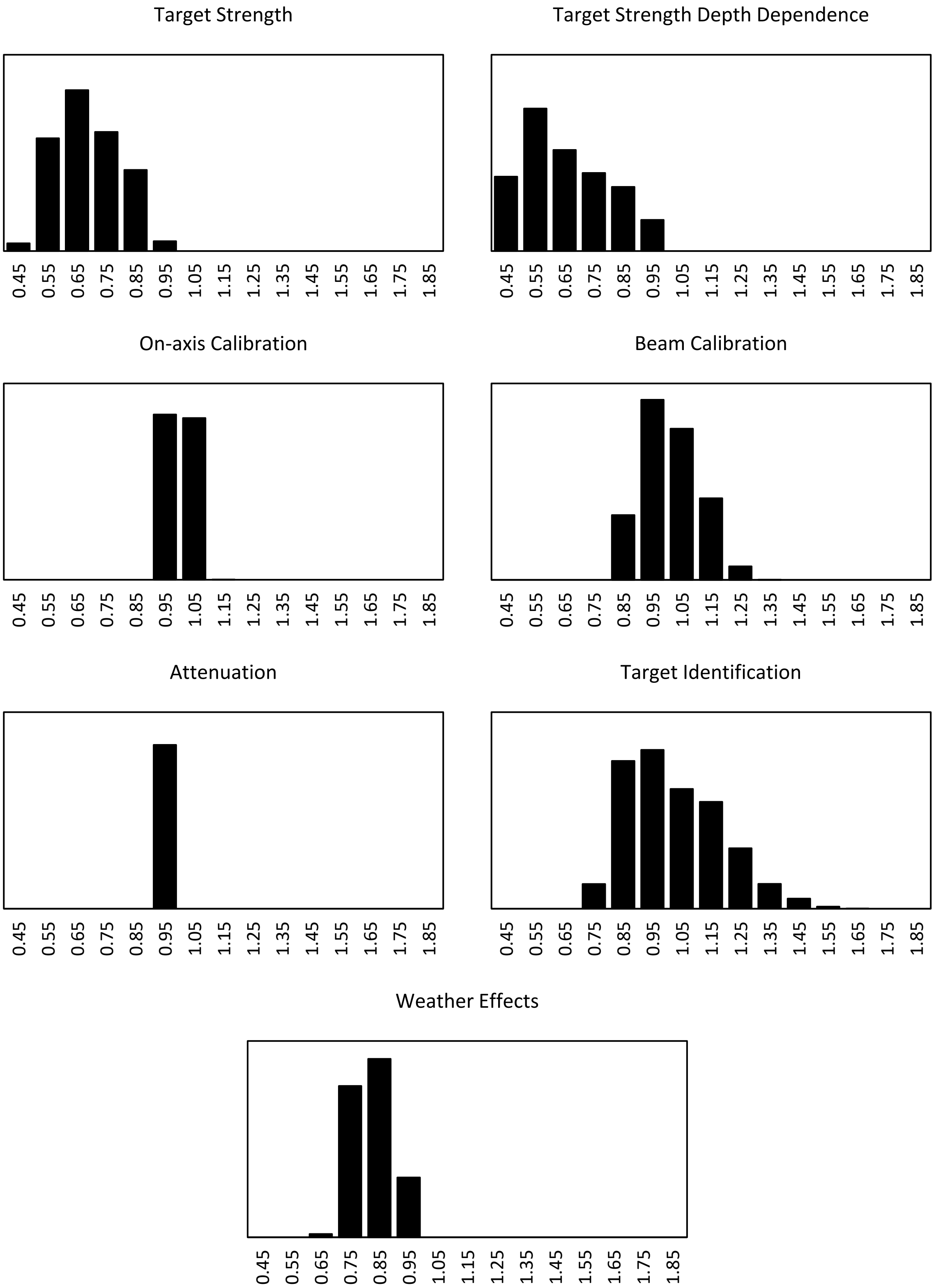
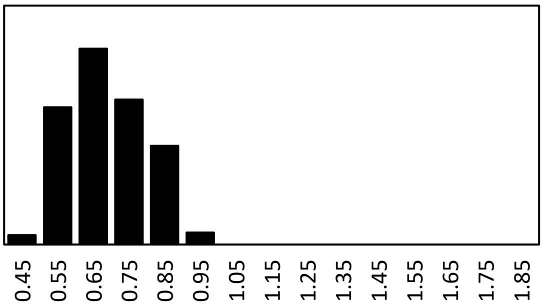
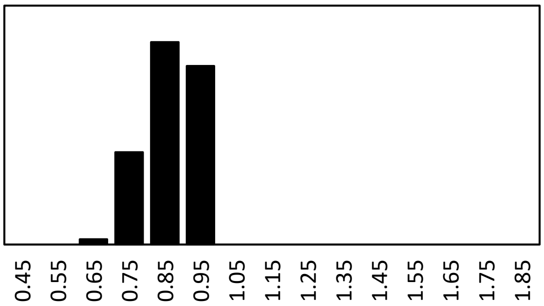


Figure 2a. The histograms of 10 000 samples of the individual error factors C_j^k and V_j^k for **November** survey errors.

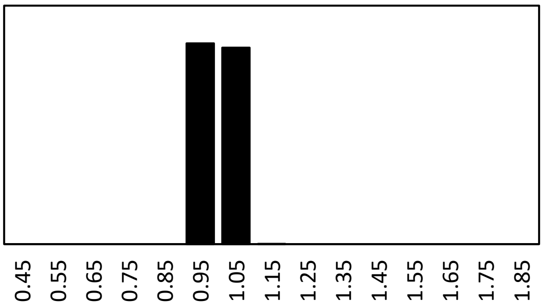
Target Strength



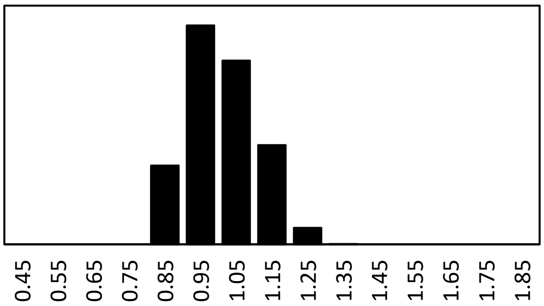
Target Strength Depth Dependence



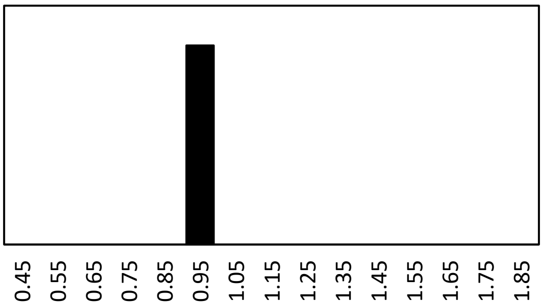
On-axis Calibration



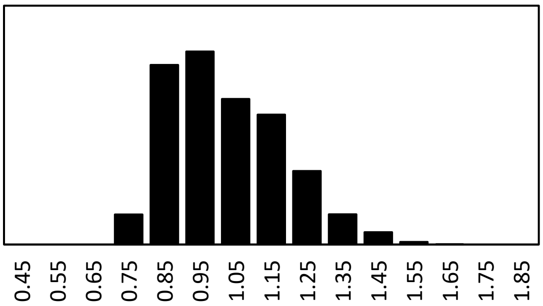
Beam Calibration



Attenuation



Target Identification



Weather Effects

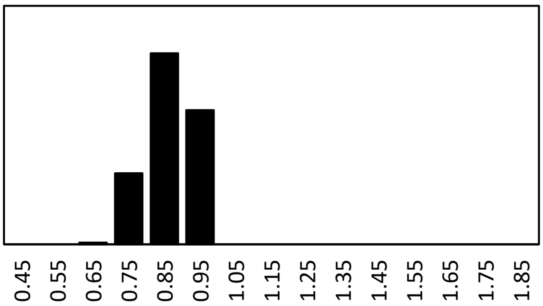


Figure 2b. The histograms of 10 000 samples of the individual error factors C_j^k and V_j^k for **May/June** survey errors.

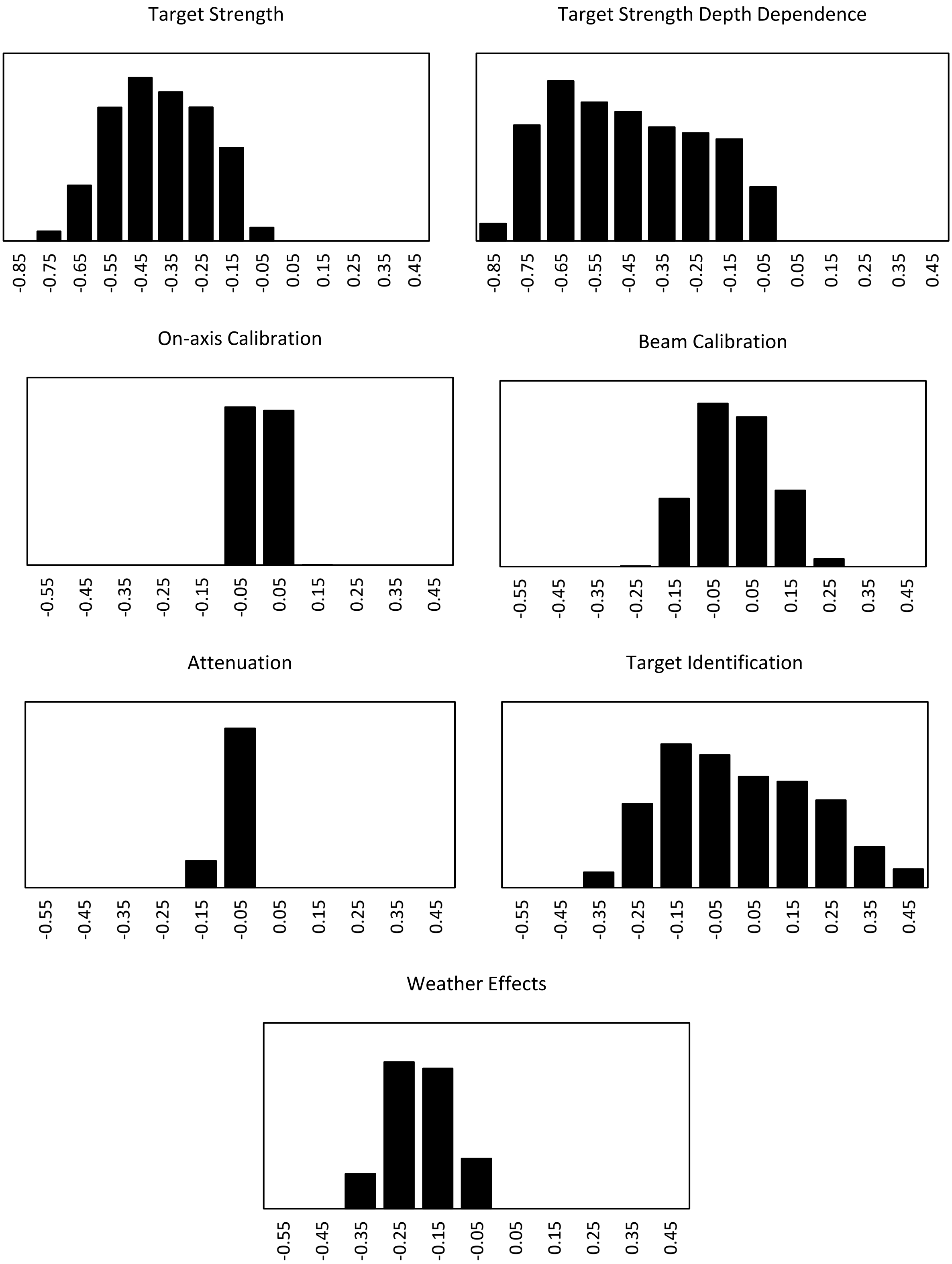
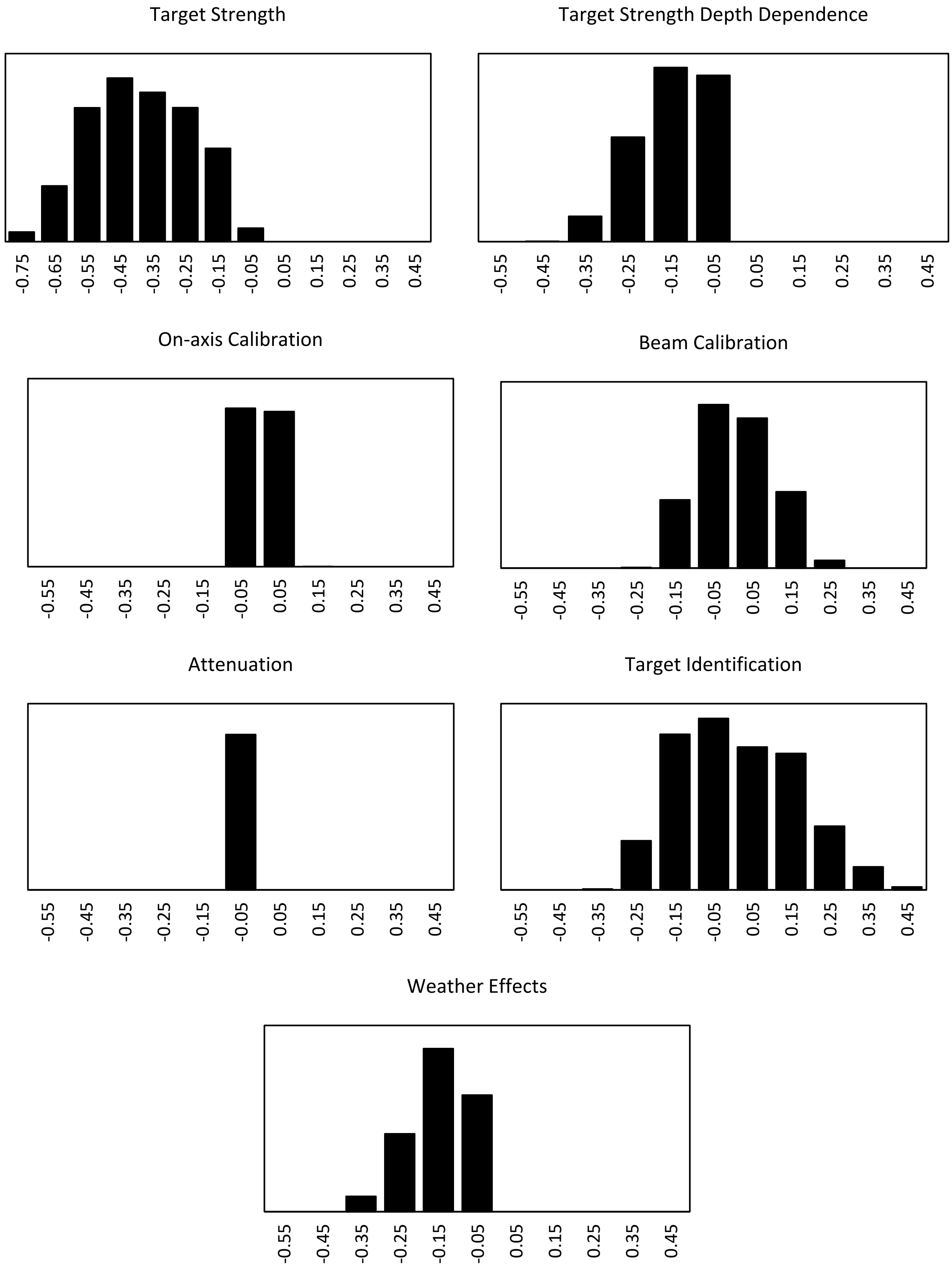


Figure 3a. The histograms of 10 000 samples of the individual error factors $\boldsymbol{ln\left(C_j^k\right)}$ and $\boldsymbol{ln\left(V_j^k\right)}$ for **November** survey errors.


Target Strength

Target Strength Depth Dependence

On-axis Calibration

Beam Calibration

Attenuation

Target Identification

Weather Effects

Figure 3b. The histograms of 10 000 samples of the individual error factors $\boldsymbol{ln}\left(\boldsymbol{C}_j^k\right)$ and $\boldsymbol{ln}\left(\boldsymbol{V}_j^k\right)$ for **November** survey errors.

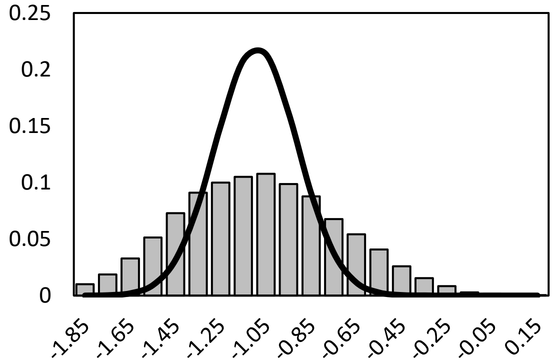


Figure 4a. The histogram of the combined errors $\sum_k ln\left(C_j^k\right) + \sum_k ln\left(V_j^k\right)$ (bars) and the assumed prior distribution for November $ln(k_{ac}) \sim N(-1.092, 0.182^2)$ (line) excluding the variable error variance which is incorporated into additional variance.

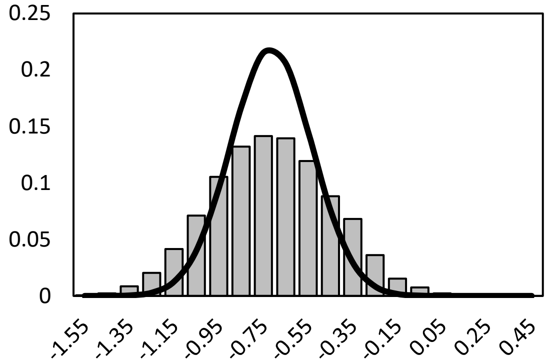


Figure 4b. The histogram of the combined errors $\sum_k ln\left(C_j^k\right) + \sum_k ln\left(V_j^k\right)$ (bars) and the assumed prior distribution for May/June $ln(k_{rec}) \sim N(-0.715, 0.182^2)$ (line) excluding the variable error variance which is incorporated into additional variance.