Available online at www.eccomasproceedia.org
Eccomas Proceedia COMPDYN (2023) 3636-3646

COMPDYN 2023
9th ECCOMAS Thematic Conference on
Computational Methods in Structural Dynamics and Earthquake Engineering
M. Papadrakakis, M. Fragiadakis (eds.)
Athens, Greece, 12-14 June 2023

KURTOSIS AND SKEWNESS BASED COMPARISON OF HERMITE POLYNOMIAL AND JOHNSON TRANSFORMATIONS

Ayse S. Giz and Ata Mugan

Istanbul Technical University, Department of Mechanical Engineering
34437 Beyoglu, Istanbul, Turkey

{giza, mugan}@itu.edu.tr

Abstract

In fatigue analyses and structural health monitoring applications, frequency domain theoretical fatigue life estimation models work best on Gaussian data, while many natural occurrences that influence the stress history have non-Gaussian properties. In order to handle realistic stress data, various transformations from non-Gaussian to Gaussian distributions are used. In this study, two transformation methods namely moment based Hermite polynomial model and Johnson transformation model were examined that have both a non-Gaussian to Gaussian transformation and a Gaussian to non-Gaussian transformation. These established transformation methods were compared for their distortion with reference to increasingly leptokurtic behavior and skewness.
Normally distributed Gaussian data were distorted to a non-Gaussian form and then re-stored back to Gaussian form using parameters estimated from the non-Gaussian data were compared according to their kurtosis and skewness values. It was found that the restored da-ta contains significant deviation from Gaussian form the deviation depending on the distortion in the initial transformation and the deviations are more prominent in the Hermite Polynomial Transformation. These deviations in turn affect fatigue life estimations.

Keywords: Non-Gaussian, Gaussian Transformation, Hermite, Johnson Transformation Method, Fatigue Analysis.

ISSN:2623-3347 © 2023 The Authors. Published by Eccomas Proceedia.
Peer-review under responsibility of the organizing committee of COMPDYN 2023.
doi: 10.7712/120123.10667.20286

1 INTRODUCTION

The analysis of service life with respect to fatigue failure gains more importance as structures are designed, primarily for economic reasons, to have a finite life. Obtaining a realistic estimation of fatigue lifetime of mechanical systems is of crucial importance. Various estimation methods have been developed based on stress history. These methods work in time domain or frequency domain. Frequency domain methods based on fast Fourier transform are less computationally intensive and are preferred for long data streams. They have been developed [1] and found to give results consistent with more precise methods such as cycle counting for stress distributions in Gaussian form. Normally distributed, also known as Gaussian random data is symmetric, with a skewness of zero and has kurtosis of three. These methods have limitations on them in regards to the sampling rate [2] and Gaussianity of the data. On the other hand, many natural phenomena like wind pressure or wave pressure, which are important for wind turbines and off-shore structures have non-Gaussian properties.

To treat non-Gaussian data in the frequency domain it is commonly converted to Gaussian form by a transformation technique which preserves time domain properties such as peak count and upcrossing rate. The fatigue lifetime is calculated for the Gaussian data for which analytical formulae are available [3–5]. The estimated lifetime is corrected using a conversion factor which depends on the kurtosis and skewness of the non-Gaussian data.

Various methods have been developed to transform non-Gaussian load histories to Gaussian data. Most prominent of these are based on Hermite Polynomials [5], power law [6], series expansions, logarithms and hyperbolic functions [7], and empirical methods as well [8,9]. Most of these are one-way transformations [10], turning the non-Gaussian to Gaussian or vice versa, and few are transformations to both transform the data into Gaussian or into non-Gaussian. The two most prominent of which, the Hermite polynomial based Winterstein's transformation [1,11] and Johnson transformation will be studied here.[12]

In life estimation studies using simulations, and initial Gaussian random is transformed into a non-Gaussian data for analysis. The usual procedure is to start with Gaussian data, convert it to non-Gaussian data form using the inverse of the transform then attempt to go back to Gaussian data by estimating the transform parameters.[10] Obviously if the inverse transform is performed using the exact same parameters the original data would be recovered exactly. However, since the underlying Gaussian would not be available in real cases, the parameters for the restoring transformation are estimated based on kurtosis and skewness of this created non-Gaussian data or its percentile rankings.

Liu and group have developed a modified version of the Hermite polynomial method [13] and compared it with original Hermite method and the Johnson Transformation method [14–16] (with parameters estimated by the moments) in life estimation using simulated and wind data. They found that both the Johnson transformation and the modified Hermite polynomial methods gave better results than the original Hermite polynomial method.

Since the correction factor is based on the assumption that the transformed function is completely Gaussian, the level of distortion between transformations can significantly affect the accuracy measure in fatigue calculations. When the transformation affects the subsequent calculations and assumptions dramatically, it becomes difficult to make accurate fatigue life evaluations.

Here, we compare the Hermite polynomial method and Johnson transformation on how they perform with different distorted skewness and kurtosis values. Random Gaussian data is distorted to a non-Gaussian form of various kurtosis and skewness values using both the Hermite and the Johnson methods. The distorted data are then brought back to the Gaussian form. Parameters for the back transform are estimated using the formulae given in [10] for the Hermite method and by Slifker and Shapiro for the Johnson method in [12]. The kurtosis and skewness of the final form of the data is discussed and accuracy of fatigue calculations of transformed data is compared with that of the initial load, using Dirlik's method [1,17]. This is done to obtain a more realistic assessment of both transformation methods' performance in transforming non-Gaussian data to Gaussian form.

2 METHOD

In order to compare transformation methods, the random data was simulated with a sampling rate of 500 samples per second and a frequency range between 0 to 60Hz. Three samples were prepared and results were averaged for statistical significance. All three had a zero mean, and a kurtosis of three and skewness of zero as per Gaussian standards. For both methods the skewness between -2.5 and 2.5 with 50 intervals. Softening cases were analyzed, kurtosis was scanned between 4 and 18, in 140 intervals.

Non-Gaussian distorted data were obtained by following the Johnson and Hermite polynomial transformations described above. Lifetime estimation was performed by the Dirlik method. The average of the three datasets results is quoted and plotted in the figures.

2.1 Johnson transformation method

N.L. Johnson in 1949 suggested three families of transformations to transform non-Gaussian data, which can be unbounded, bounded on one side or bounded on both sides, to Gaussian form. All three families involve a shift and scale operation before and after the transform by a monotonically increasing function. [7]

To obtain a Gaussian X_2, a non-Gaussian Z is first shifted and scaled to form an intermediate variable:

$$u = (Z - \xi)/\lambda \tag{1}$$

Then the transformation function $f(u)$ and finally another scaling and shifting is applied.

$$X_2 = \gamma + \delta f(u) \tag{2}$$

The stress distribution is expected to be unbounded. So, for our purposes the unbounded distribution, with the transformation function $f_U(u) = \sinh^{-1}(u)$ is the more relevant.

Thus, the Johnson transform for unbounded Z is:

$$X_2 = \gamma + \delta \sinh^{-1}((\mathrm{Z} - \xi)/\lambda) \tag{3}$$

To estimate Johnson's parameters one can follow three strategies, namely; percentage and moment fitting and fitting by non-linear least squares (NLLS).

Slifker and Shapiro [12] developed an estimation method based on percentage fitting. Their method takes a real x_p and take four percentages of a normal distribution corresponding to; $\zeta = [3x_p, x_p, -x_p, -3x_p]$. Letting these percentages to be $[Z_{3xp}, Z_{xp}, Z_{-xp}, Z_{-3xp}]$ they defined the quantities:

$$p = Z_{xp} - Z_{-xp},\ m = Z_{3xp} - Z_{xp},\ n = Z_{-xp} - Z_{-3xp} \text{ and } d = mn/p^2. \tag{4}$$

The parameter estimates for Johnson's U distribution are:

$$\delta_e = \frac{2x_p}{\operatorname{arccosh}\left(\frac{1}{2}\left(\frac{m}{p}+\frac{n}{p}\right)\right)}$$
$$\gamma_e = \delta_e \operatorname{arcsinh}\left(\frac{\frac{n}{p}-\frac{m}{p}}{2\sqrt{d-1}}\right) \tag{5}$$
$$\lambda_e = \frac{2p\sqrt{d-1}}{\left(\frac{m}{p}+\frac{n}{p}-2\right)\left(\frac{m}{p}+\frac{n}{p}+2\right)^{1/2}}$$

$$\xi_e = \frac{(x_{zp}+x_{-zp})}{2} + \frac{p\left(\frac{n}{p}-\frac{m}{p}\right)}{2\left(\frac{m}{p}+\frac{n}{p}-2\right)}$$

Simulated non-Gaussian data can be generated from a random Gaussian data X by applying the inverse of the transformation. For the unbounded case the inverse transform is:

$$\mathrm{Z} = \xi + \lambda \sinh((\mathrm{X}-\gamma)/\delta) \tag{6}$$

This non-Gaussian data is then transformed to Gaussian form via the Johnson transformation using the estimated parameters.

$$\mathrm{X}_2 = \gamma_e + \delta_e \sinh^{-1}((\mathrm{Z}-\xi_e)/\lambda_e) \tag{7}$$

2.2 Hermite Polynomial Method

For Hermite Polynomial method, the formulations developed by Winterstein for softening processes [5] were used. To get the distorted non-Gaussian version (Z) from the initial random generated Gaussian (X) the below formulae were used, as it is stated in [10], with γ_3 and γ_4 are the skewness and kurtosis, μ is the mean and σ^2 is the variance.

$$\begin{gathered}
Z = \mu_z + \sigma_z K\left[X + \widetilde{h_3}(X^2-1) + \widetilde{h_4}(X^3-3X)\right] \\
\widetilde{h_4} = \widetilde{h_{40}}\left[1 - \frac{1.43\gamma_3^2}{\gamma_4-3}\right]^{1-0.1\gamma_4^{0.8}} \\
\widetilde{h_3} = \frac{\gamma_3}{6}\left[\frac{1-0.15|\gamma_3|+0.3\gamma_3^2}{1+0.2(\gamma_4-3)}\right] \\
K = \frac{1}{\sqrt{1+2\widetilde{h_3}^2+6\widetilde{h_4}^2}} \\
\widetilde{h_{40}} = \frac{[1+1.25(\gamma_4-3)]^{\frac{1}{3}}-1}{10}
\end{gathered} \tag{8}$$

And to achieve the underlying Gaussian (X_2) of the non-Gaussian (Z) the inverse of the previous expression is used, with the following formula for the transformation;

$$\begin{gathered}
X_2 = \left[\sqrt{\xi^2(Z)+c} + \xi(Z)\right]^{\frac{1}{3}} - \left[\sqrt{\xi^2(Z)+c} - \xi(Z)\right]^{\frac{1}{3}} - a \\
\xi(Z) = 1.5b\left(a + \frac{Z-\mu_z}{K\sigma_z}\right) - a^3 \\
a = \frac{\widetilde{h_3}}{3\widetilde{h_4}} \\
b = \frac{1}{3\widetilde{h_4}} \\
c = (b-1-a^2)^3
\end{gathered} \tag{9}$$

2.3 Fatigue Life Calculation

Properties of s355j10 steel were used in fatigue life analyses. The Gaussian data before and after the transformation procedures were analyzed with Dirlik's fatigue damage estimation method, which was proposed in 1985 and is considered as one of the most popular method among frequency domain estimation methods for wide band Gaussian data [1]. The fatigue life is computed from the damage, with T=1/D and the damage formula as;

$$D = K^{-1}\upsilon_p m_0^{\frac{k}{2}}\left[G_1 Q_k \Gamma(1+k) + \left(\sqrt{2}\right)^k \Gamma\left(1+\frac{k}{2}\right)(G_2|R|^k + G_3)\right] \tag{10}$$

The percent difference in estimated life was calculated with 100*(T_2-T_0)/T_0. All three results for each 7000 points in plot, was averaged.

3 NUMERICAL RESULTS

3.1 Numerical results using the Johnson transformation method

The figure 1 shows the kurtosis (in color code) of data distorted to a wide range of skewness values from -2.5 to + 2.5 and kurtosis values from 3 to 18. As can be seen from the color code the kurtosis values of the restored data were between 2.8 and 2.9.

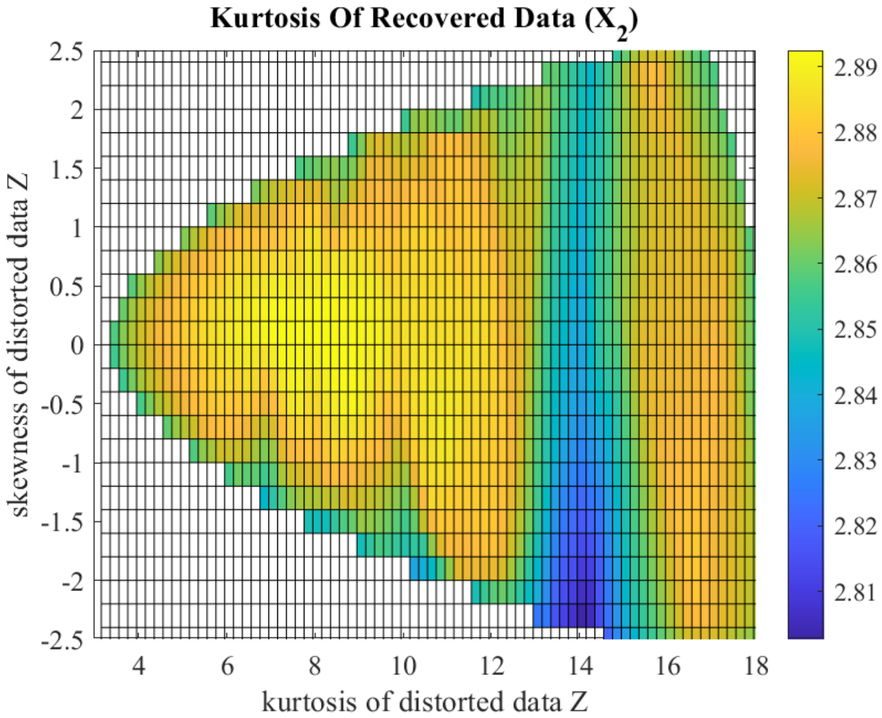

Figure 1 Kurtosis of the "Gaussian" data after distortion and transformation via Johnson transformation method, plotted against the skewness and kurtosis values the data were distorted to

The skewness results shown in Figure 2 were very close to the Gaussian value of zero; that is, they were between 0.007 and 0.028 for the average, and within the dataset never got values above 0.036 or below -0.007

Distorting the data to negative skewness was slightly less robust, resulting in relatively higher skewness values for the final Gaussian distribution. Whereas for positive skewness distortions in the non-Gaussian transformation resulted in very close to zero values, all below 0.02 across all kurtosis values. The resulting skewness was only higher in cases with distortions to extreme kurtosis values and negative skewness.

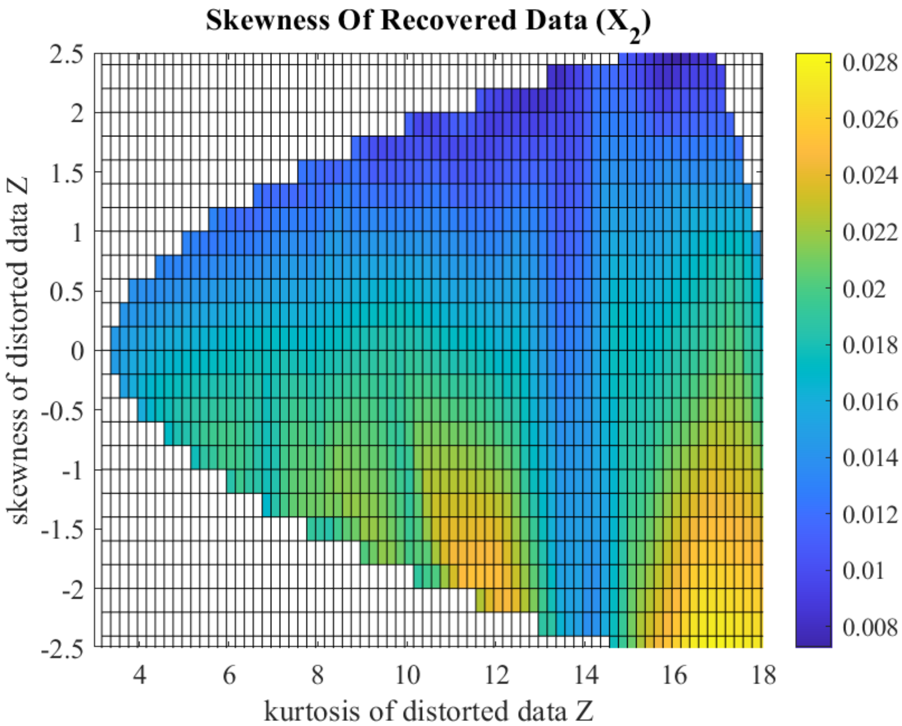

Figure 2 Skewness of the Gaussian data after distortion and transformation via Johnson transformation method, plotted against the skewness and kurtosis values the data were distorted to

The fatigue life estimation values after the process were around 28% lower compared with that of the before transformation, this can be considered because of the decrease in kurtosis, as a result of around 0.2 decrease in the kurtosis values, since it made the final loading slightly smoother with less extremities.

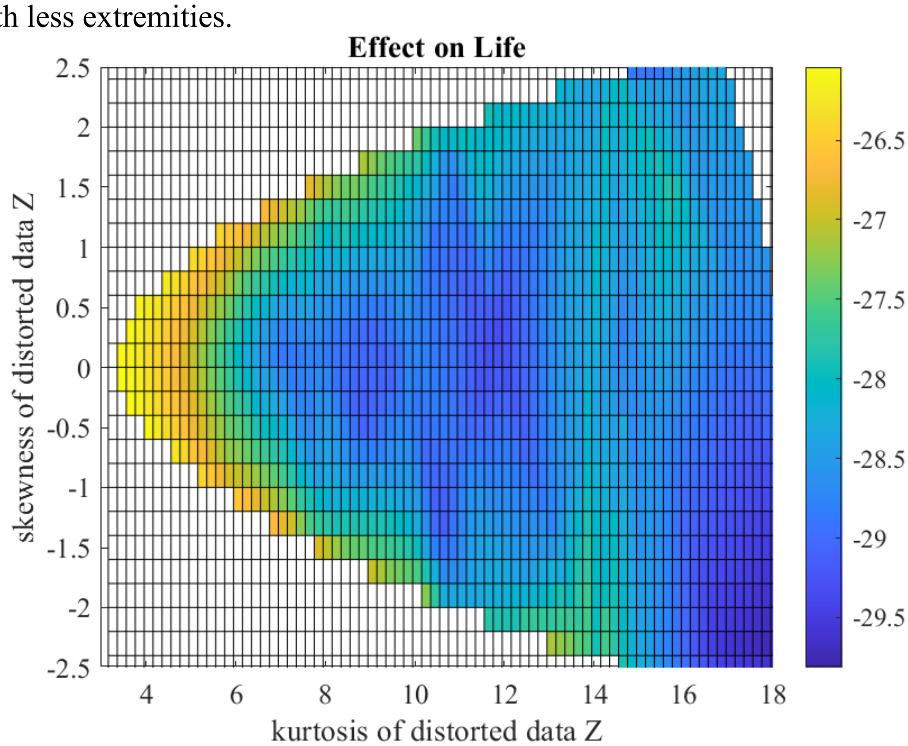

Figure 3 Percent change in estimated life, between the original data and the restored data, plotted against the skewness and kurtosis values of the distorted data

The difference in fatigue estimation results here depend mostly on the kurtosis change, as the graph follows a similar pattern and does not seem to be affected by the same trends visible in the skewness graph. This is also because the change in skewness before and after the process was minimal, especially compared the kurtosis.

3.2 Numerical results using Hermite transformation method

Figure 4 shows the kurtosis of the data restored by the Hermite polynomial method. It is seen that the kurtosis values range from less than 3 to 3.5 for most of the parameter space but varied on the edges of the applicable zone with negative skewness values, reaching up to 6.973 for the average and 12.251 in one of the datasets.

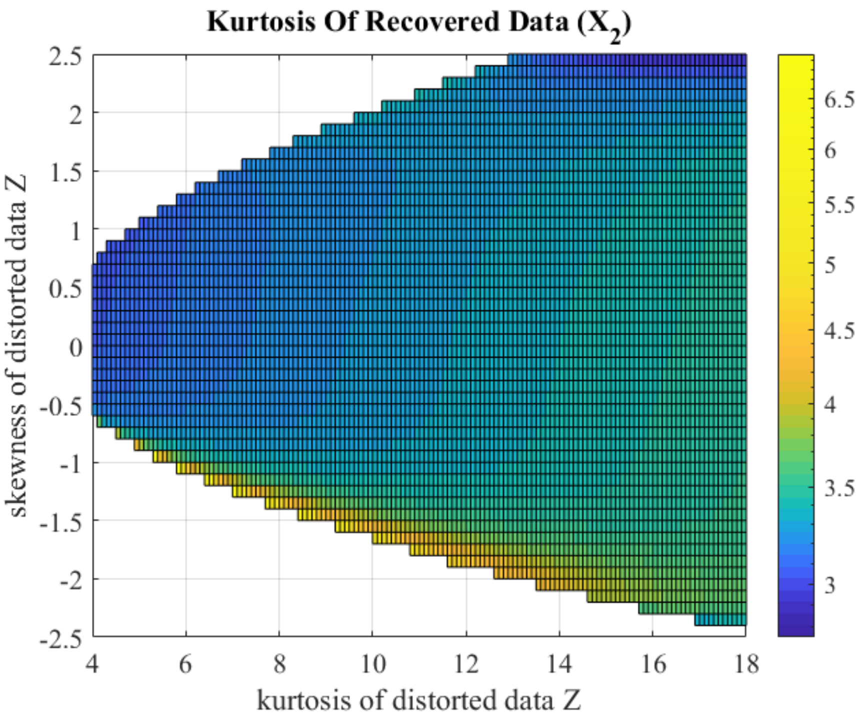

Figure 4 Kurtosis of the Gaussian data after distortion and transformation via Hermite transformation method, plotted against the skewness and kurtosis values the data were distorted to

Though the edges showed great change in kurtosis, and not enough recovery from the distortion procedure, the rest of the data was less varied. There was an increase in final kurtosis values with increasing distorted kurtosis values. Smaller kurtosis values were found as the distortion in skewness increases.

The data was mostly higher than 3, only going below down 2.756 in the cases of extreme kurtosis and highest skewness values. In individual datasets, these points went down to 1.958.

For almost zero skewness distortion, the resulting kurtosis slowly increased from 3.081 to 3.504 as the distorted kurtosis increased.

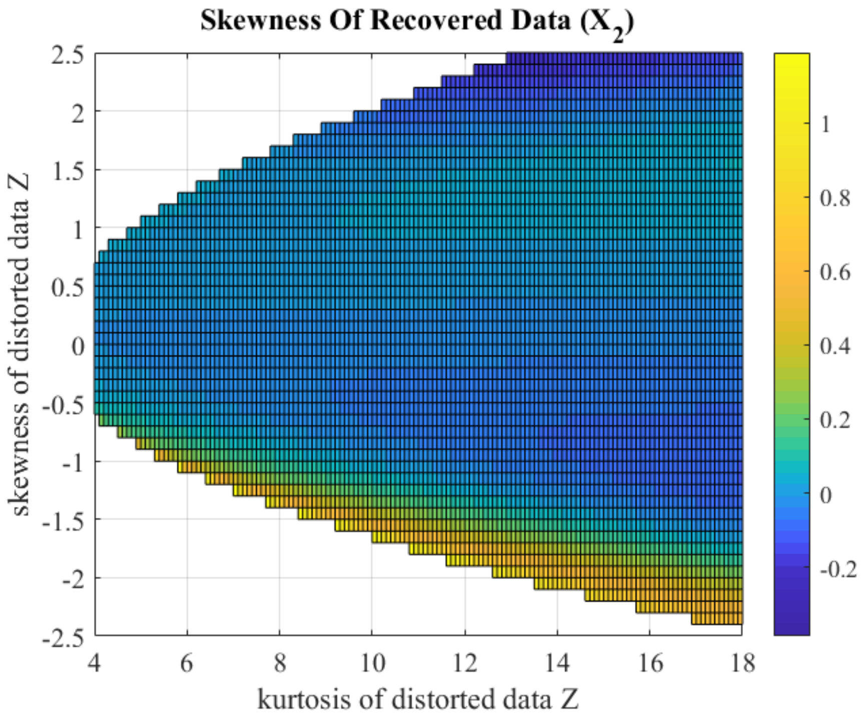

Figure 5 Skewness of the Gaussian data after distortion and transformation via Hermite transformation method, plotted against the skewness and kurtosis values the data were distorted to

Less monotonic behavior was seen from the skewness values, though the edges were once again heavily affected by the transformation sequence. Skewness values as low as -0.382 were observed in with higher skewness and extreme kurtosis values. Considering the individual datasets, the resulting skewness went as low as -1.042 and as high as 2.214. The highest average value is 1.189, still significantly higher than the values produced by using the Johnson transformation method. For no skewness distortion, the recovered skewness decreased monotonically from -0.008 to -0.03. Positive skewness distortion resulted in mostly positive skewness until the edges, where the behavior was the opposite with negative results. And negative skewness distortion resulted in mostly negative resulting skewness, however at the edges once again the sign quickly changed to higher positive values.

The change in fatigue life estimation values with the transformation sequence were between 70% less and 20% more, a large variance depending on both the kurtosis and skewness values as can be seen from the graph patterns. The edges with both higher and lower skewness show a heavily decreased fatigue result. And a higher than before estimation at low absolute skewness distortion with the gradual soft increase with kurtosis, similar to the resulting kurtosis graph.

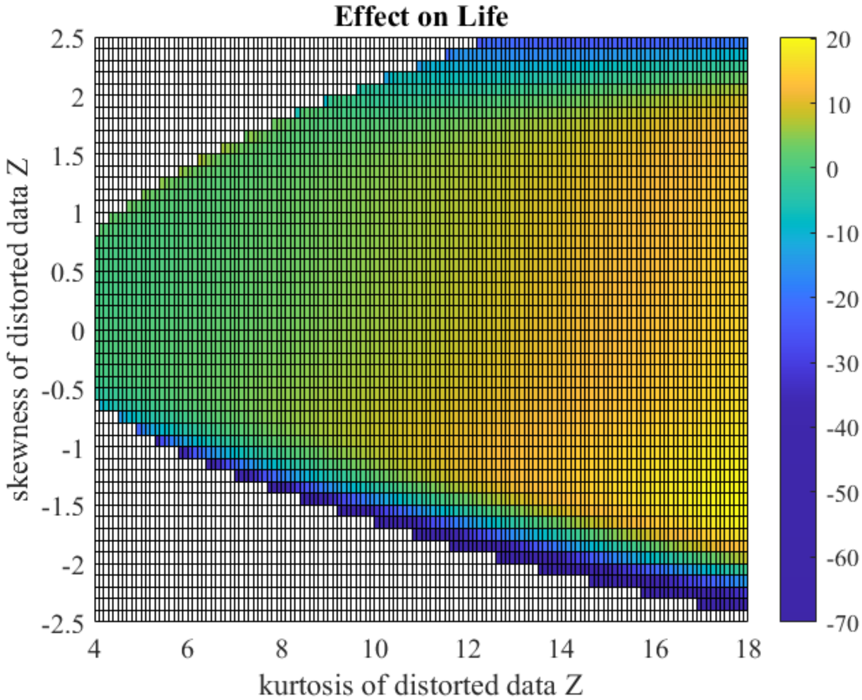

Figure 6 Change in estimated life of load, before and after distortion and Gaussian transformation via Hermite transformation, plotted against the skewness and kurtosis values the data were distorted to

4 DISCUSSION

The variation of both kurtosis and skewness with Johnson transformation were much smaller in range than that of Hermite transformation.

The kurtosis values after Johnson transformation all skewed towards 2.8, regardless of the values, it was distorted to original values. The average values were between 2.802 and 2.892 and considering the load samples individually, the secondary Gaussian kurtosis values had a range between 2.77 and 2.92

The edges of the viable range had a lower kurtosis values and the closest to 3 was towards the center, where no skewness change was applied to the data to begin with. However, kurtosis consistently undershooting to below 3, which also affected the fatigue calculations.

The resulting kurtosis and skewness values from distorting the loads to non-Gaussian and then using the inverse function to retransform them to Gaussian by the Hermite method were much more varied. For both the kurtosis and skewness of the final load, the edges of the data behaved wildly differently than the majority, while the rest of the data behaved more monotonously than those of the Johnson transformation method

The effect of this series of Hermite polynomial transformations on life estimation was very varied and followed a similar trend as the kurtosis and skewness results, showing extremities on the edges and a monotonically increasing pattern with increasing kurtosis in the center of the graph where skewness was changed less. The changes in fatigue mildly followed a merge of the patterns seen in kurtosis and skewness graphs. As the kurtosis and skewness changes were more varied for this transformation, the fatigue calculation was also more varied than its Johnson transformation counterpart. The part of the graph with smaller kurtosis values were more accurate, with less than 5% difference up to kurtosis of 8, and 10% up to a kurtosis of 11 for the majority of skewness spectrum.

5 CONCLUSIONS

When a Gaussian time series is distorted by transforming into a highly leptokurtic form and back using the Slifker formulation for the Johnson transformation and Winterstein's formula for the Hermite polynomial transformation, the final result is not exactly Gaussian in that the kurtosis as differs from 3.0.

Distortion and recovery by the Johnson Method resulted in skewness very close to zero and kurtosis consistently between 2.8 and 2.9. As a consequence, lifetimes estimated by the Dirlik method were consistently 28 – 30 % less than the lifetime of the original data.

Hermite polynomial method gave good results for most of the parameter space but at the edge of maximal skewness the results of this method showed large variations. Skewness, kurtoses and lifetimes of the restored data for this part of parameters were far from the original data. Kurtoses as large as 10 and skewness with absolute values as large as 0.8. These deviations from normal distribution resulted in expected fatigue life to vary by up to 60 % for data distorted to kurtosis of 18 and skewness of ± 2.

6 ACKNOWLEDGEMENT

This study is supported by the Scientific and Technological Research Council of Türkiye (TUBITAK) under the grant number 120M214

REFERENCES

[1] T. Dirlik, D. Benasciutti, Dirlik and Tovo-Benasciutti Spectral methods in vibration fatigue: a review with a historical perspective. *Metals 2021,* **11**, 1333, 2021.

[2] A. Trapp, Q. Hoesch, P. Wolfsteiner, Effects of insufficient sampling on counting algorithms in vibration fatigue. *Procedia Structural Integrity*, **38**, 260–270, 2021.

[3] M. Grigoriu, Crossings of non-Gaussian translation processes. *Journal of Engineering Mechanics*, **110**, 610–620, 1984.

[4] S.R. Winterstein, Non-normal responses and fatigue damage. *Journal of Engineering Mechanics*, **111**, 1291–1295, 1985.

[5] S.R. Winerstein, Nonlinear vibration models for extremes and fatigue. *Journal of Engineering Mechanics*, **114,** 1772–1790, 1988.

[6] G.E.P. Box, D.R Cox, An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, **26**, 211–243, 1964.

[7] N.L. Johnson, Systems of frequency curves generated by methods of translation. *Biometrika*, **36,** 149, 1949.

[8] C. Braccesi, F. Cianetti, G. Lori, D. Pioli, The frequency domain approach in virtual fatigue estimation of non-linear systems: The problem of non-Gaussian states of stress. *International Journal of Fatigue,* **31**, 766–775, 2009.

[9] F. Cianetti, M. Palmieri, C. Braccesi, G. Morettini, Correction formula approach to evaluate fatigue damage induced by non-Gaussian stress state. *Procedia Structural Integrity*, **8**, 390–398, 2018.

[10] D. Benasciutti, Fatigue analysis of random loadings. 2004.

[11] D. Benasciutti, R. Tovo, Frequency-based analysis of random fatigue loads:Models, hypotheses, reality. *Materialwissenschaft und Werkstofftechnik,* **49**, 345–367, 2018.

[12] J.F. Slifker, S.S. Shapiro, The Johnson system: selection and parameter estimation. *Technometrics,* **22**, 239–246, 1980.

[13] M. Liu, X. Chen, Q. Yang, Estimation of peak factor of non-Gaussian wind pressures by improved moment-based Hermite model. *Journal of Engineering Mechanics,* **143,** 2017.

[14] F. Wu, G. Huang, M. Liu, Simulation and peak value estimation of non-Gaussian wind pressures based on Johnson transformation model. *Journal of Engineering Mechanics,* **146**, 2020.

[15] Q. Yang, X. Chen, M. Liu, Bias and sampling errors in estimation of extremes of non-Gaussian wind pressures by moment-based translation process models. *Journal of Wind Engineering and Industrial Aerodynamics*, **186,** 214–233, 2019.

[16] L. Peng, M. Liu, Q. Yang, G. Huang, B. Chen, An analytical formula for Gaussian to non-Gaussian correlation relationship by moment-based piecewise Hermite polynomial model with application in wind engineering. *Journal of Wind Engineering and Industrial Aerodynamics*, **198**, 2020.

[17] T. Dirlik, Application of computers in fatigue analysis. 1985.