Shahrood University of Technology

Journal of Mining and Environment (JME)

Journal homepage: www.jme.shahroodut.ac.ir

Iranian Society of Mining Engineering (IRSME)

Selection of Appropriate Probability Distributions for Rock Analysis using Laser-induced Breakdown Spectroscopy

Kamran Abbas[1], Adeel Nawazish[1], Navid Feroze[1], and Nasar Ahmed[2*]

1. Department of Statistics, King Abdullah Campus Chatter Kalas, The University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan
2. Department of Physics, King Abdullah Campus Chatter Kalas, The University of Azad Jammu & Kashmir, Muzaffarabad, Pakistan

Article Info

Received 7 October 2022
Received in Revised form 25 November 2022
Accepted 12 December 2022
Published online 12 December 2022

DOI:10.22044/jme.2022.12327.2237

Keywords

LIBS
Rock Analysis
Statistical Distributions
Akaike Information Criterion
Amplification factor

Abstract

In this work, an attempt is made to fit and identify the most appropriate probability distribution(s) for the analysis of seventeen rock samples including diorite, gypsum, marble, basalt, sandstone, limestone, apatite, slate, dolomite, granite-II, schist, gneiss, amphibolite, hematitle, magnetite, Shale, and granite-I using laser-induced breakdown spectroscopy. The graphical assessment and visualization endorse that the rock dataset series are positively skewed. Therefore, Frechet, Weibull, log-logistic, log-normal, and generalized extreme value distributions are considered as candidate distributions, and the parameters of these distributions are estimated by maximum likelihood and Bayesian estimation methods. The goodness of fit test and model selection criteria such as the Kolmogorov-Smirnov test, Akaike Information Criterion, and Bayesian Information Criterion are used to quantify the accuracy of the predicted data using theoretical probability distributions. The results show that the Frechet, Weibull, and log-logistic distributions are the best-fitted probability distribution for rock dataset. Cluster analysis is also used to classify the selected rocks that share common characteristics, and it is observed that diorite and gypsum are placed in one cluster. However, slate, dolomite, marble, basalt, sandstone, schist, granite-II, and gneiss rocks belong to different clusters. Similarly, limestone and apatite appeare in one cluster. Likewise, shale, granite-I, magnetite, amphibolite, and hematitle appeare in a different cluster. The current work demonstrate that coupling of laser-induced breakdown spectroscopy with suitable statistical tools can identify and classify the rocks very efficiently.

1. Introduction

Rocks and minerals are of great importance in the universe. They have a wide range of applications that make them important for human. Usually, rocks are used in construction, for manufacturing substances, making medicines, and for the extraction of precious elements. Rocks provide clues about the earth's history, and therefore, they are extremely interesting for the scientists and researchers. Rocks tell us about the history of the earth's surface because they are the primary storyteller of the past climate, life, and major events at the earth's surface. As the rock gradually breaks down, release minerals that end up in the water of oceans, lakes, and the soil.

A rock is a naturally occurring solid cohesive aggregate of one or more mineral or mineral materials. An unlimited variety of rocks are present in Pakistan that employ effects on the properties of soil. The most common rock types are metamorphic that are found in Himalayan regions. It includes gneisses, schist, slates, and phyllites with some quartzite and marble. Small outcrops of phyllites and quartzites are also found in the northern part of the Indus plain. Granite, diorite, dolerite, and peridotite are the more common types of igneous rocks that occur in Dir, Swat, Chitral, Gilgit, Zhob, and Chagai. The gemstones, marbles, and much other economic mineralization are found in Azad Jammu and Kashmir (AJ&K) and Gilgit-Baltistan. The AJ&K region has also a share of gemstones and granite, especially from the upper areas of Neelam valley, whereas marble,

✉ *Corresponding author: nasar.ahmed@ajku.edu.pk (N. Ahmed)*

construction materials, coal, clays, and other minerals are found in the different areas of AJ&K. In general, rocks can be classified into three main types on the process of their embodiment. These are igneous rocks, sedimentary rocks, and metamorphic rocks [1-3].

However, it is a desired to characterize the rocks samples from the statistical viewpoint. Therefore, it is mandatory to identify the best suitable distribution(s) for the rock samples because the choice of probability distributions is of essential significance. The choice of suitable probability distributions for a given sample cannot be made on a physical basis, and therefore, statistical inference and practical appropriateness play a much superior role in the distributional choice than physical reasoning [4].

In this regard, a little contribution is found in the literature. For instance, Azizi et al. applied different probability distribution functions including normal, lognormal, beta, and gamma, along with the Kolmogorov Smirnov test [5]. It was concluded that normal distribution was most suitable for the rocks samples. Ghazdali et al. conducted a statistical analysis about the rock mass of the mine in Morocco [6]. Malkowski et al. checked the variability of rock properties in the roadways' roofs, and also analyzed the effect of geomechanical data on numerical modeling of the stability [7]. Teymen and Manguc applied different statistical techniques for the prediction of the uniaxial compressive strength of rocks [8]. Gent et al. examined the stability of rocks slopes to examine the damage in the design of rock armored slopes [9]. Cai et al. investigated the water saturation effects on the mechanical behavior of different rocks [10]. Salih and Alshkane determined the relationship between the physical and mechanical properties of igneous rocks [11]. Mayer et al. used the application of statistical approaches to analyze the geological, geotechnical, and hydrogeological data at fractured rock mine sites in northern Canada [12]. Karakul and Ulusay carried out a study to correlate the strength properties of rocks with a p-wave velocity of many rocks under different degrees of saturation [13]. Ceryan et al. applied generalized regression neural networks to establish predictive models for the unconfined compressive strength of carbonate rocks in Turkey [14]. Ghazvinian and Hadei explored the effects of discontinuity orientation and confinement on the strengths of rocks [15]. Huang et al. investigated the dependence of tensile strength softening of the sandstone on loading rate [16].

Similarly, Huang et al. conducted a study on sedimentary rocks' dynamic characteristics under creep state using a new type of testing equipment [17]. Further, Liu et al, studied mechanical parameters with a statistical methods [18]. G. Mibei introduced and classified the different rock samples [19]. Rybar et al. studied the physical-mechanical properties of rocks [20]. Singh et al. detected a correlation between point load index and uniaxial compressive strength for different rock types [21]. Wang et al. predicted uniaxial compressive strength of rocks from simple index tests using a random forest predictive model [22]. Recently, Probability Distribution Functions have also been selected for Rock Joint Geometric Properties by Jamal et al. [23].

Here, we present a new work to distinguish the rock samples by comparing several distributions and determining the best probability distribution for the selected rocks available in Pakistan based on AIC and BIC. The optical emission data of these rock samples was taken using laser-induced breakdown spectroscopic (LIBS) setup. This emission data was utilized to get the best probability distributions for accurate investigations. The present work will be interesting and beneficial for a wide range of audiences working in the field of spectroscopy, geology, and statistics. Including this introduction section, the remaining paper unfolds as what follows. Section 2 introduces the methodology. Results and discussions are presented in Section 3, and finally, conclusion is given in Section 4.

1.1. Data description

The seventeen rock samples were collected from different locations of Pakistan. Seventeen rock samples were considered including Diorite (D), Gypsum (G), Marble (M), Basalt (B), Sandstone (S), Limestone (L), Apatite (A), Slate (SL), Dolomite (DO), Granite-II (GR-II), Schist (SC), Gneiss (GN), Amphibolite (AM), Hematitle (H), Magnetite (MA), Shale (SH), and Granite-I (GR-I), as shown in Figure 1.

Figure 1. Rock samples collected from Pakistan.

2. Methodology

The experimental setup used to get data of optical emission spectra is the same as discussed in our earlier papers [24-29], as shown in Figure 2. In brief, it consists of second harmonic Q-Switched Nd: YAG laser having 532 nm wavelength and a focusing lens of 20 cm was used to focus the laser beam on the target. Sample was placed on a motorized sample holder to provide fresh surface to each shot and to prevent deep craters on the surface of the sample. An optical fiber (high-OH, core diameter: 600 µm) with a collimating lens (0-45°) coupled to Avantes spectrometer that covered the wavelength range from 250 nm to 870 nm to record the emission spectra from the plasma plume. The laser delivers 850 mJ pulse energy at 1064 nm and 400 mJ at 532 nm. At the target surface, the spot diameter was calculated as 0.5 mm, and the corresponding power density would be about $2 \times 10^{10}\ Wcm^{-2}$ at 200 mJ laser energy. The laser pulse energy was varied with the delay of the flash lamp switch, and an energy-meter that was used to measure the laser pulse energy. When laser was fired on the target surface, the incident photons were absorbed by the sample, which leads to an excited state, and for a very short time, the plasma was produced, and emission spectra were recorded using the spectrometer. A physical diagram of LIBS setup is shown in Figure 3.

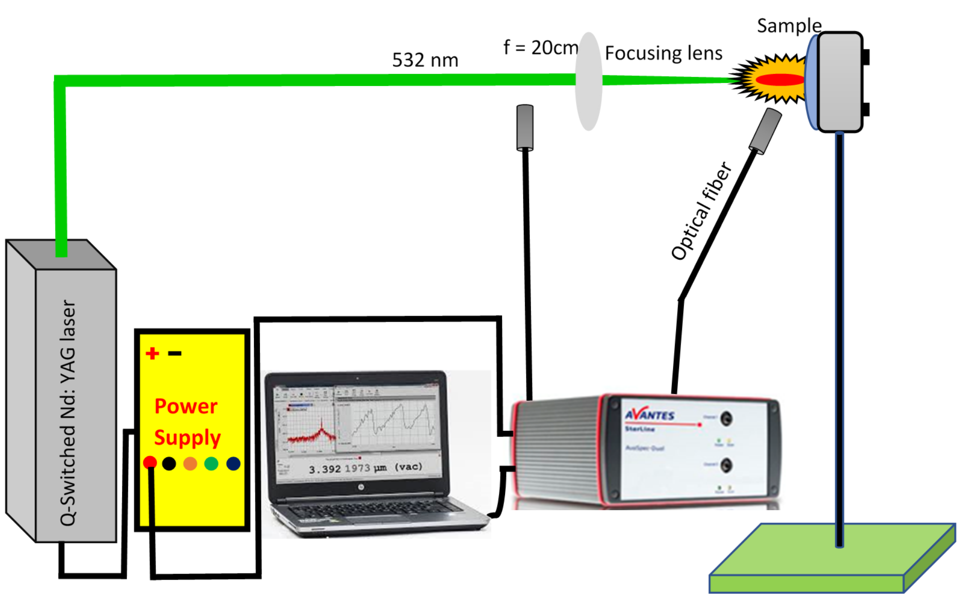

Figure 2. Schematic diagram of LIBS setup used for rock analysis.

After formation of plasma plume, it starts cooling after spreading, and then it emits spetra of light rays having different wavelengths that were collected using a spectrograph having a charge coupled device (CCD) that records all wavelengths simultaneously. After a careful identification of the spectral lines for all the rock samples, the major lines of those elements that were present in all the samples were selected as the input data.

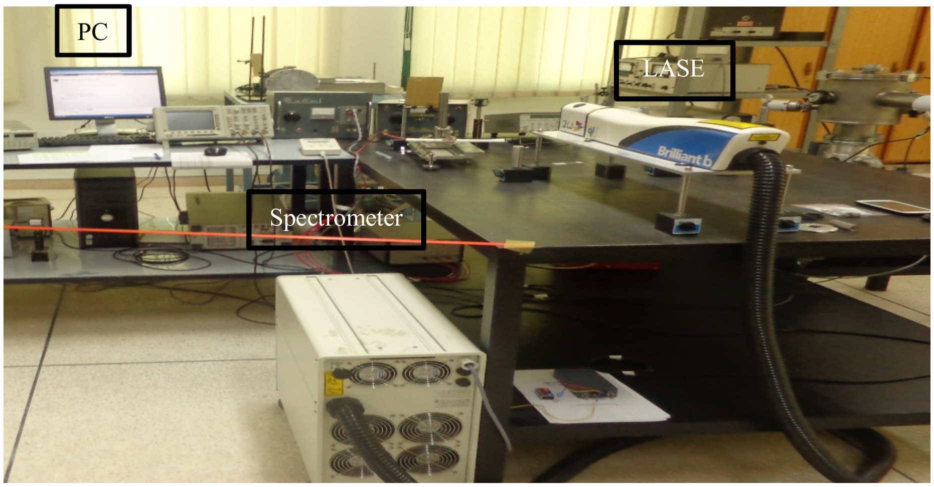

Figure 3. Physical diagram of LIBS setup used for rock analysis.

2.1. Emission studies and statistical distributions

The optical emission spectra were collected by focusing the laser beam on the rock samples. Figure 4 shows the optical emission spectra of all the rock samples collected using Aventes spectrometer at 200 mJ laser energy in the wavelength ranges from 250-870 nm.

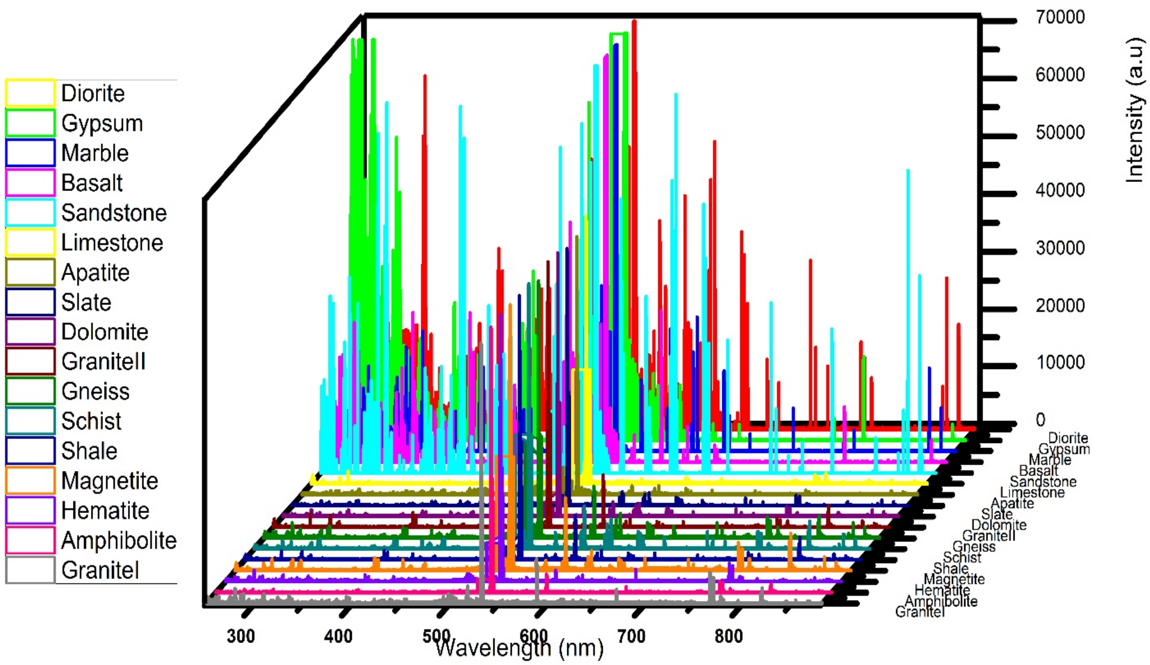

Figure 4. Optical emission spectra of rock samples in the range 250-870 nm.

The emission data of the major lines of those elements that were present in all the samples were selected for the different statistical distributions. Integrated line intensities of all the selected elements were selected as the input data, and many probability distribution functions (PDFs) have been proposed in the recent past but in the present study, Frechet distribution (FD), Weibull distribution (WD), log-logistic distribution (LLD), lognormal distribution (LND), and generalized extreme value distribution (GEVD) are used in the current study to describe the characteristics of the selected rocks. The PDF of these distributions is presented in Table 1. Maximum likelihood (ML) and Bayesian estimation (BE) methods are used to estimate the parameters of FD [30], WD [31], and LLD [32]. However, the parameters of LND and GEVD are estimated by only the ML method.

Table 1. PDF of five distributions and its parameters.

Distribution	PDF	Parameter
FD	$f(x;\ \alpha,\beta)=\left(\frac{\alpha}{\beta}\right)\left(\frac{\beta}{x}\right)^{\alpha+1}e^{-\left(\frac{\beta}{x}\right)^{\alpha}}\ ,\quad x>0,\quad \alpha,\beta>0$	$\alpha = Shape$ $\beta = Scale$
WD	$f(x;\ \alpha,\beta)=\left(\frac{\alpha}{\beta}\right)\left(\frac{x}{\beta}\right)^{\alpha-1}e^{-\left(\frac{x}{\beta}\right)^{\alpha}},\quad x>0,\quad \alpha,\beta>0,$	$\alpha = Shape$ $\beta = Scale$
LLD	$f(x;\ \alpha,\beta)=\frac{\left(\frac{\beta}{\alpha}\right)\left(\frac{x}{\alpha}\right)^{\beta-1}}{\left\{1+\left(\frac{x}{\alpha}\right)^{\beta}\right\}^{2}}\ ,x>0,\alpha,\beta>0$	$\alpha = Scale$ $\beta = Shape$
LND	$f(x;\ \mu,\sigma)=\frac{1}{x\sigma\sqrt{2\pi}}\exp\left\{\frac{-(lnx-\mu)^{2}}{2\sigma^{2}}\right\},x>0,-\infty<\mu<\infty,\sigma^{2}>0$	$\mu = Shape$ $\sigma^2 = Scale$
GEVD	$f(x;\ \theta,\alpha,\eta)=\frac{1}{\alpha}\left\{1+\eta\frac{(x-\theta)}{\alpha}\right\}^{-\frac{1}{\eta}-1}exp\left[-\left\{1+\eta\frac{(x-\theta)}{\alpha}\right\}^{-\frac{1}{\eta}}\right],x>0$	$\eta = Shape$ $\alpha = Scale$ $\theta = Location$

2.1 Model selection

The following goodness of fit tests are used for the selection of best-fitted distribution for the rock series:

2.1.1. Kolmogrove Smirnove test

KS (Kolmogrove, 1933) test was performed under the null hypothesis to check whether the rock samples originate from a hypothesized continuous distribution [33]. The KS test statistic (D) can be expressed as:

$$D=\max\left|F(x_i)-\frac{i-1}{n},\ \frac{i}{n}-F(x_i)\right|$$

where x_i represents the rocks samples, $i = 1,2,\dots,17$.

2.1.2. AIC and BIC

AIC and BIC are used to pick and endorse the most applicable distribution for describing the behavior of selected rocks based on the minimum AIC and BIC values. The AIC and BIC values can be calculated as:

$$AIC=2p-2ln(\hat{L})$$

$$BIC=Pln(n)-2ln(\hat{L})$$

where $\hat{L}$ is the maximum value of the likelihood function, and ‘p’ is the number of parameters estimated.

2.1. Kruskal Wallis test

The Kruskal and Wallis (1952) test does not make any assumptions about normality, and in the current study, it is used under the null hypothesis for testing whether the rock samples emanate from the same distribution at a 5% level of significance [34].

2.2. Cluster analysis

Cluster analysis is used to classify the rocks that share common characteristics and the groups are initially not known. The cluster of rocks was grouped based on the similarity level. The higher the similarity level, the more similar rocks are in each cluster. The lower the distance level, the closer the rocks are in each cluster. A dendrogram is constructed to visualize the clustering results at each step.

3. Results And Discussion

The descriptive statistics such as mean, median, coefficient of variation (CV), minimum (Min.), maximum (Max.), coefficient of skewness, and coefficient of kurtosis for rocks samples are provided in Table 2. The maximum standing varies from 69526 to 8910, and the minimum is between 1 and 210. On average, the AM rock has maximum standing of 14048.2 arbitrary units (arb.u.), where A rock has a minimum standing of 389.30 arb. u.

Table 2. Descriptive statistics for rock samples with sample size (n = 32).

Rocks	Mean	Median	CV	Min.	Max.	Skewness	Kurtosis
D	870.04	199.95	200.43	81.83	7669	3.28	12.61
G	1047.32	214.13	195.54	73.63	8910	3.13	11.87
M	570.78	408.76	92.80	92.27	2258	1.83	6.12
B	528.73	352.50	101.20	82.83	2134	1.75	5.26
S	465.49	126.31	164.40	41.57	30191	2.40	7.78
L	737.12	503.63	96.23	62.57	26872	1.49	4.51
A	389.30	226.90	96.99	97.50	1997	2.71	11.56
SL	1559.20	464.20	170.69	177.70	11602	3.00	11.22
DO	1464.90	940.00	133.53	210.00	9017	2.91	10.99
GR-II	1143.10	527.00	125.71	157.50	6082	2.41	8.43
GN	544.40	379.10	108.99	59.20	3230	3.09	14.32
SC	1406.30	727.90	126.13	117.50	7686	1.25	12.41
U	14048.2	11672.0	93.91	169.9	61370	2.0491	7.7679
AM	14048.20	11672.0	93.91	169.90	61370	2.05	7.76
H	11455.60	3247.30	181.17	1.00	69526	2.23	6.60
MA	5415.34	2462.80	124.87	3.10	24245	1.22	3.65
SH	4383.80	2996.10	97.74	6.10	17961	1.45	4.96
GR-I	24882.40	22955.00	67.05	24.60	64240	0.63	3.42

The coefficient of skewness varying from 0.63 to 3.28 shows that distributions of selected rocks are positively skewed. Therefore, it would be appropriate to select positively skewed distribution(s) as a candidate for the observed data series of selected rocks. Similarly, the range of CV varies from 67.05 to 200.43 that means that there is a significant variation in the materialization of rocks. Further, all coefficients of kurtosis for the rocks data are greater than three, revealing that distributions of observed datasets are leptokurtic having a wider or flatter shape with fatter tails than the normal distribution.

3.1. Parameter estimates for LND and GEVD

In this work, LND and GEVD are considered for the analysis of selected rocks, and their parameters are estimated by only the ML method. The estimates of parameters are provided in Table 3.

3.2. KS test for LND and GEVD

To examine the suitability of LND and GEVD for the rock dataset, the values of the KS test and P-values are shown in Table 4. Based on the p-values of KS test, the LND provides a good fit to all rocks except GR-I. Similarly, p-values of KS test reveals that GEVD is good fit to all rocks excluding SH and H rocks at 5% level of significance based on the ML estimation method. These distribution functions can be used for the characterization of the selected rocks. However, AIC and BIC are considered to pick the most suitable distribution for the remaining rocks, and the result are presented in Table 5.

Table 3. Estimates of parameters for LND and GEVD.

	LND		GEVD		
Rocks	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\eta}$	$\hat{\alpha}$	$\hat{\theta}$
D	5.8612	1.2482	2.0457	177.4233	164.0803
G	5.8881	1.3728	2.3801	304.7810	199.3290
M	6.0055	0.8209	0.6257	204.2494	282.3678
B	5.8612	0.8927	0.7538	184.7516	236.5667
S	5.2855	1.2006	1.2358	89.4480	103.9812
L	6.1600	0.9854	0.5783	314.1833	350.6152
A	5.6786	0.7061	0.6180	117.2347	208.4598
SL	6.5563	1.1610	1.8278	451.7178	409.1760
DO	6.7865	0.9399	0.8528	502.0804	599.2422
GR-II	6.5245	0.9647	0.9581	410.1837	473.1360
GN	5.9006	0.8918	0.5269	217.0383	269.1082
SC U	6.6372	1.1035	0.9055	527.9568	522.3487
AM	9.0397	1.2572	0.2535	7588.4837	7920.6838
H	6.9624	3.1104	2.0155	4619.3278	2132.2987
MA	6.7111	2.8486	1.6992	3427.2400	1790.5409
SH	7.5520	1.8718	0.3511	2568.4576	2222.6363
GR-I	9.6268	1.5663	0.1831	2655.3655	3193.7632

Table 4. KS test for LND and GEVD

	LND		GEVD	
Rocks	**KS**	**P-values**	**KS**	**P-values**
D	2.0457	0.6611	0.1644	0.3171
G	2.3801	0.3191	0.1953	0.1520
M	0.6257	0.7527	0.1015	0.8634
B	0.7538	0.5040	0.1081	0.8099
S	1.2358	0.2072	0.1037	0.8467
L	0.5783	0.9372	0.0903	0.9355
A	0.6180	0.2608	0.1262	0.6416
SL	1.8278	0.3368	0.1494	0.4305
DO	0.8528	0.8163	0.1058	0.8301
GR-II	0.9581	0.4443	0.1325	0.5822
GN	0.5269	0.9479	0.1006	0.8707
SC	0.9055	0.9807	0.0840	0.9633
AM	0.2535	0.4216	0.0995	0.8787
H	2.0155	0.1376	0.2883	**0.0106**
MA	1.6992	0.2164	0.2758	0.0604
SH	0.3511	0.1108	0.2636	**0.0381**
GR-I	0.1831	**0.0002**	0.7740	0.5456

The lowest AIC and BIC values nominate that LND is the best-fitted distribution for all rocks excluding GR-I rock because the KS test confirm that LND is inappropriate for the GR-I rock. Hence, the remaining rocks favor LND.

3.3. Parameter estimates of FD, WD, and LLD

The estimates of parameters for FD, WD, and LLD are presented in Table 6 and KS test along with p-values are shown in Table 7. The values of AIC and BIC are listed in Table 8 for comparison purposes.

Table 5. AIC and BIC values for LND and GEVD

	LND		GEVD	
Rocks	**AIC**	**BIC**	**AIC**	**BIC**
D	77.0011	79.9326	466.78	471.17
G	83.0907	86.0222	481.60	486.00
M	50.1815	53.1130	468.93	473.33
B	55.5478	58.4793	464.25	468.65
S	74.5127	77.4442	435.08	439.48
L	61.8708	64.8023	492.42	496.82
A	40.5402	43.4717	433.81	438.21
SL	72.3662	75.2976	514.96	519.36
DO	58.8453	61.7768	525.53	529.93
GR-II	60.5121	63.4436	509.36	513.76
GN	55.4833	58.4148	468.53	472.93
SC	69.1153	72.0468	528.82	533.21
AM	77.4609	80.3924	683.76	688.16
H	127.2215	130.0239	608.02	612.23
MA	94.4253	96.6963	451.11	454.51
SH	87.4754	90.0670	519.87	523.76
GR-I	72.3828	74.8205	456.96	459.40

Table 6. Estimates of parameters for FD, WD, and LLD

		FD		WD		LLD	
Rocks	**Methods**	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$
D	ML	1.0865	187.9471	0.7123	651.6142	292.9720	1.3833
	Bayesian	1.0710	183.7022	0.6955	599.3741	278.8408	1.3576
G	ML	0.9569	190.7382	0.6755	753.7369	321.3974	1.2304
	Bayesian	0.9439	185.2517	0.6629	691.1234	301.3821	1.2069
M	ML	1.3802	271.7232	1.2322	615.6274	398.0460	2.0514
	Bayesian	1.3625	268.2754	1.2107	598.1170	388.7729	2.0089
B	ML	1.3015	228.0841	1.1200	551.1799	339.6777	1.8889
	Bayesian	1.2850	224.7495	1.0908	533.8703	330.2042	1.8496
S	ML	1.1484	114.8727	0.7548	374.5018	171.3085	1.4535
	Bayesian	1.1308	112.5649	0.7316	347.2040	164.0625	1.4267
L	ML	1.0521	287.7802	1.1142	766.0499	480.5195	1.7161
	Bayesian	1.0416	280.9123	1.0829	736.8644	464.5011	1.6773
A	ML	1.7773	210.8105	1.2746	423.5877	275.5477	2.4527
	Bayesian	1.7501	209.4598	1.2476	412.7836	271.4038	2.4056
SL	ML	1.1407	412.3683	0.7856	1313.2173	634.8851	1.4737
	Bayesian	1.1240	404.1713	0.7605	1196.0502	607.7080	1.4450
DO	ML	1.2716	566.8772	0.9815	1433.9529	848.6716	1.8491
	Bayesian	1.2551	558.1840	0.9607	1386.1132	824.8620	1.8101
GR-II	ML	1.2809	432.9737	0.9773	1142.3571	638.2448	1.7731
	Bayesian	1.2627	426.3139	0.1780	1069.7654	618.8390	1.7377
GN	ML	1.1968	234.4492	1.1162	569.4417	368.3003	1.9173
	Bayesian	1.1829	230.3673	1.0998	551.6636	358.1094	1.8745
SC	ML	1.0267	445.7392	0.9172	1332.8083	738.9358	1.5406
	Bayesian	1.0155	434.3211	0.8971	1292.4002	709.5975	1.5082
AM	ML	0.6206	4175.1757	1.1071	14352.8275	9914.3022	1.5590
	Bayesian	0.6188	3848.0437	1.0757	14082.5904	9503.1300	1.5175
H	ML	0.3110	207.0785	0.4022	4384.1927	1412.3704	0.5409
	Bayesian	0.3124	143.2378	0.1110	4149.4567	946.4194	0.5205
MA	ML	0.3357	182.1433	0.1178	1925.1442	1119.0672	0.5931
	Bayesian	0.3363	120.3467	0.1150	1810.2320	716.5817	0.5630
SH	ML	0.4148	659.9815	0.1209	3884.1256	2644.4202	1.0854
	Bayesian	0.4155	524.9093	0.8202	3778.4034	2382.0662	1.0453
GR-I	ML	0.4062	5854.9203	0.0938	4056.7080	20121.0894	1.5600
	Bayesian	0.4080	4507.8164	0.0895	3878.9654	19108.4585	1.5033

The estimated parameters of FD, WD, and LLD for seventeen rocks by using two methods of estimations are shown in Table 6. It is noted that the estimates of shape parameters of these distribution both ML and the Bayesian methods are almost the same. However, the noteworthy difference can be observed in the estimates of scale parameters may be due to dissimilar characteristics of rocks.

Table 7. KS test for FD, WD, and LLD.

		FD		WD		LLD	
Rocks	**Methods**	**KS**	**P-value**	**KS**	**P-value**	**KS**	**P-value**
D	ML	0.1429	0.4863	0.2040	0.1209	0.1533	0.3993
	Bayesian	0.1435	0.4814	0.2215	0.0737	0.1591	0.3542
G	ML	0.1449	0.4693	0.2027	0.1848	0.1403	0.5103
	Bayesian	0.1461	0.4585	0.1461	0.1250	0.1544	0.3909
M	ML	0.1148	0.7504	0.1156	0.7432	0.1140	0.7574
	Bayesian	0.1188	0.7126	0.1023	0.8576	0.1002	0.8728
B	ML	0.1015	0.8640	0.1396	0.5165	0.1356	0.5529
	Bayesian	0.1035	0.8482	0.1239	0.6644	0.1207	0.6948
S	ML	0.1259	0.6454	0.1893	0.1773	0.1440	0.4773
	Bayesian	0.1161	0.7384	0.1950	0.1532	0.1391	0.5212
L	ML	0.1115	0.7802	0.0894	0.9402	0.0917	0.9283
	Bayesian	0.1157	0.7418	0.0706	0.9938	0.1012	0.8659
A	ML	0.1335	0.5725	0.1593	0.3534	0.1543	0.3912
	Bayesian	0.1289	0.6164	0.1527	0.4038	0.1431	0.4847
SL	ML	0.1216	0.6863	0.1876	0.1849	0.1342	0.5663
	Bayesian	0.1226	0.6768	0.2090	0.1050	0.1447	0.4708
DO	ML	0.1180	0.7206	0.1408	0.5053	0.1058	0.8301
	Bayesian	0.1215	0.6873	0.1506	0.4212	0.0912	0.9306
GR-II	ML	0.0908	0.9329	0.1556	0.3811	0.1300	0.6057
	Bayesian	0.0850	0.9595	0.8088	0.4567	0.1150	0.7494
GN	ML	0.1198	0.7034	0.0908	0.7143	0.0947	0.9107
	Bayesian	0.1251	0.6530	0.0985	0.8859	0.1017	0.8616
SC	ML	0.1138	0.7592	0.1187	0.7592	0.0763	0.7592
	Bayesian	0.1212	0.6902	0.1098	0.7952	0.0745	0.9884
AM	ML	0.2203	0.0761	0.1040	0.8441	0.1324	0.5825
	Bayesian	0.2092	0.1046	0.1019	0.8610	0.1283	0.6218
H	ML	0.2281	0.0747	0.1567	0.4108	0.1727	0.2972
	Bayesian	0.2616	0.2067	0.6392	0.4786	0.2201	0.0932
MA	ML	0.2397	0.1423	0.6540	0.3214	0.1815	0.4345
	Bayesian	0.2790	0.5557	0.6544	0.4987	0.2381	0.1474
SH	ML	0.3052	0.1089	0.6953	0.4998	0.1349	0.6608
	Bayesian	0.4114	0.4321	0.1560	0.4799	0.1635	0.4207
GR-I	ML	0.3147	0.4356	0.7988	0.5783	0.1405	0.6559
	Bayesian	0.3162	0.6734	0.7912	0.6964	0.1564	0.5230

The values of KS test statistics along with P-values are listed in Table 7 as a measure of goodness. Since all the P-values of the KS test are greater than a 5% level of significance, therefore, it is determined that these distributions are seemed to be good for all selected rocks based on both ML and Bayesian estimation methods. Additionally, the AIC and BIC values are calculated and presented in Table 8 to select the preferable distribution having the smallest value of AIC and BIC respectively.

It can be seen that the values of AIC and BIC are not significantly different from each other by using the two methods of estimation. According to the AIC and BIC values, FD is selected as the best fit for A, M, B, DO, and GR-II rocks, whereas D, G, S, SL, SC, H, SH, and GR-I rocks favor WD, and some of the rocks such as MA, GN, L, and AM favor LLD. AIC and BIC placed the WD, FD, and LLD models as the first, second, and third best-fit models, respectively, for nominated rocks.

Table 8. AIC and BIC values for FD, WD, and LLD.

		FD		WD		LLD	
Rocks	**Methods**	**AIC**	**BIC**	**AIC**	**BIC**	**AIC**	**BIC**
D	ML	473.6139	476.5454	456.7162	459.6477	481.3516	484.2831
	Bayesian	484.2153	487.1468	503.8246	506.7561	493.2935	496.2250
G	ML	486.6515	489.5830	463.5745	466.5059	493.7247	496.6562
	Bayesian	497.0230	499.9545	513.4518	516.3833	505.6026	508.5341
M	ML	467.1159	470.0474	538.2613	541.1928	468.5567	471.4882
	Bayesian	478.9450	481.8764	485.0793	488.0107	481.9221	484.8536
B	ML	462.2042	465.1357	496.7739	499.7053	464.6648	467.5963
	Bayesian	473.5634	476.4948	481.3624	484.2939	477.5441	480.4756
S	ML	437.2573	440.1887	423.4384	426.3698	445.5929	448.5244
	Bayesian	446.9865	449.918	466.5572	469.4886	456.5663	459.4978
L	ML	492.4647	495.3961	516.7478	519.6792	490.0251	492.9565
	Bayesian	503.8579	506.7893	503.4029	506.3343	503.3995	506.3310
A	ML	431.9120	434.8435	530.9316	533.8631	436.8638	439.7953
	Bayesian	443.7424	446.6739	458.8592	461.7907	449.8618	452.7932
SL	ML	518.1877	521.1192	505.2545	508.186	525.2727	528.2042
	Bayesian	530.4608	533.3923	547.8733	550.8048	538.8921	541.8235
DO	ML	523.279	526.2105	531.0837	534.0152	525.9460	528.8775
	Bayesian	536.4112	539.3426	548.9479	551.8794	540.6139	543.5453
GR-II	ML	506.9453	509.8768	514.5306	517.4620	511.5045	514.4360
	Bayesian	519.5525	522.4839	607.7854	610.7169	525.5172	528.4487
GN	ML	468.1043	471.0358	497.4003	502.2101	466.5256	469.4571
	Bayesian	479.3489	482.2804	483.2575	486.189	479.5954	482.5269
SC	ML	526.4384	529.3698	517.5170	520.4485	527.6416	530.5730
	Bayesian	538.6529	541.5843	545.5812	548.5126	541.6553	544.5868
AM	ML	705.0262	707.9577	704.3107	707.2422	684.1151	687.0465
	Bayesian	720.6610	723.5924	698.1360	701.0674	703.3386	706.2700
H	ML	582.2234	585.0258	522.2999	525.1023	576.9659	579.7683
	Bayesian	590.1953	592.9977	636.4553	639.2577	589.8192	592.6216
MA	ML	432.1751	434.4461	385.5196	387.7906	427.5994	429.8704
	Bayesian	439.9985	442.2695	474.3223	476.5933	440.1165	442.3875
SH	ML	537.0968	539.6885	492.7678	495.3594	518.2645	520.8562
	Bayesian	548.0959	550.6876	525.9263	528.5180	534.0485	536.6401
GR-I	ML	599.7428	602.1805	567.9348	570.3725	569.8992	572.3725
	Bayesian	615.0365	617.4743	675.2012	677.6390	588.5219	590.9597

3.4. Evaluation of best-fit distribution model

The selection of best-fitted distributions with two methods of estimation can be explored using various graphical functions. The plots of PDF for FD, WD, and LLD have been constructed and presented in Figures 5-7. The histogram of the observed datasets superimposed the PDF of the proposed theoretical fitted distributions. It is also noticed that all rocks have a right-skewed distribution, where the tail of the distribution is longer to the right-hand side compared to the left-hand side. The plots of FD, WD, and LLD models seem to good fit the observed dataset series, and thus may be the preferred models for this dataset. Thus the WD, FD, and LLD models are the best possible choices for rocks analysis.

3.6. Kruskal Wallis test

Table 9 presents the values of the Kruskal Wallis (H) test statistic. The test statistic (H) had a p-value of 0.000, indicating that the null hypothesis could be rejected at a 5% level of significance, which is in favor of the alternative hypothesis that not all medians of rocks are the same. It ensures that rocks samples do not have identical distribution and the selected rocks differ significantly from each other concerning their characteristics.

Table 9. Kruskal Wallis test for rocks samples.

H	DF	P-value
173.74	16	0.000

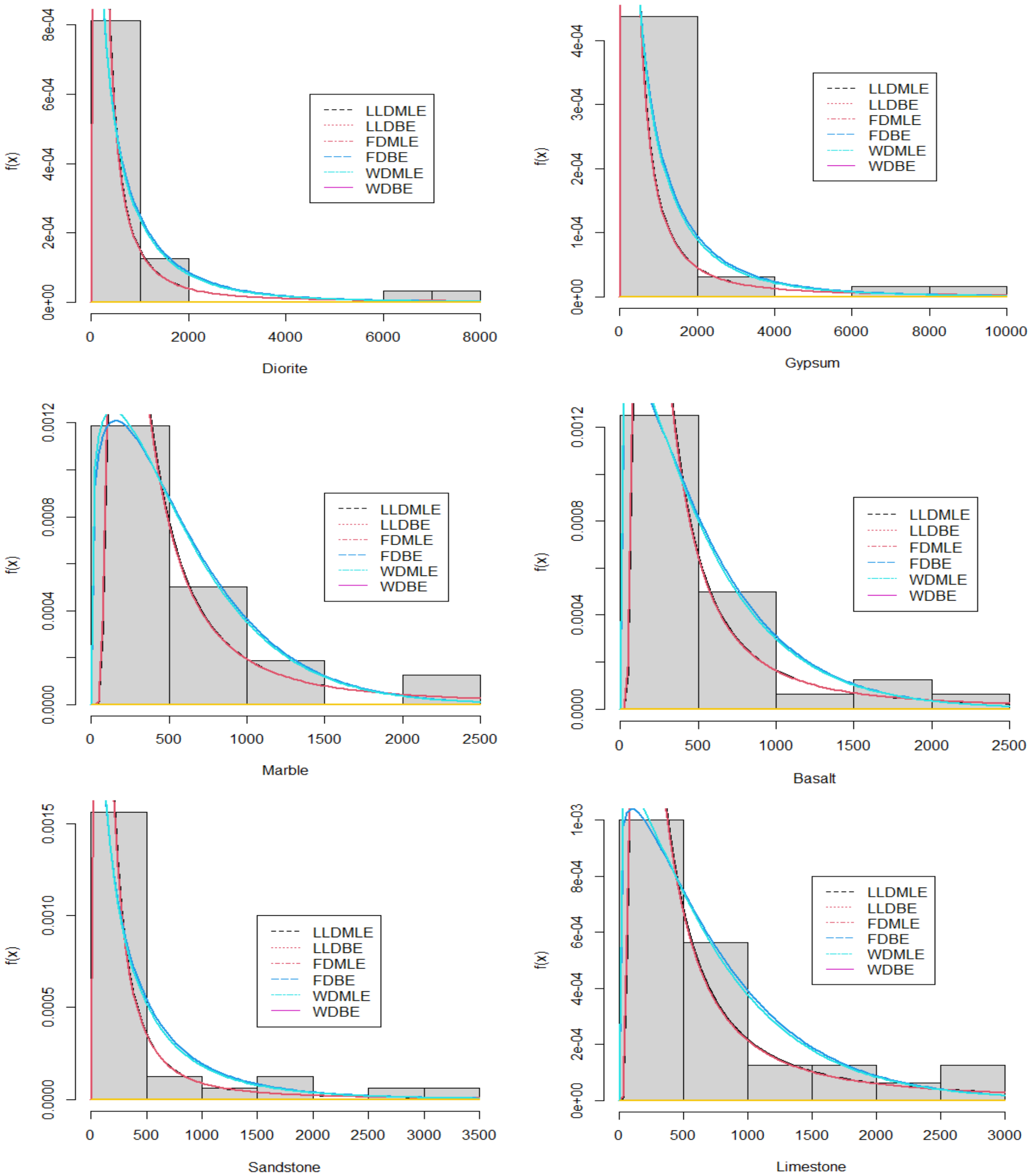

Figure 5. PDF plots of FD, WD, and LLD for D, G, M, B, S, and L.

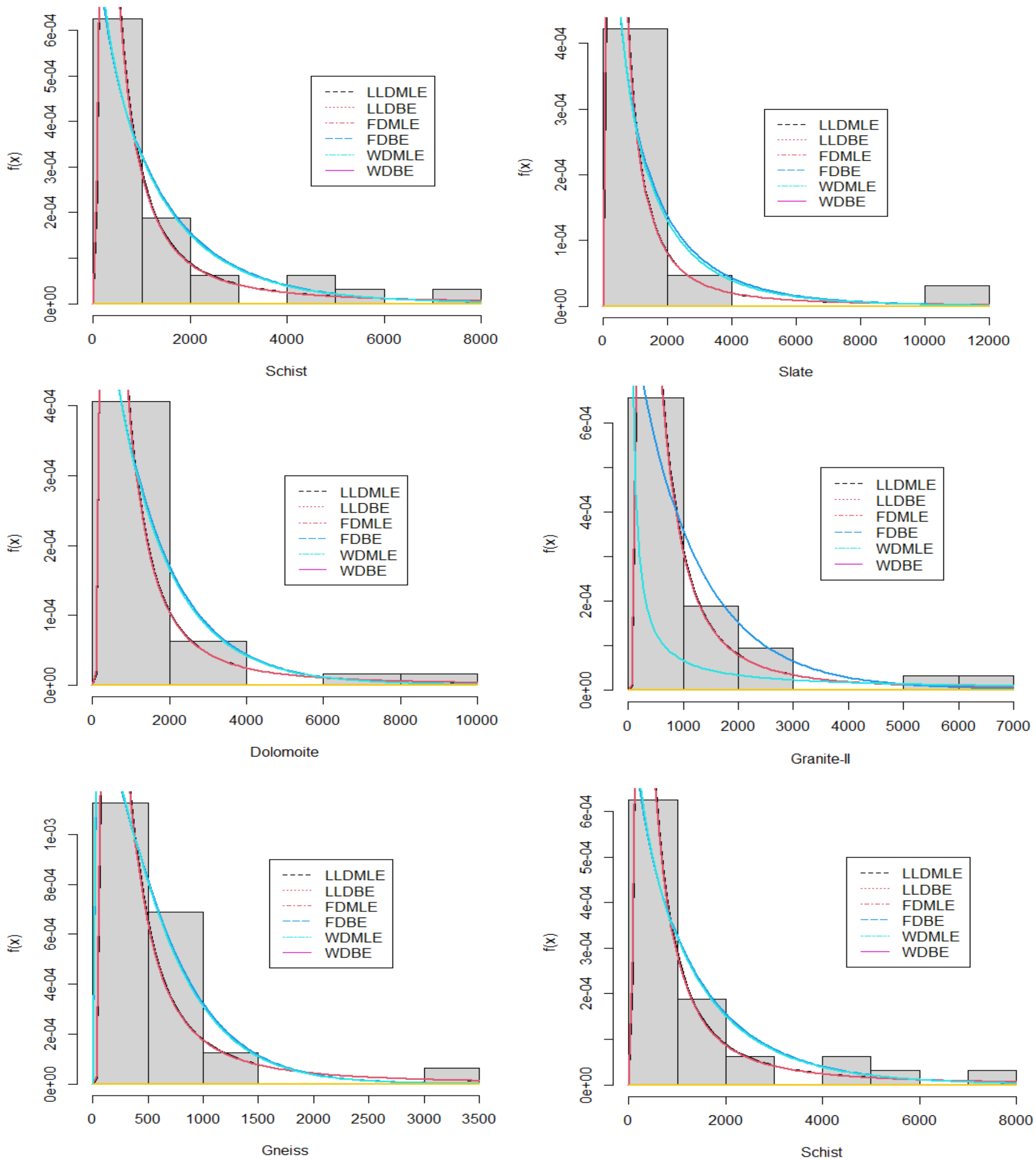

Figure 6. PDF plots of FD, WD, and LLD for A, SL, DO, GR-II, GN and SC.

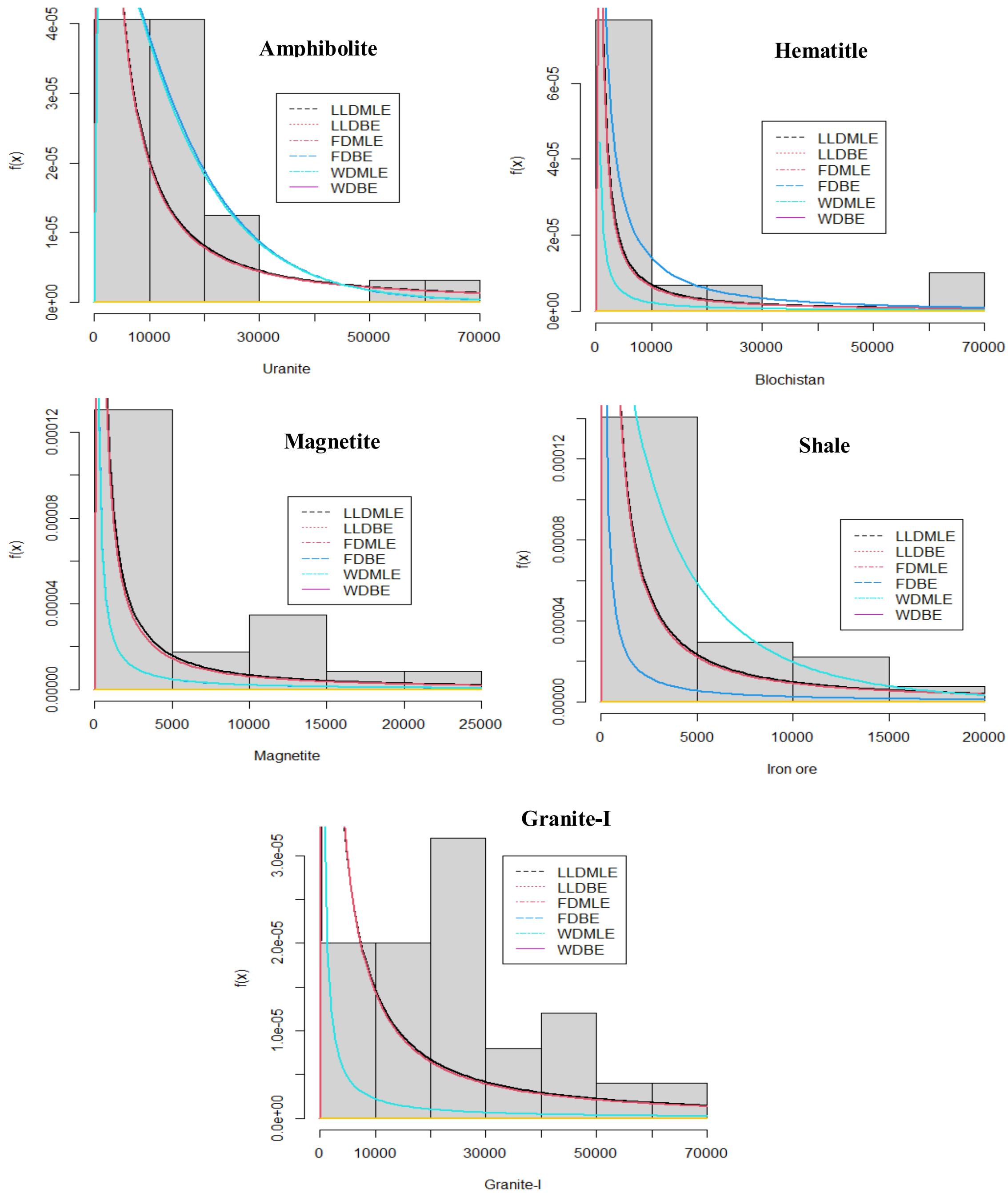

Figure 7. PDF plots of FD, WD, and LLD for AM, H, MA, SH, and GR-I rocks.

3.7. Cluster analysis

Cluster analysis classifies the number of rocks into clusters. Each cluster consists of two rocks. The number of clusters, the corresponding similarity level, the distance between them, which clusters were joined, the identification number of the new cluster, and the number of rocks in the new cluster are displayed in Table 10. In step one, two rocks are joined to form a new cluster. This step creates 16 clusters in the data with a similarity level of 100.00 and a distance level of 0.00000. The similarity level decreases slightly from step one and abruptly decreases in step nine, and the number of clusters is changed from 10 to 1. At each following step, as new clusters are formed similarity, level decreases, and the distance level increases.

Table 10. Cluster analysis of rocks

Step	No. of clusters	Similarity level	Distance level	Clusters joined	New Clusters	No. of Rocks in new clusters
1	16	100.00	0.000000	10 11	10	2
2	15	99.83	0.003389	1 2	1	2
3	14	99.173	0.016547	8 9	8	2
4	13	98.752	0.024957	1 8	1	4
5	12	98.665	0.026710	3 4	3	2
6	11	98.373	0.032533	5 12	5	2
7	10	97.915	0.041708	1 3	1	6
8	9	97.907	0.041861	1 5	1	8
9	8	88.608	0.227845	1 6	1	9
10	7	77.506	0.449881	1 7	1	10
11	6	76.419	0.471628	16 17	16	2
12	5	68.906	0.621879	1 10	1	12
13	4	60.784	0.784312	1 16	1	14
14	3	59.386	0.812272	1 15	1	15
15	2	53.974	0.920526	1 13	1	16
16	1	50.694	0.986116	1 14	1	17

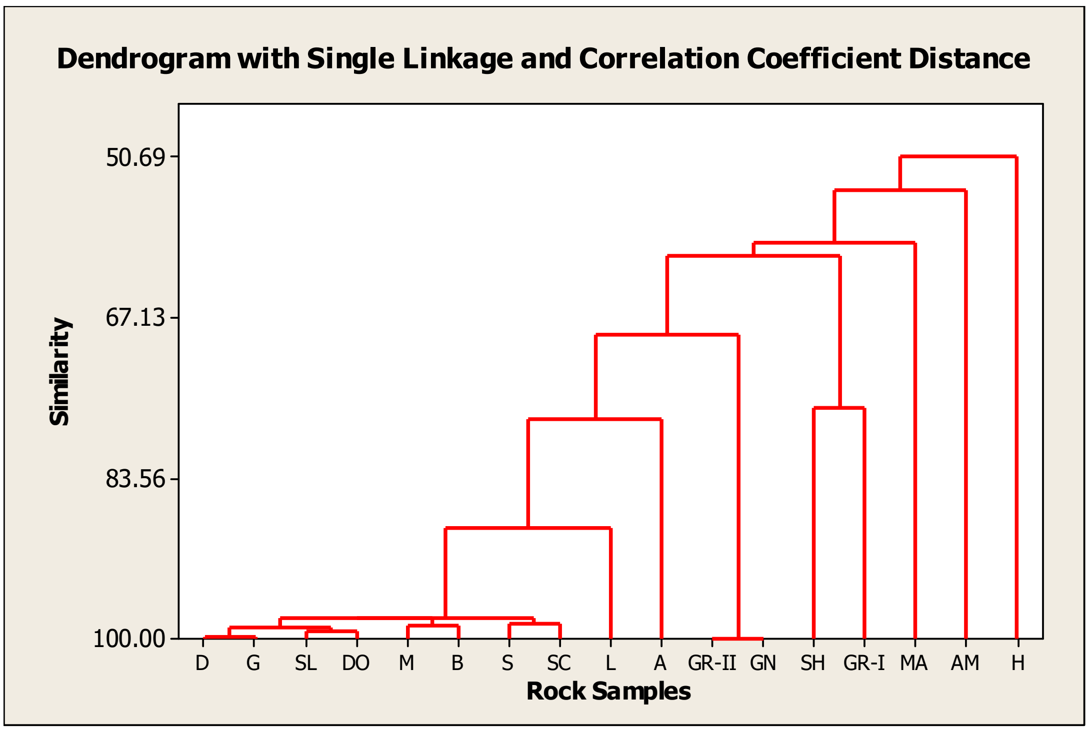

Figure 8. Dendrogram for rock samples.

In Figure 8, the horizontal axis of the dendrogram represents rocks samples, whereas the vertical axis denotes the similarity level between clusters. The dendrogram shows the information printed in the amalgamation table (Table 10) in the form of a tree diagram. In the above figure, the dendrogram proposes those rocks that are combined based on their similarity level. Figure 8 shows that D and G are placed in one cluster due to closer similarity levels. Similarly, SL and DO, M and B, S and SC, GR-II, and GN are established different clusters, respectively, while the similarity level of the remaining rocks namely L, A, SH, GR-I, MA, AM, and H are not identical and consequently these rocks constituted separate clusters.

4. Conclusions

The present work reviewed the methods of identification for the suitable probability distribution models applicable on the optical emission data of the rock samples for the selection of the best materialistic description of rocks. For the very first time, these distributions were utilized on the output data obtained using the LIBS spectroscopy. Five probability distribution models such as three-parameter distribution (GEV) and four two-parameter distributions (FD, WD, LLD, LND) were assessed using the Bayesian and ML estimation method, goodness of fit tests-based analysis to identify the most suitable distribution model for seventeen rocks. Therefore, the KS test was applied as an evaluator to judge the appropriateness of selected distributions. Moreover, AIC and BIC were used for preference and endorsement of the most appropriate distribution for the selected rocks. LND provided a good fit to all rocks except GR-I. Similarly, GEVD is a good fit for all rocks apart from SH and H rocks based on the p-values of the KS test at a 5% level of significance. It can be concluded that most of the rocks favor WD, and some of the rocks favor FD as well as LLD as the best-fitted probability distribution. Consequently, AIC and BIC positioned the WD, FD, and LLD models as the first, second, and third best-fit models, respectively, for the selected rocks. This work suggests that WD, FD, and LLD are preferable choices in modeling rocks data series in Pakistan. The results from this work will give benefits to the geologists and spectroscopists to build a better explanation about the materialistic characteristics of rocks.

Competing interests

There is not any financial or non-financial relationships/interest in this submission.

Data availability

Maximum data is given in the form of tables and software used in this study also is easily available. If any reader wants to get more details about the data or software, it can be provided on request.

Acknowledgments

We are grateful to the Atomic and Molecular Physics Laboratory and the department of physics, Quaid-i-Azam University, Islamabad, Pakistan for providing the necessary research facilities.

Disclosure statement

The authors have no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals.

References

[1]. Noori, M. Khanlari, G. Rafiei, B. Sarfarazi, and V. Zaheri, M. (2022). Correction to: Estimation of Brittleness Indices from Petrographic Characteristics of Different Sandstone Types (Cenozoic and Mesozoic Sandstones). Markazi Province, Iran, 55, 6519.

[2]. Yavari, M. D. Haeri H. Sarfarazi, V. Fatehi M. and Lazemi, H. A. (2021). Compressive Failure Analyses of Rock-like Materials by Experimental and Numerical Methods, Journal of Mining and Environment. 12 (3): 769-783.

[3]. Sarfarazi, V. Haeri, H. and Fatehi M. (2021). On Direct Tensile Strength Measuring of Anisotropic Rocks. Journal of Mining and Environment. 12 (2): 491-499

[4]. Zaman, M. A. Rahman, A. Haddad, K. and Hagare, D. (2012). Identification of the best fit probability distributions in at-site flood frequency analysis: A case study for Australia using 127 stations. Hydrology and Water Resources Symposium, 939-945.

[5]. Azizi, M.A. Karamadibrata, S. Wattimena, R.K. and Sidi, I.D. (2013). Characterization of the distribution of physical and mechanical properties of rocks at Tatupan coals mine, South Kalimanta, Indonesia. Rocks Mechanics and Resources Energy and Engineering, 79, 213-216.

[6]. Ghazdali, O. Guemouria, A. Rziki, S. and Moustadraf, J. (2021). Statistical analysis of rocks mass using data mining technique: a study case in Morocco. Journal of Physics: Conference Series. 1743 (1): 107-114.

[7]. Malkowski, P. Niedbalski, Z. and Balarabe, T. (2020). A statistical analysis of geomechanical data and its effect on rock mass numerical modeling: a case study. International Journal of Coal and Science Technology. 8 (2): 312-323.

[8]. Teymen, A. and Menguc, E.C. (2020). Comparative evaluation of different statistical tools for the prediction of uniaxial compressive strength of rocks. International Journal of Mining Science and Technology. 30 (6): 785-797.

[9]. Gent, V. Almeida, E. and Hofland, B. (2019). Statistical analysis of the stability of rock slopes. Journal of Marine Sciences and Engineering. 7 (3): 60- 75.

[10]. Cai, X. Zhou, Z. Liu, K. Du, X. and Zang, H. (2019). Water weaking effects on the mechanical behavior of different rock types: phenomenon and mechanics. Applied Sciences. 9 (20): 4450.

[11]. Salih, H. S. and Alshkane, Y. (2018). Statistical analysis of physical and mechanical properties of igneous rocks. Journal of Garmian University. 5 (2): 174-189.

[12]. Mayer, J.M. Allen, D.M. Gibson, H.D. and Mackie, D.C. (2014). Application of statistical

approaches to analyze geological, geotechnical, and hydrogeological data at a fractured- mine site in northern Canada. Hydrogeology Journal. 22 (7): 1707-1723.

[13]. Karakul, H. and Ulusay, R. (2013). Emprical correlations for predicting strength properties of rocks from P-wave velocity under different degree of saturation. Rock Mechanics and Rock Engineering. 46 (5): 981-999.

[14]. Ceryan, N. Okkan, U. and Kesimal, A. (2012). Application of the generalized regression neural networks in predicting the Unconfined Compressive Strength of carbonate rocks. Rock Mechanics and Rocks Engineering. 45 (6): 1055-1072.

[15]. Ghazvinian, A. and Hadei, M. R. (2012). Effect of discontinuity orientation and confinement on the strength of jointed anisotropic rocks. International Journal of Rock Mechanics and Mining Sciences. 55 (1): 117-124.

[16]. Huang, S. Xia, K. Yan, F. and Feng, X. (2010). An experimental study of the rate dependence of tensile strength softening of Longyou sandstone. Rock Mechanics and Rock Engineering. 43 (6): 677-683.

[17]. Huang, W. Xing, W. Chen, S. Yang, L. and Wu, K. (2017). Experimental study on sedimentary rock's dynamic characteristics under creep state using a new type of testing equipment. Advanced in Materials Sciences and Engineering, 1, 1155-1168

[18]. Liu, Z. Guo, Y. Du, S. Wu, G. and Pan, M. (2017). Research on calibrating rock mechanical parameters with a statistical method. Economic Geology. 58 (8): 1246-1266.

[19]. Mibei, G. (2014). Introduction to types and classification of rocks. Geothermal Development Company, 51(6), 143-155.

[20]. Rybar, P. Strba, L. Molokac, S. and Hvizdak, L. (2015). Study of physical-mechanical rock properties for rock disintegration purposes. Podzemni radovi. 21 (5): 131-133.

[21]. Singh, T. N. Kainthola, A. and Venkatesh, A. (2012). Correlation between point load index and uniaxial compressive strength for different rock types. Rock Mechanics and Rock Engineering. 45 (2): 259-264.

[22]. Wang, M. Wan, W. and Zaho, Y. (2019). Prediction of uniaxial compressive strength of rocks from simple index tests using a random forest predictive model. Rock Mechanics and Rock Engineering. 348 (1): 3-32.

[23]. Zadhesh, J. and Majdi, A. (2022). Determining Probability Distribution Functions of Rock Joint Geometric Properties, Journal of Mining and Environment (JME) Vol. 13, No. 1, 2022, 281-308.

[24]. Ahmed, N. Awan, J.A. Fatima, K. Iqbal, S.M.Z. Rafique, M. Abbasi, S. A. and Baig, M. A. (2022). Machine learning-based calibration LIBS analysis of aluminium-based alloys. Eur. Phys. J. Plus., 137, 671.

[25]. Ahmed, N. Ahmed, R. and Baig, M.A. (2017). Analytical Analysis of Different Karats of Gold Using Laser Induced Breakdown Spectroscopy (LIBS) and Laser Ablation Time of Flight Mass spectrometer (LA-TOF-MS). Plasma Chem Plasma Process, 38, 207-222.

[26]. Umar, Z.A. Ahmed, N. Ahmed, R. Liqat, U. and Baig, M.A. (2018). Elemental composition analysis of granite rocks using LIBS and LA-TOF-MS. Applied Optics, 57, 4985-4991.

[27]. Ahmed, N. Zeshan, A. Umar, Z.A. Ahmed, R. and Baig, M.A. (2017). On the elemental analysis of different cigarette brands using laser induced breakdown spectroscopy and laser-ablation time of flight mass spectrometry. Spectrochimica Acta Part B 136, 39–44.

[28]. Ahmed, N. Ahmed, R. Umar, Z. A. Liaqat, U. Manzoor U. and Baig, M. A. (2018). Qualitative and Quantitative Analyses of Copper Ores collected from Baluchistan. Pakistan using LIBS and LA-TOF-MS, Applied Physics B, 124, 160.

[29]. Akhtar, M. Jabbar, A. Mehmood, S. Ahmed, N. Umar, Z.A. Ahmed, R. and Baig, M.A. (2018). Magnetic Field Enhanced Detection of Heavy Metals in Soil using Laser Induced Breakdown Spectroscopy. Spectrochimica Acta Part B 148 143–151.

[30]. Abbas, K. and Tang, Y. (2015). Analysis of Frechet distribution using reference priors. Communications in Statistics-Theory and Methods. 44 (14): 2945–2956.

[31]. Sun, D. (1997). A note on non-informative priors for Weibull distributions. Journal of Statitical Planning and Inference.

[32]. Abbas, K. and Tang, Y. (2016). Objective Bayesian Analysis for Log-logistic Distribution. Communications in Statistics-Simulation and Computation, 45, 2782–2791.

[33]. Kolmogorov, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. Giornale dell, Instituto Italiano degli Attuari, 4, 83-91.

[34]. Kruskal, W.H. and Wallis, W.A. (1952). Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association. 47 (260): 583-621.

انتخاب توزیع‌های احتمالی مناسب برای آنالیز سنگ با استفاده از طیف‌سنجی شکست ناشی از لیزر

کامران عباس[1]، عادل نوازیش[1]، نوید فیروز[1] و نسار احمد[2]*

۱. گروه آمار، پردیس ملک عبدالله چتر کالاس، دانشگاه آزاد جامو و کشمیر، مظفرآباد، پاکستان

۲. گروه فیزیک، پردیس ملک عبدالله چتر کالاس، دانشگاه آزاد جامو و کشمیر، مظفرآباد، پاکستان

ارسال ۲۰۲۲/۱۰/۰۷، پذیرش ۲۰۲۲/۱۲/۱۲

* نویسنده مسئول مکاتبات: nasar.ahmed@ajku.edu.pk

چکیده:

در این کار تلاش شده است تا مناسب‌ترین توابع توزیع احتمال برای آنالیز هفده نمونه سنگ شامل دیوریت، گچ، مرمر، بازالت، ماسه سنگ، آهک، آپاتیت، slate، دولومیت، گرانیت-II، برازش و شناسایی شود. شیست، gneiss، آمفیبولیت، هماتیتل، مگنتیت، شیل، و گرانیت-I با استفاده از طیف‌سنجی شکست ناشی از لیزر، ارزیابی گرافیکی و تجسم تأیید می‌کند که مجموعه داده‌های سنگ دارای انحراف مثبت هستند. بنابراین توزیع‌های Frechet، Weibull، log-logistic، log-normal و تعمیم‌یافته به عنوان توزیع‌های کاندید در نظر گرفته می‌شوند و پارامترهای این توزیع‌ها با روش‌های حداکثر احتمال و تخمین Bayesian برآورد می شوند. خوب بودن آزمون برازش و معیارهای انتخاب مدل مانند آزمون کولموگروف-اسمیرنوف، معیار اطلاعات آکایک، و معیار اطلاعات Bayesian برای تعیین کمیت داده‌های پیش بینی شده با استفاده از توزیع‌های احتمال نظری استفاده می‌شود. نتایج نشان می‌دهد که توزیع‌های Frechet، Weibull و log-logistic بهترین توزیع احتمال برازش برای داده‌های سنگی هستند. همچنین از تحلیل خوشه‌ای برای طبقه‌بندی سنگ‌های منتخب که ویژگی‌های مشترک دارند استفاده می‌شود و مشاهده می‌شود که دیوریت و گچ در یک خوشه قرار گرفته‌اند. با این حال سنگ‌های تخته سنگ، دولومیت، مرمر، بازالت، ماسه سنگ، شیست، گرانیت-II و gneiss به خوشه‌های مختلفی تعلق دارند. به طور مشابه، سنگ آهک و آپاتیت در یک خوشه ظاهر می‌شوند. به همین ترتیب، شیل، گرانیت-I، مگنتیت، آمفیبولیت و هماتیتل در یک خوشه متفاوت ظاهر می‌شوند. کار فعلی نشان می‌دهد که جفت شدن طیف سنجی شکست ناشی از لیزر با ابزارهای آماری مناسب می‌تواند سنگ‌ها را بسیار کارآمد شناسایی و طبقه بندی کند.

کلمات کلیدی: گروه OCP، اولد عبدون، سری فسفات، مراکش مرکزی، معدن فسفات.