**Case Study 4: Measuring Snow Gauge**

Aditya Anandkumar

Andrea Cabezas

Bethan Lily Wynne-Cattanach

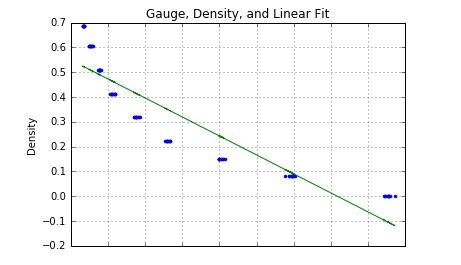
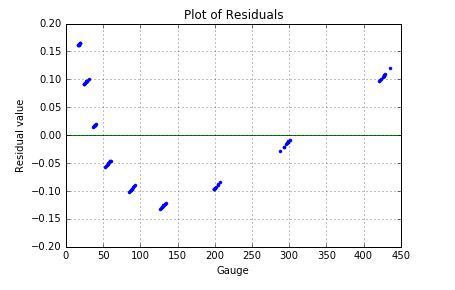
Carlo Mazzafaro

Jake Ehlers

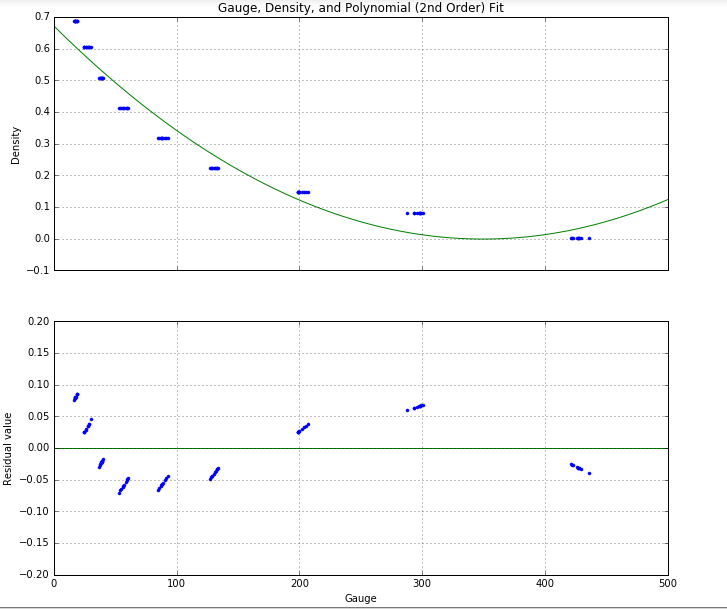
Yi Ma

**Introduction**In this study, we aim to provide a simple procedure for calibrating a gamma transmission snow gauge. The gauge does not directly measure the density of the snow, instead a measurement of the gamma ray emissions is taken and converted to density. This conversion must take into account factors such as instrument wear and radioactive source decay; hence at the beginning of each season the gauge is recalibrated. To investigate a procedure for the calibration, we performed a linear regression with least squares on the experimental data.

**Data and Method**We were given data from an experiment that uses polyethylene of known densities to simulate the snow. For each of the 9 different densities we were given 10 measurements of the amplified gamma photon count, which is referred to as the gain. First, we attempted a linear fit to the raw data.

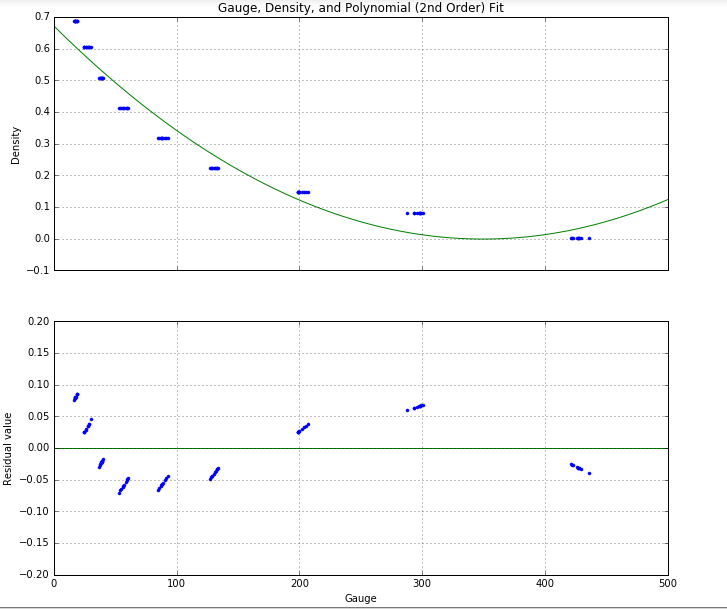


Our results, observed above, indicated that our data didn’t fit a linear distribution and hence we performed a polynomial (2nd Order Fit), shown below, and observed that our residuals didn’t deviate as much as in the former case.



We then applied a logarithmic transformation to both sides of the equation and the results from it were more linear. Hence, we used this fit to predict the gain.

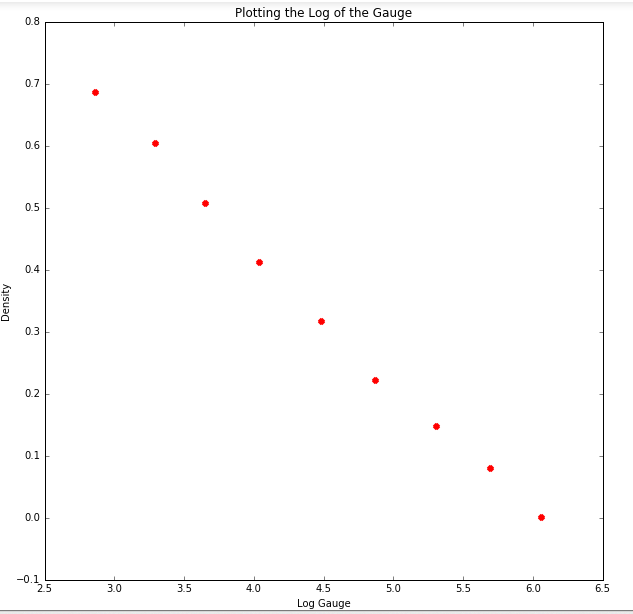
**Fitting**We first attempted a linear fit on the original data, that can be seen on the first page. We did not expect this fit to perform well given the distribution of the points. From the linear fit and the plot of residuals, we concluded that a linear fit wasn’t appropriate. A second order polynomial was a much better fit to the raw data.



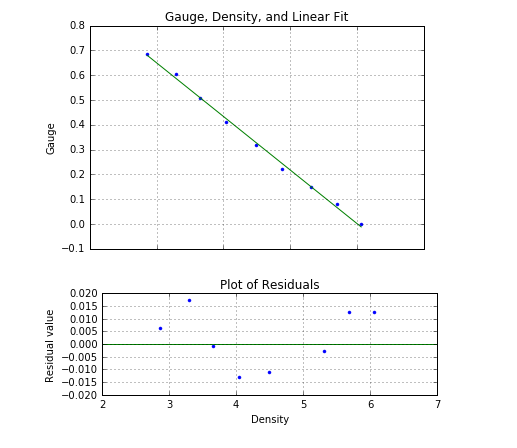
Since this fit was also unsuccessful at providing a linear approximation, we noticed that the probability of a gamma ray successfully arriving at the detector, was pm.

Where pm=em log(p)=ebx (equation1), p is the chance that a single molecule neither absorb nor bounces the gamma ray, m is the number of molecules in a straight line path from the source to the detector and x is the density, proportional to m. [refer to the derivation].

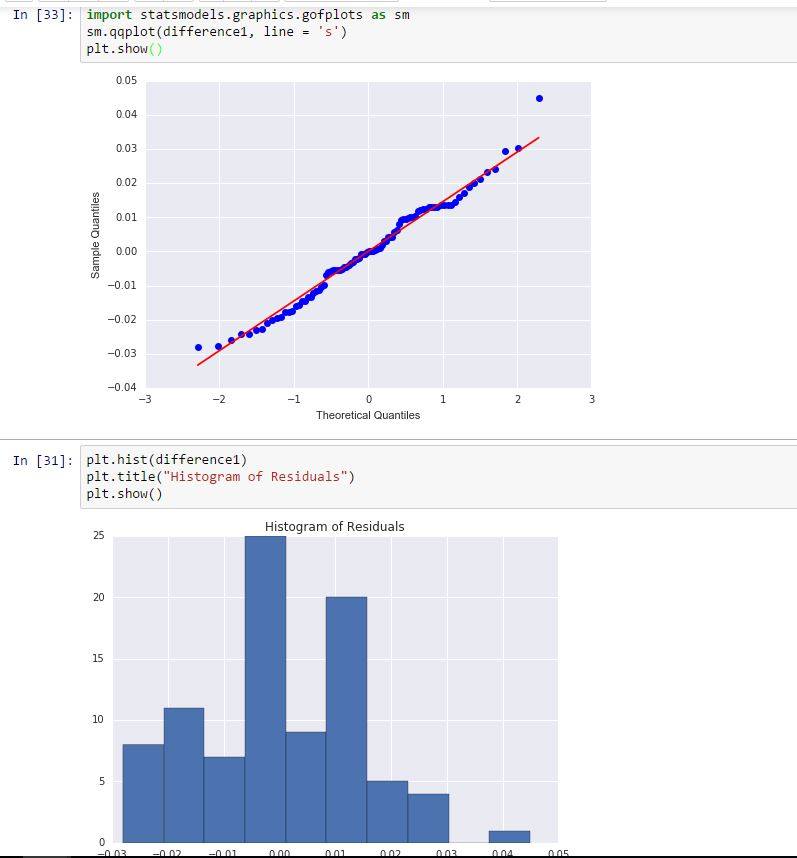
Since the log of the expectation of pm was linear, we fitted it to Y=a+bx+E where Y is the expectation, a is a constant a E is the error.   
We plotted the log of the gain, and as observed below, it was much closer to being linear.



To verify this we fit the least squares line to the log of the data, and also plotted the residuals.

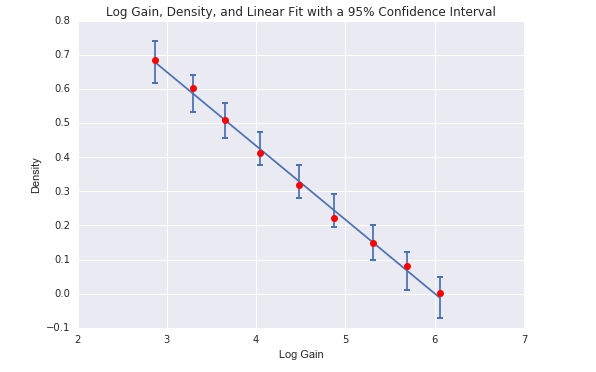


The residuals did not follow a normal distribution as demonstrated



Despite our residuals not following a normal distribution, our residuals were between -0.02 and 0.02 which show that our fit was close to being linear. If the densities of the polyethylene were not reported exactly, the distribution of the gain is broadened, and the points would appear to be more dispersed, thereby affecting the correlation.

If the blocks of polyethylene were not measured in random order this wouldn’t affect the model fit, because the relationship between the gain and density would remain the same.

**Predicting**The goal of the experiment was to create a model that could describe the relationship between gain measurements and density of the material between the sensors. In order to do so, blocks of known density were placed between the sensors and their gauge reading was taken, and the relationship between gain and density was estimated. Given the physical model described in equation 1, it is possible to create a model capable of predicting the density of the material between the sensor solely from a gauge reading. The relationship between the gauge and density can be linearized by applying a log function (again, shown in equation 1). The prediction of the density can thus be done by simply determining the parameters (slope and intercept) of the linear function that describes their relationship, and plugging in the gauge value. Mathematically, the linear function: Y = mx + c + E, describes such relationship. Y is the unknown density of the material, x is the gauge value, while m and c are the slope and intercept of the line that describes the relationship. In this case, given a reading and trained c and m values, the density Y can be estimated with error E (derivation of such value is described in the appendix).

**Cross Validation**To check how well the model actually fits the data, we omitted the data from the 7th measurement (with an average reading of 3.6519561 = log(38.6) ). This is justified since the data still maintains its log-like relationship, and also since this relationship is a function of the mode, not of the data itself. The estimate for the density was 0.509039, since the unbiased estimate for the error for this measurement is approximately 0.051196 and the true value, 0.508, lies within our error range.

We then repeated for the set of measurements at the 0.001 density for the reading with measurement average 6.05608 = log(426.7). The estimate for this density was -0.019217 with an error of 0.060254. This estimate is incorrect because it is negative, however the true value is still within the error range from this estimate.

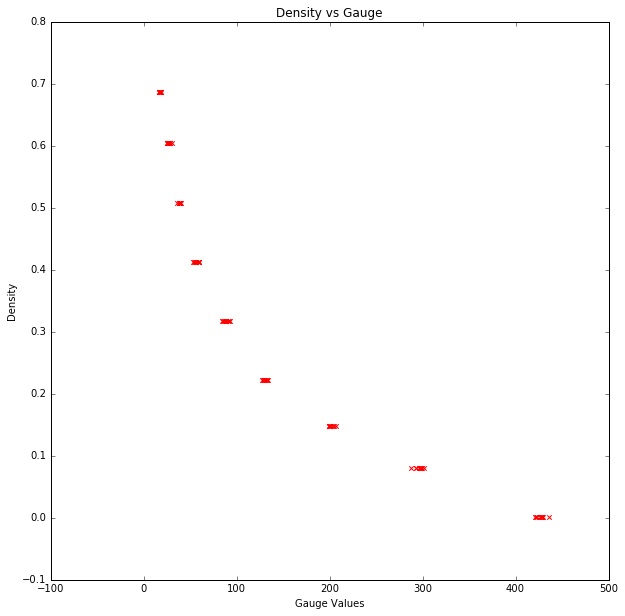
Both of our estimates suggest that our procedure is successful in this case.

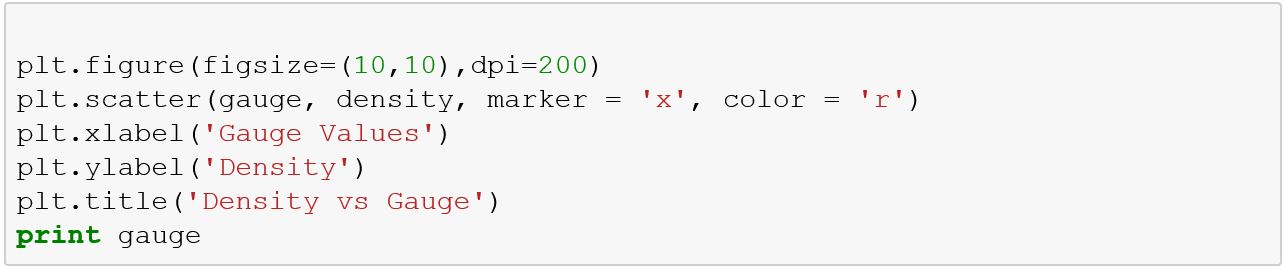
**Conclusion**

We found that it is possible to find a strong relationship between the gain and density which can be made linear by applying a logarithmic transformation, hence giving us a model that calculates the density of snow when gain is the only input.

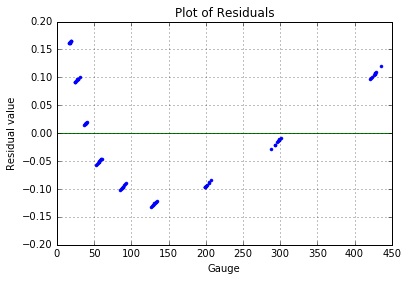
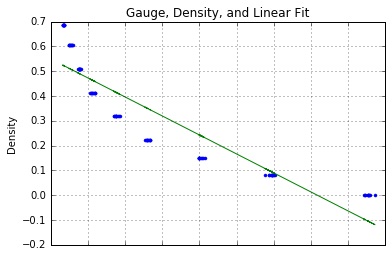
From our examples in the Cross Validation section we can see that our model is accurate given a small error margin. This is the procedure that should be used when calibrating the gauge.

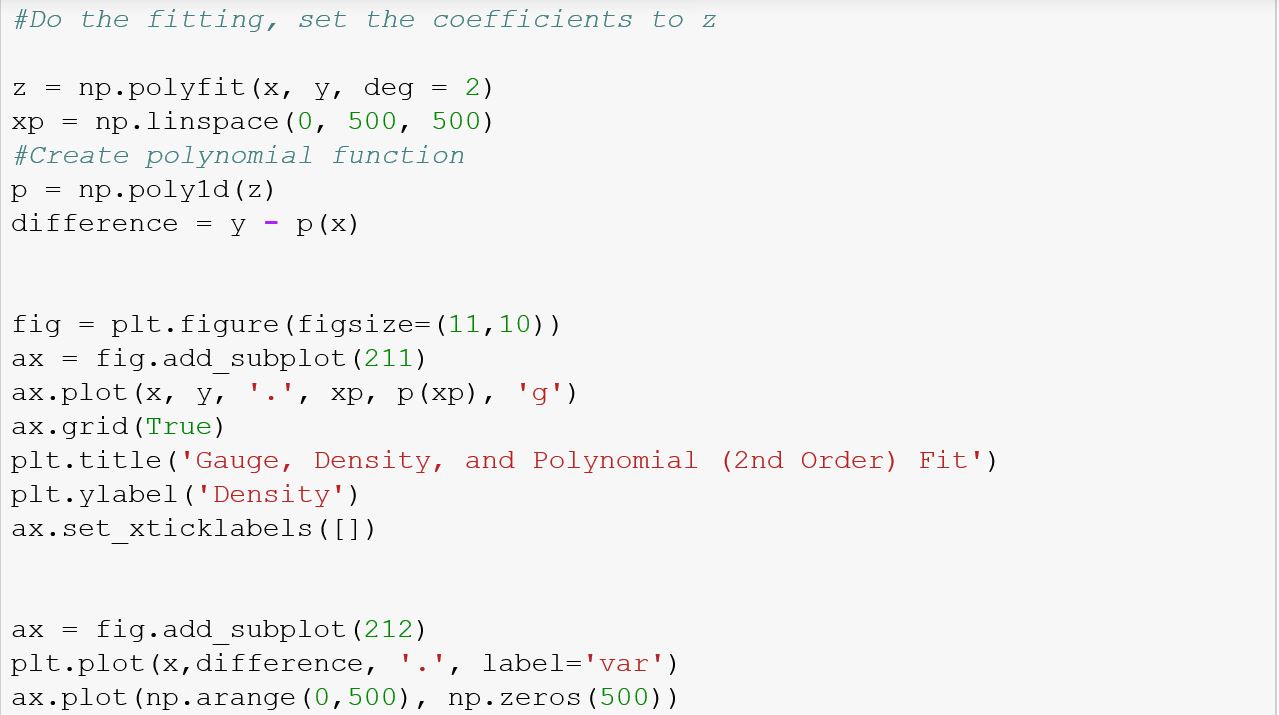
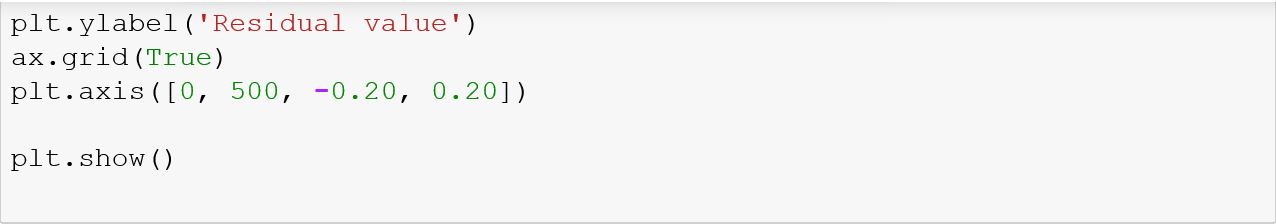
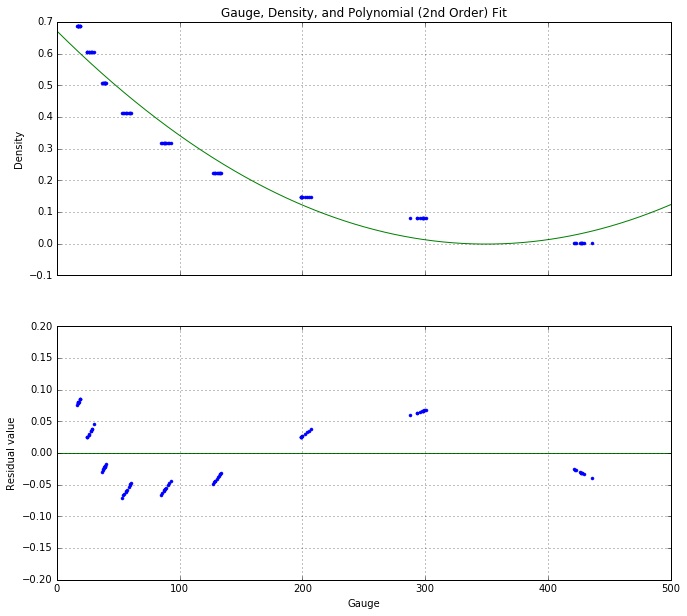
**Appendix**

1. A plot of our data points – density versus gauge.



We proceeded to attempt a linear fit, although we expected poor results given the distribution above.



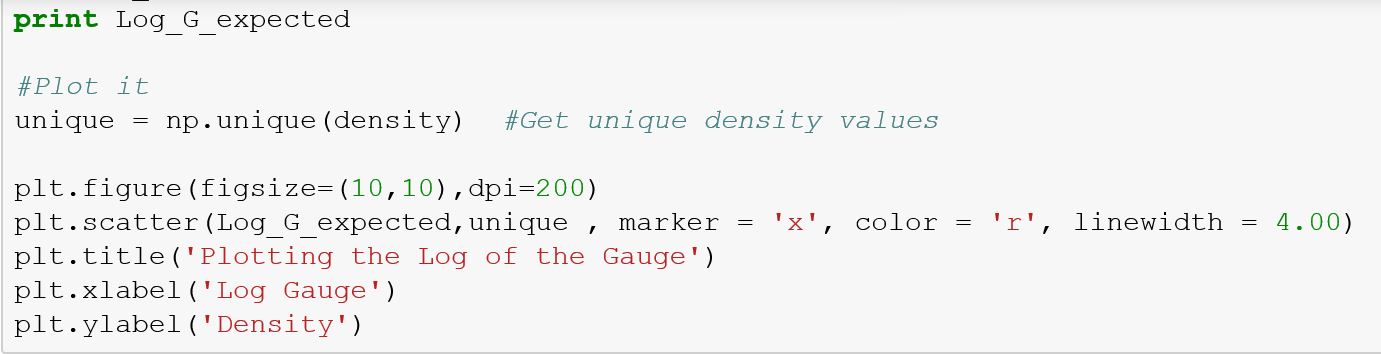


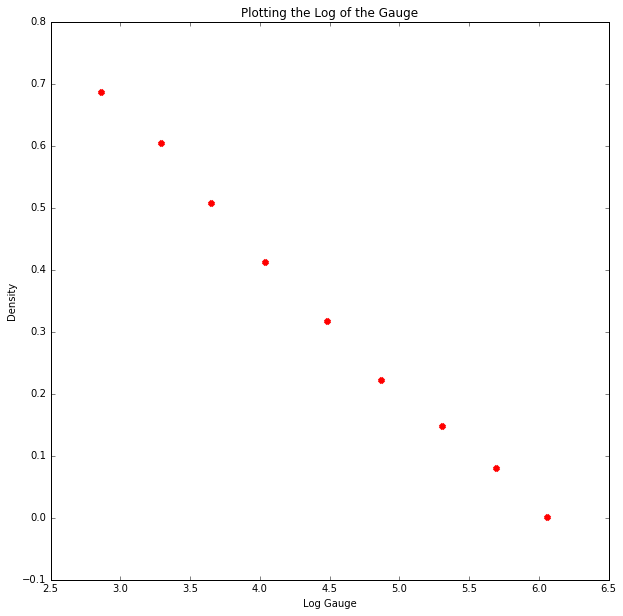
From em log (p) = ebx : equation-1, the density can be estimated as the probability of the particles going through the media (snow or ployethylene material), given by pm. Since we see that the relationship with the gauge x is not linear in the given equation, we apply log to each side.

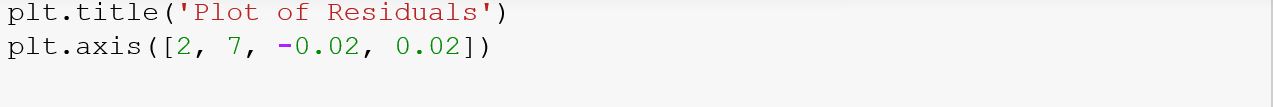
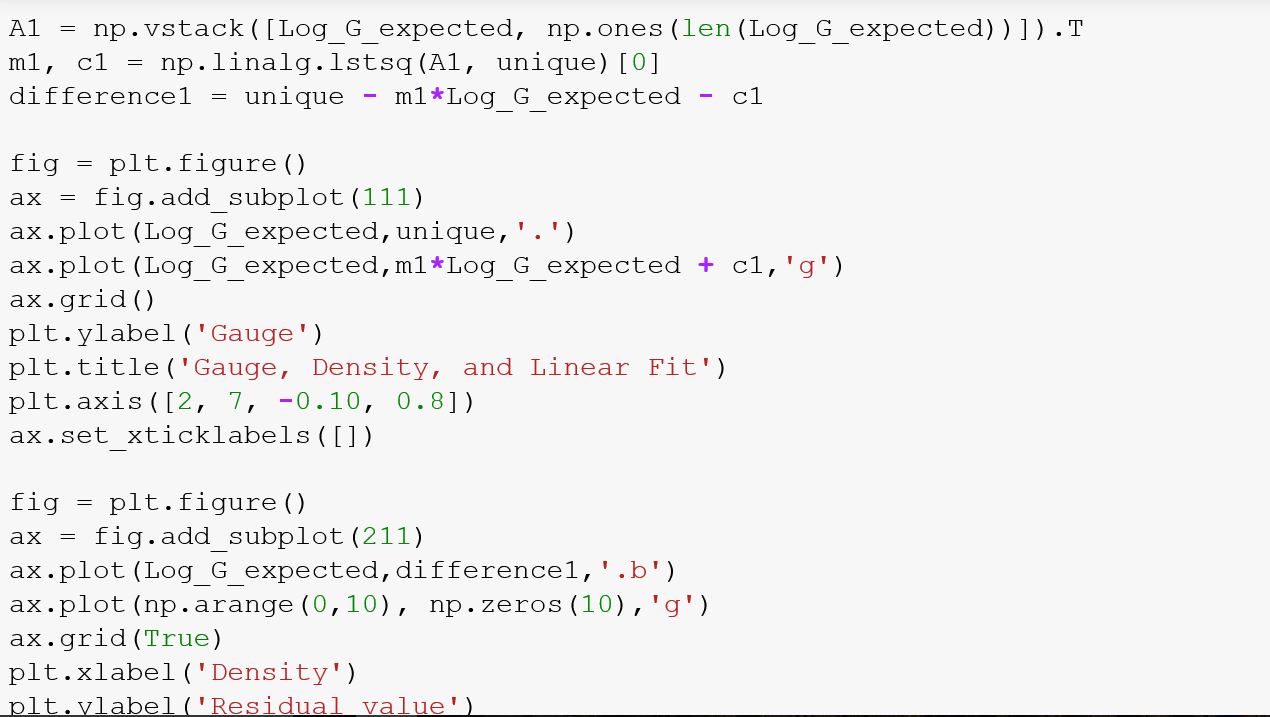
first, we note that equation-1 can be rewritten as:

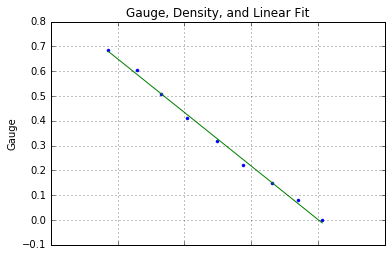
pm = ebx. Let pm be the Gauge (G). The expectation of G, E[G] = c\*ebx+ E, for some constant c and error E. To linearize, we apply: log(E(G)) = log(c) + bx + log(E), which is linear.

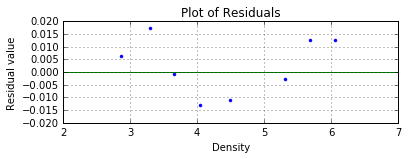
We will fit this to: Y = a + bx + E where Y is log(E(G)), a is a constant and E is the log of the error.

1. We then devised an algorithm to calculate the expectation of the gauge and its error. We will reshape the data so that we can get a 9X10 matrix containing each of the 10 measurements in each of the 9 columns.



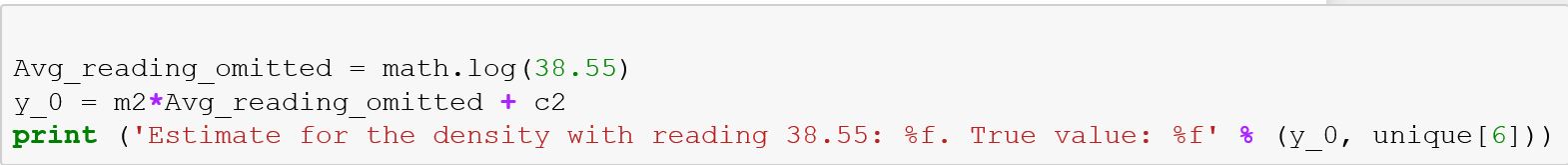
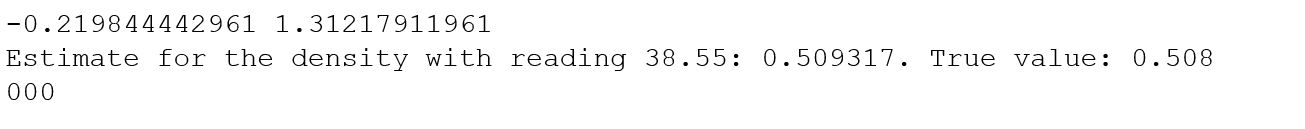


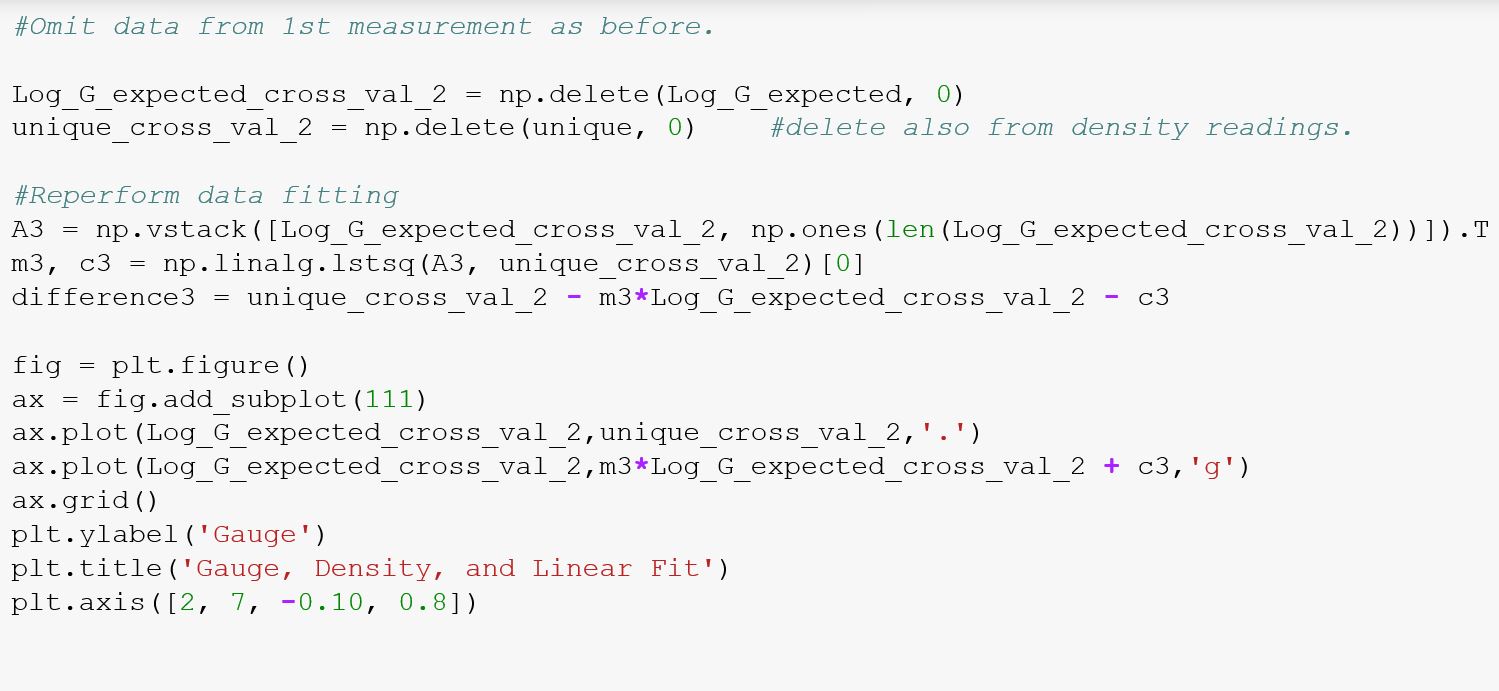




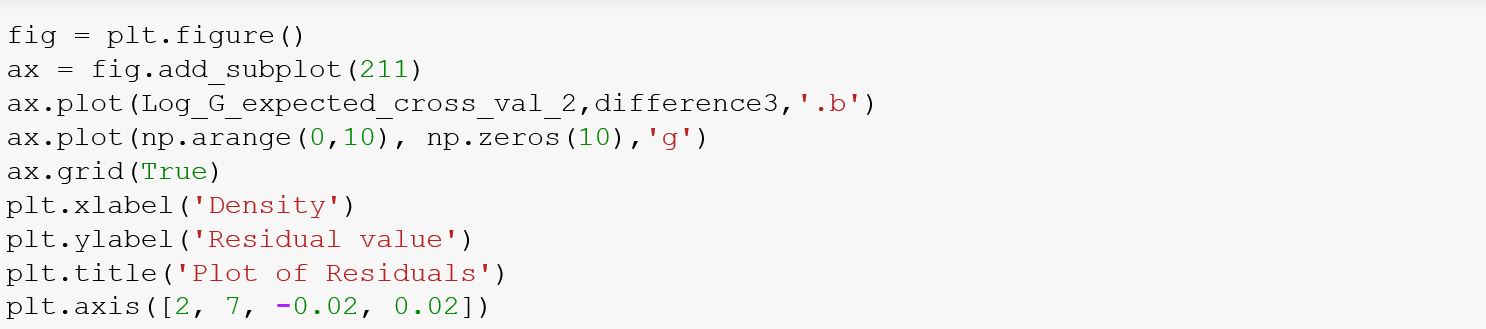
3) Our prediction can now be carried out as follows:

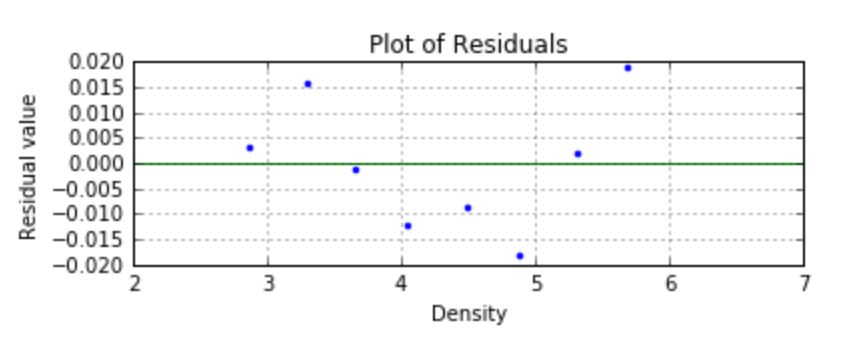
1. We have now a linear fit within reasonable errors
2. Model follows the equation Y = mx + c + E where m and c where estimated, and E is the error
3. We can now predict y by simply plugging in the values for the log gain and checking the response y.

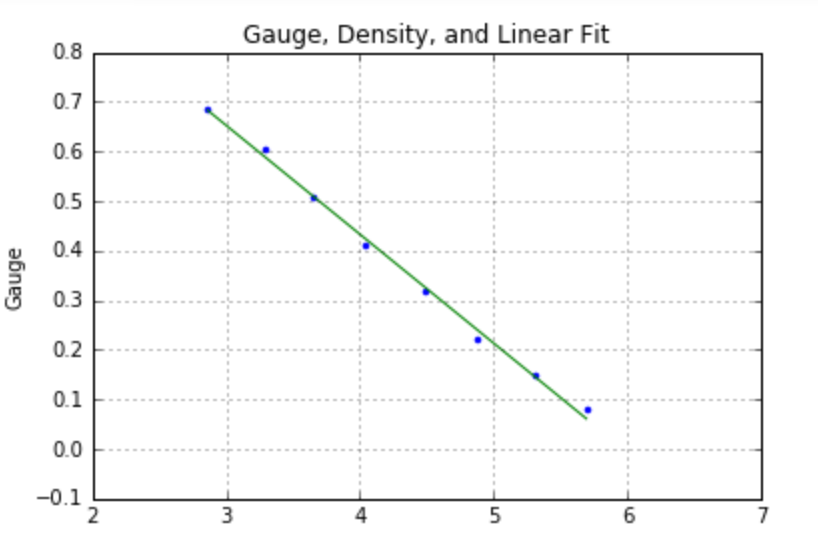


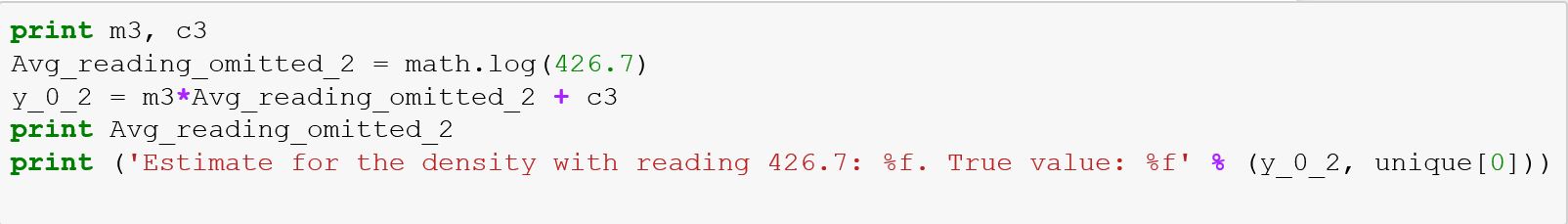
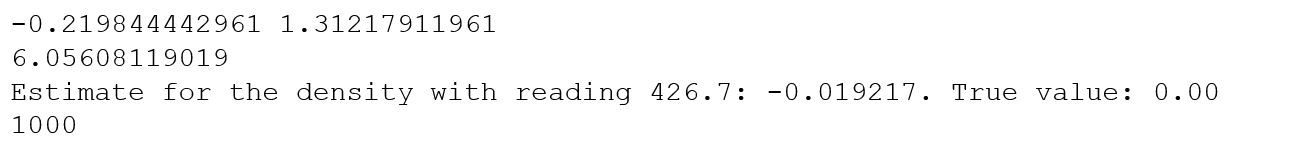


We repeated the same procedure for the reading with measurement average of 6.05608 (log(426.7)), corresponding to density of 0.001.

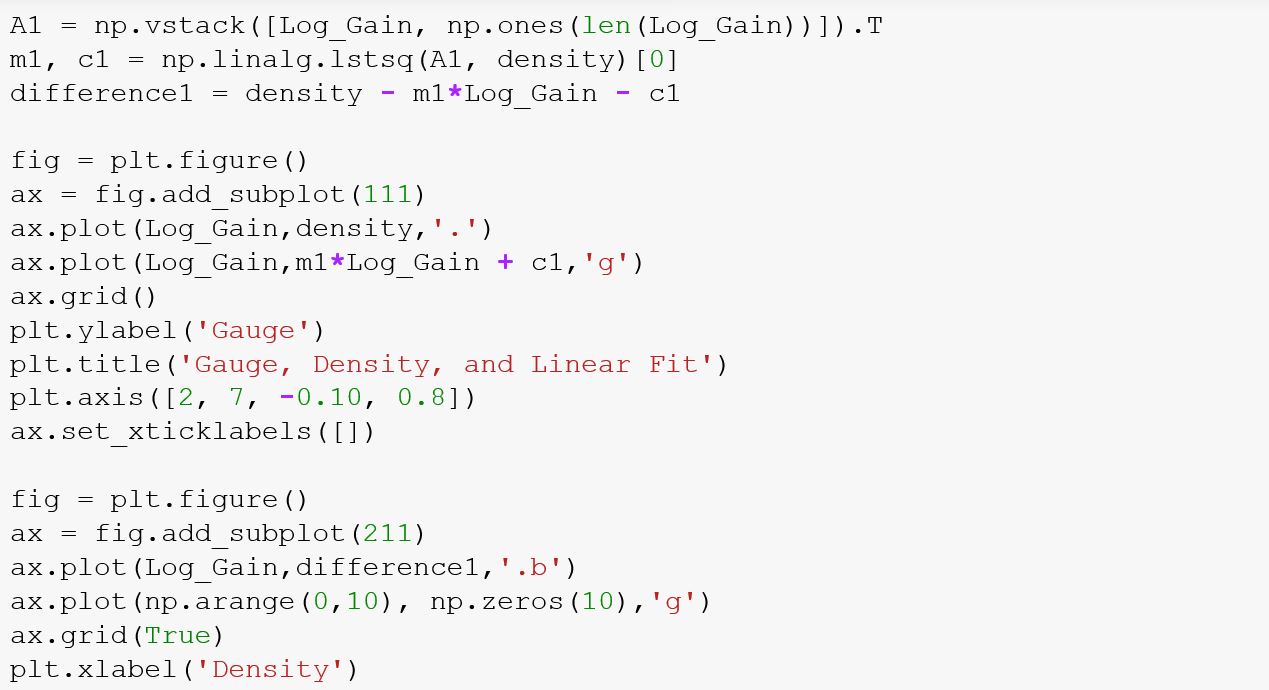
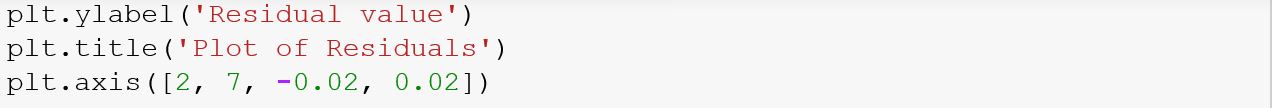




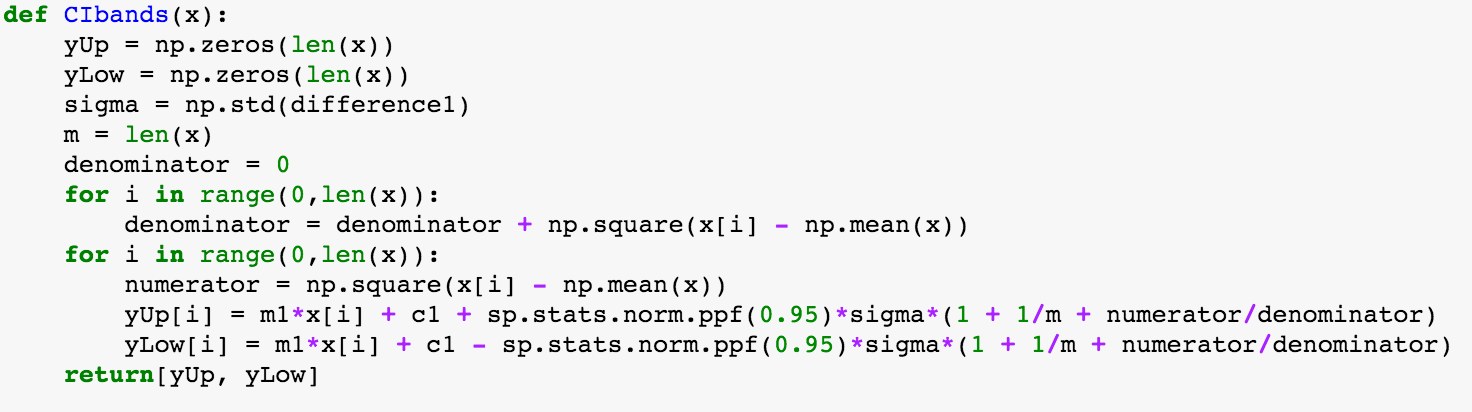




The estimate is clearly wrong as there is no such a thig as negative density. However this is justified since it is still within a very narrow error range.



1. The error bands described in the last section of the results were using the function CIbands, as shown below:



The formula is an unbiased estimation of the variance for each data point. Its derivation is included in the next section.

1. Our Derivation of the Variance:

