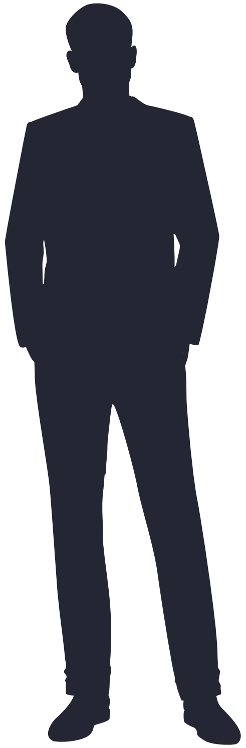
Bayesian optimization for rapid probabilistic estimations of overall level on frequency response models

Alexis Castel, Ph.D.
Rabah Hadjit, Ph.D.
Bryce Gardner, Ph.D.

SCLV 2024

Innovation Without Compromise

The story

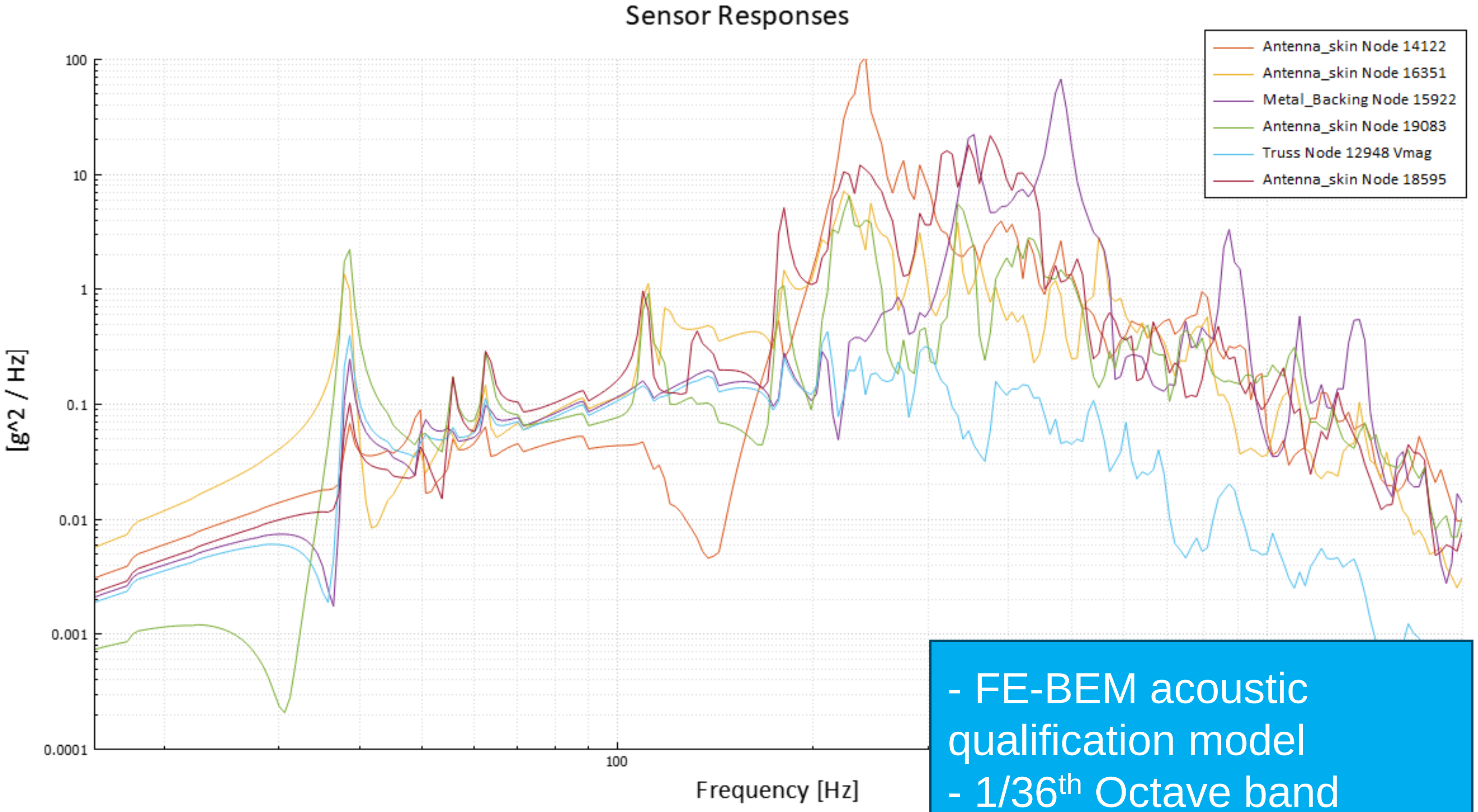


Dynamicist in charge of analyses

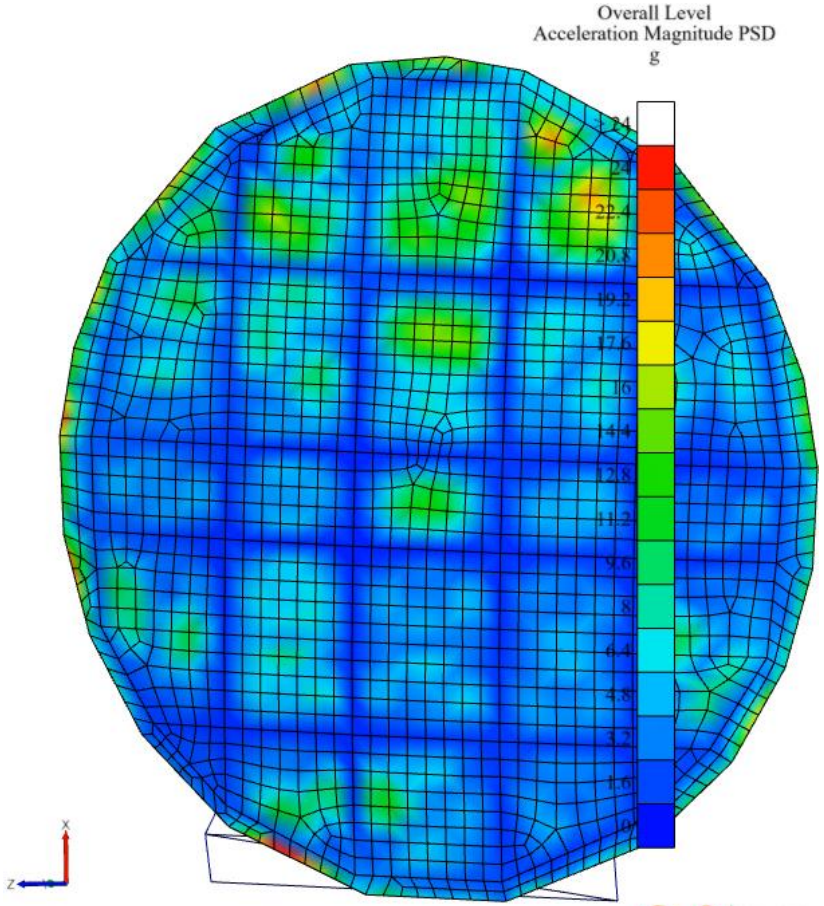
- We reviewed the vibe response results and our internal customers are really pleased with them!
    - Excellent, I am glad to hear it!
- They did have one question though.
    - Sure, what is it?
- How do we know our frequency range is capturing the overall response correctly?
    - Well, let’s look at the response curve…

ESI support engineer

The story



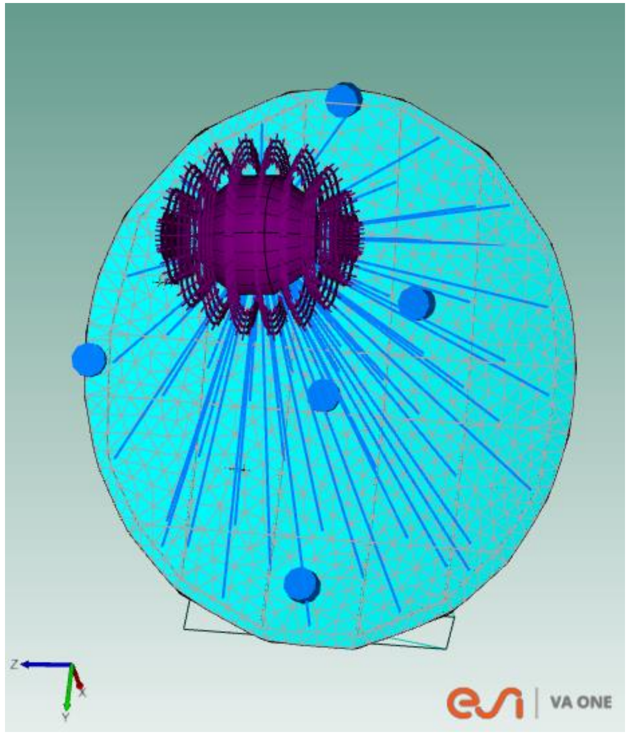
- FE-BEM acoustic qualification model
- 1/36th Octave band
- 15 – 2000 Hz



Typical FEM model for the space industry

SpaceCraft

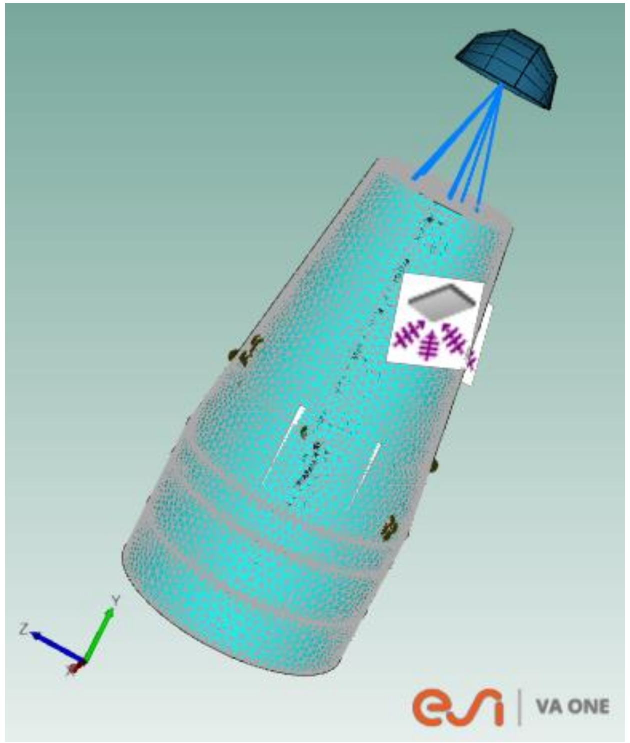
- Coupled FE-BEM model



BEM solver can be computationally expensive

Launch Vehicle

- Launch Vehicle
  - Lift-off
  - Ascent



Large number of modes (upwards of 100k modes!)

Can be Hybrid FE-SEA

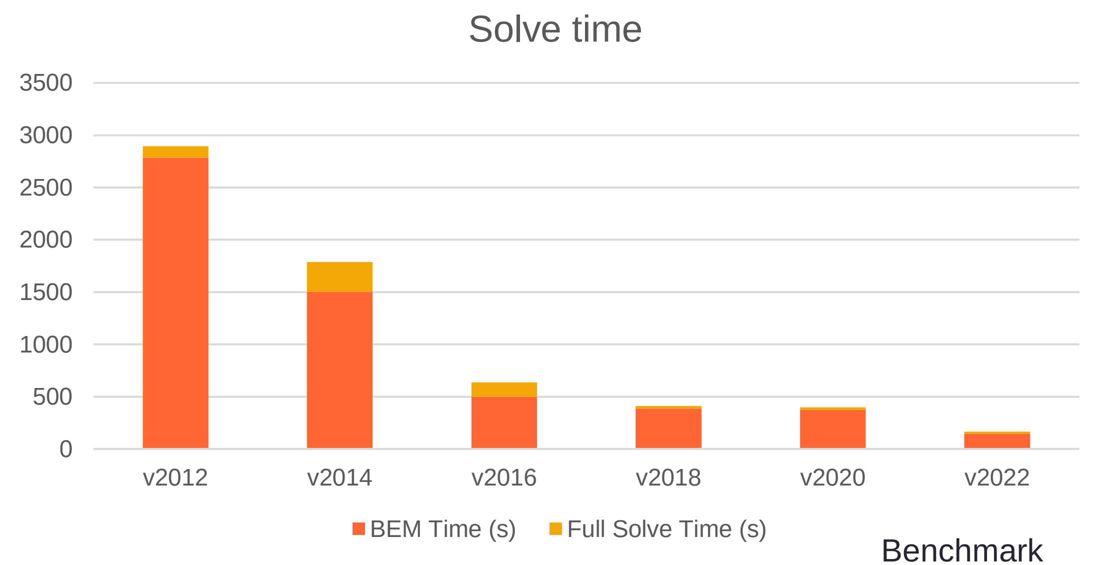
Software Trend: Solve time keeps improving

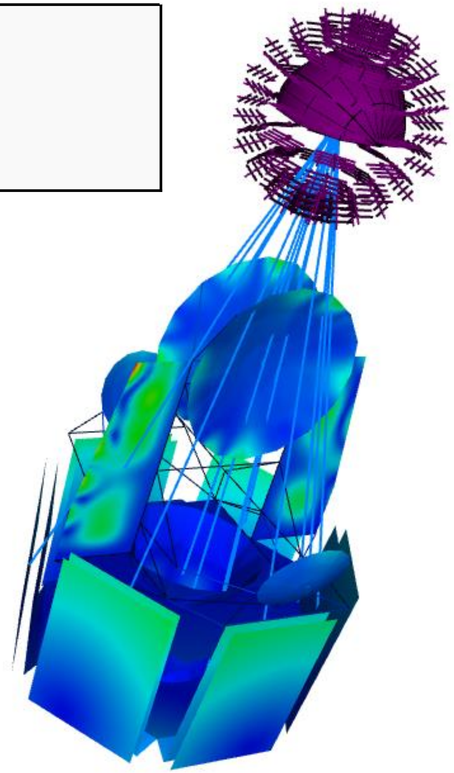
- Standard FE-BEM payload model

Intel Xeon CPU E5-2609 v3 @1.9GHz
12 Cores / 12 Logical processors
64GB RAM

VA One version	v2012	v2014	v2016	v2018	v2020	v2022
BEM Time (s)	2785	1504	499	385	375	145
Full Solve Time (s)	110	283	138	26	25	22
Total	2895	1787	638	411	400	168

5 Frequencies
8813 Wetted nodes
900 Modes
100 Plane waves





Benchmark

Industry trend: models keep growing

- Vehicle/model size is increasing
- Structural finite element solvers can handle larger models
  ➔Increases expectations for other solvers
- Prominence of “integrated models” (multiple segments in a model).
- Want/need to have finite element models “solve it all”

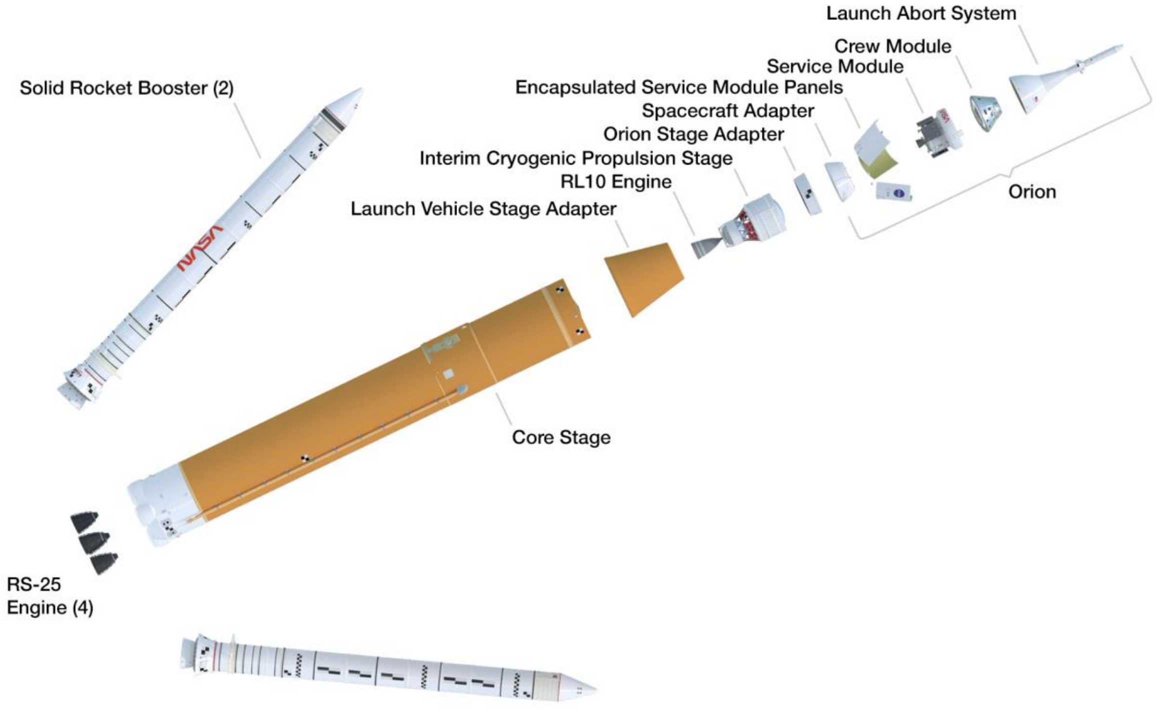


Image credit:nasa.gov

Typical analysis types

Falcon 9 user guide

Inputs:

- Acoustics
  - One third octave band

Frequency (Hz)	Cape Canaveral Acoustic Limit Levels (P95/50), 60% Fill-Factor (Third-Octave)
31.5	118
40	119.5
50	120
63	120
80	119.8
100	120.5
125	121.5
160	122
200	121.5
250	120.5
315	119
400	117
500	115

- Vibration
  - PSD spectrum

Frequency	Falcon 9/Heavy Payload Vibration MPE, (P95/50), 5.13 GRMS
20	0.0044
100	0.0044
300	0.01
700	0.01
800	0.03
925	0.03
2000	0.00644
GRMS	5.13

Frequency range and output:

- Frequency range:
  - N-th octave band (24^{th} – 48^{th} +)
- Reports overall levels and spectra

- Frequency range:
  - Regular spacing
  - N-th octave band (24^{th} – 48^{th} +)
  - Centered around eigen values, etc…
- Reports overall levels and spectra

Ideas

- If we can reliably predict the overall level and curve shape with less frequency points?

- If, as experienced analysts, we can determine that the frequency range converges in terms of overall level, can we teach a computer the same thing?

Overall level definition

- If a given band-limited RMS spectrum is expressed as a power spectral density then it must first be converted to a power-like spectrum and integrated according to the following equation:

$$O.L. = \sqrt{\sum_n S_n^{(psd)} \Delta f}$$

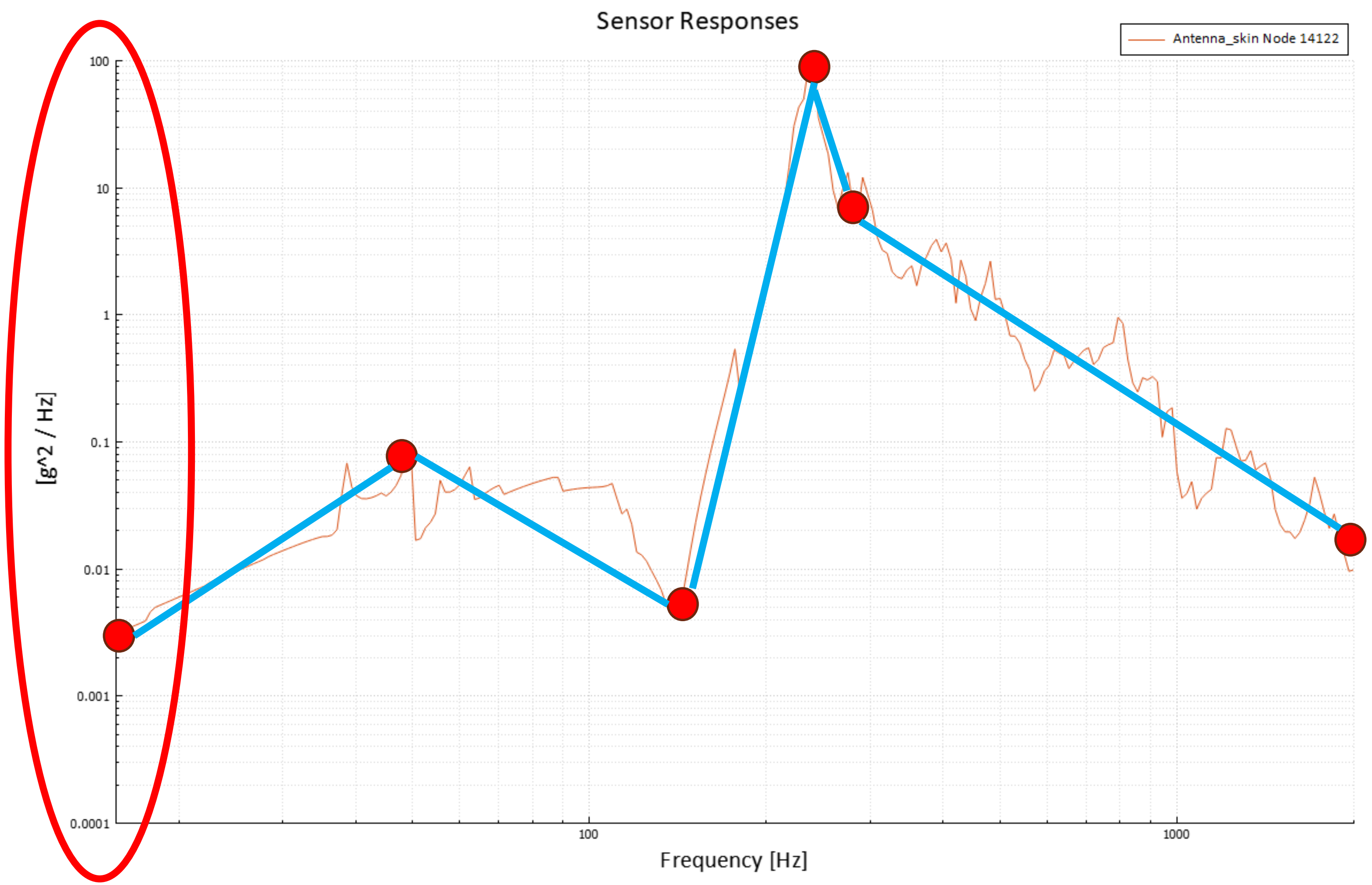
where $S_n^{(psd)}$ is the spectrum value in raw PSD form, before decibel scaling.

➔An overall level is an integral of a function

$$O.L. = \int_{\text{start band}}^{\text{end band}} S_n^{(psd)}(f)\mathrm{d}f$$

➔Can we <u>estimate</u> the integral of the PSD function while minimizing the sampling points?

Are estimations okay?



How many sampling points do we need to accurately predict the overall level of the orange curve?

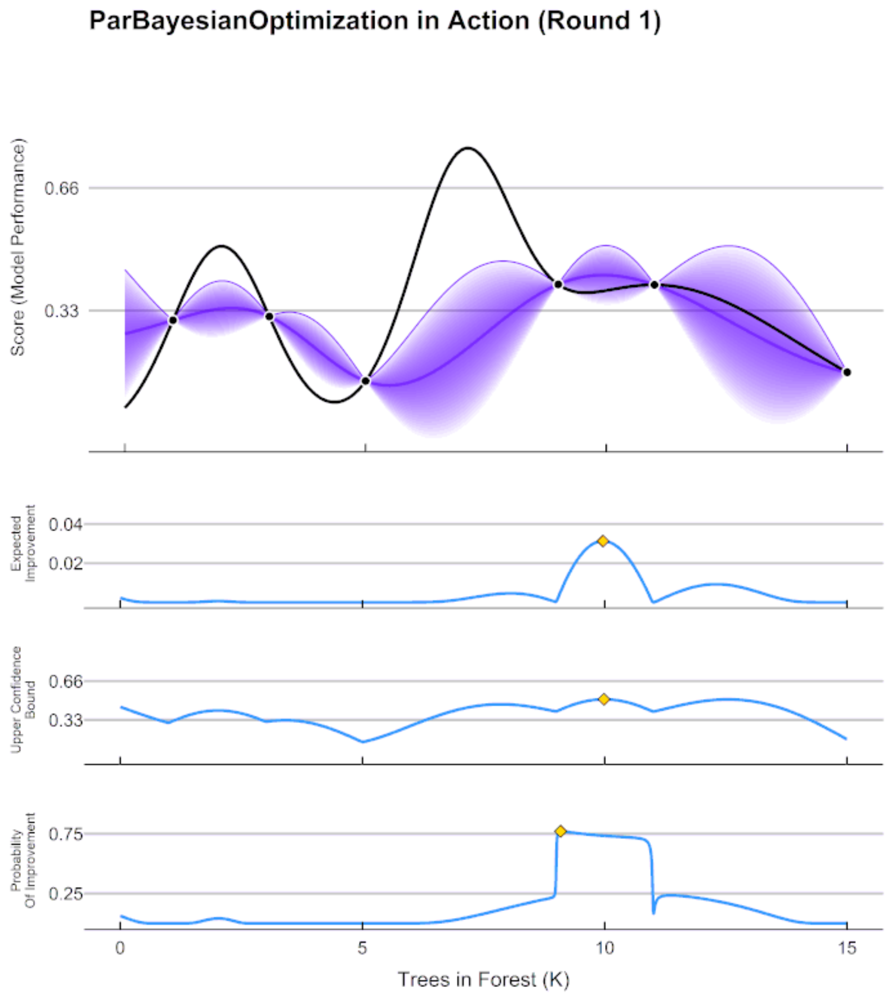
How can we control uncertainty?

Analogy with an optimization problem

- Approximating the integral of a curve can be similar to finding its maximum value: looking at optimization algorithms

➔**Bayesian optimization** is a sequential design strategy for global optimization of black-box functions that does not assume any functional forms. It is usually employed to optimize expensive-to-evaluate functions.

➔Bayesian optimization should allow us to account for uncertainty



Literature reference – Not a novel idea

Probabilistic Integration: A Role in Statistical Computation?[1]

Fran ois-Xavier Briol, Chris J. Oates, Mark Girolami, Michael A. Osborne and Dino Sejdinovic

ELSEVIER

Structural Safety

Volume 106, January 2024, 102409

Parallel Bayesian probabilistic integration for structural reliability analysis with smal failure probabilities

Zhuo Hu [a b], Chao Dang [b], Lei Wang [a], Michael Beer [b c d]

ELSEVIER

Mechanical Systems and Signal Processing

Volume 170, 1 May 2022, 108799

A Bayesian deep learning approach for random vibration analysis of bridges subjected to vehicle dynamic interaction

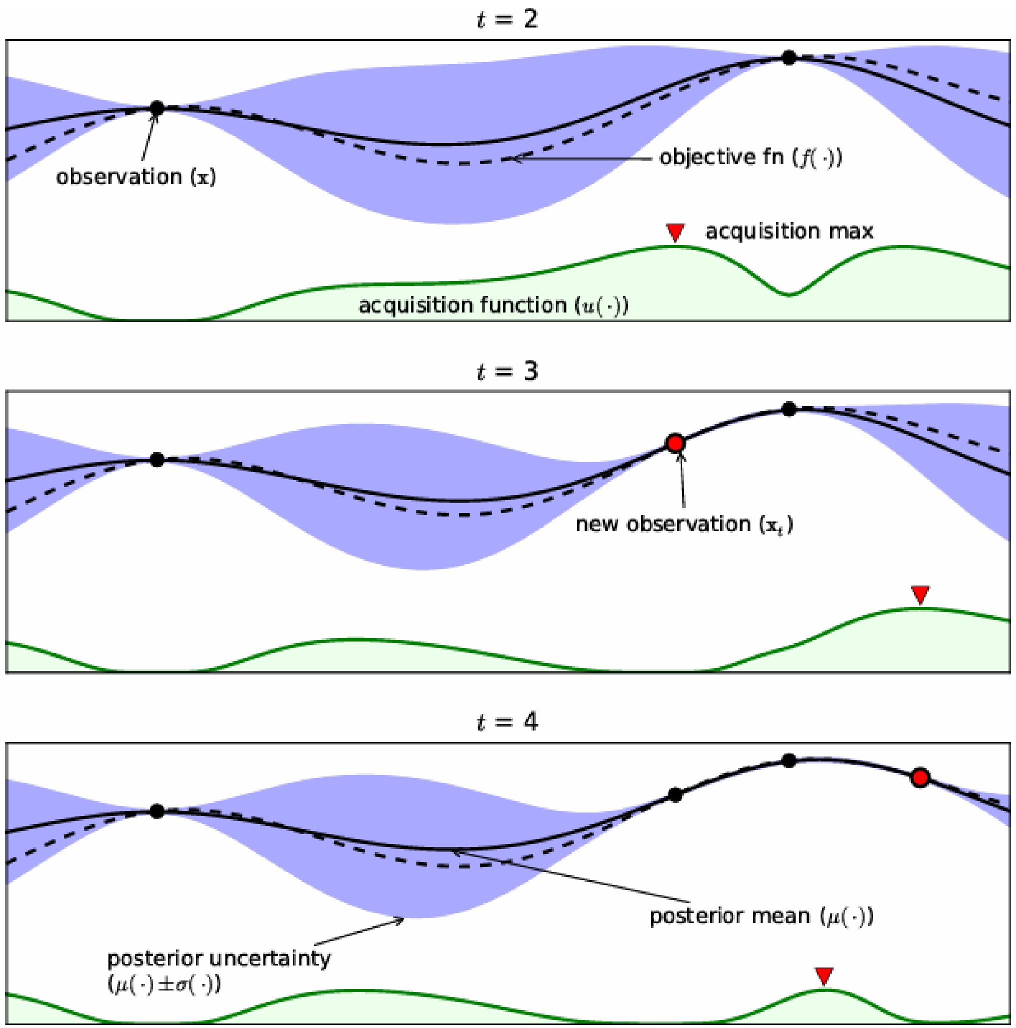
Huile Li [a b], Tianyu Wang [a b], Gang Wu [a b]

*This one is a bit different, they use the Bayesian approach to model the bridge

Is this machine learning or is this probabilities?

- **Bayesian Optimization** is an approach that uses **Bayes Theorem** to direct the search in order to find the minimum or maximum of an objective function.
  - With $\mathbf{x}_i$ the $i-\mathrm{th}$ sample of the objective function $f(\mathbf{x}_i)$
  - We accumulate observations $\mathcal{D}_{1:t} = \{\mathbf{x}_{1:t}, f(\mathbf{x}_i)\}$
  - We then write a likelihood function $P(\mathcal{D}_{1:t}|f)$

    ➔ Given what we think we know about the prior, how likely is the data we have seen?
  - We then define the posterior distribution $P(f|\mathcal{D}_{1:t}) \propto P(\mathcal{D}_{1:t}|f)P(f)$



A Tutorial on Bayesian Optimization of Expensive Cost Functions, with Application to Active User Modeling and Hierarchical Reinforcement Learning. Brochu (2010)

Is this machine learning or is this probabilities?

- We can then rewrite $f(\mathbf{x}_i)$ as a function of a Gaussian Process (GP):

$$f(\mathrm{x}) \sim \mathcal{GP}\big(m(\mathrm{x}), k(\mathrm{x}, \mathrm{x}')\big)$$

with m the mean function, and k the covariance function.

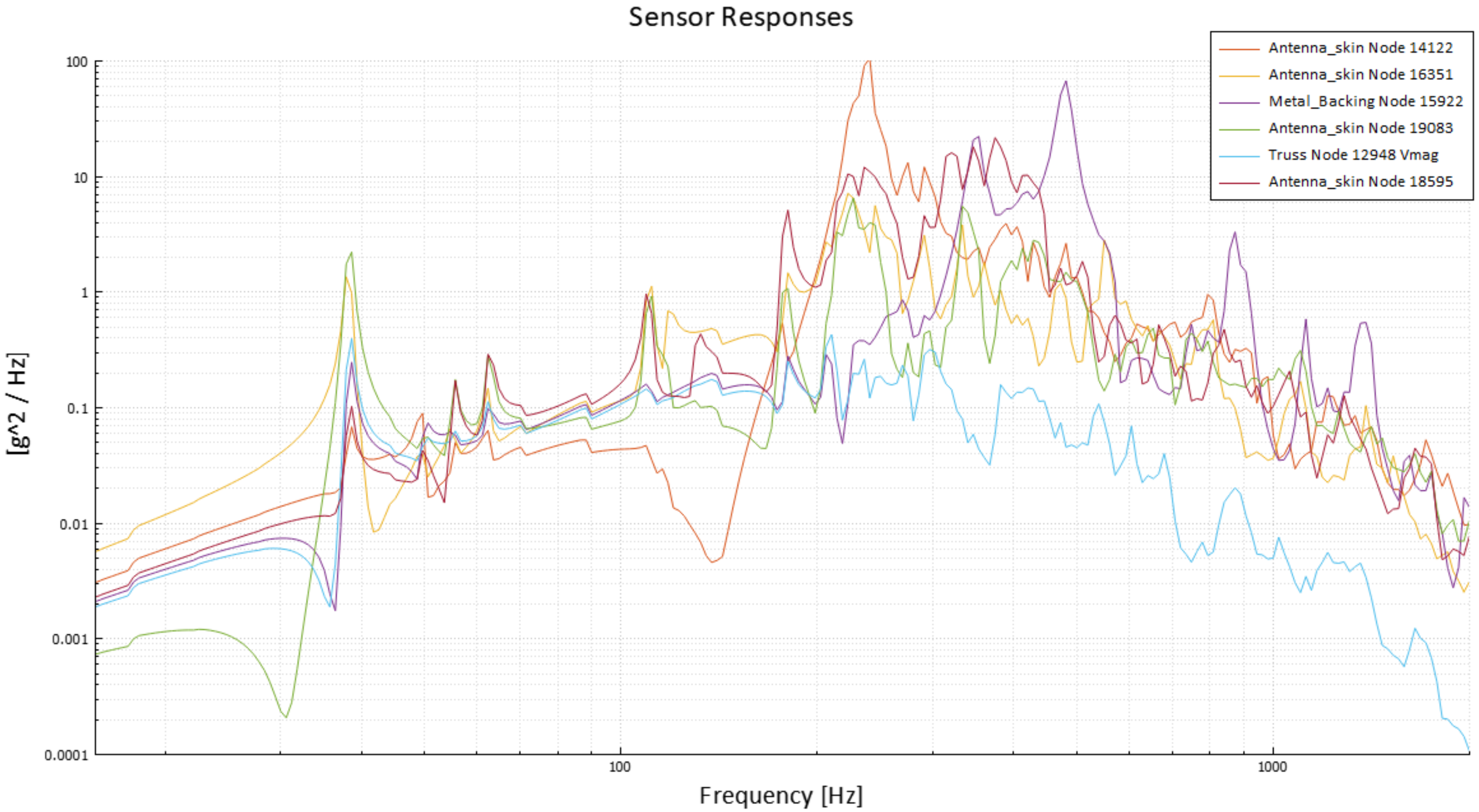
➔ This is called a kernel.

- ➔ A library of different kernels is available in the literature.
- ➔ Kernels may be better suited to represent certain physical phenomena
- ➔ Kernels have parameters
  - ➔ Parameters can be determined through machine learning (i.e., large sets of data).

It is machine learning for a probabilistic model.

Choosing a Kernel

- A well chosen and tuned Kernel will limit the number of sampling points to determine the maximum of a function.



How peaky is the function?

What is a typical trend?

How can we test kernels?

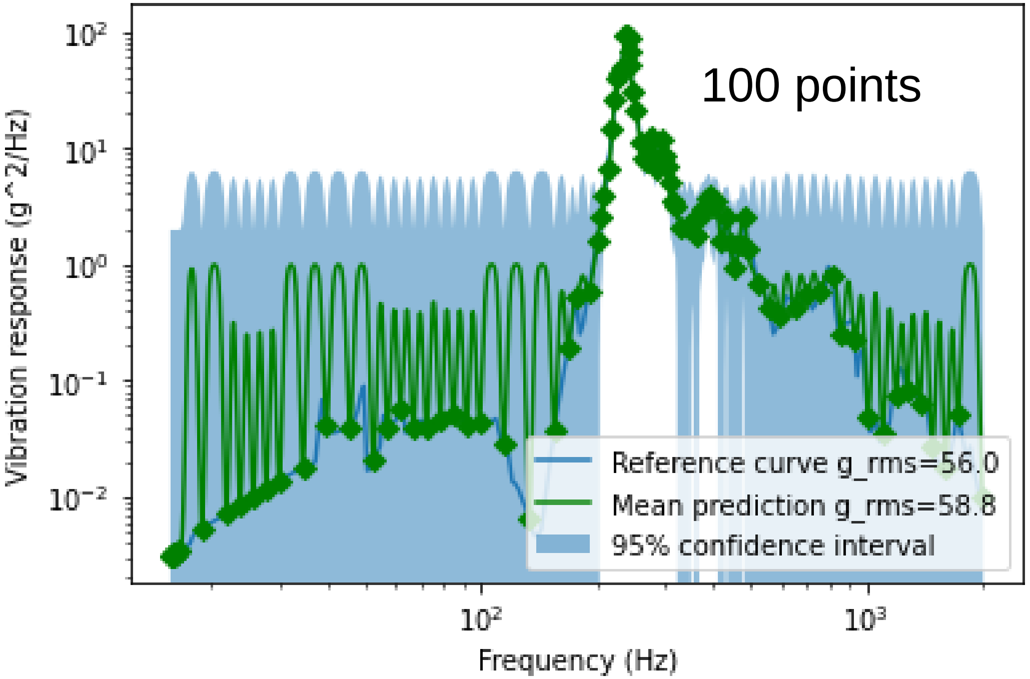
Customizing and trying the algorithm

- Random response are always positive
- Level variations are very drastic (we plot results on a log scale!)

➔Gaussian Process algorithm needs some adaptation

➔Optimization algorithm is set to estimate the overall level as accurately as possible

➔We initially sample n locations to get a global trend of the curve



Here we see the weakness of non-tuned model.

Radial Basis Function kernel (most basic)

$$k(x_i, x_i) = \exp\left(-\frac{d(x_i, x_i)^2}{2l^2}\right)$$

where l is the length scale of the kernel and $d(\cdot,\cdot)$ is the Euclidean distance.

Here $l = 1.0$

No tuning, no learning

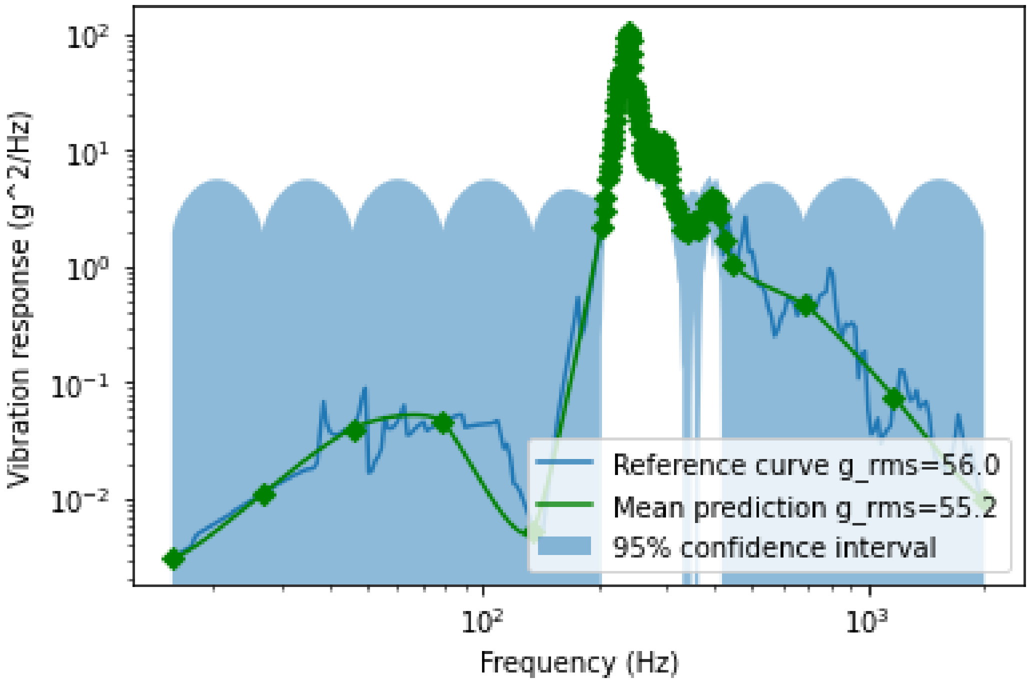
Customizing and trying the algorithm

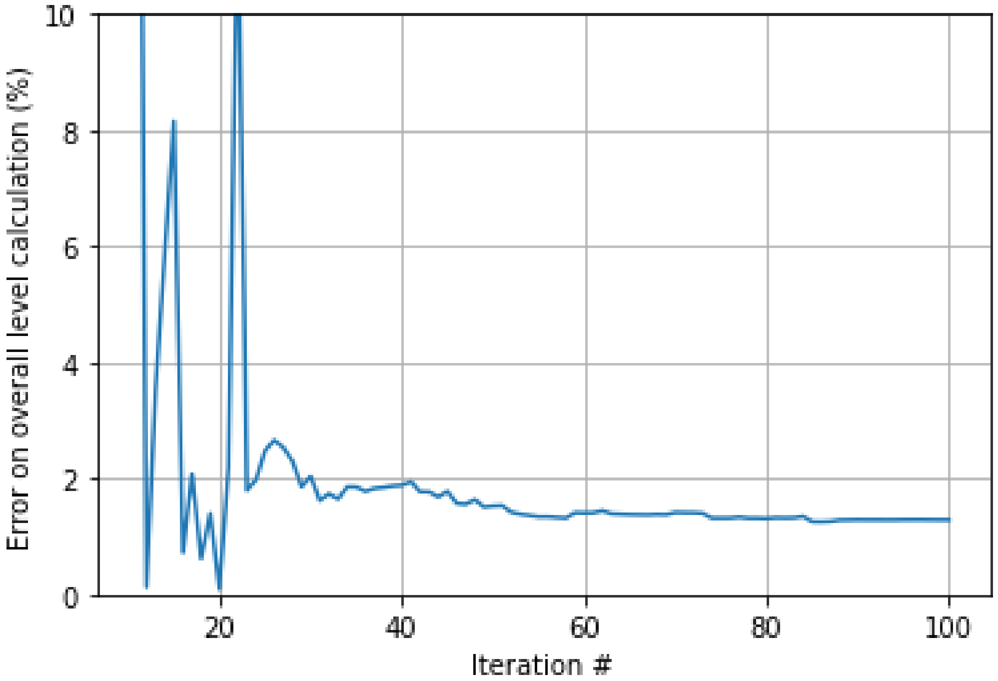
- New kernel for a model allowing for a peaky target function
- The optimization strategy (decision of the next sampled point) in the iterative process
- Increasing the number of initial points

We can the observe the convergence on the overall level estimation:

➔ How many data points to we need to estimate the overall level correctly.

Customizing and trying the algorithm



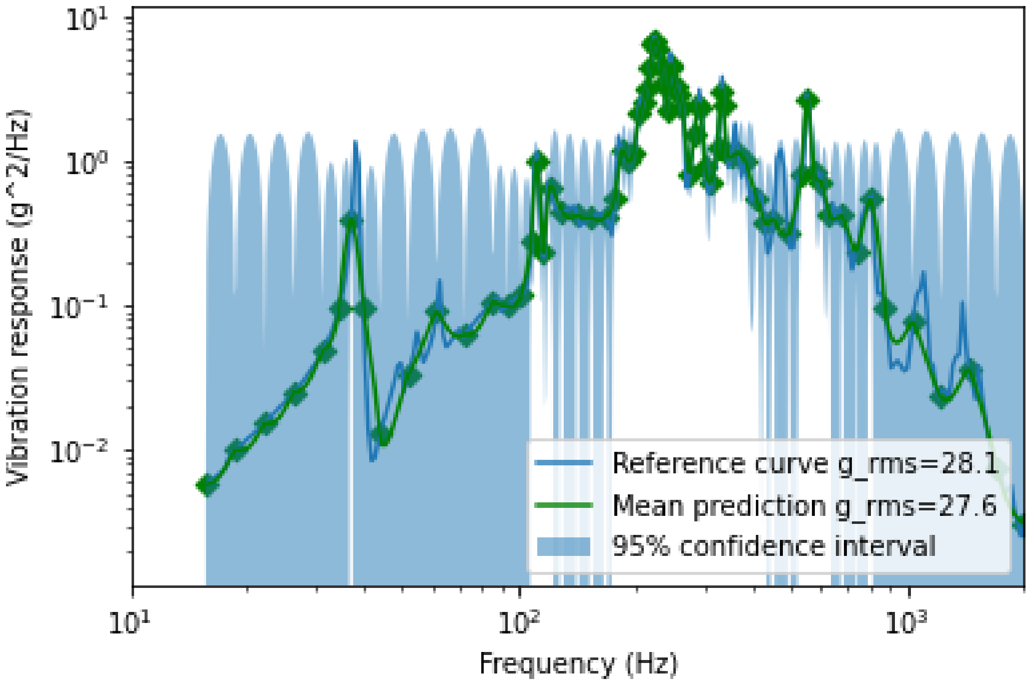


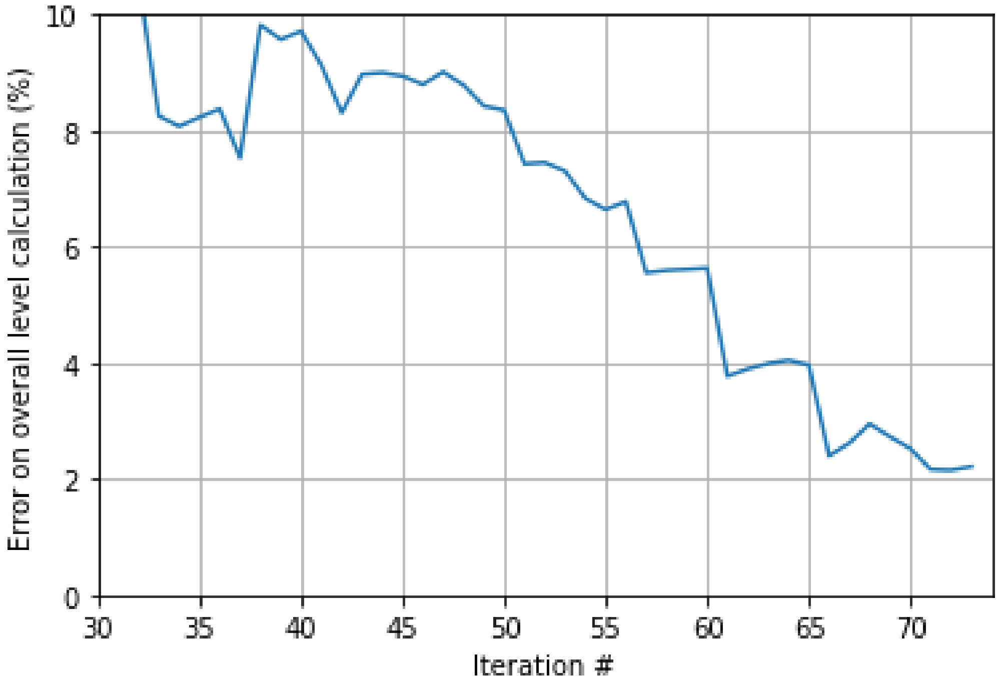
- 10 initial points
- Matern Kernel
- Next point is chosen based on both amplitude and uncertainty.

Error on overall level is less than 2% with 30 frequency points logarithmically spaced.
We observe convergence on the overall level after a while

→ Potential to reduce the number of frequency points by a factor 5x.

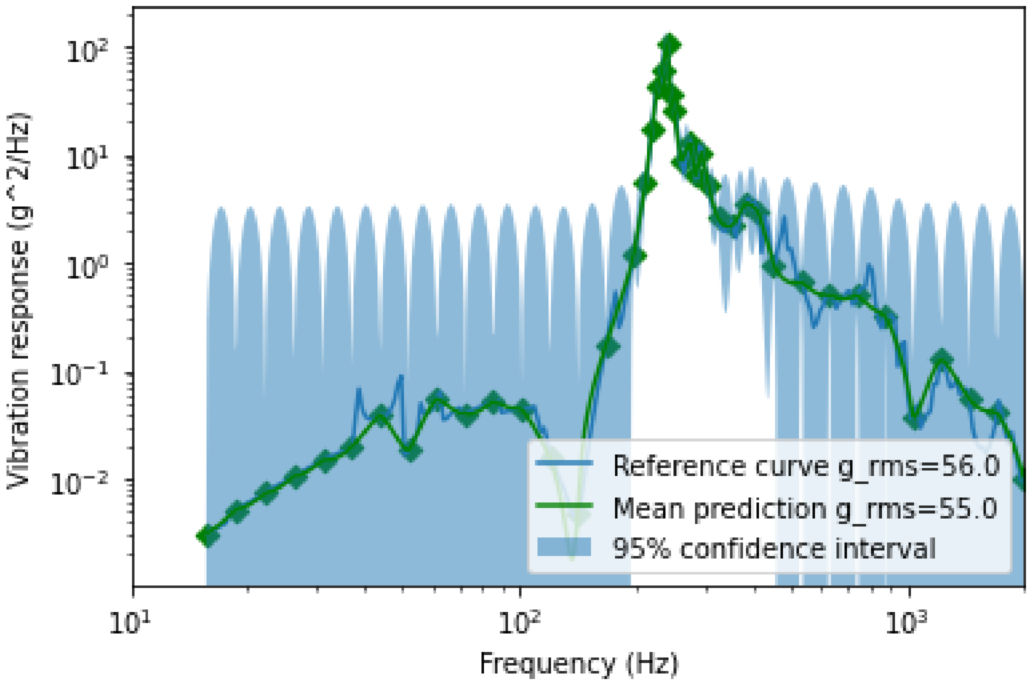
Curve 1

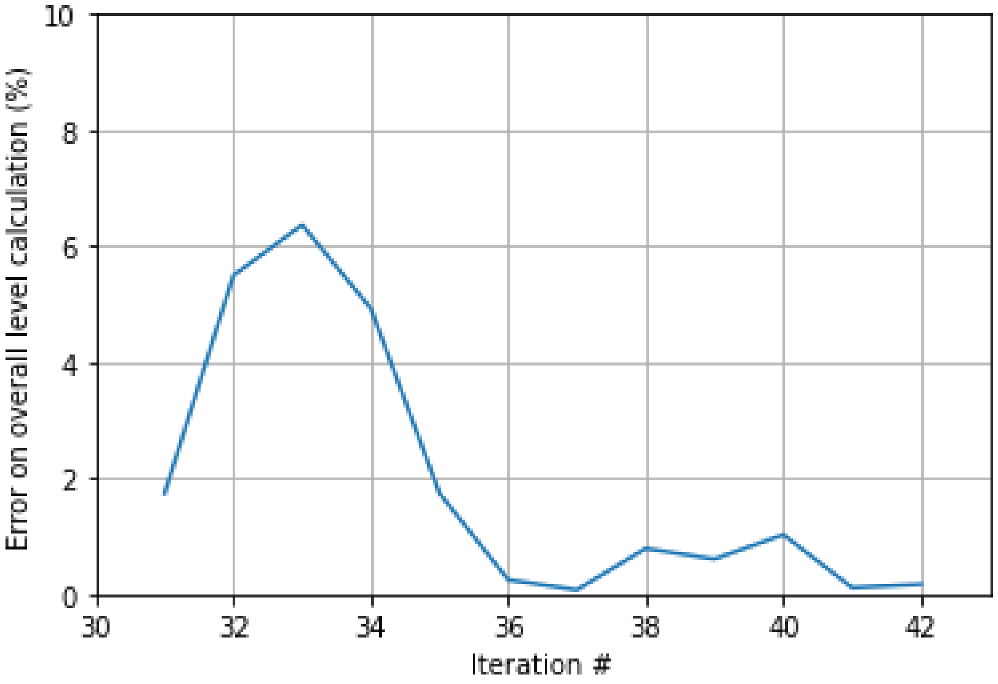




Overall level after 73 iterations: 27.6
Standard deviation: 0.8

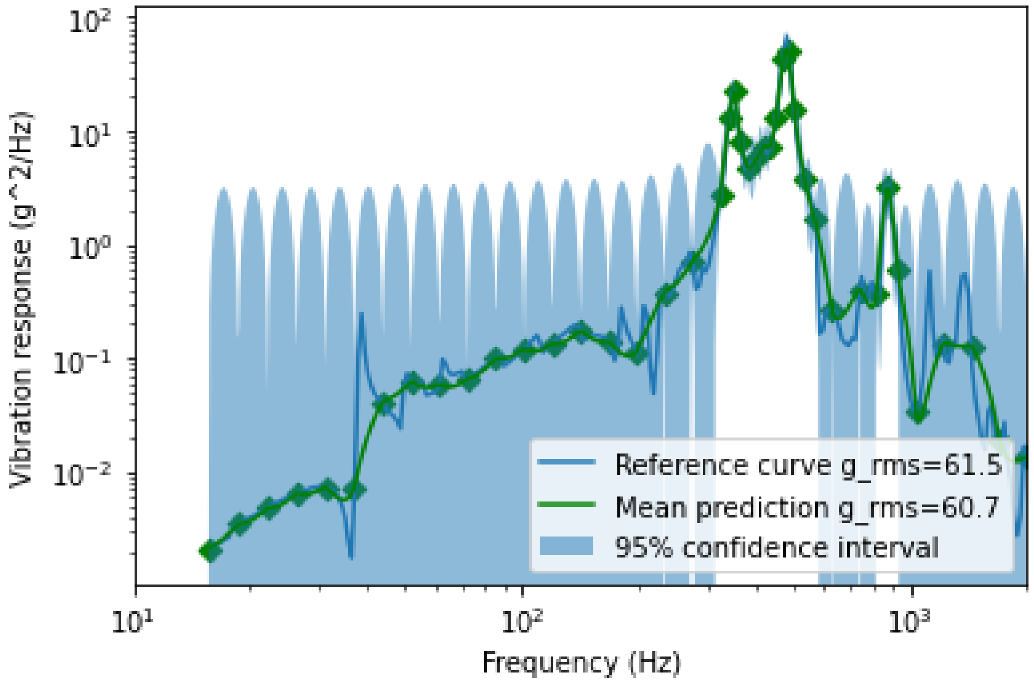
Curve 2

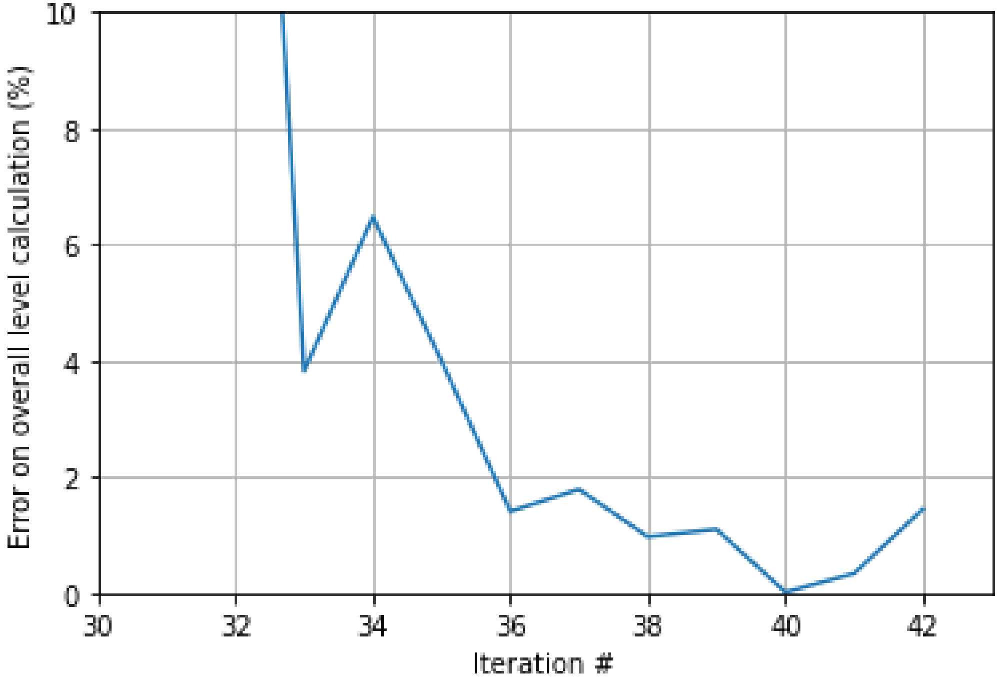




Overall level after 36 iterations: 55.0
Standard deviation: 1.1

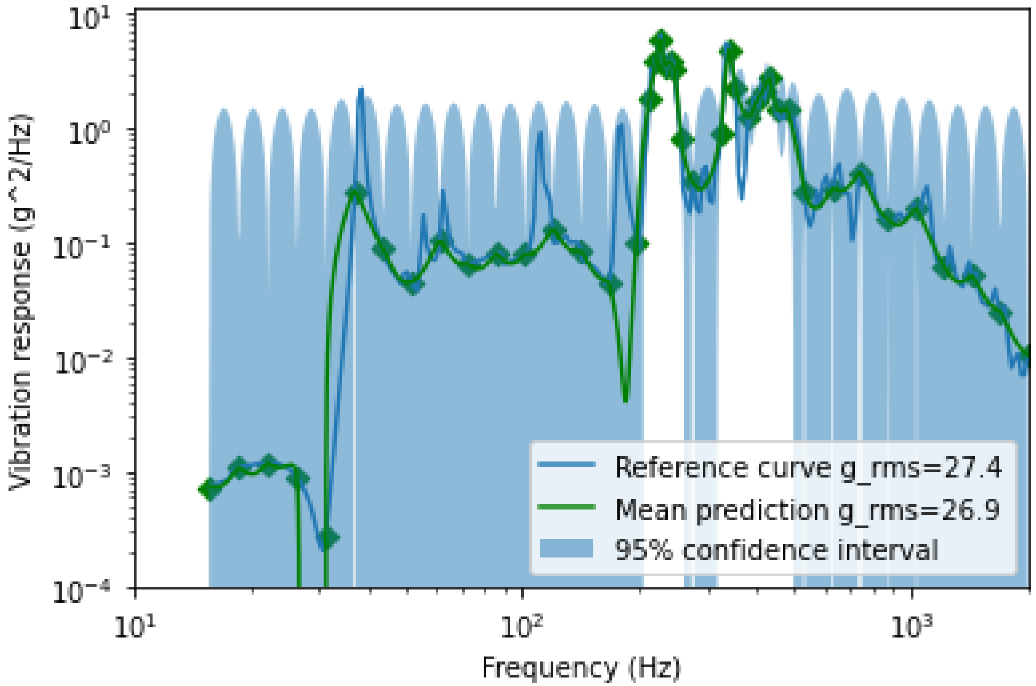
Curve 3

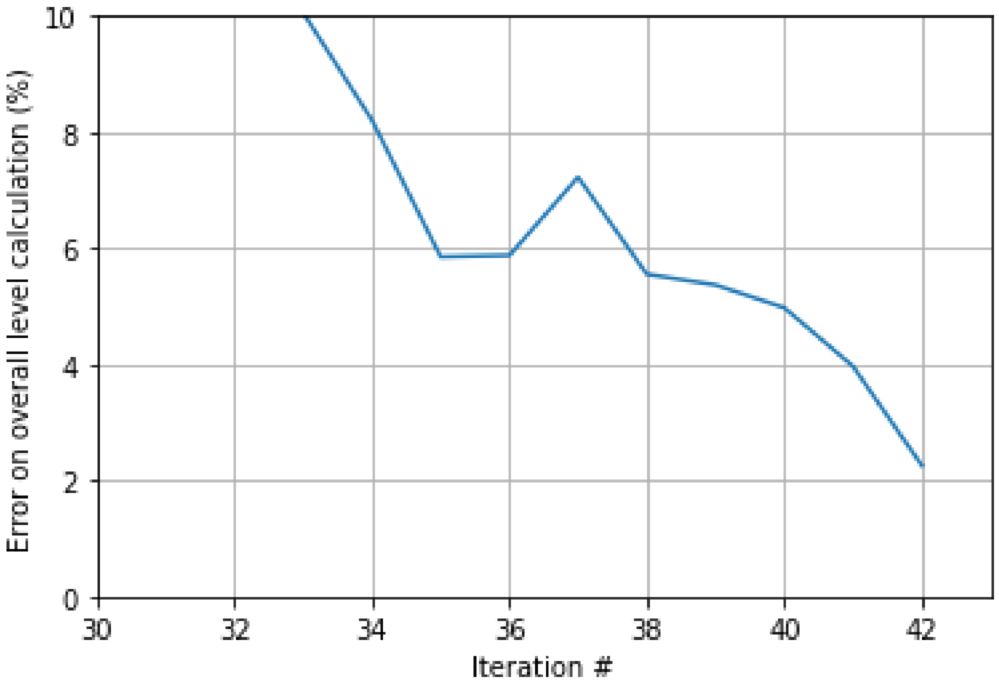




Overall level after 35 iterations: 60.7
Standard deviation: 1.1

Curve 4

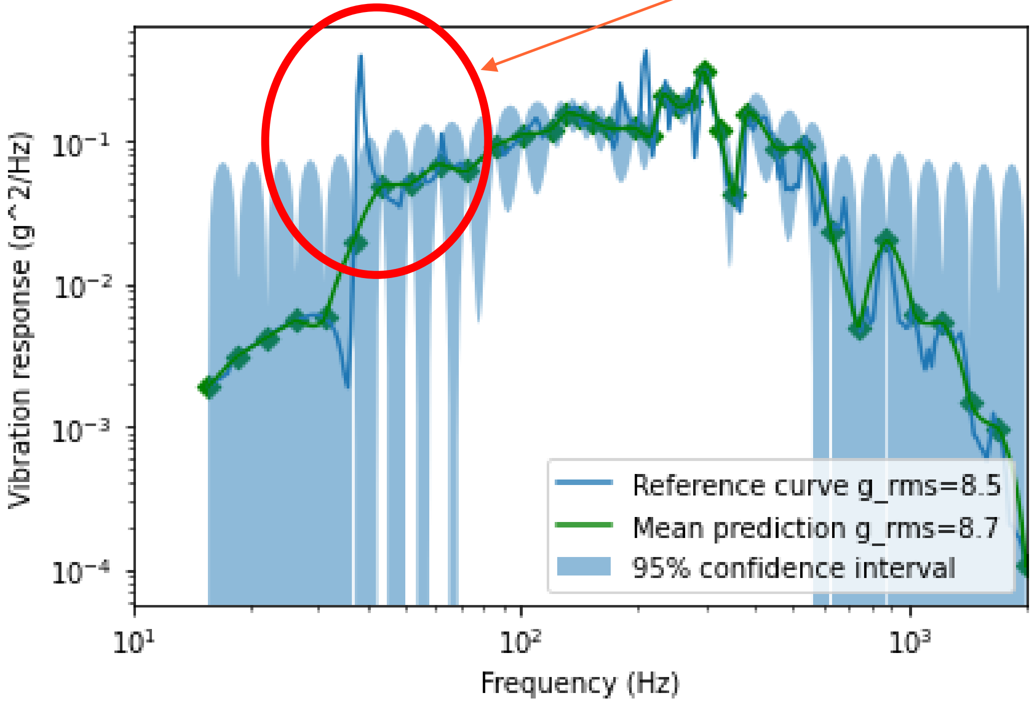


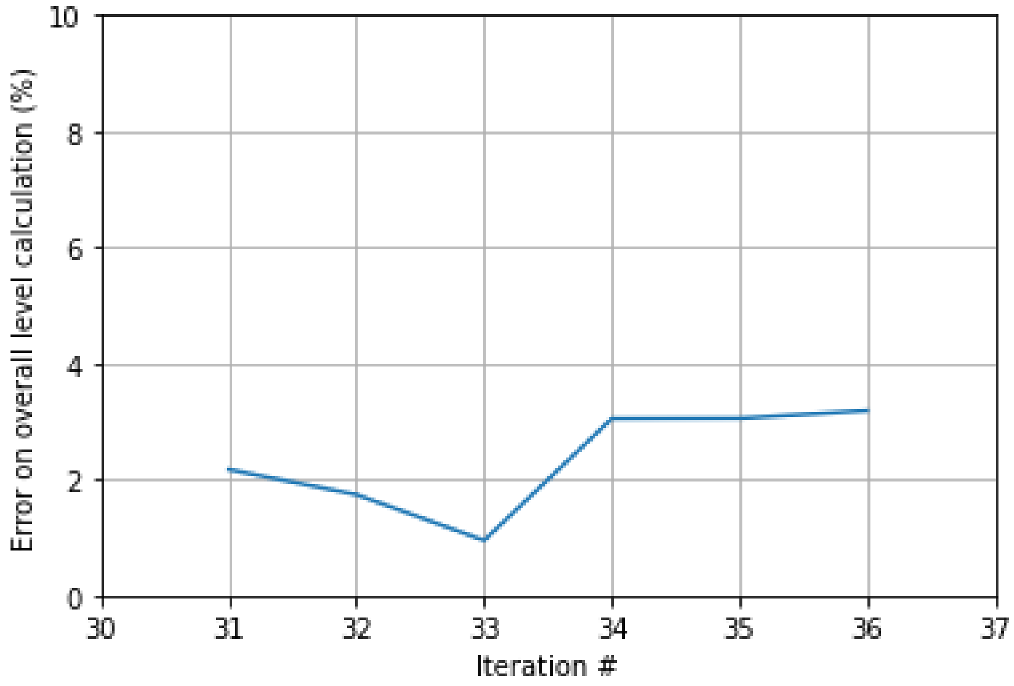


Overall level after 42 iterations: 26.9
Standard deviation: 0.9

Curve 5

Need for a denser initial sampling





Overall level after 36 iterations: 8.7
Standard deviation: 0.2

Conclusions

- A combination of statistics and machine learning can be used to limit the number of frequency points at which structural dynamics models are solved.
    - This has a drastic linear effect on solve time.

- The present model is a combination of generic Bayesian Optimization algorithms and adaptations for the type of data we are working with.

- Additional investigations could enable the selection of a smarter initial set of frequency points. These could be based on the resonant modes of the finite element model and the chosen damping.

- Additionally, the kernel parameters, here tuned based on a reduced dataset could be adjusted with a larger dataset.

Thank you!

esi-group.com

get it right

esi-group.com