NEUQUÉN VACA MUERTA PRODUCTION DATA ANALYTICS: IMPACT OF GEOLOGY AND WELL COMPLETION DESIGN PARAMETERS ON WELL PRODUCTION

Peter Zannitto[1], Hugh Grier[1], Hugo Aguirre[2], Jack Trueblood[1], David Delgado[2], Wasim Nasir[3], Francisco García-Goya[2], Brad Speidel[2], Rubén Domínguez[3], John Tolle[1], Scott McEntyre[4], Charles Kosa[4], Ted Cross[4]

1: Shell Exploration & Production Company. peter.zannitto@shell.com, hugh.grier@shell.com, jack.trueblood@shell.com, john.tolle@shell.com

2: Shell Compañía Argentina de Petroleo S.A. hugo-damian.Aguirre@shell.com, david.delgado@shell.com, f.garciagoya@shell.com, b.speidel@shell.com

3: Shell Canada Services Limited. wasim.nasir@shell.com, ruben.dominguez@shell.com

4: Novi Labs. smcentyre@novilabs.com, ckosa@novilabs.com, tcross@novilabs.com

Keywords: Vaca Muerta, black oil, production data, data analytics, data model, type-curve, production forecasting, shale, scaling factors, production normalization

RESUMEN

Si bien la Cuenca Neuquina tiene todos los elementos de una Super Basin, cuando se la compara con otros plays no convencionales norteamericanos, tiene datos limitados de producción en varias ventanas de fluido. Como comparación, hay menos de 100 pozos productores en la ventana de petróleo volátil de Vaca Muerta (VM), mientras que hay más de 6.000 en Eagle Ford. Dentro de los bloques operados por Shell en la ventana de petróleo negro, el control de pozos es todavía menor. Además de modelado dinámico y análisis estándar de datos de producción de pozos, para acelerar las lecciones aprendidas y desrisquear la Fm VM más rápidamente, Shell trabaja con la empresa Novi Labs para construir modelos de análisis de datos para estudiar el impacto del subsuelo y el diseño de parámetros de pozo en la productividad del reservorio. Vaca Muerta presenta variaciones en propiedades del subsuelo como porosidad, saturación de hidrocarburos, contenido de carbono orgánico, madurez térmica, contenido de arcillas, espesor de pay zone, y propiedades de fluidos, además de parámetros de diseño de pozo como longitud lateral, espaciamiento de etapas, número de clusters de perforación por etapa, y carga de fluido y propante. A diferencia de los plays convencionales, se cree que no es posible tener modelos físicos que puedan predecir el rendimiento de pozos no convencionales a partir del modelado de "first principles", que no hayan sido armonizados con el historial de producción o con datos de producción real. El enfoque de análisis de datos de recursos no convencionales busca tendencias de producción en amplios grupos de datos con múltiples parámetros, aislando el impacto y fuerza relativa de cada parámetro.
El equipo técnico de Shell colaboró con Novi Labs para encarar este reto usando "machine learning" (ML) para predecir el rendimiento de producción de pozos nuevos y para profundizar en

la comprensión de los drivers de producción en VM. La colaboración conjunta empleó un enfoque iterativo para construir con éxito y validar modelos de ML que han demostrado precisión en los datos de producción de pozos. Con este esfuerzo, los expertos técnicos de Shell han creado bases de datos completas, configurado modelos de ML, y generado previsiones derivadas de ML para múltiples escenarios de desarrollo de activos. Varias combinaciones de aspectos de subsuelo y de completación fueron utilizadas para entrenar los modelos de ML. Este artículo muestra el proceso colaborativo e iterativo empleado por el equipo, provee ejemplos de la precisión de pronósticos generados con ML comparados con datos de producción real, y resume lecciones útiles claves para repetir y continuar el éxito utilizando ML como una herramienta útil para desarrollo temprano y de avanzada.
Este artículo también expone como Shell utiliza modelos de datos analíticos para generar estudios de adaptación específicos a bloques de petróleo negro de VM que son actualizados anualmente y se utilizan para el Workflow de evaluación de curvas tipo del activo, diseño de pozos de "factoría", y finalmente para "business planning". Los modelos de datos analíticos también han proveído una visión útil para optimizar el Plan de Desarrollo del Activo, en particular, utilizando los modelos de datos para ensayar diseños de fracturación hidráulica de mayor intensidad.

INTRODUCTION

The Neuquén Basin is emerging as a "Super Basin" for unconventional oil and gas development (shown in Figure 1). As such, the Neuquén Basin presents an unprecedented opportunity to apply the best evaluation methods learned in other North American shale plays early in the life of this basin to maximize its full development potential. However, development teams for the Neuquén Basin are faced with the common appraisal challenge of sparse well control. Shell is working with third-party Novi Labs to apply machine learning models and data analytics in combination with physics-based dynamic modeling and standard well performance analysis to simultaneously de-risk and optimize field development planning across the Vaca Muerta acreage. The machine learning techniques identify co-dependent impacts of large variations in subsurface properties with spacing and completion designs to directly predict actual well production. The complex interaction of the variables on shale well production provides a key piece of analysis that increases confidence in full-field optimization.

The combined team Shell technical staff and Novi data scientists applied machine learning (ML) models to generate scaling factors that tie multiple co-dependent variables directly to individual well production. The machine learning models were based on random forest regression algorithms that have been tuned for oil and gas forecasting. The ML models included subsurface features including total porosity, water saturation, organic matter maturity, gross formation thickness, reservoir fluid properties, and total organic carbon, combined with well completion parameters to target oil, water, and gas monthly production rates. The subsurface features represented a blend of reservoir and geology parameters. Additional subsurface parameters including depth, reservoir pressure, and reservoir temperature were explored but ultimately not used in the final machine

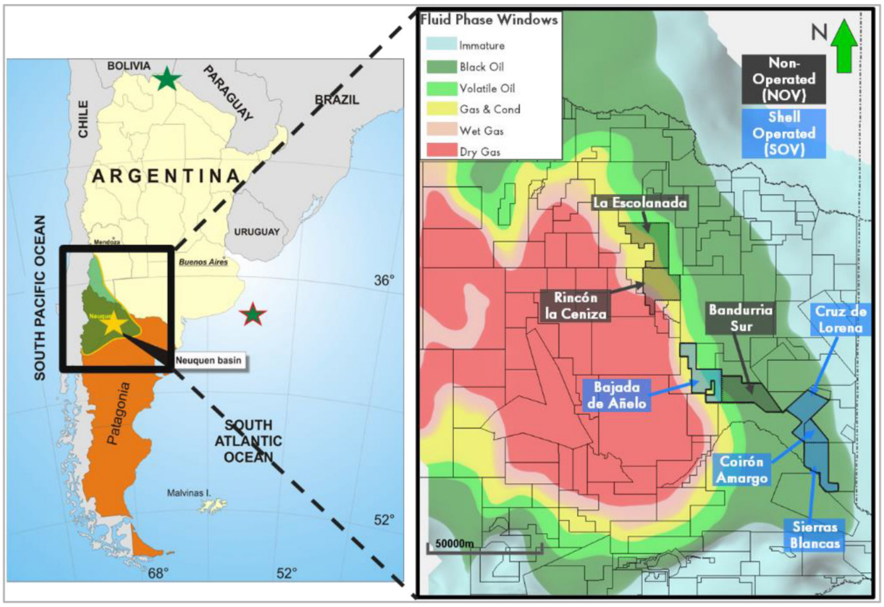


Figure 1. Map on the left displays the Neuquén basin with respect to Argentina. Map on the right shows the Shell-operated venture (SOV) and non-operated venture (NOV) blocks in the Neuquén basin, with respect to Vaca Muerta hydrocarbon fluid windows, ranging from black oil to dry gas.

learning model due to a high degree similarity in the dataset and co-linearity with other features in the dataset. Completion design parameters included lateral length, fluid and proppant intensity, and stage spacing. The production predictions from the machine learning model were validated against actual production using established data analytics practices and an 80% train to 20% test split. Additionally, the machine learning model identified production drivers and analog wells to explain how it predicted well production.

The multi-disciplinary team included Vaca Muerta Subject Matter Experts in Geology, Reservoir Engineering, and Completions Engineering. The iterative approach that was applied is summarized in Figure 2. The team first merged and enriched well data and subsurface data to create a complete ML- ready data set. Next, the team configured multiple ML models, and validated the model forecasts by comparison to actual production and using ML outputs. Finally, the validated ML models were used to forecast specific future well pad locations and cross-basin well inventories to investigate specific scaling factors. This iterative approach integrated machine learning capability with the team's subject matter expertise. Drawing from current well observations and traditional analysis, the team evaluated multiple combinations of features and training data sets.

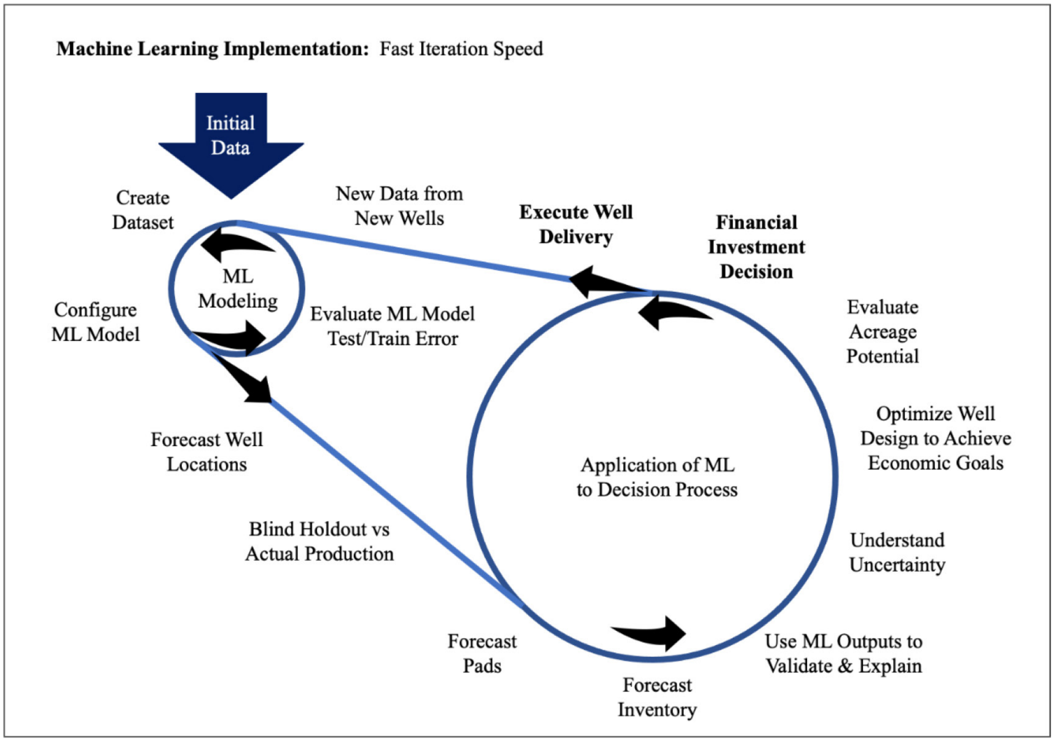


Figure 2. Collaborative and Iterative process used to create, validate, and apply machine learning models on sparse Vaca Muerta well data.

The team took two key steps to validate a ML model that could be used for predicting future well locations. First, the team evaluated the trained ML model accuracy error to the test dataset. To establish credibility in the ML model predictions, the team held out a specific set of wells from the training data set and compared forecasts to both actual production and internal type-curves generated by traditional manual methods. Confirming accuracy to actual cumulative oil production is critical to establish credibility for ML- generated forecasts, especially when also comparing to accuracy of traditional type-curves. Accuracy and credibility are critical to use ML-generated forecasts for informing high value, high impact final investment decisions. Examples of model error measured between test datasets, and representative blind hold out tests are summarized in Table 1 below. Median Absolute Percent Error (MAPE) evaluates model forecast accuracy on an individual well level. A MAPE of 15%, for example, indicates that half of the individual well forecasts are within ±15% of actual production. This is good performance compared to traditional type- curve estimates on smaller datasets. The blind hold-out evaluation most closely resembles how ML models are used to predict future wells. This evaluation involves removing specific wells of interest from the ML training dataset and then measure accuracy of forecasts against actual cumulative production. The representative blind hold-out error is presented in percent error bands to enable comparison to percent error bands normally reported with type curves generated by traditional methods. These blind hold-out error bands vary for different areas of interest and completion designs across the dataset.

Summary table of example ML model errors to actual cumulative production

	90 days	180 days	360 days
MAPE of ML forecasts to test dataset	23%	20%	15%
Representative aggregate error band of ML to blind hold-out areas of interest	+/- 10 to 15%	+/- ~10%	+/- 5 to 10%

Table 1. Summary of Median Absolute Percent Error (MAPE) and percent error from ML-generated forecasts and test or blind hold-out data sets. Error was calculated using cumulative oil production at 90, 180, and 360 days from initial production date.

As a second step to validate the ML models, the team used output data generated by the ML process to investigate and explain the resulting forecasts and gain insight into and confidence in the forecasts. The following example ML model output, shown in Figure 3 below, was used to identify the relative importance and impact of features in this second step. Once accuracy, credibility, and qualitative confidence were confirmed, the team progressed to using the ML model to forecast new well locations in varied rock quality with a varied combination of completion designs.

By using similar model output, the team evaluated which features the model determined to have the highest empirical impact on the forecasts, as well as confirm the directional behavior of each feature. In this model output, each dot represents a specific well value for the feature and the position of the dot indicates the positive or negative impact to the forecast, with 0.0 representing an average well in the data set. In this example, it makes sense that higher fluid and proppant intensities tend to positively impact predicted well oil production. The example in Figure 3 was produced by a model intended to investigate the co-dependent impacts to production from a specific set of geologic and completion features. The dataset used in this model was filtered for lateral lengths. While lateral length was found to be an impactful driver of well performance overall, it was useful to the team to adjust the dataset to focus investigation on the contribution from completion and geologic features for a narrow band of representative lateral lengths for future well planning. This practice demonstrates that application of machine learning models is most useful when it is carefully focused to investigate specific objectives. These model outputs helped to find the most impactful features that created the most explainable ML model with the best accuracy. Further, these model outputs were used during the later exercises to explain how the ML model arrived at well-specific forecasts for future locations.

The machine learning method also provided insight into feature-specific impact on forecasts. Figure 4 shows an example how completion fluid intensity impacts the predicted cumulative oil production volume. By using random forest machine learning methods combined with Shapley production drivers, the team reliably isolated the impact from co-dependent features. Determining the isolated impact of specific features is often non-linear as shown in this example and is perfectly suited to the enhanced random forest type algorithms used by the ML model.

The team then used the results of the multiple forecasted scenarios in a proprietary Shell post-processing parametric process to generate scaling factors for each co-dependent subsurface and completion design feature.

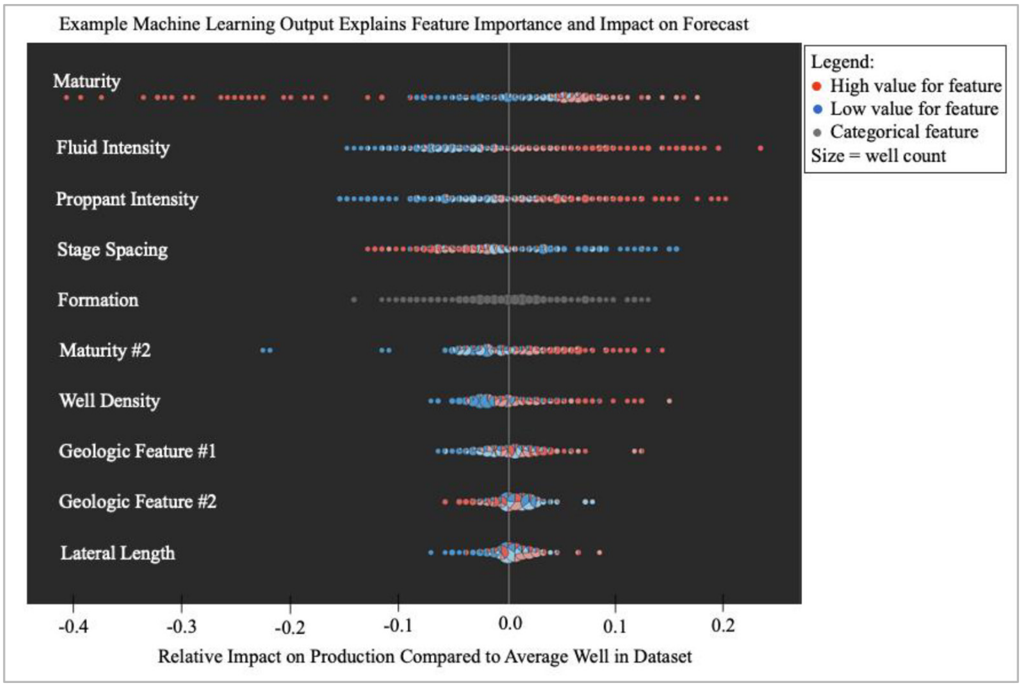


Figure 3. An initial example of Shapley Production Drivers produced by machine learning methods using sparse Vaca Muerta well and subsurface data.

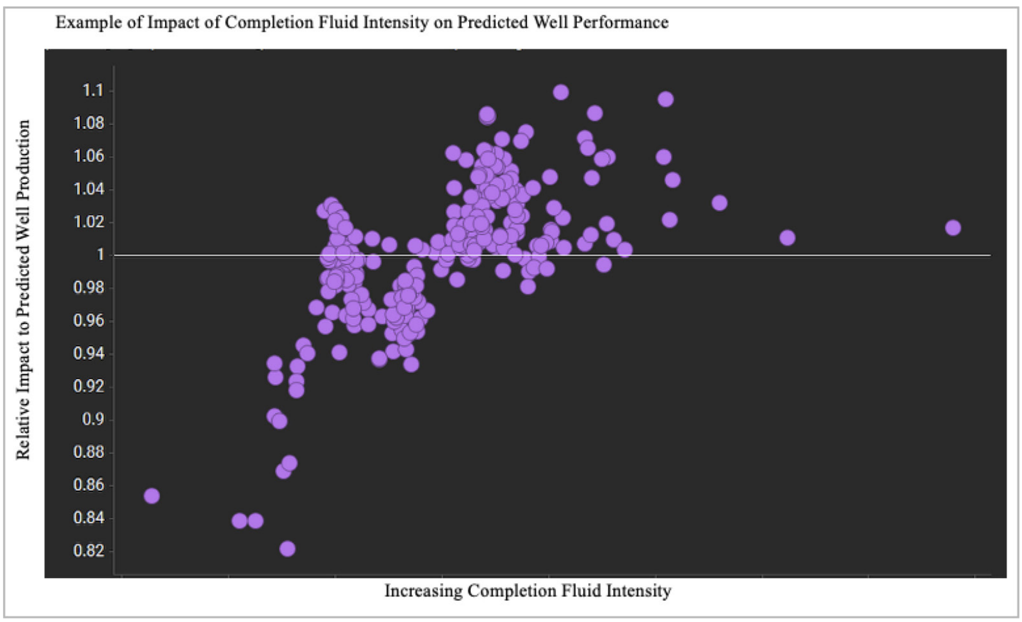


Figure 4. Example of feature-specific impact to production. The machine learning methods produces this data to help explain and validate the ML model.

RESULTS

We investigated the underlying model relationship between the individual subsurface parameters and cumulative oil production in the Vaca Muerta black-oil window. The data underpinning the model was reviewed to ascertain the specific range of well and reservoir parameters where information existed. Since there is higher confidence in data model predictions when inputs are interpolated between existing data points, we narrowed the model input ranges for each parameter to only those seen in the data model. In this way, a Design of Experiments was constructed. The team then ran hundreds of simulations in forecast mode, isolating the impact of each subsurface parameter. Simulations were run across all the parameter ranges using a controlled approach, i.e., keeping all the parameters invariant and varying one at a time over the range of interest.

To quantify the impact and relative strength of each parameter versus oil production, we apply an approach that assumes the subsurface parameters can be decoupled from each other, in the following way:

$$\frac{3\ Year\ Cum\ Oil\ Prod'}{3\ Year\ Cum\ Oil\ Prod} \propto \left(\frac{x_1'}{x_1}\right)^{\beta_1} \times \left(\frac{x_2'}{x_2}\right)^{\beta_2} \times \left(\frac{x_3'}{x_3}\right)^{\beta_3} \times \ \dots \ \times \left(\frac{x_n'}{x_n}\right)^{\beta_n}$$

As mentioned earlier, we limited the Vaca Muerta data model to black-oil wells and only those with at least three years of production history. The equation above relates the ratio of cumulative oil production between two different well/subsurface configurations, using scaling exponents (β_1, β_2 , β_3, β_n) for each of the subsurface parameters (x_1, x_2, x_3, x_n). Here, the "unprimed" variables x_1, x_2, x_3, and x_n all describe a subsurface reference case with a three-year cumulative oil production equal to "3 Year Cum Oil Prod". The "primed" variables x_1', x_2', x_3', and x_n' represent another subsurface realization with a three-year cumulative

oil production equal to 3 Year Cum Oil Prod' that we seek to scale to. The scaling exponent for each of the parameters is given by β_1, β_2, β_3, and β_n.

Using the controlled Design of Experiments mentioned earlier, a log-log (or natural log) plot of model predicted 3-year cumulative oil production versus individual parameters can be generated. The slope of the log-log for each parameter is the β exponent. The process workflow for determining β for each parameter is given in Figure 5. The equation above for scaling between different subsurface/completion realizations may not be exactly correct, since it assumes that all key parameters are independent and can be decoupled from one another. However, it is a good way to approximate the relationships observed in the data model from the Design of Experiments.

Scaling exponents for well lateral length, sand proppant intensity, frac fluid intensity, fracture stage spacing, reservoir thickness (height), saturation, porosity, and total organic carbon (TOC) allowed for normalized production comparisons between otherwise dissimilar black-oil well and Vaca Muerta geologic settings.

Data analytics modelling in other shale plays points to common finding that that well completion parameters are the most strongly impactful to well production. The team first included the impact of the static geologic model parameters and ran simulations for multiple Vaca Muerta black-oil areas. In these cases, Shell's view of the basin geologic model was used in the data models. The data model was run in 4 distinct black- oil areas; Figure 6 shows the forecasted oil production plotted against (1.) proppant intensity, (2.) fluid intensity, and (3.) stage intensity, i.e. number of frac stages in 1 km. When the geologic model was included, a strong correlation with fracture stage intensity (or the inverse of stage spacing) was observed, evidenced by β ranging from 0.7 to 1.2. However, the team was surprised to see no consistent correlation (or only weak correlation) against proppant and fluid intensities when the geologic model was factored into the data model. In these cases, the proppant intensity β ranged from –0.1 to 0 and fluid intensity β hovered around 0.1.

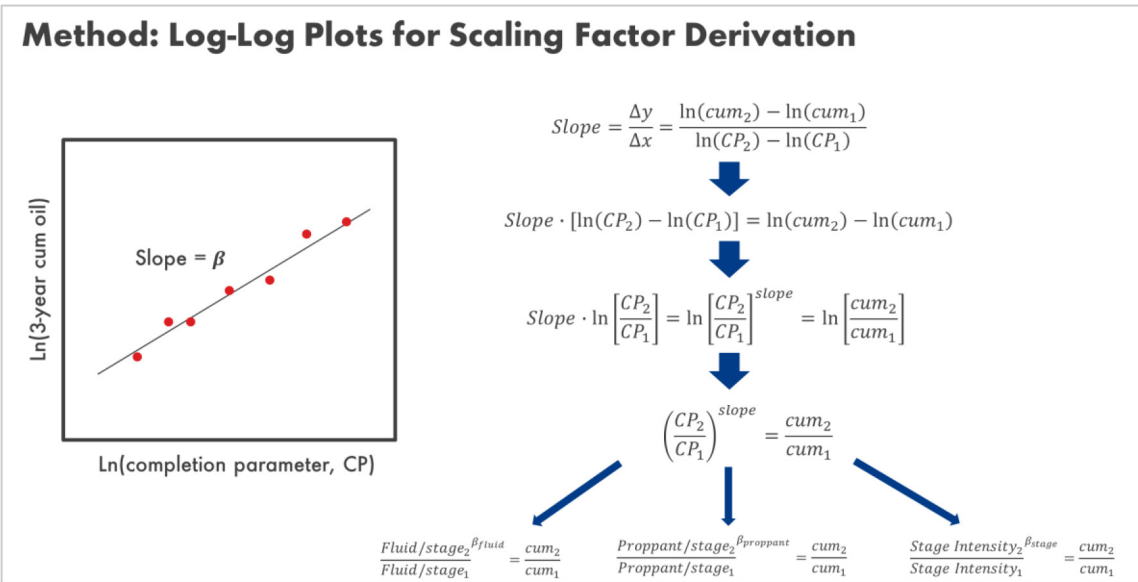
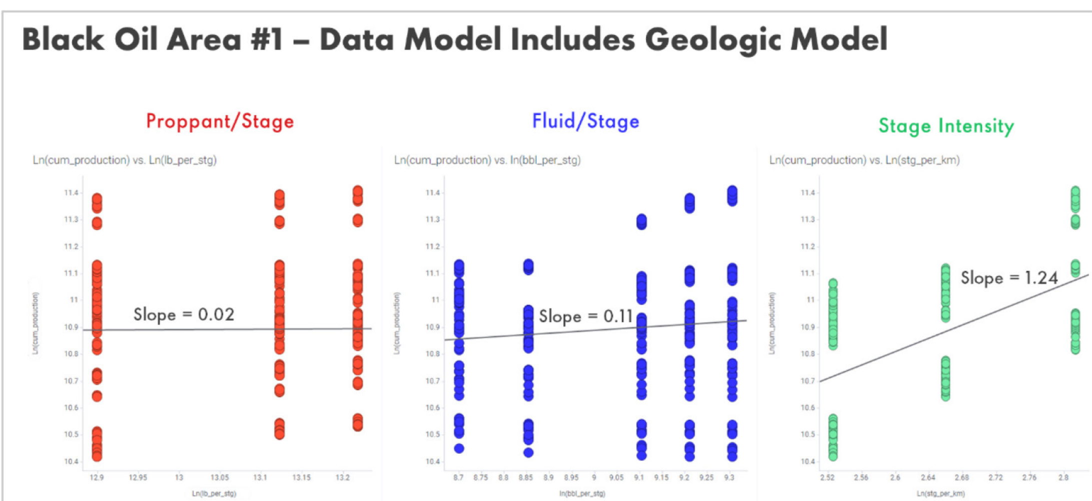


Figure 5. Workflow for Determining the Scaling Exponent β for Each Subsurface Parameter. In the plot on the left-hand side, all of the red points are data model predictions assuming one parameter is varied, and all others are held invariant.



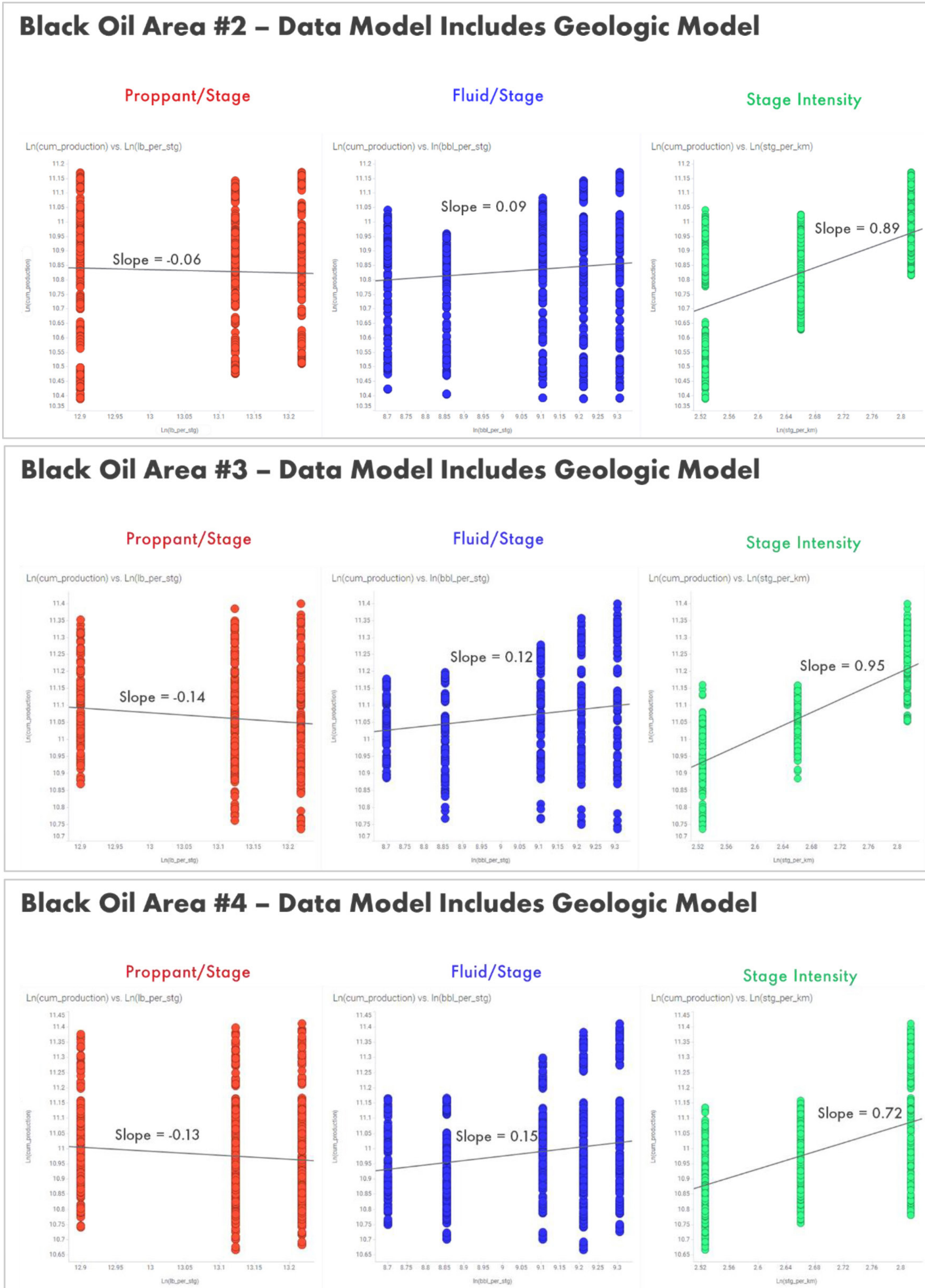


Figure 6. Plot of 4 black oil areas where the geologic model parameters were factored into the data model. For each area, the simulated cumulative oil production is plotted against proppant, fluid, and stage intensities. The slopes on each plot are the scaling exponent β for each parameter. All simulations are run with the same reference case well lateral length.

The team also considered a data model that excluded the impact of geologic model parameters, i.e. porosity, saturation, interval thickness, and TOC. The cumulative oil production variance for the observed range of proppant, fluid, and stage intensities is given in Figure 7. When the geologic model was excluded from the data model's parameter space, we found that the relative strength of proppant and fluid intensity on oil production increased. Here, the best-fit stage intensity β was 0.9, similar to runs that included the geologic model, while both the proppant and fluid intensity β increased to 0.2.

Although, not presented in this paper, these scaling factor findings are aligned with results from data analytics modelling we have undertaken in our other shale developments in North America. Our choice to exclude the impact of geologic model parameters on oil production variance is also based on evidence from our experience in other shale plays. By removing the impact of geologic model parameters and limiting the extent of the data model to a single hydrocarbon fluid window (e.g. black-oil), we have found that the Vaca Muerta production scaling relationships (β) for fracture stage spacing and proppant/fluid intensity are comparable to our findings in other North American shales.

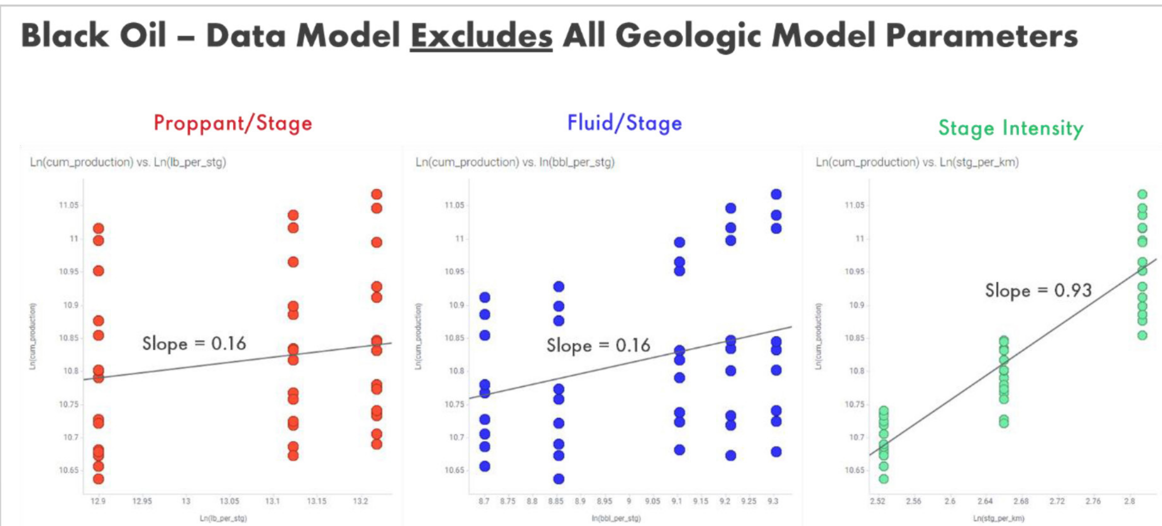


Figure 7. Plot of black-oil data model's 3-year cumulative oil production prediction versus completion design parameters, excluding geologic model parameters. Since the geologic model is excluded from the parameter space, the model predictions are not dependent on location in the basin. The slope for each plot is the scaling exponent β for each parameter. All simulations are undertaken with the same reference case lateral length.

After reviewing the data model results with and without the geologic model parameters, we deemed it more prudent to proceed without the geologic model. There is significant uncertainty in the geologic model properties for the basin, particularly in areas where Shell does not have a working interest. The team also found it surprising that by including the geologic model, the fracture proppant and fluid intensities weakly correlated with oil production. The team intends to update the data analytics model annually and revisit the impact of future updates to geologic model parameters on black-oil production, especially after we add more wells through competitor data trades.

Overall, the Vaca Muerta data model points to proppant, fluid, and stage intensities as being the most impactful parameters that controls oil production. Additionally, we have observed that oil production is correlated to well lateral length and scales linearly for lateral lengths out to ~3000 m, beyond which there is limited data to extrapolate.

In addition to the log-log "slope" regression method to determine parameter scaling exponents, we also evaluated via a different method. Shell has type-curve decline function that is based on a combination of historical well performance, rate-transient analysis, and numerical reservoir simulation. In the alternative approach, we regressed the β exponent for proppant, fluid, and stage intensities for various (simulated) well completion designs to force a match between the data model predictions and our scaled type-curve (that are dependent on β-values). We considered a reference case with proppant intensity PI, fluid intensity FI, and a stage intensity SI that yields a type-curve 3 Year Cum Oil Prod per the equation below.

When we consider other completion designs with a new proppant intensity PI', new fluid intensity FI', and stage intensity SI', we can determine the scaled three year cum oil production (3 Year Cum Oil Prod'). When we consider alternative completion designs, we can tune the β factors for all three parameters to match the scaled type-curve's 3-year cumulative oil production with the data analytics model results. An example of this is demonstrated in Figure 8.

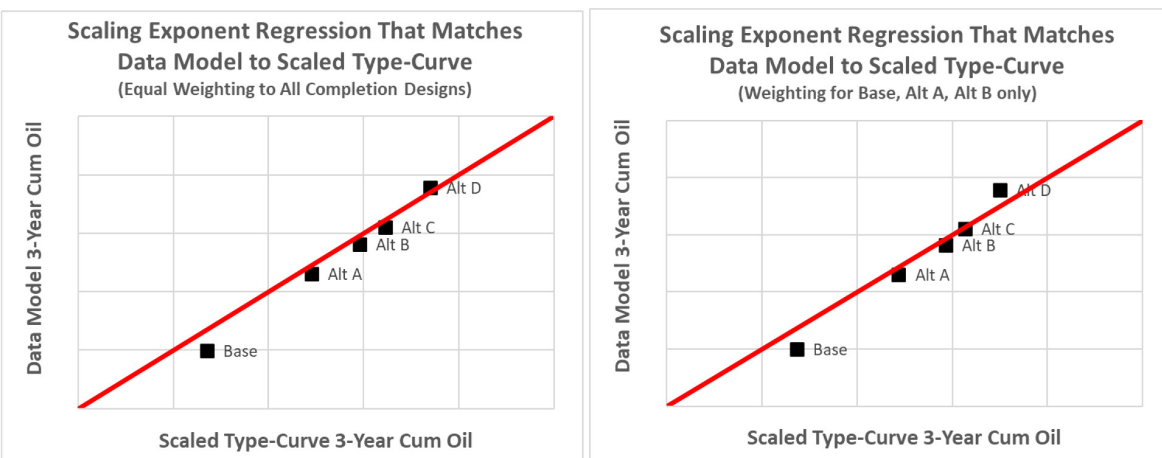


Figure 8. Comparison of β-values regressed to the force a match between the data model cumulative oil production and Shell's scaled type-curve cumulative oil production (dependent on β). Shell's type-curve model was originally based on well performance with the "Base" completion design.

In Figure 8, the red line in both plots is a y = x line; perfect agreement between the scaled type-curve and the data model would result in all of the black points falling on the red line. Each black point in the plots is a specific well completion design that the team analyzed. The plot on the left shows the impact of the β- regression by weighting all completion designs equally. The plot on the right shows the impact of the β- regression by equally weighting only the "Base", "Alt A", and "Alt B" completion designs, and ignoring the "Alt C" and "Alt D" data. The motivation for the weighting factors on the right plot is that three are being considered and trialed by the team, while the other two are not; the team wants to have the β-value regressed to the data model for the three completion designs of interest to greatest degree. Table 2 shows a summary of the results from the right-side plot.

Parameter	Log-Log Linear Regression	Scaled Type-Curve Regressed to Data Model	Average of Two Methods	Applicable Range
Proppant Intensity β	0.2	0.2	0.2	400 - 550 klb/stage
Fluid Intensity β	0.2	0.2	0.2	6 - 11 kbbl/stage
Stage Intensity β	0.9	0.7	0.8	12.5 - 16.7 stages/km

Table 2. Comparison of data model scaling exponents for proppant, fluid, and stage intensities. The average of the two methods, given in the fourth column, is used in all further analysis.

The applicable range in Table 2 is based on the team's confidence in interpolating between existing well completion designs in the available data set. Extrapolating outside these ranges carries greater risk that is not assessed in the model.

A comparison of the two techniques yields similar scaling exponents for proppant and fluid intensity, both equal to 0.2. The data model points to stage intensity as the most impactful parameter, with $\beta = 0.8$. With these results, we can apply scaling factor learnings to determine the corresponding 3-year cumulative oil production for other alternative completion designs, within the data model's existing ranges (see Table 3).

Completion Design	Proppant Intensity (klb/stage)	Fluid Intensity (kbbl/stage)	Stage Intensity (# stages/km)	Ratio of 3-Year Cum Oil to Base Case
Base Case	492	6.9	13.3	--
Alternative X	531	7.9	16.7	1.25
Alternative Y	531	10.2	16.7	1.31

Table 3. Scaling factor impact to 3-year cumulative oil production for three selected well completion designs.

Using the scaling factors from the data model, alternatives completion designs X and Y show a cumulative oil uplift of 25% and 31%, respectively, over the base case design. Assuming that the type-curve for the base case can be scaled to the alternative completion design (ala production multipliers), we can also generate type-curves for other cases. (Note: every new case carries the uncertainty associated with using type-curve decline parameters beyond which they were constructed for.)

Figures 9 and 10 show a comparison of the team's Vaca Muerta black-oil type-curve range of uncertainty, scaled from the "Base Case" completion design to "Alternative X" using the model-derived scaling factors, compared to actual production from wells with "Alternative X" completion design.

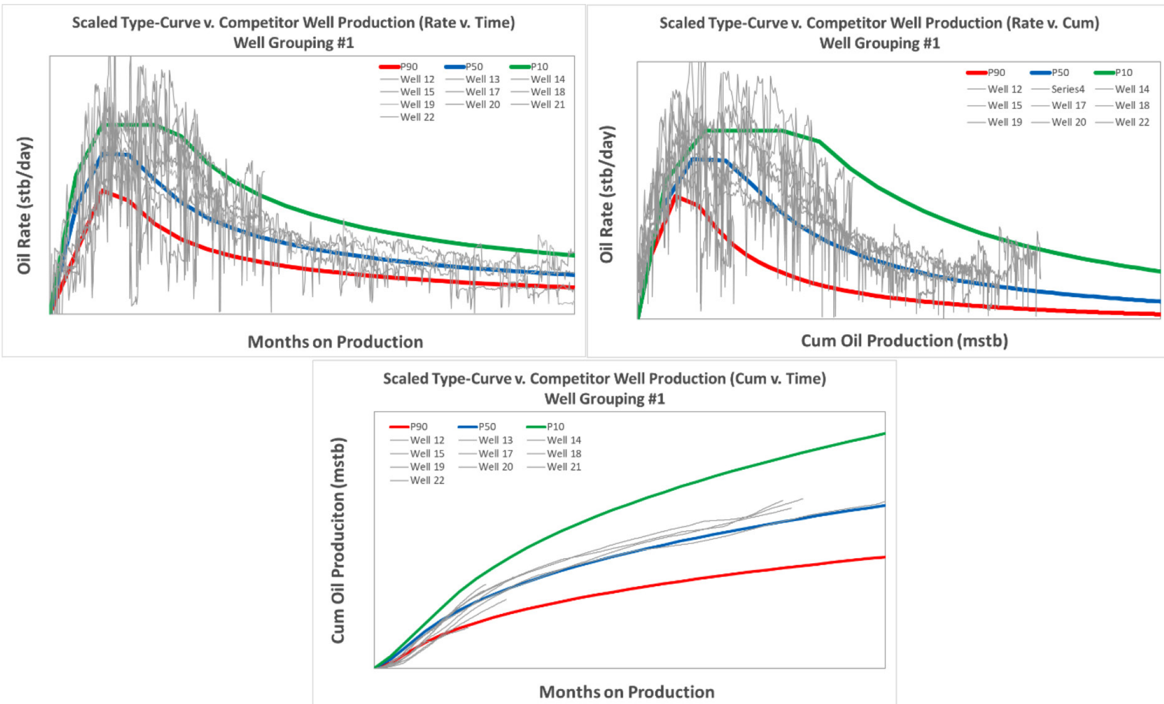


Figure 9. Comparison of "Alternative X" scaled type-curve P90-P50-P10 range of uncertainty against production data for Well Grouping 1. Well Grouping 1 is defined as all the black-oil wells in a distinct geographic area with "Alternative X" completion design.

There is overall good agreement between the data model driven type-curve range of uncertainty and well production with the same completion design. Actual production data for wells in black-oil Grouping 1 is within the P90-P10 uncertainty range of the data model and is centered about the P50 prediction. It is also evident in Grouping 2 (see Figure 10) that the well production is within the model's P90-P10 range, however in this case, the well production exhibits slightly higher oil productivity than the P50 forecast. Further work will be needed to better understand the variance in this case.

Although the data model does a good job describing Vaca Muerta black-oil production ranges across the Basin and for a wide range of well completion designs, potential shortcomings could include: (1.) limited well production (larger sample-sizes needed) and (2.) missing the impact of other key variables that are currently not factored into the data model, e.g. well spacing, geologic model parameters, parent-child well pair attributes, and perforation cluster spacing. These are issues that will be addressed in future models and as more basin well data becomes available.

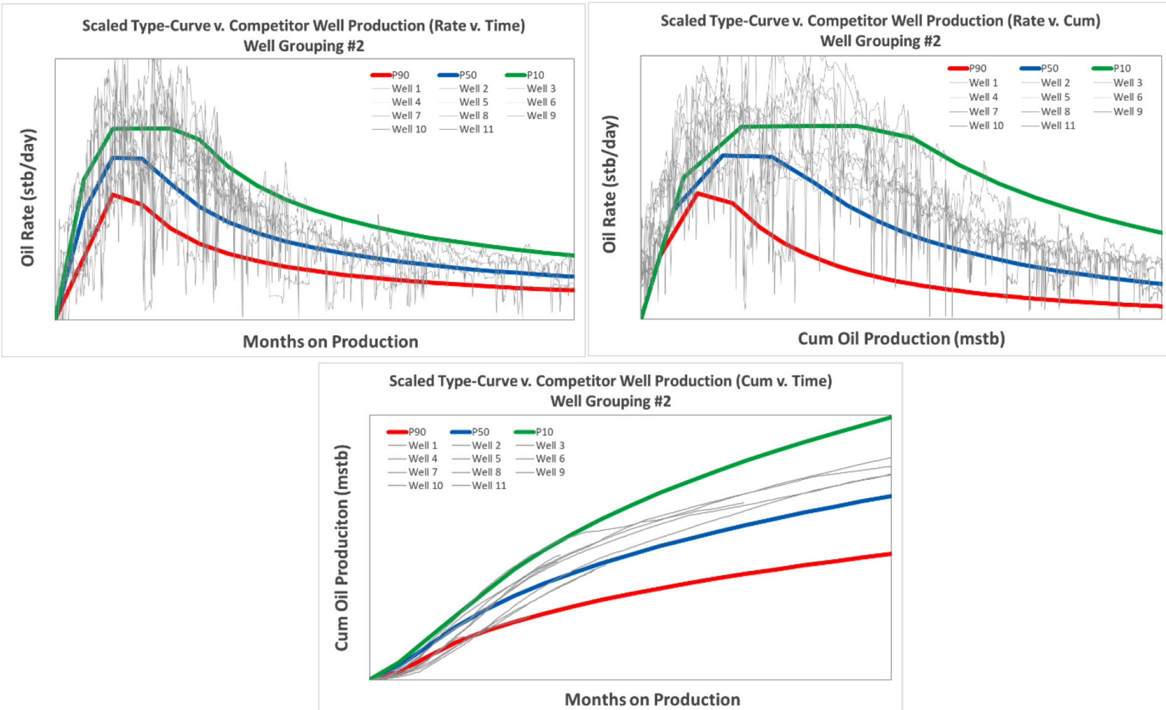


Figure 10. Comparison of "Alternative X" scaled type-curve P90-P50-P10 range of uncertainty against production data for Well Grouping 2. Well Grouping 2 is defined as all the black-oil wells in a separate geographic area with "Alternative X" completion design.

Overall, production scaling (or normalization) that is made possible through data analytics, allows for larger groupings of wells to be analyzed on the same production basis, even if they have dissimilar reference conditions. Good data models can lead to enlarged inventories for performance-based estimated ultimate recovery (EUR) well statistics, provided there is actual data to support interpolating between the various well attributes. Scaling factors derived from data analytics can also assist in type-curve generation process for future undeveloped wells, mitigating the costly learnings and lengthy lead times a single company's go- it-alone appraisal drilling would incur.

CONCLUSIONS

Although the Neuquén Basin has all the elements of a "Super Basin," when compared to other North American shale plays, there is limited well production data across the Vaca Muerta's various fluid regimes. In addition to dynamic modelling and standard well performance analysis, to further accelerate learnings and derisk the Vaca Muerta more rapidly, Shell is working with third-party Novi Labs to build data analytics models to study the impact of subsurface and well design parameters on reservoir productivity. Shell's technical team collaborated with Novi Labs to address this challenge by using machine learning to predict new well performance and deepen insight into production drivers in the Vaca Muerta. The joint collaboration employed an iterative approach to successfully build and validate machine learning (ML) models that have demonstrated accuracy to actual well production.

This paper discussed how Shell is using data analytics models to generate Vaca Muerta black-oil specific scaling studies that are updated annually and feed into the asset type-curve evaluation workflow, well "factory" design, and eventually business planning. The data analytics models have also provided valuable insight toward optimizing the Field Development Plan, in particular, using the data models to trial larger intensity hydraulic fracture designs.

During this collaborative engagement, the team's conclusions can be summarized as follows: (1.) Machine Learning is a useful tool for gaining insight and information from small datasets; (2.) Iterative approach using ML models can accelerate development optimization in new basins; (3.) Empirical ML processes complement traditional analysis to create higher confidence; and (4.) Possible to achieve accurate ML- generate forecasts with technical focus and fast iteration.

The early results from this effort demonstrate the usefulness of applying machine learning with sparse data sets. As a tool, ML provided the capability to extract maximum information and insight from all available well and subsurface data. Additionally, the iterative process based around ML methods both improved early-time well performance evaluation as well as enhanced the ability to quickly assimilate new well production into the process, thus accelerating efforts to optimize well designs and understand production drivers in the Neuquén Basin.

Future work will involve filling in perceived data model gaps, for example: (1.) horizontal and vertical spacing distance between existing wells, (2.) number of perforation clusters per frac stage, (3.) parent-child attributes, (3.) actual well landing zone information, (4.) vertical frac height, and (5.) improved geologic static models. The team envisions annual data model updates to capture internal and competitor learnings to better guide Shell's field development plan.

As the Vaca Muerta progresses from early appraisal phase toward full-scale development, an emerging future challenge is to optimize well density. In the future, building on the validated application of ML methods, the technical team can potentially determine optimal well spacing faster by experimenting with data instead of expensive trial-and-error drilling and completion pilot studies. ML methods show promising results to accelerate this next step in Vaca Muerta development. Overall, this approach can enhance total value potential in the Neuquén Basin.

REFERENCES

Al-Alwani, M., Dunn-Norman, S., Britt, L., Alkinani, H., Al-Hameedi, A., Al-Attar, A., Trevino, H., and Al-Bazzaz, W., 2019, Descriptive Data Analytics for the Stimulation, Completion Activities, and Wells' Productivity in the Marcellus Shale Play, Asia Pacific Unconventional Resources Technology Conference, Brisbane, Australia, November 18-10, Conference Paper, https://doi.

org/10.15530/AP-URTEC-2019- 198290

Argentina Government Secretariat for Energy. Non-Renewable Resources and Fuel Market Secretariat. Undersecretary for Hydrocarbons and Fuels. National Directorate for Exploration and Production, Public Data Set "Oil and Gas Production per Well (Chapter IV)" website, access at: http://datos.minem.gob.ar/dataset/produccion-de-petroleo-y-gas-por-pozo

Barree, R., Cox, S., Miskimins, J., Gilbert, J., and Conway, M., 2014, Economic Optimization of Horizontal Well Completions in Unconventional Reservoirs, SPE Hydraulic Fracturing Technology Conference, February 4-6, The Woodlands, Texas, SPE-168612-MS, https://doi.org/10.2118/168612-MS

Blackwood, K., Handren, P., Chapman, M., Palisch, T., Godwin, J., 2011, Production Acceleration or Additional Recovery? A Look Back at Three Published Field Trials to Determine the Long Term Benefits of Improved Fracture Treatments, North American Unconventional Gas Conference and Exhibition, SPE- 144349-MS, https://doi.org/10.2118/144349-MS

Curtis, T., and Montalbano, B., 2017, Completion Design Changes and the Impact on US Shale Well Productivity, Oxford Institute for Energy Studies Energy Insights, v. 21

Kolawole, O., and Ispas, I., 2020, Interaction between hydraulic fractures and natural fractures: current status and prospective directions, Journal of Petroleum Exploration and Production Technology, v. 10, p. 1613–1634 https://doi.org/10.1007/s13202-019-00778-3

Liang, F., Sayed, M., Al-Muntasheri, G., Chang, F., and Li, L., 2016, A comprehensive review on proppant technologies, Petroleum, v. 2, no. 1, p. 26-39, https://doi.org/10.1016/j.petlm.2015.11.001

Lundberg, S., Nair, B., Vavilala, M., Horibe, M., Eisses, M., Adams, T., Liston, D., Low, D., Newman, S., Kim, J., and Lee, S., 2018, Explainable Machine-Learning Predictions for the Prevention of Hypoxemia During Surgery, Nature Biomedical Engineering, v. 2, no. 12, p. 749-760, https://doi 10.1038/s41551-018-0304-0

Lundberg, S., Erion, G., and Lee, S., 2018, Consistent Individualized Feature Attribution for Tree Ensembles, arXiv, accessed May 25, 2020, https://arxiv.org/pdf/1802.03888.pdf

Srinivasan, K., Ajisafe, F., Alimahomed, F., Panjaitan, M., Makarychev-Mikhailov, S., and Mackay, B., 2018, Is There Anything Called Too Much Proppant? SPE Liquids-Rich Basins Conference, Midland, Texas, Conference Paper, SPE-191800-MS, https://doi.org/10.2118/191800-MS

Williams-Derry, C., Hipple, K., and Sanzillo, T., 2020, Shale Producers Spilled $2.1 Billion in Red Ink Last Year, Institute for Energy Economics and Financial Analysis report, accessed May 25, 2021 at ttps://ieefa.org/wp-content/uploads/2020/03/Shale-Producers-Spilled-2.1-Billion-in-Red-Ink-Last- Year_March-2020.pdf