Calculating an informative prior distribution for the bias of hydro-acoustic survey estimates of the biomass of the South African anchovy

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

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

A prior distribution for the bias of the hydro-acoustic survey estimates of the biomass of the South African anchovy is developed by considering the bias associated with the individual sources of error.

Introduction

The quantitative assessment of South African anchovy resource has traditionally assumed that the time-series of Daily Egg Production Method (DEPM) indices of abundance provide a time-series of absolute estimates of anchovy spawner biomass (e.g. de Moor 2020a). Robustness to this assumption (assuming either over- or under-estimation by the DEPM-based indices of abundance) is typically tested during OMP development (e.g. de Moor 2019). By fixing the assumed bias of this DEPM time-series, the assessment is able to estimate the bias associated with the hydro-acoustic survey estimates of abundance from the November survey (k_{ac}^{A}) with an uninformative prior distribution. The most recent assessment of anchovy estimated this bias at 0.68 for the baseline model and 0.98 for the model with an alternative maturity ogive (de Moor 2020a, b).

A 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 (Table 1). 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.

This document first presents an update to the sources of error inherent in the hydro-acoustic surveys for South African anchovy; it then provides a calculation of an informative prior distribution for the bias in hydro-acoustic survey estimates of anchovy, using a similar method to that previously employed for South African sardine (de Moor and Butterworth 2016).

Sources of survey bias

For each source of error (bias), five values of limits were previously agreed (Table 1) 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 (Figure 1). The agreed “likely lower” and “likely upper” values were equally distributed 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

---

[*] 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.

for negative biases and negative errors for positive biases. For South African anchovy, the main sources of the November hydro-acoustic survey bias as estimated by Anon. (2000) were dominated by target strength uncertainty (Table 1).

These minimum, maximum and likely limits have now been reconsidered for each source of bias as well as the shape of the target strength distribution (Table 2). Following revision of the distribution for target strength bias (see below) used to estimate the biomass of South African anchovy (Coetzee *et al*. 2008), the overall bias is now dominated by a negative bias due to weather effects and variable target identification uncertainty.

Target strength

The target strength regression used for anchovy has been changed since that applied at the time of Anon. (2000). The previous “likely” error of 1.4 was based on the difference in biomass that would be expected from applying the older herring target strength expression (Reynisson 1993), rather than the Barange *et al.* (1996) target strength regression for anchovy. The anchovy biomass is now estimated using the Barange *et al*. target strength expression and hence the “likely” error should be changed to 1. Analyses of target strength data collected during subsequent surveys of anchovy off South Africa provide support for the continued use of the Barange *et al*. target strength expression. However, due to the high variability in South African anchovy target strength estimates (Hampton *et al*. 2012), there is a need for further target strength data collection under controlled conditions.

Apart from the Barange *et al*. target strength expression, which is also used by other fisheries management organisations (e.g., Zwolinski *et al*. 2012), most ICES-affiliated countries use a higher target strength expression (ranging from 0.8dB to 4.9 dB higher). Many of these higher target strength regressions are still based on data collected for herring and sprat in the European region (Masse *et al.* 2018). A few recent, direct estimates of the target strength for *Engraulis encrasicolus* and other anchovy species have been published (Madirolas *et al*. 2017, Zhao *et al.* 2008, Sawada *et al*. 2009) but as far as could be determined, none of these are routinely applied in acoustic surveys of anchovy resources. Setting limits for the target strength error, based on other species’ target strength seems counterintuitive as those other species’ target strength expressions remain in use mainly to ensure biomass estimates remain comparable. Similarly, it would seem inappropriate to base the South African anchovy target strength error range on published estimates of the same species if those target strength estimates are not used routinely elsewhere.

Fish density is inversely related to the average backscattering cross section ($\overline{\sigma}_{bs}$) where $\overline{\sigma}_{bs} = 10^{TS/10}$. Rather than estimating the bias associated with target strength using a trapezium-type pdf, it is therefore probably more defensible to estimate target strength uncertainty by assuming a normal distribution in log$_{10}$ space about the centroid of the Barange *et al*. target strength estimate and its standard error, $S_{y.x}$:

$$S_{y.x} = \sqrt{S_y^2(1-r^2)\frac{n-1}{n(n-2)}} = 0.0259\text{dB}$$

where

S_y^2 is estimated from the original experimental weight-normalised target strength at length data used to estimate the regression between TS_{kg} and $\log_{10}$ total length (cm), and

r^2 is the correlation coefficient squared, i.e. target strength error is drawn from 10^x, where $x \sim N(0, 0.0259^2)$. In normal space, therefore, the bias correction is centred on 1 with a standard error of 0.061. This differs from the Barange *et al*. (1996) published standard error of the predicted target strength of 0.7 dB or 0.17 in normal space, which does not seem appropriate to the circumstances of use here.

Calibration

Although the acoustic system has been upgraded from the Simrad EK500 used at the time of the workshop in 2000 to the Simrad EK60, no further work on calibration sources of error has been conducted locally. Calibrations continue to be conducted in the same way by the same acoustic engineer from Fisheries Resource Surveys, and although calibration accuracy is likely to have improved slightly with the newer generation echo-sounders and software advances, their accuracy is dependent on weather conditions during the actual calibration. There is thus no basis to modify the minimum, maximum or limits of this bias distribution at present.

Target identification

Without further analyses into the recent efficiency of identification trawls to classify small pelagic fish, the minimum, maximum and likely limits specified for this bias remain unchanged. For mixed species aggregations, the error in identification is likely related to the relative biomasses of the co-occurring species. For this reason, there may be a defensible argument that for South African anchovy at least, the error range could be smaller than the minimum or larger than the maximum during periods of very low or very high sardine or round herring biomass, respectively. Off South Africa, however, periods of low sardine biomass have often corresponded with high round herring biomass.

Weather effects

The values for the survey bias distribution parameters in Table 1 related to weather effects were based on work conducted in the 1980s (Dalen and Løvik 1981). Recent estimates of the error associated with bad weather and increased aeration and transducer pitch and roll, however, suggest that the maximum error is unlikely to be greater than 1.5 and that the minimum error is 1 (Shabangu *et al*. 2014). However, because surveys are never conducted entirely in good conditions, avoiding any wind or swell-induced air bubble attenuation of the acoustic signal is unrealistic. This is the reason that the minimum value was previously set at 1.01 and this value should be retained. Estimates of average error at wind speeds of 20 and 30 knots were 1.02 and 1.28 respectively (Shabangu *et al*. 2014). Summer surveys off South Africa are frequently conducted during adverse weather conditions with acoustic surveys typically suspended only when the wind speed exceeds 30 knots. Wind speeds of 20 to 30 knots are experienced frequently, so that an average error between these limits could be used as the likely value (1.15). This corresponds with what was previously used. The likely lower and upper values could then be set at 1.02 and 1.28 and the min and max at 1.01 and 1.5 respectively.

Sampling error

A further source of survey error is the survey sampling CV, which is assumed to be lognormally distributed. Anon. (2000) listed the average 1984 – 1999 CV of November hydro-acoustic surveys of 18%[1] to be included in the calculation of anchovy survey bias. However, in the anchovy assessment, the annual CVs are used directly in the likelihood when conditioning the model to the survey data and thus the survey bias, k_{ac}^{A}, need only account for errors other than this sampling error.

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 function (pdf) for the bias in the hydro-acoustic survey that relates directly to the acoustic survey (rather than, for example, the coverage of the stock), k_{ac}^{A}, is based on all these errors. The CV associated with variable error factors only, ϕ_{ac}^{A}, is 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 (or $\log_{10}$-Normal distribution for target strength) for each source of error. These sampled numbers were then all inverted so as to correspond directly to the model parameter k_{ac}^{A} which accounts 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 2020), the distributions of k_{ac}^{A} and ϕ_{ac}^{A} 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 $ln(k_{ac}^{A})$ 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 -0.158. The standard deviation of $ln(k_{ac}^{A})$ 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 0.112. The prior distribution for $ln(k_{ac}^{A})$ is accordingly taken to be normally distributed, i.e. $ln(k_{ac}^{A}) \sim N(-0.158, 0.112^2)$ (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 $\phi_{ac}^{A} = 0.197$.

References

---

[1] The average CV from 1984-2019 is 17%.

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 and Soule MA. 1996. Empirical determination of in situ target strengths of three loosely aggregated pelagic fish species. *ICES Journal of Marine Science.* 53: 225 – 232.

Coetzee JC, Merkle D, Cunningham CL, Twatwa NM, Barange M, Butterworth DS. 2008. Refined estimates of South African pelagic fish biomass from hydro-acoustic surveys: quantifying the effects of target strength, signal attenuation and receiver saturation. *African Journal of Marine Science.* 30(2): 205-217

Dalen J and Løvik A. 1981. The influence of wind induced bubbles on echo integration. *Journal of the Acoustical Society of America.* 69(3): 1653–1659.

de Moor CL and Butterworth DS. 2016. Assessment of the South African sardine resource using data from 1984-2015: Results at the joint posterior mode for the two mixing-stock hypothesis. Department of Agriculture, Forestry and Fisheries: Branch Fisheries Document FISHERIES/2016/JUL/SWG-PEL/22REV2.

de Moor CL. 2019. Is OMP-18 robust to alternative assumptions about anchovy dynamics?. Department of Agriculture, Forestry and Fisheries: Branch Fisheries Document FISHERIES/2019/JUL/SWG-PEL/17.

de Moor CL. 2020a. Assessment of the South African anchovy resource using data from 1984-2019: initial results. Department of Forestry, Fisheries and the Environment: Branch Fisheries Document FISHERIES/2020/JAN/SWG-PEL/03.

de Moor CL. 2020b. Further results pertaining to the South African anchovy assessment. Department of Forestry, Fisheries and the Environment: Branch Fisheries Document FISHERIES/2020/MAR/SWG-PEL/15.

Hampton I, Soule MA and Nelson JC. 2012. In situ estimates of sardine and anchovy target strength derived from data collected in South Africa, 2004-2009. Benguela Current Commission report: BCC Project: LMR/PEL/09/07 Acoustic survey methods and associated errors. April 2012.

Madirolas A, Membiela FA, Gonzalez JD, Cabreira AG, dell'Erba M, Prario IS and 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

Massé J, Uriarte A, Angélico MM and Carrera P (Eds.) 2018. Pelagic survey series for sardine and anchovy in ICES subareas 8 and 9 – Towards an ecosystem approach. *ICES Cooperative Research Report* No. 332. 268 pp. https://doi.org/10.17895/ices.pub.4599

Reynisson P. 1993. In situ target strength measurements of Icelandic summer spawning herring in the period 1985–1992.ICES 1993 CM/B: 40. 15 pp.

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

Shabangu FW, Ona E and Yemane D. 2014. Measurements of acoustic attenuation at 38 kHz by wind-induced air bubbles with suggested correction factors for hull-mounted transducers. *Fisheries Research.* 151: 47– 56.

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.

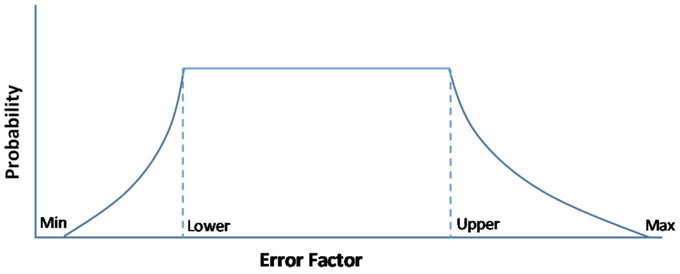
Zwolinski JP, Demer DA, Byers KA, Cutter GR, Renfree JS, Sessions TS and Macewicz BJ. 2012. Distributions and abundances of Pacific sardine (*Sardinops sagax*) and other pelagic fishes in the California Current Ecosystem during spring 2006, 2008, and 2010, estimated from acoustic–trawl surveys. *Fishery Bulletin.* 110:110–122.

Table 1. Individual error factors as estimated by Anon. (2000) for hydro-acoustic surveys of anchovy biomass, where the values define trapezium-type pdfs (Anon. 2000). Note that these error factors apply to the observed biomass, i.e. they reflect the inverse of the multiplicative bias factor k_{ac}^{A} in the assessment model.

Error	Minimum	Likely (lower)	Likely (midpoint)	Likely (upper)	Maximum	Nature
Target Strength	0.80	1.15	1.40	1.65	2.00	Constant
Calibration						
(On-axis sensitivity)	0.90	0.95	1.00	1.05	1.10	Variable[3]
(Beam factor)	0.75[2]	0.90	1.00	1.10	1.25	Constant
Target Identification	0.50	0.90	1.00	1.10	1.50	Variable[3]
Weather Effects	1.01	1.05	1.15	1.25	2.00	Variable

Table 2. Revised individual error factors for hydro-acoustic surveys of anchovy biomass, where the values define trapezium-type pdfs (Anon. 2000). Note that these error factors apply to the observed biomass, i.e. they reflect the inverse of the multiplicative bias factor k_{ac}^{A} in the assessment model.

Error	Minimum	Likely (lower)	Likely (midpoint)	Likely (upper)	Maximum	Nature
Target Strength	0.84	0.94	1.00[4]	1.06	1.19	Constant
Calibration						
(On-axis sensitivity)	0.90	0.95	1.00	1.05	1.10	Variable[3]
(Beam factor)	0.75	0.90	1.00	1.10	1.25	Constant
Target Identification	0.50	0.90	1.00	1.10	1.50	Variable[3]
Weather Effects	1.01	1.02	1.15	1.28	1.50	Variable



---

[2] This was originally reported as 0.8 in Anon. 2000, but subsequently corrected (I. Hampton pers. comm.).

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

[4] Normal distribution $N(0, 0.0259^2)$ about the centroid of the Barange *et al*. target strength estimate. In linear ($\overline{\sigma}_{bs}$) space, this corresponds to a centroid of 1, with likely values corresponding to $10^{\mp 0.0259}$, i.e. 1 standard error, and minimum and maximum values corresponding to $10^{\mp 0.077}$, where 0.077 is 3 standard errors.

Figure 1. Trapezium-type distribution function used in error model.

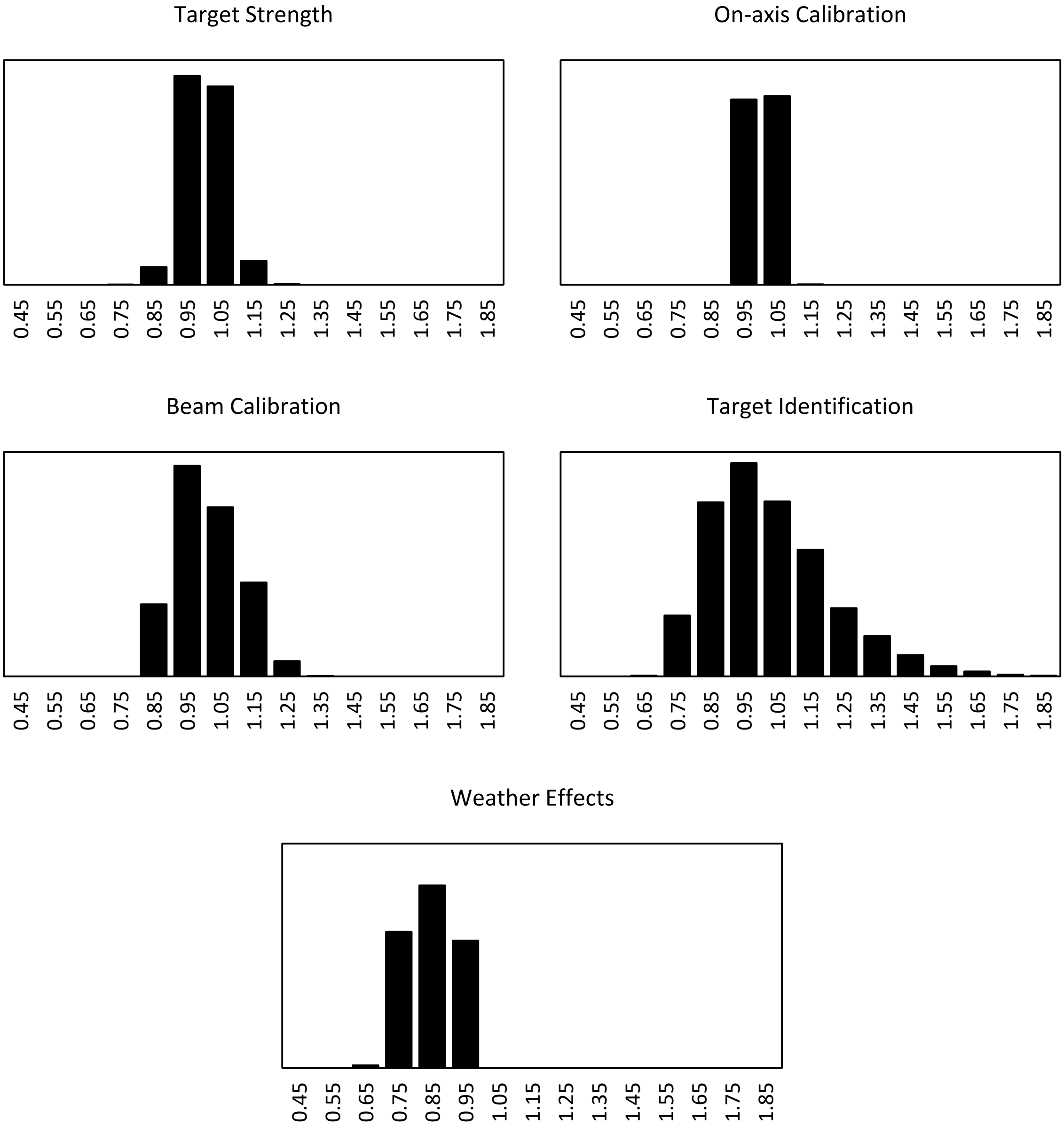


Figure 2. The histograms of 10 000 samples of the individual error factors C_j^k and V_j^k.

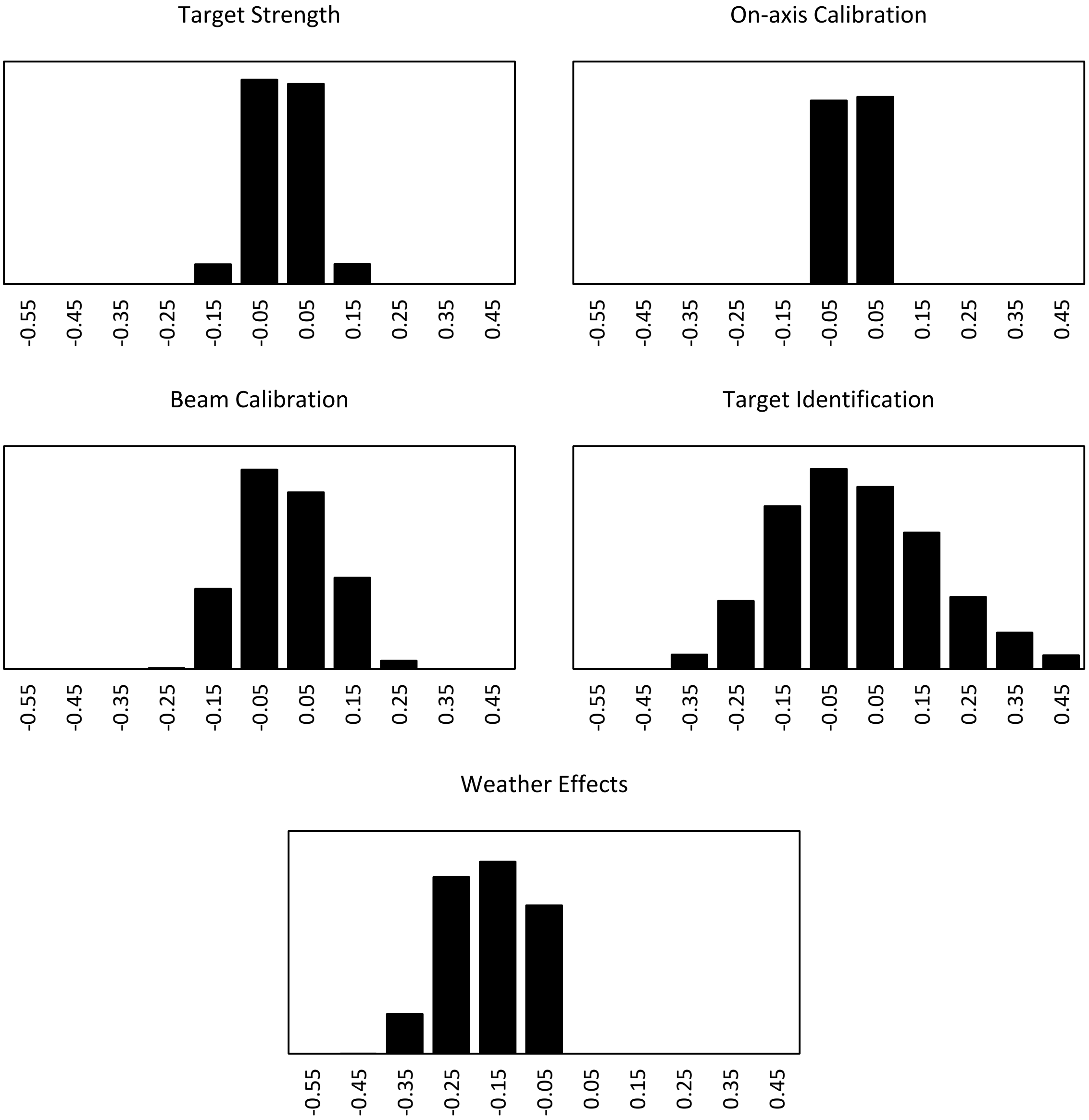


Figure 3. The histograms of 10 000 samples of the individual error factors $ln(C_j^k)$ and $ln(V_j^k)$.

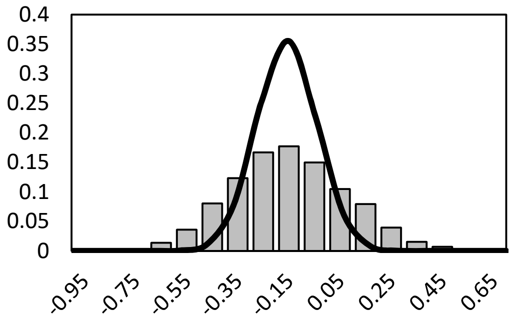


Figure 4. The histogram of the combined errors $\sum_k ln(C_j^k) + \sum_k ln(V_j^k)$ (bars) and the assumed prior distribution for $ln(k_{ac}^A) \sim N(-0.158, 0.112^2)$ (line) excluding the variable error variance which is incorporated into additional variance.