

Application of tests of goodness of fit in determining the probability density function for spacing of steel sets in tunnel support system

Farnoosh Basaligheh* and Ali Keyhani

Department of Civil Engineering, Shahrood University of Technology, Shahrood, Iran

Received 25 Jul. 2014; Received in revised form 23 Dec. 2015; Accepted 3 Jan. 2016

*Corresponding author Email: f_basaligheh@yahoo.com, f_basaligheh@shahroodut.ac.ir

Abstract

One of the conventional methods for temporary support of tunnels is to use steel sets with shotcrete. The nature of a temporary support system demands a quick installation of its structures. As a result, the spacing between steel sets is not a fixed amount and it can be considered as a random variable. Hence, in the reliability analysis of these types of structures, the selection of an appropriate probability distribution function of spacing of steel sets is essential. In the present paper, the distances between steel sets are collected from an under-construction tunnel and the collected data is used to suggest a proper Probability Distribution Function (PDF) for the spacing of steel sets. The tunnel has two different excavation sections. In this regard, different distribution functions were investigated and three common tests of goodness of fit were used for evaluation of each function for each excavation section. Results from all three methods indicate that the Wakeby distribution function can be suggested as the proper PDF for spacing between the steel sets. It is also noted that, although the probability distribution function for two different tunnel sections is the same, the parameters of PDF for the individual sections are different from each other.

Keywords: *probability density function, random variable, reliability, steel sets, temporary support, tunnel.*

1. Introduction

In tunnel engineering, a temporary support structure is needed in many cases. Several type of temporary supports have been proposed and used in tunnels, including wood frames, rock bolts, cable bolts, steel frames, shotcrete, lining, and segments [1-3]. Generally, the most widely used type of temporary support system is shotcrete. Indeed, shotcrete is considered as the standard temporary support during design and

construction stages of tunnels. However, if the magnitude of loads transmitted by the ground to the support is too large to be carried by shotcrete alone or if squeezing or raveling behavior requires complete surface coverage, steel sets are commonly used in combination with shotcrete. This combination can be in the form of a complete composite annulus or may be a semi-circular or partial arch configuration [4].

Sometimes, lattice girders are used instead of steel sets [5-8]. In designing steel sets as the primary support system, two important parameters are the spacing between sets and the thickness of the shotcrete. Wong et.al have proposed and compared different design methods for steel sets [9]. Due to the shortage of time after the excavation, the installation of the steel sets and the casting of the concrete layer are generally performed very quickly. Therefore, it is observed that the actual spacing of the steel sets and the thickness of the shotcrete are unexpectedly different from the designed values. This difference may also occur as a result of special working conditions.

On the other hand, when the New Austrian Tunneling Method (NATM) is used, the spacing and the thickness of the shotcrete are changed according to the ground condition [10]. As a result, both of these parameters are not fixed and can be considered as random variables. Therefore, a probabilistic and reliability analysis of the system seems to be necessary for understanding the behavior of the support structure.

One of the significant problems of reliability analysis is the selection of an appropriate Probability Density Function (PDF) for random variables. During construction process of a project, there are many uncertainties such as natural and human-related uncertainties that affect the spacing between steel sets and thickness of the shotcrete. The recognition of each one of these uncertainties needs experience and one cannot properly know all of them [11]. The objective of the present research is to study the probability distribution function for spacing of steel sets used for primary support of the tunnels.

In this regard, the Goodness of Fit Tests are utilized to determine the proper density function. The common methods of goodness of fit are introduced and discussed in section 2. In section 3, some information about the spacing of steel sets in a tunnel in Iran is provided. The tunnel is excavated in moderate to highly weathered Andesite rocks. The nature of rock mass is blocky with maximum joint spacing equal to 60 cm (classified as fair rock with RMR between 38 and 53). The tunnel is excavated by drilling and blasting methods. It has two horseshoe excavation

sections: a large section with dimension of 8 x 8 m² and a small section with dimension of 5.1 x 5 m². The large section was excavated in two stages while the small section was excavated in only one. The spacing of steel sets was collected in 82 different locations for the larger tunnel section (8 x 8 m²) and in 108 different locations for the smaller section (5.1 x 5 m²). Based on this information, the proper density function for the spacing of the steel sets is selected and discussed. Section 4 includes conclusions and summary of the results.

2. Goodness of fit tests and estimation of probability distribution parameters

A Goodness of Fit Test is used to measure how well a sample of observed data follows a distinctive distribution function. It is noteworthy that no statistical distribution can precisely show a perfect fit with observed data. Therefore, one distribution is selected as the best one based on comparison with the other distributions [12]. In general, three tests of Goodness of Fit are used for determining the best probability distribution function. These three methods are Chi-Squared, Kolmogorov-Smirnov, and Anderson-Darling.

Each one of these methods has its advantages and disadvantages and one cannot simply prefer one of them. Therefore, in the present study, all three methods are used to determine the most suitable density function. Then, the desired result is obtained through scoring each one of these methods and averaging them [13-15]. In the following, each method is concisely introduced and discussed.

2.1. Chi-squared test

The Chi-squared test, like other tests, uses a test statistics for comparing the observed data with a distribution function. In this method, the observed data is partitioned to several categories and the test statistic is estimated from Equation (1).

$$\chi^2 = \sum_{i=1}^{NC} \frac{(f_i - \bar{f}_i)^2}{\bar{f}_i} \quad (1)$$

In the above equation, NC represents the number of categories. f refers to the frequency (number) of the observed (measured) samples in category i and \bar{f} represents the frequency of

category i as calculated by the assumed distribution function.

Based on this test, the numbers of existing samples in each category is recommended to be at least 5. χ^2 is the Chi-squared distribution with (NC-NP-1) degree of freedom in which NP is the number of parameters to be estimated. The above statistic is compared with the value of Chi-squared distribution with a definite level of acceptance and degree of freedom. If the calculated value of χ^2 in the above equation is less than the value of Chi-squared distribution, the hypothesis that the data follow the assumed distribution is accepted [16, 17].

2.2. Kolmogorov-Smirnov test

The Kolmogorov-Smirnov test is another test used to measure how well a sample follows a distinctive distribution. The Chi-squared test acts properly when the number of samples is large enough to have five data in each category. Otherwise, it is better to use Kolmogorov-Smirnov test. The statistic of this test is the highest difference between expected and real measured frequencies (in absolute value) in different groups. This statistic is determined through the following equation:

$$D = \max |F(x) - \hat{F}(x)| \tag{2}$$

In the above equation, F is the observed relative accumulative frequency and \hat{F} is the expected relative accumulative frequency. The following steps are taken to do this test.

1. The relative accumulative frequency of a sample is measured for different ranges (categories).
2. The relative accumulative frequency for different categories is obtained through theoretical statistical distribution or another instance of information.
3. The absolute value of difference between the two frequencies of the above two steps is calculated.
4. The highest value of difference obtained from step 3 is defined as the value of statistic for test D.
5. A value of α is selected as error and D_α is read from the associated tables.
6. If the value of D is more than D_α , the hypothesis that the sample follows the assumed distribution is accepted. Otherwise, it is rejected [16].

2.3. Anderson-Darling test

This test is another method for determining the fitness of an assumed probability density function to a given set of data. This test assigns higher weight to sequences in comparison with other tests, as a result, it has higher accuracy. For the variable x and presumed distribution $F^0(\cdot)$, the random variable $nF_n(x)$ is a binomial distribution with probability of $F^0(x)$. The expected value of $nF_n(x)$ is $nF^0(x)$ and its variance is $nF^0(x)[1-F^0(x)]$ and the weight function W_n^2 is calculated as:

$$W_n^2 = n \int_{-\infty}^{\infty} [F_n(x) - F^0(x)]^2 \Psi[F^0(x)] dF^0(x) = \tag{3}$$

$$n \int_{-\infty}^{\infty} [F_n(x) - F^0(x)]^2 \Psi[F^0(x)] f^0(x) dx$$

$$\Psi(u) = \frac{1}{u(1-u)} \tag{4}$$

For the values of x , we have:

$$\sqrt{n} \frac{F_n(x) - F^0(x)}{\sqrt{F^0(x)[1-F^0(x)]}} \tag{5}$$

The mean 0 and variance 1 are obtained when the null hypothesis is correct. The Anderson-Darling statistic is defined in the following:

$$A_n^2 = n \int_{-\infty}^{\infty} \frac{[F_n(x) - F^0(x)]^2}{F^0(x)[1-F^0(x)]} dF^0(x) \tag{6}$$

The above equation can also be written as:

$$A_n^2 = -n - \frac{1}{n} \sum_{j=1}^n (2j-1) [\log u_{(j)} + \log(1-u_{(n-j+1)})] \tag{7}$$

The Anderson-Darling test defines the distribution limit A_n for the weight function of Equation (4).

In the above equation, $u_j = F^0(x_{(j)})$ and $x_{(1)} < x_{(2)} < \dots < x_{(n)}$ is the sequence of ordered samples [18,19].

3. Random variable of applied spacing for steel sets

As mentioned before, the drilling of tunnels requires a quick covering and this is done by shotcreting. If the intensity of loads transferred from ground to bearing structure is high, then

the steel sets are used in combination with shotcrete for temporary support of the tunnel. The spacing between steel sets is a variable due to the quick installation. As a result, the distance between steel sets can be regarded as a random variable. In the present study, the data gathered from one of the tunnels in Iran is used to determine the distribution function of this random variable. The tunnel has two drilling horseshoe section as shown in Figures 1 and 2.

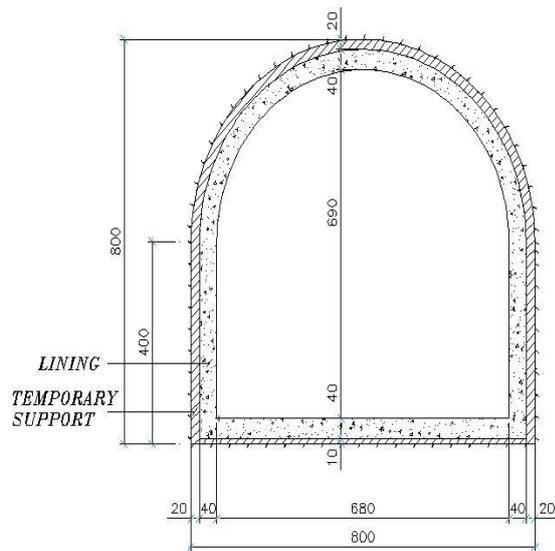


Fig. 1. Large section of tunnel with dimension of 8 x 8 m²

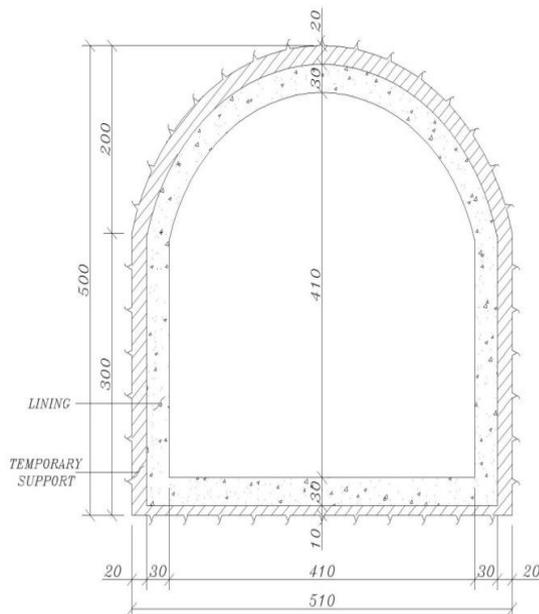


Fig. 2. Small section of tunnel with dimension of 5.1 x 5 m²

For the larger tunnel section (8 x 8 m²), the spacing of steel sets are collected in 82 different locations as shown in Table 1. The spacing of the steel sets for the smaller section (5.1 x 5 m²) are collected for 108 different locations as shown in Table 2.

Table 1. Spacing of the steel sets in the large section

Frame number	Location (m)	Distance to next frame (cm)	Frame number	Location (m)	Distance to next frame (cm)
1	0.00	70	42	36.36	100
2	0.70	70	43	37.36	104
3	1.40	82	44	38.40	100
4	2.22	80	45	39.40	100
5	3.02	82	46	40.40	100
6	3.84	83	47	41.40	100
7	5.67	82	48	42.40	98
8	5.49	82	49	43.38	103
9	6.31	82	50	44.41	102
10	7.13	82	51	45.42	102
11	7.95	82	52	46.45	95
12	8.77	83	53	47.40	100
13	9.60	82	54	48.40	100
14	10.42	83	55	49.40	101
15	11.25	80	56	50.41	102
16	12.05	85	57	51.43	98
17	12.90	81	58	52.41	99
18	13.71	81	59	53.40	103
19	14.52	101	60	54.43	82
20	15.53	62	61	55.25	100
21	16.15	83	62	56.25	98
22	16.98	82	63	57.23	99
23	17.80	80	64	58.22	100
24	18.60	84	65	59.22	102
25	19.44	80	66	60.25	97
26	20.24	108	67	61.22	101
27	21.32	101	68	62.23	102
28	22.33	100	69	63.25	98
29	23.33	102	70	64.23	103
30	24.35	102	71	65.26	101
31	25.37	104	72	66.27	99
32	26.41	101	73	67.26	104
33	27.42	102	74	68.30	100
34	28.44	100	75	68.30	100
35	29.44	104	76	70.30	101
36	30.48	96	77	71.31	72
37	31.44	95	78	72.03	102
38	32.39	99	79	73.05	95
39	33.38	95	80	74.00	100
40	34.33	102	81	75.00	100
41	35.35	101	82	76.00	118

Table 2. Spacing of steel sets in small section

Frame number	Location (m)	Distance to next frame (cm)	Frame number	Location (m)	Distance to next frame (cm)	Frame number	Location (m)	Distance to next frame (cm)
1	4.00	42	37	33.80	100	73	70.42	103
2	4.42	45	38	34.80	100	74	71.45	102
3	4.87	43	39	35.80	103	75	72.47	103
4	5.30	73	40	36.83	102	76	73.50	102
5	6.03	72	41	37.85	106	77	74.52	101
6	6.75	60	42	38.91	99	78	75.53	101
7	7.35	50	43	39.9	100	79	76.54	101
8	7.85	79	44	40.9	98	80	77.55	83
9	8.64	80	45	41.88	102	81	78.38	83
10	9.44	81	46	42.90	101	82	79.21	82
11	10.25	82	47	43.91	99	83	80.03	80
12	11.07	80	48	44.90	102	84	80.83	87
13	11.87	109	49	45.92	102	85	81.70	80
14	12.96	53	50	46.94	106	86	82.50	80
15	13.49	83	51	48.00	100	87	83.30	84
16	14.32	92	52	49.00	97	88	84.14	79
17	15.24	90	53	49.97	103	89	84.93	83
18	16.14	88	54	51.00	103	90	85.76	89
19	17.02	73	55	52.03	102	91	86.65	91
20	17.75	62	56	53.05	101	92	87.56	91
21	18.37	75	57	54.06	99	93	88.47	101
22	19.12	78	58	53.05	101	94	89.48	102
23	19.90	84	59	56.05	100	95	90.50	105
24	20.74	83	60	57.05	106	96	91.55	87
25	21.57	85	61	58.11	100	97	92.42	81
26	22.42	98	62	59.11	109	98	93.23	85
27	23.40	108	63	60.20	110	99	94.08	80
28	24.48	102	64	61.30	98	100	94.88	82
29	25.50	100	65	62.28	100	101	95.70	84
30	26.50	105	66	63.28	105	102	96.54	85
31	27.55	98	67	64.33	97	103	97.39	81
32	28.53	104	68	65.30	106	104	98.20	84
33	29.57	101	69	66.36	100	105	99.04	82
34	30.58	119	70	67.36	101	106	99.86	80
35	31.77	98	71	68.37	105	107	100.66	82
36	32.75	105	72	69.42	100	108	101.48	84

The probability distribution function of this variable was obtained by Easy Fit 5.5 Software and through the above mentioned three methods. From 65 distribution functions available in the software, 32 functions which best fit the data are selected and their rank (based on 3 mentioned tests of Goodness of Fit) are shown in Tables 3 and 4. The sum of scores for each function in different methods is provided and is used to define the best fit distribution function. The results are shown in Tables 5-7. As these tables indicate, the

Gumbel Min Function has the first rank in Kolmogorov-Smirnov and Anderson-Darling while the Chi-squared method shows rank of 21 for this function. This significant difference between the evaluations by different tests for Goodness of Fit indicates an irregular data that requires more examination and understanding before analysis. This difference can be attributed to the fact that the number of samples in some categories is low.

The Gumbel Min distribution function for the large section is shown in Figure 3. As it

can be seen in this Figure and Table 1, the number of samples in the interval of 86 to 94 centimeters is significantly less than the other intervals. The Gumbel Min distribution function for the small section is illustrated in Figure 4. This Figure, together with the data

indicated in Table 2, show that the number of samples in the interval of 86 to 97 centimeters is much less than the other intervals. From Figures 3 and 4, it is evident that these empty intervals are at odds with the fitted functions.

Table 3. Statistics and ranks of distribution functions for large section

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.21754	8	4.9794	6	43.841	23
2	Burr	0.22023	9	5.2994	8	47.975	27
3	Burr (4P)	0.20889	5	4.7874	3	46.794	26
4	Cauchy	0.25584	16	11.036	32	31.318	5
5	Chi-Squared	0.24138	11	6.237	11	21.377	2
6	Chi-Squared (2P)	0.27695	31	6.8867	27	39.714	20
7	Erlang	0.26923	29	6.9436	28	33.815	13
8	Erlang (3P)	0.26197	19	6.4184	17	33.524	10
9	Error	0.25443	13	6.3292	13	34.18	16
10	Fatigue Life	0.26456	22	6.7096	23	32.733	8
11	Fatigue Life (3P)	0.25448	14	6.3491	14	34.333	18
12	Gamma	0.26478	23	6.8443	26	33.755	12
13	Gamma (3P)	0.26773	28	6.5688	18	26.85	4
14	Gen. Gamma	0.26138	18	6.5998	19	33.385	9
15	Gen. Gamma (4P)	0.25873	17	6.404	16	34.105	15
16	Gumbel Min	0.20251	1	4.6156	2	46.507	25
17	Kumaraswamy	0.20293	2	4.5911	1	42.932	21
18	Log-Gamma	0.26576	25	6.7841	24	32.46	6
19	Log-Logistic	0.26528	24	6.6069	20	37.081	19
20	Log-Logistic (3P)	0.21021	6	5.6965	10	48.387	28
21	Logistic	0.27007	30	7.1974	30	92.302	32
22	Lognormal	0.26392	21	6.701	22	32.711	7
23	Lognormal (3P)	0.25582	15	6.393	15	34.028	14
24	Nakagami	0.26375	20	7.0364	29	50.64	30
25	Normal	0.25275	12	6.2615	12	34.259	17
26	Pearson 5	0.26697	27	6.8315	25	24.976	3
27	Pearson 5 (3P)	0.26678	26	6.6461	21	33.557	11
28	Pert	0.22378	10	5.338	9	55.147	31
29	Power Function	0.33431	32	9.0577	31	3.473	1
30	Triangular	0.20564	3	4.9889	7	45.901	24
31	Weibull	0.21481	7	4.9002	5	48.479	29
32	Weibull (3P)	0.20825	4	4.7932	4	43.445	22

Table 4. Statistics and ranks of distribution functions for small section

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.16695	6	2.7185	2	29.952	9
2	Burr	0.18907	8	3.4526	5	48.294	19
3	Burr (4P)	0.24652	31	6.951	29	61.609	21
4	Cauchy	0.22896	28	9.9515	32	90.588	30
5	Chi-Squared	0.22916	29	5.9372	23	14.388	2
6	Chi-Squared (2P)	0.20316	21	5.5468	21	45.063	17
7	Erlang	0.22068	27	5.7262	22	15.592	3
8	Erlang (3P)	0.20386	22	4.9525	15	64.749	22
9	Error	0.2417	30	5.3593	17	31.334	11
10	Fatigue Life	0.20113	17	6.5355	26	106.24	31
11	Fatigue Life (3P)	0.19633	10	4.2818	9	66.017	25
12	Gamma	0.21317	23	5.4334	18	13.718	1
13	Gamma (3P)	0.20174	18	4.827	14	67.938	27
14	Gen. Gamma	0.2018	19	5.5434	20	30.18	10
15	Gen. Gamma (4P)	0.20093	15	4.191	8	65.302	23
16	Gumbel Min	0.1582	2	2.585	1	38.057	14
17	Kumaraswamy	0.16621	5	2.7641	4	28.513	8
18	Log-Gamma	0.20189	20	6.6463	28	33.235	12
19	Log-Logistic	0.19429	9	6.5374	27	41.547	15
20	Log-Logistic (3P)	0.17756	7	3.8033	6	69.704	28
21	Logistic	0.21811	26	4.4279	12	20.992	6
22	Lognormal	0.19865	13	6.3035	25	114.75	32
23	Lognormal (3P)	0.20098	16	4.535	13	66.377	26
24	Nakagami	0.21773	25	5.1817	16	20.133	5
25	Normal	0.19794	11	4.3282	10	65.998	24
26	Pearson 5	0.19835	12	7.3541	30	42.835	16
27	Pearson 5 (3P)	0.21628	24	5.4897	19	18.496	4
28	Pert	0.16006	3	4.0332	7	37.88	13
29	Power Function	0.25049	32	8.1503	31	83.49	29
30	Triangular	0.1998	14	6.2101	24	49.583	20
31	Weibull	0.15709	1	4.3586	11	47.57	18
32	Weibull (3P)	0.16521	4	2.7592	3	28.51	7

Table 5. Rank of distribution functions in both sections by Kolmogorov-Smirnov test

Distribution Function	Sum of Scores for Rank	Rank	Distribution Function	Sum of Scores for Rank	Rank
Beta	14	7	Kumaraswamy	7	2
Burr	17	8	Log-Gamma	45	24
Burr (4P)	36	17	Log-Logistic	33	15
Cauchy	44	23	Log-Logistic (3P)	13	5
Chi-Squared	30	12	Logistic	56	29
Chi-Squared (2P)	52	29	Lognormal	34	16
Erlang	56	30	Lognormal (3P)	31	13
Erlang (3P)	41	21	Nakagami	45	24
Error	43	22	Normal	23	10
Fatigue Life	39	19	Pearson 5	39	19
Fatigue Life (3P)	24	11	Pearson 5 (3P)	50	28
Gamma	46	26	Pert	13	5
Gamma (3P)	46	26	Power Function	64	32
Gen. Gamma	37	18	Triangular	17	8
Gen. Gamma (4P)	32	14	Weibull	8	3
Gumbel Min	3	1	Weibull (3P)	8	3

Table 6. Rank of distribution functions in both sections by Anderson-Darling test

Distribution Function	Sum of Scores for Rank	Rank
Beta	8	4
Burr	13	5
Burr (4P)	32	15
Cauchy	64	32
Chi-Squared	34	18
Chi-Squared (2P)	48	26
Erlang	55	29
Erlang (3P)	32	15
Error	30	13
Fatigue Life	49	27
Fatigue Life (3P)	23	10
Gamma	44	22
Gamma (3P)	32	15
Gen. Gamma	39	19
Gen. Gamma (4P)	24	11
Gumbel Min	3	1
Kumaraswamy	5	2
Log-Gamma	52	28
Log-Logistic	47	24
Log-Logistic (3P)	16	6
Logistic	42	21
Lognormal	47	24
Lognormal (3P)	28	12
Nakagami	45	23
Normal	22	9
Pearson 5	55	29
Pearson 5 (3P)	40	20
Pert	16	6
Power Function	62	31
Triangular	31	14
Weibull	16	6
Weibull (3P)	7	3

Table 7. Rank of distribution functions in both sections by Chi-squared test

Distribution Function	Sum of Scores for Rank	Rank
Beta	32	13
Burr	46	29
Burr (4P)	47	30
Cauchy	35	16
Chi-Squared	4	1
Chi-Squared (2P)	37	18
Erlang	16	4
Erlang (3P)	32	13
Error	27	8
Fatigue Life	39	21
Fatigue Life (3P)	43	26
Gamma	13	2
Gamma (3P)	31	12
Gen. Gamma	19	6
Gen. Gamma (4P)	38	19
Gumbel Min	39	21
Kumaraswamy	29	9
Log-Gamma	18	5
Log-Logistic	34	15
Log-Logistic (3P)	56	32
Logistic	38	19
Lognormal	39	21
Lognormal (3P)	40	24
Nakagami	35	16
Normal	41	25
Pearson 5	19	6
Pearson 5 (3P)	15	3
Pert	44	27
Power Function	30	11
Triangular	44	27
Weibull	47	30
Weibull (3P)	29	9

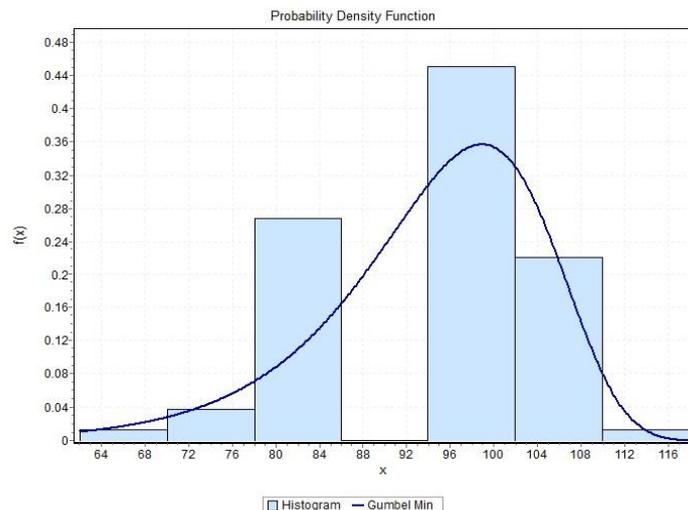


Fig. 3. Fitted diagram of gumbel min for large section

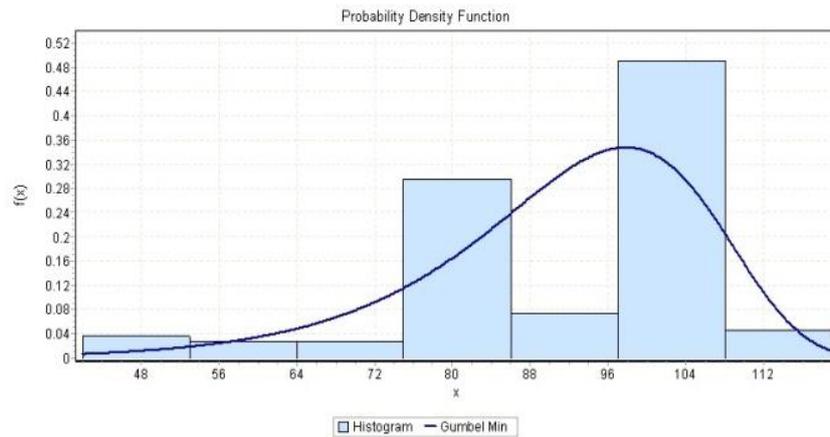


Fig. 4. Fitted diagram of gumbel min for small section

To examine whether these almost empty intervals are just created randomly or if there are any reasons for this phenomena, the observed data was investigated to find whether any order can be discovered in the data.

In Table 1 for the large section, it was observed that in the first part of the tunnel (from Frame No. 1 to Frame No. 25), the distances of steel sets are generally under 85 cm and in the second part of the tunnel (from Frame No. 26 to Frame No. 82) the distance of the frames are generally more than 95 cm.

It can also be observed from Table 2 that in the first part of the tunnel in the small section, (from Frame No. 1 to Frame No. 26), the distances of steel sets are generally under 90 cm, in the second part, (from Frame No. 27 to Frame No. 79) these distances are generally more than 95 cm and in the third part (from Frame No. 80 to Frame No. 108) the distances are generally under 90 cm.

Therefore, it can be concluded that the empty intervals in the data are not random and there should be a reason behind the lack of data in the empty intervals. The reason may be a change in the working condition including:

1. The condition of the ground has changed along the tunnel and the spacing of steel sets has altered accordingly.
2. The different construction teams have done their job with different levels of risk and decisions.
3. Different judgment of supervisors from the safety point of view may lead to variations in distance between steel sets.

Regardless of which reason is correct, one

can divide the observed (measured) data into two categories of A (low spacing of steel sets) and B (high spacing of steel sets). With such analyses, the difference between Chi-squared method and other methods can be justified. Based on the above analysis, the data for both large and small sections of the tunnel (Tables 1 and 2) was divided into two types of A and B and the distribution function for each type of data was obtained separately by Easy Fit 5.5 Software. From 65 distribution functions available in the software, 32 best fitted functions were selected. The list of these distribution functions, the statistic values, and the rank of each function by three methods of Goodness of Fit are shown in Tables 8-11. Then, the summation of scores of each method was obtained and finally, the total score was used to determine the best function.

A summary of these values is shown in Tables 12-14. As these tables show, the Wakeby Distribution Function has rank (1) for all three methods. The fitted Wakeby Distribution Function for all cases are shown in Figures 5-8. From these Figures, it can be seen that Wakeby function, which is an advanced distribution function with 5 parameters, is the appropriate distribution function for the random variable of spacing of the steel sets.

From the above discussions, it can also be concluded that the significant difference between the evaluation results performed by different tests for Goodness of Fit indicates an irregularity in the data. In order to find the source of this irregularity and possible correction of data, more examinations and understanding is needed.

Table 8. Statistics and rank of distribution function for large section (type A)

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.59703	32	12.444	32	1.528	16
2	Cauchy	0.22061	2	1.0596	2	1.7305	24
3	Chi-Squared	0.40103	30	4.6	29	1.6645	22
4	Chi-Squared (2P)	0.38767	28	3.8098	26	5.6276	27
5	Erlang	0.36026	25	3.4525	24	1.515	12
6	Erlang (3P)	0.33775	19	3.2912	15	1.5122	11
7	Error	0.28749	5	2.446	4	5.963	30
8	Fatigue Life	0.33902	21	3.3452	20	1.54	17
9	Fatigue Life (3P)	0.32065	10	3.2015	8	1.4999	5
10	Frechet	0.43752	31	4.6834	30	8.5242	31
11	Frechet (3P)	0.36591	26	3.8488	27	4.6539	25
12	Gamma	0.33308	17	3.2892	13	1.5239	14
13	Gamma (3P)	0.33593	18	3.2899	14	1.5167	13
14	Gumbel Max	0.38869	29	4.7556	31	5.6253	26
15	Gumbel Min	0.31483	8	3.8652	28	1.7102	23
16	Hypersecant	0.30828	6	2.7547	6	1.511	9
17	Inv. Gaussian	0.32215	11	3.2914	16	1.544	19
18	Inv. Gaussian (3P)	0.32035	9	3.2035	9	1.5013	6
19	Laplace	0.28749	4	2.446	3	5.963	29
20	Log-Gamma	0.34189	23	3.3731	21	1.5551	20
21	Log-Logistic	0.38319	27	3.6705	25	5.6739	28
22	Log-Logistic (3P)	0.24813	3	2.5858	5	9.0273	32
23	Logistic	0.31478	7	2.926	7	1.5063	7
24	Lognormal	0.33854	20	3.3406	19	1.5401	18
25	Lognormal (3P)	0.3273	13	3.2419	11	1.5111	10
26	Nakagami	0.32784	14	3.2465	12	1.5099	8
27	Normal	0.32249	12	3.213	10	1.4977	4
28	Pearson 5	0.34459	24	3.4031	22	1.5622	21
29	Pearson 5 (3P)	0.34036	22	3.3224	17	1.5255	15
30	Wakeby	0.19716	1	0.74762	1	1.3338	2
31	Weibull	0.32828	15	3.4233	23	1.3069	1
32	Weibull (3P)	0.32884	16	3.3289	18	1.4717	3

Table 9. Statistics and rank of distribution function for large section (type B)

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.2074	4	4.8111	8	34.266	8
2	Cauchy	0.15086	2	0.81021	1	0.95324	1
3	Chi-Squared	0.36983	32	12.251	31	130.31	32
4	Chi-Squared (2P)	0.2955	28	8.8653	28	65.408	30
5	Erlang	0.24217	19	5.6122	18	59.634	17
6	Erlang (3P)	0.23443	17	5.8147	19	59.735	21
7	Error	0.23232	14	3.1364	5	5.207	3
8	Fatigue Life	0.24751	23	6.0574	23	59.805	24
9	Fatigue Life (3P)	0.22843	10	5.301	12	59.59	15
10	Frechet	0.35595	31	11.288	30	62.76	28
11	Frechet (3P)	0.3208	30	9.9406	29	94.606	31
12	Gamma	0.2345	18	5.4993	17	59.641	18
13	Gamma (3P)	0.24592	21	5.8719	20	59.689	20
14	Gumbel Max	0.2973	29	13.146	32	38.731	11
15	Gumbel Min	0.22871	12	4.9164	9	20.844	6
16	Hypersecant	0.22862	11	3.8525	6	22.108	7
17	Inv. Gaussian	0.22696	6	5.4727	16	59.686	19
18	Inv. Gaussian (3P)	0.23236	16	5.3572	14	59.59	14
19	Laplace	0.23232	15	3.1364	4	5.207	4
20	Log-Gamma	0.24906	24	6.1295	24	59.851	26
21	Log-Logistic	0.28051	27	7.3507	27	62.917	29
22	Log-Logistic (3P)	0.17945	3	2.1616	3	15.542	5
23	Logistic	0.22776	7	4.3658	7	59.839	25
24	Lognormal	0.24653	22	6.0286	22	59.803	23
25	Lognormal (3P)	0.21409	5	5.2976	11	34.346	10
26	Nakagami	0.22961	13	5.0416	10	59.572	13
27	Normal	0.22818	9	5.3229	13	59.594	16
28	Pearson 5	0.2536	25	6.3294	25	60.963	27
29	Pearson 5 (3P)	0.2452	20	5.9466	21	59.747	22
30	Wakeby	0.14792	1	0.98525	2	1.2573	2
31	Weibull	0.25526	26	6.4967	26	46.281	12
32	Weibull (3P)	0.22812	8	5.4545	15	34.276	9

Table 10. Statistics and rank of distribution function for small section (type A)

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.21199	6	2.8112	8	11.887	10
2	Cauchy	0.07607	2	0.53788	2	2.1722	6
3	Chi-Squared	0.29894	29	5.1518	26	31.632	17
4	Chi-Squared (2P)	0.28843	25	4.52	21	48.678	21
5	Erlang	0.28241	22	4.5923	23	49.836	22
6	Erlang (3P)	0.26405	20	3.8826	18	51.219	28
7	Error	0.24156	8	2.7296	7	24.41	15
8	Fatigue Life	0.29247	26	5.0929	25	0.3688	2
9	Fatigue Life (3P)	0.24479	11	3.4882	11	50.842	24
10	Frechet	0.35623	32	8.4124	32	86.158	32
11	Frechet (3P)	0.28254	23	5.6972	30	0.66478	4
12	Gamma	0.27184	21	4.4063	20	49.965	23
13	Gamma (3P)	0.25932	19	3.8438	16	51.505	29
14	Gumbel Max	0.31812	31	8.4001	31	12.216	13
15	Gumbel Min	0.18491	4	2.6527	5	9.4236	8
16	Hypersecant	0.242	10	2.8445	9	38.078	18
17	Inv. Gaussian	0.24902	13	4.5891	22	52.185	31
18	Inv. Gaussian (3P)	0.25228	17	3.5832	13	50.933	26
19	Laplace	0.2417	9	2.6242	4	23.138	14
20	Log-Gamma	0.29409	27	5.2505	28	0.36155	1
21	Log-Logistic	0.29863	28	5.185	27	10.664	9
22	Log-Logistic (3P)	0.17615	3	2.4925	3	12.115	12
23	Logistic	0.2453	12	3.0722	10	39.536	20
24	Lognormal	0.28812	24	4.9858	24	2.9995	7
25	Lognormal (3P)	0.25558	18	3.6985	14	50.942	27
26	Nakagami	0.25039	15	3.9165	19	38.496	19
27	Normal	0.24919	14	3.5302	12	50.863	25
28	Pearson 5	0.30139	30	5.6013	29	0.37492	3
29	Pearson 5 (3P)	0.25217	16	3.8459	17	51.776	30
30	Wakeby	0.07214	1	0.26787	1	2.0788	5
31	Weibull	0.2368	7	3.7601	15	29.367	16
32	Weibull (3P)	0.20699	5	2.7058	6	11.924	11

Table 11. Statistics and rank of distribution function for small section (type B)

No	Distribution	Kolmogorov-Smirnov		Anderson-Darling		Chi-Squared	
		Statistic	Rank	Statistic	Rank	Statistic	Rank
1	Beta	0.13295	10	0.58419	9	4.2656	8
2	Cauchy	0.1413	13	1.5249	15	5.3011	14
3	Chi-Squared	0.37863	32	11.79	32	112.97	32
4	Chi-Squared (2P)	0.2387	30	2.5615	28	6.2896	15
5	Erlang	0.16974	18	1.677	21	9.7442	23
6	Erlang (3P)	0.20935	29	2.9752	29	7.6403	16
7	Error	0.18653	26	1.7549	24	8.9962	17
8	Fatigue Life	0.16979	19	1.5503	20	9.4049	21
9	Fatigue Life (3P)	0.13199	9	0.5606	8	4.3182	9
10	Frechet	0.13163	7	0.5478	6	2.9949	6
11	Frechet (3P)	0.12417	3	0.5094	3	2.893	5
12	Gamma	0.17154	20	1.6969	22	14.83	29
13	Gamma (3P)	0.13723	12	0.60896	10	4.2479	7
14	Gumbel Max	0.13352	11	0.70654	11	4.847	13
15	Gumbel Min	0.24705	31	5.5136	31	15.403	31
16	Hypersecant	0.17893	25	1.536	16	14.431	26
17	Inv. Gaussian	0.17646	23	1.7388	23	13.955	24
18	Inv. Gaussian (3P)	0.13187	8	0.55587	7	4.3229	10
19	Laplace	0.18653	27	1.7549	25	8.9962	18
20	Log-Gamma	0.16859	16	1.5482	19	9.4534	22
21	Log-Logistic	0.13	6	0.97782	13	4.7364	11
22	Log-Logistic (3P)	0.11206	2	0.49141	2	2.7435	3
23	Logistic	0.17779	24	1.5461	17	14.555	28
24	Lognormal	0.16964	17	1.5475	18	9.4046	20
25	Lognormal (3P)	0.12971	5	0.53476	5	1.2767	1
26	Nakagami	0.17295	21	1.8473	27	14.453	27
27	Normal	0.17646	22	1.7978	26	14.167	25
28	Pearson 5	0.16734	15	1.4795	14	9.3157	19
29	Pearson 5 (3P)	0.12617	4	0.51629	4	1.2812	2
30	Wakeby	0.10513	1	0.45257	1	2.8143	4
31	Weibull	0.19563	28	4.7543	30	15.08	30
32	Weibull (3P)	0.15192	14	0.77512	12	4.7765	12

Table 12. Rank of distribution functions in all cases by Kolmogorov-Smirnov test

Distribution Function	Sum of Scores for Rank	Rank
Beta	52	9
Cauchy	19	3
Chi-Squared	123	32
Chi-Squared (2P)	111	31
Erlang	84	23
Erlang (3P)	85	24
Error	53	11
Fatigue Life	89	26
Fatigue Life (3P)	40	4
Frechet	101	30
Frechet (3P)	82	21
Gamma	76	19
Gamma (3P)	70	18
Gumbel Max	100	29
Gumbel Min	55	13
Hypersecant	52	9
Inv. Gaussian	53	11
Inv. Gaussian (3P)	50	7
Laplace	55	13
Log-Gamma	90	27
Log-Logistic	88	25
Log-Logistic (3P)	11	2
Logistic	50	7
Lognormal	83	22
Lognormal (3P)	41	5
Nakagami	63	17
Normal	57	15
Pearson 5	94	28
Pearson 5 (3P)	62	16
Wakeby	4	1
Weibull	76	19
Weibull (3P)	43	6

Table 13. Rank of distribution functions in all cases by Anderson-Darling test

Distribution Function	Sum of Scores for Rank	Rank
Beta	57	12
Cauchy	20	3
Chi-Squared	118	32
Chi-Squared (2P)	103	30
Erlang	86	22
Erlang (3P)	81	20
Error	40	7
Fatigue Life	88	23
Fatigue Life (3P)	39	6
Frechet	98	29
Frechet (3P)	89	24
Gamma	72	17
Gamma (3P)	60	14
Gumbel Max	105	31
Gumbel Min	73	18
Hypersecant	37	5
Inv. Gaussian	77	19
Inv. Gaussian (3P)	43	10
Laplace	36	4
Log-Gamma	92	26
Log-Logistic	92	26
Log-Logistic (3P)	13	2
Logistic	41	8
Lognormal	83	21
Lognormal (3P)	41	8
Nakagami	68	16
Normal	61	15
Pearson 5	90	25
Pearson 5 (3P)	59	13
Wakeby	5	1
Weibull	94	28
Weibull (3P)	51	11

Table 14. Rank of distribution functions in all cases by Chi-squared test

Distribution Function	Sum of Scores for Rank	Rank
Beta	42	3
Cauchy	45	4
Chi-Squared	103	32
Chi-Squared (2P)	93	29
Erlang	74	24
Erlang (3P)	76	25
Error	65	13
Fatigue Life	64	12
Fatigue Life (3P)	53	7
Frechet	97	31
Frechet (3P)	65	13
Gamma	84	28
Gamma (3P)	69	19
Gumbel Max	63	11
Gumbel Min	68	17
Hypersecant	60	10
Inv. Gaussian	93	29

Distribution Function	Sum of Scores for Rank	Rank
Inv. Gaussian	93	29
Inv. Gaussian (3P)	56	8
Laplace	65	13
Log-Gamma	69	19
Log-Logistic	77	26
Log-Logistic (3P)	52	6
Logistic	80	27
Lognormal	68	17
Lognormal (3P)	48	5
Nakagami	67	16
Normal	70	22
Pearson 5	70	22
Pearson 5 (3P)	69	19
Wakeby	13	1
Weibull	59	9
Weibull (3P)	35	2

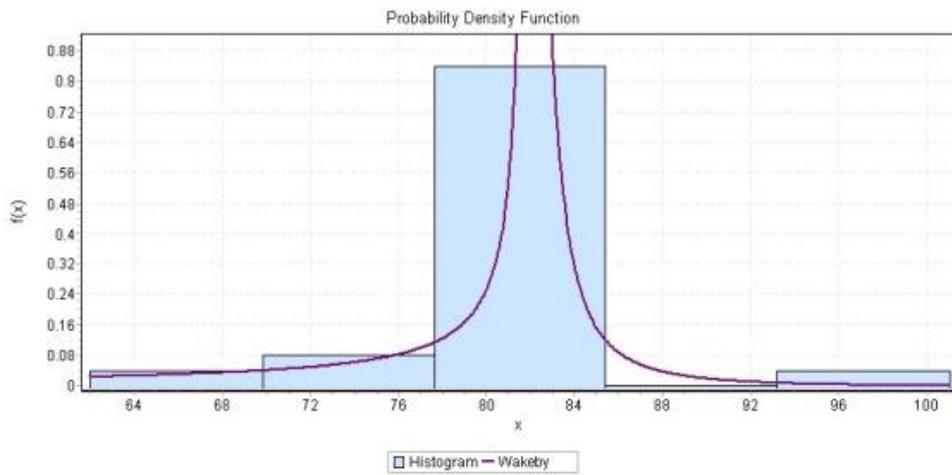


Fig. 5. Fitted diagram of Wakeby for large section (type A)

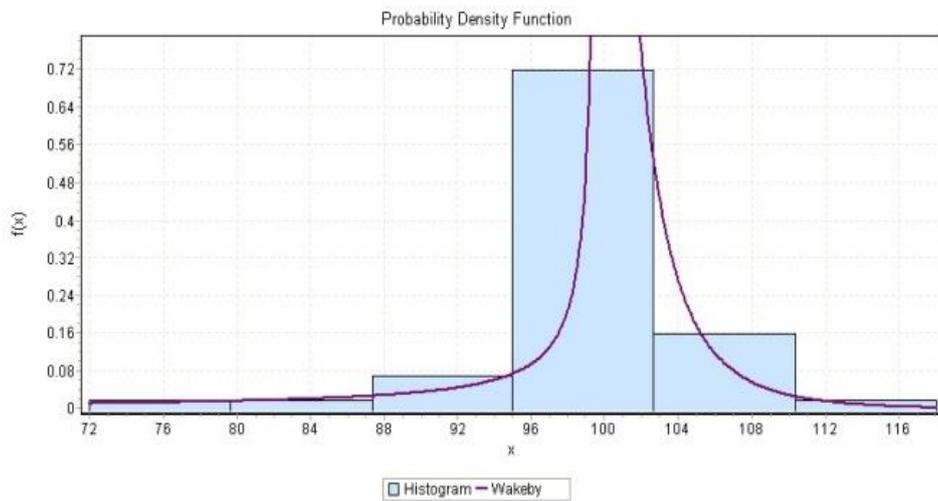


Fig. 6. Fitted diagram of Wakeby for large section (type B)

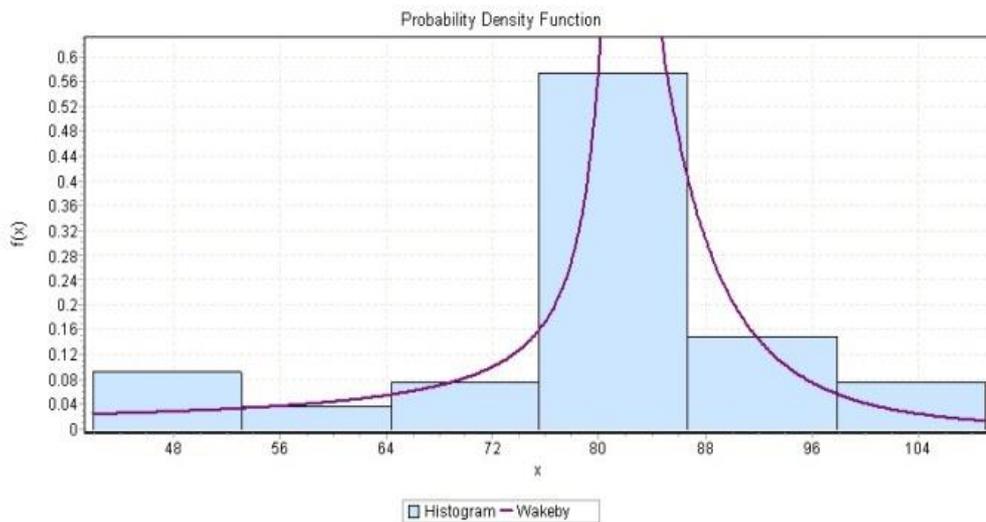


Fig. 7. Fitted diagram of Wakeby for small section (type A)

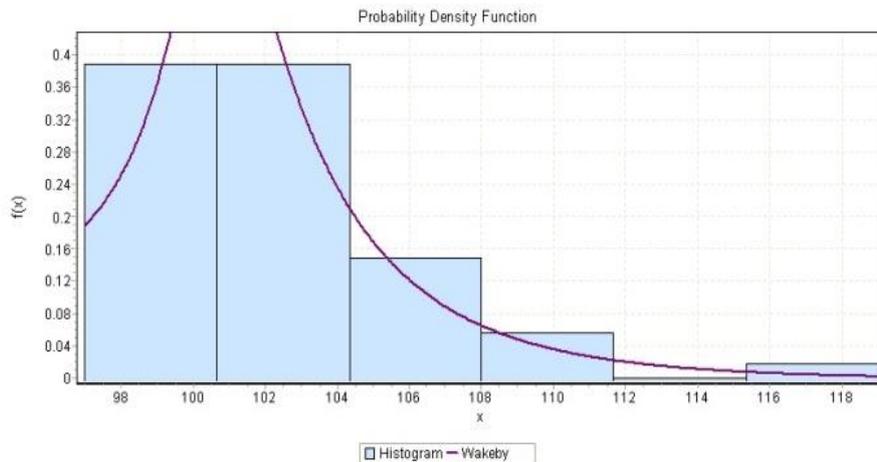


Fig. 8. Fitted diagram of Wakeby for small section (type B)

4. Conclusion

In the present study, the spacing of the steel sets for temporary support of tunnels was considered as a random variable. To determine the distribution function of this random variable, the real data gathered from a tunnel in Iran with two drilling sections was used. The proper distribution function for both sections was evaluated by three methods of Goodness of Fit (Kolmogorov-Smirnov, Anderson-Darling, and Chi-squared). The analysis of data was performed by the Easy Fit Software and the following results were obtained:

- Wakeby Distribution function is the best probability distribution function that can be fitted to the spacing of the steel sets.
- A significant difference between evaluations by different tests for Goodness of Fit indicates an irregular data that requires more examination and understanding in order to lead us to find the source of irregularity and possible corrections.
- The studies in this research show that a change in the tunnel section or a change in working conditions does not affect the proper distribution function. However, it affects the parameters of the PDF. Of course, this result needs to be verified by more case studies.

References

[1] Madani, H. (2005). Tunneling (Volume 4): Design and construction of support systems. ISBN: 964-463-135-8.

[2] Biron, C., Ariogulu, E., & Lucas, J. R. (1983). Design of supports in mines, Wiley, ISBN-10: 0471867268. ISBN-13: 978-463-135-8. 0471867265.

[3] Hjálmarsson, E.H. (2011). Tunnel Support, Use of lattice girders in sedimentary rock, MSc Thesis, University of Iceland.

[4] Carranza-Torres, C., & Diederichs, M. (2009). Mechanical analysis of circular liners with particular reference to composite supports. For example, liners consisting of shotcrete and steel sets. *Tunnelling and Underground Space Technology*, 24, 506–532. DOI: 10.1016/j.tust.2009.02.001

[5] Basaligheh, F. (2003). The advantages of using lattice girder with shotcrete instead of steel sets for temporary support of tunnels (in Persian). *Proceedings of the Sixth International Conference on Civil Engineering*, 5, 265-270.

[6] Basaligheh, F. (2006). Evaluation of two methods for the temporary support of the tunnels from economic point of view (in Persian). *Journal of Science and technology*, 12, 11-17.

[7] Basaligheh, F., & Keyhani, A. (2013). The advantages of using the method of “equivalent section” in design of shotcrete and steel sets for temporary support of tunnels (in Persian). *Proceedings of Seventh National Congress on Civil Engineering*. Shahid Nikbakht Faculty of Engineering, Zahedan-Iran.

[8] Basaligheh, F., & Keyhani, A. (2013). New approaches for modelling of two common methods of temporary support for tunnels. *Proceedings of Tenth Iranian Tunnelling Conference*.

- [9] Wong, L.N.Y., Fang, Q., & Zhang, D. (2013). Mechanical analysis of circular tunnels supported by steel sets embedded in primary linings. *Tunnelling and Underground Space Technology*, 37, 80–88. DOI: 10.1016/j.tust.2013.03.011
- [10] Schwingenschloegl, R., & Lehmann, C. (2009). Swelling rock behaviour in a tunnel: NATM-support vs. Q-support – A comparison. *Tunnelling and Underground Space Technology*, 24, 356–362. DOI: 10.1016/j.tust.2008.08.007
- [11] Nowak, A. a. (2000). *Reliability of structures*. McGraw Hill, International Edition. ISBN: 0-07-116354-9.
- [12] Rezaee-Pajand, H. (2002). *Application of statistics and probability for water reservoirs*. Sokhan-Gostar Press. ISBN: 964-6932-35-5.
- [13] Sanchidrian, J, O. A. (2014). Size distribution functions for rock fragments. *International Journal of Rock Mechanics & Mining Sciences*, 71, 381-394. DOI:10.1016/j.ijrmmms.2014.08.007.
- [14] Milford, R. (1987). Annual maximum wind speeds from parent distribution functions. *Journal of wind Engineering and Industrial Aerodynamics*, 25, pp. 163-178.
- [15] Ghamishon, R., & Malekian, A. (2011). Determination of the most suitable statistical distribution functions for regional floods (Case study of southwest Kerman Province). *Proceedings of Sixth National Watershed Sciences and Engineering Conference and forth national congress on soil erosion and sediment*, Tarbiat-Modares University.
- [16] Meftah-Halghi, M., Zangale, M. E., & Aghili, R. (2012). A Comparison of the most suitable statistical distribution functions for maximum daily flow rate and rainfall (Case study: Hydrometry station of Gonbad Kavos)., *Fifth National Watershed Sciences and Engineering Conference*, Iran. University of Agricultural Sciences of Gorgan. COI: WATERSHED05_224.http://www.civilica.com/Paper-WATERSHED05-WATERSHED05_224.html
- [17] Cochran, W.G. (1952). The χ^2 test of goodness of fit. *Institute of Mathematical Statistics*, Vol 23, No. 3, pp. 315-345. <http://www.jstor.org/stable/2236678>
- [18] Anderson, T, A. (1952). Asymptotic theory of certain goodness of fit criteria based on stochastic processes. *Ann. Math. Statist*, 23, pp. 193-212.
- [19] Darbandi, S., Mahmudi, S., & Ebrahimi, S., & Shoeybi-Nobarian, M. R. (2012). Introduction and application of Anderson-Darling in river engineering of East Azerbaijan Province. *Fifth National Conference on Watersheds and management of soil and water resources*.