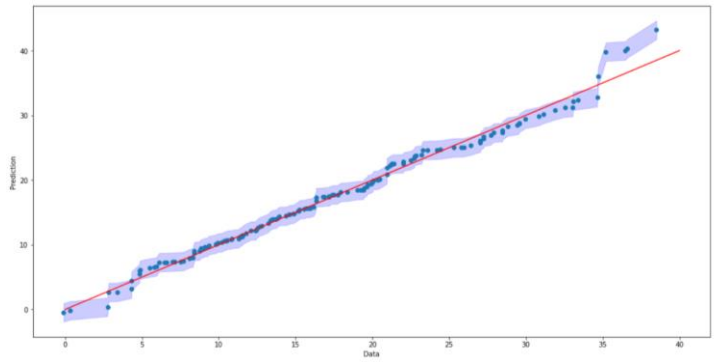
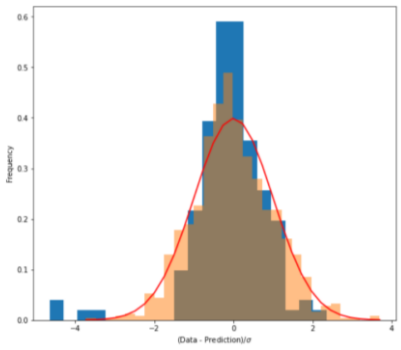
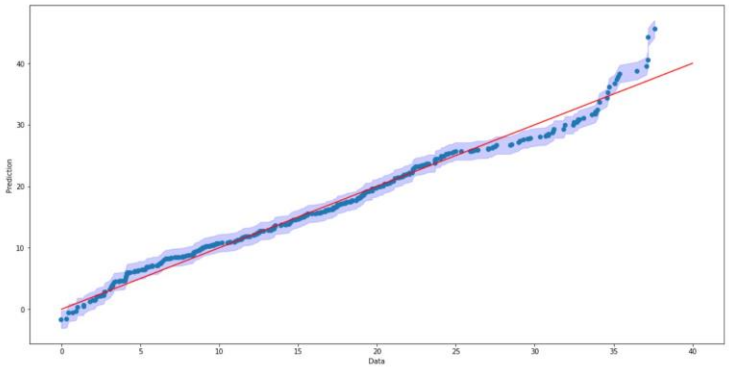
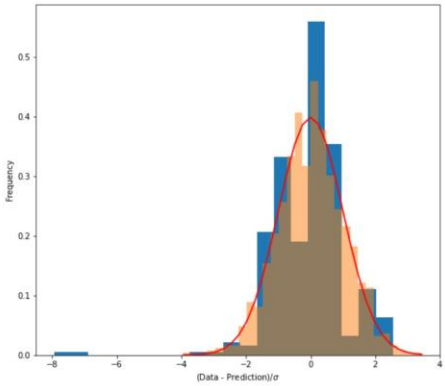
Semester 2 Week 1&2 Results



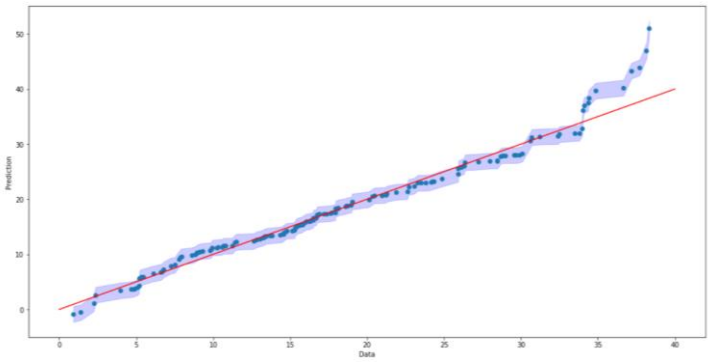


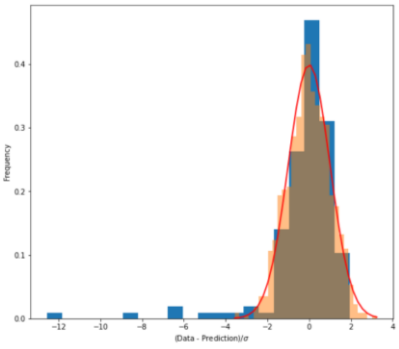
Figures 1 (a) & (b): Run with 500 samples for 289 data points



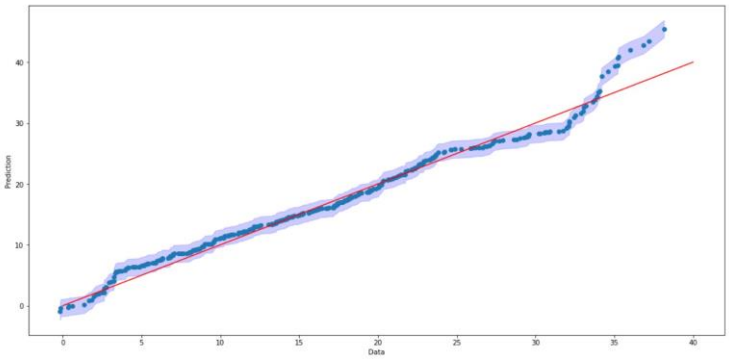


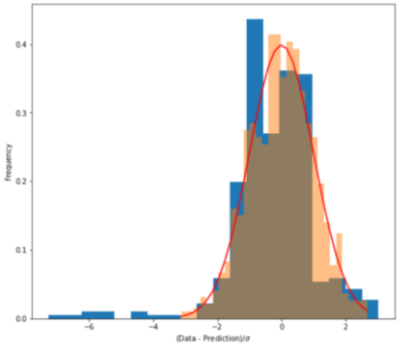
Figures 3 (a) & (b): Run with 500 samples for 722 data points



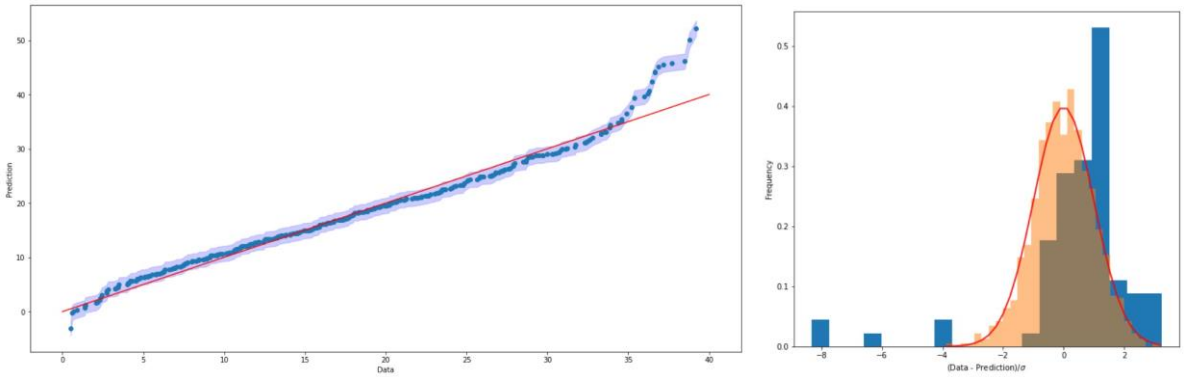


Figures 3 (a) & (b): Run with 500 samples for 293 data points





Figures 4 (a) & (b): Run with 500 samples for 719 data points



Figures 5 (a) & (b): Run with 500 samples for 718 data points

Notes one results:

The runs were performed on randomly selected date points form 'Data 1.csv' . All runs were satisfactory creating a Normal distribution for the z-value with the exception of 'Figure 5' which seemed to be centred at 1. As the data were randomly selected it's not possible repeat the run and understand what the problem was. It was labelled as a 'bad run' as the problem was not repeated.

Autocorrelation of Samples:

Precautions have been takes to avoid autocorrelation of samples by introducing 'leap frog' steps which the samples takes three steps between recorded steps. However, a test was run to verify that the samples were not autocorrelated. An example diagram is shown below from the run of Figure 4.

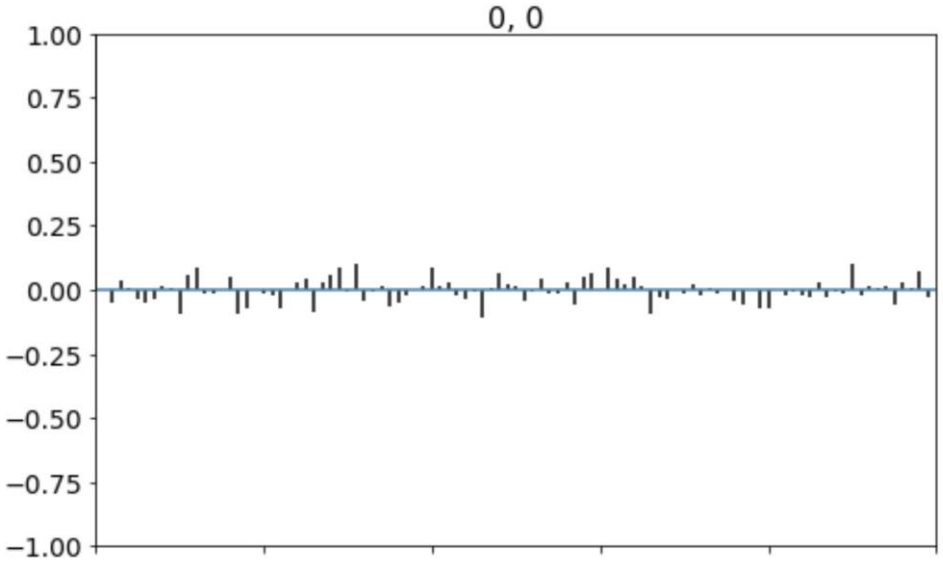
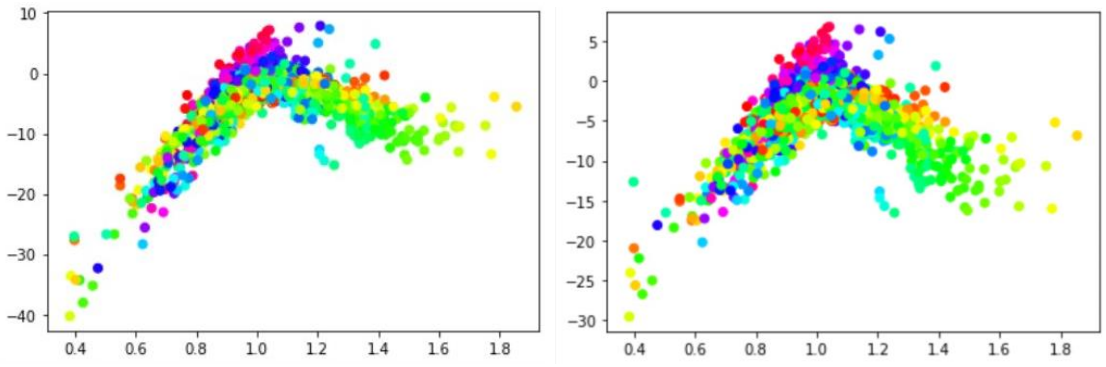


Figure 5: Example autocorrelation between samples.

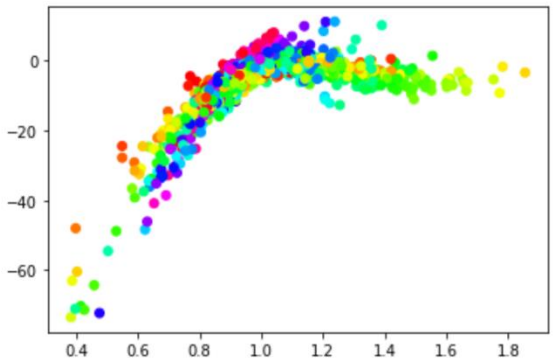
The diagram was created using Arviz library.

Mean function:

Inconsistencies on the power dependence of mass in the mean function between me and Henry (values of 0,3 and -0,6) has lead to further investigation and decided that the mean function does not need to include the mass. Figure 6 shows the data – mean function graph.



(a) (b)



(c)

Figure 6: Data – Mean Function (a) No mass dependence, (b) Mean function prop m^3, (c) Mean function proportion to $m^{-0.6}$

Using m^3 has very little improvement from the Barnes 2007 mean function (no mass dependence) as the extreme values as less extreme (close to 0). The $m^{-0.6}$ model has more points close to zero but also has higher extreme values. Gaussian Processes work best when the values re close to zero with less extremes, therefore there is no point in introducing a dependence on mass.

Notes for next week:

- Apply the model on real data
- If time permits: Revisit the Neural Network idea for the mean function
    - Specifically: why did it not work lass time
    - Possible: introduce higher dimensions
    - Reason: CNN don't give good uncertainties like GPs do (they tend to overestimate the quality of their results). However a basic model for mean function might be a better performer than a deterministic one.