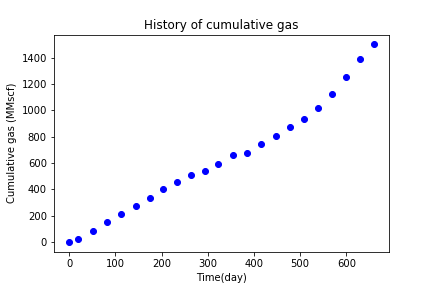
**Predicting the gas production**

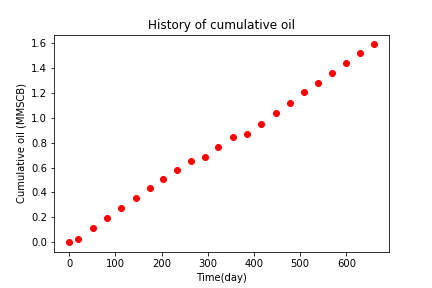
Vahid Mohammadnia, August 2018

Predicting gas production from oil wells is a crucial factor in predicting how long a well is capable to produce oil. The main force that pushes the oil out of well is the gas dissolved in the oil because gas is compressible. That is the reason why in production, engineers try to keep the gas in place as long as possible to provide the necessary force for the oil.

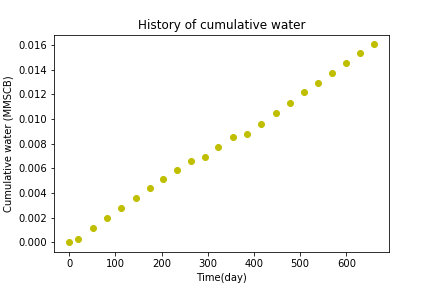
In this project the production of three wells in a small reservoir is given. Figure 1, 2 and 3 show the cumulative production of gas, oil and water within 690 days in all the reservoirs.



**Figure 1**: History of cumulative gas production in 690 days.



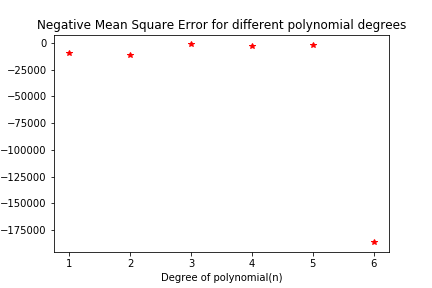
**Figure 2**: History of cumulative oil production in 690 days.



**Figure 3**: History of cumulative oil production in 690 days.

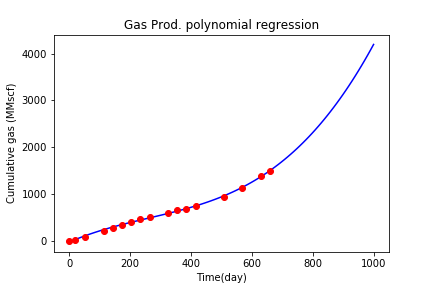
These figures present the production history for all the reservoir. In this project different polynomial regressions are developed to predict the gas production on the one of the given wells, named Well\_x05. The developed algorithm can be applied to other wells too. Time (in day) is used as the feature of the model. The model is small and I did not use sophisticated algorithms like SVM, NN or decision tree. Figure 4 shows negative mean square errors for different values of polynomial degrees (n= 1 to 7) with cross validation is performed based on k-fold evaluation method. This score is calculated based on estimation of cross validation data not the training data. As seen in this figure, for n = 3 and n = 5, the model presents the highest score.

Figures 5 and 6 are the polynomial regression models suggested by the two models for the training data for well\_x05. However, the prediction capability of models for after 900 days is different for these models.

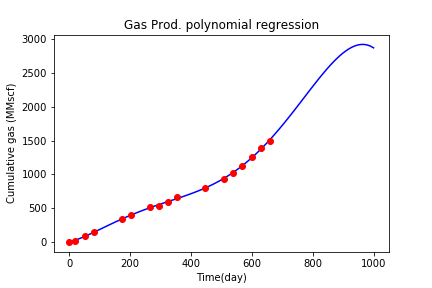


**Figure 4**: Negative Mean Square Error for different values of polynomial degrees (n= 1 to 7),

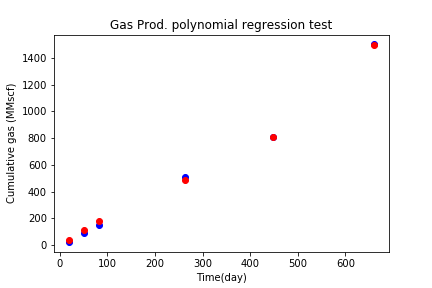
As seen at n= 6, the error rises up sharply as overfitting happens for degrees higher than 5.



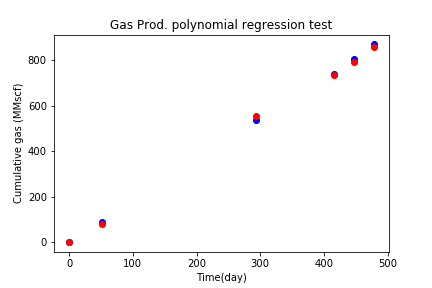
**Figure 5**: Polynomial regression for gas production (degree = 3), red circles represent the observed training data, mean square error for the test data is: 370.74716



**Figure 6**: Polynomial regression for gas production (degree = 5), red circles represent the observed training data, mean square error for the test data is: 26.81522



**Figure 7**: Polynomial regression for gas production (degree = 3), red circles represent the predicted values the observed value indicted by blue circles.



**Figure 8**: Polynomial regression for gas production (degree = 5), red circles represent the predicted values the observed value indicted by blue circles.

Figure 7 and 8, show the predicted values suggested by the models for test data. Both models present decent results but still they behave differently for large number of days. As seen, the models almost predict similar values for days earlier than 900 but for higher days, more data is needed for verification.

**Recommendation**

Any of the models (n = 3 and n = 5) can be used to predict the cumulative gas production (more reliable for days earlier than 900th day). But it is strongly recommended as the real data started to deviate from the predicted values, the models are regenerated for more accurate results.