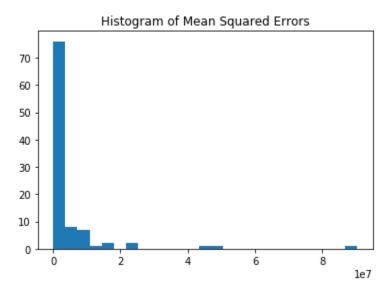
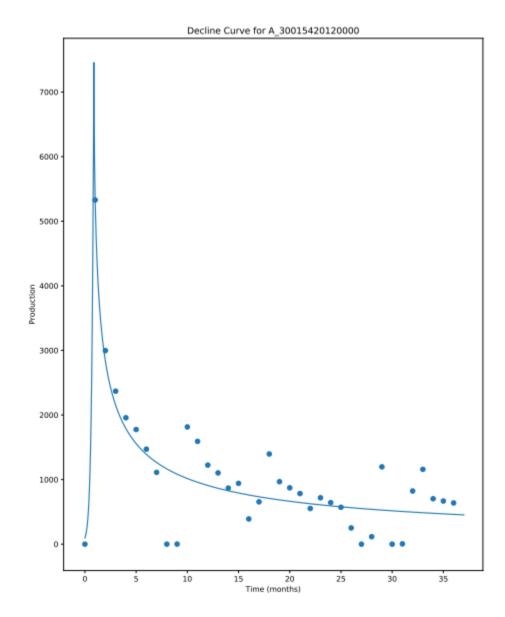
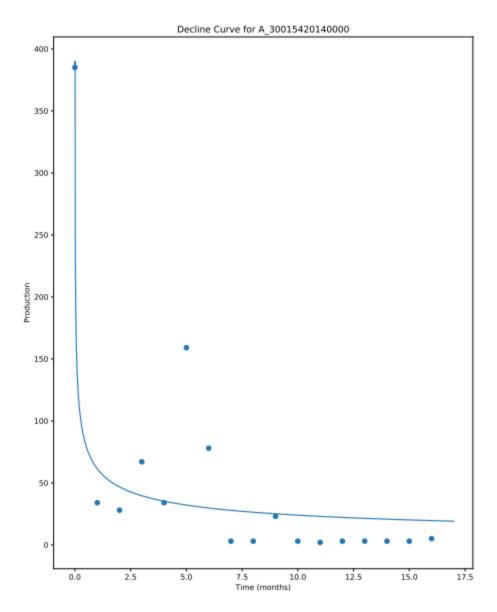
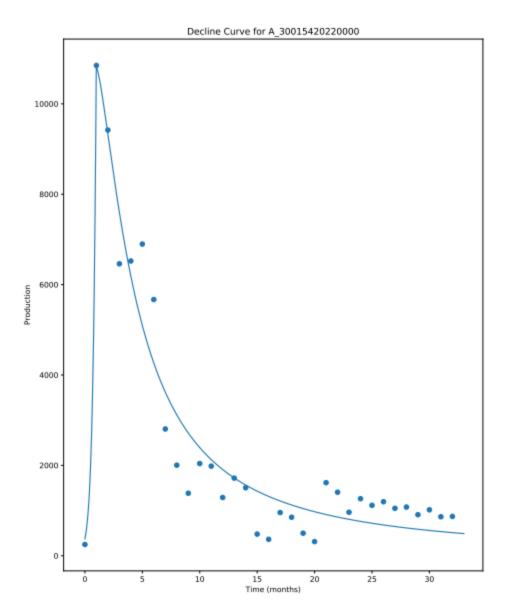
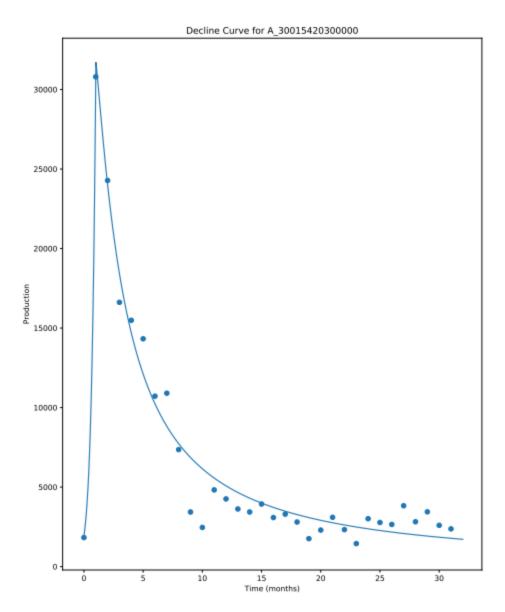
2.











- 4. I chose MSE for the loss function because outliers were already removed and to discourage large errors in the regression.
- 5. If a production point data was more than a standard deviation away from the mean of its neighbors, it was removed.
- 6. The evaluation schema/prediction power was determined by seeing how much improvement there was from using the curve rather than just extending the last production data. This was done by splitting the production data into two parts, training set and test set. Using the training set, a curve was created using the same method as fitting. Then it was compared with a flat extended line from the last point on the training data. The residual from the curve and the residual from the extended line were used to calculate a r2 coefficient for the curve. This test can be run for multiple wells at various split points using the python code. The curve proved useful as it had a r2 greater than 0 for most cases which means it is better than the null hypothesis line drawn from the last point.