Modelling hydrocarbons flow from reservoir rocks using deep learning

Raisa Energy

Hadeer El-Ashahb – Salma Amr – Mohamed Salem – Youssef Hossam
Motaz El-Saban – Curtis Caile – Paul Schietinger

Forecasting production

Problem description

- Understand how hydrocarbon flows out of the earth using **geological** and **engineering** features
- Predict Oil&Gas production until the end of a well’s life (30-40 years) for unconventional horizontal wells using deep-learning
- The approaches presented here is intended to forecast production for both producing locations **(PLs)** and non-producing locations **(NPLs)**
- Dataset used in all our experiments contained **100k PLs**

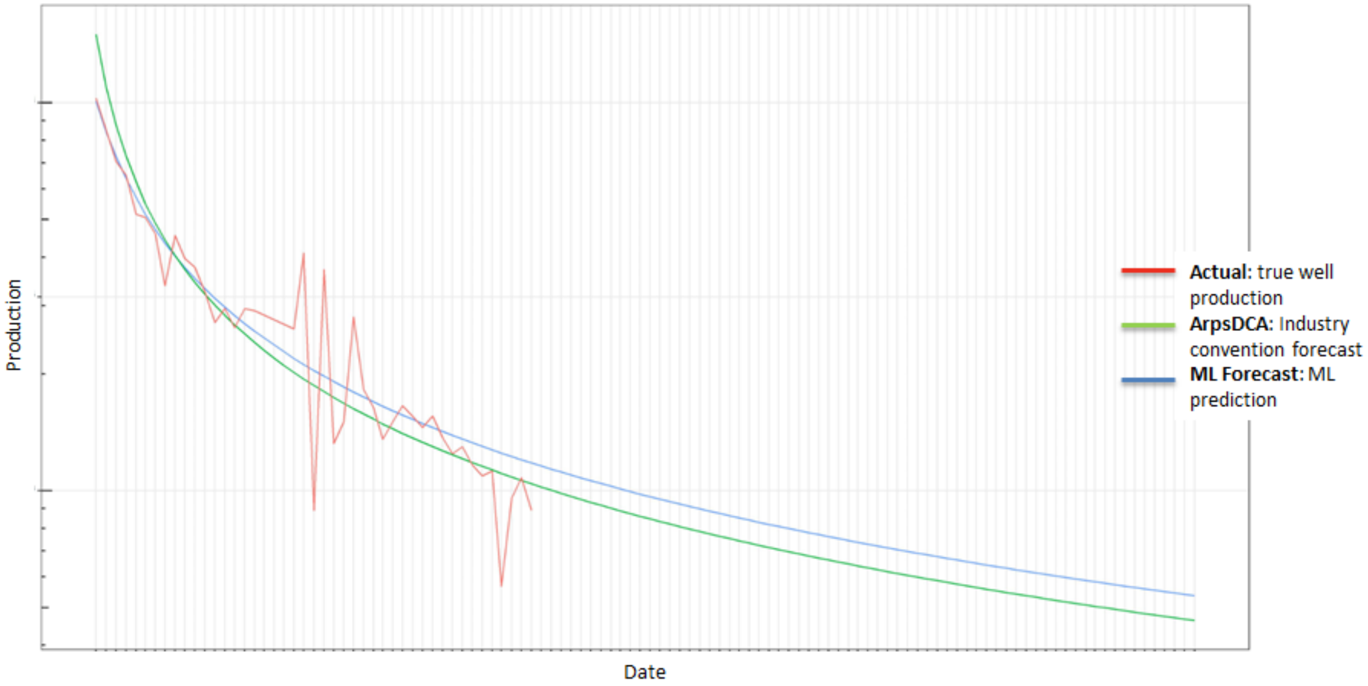
Forecasting production

Underlying problems

- Forecasting production for PLs
- Forecasting production for NPLs
- Downtime prediction
- Pre-peak modelling
- Multi-segment modelling
- Parent-Child relationship

Forecasting production for PLs

- In decline curve analysis, Arps equation is used to fit the productior starting from the peak
- In order to locate a hyperbola in space one must know the following:
  - The starting point on Y axis – qi
  - Initial decline rate (di)
  - The degree of curvature of the line (b)

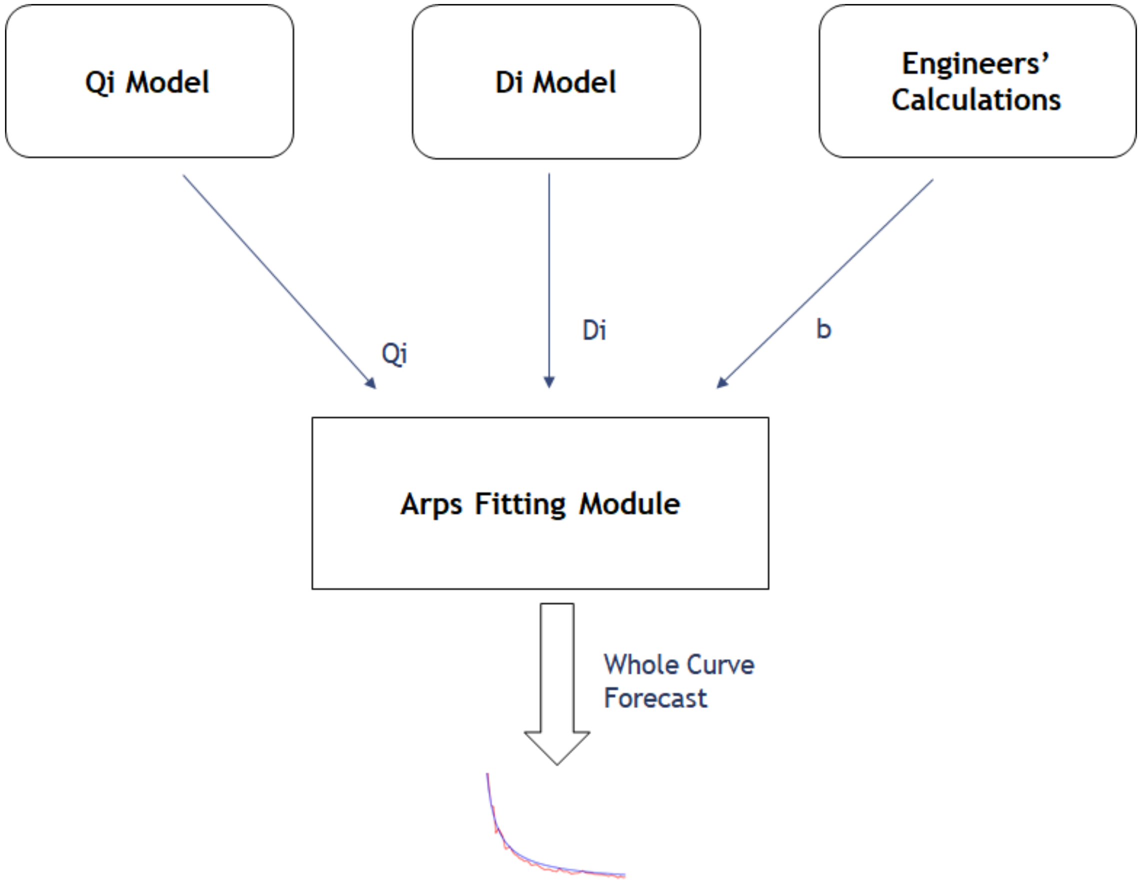


$$q(t) = \frac{q_i}{(1 + bD_i t)^{1/b}}$$

Forecasting production for PLs

One Month ML model

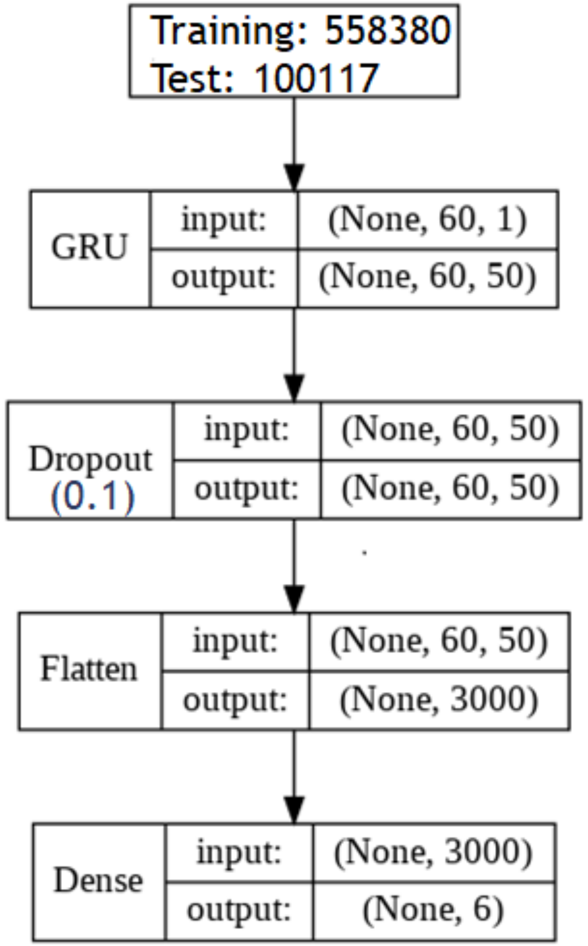
- Two Quantile Random Forests predicting the (qi) & (di)
- Features used in both models were:
  - Longitude
  - Latitude
  - Peak Production
- Training Set: 21543, Test Set: 3544
- n_estimators=100, min_samples_leaf=5



Forecasting production for PLs

Variable Month ML model

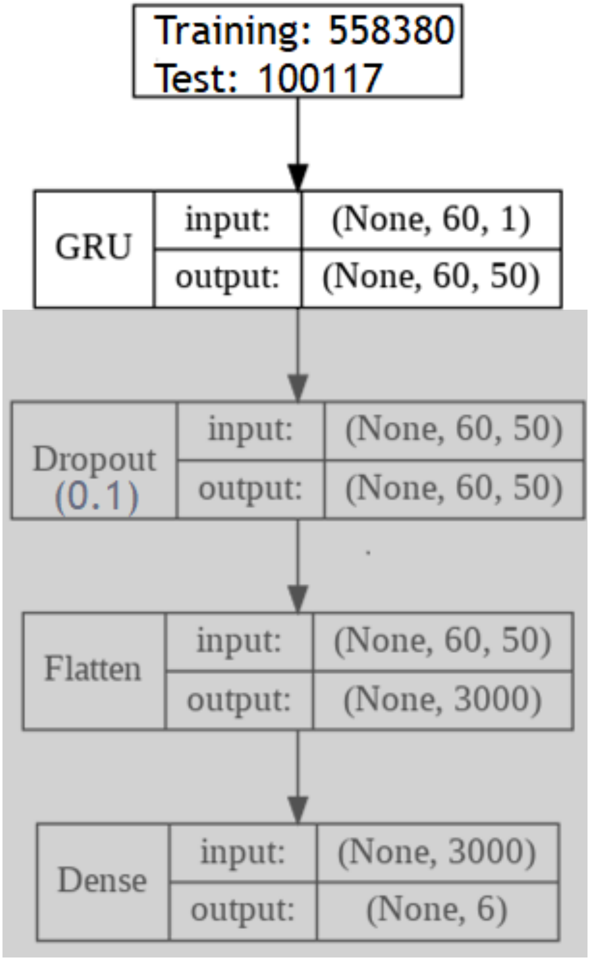
- A vanilla GRU neural network was used for predicting the next 6 months of production given all the available months of production
- Each well was used to generate (N-6) examples where N is the number of months of data available for that well



Forecasting production for PLs

Variable Month ML model

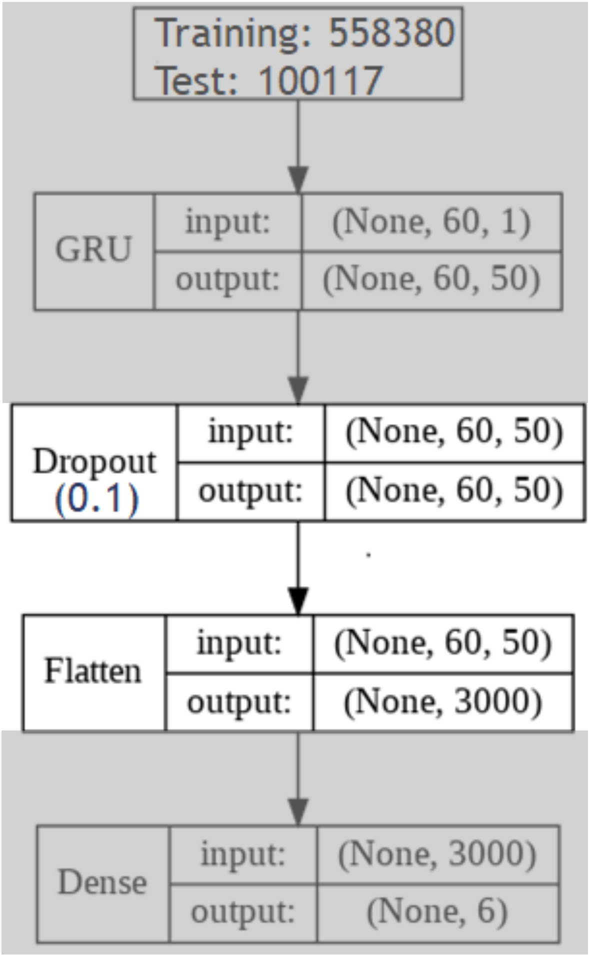
- The input to the model is all the available months of production (up to 60) padded with a dummy negative value
- The output of the GRU is the next 6 months of production



Forecasting production for PLs

Variable Month ML model

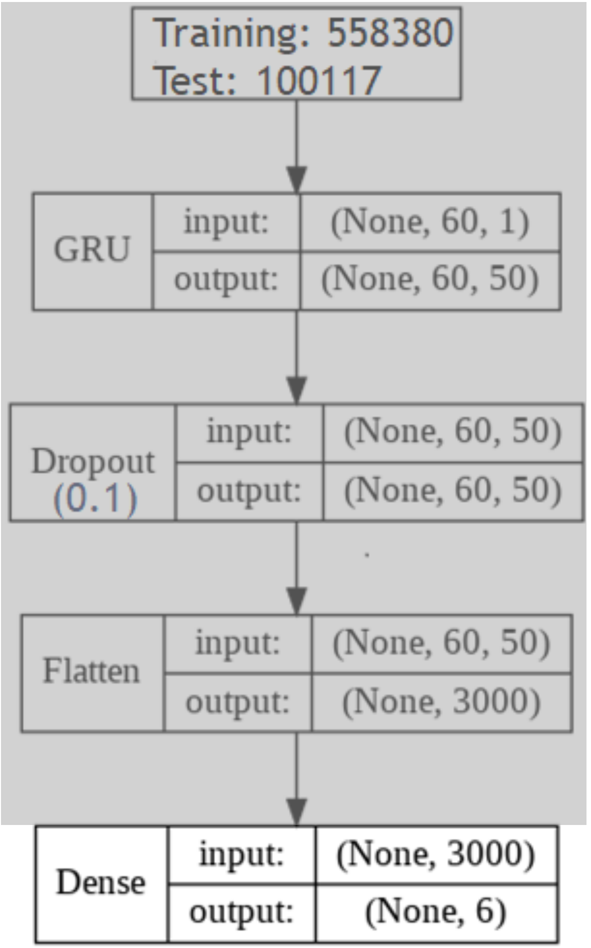
- Dropout used = 0.1
- A fully connected layer was added after the GRU layer



Forecasting production for PLs

Variable Month ML model

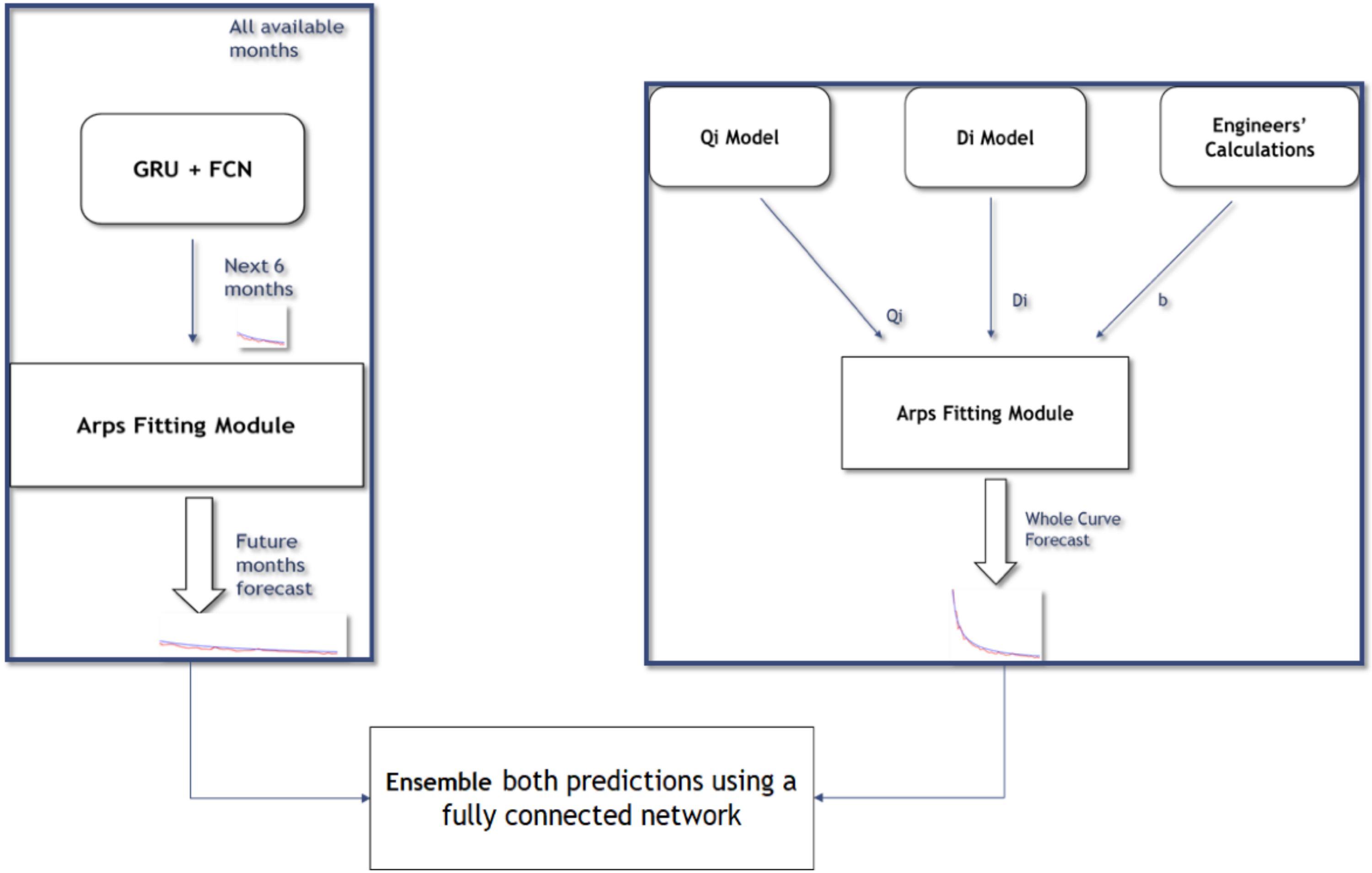
- The output of the model is fed into the Arps fitting module to generate qi, di, b for the predicted production curve in the coming months



Forecasting production for PLs - Results

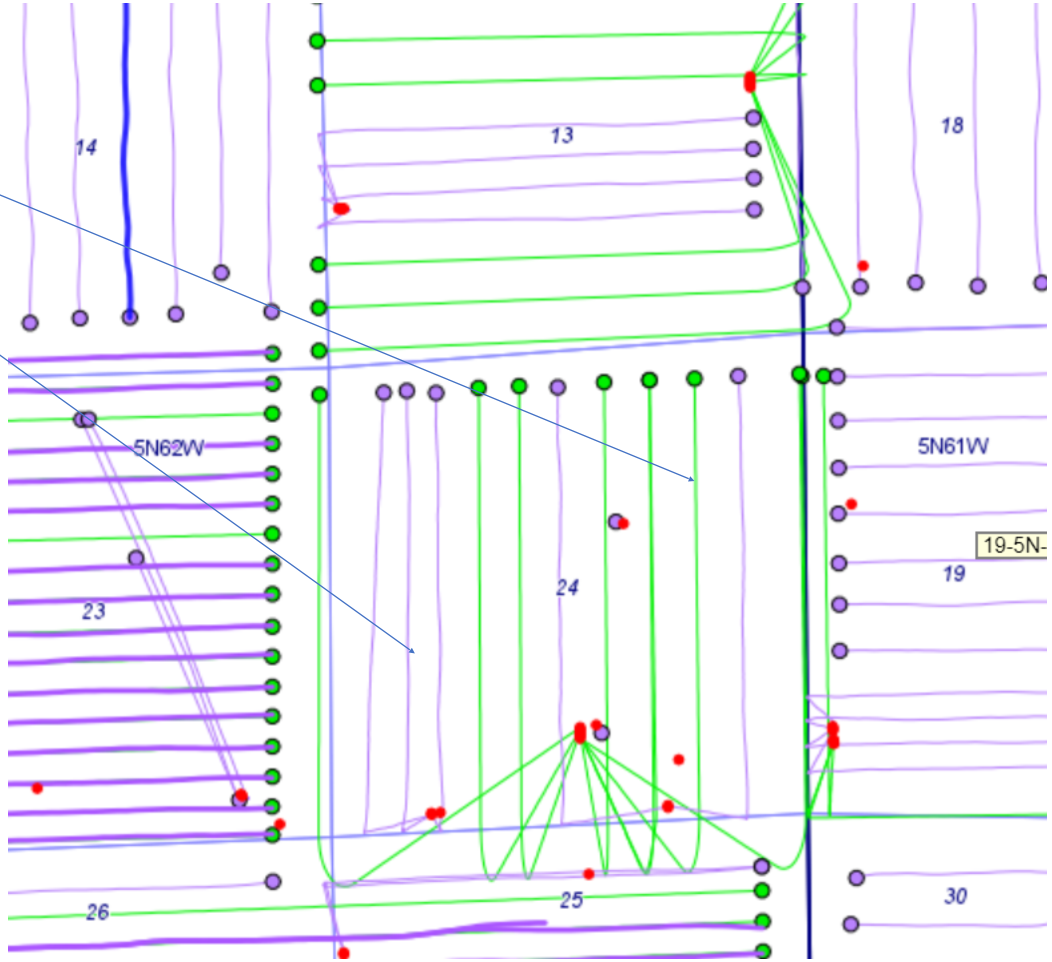
Combining two models

- Improves Median Relative Error from [–0.34 -> 0.57] using [1 -> 30] monthly data point, to predict next 6 months



Forecasting production for NPLs

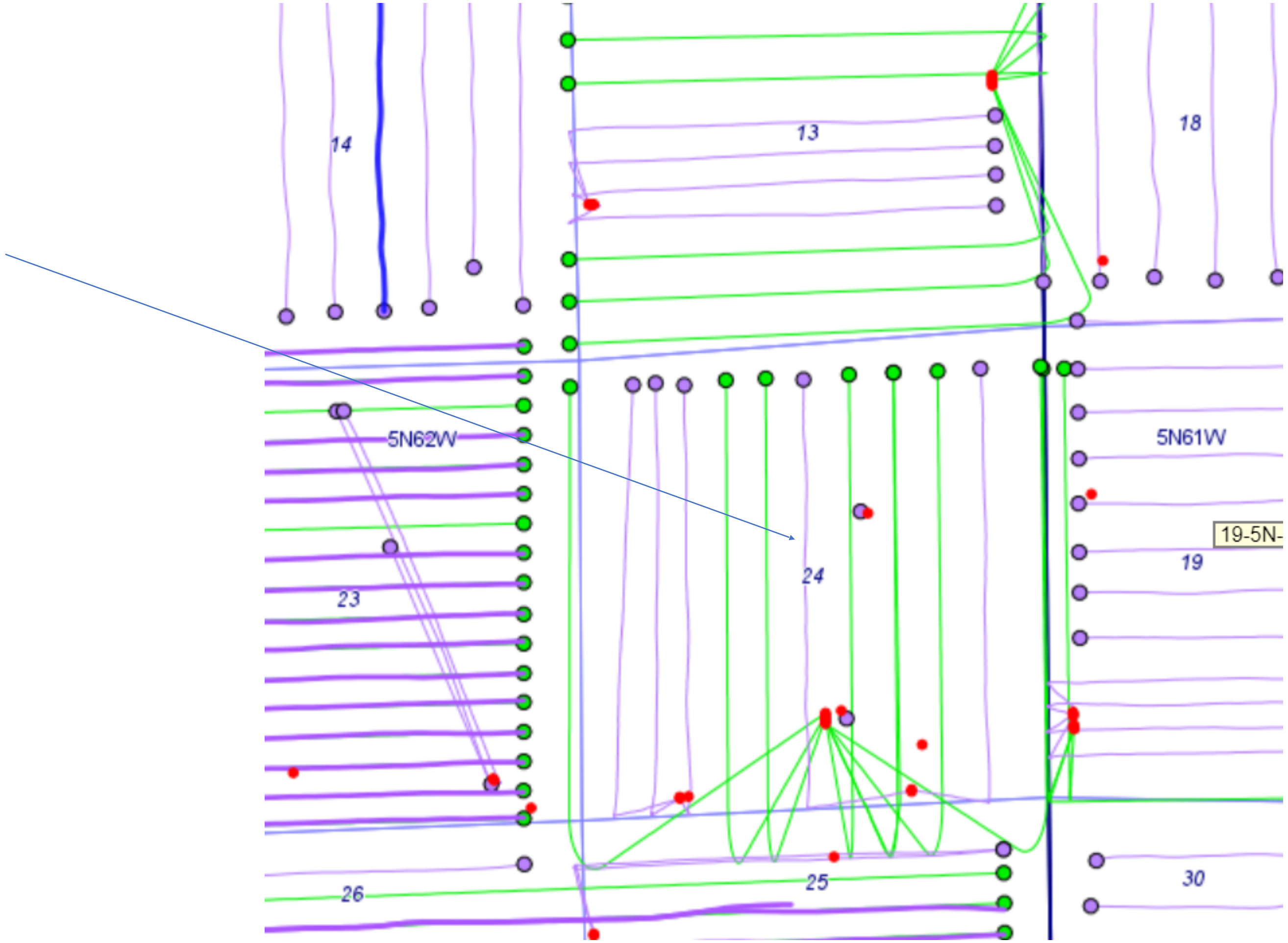
Imagine an area containing already producing wells



Forecasting production for NPLs

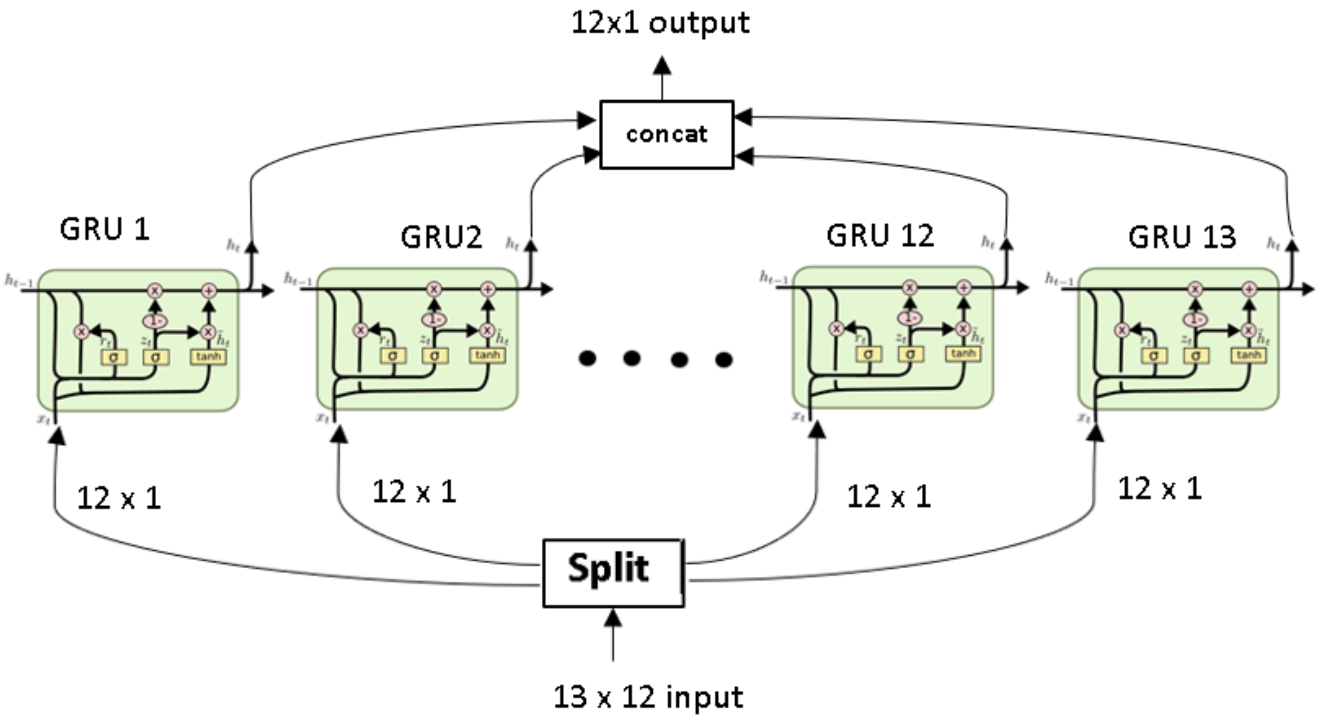
Imagine a well is being drilled in the center

We want to predict this well's production given its neighbours



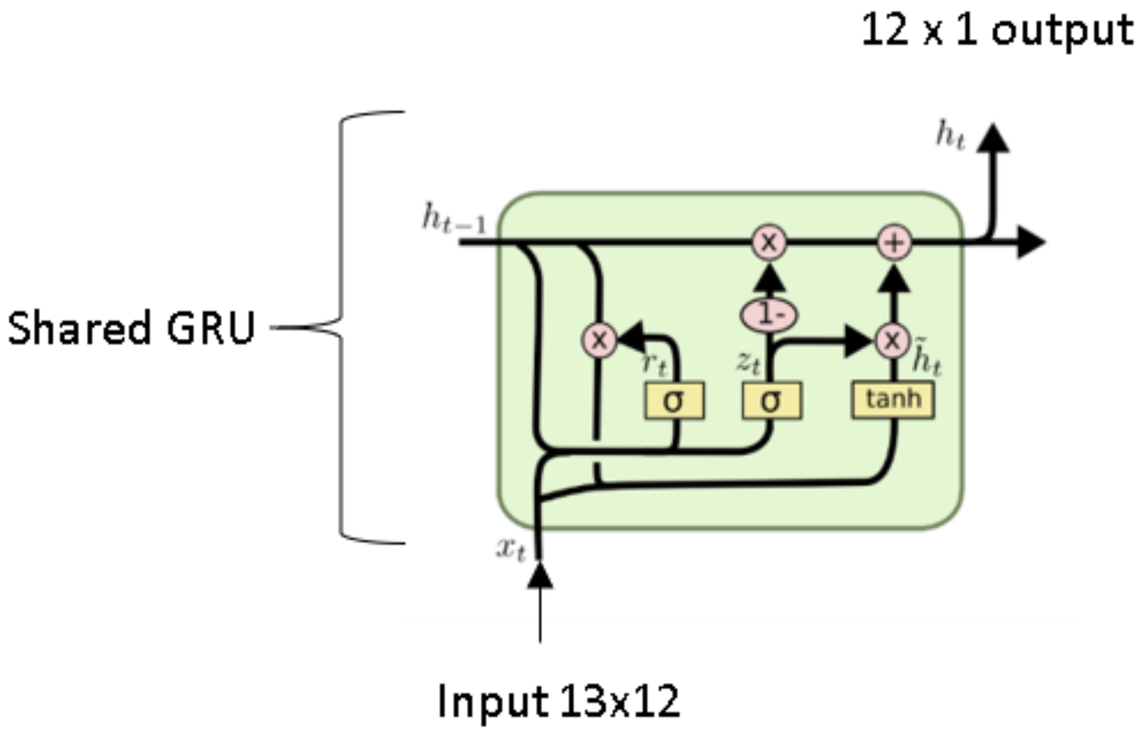
Forecasting production for NPLs

- The first model we tried is based on 13 GRU encoders each of 100 hidden size and ReLU activation.
- The first 10 GRU encoders represent the nearest 10 neighbors within 1 Mile Radius to the potential well respectively.
- The last 3 encoders represent the 10th , 50th, and 90th percentiles of all neighboring wells within 1 Mile Radius.



Forecasting production for NPLs

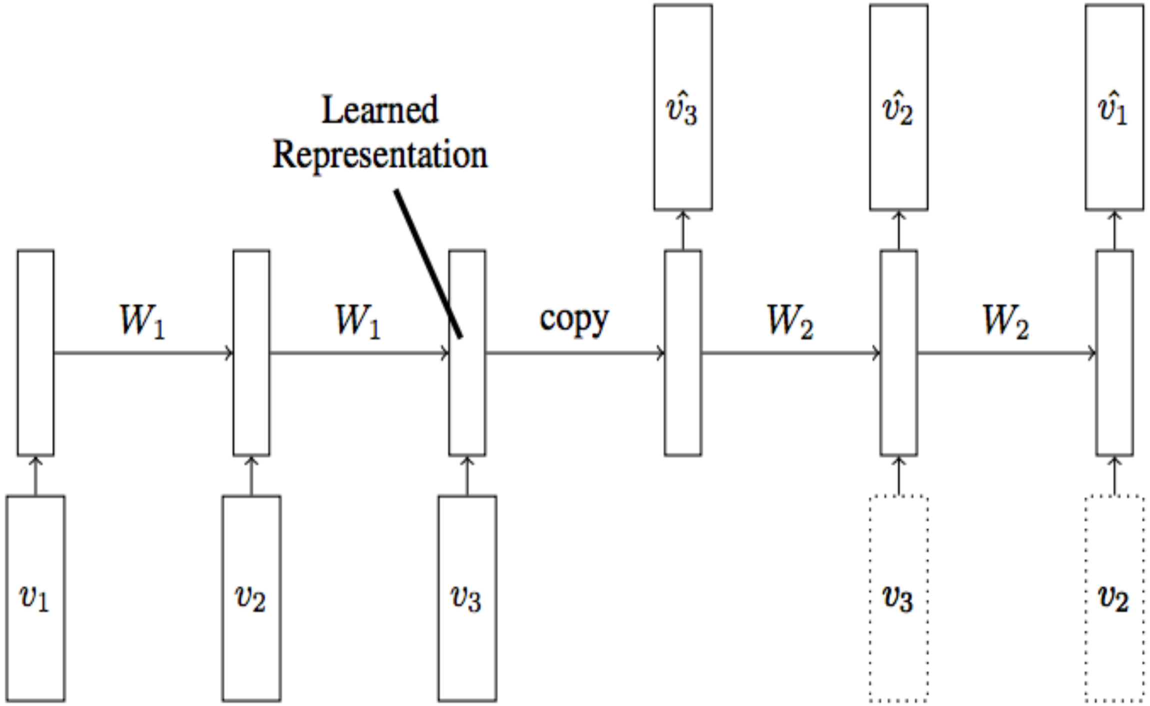
- Instead of using 13 different encoders for each input type, we share one GRU on the different inputs.
- Our hypothesis is that neighbors can share information between each other.
- In addition, it also reduces the parameters learned in the network.



Forecasting production for NPLs

Pretraining the encoder

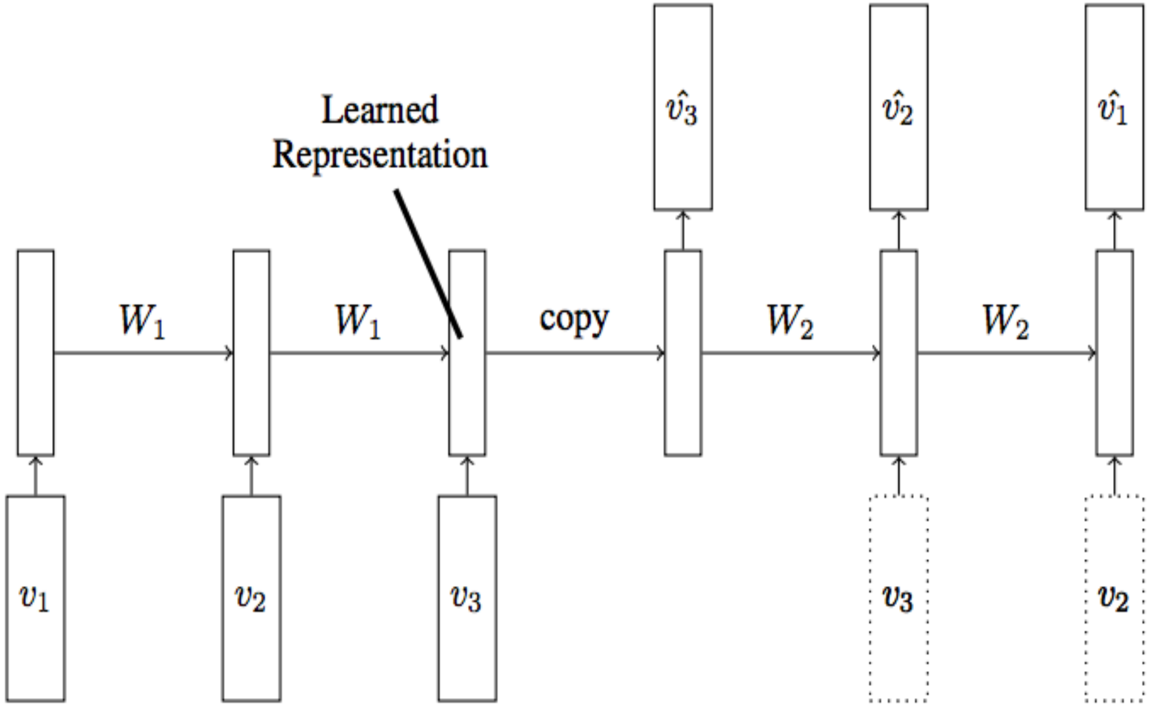
1. The number of training points that satisfies the condition of 10 nearest neighbors are only 15k/100k PLs.



Forecasting production for NPLs

Pretraining the encoder

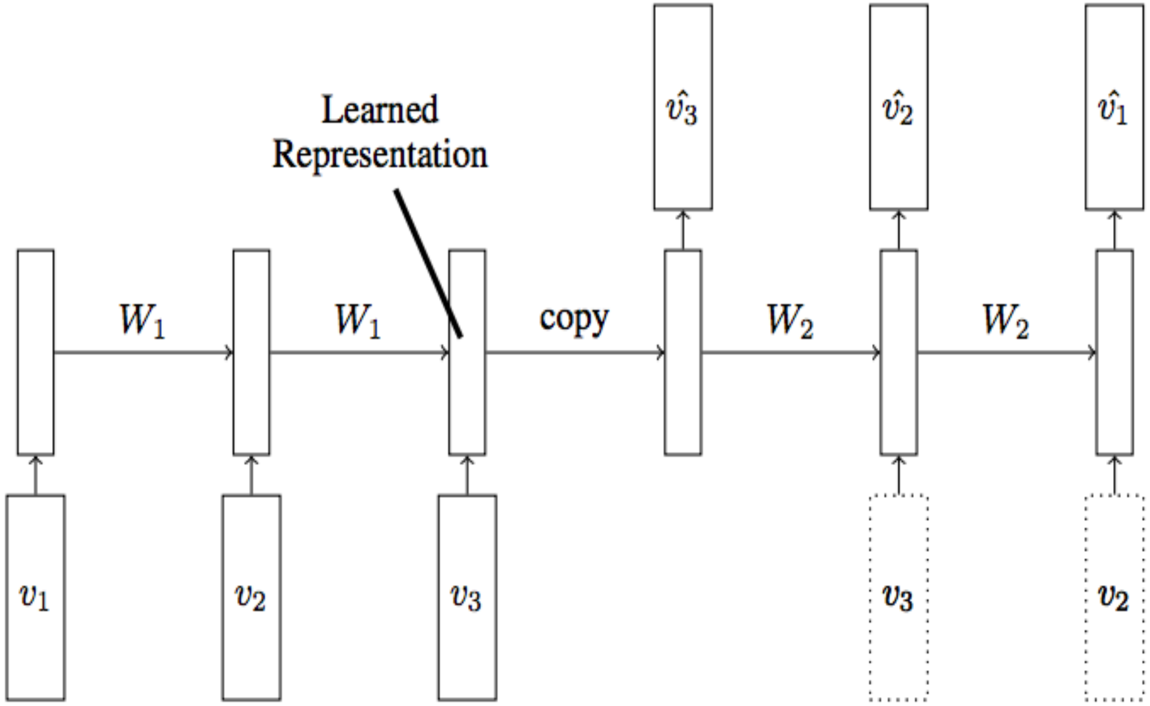
2. In order to make use of the whole dataset, we use self-supervised learning via a sequence to sequence autoencoder trained on the whole producing wells, excluding the wells we are trying to predict.



Forecasting production for NPLs

Pretraining the encoder

3. We initialize the GRU encoder by the learned weights by the learnt weights of the encoder part in the seq2seq autoencoder



Forecasting production for NPLs - Results

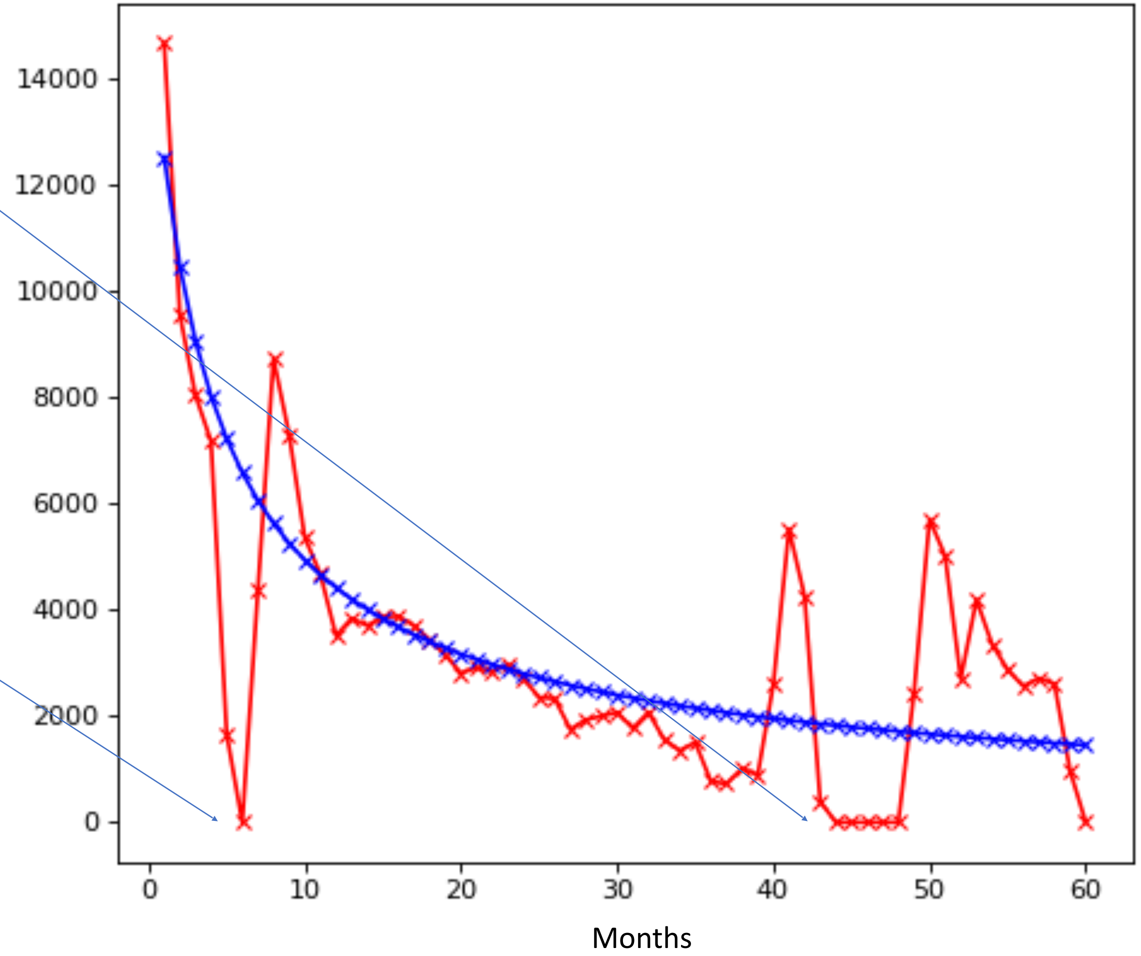
- Our evaluation criteria is the curve to curve Mean Absolute Error between actual production and predicted ones
- We compare our models against standard industrial baseline convention which is the 50th percentile of production of neighboring wells in 1 Mile
- We also compare our results versus a “Random Forest” model trained on a mixture of engineering and industrial features

Model	Mean Absolute Error
50^{th} Percentile	115
Random Forrest Regressor	110.48
GRU (unshared)	103.4
GRU (shared)	**98.04**
Pretrained -GRU (Shared)	98.7

Our best results obtained by sharing the GRU across the input

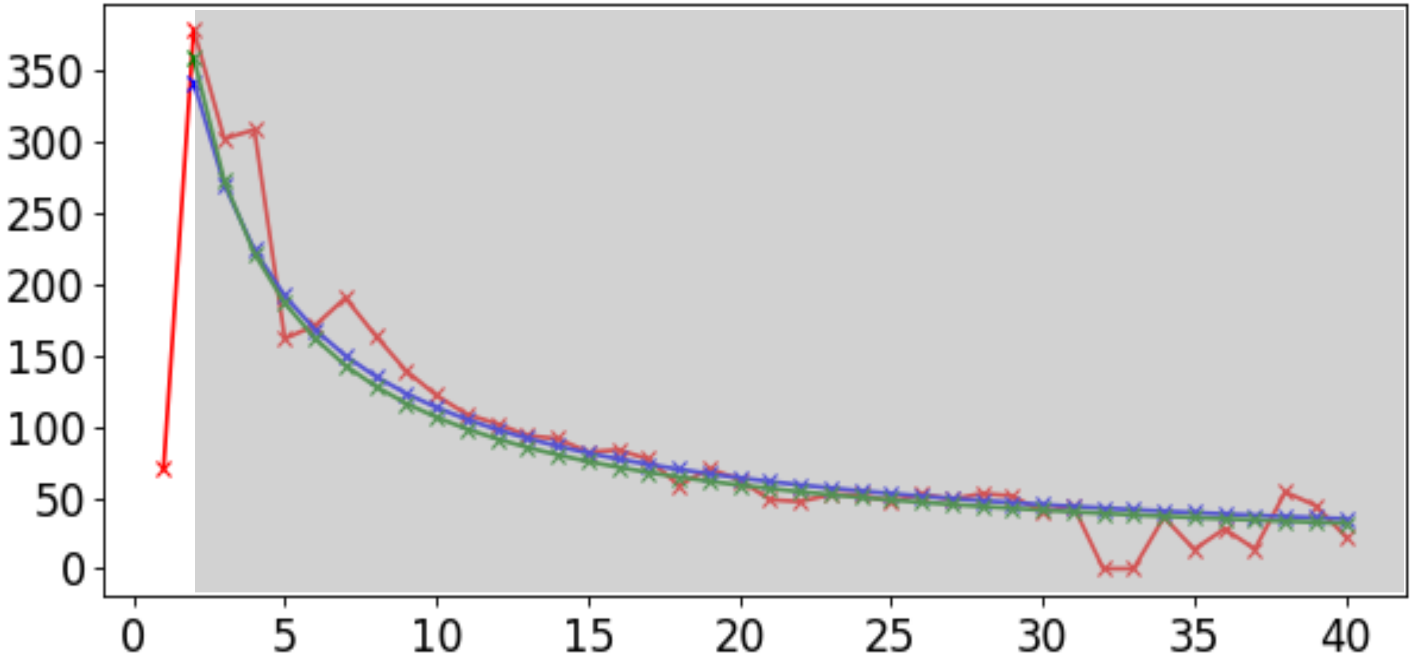
Downtime prediction

- Downtime is defined as the time (in months) required for shutdowns - not specified in advance - that interrupts the production stream of PLs, might be due to maintenance or mechanical issues
- Predicting downtime is important for cashflows
- Results
  - Downtime accounts for 25-40% of our median relative error



Pre-peak modelling

- Modelling pre-peak production by predicting the production volume and number of months before the production peak
- Results
  - 8% improvement on the median of the distribution over 5 years
  - 40% improvement over the first 3 months



Thank you