THE MODAL DISTRIBUTION METHOD: A NEW STATISTICAL ALGORITHM FOR ANALYZING MEASURED RESPONSE

A Dissertation

by

MYOUNG KEUN CHOI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2009

Major Subject: Civil Engineering

THE MODAL DISTRIBUTION METHOD: A NEW STATISTICAL ALGORITHM FOR ANALYZING MEASURED RESPONSE

A Dissertation

by

MYOUNG KEUN CHOI

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Approved by:

Chair of Committee,	Bert Sweetman
Committee Members,	Richard Mercier
	Luciana Barroso
	Ayal Anis
Head of Department,	David Rosowsky

May 2009

Major Subject: Civil Engineering

ABSTRACT

The Modal Distribution Method: A New Statistical Algorithm for Analyzing Measured Response. (May 2009)

Myoung Keun Choi, B.S., Seoul National University;

M.S., Seoul National University

Chair of Advisory Committee: Dr. Bert Sweetman

A new statistical algorithm, the "modal distribution method", is proposed to statistically quantify the significance of changes in mean frequencies of individual modal vibrations of measured structural response data. In this new method, a power spectrum of measured structural response is interpreted as being a series of independent modal responses, each of which is isolated over a frequency range and treated as a statistical distribution. Pairs of corresponding individual modal distributions from different segments are compared statistically.

The first version is the parametric MDM. This method is applicable to well-separated modes having Gaussian shape. For application to situations in which the signal is corrupted by noise, a new noise reduction methodology is developed and implemented. Finally, a non-parametric version of the MDM based on the Central Limit Theorem is proposed for application of MDM to general cases including closely spaced peaks and high noise. Results from all three MDMs are compared through application to simulated clean signals and the two extended MDMs are compared through application to simulated noisy signals. As expected, the original parametric MDM is found to have the best performance if underlying requirements are met: signals that are clean and have well-separated Gaussian mode shapes. In application of nonparametric methods to Gaussian modes with high noise corruption, the noise

reduction MDM is found to have lower probability of false alarms than the nonparametric MDM, though the nonparametric is more efficient at detecting changes.

In closely related work, the Hermite moment model is extended to highly skewed data. The aim is to enable transformation from non-Gaussian modes to Gaussian modes, which would provide the possibility of applying parametric MDM to well-separated non-Gaussian modes. A new methodology to combine statistical moments using a histogram is also developed for reliable continuous monitoring by means of MDM.

The MDM is a general statistical method. Because of its general nature, it may find a broad variety of applications, but it seems particularly well suited to structural health monitoring applications because only very limited knowledge of the excitation is required, and significant changes in computed power spectra may indicate changes, such as structural damage.

To my family

ACKNOWLEDGMENTS

This dissertation arose in part out of years of research that has been done since I came to Texas A&M University. It is a pleasure to convey my gratitude to my committee in my humble acknowledgment. In the first place I would like to record my gratitude to Prof. Sweeetman for his supervision, advice, and guidance from the very early stage of this research as well as giving me extraordinary experiences throughout the work. Above all and the most needed, he provided me unflinching encouragement and support in various ways. His truly engineering intuition has made him as a constant oasis of ideas and passions in engineering, which exceptionally inspired and enriched my growth as a student, a researcher and an engineer I want to be. I am indebted to him more than he knows. I gratefully thank Prof. Mercier, Prof. Barroso, and Prof. Anis for their constructive comments on this dissertation. I am thankful that in the midst of all their activity, they agreed to be members of this committee.

TABLE OF CONTENTS

CHAPTER Page

CHAPTER Page

CHAPTER Page

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

CHAPTER I

INTRODUCTION

A new statistical method, called Modal Distribution Method (MDM), was developed to compute the significance of changes in modal response frequencies, not the significance of structural damage. The present method monitors a global condition of the system via various response time-histories, such as stress, strain, acceleration, and displacement. The parametric MDM developed for well-separated Gaussian modal distributions was extended for application to noise-corrupted modal distributions and closely-spaced modes. A new noise reduction technique, a modified version of power subtraction method, was implemented into the parametric MDM to detect changes in modal frequencies corrupted by high noise levels, and the central limit theorem was implemented to detect changes in severely corrupted modal distributions like closely-spaced modes.

The application of the MDM to the non-Gaussian shape of modal distributions requires transforming them into the Gaussian shape of modal distributions and using a relatively powerful parametric statistical comparison. The Hermite moment model was completed by overcoming its monotone limit to be used as the transformation. In addition, the features representing all previous measurements can be combined after those measurements have been determined to statistically be the same. The MDM uses statistical moments of modal distributions as the features. The combined statistical moments are believed to be more reliable features since they include more data than the feature extracted from single measurement. More reliable continuous monitoring is expected when the combined statistical moments are compared with

---

The journal model is *IEEE Transactions on Automatic Control.*

the statistical moments of modal distributions extracted from the new measurement. A new efficient statistical moment combination method applicable to the frequency domain analysis was developed in the present study.

A. Background

All structures are designed to perform their functions while sustaining the applied loads over a pre-established period. There are so many old structures operating over their pre-established lifetime, because their inspection and maintenance costs less than their demolition and reconstruction. The strength reduction due to the fatigue accumulation by environmental and service loads increases the risk of structural destruction unless appropriate monitoring and maintenance are accomplished, as learned from the case of the I-35W Mississippi River bridge collapse. In addition, modern structures are optimally designed by minimizing the structural weight together with performing the required service and providing minimum protection against failure. Consequently, the optimized structures can be seriously damaged when the structures are exposed to unexpected severe loading conditions or deterioration. In other words, optimal design of new structures can include vulnerability to damage.

Unchecked structural damages in both old and new structures diminish their capacities to perform intended functions and lead to structural failure, which can cause economic and life losses. Thus, it is essential to continuously monitor the structures for the purpose of maintaining adequate safety and detecting damages at the initial stage, resulting in lower repair costs.

Rytter [1] summarizes various non-destructive inspection methods through simple explanation and comparison to detect damage and warn before structures fail.

Conventional visual inspection can be used to detect only surface damage and requires accessibility to that surface. Most other local methods such as acoustics, eddy currents, magnetic fields and radiography also require access to the surface, and frequently require a shut down of production or service. A global monitoring method is necessary to specify the suspect sections, and identify failure modes of the structure since the local methods mentioned above are time-consuming and expensive to inspect the entire structure, especially for a large structure.

The basic idea of global monitoring is to obtain information on the soundness of a structure using the modal parameters extracted from the measured responses. A new algorithm quantifies the significance of observed changes in the modal parameters from the structural responses to either artificial or operational loads. The hypothesis is that the changes in the modes of a structure are sensitive indicators of changes in its physical properties. In other words, the damage will change the structural properties, which leads to changes in the modal parameters such as natural frequency, damping factors, and mode shapes.

However, changes in other conditions except structural properties can also contribute to changes in the modal parameters. Thus, the new algorithm could eventually be used as a part of a structural health monitoring system by discriminating between the changes in modal frequencies by variation of structural properties and those by variation of other conditions such as environmental loading or temperature. The main advantage is that the measurements at one location are sufficient to perform an assessment of the entire structure.

B. Historical Review

Integrity monitoring was first developed in the oil and gas industry in the 1970's. Vandiver [2] [3] investigated the changes in resonant frequencies of an offshore light station tower to identify damage based on numerical simulations. Begg et al. [4] described that the severing of offshore structural members in model scale could produce about 5% to 30% change in resonant frequencies. Kenley and Dodds [5] concluded that 5% change in overall stiffness was required to detect changes in the resonant frequencies, based on a decommissioned offshore platform. The resonant frequency can be found within 1% change for a global mode, but the error increases to 2% to 3% for local modes because the peaks in the power spectrum are not well defined. Whittom and Dodds [6] monitored British Petroleum's Alpha Forties platform over 2.5 years, and concluded that the changes in the resonant frequencies produced by damage or foundation deterioration were greater than the observed variations in resonant frequencies of the undamaged platform over time.

Research on integrity monitoring, started in the offshore industry, has continued in the field of rotary machinery since the 1980's. So called"Condition Monitoring" has come from laboratory research topics, and it has been widely used in practical applications. Mature integrity monitoring in the rotary machinery field encouraged civil engineers to adapt the idea to a large civil structure. "Structural Health Monitoring" started being a new research area in the 1990's. Doebling et al. [7] extensively reviewed the research on structural health monitoring. In their review, the various methods were categorized based on the types of measured data and the techniques identifying the damage from the data. Many of the techniques were based on changes in modal frequencies and mode shapes, which could be yielded by modal analysis.

One of the issues in structural health monitoring is environmental variability,

which can contribute changes in an extracted feature. When modal parameters are used, the changes in modal parameters by environmental variability should be discriminated from those by structural damage for reliable and accurate monitoring. To address the environmental variability, the necessity of a statistical algorithm has been recognized. However, almost none of the studies summarized in Doebling et al. [7] [8] make use of any statistical methods to assess if the changes in the selected features are statistically significant. The MDM aims at being used as a part of the structural health monitoring that uses a statistical method.

C. Motivation

The large amount of vibration data and the associated heavy computational loads often require data compaction before analyzing. Modal parameters are generally used as the compact feature representing dynamic characteristics of the structure. Some derivatives or indices from the basic modal parameters might be more sensitive to damage, but are also more sensitive to environmental conditions and measurement noise. Natural frequencies are measured more accurately than mode shapes; typical resolution for the natural frequencies of a lightly damped structure is 0.1% whereas typical mode shape error is 10% or more [9]. The uncertainty of the extracted modal parameters requires a statistical approach.

Information about the modal properties can be found in the frequency response function resulting from the forced response measurements with controlled or measured excitation. Many large structures such as bridges and offshore platforms, however, are not easily excited by artificial loads, but are often excited by natural loads such as wind and wave loads. Therefore, it is necessary to reduce the dependence on artificial excitation mechanism in monitoring of structures. The ability to use the vibrations

induced by environmental or operating loads for the assessment of structural integrity has advantages: 1) testing is less time-consuming since the equipment for exciting the structure is not needed, 2) testing does not interrupt the operation of structure, and 3) the measured response represents the real operating condition of the structure. However, use of the vibrations induced by natural loads is more complicated since there is usually no information about the loads [10]. Again a statistical approach is imperative to account for variability in environmental and operational loads. Sohn et. al [11] recommended the normalization of the measured responses by the measured inputs when environmental or operating-condition variability is an issue.

Many of the proposed monitoring methods depend on prior analytic models or prior test data to detect damage. Some methods need a detailed finite element model of the structures being monitored while others use data sets measured from the undamaged structure. The dependence on this type of data should be reduced since the lack of such data can make them impractical for application. Additionally, noise corruption should be addressed since the measured modal properties are inevitably corrupted with measurement noise no matter how precise the instrumentation [12].

Conventionally, modal parameters are extracted from the frequency response functions obtained by measuring both excitation and response of the structure. The relation is simple; the power spectral density of the response is a product of the square of the frequency response function and the power spectral density of the excitation. The MDM should extract the modal parameters without measuring the input force. The input force is typically assumed as white noise. Under the white noise excitation force, the power spectral density of a response becomes the product of the constant white noise level and the square of the frequency response function. Thus, the modal parameters can be extracted directly from the power spectral density of the response of a structure. In the MDM , stationarity of the excitation is more important than the

shape of the power spectral density of the excitation. If excitation is non-stationary, the excitation would change the modal parameters. Changes in modal parameters by the non-stationary excitation would be included in the MDM analysis unless the changes are differentiated from those by structural damage or full information of the non-stationary excitation is available.

D. Significance of Work

The objective of this work is to develop a new algorithm quantifying statistical significance of observed changes in modal frequencies with consideration of its uncertainty, which could play a critical role in future structural health monitoring applications. The new algorithm periodically measures the vibrations induced by natural loads while the structure is in service; then, the modal frequencies are extracted from the measurements, and are statistically compared with those from previous measurements. Hence, a statistical comparison of modal frequencies is important in reliable monitoring since the statistical comparison accounts for the variability of modal frequencies.

The new statistical algorithm should quantify the significance of changes in modal parameters using vibration responses induced by natural loads. Considering modal frequency as a random variable, an individual modal distribution is interpreted as a probability distribution of the modal frequency. This interpretation enables application of a statistical comparison to address the uncertainty associated with the modal frequency. The resulting P-value indicates the quantitative significance of changes in the modal frequencies, and help to indicate whether the system has changed or not. Implementation of this algorithm into structural health monitoring for automatic detection requires differentiation of changes in a modal distribution caused by structural

damage from those caused by environmental variability, which is outside the scope of this work.

The new algorithm may detect small levels of change by monitoring the structure over a long time period. Implementation of a newly proposed noise reduction method and a nonparametric method can address the problem of white noise corruption and unknown modal distribution shape, respectively.

E. Objective: The Modal Distribution Method

The main purpose of this work is to develop a new statistical method for continuous monitoring called the "Modal Distribution Method." In the Modal Distribution Method (MDM), a power spectrum of measured structural response is interpreted as a series of independent modal responses, and each of these responses is isolated over a frequency range and treated as a statistical distribution. The MDM uses the modal distributions of vibratory response generated by natural loads, and quantifies the significance of observed changes in those modal distributions. The introduction of a modal distribution representing the individual modal response enables statistical comparison. Rough estimates of modal frequencies from a previous analysis or measurement are required to separate those modal distributions.

For example, the simplest system with a single modal response is considered. The whole power spectrum of a measurement is then considered the "modal distribution," representing the single modal response, since there is only one mode. The power spectrum is compared with the power spectrum of some other measurement. A statistical hypothesis test is then applied to those two power spectra. In the case of well-separated Gaussian modes, individual modal distributions from different measurements can be compared by T-statistics. The resulting significance level indicates

the likelihood that observed difference is caused by chance alone.

Real application of the MDM to complex structures needs to consider multiple modes. Each modal distribution in the multi-modal spectrum exists inside the frequency range between two troughs adjacent to its peak. Every trough is found by a newly developed penalty method between rough guesses of two consecutive modal frequencies. Rough guesses can be given by pre-analysis or simply by the selection of peaks in a previous sample spectrum. The modes are separated at the trough identified by the penalty method. The resulting individual modal distributions from different measurements are compared by a statistical hypothesis test. The combined statistic detects changes in an entire multi-modal spectrum, and is expected to be more sensitive since it uses more data. Relative sensitivity of detection ability between the modal comparison and the spectral comparison using combined statistic is investigated in this work. The MDM is intended to detect changes in measured structural response, which are believed to be indicative of changes in the structural condition only if other conditions are constant. This method uses only output response and separates the multi-modal spectrum into each modal distribution. If separated modal distributions from different segments are Gaussian, two corresponding modal distributions are compared statistically by a relatively powerful parametric test, the T-test, which is commonly used to compare two Gaussian distributions. Comparison of all individual modes are then combined in order to detect more subtle changes in structural conditions.

In the case of high noise corruption, a noise reduction method is implemented to enhance the signal. It is shown that a simple power subtraction method generally underestimates the variance of the modal distribution. A modified noise reduction method is newly developed and imbedded into the parametric MDM for application to noisy measurements. This approach is verified by application to simulations based

on field riser Vortex Induced Vibration (VIV) measurement. As two adjacent modes become closer and closer, the original shapes of the modal distributions are not preserved, whether they are Gaussian or non-Gaussian. In this case, a distribution-free statistical comparison is necessary. The MDM is generalized by removing the assumption of modal distribution shape with aid of the central limit theorem. The central limit theorem guarantees that the distribution of sample averages approaches Gaussian regardless of underlying modal distribution shape [13].

A parametric comparison requires a distributionship, typically a Gaussian shape. Even when the requirement is not met, transformation of the non-Gaussian distribution to Gaussian and the use of a parametric comparison are more powerful than nonparametric comparison [14]. Hence, the Hermite moment model, which is popular for transforming a Gaussian extreme to a non-Gaussian one, is extended to a highly skewed distribution with kurtosis near three. When all statistical moments of modal distributions in previous segments are combined and compared with those in the next segment, the resulting power spectrum better represents the process. Conventional methods of computing higher moments are available in the time domain. A new efficient statistical moment combination method presented here uses a histogram and is equally applicable in the frequency domain, which is necessary for the MDM. The efficiency of computation and storage of the new method is verified in the time domain.

F. Organization of Dissertation

This dissertation is organized into seven main chapters.

Chapter I contains the background, motivation, significance of work, and objectives.

Chapter II introduces a newly developed method for comparing measured time histories statistically, coined the "Modal Distribution Method (MDM)." Modal distributions are assumed to have Gaussian shapes and to be well separated. Therefore, a parametric hypothesis test is imbedded for a statistical comparison of two Gaussian modal distributions.

Chapter III describes the extension of the original four moment Hermite moment model into a highly skewed distribution with kurtosis near three and the application of this development in time domain analysis.

Chapter IV proposes a new method using a histogram to combine statistical moments. This new method is applied in the time domain, and is proven to take less computational power and storage.

In Chapter V, noise reduction is addressed, and the method is imbedded within the parametric MDM in order to extend the MDM application to noisy measurements. Validation is presented by applying the method to noisy simulations based on the field riser VIV measurement.

In Chapter VI, a modal distribution method independent of distribution shape is described. The central limit theorem guarantees that sample averages of modal distribution have Gaussian distribution regardless of modal distribution shape. Parametric hypothesis test then compares distributions of a sample average.

Chapter VII presents the conclusion, contributions of this dissertation, and recommendations for future research.

CHAPTER II

THE MODAL DISTRIBUTION METHOD FOR STATISTICAL ANALYSIS OF MEASURED STRUCTURAL RESPONSE

The newly proposed modal distribution method statistically quantifies overall differences between measured time-histories. In this method, power spectra of measured structural response are interpreted as collections of independent modal responses. Each modal response is isolated, re-scaled, and interpreted as a statistical distribution. Data-sets (windows in a measured time-history) are then compared using standard statistical methods, resulting in a quantitative significance level of the differences between power spectra. An example is presented to validate the new method and to quantify how long a time-history is required for the new method to meet confidence level requirements. The modal distribution method is found to be very effective at detecting subtle changes of mean modal frequencies, which may be used to infer changes in structural condition. The method is general and may find a broad variety of applications, but seems particularly well suited for structural health monitoring because it can be used to infer changes in structural condition from measured response data with only limited knowledge of the excitation.

A. Introduction

The use of sensors in structural applications has expanded dramatically in recent years. The cost and complexity of data collection has decreased to the point that sensor networks intended to detect structural failure are expected to eventually become a part of every major new engineered structure. Acceleration remains the preferred measurement because of low sensor cost and high reliability. Detecting variations in structure response, making decisions, and taking appropriate action all

rely on interpreting measured data, which is inherently a statistical problem. Simply detecting that a measured quantity has changed is not sufficient for decision making: a method to quantify the statistical significance of observed changes in measured data is necessary.

The earliest work for damage detection through modal analysis was on offshore structures. Early investigators e.g. [15, 16] concluded that changes in modal frequencies due to member damage of offshore structures are detectable based on numerical simulation and a scale model, respectively. Three North Sea platforms were monitored for six to nine months and it was concluded that changes in the response spectrum can be used to monitor structural integrity [17]. Since that time, numerous investigators have measured accelerations on various civil structures and used computed power spectra to make inferences about the condition of these structures. Conventional methods require detailed knowledge of both the excitation force and vibration response [18, 19]. In a different approach, vibration signature analysis and advanced statistical methods are applied to structural health monitoring [20], but in this work detailed knowledge of the excitation is still required. White noise excitation is often assumed in practical applications in which the excitation force is unknown e.g. [21, 22], but in some applications such an assumption may be unreasonable, especially if there is known to be a dominant frequency of excitation.

A major state of the art review for structural health monitoring (SHM) conducted as Sandia National Labs noted that little progress has been made in application of statistical methods to SHM [23]. The work presented here addresses the need for additional statistical methods for interpretation of power spectra computed from measured accelerations in SHM applications.

The proposed method compares modal response characteristics computed for separate segments of a measured time-history. The new method has important ad-

vantages over more conventional analyses: 1) The only measured time-history required as input is that of the response, 2) the acceleration data is separated into individual vibrational modes of response, enabling detection of vibrational changes in individual modes even if the overall response is relatively unaffected, 3) if there is a dominant excitation frequency which varies over time, the mode containing that frequency can be excluded from the analysis such that a shift in the excitation frequency is not misinterpreted as a change in structural behavior, and 4) the result is a numerical prediction of the confidence level that there has been some change in the response of the structure. The method presented here is intended only to quantify the statistical significance of shifts in modal frequencies, and leaves the any inference regarding structural parameters such as mass, stiffness, or damage to others.

B. Theory: The Modal Distribution Method

1. Overview

This section gives an overview from a conceptual standpoint; computational detail is presented in Section C. In the newly proposed method, a measured acceleration time-history is divided into a series of segments, or windows, which are to be sequentially analyzed. Each segment is first converted into a power spectrum through use of the FFT. As always, use of the power spectrum implies stationarity within any window of the time-history, and use of the Fourier transform implies ergodicity. It is implicitly assumed that the dynamic behavior of a structure in a given frequency range can be considered as a set of individual modes of vibration and that these vibrational modes can be considered to be independent. Accordingly, the power spectrum is divided into a series of response frequency ranges, each of which includes a distribution of energy around a single peak frequency. Hereafter, this distribution of energy around a single

peak frequency is referred to as a modal distribution. The process of dividing the modal distributions is based on using a penalty method to find local minima between energy peaks (Section 1). After this initial division, each modal distribution is treated individually. Any frequency ranges excepting those including a dominant frequency of excitation are presumed to be modal responses; those ranges including dominant frequencies of excitation can be explicitly excluded from the analysis.

Each remaining mode is mapped into a statistical distribution in which the random variable is the modal frequency. The theory underlying this new methodology is further discussed in Section 2. The overall goal is to quantify the statistical significance of subtle changes in the state of a vibrating structure by comparing modal frequency distributions representing two different segments of a measured response time-history. The mean of each sequential modal distribution in the first window is statistically compared with that of the second window, i.e, the mean of the first mode of the first window is compared with the mean of the first mode of the second window, as are each of the subsequent modal means. In the implementation presented here, the T-statistic is applied to quantify the significance of observed differences. The significance of each mean difference is computed independently, and then results are weighted by the magnitude of the energy associated with that mode to compute an overall combined T-statistic. Applicability of the T-statistic and computation of degree of freedoms are further discussed in Section 3. The statistical significance of any difference between individual modal means is reported, as well as a combined statistic quantifying differences between all modal means. The entire procedure could be repeated over consecutive segments (windows) of a time-history to monitor for changes in modal response.

2. Interpretation of a Power Spectrum as a Statistical Distribution

The heart of the new method lies in interpretation of the vibrational response spectrum as a statistical distribution of modal frequencies. The individual frequency bars making up the discrete power spectrum represent the energy occurring at that frequency, so the discrete power spectrum can be considered as a discrete distribution in which the abscissa is the frequency and the ordinate is the amount of energy at that frequency. The region of the power spectrum associated with any single mode is then normalized to unit area, such that each bar represents the fraction of total energy of the individual mode. The resulting distribution has the same properties as a conventional relative histogram, in which any bar commonly represents the fraction of total occurrences. This interpretation of the power spectrum enables use of conventional statistical tools to compute the significance of changes in mean modal frequencies.

Interpretation of power spectra as statistical distributions for use in statistical inference is quite novel, but the inverse of this interpretation is commonly used in simulation. An infinite time-history of a stationary process would result in a theoretically perfect (statistical) distribution of the energy of the response, typical of a target power spectrum used in simulation. Mapping of a power spectrum into statistical space is precisely the inverse of simulating a time series using a power spectrum with random phase. In simulation, a time-history is commonly realized from a target power spectrum as:

$$x(t) = \sum_{n=1}^{N} H_n cos[2\pi f(n)t + \theta_n] \tag{2.1}$$

where $x(t)$ is the time-history of the simulation. H_n is the amplitude of each frequency component, the square of which is directly proportional to energy density at that frequency, i.e., the height of an individual bar of the power spectrum. $f(n)$ is the n'th frequency (Hz), and θ_n is a random phase angle.

The power spectrum is appropriate for interpretation as a statistical distribution, rather than the Fourier amplitude spectrum, because the total area under the power spectrum has physical interpretation as the variance of the process. Separate areas under the power spectrum can be added directly. The combined area of two frequency bars of the power spectrum represents the total power of the process at those two frequencies, analogous to combining frequencies of occurrence in a traditional frequency histogram; the combined area of two frequency bars of a Fourier spectrum has no particular physical interpretation.

3. Comparison of Statistical Distributions

An infinite time-history would result in an ideal statistical distribution of energy, but in practice a finite time-history is applied to approximate the true underlying distribution. Various statistical methodologies have been developed to compare the means of two data-sets including corrections for relatively small sample sizes. In the response spectra considered here, the Gaussian distribution provides a reasonable fit to the power spectra being compared, and so the T-statistic can be used for comparison of the two distributions. The Gaussian assumption is strongly violated in cases of very high noise or closely spaced peaks and so the T-statistic should not be used. Other statistical tests including nonparametric methods could alternatively be used in place of the T-test with no fundamental change to this underlying methodology. Under the Gaussian assumption, the mean and mode of each distribution occur at the same frequency. Use of the mean is preferred over the mode in this analysis because it provides a more robust and stable estimation of the peaks of a raw power spectrum, i.e., a power spectrum that has not been smoothed across frequency bars.

In a more typical application, the T-statistic is used with a histogram by first selecting bin widths for the raw data, then binning the data into a frequency histogram

to form a statistical distribution. The standard deviation of the raw data can be closely estimated from the geometry of the resulting histogram. The purpose of the T-statistic is to compensate for the effect of small sample size on the standard deviation as compared with an ideal underlying Gaussian distribution.

Here, the frequency intervals implicit to the FFT are equivalent to the bin widths. The energy associated with any frequency bar relates to the energy of each sinusoidal cycle at that frequency. As in a more typical application, the standard deviation of the distribution is closely estimated from its geometry; note, however, that this standard deviation is not equivalent to the square root of the variance of the process, but instead represents the distribution of energy about the mean frequency of an individual modal response. The number of degrees of freedom must also be estimated for use with the T-distribution. Here, the number of degrees of freedom is estimated as the number of cycles over which the energy is averaged, which is approximated by the mean frequency in Hertz times the duration of the time-history in seconds. This procedure is equivalent to dividing the total duration by the peak period since the mean and peak frequencies coincide under the Gaussian assumption.

C. Computational Details

1. Separation of Modal Distributions: The Penalty Method

The power spectrum is computed from the Fourier transform of a window of the measured time-history. Power spectra of measured data generally are not smooth, and identifying the correct minima to use as dividing points between modal distributions is non-trivial. The methodology presented here relies on an initial estimation of the frequency of each mode of vibration for the structure. This initial value would be expected to come from structural analysis, though it could also come from visual inspec-

tion of the power spectrum. This initial value need not be of high accuracy because it is only used to specify the endpoints of a penalty method. The penalty method searches for one local minimum between every two adjacent modal frequencies, plus one additional minimum above the highest and one below the lowest specified frequencies. The penalty method has been found to be robust in applications with high noise or closely spaced frequencies: it always finds the local minimum, regardless of the distance from zero on the spectral plot. The user is cautioned, however, that in these applications the Gaussian assumption is strongly violated and the T-statistic should not be applied. Basing the modal separation on local minima also provides additional robustness against noise in the signal because the squaring inherent to the calculation of the power from the Fourier spectrum amplifies maxima in the spectrum more than minima.

To find each local minimum, the method searches the region between two adjacent modal frequency estimates by progressively seeking the minimum of increasingly fine divisions of the region. First, the region is divided into three equal frequency intervals as shown in the second frame of Figure 1. The area under the power spectrum within each of the three intervals is then computed; every frequency in that interval having the lowest average energy is assigned a penalty value of 1.0. Next, the number of divisions between initial modal frequency guesses is increased to four. The same penalty approach is applied: again, every frequency in the region having the smallest average energy receives a penalty; on this second iteration the penalty is less than 1.0. The process is repeated with the number of divisions increasing by 1 and the penalty decreasing linearly for each subsequent penalty assessment until the average of each region includes not less than 25 of the original bars. Beyond four intervals the method is modified: a penalty is applied to every frequency in each interval having lower average energy (lower average bar height) than both of its contiguous

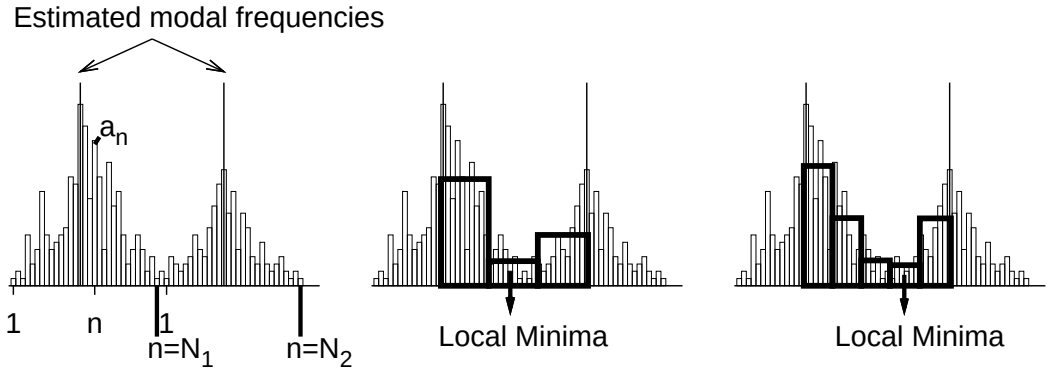


Fig. 1. Separation of Modal Distributions: The Penalty Method

neighbor intervals. After the final iteration, the total penalty is calculated for each of the frequencies between initial modal frequency estimates. The frequency having the largest total penalty is concluded to be the local minimum between energy peaks and is subsequently used as the dividing point. All energy between adjacent local minima is assumed to be associated with a single vibrational mode. Once these modes have been isolated and the peak associated with any dominant frequency in the excitation (if any) has been removed, distribution parameters are calculated for each mode.

2. Modal Distributions and Statistical Distributions

The first step in calculating distribution parameters is normalizing each modal distribution to have unit energy such that it can be treated as a frequency histogram, as outlined in section 2. Within each modal distribution, the fraction of energy associated with each frequency bar is computed as:

$$P_i(e_n) = \frac{a_n}{A_i} \tag{2.2}$$

where a_n is the area of the n'th frequency bar and A_i is the total energy associated with the i'th modal frequency (Figure 1), or the total area between relative minima:

$$A_i = \sum_{n=1}^{N_i} a_n = \sum_{n=1}^{N_i} S(n)df \tag{2.3}$$

in which $S(n)$ is the offset at the n'th frequency of the power spectrum, df is the frequency spacing and N_i is the total number of bars within the i'th modal distribution. The mean and variance are computed directly from the offsets of the power spectrum using conventional definitions of moments and central moments, e.g. [24, 25].

$$\mu_{i,w} = \frac{m_i}{A_i} = \frac{1}{A_i}\sum_{n=1}^{N_i} S(n)f(n)df \tag{2.4}$$

$$s^2_{i,w} = \frac{\theta_i}{A_i} = \frac{1}{A_i}\sum_{n=1}^{N_i} S(n)(f(n)-\mu_{i,w})^2 df \tag{2.5}$$

where $\mu_{i,w}$ and $s^2_{i,w}$ are the mean frequency and sample variance of the modal distribution, m_i is the first geometric moment, θ_i is the second central geometric moment and $f(n)$ is the n'th frequency.

Two important sources of variations in the mean and variance are implicitly considered. The first source is that the data-set being sampled is finite, so any computed mean and variance are sample statistics and not necessarily those of the underlying population; this source of variation is accounted for in the T-statistic. In this application to random vibrations there is a second important source of variation caused by non-ergodicity; it has been found here that the discontinuous ends of a time-history can have a meaningful impact on computed statistical moments that is not considered in the T-statistic. To minimize apparent differences between windows caused by this non-ergodic effect, power spectra are averaged over successive data-points. Specifically, for each window a power spectrum is computed 100 times, with both the start and end points shifted two points forward in time; the resulting 100

spectra are averaged and treated as a single representative spectrum for that window.

3. Comparison: Statistical Hypothesis Testing

Spectral peaks of field data generally appear as a cluster of energy around some frequency. Here, mean frequencies of individual modes are treated as random variables and observed clusters of energy are treated as statistical distributions. The benefit is that rigorous statistical analysis can then be applied to assess the significance of apparent differences between distributions.

The T-test is a conventional method to determine the statistical significance of a difference between two sample means from populations sampled from underlying Gaussian distributions. The test can be used when the number of samples is too small for the central limit theorem to apply and where the true variances of the underlying processes are not known to be equal. Use of the T-test in this application is not strictly justified because the underlying distributions are not known to be Gaussian, though in the example presented here the energy peaks are broadly spaced and damping is sufficient to have "bell-shaped" modal peaks, so use of the T-test may be reasonable. As noted in Section 3, for some cases the Gaussian assumptions are strongly violated and use of the T-test is not appropriate. If the T-test is deemed suitable, the T-statistic can be computed as e.g. [26]:

$$T_i = \frac{\Delta\mu_i}{s_i} \tag{2.6}$$

where

$$\Delta\mu_i = \mu_{i,1} - \mu_{i,2} \tag{2.7}$$

$$s_i^2 = s_{i,1}^2/N_{i,1} + s_{i,2}^2/N_{i,2} \tag{2.8}$$

$$N_{i,w} = D_w\mu_{i,w} \tag{2.9}$$

in which $\mu_{i,w}$ is the mean of the i'th mode of the w'th segment, $s^2_{i,w}$ is the sample variance, and the number of samples $N_{i,w}$ is estimated as the number of cycles expected at each modal frequency. D_w is the duration of the w'th data window.

T_i is distributed approximately as a Student's T with the number of degrees of freedom for the i'th mode equal to e.g. [26]:

$$DOF_i = \frac{(s^2_{i,1}/N_{i,1} + s^2_{i,2}/N_{i,2})^2}{(s^2_{i,1}/N_{i,1})^2/(N_{i,1}-1) + (s^2_{i,2}/N_{i,2})^2/(N_{i,2}-1)} \tag{2.10}$$

4. Overall Comparison of Response

Accurate assessment of the significance of changes in the observed vibrational response requires consideration of all modes. Here, the ensemble of observed differences between each pair of mean modal frequencies is treated as a set of repeated measurements with differing uncertainties, i.e., observed differences between the first pair of modal distributions are combined with observed differences between pairs of sequentially higher modes. An overall P-value is calculated by computing a combined T-statistic, which weights differences between the means by the fraction of energy represented for each individual mode. The fraction of energy associated with the i'th mode for each of the two windows is:

$$E_{i,w} = \frac{A_{i,w}}{\sum_{i=1}^{I} A_{i,w}} \tag{2.11}$$

where I is the total number of modes in the power spectrum and w is either 1 or 2, denoting either the first or second window of the time-series. The overall T-statistic is found using Equation 6.9, in which $\Delta\mu$ and s^2 are computed as averages weighted by the fraction of energy associated with each of the I modes:

$$T = \frac{\Delta\mu}{s} \tag{2.12}$$

where

$$\Delta\mu = \sum_{i=1}^{I} 0.5(E_{i,1} + E_{i,2})\Delta\mu_i \tag{2.13}$$

$$s^2 = \sum_{i=1}^{I} 0.5(E_{i,1} + E_{i,2})s_i^2 \tag{2.14}$$

The T-statistic resulting from Equation 6.9 represents the weighted average of differences between the means of each pair of modes. The number of degrees of freedom for use with a standard T-distribution is estimated as the total number of vibration cycles in all vibrational modes. This number is expected to be larger than the observed number of peaks in the time-history because modal frequencies with small fractions of the total energy will not contribute to the number of directly measurable peaks. The total degrees of freedom to be used in conjunction with the T-statistic is estimated as the sum of individual degrees of freedom:

$$DOF = \sum_{i=1}^{I} DOF_i \tag{2.15}$$

D. Type I and Type II Errors: Minimum Data Requirements

One significant issue that must be addressed prior to any practical implementation of any data analysis technique is assessing how much data is required for the method to be effective. There are exactly two ways a method can fail: 1) it could detect a difference between the two power spectra when there is in fact no underlying difference, or 2) it could fail to detect a difference that does in fact exist. The null hypothesis (H_0) implicitly underlying use of the T-test is that there is no difference between the two distributions, which makes category 1) a Type I error (reject H_0 when it is true) and category 2) a Type II error (fail to reject H_0 when it is false).

For many practical applications, it is desirable to calculate how long a dataset is required to detect a modal shift of a predetermined size with a pre-specified confidence level while maintaining an acceptably low probability of false alarms. To

compute the required length of the data-set, the P-value is itself treated as a random variable. For any specific number of vibration cycles, a histogram of predicted P-values could be created and normalized to a PDF. As part of this work, numerous sets of P-values have been generated by application of the modal distribution method to pairs of time-histories, each independently realized from a pair of target spectra with known modal differences; the distribution of predicted P-values has been found to be well described by a theoretical gamma distribution. The PDF and CDF (g and G) can be expressed in terms of the gamma function e.g. [26]:

$$g(x;\alpha,\theta) = \frac{x^{\alpha-1}e^{-\frac{x}{\theta}}}{\Gamma(\alpha,\infty)\theta^{\alpha}} \tag{2.16}$$

$$G(x;\alpha,\theta) = \frac{\Gamma(\alpha,\frac{x}{\theta})}{\Gamma(\alpha,\infty)} \tag{2.17}$$

in which $\Gamma(\alpha,y)$ denotes the gamma function: $\Gamma(\alpha,y) = \int_0^y t^{\alpha-1}e^{-t}dt$. The shape parameter ($\alpha$) and the scale parameter (θ) define the gamma distribution completely, and can be calculated from the mean (μ) and sample variance (s^2) of an ensemble of P-values as $\alpha = \mu^2/s^2$ and $\theta = s^2/\mu$.

The area under the gamma PDF below the pre-specified confidence level corresponds to the probability that any future prediction of a P-value will be at or below that confidence level, i.e., the probability that the method will detect changes in modal frequencies with the required confidence level. The probability of a Type I error equals this area if there is no actual difference between the underlying distributions; the probability of a Type II error can be computed as one minus this area if there is a difference between distributions.

It is relatively straightforward to determine the minimum required time-history for some new application using the gamma distribution. First, the mean and standard deviation of the predicted P-values must be known. These values can be determined

by simulation from target spectra with spectral shapes comparable to those expected for the physical system to be studied and with modal frequency shifts equal to those sought to be detected. The mean of the P-values is the expected value of future P-value predictions; the standard deviation can be used in conjunction with the gamma distribution to assign confidence intervals to this mean. A worked example of each error type is provided.

E. Example: Minimum Data Requirements

This example quantifies the probability of both Type I and Type II errors, then compares the results of this method with conventional tests of stationarity. The example is based on simulated data for which the actual underlying target spectra are precisely known so the effectiveness of the new method can be assessed. The underlying spectra are based on physical data of practical importance, vortex induced vibrations of marine drilling risers. Here, the riser is undamaged in all cases, and the dominant response frequencies are not related to environmental excitation frequencies: any change in modal response parameters is believed to be caused by hydrodynamic interactions, in particular, to be caused by the mass of the water entrained with the vibration.

Marine risers are the pipes transmitting fluids between the sea-floor and a floating production platform. A riser, or any other slender structure subject to strong flow across its axis, may interact with the current to create vortex induced vibrations (VIV). Prediction of this complicated fluid-structure interaction has been a historically intractable problem, in part because two separate quantities are unknown: the excitation force due to fluid-structure interaction and the effective mass of the riser which is highly influenced by the mass of the water entrained with the riser motion,

usually referred to as the "added mass."

1. Underlying Spectra and Data Simulation

In order to quantify the effectiveness of the new method, it is tested against time-histories simulated from ideal target spectra, which are precisely known. The initial target power spectrum is generated from measured field data by first computing the power spectrum from a measured time-history and then smoothing the first two modes into the ideal target spectrum S (Figure 2). Known changes in modal properties are then introduced to the smoothed spectrum (Figure 3). Time-histories needed for the example are then simulated from these spectra using the conventional method as in Equation 2.1. In this example, the simulation includes as many frequency components as there are time-steps so the series does not repeat within a simulation. Finally, a Gaussian white-noise signal is added to each time-history. The level of noise added is 1.9 mg rms, which is typical of the noise level observed in the actual field data underlying this example; noise inherent to a high-quality ± 1 g accelerometer alone is typically around 0.5 mg rms.

2. Application of the Method

To assess the minimum amount of data necessary, the new method is applied to each pair of simulated time-histories for increasingly long data windows. The shorter time-histories are the early parts of the longer histories, i.e., the 30 minute data window test for target spectrum S_1 is the first part of the time-history used in the 60 minute test. Unless otherwise noted, all cases include the 1.9 mg Gaussian white noise and all plots are generated by performing 25 independent tests and averaging the results.

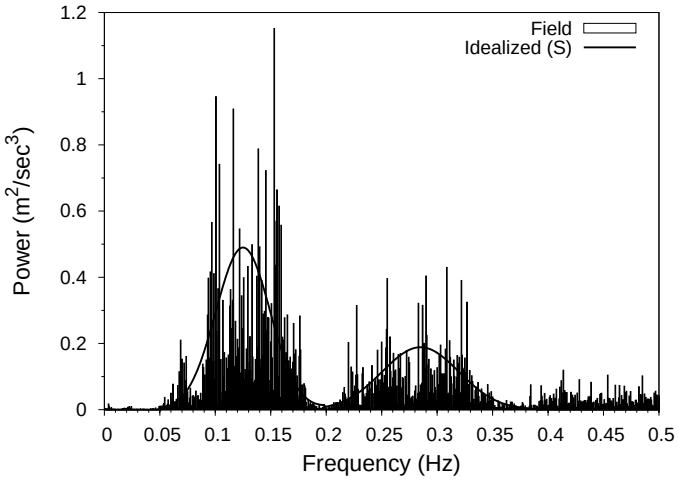


Fig. 2. Field Data Spectrum and Idealized Target Spectrum S

a. Type I Errors: No Change in True Underlying Spectra

To quantify the probability of having a Type I error, 50 realizations of the original bimodal target spectrum, S, are simulated and the modal distribution method is applied to the resulting 25 independent pairs of time-histories. Figure 4 shows the average of the P-values from these tests, all of which are above 50%. These very high P-values correspond to the fact that observed differences result from random chance and not from differences in the underlying spectra. The error bars correspond to a double-sided 90% confidence interval; there is almost no chance of a false-alarm when at least 30 minutes of data is considered.

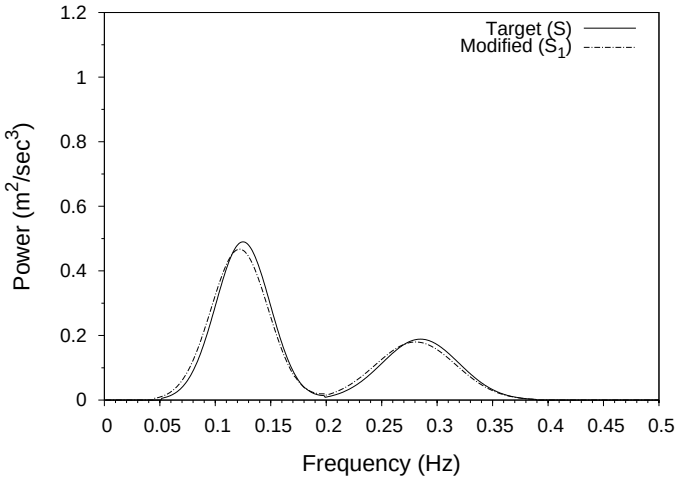


Fig. 3. Original Spectrum (S) and Modified Spectrum (S_1)

b. Type II Errors: Known Changes in Mean Modal Frequencies

In this part of the example, the mean of each mode is shifted to the left. Here, the change to the mean of each mode ($\Delta\mu_i$) is chosen such that each mean of each mode is shifted an amount consistent with a constant $\Phi(z)$ of a standard normal distribution:

$$\Delta\mu_i = zs_i; \qquad \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \tag{2.18}$$

For example, applying Equation 2.18 with $\Phi(z) = 5\%$ of the area under a standard normal distribution yields $\Delta\mu$ = -0.1257 s_i. Transforming this value to each modal distribution of the original target spectrum, S, results in the new mean frequencies shown in Table I. The resulting spectra S_1, S_2 and S_3 are target spectra with progressively greater shifts in mean-modal frequencies. In addition to this mean change, the variance of each mode of each target spectrum is increased from the orig-

inal by 10% while holding the total energy of the spectra constant. Figure 3 shows the original target spectrum, S, with one of the modified spectra, S_1. The resulting magnitudes of each frequency shift is particulary convenient to illustrate calculation of the amount of data required for the new method to attain a pre-determined level of effectiveness. Specifically, to the extent the Gaussian assumption is not violated, the resulting confidence levels should be a function of only the number of cycles, without regard for vibrational mode with which these cycles are associated.

Table I. Target Spectra Parameters with Percent Change from Original Idealized Spectrum

Target Spectrum	1st Mode			
	Mean	Change	Variance	Change
S	0.12500	N/A	0.00060	N/A
S_1	0.12192	-2.46%	0.00066	10%
S_2	0.12037	-3.70%	0.00066	10%
S_3	0.11879	-4.96%	0.00066	10%
Target Spectrum	2nd Mode			
	Mean	Change	Variance	Change
S	0.28500	N/A	0.00120	N/A
S_1	0.28064	-1.53%	0.00132	10%
S_2	0.27845	-2.30%	0.00132	10%
S_3	0.27622	-3.05%	0.00132	10%

c. Minimum Required Time-history Duration

After the modified target spectrum has been developed, the new S_1 and original S are used to simulate a total of fifty sets of time-histories. The modal distribution method is then applied to each resulting twenty-five pairs of time-histories to calculate a P-

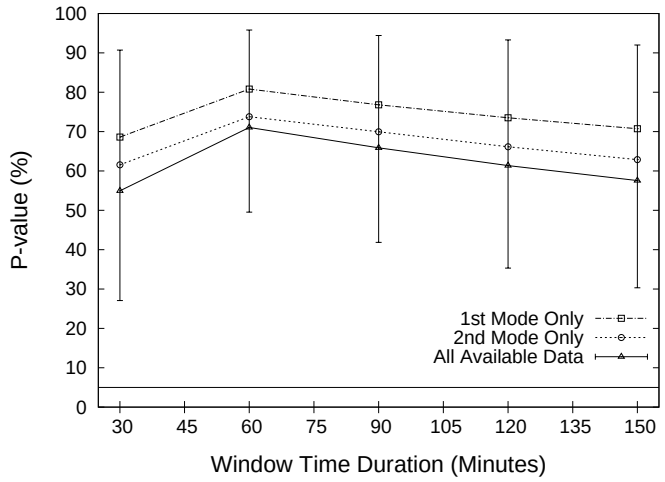


Fig. 4. P-value that the Observed Differences Between the Time Series Are Due To Differences Between Underlying Spectra When in Fact No Such Differences Exist

value for each. These results are averaged and presented as Figure 5. The monotonic decrease shown in Figure 5 shows greater confidence levels (smaller P-values) are associated with longer time-histories. Confidence levels of 95% (5% P-values) are often considered to indicate statistical significance. The upper two lines on the figure indicate confidence levels based on each of the first two modes. The lowest line on the figure indicates the method's capability to detect the change in the underlying spectrum using all available data, which in this case means both modes. The very subtle shift shown in Figure 3 can be detected with about 95% confidence (5% P-value) using only 30 minutes of data.

Figure 6 shows the probability of failing to detect the known shift in modal frequencies from Figure 3. These probabilities are calculated by treating the P-value as a random variable (Equation 2.17). Using only 30 minutes of data, there is about an 18% chance the method will not detect the change in the underlying spectrum with a P-value of 5% or less. However, given 90 minutes of data, the method is almost certain to detect even the very subtle change indicated in Figure 3.

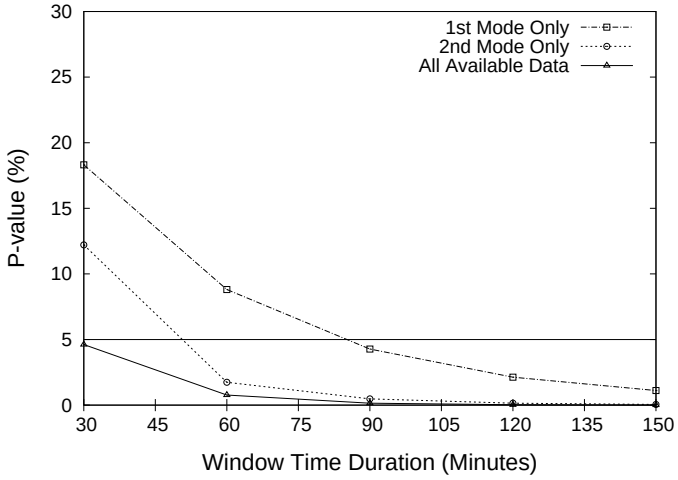


Fig. 5. Ability to Detect Underlying Changes in the Target Spectrum as a Function of Time-history Duration (Smaller P-values Indicates Greater Ability)

d. Minimum Required Number of Cycles

The main factor for determining the necessary data set length is in fact the number of vibrational cycles rather than the temporal duration. Systems with higher frequencies or with more vibrational modes have more cycles in a fixed duration. The data of Figure 5 is again presented as Figure 7, but with the x-axis presented as cycles. The number of cycles is estimated as the simulation duration times the sum of mean modal frequencies. The fact that the P-value predictions based on the first mode, second mode and combination nearly collapse to a single line when plotted against cycles is a property of the relative shift of each of the modal frequencies corresponding to a fixed percent of an ideal Gaussian distribution (Equation 2.18) and not of the method itself.

Figure 7 shows that approximately 1,500 cycles are needed to obtain a statistically significant detection of the very subtle shift between the S and S_1 power spectra. The figure shows results both for an ideal noise-free signal and for a signal containing a realistic 1.9 mg of white noise. The added noise has only minimal effect on the

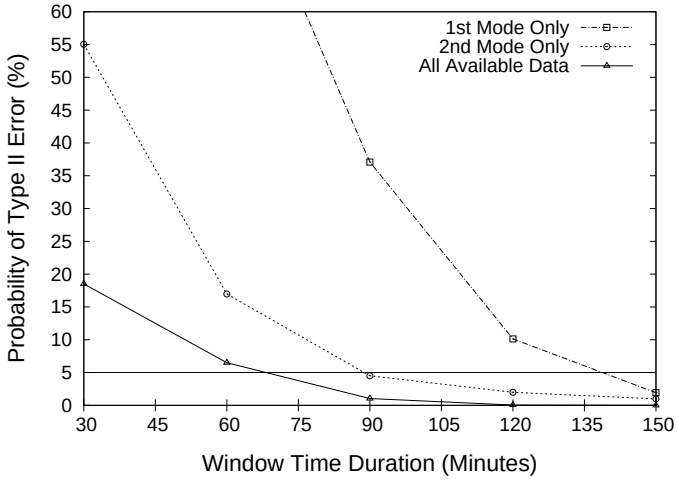


Fig. 6. Probability of Failing to Detect a Known Difference Between Underlying Spectra as a Function of Time-history Duration

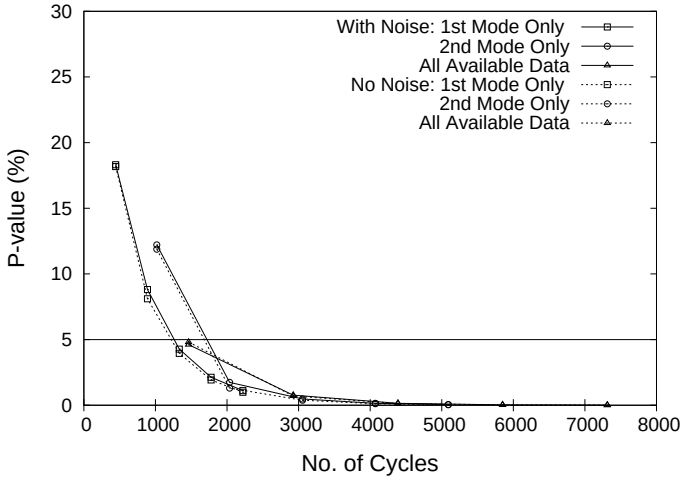


Fig. 7. Ability to Detect Underlying Changes in the Target Spectrum as a Function of Number of Cycles With and Without Noise

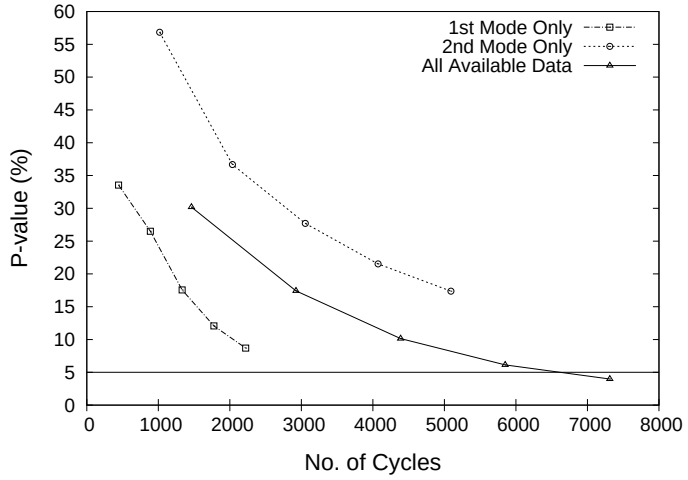


Fig. 8. Very High Noise Case (19 mg): Detection of Underlying Changes in the Target Spectrum With Gaussian Assumption Strongly Violated

P-values predicted for each mode and the overall comparison. Results for each mode and combined results all tend to collapse toward a single line as would be expected for a Gaussian distribution. Inclusion of the noise creates only a slight change in computed non-Gaussianity for each mode: average skewness for both modes is reduced from 0.082 to 0.067 and average kurtosis is increased from 3.10 to 3.15, compared with the ideal Gaussian values of 0 and 3. In Figure 8, however, results do not collapse to a line, which qualitatively shows the effect of strong violations of the Gaussian assumption. The level of noise included here is 19 mg, resulting in average skewness of -0.52 and kurtosis of 4.51. The T-statistic is clearly not valid for this strong violation of the Gaussian assumption. Returning to the low-noise (1.9 mg) case, Figure 9 shows that at least 3,000 cycles are needed to obtain a 5% probability of failing to detect such a shift, based on the gamma distribution. Similar results are expected for predictions using three or more modes.

Thus far, the example has sought to detect extremely subtle shifts in the underlying power spectra and detecting these shifts has required considerable time durations.

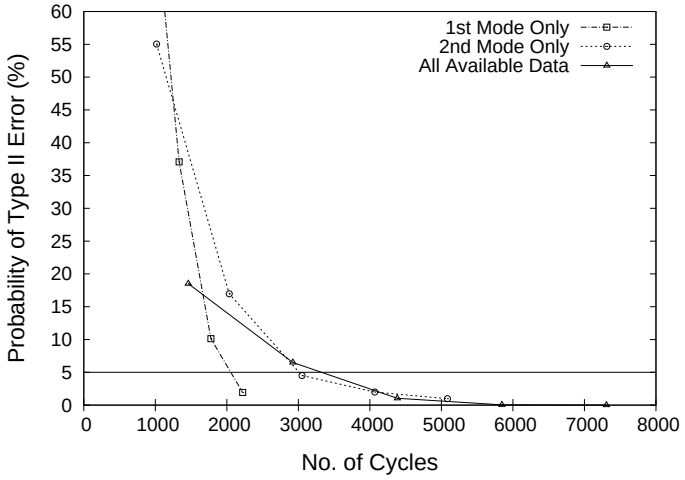


Fig. 9. Probability of Failing to Detect a Known Difference Between Underlying Spectra as a Function of Number of Cycles

Detecting larger changes between underlying spectra requires less data. Figure 10 shows the amount of data required to detect increasingly large shifts in modal frequencies. All shifts are made to both modes in accordance with Equation 2.18; the details of the underlying target spectra are shown in Table I. The curve corresponding to shifts in mean modal frequencies corresponding to 5% of the area under the standard normal is fit through the same data as presented in Figure 7. The very subtle shift corresponding to 5% of the area under a normal distribution requires around 1,400 cycles to detect at a 5% P-value; if the size of the modal shifts are approximately doubled, the required number of cycles drops to around 400 cycles.

3. Comparison with Conventional Tests

The newly proposed method is compared with two conventional tests of statistical stationarity: the reverse arrangements test and the runs test e.g. [27] and [28]. Numerous pairs of 150 minute time-series were used to test these conventional methods. These tests were made for both Type I and Type II errors using the same data as

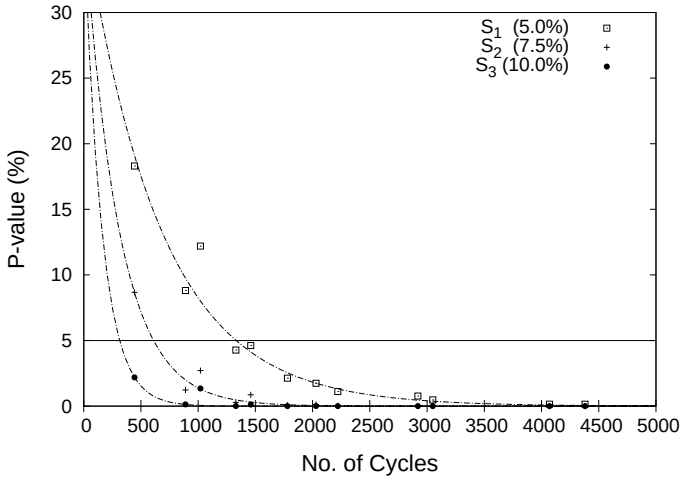


Fig. 10. Number of Cycles Required to Detect Increasingly Large Shifts in Mean Modal Frequencies. Percents Refer to Fractions of the Area Under a Standard Normal Distribution.

previously used to verify the new method. Neither of these conventional methods was found to be able to meaningfully detect the change in the target spectrum

The reverse arrangements test never detected a statistical change in the time-series, regardless of whether or not there is a change between underlying spectra. Results from the runs test varied with the number of segments into which the time-series was divided as part of the test, but in no case were the results useful. When 40 segments were used, the probability of a Type I error (false alarm) was found to be 24% and the probability of a Type II error 76%. Using 100 segments, the probability of a Type I error was found to be 52% and probability of a Type II error 44%. Recall that the new modal distribution method applied to these same time-series yielded a probability of a Type I error and of a Type II error near zero (Figures 4 and 6).

F. Conclusions

A newly proposed modal distribution method of quantifying the significance of subtle changes in modal vibrations based on power spectra has been presented. The new method is very effective at detecting changes in mean frequencies of individual modal vibrations (Figure 5–10), even if the data is simulated from underlying target spectra with only very subtle differences (Figure 3). The method is also very robust against false-alarms. The amount of data required decreases considerably if larger changes in underlying spectra are to be detected (Figure 10). The new method is considerably more effective than the conventional runs and reverse arrangements tests.

The modal distribution method is based on transforming individual modes of a power spectrum into a statistical distribution, which enables application of conventional statistical tools. Response modes containing frequencies of dominant excitation can be explicitly excluded from the analysis. After transformation into probability space, each modal response is treated as an independent statistical distribution for which the spectral moments are directly calculated from geometry. The mean and variance are calculated independently for each individual mode of each of two segments of the time-history and then used in a statistical comparison. Combining differences between the means of individual modal frequency pairs results in an overall quantitative significance level of the difference between power spectra. An example is presented in which minimum data requirements for the new method investigated in detail, and a method is outlined to determine minimum data requirements for new applications.

The newly proposed modal distribution method is general and is applicable to any number of modes of vibration. The development of the method presented here makes use of the T-statistic, so this presentation is directly applicable to only well-

separated modes with a generally Gaussian shape and relatively low levels of noise in the signal. In concept, any distribution could be used in place of the Student-T. The new method may find a broad variety of applications, though it seems particularly well suited for structural health monitoring because detailed knowledge of the excitation is not required as input.

CHAPTER III

THE HERMITE MOMENT MODEL FOR HIGHLY SKEWED RESPONSE WITH APPLICATION TO TENSION LEG PLATFORMS

A new addition to the statistical Hermite moment model of extremes is introduced for use on processes with high skewness and near-Gaussian kurtosis. The monotone limits of the existing model are expressed as ellipses in response moment space and a new methodology is introduced that combines hardening and softening models to overcome these limits. The result is that any fractile of a distribution described by its first four statistical moments can be transformed to or from the Gaussian, subject only to a theoretical orthogonality limit. An example application to a Tension Leg Platform is presented.

A. Introduction

Many irregular time-series processes exhibit non-Gaussian behavior. In some specific cases, response of systems subject to irregular excitation have been shown to be well represented by a Gaussian random process. However, other structural systems often have response processes with meaningfully larger or smaller extreme values than would be predicted directly from Gaussian theory. This non-Gaussian behavior may result from nonlinear structural response characteristics, nonlinear excitation loads, or both. For such systems, direct use of Gaussian statistics is not appropriate. The most rigorous alternative is to predict the characteristics of a non-Gaussian response using numerical time stepping direct integration methods. These methods generally include development of a full probability distribution for use in prediction of upcrossing rates, extreme values and fatigue damage; unfortunately, these methods require extensive data sets and heavy computation.

Less computationally intensive alternatives to direct integration methods are also available, including various approximation approaches such as equivalent linearization techniques and series distribution methods. Equivalent linearization techniques [29, 30, 31] are only useful to estimate gross response statistics such as root mean square levels. Series distribution methods [32, 33] can be used to transform known statistical results such as mean upcrossing rates and extreme values of a Gaussian process into those of a non-Gaussian process by finding a simple functional transformation of the equivalent Gaussian statistics.

Full probability distributions of a non-Gaussian process can be estimated using Gram-Charlier series, Edgeworth series, and Longuet-Higgins series from those of a Gaussian process e.g. [34, 32, 35]. These series distribution methods have a common weakness: they all tend to exhibit oscillating and negative tail behavior. Overcoming this weakness, the Hermite moment model was developed and proven more flexible over a wider range of skewness and kurtosis [36, 37]. This Hermite moment model offers equivalent fractile mapping between Gaussian and non-Gaussian processes and has been widely applied to a variety of areas: non-Gaussian excitation and structural response estimation e.g. [38], extreme response estimation e.g. [39], non-Gaussian wave kinematics estimation e.g. [40], and non-Gaussian simulation e.g. [41]. This paper presents a review of the theory underlying the Hermite moment model with special emphasis on two limitations. First, a theoretical limitation is imposed by the requirement for orthogonality of individual terms of the Hermite polynomial expansion underlying the model. Second, a computational limit arises for the transformation to have a simple functional form, and this limit prevents use of the transformation on processes with high skewness but near-Gaussian kurtosis.

B. Background: The Four-moment Hermite Model

The four-moment Hermite model uses skewness and kurtosis to transform between standardized Gaussian and equivalent non-Gaussian fractiles. This paper addresses both forward and backward transformations: the forward transformation maps fractiles of a Gaussian process to those of a non-Gaussian, and the backward transformation maps fractiles of a non-Gaussian process to Gaussian equivalents. Additionally, any non-Gaussian processes can be softening, hardening, or a combination of the two. The probability distribution of a softening process has thicker tails than a Gaussian distribution. An example softening process would be the irregular dynamic response of a spring-mass system to a Gaussian excitation if the spring is non-linear, with stiffness that decreases for greater positive displacement (softens). Such a system would have relatively greater maxima and smaller minima than would an equivalent system with a linear spring. Conversely, the probability distribution of a hardening process has thinner tails than the Gaussian distribution.

The four-moment Hermite model for transformation of a process that is relatively thin-tailed to a process that is thicker-tailed over the entire fractile range is a cubic polynomial. The polynomial applies to either forward transformation of a softening process (from Gaussian to non-Gaussian) or backward transformation of a hardening process (from non-Gaussian to Gaussian). Transformations from processes with relatively thick-tailed distributions to thinner-tailed distributions require inversion of the cubic polynomial. The model developed by Winterstein [36, 37] has two limitations: 1) the orthogonality limit, which is innate to the underlying theory and 2) the monotone limit, which enables a simple inversion of the cubic polynomial over the entire fractile range. This paper provides an alternative inversion methodology that is not subject to the monotone limit.

1. The Functional Form of Transformation

The Hermite moment model transforms any fractile between standardized Gaussian and non-Gaussian processes, such that the cumulative density function (CDF) of a standard non-Gaussian process matches that of the standard Gaussian. The model for mapping a relatively thin-tailed process to a thicker-tailed process has a simple form resulting from an infinite series polynomial expansion. The first few terms of the expansion can be conveniently and appropriately applied to transform between specific fractiles.

$$x = a_0 + a_1 P_1(u) + a_2 P_2(u) + a_3 P_3(u) + ... \tag{3.1}$$

$$u = b_0 + b_1 P_1(x) + b_2 P_2(x) + b_3 P_3(x) + ... \tag{3.2}$$

where x and u are values of random variables X and U. Standardized random variables are used throughout this paper: $X = \frac{Y-\mu_Y}{\sigma_Y}$ and $U = \frac{V-\mu_V}{\sigma_V}$ represent standard non-Gaussian and standard Gaussian processes, respectively; argument t is omitted for brevity. μ_Y, σ_Y, α_3 and α_4 are the mean, standard deviation, skewness and kurtosis of a non-Gaussian process Y. μ_V and σ_V are the mean and standard deviation of a Gaussian process V.

Winterstein [36, 37] provides an infinite-series polynomial expansion and a representation of the functional transformation with four terms, whose coefficients are directly related to the response moments of the non-Gaussian process. A non-Gaussian softening process can be expressed by a Hermite polynomial series expansion of the Gaussian process:

$$x = \kappa\Big\{u + \sum_{n=3}^{N=\infty} h_n \mathrm{He}_{n-1}(u)\Big\} \tag{3.3}$$

$$\simeq \kappa\Big\{u + h_3(u^2-1)\Big) + h_4\Big(u^3 - 3u)\Big\} \tag{3.4}$$

where κ is a scale factor that ensures x has unit variance. He_n is the n^{th} Hermite polynomial: $\mathrm{He}_1(\xi) = \xi$, $\mathrm{He}_2(\xi) = \xi^2 - 1$, $\mathrm{He}_3(\xi) = \xi^3 - 3\xi$, etc., where ξ is any value. Coefficients h_3 and h_4 are directly calculated from the skewness and kurtosis of the non-Gaussian process. Similarly, a Gaussian response can be expressed by a Hermite polynomial series expansion of a hardening response process:

$$u = x - \sum_{n=3}^{\infty} h_n \mathrm{He}_{n-1}(x) \tag{3.5}$$

$$\simeq x - h_3(x^2 - 1) - h_4(x^3 - 3x) \tag{3.6}$$

In both transformations, the Hermite polynomial coefficients h_n give shape to the resulting distributions. Taking only first-order terms of the expansion yields direct relationships between the first n Hermite coefficients and the first n statistical moments of the standardized non-Gaussian variable [36, 37].

$$h_1 = h_2 = 0, \qquad h_3 = \frac{\alpha_3}{6}, \qquad h_4 = \frac{\alpha_4 - 3}{24} \tag{3.7}$$

Truncating the infinite expansion at the first four response moments makes the implicit assumption that terms higher than h_4 are negligible. In practice, use of response moments higher than the kurtosis is difficult because of high variability in sampling these higher moments. In addition to these first-order coefficients, Winterstein developed second-order coefficients for the forward softening case by taking all the first- and second-order terms in the expansion [36, 37]. Other second-order representations are proposed in [42, 43]. In later work, Winterstein and Lang [44, 45] proposed computational methods to find optimized coefficients by minimizing the sum of squared errors.

2. Inversion

Equations 3.4 and 3.6 directly map thin- to thick-tailed distributions; mapping from thick- to thin-tailed distributions requires inversion of the polynomial transformation. Winterstein does not provide an inversion of Equation 3.6; he does provide an inversion of Equation 3.4 which is applicable only inside the monotone limits (Section C):

$$u = \left[\sqrt{\xi^2(x)+c}+\xi(x)\right]^{1/3} - \left[\sqrt{\xi^2(x)+c}-\xi(x)\right]^{1/3} - a \tag{3.8}$$

$$\text{where} \quad \xi(x) = 1.5b\left(a+\frac{x}{\kappa}\right) - a^3$$

$$a = \frac{h_3}{3h_4}, \qquad b = \frac{1}{3h_4}, \qquad c = \left(b-1-a^2\right)^3$$

In addition to the monotone limit, a square root in Equation 3.8 requires that $\xi^2(x)+c$ be non-negative.

Figure 11 shows the various regions where each of the transformation pairs is applicable. The curved lines are the orthogonality and monotone limits, which divide space into various regions. In regions I and II the original polynomials are applicable (Equations 3.4 and 3.6). The existing inversion was only applicable in region I (Equation 3.8). Here, a new inversion is offered for use in region II as is a new methodology for use in regions III and IV. The new method is applicable to either forward or backward transformations. Outside the orthogonality limit, use of the Hermite model is not theoretically justified (I_X, II_X, III_X, and IV_X).

3. Limitations of the Original Hermite Model of Extremes

The parabolic curve shown with a solid line on Figure 11 represents the orthogonality limit.

$$\alpha_4 \geq \alpha_3^2 + 1 \tag{3.9}$$

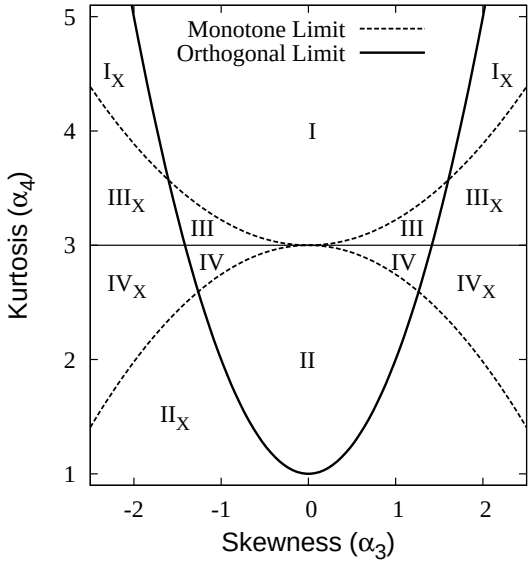


Fig. 11. Various Regions of Model Applicability

Outside the limit (all regions with subscript X), the transformations are generally smooth and well-behaved, but the underlying theory is violated and use of the model is on weak theoretical ground. The original development by Winterstein [36] includes a cogent explanation, some of his most salient points are closely paraphrased here:

> Various results of any random process $Z(t)$ can be expressed as a sum of n polynomials, $P_n[Z(t)]$, that possess orthogonal properties:
>
> $$E[P_n(Z)] = E[P_n(Z)P_m(Z)] = 0 \qquad (n \neq m; \quad n, m > 0) \tag{3.10}$$
>
> The n^{th} such polynomial can be constructed from the simple power law

$Z^n(t)$ by removing its correlation with all lower order polynomials:

$$P_n(Z) = Z^n - \sum_{k=0}^{n-1} c_{nk} P_k(Z); \qquad c_{nk} = \frac{E[Z^n P_k(Z)]}{E[P_k^2(Z)]} \tag{3.11}$$

If $Z(t)$ has been standardized to have zero mean and unit variance, applying Equation 3.11 with $P_0(Z) = 1$ leads to the polynomials

$$P_1(Z) = Z; \qquad P_2(Z) = Z^2 - \alpha_3 Z - 1; \tag{3.12}$$

$$P_3(Z) = Z^3 - c_{32} P_2(Z) - \alpha_4 Z - \alpha_3; \qquad ... \tag{3.13}$$

in which $c_{32} = E[Z^3 P_2(Z)]/E[P_2^2(Z)] = (\alpha_5 - \alpha_3 - \alpha_3\alpha_4)/(\alpha_4 - \alpha_3^2 - 1)$. Noting that the $E[P_2^2(Z)] = \alpha_4 - \alpha_3^2 - 1$ cannot be negative, the kurtosis value, α_4 cannot be less than $\alpha_3^2 + 1$.

Winterstein goes on to show that various generalizations of this model can be formed from these orthogonal polynomials, including the ability to model a new random process from only its first N moments; one such example is the Hermite model of extremes. As the distribution of a non-Gaussian process approaches a Gaussian through the infinite-series approximation, the orthogonality condition of Equation 3.10 is used to ensure several important properties. Most importantly, 1) it ensures that the polynomial moments are the central moments of the process and 2) that polynomials of order higher than N are equal to zero or at least negligible. The orthogonality of the polynomials also guarantees uniqueness of the Hermite coefficients. Thus, the orthogonality limit (Equation 3.9) is a hard theoretical limit imposed by use of the Hermite polynomial. In the next section, the monotone limits are fully explained and represented geometrically as ellipses in moment space. Alternate solutions are then offered that yield results identical to those of Winterstein in the monotone regions, and also enable application of the model over the entire range

within the orthogonality limit, i.e., the entire range for which the underlying theory is valid.

C. Overcoming The Monotone Limitation

One of the principal contributions of this paper is an alternative implementation of the Hermite polynomial that overcomes the monotone limitation. First, the monotone limits are investigated numerically and represented geometrically, then the results are used to develop a new inversion and a new alternative methodology to construct monotonic transformations for non-monotone regions.

1. Understanding the Monotone Limits

Figure 12 is developed later in this section. The horizontal and vertical axes are the coefficients of the Hermite polynomials: $h_3 = \alpha_3/6$, and $h_4 = (\alpha_4 - 3)/24$. The Hermite polynomial transformation is shown to be monotone inside each ellipse and non-monotone outside. The regions in Figure 11 depict the five possible cases:

I. The two coefficients, h_3 and h_4, are inside the upper ellipse ($D_s \leq 0$), which indicates the transformation for softening response increases monotonically.

II. The two coefficients are inside the lower ellipse ($D_h \leq 0$), which indicates the transformation for hardening response increases monotonically.

III. The two coefficients are outside the upper ellipse as well as above the horizontal axis ($h_4 > 0$ and $D_s > 0$). For this region, the transformation for softening response increases continuously only for fractiles beyond either of the two critical points (Equation 3.15) and decreases continuously between. The transformation for hardening response increases continuously for fractiles between the two critical points (Equation 3.19) and decreases continuously beyond.

IV. The two coefficients are outside the lower ellipse as well as below the horizontal axis ($h_4 < 0$ and $D_h > 0$). For this region, the transformation for hardening response increases continuously only for fractiles beyond either of the two critical points (Equation 3.19) and decreases continuously between. The transformation for softening response increases continuously for fractiles between the two critical points (Equation 3.15) and decreases continuously beyond.

V. The coefficients lie on the line $h_4 = 0$. The process is non-monotonic and the hermite polynomial reduces to a quadratic. If $h_3 = 0$ then the process is Gaussian.

Three solutions for cases in regions I and II were developed in the original work by Winterstein. A missing inversion for the monotone backward hardening Case (II) is offered here, as are solutions for Cases III and IV (above the orthogonality limit but outside the ovals).

a. The Monotone Limit for a Softening Response (Upper Ellipse)

In the model for softening response, the non-Gaussian variable is expressed as a cubic function of the standard Gaussian variable. The slope of Equation 3.4 is

$$\frac{dx}{du} = \kappa\Big\{3h_4u^2 + 2h_3u + (1 - 3h_4)\Big\} \tag{3.14}$$

which is a quadratic function. The solutions of the quadratic equation $\frac{dx}{du} = 0$, if any, are the critical points of the cubic function, and play a key role in determining the monotone range. The resulting critical points are

$$u_c = \frac{-2h_3 \pm \sqrt{4h_3^2 - 12h_4(1 - 3h_4)}}{6h_4} \tag{3.15}$$

The existence of these solutions is determined by the discriminant, D_s for the softening

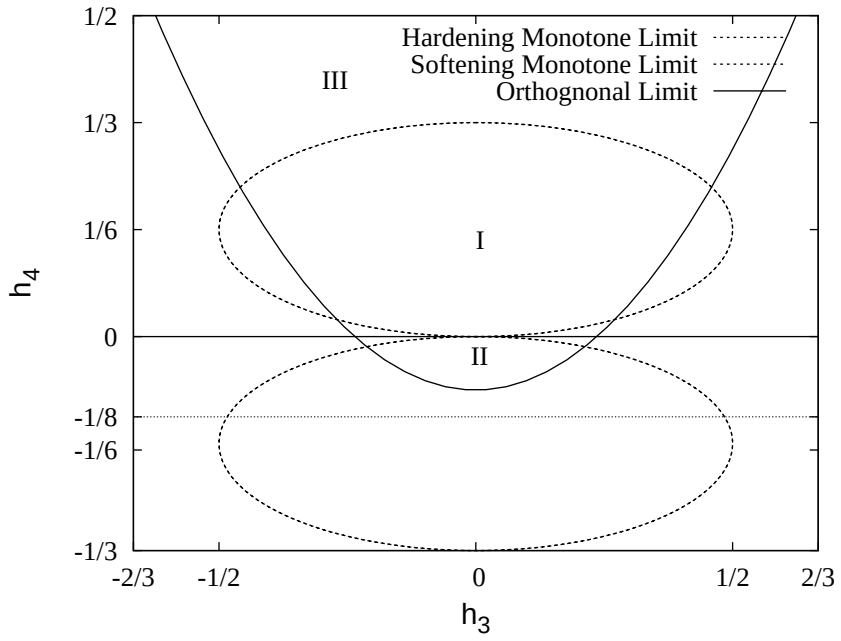


Fig. 12. Orthogonal, Softening Monotone and Hardening Monotone Limits in Hermite Moment Space

case, which is equivalent to the equation inside square root in Equation 3.15.

$$D_s = 4\Big[h_3^2 - 3h_4(1 - 3h_4)\Big] \tag{3.16}$$

A positive discriminant produces two solutions, which implies the curve is non-monotonic, with transitions between the increasing and decreasing parts of the curve at the critical points. Zero discriminant produces one solution (two identical solutions), and a negative discriminant produces no real solution. Since no solution or a single solution to the quadratic equation indicates no critical point in the cubic

function, the resulting cubic function with non-positive discriminant is monotone if

$$h_3^2 \leq 3h_4(1 - 3h_4) \tag{3.17}$$

The zero discriminant case, $D_s = 0$, can be reorganized to represent the upper ellipse in (h_3, h_4) plane of Figure 12.

$$\frac{h_3^2}{(\frac{1}{2})^2} + \frac{(h_4 - \frac{1}{6})^2}{(\frac{1}{6})^2} = 1 \tag{3.18}$$

b. The Monotone Limit for a Hardening Response (Lower Ellipse)

Similarly, the standard Gaussian variable can be expressed as a cubic function of a standardized non-Gaussian variable in the model for hardening response. The same procedure of softening response using Equation 3.6 results in two critical points, and the discriminant and monotone limit for the hardening case:

$$x_c = \frac{-2h_3 \pm \sqrt{4h_3^2 + 12h_4(1 + 3h_4)}}{6h_4} \tag{3.19}$$

$$D_h = 4[h_3^2 + 3h_4(1 + 3h_4)] \tag{3.20}$$

$$h_3^2 \leq -3h_4(1 + 3h_4) \tag{3.21}$$

Setting $D_h = 0$ in Equation 3.20 and reorganizing defines the lower ellipse in Figure 12.

$$\frac{h_3^2}{(\frac{1}{2})^2} + \frac{(h_4 + \frac{1}{6})^2}{(\frac{1}{6})^2} = 1 \tag{3.22}$$

Values of $h_4 < -\frac{1}{8}$ are not possible in application of the model (Figure 12) because kurtosis α_4 is always non-negative in Equation 3.7.

2. An Alternative Inversion of the Hermite Polynomial

Here, a new inversion methodology of the original Hermite polynomials (Equations 3.4 and 3.6) is offered to enable use of the Hermite model of extreme values in the Cases III and IV, where the polynomial is orthogonal but non-monotone. The orthogonality limit described by Winterstein (Equation 3.9) still remains. This new inversion has an important advantage over the original: it enables inversion outside the monotone limit, where multiple solutions are possible. Choosing between these solutions is addressed in Section 3. To develop the alternate inversion, the forward transformation for a softening response (Equation 3.4) is rearranged to a more convenient form:

$$\kappa h_4\Big(u^3 + a_2u^2 + a_1u + a_0\Big) = 0 \tag{3.23}$$

where $a_0 = \frac{-x-\kappa h_3}{\kappa h_4}$, $a_1 = \frac{1-3h_4}{h_4}$, $a_2 = \frac{h_3}{h_4}$ and x is the value of a non-Gaussian response to be transformed to the Gaussian. The Gaussian equivalent of the specified x is a value of u solving the cubic equation. Among many existing solution approaches, the most practical here is to find its real roots through use of a trigonometric identity [46, 47]. In this solution, the real roots are selected based on computed parameters p and C.

$$p = \frac{3a_1 - a_2^2}{3}; \qquad C = \frac{q}{2}\left(\frac{3}{|p|}\right)^{\frac{3}{2}} \tag{3.24}$$

where $q = (9a_1a_2 - 27a_0 - 2a_2^3)/27$.

(1) if $p \geq 0$, the equation is monotone and this alternate inversion yields numerical results identical to those of Winterstein (Equation 3.8), but without restrictions imposed by having a square root of a potentially negative quantity in the solution.

$$u \;=\; -\frac{1}{3}a_2 + 2\sqrt{\frac{p}{3}}\sinh\left(\frac{1}{3}\sinh^{-1}C\right) \tag{3.25}$$

(2) if $p < 0$, the equation is non-monotone and the root selection depends on the value of C.

$$u = -\frac{1}{3}a_2 + 2\sqrt{\frac{|p|}{3}}\cosh\left(\frac{1}{3}\cosh^{-1} C\right) \quad \text{for} \quad C \geq 1 \tag{3.26}$$

$$u = -\frac{1}{3}a_2 - 2\sqrt{\frac{|p|}{3}}\cosh\left(\frac{1}{3}\cosh^{-1} |C|\right) \quad \text{for} \quad C \leq -1 \tag{3.27}$$

$$u = -\frac{1}{3}a_2 + 2\sqrt{\frac{|p|}{3}}\cos\left(\frac{1}{3}\cos^{-1} C\right) \quad \text{for} \quad |C| < 1 \tag{3.28}$$

If the kurtosis of a process is exactly 3, then Equation 3.4 reduces to a quadratic, and $h_4 = 0$, which creates a divide by zero in the cubic inversion. The inversion for the quadratic is:

$$u = \frac{-1 \pm \sqrt{1 + 4h_3(h_3 + x/\kappa)}}{2h_3} \tag{3.29}$$

which increases monotonically if $(\pm)$ is taken as $(+)$ for $h_3 > 0$ and $(-)$ for $h_3 < 0$.

The inversions for the backward transformation for a hardening response is developed equivalently by rearranging Equation 3.6.

$$h_4\Big(x^3 + a_2x^2 + a_1x + a_0\Big) = 0 \tag{3.30}$$

where $a_0 = \frac{-u-h_3}{h_4}$, $a_1 = \frac{1-3h_4}{h_4}$, and $a_2 = \frac{h_3}{h_4}$. Equation 3.30 is in exactly the same form as Equation 3.23, so the same solution technique is applied to invert the backward transformation of the hardening response, and the results are nearly identical.

(1) If $p \geq 0$, then the transformation is monotone:

$$x = -\frac{1}{3}a_2 + 2\sqrt{\frac{p}{3}}\sinh\left(\frac{1}{3}\sinh^{-1} C\right) \tag{3.31}$$

(2) if $p < 0$, the equation is non-monotone and the root selection depends on the

value of C.

$$x = -\frac{1}{3}a_2 + 2\sqrt{\frac{|p|}{3}}\cosh\left(\frac{1}{3}\cosh^{-1} C\right) \quad \text{for} \quad C \geq 1 \tag{3.32}$$

$$x = -\frac{1}{3}a_2 - 2\sqrt{\frac{|p|}{3}}\cosh\left(\frac{1}{3}\cosh^{-1} |C|\right) \quad \text{for} \quad C \leq -1 \tag{3.33}$$

$$x = -\frac{1}{3}a_2 + 2\sqrt{\frac{|p|}{3}}\cos\left(\frac{1}{3}\cos^{-1} C\right) \quad \text{for} \quad |C| < 1 \tag{3.34}$$

If the kurtosis is exactly 3, the quadratic inversion is:

$$x = \frac{1 \pm \sqrt{1 + 4h_3(u - h_3)}}{2h_3} \tag{3.35}$$

which increases monotonically if $(\pm)$ is taken as $(-)$ for $h_3 > 0$ and $(+)$ for $h_3 < 0$.

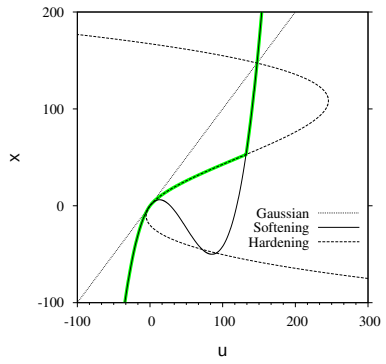
3. Piecewise Construction of a Monotone Transformation

In this section, a method is outlined to construct a monotonic transformation by combining the results of non-monotonic hardening and softening transformations over their individual regions of applicability. If two continuous CDF's exist, then there must also exist a final combined transformation that is monotonic at all fractiles, because there must exist a one-to-one mapping between the two distributions such that for any value of u there is exactly one equivalent value of x, and vice versa. The polynomials are non-monotone in regions III and IV, so direct inversion would result in multi-valued transformations at some fractiles. In the cumulative density function (CDF), constant fractiles appear as horizontal lines; a transformation between the Gaussian and non-Gaussian at a single fractile appears as a horizontal shift along the constant fractile line. If the transformation is multi-valued, the implication is that there exist multiple crossings of a single fractile of the CDF. The only way multiple crossings can occur is if there is a downward slope in the CDF. By definition, probabilities must be non-negative, so transformations must be monotonic.

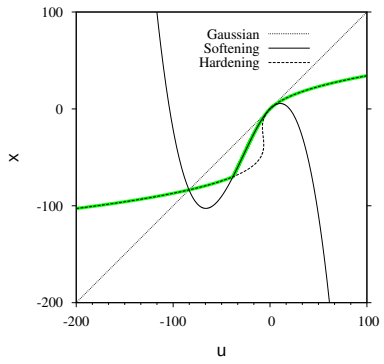
The transformation for monotone cases is either hardening or softening over the entire distribution, and either the forward or backward transformation is a cubic polynomial. Non-monotone cases, however, are hardening over some fractiles of the distribution and softening over others. At extreme fractiles (very high or very low), the cubic transformation is dominated by the kurtosis because that term is cubed in Equations 3.4 and 3.6; for the transformation to be non-monotone, the skewness term must dominate nearer to the body of the distribution. Such a distribution is either hardening in the kurtosis-controlled extreme tails and softening in the skewness-dominated fractiles nearer to the body of the distribution or vise-versa. Thus, any fractile of a distribution that can be described by its first four statistical moments can be transformed either to or from a Gaussian using a cubic polynomial because that fractile must be dominated by either the skewness or kurtosis of the process.

A piecewise monotone transformation can be constructed over the entire fractile range by determining the correct fractiles, or points in the x-u transformation space, to transition between the hardening and softening models such that only regions of each model with positive slope are used. The intersection of two non-monotonic cubic polynomials is shown in Figure 13. For forward softening, x is a cubic function of u and for backward hardening, u is a cubic function of x. There are in general as many as nine points of intersection between the two polynomials. Optimal transition points can be readily selected from a plot of the two non-monotone cubic polynomials or these points can be selected numerically. For numerical selection, the forward softening model (Equation 3.4) and the backward hardening model (Section 2) are computed and the critical points for each model (Equations 3.15 and 3.19) are identified analytically. These critical points are used to detect a total of three segments on the two curves in which both x and u are increasing, i.e, $\frac{dx}{du}$ and $\frac{du}{dx}$ are both positive.

The coefficients of the cubic terms in Equations 3.4 and 3.6 are opposite in sign,



(a) Region III: Softening-Hardening-Softening Required for Piecewise Monotone (α_3 = -0.2651, α_4 = 3.0072, t = 4,400 sec)



(b) Region IV: Hardening-Softening-Hardening Required for Piecewise Monotone (α_3 = -0.2402, α_4 = 2.9885, t = 4,700 sec)

Fig. 13. Combination of Non-monotone Hardening and Softening Models into a Single Monotonic Transformation

which guarantees for any $\alpha_4 \neq 3$ that one of the models is increasing at extreme fractiles and the other is decreasing, and non-monotone cubic behavior guarantees opposite slope at more moderate fractiles. There are two overlapping regions of the three increasing segments with an intersection between those segments in each region; these increasing regions must overlap because their must exist a monotonic transformation. Theoretically, the extended model consists of three parts: softening-hardening-softening or hardening-softening-hardening, though quite frequently a two part model is sufficient (softening-hardening or hardening-softening). In practice, a very large but finite range of fractiles must be pre-specified for a numerical search, and in the event a critical point is found to be outside the range, that point is replaced by the end-point of the range. Also, the intersection between the hardening and softening

models often creates a discontinuity in the CDF, which would create a jump in the PDF. To avoid this unrealistic behavior, the combined monotonic transformation is smoothed near the intersection. After a piecewise monotonic transformation has been developed, it can be used both forwards and backwards.

In the unlikely event the kurtosis is exactly three, the cubic transformations reduce to quadratics, which are necessarily non-monotone. Monotonic transformations can be easily constructed using the methodology presented here to combine monotonic segments of Equations 3.29 and 3.6 or Equations 3.35 and 3.4.

D. Application of the Theoretical Developments

The strengths of the original development of the Hermite moment model that helped lead to its widespread use in practical applications is that it is based on a solid theoretical foundation and the resulting equations are extremely simple to implement. The expanded methodology developed in this paper is also reasonably straightforward to implement in a practical application, but not nearly as straightforward as the original, mainly because of the need to identify the points of transition between the hardening and softening models. The numerical scheme described in Section 3 has been implemented in Matlab and it is the intent of the authors to make the source code generally available. Two typical applications of the Hermite model are transformations between Gaussian and non-Gaussian equivalents.

1. Transformation of a Gaussian Fractile to Its Non-Gaussian Equivalent

In this application, the Hermite moment model is used to predict a pre-specified extreme value of a non-Gaussian process. In such an application, the skewness and kurtosis are calculated directly from the time-history and used to compute the co-

efficients for the Hermite polynomials (Equations 3.7). Next, the process is checked against the orthogonality limit (Equation 3.9) to confirm the process is within the range of applicability. The kurtosis is checked to determine if the process is hardening or softening, and the monotone limit is then checked (Equation 3.17 or 3.21). If the process is monotone and hardening, Equation 3.31 is applicable to calculate the non-Gaussian equivalent to the desired Gaussian fractile; if softening and either monotone or non-monotone, Equation 3.4 is applicable. If non-monotone hardening, a piecewise monotone transformation must be constructed as outlined in Section 3 using Equations 3.4 and 3.32–3.34. Once the transformation has been computed, it can be applied to transform any Gaussian fractile (e.g., the 100-year event) to its non-Gaussian equivalent. The model is valid at every point in the process or distribution, and so can also be used to transform a Gaussian distribution into a general reference shape CDF or PDF conforming to the first four statistical moments of the process.

2. Transformation of a Non-Gaussian Response to Its Gaussian Equivalent

In this application, the Hermite moment model is used to map an equivalent fractile of a process to enable use of traditional statistical tools such as the Student's T-test to compare means of the processes. In such an application, the skewness and kurtosis are again calculated directly from the time-history and used to compute the coefficients for the Hermite polynomials. The process is checked against the orthogonality limit to confirm the Hermite model is applicable as in Section 1. The kurtosis is checked to determine if the process is hardening or softening, and the monotone limit is then checked (Equation 3.17 or 3.21). If the process is hardening and either monotone or non-monotone, Equation 3.6 is applicable to calculate the Gaussian equivalent; if monotone and softening, Equation 3.25 is applicable. If non-monotone softening, a

piecewise monotone transformation must be constructed as outlined in Section 3 using Equations 3.6 and 3.26–3.28. Once the transformation has been developed, it can be applied to the complete distribution or to every point in the process to transform it to the Gaussian equivalent such that conventional Gaussian statistical tools can be used in place of typically less-powerful non-parametric techniques.

E. Example

The Hermite moment model is applied to simulated data resulting from a time-domain solution of a simple numerical model of a Tension Leg Platform (TLP) subject to irregular seas. A TLP is a compliant offshore structure used for production of oil and gas in deep ocean waters that has highly non-Gaussian surge response (horizontal translation in the direction of the environmental loading). The platform is vertically moored by tendons at each of its corners (Figure 14). Surge response time-histories have been simulated using a simplified 2-dimensional nonlinear numerical model, which approximates the Snorre TLP in the Norweigian sector of the North Sea. Environmental conditions were intentionally selected such that the simulated response is very near the intersection of the monotone and non-monotone regions.

1. Numerical Model and Simulation of TLP Surge Response

In the numerical model, wave, wind and current forces are applied to a single degree of freedom system including nonlinear restoring force and the equation of motion is solved in the time domain. The two main sources of non-Gaussianity in the model are 1) the non-linear mooring restoring force caused by the changing angles of the tendons with increased offset and the increased buoyancy of the hull caused by being pulled downward by the tendons, and 2) the non-Gaussian wave forcing caused by

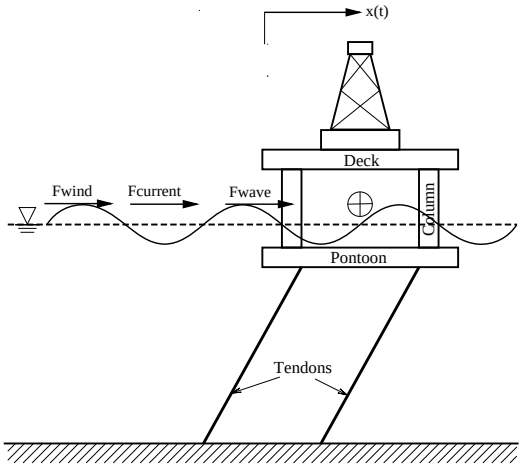


Fig. 14. Tension Leg Platform

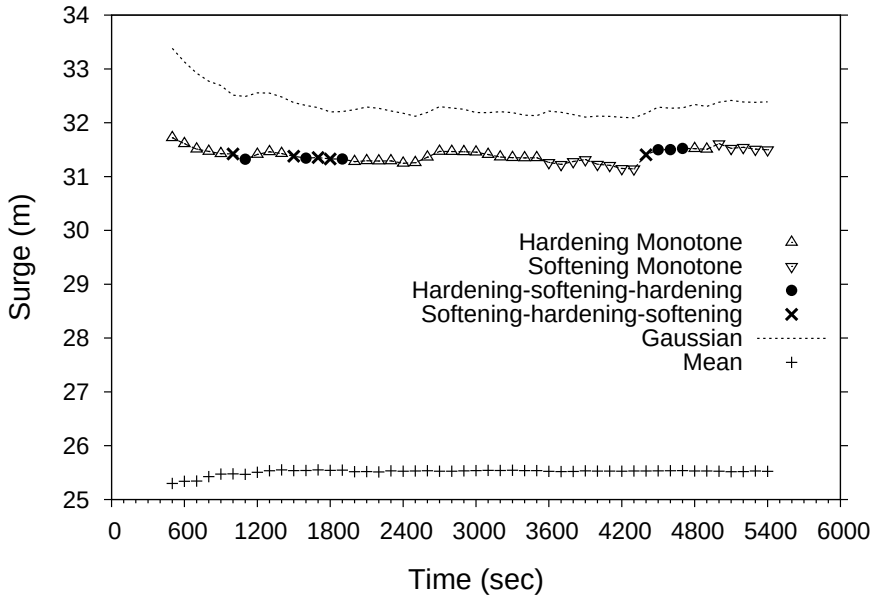


Fig. 15. Hermite Extreme Value Estimation: A Single Realization with Statistical Moments Shifting Between Monotone and Non-monotone Models Used to Predict 90 Minute Maximum.

the highly nonlinear drag term in the Morison Equation e.g., [48].

2. Application of the Hermite Model to Simulation in the Time Domain

Simulations are for a significant wave height of 10.3 m with spectral peak period of 15.5 sec. Wind force and current velocity are held constant at 3×10^7 N and -0.1 m/s, respectively. The time step of integration is 0.05 sec and total time duration for the long time-histories is 90 minutes; in this environment, the natural period is around 72 seconds. Irregularity in the response is due to the random phasing in the wave simulation from a JONSWAP spectrum e.g., [48]. The spectrum was decomposed into a large number of frequency components and wave particle kinematics were computed for each frequency component using Airy theory. Wave kinematics are combined with current velocities at each time-step and hydrodynamic forces are then estimated using the Morison Equation and a constant wind force is added. Finally, the nonlinear equation of motion is solved by time-step integration (Newmark method) [49].

The first four statistical moments of the response ($\mu_Y, \sigma_Y, \alpha_3$ and α_4) have been calculated for the beginning of the time-history up to specified cut-off times. The first window is from time, $t = 0$ until $t = 200$ seconds; the next window is from $t = 0$ until $t = 300$ seconds, with each successive window using the entire previous window plus an additional 100 seconds until the final time of 5,400 seconds (90 minutes). The resulting statistical moments are then used with the Hermite model to predict the 1.5-hour mean maximum. In this example, the kurtosis is estimated from the first M points of the time-history using the correction for small samples (e.g., [50]) in which $\alpha_4 \approx [(M+1)\hat{\alpha_4} - 3(M-1)](M-1)/[(M-2)(M-3)] + 3$ where $\hat{\alpha_4} = 1/(M\sigma_X^4)\sum_{m=1}^{M}[X(m) - \mu_X]^4$.

Figure 15 shows prediction results for the 90 minute mean maxima as the process shifts between solution cases compared with the equivalent maxima predicted using

Gaussian theory only. Table II shows the skewness, kurtosis and the behavior of the process. Variation in the moments causes the single time-history to shift between the four different cases (Table II); the relatively consistent prediction of the 1.5-hour maximum shown in Figure 15 indicates the consistency between four different solution regions of the Hermite model. Prior to this work there was no Hermite transformation available in the cubic non-monotone regions (Region III or IV in Figure 11).

Table II. Resulting Response Moments

Time(sec)	Skewness	Kurtosis	Behavior*
4400	-0.2651	3.0072	S-H-S
4700	-0.2402	2.9885	H-S-H
4900	-0.2417	2.9670	H
5000	-0.2859	3.1169	S

* S = Softening; H = Hardening

Figure 16 shows that the extended model gives consistent results near the kurtosis of 3.0, where the model changes between softening-hardening-softening (S-H-S, 4,400 sec) and hardening-softening-hardening (H-S-H, 4,700 sec,) and softening only (S, 5,000 sec). Results from the three models cannot be differentiated by eye over the entire PDF.

Figure 17 shows a histogram of probability density from a simulation case ending at 4,400 seconds (TableII) compared with two theoretical PDF's. The non-Gaussian PDF is seen to have a much better fit than the Gaussian, especially in the extreme tail, which is generally of highest importance for engineering applications. For this case, the tail appears to be a good fit down to fractiles around 0.001.

Figure 18 compares Hermite model results with extreme value predictions of the 90 minute mean maxima with mean maxima based on Monte Carlo simulations. Each

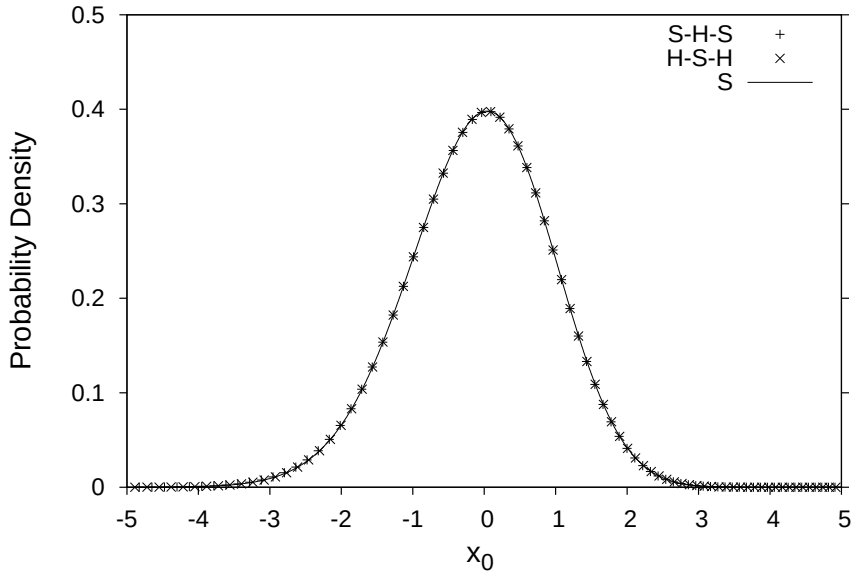


Fig. 16. PDF Comparison: Standardized Softening-Hardening-Softening, Hardening–Softening-Hardening and Monotone Responses (4,700 sec, 5,000 sec and 5,300 sec in Table II)

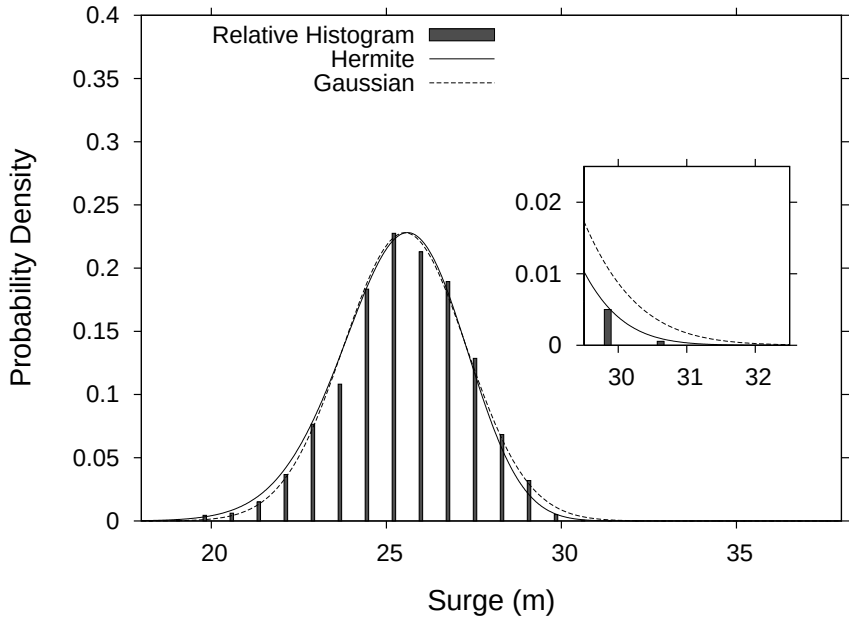


Fig. 17. Histogram of 4,100 Second Simulated Response Compared with Theoretical PDF's Based on α_3 = -0.2654, α_4 = 3.0360

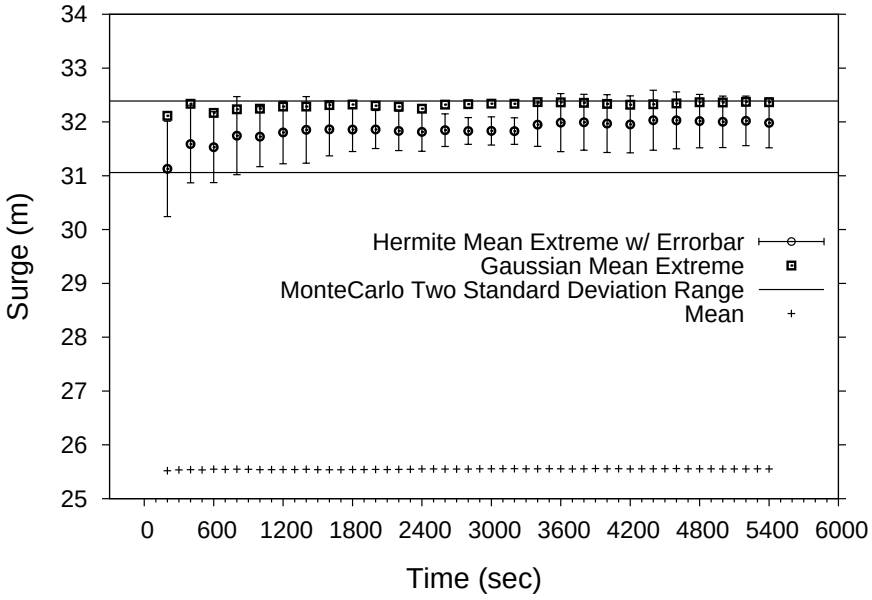


Fig. 18. Statistical Moments Based on Short Durations Used to Predict 90 Minute Maximum. All Statistics Based on 20 Simulations

data point is a prediction of the 90 minute mean-maximum based on increasing simulation times, i.e., the point at 1,200 seconds shows a prediction of the 90 minute maximum based on only 20 minutes of measured data. The irregular line with the error bars is the predicted mean-maximum based on the response moments ± one standard deviation. These bars are constructed by computing the statistical moments of each of the 20 realizations and applying the Hermite model to predict a mean-maximum for each one. The two solid lines are the average of the twenty observed 90-minute maxima ± one standard deviation. The error bars' being mainly between the horizontal lines indicates reduced variability in the predicted versus observed maxima. The consistent over-prediction of the Gaussian results compared with

Hermite and Monte Carlo results (about 0.5 meter) shows that using Gaussian results is overconservative for this negatively skewed case. Finally, the relatively steady value of the predicted mean maximum indicates that only a relatively small amount of data is needed to compute the statistical moments with sufficient accuracy to predict the expected maximum, and for these cases, only about five minutes of data is necessary for reasonably accurate predictions.

F. Discussion and Conclusions

The Hermite moment model is widely used to transform extreme fractiles between Gaussian and non-Gaussian processes. Historically, the model has been constrained by two theoretical limits: the orthogonality limit and the monotone limit. Outside the orthogonality limit (Equation 3.9) the Hermite polynomials are not guaranteed to be orthogonal, which has significant theoretical implications. Here, the monotone limit is investigated in detail and mathematically expressed as two ellipses in Hermite-coefficient space (Figure 12). An alternate solution technique is offered that overcomes the monotone limitations of the original Hermite moment model, along with one additional monotone case. The work presented here completes the original Hermite moment model for all values of skewness and kurtosis for which the underlying theory is valid, and includes mapping between Gaussian and non-Gaussian processes in either direction.

The major innovation offered in this paper is a new methodology by which the Hermite model can be applied to highly skewed cases with near-Gaussian kurtosis, which are associated with non-monotone cubic polynomial transformations (Section 3). In non-monotone regions, the new methodology combines monotonic segments of a cubic polynomial with monotonic segments of the inversion of another

cubic polynomial to create one complete monotonic transformation over the entire fractile range. For the forward model, the increasing parts of the cubic polynomial (Equation 3.4) are combined with the increasing parts of the inversion (Equations 3.32–3.34). For the backward model, the increasing parts of Equation 3.6 are combined with the increasing parts of Equations 3.26–3.28.

Additionally, a new inversion of the backward model is developed as a solution for the forward hardening monotone case (Equation 3.31). This inversion was excluded from the original work by Winterstein [36, 37]. For consistency, an equivalent inversion is also proposed for monotone softening, which yields the same result as the original formulation.

The new technique is demonstrated through application to simulated irregular motions of a tension leg platform (TLP). The simulated process is non-Gaussian and the environmental conditions were selected such that variability of sample response moments caused the response to shift between regions in which the monotone, softening-hardening-softening and hardening-softening-hardening solutions apply. Results show that the model transitions smoothly between these regions. Predictions of extreme responses are shown to be more accurate than application of a Gaussian model alone, and require less data and show less variability than averaging the observed maxima of twenty simulations.

CHAPTER IV

EFFICIENT CALCULATION OF STATISTICAL MOMENTS FOR STRUCTURAL HEALTH MONITORING

Wireless networks of smart sensors with computations distributed over multiple sensor packages have shown considerable promise in providing low-cost Structural Health Monitoring (SHM). In these networks, microprocessors are typically embedded in individual smart sensor packages. The efficiency of embedded computational algorithms is of critical importance because the size,cost, and power requirements of the sensor arrays are central concerns. Here, very efficient methodologies are presented to compute statistical moments of a measured response time-history. These moments: the mean, standard deviation, skewness and kurtosis, are often used to characterize a measured irregular response.

Two alternative approaches are presented, each of which can save substantial computer memory requirements and CPU time in certain applications. The first approach reconsiders the computational benefits of computing statistical moments by separating the data into bins and then computing the moments from the geometry of the resulting histogram. One benefit is that the statistical moment calculations can be carried out to arbitrary accuracy such that the computations can be tuned to the precision of the sensor hardware. The second approach is a new analytical methodology to combine statistical moments from individual segments of a time-history such that the resulting overall moments are those of the complete time-history. The computed moments are the same as if the segments had been concatenated prior to moment calculations.

A worked example is presented comparing two implementations of the new methodologies with conventional calculations in monitoring the global performance of an

offshore Tension Leg Platform (TLP). Accuracy, efficiency and storage requirements of the calculation methods are compared with those of conventional methods. The results show that substantial CPU and memory savings can be attained with no loss in accuracy and that more dramatic savings can be attained if a slight reduction in accuracy is acceptable.

A. Introduction

Health monitoring of either old or new structures is desirable for both economic and public safety concerns. The overall goal of these systems is to monitor structural performance such that structural integrity can be regularly assessed to detect any deficiencies before minor structural defects worsen and ultimately result in catastrophic failures. Conventional visual inspection techniques are labor intensive, time-consuming, and costly. Networks of smart sensors that include embedded computers and wireless connectivity have been proposed as a low-cost alternative inspection methodology. Extensive reviews of Structural Health Monitoring (SHM) techniques have been done by others, e.g. [8, 7, 51, 11].

A significant cost to field application of SHM systems on large civil structures can be that of installing the cables; the cabling and its various termination points may also present likely failure points for civil structures in harsh environments such as offshore structures. Finally, some types of structures with rotating components, such as the blades on wind turbines, cabling is nearly impossible. For these reasons and others, active development is ongoing for wireless sensors, often including some degree of distributed computing.

A major challenge for wireless structural monitoring systems is availability of power to run the sensor network for extended periods of time. In most applications,

wireless communication between sensors consumes more power than any other operation, which leads designers of these networks to process the raw data locally near the sensor and wirelessly transmit only the results. For example, Lynch [52] embedded damage identification algorithms into wireless sensing units to execute data interrogation. In a sensor network with distributed computing, the wireless sensing units include an embedded microcomputer and are responsible for acquiring sensor measurements, analyzing the measured data, and transmitting the results to a central server or to another sensor package in the network. The computational demands on the embedded microcomputer for real-time local data interrogation can be quite substantial. Employment of very efficient computational algorithms, such as those presented here, will enable lower cost sensors due to reduced CPU time, data storage and power requirements.

Similarly, in a field application of a significant SHM network, long-term storage of actual time-histories may be cost-prohibitive because an enormous amount of data are measured; statistical moments offer a compact way to characterize time-history data for wireless transmission and future storage. The methodologies to combine statistical moments from different segments of a time-history offer a mechanism to combine and reanalyze data to, e.g., detect long-term trends in the underlying process. Sohn et al [53] investigate novelty detection for non-Gaussian response to more accurately assess tail behavior using known extreme-value distributions. Efficient calculation of skewness and kurtosis could enable use of non-Gaussian distributions that more precisely match the tail behavior of measured data, e.g. [37].

In general, statistical moments can be used to represent the characteristics of any random data, e.g., [54, 55]. Statistical moments have found a broad range of application including: blind decomposition [56], asymptotic probability of detection criterion in the frequency domain [57, 58], non-Gaussian noise modeling [59], and

condition monitoring and diagnosis of rolling element bearings. Dyer and Stewart first proposed the use of the kurtosis for rolling element bearing defect detection in 1978 [60], and use of statistical moments remains important in this area [61, 62]. These techniques have advantages over traditional time and frequency analysis: lower sensitivity to the variations of load and speed, and easy and convenient analysis of the results [62].

Conventionally, the statistical moments of a set of discrete data, x_i, are computed directly using a two-pass algorithm (e.g. [54, 55]):

$$\mu = \frac{1}{I}\sum_{i=1}^{I} x_i = E[x] = m_1 \tag{4.1}$$

$$\sigma^2 = \frac{1}{I}\sum_{i=1}^{I}(x_i - \mu)^2 = \theta_2 \tag{4.2}$$

$$\alpha_3 = \frac{1}{I\sigma^3}\sum_{i=1}^{I}(x_i - \mu)^3 = \frac{\theta_3}{\sigma^3} \tag{4.3}$$

$$\alpha_4 = \frac{1}{I\sigma^4}\sum_{i=1}^{I}(x_i - \mu)^4 = \frac{\theta_4}{\sigma^4} \tag{4.4}$$

where I is the number of points in the sample; μ, σ^2, α_3, and α_4 are the mean, variance, skewness, and kurtosis of the data x_i, and θ_2, θ_3 and θ_4 are the central moments. Such algorithms are called two-pass because the mean must first be computed and that mean is subsequently used in the computation of the remaining moments, which implies the entire dataset must be retained. The computational demands imposed by calculation of these moments is a strong function of the number of times a quantity is raised to a power. For example, in calculating the α_4, the quantity $(x_i - \mu)$ is raised to a power of 4, I times: resulting in one power computation per data-point per statistical moment. Here, computational efficiency of an alternative method to compute statistical moments is re-investigated: the data are first binned to create a histogram from which the desired moments can be calculated. In this method, the

number of times a quantity must be raised to a power is a function of the number of bins rather than the number of data points, and the width of each bin can be specified as a function of the required accuracy. Setting the bin width to the precision of the original sensed data yields exact results. The computational savings of this alternative method can be substantial, but perhaps more importantly, dramatic memory savings can also be realized if the raw data are binned real-time, and so the complete time-history does not need to be retained. Avoiding the need to retain a the time-history makes this new methodology competitive with existing one-pass algorithms for the variance, except that here the statistical moments are not computed every time-step. One-pass or on-line algorithms for the mean and variance have been known for some time (e.g., [63]), and have been implemented on real-world hardware (e.g. [64]).

Methodologies for one-pass algorithms for the higher moments, however, are less common. Terriberry [65] offers pairwise updating formulas for the skewness and kurtosis (without derivation). Pebay [66] explains how Terriberry's updating formulae could be implemented as a one-pass algorithm. The one-pass histogram-based algorithm presented here is unique from earlier work in that it allows the user to specify arbitrary accuracy, such that the accuracy of the calculations can be made equivalent to that of the measurement equipment, enabling some savings in computational and memory requirements. Computation of statistical moments from a histogram is generally well-established in the statistics community, but investigation of the methodology as an efficient one-pass algorithm with arbitrarily specified accuracy is unique. As such, this part of the paper is of little theoretical interest to the statistics community, but may be of considerable practical interest to the structural health monitoring community. The technique presented here is less computationally intensive for very large data sets than a true one-pass algorithm because here the data is binned real-time and the moments are computed from the binned data, rather than computing an updated

skewness and kurtosis every time-step. The Terriberry/Pebay one-pass algorithm is derived as a special case of a pairwise updating algorithm, which is substantially different than the derivation presented here.

In a related calculation, it is often convenient to use an updating algorithm to combine statistical moments from individual segments of a data set. This combination can be useful such that various segments of the data can be processed in parallel, or such that segments of a time-history can be processed sequentially as they are measured. In SHM problems, it can be useful to compute statistical moments of new measured data for comparison with those of previously measured historic data. If no differences are observed, then the new data are merged into the historic data to give a more robust estimate of the true underlying moments. The most simple, though inefficient, way to revise the historic statistical moments would be to concatenate the new data onto the old, then compute the moments of the complete history. Here, a new pairwise method is presented to combine the statistical moments computed from individual segments of a time-history, without the need to recompute any of these statistical moments. Using this technique, only five historic values need to be retained: the first four statistical moments and the amount of data on which they are based.

In addition to explaining the use of Terriberry's [65] results as a one-pass algorithm, Pebay [66] also notes that Terriberry's results are special cases of Pebay's arbitrary-order update formulae. The implementation offered by Pebay is substantially different from that suggested here, and the derivation offered by Pebay also differs meaningfully from that offered here, though both derivations hinge on the commutativity of summations over finite sets as applied to statistical moments. In his report, Pebay echoes Terriberry's thoughts, noting that: "To our knowledge, there are currently no published formulas for parallel updates of higher-order moments."

An important capability of the updating formulae offered here compared with those of Terriberry/Pebay is that this methodology can be readily applied to distributions of data that are not specifically countable, such as the moments of a probability distribution estimated from e.g. a power spectrum (e.g., [67]). Also of potential interest to the structural health monitoring community is that this new methodology is easily modified to allow a user to assign different importance to specific segments of the data, such as newer data being more important than older data, or one sensor being more reliable or accurate than another. Finally, the Terriberry formulae apply to combination of only two sets of higher-order moments, and would need to be used recursively to combine multiple sets of moments. In contrast, the new methodology presented here can be used to combine any number of sets of moments, which could be relevant in the case of massively parallel computations.

B. Theory

Derivation of the new methodology requires some knowledge of moments, which is presented here to ensure consistent notation. The background is followed by an investigation of computing statistical moments of discrete data from a histogram, with emphasis on the relative efficiency and accuracy compared with conventional methods. Finally, the proposed method to combine statistical moments is derived and presented. Throughout most of this text, the data set is referred to as an irregular time-history because that is the most common application to structural health monitoring. In general, the equations and techniques presented here are equally valid for any random variable.

1. Background

a. Calculation of Moments from Discrete Data

The moments of a random variable about zero and about its mean are referred to as the raw and central moments, respectively. The n^{th} moment of a discrete random variable $x(t)$ about value r with a finite range is defined as (e.g. [54, 55]):

$$M_{n,r} = \sum_{i=1}^{I} \Big(x(t_i) - r\Big)^n \Delta t_i \tag{4.5}$$

which can be expressed as follows if the continuous distribution function, $f(x)$, is known:

$$M_{n,r} = \sum_{k=1}^{K} (x_k - r)^n f(x_k) \Delta x_k \tag{4.6}$$

in which I represents the number of data points $x(t_i)$ and K represents the number of base x_k in its discrete distribution function $f(x_k)$, i.e., the number of bins in the discrete distribution.

The n^{th} raw moment of a discrete time-history (Equation (4.5) with $r = 0$) can be normalized by the time duration, with the result equal to the expected value of the n^{th} power of x

$$m_n^{(t)} = \frac{1}{T} M_{n,0} = \frac{\sum_{i=1}^{I} x(t_i)^n \Delta t_i}{\sum_{i=1}^{I} \Delta t_i} = E[x^n] \tag{4.7}$$

where $T = \sum_{i=1}^{I} \Delta t_i$ is the time duration and the superscript $^{(t)}$ indicates moments are calculated directly from the time-history. For constant Δt, the duration is $T = I\Delta t$, which enables Equation (4.7) to be simplified:

$$m_n^{(t)} = \frac{1}{I} \sum_{i=1}^{I} x(t_i)^n \tag{4.8}$$

The first normalized raw moment ($n = 1$) is the sample mean, which is often used to

estimate the true mean of the process for normalization of other central moments.

$$\theta_n^{(t)} = \frac{1}{T} M_{n,m_1^{(t)}} = \frac{\sum_{i=1}^{I} \left(x(t_i) - m_1^{(t)}\right)^n \Delta t_i}{\sum_{i=1}^{I} \Delta t_i} = E\left[\left(x - m_1^{(t)}\right)^n\right] \tag{4.9}$$

which for constant time interval Δt is:

$$\theta_n^{(t)} = \frac{1}{I} \sum_{i=1}^{I} \left(x(t_i) - m_1^{(t)}\right)^n \tag{4.10}$$

The second normalized central moment ($n = 2$) is the sample variance (Equation (4.2)).

b. The Relationship Between Raw Moments and Central Moments

The first four raw moments and central moments have the following well-known mathematical relationships, e.g. [54, 55]:

$$m_1 = E[x] = \mu \tag{4.11}$$

$$m_2 = E[x^2] = \theta_2 + m_1^2 \tag{4.12}$$

$$m_3 = E[x^3] = \theta_3 + 3m_1\theta_2 + m_1^3 \tag{4.13}$$

$$m_4 = E[x^4] = \theta_4 + 4m_1\theta_3 + 6m_1^2\theta_2 + m_1^4 \tag{4.14}$$

$$\theta_2 = E[(x-\mu)^2] = m_2 - m_1^2 \tag{4.15}$$

$$\theta_3 = E[(x-\mu)^3] = m_3 - 3m_1m_2 + 2m_1^3 \tag{4.16}$$

$$\theta_4 = E[(x-\mu)^4] = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 \tag{4.17}$$

2. Calculation of Moments Using a Relative Histogram

A relative histogram of a random variable can be constructed in the conventional way. The range of potential values is divided into bins and the number of occurrences within each bin are counted and plotted such that the area of each rectangle equals

the portion of the sample values within that bin (e.g. [54, 55]):

$$H(x_k) = \frac{h(x_k)}{A} \tag{4.18}$$

where $h(x_k)$ and $H(x_k)$ represent the frequency and the relative frequency at bin x_k, and $A = \sum_{k=1}^{K} h(x_k)\,\Delta x_k$ is the total area of the histogram. After this normalization, the n raw moments and central moments of $x(t)$ can be calculated from the relative histogram, similar to Equation (4.6):

$$m_n^{(h)} = \sum_{k=1}^{K} x_k^n\, H(x_k)\Delta x_k = \frac{1}{A}\sum_{k=1}^{K} x_k^n\, h(x_k)\Delta x_k \tag{4.19}$$

$$\theta_n^{(h)} = \sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n H(x_k)\Delta x_k = \frac{1}{A}\sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n h(x_k)\Delta x_k \tag{4.20}$$

where the superscript $^{(h)}$ indicates the moments are calculated from the histogram. For constant bin width $\Delta x_k = \Delta x$ these two expressions can be simplified with $A = \sum_{k=1}^{K} h(x_k)\Delta x = I\Delta x$:

$$m_n^{(h)} = \frac{1}{I}\sum_{k=1}^{K} x_k^n\, h(x_k) \tag{4.21}$$

$$\theta_n^{(h)} = \frac{1}{I}\sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n h(x_k) \tag{4.22}$$

a. Precision of Moments Calculated from a Histogram

Calculation of sample moments directly from the time-history results in optimal accuracy since all data are used directly and there is virtually no opportunity for round-off error. The histogram method also yields perfect accuracy if the bin width is equal to the precision with which the data has been measured. If computational efficiency is of greater importance than absolute precision, then the bins can be made arbitrarily wider, trading a reduction in computational demands for reduced accuracy. The resulting difference between moments calculated using a time-history and using its

histogram represents the error:

$$m_n^{(h)} - m_n^{(t)} = \frac{1}{A}\sum_{k=1}^{K} x_k^n \, h(x_k)\Delta x_k - \frac{1}{T}\sum_{i=1}^{I} x(t_i)^n \Delta t_i \tag{4.23}$$

which for both constant time interval Δt and constant bin width Δx are:

$$m_n^{(h)} - m_n^{(t)} = \frac{1}{I}\left(\sum_{k=1}^{K} x_k^n \, h(x_k) - \sum_{i=1}^{I} x(t_i)^n \right) \tag{4.24}$$

The trade-off between precision and computational savings is investigated as part of the example later in this paper.

3. Efficient Combination of Statistical Moments

Here, a new computational method is proposed to combine multiple sets of statistical moments. An example application would be combining moments from several individual segments of a long time-history, with each segment possibly being processed by a separate processor. If sample statistical moments describing several separate segments of an irregular time-history have been computed from measured data, statistical moments describing a single concatenated time-history of all data can be calculated directly from the existing statistical moments. This proposed computational technique uses the first four statistical moments of each segment to compute the four raw moments, which are then transformed into new variables (γ_n) that are easily combined by addition. After combination, the new variables are inversely transformed back to four raw moments now describing all the data, from which the statistical moments are easily calculated.

a. Moment Addends, γ_n

New moment addend variables, γ_n, are introduced to enable straightforward combination of the statistical moments of multiple time-histories. For an irregular time-history

$x(t)$ with variable time interval Δt_i:

$$\gamma_n = \sum_{i=1}^{I} (x(t_i))^n \Delta t_i \tag{4.25}$$

where γ_0 is the duration of each time-history. For constant time interval $\Delta t_i = \Delta t$:

$$\gamma_n = \Delta t \sum_{i=1}^{I} (x(t_i))^n \tag{4.26}$$

The same values of γ for the histogram form of computing the moments can be expressed in terms of the frequency of occurrence at the x_k bin, $h(x_k)$, with variable bin width Δx_k.

$$\gamma_n \quad = \quad \sum_{k=1}^{K} x_k^n h(x_k) \Delta x_k \tag{4.27}$$

yielding γ_0 as the area of the histogram. For constant bin width $\Delta x_k = \Delta x$:

$$\gamma_n = \Delta x \sum_{k=1}^{K} x_k^n h(x_k) \tag{4.28}$$

b. Combination of Statistical Moments

If Q sets of statistical moments are known: $(\gamma_{0,q}, \mu_q, \sigma_q^2, \alpha_{3,q}, \alpha_{4,q}) \quad \text{for} \quad q = 1, 2, ..., Q$, then each γ_n can be expressed in terms of the equivalent n raw moments (Equations 4.7, 4.8, 4.19, and 4.21).

$$\gamma_{n,q} = m_{n,q} \gamma_{0,q} \qquad \text{for} \quad n = 1, 2, 3, 4 \quad \text{and} \quad q = 1, 2, ..., Q \tag{4.29}$$

where $\gamma_{0,q}$ is generally taken to be the duration of the q^{th} time-history, or the number of points if Δt is constant. It is worth noting, however, that $\gamma_{0,q}$ is a weighting factor only, and its interpretation can be flexible depending on the application. Importantly, in this method the statistical moments are not required to be those of a quantity that is countable: these moments could be computed directly from e.g., a probability

distribution, in which case the value of $\gamma_{0,q}$ would represent the relative importance of the moment estimate. There is no theoretical limitation on the maximum order of the moments (the value of n), though higher-order equivalents to Equations (4.11–4.17) would be needed in practical applications. The benefit of expressing the statistical moments in terms of γ is that the Q sets can be combined by addition, and there is no upper limit on the value of Q.

$$\gamma_{n,c} = \sum_{q=1}^{Q} \gamma_{n,q} \qquad \text{for} \quad n = 0, 1, 2, 3, 4 \tag{4.30}$$

where the subscript $_c$ represents the concatenated time-history or combined γ. These combined values of γ can then be inversely transformed into raw moments representing the concatenated time-history by inverting Equation (4.29).

$$m_{n,c} = \frac{\gamma_{n,c}}{\gamma_{0,c}} \quad \text{for} \quad n = 1, 2, 3, 4 \tag{4.31}$$

The relationship between raw moments and central moments (Equations 4.15–4.17) are then used to compute the central moments of the concatenated time-history. Finally, the statistical moments of the concatenated history are computed as in Equations (4.1–4.4)

$$\mu_c = m_{1,c} \qquad \sigma_c^2 = \theta_{2,c} \qquad \alpha_{3,c} = \frac{\theta_{3,c}}{\sigma_c^3} \qquad \alpha_{4,c} = \frac{\theta_{4,c}}{\sigma_c^4} \tag{4.32}$$

C. Application

1. Estimation of Statistical Moments of a Concatenated Time-history

In general, sample moments approach the true moments of the underlying process as the time-history increases in length. If very long time-histories are available, moments of the true process can often be estimated by those of the long sample. However, if individual segments of the long history are available, several possible approaches to

finding moments of the concatenated history are available. Three of these possibilities are presented in Figure 19. Each of these methods can be applied repeatedly to compute statistics of a long time-history made up of any number of segments.

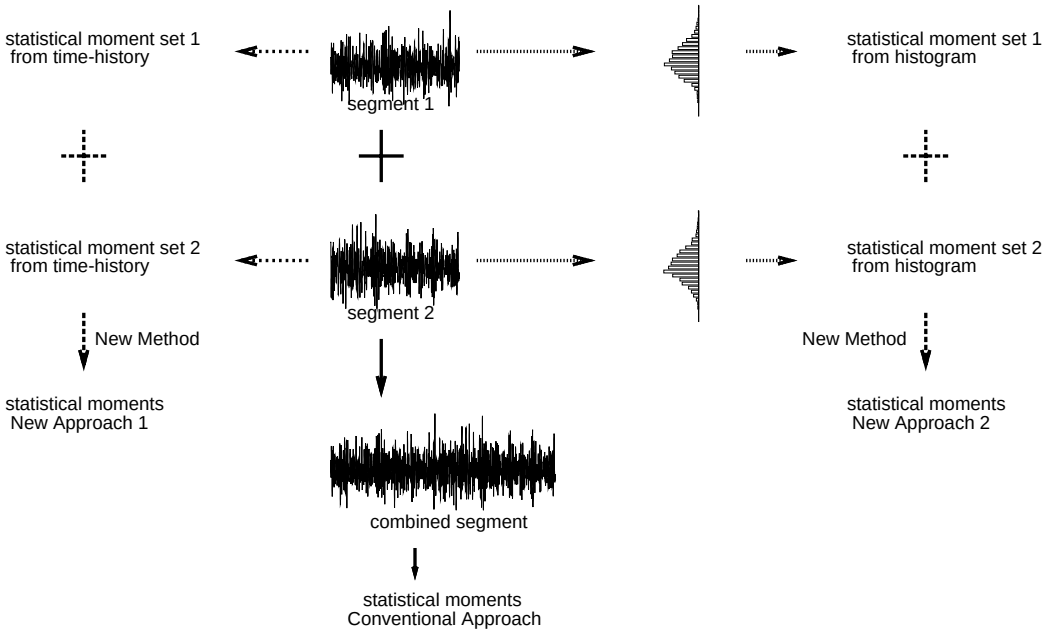


Fig. 19. Three Approaches to Calculate Statistical Moments of a Concatenated Time-history

The most conventional method is shown in the center of the figure: the first segment and the second segment are concatenated and the statistical moments are calculated directly from the resulting long time-history. This simple approach requires considerable computational resources and substantial storage space to retain the previous time-histories, both of which may be problematic in field applications using micro-computers in distributed sensor arrays.

Two alternative approaches are shown progressing down the left and right sides of Figure 19. The principle benefit compared with the conventional approach is a significant decrease in storage requirements since only the statistical moments need to be retained, rather than the complete time-histories. In both cases, the methods of Section b are applied to combine these moments. The two approaches differ only in how the statistical moments for each segment are computed before concatenation. The method on the left side of the figure is based on calculation of the statistical moments of each time-history by the conventional method of Equations 4.1–4.4; the approach shown on the right side of the figure is based on calculation of the statistical moments through use of the histogram as in Section 3.

a. Calculation of Statistical Moments from a Histogram

One method to calculate the statistical moments is by calculation of the central moments from the histogram (Equations 4.20 or 4.22) and then converting these to the central statistical moments through the relationships of (Equations 4.1–4.4). If the bin width of the histogram is equivalent to the smallest decimal place of the measured data, then the results will be identical to those calculated by conventional means. As previously noted, for very long time-histories the histogram method will generally be more efficient; for shorter time-histories the more conventional method will generally be more efficient. However, the histogram method also offers an additional option to reduce computational demands: the bin width can be made larger than the smallest decimal place of the measured data. The effects of increased bin width versus computational demands are investigated in a later example.

b. Combination of Statistical Moments

Regardless of how the statistical moments of individual segments of a time-history were calculated, these moments can be combined efficiently. The procedure is straightforward: First, the $n = 4$ statistical moments for each (q) of the Q segments to be combined are transformed into the mean and three central moments by inverting the definitions of the statistical moments (Equations 4.1–4.4) to $m_{1,q} = \mu_q$, $\theta_{2,q} = \sigma_q^2$, $\theta_{3,q} = \alpha_{3,q}\sigma_q^3$, $\theta_{4,q} = \alpha_{4,q}\sigma_q^4$. Second, the resulting Q sets of three central moments are transformed into Q sets of three raw moments using the well-known relationships between raw and central moments (Equations 4.12–4.14). Third, the resulting $4Q$ raw moments (including Q means) are then transformed into $4Q$ values of $\gamma_{n,q} = m_{n,q}\gamma_{0,q}$ (Equation (4.29)). Fourth, each of the 5 sets of Q values of $\gamma_{n,q}$ are combined as in Equation (4.30), $\gamma_{n,c} = \sum_{q=1}^{Q} \gamma_{n,q}$ ($n = 0$ to 4). Finally, the transformation process is reversed for the resulting 5 values of $\gamma_{n,c}$ to produce the desired four central statistical moments as in Equations 4.31–4.32.

D. Example

The new methodologies are applied to simulated data resulting from a time-domain solution of a simple numerical model of a Tension Leg Platform subject to irregular seas. A TLP was selected for this example because of its highly non-Gaussian surge response (horizontal translation in the direction of the environmental loading). Three segments of a surge time-history, each having significantly different statistical moments, are created without modifying the structural model, but only changing the peak period of the incident waves, the current velocity and the wind force. Calculation and combination of the resulting statistical moments are investigated and compared.

1. Numerical Simulations

A Tension Leg Platform (TLP) is a compliant offshore structure used for production of oil and gas in deep ocean waters. The platform is vertically moored by tendons at each of its corners (Figure 20). Surge response time-histories of the Snorre TLP, which is located in the Norweigian sector of the North Sea, have been simulated by a simplified 2-dimensional nonlinear numerical model.

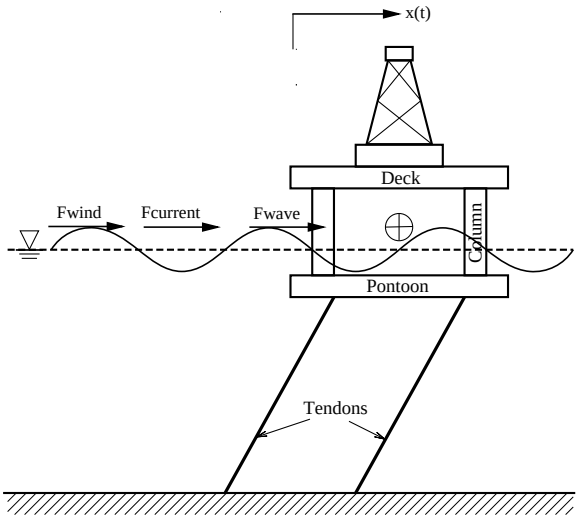


Fig. 20. Schematic View of Tension Leg Platform

The wave and current forces are non-linear and non-Gaussian. These forces are applied to a single degree of freedom system including nonlinear restoring force. The equation of motion is solved in the time domain using the Newmark Beta Method (e.g., [49]). The time step of integration (Δt) is 0.01 sec and total time duration for each of the three time-histories is one hour. The two main sources of non-Gaussianity in the model are 1) the non-linear mooring restoring force caused by the changing

angles of the tendons with increased offset and the increased buoyancy of the hull caused by being pulled downward by the tendons, and 2) the non-Gaussian wave forcing caused by the highly nonlinear drag term in the Morison Equation [48, 68].

2. Results

The statistical moments of each response time-history simulated under three environmental conditions are summarized in Table III. Changing environmental conditions result in substantial changes in the statistical moments of the simulated response.

Table III. Environmental Conditions and Associated Statistics of The Response

Condition	$H_s(m)$	$T_p(sec)$	$U_c(m/s)$	$F_w(kN)$	μ	σ^2	α_3	α_4
1	14.5	15.5	1.5	-3000	-1.3833	0.2742	0.1236	2.9969
2	14.5	12.5	2.0	-5000	-2.2512	0.2447	0.1203	2.3685
3	14.5	18.5	1.0	-2000	-0.9154	0.2510	0.0411	2.1207

a. Calculation of Statistical Moments

Calculation of statistical moments by the conventional methods of Equations 4.1–4.4 are compared with application of the relative histogram (Section a) in Figure 21. The vertical axis on the left compares the relative CPU time needed to compute all four of the statistical moments, with the conventional methodology defined as 100% CPU time. For this one-hour time-history (360,000 data points) with a measurement precision of 0.001, the histogram method takes about 75% of the CPU time as the conventional method if the bin width is set to the precision (zero error in binning leading to exact sample statistical moments). CPU usage figures result from binning and computations performed using MatLab.

Progressing from left to right on the horizontal axis shows gradual increases in the bin width, and the vertical axis on the right shows the associated error in the statistical moments. The error is the difference between the exact and approximate statistical moments divided by the exact statistical moments (calculated conventionally). The plot shows that for this time-history, increasing the bin width to ten times as large as the precision of the data has a savings of about half of the CPU time, with virtually no noticeable increase in error in the statistical moments. The CPU time does not drop by a factor of ten because the same 360,000 data points must be binned.

Calculation using a histogram with bin width of 0.001 yields results identical to the conventional method (Table III), with using 74.7% of the CPU time and 2.78% of the storage requirements. Calculation using a histogram with bin width of 0.01 results in statistical moments with error less than 0.1% using 33.9% CPU time and 0.25% storage space relative to conventional calculation. The dramatic savings in storage space results from assuming the data are binned as they are collected such that the time-history need not be saved. Computation of statistical moments by the conventional two-pass method requires saving the complete time-history so the mean can be computed in the first pass and applied to computation of the moments in the second pass through the data.

b. Calculation of Statistical Moments of the Concatenated Time-history

Three approaches are compared for calculating the statistical moments of a concatenated time-history made up of the three independent time-histories described by Table III.

Method A (Conventional): The conventional method of calculating the statistical moments defines 100% CPU and 100% Storage requirements. This method is to store both of two one-hour time-histories as they are collected, to compute the statistical

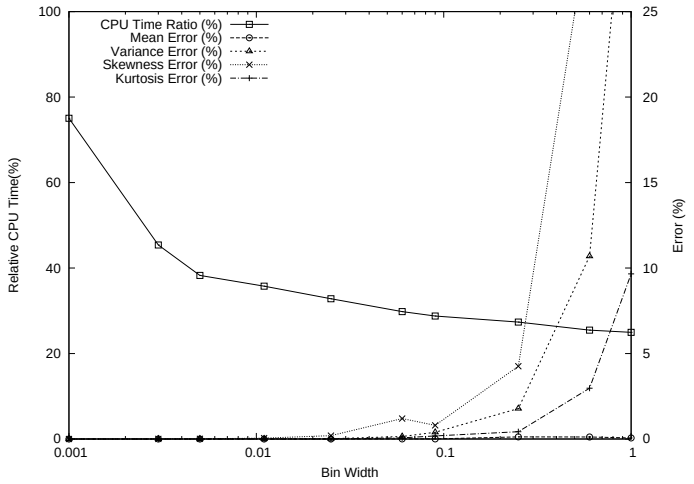


Fig. 21. Effect of Bin Width on Calculation of Statistical Moments. Time Duration: 1 hour, Δt: 0.01 sec, Data Precision: 0.001

moments using the conventional means of Equations 4.1–4.4, then to concatenate the two histories, and then finally to compute the statistical moments of the concatenated time-history.

Method B (Statistical Concatenation): In this application, it is assumed that the first time-history is stored and conventional methods are applied to calculate the statistical moments; the first time-history is then deleted and the second is stored and its statistical moments are calculated. These two sets of statistical moments are then combined using the methods of Section 1. The approximately 50% savings in both CPU usage and memory requirements results from not having to save both time-histories and recompute the moments for the concatenated history.

Method C (Histogram with Statistical Concatenation): The method applied here is to calculate the moments using both the histogram method of Section a and the

statistical method of Section 1 to combine moments. Results from Method (C) are presented both for $\Delta x = 0.001$, which yields exact results, and for $\Delta x = 0.01$, which introduces some error in return for greater CPU and storage savings. Method (C) with adequate bin width for an exact solution requires only about 27% of the CPU time and around 1.4% of the required memory. The dramatic memory savings result from the assumption that the time-history will be binned on the smart sensor package as the data are collected, such that no time-history ever needs to be stored. Greater savings are also shown if the bin width is made ten times as large, though the resulting moments are not exact.

Methods (D), (E) and (F) are equivalent, but the single concatenated segment resulting from (A), (B) and (C) is now combined with the third segment.

E. Conclusions

Two methodologies have been presented and demonstrated through an example. The first methodology is effectively a single-pass methodology for computation of higher statistical moments through use of a histogram. This methodology is based on a simple combination of methods well-known in the statistical community and is therefore of little theoretical interest. It may, however, be of considerable practical interest in the field of structural health monitoring because it enables computation of the skewness and kurtosis to arbitrarily selected accuracy, offering a means to effectively trade computational intensity against accuracy. Generally, for very long time-histories with low-precision data it is more efficient to use the histogram for statistical moment calculation since the number of power calculations is smaller, but for relatively short time-histories with high-precision data the more conventional method will require less CPU time.

The second methodology enables direct combination of the skewness and kurtosis of any number of data sets. The method and its derivation are new, and may be of theoretical interest. One unique aspect of this new methodology is that it can be used to combine statistical moments of data that is not countable, e.g., moments extracted directly from a probability density spectrum.

An example is presented in which both the histogram approach and the updating methodology for the skewness and kurtosis are verified. The relative efficiency and accuracy of the histogram approach are examined in some detail. In the example, statistical moments for time-histories are computed using both the conventional approach and the histogram approach, including setting varying degrees of accuracy. Setting the bin width equal to the precision of the measured data gives perfect accuracy and potential computational savings. A noticeable decrease in computational demands accompanied by some decrease in accuracy caused by increasing the bin width is shown in Figure 21. In the example, CPU savings on the order of 75% of that required for a conventional two-pass algorithm are realized with only a minor decrease in accuracy. Savings in memory requirements can also be quite substantial (Table IV). In the example, applying the histogram as a one-pass algorithm by binning the data as it is collected and then computing the moments from the resulting histogram uses around 1% of the memory compared with a more conventional two-pass algorithm.

Table IV. Statistical Moments of Concatenated Irregular TLP Response Time-histories; Comparison of Calculation Methodologies

Method	Segment	μ	σ^2	α_3	α_4	CPU (%)	Storage (%)
Conv. (A)	1+2	-1.8173	0.4478	0.1180	2.5684	100.0	100.0
Statistical	1+2	-1.8173	0.4478	0.1180	2.5684	51.86	50.00
Concat. (B)	Error(%)	0.0000	0.0000	0.0000	0.0000	-	-
Histogram	$\Delta x = 10^{-3}$	-1.8173	0.4478	0.1180	2.5684	26.58	1.39
with	Error(%)	0.0000	0.0000	0.0000	0.0000	-	-
Statistical	$\Delta x = 10^{-2}$	-1.8177	0.4478	0.1183	2.5687	20.86	0.14
Concat. (C)	Error(%)	0.0021	0.0054	0.0199	0.0034	-	-
Conv. (D)	12+3	-1.5167	0.5629	-0.0917	2.3856	100.0	100.0
Statistical	12+3	-1.5167	0.5629	-0.0917	2.3856	34.99	33.33
Concat.(E)	Error(%)	0.0000	0.0000	0.0000	0.0000	-	-
Histogram	$\Delta x = 10^{-3}$	-1.5167	0.5629	-0.0917	2.3856	15.63	0.93
with	Error(%)	0.0000	0.0000	0.0000	0.0000	-	-
Statistical	$\Delta x = 10^{-2}$	-1.5171	0.5630	-0.0913	2.3856	11.82	0.09
Concat.(F)	Error(%)	0.0279	0.0185	-0.3742	-0.0008	-	-

CHAPTER V

NOISE REDUCTION METHODOLOGY

A. Introduction

Vibration-based Structural Health Monitoring (SHM) systems use measured response data that conveys information of structural properties. It is unavoidable that field measurements of operational signals are distorted by noise from many sources including the sensors themselves, electromagnetic interference, wiring problems, and environmental conditions. Reducing this noise allows a more accurate assessment of the original "clean" signal and improves analysis results. Noise reduction is the process of extracting the original clean signal from a measured noisy signal, and is a popular research topic in the areas of speech signal processing and image processing. Numerous noise reduction strategies exist. Perhaps the simplest is to apply a series of bandpass filters, which remove specific frequencies from the signal. This method is obviously most effective if the noise occurs at a relatively narrow and well defined range of frequencies. In case of a very broadband spectrum such as white noise, energy is distributed over a wide range of frequencies and bandpass filtering is ineffective.

More complicated methods of noise reduction in power spectra have been widely used in speech signal processing to estimate the Power Spectrum Density (PSD) of the original pure signal which has been corrupted by broadband noise. Perhaps the next most simple was proposed by Boll [69]. The spectral subtraction method is a noise reduction technique that is popular due to its simple underlying concept and its effectiveness for enhancement of a signal that had been degraded by additive broadband noise. The underlying assumption is that the total (noisy) signal consists of the clean signal plus a theoretical spectrum of white noise. The basic principle of

the spectral subtraction method is to subtract a constant magnitude of noise, that of a pure white-noise spectrum, from the total noisy signal. The noise is assumed to be uncorrelated and additive to the signal. After substraction, any frequencies having negative power are set to a floor value, which is generally zero. This zeroing off of negative areas causes the method to be non-conservative: the energy of the clean signal plus the energy of the noise is generally greater than the energy of the measured signal.

Unfortunately, the non-conservative nature of this popular method makes it inappropriate for use with the modal distribution method (MDM). The incompatibility stems from the fact that the MDM uses the power associated with each mode to weight the relative importance of the individual modes in assessing the statistical significance of observed changes in the PSD. Implementing the spectral subtraction method to the MDM generally results in the noise-reduced modal distribution having smaller energy (area) than the underlying modal distribution of the true (noiseless) signal, and the reduction occurs in an absolute, rather than proportional way. Modes having relatively small energy are reduced by about the same amount of energy as more prevalent modes, rather than by the same fraction of energy. This absolute subtraction can dramatically affect the statistical importance when combining the contributions from individual modes. Problems can include having modes with very small energies dominating results because all but a few frequencies have been zeroed off, or at the other extreme, eliminating small modes entirely. In this work, a new methodology is introduced that enables a conservation of total energy. This new method is a variation on the modal subtraction method. The main difference is that here, any negative areas (bars with a height below zero) are not set to zero, but instead the height of these bars is set to the absolute value of their height after subtraction. This relatively simple modification allows conservation of total energy.

The physical interpretation is that any negative heights in the power spectrum are caused by phase differences.

Here, the original power spectrum subtraction method is explained in detail and the proposed modifications to this original method are mathematically outlined. The new method is then demonstrated with implementation to the modal distribution method. In the MDM, the magnitude of the noise profile is estimated from the magnitude of the PSD at frequencies higher than those expected to have any true signal associated with structural vibrations. An example is presented that demonstrates the effectiveness of this new noise reduction technique.

B. Review of Noise Reduction Methods for Speech Enhancement

Noise reduction methodologies are general in concept, and can be implemented to nearly any measured signal that has been corrupted by noise. A variety of approaches have been proposed to reduce noise for speech enhancement: Wiener filtering, dynamic comb filtering, short-time spectral modification techniques, and others [70, 71]. In this section, two of these techniques are reviewed that are popular in the field of acoustics.

1. Signal with Uncorrelated Additive White Gaussian Noise

Suppose that a measured discrete time-history y_m is composed of a clean, noiseless signal x_m, additively combined with noise n_m:

$$y_m = x_m + n_m \tag{5.1}$$

This additive relationship is conserved in Discrete Fourier Transform (DFT) space.

$$Y_k = X_k + N_k \tag{5.2}$$

where Y_k, X_k and N_k are DFTs of y_m, x_m and n_m, respectively. Subscripts m and k will be omitted for convenience. These DFTs are complex, which are expressed in terms of magnitude and phase:

$$|Y|e^{i\theta_Y} = |X|e^{i\theta_X} + |N|e^{i\theta_N} \tag{5.3}$$

Multiplying both sides of Equation 5.2 by their complex conjugates gives:

$$YY^* = XX^* + NN^* + XN^* + NX^* \tag{5.4}$$

$$|Y|^2 = |X|^2 + |N|^2 + |X||N|e^{i(\theta_X - \theta_N)} + |X||N|e^{-i(\theta_X - \theta_N)} \tag{5.5}$$

Consequently, the power spectral density (PSD) of measured signal including additive noise consists of the auto-spectrum and cross-spectrum of clean signal and noise:

$$S_Y = S_X + S_N + S_{XN} + S_{NX} \tag{5.6}$$

The PSD of the clean signal S_X can be recovered by approximating the last three components in the Equation 5.6. If the noise is uncorrelated with the signal, which is implied by its being independent, the cross spectra are theoretically zero: $S_{XY} = 0$ and $S_{YX} = 0$. Thus, if the signal and noise are stationary and independent, the power spectrum of the noise-corrupted signal, S_Y, is simply the sum of the power spectra of the signal and noise:

$$S_Y = S_X + S_N \tag{5.7}$$

Many noise reduction methodologies are based on this simple equation.

2. Spectral Power Subtraction and Wiener Filtering

Among methods to reduce the effect of additive broadband noise, the spectral subtraction method is the most popular due to its robustness and simplicity. The PSD

estimate of the clean signal is easily obtained by subtracting the power spectrum of the noise from that of the total observed signal, which includes noise:

$$\widehat{S_X} = S_Y - S_N \tag{5.8}$$

Though trivial in concept, this fundamental spectral power subtraction relation forms the basis for many noise reduction methods.

Another powerful tool for reducing additive noise is the Wiener filter [72], which has a fundamental relation to all past and modern noise reduction methods. It makes use of classical linear estimation theory to estimate x; the estimation $\widehat{x}$ of x minimizes the mean-squared error $\|x - \widehat{x}\|$.

$$\widehat{X} = HY \tag{5.9}$$

where $\widehat{X}$ is the Fourier transform corresponding to the optimum $\widehat{x}$, Y is the Fourier transform of y, and

$$H = \frac{S_X}{S_Y} = \frac{S_X}{S_X + S_N} \tag{5.10}$$

is a form of the filter frequency response function derived by Norbert Wiener .

The Wiener filter can be used as a spectral modification method by combining Equations 5.8-5.10. The least-mean-square estimate for Fourier transform of the signal is acquired simply by applying the following frequency dependent gain function to the spectrum of the noisy signal.

$$\widehat{X} = \frac{S_Y - S_N}{S_Y} Y \tag{5.11}$$

This is a form of Wiener solution that utilizes a spectral subtraction operation.

3. Stationary Signal

Using the principle of ensemble averaging for stationary signals, the power spectra of signal and noise are given by the expected value of the squared-modulus of their respective Fourier transforms.

$$S_X = E\{|X|^2\} \tag{5.12}$$

$$S_N = E\{|N|^2\} \tag{5.13}$$

Consequently, the observed noisy signal has the same relation, $S_Y = E\{|Y|^2\}$. Combining these ensemble averages with spectral power subtraction (Equation 5.8) and the Wiener filter using spectral power subtraction (Equation 5.11) yields:

$$E\{|X|^2\} = E\{|Y|^2\} - E\{|N|^2\} \tag{5.14}$$

$$\hat{X} = \frac{E\{|Y|^2\} - E\{|N|^2\}}{E\{|Y|^2\}} Y \tag{5.15}$$

When these ensemble averages are unknown, Fourier transforms of the observed signals can be used as sample estimates of the unknown ensemble averages, leading to:

$$|X|^2 = |Y|^2 - |N|^2 \tag{5.16}$$

$$\widehat{X} \simeq \frac{|Y|^2 - |N|^2}{|Y|^2} Y \tag{5.17}$$

The Fourier transform is expressed in terms of its magnitude and phase components, namely,

$$X = |X| e^{i\theta_X} \tag{5.18}$$

The square root of Equation 5.16 provides an estimate of the magnitude of the signal spectrum. It is reasonable to estimate the phase of the signal, θ_X, by the phase of

the noisy signal, θ_Y, if the signal-to-noise ratio is reasonably high.

$$\widehat{X} = \sqrt{|Y|^2 - |N|^2} e^{i\theta_Y} \tag{5.19}$$

This is an alternative form of the spectral power subtraction method for a stationary signal when the ensemble average is unknown.

C. Methodology

1. Review: Derivation of Spectral Power Subtraction Method

A firm theoretical understanding of the pre-existing spectral spectral subtraction method is necessary to fully justify the newly proposed modifications to this well-known method. There are two different ways to derive the power subtraction method. In the first derivation, the imaginary part of Equation 5.3 results in trigonometric functions, which are expressed as sine functions.

$$|Y| \sin \theta_Y = |X| \sin \theta_X + |N| \sin \theta_N \tag{5.20}$$

The phase differences, ϕ_Y, between the noisy and clean signals caused by the noise and the phase difference, ϕ_N, between the clean signal and the noise itself are defined by:

$$\theta_Y = \theta_X + \phi_Y \tag{5.21}$$

$$\theta_N = \theta_X + \phi_N \tag{5.22}$$

A trigonometric identity enables expression of the magnitude of the noisy signal (Equation 5.23), and the phase difference between the noisy signal and the clean signal (Equation 5.24)

$$|Y| = \sqrt{|X|^2 + |N|^2 + 2|X||N|\cos\phi_N} \quad (5.23)$$

$$\phi_Y = \arctan\left(\frac{|N|\sin\phi_N}{|X| + |N|\cos\phi_N}\right) \quad (5.24)$$

The irregular phase difference between the noisy signal and the clean signal in each frequency component can be any value between $-\pi$ and π radians. If the distribution of these differences is uniform, the best estimation for the squared magnitude and the phase difference can be computed by averaging from $-\pi$ to π. Integrating over all possible values of the continuous random variable ϕ_N:

$$\begin{aligned} E\{|Y|^2\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |Y|^2 d\phi_N \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \left\{|X|^2 + |N|^2 + 2|X||N|\cos\phi_N\right\} d\phi_N \\ &= E\{|X|^2\} + E\{|N|^2\} \quad (5.25) \\ E\{\phi_Y\} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi_Y d\phi_N = 0 \quad (5.26) \end{aligned}$$

This result is the same as the power subtraction method (Equation 5.19), in which the square of the DFT magnitude is used in place of the power spectrum and the noisy signal phase is used as the clean signal phase.

Alternatively, taking the real part of Equation 5.5 results in the following:

$$|Y|^2 = |X|^2 + |N|^2 + 2|X||N|\cos\phi_N \quad (5.27)$$

where $\phi_N = \theta_X - \theta_N$. This equation is the same as the square of Equation 5.23. Taking the expected value of both sides results in:

$$\begin{aligned} E\left\{|Y|^2\right\} &= E\left\{|X|^2\right\} + E\left\{|N|^2\right\} + E\left\{2|X||N|\cos\phi_N\right\} \\ &= E\left\{|X|^2\right\} + E\left\{|N|^2\right\} + 2E\left\{|X|\right\}E\left\{|N|\right\}E\left\{\cos\phi_N\right\} \quad (5.28) \end{aligned}$$

Three assumptions are implied by the transition from the first to the second line of Equation 5.28: First, the magnitude of the spectral offsets of the noise and signal are independent of each other, Second, the phase of the noise and the phase of the signal are independent of each other, and Third, these phases are independent of the associated magnitudes. If $E\{\cos\phi_N\} = 0$, the power subtraction method results in:

$$E\left\{|Y|^2\right\} = E\left\{|X|^2\right\} + E\left\{|N|^2\right\} \tag{5.29}$$

$$|X|^2 = |Y|^2 + E\{|N|^2\} \tag{5.30}$$

Additionally, Equation 5.23 and 5.27 provide the maximum and minimum magnitude of the clean signal, because the cosine function is bounded : $-1 \leq \cos\phi_N \leq 1$.

$$|X|_{min} = \begin{cases} 0, & |Y|^2 > |N|^2 - 2|X||N| \\ \sqrt{|Y|^2 - |N|^2 + 2|X||N|}, & \text{otherwise} \end{cases} \tag{5.31}$$

$$|X|_{max} = \sqrt{|Y|^2 - |N|^2 + 2|X||N|} \tag{5.32}$$

Dividing the divisor and dividend in Equation 5.24 by $|N|$, the phase difference between the noisy and clean signals can be expressed by the signal to noise ratio ($SNR = |X|^2/|N|^2$).

$$\phi_Y = \arctan\left(\frac{\sin\phi_N}{\sqrt{SNR} + \cos\phi_N}\right) \tag{5.33}$$

2. Shortcomings of the Spectral Power Subtraction Method

The newly proposed method is proposed to overcome some specific shortcomings of the well-known spectral power subtraction method. Application to the MDM of speech enhancement methodologies based on spectral subtraction has two shortcomings: First, it is based on the direct estimation of the short-term spectral magnitude, which may be less accurate than a long-term spectral magnitude for a steady signal.

This inaccuracy may introduce distortion and large variation to the estimate of the enhanced spectrum. The second disadvantage of spectral subtraction is the possibility of negative estimates of power at specific frequencies in the enhanced spectrum. In such cases, the negative spectral components are conventionally floored to some values by linear or nonlinear methods. More complex methods to reduce the level of residual noise have been devised for cases in which the noise spectrum has been overestimated (leading to negative areas under the power spectrum), including an early implementation by Berouti et al. [73]. In these methods, the portion of noise that is subtracted from the signal is adaptively adjusted according to the signal-to-noise ratio. The adaptive floor is commonly computed for spectral magnitude of noisy signal less than one of noise. Several implementations of this algorithm are presented in the literature [73] [74].

In order to apply any algorithm for noise reduction, an estimate of the noise must first be established. In speech signal processing, the noise can be well-approximated by measurement of the signal at times when no voice is present. Unfortunately, in structural health monitoring of civil structures, the system cannot be restrained to a no-response condition such that the noise can be estimated. Instead, it is here proposed that a white-noise power spectrum be applied and the level of the white noise be approximated by the level detected outside the range of frequencies in which any structural response is expected or observed. However the noise is estimated, there is generally some difference between the true noise and the estimated noise in both magnitude and phase. This difference can result in negative offsets in the power spectrum. Since the offset of the power spectrum is computed as a real magnitude squared, these negative offsets are not physical. Any spectral subtraction method that is required to conserve energy must appropriately handle these negative densities.

The most simple, though non-conservative, solution is to set any negative values

to zero:

$$S_X = \begin{cases} S_Y - S_N, & S_Y > S_N \\ 0, & \text{otherwise} \end{cases} \tag{5.34}$$

where S_Y is estimated from measured y and S_N is approximated. Whenever S_Y is less than S_N or the resulting S_X is less than zero, it is assumed that S_Y is equal to S_N or S_X is zero.

The conventional spectral subtraction method (Equation 5.19) assumes that the phase of the clean signal is the same as that of noise and of the noisy signal. If any two of three phases (θ_Y, θ_X, and θ_N) are the same, then Equations 5.21-5.22 imply that all three are the same.

3. New Methodology

The new methodology presented here overcomes the main limitation of the spectral subtraction method: conservation of energy. In practice, this quantity may or may not be energy: it generally has the units of the measurement squared per frequency. These techniques can be applied whether or not the specific units are those of energy density. The theoretical modifications proposed here are based on physical interpretation: the negative components that appear in a noise-reduced PSD are considered to have been caused by phase differences between the signal and noise. If $|X|^2$ is estimated by $\widehat{|X|^2} = |Y|^2 - |N|^2$, then the difference between actual spectrum and estimated spectrum of the signal results from Equation 5.27:

$$\widehat{|X|^2} = |X|^2 + 2|X||N|\cos\phi_N \tag{5.35}$$

If the phase difference between the noise and clean signal ϕ_N causes the cosine to be negative, it will cause the estimated magnitude of that frequency component of the

the signal power spectrum to be smaller than actual (clean) magnitude. For a single frequency with a relatively high noise component ($|N|$), the magnitude of $|\widehat{X}|^2$ can be negative. The worst case happens when $|N|$ is large and the phase difference between the noise and signal is $-\pi$.

Any negative components can contribute to the total power (area under the PSD) by taking its absolute value. An individual frequency component of the power spectrum is in general directly proportional to the square of the Fourier amplitude computed from the time-series. Computing the Fourier amplitude, therefore, from a negative PSD component would seem to imply an imaginary Fourier amplitude. The PSD component is of course computed as proportional to the square of the magnitude of the Fourier amplitude, which would be equivalent to taking the absolute value of any negative PSD component. Taking absolute values of areas under the power spectrum is not completely novel: a single-sided power-spectrum is computing by integrating over all frequencies, from negative to positive infinity, and the amplitude of the spectral offsets of the double-sided is simply twice those of the single-sided computed by integrating over only positive frequencies. This observation implies taking absolute values of the negative areas (associated with negative frequencies) as they are reflected about zero-frequency.

Expressing this absolute value process in equation form: the signal power spectrum estimated from the noisy spectrum is simply:

$$S_X = |S_Y - S_N| \tag{5.36}$$

The resulting noise reduction function has the form:

$$\hat{X} = \frac{|S_Y - S_N|^{1/2}}{S_Y^{1/2}} Y \tag{5.37}$$

This modification results in conservative total energy of the PSD of cleaned signal

when it is separated from noise.

D. Example

In this section, the newly developed noise reduction methodology is implemented with the parametric MDM for well-separated modes. Its application to a noisy signal is compared with that of the original parametric MDM. In this example, the signal and noise are simulated separately based on pre-specified target power spectra, with phase randomly drawn from a uniform distribution. The resulting realizations of signal and noise are added point-by-point in the time domain to construct a noisy signal. Here, the noise level is detected from a region of the power spectrum above the frequency range of the clean signal and that estimate is used to reduce the noise using this newly proposed methodology.

1. Simulations

Target power spectra of of both the signal and noise are shown in Figure 22. The target power spectrum of signal is bi-modal and has been estimated from real measurements of vortex induced vibration of a marine riser. A white-noise power spectrum is used to simulate noise for addition to the signal over the frequency range from 0.0 to 0.7 Hz, which is larger than the range of signal: 0.0 to 0.4 Hz. An alternative case used to evaluate the effectiveness of the method for detection of a known frequency shift is generated separately. The signal power spectrum is modified as if the stiffness of the measured system has been decreased. Both modal frequencies are decreased by 10% of the area under a normal probability density function. Applying Equation 6.13 with $\Phi(z) = 10\%$ of the area under a standard normal distribution yields

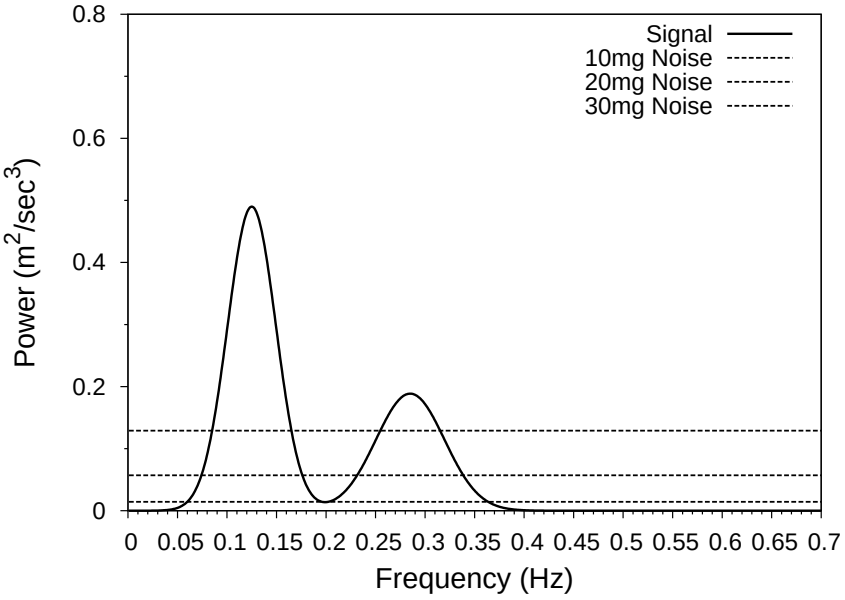


Fig. 22. Target Signal Spectra and Noise Spectrum

$\Delta\mu$ = -0.1257 s_i.

$$\Delta\mu_i = zs_i; \qquad \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \tag{5.38}$$

Transforming this value to each modal distribution of the original target spectrum (S) results in the new mean frequencies of the modified target spectrum, S_3, shown in Table I. In addition to this mean change, the variance of each mode of the modified target spectrum is increased from the original by 10% while holding the total energy of the spectra constant.

Four families of realizations based on the original target spectrum of the signal or the modified spectrum of the signal are corrupted by realizations of the noise spectrum with four different RMS values; 1, 10, 20, and 30 mg. Consequently, a

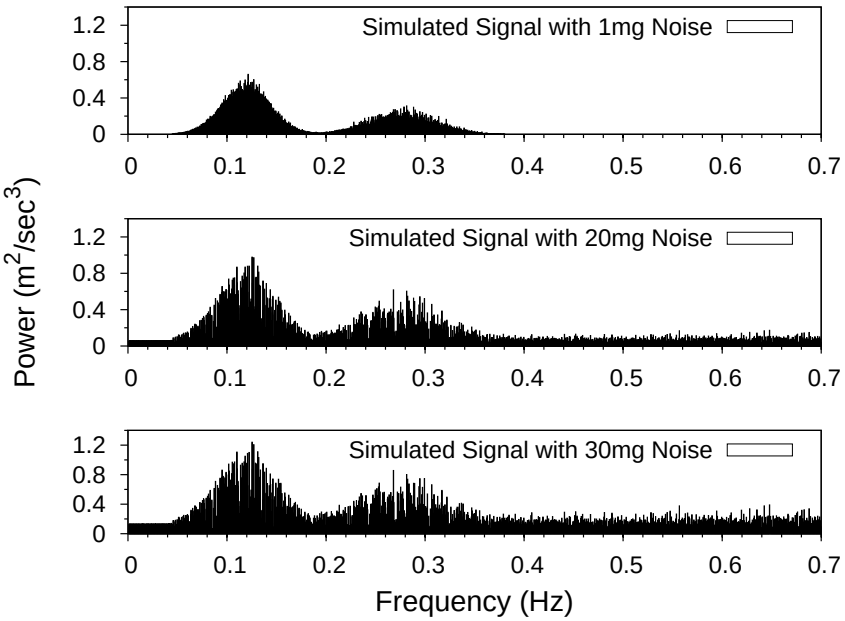


Fig. 23. Sample Signal Spectra with 1, 20, and 30 mg RMS Noise

total of five cases are tested, including realizations of the clean signal. Each case is statistically analyzed through 25 independent realizations and subsequent tests of the noise reduction methodology coupled with the MDM. All realizations have a time step of 0.2 sec. Sample power spectra of realization for signal corrupted by 1, 20, and 30 mg RMS noises are shown in Figure 23.

2. Application of MDM to Signal Corrupted by Noise

The original parametric MDM without noise reduction is applied to the simulated signals. Later, these results are are compared with the results including application of the new noise reduction methodology. The probability of Type I error (false alarm) is also investigated.

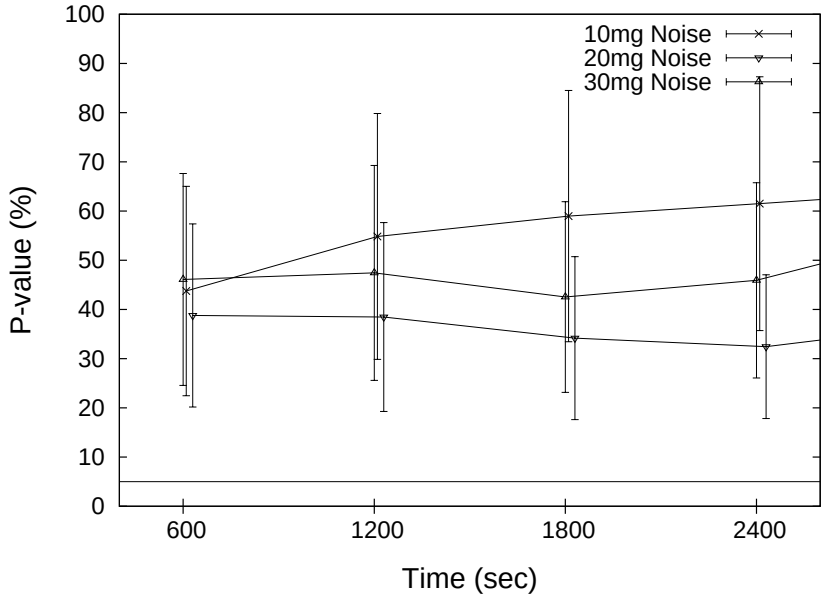


Fig. 24. Mean and Standard Deviation of 25 Significance Levels Resulting from Parametric MDM Applied to Signals Whose Underlying Modal Distributions Are Changed

Results of application of the original MDM without noise reduction are shown in Figure 24. The development of the original parametric MDM is based on the assumption of well-separated Gaussian modes. Increasing levels of noise challenge both of these assumptions: the peaks become less well-separated and the distribution becomes less Gaussian. As such, the original method would not be expected to be appropriate for a signal corrupted by large amounts of noise. The sensitivity to noise is investigated by application of the parametric MDM to the simulated noisy signals. The Figure 24 shows that high noise corruptions (10 to 30 mg RMS) prevent the parametric MDM from detecting the same changes in underlying modal distribution of signal. There are almost no decrease or increase in the significance level with

increasing time duration. In all the following figures, the symbols represent the mean of 25 computed significance levels and the error bars indicate one standard deviation. In very high noise cases, neither the significance level nor its uncertainty decreases with time duration. This trend indicates the method will be ineffective regardless of the amount of data available.

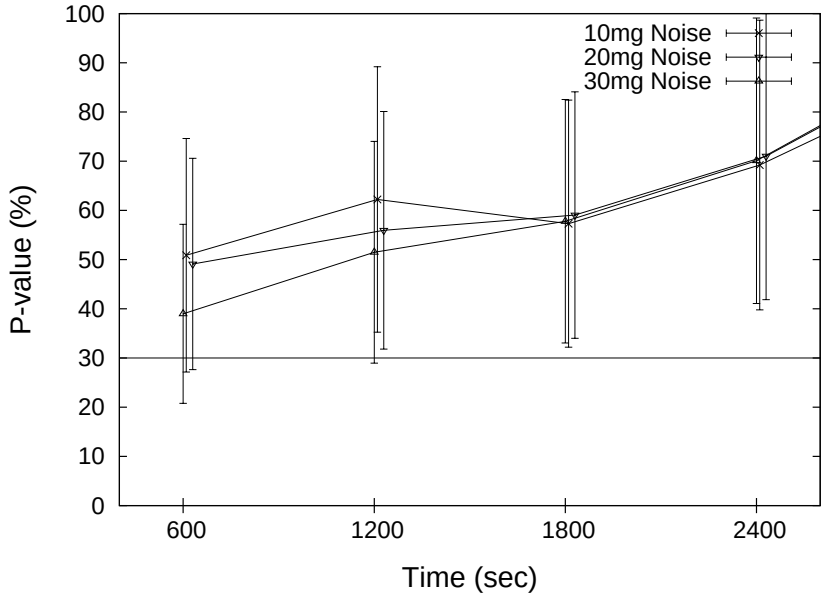


Fig. 25. Mean and Standard Deviation of 25 Significance Levels Resulting From Parametric MDM Applied to Signals Whose Underlying Modal Distributions Are the Same

In high noise situations, the individual modes resulting from the penalty method process have very non-Gaussian shape because the would-be tails are trimmed vertically at the separation point. Thus, the most possible explanation for the break-down of the method with high noise is the underlying requirement of parametric test, the Gaussian distribution, is increasingly violated, as shown in Figure 23.

To test the robustness of the original MDM without noise reduction against Type I errors (false alarms), an additional 25 realizations of the original, unmodified target signal spectrum are generated and compared with the previous 25 realizations of the target signal spectrum. A Type I error (false alarm) would be indicated if the method detects a statistically significant change when in fact there is no real change in the underlying modal distributions. The results of application of the MDM to these statistically similar time-histories are shown in Figure 25. The resulting probability of false alarms decreases as time duration increases. Considering both Figures 24 and 25, the noise corrupts the signal sufficiently that there is no appreciable difference in P-value predicted for signals that do or do not have a frequency shift in the underlying power spectra. The original MDM is clearly ineffective when applied to high-noise signals.

3. Application of Noise Reduction Methodology

Noise reduction methodologies are applied to decrease the noise corruption shown in Figure 23. The conventional and new noise reduction methods are applied to the same sample noisy spectra, resulting in the cleaned spectra shown in Figures 26 and 27, respectively. Both of the cleaned spectra appear very similar. One important difference is that the spectra cleaned using the new method includes more energy preserved from the original signal. The additional components present in results computed using the new method enable accurate variance computations for each modal distribution. The conventional methodology is known to underestimate the variance of each mode and underestimation of this variance results in under-prediction of the P-value by the T-statistic. However, this trend is not realistic: it can not be that the higher noise (somehow) enables the better signal processing. The new method developed here has the advantage that the area under the PSD is preserved, so the variance will not be

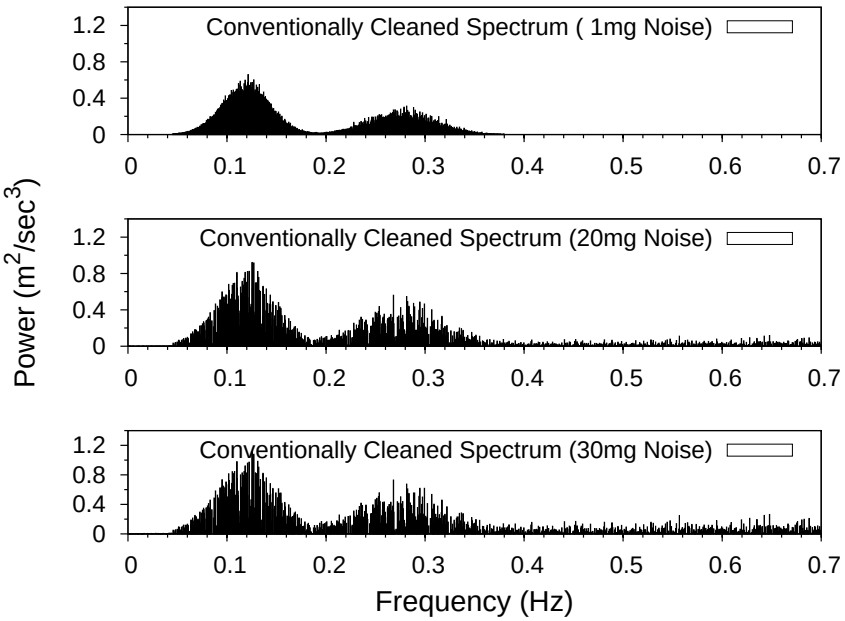


Fig. 26. Sample Cleaned Spectra of Conventional Noise Reduction Method

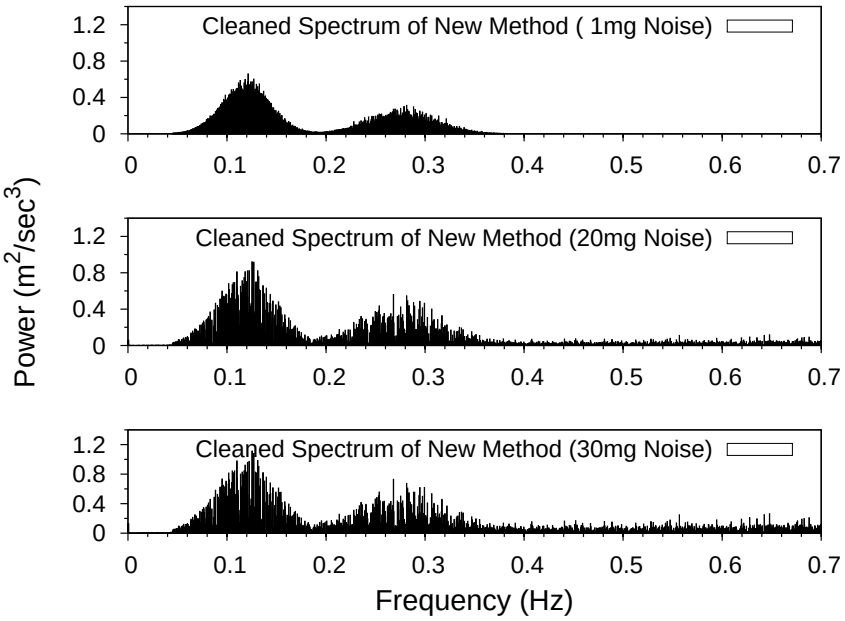


Fig. 27. Sample Cleaned Spectra of New Noise Reduction Method

systematically underestimated. P-values resulting from application of the conventional method are smaller in Figure 28. Additionally, underestimating the variance increases the probability of Type I errors in the conventional method (Figure 29).

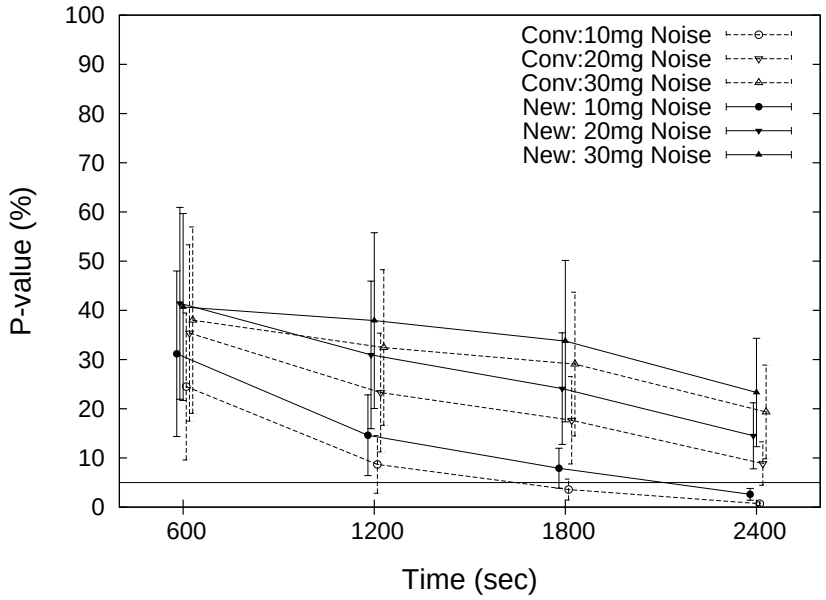


Fig. 28. Comparison of Detection Ability of MDMs Using Conventional and New Noise Reduction Methods

Now, the parametric MDM is again tested against the same noisy signals, but this time the new noise reduction methods described in Sections C. 3 are implemented and applied. The same time-histories are used, and results are available for comparison with the original parametric MDM. Figure 30 shows the resulting P-values. Visually comparing with Figure 24 shows that the noise reduction method markedly improves the MDM results: the computed significance level and its uncertainty are both decreased for all noise levels. Similarly, both the significance and its uncertainty now

decrease as time duration increase, with the implication that sufficient data might enable statistical detection of this change despite the very high noise levels, and in the case of 10 mg RMS noise, it has already been overcome with application of 2,400 seconds of data. Even with very high noise levels, there still remains an inverse proportional relationship between significance level and time duration although longer realizations or monitoring is required for detecting the same changes.

Figure 31 is used to investigate the probability of a Type I error being made by the MDM on very high-noise signals, e.g., the likely-hood that a statistically significant change is detected when in fact there is no change in the underlying target spectrum. The figure shows the probability is low and continues to drop with increasing amounts of data.

a. Signal to Noise Ratios and Minimum Data Requirements

Signal to Noise Ratio (SNR) is a conventional measure of the amount of noise present in a measured signal. This ratio is meaningful in an MDM analysis only if applied to single modes. Table V shows the SNR for each mode, with the noise and signal both considered only in the frequency range of the signal: 0.05 to 0.2 Hz for the first mode and 0.2 to 0.39 Hz for the second mode (Figure 22). Figure 32 shows the significance levels for the same noisy data, but this time each mode is considered individually. The white noise assumption applies a constant noise level over a broad range of frequencies. Since the second mode is smaller than the first, or has relatively less signal, the SNR of the second mode is smaller and therefor the noise degradation is comparatively worse. As expected from the smaller SNR, the significance level and its uncertainty for the second mode are both larger than for the first.

Moreover, the figure indicates that both significance level and its uncertainty (standard deviation) decrease as the time duration increases, except for the case of

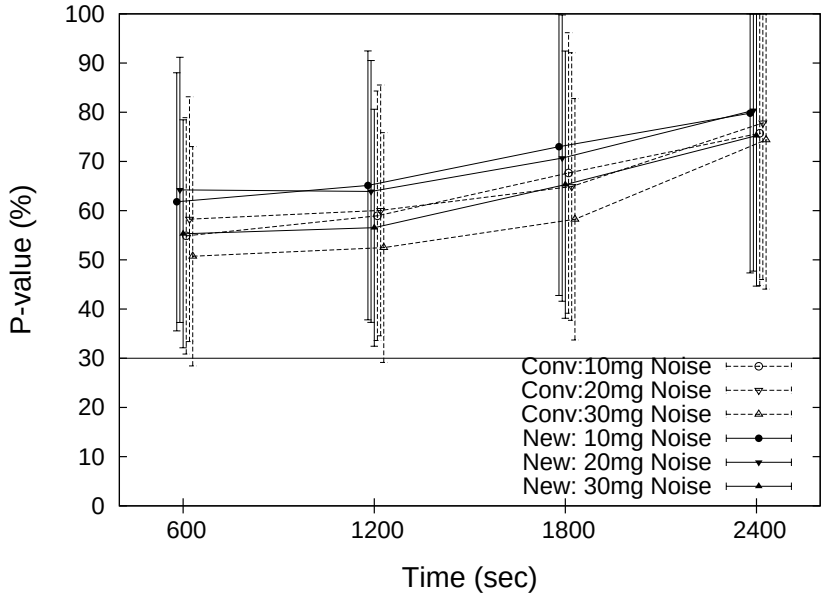


Fig. 29. Comparison of Probability of Type I Errors Between MDMs Using Conventional and New Noise Reduction Method

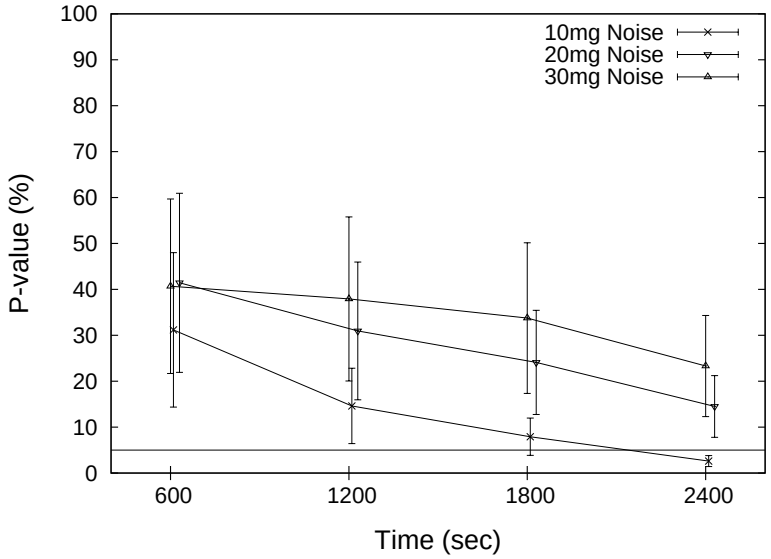


Fig. 30. Mean and Standard Deviation of 25 Significance Levels Resulting From Noise Reduction Method Implemented and Applied with the Parametric MDM.

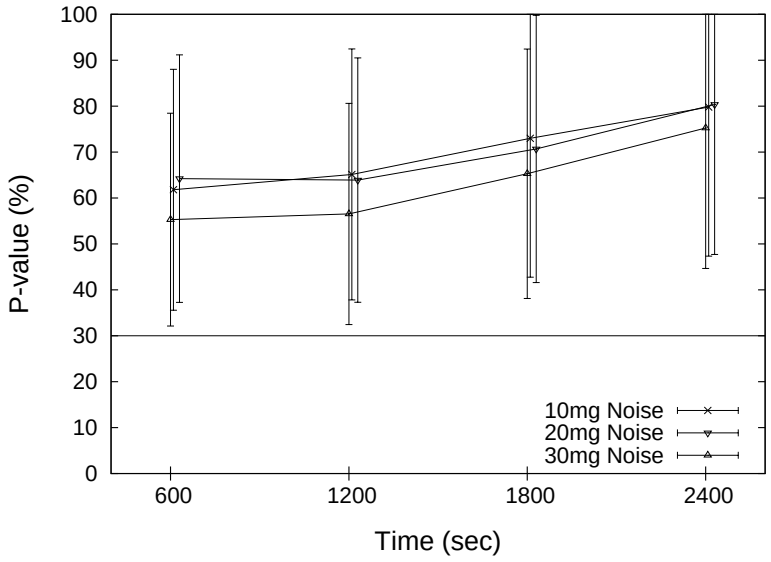


Fig. 31. Mean and Standard Deviation of 25 Significance Levels Resulting From Noise Reduction Method Implemented Parametric MDM Applied. No Change in the Underlying Modal Distributions of the Clean Signals.

the second mode with 30 mg RMS noise, whose SNR is below one. This result implies that the noise reduction method is ineffective regardless of time duration when the energy of the noise is greater than that of the signal: the noise level is so high that it corrupts signal beyond recovery.

Finally, Figure 33 shows the probability of false alarms. False alarms are not a concern for any of the cases presented here if the time duration is longer than 1,200 seconds.

Figures 30 and 31 can be used to investigate minimum data requirements necessary to obtain acceptable significance levels considering the possibility of both Type I and Type II errors. Using the full predictive power of both modes and an ideally

Table V. Signal to Noise Ratio of Each Mode

Mode	10 mg RMS	20 mg RMS	30 mg RMS
1st	13.69	3.56	1.06
2nd	5.51	1.43	0.43

clean signal results in a minimum monitoring time of 1,800 seconds. Relatively high noises require longer monitoring time: 2,400 sec for a noise level of 10 mg RMS. For noise levels higher than 10 mg, these methods would not be expected to yield useful results and are not recommended unless more than 2,400 seconds of data is available.

E. Discussion

Undesirable noise in field measurements is unavoidable, though it should be reduced as much as practical to minimize degradation of analysis results. Spectral subtraction methods offer simple and effective methodologies to reduce additive white noise. The new noise reduction methodology introduced here is a variation on these methods. That difference is that negative energy density is interpreted as being caused by phase differences and recognizes that the absolute value of that area is meaningful as energy of the signal. This interpretation avoids possibility of unreasonable P-value due to underestimation of variance in conventional method. Additionally, the noise level is identified outside frequency range of signal in this methodology.

Without noise reduction, the original parametric MDM proved to be less effective in the presence of high noise. With very high noise, the significance level and its uncertainty are not observed to decrease with increasing time duration (Figure 24), which implies the method will not work on this noisy data regardless of the amount of data available. Implementation of the new noise reduction methodology to the para-

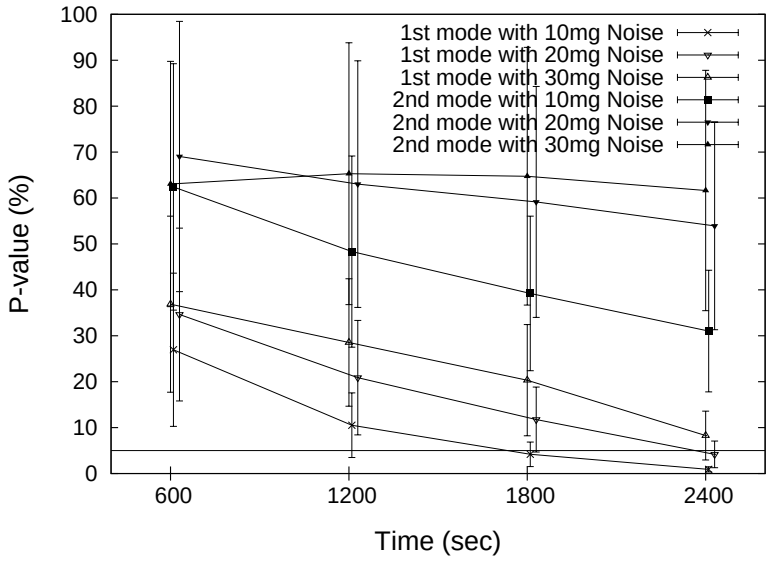


Fig. 32. Probability of Detecting a Known Shift in the Target Spectra. First and Second Modes Considered Individually.

metric MDM meaningfully improves analysis results for noisy measurements. With noise reduction applied, the desired inverse relationship between significance level and time duration (Figure 30) is observed. Higher noise requires longer monitoring time, but there remains a limit of maximum noise level, which is apparently around SNR of one (Table V). Excessive noise not only reduces the probability of detecting a known shift in the underlying power spectrum, but also is observed to increase the probability of false alarms (Figure 31). The method described in this chapter still requires that the shape of the peaks after noise reduction be Gaussian. Unfortunately, that requirement is increasingly poorly met as the noise level increases (Figure 27). Another methodology, the nonparametric MDM is introduced in the following chapter. This alternative is intended for use on non-Gaussian distribution shapes.

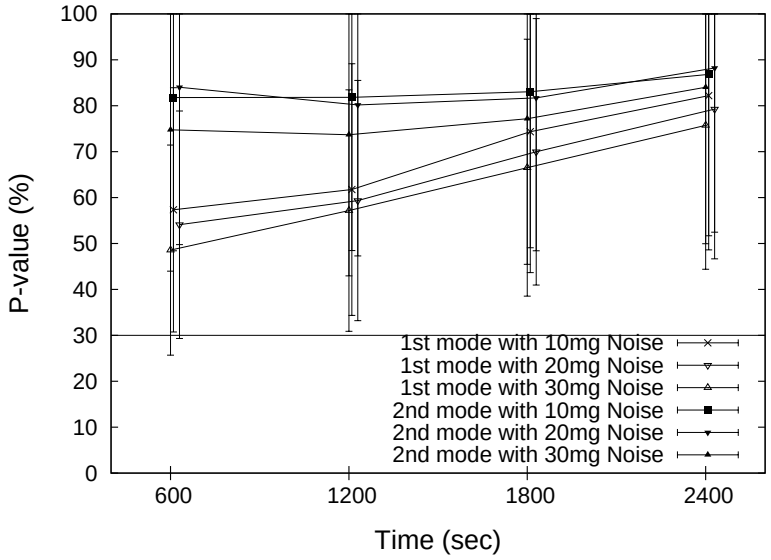


Fig. 33. Probability of Type I Errors. First and Second Modes Considered Individually. No Change to Underlying Target Spectra.

CHAPTER VI

NONPARAMETRIC MODAL DISTRIBUTION METHOD

In previous chapters, it is assumed that individual peaks in the power spectrum associated with discrete vibrational modes are well-separated, whether these peaks have a Gaussian or non-Gaussian shape. It has been shown that for well-separated peaks, the statistical significance of observed changes to separated individual modal peaks can be determined using the parametric statistical T-test, and that this parametric method is more powerful than conventional non-parametric comparisons of the data. In the event that the peaks do not have a Gaussian shape, the resulting non-Gaussian modal distributions can be transformed into their Gaussian equivalents prior to use of the T-test.

Unfortunately, for closely spaced modal frequencies the distance between modal peaks is not adequate for including sufficient distribution tails. In these cases, parametric tests are less reliable than nonparametric tests because distributions being compared are not Gaussian and cannot be readily transformed into Gaussian shapes. In this chapter, an alternative implementation of the modal distribution method is developed such that the parametric T-test is applied to the distribution of the means of modal frequencies rather than to the underlying distribution associated with the power spectrum itself; this modification enables the central limit theorem to be realistically invoked. The resulting method is equally applicable to modal distributions with Gaussian or non-Gaussian shapes as well as to those with very closely spaced peaks because the central limit theorem guarantees that the distribution of mean modal frequencies approaches the Gaussian regardless of the modal distribution shape.

A. Introduction

In a typical statistical comparison of experimental results, data collected from a treatment group is compared to that collected from a control group with the aim of determining if the treatment produces "meaningfully" different results from the control group. If the outcomes in the treatment group are not found to be significantly different from those of the control group, then there is no statistical reason to believe the treatment has had any effect. Similarly, MDM compares data measured at one time with data measured at another time with the aim of detecting any system changes. If the measurements are not significantly different, the state of the monitored system is assumed to be unchanged.

The individual bars of a discrete power spectrum of field data generally appear as a cluster of energy around a specific frequency. MDM treats the mean frequency of an individual mode as a random variable and observed cluster of energy as statistical distributions. The benefit to this interpretation is that rigorous statistical analysis can be applied to assess the significance of apparent differences between modal distributions. For underlying distributions having a Gaussian shape, the parametric T-test is applicable and has been shown to be generally more powerful than non-parametric tests (Section II). The T-test is a parametric statistical methodology that has been developed to compare the means of two data-sets sampled from an underlying normal distribution, and it includes a correction for small sample size.

Application of any parametric test, however, requires that the underlying distribution of the data meets the requirement of the test, e.g., a normal distribution for the Student T-test. Use of a parametric test on data not conforming to the appropriate underlying distribution is not appropriate and will generally result in either a increase in the number of Type I or Type II errors. Nonparametric statistical tests

provide a viable alternative. These tests can be applied directly to a non-Gaussian modal distribution or one with very closely spaced peaks with no fundamental change to the methodology because non-parametric tests are independent of the underlying distribution shapes.

Here, the mean frequency of each mode is computed for sequential estimates of the underlying power spectrum such that the requirements of the central limit theorem are met and Gaussian statistics can then be applied to these collections of mean values. Consequently, the T-statistic can be appropriately applied to compare the distributions of sample averages to detect system changes. For rigorous application of the central limit theorem, it is important that consistent frequency ranges be used such that the mean of each mode is consistent, i.e., that it is a collection of mean values of the energy detected between two fixed frequencies.

B. Review: Statistical Comparison

The overall objective of a statistical comparison is to determine whether or not two sets of data are drawn from the same population or from different populations. In statistical terms, the null hypothesis is that the two data sets are drawn from the same population; the statistical tests are then applied to assess whether or not there is adequate statistical reason to believe otherwise, i.e., whether or not the observed statistical differences exceed some pre-specified level of statistical significance, or P-value. Specifically, the P-value represents the probability that the observed differences between the data sets are due to random chance alone. Small P-values indicate the observed differences between the data-sets is very unlikely to have been caused by random chance alone and it is therefore unlikely that the null-hypothesis is correct.

1. Underlying Theory: Central Limit Theorem

The Central Limit Theorem (CLT) underlies most non-parametric statistical methods. Simply stated, the CLT shows that if a large enough number of sample sets is drawn from a population, then the distribution of the sample averages approaches a normal (Gaussian) with increasing sample size, regardless of the distribution of the underlying populations [13]. Here, as in so many non-parametric methods, the CLT is applied to allow use of a parametric test on non-parametric data. In the MDM, the distribution of mean frequencies meets the requirement of the CLT because each mean itself represents a sample average: here, the average of the energy surrounding the modal mean. This observation enables appropriate application of the T-statistic regardless of the normality of the shape of each modal distribution. For rigorous application of CLT, it is important that consistent frequency ranges be used such that the mean of each mode is consistent, i.e., that it is a collection of mean values of the energy detected between two fixed frequencies.

2. Nonparametric Comparison of Two Histograms

Comparing two histograms is the same as comparing two distributions of grouped data, or, in statistical terms, investigating the homogeneity of two samples of grouped data. The most commonly described nonparametric statistic for comparing two distributions is Kolmogorov-Smirnov (or K-S) test. The test is applicable to data sets where each data point can be associated with a single number. In such a case, the list of data points can be converted to a cumulative distribution function estimator $S_N(x)$ of the probability distribution from which it was drawn: If the N data points have values $x_i, i = 1, ..., N$, then $S_N(x)$ is the function giving the fraction of data points to the left of a given value x. This function has step function shape. The K-S statistic is

the maximum value of the absolute difference between two cumulative distributions. The K-S static for comparing two cumulative distributions $S_{N_1}(x)$ and $S_{N_2}(x)$ is

$$D = max|S_{N_1}(x) - S_{N_2}(x)| \tag{6.1}$$

The significance level of an observed value of D is given approximately by the formula [75].

$$\text{Probability}(D > observed) = Q_{KS}\left(\left[\sqrt{N_e} + 0.12 + 0.11/\sqrt{N_e}\right] D\right) \tag{6.2}$$

$$Q_{KS}(y) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2y^2} \tag{6.3}$$

with limiting values of $Q_{KS}(0) = 1$ and $Q_{KS}(\infty) = 0$. The effective number of data points is

$$N_e = \frac{N_1 N_2}{N_1 + N_2} \tag{6.4}$$

where N_1 is the number of data points in the first distribution and N_2 is the number in the second.

Scholz and Stephens [76] proposed the Anderson-Darling statistic, which is a modification of the K-S test. This modification gives more weight to the tails than does the original K-S test. The Anderson-Darling test makes use of the specific distribution in calculating critical values. The modification has the advantage of allowing a more sensitive test, but has the disadvantage that critical values must be calculated for each distribution. The two-sample version is as follows:

$$A^2_{mn} = \frac{mn}{N}\int_{-\infty}^{\infty}\frac{(F_m(x) - G_n(x))}{H_N(x)\,(1 - H_N(x))}dH_N(x) \tag{6.5}$$

where $F_m(x)$ and $G_n(x)$ are the empirical cumulative distribution function of the first sample and the second sample, respectively. $H_N(x) = (mF_m(x) + nG_n(x))\,/N$ with

$N = m + n$, is the empirical distribution function of the pooled sample. This can be used to test the hypothesis that $F = G$ without specifying the common continuous distribution function.

Both the Anderson-Darling and K-S tests are based on cumulative density functions and are appropriate for comparing two distributions of one variable (one mode among possible multi-modes). The MDM, however, needs multi-modal comparison requiring an equivalent to the combined T statistic; as far as the author is aware, combined statistics for these CDF-based methods have not been developed. Further, any CDF-based method will be extremely dependent on the precision of the division between modes: inappropriately cutting or including even small amounts of energy from the left tail could greatly affect the CDF. It has been found in the work on the penalty method that very precise divisions between modes is either difficult or impossible, so it seems unlikely that any CDF-based statistical comparison will prove to be an effective part of the MDM.

3. Bootstrap Resampling

The Bootstrap is a popular alternative to parametric statistical inference for cases in which the conditions necessary for parametric methods are in doubt. Briefly stated, the bootstrap methodology assumes that the original sample represents the population from which it was drawn. Thus, re-sampling from the original sample should statistically represent what would be obtained if many samples were taken from the original population. Key to the bootstrap method is that the re-sampled data is not compared directly, instead, the means of groups of these samples are compared. This subtlety enables invocation of the central limit theorem and the associated parametric statistical comparisons. The parametric bootstrap is the practice of estimating properties of an estimator (such as its variance) by randomly sampling (with replace-

ment) from an approximate distribution and using these samples to estimate the desired parameter. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples of the observed dataset (and of equal size to the observed dataset), each of which is obtained by random sampling with replacement from the original dataset.

The advantage of bootstrapping over analytical methods is its great simplicity: it is straightforward to apply the bootstrap to derive estimates of standard errors and confidence intervals for complex estimators of complex parameters of the distribution, such as percentile points, proportions, odds ratio, and correlation coefficients. The disadvantage of bootstrapping is that while (under some conditions) it is asymptotically consistent, it does not provide general finite-sample guarantees, and has a tendency to be overly optimistic. The apparent simplicity may conceal the fact that important assumptions are being made when undertaking the bootstrap analysis (e.g. independence of samples) where these would be more formally stated in other approaches.

The details of the parametric bootstrap procedure are relatively straightforward: A distribution is created based on random sampling of the the underlying population. The method is termed parametric because a future re-sampling is based on this distribution, rather than on raw data. New (resampled) data sets are then created by repeatedly sampling with replacement from this one distribution, and the resulting data-sets are treated as if many sets had been sampled from the original population. Sampling from the distribution with replacement means that after a sample is drawn from the original sample set, it is put back before drawing the next sample. As a result, any number can be drawn more than once, as if sampling from an infinite data set.

Application of the parametric bootstrap to the MDM would be a straightforward extension of a well-known statistical method to this unique methodology, and how this implementation would be put into the MDM is described in detail in the discussion section of this chapter, Section 2. However, it is here observed that the underlying principle, the CLT, can be applied without re-sampling, and with the added advantage of maximizing the use of "original" data. The observation goes back to Chapter II and the implementation of the penalty method. In that section, it was observed that accuracy could be increased by using a variation on Welch's method of spectral smoothing: many power spectra are computed and the bars corresponding to equivalent frequencies are averaged. In the case of the penalty method, 100 separate power spectra computations were found to be adequate. In the new non-parametric method proposed here, these 100 spectra are treated as separate samples of the same quantity, and the mean of each spectral peak is treated as an estimate of the mean of the modal frequency. Since each estimate is itself a mean, the CLT can be invoked and the T-statistic can be meaningfully applied regardless of noise or closeness of the modal peaks. This new methodology has all the theoretical advantages of the parametric bootstrap but does not include the difficulty, computational inefficiency and inelegance of the re-sampling process.

C. Implementation to MDM

The central limit theorem implementation of the modal distribution method is a non-parametric variation on the original implementation. This implementation still divides a measured acceleration time-history into a series of segments; each segment is converted into a power spectrum through use of the FFT, and the resulting power spectrum is divided into a series of response frequency ranges by the same penalty

method. At this stage is where the methods diverge. In order to apply the central limit theorem, a collection of N mean values of samples of the same population is necessary. To that end, each one of the N mean values is computed as the mean (centroid) of the area under the power spectrum associated with an individual mode. In this case, the process is assumed to be constant over a very short time duration, for example, 200 data points, which in practice may be a few seconds depending on application.

The mechanics of computing each of the N means are to first compute the power spectrum of a segment of B points in the time-history, where $B >> N$. As in the original MDM, the penalty method is applied to separate the modes. The separation points between individual modes are recorded. Then the time-history is shifted two points forward in time and a new power spectrum is computed. A second set of separation points between modes is recorded and the process is repeated to develop N sets of $M+1$ separation points, where M is the number of modes. For each of the $(M+1)$ separation points, the N points are then averaged to find a single set of modal divisions that will be used for all of the N power spectra. These averaged values are then used as fixed frequency ranges for computation of the N mean frequencies, each of which is computed by averaging the area under the power spectrum within the upper and lower bounds found by averaging the modal separation points from the N power spectra, each of which is computed from B data points. The resulting N predictions of the mean modal frequency then constitute N estimates of the true mean frequency, which is the random variable in this method.

The purpose of this process is to have a collection of mean values of essentially the same population, thereby conforming to the central limit theorem. The resulting distribution of sample averages approaches the Gaussian with increasing length of the segment of the time-history, B. It is common to apply Gaussian statical methods to

distributions resulting from the CLT with sample size as small as 30 regardless of the distribution of the underlying data [26].

The overall goal of the method is to quantify the statistical significance of subtle changes in the state of a vibrating structure by comparing sampling distributions of sample average representing two different segments of a measured response time-history. Application of the CLT guarantees a near-Gaussian distribution for sufficiently large sample size, so the T-test can be rigorously applied regardless of mode shape or modal spacing.

At this stage, the CLT implementation of the MDM returns to that of the original. The sampling distribution of the sample averages in the first segment is compared with that in the second segment to detect subtle change in mean modal frequency, if any. The mean of each sequential sampling distribution in the first segment is statistically compared with that of the second segment, i.e, the mean of sample average of the first mode in the first segment is compared with that of the first mode in the second segment, as are each of the subsequent means of sample average. The T-statistic is applied to quantify the significance of observed differences of means of sample average. The significance of each mean difference between sample averages is computed independently, and then each mean difference is weighted by the magnitude of the energy associated with that mode to compute an overall combined T-statistic. The statistical significance of any difference between means of sample average is reported, as well as a combined statistic quantifying differences between all modal mean frequencies. The entire procedure could be repeated over consecutive segments of a time-history to monitor for changes in modal response.

1. Statistical Comparison of Means: The T-test

At this stage, the CLT implementation of the modal distribution method reverts to the original. A combined T-statistic is computed for comparison between separate segments of the time-history. The T-statistic for modal comparisons is computed as e.g. [26]:

$$T_i = \frac{\Delta\mu_i}{s_i} \tag{6.6}$$

where $\Delta\mu_i = \mu_{i,1} - \mu_{i,2}$ and $s_i^2 = s_{i,1}^2/N_{i,1} + s_{i,2}^2/N_{i,2}$. $\mu_{i,w}$ is the mean of sampling distribution of sample average, $s_{i,w}^2$ is the sample variance of sampling distribution of sample average, and $N_{i,w}$ is the number of sample averages in the i'th mode of the w'th segment. T_i is distributed approximately as a Student's T with the number of degrees of freedom for the i'th mode equal to e.g. [26]:

$$DOF_i = N_{i,1} + N_{i,2} - 2 \tag{6.7}$$

Accurate assessment of the significance of changes in the observed vibrational response requires consideration of all modes. The ensemble of observed differences for each pair of means (one mean for each mode from each of the two segments of the time-history) is treated as a set of repeated measurements with differing uncertainties, i.e., observed differences between the first pair of sampling distributions are combined with observed differences between pairs of sequentially higher modes. An overall P-value is calculated by computing a combined T-statistic, which weights differences between the means by the fraction of energy represented for each individual mode. The fraction of energy associated with the i'th mode for each of the two windows is:

$$E_{i,w} = \frac{A_{i,w}}{\sum_{i=1}^{I} A_{i,w}} \tag{6.8}$$

where I is the total number of modes in the power spectrum and w is either 1 or 2, denoting either the first or second window of the time-series. The overall T-statistic is found using Equation 6.9, in which $\Delta\mu$ and s^2 are computed as averages weighted by the fraction of energy associated with each of the I modes:

$$T = \frac{\Delta\mu}{s} \tag{6.9}$$

where

$$\Delta\mu = \sum_{i=1}^{I} 0.5(E_{i,1} + E_{i,2})\Delta\mu_i \tag{6.10}$$

$$s^2 = \sum_{i=1}^{I} 0.5(E_{i,1} + E_{i,2})s_i^2 \tag{6.11}$$

The T-statistic resulting from Equation 6.9 represents the weighted average of differences between the means of sampling distribution of sample average. The number of degrees of freedom for use of a standard T-distribution is estimated as the total number of sample averages. The total degrees of freedom to be used in conjunction with the T-statistic is estimated as the sum of individual degrees of freedom:

$$DOF = \sum_{i=1}^{I} (N_{i,1} + N_{i,2}) - 2 \tag{6.12}$$

D. Example of Well-Separated Gaussian Modes with Noise

As noted earlier, the CLT implementation of the MDM is intended to be more general than the original parametric MDM: specifically, it is applicable to closely spaced peaks and to modes with non-Gaussian shape. However, for broadly spaced peaks and power spectra having roughly Gaussian shape, the original parametric MDM would be expected to offer superior performance. Here, the relative performance of the two methodologies is investigated. Time-histories of both acceleration data and of noise are simulated, and the three methodologies are applied to ascertain their effectiveness

on noisy signals. The three competing methodologies are: non-parametric MDM (this Chapter), noise reduction MDM (Chapter V), and parametric MDM (Chapter II).

1. Simulation

Target power spectra for both of the signal and of the noise are shown in Figure 34. The target bi-modal power spectrum of signal has been estimated from real measurements of vortex induced vibration of a marine riser. The power spectra of the white noise being added to the signal are applied over the frequency range from 0.0 to 0.7 Hz, which is larger than the frequency range of signal: 0.0 to 0.4 Hz. For this example, the signal corrupted by noise is generated as follows: a time-history of the signal is simulated and stored; separately, a time-history of noise is simulated, and the resulting time-series of signal and of noise are added together point-by-point in the time-domain. In application of the MDM, the power spectrum is calculated from the time-history of the combined signal. Five different RMS noise levels are used: 0, 1, 10, 20, and 30 mg, and each of these noise levels is applied to two different "clean" power-spectra (an "original" and a "modified" target spectrum). All realizations have time step of 0.2 sec. Sample power spectra of signal combined with 1 mg, 20 mg, and 30 mg RMS noises are shown in Figure 35. Note that same signal is used for all cases. In the figure, underlying modal distribution shape is destroyed as the noise increases. 20 mg and 30 mg RMS noise changes modal distribution shapes, around both peaks and tails.

The "modified" case used to evaluate the effectiveness of the method for detection of a known frequency shift is generated using a slightly modified target spectrum. The target power spectrum is modified as if the stiffness of the measured system is decreased. Both modal frequencies are decreased by 10% of their area, based on normal probability density function. Applying Equation 6.13 with $\Phi(z) = 10\%$ of the

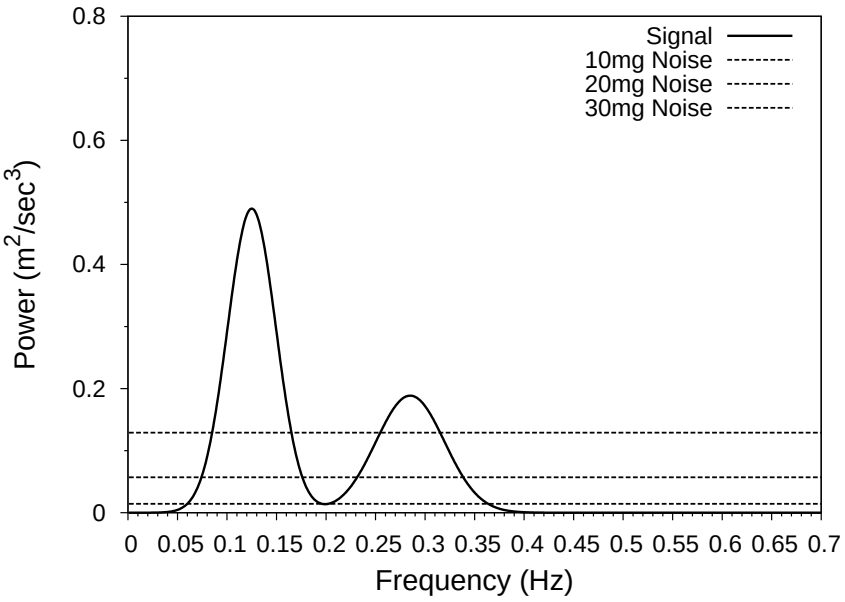


Fig. 34. Target Signal Spectra and Various Levels of Noise Spectrum

area under a standard normal distribution yields $\Delta\mu$ = -0.2533 s_i.

$$\Delta\mu_i = zs_i; \qquad \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du \tag{6.13}$$

Transforming this value to each modal distribution of the original target spectrum (S) results in the new mean frequencies of the modified target spectrum (S_3) shown in Table VI of Chapter II. Some parts are repeated here for convenience.

2. Results of Comparison

The three variations in the method are exercised on various cases and results are critically compared to assess the comparative effectiveness of each method. In each comparison, the results shown are based on 25 realizations of each case to obtain

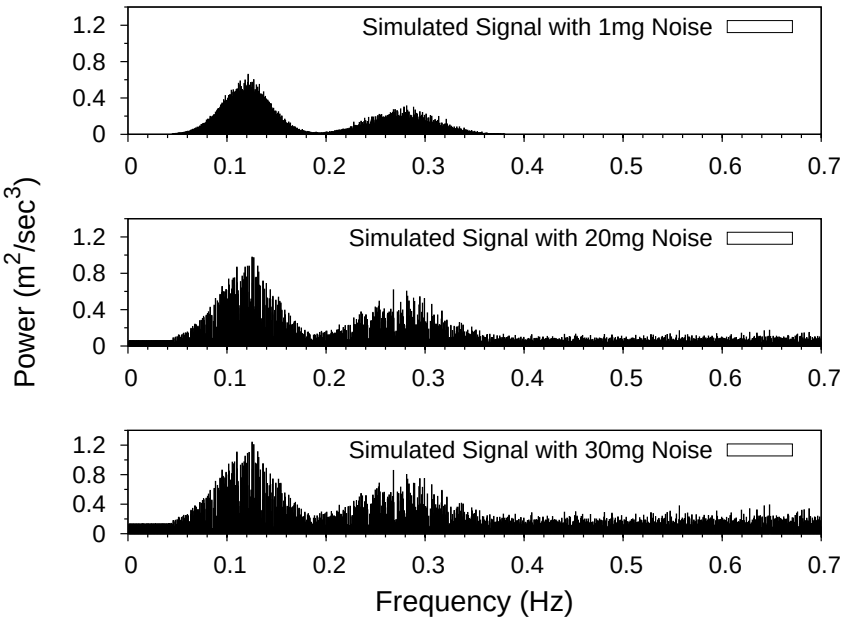


Fig. 35. Sample Signal Spectrum Corrupted by Various Levels of Noise

statistically significant results. There are exactly two ways a statistical test can fail: it can reject the null hypothesis when it is in fact true (Type I error), or it can fail to reject the null hypothesis when it is in fact false (Type II error). The null hypothesis underlying the statistical methods presented here is that there is no difference between the two segments in the time history, so a Type II error is failing to detect a change when in fact the underlying target spectrum has been changed, and a Type I error is wrongly detecting a change when in fact there has been no change to the underlying spectrum (a false alarm).

Table VI. Target Spectra Parameters with Percent Change from Original Idealized Spectrum

Target Spectrum	1st Mode			
	Mean	Change	Variance	Change
S	0.12500	N/A	0.00060	N/A
S_3	0.11879	-4.96%	0.00066	10%
Target Spectrum	2nd Mode			
	Mean	Change	Variance	Change
S	0.28500	N/A	0.00120	N/A
S_3	0.27622	-3.05%	0.00132	10%

a. Non-parametric MDM vs Parametric without Noise Reduction - Low Noise

In this first comparison, the significance level of the non-parametric MDM is compared with that of the original parametric MDM. This first comparison is a test of robustness against Type II errors. Here, both methods are applied to a clean signal, or to one with very low noise (1 mg). The noise level typical of a high-quality 1 g accelerometer is around 0.5 mg RMS. Here, additional surrounding noise is assumed as 0.5 mg, resulting in total of 1 mg RMS noise.

Figure 36 is used to compare the original (parametric) version of the MDM with noise reduction (Chapter V) with the non-parametric MDM based on the CLT. For the low-noise cases, the parametric version is seen to be more effective. This result is not surprising because it follows the general observation that parametric statistics are more effective if the data conforms to the requirements of the methodology. In this example, the target spectrum was created with a near-Gaussian shape and widely spaced peaks: exactly the situation for which the parametric method would be most

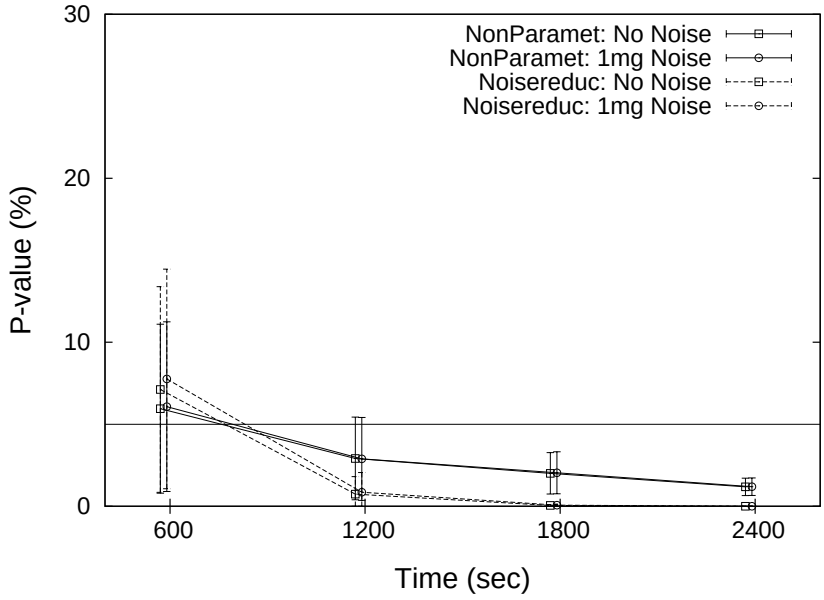


Fig. 36. Comparison of Detection Ability Between Parametric MDM and Non-parametric MDM When the Underlying Target Has Been Changed (Smaller P–values Indicates Greater Ability).

effective. The non-parametric MDM requires longer than 1200 seconds monitoring time to detect the known change with very high confidence ($\pm 1\sigma$ error bars fully below 5% P-value), while the parametric MDM requires less than 1200 sec monitoring time. Surprisingly, the non-parametric method is more effective with a 600 second signal. One possible explanation may be that the power-spectrum computed from such a short time-history does not have sufficient data to reconstruct the underlying Gaussian modal distribution very well.

The probability of Type I error (a false alarm) is shown in Figure 37. The probability of a false alarm is nearly the same for both methods and therefor does

not affect required minimum monitoring time. All methods have the pleasing trend of decreasing probability of a false alarm with increasing window size (number of data-points).

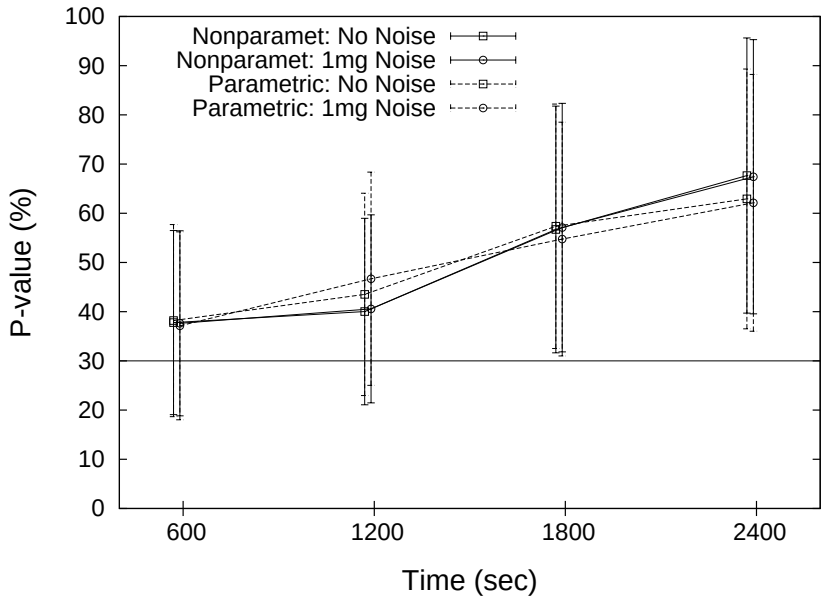


Fig. 37. Comparison of P-values Between Parametric MDM and Non-parametric MDM When Underlying Modal Distribution of Signal is Not Changed. Bigger Numbers Indicate Less Probability of a False Alarm.

b. Non-parametric MDM vs Parametric without Noise Reduction - High Noise

Very high levels of noise added to the clean signal cause the originally Gaussian shape of the distribution predicted from the power spectrum to be corrupted to a non-Gaussian. In the presence of such noise, the underlying assumptions inherent to the original parametric MDM are strongly violated as shown in Figure 35. In this example, the effects of this violation are investigated.

Figures 38 and 39 show comparisons of the effectiveness of the two methods. Figure 38 shows that parametric MDM is ineffective at detecting the change in the original signal in the presence of high noise. The significance level found using the parametric MDM is very high (generally around 50%) and there is no compelling downward trend with increasing window size; further, there is not even a downward trend in the size of the error bars, which represent the uncertainty. This is compelling evidence that the original MDM is not robust against very high levels of noise in the signal.

Compared with the parametric MDM, the nonparametric MDM remains reasonably effective at detecting the known change in the signal target spectrum in the presence of high noise. With 10 mg RMS noise, 2,400 sec monitoring time results in detection of given changes in underlying modal distributions. 20 mg RMS noise corrupts signal to the point that detection of the changes requires a window longer than 2,400 sec. Corrupting the signal with 30 mg RMS noise creates a signal which may or may not be detectable, though there is a promising downward trend with increasing window size.

In Figure 39, the probability of a false alarm is investigated. The presence of high noise increases the likelihood that the nonparametric MDM will create a false alarm. The observed P-values are much lower than those of the parametric or low-noise scenarios, but the performance of the method may still be acceptable for practical applications. The portion of the error bar below the 5% P-value is a rough indicator of the probability of a false alarm, and that portion is relatively small for long time-series: above about 1800 seconds the likelihood may be acceptable for even the 30 mg case.

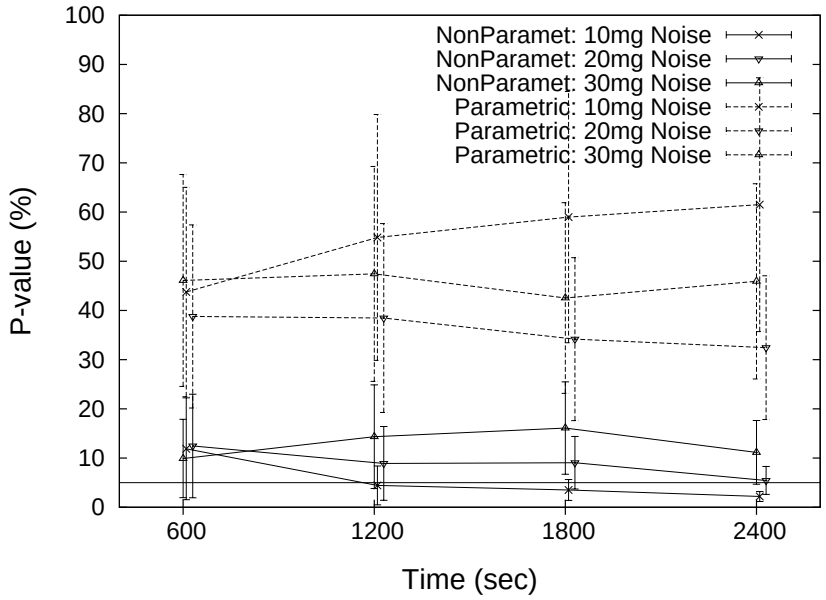


Fig. 38. Comparison of Detection Ability Between Parametric MDM and Non-parametric MDM When Underlying Modal Distribution of the Signal is Changed and the Signal is Corrupted by High Noise (Smaller P-values Indicates Greater Ability)

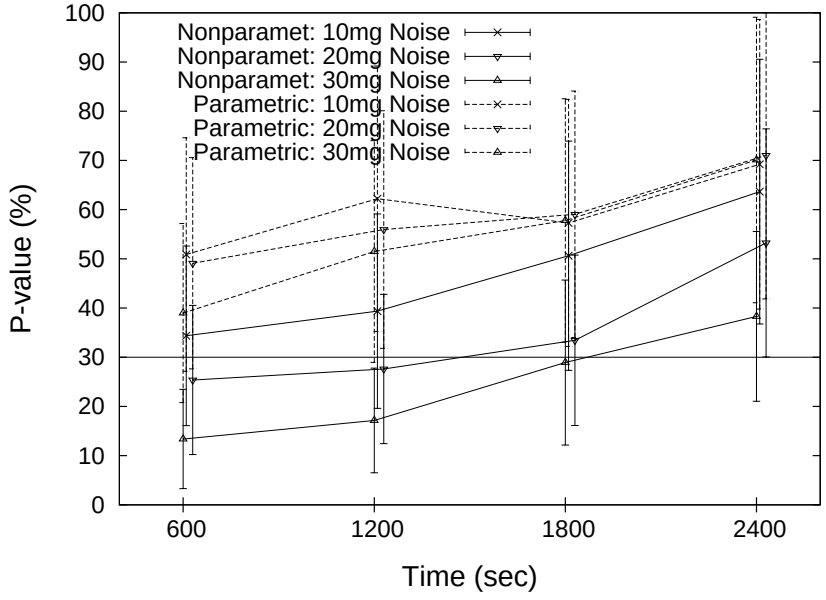


Fig. 39. Comparison of P-values Between Parametric MDM and Non-parametric MDM When Underlying Modal Distribution of Signal is Not Changed

c. Non-parametric MDM vs Parametric with Noise Reduction - High Noise

As in the last example, very high levels of noise added to the clean signal cause the originally Gaussian shape of the distribution predicted from the power spectrum to be corrupted to a non-Gaussian as shown in Figure 35. It was shown that the non-parametric adaptation of the MDM is far more effective at handling these noisy signals. Methods to remove the noise from a signal were investigated in Chapter V. Thus, two of the methods that are available to handle noisy signals are the non-parametric MDM or de-noising the signal and applying the parametric MDM. These two possibilities are compared in Figures 40 and 41.

The results shown in Figure 40 show that the detection ability of the non-parametric MDM is better than the parametric MDM even if noise removal is applied to the signal. Parametric MDM with noise reduction assumes modal distribution after noise reduction is not far from Gaussian but it is not. Unfortunately, Figure 41 shows that there is also a lower probability of false alarms if the noise reduction is applied rather than the non-parametric MDM. This example is therefor inconclusive as to which method would be preferred: here, it would depend on the relative importance of detection ability vs likeliness of false alarms.

E. Discussion

1. The Central Limit Theorem Implementation of the MDM

In this section and elsewhere in the thesis, it has been shown that if the requirements of original parametric MDM is satisfied, most notably that the power spectrum computed from the time-series is in fact Gaussian in shape, then the parametric MDM is reliable to detect given changes in underlying modal distributions. However, when that assumption is violated by high noise, the MDM modified to include application of

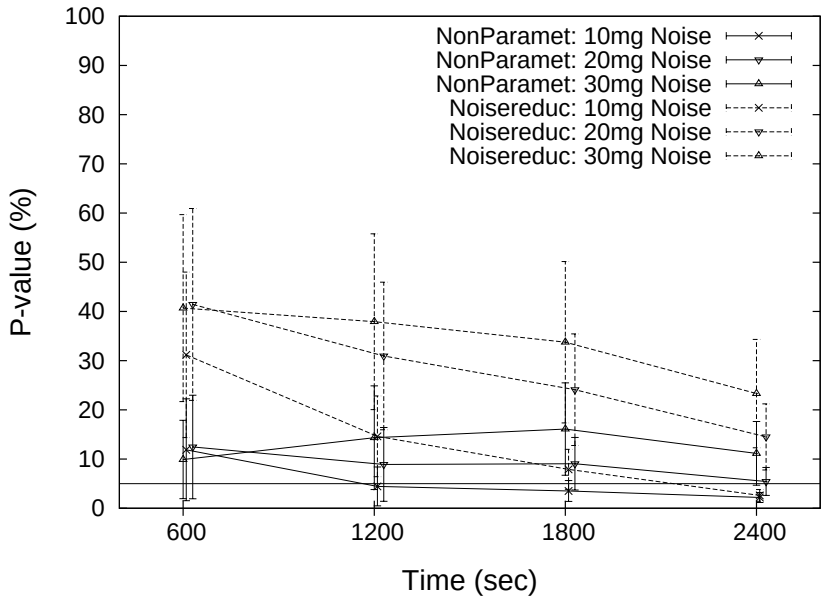


Fig. 40. Comparison of Detection Ability Between Non-parametric MDM and Noise Reduction MDM When Underlying Modal Distribution of the Signal is Changed and the Signal is Corrupted by High Noise (Smaller P-values Indicates Greater Ability)

the CLT provides a strong alternative. Indeed, the non-parametric MDM is observed to be a versatile method because it does not have any specific requirements regarding underlying distributions and in theory should be equally applicable to cases such as closely-spaced modes, noise corruption, and non-Gaussian modal distributions.

When the requirements of parametric MDM are met, the parametric MDM is recommended. For closely spaced peaks, the CLT MDM is expected to be the only viable option at this time, though definitive testing remains underway. For applications with broadly-spaced peaks but very high noise, either the parametric MDM with noise reduction or the CLT MDM provides a viable option. There is no general rec-

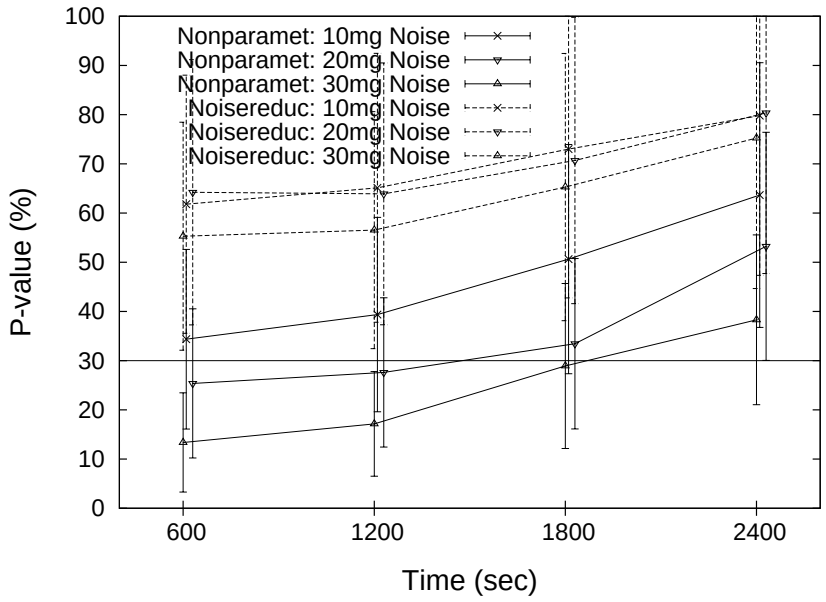


Fig. 41. Comparison of P-values Between Noise Reduction MDM and Non-parametric MDM When Underlying Modal Distribution of Signal is Not Changed

ommendation that one of two methods is better than the other. The non-parametric method is more effective at detection of actual changes in the underlying spectra, but unfortunately also has a higher probability of a false alarm. Which of these two methods should be applied would depend on the application: the non-parametric MDM for applications where detection of changes is more important and noise reduction MDM for applications where false alarms are more serious.

Overall, the non-parametric MDM is preferred because it is more general, while the noise reduction MDM is specific for signals where an underlying Gaussian shape has been corrupted by noise. At some time in the future, applicability of the non-parametric MDM to strongly non-Gaussian mode shapes or to closely-spaced modes

will be investigated to verify the generality of the non-parametric MDM. It would also be possible to combine those two methods.

2. MDM and a Semi-Parametric Bootstrap Hypothesis Test

The effectiveness of the CLT implementation to the MDM, particularly in the presence of high noise, was found to be something of a pleasant surprise. Unfortunately, this pleasant surprise was discovered well-into the Doctoral research process. The effectiveness, relative ease of implementation, and theoretical robustness for unusual mode shapes and high-noise situations strongly suggests non-parametric methods deserve further investigation at some time in the future. One avenue that seems particularly relevant is the Bootstrap re-sampling [77, 78]. One existing implementation of the bootstrap is to sample data by simulation from an existing distribution, and then using that data with conventional statistical tools. Here, the two distributions to be compared would be the two power spectra computed from contiguous segments of the time-history. This combination of the newly developing MDM and existing bootstrap statistical theory is a very logical extension of the present work and would be relatively easy to implement at some future time. The theory is outlined here, but the concept was developed relatively late in the Doctoral studies and so no worked example is included.

As in the other implementations, the bootstrap implementation of the modal distribution method divides a measured acceleration time-history into a series of segments. Each segment is first converted into a power spectrum through use of the FFT. After getting the first power spectrum, the next power spectrum is calculated from data window shifted slightly forward in time. Power spectra from a large number of windows are then averaged, which has been found to reduce variability. As noted in Chapter II, this reduction in variability is due to minimizing the effects of

incomplete cycles in the time-history leading to excessive noise in the modal response. The resulting averaged power spectrum is then divided into a series of response frequency ranges by penalty method. Each separated modal distribution is mapped into a statistical distribution in which the random variable is the modal frequency. In the bootstrap application, the separated modal peaks are treated as individual distributions. New data-sets are then sampled from these distributions as is typical in a parametric bootstrap application.

The details of the data-sampling is as follows: A cumulative density function (CDF) is constructed from each of the two modal peaks to be compared. The cumulative density (the vertical axis of the CDF) is then assumed to be uniformly distributed between zero and one. The CDF is then used to directly map the fractiles of the uniform distribution to those of the target distribution based on the power spectrum for each of the modal peaks. The result is a simulated data-set following the distribution extracted directly from the power spectrum. At this stage two options exist: collect a large number of sub-samples, e.g., 10,000 samples of six data points each such that the T-statistic can be applied, or alternatively use other conventional non-parametric tests directly on the simulated data-set.

Whichever method is selected, the overall goal is to quantify the statistical significance of subtle changes in the state of a vibrating structure by comparing simulated distributions sampled from the distribution shape computed from the data as a power spectrum.

CHAPTER VII

SUMMARY

A. Conclusions

A new statistical method for continuous monitoring of measured vibration data, called the Modal Distribution Method (MDM), has been developed. This new technique applies statistical methods in order to compare power spectra computed from measured structural response and to quantitatively assess the significance of observed changes in the power spectra. This new method differs from conventional methods since the statistical significance of changes in the response is assessed directly without consideration of the excitation; conventional methods typically either assume white-noise excitation or measure excitation directly to compute a structural response function. The benefit of using only the measured response is that there is no need to measure the excitation or to make assumptions about its spectral shape.

The only requirement for the MDM to be useful is a stationary excitation process. The shape of the excitation power spectrum is believed to be less critical than the stationarity because it is reasonable to assume that any changes in the response for stationary input are due to changes in structural parameters. However, some knowledge of the excitation can be applied in the method: if the response at some frequencies is known to be dominated by a narrow-banded excitation, which may be non-stationary, then the associated response within that known frequency range can easily be excluded from the MDM analysis. As part of the method, physical modes of vibration observed in the response power spectrum are separated and treated individually to assess the statistical significance of observed changes. Statistically combining the results of multiple modes enhances the ability to detect known changes

in the response, and wide-band excitation will enhance the participation of multiple modes.

The MDM starts from the idea that the measured vibratory responses of a complex structure include a series of modal responses, and each modal response has its own peak in power spectrum. One global minimal point exists between two adjacent peaks and a penalty method finds the global minimum point between the roughly guessed two peaks. The frequency range between two adjacent global minimal points found in the power spectrum is considered as a statistical distribution; the distribution is also called modal distribution, and it represents the individual modal response including the peak inside the frequency range. Finally, two corresponding modal distributions obtained from two measured responses are compared by a statistical test. For instance, the modal distribution representing the first mode of the first measurement is compared with the one representing the first mode of the second measurement.

The resulting significance indicates whether or not the underlying modal distribution has changed. Changes in the underlying modal distribution imply changes in the system being monitored, such as changes in the structure or in the excitation load. The statistic to combine the observed changes in all modal distributions was introduced for a more sensitive detection. A parametric MDM was developed to detect the changes in well-separated Gaussian modal distributions. This MDM was extended to apply in cases of Gaussian modal distributions corrupted by noise, a common occurrence in field measurement. To reduce the noise, a modified version of the power subtraction method was newly introduced. Furthermore, a non-parametric MDM does not require the assumptions on modal distribution shape so that the method can theoretically be applied to any shapes of modal distribution.

1. Parametric MDM

If the individual modal distributions separated by the global minimal points have Gaussian shapes, a parametric statistical comparison like the T-test can compare two corresponding Gaussian modal distributions obtained from two different measurements. A parametric test is preferred to a non-parametric test since the first is more powerful if the underlying distribution has Gaussian shape. In the results, the MDM using a parametric test is the simplest and the most powerful in the special case of well-separated Gaussian modal distributions. It was applied to simulations of smoothed power spectra based on riser VIV field measurement. Moreover, this method was extended to more complex modal distribution shapes such as modal distributions corrupted by noise and closely spaced modal distributions.

2. Noise Reduction MDM

Noise is inevitable in field vibratory response measurement during operation, and it may corrupt original modal distributions such that the parametric MDM cannot detect subtle changes in the original underlying modal distributions. A modified version of the popular spectral subtraction method was implemented into the parametric MDM for detecting changes in the underlying Gaussian modal distributions, even when they are corrupted by high level of noise. Various noise levels were simulated separately and added to the simulations of signals in the time domain to generate noisy signals. The noise reduction MDM was applied to the resulting simulated noisy signals.

3. Non-parametric MDM

Besides corruption by high levels of noise, other complications due to more complex situations such as closely spaced peaks may prevent the parametric MDM from detecting subtle changes in underlying modal distributions. Noise corruption can be reduced while reducing the complication due to closely spaced modal distributions is unfeasible. For a more general application of the MDM in such complicated situations, a non-parametric statistical test was introduced to the MDM which does not require knowledge of the underlying modal distribution shape a priori. The central limit theorem was imbedded where a parametric statistical test is applied to the distribution of sample means, since the theorem guarantees distribution of sample means approaches Gaussian. For verification purposes, the non-parametric MDM was applied to both pure and noisy signals. The results from two applications were compared with those obtained from the parametric MDM and the noise reduction MDM, respectively.

4. Extension to Well-separated Non-Gaussian Modal Distributions

Even though the modal distributions are non-Gaussian, a parametric statistical test is still recommended. However, application of a parametric test to non-Gaussian data is unreliable, and transformation from non-Gaussian distributions to Gaussian ones is required for reliable results. The Hermite moment model has been widely applied to estimate extremes of a non-Gaussian process using extremes of the theoretically well-developed Gaussian model. However, this model has two limits: orthogonality and monotone limits. The Hermite moment model was extended to overcome the monotone limit. The extended model was verified by comparing the extreme estimations obtained from both the extended and the original model near the monotone

limit in the time domain.

5. Extension to Continuous Monitoring for Well-separated Modal Distributions

When the resulting P-value is smaller than the predetermined significance level, the underlying modal distributions of two measurements are statistically the same. The MDM compares one of them and the next measurement for continuous monitoring, while more reliable continuous monitoring compares the combined measurement and the next one. However, concatenating the two measurements and calculating the power spectrum of the concatenated longer measurement are time-consuming. Therefore, a new method was developed to combine directly the statistical moments representing modal distributions, and it is applied to time domain analysis for verification of computational efficiency and storage save.

6. Comparison of Parametric MDM and Non-parametric MDM

First, both the parametric and the non-parametric MDM were applied to the signal simulations whose modal distributions satisfy the requirements of the parametric MDM; modal distributions are well separated Gaussian. As expected, the parametric MDM is more powerful than the non-parametric MDM because the requirements of the parametric MDM are fully met. Next, both methods were applied to the generated noisy signals that did not meet the requirement of the parametric MDM. The non-parametric MDM was able to detect the given changes while the parametric MDM was not.

7. Comparison of Non-parametric MDM and Noise Reduction MDM

The non-parametric MDM was validated through application to the noisy signal and comparison to the noise reduction MDM. Each method has its own advantages; the

noise reduction MDM has a low probability of Type I error while the non-parametric MDM has a low probability of Type II error. However, the noise reduction MDM is limited to the signal corrupted by a high level of noise, while the non-parametric MDM is applicable to other complicated situations including closely spaced peaks. This applicability of the non-parametric MDM should be verified before using it.

8. Recommendation

Before the use of the Modal Distribution Method (MDM), measured structural vibration responses should be confirmed to have separate modal responses and stationary excitation. Then the MDM uses only rough guesses of modal frequencies and measured responses to quantify the significance of observed changes in modal frequencies. The simplest way to roughly guess modal frequencies is to select peaks with the human eye in power spectrum of a measurement. The parametric MDM is highly recommended only when all of the following four requirements are met:

- underlying modal distributions are known before monitoring,
- the modal distributions are Gaussian,
- they are well separated, and
- the noise levels are as low as a good commercial off-the-shelf sensor.

The noise reduction MDM overcomes the fourth requirement, but it still requires responses to meet the remaining three requirements. The non-parametric MDM is theoretically applicable to any case, even when any of the four requirements are not met. The applicability of the non-parametric MDM to other complex situations needs to be verified before use.

B. Contribution

The original contributions of this dissertation work are the following:

- The modal distribution was introduced as a modal parameter representing modal response of vibratory structural response. It is considered as a statistical distribution after normalization, which makes the statistical method applicable. Many conventional methods select the peaks of spectrum as the characteristics of modal responses, while a Modal Distribution Method (MDM) uses spectral densities between two frequencies. Use of the modal distribution is more reliable than the single value of modal frequency since it uses more information.

- The application of a statistical test to modal distributions results in quantitative P-values, which indicate that observed changes can be ascribed to chance alone. P-values smaller than a given significance level imply that the observed changes are due to changes in underlying modal distributions. Moreover, the combined statistic including all observed changes in the modal distributions allows for a more sensitive detection of changes in response spectrum.

- The MDM is general and can be applied to any time-history such as stress, strain, motion, and acceleration if those have spectrums with separate modal responses which are statistically independent.

- The newly developed penalty method can find a global minimum as a separation point between roughly guessed two natural frequencies.

- A new modified version of the power subtraction method enables the MDM to detect changes in underlying modal distributions under a high noise corruption, while the original parametric MDM cannot detect them.

- Non-parametric MDM using Central Limit Theorem (CLT) in Chapter VI is free from underlying modal distribution shape. Theoretically, it can be applied to any shape of modal distributions including signal corruption by closely spaced modal peaks.

- The extension to the Hermite moment model enables transformation of non-Gaussian distributions to Gaussian ones in a wider range of skewness and kurtosis. Its implementation into the MDM is expected to detect changes in well-separated non-Gaussian modal distributions.

- The new, efficient algorithm that combines statistical moments can be implemented into the MDM since it uses distribution or histogram.

C. Future Work

The non-parametric MDM will be applied to other complicated conditions such as closely-spaced modes. The extended Hermite moment model will be implemented to the parametric MDM in order to analyze well-separated non-Gaussian modes. After the model transforms the isolated non-Gaussian modal distributions into the equivalent Gaussian ones, the transformed Gaussian distributions can be compared in the same way the parametric MDM does. The new efficient algorithm combines statistical moments using histogram, and will be also implemented into the MDM for a more reliable continuous monitoring.

In addition, other future researchers may apply the MDM to experimental or numerical responses of a structure whose damages increase progressively. This application can determine how much damage is necessary to be detected by the MDM. The MDM is based on an auto-spectrum of single sensor measurement to detect a global change of modal frequencies in a monitored system. In the future, the MDM

approach may be applied by other researchers to detect local change between any two sensors through a cross-spectrum between those two sensor measurements.

REFERENCES

[1] A. Rytter, "Vibration Based Inspection of Civil Engineering Structures," Ph.D. dissertation, Aalborg University, Denmark, 1993.

[2] J. K. Vandiver, "Detection of Structural Failure on Fixed Platforms by Measurement of Dynamic Response," in *Proc. of the 7th Annual Offshore Technology Conference*, Dallas, 1975, vol. 2, pp. pp.243–252.

[3] J. K. Vandiver, "Detection of Structural Failure on Fixed Platforms by Measurement of Dynamic Response," *Journal of Petroleum Technology*, vol. XXIX, pp. pp 305–310, 1977.

[4] R. D. Begg, A. C. Mackenzie, C. J. Dodds, and O. Loland, "Integrity Monitoirng Using Digital Processing of Vibration Signals," in *Proc. of the 8th Annual Offshore Technology Conference*, 1976, pp. 305–311.

[5] R. M. Kenley and C. J. Dodds, "West Sole WE Platform: Detection of Damage by Structural Response Measurements," in *Proc. of the 12th Annual Offshore Technology Conference*, 1980, pp. 111–118.

[6] T. R. Whittome and C. J. Dodds, "Monitoring Offshore Structures by Vibration Techniques," in *Proc. of Design in Offshore Structures Conference*, 1983, pp. 93–100.

[7] S.W. Doebling, M.B. Farrar, M.B. Prime, and D.W. Shevitz, "Damage Identification and Health Monitoring of Struyctural and Mechanical Systems from Changes in their Vibrational Characteristics: A Literature Review," Tech. Rep. LA–13070–MS, Los Alamos National Laboratory, 1996.

[8] S. Doebling, C. Farrar, and M. Prime, "A Summary Review of Vibration-Based Damage Identification Methods," *The Shock and Vibration Digest*, vol. 30, no. 2, pp. 91–105, 1998.

[9] M. Friswell and J. E. T. Penny, "The Practical Limits of Damage Detection and Location Using Vibration Data," in *Proc. of the 11th VPI & SU Symposium on Structural Dynamics and Control*, 1997.

[10] C. R. Farrar and T. A. Duffey, "Vibration-Based Damage Detection in Rotating Machinery and Comparison to Civil Engineering Applications," in *Proc. of Damage Assessment of Structures*, 1999.

[11] H. Sohn, C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, and B. R. Nadler, "A Review of Structural Health Monitoring Literature: 1996-2001," Tech. Rep. LA-13976-MS, Los Alamos National Laboratory, 2003.

[12] M. Sanayei, S. Wadia-Fascetti, B. Arya, and E. M. Santini, "Significance of Modeling Error in Structural Parameter Estimation," *Computer Aided Civil and Infrastructure Engineering*, vol. 16, pp. 12–27, 2001.

[13] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 2, New York, Wiley, 3rd edition, 1971.

[14] J. L. Rasmussen and W. P. Dunlap, "Dealing with Nonnormal Data: Parametric Analysis of Transformed Data vs Nonparametric Analysis," *Educational and Psychological Measurement*, vol. 51, pp. 809–820, 1991.

[15] J. K. Vandiver, "Detection of Structural Failure on Fixed Platforms by Measurement of Dynamic Response," in *Proc. of the 7th Annual Offshore Technology Conference*, 1975, pp. 243–252.

[16] J. K. Vandiver, "Detection of Structural Failure on Fixed Platforms by Measurement of Dynamic Response," *Journal of Petroleum Technology*, pp. 305–310, 1977.

[17] O. Loland and J. C. Dodds, "Experience in Developing and Operating Integrity Monitoring System in the North Sea," in *Proc. of the 8th Annual Offshore Technology Conference*, 1976, pp. 313–319.

[18] L. B. Crema and F. Mastroddi, "A Direct Approach for Updating and Damage Detection by Using FRF Data," in *Proc. of ISMA23, Noise and Vibration Engineering*, Leuven, Belgium, 1998.

[19] C. M. Harris and C. E. Crede, *Shock and Vibration Handbook*, New York, McGraw-Hill, 1995.

[20] A. W. Smyth, S. F. Masri, T. K. Caughey, and N. F. Hunter, "Surveillance of Mechanical Systems on the Basis of Vibration Signature Analysis," *ASME Journal of Applied Mechanics*, vol. 67, pp. 540–551, 2000.

[21] L. S. Katafygiotis and H. F. Lam, "A Probabilistic Approach to Structural Health Monitoring Using Dynamic Data," in *Proc. of Structural Health Monitoring, Current Status and Perspectives*, Stanford University, Palo Alto, California, 1997, pp. 152–163.

[22] M. Abe, Y. Fujino, M. Kajimura, M. Yanagihara, and M. Sato, "Monitoring of a Long Span Suspension Bridge by Ambient Vibration Measurement," in *Proc. of Structural Health Monitoring 2000*, Stanford University, Palo Alto, California, 1999, pp. 400–407.

[23] H. Sohn, C. R. Farrar, F. M. Hemez, D. D. Shunk, D. W. Stinemates, and B. R.

Nadler, "A Review of Strucutural Health Monitoring Literature," Tech. Rep. LA–13976–MS, Los Alamos National Laboratory, 2003.

[24] R. M. Bethea, *Statistical Methods for Engineers and Scientists*, New York, Marcel Decker Inc., 1984.

[25] S. Brandt, *Statistical and Computational Methods in Data Analysis*, Amsterdam, North-Holland Pub. Co., 1970.

[26] W. Navidi, *Statistics for Engineers and Scientists*, New York, McGraw-Hill Higher Education, 2006.

[27] D. E. Newland, *An Introduction to Random Vibrations, Spectral & Wavelet Analysis*, New York, Dover Publications, third edition, 2005.

[28] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procesures*, New York, John Wiley & Sons, 1991.

[29] T. K. Caughey, "Equivalent Linearization Techniques," *Journal of the Acoustical Society of America*, vol. 35, no. 11, pp. 1706–1711, 1963.

[30] T. S. Atalik and S. Utku, "Stochastic Linearization of Multi-degree-of-freedom Nonlinear Systems," *Earthquake Engineering & Structural Dynamics*, vol. 4, no. 4, pp. 411–420, 1976.

[31] P. Spanos, "Stochastic Linearization in Structural Dynamics," *Applied Mechanics Review*, vol. 34, no. 1, pp. 1–8, 1981.

[32] M.K. Ochi, "Non-Gaussian Random Processes in Ocean Engineering," *Probabilistic Engineering Mechanics*, vol. 1, no. 1, pp. 28–39, 1986.

[33] M. Grigoriu, "Crossings of Non-Gaussian Translation Processes," *Journal of Engineering Mechanics*, vol. 110, no. 4, pp. 610–620, 1984.

[34] Y. Murotsu, H. Okada, M. Kishi, M. Yonezawa, and K. Niwa, "Fourth Order Moment Approximation to Reliability of Non-linear Structure," in *Proc. of the 6th International Conference on Structural Mechanics in Reactor Technology*, Paris, France, 1981, number M12.

[35] C. Soize, "Gust Loading Factors with Nonlinear Pressure Terms," *Journal of the Structural Division, ASCE*, vol. 104, pp. 991–1007, 1978, No. ST6.

[36] S. R. Winterstein, "Moment-based Hermite Models of Random Vibration," Tech. Rep. 219, Department of Structural Engineering, Technical University of Denmark, March 1987.

[37] S. R. Winterstein, "Nonlinear Vibration Models for Extremes and Fatigue," *Journal of Engineering Mechanics*, vol. 114, no. 10, pp. 1772–1790, 1988.

[38] J. J. Jensen, "Dynamic Amplification of Offshore Steel Platform Responses due to Non-Gaussian Wave Loads," *Marine Structures*, vol. 7, pp. 91–105, 1994.

[39] J. M. Peeringa, "Extrapolation of Extreme Responses of a Multi Megawatt Wind Turbine," Tech. Rep. ECN-C–03-131, Energy Research Centre of The Netherlands, 2003.

[40] M. R. Moarefzadeha and R. E. Melchers, "Nonlinear Wave Theory in Reliability Analysis of Offshore Structures," *Probabilistic Engineering Mechanics*, vol. 21, no. 2, pp. 99–111, 2006.

[41] B. Puig and J. Akian, "Non-Gaussian Simulation Using Hermite Polynomials

Expansion and Maximum Entropy Principle," *Probabilistic Engineering Mechanics*, vol. 19, pp. 293–305, 2004.

[42] S. R. Winterstein, T. C. Ude, and G. Kleiven, "Springing and Slow-drift Responses: Predicted Extremes and Fatigue vs. Simulation," in *Proc. of The International Conference: Behaviour of Off-Shore Structures 1994*, MIT, Cambridge, USA, 1994, vol. 3, pp. 1–15.

[43] A. E. Mansour and J. J. Jensen, "Slightly Non-linear Extreme Loads and Load Combinations," *Journal of Ship Research*, vol. 39, no. 2, pp. 139–149, 1995.

[44] S. R. Winterstein and C. H. Lange, "Moment-Based Probability Models for Wind Engineering Applications," in *Proc. of the 10th Engineering Mechanics Speciality Conference, ASCE*, 1995, vol. 1, pp. 159–162.

[45] C. H. Lange, "Probabilistic Fatigue Methodology and Wind Turbine Reliability," Ph.D. dissertation, Stanford University, 1996.

[46] L. E. Dickson, *Elementary Theory of Equations*, New York, Wiley, 1914.

[47] G. Birkhoff and S. MacLane, *A Survey of Modern Algebra*, New York, Macmillan, 5 th edition, 1996.

[48] T. Sarpkaya and M. Issacson, *Mechanics of Wave Forces on Offshore Structures*, New York, Van Nostrand Reinhold, 1981.

[49] A. K. Chopra, *Dynamics of Structures: Theory and Applications to Earthquake Engineering*, Englewood Cliffs, NJ, Prentice Hall, 2001.

[50] D. N. Joanes and C. A. Gill, "Comparing Measures of Sample Skewness and Kurtosis," *Journal of Royal Statistical Society*, vol. 47, pp. 183–189, 1988.

[51] H. Sohn, C. Farrar, N. Hunter, and K. Worden, "Structural Health Monitoring Using Statistical Pattern Recognition Techniques," *Journal of Dynamic Systems, Measurement, and Control*, vol. 123, pp. 706–711, 2001.

[52] J. P. Lynch, A. Sundararajan, K. H. Law, A. S. Kiremidjian, and E. Carryer, "Embedding Damage Detection Algorithms in a Wireless Sensing Unit for Operational Power Efficiency," *Smart Materials and Structures*, vol. 13, pp. 800–810, 2004.

[53] H. W. Park and H. Sohn, "Parameter Estimation of the Generalized Extreme Value Distribution for Structural Health Monitoring," *Probabilistic Engineering Mechanics*, vol. 21, pp. 366–376, 2006.

[54] S. Ehrenfeld and S. B. Littauer, *Introduction to Statistical Method*, New York, McGraw-Hill, 1964.

[55] C. Mack, *Essentials of Statistics for Scientists and Technologists*, New York, Plenum Press, 1967.

[56] J. A. Cadzow and X. Li, "Blind Decomposition," *Digital Signal Processing*, vol. 5, pp. 3–20, 1995.

[57] R. F. Dwyer, "Use of The Kurtosis Statistics in The Frequency Domain as an Aid in Detecting Random Signals," *IEEE Journal of Oceanic Engineering*, vol. 9, pp. 85–92, 1984.

[58] R. F. Dwyer, "Asymptotic Detection Performance of Discrete Power and Higher-order Spectra Estimates," *IEEE Journal of Oceanic Engineering*, vol. 10, pp. 303 – 315, 1985.

[59] A. Tesei and C. S. Regazzoni, "Signal Detection in Non-Gaussian Noise by a Kurtosis-based Probability Density Function Model," in *IEEE Workshop on HOS*, 1995, pp. 16216–5.

[60] D. Dyer and R. M. Stewart, "Detection of Rolling Element Bearing Damage by Statistical Analysis," *ASME Journal of Mechanical Design*, vol. 100, pp. 229235, 1978.

[61] R. B. W. Heng and M. J. M. Nor, "Statistical Analysis of Sound and Vibration Signals for Monitoring Rolling Element Bearing Condition," *Applied Acoustics*, vol. 53, pp. 211–226, 1998.

[62] H. R. Martin and F. Honarvar, "Application of Statistical Moments to Bearing Failure Detection," *Applied Acoustics*, vol. 44, no. 1, pp. 67–77, 1995.

[63] T. F. Chan, G. H. Golube, and R. J. LeVeque, "Updating Formulae and a Pairwise Algorithm for Computing Sample Variances," Tech. Rep. STAN-CS-79-773, Department of Computer Science, Stanford University, 1979.

[64] N. A. Tanner, J. R. Wait, C. R. Farrar, and H. Sohn, "Structural Health Monitoring Using Modular Wireless Sensors," *Journal of Intelligent Material Systems and Structures*, vol. 14, pp. 43–56, 2003.

[65] "Computing Higher-Order Moments Online," 2008, [online] http://people.xiph.org/ tterribe/notes/homs.html.

[66] P. Pebay, "Formulas for Robust, One-pass Parallel Computation of Covariances and Arbitrary-Order Statistical Moments," Tech. Rep. SAND2008-6212, Livermore, CA, Sandia National Laboratories, 2008.

[67] B. Sweetman and M. Choi, "The Modal Distribution Method: A New Statistical Algorithm for Analyzing Measured Data," in *Proc. of SPIE*, San Diego, California, February-March 2006, vol. 6174.

[68] J. F. Wilson, Ed., *Dynamics of Offshore Structures*, New York, John Wiley & Sons, 2002.

[69] S. Boll, "Suppression of Acoustic Noise in Speech Using Spectral Subtraction," in *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, pp. 113–120. 1979.

[70] S. H. Godsill and P. J. Rayner, *Digital Audio Restoration: A Statistical Model Based Approach*, Secaucus, NJ, USA, Springer-Verlag Inc., 1998.

[71] P. C. Loizou, *Speech Enhancement: Theory and Practice*, Boca Raton, FL, CRC Press, 2007.

[72] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, Cambridge, MA, The MIT Press, 1964.

[73] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," in *Proc. of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1979, vol. 4, pp. 208–211.

[74] V. Schless and F. Class, "SNR-dependent Flooring and Noise Overestimation for Joint Application of Spectral Subtraction and Model Combination," in *Proc. of the 5th International Conference on Spoken Language Processing*, 1998.

[75] R. V. Mises, *Mathematical Theory of Probability and Statistics*, New York: Academic Press, 1964.

[76] F. W. Scholz and M. A. Stephens, "K-Sample Anderson-Darling Tests," *Journal of the American Statistical Association*, vol. 82, pp. 918–924, 1987.

[77] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, New York, Chapman & Hall, 1993.

[78] D. S. Moore and G. P. McCabe., *Introduction to the Practice of Statistics*, New York, W.H. Freeman and Company, 2006.

APPENDIX: EFFICIENT CALCULATION OF HIGHER-ORDER STATISTICAL MOMENTS

Efficient methodologies to compute skewness and kurtosis are developed and presented. First, a new methodology is developed for combination of the statistical moments of individual segments of a data-set into those of the complete set. Next, a one-pass methodology is presented for calculation of higher moments based on the properties of a histogram, which allows computations to be carried out to accuracy matching the precision of measured data. An example is presented in which the histogram method is shown to require meaningfully less CPU time than existing on-line algorithms

A. Introduction

Statistical moments can be used to represent the characteristics of any irregular data, e.g., [54, 55]. Statistical moments have found a broad range of application including: blind decomposition [56], asymptotic probability of detection criterion in the frequency domain [57, 58], non-Gaussian noise modeling [59], and use of non-Gaussian distributions that more precisely match the tail behavior of measured data, e.g. [37].

Conventionally, the statistical moments of a set of discrete data, x_i, are computed directly using a two-pass algorithm (e.g. [54, 55]):

$$\mu = \frac{1}{I}\sum_{i=1}^{I} x_i = E[x] = m_1 \tag{A-1}$$

$$\sigma^2 = \frac{1}{I}\sum_{i=1}^{I}(x_i - \mu)^2 = \theta_2 \tag{A-2}$$

$$\alpha_3 = \frac{1}{I\sigma^3}\sum_{i=1}^{I}(x_i - \mu)^3 = \frac{\theta_3}{\sigma^3} \tag{A-3}$$

$$\alpha_4 \;=\; \frac{1}{I\sigma^4}\sum_{i=1}^{I}(x_i-\mu)^4 = \frac{\theta_4}{\sigma^4} \tag{A-4}$$

where I is the number of points in the sample; μ, σ^2, α_3, and α_4 are the mean, variance, skewness, and kurtosis of the data x_i, and θ_2, θ_3 and θ_4 are the central moments. Such algorithms are called "two-pass" because the mean must first be computed and that mean is subsequently used in the computation of the remaining moments, which implies the entire data-set must be retained. Computational methodologies requiring retention of the entire data-set are termed "off-line" methods because the statistical moments can not be computed real-time; methodologies that enable computation of the moments at each successive data-point with only one pass through the data are termed "on-line" methods. These on-line methods have the obvious advantage that the data-set need not be stored. One-pass on-line algorithms for the mean and variance have been known for some time (e.g., [63]), and have been implemented on real-world hardware (e.g. [64]), but one-pass algorithms for higher moments are less common.

Combination of the statistical moments from various segments of a data-set into overall statistical moments shares the same complication as the off-line methods: the individual segments do not in general share a common mean. Methodologies to combine these statistical moments are sometimes referred to as updating methodologies because known statistical moments of earlier segments are "updated" by the statistical moments of an additional segment of the data-set. Terriberry [65] offers pairwise updating formulas for the skewness and kurtosis (without derivation) and Pebay [66] explains how Terriberry's updating formulae could be implemented as a one-pass algorithm by applying the update formulae to only one additional point at a time, and also notes that Terriberry's results are special cases of Pebay's arbitrary-order update formulae. The Terriberry/Pebay updating formula for the third and fourth central

moments are:

$$M_3 = M_{3,A} + M_{3,B} + n_A n_B (n_A - n_B)\frac{\delta^3}{n^2} + 3(n_A M_{2,B} - n_B M_{2,A})\frac{\delta}{n} \quad \text{(A-5)}$$

$$M_4 = M_{4,A} + M_{4,B} + n_A n_B (n_A^2 - n_A n_B + n_B^2)\frac{\delta^4}{n^3} + 6(n_A^2 M_{2,B} + n_B^2 M_{2,A})\frac{\delta^2}{n^2} + 4(n_A M_{3,B} - n_B M_{3,A})\frac{\delta}{n} \quad \text{(A-6)}$$

in which n represents the number of data points in the set made up of segments A and B having number of points n_A and n_B respectively; $M_{3,*}$ and $M_{4,*}$ are the third and fourth central moments, and δ is the difference between means: $\delta = \mu_B - \mu_A$. These updating formula are readily specialized into a one-pass online algorithm by setting n_A or n_B equal to one, and thereby become the only pre-existing one-pass algorithms for skewness and kurtosis known to the authors.

Here, alternatives to the Terriberry/Pebay skewness and kurtosis update formulae are derived and presented and then a less theoretically complex alternative to the Terriberry/Pebay on-line methodology is investigated. The implementation offered by Pebay is substantially different from that suggested here, and the derivation offered by Pebay also differs meaningfully, though both derivations hinge on the commutativity of summations over finite sets as applied to statistical moments. In his report, Pebay echoes Terriberry's thoughts, noting that: "To our knowledge, there are currently no published formulas for parallel updates of higher-order moments." The authors of this paper agree with that assessment, and developed the methodologies presented here because they had use for these methodologies in ongoing work in the engineering field of structural health monitoring.

The methods presented here have specific advantages over those of Terriberry/Pebay. First, the Terriberry/Pebay formulae apply to combination of only two sets of mo-

ments, and would need to be used recursively to combine multiple sets. In contrast, the new methodology presented here can be used to combine any number of sets of moments, which could be relevant in the case of massively parallel computations. Second, an important capability of the updating formulae offered here compared with those of Terriberry/Pebay is that this methodology can be readily applied to distributions of data that are not specifically countable, such as the moments of a probability density function estimated from a power spectrum (e.g., [67]). Third, an important capability for practical applications is that this new methodology is easily modified to allow a user to assign different importance to specific segments of the data, such as newer data being more important than older data, or one sensor being more reliable or accurate than another.

Finally, this paper offers a one-pass algorithm to compute statistical moments of a data set. This method is shown in an example to be more computationally efficient than the one-pass method offered by Terriberry/Pebay. In the proposed method, the data are first binned to create a histogram from which the desired moments can be calculated; the width of each bin can be specified as a function of the required accuracy. Setting the bin width to the precision of the original data yields exact results. The computational savings of this alternative method can be substantial, and significant memory savings can also be realized if the raw data are binned real-time so the complete time-history need not be retained. This new methodology is competitive with existing one-pass algorithms for the statistical moments in that the raw data need only be passed through one time, but it is not technically an on-line algorithm because the data is only binned real-time and the statistical moments are not computed every time-step. Computation of statistical moments from a histogram is generally well-established in the statistics community, but investigation of this methodology as an efficient one-pass algorithm with arbitrarily specified accuracy

is unique. As such, the latter part of the paper is of mainly practical rather than theoretical interest, whereas the updating methodologies are derived directly from theory.

B. Theory

Derivation of the new updating methodology requires some knowledge of moments, which is presented here to ensure consistent notation. This background information is followed by the derivation of the proposed method to combine statistical moments, and then an investigation of computing statistical moments of discrete data from a histogram. The derivation and example refer to data collected as a function of time, but the methods are generally applicable to any type of data.

1. Background

a. Calculation of Moments from Discrete Data

The moments of a random variable about zero and about its mean are referred to as the raw and central moments, respectively. The n^{th} moment of a discrete random variable $x(t)$ about value r with a finite range is defined as (e.g. [54, 55]):

$$M_{n,r} = \sum_{i=1}^{I} \Big(x(t_i) - r\Big)^n \Delta t_i \tag{A-7}$$

which can be expressed as follows if the continuous distribution function, $f(x)$, is known:

$$M_{n,r} = \sum_{k=1}^{K} (x_k - r)^n f(x_k) \Delta x_k \tag{A-8}$$

in which I represents the number of data points $x(t_i)$ and K represents the number of base x_k in its discrete distribution function $f(x_k)$, i.e., the number of bins in the

discrete distribution.

The n^{th} raw moment of a discrete time-history (Equation (A-7) with $r = 0$) can be normalized by the time duration, the result if which is the expected value of the n^{th} power of x

$$m_n^{(t)} = \frac{1}{T} M_{n,0} = \frac{\sum_{i=1}^{I} x(t_i)^n \Delta t_i}{\sum_{i=1}^{I} \Delta t_i} = E[x^n] \tag{A-9}$$

where $T = \sum_{i=1}^{I} \Delta t_i$ is the time duration and the superscript (t) indicates moments are calculated directly from the time-history. For constant Δt, the duration is $T = I\Delta t$, which enables Equation (A-9) to be simplified:

$$m_n^{(t)} = \frac{1}{I} \sum_{i=1}^{I} x(t_i)^n \tag{A-10}$$

The first normalized raw moment ($n = 1$) is the sample mean, which is often used to estimate the true mean of the process for normalization of other central moments.

$$\theta_n^{(t)} = \frac{1}{T} M_{n,m_1^{(t)}} = \frac{\sum_{i=1}^{I} \left(x(t_i) - m_1^{(t)} \right)^n \Delta t_i}{\sum_{i=1}^{I} \Delta t_i} = E\left[\left(x - m_1^{(t)} \right)^n \right] \tag{A-11}$$

which for constant time interval Δt is:

$$\theta_n^{(t)} = \frac{1}{I} \sum_{i=1}^{I} \left(x(t_i) - m_1^{(t)} \right)^n \tag{A-12}$$

The second normalized central moment ($n = 2$) is the sample variance (Equation (A-2)).

b. The Relationship Between Raw Moments and Central Moments

The first four raw moments and central moments have the following well-known mathematical relationships, e.g. [54, 55]:

$$m_1 = E[x] = \mu \tag{A-13}$$

$$m_2 = E[x^2] = \theta_2 + m_1^2 \tag{A-14}$$

$$m_3 = E[x^3] = \theta_3 + 3m_1\theta_2 + m_1^3 \tag{A-15}$$

$$m_4 = E[x^4] = \theta_4 + 4m_1\theta_3 + 6m_1^2\theta_2 + m_1^4 \tag{A-16}$$

$$\theta_2 = E[(x-\mu)^2] = m_2 - m_1^2 \tag{A-17}$$

$$\theta_3 = E[(x-\mu)^3] = m_3 - 3m_1m_2 + 2m_1^3 \tag{A-18}$$

$$\theta_4 = E[(x-\mu)^4] = m_4 - 4m_1m_3 + 6m_1^2m_2 - 3m_1^4 \tag{A-19}$$

2. Updating Formulae for Statistical Moments

Here, a new computational method is proposed to combine multiple sets of statistical moments, each representing a segment a time-history, into a single set describing a single concatenated time history. An example application might be combining moments from several individual segments of a long time-history, perhaps with the moments of each segment having been computed on a separate processor. This proposed computational technique uses the first four statistical moments of each segment to compute four raw moments, which are then transformed into new variables (γ_n) that are easily combined by addition. After combination, the new variables are inversely transformed back to four raw moments now describing all the data, from which the statistical moments are easily calculated.

a. Moment Addends, γ_n

New moment addend variables, γ_n, are introduced to enable straightforward combination of the statistical moments of multiple time-histories. For an irregular time-history $x(t)$ with variable time interval Δt_i:

$$\gamma_n = \sum_{i=1}^{I} (x(t_i))^n \Delta t_i \tag{A-20}$$

where γ_0 is the duration of each time-history. For constant time interval $\Delta t_i = \Delta t$:

$$\gamma_n = \Delta t \sum_{i=1}^{I} (x(t_i))^n \tag{A-21}$$

The same values of γ for the histogram form of computing the moments can be expressed in terms of the frequency of occurrence at the x_k bin, $h(x_k)$, with variable bin width Δx_k.

$$\gamma_n \quad = \quad \sum_{k=1}^{K} x_k^n h(x_k) \Delta x_k \tag{A-22}$$

yielding γ_0 as the area of the histogram. For constant bin width $\Delta x_k = \Delta x$:

$$\gamma_n = \Delta x \sum_{k=1}^{K} x_k^n h(x_k) \tag{A-23}$$

b. Combination of Statistical Moments

If Q sets of statistical moments are known: $\gamma_{0,q}, \mu_q, \sigma_q^2, \alpha_{3,q}, \alpha_{4,q}$ for $q = 1, 2, ..., Q$, then each γ_n can be expressed in terms of the equivalent n raw moments (Equations 4.7, 4.8, 4.19, and 4.21).

$$\gamma_{n,q} = m_{n,q} \gamma_{0,q} \qquad \text{for} \quad n = 1, 2, 3, 4 \quad \text{and} \quad q = 1, 2, ..., Q \tag{A-24}$$

where $\gamma_{0,q}$ is generally taken to be the duration of the q^{th} time-history, or the number of points if Δt is constant. It is worth noting, however, that $\gamma_{0,q}$ is a weighting factor only, and its interpretation can be flexible depending on the application. Importantly, in this method the statistical moments are not required to be those of a quantity that is countable: these moments could be computed directly from e.g., a probability distribution, in which case the value of $\gamma_{0,q}$ would represent the relative importance of the moment estimate. There is no theoretical limitation on the maximum order of the moments (the value of n), though higher-order equivalents to Equations 4.11–4.17

would be needed. The benefit of expressing the statistical moments in terms of γ is that the Q sets can be combined by addition, and there is no upper limit on the value of Q.

$$\gamma_{n,c} = \sum_{q=1}^{Q} \gamma_{n,q} \qquad \text{for} \quad n = 0, 1, 2, 3, 4 \tag{A-25}$$

where the subscript c represents the concatenated time-history or combined γ. These combined values of γ can then be inversely transformed into raw moments representing the concatenated time-history by inverting Equation (4.29).

$$m_{n,c} = \frac{\gamma_{n,c}}{\gamma_{0,c}} \quad \text{for} \quad n = 1, 2, 3, 4 \tag{A-26}$$

The relationship between raw moments and central moments (Equations A-17–A-19) are then used to compute the central moments of the concatenated time-history. Finally, the statistical moments of the concatenated history are computed as in Equations (A-1–A-4)

$$\mu_c = m_{1,c} \qquad \sigma_c^2 = \theta_{2,c} \qquad \alpha_{3,c} = \frac{\theta_{3,c}}{\sigma_c^3} \qquad \alpha_{4,c} = \frac{\theta_{4,c}}{\sigma_c^4} \tag{A-27}$$

3. One-pass Calculation of Statistical Moments

One-pass algorithms can be developed from updating formulae, but it is here proposed that it is theoretically simpler and in practice more computationally efficient to compute these moments directly from a histogram. In this method, a relative histogram of the random variable is first constructed in the conventional way: the range of potential values is divided into bins and the number of occurrences within each bin are counted and plotted such that the area of each rectangle equals the portion of the sample values within that bin (e.g. [54, 55]):

$$H(x_k) = \frac{h(x_k)}{A} \tag{A-28}$$

where $h(x_k)$ and $H(x_k)$ represent the frequency and the relative frequency in bin x_k, and $A = \sum_{k=1}^{K} h(x_k)\,\Delta x_k$ is the total area of the histogram. If the bin-width is pre-specified but the data-range is not known a-priori, it is straightforward to program a binning methodology in which bins are added dynamically as required by the data. After the normalization by A in Equation A-28, the n raw moments and central moments of $x(t)$ can be calculated from the relative histogram, similar to Equation (A-8):

$$m_n^{(h)} = \sum_{k=1}^{K} x_k^n\, H(x_k)\Delta x_k = \frac{1}{A}\sum_{k=1}^{K} x_k^n\, h(x_k)\Delta x_k \tag{A-29}$$

$$\theta_n^{(h)} = \sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n H(x_k)\Delta x_k = \frac{1}{A}\sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n h(x_k)\Delta x_k \tag{A-30}$$

where the superscript (h) indicates the moments are calculated from the histogram.

For constant bin width, $\Delta x_k = \Delta x$, these two expressions can be simplified with $A = \sum_{k=1}^{K} h(x_k)\Delta x = I\Delta x$:

$$m_n^{(h)} = \frac{1}{I}\sum_{k=1}^{K} x_k^n\, h(x_k) \tag{A-31}$$

$$\theta_n^{(h)} = \frac{1}{I}\sum_{k=1}^{K} \left(x_k - m_1^{(h)}\right)^n h(x_k) \tag{A-32}$$

C. Example

The new methodologies are applied to data resulting from a time-domain solution of a simple numerical model of a floating offshore oil production platform subject to irregular seas. In the first part of the example, the statistical moments of the response are combined, and in the second part the histogram methodology of computing statistical moments is critically compared with one- and two-pass algorithms. The response selected for the example is surge motion (horizontal translation in the direction of the

Table A-1. Environmental Conditions

Condition	$H_s(m)$	$T_p(sec)$	$U_c(m/s)$	$F_w(kN)$
1	14.5	15.5	1.5	-3000
2	14.5	12.5	2.0	-5000
3	14.5	18.5	1.0	-2000

waves) of a tension leg platform (TLP), which is known to be highly non-Gaussian. Three segments of a surge time-history, each having significantly different statistical moments, are created without modifying the structural model, but only changing the peak period of the incident waves, T_p, the current velocity, U_c, and the wind force, F_w. In the example, the wave time-histories are simulated from a JONSWAP sea spectrum (e.g., [48]) with a 14.5 m significant wave height, H_s. Environmental forcing is computed using Airy wave theory and Morison's equation (e.g., [48, 68]. The three environmental conditions used in the dynamic simulation are summarized in Table A-1. The one-dimensional dynamic equation of motion is solved in the time domain using the Newmark Beta Method (e.g., [49]). The time step of integration (Δt) is 0.01 sec and the total time duration for each of the three time-histories is one hour. The resulting time-histories and associated statistical moments are used here to investigate the new methodology and to compare with existing methodologies.

1. Combination of Statistical Moments

In this first part of the example, statistical moments are first calculated using the conventional two-pass method (Equations A-1–A-4). Results are shown on the first three lines of Table A-2. The fourth and fifth lines of the table show a statistical summary of a combination of the first two segments. The results on the fourth line were computed by concatenating the two complete one-hour time-histories and then

applying the conventional two-pass method to compute the summary statistics. The fifth line was computed using the new methodology derived in Section 2. The specific steps to compute these values are as follows: First, the $n = 4$ statistical moments for each (q) of the $Q = 2$ segments to be combined are transformed into the mean and three central moments by inverting the definitions of the statistical moments (Equations A-1–A-4) to $m_{1,q} = \mu_q$, $\theta_{2,q} = \sigma_q^2$, $\theta_{3,q} = \alpha_{3,q}\sigma_q^3$, $\theta_{4,q} = \alpha_{4,q}\sigma_q^4$. Second, the resulting $Q = 2$ sets of three central moments are transformed into $Q = 2$ sets of three raw moments using the well-known relationships between raw and central moments (Equations A-14–A-16). Third, the resulting $4Q$ raw moments (including $Q = 2$ means) are then transformed into $4Q$ values of $\gamma_{n,q} = m_{n,q}\gamma_{0,q}$ (Equation (A-24)). Fourth, each of the 5 sets of $Q = 2$ values of $\gamma_{n,q}$ are combined as in Equation (A-25), $\gamma_{n,c} = \sum_{q=1}^{Q} \gamma_{n,q}$ ($n = 0$ to 4). Finally, the transformation process is reversed for the resulting 5 values of $\gamma_{n,c}$ to produce the desired four central statistical moments as in Equations A-26–A-27. The sixth and seventh lines are the equivalent to the forth and fifth, excepting that three segments are combined ($Q = 3$). As can be seen in the table, combining the statistical moments using the newly proposed methodology yields results identical to those computed directly from a concatenated time-series using the conventional two-pass method.

2. Calculation of Statistical Moments

In this second part of the example, the relative efficiency of calculating statistical moments from time-series data is investigated through comparison of three different computational methodologies. First, calculations are performed using the conventional two-pass method (Equations A-1–A-4). Next, equivalent calculations are performed using the methods of Terriberry/Pebay (Equations A-5–A-6). Finally, a third set of equivalent calculations are performed by binning the data and computing

Table A-2. Summary Statistics of Response

Condition	Method	μ	σ^2	α_3	α_4
1	Two-Pass	-1.3833	0.2742	0.1236	2.9969
2	Two-Pass	-2.2512	0.2447	0.1203	2.3685
3	Two-Pass	-0.9154	0.2510	0.0411	2.1207
1+2	Two-Pass	-1.8173	0.4478	0.1180	2.5684
1+2	Updating	-1.8173	0.4478	0.1180	2.5684
1+2+3	Two-Pass	-1.5167	0.5629	-0.0917	2.3856
1+2+3	Updating	-1.5167	0.5629	-0.0917	2.3856

the statistical moments from the resulting histogram as outlined in Section 3. The amount of CPU time and computer memory requirements are presented and compared in Figures A-1 and A-2. In both comparisons, the baseline for comparison is the conventional two-pass method, which defines the 100% CPU and the 100% memory conditions. In all comparisons, CPU requirements result from computations and binning performed using MatLab; memory usage is based on computing how many floating-point numbers must be stored. The reader is cautioned that the relative efficiencies presented in the example are not general, and different lengths and ranges of data-sets may meaningfully affect these comparisons.

First considering Figure A-1, the vertical axis on the left compares the relative CPU time needed to compute all four of the statistical moments. Application of the online method of Terriberry/Pebay to this one-hour time-history (360000 data points) with a measurement precision of 0.001, requires about 70% more CPU time than the conventional two-pass methodology, though it has the obvious advantage that the data need not be stored as part of the calculation process. The histogram method with the bin width set to the precision of the data (zero error in binning leading to exact

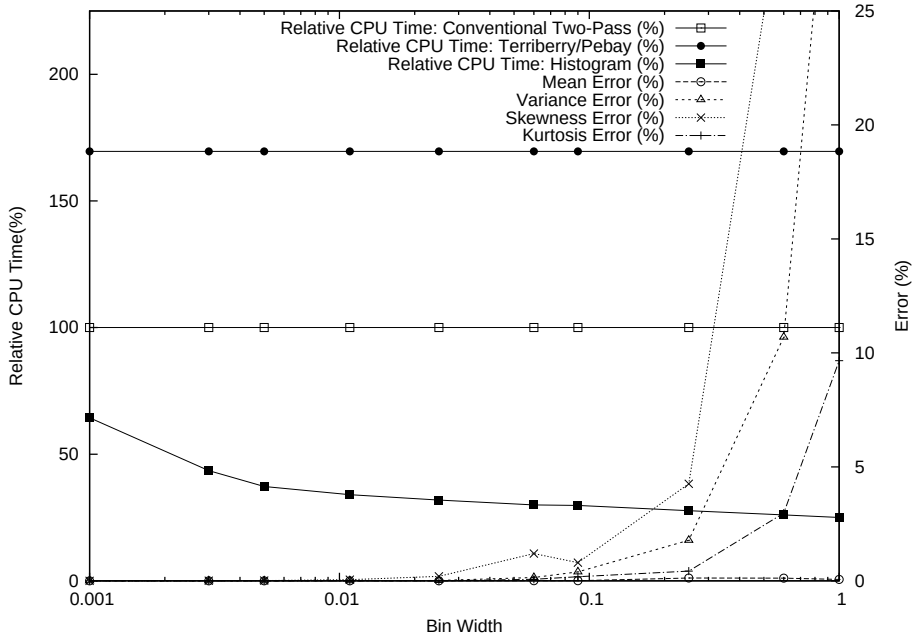


Fig. A-1. Effect of Bin Width on CPU Time Required for Calculation of Statistical Moments. Conventional Two-pass Algorithm Defines 100%. Time Duration: 1 hour, Δt: 0.01 sec, Data Precision: 0.001

sample statistical moments) takes about 64% of the CPU time as the conventional method which is about 38% of the CPU time required for the Terriberry/Pebay on-line method. Progressing from left to right on the horizontal axis shows gradual increases in the bin width, and the vertical axis on the right shows the associated error in the statistical moments. The error is computed as the difference between the exact and approximate statistical moments divided by the exact statistical moments. The plot shows that for this time-history, increasing the bin width to ten times as large as the precision of the data reduces the CPU time to 34% of the conventional methodology, with virtually no noticeable increase in error in the statistical moments.

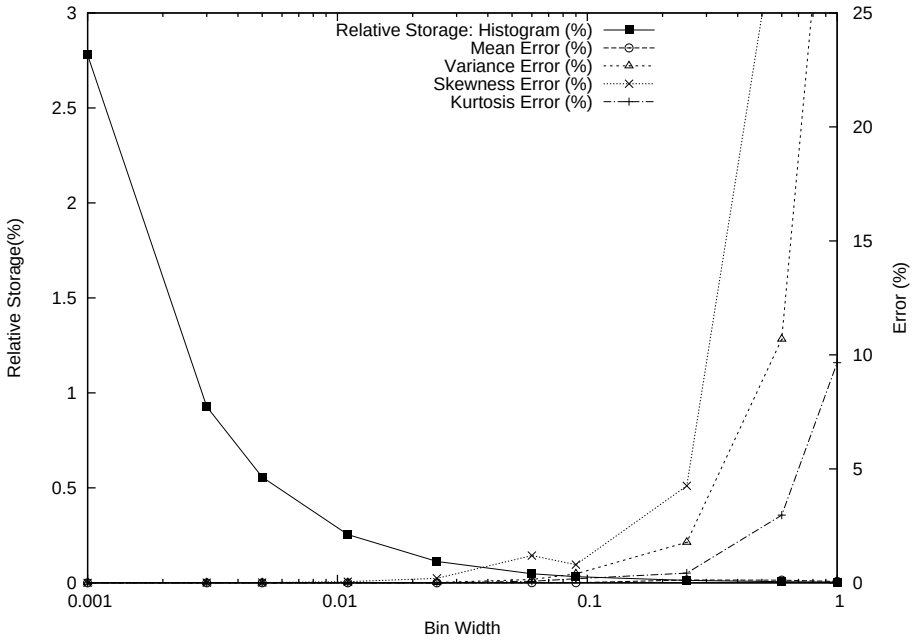


Fig. A-2. Effect of Bin Width on Computer Memory Required for Calculation of Statistical Moments. Conventional Two-pass Algorithm Defines 100%. Time Duration: 1 hour, Δt: 0.01 sec, Data Precision: 0.001

Next, considering Figure A-2, the vertical axis represents the amount of memory required to compute the statistical moments, with 100% being defined by the two-pass method in which the entire time-history must be stored. Only the memory requirements for the histogram method is shown because of the extreme scale differences: the two-pass method requires about 36 times as much storage, and the method of Terriberry/Pebay requires only about 0.0033% of that required for the two-pass method, or 0.12% of that required for the histogram. The dramatic savings in memory available to the true one-pass algorithm results from having no need to store or bin any raw data.

D. Conclusions

Two methodologies have been presented and demonstrated through an example. The first methodology is an updating algorithm for computing the skewness and kurtosis of segmented data sets. This methodology has some advantages over the existing methods: First, it enables direct combination of the skewness and kurtosis of any number of data sets, rather than combining only two at a time. Second, this new methodology can be used to combine statistical moments of data that is not countable, e.g., moments extracted directly from a probability density function, and third, the new methodology can be easily modified to assign different weightings to specific segments of the data if the user believes some of the data is of greater importance to estimating the statistical moments.

The second methodology is a one-pass algorithm for computation of higher statistical moments through use of a histogram. While it is a one-pass algorithm in that the data is only passed through once, it is not technically an on-line algorithm because the moments are not computed every time-step. The method investigated also has distinct advantages over other methods: First, it uses significantly less CPU time than conventional two-pass algorithms or an existing one-pass algorithm. Second, it uses significantly less computer memory than a two-pass algorithm, though much more than an existing on-line algorithm. Third, the method can be implemented to arbitrary accuracy, such that computations can be performed to the same accuracy as the measured data, or to effectively trade required accuracy against CPU and memory requirements. The proposed methodology is of more practical than theoretical interest because computation of moments from a histogram is generally well-known in the statistics community, but it is here investigated as a viable alternative to true on-line methods.

An example is presented in which both the updating methodology for the skewness and kurtosis and the histogram approach are verified. The relative efficiency and accuracy of the histogram approach are compared with existing two-pass and online methods, and the savings in CPU time and memory requirements are quantified (Figures A-1 and A-2).

VITA

NAME: Myoung Keun Choi

ADDRESS: Zachry Department of Civil Engineering
Texas A& M University
College Station, TX 77843

EMAIL: mkchoi@tamu.edu

EDUCATION: B.S., Naval Architecture and Ocean Engineering,
Seoul National University, South Korea, 1997

M.S., Naval Architecture and Ocean Engineering,
Seoul National University, South Korea, 1999

Ph.D., Civil Engineering,
Texas A&M University, USA 2009

RELEVANT PUBLICATIONS:

[1] B. Sweeetman and M. Choi, "The Modal Distribution Method for Statistical Analysis of Measured Structural Response", *Journal of Probabilistic Engineering Mechanics*, Accepted with Revision, April 2009

[2] M. Choi and B. Sweeetman, "Efficient Calculation of Statistical Moments for Structural Health Monitoring", *An International Journal of Structural Health Monitoring*, Submitted for Review, November 2008

[3] M .Choi and B. Sweeetman, "The Hermite Moment Model for Highly Skewed Response with Application to Tension Leg Platforms", *Journal of Offshore Mechanics and Arctic Engineering*, Submitted for Review, August 2008

[4] B. Sweeetman and M .Choi, "The Modal Distribution Method: A New Statistical Algorithm for Analyzing Measured Acceleration Data", in *Proc. of Smart Structures and Materials NDE for Health Monitoring and Diagnostics*, San Diego, California, February 2006