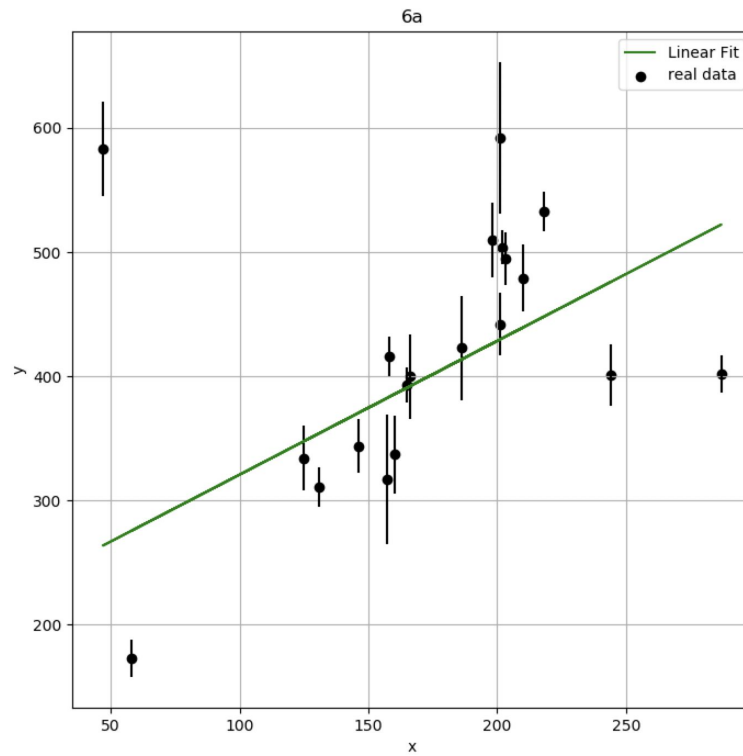


6a.

1. (plot)



a.

2. (values)

```

X:
[[213.27349198]
 [ 1.07674752]]
b:
[213.27349198]
m:
[1.07674752]
chi2:
[[289.96372278]]
cov:
[[ 2.07188189e+02 -1.05427206e+00]
 [-1.05427206e+00  5.99181008e-03]]
sigb:
14.394033107162233
sigm:
0.07740678316575621
q:
[[3.98450958e-118]]

```

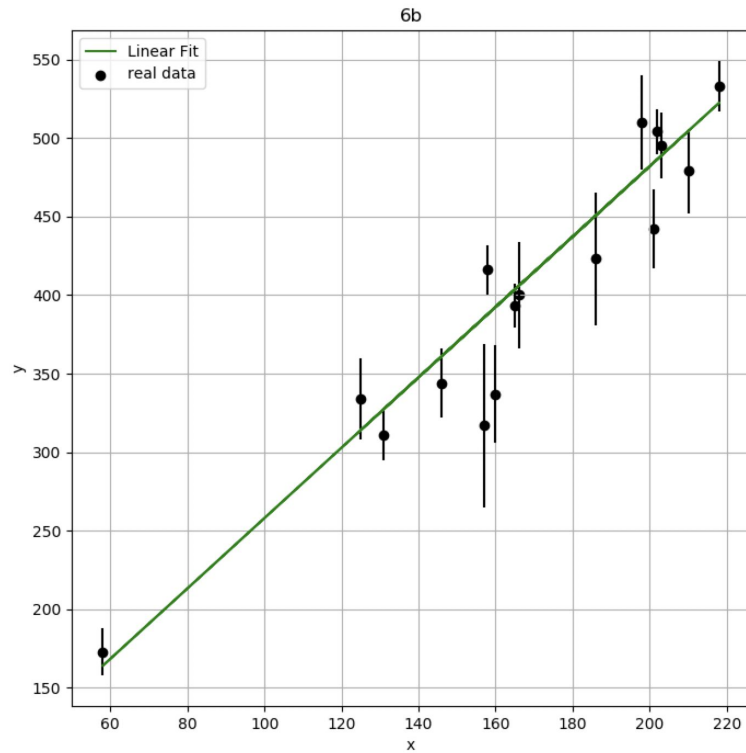
a.

3. The value for chi-squared is very large, so our q value is extremely small. This indicates that our linear fit is not a very good fit for the data. This can be seen in

the plot due to the extreme outliers in our data. Our values match with the values in H10.

6b.

1. (plot)



a.

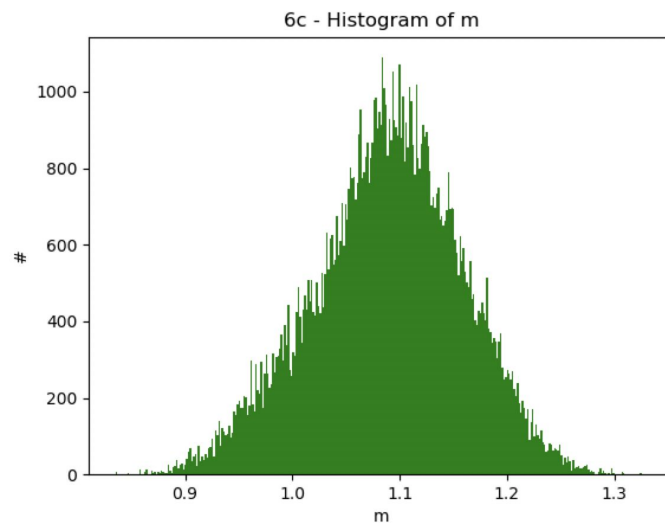
2. (values)

```
X:
[[34.04772776]
 [ 2.23992083]]
b:
[34.04772776]
m:
[2.23992083]
chi2:
[[18.68076991]]
cov:
[[ 3.32922601e+02 -1.88954491e+00]
 [-1.88954491e+00  1.16166311e-02]]
sigb:
18.24616674926818
sigm:
0.1077804765405008
q:
[[0.2329]]
```

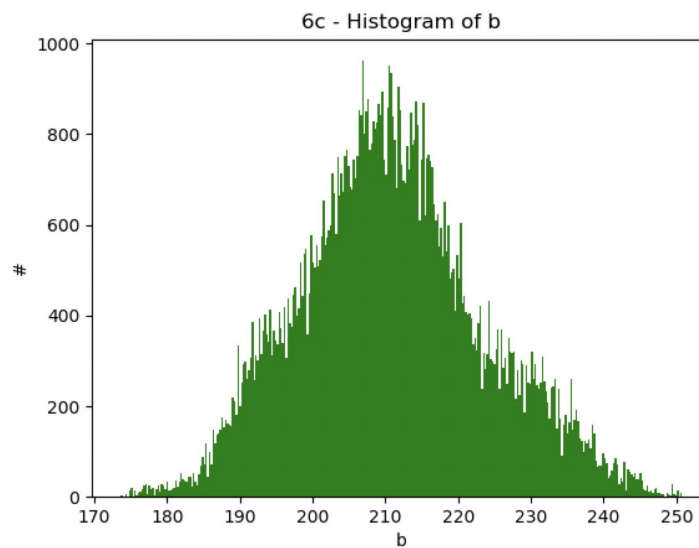
a.

3. The value for chi-squared is less large, so our q value is _____. This indicates that our linear fit is a much better fit for the data. This can be seen in the plot due to the lack of extreme outliers in our data. Our values match with the values in H10.
- 6c.

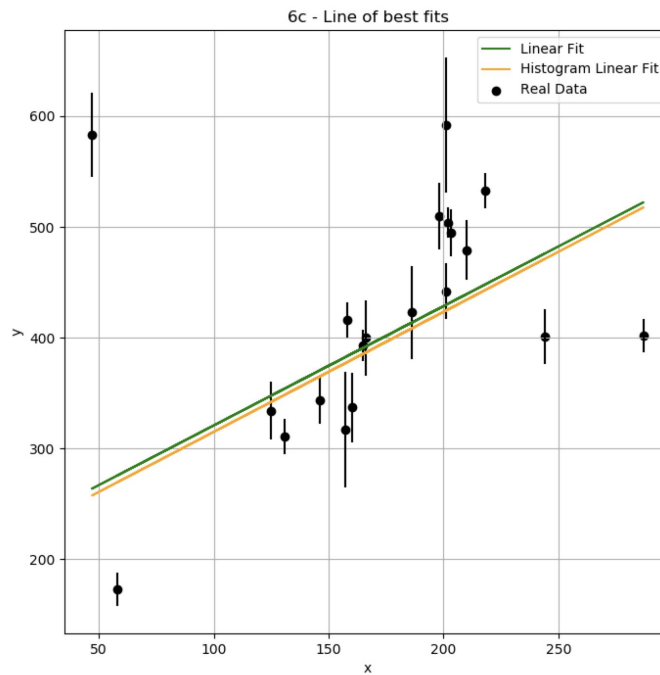
1. (plot)



a.



b.



C.

2. (values)

```

X:
[[213.27349198]
 [ 1.07674752]]
chi2:
[[289.96372278]]
cov:
[[ 2.07188189e+02 -1.05427206e+00]
 [-1.05427206e+00  5.99181008e-03]]
q:
[[3.98450958e-118]]
sigb:
14.394033107162233
sigm:
0.07740678316575621
L:
[1.92927157e-99]
<m>:
1.0833806137687567
<b>:
206.7784440233994

```

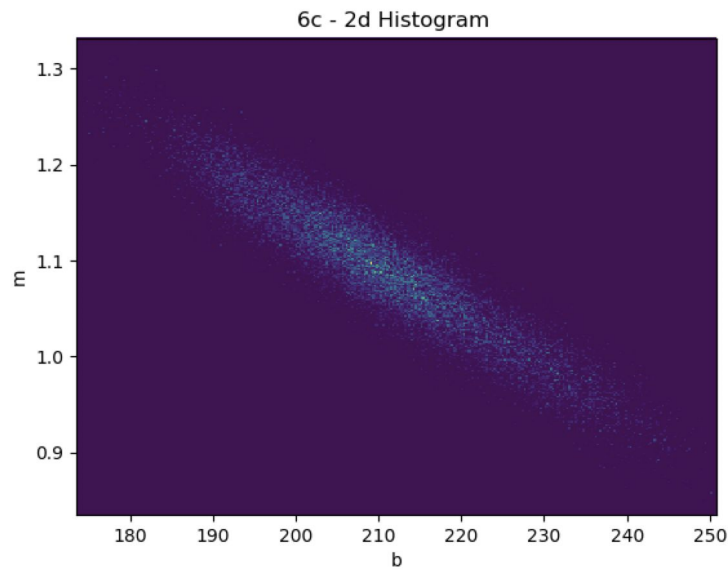
a.

3. (compare)

- a. The values we calculated in part a were very similar to the values we calculated by using the peaks of the histograms we created for linear fit. The values we calculated in part a were:
 - i. $m = 1.08 \pm 0.08$ and $b = 213 \pm 14$.
- b. Our peak values we found from the histograms were:
 - i. $m_p = 1.08$ and $b_p = 207$
- c. The values found from our histogram are probabilistic so they are not consistent between runs, however they are always similar to the m and b

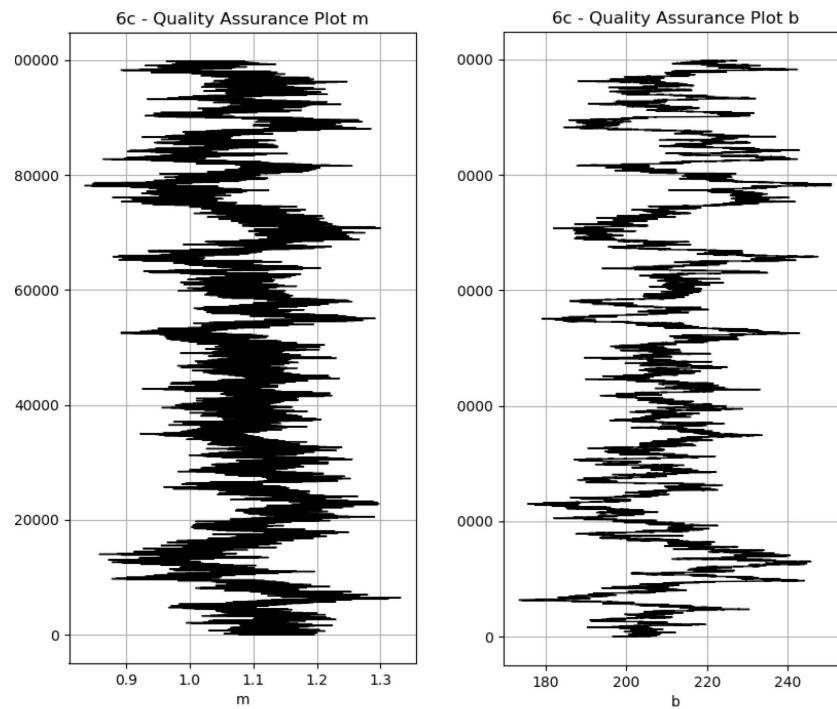
values we calculated in part a (and within the uncertainty values we calculated in part a).

4. (2D)



a.

5. (quality plot)



a.

6. (consistency)

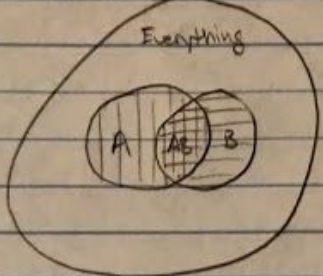
- a. The values do not stay consistent between runs. This is due to the probabilistic nature of the method we are using (the metropolis hasting's sampling method) combined with the step size ($\delta m = 0.1$, $\delta b = 1$) and

Markov chain length ($T = 100,000$). If I chose to tune the parameters (by increasing the Markov chain length and optimizing the step size) the method would become significantly slower, but would also become more consistent.

6d.

1. (Determine reasonable priors)
 - a. Since we do not know too much about the priors for the five parameters (m, b, P_b, Y_b, V_b), we should choose simple priors. Additionally, in H10 it states that "Suce it to say, in this situation—pruning bad data—the setting of priors is the least of one's problems." Thus, the reasonable priors I chose were that $p(\theta) = 1$, when the values of P_b and V_b are valid (P_b is between 0 and 1, and V_b is not negative) and $p(\theta)=0$ when the values of P_b and V_b are invalid. Probabilities must be between 0 and 1, so that is why P_b must be between 0 and 1. Variances are also always positive, so that is why $V_b < 0$ is invalid.
2. (Show bayes theorem, prove bayes theorem)
 - a. Prove Bayes Theorem

Bayes Theorem Proof (With Pictures)


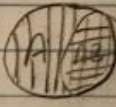


$$P(A) = \frac{|A|}{|\text{Everything}|}$$

$$P(B) = \frac{|B|}{|\text{Everything}|}$$

$$P(AB) = \frac{|AB|}{|\text{Everything}|}$$

$$P(A|B) = \frac{|AB|}{|B|} = \frac{\frac{|AB|}{|\text{Everything}|}}{\frac{|B|}{|\text{Everything}|}} = \frac{P(AB)}{P(B)}$$

$$P(B|A) = \frac{P(AB)}{P(A)}$$



$$\Rightarrow P(A|B)P(B) = P(B|A)P(A)$$

$$\Rightarrow P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

b. Since we know that:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

i.

ii. We know that $p(\theta|y) = p(\theta|y)p(\theta)/p(y)$

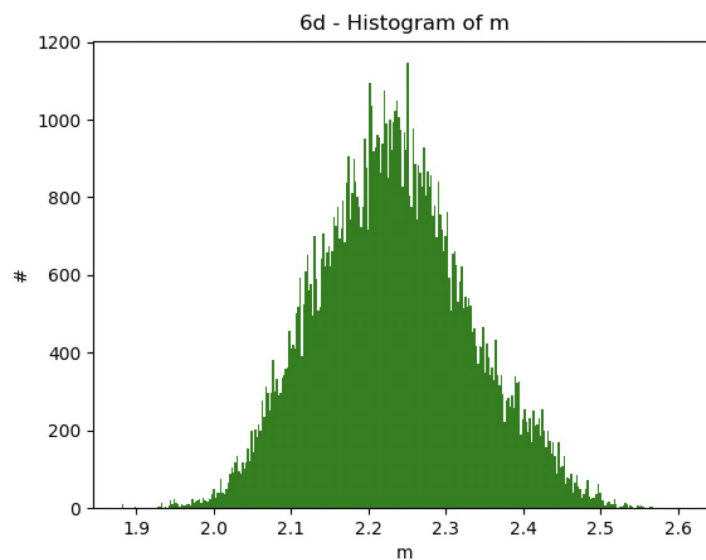
iii. This is the same as $p(\theta|y) = p(\theta|y)*p(\theta)$ without the constant $1/p(y)$

3. (irrelevant constant)

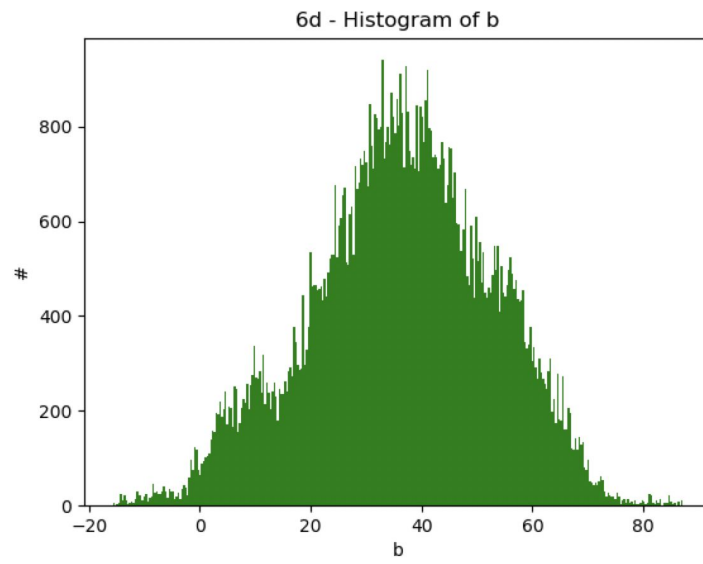
a. The normalization constant being missing in eq. 7 is irrelevant for our purposes since in our metropolis hasting's algorithm, the acceptance ratio is the ratio of $p(\theta|y)$ with our candidate values of theta and the previous values of theta. Since both $p(\theta|y)$ values would have been divided by the same constant ($p(y)$), it can be ignored in our calculations (it would cancel out when finding the acceptance ratio).

4.

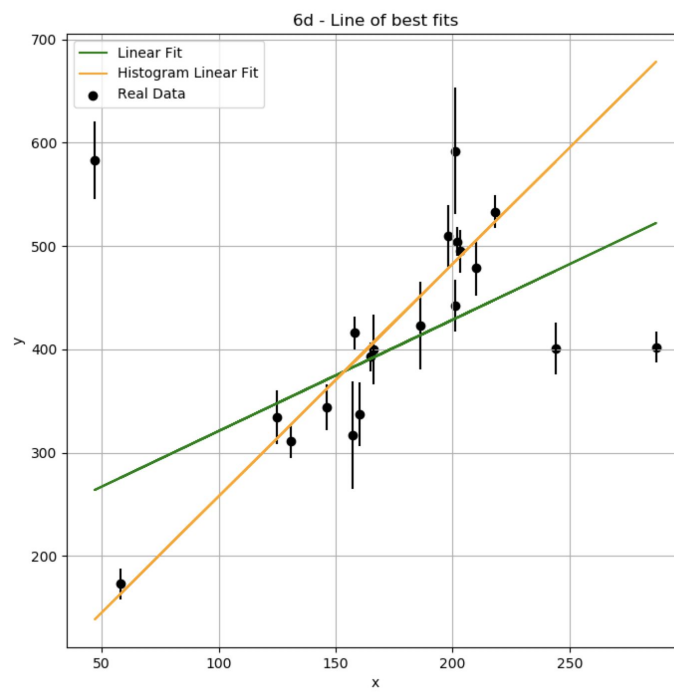
a. (plot)



i. _____



ii.



b. (values)


```

X:
[[213.27349198]
 [ 1.07674752]]
chi2:
[[289.96372278]]
cov:
[[ 2.07188189e+02 -1.05427206e+00]
 [-1.05427206e+00  5.99181008e-03]]
q:
[[3.98450958e-118]]
sigb:
14.394033107162233
sigm:
0.07740678316575621
<m>:
2.2498247987029325
<b>:
32.79863305676257

```

i.

c. (compare)

i. The values we calculated in part b were very similar to the values we calculated by using the peaks of the histograms we created for linear fit. The values we calculated in part b were:

$$1. \quad m = 2.24 \pm 0.11 \text{ and } b = 34 \pm 18.$$

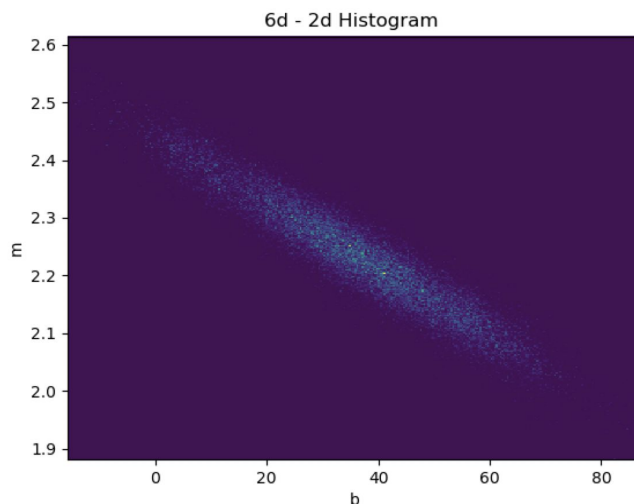
ii. Our peak values we found from the histograms were:

$$1. \quad m_p = 2.25 \text{ and } b_p = 33$$

iii. The values found from our histogram are probabilistic so they are not consistent between runs, however they are always similar to the m and b values we calculated in part b (and within the uncertainty values we calculated in part b). Also since we removed outliers by modeling that data, there is a chance that we removed different outliers than when we removed outliers by hand. This would cause the expected m_p and b_p to be slightly different from the m and b we calculated in part b.

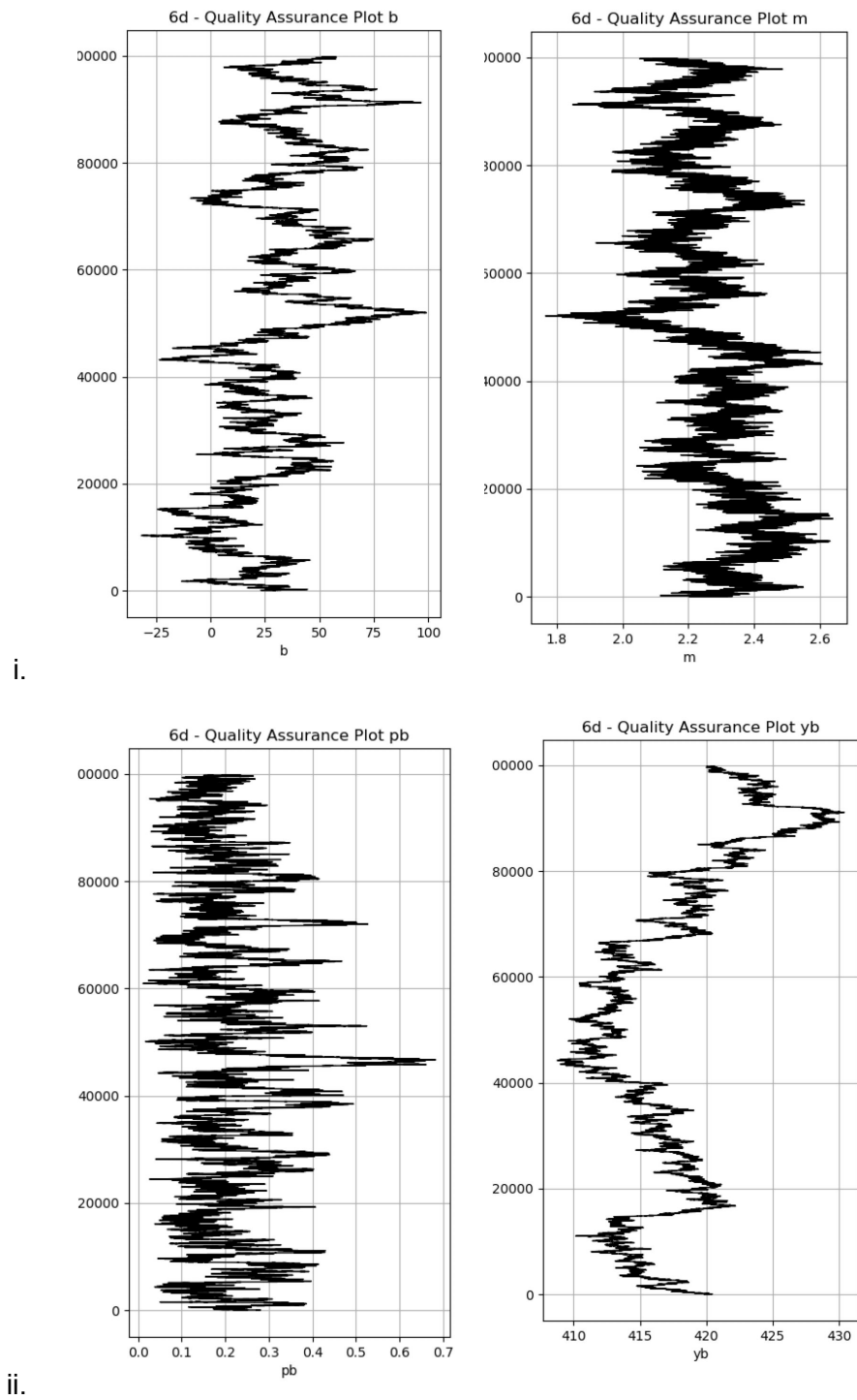
iv.

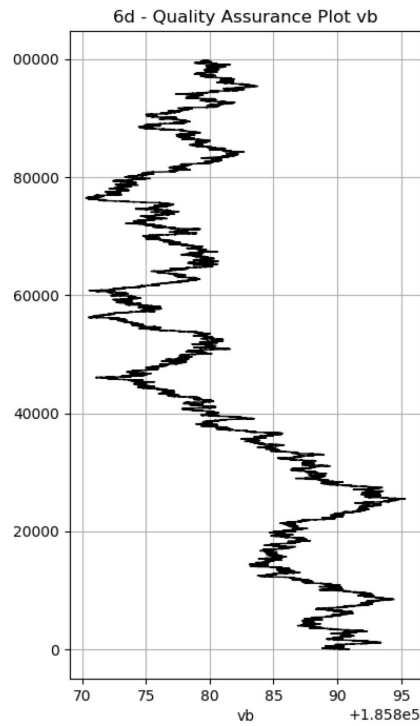
d. (2D)



i.

e. (quality plot)





- iii.
- f. (consistency)
- i. The exact same as my answer to 6c-6.