https://arxiv.org/pdf/1703.04977.pdf
http://www.ce.memphis.edu/7137/PDFs/Abrahamson/C05.pdf

Total Uncertainty = Data Uncertainty + Model Uncertainty = Aleatoric Uncertainty + Epistemic Uncertainty

Aleatoric Uncertainty - comes from our data

- It is the noise that is in our data
- We cannot lower this noise even if we collect more data
- It is the noise in sensors, e.g. noise present on images
- **Homoscedastic Uncertainty -** stays constant for different inputs
- **Heteroscedastic Uncertainty -** noise that differs between inputs
    - E.g. textured images tend to have more confident predictions than featureless walls

Epistemic Uncertainty - knowledge/model uncertainty in the parameters

- As we get more and more data to train with, the model's epistemic uncertainty is lowered

Simplex Plane - the plane on 3D axes where the length of the norm from the origin to the plane is always 1

Ensemble - a collection of trained models that were trained with a different seed to initialise their weights, that we can use to get sample uncertainty on our data

- Each model in an ensemble is called an **ensemble member**

Learned Attenuation -

Well-Calibrated Uncertainty - it's when a model can predict it's uncertainty in a result

- A lot of models, when wrong, are overly confident in their incorrect solution
    - This can happen irrespective of the accuracy of the model
- A model with well-calibrated uncertainty provides a confidence matrix which lines up with the model's accuracy
- This is particularly useful for safety-critical tasks