Statistics of seafloor backscatter measured with multibeam sonar systems

John Penrose, Alexander Gavrilov and Iain Parnum

Curtin University of Technology, Centre for Marine Sci & Tech, GPO Box U1987, 6845 Perth, WA, Australia
j.penrose@cmst.curtin.edu.au

A number of theoretical models for seafloor backscatter statistics developed for the recent years show a good agreement with experimental measurements made with sonar systems. However, methods of data collection used in multibeam systems are commonly not taken into consideration when analysing backscatter statistics. Using data collected with a Reson SeaBat 8125 system and based on theoretical considerations, it is shown that the seafloor backscatter strength derived from the peak intensity measured as a single value for each beam leads to considerable backscatter overestimation at oblique angles of incidence when the beam footprint is much larger than the insonification area. This occurs because fluctuations of the peak intensity are extreme value distributed, which can be approximated by the Fisher-Tippet distribution. The mean value of the Fisher-Tippet distribution depends on the number of statistically independent samples taken to detect the peak intensity, i.e. approximately on the ratio of the footprint and insonification areas. On the other hand, the average backscatter energy provides an almost unbiased estimate of the seafloor backscatter strength. The Gamma distribution is demonstrated to be a good approximation for statistics of the backscatter energy, especially for oblique angles of incidence.

1   Introduction

Modern multibeam sonar (MBS) systems collect seafloor backscatter data in addition to high-resolution bathymetry, which enhances substantially their capability as a means for seafloor classification. Seafloor backscatter characteristics, such as the backscatter strength, its angular dependence and statistical moments derived from MBS data can be used to characterize seafloor properties by comparing to theoretical predictions. However, in order to make an adequate comparison, it is essential to understand the relationship between the backscatter characteristics expected from physical models and those measured with a MBS system. Older MBS systems provided one backscatter value per each beam and each ping. Newer sonars, such as Reson SeaBat 81 and 71 series models, also collect backscatter waveforms. There are basically three ways to process the backscatter waveform in individual beams: 1) to pick a single value, which is usually the peak amplitude, 2) to calculate a single value, which can be the average amplitude or backscatter energy, and 3) to take every sample from the waveforms and treat coinciding samples of overlapping signals from adjacent beams in one way or another to form a sidescan-like signal. These coinciding samples have generally different amplitudes in different beams because the backscatter envelopes are distorted by the receive beam pattern. The most appropriate way to exclude overlapping samples is to accurately correct the backscatter envelope for the beam pattern. However, this is not trivial to implement, because the angular dependence of the beam pattern should be correctly projected into the time domain, which requires knowing the bottom relief within the footprint. All the above-mentioned procedures may affect statistical characteristics of backscatter measured with a MBS system and, hence, distort estimates of the seafloor backscatter strength and its dependence on incidence angle. Effects of MBS backscatter data processing on the estimates of seafloor backscatter strength are considered in this paper using numerical modelling and data collected with a Reson SeaBat 8125 system over a coastal shelf area off Esperance in Western Australia.

2   Effects of beam pattern

To derive the seafloor backscattering coefficient from MBS data, the instantaneous backscatter intensity measured on the receive array must be corrected for the transmission loss and the effective insonification area, which is an integral of the product of transmit and receive beam patterns projected onto the seafloor surface, taken over the area insonified by the transmitted pulse [1]. Based on the energy conservation principle, the energy of backscatter signals in individual beams is expected to be proportional to the whole surface integral of the beam patterns. Corrections for the insonification and footprint areas, applied to the instantaneous intensity and energy respectively, are often made assuming a rectangular shape of the beam patterns, which may result in certain errors. For a narrow along-track pattern of the transmit beam, the main lobe can be well approximated by a Gaussian shape $\Psi(\varphi) = \exp\{\varphi^2/2\sigma_\varphi^2\}$, where φ is the elevation angle, $\sigma_\varphi = [12\log(10)/5]^{-1/2}\Delta\varphi$ and $\Delta\varphi$ is the -3-dB full width of the beam. Integration of $\Psi(\varphi)$ gives $[5\pi/6\log(10)]^{1/2}\Delta\varphi \approx 1.066\Delta\varphi$, i.e. a value of about 0.28 dB higher than that obtained for a rectangular beam pattern. A similar error was predicted in [2].

The effect of the across-track beam pattern of MBS receive beams on the backscatter energy and instantaneous intensity is more complex, because the insonification area is limited by both the length of the transmitted pulse and the across-track width of the beam footprint, which depend on the incidence angle, transmitted pulse length and sea depth. A number of different approaches have been suggested to model the influence of insonification area on the backscatter intensity at small and moderate incidence angles, when the width of the insonification area is larger than or comparable to the across-track width of the beam footprint [2]. To model the effect of receive beam pattern on the backscatter energy and instantaneous intensity, we used an approach similar to the method suggested in [2] for nearly vertical incidence, but formulated in way suitable for modelling beam pattern effects at an arbitrary steering angle of receive beams. Assuming the backscattering coefficient to be uniform within the footprint of a narrow receive beam and ignoring the transmission loss, which can be added later in the model, the instantaneous backscatter intensity can be approximated in the time domain by the following function:

$$I(t_r) = \pi D^2 \sigma_\theta \sigma_\varphi S(\theta)\left\{\mathrm{erf}(r')\big|_{r_1(\theta_S)} - \mathrm{erf}(r')\big|_{r_2(\theta_S)} + \mathrm{erf}(r')\big|_{r_1(-\theta_S)} - \mathrm{erf}(r')\big|_{r_2(-\theta_S)}\right\} \quad , \quad (1)$$

where the transmit and receive beam patterns are approximated by a Gaussian shape with half-widths σ_θ and σ_φ respectively, D is the sea depth, θ_S is the beam steering angle, and $S(\theta_S)$ is the backscattering coefficient. The argument

$$r' = \frac{(r - D\tan\theta_S)}{2^{1/2} D\sigma_\theta \left(1 + \tan^2\theta_S\right)}$$

of the error functions in Eq.(1) is calculated at

$$r = r_1 = \left[(Ct_r/2)^2 - D^2\right]^{1/2} \quad \text{and}$$

$$r = r_2 = \text{Re}\left[(Ct_r/2 - CT/2)^2 - D^2\right]^{1/2},.$$

where C is the sound speed and t_r is the elapsed time from the moment of transmission. Eq.(1) provides a reasonably accurate approximation for the distortion of backscatter envelope due to the effect of insonification area and beam pattern at steering angles $\theta_S > \Delta\varphi$.

The Reson 8125 MBS system operates at 455 kHz. Its transmit beam is 1 degree wide along ship's track. The sonar forms 240 receive beams with the across-track width varying from 0.5 degree for the innermost (vertical) beams to 1 degree for the outermost beams. Figure 1 shows the error of backscatter strength estimates obtained from the backscatter energy and peak intensity modelled for each beam of the Reson MBS at sea depth of 35 m and pulse length of 50µs, 100µs and 200µs. The peak intensity is corrected for the insonification area and the energy is corrected for the footprint area, assuming an ideal rectangular model for the beam patterns. At oblique angles, where the width of the receive beam footprint is much larger than the across-track width of the insonification area, the error of the backscatter strength estimate tends to 0.28 dB expected from the approximation of the transmit beam pattern by a rectangular shape. At small angles, where footprint is much smaller than the insonification area, the error tends to 0.56 dB, which is a result of the rectangular shape approximation applied to both transmit and receive beam patterns. At moderate angles, the receive beam pattern changes significantly within the insonification area, which leads to a noticeable underestimation of the backscatter strength obtained from the peak intensity.
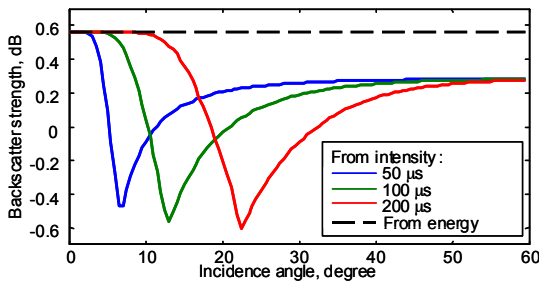

Fig.1 Errors of backscatter strength estimates from the energy (dashed line) and peak intensity (solid lines corresponding to 3 different pulse length) due to effects of insonification area and beam pattern.

The estimates of backscatter strength from the backscatter energy do not suffer from such angle dependent errors. An error of about 0.56 dB, resulting from the rectangular shape approximation for the Gaussian beam pattern of both transmit and receive beams, is nearly independent of incidence angle and pulse length. It is necessary to note, that a finite sampling interval (~35 µs in Reson 8125) will cause additional estimation errors (e.g. underestimation of energy), which can be significant at small incidence angles and short transmitted pulses.

3 Statistics of peak backscatter intensity

Figure 2 shows the angular dependence of the seafloor backscatter strength derived from the average peak intensity and energy of backscatter signals received from relatively homogeneous sand and rhodolith seabeds over ship's track sections of about 300 m long. Sea depth was about 35 m, and the transmitted pulse length was 101 µs. Backscattering from rhodolith, which is hard and unattached red coralline algae densely covering the seabed, is much stronger than that from sand and weakly dependent on incidence angle. The backscatter strength estimates from the peak intensity and energy are different. The difference is small at near vertical incidence. As the angle increases, the estimates diverge up to an angle of about 12^0, where the peak intensity estimates are about 1 dB lower than those obtained from the backscatter energy. This occurs at the incidence angle, where the footprint width becomes larger than the width of the insonification area, which is consistent with the prediction for the beam pattern effect shown in Fig.1.
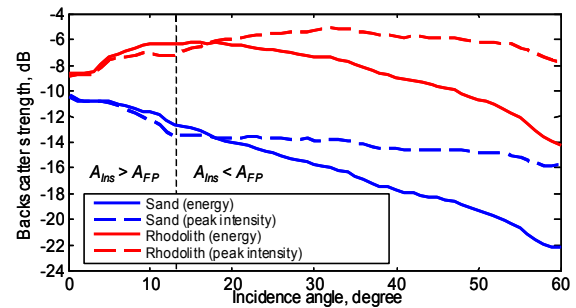

Fig.2 Angular dependence of backscatter strength from sand (blue) and rhodolith (red) derived from backscatter energy (solid) and peak intensity (dashed).

However, as the incidence angle further increases, the estimates become diverging in an opposite way - the backscatter strength derived from the peak intensity rapidly increases relative to the estimate from the backscatter energy. This effect can be explained based on the theory of extreme value statistics. At oblique incidence angles, the beam footprint contains a number of non-overlapping insonification areas, and this number increases with angle. If the correlation length of the seafloor roughness is smaller than the spatial separation between the adjacent insonification areas, then backscatter from these areas is expected to be statistically independent. Uncorrelated backscatter signals from different insonification areas can be referred to as elementary backscatter returns. Intensity variations of each elementary return with time (or with the ping number) can be considered as a stochastic process with a certain distribution function. Let the number of elementary returns from the footprint of an individual beam be M, so that the full backscatter signal comprises a series of M stochastic processes. If these processes are statistically independent and has a Gaussian distribution, then the backscatter intensity I_m has an exponential distribution. Let I_M be a process constituted from the peak values of M statistically independent and identically exponentially distributed processes with a unit variance. For the probability of I_M, one can write the following equation [3]:

$$P(I_M \le I + \ln M) = \left[1 - \exp(-I)/M\right]^M. \qquad (4)$$

The right-hand side of Eq.4 rapidly trends to $\exp[-\exp(I)]$ with $M \to \infty$, which is the Gumbel distribution widely used in the extreme value statistics. If the original processes have the mean value α, then a generalized Gumbel distribution, commonly referred to as the Fisher-Tippet (*F-T*) distribution, can be obtained from Eq.4:

$$F_{F-T} \equiv P(I_M \le I)\big|_{M\to\infty} = \exp\left[-\exp\left(\frac{I-\beta}{\alpha}\right)\right],$$

where $\beta = \alpha \ln M$ is the location parameter and α is called the scale parameter. The probability density function (PDF) of this distribution is

$$P(x) = \frac{1}{\alpha}\exp\left[\frac{x-\beta}{\alpha} - \exp\left(\frac{x-\beta}{\alpha}\right)\right],$$

the mean value $\mu_M = \beta + \zeta\alpha$, and the variance $\sigma_M = 6^{-1/2}\pi\alpha$, where $\zeta \approx 0.577$ is the Euler-Mascheroni constant. The *F-T* distribution is also expected for peak values of other distributions with an exponential fall-off rate in the PDF tail, such as the Rayleigh, Gamma (Γ), and *K* distributions, however, there is no closed-form expression for the relationship between the *F-T* shape parameter and the number of samples *M* for these distributions.
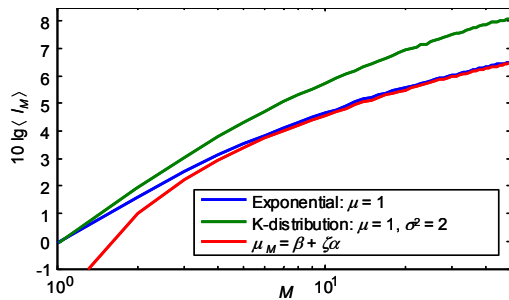

Fig.3 Mean value of the peak value distribution expected for series of exponentially (blue) and *K*- (green) distributed processes. The red line shows the prediction for exponential processes from the *F-T* model distribution.

The variation of the mean value of I_M (expressed in dB) with the number *M* shown in Fig.3 was numerically modelled for the exponential distribution of instantaneous intensity with the mean value $\mu = 1$ and for the *K*-distribution with $\mu = 1$ and the variance $\sigma = 2$. The red line shows the mean value of the *F-T* distribution with the location parameter $\beta = \alpha \ln M$. The *F-T* mean value rapidly tends to the actual mean at $M > 2$.

To compare the experimental results with the expectation from the extreme value theory, we calculated the difference between the backscatter strength derived from the signal energy and peak intensity. The distributions of this difference at different incidence angles should have the same scale parameter $\alpha \approx 1$, providing that the backscatter strength estimated from energy has the same angular dependence as that derived from the mean intensity. If the backscatter process is nearly Gaussian, the mean value of this difference should depend primarily on the number of individual scattering cells *M* within the beam footprint.

Figure 4 shows the difference of the backscatter strength estimated from the peak intensity and energy, as a function of the number of non-overlapping insonification areas within the beam footprint. The difference is remarkably

similar for backscatter from sand and rhodolith. The dashed line demonstrates the prediction for a Gaussian backscatter process, assuming $M = A_{FP}/A_{Ins}$. At small *M*, the experimental values are about 1 dB lower than the prediction. At larger *M*, the difference between measured and modelled values stays nearly the same up to $M \approx 12$ at about 50^0 and then rapidly reduces. There are a number of factors that may influence this dependence. First of all, the backscatter process may not be Gaussian and, hence the distribution of backscatter intensity is not exponential. Moreover, the *M* samples of I_m observed within the footprint are distorted in amplitude by the beam pattern, which disrupts the *F-T* distribution and affects the mean value. However, it is evident that the backscatter strength derived from the mean value of peak intensity increases with the number of scattering cells contained in the beam footprint. The mean of peak intensity equals the mean intensity only when the footprint contains one scattering cell.
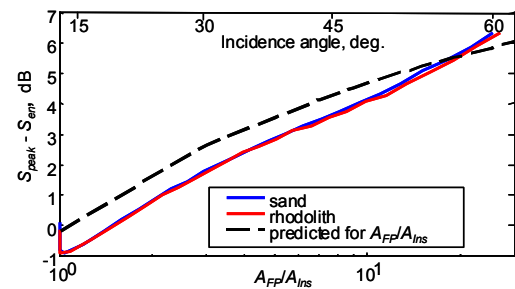

Fig.4 Difference between the backscatter strength estimates from energy and peak intensity as a function of the ratio of footprint A_{FP} and insonification A_{ins} areas. The dashed line shows prediction from the extreme value statistics model.
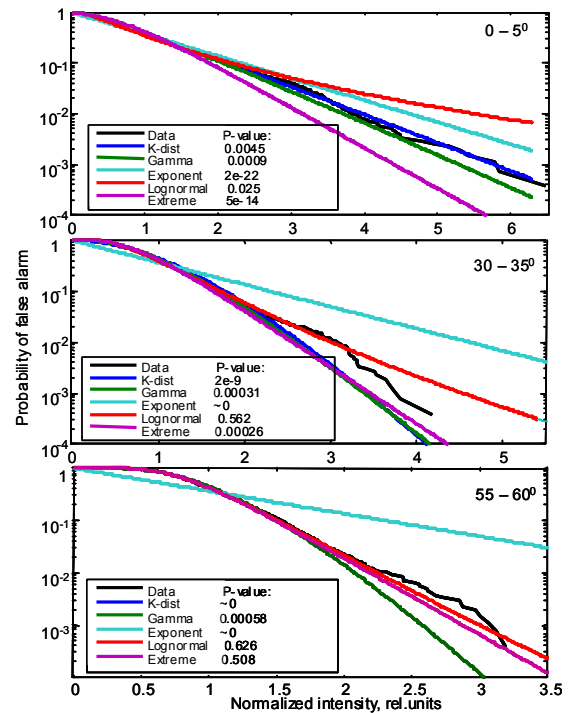

Fig.5 Probability of false alarm of the peak intensity distribution of backscatter measured from rhodolith at different incidence angles: 0-5^0 (top), 30-35^0 (middle) and 55-60^0 (bottom). Numbers in the legend indicate *p*-values of the Kolmogorov-Smirnov goodness-of-fit test.

To examine the peak intensity distribution, a number of model distributions were compared with the experimental data, using maximum-likelihood estimates of distribution parameters at different incidence angles. The *K*-distribution parameters were estimated by the method of moments. Figure 5 shows the probability of false alarm (1 - CDF) of the peak backscatter intensity measured from rhodolith in three different angular domains: near-vertical, moderate and large incidence angles, and different model fits for the experimental distribution. The *p*-values of the Kolmogorov-Smirnov goodness-of-fit test are given in the legends. At vertical angles of incidence, the lognormal distribution has the maximum *P*-value, because it fits best the PDF shape. However, at small incidence angles, the lognormal distribution does not satisfactorily model the tail of experimental distributions. The *K* and Γ distributions approximate both the shape of experimental histograms and the distribution tail with an acceptable accuracy better than that of the exponential model. At moderate angles, where the footprint contains a few insonification areas, the lognormal distribution provides the best fit for the PDF shape and the tail, except for the extreme values. The *F-T* distribution is a reasonable approximation for the PDF shape and for the extreme values. The *K*-distribution is not a satisfactory model for this angular domain. As the ratio A_{FP}/A_{Ins} becomes large at more oblique incidence angles, the *F-T* distribution model becomes as accurate as the lognormal model for both the histogram shape and the tail, which is also indicated by large *P*-values.

Although there are several explanations suggested for modelling backscatter statistics with an approximately lognormal distribution, especially for volume backscatter fluctuations (for example, see [4]), the *F-T* distribution appears to be a more reasonable model for the distribution of peak intensity fluctuations at large numbers *M* with respect to the underlying physical phenomenon.

4   Statistics of backscatter energy

Middleton [5] demonstrated that fluctuations of the average backscatter intensity tend to be Γ-distributed, if the scattering process is Gaussian. This follows from the known fact that a sum of *M* statistically independent and identically exponentially distributed processes I_m has a Γ-distribution, i.e. if the PDF of I_m is $f(I) = \exp(-I/I_0)/I_0$, where $I_0 = \langle I \rangle$, then

$$f\left(I_\Sigma \equiv \frac{1}{M}\sum_{m=1}^{M} I_m\right) = (I_\Sigma/\lambda)^{\beta-1}\frac{\exp(-I_\Sigma/\lambda)}{\lambda\,\Gamma(\beta)}, \quad (5)$$

where $\beta = M$ is the shape parameter and $\lambda = I_0/M$ is the scale parameter. The mean value of the Γ-distribution is $\mu \equiv \langle I_\Sigma \rangle = \beta\lambda = I_0$, i.e. it is equal to the mean value of I_m. Backscatter signals received by a sonar consist of a series of backscatter returns from different scattering sections of the seafloor. As in the previous section, non-overlapping insonification areas can be considered as statistically independent scattering cells, if the correlation length of the seafloor roughness is smaller than the width of the insonification area. If the beam footprint contains a number *M* of such scattering cells, then the backscatter signal energy is approximately a sum of statistically independent

backscatter processes I_m. Fluctuations of this sum normalized by the footprint size $A_{FP} \approx MA_{Ins}$ are expected to be Γ-distributed with the mean value $\mu \approx I_0/A_{Ins}$. This means that the backscattering coefficient estimated from the mean backscatter energy is approximately equal to the backscattering coefficient determined from the mean intensity. At the same time, the variance $\sigma^2 = \beta\lambda^2 \approx \mu^2/M$ of the Γ-distribution decreases with *M*, which means that the variations of the backscattering coefficient derived from the backscatter energy decreases with incidence angle when *M* > 1. This is an expected result of averaging.

If the intensities I_m are *K*-distributed, then the average intensity $\langle I_m \rangle$ is also expected to be Γ-distributed. However, the shape parameter of this distribution will not equal *M*, because the standard deviation and mean of the *K*-distribution are different.
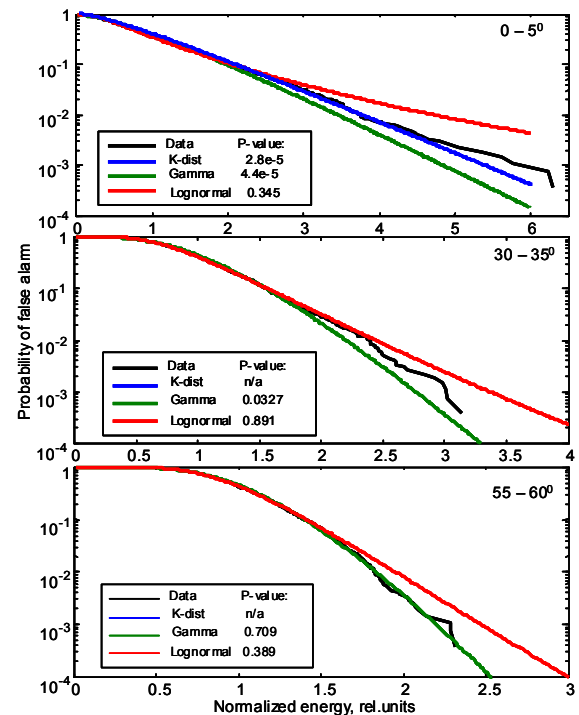

Fig.6 Probability of false alarm of backscatter energy fluctuations measured from rhodolith at different incidence angles: 0-5^0 (top), 30-35^0 (middle) and 55-60^0 (bottom). Colour lines show the best fit by different distribution models. Numbers in the legend indicate *P*-values of the Kolmogorov-Smirnov goodness-of-fit test.

Figure 6 shows the probability of false alarm of backscatter energy fluctuations measured within different angular domains, and the best fits of the different distribution models: *K*-distribution, Γ-distribution and lognormal distribution. Other distribution models did not demonstrate reasonable fit and therefore are not shown. At small incidence angles, the lognormal distribution provides the best approximation for the experimental histograms. However, it fails when predicting the experimental distribution tails. The *K*-distribution fits relatively well the distribution tail. The fit by the Γ-distribution is not as good in this angular domain. Such a result could be expected for vertical incidence, when the beam footprint contains a single insonification area. In this case, the distribution of backscatter energy is similar to that of the peak intensity. At moderate and large angles of incidence, the relationship

between the mean and variance of the experimental distribution is such that the *K*-distribution cannot be found through the method of moments. The lognormal distribution provides the best fit at moderate angles, although the prediction accuracy for extreme values by the Γ-distribution is similar. At oblique angles, when the number *M* is large, the Γ-distribution model is the best approximation for the distribution of backscatter energy fluctuations.
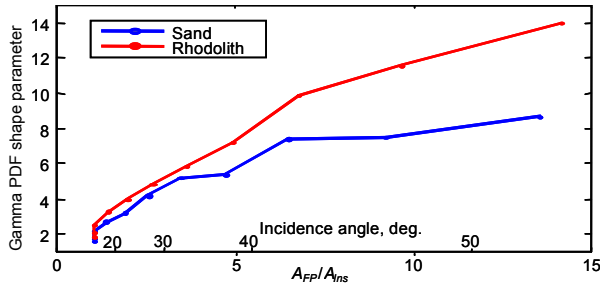

Fig.7 The relationship between the shape parameter β of the Γ-distribution model fit to the backscatter energy fluctuations and the ratio A_{FP}/A_{Ins} of beam footprint and insonification areas measured for backscatter from rhodolith (red) and sand (blue) beds.

The shape parameter β of the Γ-distribution best-fitted to the experimental data was compared with the ratio $M = A_{FP}/A_{Ins}$. If backscatter from adjacent non-overlapping insonification areas is statistically independent, then the shape parameter β is expected to be a linear function of *M*. Figure 7 demonstrates that the relationship between β and *M* is nearly linear for angles of incidence below approximately 40^0, but the shape parameter is larger than the ratio *M* by an increment varying from 1 to 2. At larger *M*, the Γ-shape parameter estimated for backscatter from rhodolith tends to the ratio *M*, while the estimates of β made for sand are noticeably lower. There are several possible factors that could cause such behaviour of this relationship. First of all, the backscatter intensity, as a function of time, is not a series of statistically independent samples. As discussed in [5], the correlation between samples depends on the bandwidth of sonar signals and the Γ-distribution may not fit well the average intensity, when the number of samples is small. Another important factor is that the distribution of instantaneous backscatter intensity is not exponential, which is clearly seen in the backscatter energy and peak intensity distributions observed at small incidence angles. Finally, the insonification area varies with incidence angle, so that its width at oblique angles may become smaller than the correlation length of the seafloor roughness. Consequently, the number of statistically independent scattering cells within the beam footprint will be smaller than that predicted from the ratio *M*.

5   Conclusion

The measurement scheme and backscatter data collection procedures implemented in MBS systems affect substantially the measurement results so that a simple direct comparison with the seafloor backscatter characteristics expected from theoretical models is no longer possible. The backscatter instantaneous intensity observed in individual

beams is distorted by the joint effect of a limited insonification area and directivity pattern of receive beams. This effect varies with the beam steering angle, which should be taken into account when estimating the seafloor backscatter strength and its angular dependence from MBS data.

The peak backscatter amplitude or intensity data, provided by some MBS systems as one value per each beam, are demonstrated to be extreme value distributed, when the beam footprint is larger than the insonification area. This results in significantly higher estimates of the seafloor backscatter strength derived from the mean value of the peak intensity than those expected for the instantaneous intensity. The overestimation effect depends on the ratio of footprint and insonification areas, which varies with incidence angle. As a result, the angular dependence of seafloor backscatter strength obtained from the peak intensity is considerably distorted at oblique angles.

Estimates of the seafloor backscatter strength and its angular dependence derived from the backscatter energy measured in individual beams are more robust with respect to distortion due to effects of insonification area and beam pattern. It is shown that the backscatter strength estimated from the backscatter energy is similar to the values derived from the instantaneous intensity. However, in contrast to the intensity, the variance of backscatter energy decreases with incidence angle. Fluctuations of the backscatter energy can be satisfactorily modelled by the Γ-distribution at oblique incidence angles, where the ratio *M* of the footprint and insonification areas is much larger than unity. Change in the variance of backscatter energy with incidence angle is primarily due to the variation of the Γ-shape parameter with the ratio *M*. This makes the seafloor backscatter images, derived from backscatter energy and corrected for the angular dependence, considerably less noisy at the edges of swath tracks than at their middle.

Acknowledgments

References

[1] H. Medwin, C.S. Clay, *Fundamentals of Acoustical Oceanography* (Academic Press, New York, 1998)

[2] D.R. Jackson, M.D. Richardson, *High-Frequency Seafloor Acoustics* (Springer Science, New York, 2007)

[3] P. Embrechts, C. Kluppelberg C., T. Mikosch, *Modelling Extremal Events in Insurance and Finance* (Springer-Verlag, Berlin, 1997)

[4] T. Gallaudet, C. de Moustier, "High-frequency volume and boundary acoustic backscatter fluctuations in shallow water", *J. Acoust. Soc. Am.* 114(2), 707-725 (2003)

[5] D. Middleton, "New physical–statistical methods and models for clutter and reverberation: the KA-distribution and related probability structures", *IEEE J. Ocean Eng.* 24(3), 261-284 (1999)