Investigating The Relationship Between Wind Speed and LIGO Strain Data Noise

Xuanming (Rudy) Zhang

Mentor: Prof. Eric Myers

Gravitational Waves: Section 1

Pioneer Academics

(Dated: August 24, 2020)

Abstract

The Laser Interferometer Gravitational-wave Observatories in Hanford and Livingston are susceptible to reductions of sensitivity due to vibrations from the ground or the instruments' walls. Sources of vibrations include seismic activity, strong wind, and thunderstorms, to name a few. This paper describes the use of numerous matched pair data samples from both LIGO locations to determine wind's effect on the instruments' sensitivity. While wind has a moderate effect in elevating noise in lower frequencies such as 7Hz and 13Hz, its impacts on higher frequencies, such as 500Hz or 1000Hz, are much weaker.

I. Introduction

Gravitational waves are ripples in space-time that cause the space it passes through to become stretched in one dimension and compressed in another dimension. Because of this property of gravitational waves, and the fact that gravitational waves signals are so small when they reach earth, the Laser Interferometer Gravitational-wave Observatories (LIGO) were designed and operated to detect such small signals. Taking advantage of the Michealson interferometer inside to measure the phase shift of two laser beams shined at two perpendicular directions, humans were able to make their first direct detection of gravitational waves in 2015. Nevertheless, various sources of noise, such as thermal noise and seismic noise, add difficulty to detections as they interfere with gravitational wave signals.

Perhaps the most common form of noise for LIGO detection is the weather, as different weather features such as wind and rain cause vibrations on the ground and walls of the observatory. Weather is composed of different variables, such as temperature, wind speed, and precipitation, and each of these factors affects the detection of gravitational waves in their own way. Because of this complexity of components, it is difficult to determine the overall effect of the local climates in Livingston and Hanford on creating noise on the latest installation of LIGO that has available data- advanced LIGO during observational run 2 (O2). The purpose of this research is to isolate some of these variables, specifically wind speed, and use a controlled observational study to compare the noise created by different levels of wind on advanced LIGO strain data. The results of this study will be helpful for gravitational wave scientists aiming to reduce noise caused by the surrounding weather. As climate scientists have developed a sophisticated system of forecasting weather, the findings from this research may lay foundations for predictions of advanced LIGO noise levels in the future based on weather forecasts and also help to develop solutions to complement such sources of noise.

II. Background

Gravitational waves are ripples of spacetime generated by huge masses accelerating with respect to each other in space that results in quadropole distortions in space-time[1]. The reason they are called "waves" is that such fluctuations in spacetime propogate from its source in all directions at the speed of light[1], similar to waves in a water pond after a stone is dropped inside. These ripples in spacetime mainly cause mass experiencing them to accelerate, as the space around them is squeezed in one dimension while stretched in another dimension[2]. Albert Einstein predicted the existence of gravitational waves in 1916 in his general theory of relativity, but because of the theoritical nature of his prediction, even he himself was not certain about their existance at the time [1]. Events that cause gravitational waves include collisions of two neutron stars, which are collapsed cores of large stars and thus very dense; or two black holes, massive bodies that attract all mass and even light. Other examples are asymmetric spinning neutron stars and asymmetric explosions of stars[2]. As a result, when scientists detect gravitational waves, they can perceive and further understand cosmic events that can be hard to detect through optical methods, like blackhole merger events, as blackholes absorbs all electromagnetic radiation.

While the causes and effects of gravitational waves are quite significant, they originate so far from earth that by the time they reach earth their energy has been reduced by thousands of times, making detection very challenging. To make detection possible, the LIGO was created and operated in Louisiana and Washington, and it succeeded in finding evidence of gravitational waves in 2015, ending scientists', and Einstein's, doubt of their existence [3]. As the name

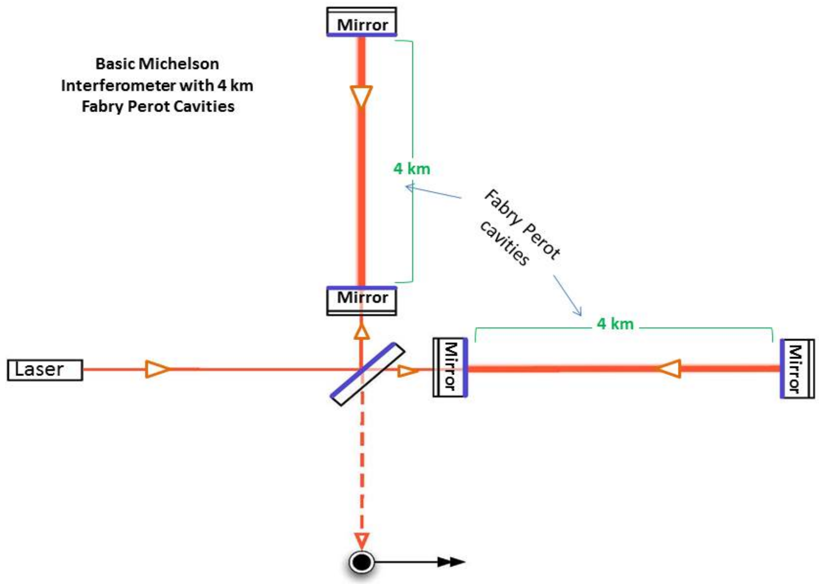


Figure 1.A brief diagram of LIGO's structure[4]

implies, LIGO uses a Michelson Interferometer to detect gravitational waves, which is shown in figure 1. Inside, the laser shines into a beam splitter that splits the light into two directions, which are reflected between two mirrors separated by 4 kilometers. After being reflected for hundreds of times, the light exits the mirror, passes through the beam splitter, and ends up on the photodetector at the bottom. Because gravitational waves stretch space in one dimension and compress it in another dimension, when the two beams merge at the photodetector, they would be out of phase since they travelled different distances, creating an interference pattern [5]. Scientists can use this pattern to calculate strain (h):

$$h = \frac{L}{\Delta L} \tag{1}$$

where L is the length of one arm of the interferometer, and ΔL is the change in that length. The strain data represent signals, both from gravitational waves and from other sources, detected by the interferometer, and researchers can transform these data to represent other information. For example, applying a windowing function reduces the amplitude of strain data at the beginning and end of a specific time segment. This allows the series of data within this time segment to be periodical and continuous, which is required for Fast Fourier Transforms (FFTs) to break down the strain time-series into its constituent frequencies, shown in figure 2b [6]. Additionally, taking the averages of the square of the FFT amplitudes of the data creates power spectral densities (PSD) for the time interval, which is shown in figure 2c; and taking the square root of the PSD creates the amplitude spectral density (ASD) shown in figure 2d, which is "an estimate of the strain-equivalent noise of the detectors versus frequency" for the time sampled[8].

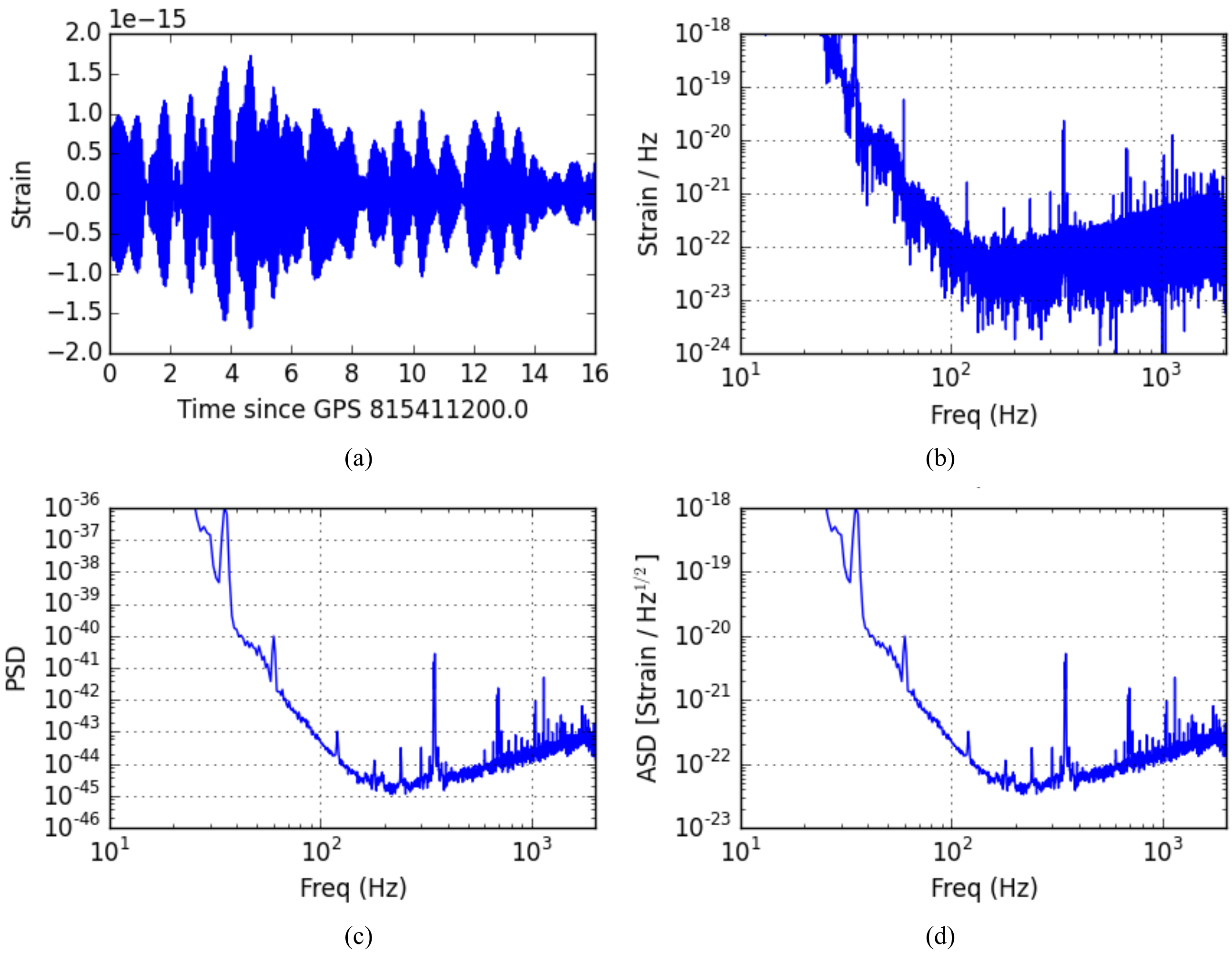


Figure 2. diagrams of strain time series (a), FFT (b), PSD (c), ASD (d) [7]. The diagram demonstrates that ASD yields lower variances than FFT on the same order of magnitude, making it a much better estimate for non-gravitational wave strain signal levels.

The construction of Initial LIGO began in 1995. While designed to search for gravitational waves, Initial LIGO also tested the concepts and instruments that would help developing more sensitive technologies for LIGO. It set the foundation for more advanced detectors in its 9 years of operation, and in 2007, Initial LIGO received its first major upgrade and became enhanced LIGO, having a increase its sensitivity during observations in 2009. The detectors were replaced in 2010 by Advanced LIGO, and when completed in 2015, Advanced LIGO went to 10 times the sensitivity of Initial LIGO, ultimately leading to its discovery of gravitational waves in September 2015 from two colliding black holes [9]. LIGO observations were made in runs, and between each runs maintenance and upgrades were made, making the conditions of the equipment slightly different between each run. After its first observing run (O1) from September 2015 to January 2016, advanced LIGO made its second observing run (O2) from November 2016 to August 2017, whose data are the latest LIGO data available yet. Although LIGO has started its third Observing run (O3) since April 1, 2019, and so far has yielded 56 gravitational-wave detections, due to COVID-19, the O3 run has been suspended since March 27, 2020, resulting its data to be delayed for public use [10].

The sensitivity of a LIGO detector is often reduced due to a variety of noise sources. Seismic noise, thermal noise, and electronic noise are some common interferences that are often studied [11]. As the whole LIGO apparatus is placed on the ground, these noises could shift the position of the apparatus, including the mirrors and photodetectors. This shift results in the lasers traveling for a different distance than they would if they were just affected by gravitational waves. Eventually, strain data will be shifted as well, as they are dependent on how the laser beams travel, resulting in the so-called “noise” in the strain data collected by LIGO [11].

III. Methods

The primary research method is a matched-pair observational study, in which each experimental sample is matched with a control sample, and they differ in only one variable. For the observational study's design, the independent variable is wind speed, and the dependent variable is the amount of noise in LIGO strain data, represented by ASD values. The higher the ASD value at a certain frequency, the higher the noise level will be at that frequency. Because this research did not include random assignment of treatments, in which the wind speeds would have been randomly assigned to LIGO data, this study is limited to being completely observational, not experimental and manipulative. First, historical hourly weather data were gathered from the National Centers for Environmental Information [12], part of the National Oceanic and Atmospheric Administration. Baton Rouge and Hammond's weather data were averaged to approximate Livingston's conditions. Both locations have historical weather data available throughout O2 (Observational Run 2), and because Baton Rouge is located to the west of Livingston, while Hammond is located to the east, both locations' data together can be an accurate approximation. For Hanford, the data for Richmond, WA, was used, because Richmond is very close to Hanford and it was reasonable that the weather in both places is the same. The weather data spanned from the beginning to the end of the second observation run, and they were in the form of hourly records. To avoid confounding other weather variables with wind speed that could potentially be a noise source for LIGO, the hourly data were filtered so that the time points left have no rainfall, no thunder, no wind gusts, and have "clear" for sky conditions. Next, the filtered data went through another filter to eliminate the hours in samples when LIGO in their corresponding locations did not produce strain data. Finally, for every data entry left, the closest time point within these data that has zero wind speed was found and paired with it. This created the pools of thousands of matched-pairs, one as an experimental data that varies in wind speed from pair to pair, and the other as a control, zero wind speed data, for Hanford and Livingston separately. A simple random sample of the size of 30 was drawn from the data pools of Hanford and Livingston separately, which determines 30 pairs of time points during O2 as well as their corresponding wind speed present. After sampling, strain data for both timepoints in each matched pair sample were downloaded from the GW Open Science Center Strain Data page [13].

To read in all the obtained strain data and plot the noise level, the "lotsofplot" python script, along with "readligo" module, were downloaded from the GW Open Science Center "Lots of Plots Tutorial" page [14]. The "lotsofplots" script was then modified so that it only graphed the ASD, excluding graphs of strain time series, FFT, PSD, ASD, and spectrums. Then, variables for the file name and starting time for both data within a matched pair were created, so the script knew which file has the experimental data or control data, and to trim the strain data to a 16-second interval centered at the wanted timepoint for each sample using the slicing function in Python. Once the 16-second time-series was established, the script calculated the ASDs, one for the experimental sample, another for the control sample, for each pair. To reduce the effect of other transient noise sources that can confound with wind's effect on creating noise, stacking was applied to every experimental data withing each matched-pair. Specifically, nine 16-second intervals, each 1 minute apart, four before the experimental sample time point, one at the time point, and 4 after the time point, were separately read and calculated into ASD. These nine ASDs were then averaged to produce one ASD to represent the experimental time point with fewer variances between individual ASD data. The result of stacking, as shown in figure 3, was that the variances over all frequency ranges were generally decreased, with a stronger decrease over lower-frequency ranges than higher ones. Finally, the control time point's ASD was subtracted

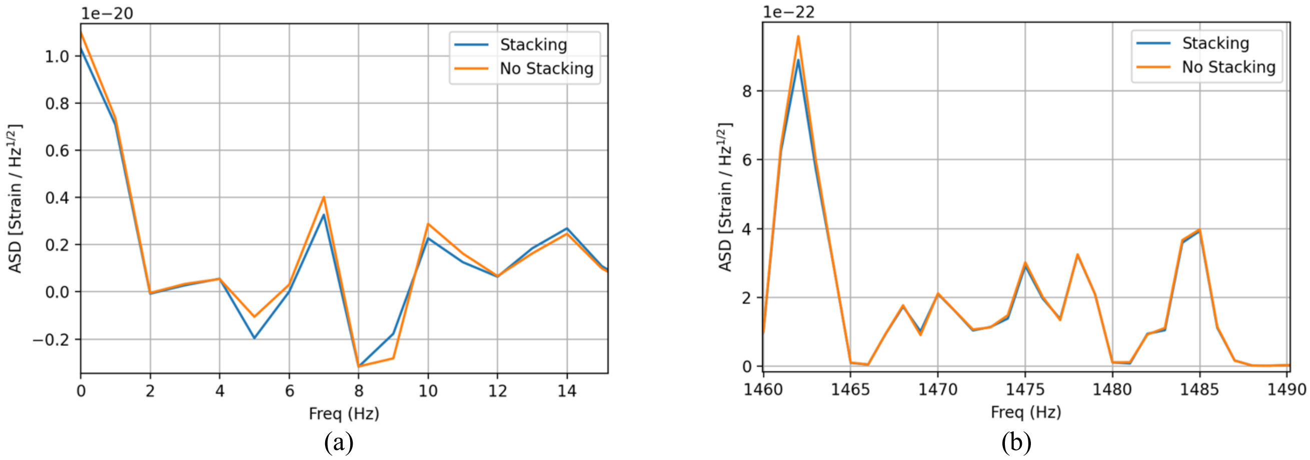
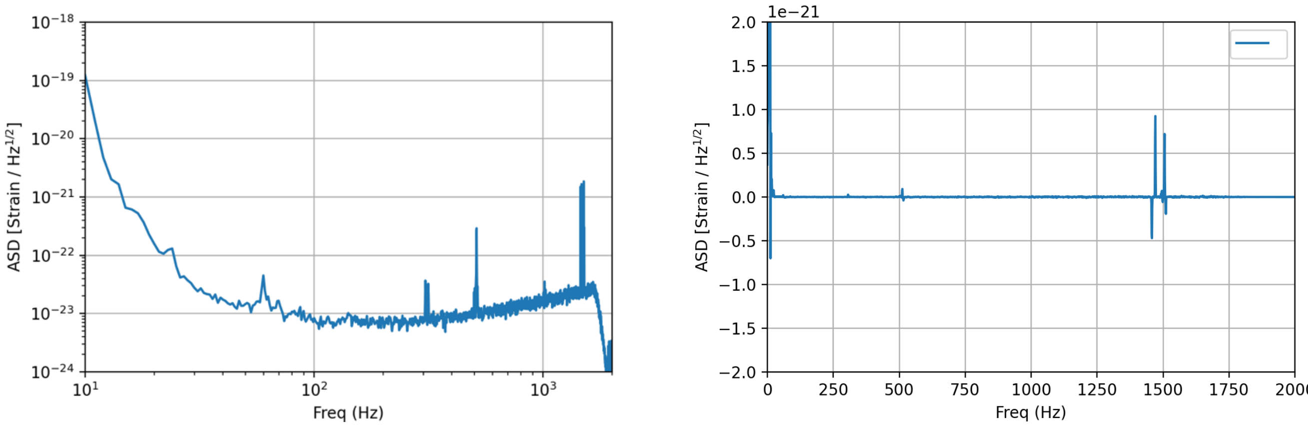


Figure 3. diagrams of differences in ASD between an experimental sample and its paired control sample from Hanford at low-frequency range (a) and high-frequency range (b). The experimental time point, in GPS time, is 1177282398, and the control time point, in GPS time, is 1177318398. Applying stacking to the experimental data (blue line) results in a generally smaller variance than not applying stacking (orange line) at the low-frequency range, but the difference is less significant at the high-frequency range.

from the experimental point's ASD to calculate the differences in between, which are measurements of noise levels caused by wind. The advantage of using the differences of pairs instead of reading directly from the experimental sample's ASD is that the "differences" graph produces ASD data points that are caused mainly by differences in wind speed, whereas the entire ASD for an experimental sample contains multiple factors that can contribute to high ASD values in certain frequencies. Using the "differences in ASD" graph, compared with using ASD plots, allows us to not only pinpoint frequencies that wind has a major effect but also determine the rise in ASD that the wind in the experimental sample causes, which is shown in figure 4.

Using the difference-of-ASD graphs for all thirty time points, an overall graph was created, which contains all thirty time points of a variety of wind speeds. This helped to find the most common frequencies that have high ASD values, or the frequencies that the wind has the most impact on. With these frequencies chosen, each sample's differences-in-ASD value at these frequencies were read, and the readings were plotted into a wind speed v. differences-in-ASD plot for each frequency determined. Because wind may behave differently on Hanford's LIGO



(a) (b)

Figure 4. A plot of a regular ASD of an experimental time point from Livingston (a), and a plot of the difference of ASD between the same experimental time point and its paired control time point (b). The experimental time point, in GPS time, is 1175457198, and the control time point in its pair, in GPS time, is 1175403198. The right plot has high-value ASD points at fewer frequencies, allowing wind's impact frequencies to be more obvious.

and Livingston's LIGO, the 30 samples for Hanford have had their differences in ASD graphed, read, and analyzed separately from Livingston, and the frequencies being read differ between these two locations.

IV. Results

I. Hanford

By plotting the overall differences-in-ASD plot for all thirty sample pairs (shown in figure 5 and 6), the following frequencies showed the greatest concentration of high differences-in-ASD values, which means that the differences between the experimental, wind-affected sample and the controlled, without wind sample at these frequencies are generally the largest: 5Hz, 8 Hz, 14Hz at low frequencies; =502Hz, 506Hz around the 500Hz mirror suspension violin mode; 997Hz, 1009Hz around the 1000Hz mirror suspension violin mode; and 1457Hz, 1462Hz, 1484Hz around the 1500Hz mirror suspension violin mode. The possible reason why these frequencies are so close to mirror suspension violin mode frequencies is that the violin modes are

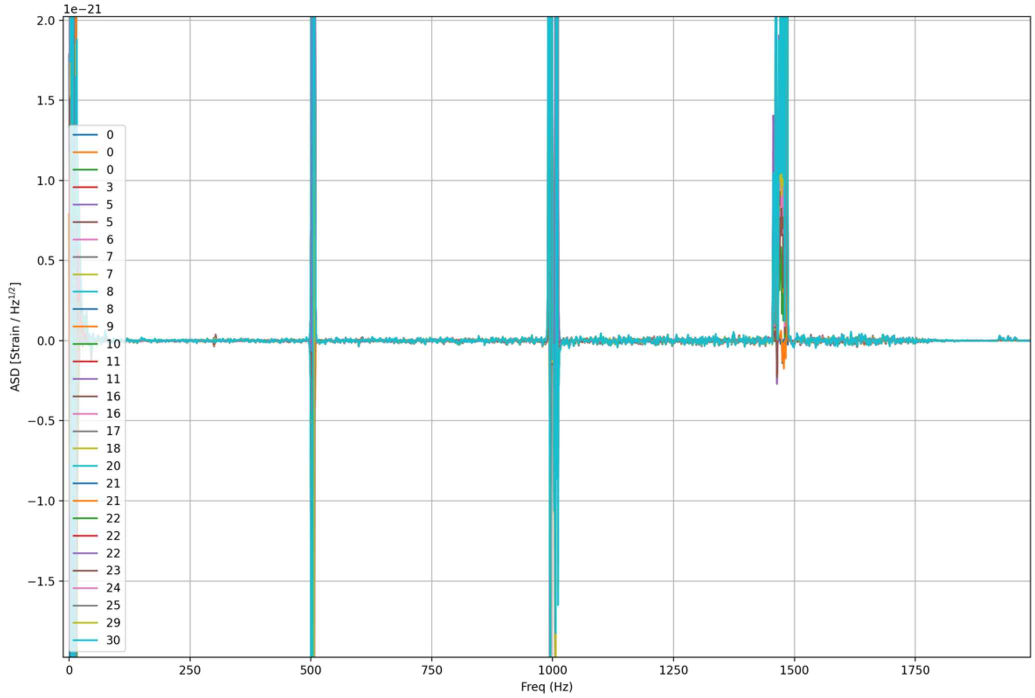


Figure 5. A plot of all thirty sample pair's differences-in-ASD for Hanford. The numbers in the legends show the corresponding line's wind speed. The majority of the peaks of differences-in-ASD occur in the aforementioned frequencies.

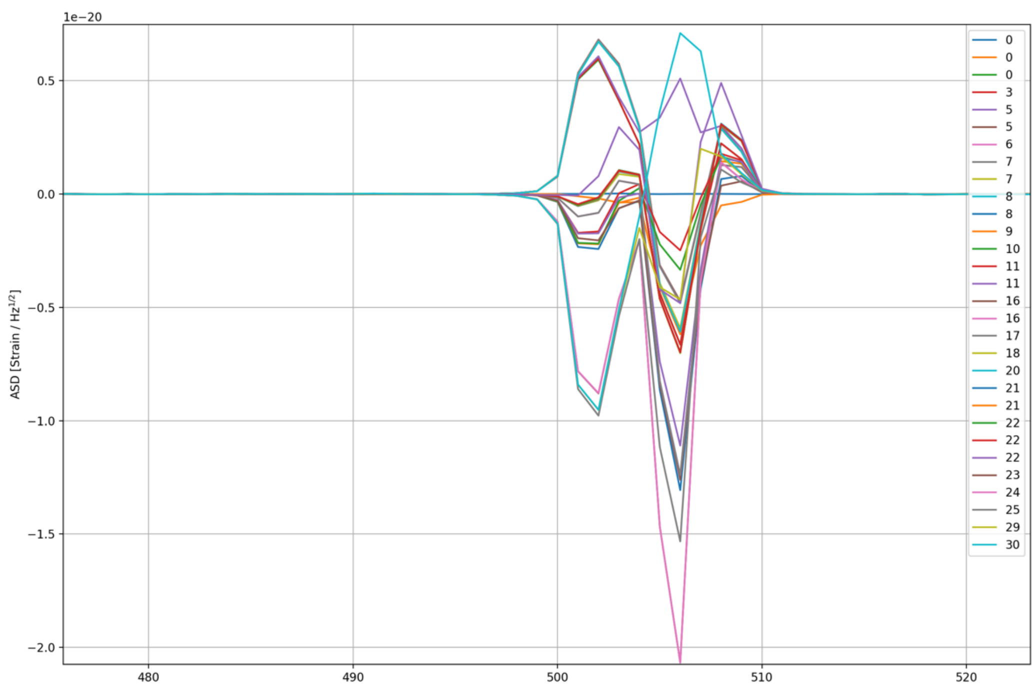


Figure 6. A plot of all thirty sample pair's differences-in-ASD data for Hanford between 480Hz and 520Hz. The numbers in the legends show the corresponding line's wind speed. There are multiple peaks around the 500Hz suspension fiber violin mode (one at 502Hz, another at 506Hz). Similar patterns can be found at other violin modes and low-frequency range in both Hanford and Livingston.

caused by the instrument's suspension fibers, the fibers and wires that hold all the lasers, mirrors and detectors in place. Any vibration that can cause significant changes in the suspension of the mirrors inside the instrument, such as wind, can cause noise in the data at the suspension violin mode frequencies[15]. In other words, the wind served as a "bow" on a "violin" named LIGO to produce signals at 500Hz, 1000Hz, and 1500Hz.

Each sample pair's differences in ASD at the aforementioned frequencies were read and plotted into wind speed v. differences-in-ASD plot for each frequency with a linear regression, some of which are shown below as figure 7-11. The y-intercept of the regression was set to 0 because the wind cannot cause any noise when it is absent. The coefficient of determination (R-squared) is also shown. This coefficient, when taken the square root, yields the correlation coefficient (r), a value between 1 and -1.

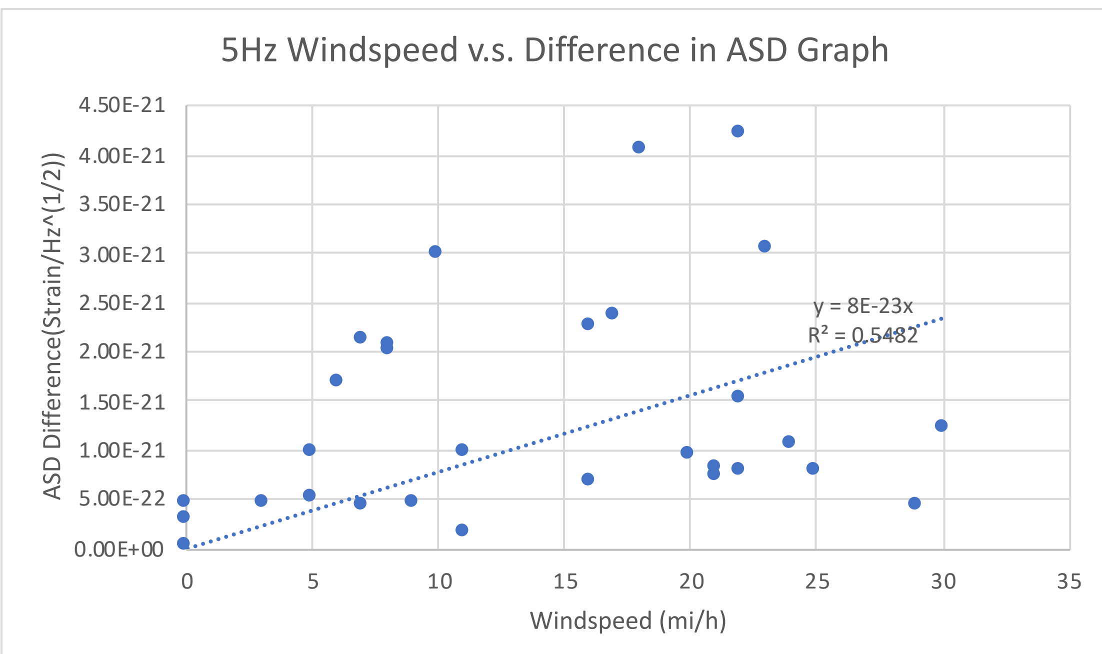


Figure 7. The wind speed v. differences-in-ASD plots for 5Hz signals at LIGO Hanford. The plot is displayed with a linear regression, its equation and the coefficient of determination(R-squared).

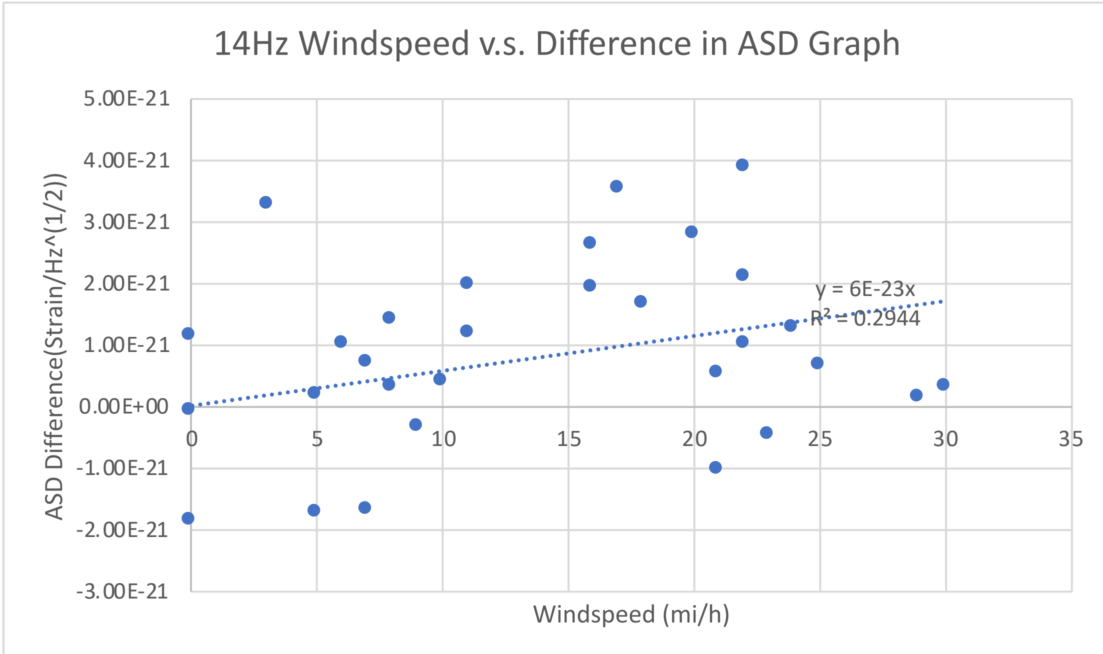


Figure 8.The wind speed v. differences-in-ASD plots for 14Hz signals at LIGO Hanford. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 8Hz follow a similar distribution with a R-squared value of 0.395.

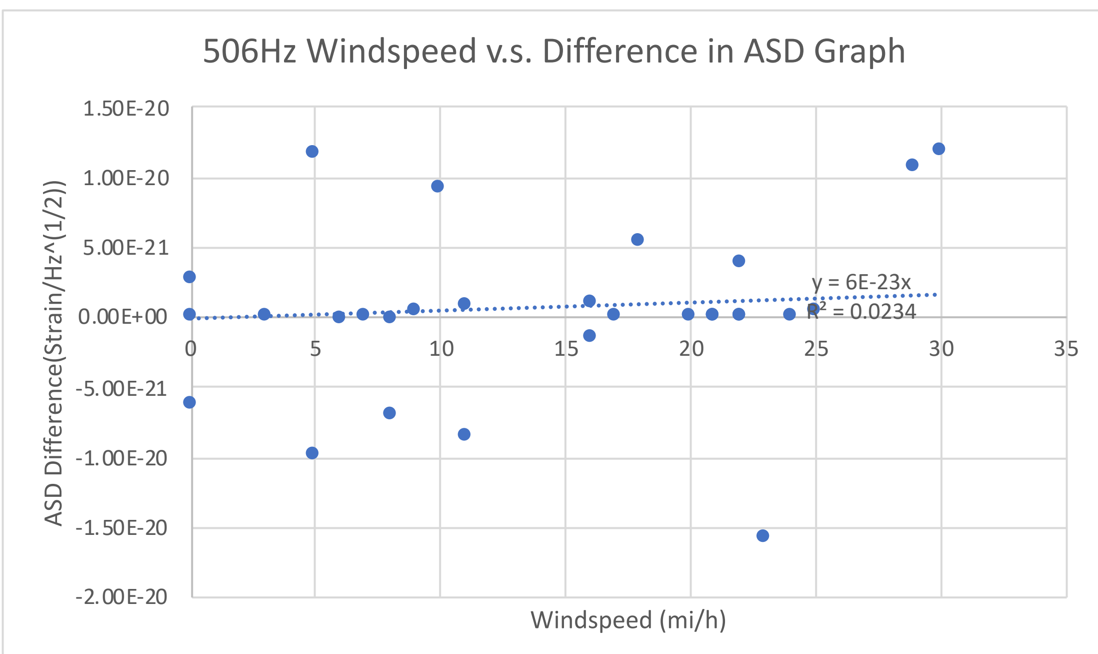


Figure 9.The wind speed v. differences-in-ASD plots for 506Hz signals at LIGO Hanford. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 502Hz follow a similar distribution with a R-squared value of 0.010.

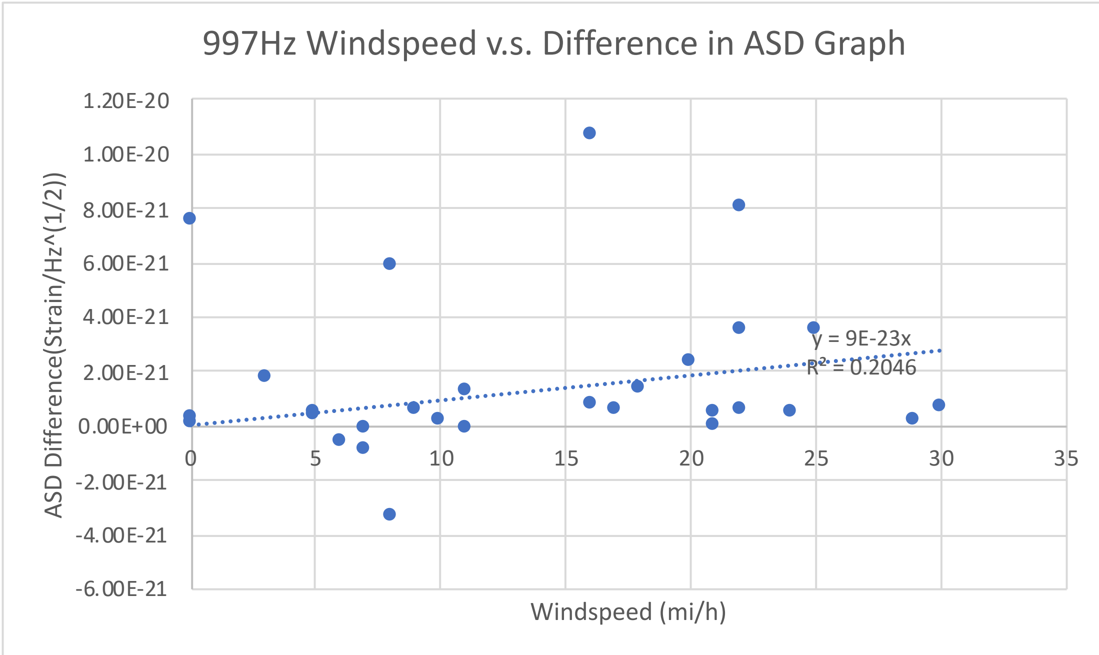


Figure 10.The wind speed v. differences-in-ASD plots for 997Hz signals at LIGO Hanford. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 1009Hz follow a similar distribution with a R-squared value of 0.201.

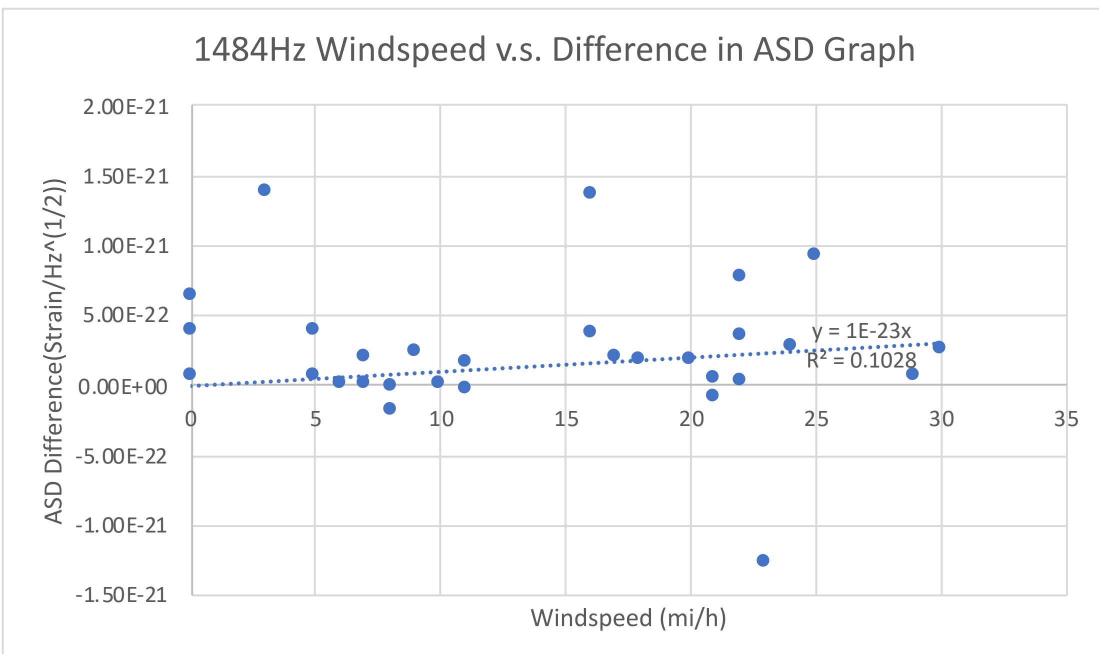


Figure 11.The wind speed v. differences-in-ASD plots for 1484Hz signals at LIGO Hanford. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 1462Hz follow a similar distribution with a R-squared value of 0.052. Data at 1457Hz follow a similar distribution with a R-squared value of 0.150.

Noise Frequency (Hz)	Coefficient of Determination (R-squared)	Correlation Coefficient (r)
5	0.548	0.740
7	0.395	0.628
14	0.294	0.542
502	0.010	0.100
506	0.023	0.152
997	0.205	0.453
1009	0.201	0.448
1457	0.150	0.387
1462	0.052	0.228
1484	0.103	0.321

Table 1. The Coefficient of Determination (R-squared) and Correlation Coefficient (r) for each noise frequency for LIGO Hanford. The Coefficient of Determination is obtained directly from the plots of differences-in-ASD data, and the Correlation Coefficient is calculated by taking the square root of the Coefficient of Determination. Because the linear regressions for all test noise frequencies have a positive slope, the Correlation Coefficient is positive for all test noise frequencies in Hanford.

A positive correlation coefficient means the dependent variable increases as the independent variable increases, and negative means the opposite. The larger the absolute value of the correlation coefficient, the stronger relationship there is between the two variables. When the correlation coefficient's absolute value is greater than 0.7, it means that there is a strong correlation between the two variables. Any absolute value between 0.7 and 0.3 means a medium correlation. Having the absolute value below 0.3 means there's little to no correlation. The 5Hz data show a correlation coefficient of 0.740, which is larger than 0.7. This means that there is a strong correlation between wind speed and elevated ASD values at 5Hz in Hanford, which implies that wind speed is positively related to the amount of noise at 5Hz. In the same lower frequency range as 5Hz, data at 8Hz and 13Hz show correlation coefficients of 0.628 and 0.542. This means that although the correlation between wind speed and noise level is not as strong as even lower frequencies, there is a medium-strong correlation between the two at 8Hz and 14Hz in Hanford.

At frequencies around the 1000Hz fiber suspension violin mode (997Hz, 1009Hz), the correlation coefficients are 0.453 and 0.448, yielding a medium correlation between wind speed and its impact on ASD values around 1000Hz. This correlation, however, was not seen around other suspension violin modes. The frequencies around the rest of fiber suspension violin modes all demonstrated a weak correlation by having a correlation coefficient of below or not significantly above 0.3, also meaning that wind's impact on these frequencies around violin modes was less significant.

Therefore, the wind's impact on LIGO Hanford detector noise is the strongest in lower frequencies such as 5Hz, 8Hz, and14Hz by showing a strong correlation between wind speed and differences in ASD, and a moderate correlation at frequencies around the 1000Hz violin mode. At other frequencies, wind has little to almost no influence in elevating detector noise. The elevated ASD values at these frequencies for the experimental sample compared to its paired, control sample could be caused by other sources of random vibrations, which did not increase significantly as wind speed increases.

2. Livingston

By plotting the overall differences-in-ASD plot for all thirty sample pairs, which is shown in figure 12, the following frequencies showed the greatest concentration of high differences-in-ASD values, which means that the differences between the experimental, wind-affected sample and the controlled, without wind sample are generally the largest: 7Hz, 15Hz at low frequencies; 500Hz, 513Hz around the 500Hz violin mode; 997Hz, 1012Hz around the 1000Hz violin mode; and 1458Hz, 1471Hz, 1505Hz around the 1500Hz violin mode. These frequencies are generally close to those from Hanford.

Each sample pair's differences in ASD for each of the aforementioned frequencies were then read and plotted into a wind speed v. differences-in-ASD plot for each noise frequency, some of which are shown below as figure 13-17. The y-intercept was set to zero because the wind cannot cause any noise when it is absent. The coefficient of determination (R-squared) is also shown, which was later calculated into the correlation coefficient (r) by taking its square root.

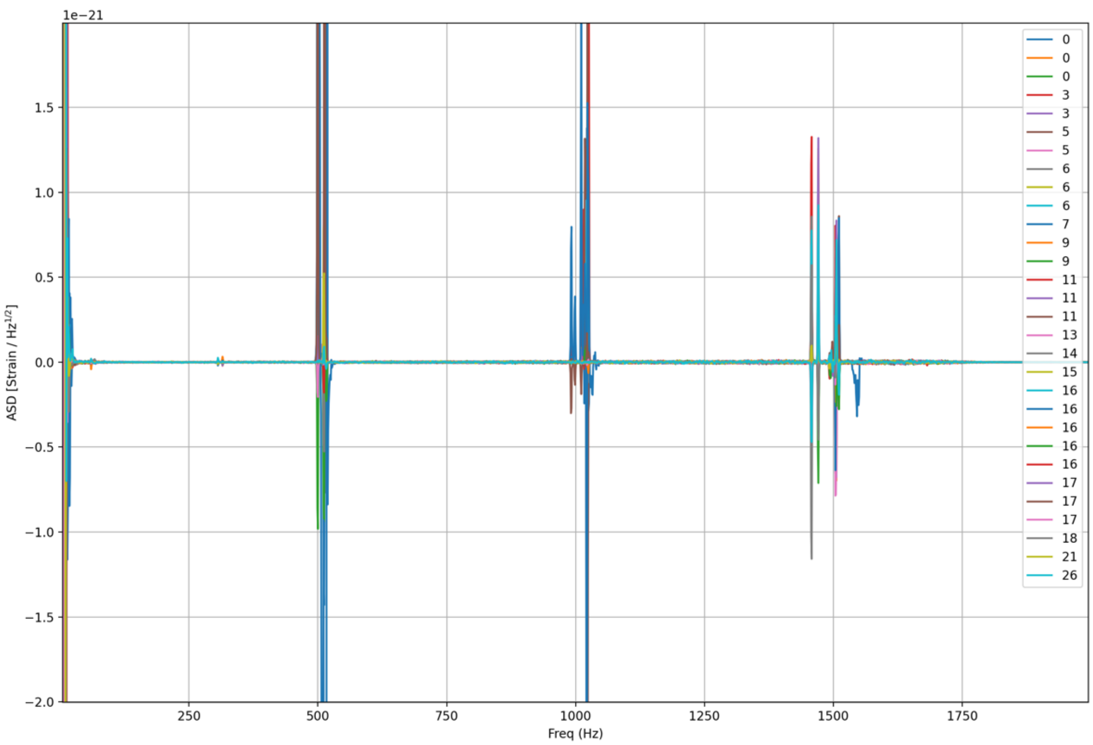


Figure 12. A plot of all thirty sample pair's differences-in-ASD for Livingston. The numbers in the legends show the corrresponding line's wind speed. The majority of the peaks of differences-in-ASD occur in the aforementioned frequencies.

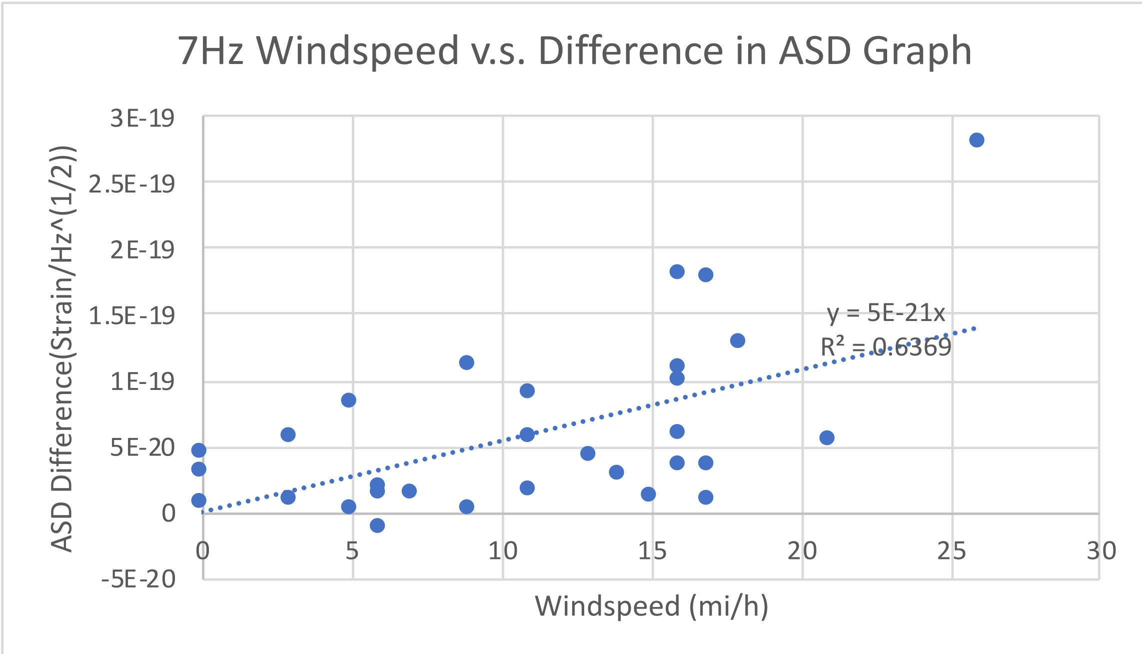


Figure 13. The wind speed v. differences-in-ASD plots for 7Hz signals at LIGO Livingston. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared).

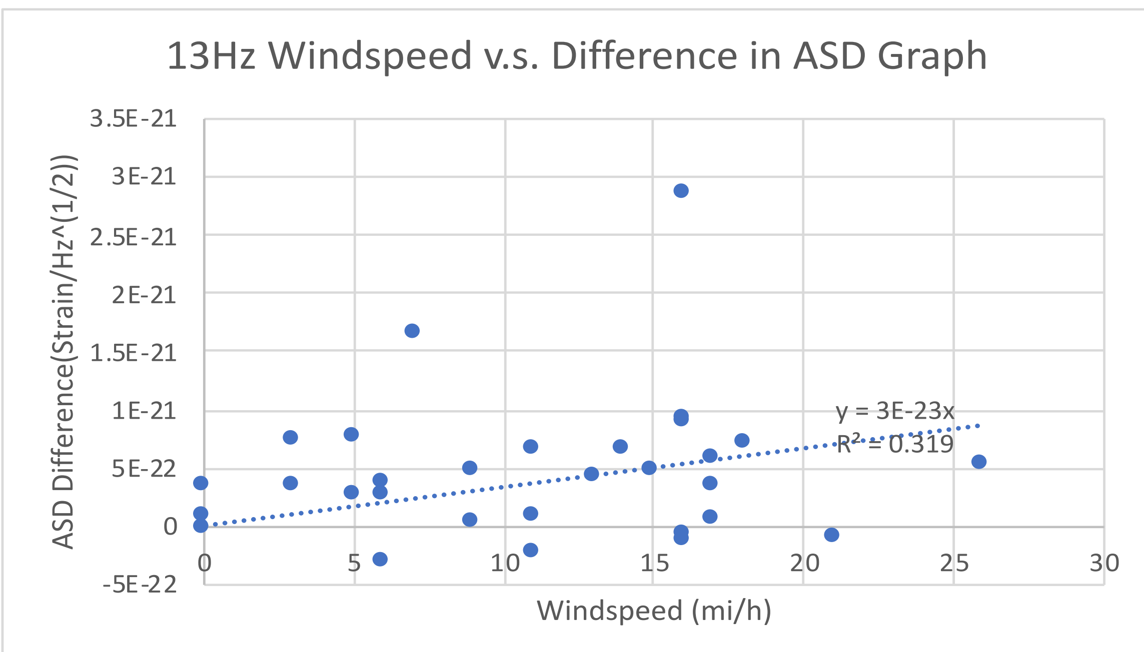


Figure 14. The wind speed v. differences-in-ASD plots for 13Hz signals at LIGO Livingston. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared).

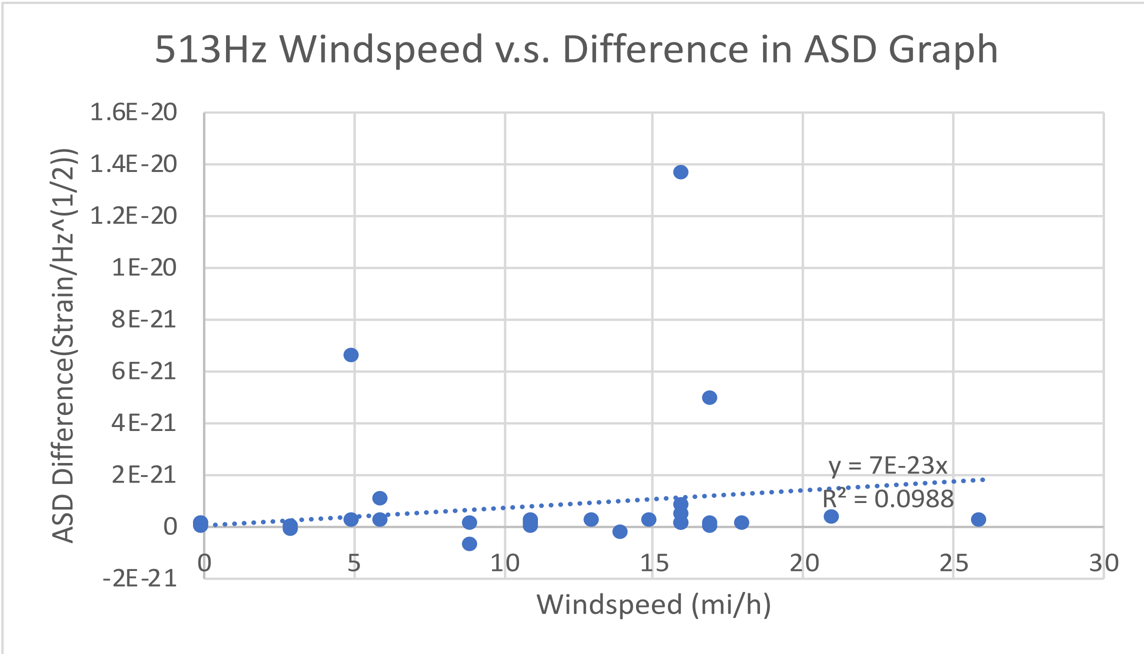


Figure 15. The wind speed v. differences-in-ASD plots for 513Hz signals at LIGO Livingston. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 500Hz follow a similar distribution with a R-squared value of 0.136.

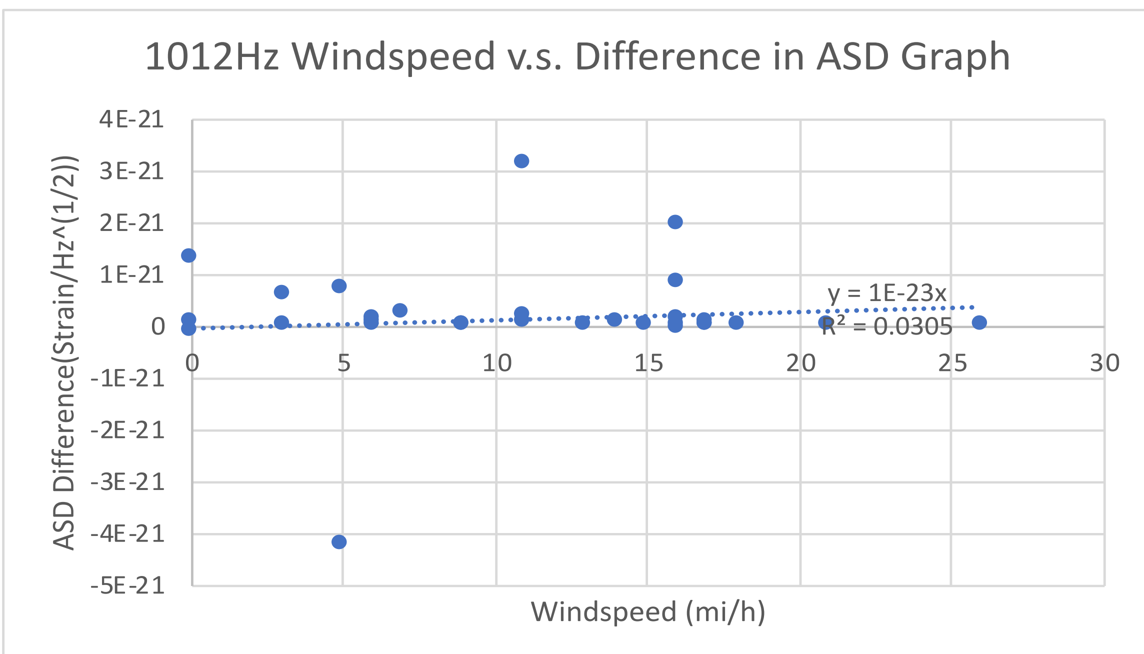


Figure 16. The wind speed v. differences-in-ASD plots for 1012Hz signals at LIGO Livingston. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 997Hz follow a similar distribution with a R-squared value of 0.044

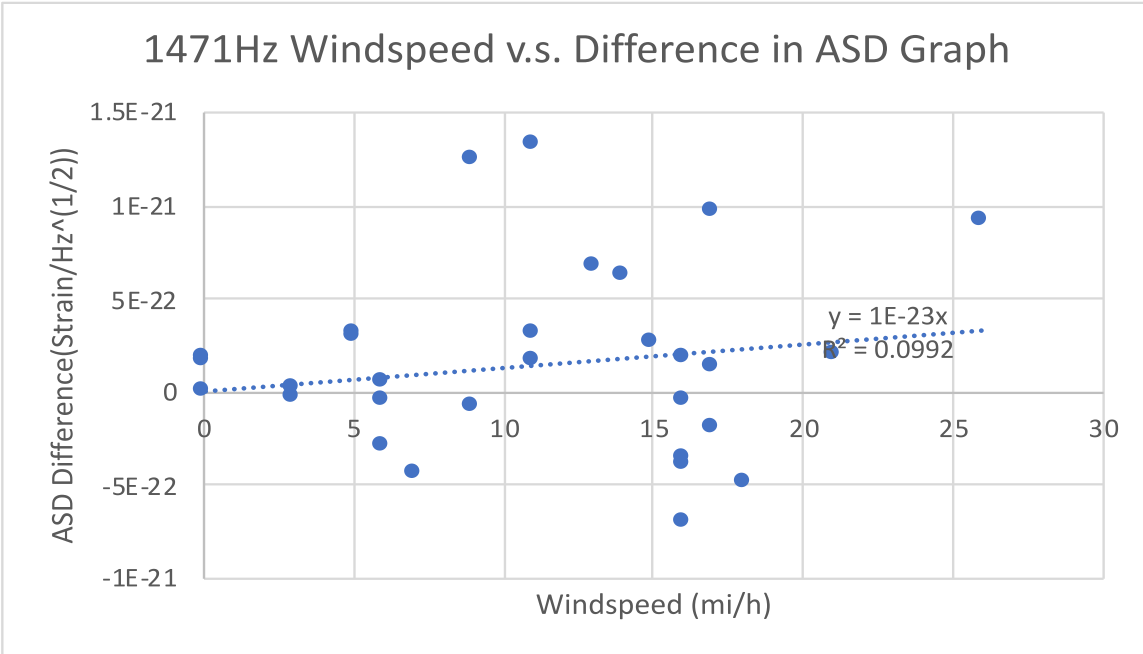


Figure 17. The wind speed v. differences-in-ASD plots for 1471Hz signals at LIGO Livingston. The plot is displayed with an linear regression, its equation and the coefficient of determination(R-squared). Data at 997Hz follow a similar distribution with a R-squared value of 0.044

Noise Frequency (Hz)	Coefficient of Determination (R^2)	Correlation Coefficient (r)
7	0.637	0.798
13	0.319	0.565
500	0.136	0.369
513	0.099	0.314
997	0.044	0.210
1012	0.031	0.175
1458	0.038	0.195
1471	0.099	0.315
1505	0.181	0.426

Table 2. The Coefficient of Determination (R-squared) and Correlation Coefficient (r) for each noise frequency for LIGO Livingston. The Coefficient of Determination is obtained directly from the plots of differences-in-ASD data, and the Correlation Coefficient is calculated by taking the square root of the Coefficient of Determination. Because the linear regressions for all test noise frequencies have a positive slope, the Correlation Coefficient is positive for all test noise frequencies in Livingston.

The 7Hz data show a correlation coefficient of 0.8. This means that there is a strong correlation between wind speed and elevated ASD values at 7Hz, which implies that wind speed is positively related to the amount of noise at 7Hz. In the same lower frequency range as 7Hz, data at 13Hz show a coefficient of determination of 0.407 or a correlation coefficient of 0.64. This means that although the correlation between wind speed and noise level is not as strong as even lower frequencies, there is a moderate correlation between the two at 13 Hz.

The rest of the frequencies around the violin modes all demonstrated a weak correlation by having a correlation coefficient of below or not significantly above 0.3 after taking the square root of their R-squared values. Among them, frequencies around 500Hz violin mode (500Hz, 513Hz) and 1500Hz violin mode (1471Hz, 1505Hz) demonstrated a slightly stronger correlation, which is different from Hanford.

Therefore, wind's impact on LIGO detector noise is the strongest in lower frequencies such as 7Hz and 13Hz by showing a strong correlation between wind speed and differences in ASD, while at other frequencies wind has little to almost no influence in elevating detector noise. This is generally consistent with the results from Hanford, with one difference: while LIGO Hanford's sensitivity around 1000Hz tends to be more affected than Livingston during the presence of wind, represented by correlation coefficients at these frequencies that are 0.2 higher than those of Livington, LIGO Livingston's sensitivity around 500Hz and 1500Hz tend to be more affected by having correlation coefficients around these frequencies that are 0.1 - 0.2 higher than those of Hanford. Besides, relatively medium-high correlation coefficients are clustered around 1000Hz in Hanford, excluding low frequencies; while they are clustered around 500Hz and 1500Hz in Livingston, excluding low frequencies.

V. Summary

After calculating the difference of ASD between various samples that vary in wind speed with their controlled, zero-wind-speed pairs, the wind was found to have the largest impact on LIGO sensitivity at low frequencies in both LIGO Hanford and LIGO Livingston. Higher frequencies, such as those around 500Hz, 1000Hz and 1500Hz, have little to no elevated noise that could be explained by wind speed alone. The difference between Hanford and Livingston is that while LIGO Hanford's sensitivity around 1000Hz tends to be more affected than Livingston during the presence of wind, LIGO Livingston's sensitivity around 500Hz and 1500Hz tend to be more affected.

VI. Recommendations for Further Research

Due to the limitations of the scope of this research, further research using more weather data and LIGO data is recommended to analyze a more precise quantitative relationship between wind and LIGO strain data noise. One other aspect that deserves attention is that due to limitations of the accuracy of the data and python script, this research can only look at frequencies at the order of magnitude of 1Hz. However, a recent paper by Michael P. Ross et al. wrote:" [LIGO] are susceptible to contamination from a ground tilt at frequencies below 0.1 Hz, particularly arising from wind-pressure acting on building walls"[16].This research could not examine the correlation between wind speed and differences-in-ASD at 0.1Hz. Therefore, to further understand wind's impact on low-frequency noise levels, similar researches that can read in strain data and ASDs more accurately is necessary.

One particular weather event that deserves further investigation is Hurricane Barry, which went over Livingston between July the 13th and 15th. Because the eye of this hurricane passed directly over LIGO Livingston, the detector was able to experience extreme weather in both directions: both extremely strong wind with heavy rain, and mild breeze with no rain. Such shift between these two extremes within a short period of time means that the strain data during Hurricane Barry can be valuable in analyzing the influences of extreme weather, such as wind speed or rain, on LIGO sensitivity as well as how such quick shift between these extremes can impact LIGO strain data differently than normal wind speed or rainfall. Unfortunately, the strain data during Hurricane Barry, which is part of Observation Run 3, has not been available throughout this research, and as a result there are no concrete evidence that LIGO Livingston has recorded data during the hurricane, as it may has been shut down temporarily. I would like to encourage future researches in the area of analyzing noise sources for LIGO to take into account similar hurricanes and reveal the impact of weather on LIGO sensitivity during weather conditions that are unachievable normally. This will help us to develop a more precise model to track LIGO noise levels caused by weather conditions.

VII. Acknowledgments

I would like to thank Professor Eric Myers for providing guidance on creating and focusing my ideas on the topic of gravitational waves, as well as his valuable comments on this paper. This research was made possible thanks to the opportunity and resources provided by Pioneer Academics. The Gravitational Wave Open Science Center (GWOSC) is acknowledged for providing access to LIGO data and part of the python codes, and the National Oceanic and Atmospheric Administration (NOAA) is also acknowledged for providing access to past weather data for areas of both LIGO locations.

VIII. References

[1] "What Are Gravitational Waves." LIGO Caltech. Accessed July 29, 2020. https://www.ligo.caltech.edu/page/what-are-gw.

[2] "What Is a Gravitational Wave?" NASA Space Place. Last modified June 4, 2020. Accessed July 29, 2020. https://spaceplace.nasa.gov/gravitational-waves/en/.

[3] Choi, Charles Q. "With New Gravitational-wave Detectors, More Cosmic Mysteries Will Be Solved." Space.com. Last modified May 8, 2019. https://www.space.com/gravitational-waves-future-discoveries.html.

[4] *LIGO's Interferometer*. Photograph. https://www.ligo.caltech.edu/page/ligos-ifo.

[5] "What Is an Interferometer." LIGO Caltech. Accessed July 29, 2020. https://www.ligo.caltech.edu/page/what-is-interferometer.

[6] Lyons, Richard. "Windowing Functions Improve FFT Results, Part I." EDN. Last modified June 1, 1998. Accessed July 29, 2020. https://www.edn.com/windowing-functions-improve-fft-results-part-i/.

[7] *Lots of Plots*. Image. Gravitational wave open science center. https://www.gw-openscience.org/static/images/lotsofplots.png.

[8] "Plot the Amplitude Spectral Density." In *Binary Black Hole Signals in LIGO Open Data*. Version 1.63 ed. 2017. https://notebooks.azure.com/api/user/losc/project/tutorials/html/LOSC_Event_tutorial.ipynb#x_Plot-the-Amplitude-Spectral-Density-(ASD).

[9] "About 'aLIGO.'" LIGO Caltech. Accessed July 29, 2020. https://www.ligo.caltech.edu/page/about-aligo.

[10] "LIGO Suspends Third Observing Run (O3)." LIGO Caltech. Last modified March 26, 2020. Accessed August 19, 2020. https://www.ligo.caltech.edu/news/ligo20200326.

[11] "The Sensitivity of the Advanced LIGO Detectors at the Beginning of Gravitational Wave Astronomy." Accessed July 29, 2020. https://www.ligo.org/science/Publication-O1Noise/index.php.

[12] "Data Tools: Local Climatological Data." National Center for Environmental Information. Accessed July 29, 2020. https://www.ncdc.noaa.gov/cdo-web/datatools/lcd.

[13] "LIGO and Virgo Data." Gravitational Wave Open Science Center. Accessed July 29, 2020. https://www.gw-openscience.org/data/.

[14] "Tutorial: Lots of Plots." Gravitational Wave Open Science Center. Accessed July 29, 2020. https://www.gw-openscience.org/tutorial05/.

[15] "O2 Instrumental Lines." Gravitational Wave Open Science Center. Accessed July 29, 2020. http://compliance.fifoundry.net.

[16] Ross, Michael P., Krishna Venkateswara, Conor Mow-Lowry, Sam Cooper, Jim Warner, Brian Lantz, Jeffrey Kissel, Hugh Radkins, Thomas Shaffer, Richard Mittleman, Arnaud Pele, and Jens Gundlach. *Towards Windproofing LIGO: Reducing the Effect of Wind-driven Floor Tilt by Using Rotation Sensors in Active Seismic Isolation.* Research report no. 2003.06447v2. 2020. https://arxiv.org/pdf/2003.06447.pdf.

IX. Appendix: Python script used for this paper

```
#Rudy Zhang, zxm011022@gmail.com, 7/29/2020
#This script calculates the difference in ASD between experimental LIGO data
#samples and its paired control samples, which are available from GW Open
#Science Center. Each sample requires a starting time, in GPS time, to start
#reading strain data and calculate the ASD.
#The number of sample pairs, the time series reading length, number of times
#for stacking, and stacking time intervals are all adjustable.
#The script will produce a graph recording the difference in ASD between each
#sample pairs (shown in colors and with legends) across frequencies of 0-2000Hz.
#Please go to https://www.gw-openscience.org/tutorials/, follow the steps in the
#"Introduction to GWOSC Data Files" section to download the necessary softwares
#before using this python script.

# Import needed modules
import numpy as np
import h5py
import matplotlib.pyplot as plt
import matplotlib.mlab as mlab
import readligo as rl

dataList = [] #This is the list object for all the lables for legends, which are
              #Windspeed data that you will imput for each sample pair

#The definition of the function to calculate ASD values for each sample time
def calcAsd(fileName1,sliceStart1):
    #Read in strain data of the inputed file
    fileName = fileName1
    strain, time, channel_dict = rl.loaddata(fileName, 'H1')
    ts = time[1] - time[0] #Time between samples
    fs = int(1.0 / ts)     #Sampling frequency

    #Find a good segment with available strain data and get the entire data
    segList = rl.dq_channel_to_seglist(channel_dict['DEFAULT'], fs)
    length = 10000
    strain_seg = strain[segList[0]][0:(length*fs)]
    time_seg = time[segList[0]][0:(length*fs)]

    #Slice the data to the wanted range
    sliceStart = sliceStart1 #The start of slicing is a constant you input later
    sliceStart = int((sliceStart - time_seg[0])*fs) #converting time to frames
    sliceEnd = int(sliceStart + 16*fs) #The end is 16 seconds after your input
                                       #You can change the 16 to your ideal length
    slice_object = slice(sliceStart,sliceEnd,1)
    strain_seg = strain_seg[slice_object]

    #Apply a Blackman Window, and plot the FFT
    window = np.blackman(strain_seg.size)
    windowed_strain = strain_seg*window
    freq_domain = np.fft.rfft(windowed_strain) / fs
    freq = np.fft.rfftfreq(len(windowed_strain))*fs

    #Make PSD for first chunk of data
    PSD, freqs = mlab.psd(strain_seg, Fs = fs, NFFT=fs)

    #Converting the PSD tot ASD by taking the square root
    ASD = np.sqrt(PSD)
    return ASD, freqs #Returning the list of ASD values and frequncy to stack

#The definition of the function to apply stacking for the file
```

```
def stackAsd(fileName1,sliceStart1,nStack,tStack):
    ASDTotal = 0

    #The stacking loop. nStack is the number of ASDs taken from the file for
    #stacking, tStack is the time interval between each take of ASD
    for n in range (0,nStack):
        ASD, freqs = calcAsd(fileName1,sliceStart1) #Calculating the ASD
                                                    #from the start time
        ASDTotal = ASDTotal + ASD #Summing the ASDs calculated already
        sliceStart1 = sliceStart1+tStack #adding tstack second to the start time

    ASDAvg= ASDTotal / nStack #Averaging all the ASD takes for the file
    return ASDAvg, freqs #Returning the averaged of ASD values and frequncy

#Setting all the plot properties (size, grid, lables for axis)
fig = plt.figure(figsize=(15,10))
plt.axis([0, 2000, -2e-20, 2e-20])
plt.grid('on')
plt.xlabel('Freq (Hz)')
plt.ylabel('ASD [Strain / Hz$^{1/2}$]')

#Repeating stacking, graphing for the number of sample pairs you intend, which is
#determined by the second number in the range in the "for" statement
for i in range (0,1):
    #Asking input for the experimental sample's file name and start time for slice
    fileName1 = input("What is the experimental file name?")
    sliceStart1 = int(input("What is the starting time for the sample, in GPS time?"))

    #Asking input for the experimental sample's windspeed to be used in legends
    #You can modify the input question for your way of labeling graph entries
    data = str(input("What is the windspeed of this sample?"))
    dataList.append(data)

    #Asking input for the control sample's file name and start time for slice
    fileName0 = input("What is the control file name?")
    sliceStart0 = int(input("What is the starting time for the sample, in GPS time?"))

    #calculating and stacking ASDs for the experimental sample
    ASDAvg1, freqs = stackAsd(fileName1, sliceStart1,9,60)
    #calculating and stacking ASDs for the control sample
    ASDAvg0, freqs = stackAsd(fileName0, sliceStart0,9,60)

    #calculating the difference in ASD betweent the experimental and the control
    ASDAvgD = ASDAvg1 - ASDAvg0
    #plotting the difference in ASD graph for the pair
    plt.plot(freqs, ASDAvgD)

#showing the final graph of all sample pairs with legends
plt.legend(dataList)
plt.show()
print(ASDAvgD)
```