# Applied Machine Learning for Horizontal Wells Risk Assessment Optimization

**Ghaleb Al-Gobi[a], Neenu Achu Eapen[b], Purva Dixitkumar Desai[c], Sai charan S Pendyala[c], Sri Divya Mandadapu[c], Vom Siri Krrishna Jorige[c]**

[a]Department of Petroleum Engineering,University of Houston, Texas, United States
[b]Department of Electrical and Computer Engineering,University of Houston, Texas, United States
[c]Department of Engineering Data Science,University of Houston, Texas, United States

**Abstract**

The success of drilling new horizontal wells depends on connecting the well's surface location with the targeted formation while achieving the desired lateral length. However, the choice of the surface locations can be limited due to difficulties obtaining the permits to drill on these locations. This creates more deviated wellbore designs and decreases the optimized lateral length. To avoid this, a risk matrix will be developed to evaluate the maximum lateral length for different surface locations.

The data base has for the risk matrix has gone under rigorous geometrical analysis. The survey data can have some discrepancies. The analogy was to process the survey of each well, find any discontinuations or bad data of any sort, then find the colosest repair point to reconnect the well. In some cases, the well can be dropped if the survey cannot be repaired.

The first step to creating the risk matrix is to use a geometry-focused torque and drag model. This model will utilize the wellbore's step-out and geometrical torsion, which are resulting of both a deviated well design and a far surface location. In this paper, the torque and drag model will be used to enhance the linear regression model for machine learning. This will allow us to compare the row geometry of designs of the horizontal wells.

The used torque and drag model take a middle position between the soft-string model and the stiff-string model by capturing the geometrical torsion along the wellbore trajectory. This approach enhances geometrical optimization when calculating the axial force and the torque. The simplified mathematical derivation incorporates the step-out into the model, thus allowing for more accurate regressions for the risk matrix to optimize the lateral length based on the surface location.

Running the survey calculation along with the torque and drag model was not cumbersome. It has the simplicity of the soft-string model while taking into consideration the wellbore's geometrical torsion to capture the effect of the well's trajectory on the axial forces and the torque. This effect will partially simulate the deformation of the drill pipe using simplified mathematical expressions compared to the calculations of the stiff-string model. This methodology resulted in a direct correlation between the deviations of the well design, the step out and the lateral length with some exception when some other technical challenges influence the well design.

The use of machine learning models such as linear regression and random forest proved beneficial. The built risk matrix uses regression models and utilizes two mechanisms that determine the dominant constraint that limits the drilling process: either the maximum hook load or the maximum torque, based on the wellbore design. These limits depend on the rig's capacity and the drilling company. The risk matrix summarizes multiple designs to conveniently compare different step-out values and their respective expected later lengths. The risk assessment is

quantitative rather than qualitative and is reflected as the percentage used of the maximum lateral length. The optimum surface location with its step-out is then chosen easily from the created risk matrix.

## Introduction

The technology of horizontal well drilling allowed for efficient extraction of oil reserves comparing to vertical wells. It also allows for better management of the oil field both economically and operationally. However, this drilling technology comes with its own challenges prior to and during the drilling process. Before the drilling of a horizontal well, the well trajectory is designed. This design must show the ability to connect the surface location of the well to the targeted formation. Oil companies do not always have the rights to drill from any surface location. Due to that, the available options for the surface location might not allow to drill to the targeted formation with the desired lateral length. This also can happen when the geometry of the wellbore design is complex. Drilling the well cannot start before overcoming these problems. A geometry focused torque and drag model will allow building a risk matrix to compare the geometrical complexity of the designs with different starting surface points and the required torque and drag for each design. This risk matrix will give a better understanding of the impact of the wellbore undulation and optimize the lateral length of the well.

Many authors and engineers have worked on designing models that aims to solve the problems of the well designs and the lateral length. Most of the resources connect the problem to the torque and drag modeling. A standard notation for the torque and drag is essentially derived based on the equilibrium equations for both the torque and the drag forces (Mitchell & Miska. 2011). However, these equations do not consider the stepout explicitly to estimate the lateral length. Solving the problem numerically is hard and time consuming. This raises the need to use machine learning to study the data and find the best correlations. The step-out concept is shown on (Fig. 1):

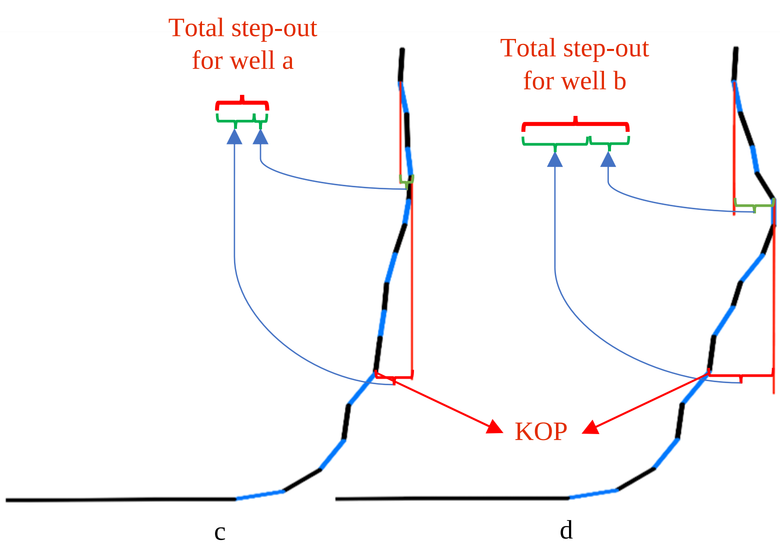



Fig. 1. The step-out distance for different well designs

The geometry of the well design is depicted through the survey of the well. This data can have a lot of discrepancies. Important question must be answered using this data. For example, what is the lateral length and the step-out of each well? What is the location of the KOP, and lateral heel? How to use this data to study the performance of the design? And how does this process impact the problem of horizontal well design and optimization? An example of various well designs is shown on (Fig. 2) . 1. The step-out distance for different well designs.

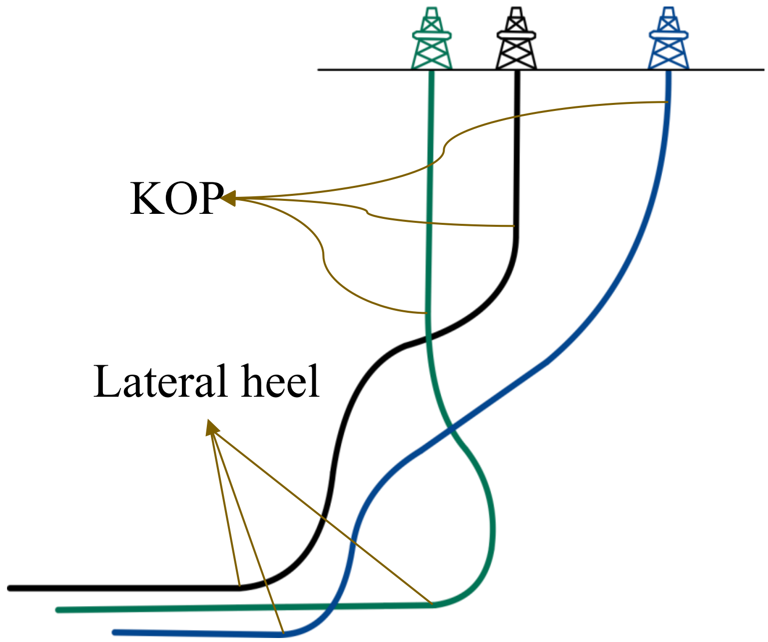



Fig. 2. Different horizontal well designs with different step-out and surface locations

**Data**

The survey data for this study was extracted from the open enverus.com. A set of 713 wells from Wattenberg field in Weld, CO were used in this project. The data had a good quality survey and provided the trajectory for the well. However, some of the wells had sidetracks or the surveys were discontinued. This prompted to review the whole data base. The data had the following columns:

| |
|---|
| WellID |
| StationNumber |
| TVD FT |
| Inclination DEG |
| Azimuth MeasuredDepth FT |
| N S |
| E W |
| GridY FT |
| GridX FT |
| DogLegSeverity DEGPer100FT |

Table 1 raw data columns

The first step to process the data was to identify the location of the KOP, lateral length and the lateral heel. To achieve this, we had to check the condition of the measured depth. west ascending value without interruption? Is there any gaps? What about duplicated and missing data? Thes questions were examined, and the data were refined every step through the way. Next the KOP was determined where the change in two or three points reaches a threshold. Similar analogy was applied for the lateral heel. Finally, the lateral length was determined. After preparing the data, it was then explored to check how the values are distributed. Fig 3. shows the range of the measured depth vs the well count. Most of the well range between 14000 and 18000 ft.

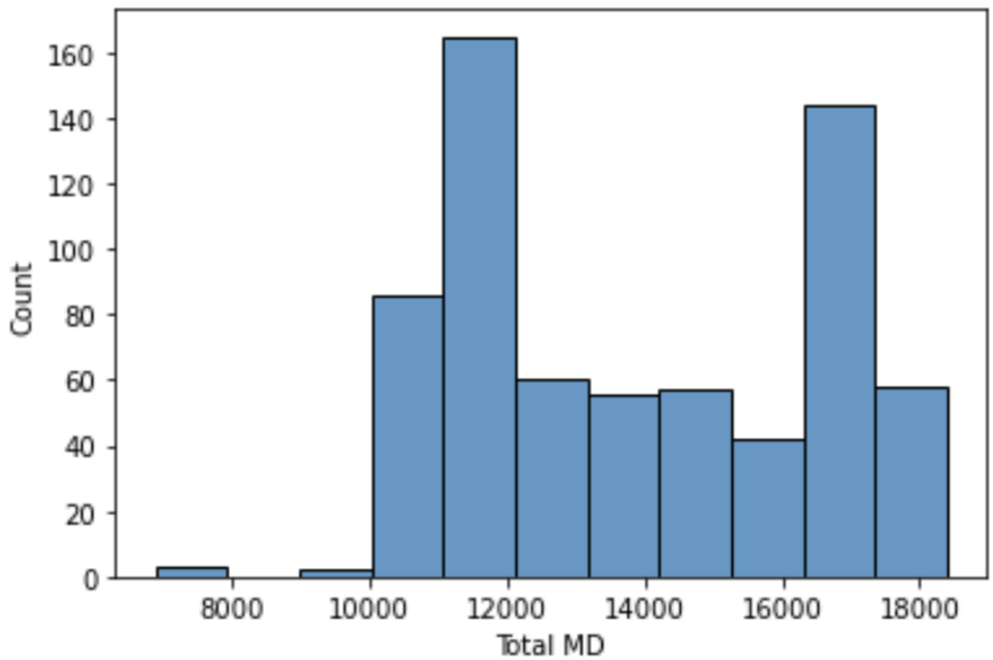



Fig 3. The range of the measured depth vs the well count

The data also show that the wells have a shallow depth. The majority of the wells are between 6500 and 6800 ft deep (Fig. 4). This is a shallow TVD. However, the lateral length sections were large for some of the well. The outliers were not removed from the data, as they did not influence the model. The features are compared then on a confusion matrix. The Correlations between the variable did not show any helpful information when plotted on

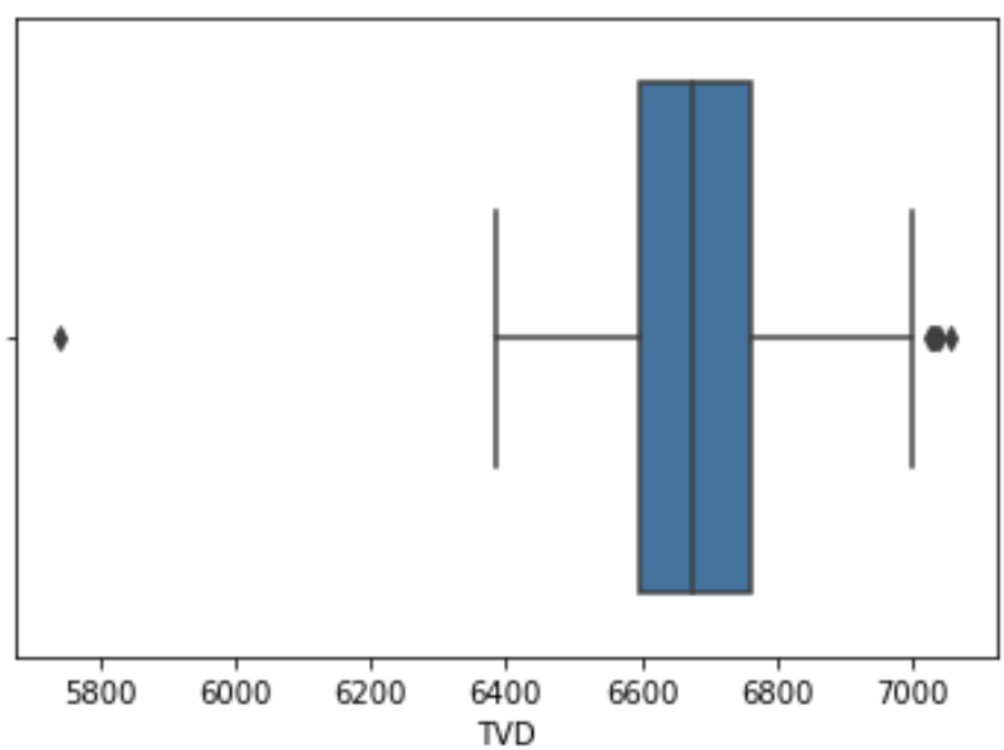



Fig 4. TVD boxplot and outliers

the confusion matrix (Fig. 5). This called for applying the torque and drag model to compare the designs.

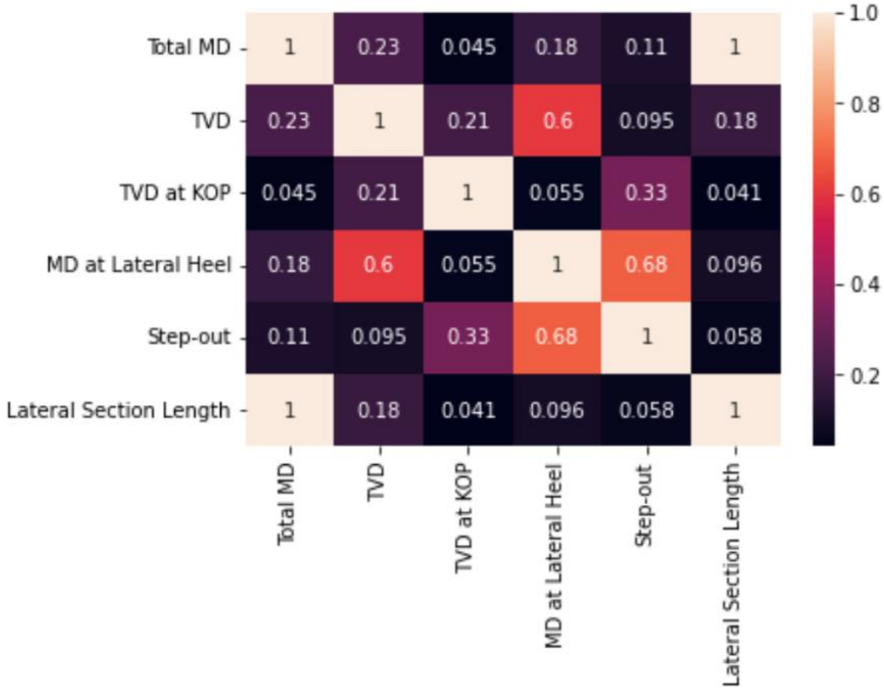
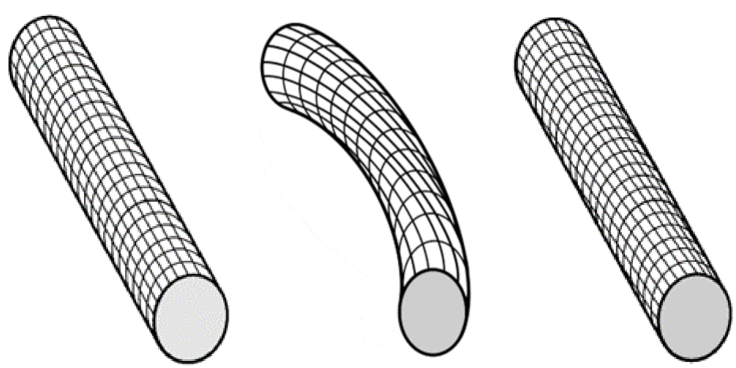



Fig 5. The range of the measured depth vs the well count

**Methodology**

As the data were extracted from one oil field, the geology of the formation was similar for all the wells. It is safe to consider a close or the same friction factor for the well in this data. It is also common to use similar muds when drilling wells in the same oil field. This means the buoyancy will be similar for all the wells. The compromising assumption here is to use the same drill pipe to evaluate the torquer and drag for this model. Doing so, provide a direct comparison of the performance of each well design under the same circumstances. The used torquer and drag model focuses on the design and the geometrical deformation of the surface mesh of the pipe (Fig.6) however it does not capture any material failure. This works well because the material deformation is not within the scope of this study.



(a) (b) (c)

Fig.6 —The geometrical deformation versus the material deformation of the drill pipe: (a) shows an undeformed pipe with undeformed surface mesh. (b) shows a pipe with deformed material and surface mesh. (c)

shows the deformation of the drill pipe represented by a deformed surface mesh to represent the wellbore curvature and torsion in that section while neglecting the material deformation.

The model's base equations are displayed below and shows the relationship between the axial force and the step-out:

$$\alpha = \pm\mu\, W_c + W_{bp}\vec{t}\cdot\vec{k} \tag{1}$$

$$\gamma = W_c\, sin\,\varepsilon + W_{bp}\vec{b}\cdot\vec{k}. \tag{2}$$

$$F_t(l_i, l_{i+1}) = \frac{(\pm\mu\, sin\,\bar{\theta} - cos\,\bar{\theta})}{\kappa^2 + \tau^2}\cdot w_{bp}\tau^2(l_{i+1} - l) \tag{3}$$

$$F_n(l_i, l_{i+1}) = \frac{(cos\,\bar{\theta} \mp \mu\, sin\,\bar{\theta})}{\kappa^2 + \tau^2}\cdot w_{bp}\kappa \tag{4}$$

$$F_b(l_i, l_{i+1}) = \frac{(\pm\mu\, sin\,\bar{\theta} - cos\,\bar{\theta})}{\kappa^2 + \tau^2}\cdot w_{bp}\kappa\,\tau\,(l_{i+1} - l) \tag{5}$$

The performance of each design was tested using the same parameters of the drill string. After that the data was split into testing and training. 135 wells were kept for testing while the rest were used for training. Two machine learning models were used. The first model was the linear regression. It provided an r score of 98% and RMSE of 420.62. This performance is very acceptable and shows a great example of the risk matrix. The picture below (Fig. 7) shows the actual lateral length vs the predicted lateral length by the linear regression:

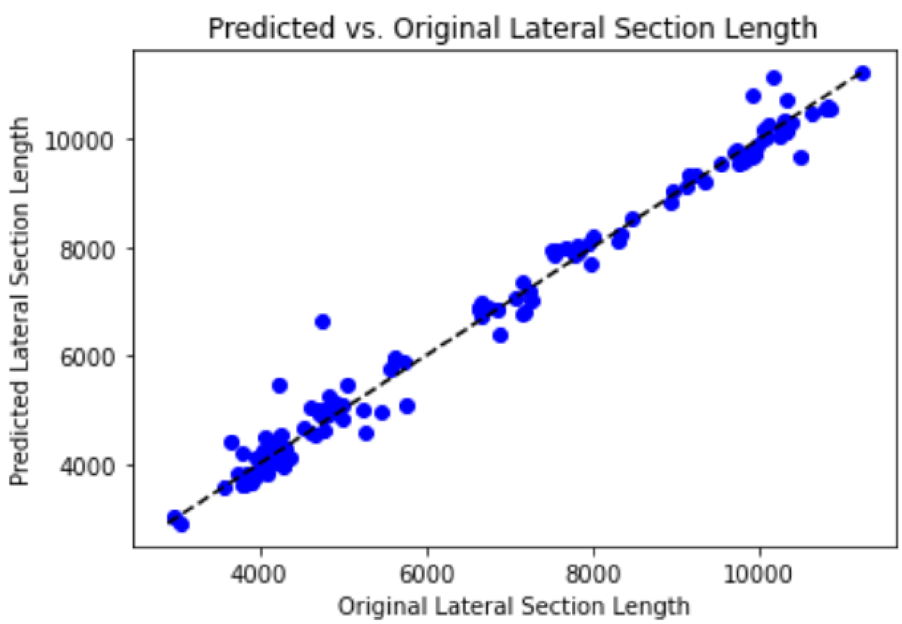



Fig 7. Cross plot of the actual lateral length vs the predicted lateral length using the linear regression.

The data reflected a high correlation between the step-out and lateral section length. There is also occasional uncorrelated data. This is interpreted as a result of technical or reservoir problems that took event during the drilling and influenced the well trajectory.

The second model was random forest which provided a lower accuracy of 94%. And RMSE of 556.26. This model performed well but became unnecessary. However, it also displayed a great performance as shown on the cross plot (Fig. 8) below:

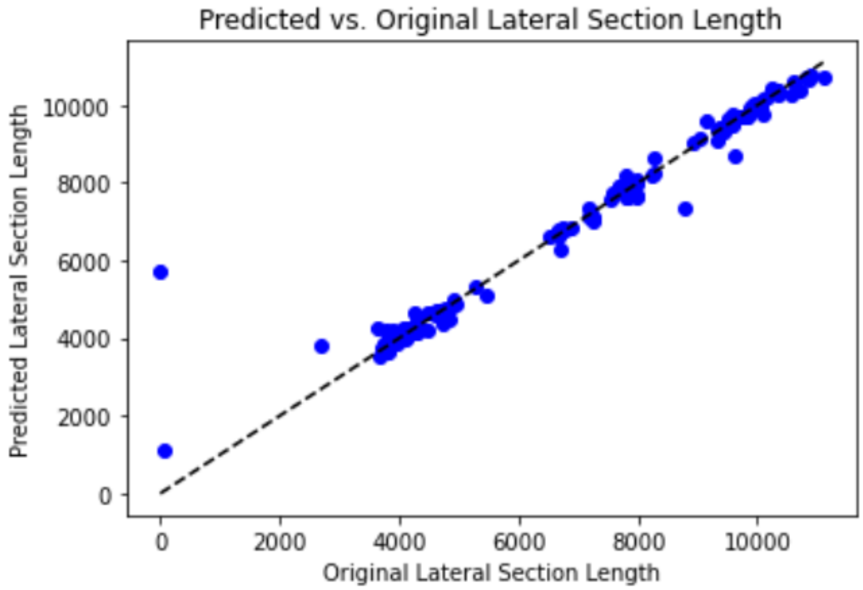



Fig 8. Cross plot of the actual lateral length vs the predicted lateral length using the linear regression.

**Results**

The risk matrix was generated by plotting correlating ranges of the step-out vs the possible ranges of the lateral length based on the performance of the designs on the linear regression that was applied on the model. This risk matrix (Fig. 9) shows the percentage of each lateral se  ction to the maximum lateral section length based on the design perfor  mance. The output displays also the length of each lateral section if desired (Fig. 10).

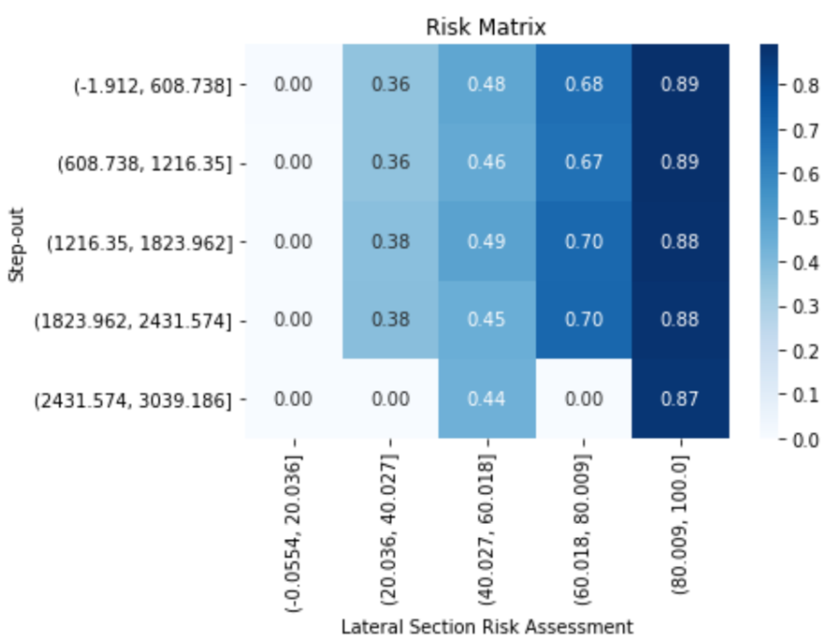



Fig 9. Percentage risk matrix of the step-out vs the predicted percentage of the lateral length.

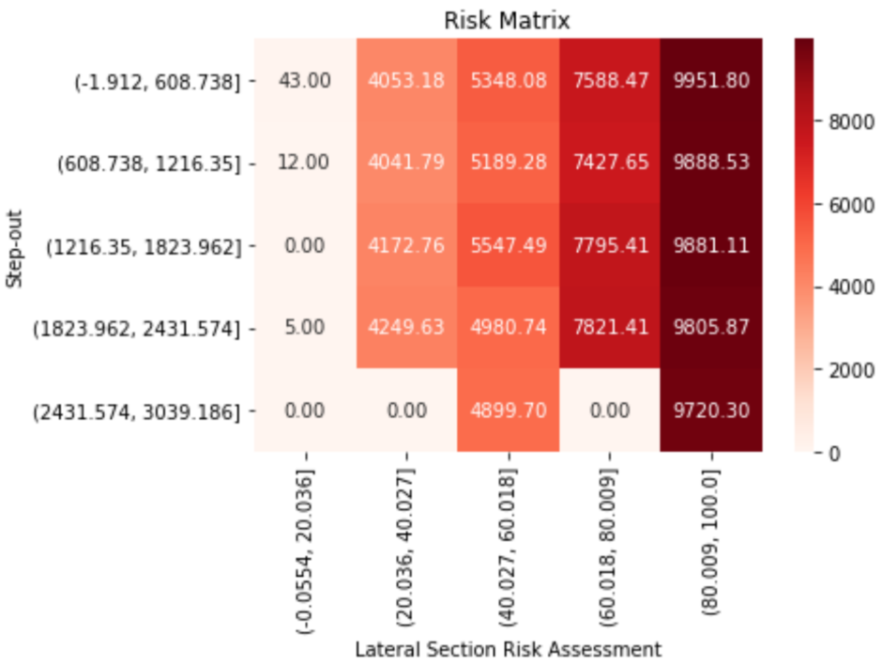



Fig 10. Risk matrix of the step-out vs the predicted lateral length.

It is worth noting that the null or zero squares in on the risk matrix can be populated if the model is trained with more data to reflect the missing range of values.

## Conclusion

The step-out shows a strong indication to reduce the possible lateral length. However the it is not the only factor. In many cases it is possible to achieve a longer lateral length with high torque and hook load limits. The survey data reflects the reaction and the deviations of well trajectory due to the effect on the forces acting on the pipe while drilling. This allowed to compare the designs with similar friction factors and rig capacity/limits. The for the geometry only, the linear regression works best when all the other physical parameters are comparable or the same.

**Nomenclature**

$\vec{\boldsymbol{b}}$ = Unit binormal vector

$\overrightarrow{F_b}$ = Binormal force

$\overrightarrow{F_n}$ = Normal force

$\overrightarrow{F_t}$ = Tangential force

$M$ = Torque

$\vec{n}$ = Unit normal vector

$l$ = Length of trajectory section or arc length

$S$ = Step-out

$\vec{t}$ = Unit tangent vector

$W_{bp}$ = Buoyant weight of the drill pipe

$W_d$ = Drag force

$W_c$ = Contact force

**References**

Mitchell, R. F. (2008, March). Drillstring Solutions Improve the Torque-Drag Model. In IADC/SPE Drilling Conference. OnePetro.

Mitchell, R. F., & Miska, S. Z. (2011). Fundamentals of drilling engineering.

Mitchell, R. F., & Samuel, R. (2009). How good is the torque/drag model?. SPE Drilling & Completion, 24(01), 62-71.

Mohamed, A. (2020). Anti-collision Planning Optimization in Directional Wells. Petroleum & Coal, 62(4).

Mohamed, A., Dahab, A. S. A., & Shokir, E. (2020). Enhancement of Water-Based Mud Rheology and Lubricity Using Silica Nanoparticles. Petroleum & Coal, 62(4).

Samuel, R., & Liu, X. (2009, October). Wellbore Tortuosity, Torsion, Drilling Indices, and Energy: What Do They Have To Do With Wellpath Design?. In Proc. of SPE Annual Technical Conference and Exhibition (Vol. 5, pp. 2831-44).