The Alpha stable distribution in ocean ambient noise modelling

Guoli Song, Xinyi Guo*, and Li Ma

Key Laboratory of Underwater Acoustic Environment, Institute of Acoustics, CAS, 100190 Beijing, China

Abstract. In view of the non-Gaussian of ocean ambient noise, the α stable distribution is applied to the statistical modelling. Firstly, the one-to-one correspondence between the four parameters of stable distribution and the sample mean, variance, skewness and kurtosis are established according to physical meaning. Then, numerical simulations are conducted to analyze the suitability of stable distribution for non-Gaussian ambient noise. In the case of white noise interference, noise is divided into Gaussian state, leptokurtic, and platykurtic separately. The parameters of stable distribution are estimated by the sample quantile and characteristic function method jointly. The simulation results show that, in the Gaussian state, α stable distribution is equivalent to normal distribution. As for leptokurtic distribution, stable distribution is much better than normal distribution, indicating absolute predominance in impulse-like data modeling. But it is not adaptive for low kurtosis state because its characteristic exponent can't be bigger than two. Finally, the result is verified by ambient noise collected in three environmental conditions, such as quiet ambient noise, airgun interference noise and ship noise. In all three cases, α stable distribution shows good adaptability and accuracy, especially for the airgun dataset it is far superior to normal distribution.

1 Introduction

In underwater acoustic signal processing, ocean ambient noise is often assumed to be Gaussian distribution. However, in practice it often shows non-Gaussian statistical characteristics due to natural or man-made disturbance[1-5].

In the literature, various noise models have been introduced to model noise with non-Gaussian characteristics[6-16]. Back in 1988, M. Bouvet[8] used Gaussian-Gaussian mixture (GGM) and Middleton Class-A models to describe underwater noise data. The two models showed better results than Gaussian model, but GGM was suitable for noise which is close to Gaussian and Middleton Class-A model was too strict to use. R.J. Webster[9] applied Champernowne model to describe underwater noise and gave the kurtosis curve of the model but without any simulation or actual data validation. In 2014, X. Guo[10-11] analyzed the data of a sea trial, and found that the noise spectrum level in the 100-1000Hz frequency band is close to Chi-square distribution with degree of freedom of 5~8. What's more, the GARCH model and two-mode model were used to simulate the actual underwater noise. The GARCH model has good adaptability without obvious interference and the two-mode model only suitable for some situations in shallow water. In recent years, the α stable distribution has attracted more and more attention. It has been used in radar signal processing[12-13], speech noise modeling[14] and also emerged in the field of underwater acoustic modeling[15-16], which is

limited to the interference conditions or reverberation fields of shrimp.

In this paper, we firstly establish the relationship between parameters of α stable distribution and numerical characteristics of variable. Then the suitability of stable distribution for ambient noise modelling is analyzed by three numerical simulations. Finally, its effectiveness verified by experimental data acquired in three different environments.

2 α stable distribution and sample numerical characteristics

α stable distribution[17-18], also known as non-Gaussian stable distribution, heavy-tailed distribution, was proposed by P. Lévy in 1925 when he studied the generalized central limit theorem.

Since the probability density function (PDF) of the stable distribution does not have a uniform and closed analytical expression except for a few special exceptions, the characteristic function is generally used to describe its distribution characteristics as follows,

$$\varphi(t) = \exp\{i\delta t - |\gamma t|^\alpha [1 + i\beta \operatorname{sgn}(t)\omega(t, \alpha)]\} \quad (1)$$

$$\omega(t, \alpha) = \begin{cases} \tan(\pi\alpha/2), \alpha \neq 1 \\ \frac{2}{\pi} \lg|t|, \alpha = 1 \end{cases} \quad (2)$$

where $0 < \alpha \leq 2$, $-1 \leq \beta \leq 1$, $\gamma > 0$, $-\infty < \delta < +\infty$,

* Corresponding author: guoxinyi@mail.ioa.ac.cn

$$\text{sgn}(t) = \begin{cases} 1, t > 0 \\ 0, t = 0 \\ -1, t < 0 \end{cases} \text{ is a symbolic function. So the stable}$$

distribution is uniquely determined by four parameters, they are location parameter δ , scale parameter γ , skewness parameter β and characteristic exponent α .

From the perspective of physical meaning, location parameter δ indicates the location of a distribution, which is similar with the mean value of a sample, representing the average level of a random variable.

The scale parameter γ represents the dispersion degree of a distribution deviating from the mean value. The larger the scale parameter is, the more obvious the deviation between the statistical data and the mean value is, and this is the same as the physical meaning of the variance.

For skewness parameter β , it is a measure of the symmetry of distribution. If a distribution has a positive skewness parameter, indicating that the tail on the right side of the distribution is longer and thicker than the left. It happens that there is a similar numerical characteristic named skewness which also describes the symmetry degree of the random variables.

As for characteristic exponent α , it determines the tail thickness of a PDF. Comparing the characteristic exponent with the kurtosis, it can be found that the stable distribution degenerated to Gaussian distribution while characteristic exponent is 2.0. In this condition, the kurtosis value is constant 3.0. In terms of the physical meaning, the smaller the characteristic exponent, the thicker the tail and the stronger the impact. On the contrary, the greater the kurtosis, the thicker the tail and the stronger the impact. In a word, kurtosis and characteristic exponent α describe the opposite physical meaning. In addition, it should be noted that the characteristic exponent ranges from 0 to 2.0, which can only describe random variable that is leptokurtic. That is to say, only when the kurtosis value is greater than 3.0, the kurtosis is related to the characteristic exponent.

Through the above analysis, it can be found that there is a corresponding relationship between the numerical characteristics of random variables and the parameters of stable distribution. Mean corresponds to location parameter, variance corresponds to scale parameter, skewness corresponds to skewness parameter one by one. The relationship between kurtosis and characteristic exponent also exists when the kurtosis value is greater than 3.0. From this point of view, the stable distribution depicts the first four order statistical properties of random variables.

3 Numerical simulation analysis

This section analyzes the feasibility of stable distribution for non-Gaussian noise modelling through theoretical simulation. Under white noise interference, the simulation experiment is divided into three cases: Gaussian state, leptokurtic, and platykurtic. Here, the four parameters of α stable distribution is estimated by

the quantile method[19] proposed by McCulloch and characteristic function method[20] jointly.

Numerical simulation I: The simulation sample is Gaussian white noise, and the sampling number is $N=5000$. Fig.1 shows the simulated noise sequence and the statistical histogram of the amplitude. On the time domain waveform, the noise is relatively stable, and the statistical histogram of it has a kurtosis value of about 3.0. Fig.2 shows the PDF curves of α stable distribution and normal distribution, they are basically coincident, and the statistical characteristics of the noise can be well described. At the same time, the estimated value of characteristic exponent α is 2.0. This further indicates that the stable distribution degenerates into a Gaussian distribution when the characteristic exponent is 2.0.

Fig. 1. Gaussian noise waveform and statistical histogram.

Fig. 2. Stable distribution modeling for Gaussian noise.

Numerical simulation II: The simulation sample is Gaussian white noise interfered by white noise. The amplitude of the interference relative to Gaussian white noise is 2.0, the kurtosis value is 8.0. The point number of synthesized noise is $N=5000$, and it has a kurtosis value greater than 3.0.

As seen from the time domain waveform in Fig.3, small shocks appear during the whole time, the stationarity is significantly worse compared with Gaussian noise. The sample kurtosis value is about 6.6 with spike statistical properties, which intuitively shows that the statistical histogram is more "high and thin" than normal distribution.

Fig.4 shows the modeling results, at this time, the estimated characteristic exponent is 1.45, indicating that the statistical histogram of the noise is leptokurtic. And the estimated value of the scale parameter γ is about 160, it means large deviation of the data from mean value.

Therefore, normal distribution is not adaptive, but the α stable distribution can well fit the peak characteristics of the noise.

Fig. 3. Leptokurtic noise waveform and statistical histogram.

Fig. 4. Stable distribution modeling for leptokurtic noise.

Numerical simulation III: On the basis of Numerical simulation II, changing the kurtosis value of interference noise from 8.0 to 2.0 and other conditions remain unchanged, the kurtosis value of the synthesized noise will become smaller than 3.0, which can be called platykurtic.

Under this circumstance, the statistical histogram becomes "short and fat", shown in Fig.5, the frequency density near the zero is reduced. The statistical kurtosis value is 2.52, which is less than the kurtosis of normal distribution.

Fig.6 shows two PDF curves of α stable distribution and normal distribution. In such a case, the two PDF curves coincide, and are more "sharp" than the data. The estimated value of the characteristic exponent is 2.0, which is equal to the theoretical value of the normal distribution. That is to say, the two distributions can not express the noise now.

The reason why the stable distribution loses adaptability in the circumstance is because when the kurtosis value is equal to 3.0, the characteristic exponent is equal to 2.0 theoretically. Since the physical meaning of kurtosis and characteristic exponent is opposite, when the kurtosis value is less than 3.0, the theoretical characteristic exponent should be less than 2.0 (if any). But the definition of the stable distribution specifies the range of the characteristic exponent that is smaller than 2.0, so that α stable distribution loses its adaptability to low-peak noise samples.

Fig. 5. Platykurtic noise waveform and statistical histogram.

Fig. 6. Stable distribution modeling for platykurtic noise.

Through the analysis above, it can be concluded that for the Gaussian noise, the modeling results of α stable distribution and normal distribution are the same, because at this time, α stable distribution degenerates into a Gaussian distribution. For the sample with strong impulses in time series, Gaussian distribution fails due to large value of kurtosis, but α stable distribution behaves well according to the PDF curve. That is to say, if the statistical histogram of a random variable is leptokurtic and only has one peak, α stable distribution is a good choice for modeling. But for the sample has a small kurtosis (less than three), α stable distribution loses its adaptability as normal distribution, which is determined by the definition of it.

4 Experiments

4.1 The dataset

The ocean ambient noise data is acquired in the northern South China Seas (shown in Fig.7) by a hydrophone placed at the bottom at 87m, the sampling frequency of the equipment is 24 kHz and the sensitivity is -164 dB.

The experiment lasts for 5 days from May 29 to June 2 in 2016. During the experiment, there are some airgun signals of oil exploration ship sometimes. They appear intermittently every 9 seconds for about 1.5 seconds, mainly concentrated in the frequency band below 550 Hz. On may 31, a ship radiation noise measurement experiment was carried out. During the period, a cooperative fishing boat sailed at a speed of 5~9 knots

near the hydrophone, and the distance from the hydrophone is changed within the range of 0.1~2.0km.

Fig. 7. Sea trial position.

Fig. 8. Time-frequency spectrum of quiet data.

Fig. 9. Time-frequency spectrum of airgun data.

Fig. 10. Time-frequency spectrum of ship data.

The measured noise data is divided into three categories according to different scenarios. (1) *quiet data*-no vehicle near the hydrophone and no other interference sources in the near field; (2) *airgun data*-noise contains a large number of regular airgun signals due to exploration activities; (3) *ship data*-the hydrophone is mainly affected by a single dominant ship. Each type of dataset consists of 200 samples with a

sample length of 9 seconds (216000 points). Among them, the ship noise data set consists of 40 samples at 5 to 9 knots. What shown in Fig.8-Fig.10 are the time-frequency spectrum of the three.

4.2 Discussion

4.2.1 Numerical characteristics

For each dataset, the mean, variance, skewness and kurtosis of the sound pressure are statistically analyzed.

It can be seen from Table 1 that in terms of mean, variance and skewness, the three datasets tend to be zero, and there is no significant difference, indicating that the three type of noise obey a symmetric distribution with a mean of zero. But there is significant difference in kurtosis statistical characteristic. For the quiet data, the kurtosis value is close to 3.0, and has a Gaussian property. As for the airgun data, the kurtosis value is much larger than 3.0, which has obvious non-Gaussian properties. And the kurtosis value of ship data is larger than that of quiet data, which is weak-Gaussian.

Table 1. Numerical characteristics of ambient noise.

	Quiet	Airgun	Ship
Mean	0.00	0.00	0.00
Variance	0.00	0.01	0.01
Skewness	-0.02	0.00	0.01
Kurtosis	3.08	11.08	3.26

4.2.2 Characteristic exponent α estimation statistics

It is known from the above analysis that the main difference between the three types of noise is reflected in the kurtosis. Therefore, this part focuses on the peak degree of the samples, which can be expressed by the characteristic exponent of α stable distribution.

Same as the part 3, the characteristic exponent α is evaluated by sample quantile and characteristic function method jointly. The statistical law of α of 200 samples is given in Table 2.

Table 2. Characteristic exponent estimation statistics of ambient noise.

	Quiet data	Airgun data	Ship data
$\alpha \leq 1.2$	0	0	0
$1.2 \leq \alpha \leq 1.6$	0	173	0
$1.6 \leq \alpha \leq 1.8$	0	26	0
$1.8 \leq \alpha \leq 1.99$	34	1	85
$\alpha \geq 1.99$	166	0	115

It can be seen from the table that the characteristic exponent of quiet data is greater than 1.8, and most of it is close to 2.0, this is the value of a Gaussian distribution. For airgun data, the evaluated value of it is mainly

between 1.2 and 1.6, indicating that the sample is leptokurtic and with heavy tail. For ship dataset, the characteristic exponent is larger than 1.8, but compared with the quiet data, the value is relatively smaller, showing weak spike characteristics.

4.2.3 Stable distribution model verification of measured noise

For the three typical ambient noise above, the normal distribution and α stable distribution are used to model them respectively. By comparing the PDF and cumulative distribution function (CDF) of the two types of models, the pros and cons of the models are analyzed.

Fig.11 and Fig.12 show one of the modeling result of quiet data, the α value is 1.9988, which is very close to 2.0. The PDF and CDF curves of stable distribution almost coincide with the normal distribution, and both models can well describe the sample.

Fig. 11. PDF of normal distribution and stable distribution (quiet data).

Fig. 12 CDF of normal distribution and stable distribution (quiet data).

The modeling results of airgun data are shown in Fig.13 and Fig.14. The statistical histogram shows obvious spike characteristics, and a large number of samples are concentrated near the sample mean. Compared with the PDF of normal distribution, it is intuitively more "high and thin", so the normal distribution completely fails. In contrast, because of the characteristic exponent of stable distribution is adjustable, and the tail has an algebraic attenuation form, so it can perfectly fit the peak and thick tail characteristic of the airgun noise. And the CDF curve of α stable distribution also remains highly consistent with the true value.

The ship noise, which is known as weak Gaussian noise according to the analysis aforementioned is shown in Fig.15 and Fig.16. At this time, the normal distribution can basically fit the data, but the stable distribution is more accurate, especially at the peak position.

Fig. 13. PDF of normal distribution and stable distribution (airgun data).

Fig. 14. CDF of normal distribution and stable distribution (airgun data).

Fig. 15. PDF of normal distribution and stable distribution (ship data).

Fig. 16. CDF of normal distribution and stable distribution (ship data).

It can be seen from the above that α stable distribution has good effectiveness in the modeling of quiet data, airgun data and ship data. It is more accurate than normal distribution in ship data modeling, and far superior to normal distribution in airgun data modelings.

5 Conclusions

Focused on the effectiveness of the α stable distribution for ocean ambient noise modeling, the mean, variance, skewness and kurtosis are cited to understand the stable distribution conveniently. In the case of white noise interference, three numerical simulations are made to analyze the modelling adaptability of it. It is outstanding in impulse-like data modelling but is not suitable for noise whose kurtosis is smaller than three. Also, the result is verified by sea trial noise data collected in three kinds of environments. For quiet data, airgun data and ship data, the α stable distribution shows good adaptability and accuracy, especially for the airgun environment it is far superior to the normal distribution.

References

1. P.L. Brocketat, M. Hinichb, G.R. Wilson, J. Acoust. Soc. Am **82**, 1386 (1987)
2. F. Traverso, G. Vernazza, A. Trucco, IEEE/MTS Oceans-Yeosu 1 (2012)
3. P.P. Parada, A. Cardenal-López, J. Acoust. Soc. Am. **135**, 3371 (2014)
4. P.T. Arveson, D.J. Vendittis, J. Acoust. Soc. Am. **107**, 118 (2000)
5. S.C. Wales, R.M. Heitmeyer, J. Acoust. Soc. Am. **111**, 1211 (2002)
6. Y. Zhou, P. Han, J. Yang, ACTA ACUSTICA **28**, 518 (2003)
7. M.A. Roch, M.S. Soldevilla, J.C. Burtenshaw, E. E. Henderson, J.A. Hildebrand, J. Acoust. Soc. Am. **121**, 1737 (2007)
8. M. Bouvet, S.C. Schwartz, J. Acoust. Soc. Am. **83**, 1023 (1988)
9. R. J. Webster, IEEE Trans. Signal Pro. **41**, 2249 (1993)
10. X. Guo, F. Li, G. Tie, Physics **43**, 723 (2014)
11. G. Tie, X. Guo, Technical Acoustics **33**, 209 (2014)
12. Z. Zheng, S. Wang, J. Electr. Inf. Technol **36**, 2963 (2014)
13. A. Zhang, T. Qiu, X. Zhang, J. Electr. Inf. Technol **27**, 1201 (2005)
14. P.G. Georgiou, P. Tsakalides, C. Kyriakakis. IEEE **1**, 291 (1999)
15. K. Pelekanakis, H. Liu, M. Chitre, Oceans-Spain **86**, 1 (2011)
16. Y. Liu, Z. Huang, C. Liu, J. Natl. Univ. Def. Technol **34**, 148 (2012)
17. G. Samorodnitsky, M.S. Taqqu, J. AM. STAT. ASSOC **90**, 805 (1995)
18. T.S. Qiu, H.Y. Wang, X.B Li, IEEE-EMBC **3**, 2635 (2003)
19. J.H. McCulloch, COMMUN. STATIST.-SIMULA **15**, 1109 (1986)
20. I.A. Koutrouvelis, J. AM. STAT. ASSOC **75**, 918 (1980)