Statistical estimation of the difference of two time-histories

Guy Nusholtz
FCA US LLC

ABSTRACT – There are many methods to determine and characterize the relative differences of two-time histories, test to test, model to test, or comparing two different systems like different vehicle responses or different dummy responses. This short communication presents a method for comparison of two-time histories (model and the average of two or more tests) using a method that estimates the minimum differences of two-time histories. The method uses: magnitude, shape, phase, and Global coefficient of variance (GCV). The comparison of the two signals is done using GCV and/ or a derivative of GCV (G-R^2) and the three values, magnitude shape and phase is used to understand the characteristics of the difference.

INTRODUCTION

Two techniques commonly used for comparing time-histories and evaluating their difference with a single parameter are CORrelation and Analysis (CORA) and the ISO/TS 18571 standard. These techniques are used to evaluate repeatability and reproducibility in testing, evaluating anthropometric tests devices and comparing models to test data. The processes includes both corridor and cross correlation method's, which are calculated independently and then combined to give an estimation of the magnitude of the difference. The two methods use different corridor and cross-correlations methods which give different values for the different components However, their algorithms can lead to discrepancies, inconsistencies and possibly contradictions when interpreting the difference between two time-histories. In addition, the two can give very different results operating on the same data. (Matthew, et all, Saunders). In addition, the ratings for both CORA and ISO are not unique, in that there are parameters that can be adjusted and so two comparisons of the same two time-histories for ISO and CORA can be different depending on the weighting factors and methods of calculating the individual components (Sanders, Davis). In addition to these two there are other methods (Nusholtz, Xu) which introduced the concept of the global coefficient of variation (GCV), showed how it can be used to evaluate repeatability and reproducibility using experimental data and showed its relationship to cross-correlation methods using magnitude and shape. This communication expands on the methods in (Nusholtz, XU), and shows how it can be used to compare two time-histories such as the average of biomechanical impact test signals to a model of the biomechanical impact tests.

METHODS

Assume a collection of transducer time histories from biomechanical impact tests to a given anatomical structure. These tests are than averaged to form a mean time-history for the biomechanical impact tests (x). Assume also that there is a model of the anatomical structure that was impacted, and the model(y) was run to imitate the biomechanical tests.

The global coefficient of variation (GCV) is the square root of the sum of the of the squared differences of the two time-histories (x and y) when the squared differences are minimized by a time shift p and divided by the sum of the square root of the sum of squares of the x time history. Therefore, GCV represents a normalized least square like value estimation of the minimum difference between two signals. In this analysis the test (experiments) time-history is considered “ground truth” and the model is being evaluated as to how it compares to ground truth.

Analysis procedures:

SSSX= Square root of the Sum of Squares of the x time history

SSSY= square root of the sum of squares of the y time history

$$SSSX = \sqrt{(\textstyle\sum_{n=1}^{last}(x_n)\text{^}2)}$$

$$SSSY = \sqrt{(\textstyle\sum_{n=1}^{last}(y_n)\text{^}2)}$$

SSMDXY= sum of squares of the minimum difference between the x and y time histories.

P= the shift in time to minimize SSMDXY. It is the same time shift that defines the maximum in the cross-correlation of two-time histories.

$$\text{SSMDXY} = \sum_{n=1}^{last}(x_{n+p} - y_n)\text{^}2$$

MXY= magnitude of the comparison of the time histories of x and y. This magnitude is somewhat different from the magnitude calculated for finding the relative magnitude between two sets of time-histories when neither is ground truth. In that case the magnitude runs from 0 to 1. In this case the model is compared to ground truth and so the magnitude can run from 0 to infinity.

MXY=(SSSY/SSSX)

XY= Sum of the product of the x and y time histories after a shift that produces the minimum difference between the two.

$$XY = \sum_{n=1}^{last}(x_{n+p} * y_n)$$

SXY= The shape comparison of the time histories of X and Y

SXY=XY/(SSSX*SSSY)

GSS= $$\sqrt{(}\sum_{n=1}^{last}(x_{n+p} - y_n)^2$$

GSS is the square root of the sum of the squared point wise difference of the two time-histories.

GCV=(GSS)/(SSSX)

$$\text{GCV} = \frac{\sqrt{(}\sum_{n=1}^{last}(x_{n+p} - y_n)^2)}{\sqrt{(}\sum_{n=1}^{last}(x_n)\text{^}2)}$$

Since the denominator of GCV represents the square root of total sum of squares for a signal from -infinity to + infinity with only a finite section different from zero (mean=0) and the numerator represents the square root of the sum of squares for the difference of the two signals, the GCV pseudo R squared approximation (G-R^2) or R^2 like is:

G-R^2= 1-GCV^2

Results:

The following results are obtained from computer idealization of both biomechanical impact tests and a model of those impact test.

Assume that a(x) is a response (Force, Acceleration, etc.) of the average of several biomechanical tests and assume that b(x) is a model of those tests (Figure 1). In both cases x represents time.

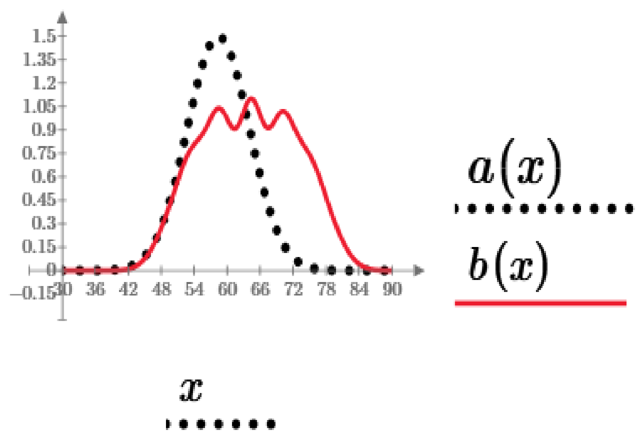

Figure 1

The first step is to shift the model signal such that the differences are minimized (figure 2). This is done through the maximum value of the crosscorrelation and it assumes that the time shift is small enough so it doses not have an effect on the sum of squares, there is significant zero value at the beginning and end of the signals.

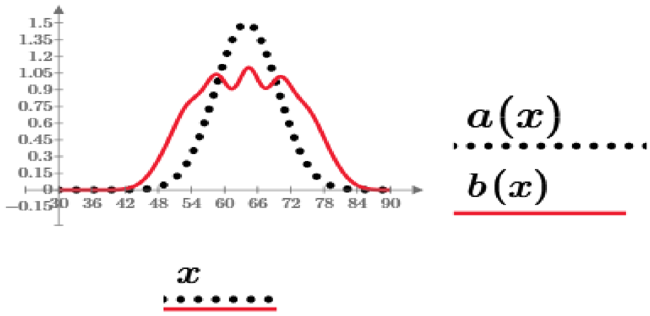

Figure 2

The magnitude=1.008 shape=.892 phase =7.3 ms GCV =.467 and G-R^2=.782

This G-R^2 is not the same as a R^2 in a logistic regression because the model is not constrained by the data as in a logistic model and so this G-R^2 does not have a value from 1 to 0 but a value from 1 to -infinity. An example to help explain this is given in figure 3 in which the model b(x) is larger than the test mean by a factor of 3.

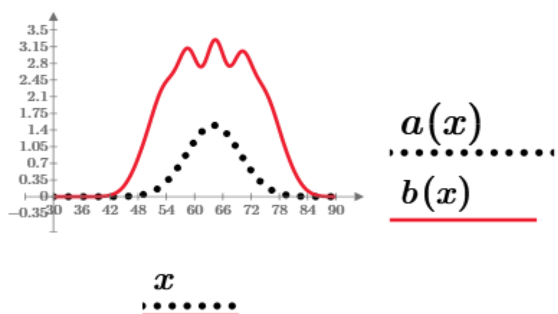

Figure 3

Mag=3.023 Shape= .892 Phase= 7.3 GCV=2.170 R^2=-3.748

Below (Figure 4) are two time histories (a1(x) and a2(x)) that are relativity close and for all practical purposes a1 can be considered as a filtered or smoothed version of a2. Mag=1.004, Shape=.999, GCV=.049 and G-R^2=.998

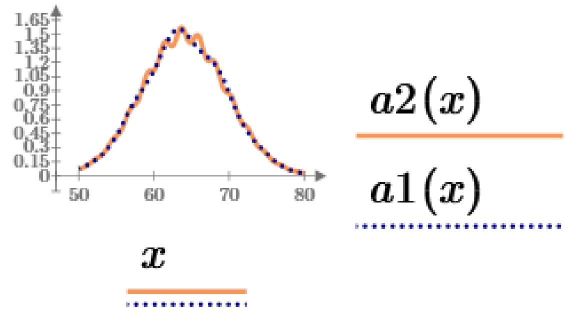

Figure 4

DISCUSSION

The results represented by figure 2 indicate that the GCV is .467 and the G-R^2 is .782. The analysis presented in (Xu), for a version of GCV in which there is no ground truth, indicates that values above .4 for GCV and below .86 for G-R^2 indicates that these time-history should not be considered similar; below .2 GCV or above .96 G-R^2 should be considered similar and between the two they are marginally similar. Although the difference in magnitude is less than 1% the shape and phase indicate a significant difference with part of the shape difference associated with the model being 14.6 ms. longer in duration. In figure 3, the model was larger by a factor of 3 indicated by the magnitude being 3 times higher. The GCV is 2.170 and the G-R^2 is -3.784 indicating that the test and the model are not related or represent the same phenomena. The shape and the phase are the same. In figure 4, the signals can be considered to be functionally the same with a G-R^2 of .998.

In short, the average of a collection of time-histories from similar tests to a model, idealization, or simulation, is done through the use of GCV or G-R^2. The magnitude, shape and phase are used to explain the difference between the test data and the model.

In this communication only the comparison between a single time-history(average) for a test series and a model were considered. However, the methods, with some modifications to the procedures and with some additional statistical procedures, can be used for other types of comparisons, such as: Determining the similarity and differences in comparing sets of signals i.e., repeatability and reproducibility between two different Anthropometric test devices. In addition, to comparison of the model to the mean of a set of test time-histories the model can be compared to each time-histories to indicate the confidence level of the model being considered one of the tests. This includes an evaluation of the test quality and scatter/variance of the test data: An alternate to a corridor evaluation.

CONCLUSION

The comparison, evaluating the magnitude of the differences of two time-histories, such as an average of test data and model, can be done using GCV and/ or G-R^2. To help understand the meaning of the differences the three values, magnitude, shape, and phase can be used to understand the characteristics of the difference.

REFERENCES

.

Saunders, James (2018) Evaluation of full and have barrier in oblick frontal tests. SAE Government Indursty meeting 2018

Davis, Matthew L., Koya Bharath, Schap Jeremy M., , Hsu Fang-Chi, Gayzik Scott (2017) Comparison of objective rating techniques vs. expert opinion in the validation of human body surrogates 25 ESV Paper Number 17-0141

Lan Xu and Guy S. Nusholtz, (2017) “Analysis of Repeatability and Reproducibility of ATD Response by the Correlation Method”, Stapp Journal, Volume 61 2017

Nusholtz Guy, Yibing Shi, (2015) Assessment of Similarity of a Set of Impact Response Time Histories, SAE international Journal of Transportation safety 3(1)28-38