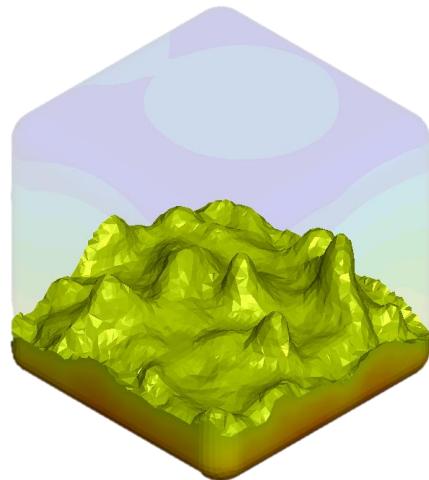


Analyzing Defective Ball Bearing Signals to Generate a Surrogate Model for
Slosh Dynamics in a Lunar Descent Environment



DALKO H. JERI

ASEN 5307-Final Project

Fall 2024

Table of Contents

Acronyms	3
Abstract	4
Introduction	4
Ball Bearing Configurations	4
Loading Cases	7
Ball Bearing Vibrational Responses	7
Region I: Low Frequency	9
Region II: High Frequency	12
Slosh Simulation	16
Downsampling	16
Input Excitations	19
LS-DYNA Runs	20
Simulation Results	21
Surface Response	25
Two-Dimensional Kriging	25
One-Dimensional Kriging	32
1D Kriging Results	36
Discussion and Conclusion	38
References	39
Appendix	40

Acronyms

Abbreviation	Meaning
ADCS	Attitude Dynamics and Control System
BPFO	Ball Pass Frequency of the Outer Race
CG	Center of gravity
ESH	Environmental, Safety, and Health
FIR	Finite Impulse Response
IIR	Infinite Impulse Response
LSF	Least Squares Fit: Assume $y_{fit}(t) = a + bt$
MLE	Maximum Likelihood Estimation
MSE	Mean Squared Error
PDF	Probability Density Function
RP-1	Rocket Propellant 1
RSS	Root Sum Squared

Abstract

Defects, such as cracks in ball bearing assemblies, are inevitable due to their demanding application. When these defects occur in the outer race, they generate undesired vibrations that can excite nearby hardware. Different crack geometries result in unique vibrational responses. This research project utilizes vibrational data from defective bearings with different crack geometries, collected by an accelerometer. The data is analyzed in both time and frequency domains and converted into exogenous excitations for a spacecraft cubic tank housing rocket propellant RP-1. The slosh dynamics of the RP-1 under a lunar gravity field and these excitations are simulated using LS-DYNA, a multiphysics simulation software. The center of gravity (CG) of the RP-1 serves as the metric for evaluating the influence of these vibrational excitations. Kriging interpolation is then employed as a surrogate model to estimate the CG under different excitation profiles.

Introduction

Defective ball bearings can generate unwanted excitations that affect nearby components. In the case of a spacecraft fuel tank housing rocket propellant for its liquid engine, these excitations can result in unexpected slosh dynamics, posing challenges to the spacecraft's Attitude Determination and Control System (ADCS). Predicting the CG movement helps contextualize the disturbances caused by these excitations. However, CG movement is often determined through hardware testing, which is expensive—not only financially, but also in terms of resources and scheduling, ESH considerations, such as ensuring human safety during tests involving highly combustible fluids, can lead to lengthy discussions and further delays.

For these reasons, if a requirement related to liquid propellant slosh is stated, it is common to address it through a Verification-by-Simulation approach. Within this framework, surrogate models can assist the program in understanding the effects of unwanted excitations. However, slosh simulations, particularly in LS-DYNA, can be computationally intensive, sometimes requiring weeks to deliver accurate results. This timeline can be impractical when a first-order assessment is needed. In this research project, Kriging interpolation is employed to predict data in the gaps between observed CG shifts obtained from LS-DYNA simulations.

Unless otherwise stated, all units are as follows: frequencies in [Hz], mass in [kg], displacement in [m], velocity in [m/s], and accelerations in [G].

Ball Bearing Configurations

The vibrational data is derived from experimental runs obtained by A. Gabrielli et al. from the University of Ferrara in Italy.¹ The figure below illustrates their experimental setup.

¹ <https://data.mendeley.com/datasets/8wdzm5gwng/1>

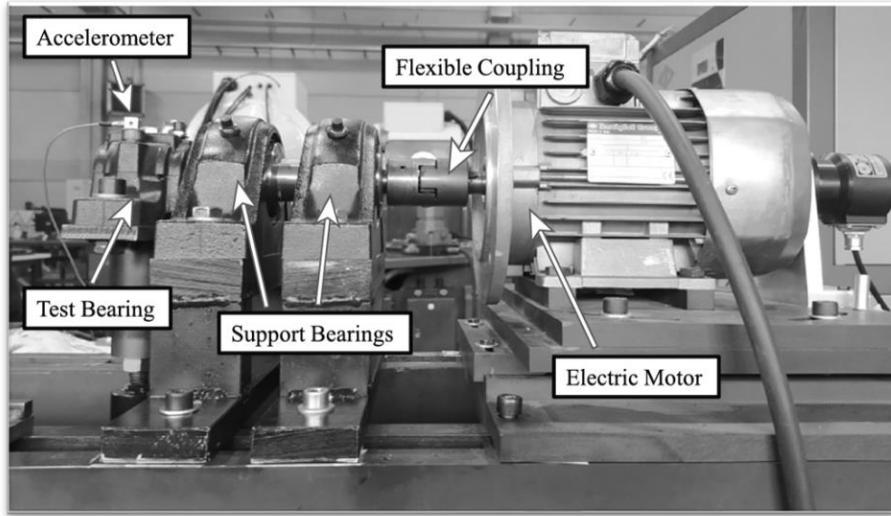


Figure 1: Rotational Setup²

A high frequency accelerometer, model PCB 356B21, measures the vertical Z-axis responses of the test bearing. The data is sampled at 51.2 kHz over a 15-second duration. Only 5 seconds of data utilized in this research project.

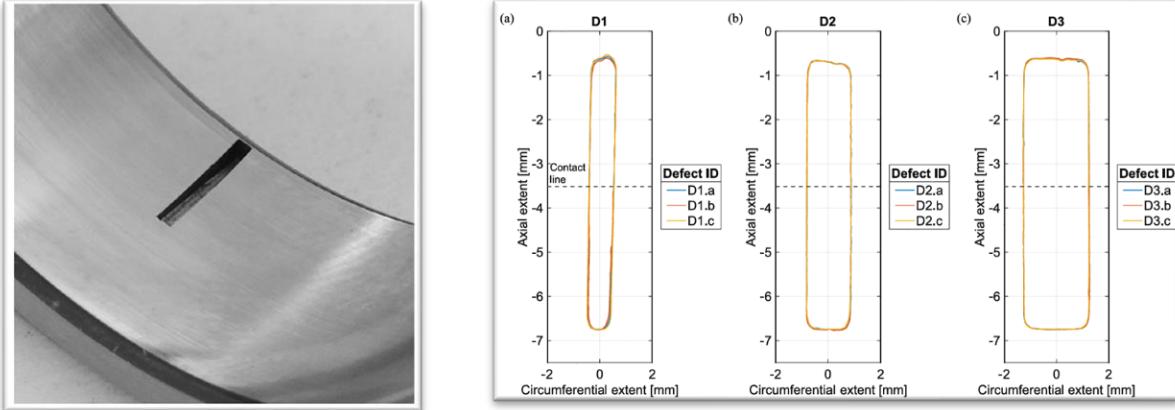
Sixteen cases with varying configurations of ball bearing defects in the outer race and inputs were selected for this project. The configuration categories include defect width, defect angular extent, applied load, and shaft rotation frequency. Each case is labeled in the format "D2c_2000N_40Hz," where the first three characters identify the defect geometry, the middle number represents the input force, and the final number indicates the rotational input frequency used to excite the ball bearings. The table below details the defect geometry measurements.

Table 1: Defect Geometry Legend³

ID	Width [mm]	Angular extent [°]
D1a	0.932	2.372
D1b	0.907	2.308
D1c	0.945	2.405
D2a	1.671	4.253
D2b	1.664	4.236
D2c	1.658	4.218
D3a	2.507	6.382
D3b	2.513	6.397
D3c	2.483	6.32

² <https://www.sciencedirect.com/science/article/abs/pii/S0888327022008512?via%3Dihub>

³ Id.



(a)

(b)

Figure 2: (a) Applied defect (b) Defect geometry⁴



(a)

(b)

Figure 3: (a) Contact area with outer race⁵ (b) contact of one ball bearing.⁶

⁴ <https://www.sciencedirect.com/science/article/pii/S0888327022008512?via%3Dihub>

⁵ <https://www.youtube.com/watch?v=5MNyLM0N6l0>

⁶ Supra footnote 4.

Loading Cases

16 loading cases were chosen for this research project, as laid out in the table below.

Table 2: Selected Cases

Case Number	Cases	Case Number	Cases
1	"D1a_1000N_20Hz"	9	"D2c_1000N_20Hz"
2	"D1a_2000N_40Hz"	10	"D2c_2000N_40Hz"
3	"D1b_1000N_20Hz"	11	"D3a_1000N_20Hz"
4	"D1b_2000N_40Hz"	12	"D3a_2000N_40Hz"
5	"D1c_1000N_20Hz"	13	"D3b_1000N_20Hz"
6	"D1c_2000N_40Hz"	14	"D3b_2000N_40Hz"
7	"D2b_1000N_20Hz"	15	"D3c_1000N_20Hz"
8	"D2b_2000N_40Hz"	16	"D3c_2000N_40Hz"

Ball Bearing Vibrational Responses

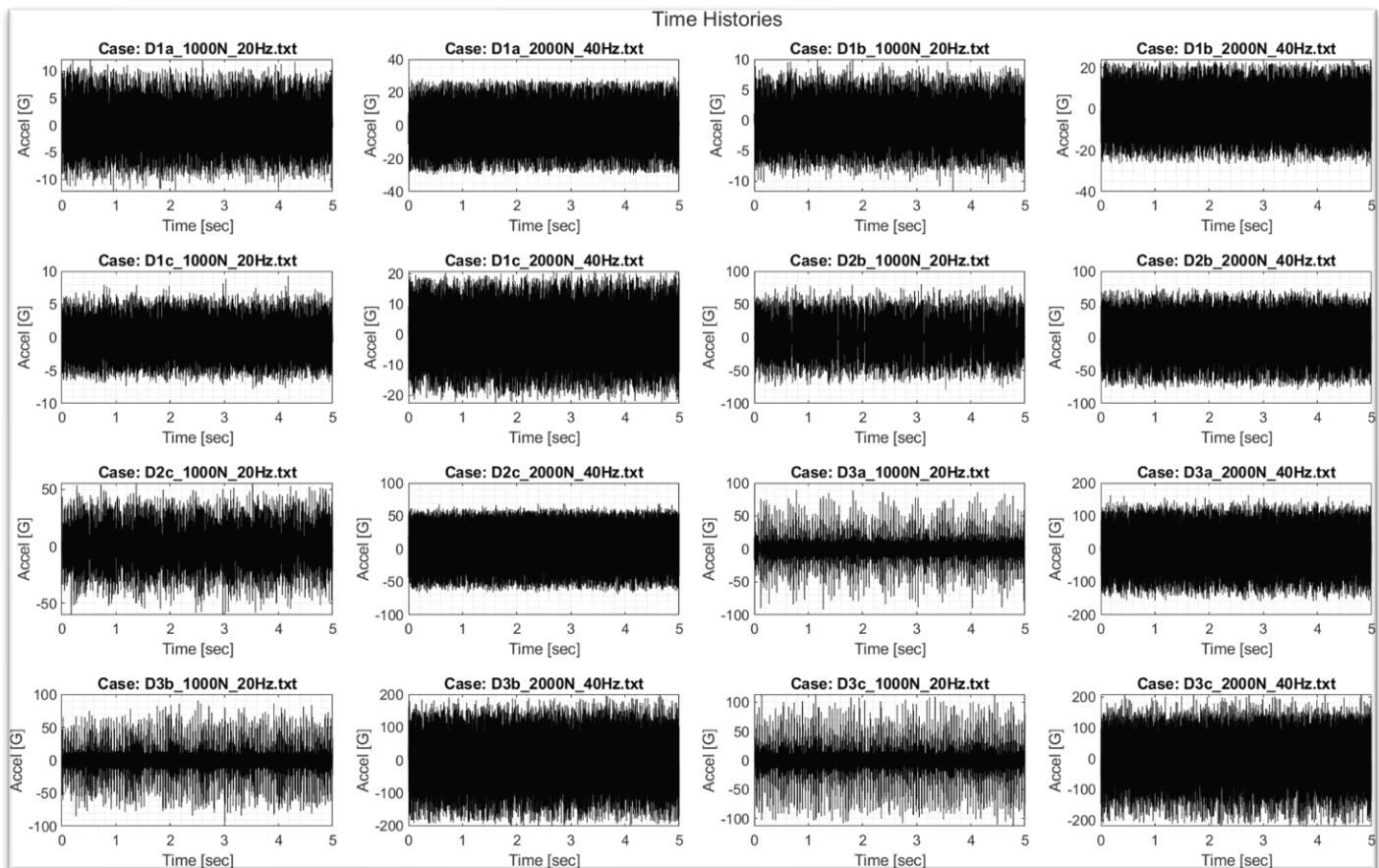


Figure 4: Raw vibrational responses for all cases.

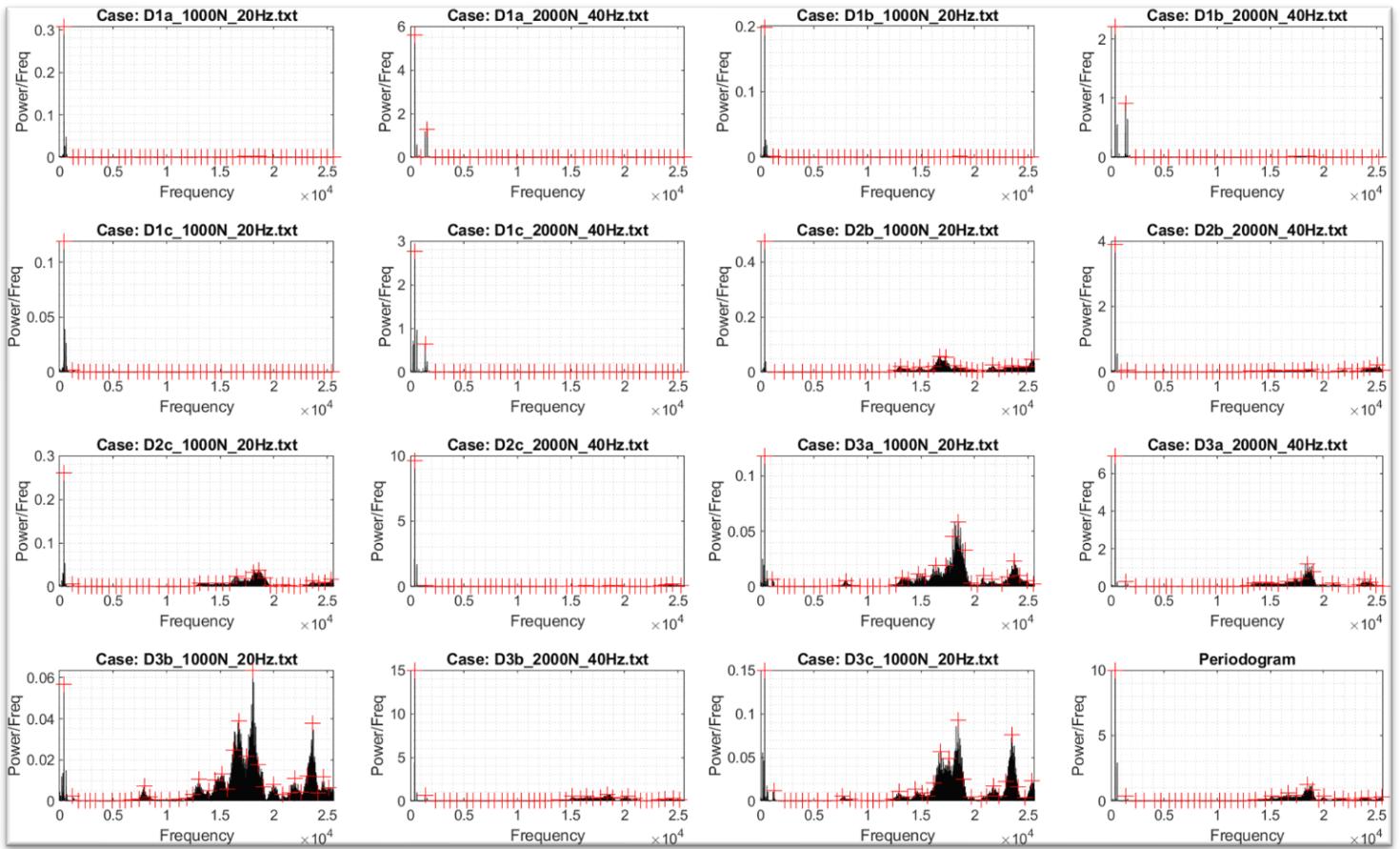


Figure 5: Periodogram of all cases acceleration

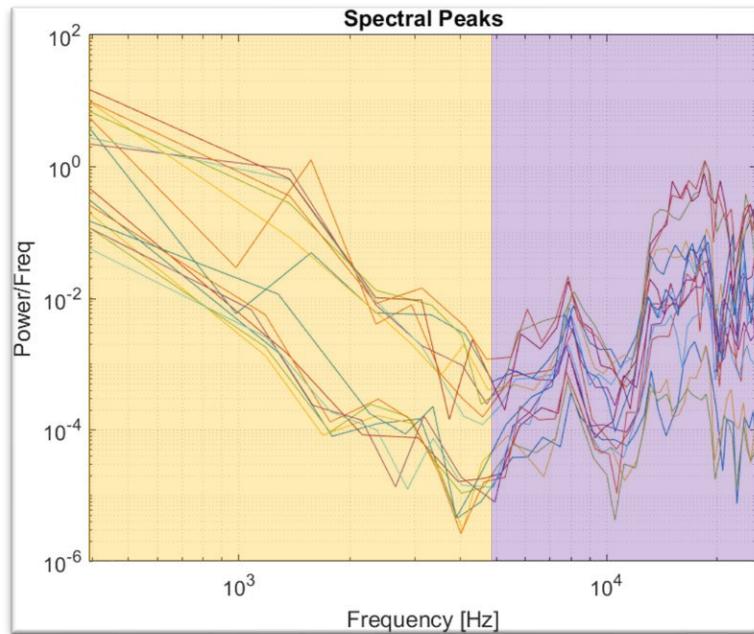


Figure 6: Frequency modes of all accelerometers

The data is divided into two regions:

- I) Low Frequency Region below 5000 Hz, shaded in orange.
- II) High Frequency region above 5000 Hz, shaded in purple.

Data was detrended by a Least Squares Fit (LSF) with the following equation to fix its unrealistic low frequency responses:

$$y_{fit}(t) = a + bt$$

This detrended data produced the following statistical parameters, which are similar to those of the raw data:

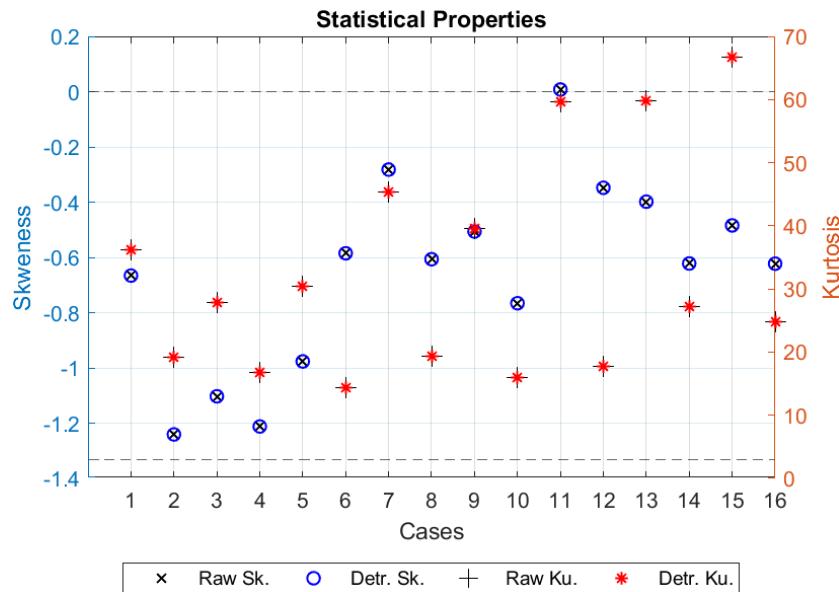


Figure 7: Kurtosis and Skewness of all cases accelerometer

As is customary with high frequency accelerations, these instruments often capture a significant noise floor that can interfere with the actual data. To effectively handle the data, it must follow a normal distribution to allow for comparison with random noise or vibrations. From *Figure 7*, the skewness of the accelerations is negative across all cases, indicating a strong bias toward positive G values. This bias is likely attributable to the installation method of the accelerometer.

Region I: Low Frequency

Frequencies above 5000 Hz in this type of data are typically corrupted by noise if not filtered appropriately. To extract the low frequency content, a lowpass filter is applied to eliminate those high frequency responses.

The following section compares two types of filters, “Chebyshev Type I” and “Butterworth”, with two different polynomial orders: 4th and 10th.

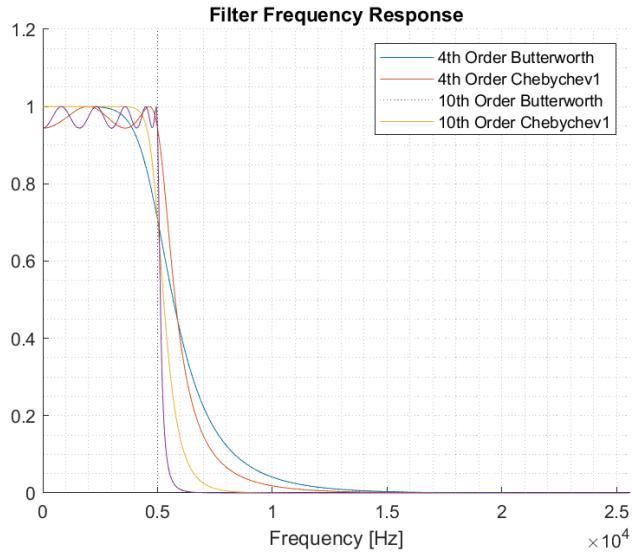


Figure 8: Different filter frequency responses

Selecting an appropriate filter requires balancing the preservation of undisturbed responses below the cutoff frequency with the attenuation of responses above it. The primary objective of the filtering scheme for this application is to minimize responses above 5000 Hz while maintaining the integrity of responses below 5000 Hz and having a sharp transition. The 10th-order Chebyshev filter was chosen for its steeper transition phase around the cutoff frequency. However, a drawback of this filter is that it introduces minor distortions to the responses below 5000 Hz, which are considered negligible for the purposes of this study.

The lowpass region is shown on the following page:

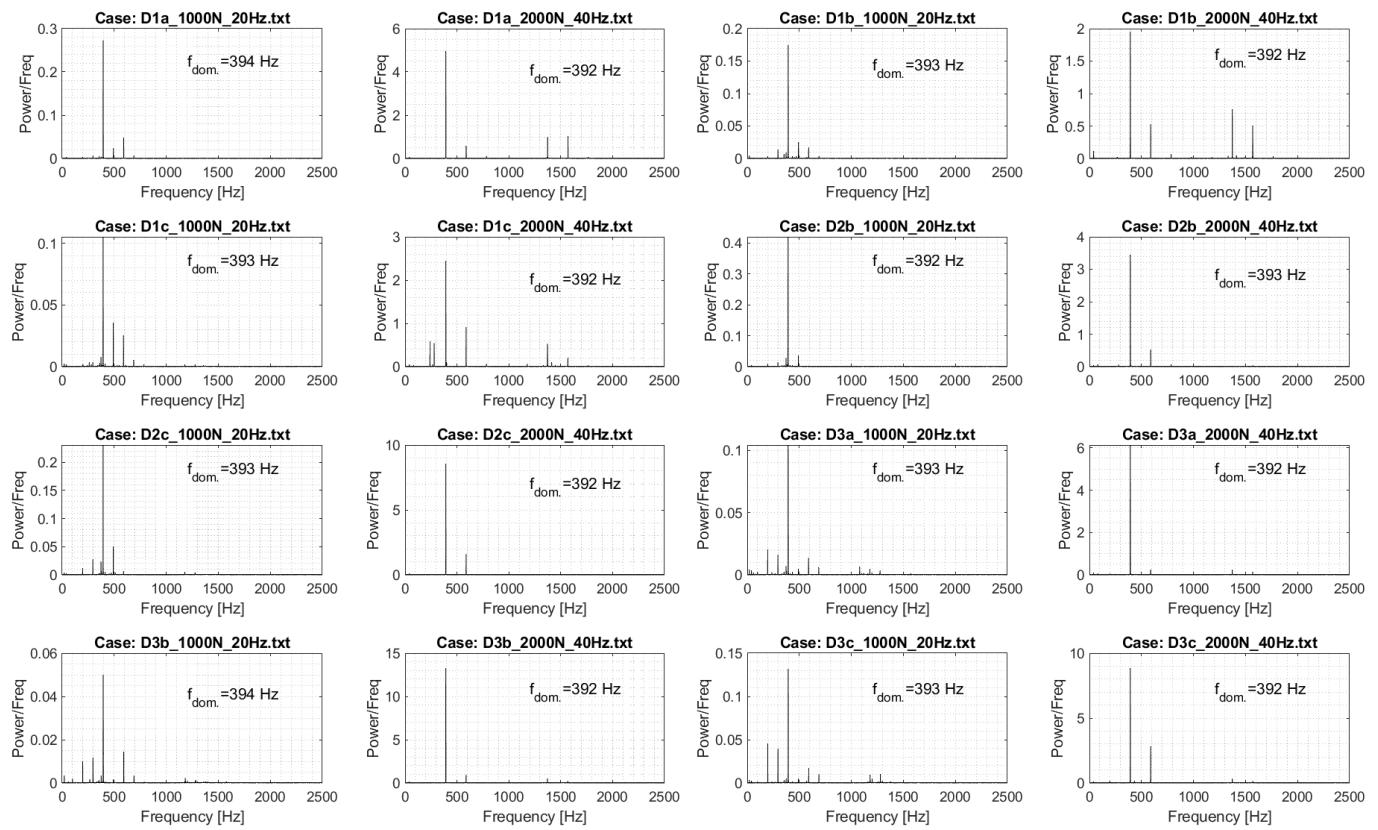


Figure 9: Periodogram of Lowpass Filtered Data

Table 3: Low Frequency Region Dominant Frequencies

Cases		Dominant Frequencies				
		1	2	3	4	5
1	"D1a_1000N_20Hz"	20	197	394	591	788
2	"D1a_2000N_40Hz"	40	240	392	589	785
3	"D1b_1000N_20Hz"	20	197	393	590	786
4	"D1b_2000N_40Hz"	40	160	265	392	588
5	"D1c_1000N_20Hz"	20	142	260	393	589
6	"D1c_2000N_40Hz"	40	240	392	588	784
7	"D2b_1000N_20Hz"	20	196	392	588	785
8	"D2b_2000N_40Hz"	40	280	393	589	785
9	"D2c_1000N_20Hz"	20	196	393	589	766
10	"D2c_2000N_40Hz"	40	196	392	588	784
11	"D3a_1000N_20Hz"	20	196	393	589	785
12	"D3a_2000N_40Hz"	40	196	392	588	784
13	"D3b_1000N_20Hz"	20	197	394	591	788
14	"D3b_2000N_40Hz"	40	196	392	588	784
15	"D3c_1000N_20Hz"	20	197	393	590	787
16	"D3c_2000N_40Hz"	40	196	392	588	785

The lowpass filter enabled the clear identification of relevant frequencies in this range. A few observations can be made from the table above. First, the rotational frequencies for all cases were correctly identified according to their naming convention, as expected. These frequencies are highlighted in green. Second, the Ball Pass Frequency of the Outer Race (BPFO) and its harmonics, as described by Gabrielli⁷, were also identified and are highlighted in blue. Additionally, a dominant mode at approximately 394 Hz was observed across all cases.

Region II: High Frequency

The high frequency region above 5,000 Hz represents the impact or contact area of the ball bearings in their natural frequency. This region is the most indicative of damage in the structure. The high frequency responses are characterized by pulses that, when converted to the frequency domain, clearly reveal defects in the outer race. The frequency window of interest is between 5,000 Hz and 22,000 Hz.

There are two methods to filter data in this region:

- 1) This frequency window can be found by:

$$R_{HF,1} = 1 - R_{LF}$$

where R_{LF} is the frequency response of the filter from Region-I, and $R_{HF,1}$ is a highpass filter above 5,000 Hz. Then, apply a lowpass filter at the upper frequency of 22,000 Hz

$$R_{HF} = R_{HF,1} - R_{LPF@22\text{ kHz}}$$

- 2) Applying a high frequency FIR bandpass filter. This is a preferred method since the lower frequency window of 12,000 Hz can be specified.

A Finite Impulse Response (FIR) with a Hamming window with different orders is shown.

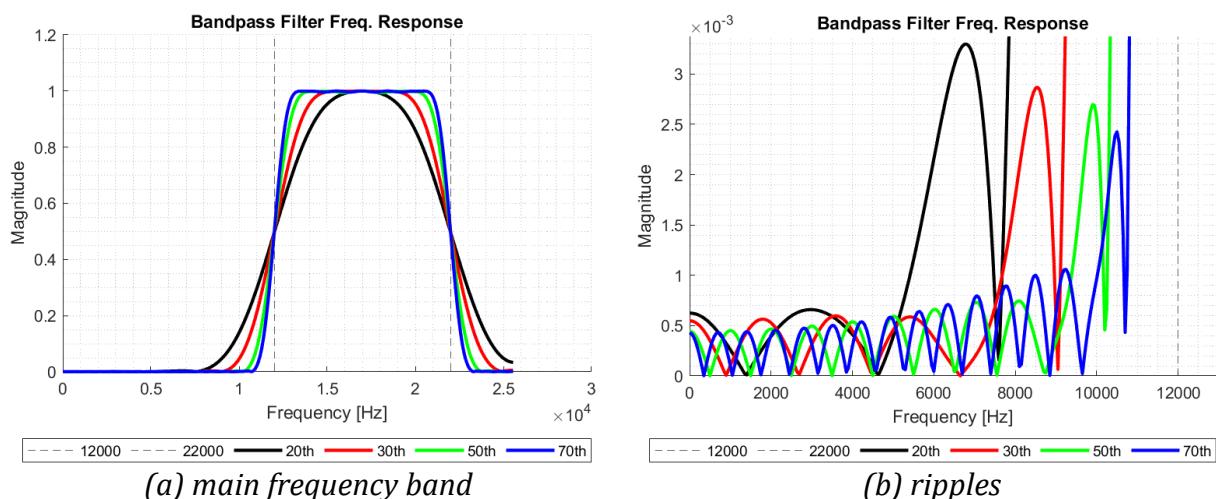


Figure 10: FIR filter using a Hamming window with different orders

⁷ <https://www.sciencedirect.com/science/article/abs/pii/S0888327022008512?via%3Dihub>

A 70th order FIR Bandpass filter was applied to the raw data, and to prevent the edge phenomenon, n/2 points at both ends are deleted from the filtered data.

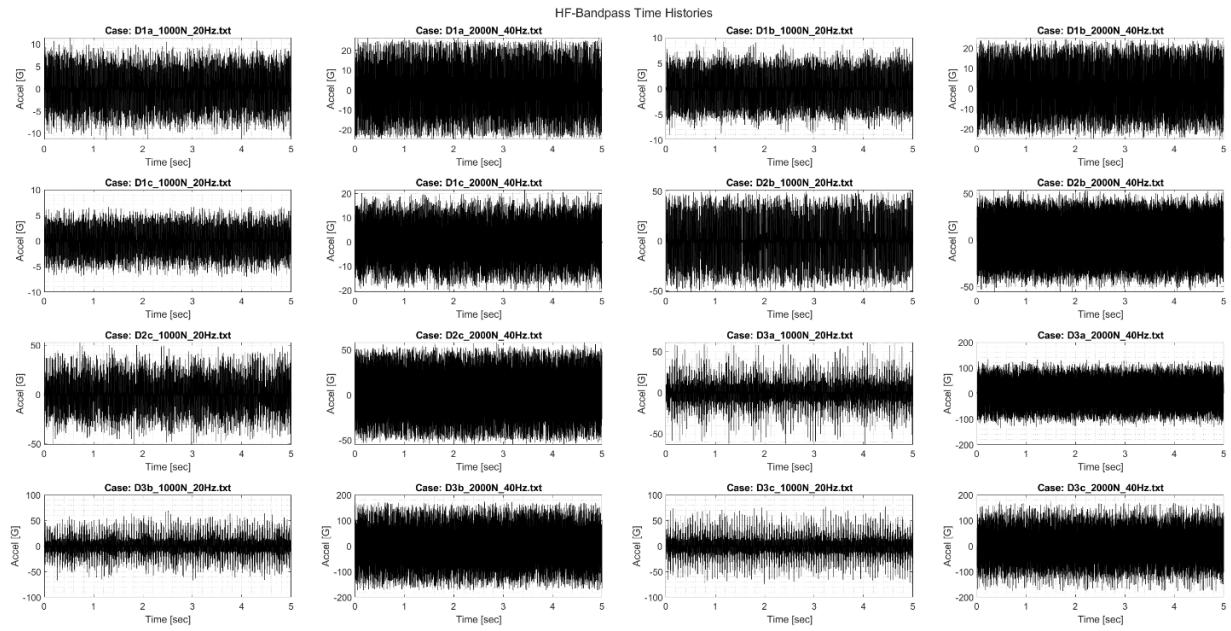


Figure 11: High Frequency Bandpass Filtered Data

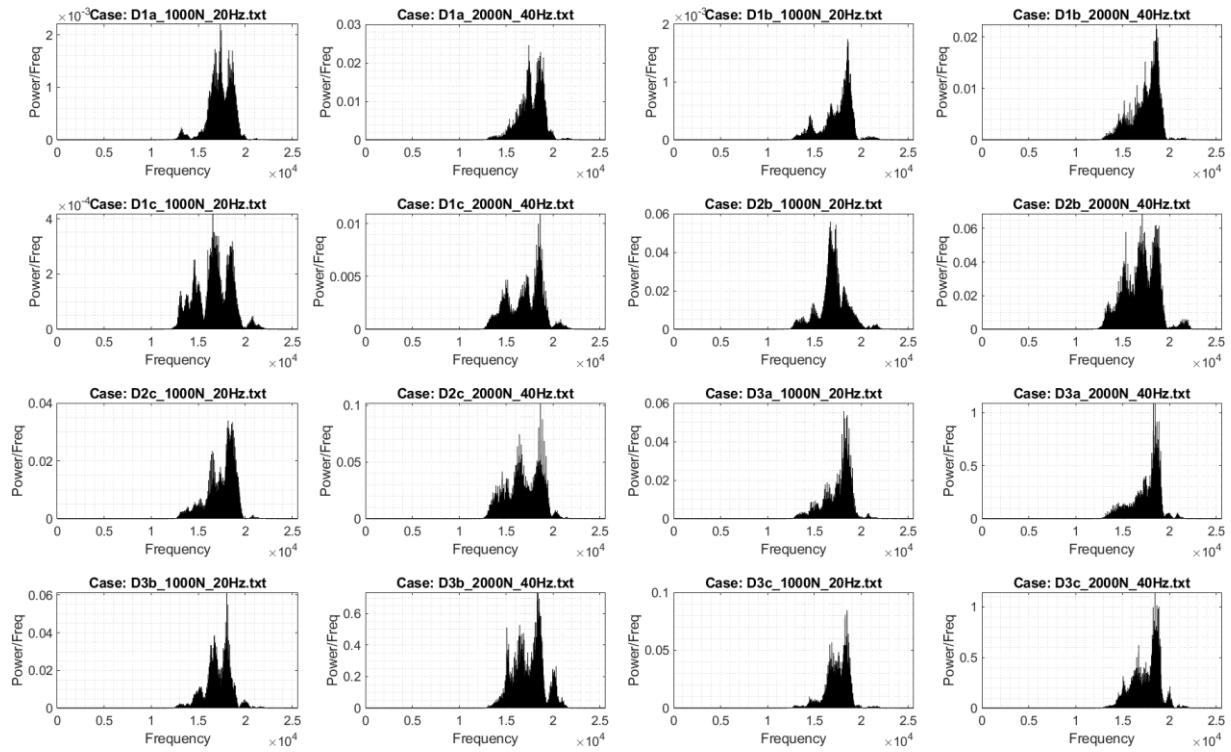


Figure 12: Periodogram of High Frequency Bandpass Filtered Data

With the high frequency data in mind, the time intervals of the high frequency peaks can be determined for each case.

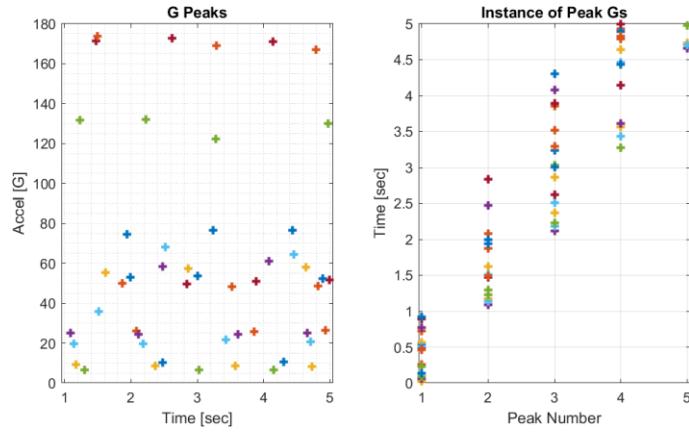


Figure 13: Time and occurrence of the peak loads

Using this plot as a reference, an algorithm was developed to find the high frequency pulses.

The characteristics pulses for each case correspond to the severity of the defect.⁸

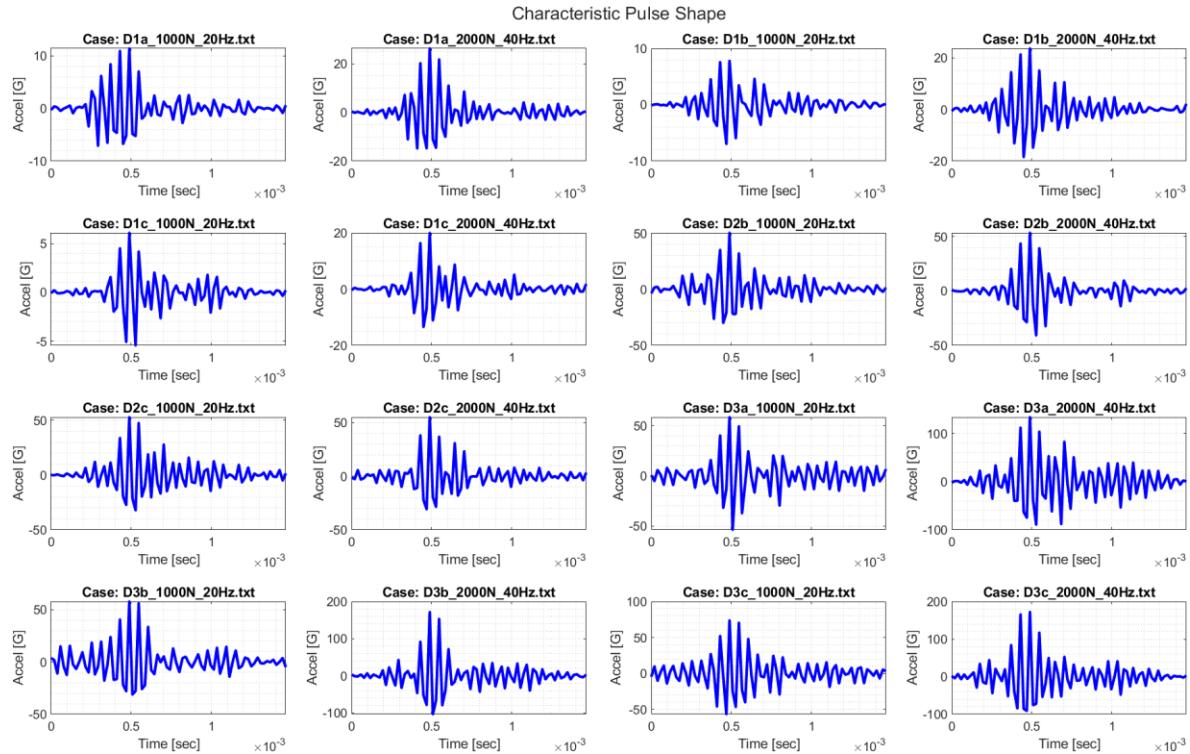


Figure 14: Characteristic pulses for each case

⁸ [Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults - ScienceDirect](#)

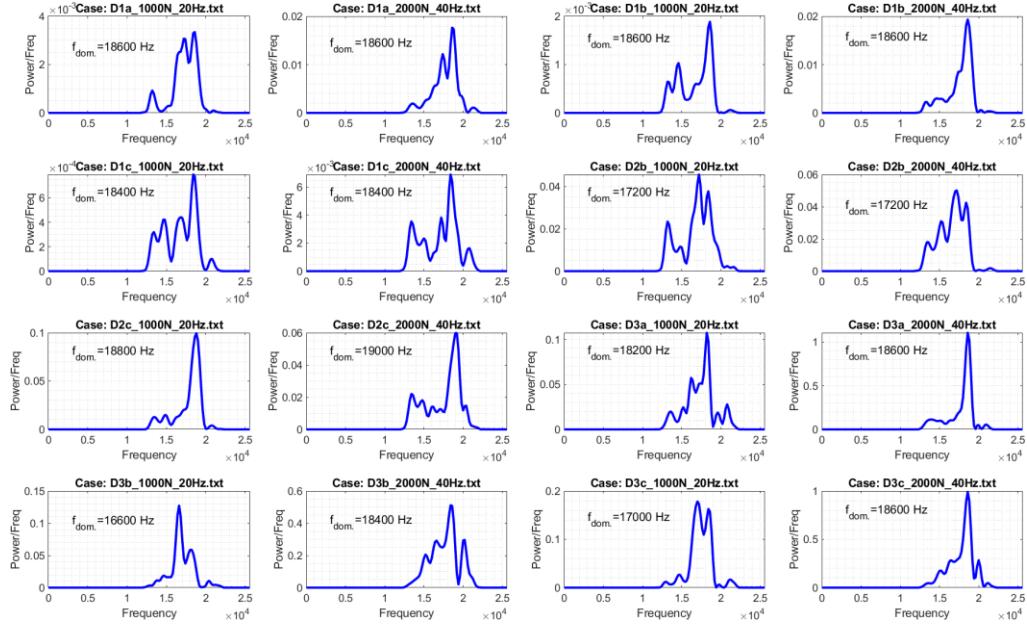


Figure 15: Periodogram of the pulses

By comparing this frequency response with that of the entire filtered data, the high frequency responses have been demodulated.⁹

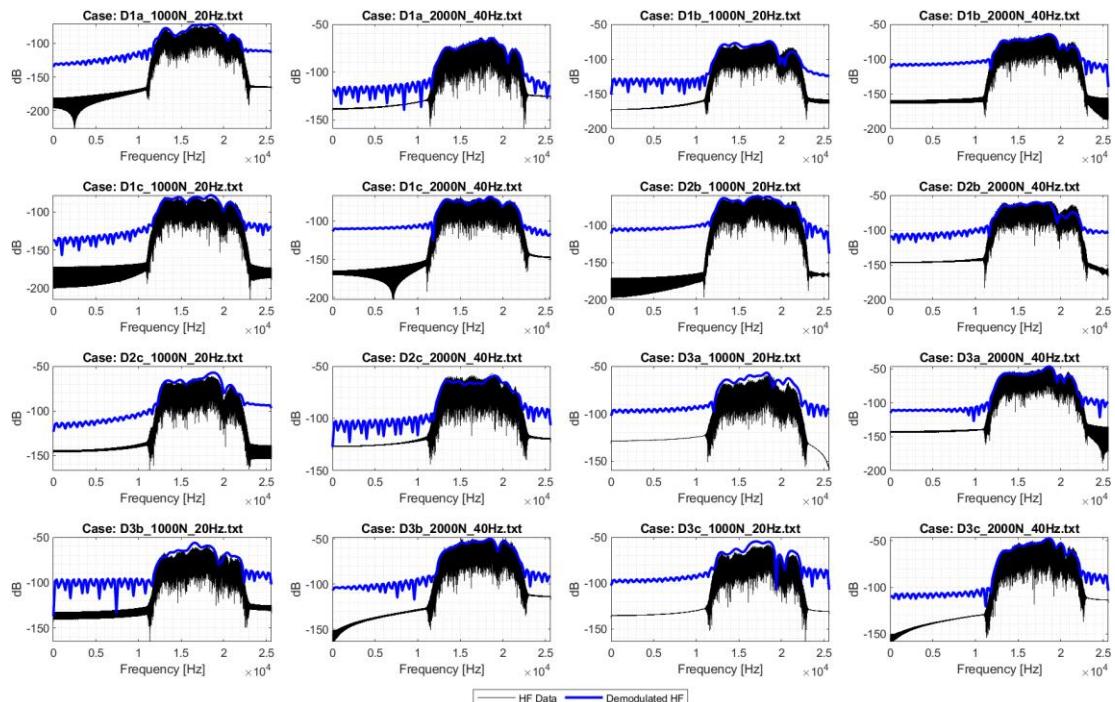


Figure 16: Comparison of High frequency data and demodulation

⁹ Simulating gear and bearing interactions in the presence of faults: Part I. The combined gear bearing dynamic model and the simulation of localised bearing faults - ScienceDirect

Slosh Simulation

The simulation of RP-1 sloshing is performed in LS-DYNA. The input excitation is derived from the low frequency region of the data, Region-I, as most of the fluid displacement is expected to be driven by inputs below 500 Hz.

The Region-I was further manipulated and filtered. Liberty was taken to manipulate the data in the most practical way to expedite the simulation process. The accuracy and fidelity of the simulation results are experimental only and do not guarantee a greater performance for physical application.

Downsampling

With the current accelerometer sampling frequency of 51.2 kHz, the simulation time step must be at least at the twice of 51.2 kHz to prevent aliasing. This would require a simulation time step of 9.8 μ sec, which would take months to generate results. Therefore, the data must be downsampled to a more practical frequency. A sampling frequency of 1000 Hz was chosen, as the frequency content below 394 Hz is most relevant in the low frequency region.

If the data is downsampled simply by selecting points at fixed intervals, it effectively acts as an unintentional lowpass filter. As a result, many of the peak responses are lost, as shown below.

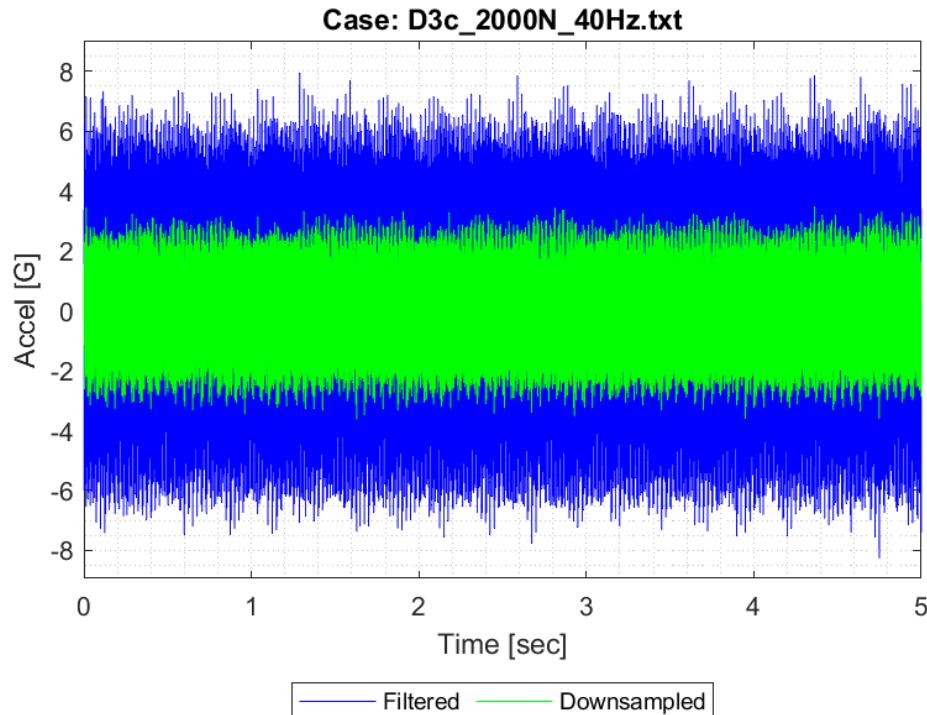


Figure 17: Downsampled Data with Points at Fixed Intervals

To maintain the downsampling rate without losing the G peaks, these peaks can be aligned with the nearest time interval to ensure their incorporation into the simulation. The tuning parameter, “*dist_t*”, represents the distance used as the ‘MinPeakDistance’ parameter within the *findpeaks.m* MATLAB function. Some results with different values of “*dist_t*”, are shown in the Appendix.

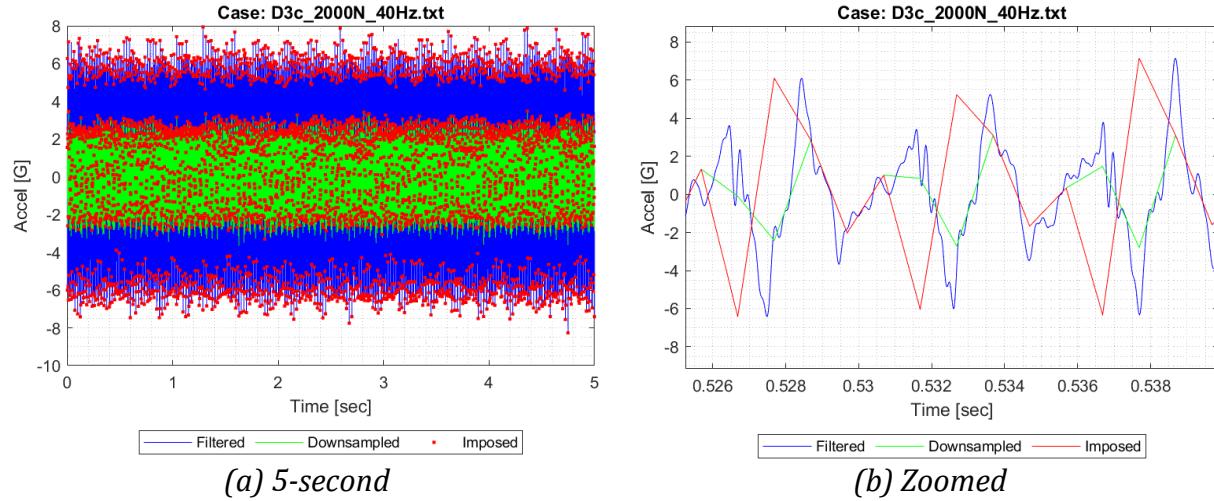


Figure 18: Imposed Peaks to Nearest Downsampled Time Interval

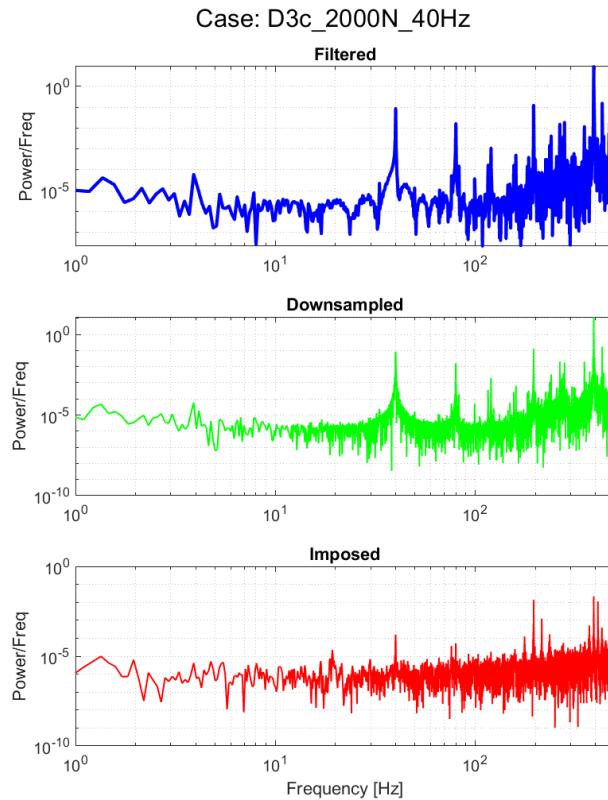
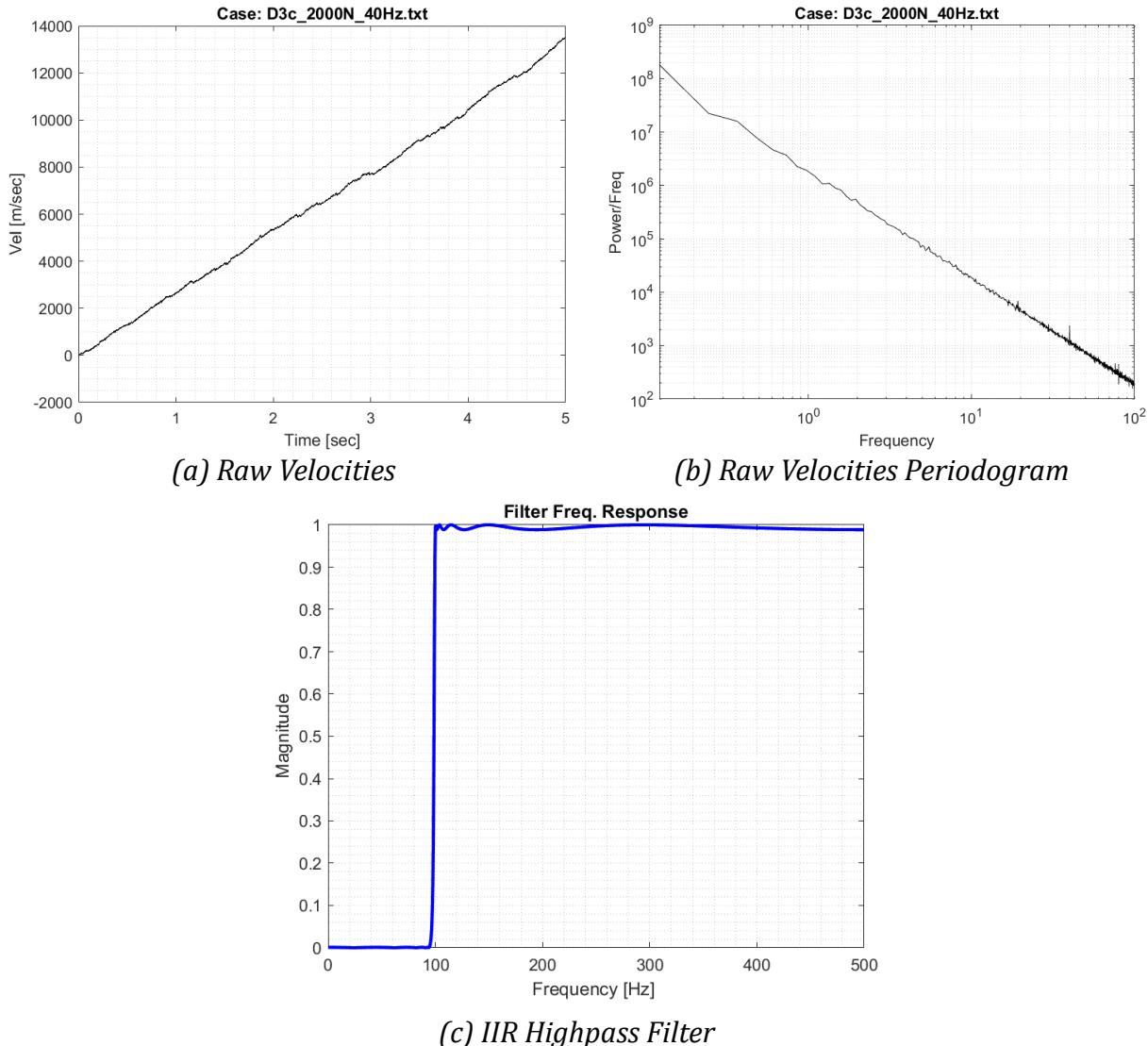


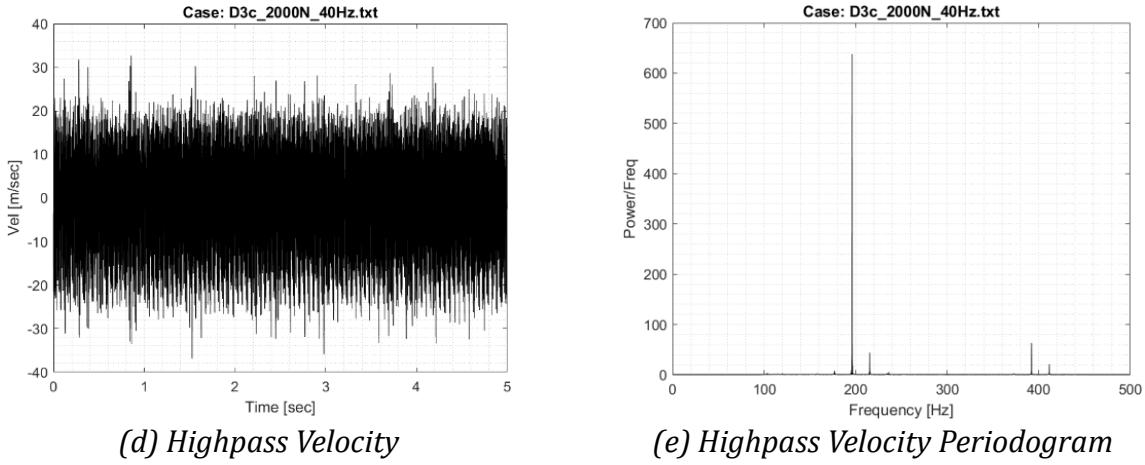
Figure 19: Different Periodograms of the Manipulated Data

As shown in *Figure 19*, the frequency content did not change dramatically. This enabled the continuation of Region-I preprocessing. The imposed low frequency accelerations from Region-I were converted into velocity as LS-DYNA processes these inputs faster. Accelerations can also be specified using the *LOAD_BODY_Z.¹⁰ The conversion from acceleration to velocity is achieved through trapezoidal integration using the *cumtrapz.m* function. The periodogram of these velocities is shown in Appendix.

As expected, all the velocities exhibit a downward trend in the frequency domain, as shown below. This trend results from the integration process. To correct this, the data is filtered using a highpass IIR filter, as shown in Figure 20.



¹⁰ <https://lsdyna.ansys.com/manuals/>



(d) Highpass Velocity

(e) Highpass Velocity Periodogram

Figure 20: Data Filtered Through Highpass IIR Filter

Input Excitations

To accommodate an inertial velocity input, the velocity profile must oscillate about zero. To achieve this, a LSF was applied to remove the residual trend in the velocities.

Due to the imposed sampling, some velocities exceeded the 3-sigma threshold. These outliers were excluded from the velocity data array to enhance simulation solver stability. In the figure below, the magenta data represents the input excitation used in the slosh simulation inside the cubic tank and the dashed lines the 3-sigma thresholds.

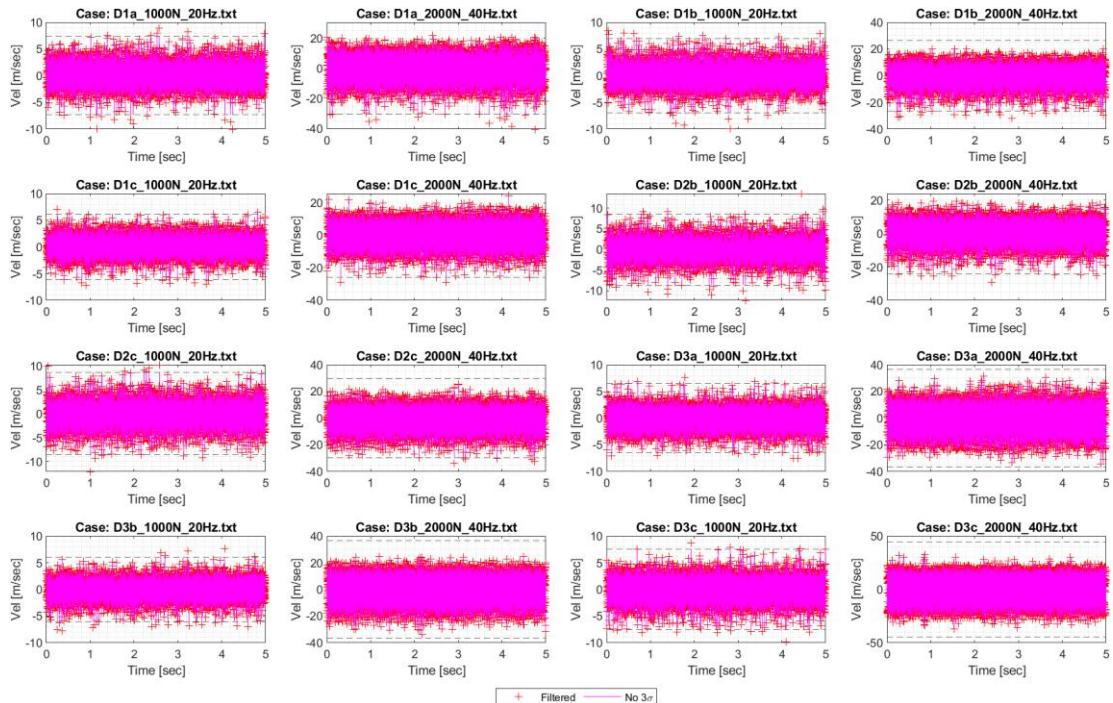


Figure 21: Input Excitation in Slosh Simulation

LS-DYNA Runs

The LS-DYNA runs simulated the following setup:

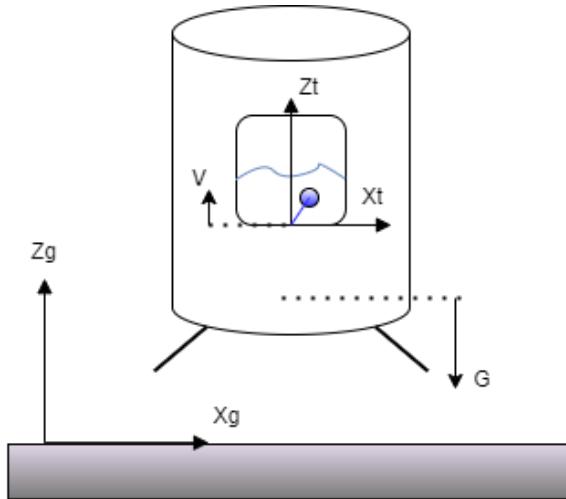


Figure 22: Model Schematic

At an unforced lunar descent with Lunar gravitational force G and input excitations to the tank as V , the CG shift of the RP-1 was treated in the local frame as shown in red. The spacecraft has $\frac{1}{4}$ tank full. At no point in the simulation did the spacecraft touch the lunar surface as this was out of the scope of the project.

The meshing for the cubic tank was inspired by a mesh model setup by DynaExamples.¹¹

The simulation used the ICFD solver of LS-DYNA, and to expedite the process, each element was automeshed. As shown in the figures below, there were three parts of the assembled mesh: 1) the filled space of the tank 2) the free space of the tank and 3) the RP-1 surface (which is hardly visible). All three parts represented boundary conditions for the RP-1. The input excitation was modeled and controlled by the *ICFD_CONTROL_IMPOSED_MOVE.¹² The input excitations are applied along the Z-Axis of the global frame.

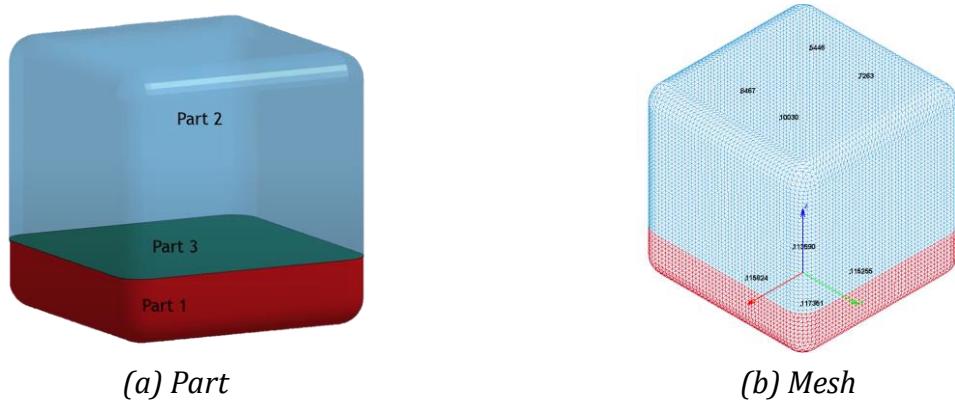


Figure 23: Cubic Tank in LS-DYNA

¹¹ https://www.dynaexamples.com/icfd/basics-examples/global_imposed

¹² <https://lsdyna.ansys.com/manuals/>

The simulation is performed using the following control parameters:

Table 4: Control Parameters

Parameter	Value	Units
Gravity	1.62	m/sec ²
Plotting time step	.05	Sec
Solver time step	.01*	Sec
End time	5.1	Sec
ICFD Boundary Nonslip	Body 1 and 2	~
Body Movement	Velocities	m/sec
Initial CG	(0,0,0.1285)	(x,y,z) [m]

* This time step was modified to this value because the solver took 12 hours to finish at a time step of .001 seconds, which would have taken a week or so to complete all simulations. The downside of this is that at a time step of 0.01, the input excitations are essentially being aliased. For the purpose of this research project this was acceptable in order to meet the project deadline.

Simulation Results

Customized *.cfiles*, LS-DYNA command files, were developed for each case in *Table 2* to extract data from the simulation output *d3plot* and *icfd_lsvol.dat* files, which contained the CG values as a function of time. Each folder size was about 10 GB. Overall, the entire simulation for all cases amounted to ~150 GB, with an average of 2.5 hours per run to complete using about 100 MB of dynamic memory for the solver. Running these simulations on a typical computer would have taken more than 10 hours.

The following shows the raw CG shifts in each orthogonal direction.

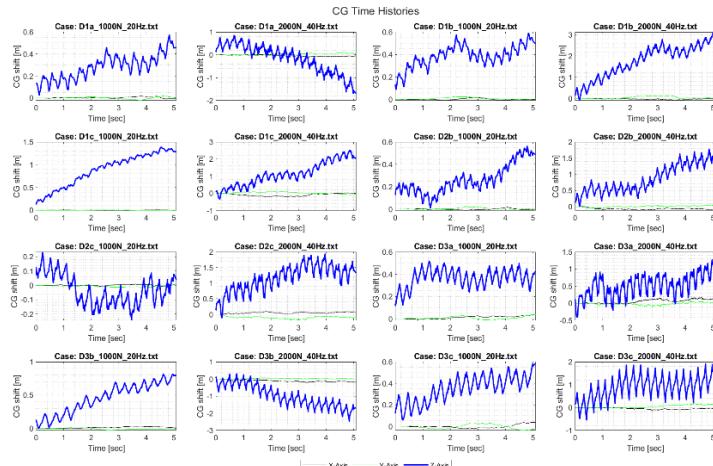


Figure 24: Raw CG Shifts

As expected, the Z-Axis CG shifts dominated the movement. It was unclear which reference frame that LS-DYNA automatically used on the raw CG data in the *icfd_lsvol.dat* file. Time was invested in finding out about that frame, but little to no documentation was found. Therefore, this data needed to be Root Sum Squared (RSS) for a single resultant magnitude, CG_r , as shown below and detrended to make it local to the tank frame with respect to the initial CG position.

$$CG_r = \sqrt{CG_x^2 + CG_y^2 + CG_z^2}$$

The detrending is performed by an IIR highpass filter as shown below and using a LSF.

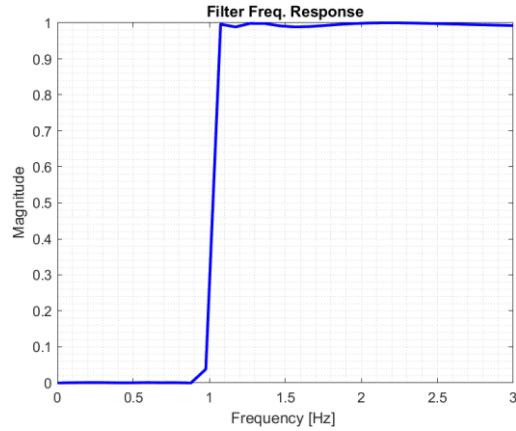
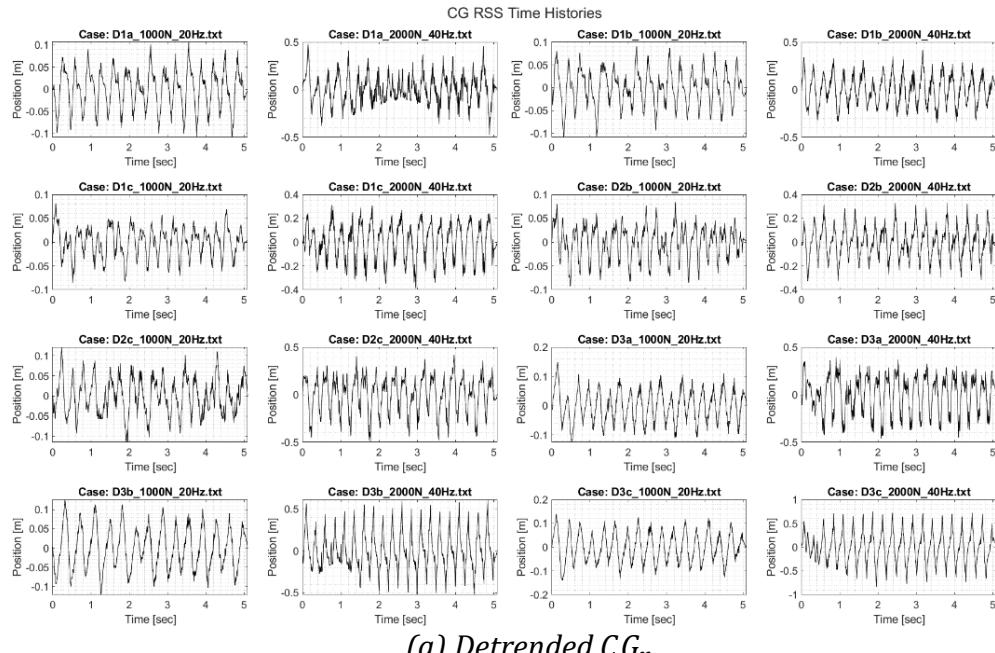
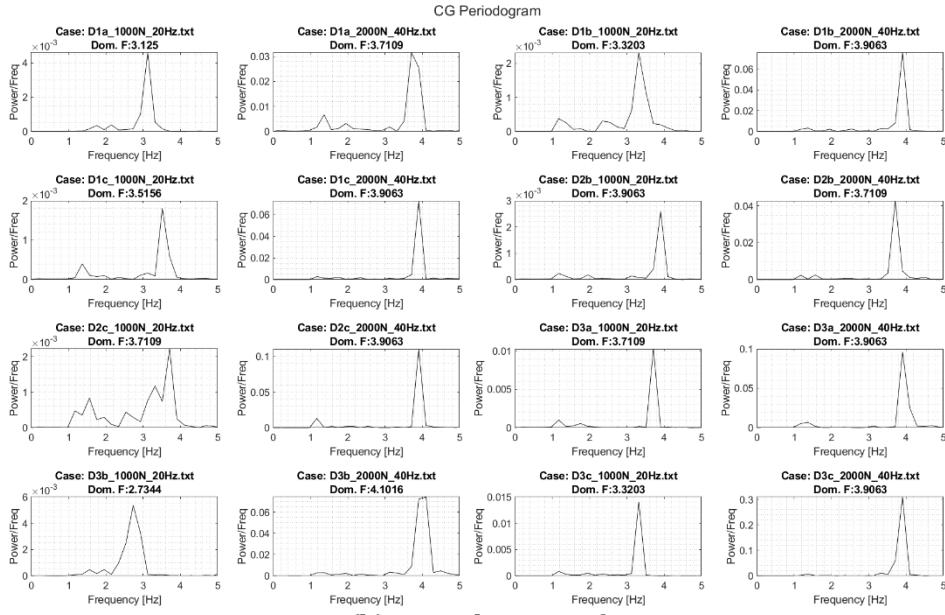


Figure 25: Detrending Performed by IIR Highpass Filter

The CG_r is plotted below:



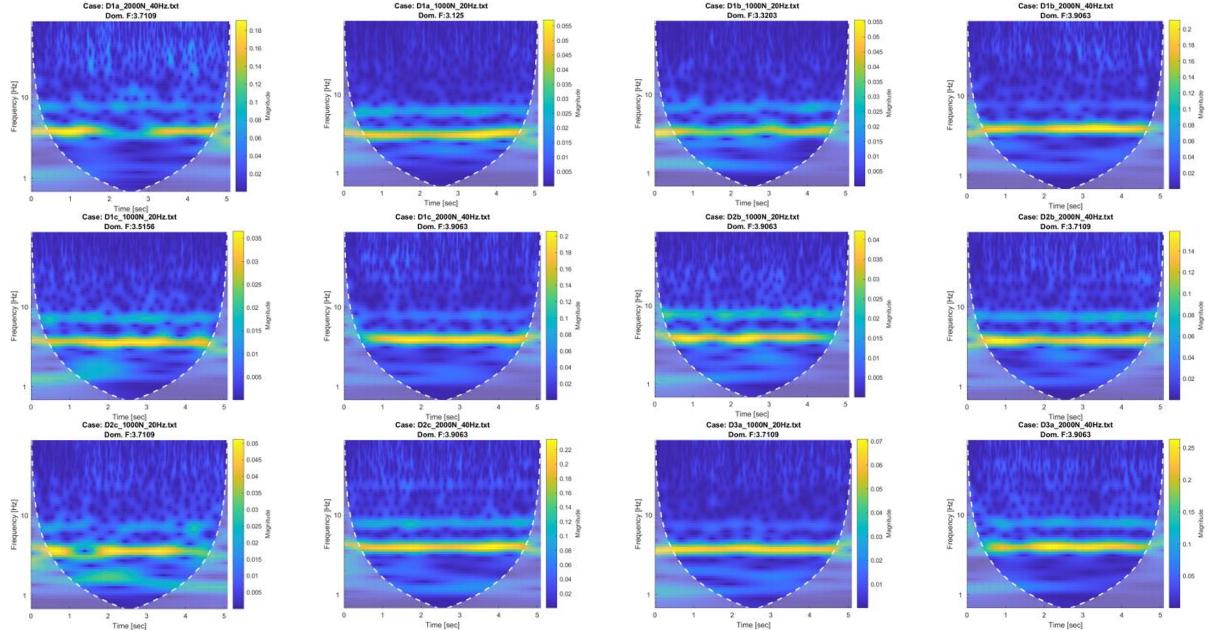
(a) Detrended CG_r



(b) Periodogram of CG_r

Figure 26: CG_r

The periodogram shows that on average the RP-1 has a displacement resonance between 3 and 5 Hz. Since these responses can also have a transient characteristic which is not detected in a simple periodogram, a Continuous Wavelet Transform scalogram is computed to identify such behaviors. These wavelet scalograms show indeed the dominant frequencies and no transient characteristics.



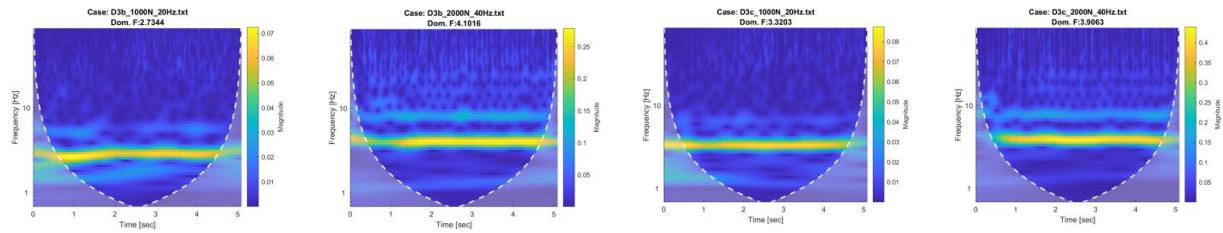


Figure 27: Continuous Wavelet Transform Scalogram of the CG_r

With the CG_r preprocessed, filtered, and briefly analyzed, one aspect is clear, the CG_r is oscillatory in nature, and the 5.1 seconds of simulation show no new frequency content as shown in the scalograms. This can be used against evidence of not using a long simulation time. The variance of the data at each 1-second interval can be computed to determine how much new information is really obtained from running a longer simulation. One detail that vertical slosh does is create Geyser slosh, which was observed in the simulation results (some plots are attached in the Appendix). The reported CG_r potentially uses a weighted average between the fluid nodal points. Time was invested into studying these calculations but turned very lengthy and out of the scope of data analysis, which is the objective of this research project. Therefore, results of the CG shifts are not guaranteed accurate, but it is sufficient for the purposes of this research project.

Figures for the simulation are shown below (More figures are located at the repository)¹³

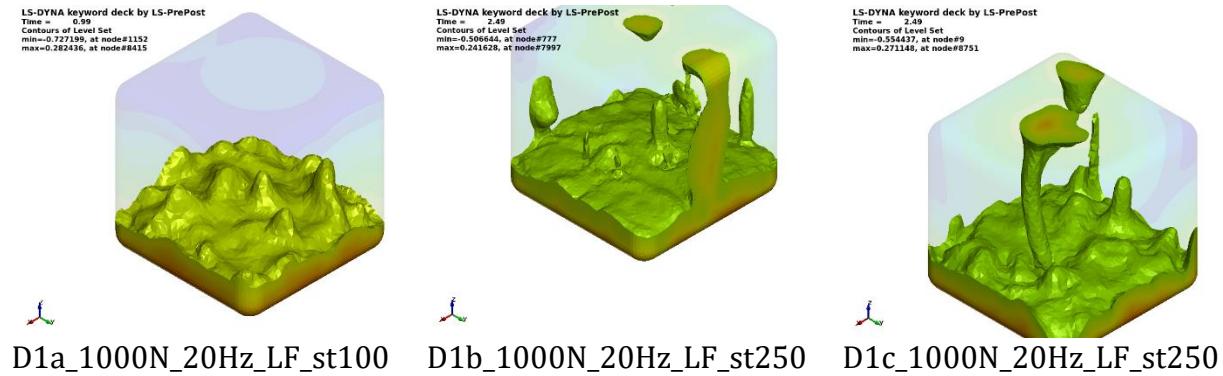


Figure 27-Extra: LS-DYNA Simulation Graphics (Shows Geyser slosh)

¹³ [repository](#)

Surface Response

Surface response is referred to a statistical technique used to model and analyze problems in which a response (output) of interest is influenced by several variables (inputs). The main idea is to develop an approximation of the true relationship between the inputs and the output using observed points known as seeds.

The Kriging interpolator uses a basis function of the form:

$$\hat{Z}(x_0) = \sum_{i=1}^N \lambda_i Z(x_i)$$

where \hat{Z} is the predicted value, $Z(x_i)$ is a seed point, λ_i is the weight at that specific location, and N is the number of available seed points. There is more mathematical background behind this interpolator which will not be explained in this research project. The Kriging interpolator is essentially a two-step process: fitting and predicting.

On the first step, a regression function and a correlation model are constructed based on the lag variogram, which is a plot of the distances between seed points. Out of these two parameters, the correlation model is the most important as it establishes the relationship between neighboring points as well as the relationship between observed points that are farther apart.

On the second step, a mesh grid with nodal values in the design space (direction of input variables) are given to the predictor to estimate values at those specific nodes using a fitter model from the first step.

Extensive literary research led to the study of a paper by Nielsen et al. from the Technical University of Denmark¹⁴ which provided a set of MATLAB functions, *dacefit.m* and *predictor.m*. These two MATLAB functions are essentially the Kriging interpolator. The fitter function uses a maximum likelihood estimation (MLE). The functions were briefly studied to compare to certain mathematical concepts such that the scripts can be used to generate interpolation results. After a few minor modifications to these two .m functions, it was decided that these functions were ready to be used in the interpolation as is.

For this research project, the design space was two dimensional: time in seconds, and velocity in m/sec. The output of the Kriging interpolation was the CG shift in meters.

Two-Dimensional Kriging

As a spatial interpolator, Kriging uses pairs of variables to define the design space. This pair is analogous to planar coordinates in the case of this project, such as latitude and longitude. In this application, the input variables are (velocity, time), and the output is the CG shifts. The figure below shows the comprehensive seed points available for the interpolation.

¹⁴ <https://orbit.dtu.dk/en/publications/dace-a-matlab-kriging-toolbox>

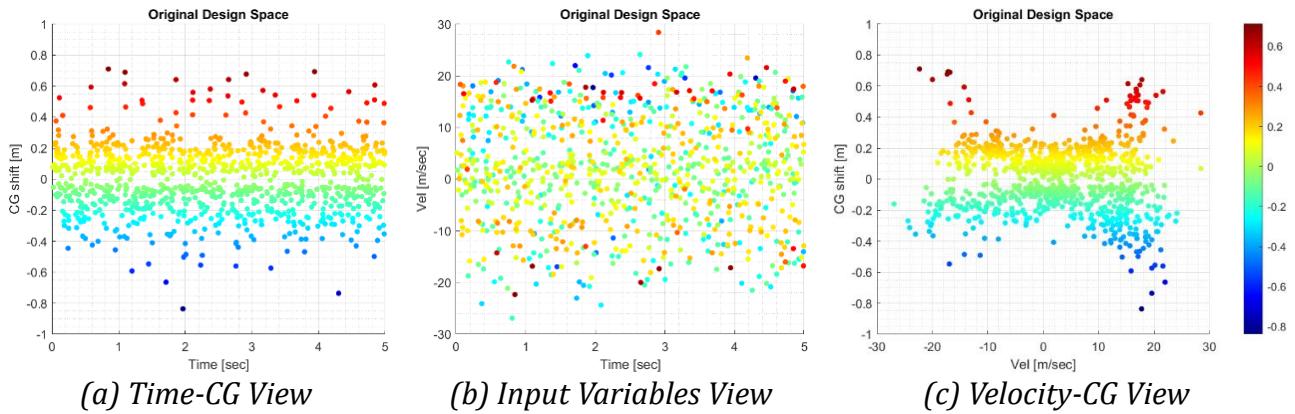


Figure 28: Comprehensive Seed Points

The respective variograms with the lag considering all 16 cases are shown below

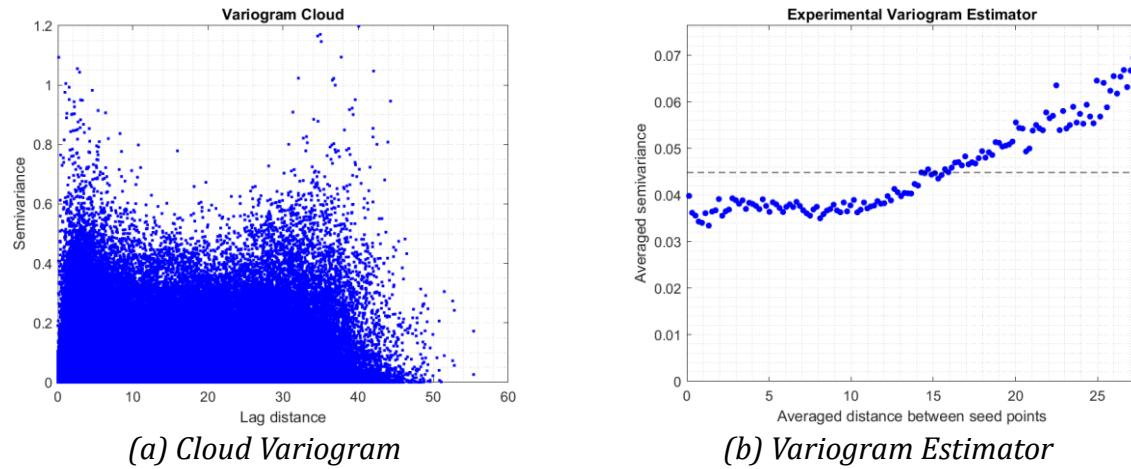


Figure 29: Variograms for All 16 Cases

The experimental variogram showed a unique trend for the seed points. The shape resembled that of a Gaussian variogram. The *dacefit.m* could then be applied using a Gaussian correlation model. Recall that the *dacefit.m* has two main inputs 1) the regression model and 2) the correlation model. After a few iterations of changing the regression model to a different polynomial order (0th, 1st, and 2nd) it was noticed that the changes in output between these regression models were negligible. Therefore, for simplicity, a 1st order polynomial was used as a standard parameter for the Kriging interpolations. Another assumption made was that the design space was anisotropic, meaning the variance for the design space directions, time, and velocity were different, but with the same degree.

For the second step, the *predictor.m*, an equidistant square mesh grid of 250 points with design space area of 0 to 5 seconds in the time direction and -30 to 30 m/sec in the velocity direction was generated for the surface response predictions.

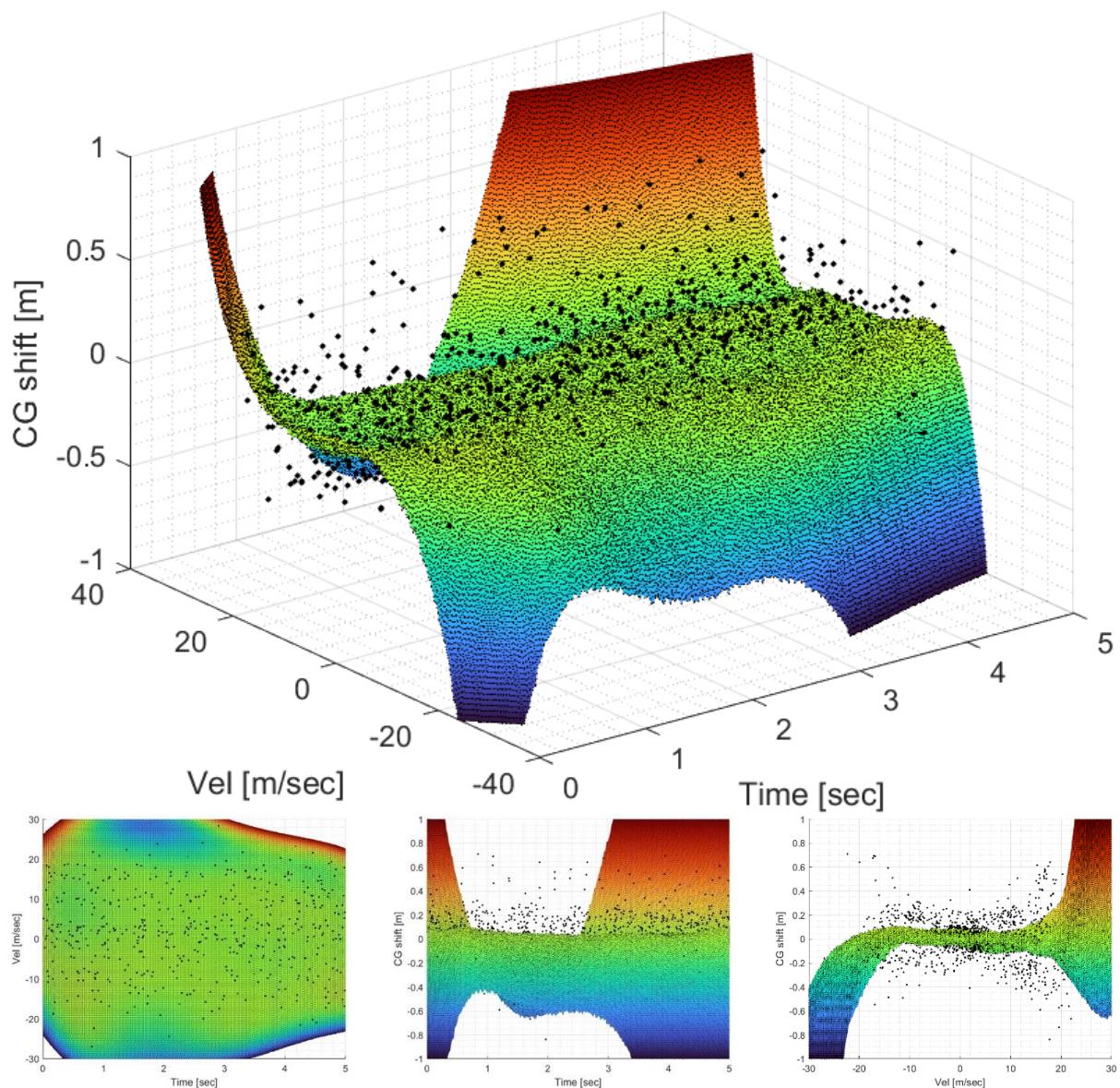


Figure 30: Surface Response Using All Available Seed Points, referred to as SurfaceResponse1

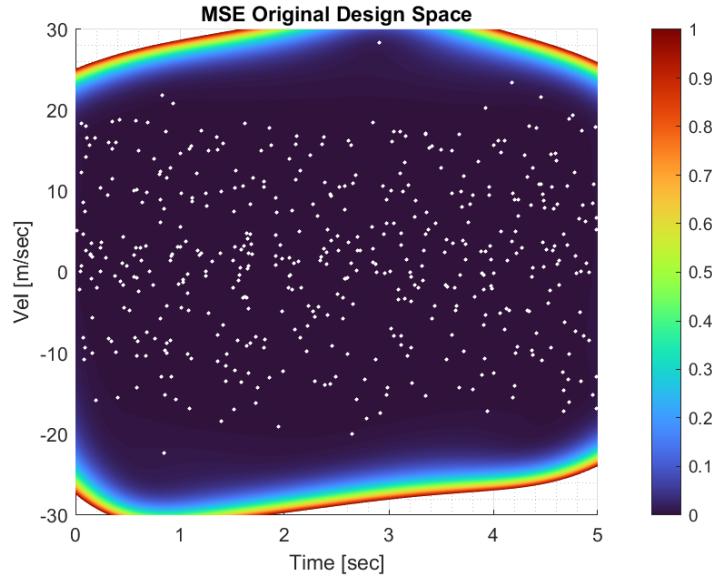


Figure 31: Mean Square Error of SurfaceResponse1

This surface response map, *SurfaceResponse1*, passed through the center of the values at a CG shift of 0. This occurred because the seed responses oscillate and cancel out within the fitter algorithm. When looking at the seed points on the velocity-CG plane, a particular trend could be seen. This trend resembled an “X” shape as shown in *Figure 32(a)*. The algorithm did not know which direction to interpolate next. The variance and distance between the seeds are what dominates the algorithm, which creates bias at points where seed clusters exist. This behavior could be seen in the fabric of *SurfaceResponse1* where elevations were generated at the positive velocities and depths in the negative velocities as shown in *Figure 32(b)*. The behavior of *SurfaceResponse1* made the map inaccurate and, therefore, *SurfaceResponse1* could not predict values within the seed points as expected.

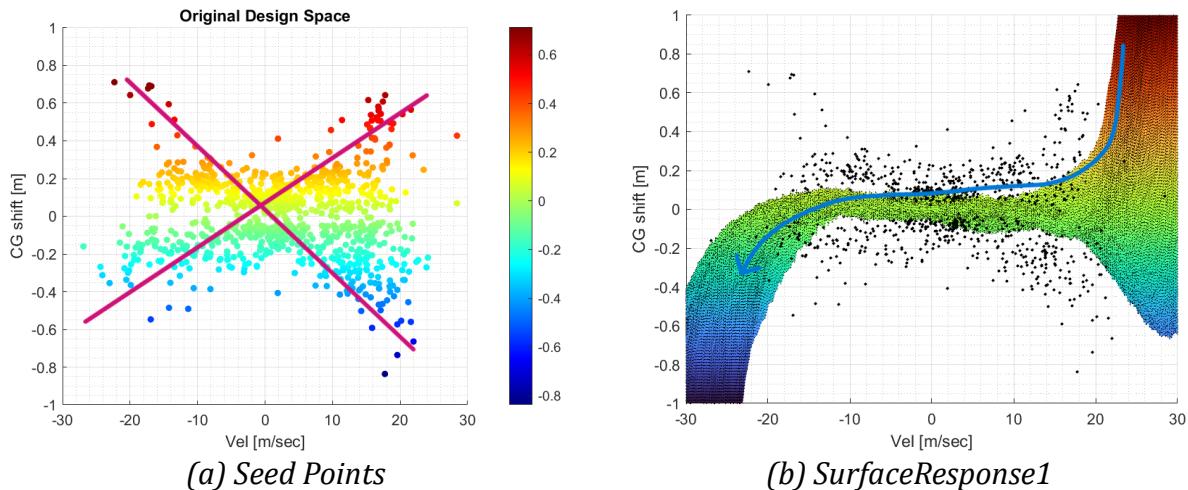


Figure 32: Trends of Seed Points and Interpolations on Velocity-CG Plane

Since data behaved in an oscillatory manner, different surface responses could be generated. One map for the positive CG shift values, and another map for the negative CG shift values. This way the interpolated fabric covers both areas as shown below:

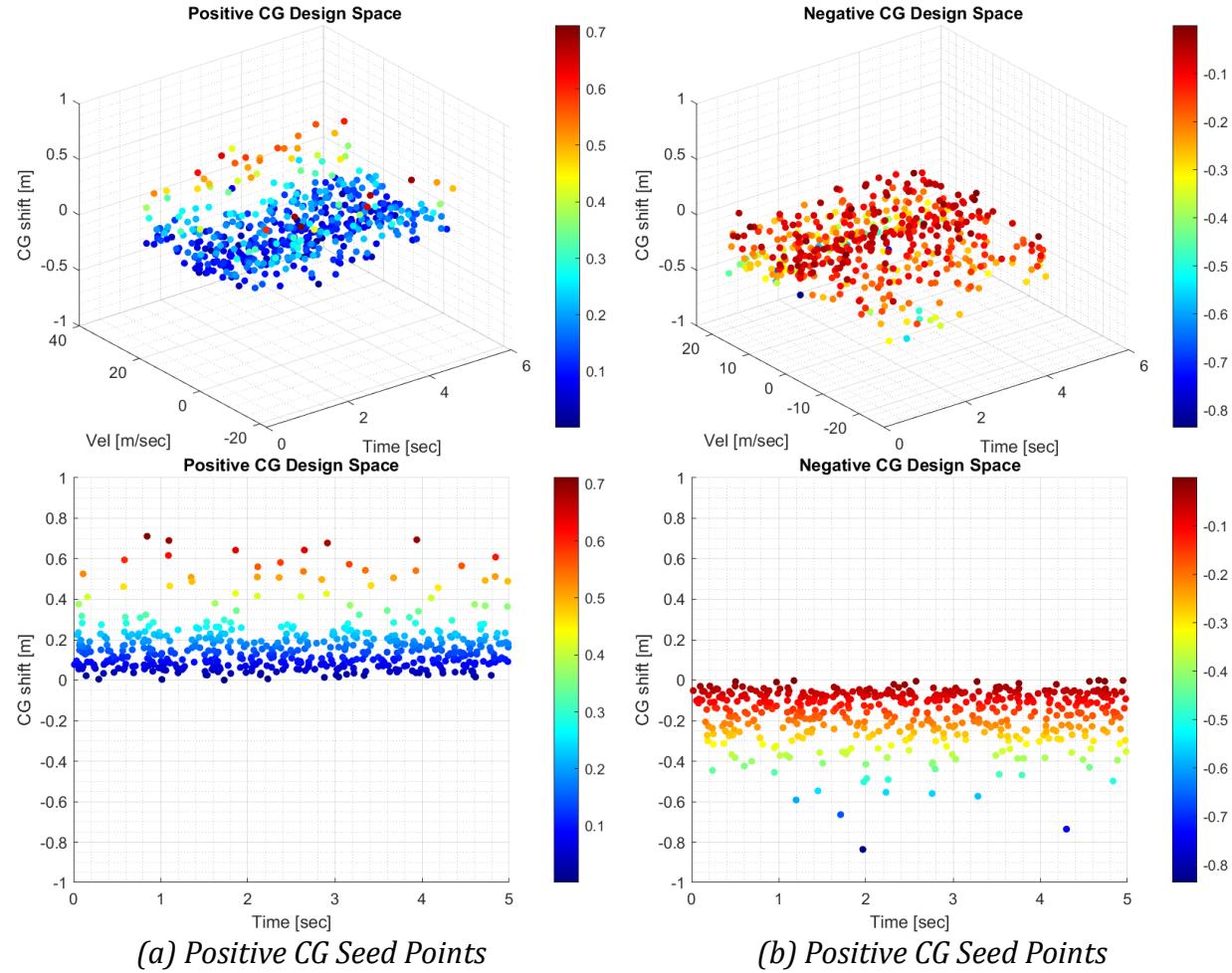


Figure 33: Seed Points Positive and Negative CG Shift Values

The results for each surface response are shown below.

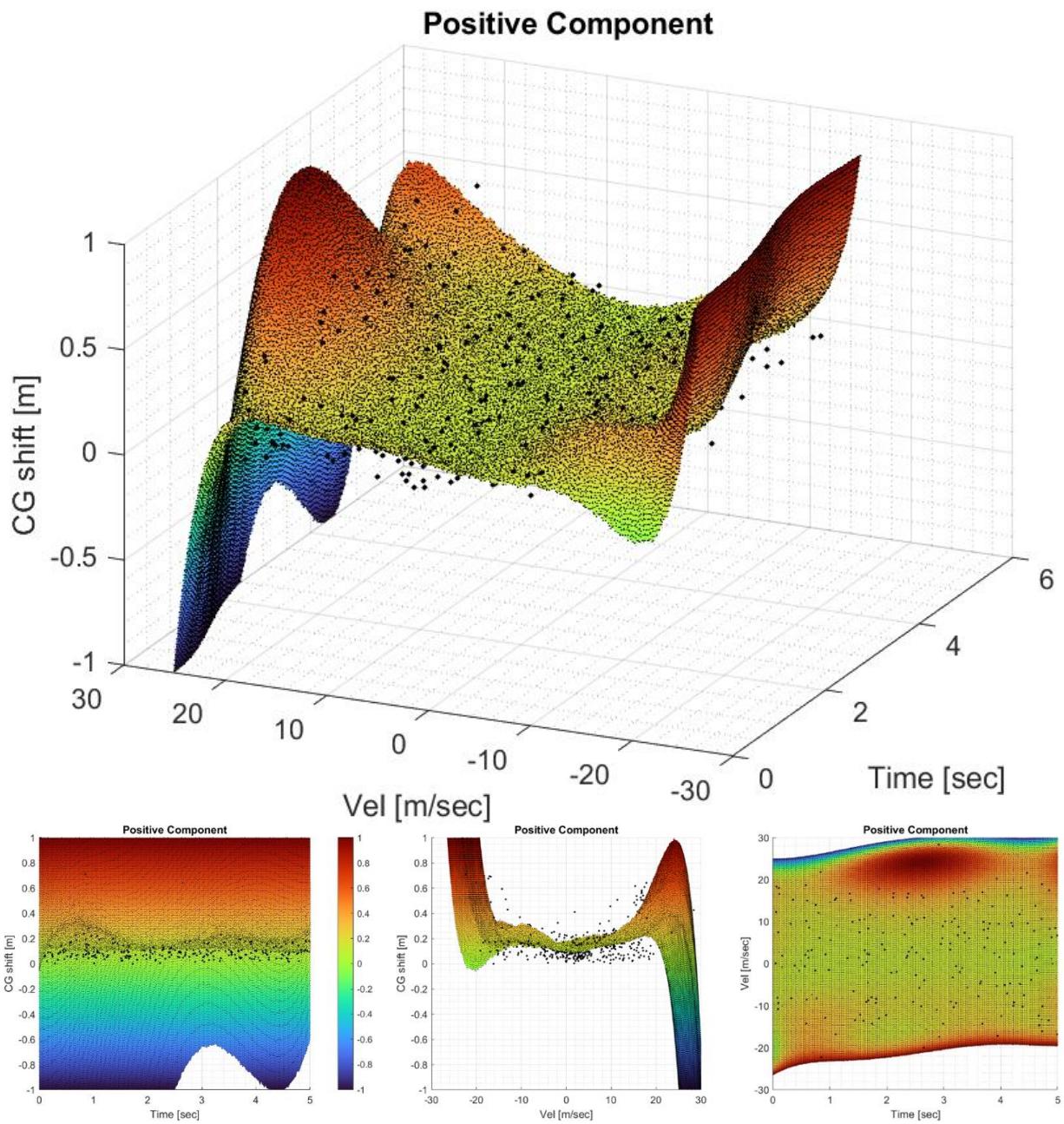


Figure 34: Positive CG Surface Response

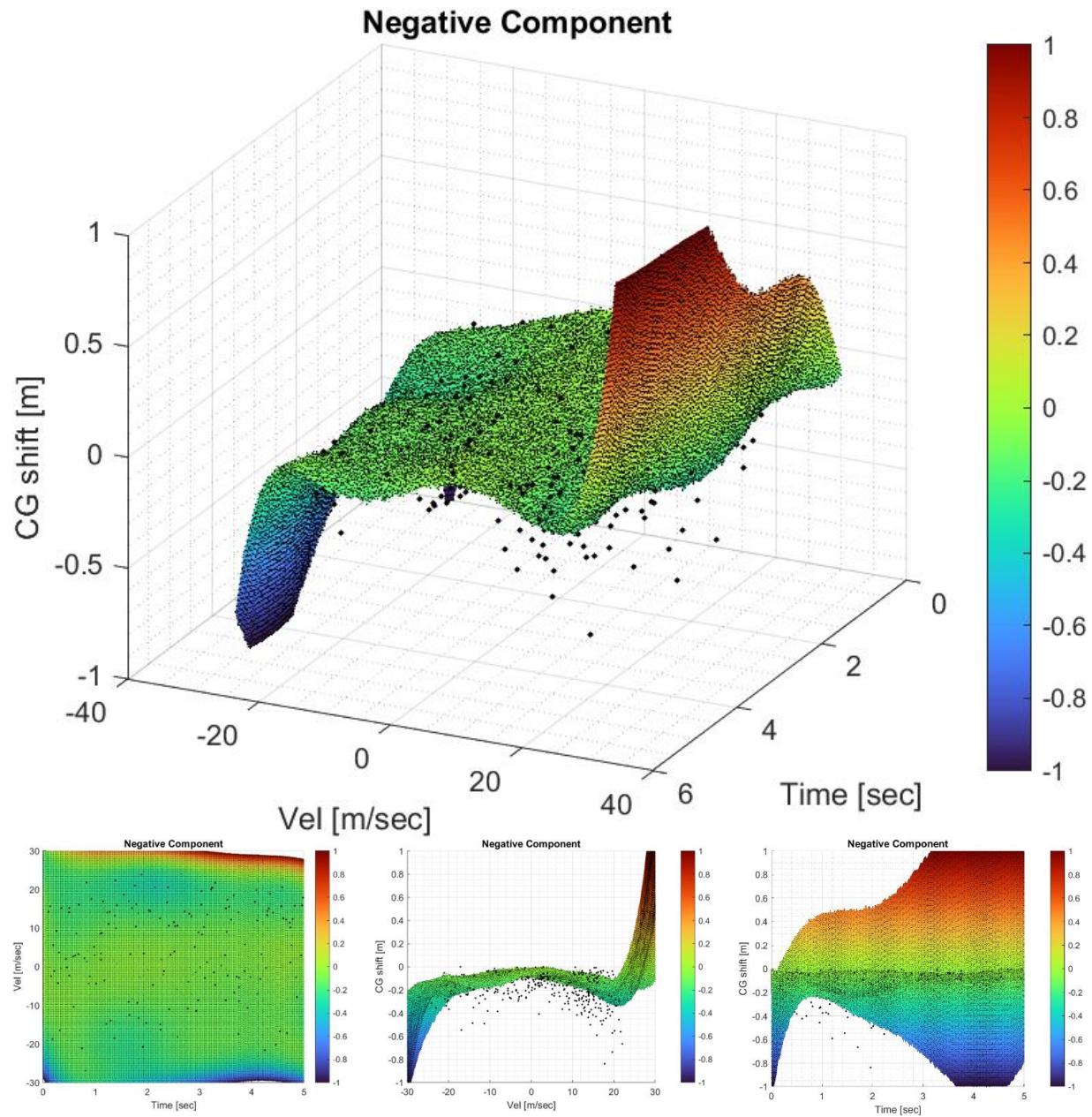


Figure 35: Negative CG Surface Response

These surface responses showed better interpolation through the selected points and regions of interest. Although this might have been the case, the maps were somewhat impractical in terms of interpolating a response based on a given set of coordinates in the design space. That is because these maps overlap each other, meaning there would be two responses (positive and negative CG shift values) at any given pair of time and velocity. The overlap is illustrated in the figure below. Therefore, a different approach must be used in order to use the Kriging as a surrogate model.

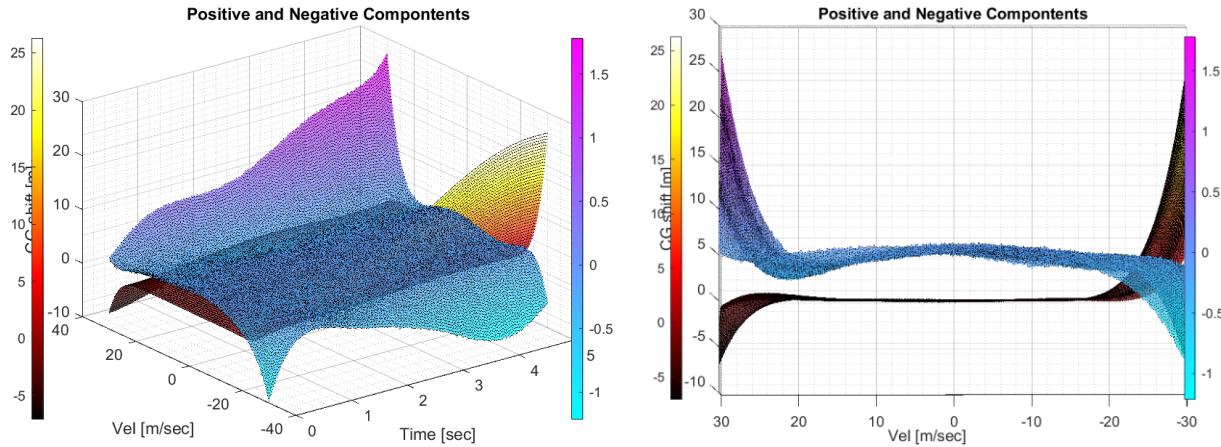


Figure 36: Overlap of Positive and Negative CG Surface Responses

One-Dimensional Kriging

There are multiple methods for determining the trend in single-input-single-output interpolation. These methods range from simple least-squares fitting in linear regression to more advanced approaches, such as kernels and other machine learning techniques. The *dacefit.m* function can be utilized to perform this interpolation.

The complexity of extracting oscillatory responses of the CG shift from the surface responses in the previous section routes the interpolation to a different approach. It opens the availability of a one-dimensional interpolation, or a one-dimensional Kriging. The interpolation is based on a single input, which is the excitation velocity, and a single output, which is CG shift, but this must be applied in consideration of the second direction, time. The *dacefit.m* function can be used to interpolate these local points at a given time step, called a state for this approach.

The one-dimensional Kriging is applied at every state. For simplicity only the first 1 second, which is a total of 201 states, is used as opposed to the full 5 seconds which would result in 1005 states. The oscillatory responses of the CG shift are captured within this time frame, and this 1-second window is also chosen arbitrarily based on available computational power. The method creates an interpolation for each state. In a 3D view, this creates a series of parallel planes at every state as shown below.

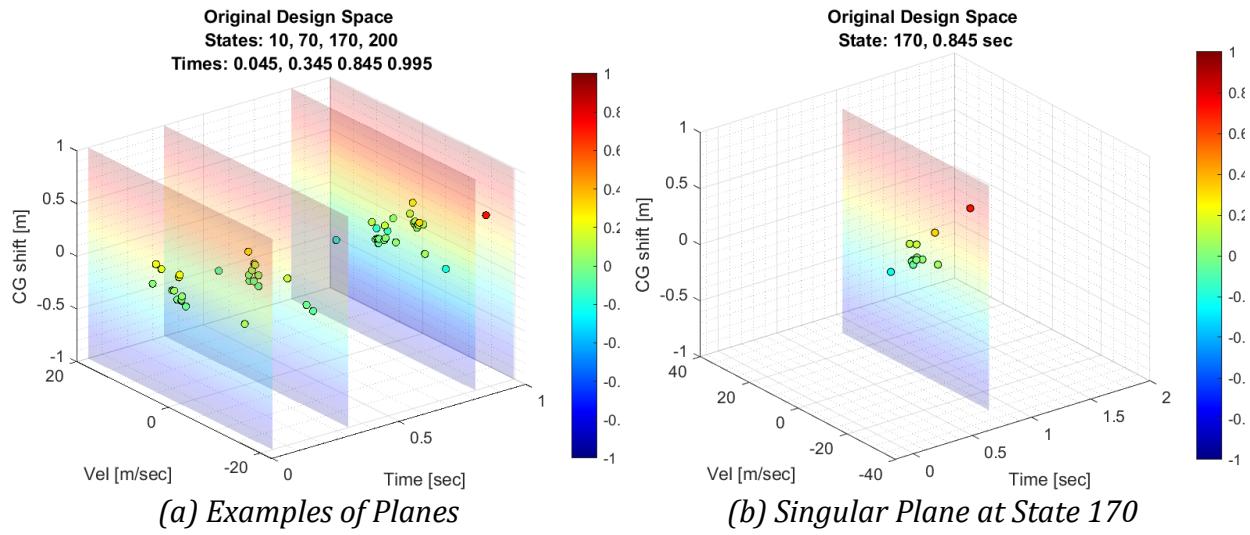


Figure 37: Examples of Parallel Planes Used for Interpolation

For this one-dimensional Kriging, two sets of interpolation maps were calculated. 1) Map with 16 cases 2) Map with 13 cases. In the latter and considering 201 states, there were a total of 2613 available seed points as shown below.

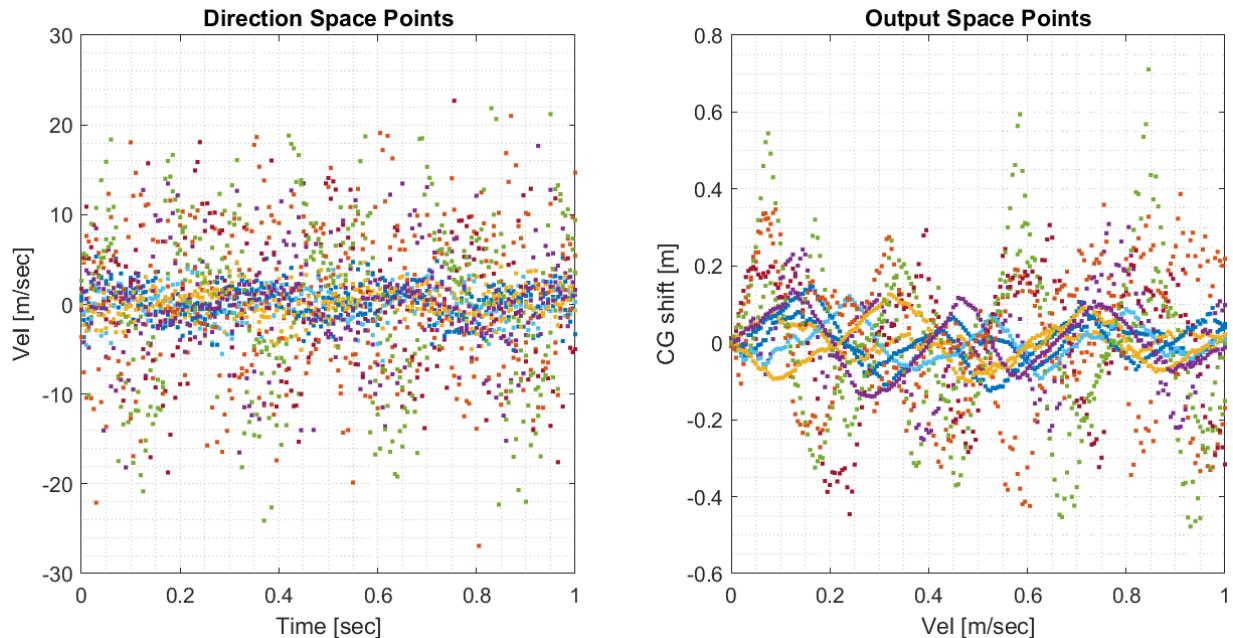


Figure 38: 13-Cases Map Seed Points

Based on all available seed data points, a cloud and an experimental variogram were calculated. These variograms are for visual reference as the interpolation happens at each state rather than using all the points available simultaneously. Each state has a different experimental variogram, which would be a total of 201. These experimental variograms can help determine the type of correlation model best suited for the available data at that state.

For simplicity, the same correlation model is used in all 201 states, but this correlation model could be a variable if the scope of the project was expanded.

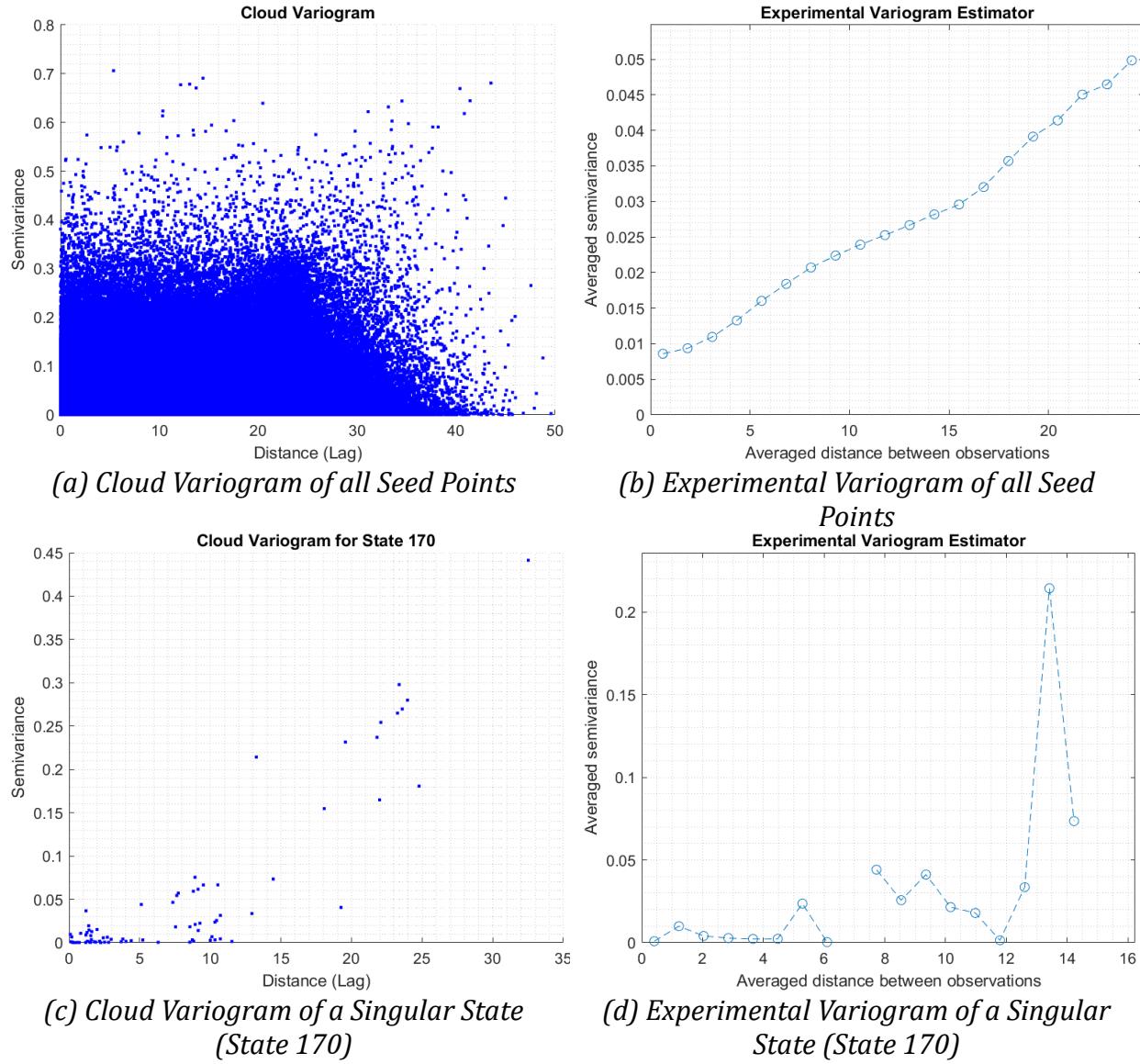


Figure 39: Cloud and Experimental Variograms Comparison

The experimental variograms at each state were different from one another, and their trends were not of typical shapes. This already was alarming as it suggested the interpolation may struggle to fit a correlation properly to the seed points given at each state. At each state, State 170 for example, the Kriging interpolation was applied based on the available points at that state, which originated from a given case (i.e. if n cases are used as seed values, then there are n points for the interpolation at that state). In this sense, there are 13 points for each state.

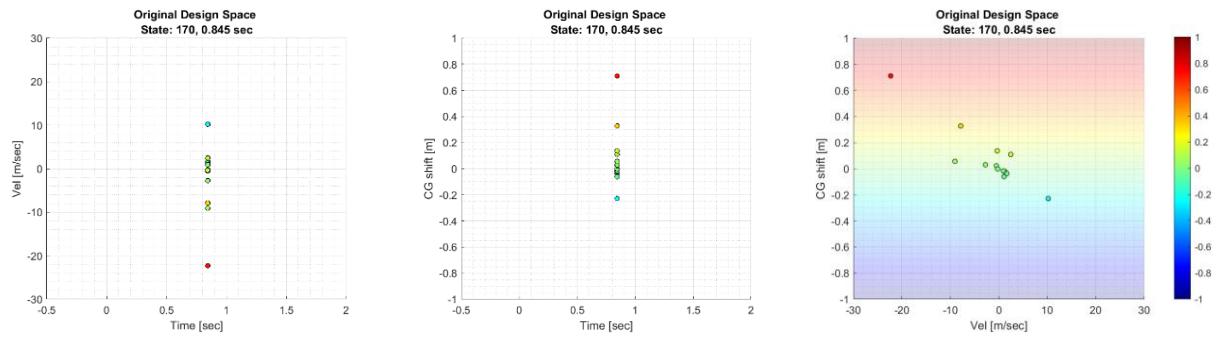


Figure 40: Example of Seed Points at A Singular State (State 170 for example)

The Kriging interpolator would then find the best fit trend based on the 13 points and a correlation model to formulate a line as shown below.

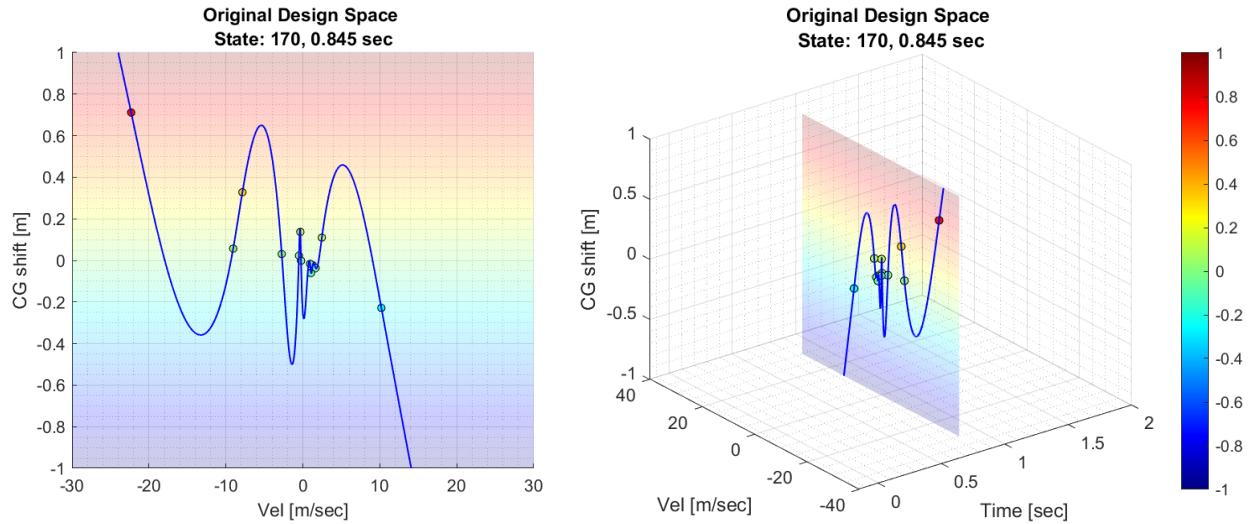


Figure 41: Best Fit Line based on a Cubic Spline Correlation Model

Different correlation models can be used to find this best fit line as shown below using State 170 as an example.

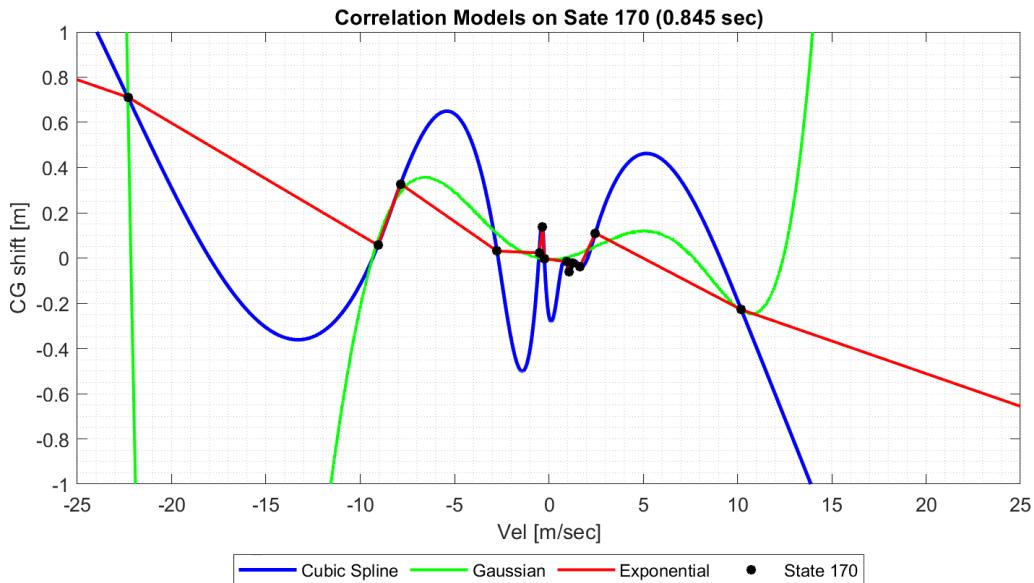


Figure 42: Different Correlation Models at State 170

The Gaussian and Cubic Spline correlation models are emphasized and their interpolation line for each 201 states is in the Appendix.

The last step is data recovery of each state. This is done by finding the nearest point equivalent to the given velocity.

1D Kriging Results

Using the same method described above, three sets of 1D Kriging maps were developed.

Map 1: A Kriging interpolator using all 16 cases as a test scenario to confirm the algorithm.

Map 2: A Kriging interpolator using 13 cases and a Gaussian correlation model.

Map 3: A Kriging interpolator using 13 cases and a Cubic Spline correlation model.

The first row of the figure below is from Map1, the second row is from Map2, and 3rd row is from Map3. Each column represents each testing case.

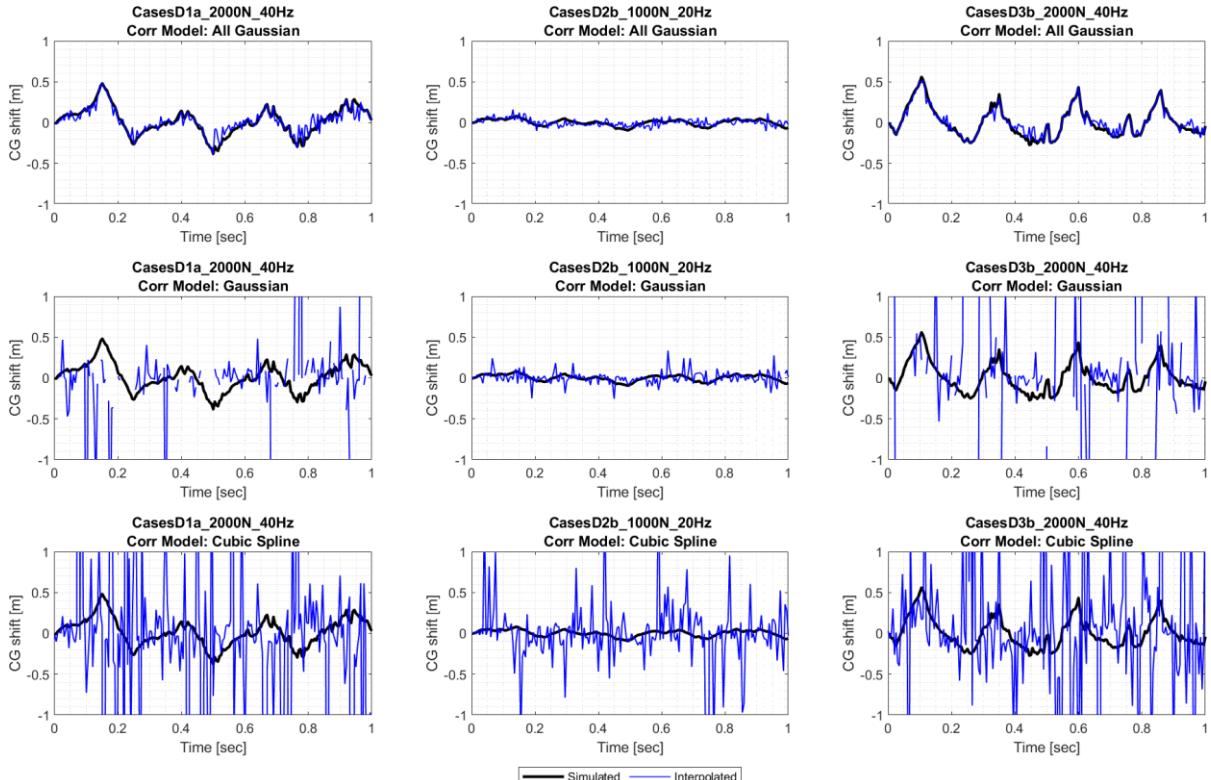


Figure 43: Results of Kriging Interpolators

By examining the results above, it is evident that the maps with 13 cases, Map1 and Map2, are not accurate. Map1 results are the best estimations. The interpolations were accurate in all three cases that used Map1. A normal fit was performed, and the residuals were analyzed. Case D2b_1000N_20Hz is the best estimate of where the residuals have Kurtosis of 3 resembling a normal distribution. This can clearly be seen in the probability density functions.

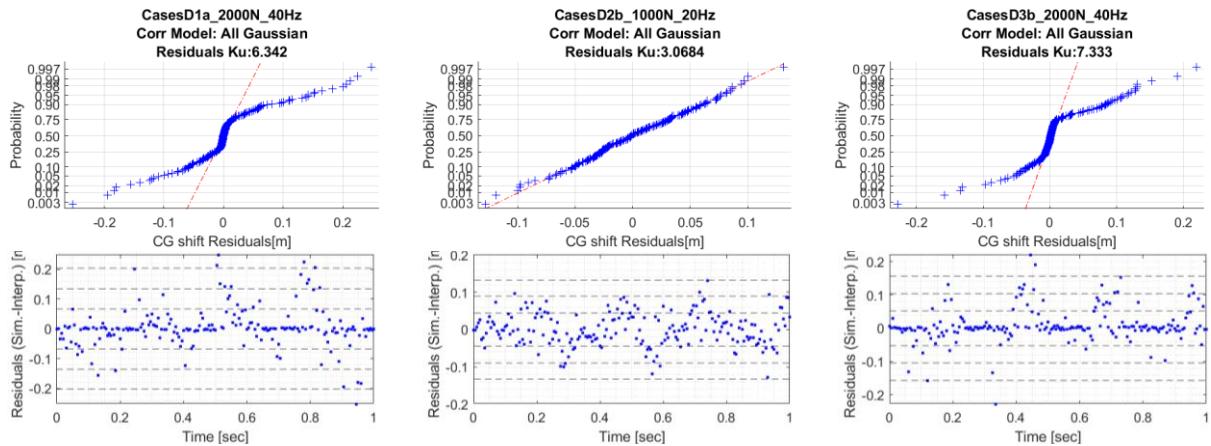


Figure 44: D2b_1000N_20Hz Normal PDF and Residuals

The 1st and 3rd column/cases have residuals above the 3-sigma thresholds, but when looking at the CG shift trends they are very close to the actual values.

The reason Map1 has better performance than Map2 and Map3, is because the interpolator in Map1 has 3 more seed points than the other maps. The Kriging interpolator is an unbiased estimator, meaning that the estimation at the same sampled and predicted location should be the same. In the case of Map1, the three testing cases were also used as seed points. Additionally, the Gaussian correlation introduces noise in the trend between the seed points and this can be seen in the trend. Similarly, Map1 also shows that the Kriging can interpolate values within the velocity gaps. This confirms that the interpolator is working as it should.

Discussion and Conclusion

Each bearing defect geometry exhibited a unique signature when driven by different loads and rotational frequencies. The responses were divided into two regions: low frequency below 500 Hz, and high frequency within 12,000 and 22,000 Hz.

In the low frequency range, the BPFO and its harmonics were identified. These responses excite the RP-1 inside the cubic tank the most. These excitations were used to analyze the effect of the RP-1 slosh dynamics inside the tank.

By applying various filtering methods to the high-frequency region, the characteristic pulses of the contact dynamics between the ball bearing and outer race were extracted. When these pulses were transformed to the frequency domain in their periodograms, the signal was modulated which enveloped other signals embedded in this frequency region as expected.

Upon comparing the different Kriging surface responses in the fabric (surrogate models), the oscillatory behavior of the CG shift made the Kriging interpolator impractical as a surface response. There could be advanced techniques outside of the scope of this research project to identify the proper Kriging map to use and interpret.

When looking at the one-dimensional Kriging interpolator, the “surface responses” (best fitted lines) in the “thread” -as opposed to fabric- were more practical. This best-fit line was heavily dependent on the correlation model used by the interpolator. The cubic spline passed through all points but at the same time produced more Mean-Square Errors than the Gaussian correlation model.

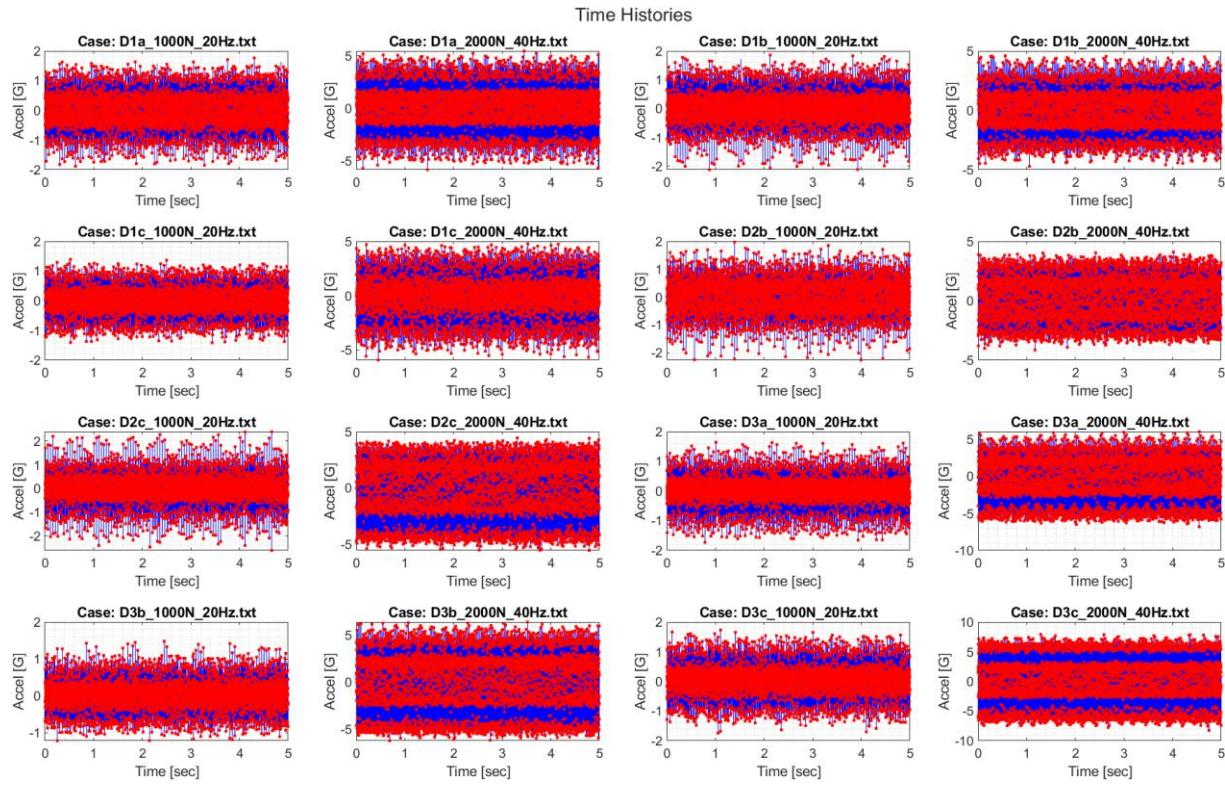
In summary, 13 cases are not sufficient to create a proper Kriging interpolator. When given a few seed points, Kriging can perform better in predicting values within the time step of the seed points, especially when using finer step size or grid points. This can be beneficial for saving computational time. For example, if a simulation is run with a coarser time step due to data availability or if a test is conducted over a longer duration with a lower sampling rate, the 1D Kriging interpolator can fill in the gaps between the seed points (observed points) and enhance the results.

References

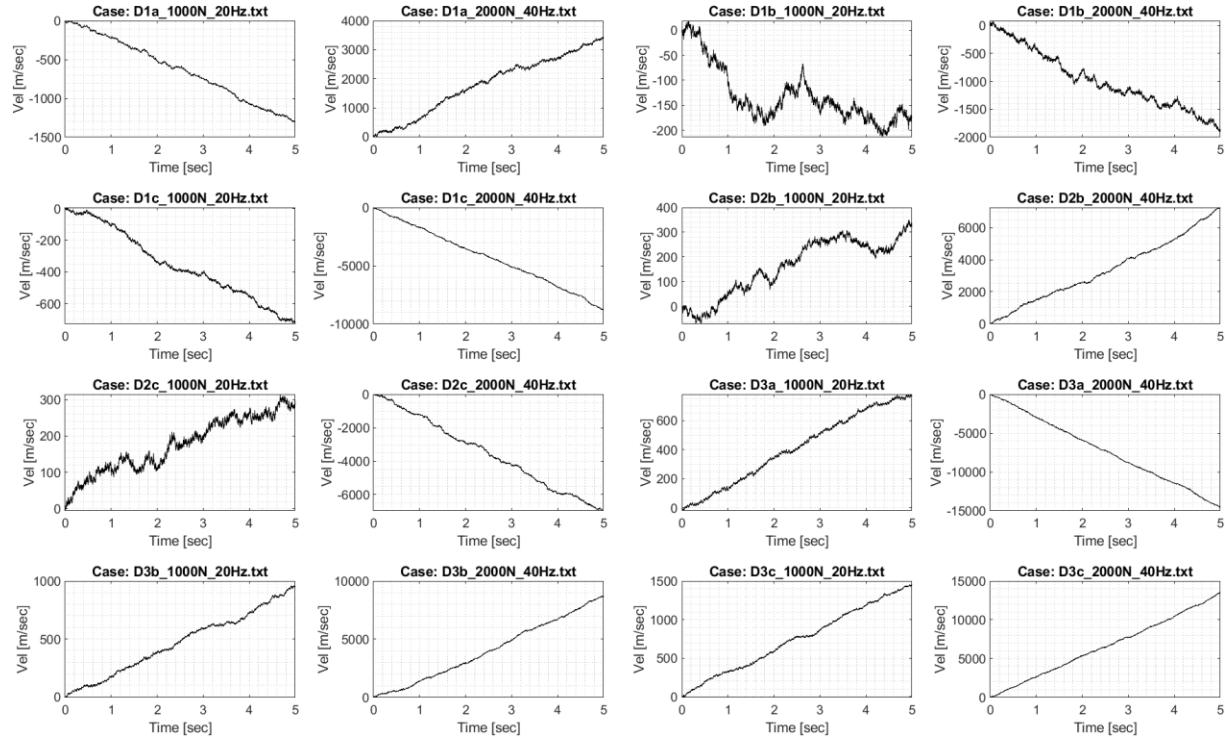
- 1) <https://data.mendeley.com/datasets/8wdzm5gwng/1>
- 2) <https://www.sciencedirect.com/science/article/abs/pii/S0888327022008512?via%3Dihub>
- 3) <https://www.sciencedirect.com/science/article/abs/pii/S0888327007002725>
- 4) <https://lsdyna.ansys.com/manuals/>
- 5) https://www.dynaexamples.com/icfd/basics-examples/global_imposed
- 6) <https://www.youtube.com/watch?v=5MNyLM0N6l0>
- 7) [DACE - A Matlab Kriging Toolbox - Welcome to DTU Research Database](#)
- 8) All data and code can be found in this [repository](#).

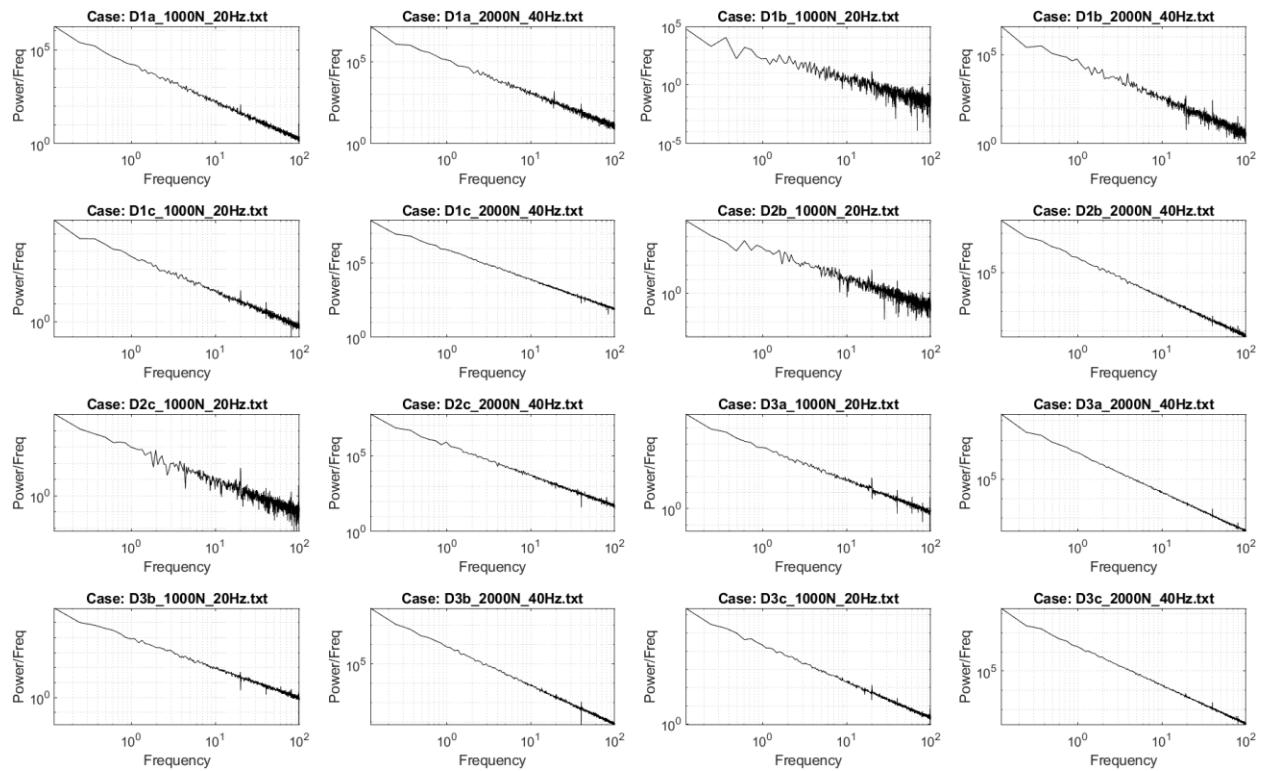
Appendix

Downsampling:

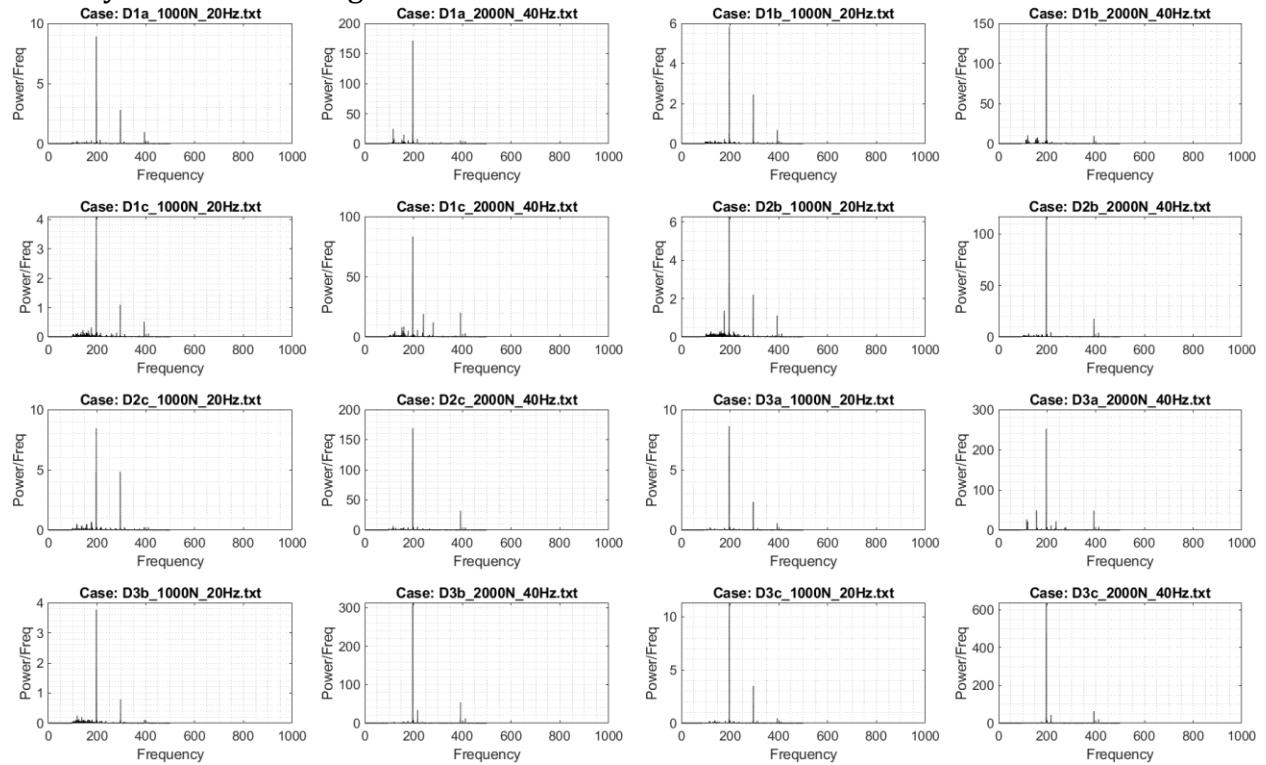


Velocities:

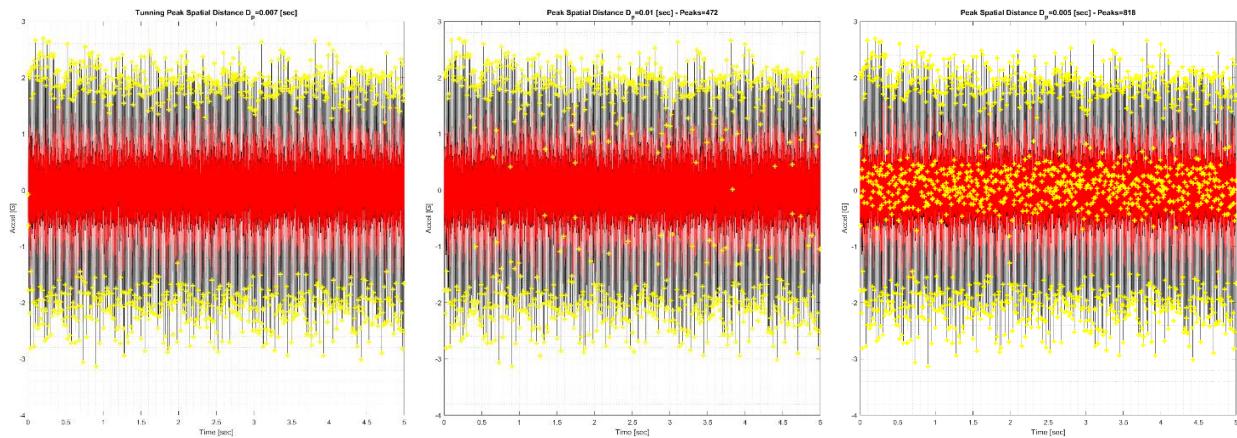




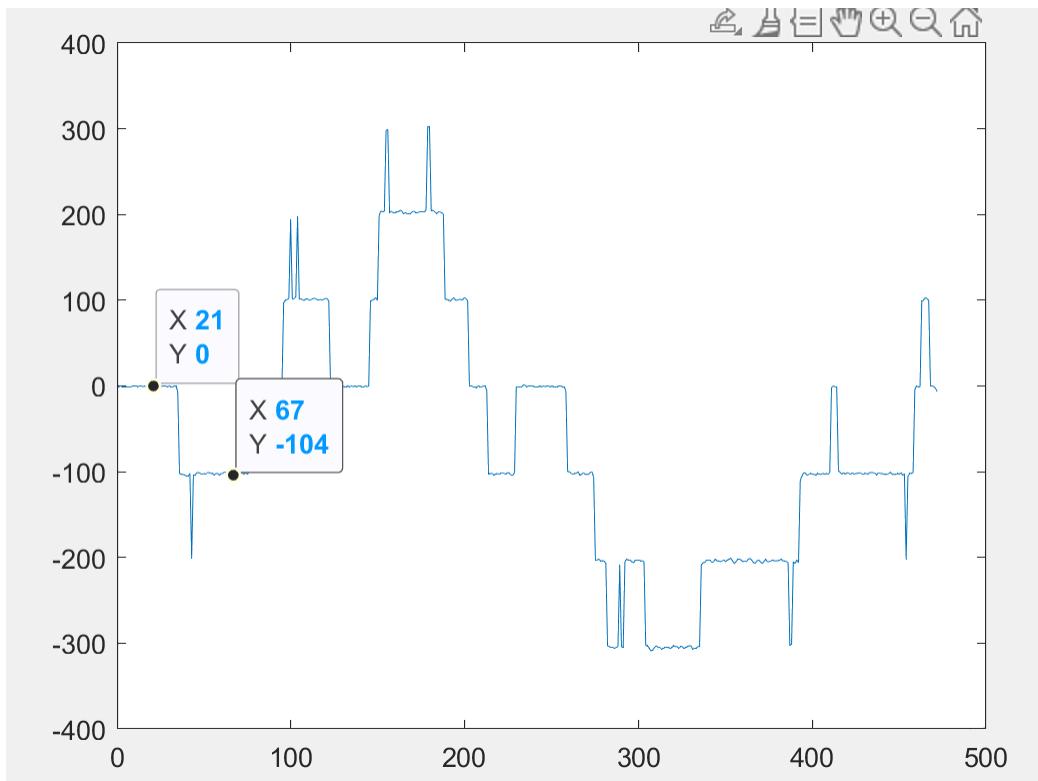
Velocity Filtered Periodogram:



Tuning Parameter *dist_t*:

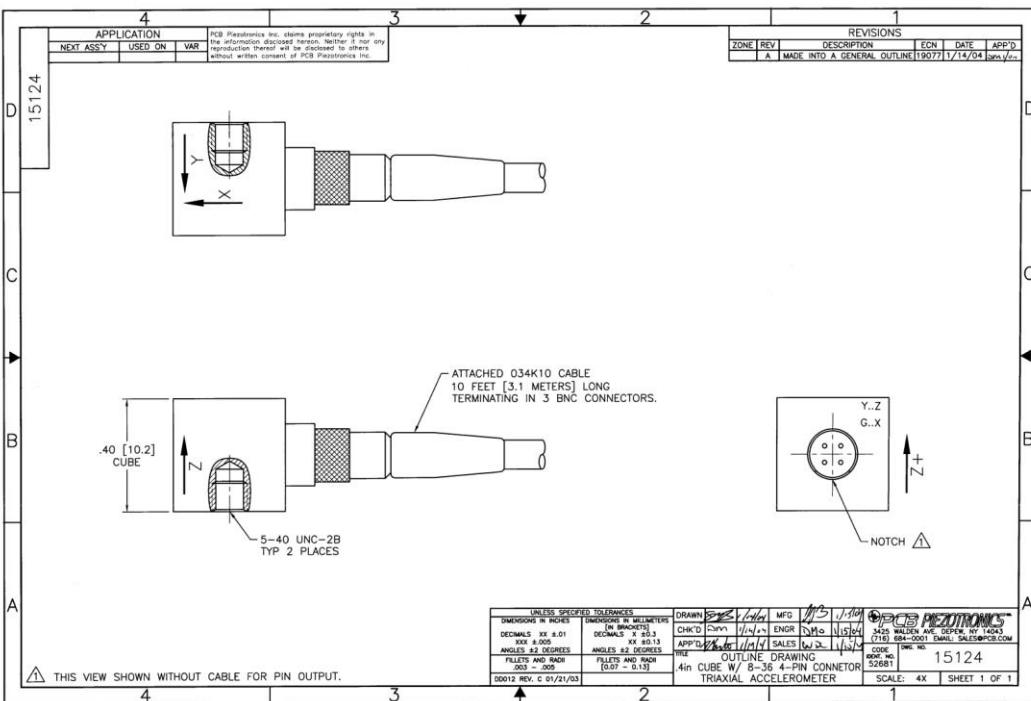


Tuning to nearest point.

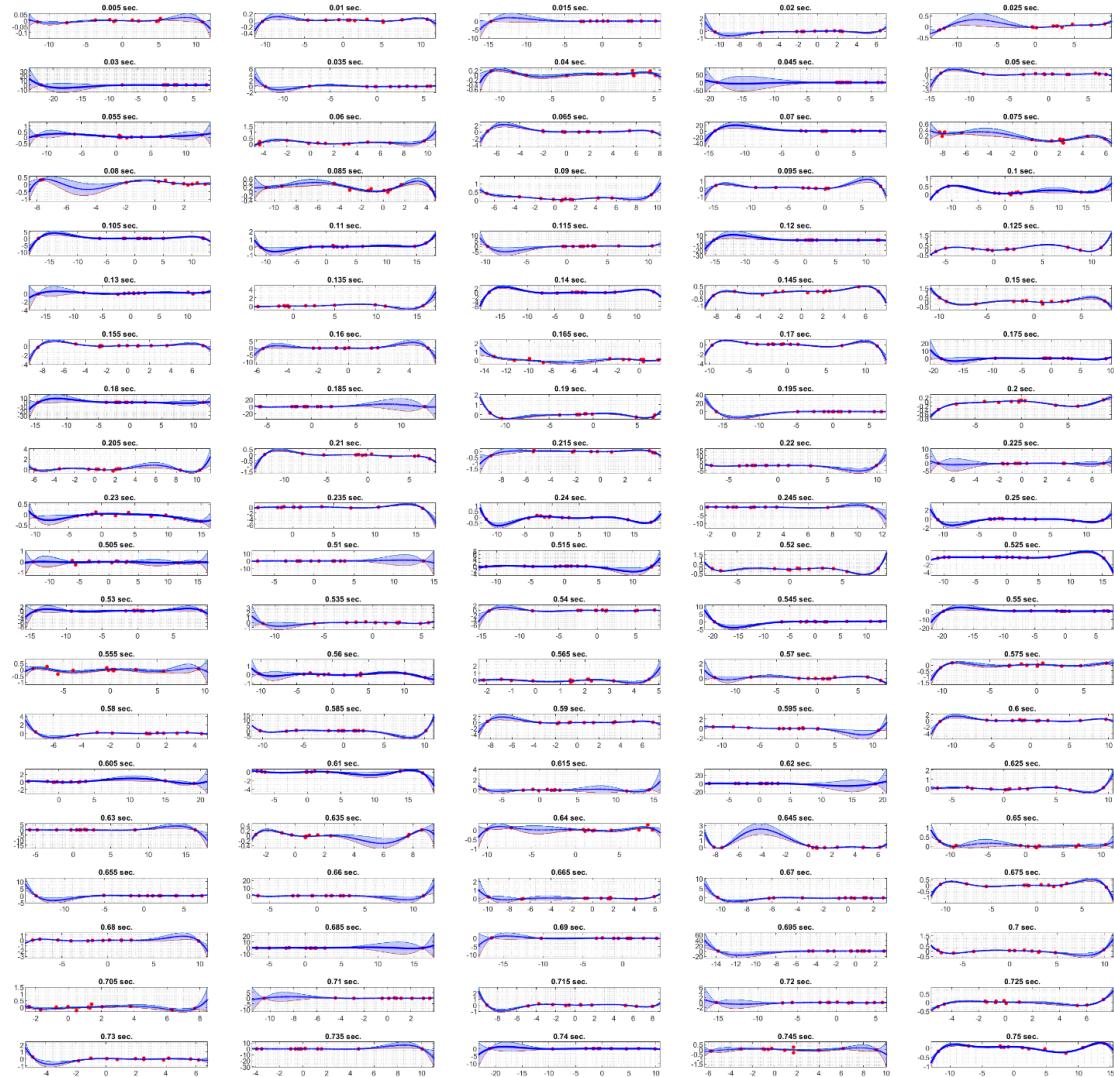


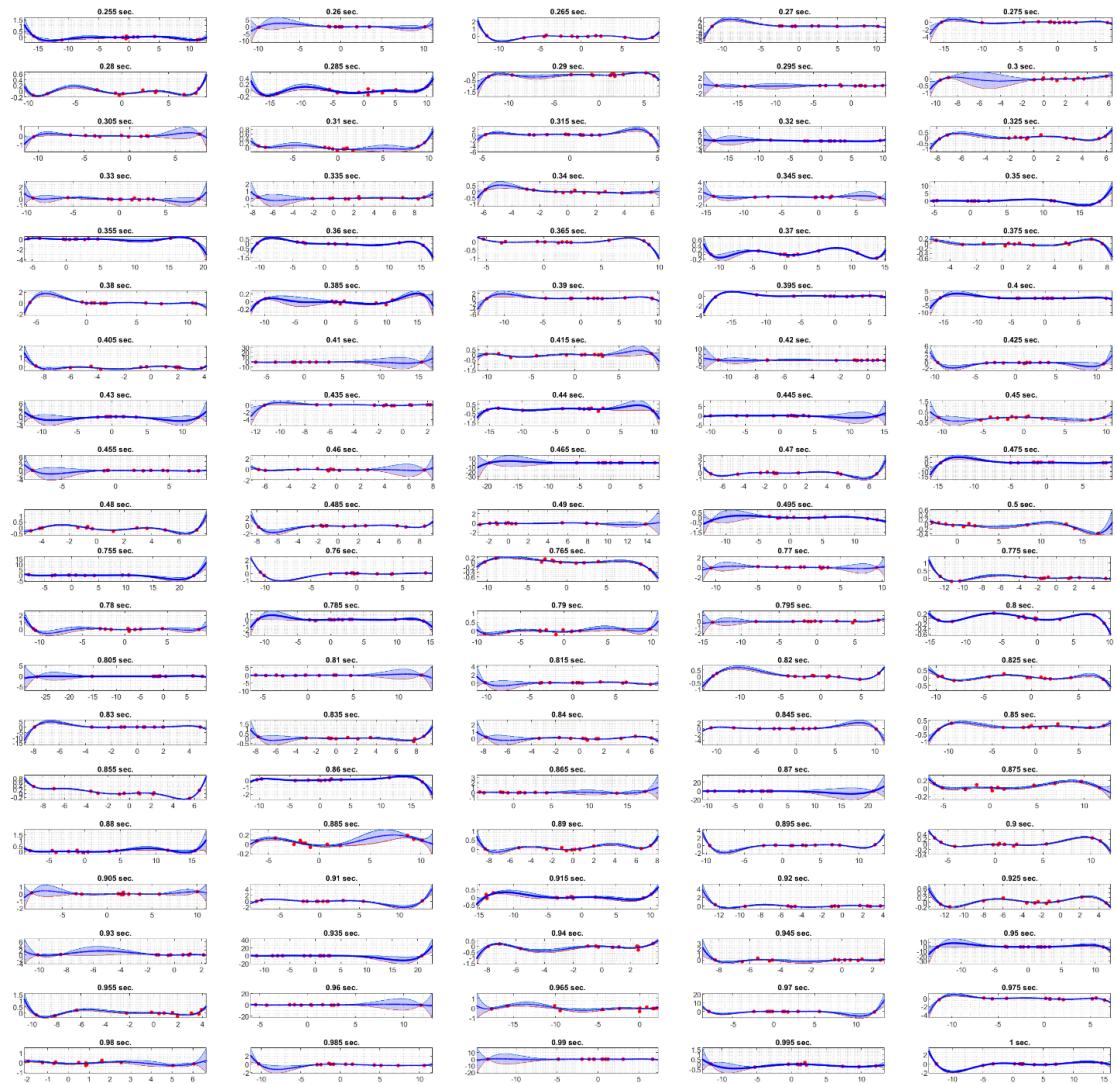
PCB 356B21 Specifications: <https://www.pcb.com/products?m=356b21>

Model Number 356B21	TRIAXIAL ICP® ACCELEROMETER			Revision: H ECN #: 42197				
Performance	ENGLISH SI							
Sensitivity(\pm 10 %)	10 mV/g $\pm 500 \text{ g pk}$ Frequency Range($\pm 5\%$)(y or z axis)							
Frequency Range($\pm 5\%$)(x axis)	2 to 1000 Hz 2 to 7000 Hz 2 to 7000 Hz							
Resonant Frequency	$\geq 55 \text{ kHz}$							
Broadband Resolution(1 to 10,000 Hz)	0.004 g rms $\leq 1\%$ $\leq 5\%$							
Non-Linearity	$\leq 1\%$							
Transverse Sensitivity	$\leq 5\%$							
Environmental								
Overload Limit(Shock)	$\pm 10,000 \text{ g pk}$							
Temperature Range(Operating)	-65 to +250 °F See Graph							
Temperature Response	-54 to +163 °C See Graph							
Electrical								
Excitation Voltage	18 to 30 VDC							
Constant Current Excitation	2 to 20 mA							
Output Impedance	$\leq 200 \text{ Ohm}$							
Output Bias Voltage	7 to 12 VDC							
Discharge Time Constant	0.3 to 1.0 sec							
Setting Time(within 10% of bias)	<3 sec							
Spectral Noise(1 Hz)	$1000 \mu\text{g}/\sqrt{\text{Hz}}$							
Spectral Noise(10 Hz)	$300 \mu\text{g}/\sqrt{\text{Hz}}$							
Spectral Noise(100 Hz)	$100 \mu\text{g}/\sqrt{\text{Hz}}$							
Spectral Noise(1 kHz)	$50 \mu\text{g}/\sqrt{\text{Hz}}$							
Physical								
Sensing Element	Ceramic							
Sensing Geometry	Shear							
Housing Material	Titanium							
Sealing	Hermetic							
Size (Height x Length x Width)	0.4 in x 0.4 in x 0.4 in							
Weight	0.14 oz							
Electrical Connector	8-36 4-Pin							
Electrical Connection Position	Side							
Mounting Thread	5-40 Female							
Typical Sensitivity Deviation vs Temperature								
<p>All specifications are at room temperature unless otherwise specified. In the interest of constant product improvement, we reserve the right to change specifications without notice. ICP® is a registered trademark of PCB Group, Inc.</p>								
NOTES: [1] Typical. [2] 250° F to 325° F data valid with HT option only. [3] Zero-based, least-squares, straight line method. [4] See PCB Declaration of Conformance PS023 for details.								
SUPPLIED ACCESSORIES: Model 034K10 Cable 10FT Mini 4 Pin To (3) BNC (1) Model 080A Adhesive Mounting Base (1) Model 080A109 Petro Wax (1) Model 081A27 Mounting Stud (5-40 to 5-40) (1) Model 081A490 Mounting Stud, 10-32 to 5-40 (1) Model ACS-11 NIST traceable triaxial amplitude response, 10 Hz to upper 5% frequency. (1) Model M081A27 Metric mounting stud, 5-40 to M3 x 0.50 long (1)								
Entered: AP	Engineer: JJB	Sales: WDC	Approved: JJB	Spec Number:				
Date: 11/8/2013	Date: 11/8/2013	Date: 11/8/2013	Date: 11/8/2013	15127				
PCB PIEZOTRONICS 3425 Walden Avenue, Depew, NY 14204 Phone: 716-684-0001 Fax: 716-684-0987 E-Mail: info@pcb.com								

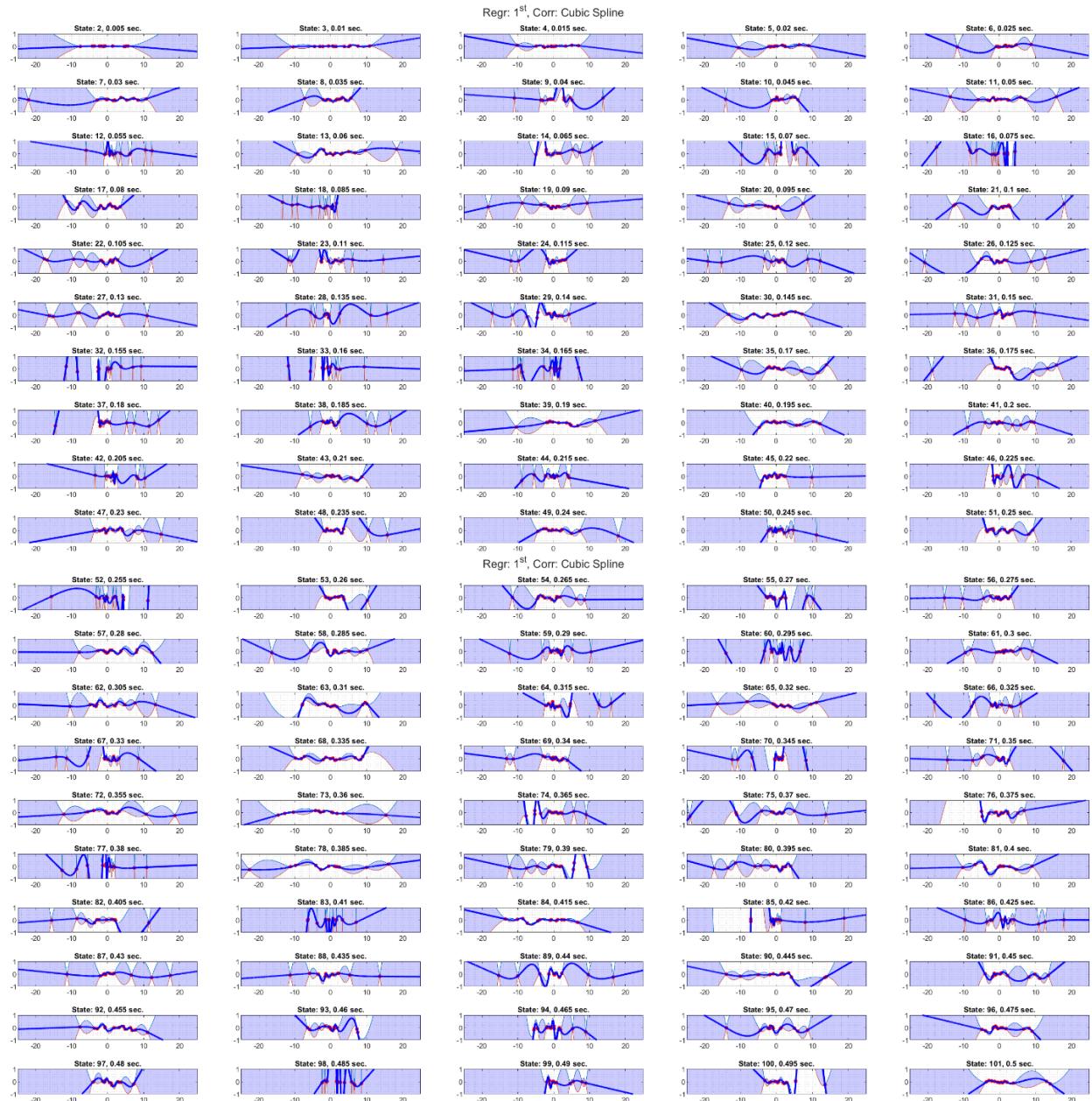


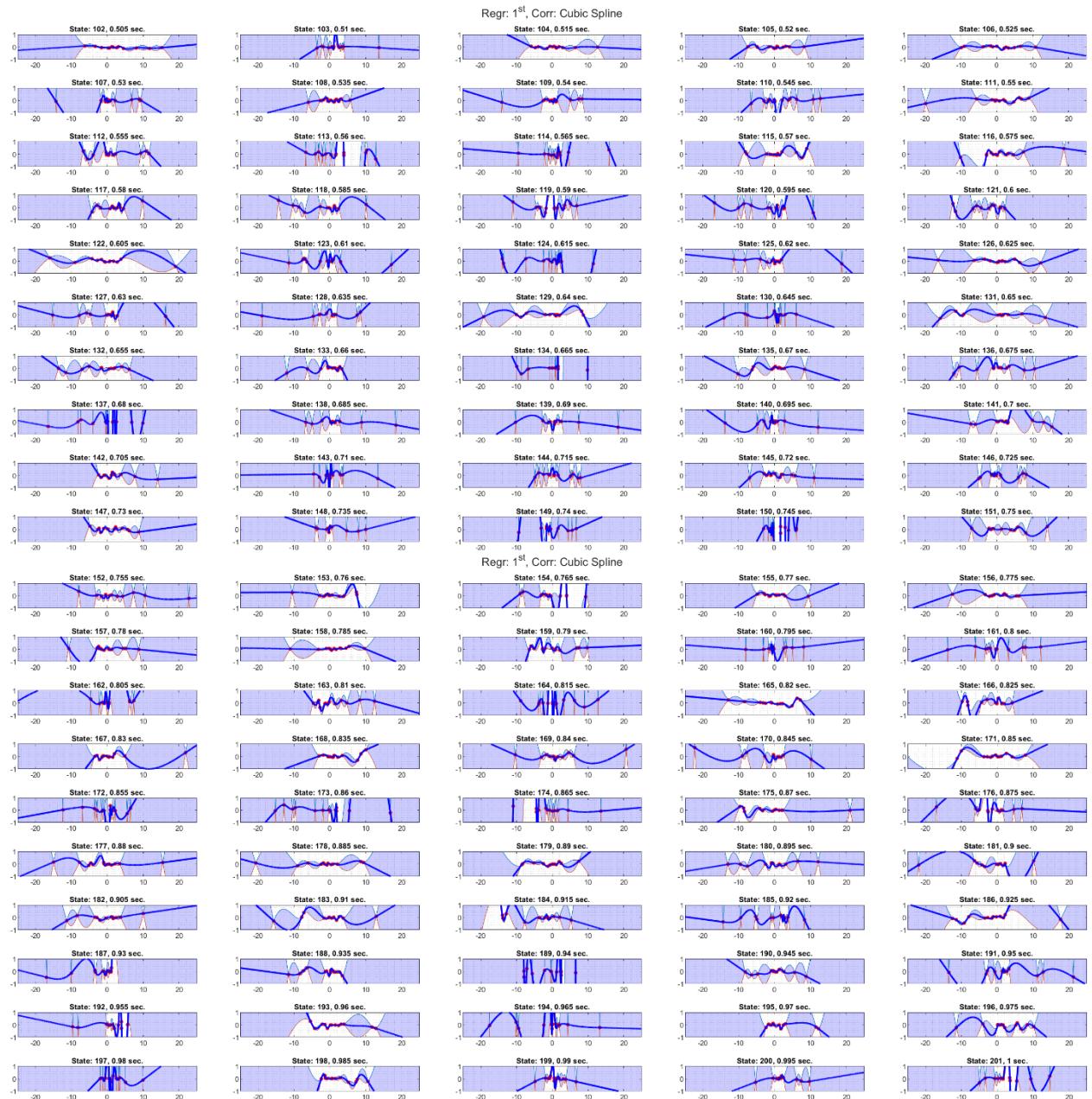
State Interpolation Gaussian Correlation (State 1 not plotted for simplicity)





State Interpolation Cubic Spline Correlation (State 1 not plotted for simplicity)





Appendix

