URTeC: 4254307

Quantifying Spacing Degradation in Northern Delaware Basin Using Machine Learning and Subsurface Properties

Charles Connell[*1], Kyle LaMotta[1], Jizhou Li[2], Matt Mayer[3], 1. PetroAI, 2. ExxonMobil, 3. TGS.

This paper was prepared for presentation at the Unconventional Resources Technology Conference held in Houston, Texas, USA, 9-11June 2025.

Abstract

This paper presents an automated modeling system designed to predict well performance in unconventional oil development, incorporating geological and geomechanical property variations in the Delaware Basin, particularly in Lea and Eddy Counties. The approach addresses the challenge of quantifying spacing degradation effects within heterogeneous reservoirs, where traditional analysis methods often reveal noisy relationships between well spacing and performance.

The methodology integrates diverse datasets including well logs, directional surveys, and production data, enabling data-driven analysis that minimizes interpretation biases. The process consists of multiple steps: first, generating machine learning ready well logs, then interpolating those well logs to create property maps, then developing a spatial clustering framework to identify distinct regimes within the Wolfcamp formation. Further analysis is done to quantify the critical distances at which well interference significantly impacts production in each cluster.

The research culminates in a spacing-aware machine learning model that achieved 78.5% accuracy in predicting normalized production outcomes. Sensitivity testing across varying well densities (1-12 wells per section) in different geological settings demonstrated variations in production degradation patterns across clusters. The findings reveal between 10% and 25% degradation at five wells per section, compared to a standalone well depending on which cluster the wells are placed. This approach equips engineers with quantitative tools to optimize well spacing decisions, prioritizing location-specific geological characteristics over basin-wide trends.

Introduction

The Delaware Basin remains one of North America's most prolific unconventional oil plays, yet optimizing field development is complicated by its complex geological characteristics. In Lea and Eddy Counties, the epicenter of development activity, rapid shifts in geological attributes, coupled with an almost 90-degree change in regional stress orientation, create significant challenges for production modeling. These complexities demand more advanced analytical approaches beyond conventional methods.

Among the many variables influencing development success, well spacing stands out as the single most critical driver of economic performance, impacting both capital efficiency and ultimate recovery. However, the relationship between well spacing and production remains frustratingly opaque, with traditional datasets displaying scattered correlations that defy simple interpretation (see Figure 1). This uncertainty often forces engineers to rely on anecdotal evidence, selective case studies, and subjective interpretations, approaches that inevitably introduce human bias into development decisions worth hundreds of millions of dollars. Purely data-driven approaches like those described by Zhou et al. 2019 and Ren et al. 2023 can provide accurate predictions, but struggle to pick up on the nuances across a range of well spacing scenarios.

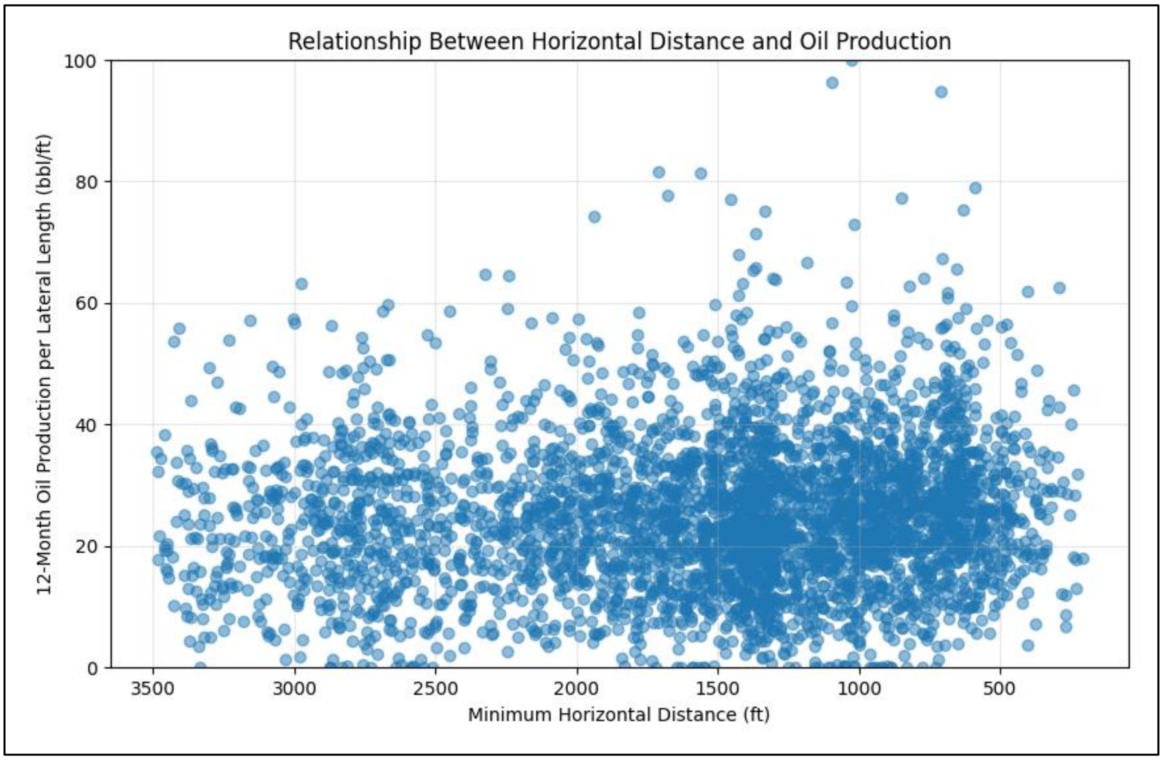

Figure 1: Production, as represented in the figure as 12-mon cumulative oil per ft has no correlation with offset well distance in the training data set.

The industry urgently requires an objective, data-driven framework for optimizing well spacing, one that isolates spacing effects from the numerous other variables influencing production. Current methodologies often involve extensive reservoir characterization studies, which are time-consuming, costly, and still leave room for subjective interpretation. Meanwhile, operators continue making high stakes spacing decisions based on incomplete information.

This paper introduces a novel, automated modeling system designed to transform well performance prediction in heterogeneous unconventional reservoirs. By integrating machine learning techniques with rigorous geological characterization, our approach quantifies spacing degradation effects with unprecedented spatial resolution. Unlike traditional basin-wide generalizations, this system identifies distinct geological regimes, each with unique spacing-performance relationships.

Our methodology deliberately minimizes interpretation biases by allowing the data to reveal natural clustering patterns and interference thresholds. The resulting model enables operators to make truly data-driven decisions about optimal well spacing; decisions that can dramatically enhance capital efficiency and ultimate recovery across the Delaware Basin. This approach represents a fundamental shift away from intuition-based development strategies, leveraging comprehensive data integration for precise spacing optimization.

Methods

To evaluate the impact of well spacing on production performance in unconventional wells within the Delaware Basin, we implemented a multi-step approach combining geospatial analysis, machine learning, and production data modeling.

This methodology involved aggregating diverse datasets to create geologic property grids, clustering wells into distinct geological regimes, and developing production degradation curves. A "spacing-naïve" machine learning model guided clustering, while breakpoint analysis quantified critical distances for well interference. Subsequently, a "spacing-aware" model integrated geological and spacing features to predict normalized production outcomes and simulate varying spacing scenarios.

The following sections detail each step of this workflow, including geologic inputs, clustering methodologies, and predictive modeling techniques.

Geologic Inputs

To ensure the model captures true geological controls on production rather than artifacts from inconsistent inputs, it integrates two key datasets: a basin-wide stratigraphic framework and a complete, standardized set of well logs.

Volumetric features from well logs were integrated into the model to represent formation characteristics and stratigraphic architecture. A significant challenge with well data is the frequent occurrence of missing curves or incomplete depth intervals, especially in critical logs like density and sonic. To address this, a basin-scale machine learning imputation workflow was developed to predict missing values across essential curves. Extensive data preprocessing, including curve categorization, unit standardization, log splicing, spatial normalization, and bad-hole detection - ensured consistent inputs across thousands of wells (Gonzalez et al. 2023). This enhanced depth coverage and property resolution for robust production modeling.

In parallel, a high-resolution stratigraphic framework was constructed, incorporating 34 regionally correlated horizons across the Bone Spring and Wolfcamp formations from over 55,000 digital well logs. Stratigraphic grids preserved vertical resolution and lateral continuity, accurately tying wells to specific benches and reservoir intervals. This framework supported volumetric and geomechanical calculations, enabling the ML model to differentiate spacing-related performance from stratigraphic variability.

Incorporating Well Data into a Modeling Pipeline

Our modeling framework integrated multiple data sources to represent the Delaware Basin's subsurface. We incorporated well header data, structural grid information (formation tops), and triple-combo well log data to capture reservoir heterogeneity. Regional geological and petrophysical property grids, which would typically enhance machine learning workflows by capturing spatial trends beyond well locations, were not available for this study.

Well header data containing surface latitude and longitude coordinates was sourced from an internal corporate database. Wells were selected based on data availability and relevance to the project scope. These surface coordinates served as the foundation for querying a proprietary subsurface model through an internal API, which retrieved structural information including surface elevation, true vertical depth (TVD), and true vertical depth subsea (TVDSS) across multiple reservoir layers. The resulting data was parsed into a structured, tabular format suitable for machine learning workflows.

For petrophysical characterization, we compiled well log data and formation tops from a regional dataset. This involved extracting log measurements, well locations, and depth information from standardized database files. To ensure spatial consistency, all coordinate reference systems were converted from EPSG:4413 (NAD27/Texas Central) to EPSG:4326 (WGS84), while depth values were standardized to true vertical depth (TVD).

Structural horizon data were interpolated using cubic interpolation and refined with Gaussian smoothing to generate continuous property surfaces from discrete well control points. We focused our analysis on key reservoir intervals, including the Leonardian, Bone Spring, Wolfcamp, and Pennsylvanian-aged units. The final dataset comprised spatially aggregated log measurements and interpolated grid attributes, providing the foundation for subsequent spatial analysis and machine learning model development.

Well Clustering Methodology

Our study required a systematic approach to identify clusters that capture both spatial distribution and key geological/completions characteristics influencing well spacing and interactions. We developed a comprehensive clustering methodology to address this challenge.

The clustering analysis incorporated a diverse set of features to capture the complex interplay between geology, completions design, and spatial positioning on well spacing. We selected eleven key parameters: petrophysical measurements including density log (DEN), neutron porosity (NEUT), resistivity (RES), and sonic log (SON) values; completions parameters including total proppant per perforation length and total fluid per perforation length; structural positioning through true vertical depth subsea (tvdss); and stress orientation represented by the angle from SHmax. Rather than use the raw values for these attributes, a machine learning model was trained on the PDP wells in the area using these features. The target variable of our clustering model was 12-mon cumulative oil, normalized by lateral length. The resulting SHAP (SHapley Additive exPlanations) value for each feature was then used. We wanted to capture how these features interacted with each other. The petrophysical features provide insights into rock quality and reservoir characteristics. The completions parameters capture the intensity and design of hydraulic fracturing treatments. The inclusion of spatial coordinates ensures geographical coherence within clusters, recognizing that proximity often correlates with similar geological conditions.

SHAP values quantify the contribution of each input feature to a model's prediction, allowing for a nuanced understanding of feature importance. By transforming raw features into SHAP values, we highlight not just individual effects but also how features interact and influence the model's output in combination, enhancing interpretability in complex, non-linear models (see Lundberg and Lee, 2017).

We began by filtering the dataset to include only wells from target intervals within the Wolfcamp formation, focusing specifically on the Wolfcamp X, Y, Shale, and B zones. This stratigraphic filtering ensures that the resulting clusters reflect meaningful geological distinctions within comparable formations.

We implemented a balanced K-means clustering approach designed to maintain statistical robustness while capturing meaningful distinctions. The clustering algorithm operates separately within each interval group, respecting the unique characteristics of different stratigraphic zones. Our method enforces a minimum cluster size of 250 wells to ensure statistical significance in subsequent analyses of spacing degradation. The algorithm targets an optimal number of clusters determined through silhouette score analysis, which measures how each well matches its assigned cluster compared to other clusters.

The inclusion of spatial coordinates (latitude and longitude) in the feature set ensures that the resulting clusters exhibit geographical coherence. This spatial awareness is critical for understanding how geological trends and completions practices vary across the study area.

We employed principal component analysis (PCA) to reduce dimensionality and visualize cluster separation in two dimensions. This visualization confirms that the identified clusters represent distinct well groups in the multidimensional feature space. We generated comprehensive cluster statistics for each feature, analyzing mean values and standard deviations to characterize the distinguishing attributes of each cluster. These statistics reveal how geological and completions factors vary between clusters, providing insights into what makes each cluster unique. Spatial mapping of the clusters verifies their geographical coherence and reveals regional patterns in well characteristics. These maps display wells colored by cluster designation, overlaid on base maps to provide geographic context. The spatial distribution of clusters often aligns with known geological trends and operational practices in different parts of the basin.

Production analysis across clusters examines how the identified geological and completions groupings relate to well performance. Box plots of normalized 12-month oil production (normalized by lateral length) reveal performance differences between clusters, validating the practical significance of the clustering approach.

The resulting clusters represent distinct combinations of geological and completions characteristics that are expected to exhibit different well interaction patterns. By grouping wells with similar attributes, we create a foundation for more nuanced analysis of spacing degradation. Each cluster effectively controls multiple variables simultaneously, allowing for more accurate isolation of spacing effects. This approach recognizes that spacing degradation is not uniform across geological settings or completion designs but rather varies based on the complex interplay of multiple factors. The clusters enable targeted spacing analysis within groups of wells that share similar characteristics, reducing confounding variables and increasing the precision of degradation estimates.

Breakpoint Analysis

Following the identification of spatially aware well clusters, we developed a systematic approach to quantify how each cluster responds to changes in well spacing. This analysis aims to identify critical threshold distances where well interactions begin to significantly impact production, representing the effective fracture half-length within each geological and completions context.

Our analysis utilizes 12-month cumulative oil production per foot as the primary productivity metric. The spatial relationship between wells is quantified using the minimum horizontal distance between offset wells. This distance metric provides a direct measure of the proximity between adjacent wellbores, capturing the potential for fracture interference and drainage area overlap.

Prior to analysis, we implemented a robust data cleaning procedure. Wells with missing values in critical fields were removed. We filtered out physically impossible values such as negative distances or production figures. To reduce the influence of outliers that could skew the analysis, we applied interquartile range filtering, removing wells with normalized production values beyond 1.5 times the interquartile range from the 5th and 95th percentiles.

Well spacing effects typically follow a non-linear pattern, with production changes becoming increasingly pronounced as wells move closer together. To capture this behavior, we employed a logarithmic function of the form:

$$Production/_{Length} = a + b * \ln(distance) \quad [1]$$

This functional form aligns with theoretical expectations of well interference: minimal impact at large distances with gradually increasing influence as wells approach each other, followed by a steep degradation at very close spacing.

For each cluster, we fit this logarithmic function to the normalized production and distance data using non-linear least squares optimization. The fitting process provides two key parameters: the intercept (a) and the slope coefficient (b). These parameters characterize the overall production level and the sensitivity to spacing changes, respectively.

The critical challenge in spacing analysis is identifying precisely where significant well interference begins. Rather than relying on visual inspection or arbitrary thresholds, we developed a derivative-based approach to objectively locate this transition point. Our method analyzes the rate of change in the fitted logarithmic curve by examining its derivatives. The first derivative (b/x) represents the instantaneous rate of production change with respect to distance. The second derivative ($-b/x^2$) captures how that rate of change itself is changing, providing insight into the acceleration of interference effects. We establish a threshold based on the range of the second derivative values, identifying regions where the curve exhibits significant curvature. The breakpoint is defined as the distance where the second derivative exceeds this threshold, representing the transition from minimal interference to significant production impact. This approach effectively filters out the noise present in raw data while highlighting the underlying relationship between spacing and production. The resulting breakpoint provides a data-driven estimate of the effective fracture half-length for each cluster, indicating the distance at which well interference becomes meaningful.

This leaves the need to select the second derivative threshold. The raw data is very noisy and there is no clear indication of where this threshold should be set. We worked with an Eddy County operator who had done a more traditional spacing study, looking at frac simulations, RTA analysis, analog wells, and reservoir models to determine the expected degradation at one location in the county. This gave us the expected answer in one of the clusters. We picked a threshold that matched this work. By honoring the spacing degradation in one cluster, we are able to generate spacing degradation curves in all cluster.

Well Performance Models

The identified breakpoints serve as critical inputs to subsequent predictive modeling of well performance. Each breakpoint represents the effective fracture half-length for its respective cluster, capturing both the geological characteristics and completions design factors that influence fracture propagation and well interaction.

We now go back to feature engineering to prepare a spacing-aware model that predicts well performance. We use the breakpoint from the previous analysis to define ellipses around each producing well. The size of these ellipses varies depending on which cluster a well is in. The area of this ellipse is multiplied by the lateral length of the well to create a volumetric feature. A conservation of volume is applied such that parent/child and sibling wells have to share the available volume. The drainage available for each well is calculated in isolation and then adjusted based on the offset wells. When the drainage from two sibling wells overlap, the area is divided using a closest well allocation logic. This approach allocates the stimulated area to each well based solely on distance to the well, not on which well generated the stimulated area. Parent wells take all available area touched by their drainage feature. The drainage available to child wells will have its area reduced by the amount already taken by the parent well. This can create a much smaller available area for child wells. More details of this drainage feature, as well as additional applications of the approach are described in Connell et al. 2023 and Connell, et al 2024. The geomechanical foundation for this approach is outlined by Zoback et al. 2022.

One step in the feature generation pipeline creates parent, child, and sibling classifications for each well-to-well pair in the data set based on user inputs. The distance between two wells and the difference between their completion dates will determine the classification. For this study, well relationships are defined as parent-child when the offset child well is within a horizontal distance of 3,500 ft, a vertical distance of 100 ft, and has a completion date that is 180 days or more after the parent well. Wells that fall within the 3,500 ft and 100 ft offsets, but that have less than 180 days between their completion dates are classified as siblings. Each well can have multiple parent or child wells and can be both a parent and a child relative to its offset wells.

This reallocation of drainage happens at three segments along the lateral. In this way, we correctly account for the impact of partially bounded wells. The drainage value gives us an engineered feature that scales with well spacing and accounts for parent-child interactions.

We then train a new machine learning model using the same features as our clustering model but instead of using 12-mon cum/ft as the target variable we use 12-mon cum/ft3. The target variable is now normalized by the drainage volume allocated to that well. This approach assumes a geometric relationship between production and available volume.

Spacing Degradation Analysis

Sensitivities were first run at three locations across Lea and Eddy Counties. Well spacing varies from 1-12 WPS using a staggered well configuration across the WCA and WCXY intervals. The completion design also varied from 2100 to 2500 lb/ft with a corresponding change to fluid intensity. Each of these runs used the drainage dimensions as defined by their corresponding cluster.

Results

Clustering Model

As described above, an initial model was trained that did not include any well spacing features. Rather, the model used log-based geologic properties and completions parameters. Even without a spacing feature, the model had 81.7% median accuracy on a five-fold cross validation test. The model is picking up on both high and low performing wells as can be seen in Figure 2.

All nine features used in the model contributed to the predictions. Figure 2 shows that density, lateral length, and proppant/ft were the three most important features.

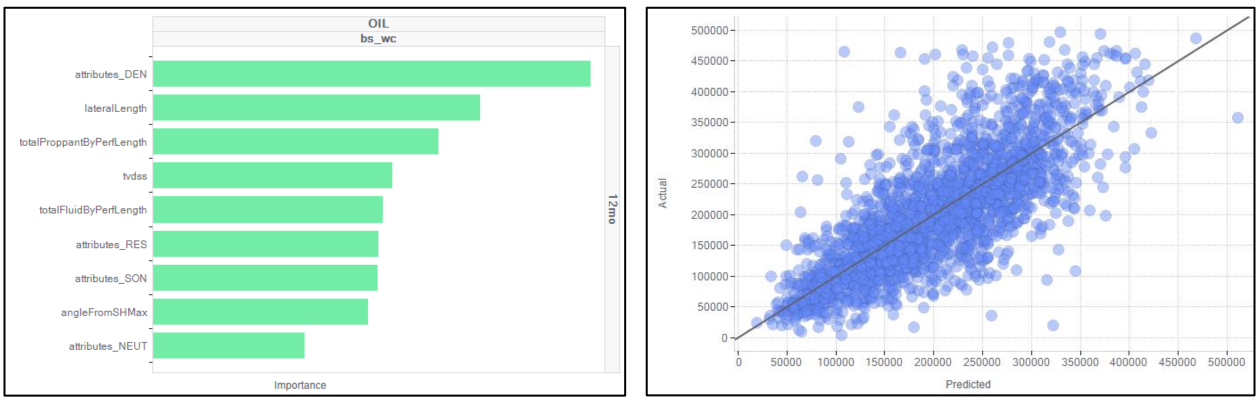

Figure 2: The relative feature importance for the clustering models is shown in the left. The scatter plot in the right shows the actual versus predicted results for the five-fold cross validation accuracy.

The spatial-interval clustering methodology successfully identified three distinct well clusters across the Wolfcamp formation, each representing unique combinations of geological properties, completions design, and spatial positioning. These clusters show clear differentiation in both feature space and geographical distribution, revealing patterns that provide insights into reservoir heterogeneity and well performance.

Principal Component Analysis (PCA) was performed on the standardized feature set comprising SHAP values for well attributes along with geographical coordinates. The first two principal components captured 49.61% of the total variance, with PC1 explaining 32.77% and PC2 explaining 16.84% of the variance.

Figure 3: The three clusters show clear separation in PCA space.

The PCA visualization in Figure 3 reveals three distinct clusters within the Wolfcamp formation wells, showing clear spatial and characteristic separation. Cluster wolfcamp_1 (green, n=601) occupies the upper portion of the feature space, while cluster wolfcamp_0 (red, n=436) is concentrated in the lower portion. Cluster wolfcamp_2 (blue, n=358) shows separation primarily along the first principal component, occupying the right side of the plot.

The clustering results demonstrate strong spatial continuity, as evidenced by the relatively tight grouping of points within each cluster, despite the high-dimensional nature of the input features. Some overlap between clusters is observed at their boundaries, particularly between wolfcamp_1 and wolfcamp_2, suggesting a gradual transition in well characteristics across these regions. The minimum cluster size constraint of 250 wells was successfully maintained, ensuring statistically meaningful groupings for subsequent analysis.

The distribution of SHAP values across the three distinct Wolfcamp clusters (wolfcamp_0, wolfcamp_1, and wolfcamp_2), providing further insights into feature importance and cluster differentiation. These distributions are shown in Figure 4 below. True Vertical Depth Subsea (tvdss) emerges as the primary discriminating feature among clusters, exhibiting clear separation in median values. Wolfcamp_0 demonstrates negative impact values, while Wolfcamp_1 and Wolfcamp_2 have progressively more positive impacts, suggesting depth-dependent clustering behavior. Petrophysical attributes display varying degrees of cluster separation. Neutron logs (NEUT) show minimal inter-cluster variation with consistently positive SHAP values. Resistivity measurements (RES) indicate moderate cluster differentiation, with all clusters maintaining positive median values and Wolfcamp_2 exhibiting slightly elevated impacts.

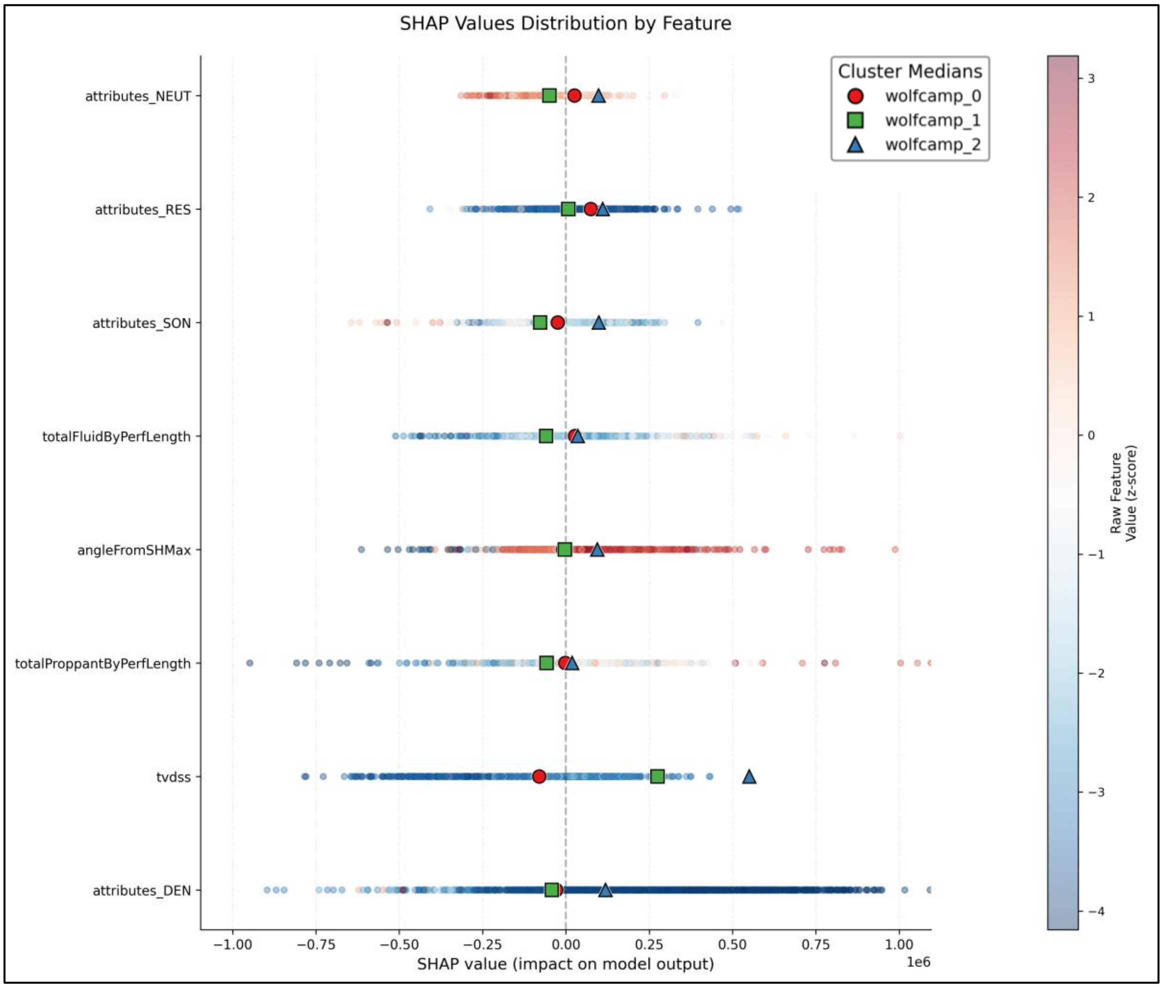

Figure 4: The SHAP distribution shows the range of SHAP values across each feature. The median value from each cluster is marked with a diamond. The three clusters have the clearest separation in TVDSS.

Completion parameters, including totalFluidByPerfLength and totalProppantByPerfLength, demonstrate similar patterns across clusters with minimal median value separation. However, these parameters show substantial intra-cluster variation, suggesting their importance in individual well performance rather than cluster differentiation. The angleFromSHMax parameter exhibits broad value distribution across all clusters, with median values concentrated near zero. This suggests that stress orientation may have localized importance but does not serve as a primary clustering factor. Density logs (DEN) show moderate cluster separation with extensive positive-side distribution, particularly in Wolfcamp_2. This indicates that density measurements contribute meaningfully to cluster characterization, though with significant overlap between clusters. Latitude and longitude were also used in the clustering but where not features in the model. These were add to further enforce geographic trends in the clustering but because they were not model features they do not appear on the SHAP distribution in Figure YY.

These findings suggest that depth and certain petrophysical properties are key determinants in Wolfcamp formation classification, while completion parameters may be more significant for individual well performance optimization.

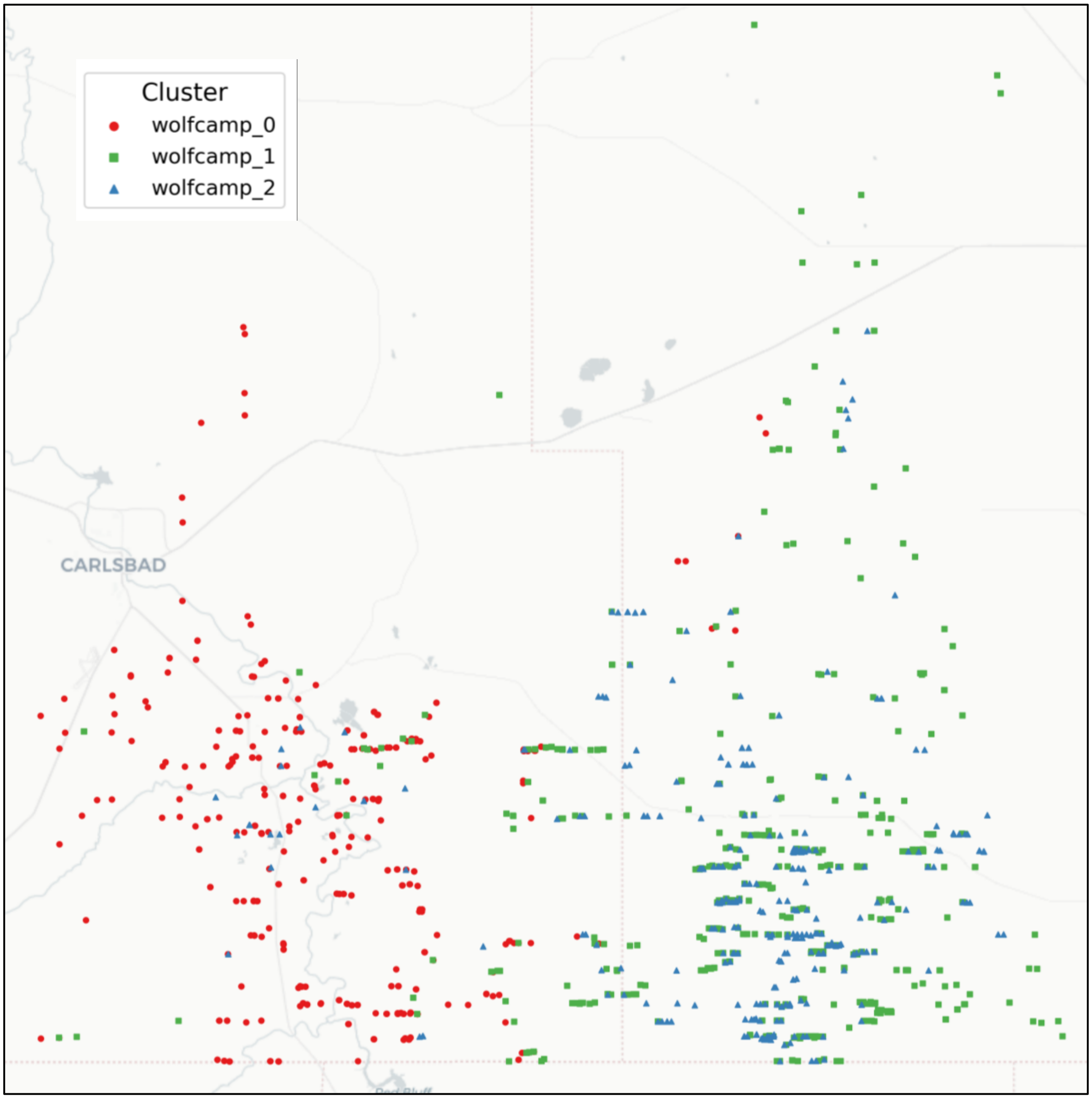

Figure 5: The location of the three clusters are shown in the map above.

The spatial distribution of Wolfcamp clusters exhibits distinct geographical segregation across the study area, with clear patterns emerging in the Delaware Basin region as is summarized by Figure 5. Wolfcamp_0 (red) demonstrates strong spatial concentration in the western portion of the study area. This cluster forms a dense, well-defined grouping, suggesting consistent geological or operational characteristics in this region. Wolfcamp_1 (green) occupies predominantly the eastern and northeastern sections of the study area. The cluster shows a more dispersed distribution pattern compared to Wolfcamp_0, with several distinct sub-groupings and a notable presence in the upper portions of the mapped region. Wolfcamp_2 (blue) displays an interesting intermediate distribution pattern, primarily occurring in a central band between Wolfcamp_0 and Wolfcamp_1 clusters. This spatial arrangement suggests a potential transitional zone or distinct geological feature that characterizes this cluster. The spatial segregation aligns with the previously observed SHAP value distributions, particularly regarding tvdss variations. The clear geographical boundaries between clusters indicate that spatial position is a strong determinant of cluster membership.

This distinct spatial organization suggests that the clustering algorithm has captured meaningful geological or operational variations that follow geographic trends across the basin. The clear spatial boundaries between clusters may have significant implications for field development strategies and well placement optimization.

The normalized 12-month oil production (Figure 6) reveals substantial performance differences between clusters, validating the practical significance of the identified groupings. Cluster WCA_1 demonstrates markedly superior productivity, with a median normalized production approximately 50% higher than other clusters and significantly higher upper quartile values. This superior performance aligns with its favorable petrophysical characteristics identified in the feature analysis.

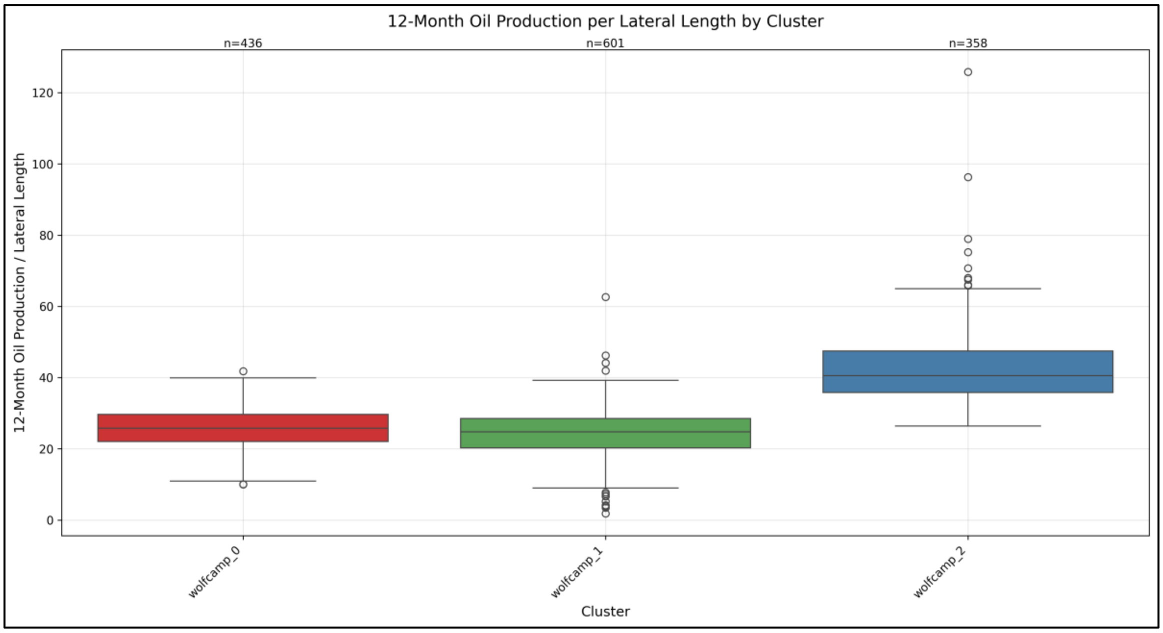

Figure 6: Production was not used to determine the clusters, but there is a clear difference in production across the three clusters, particularly in cluster 2.

Clusters WCA_0, WCA_2, WCA_3, and WCA_4 show relatively similar median production values, but with differing distributions. Cluster WCA_0 exhibits a tighter distribution, suggesting more consistent well performance. Clusters WCA_2 and WCA_3 show wider distributions with some high-performing outliers. The Wolfcamp B clusters (WCB_0 and WCB_1) demonstrate different performance profiles compared to the Wolfcamp A clusters, with WCB_0 showing higher production potential than WCB_1. These distinct production signatures across clusters confirm that the identified groupings capture meaningful differences in well performance potential. The consistency between cluster characteristics (Figure 1) and production outcomes (Figure 2) validates the clustering approach as a means to identify geological and completions regimes with practical implications for field development.

The clear differentiation in both feature characteristics and production performance across the identified clusters provides a robust foundation for the subsequent spacing analysis. Each cluster represents a distinct reservoir-completions regime that can be expected to demonstrate unique spacing-dependent behavior, enabling more nuanced optimization of well spacing across the heterogeneous asset.

Breakpoint Analysis

The spacing breakpoint analysis reveals distinct interference thresholds across the identified clusters, as illustrated in the boxplot distributions of Figure 7. Cluster wolfcamp_1 was our reference cluster and we selected the second derivative threshold to make this cluster have a breakpoint at 1400 ft. that same threshold was applied to the other clusters. This yielded a breakpoint of 1730 ft in cluster wolfcamp_0 and 1625 ft in wolfcamp_2. Although the raw data is noisy, by fitting the logarithmic curve we can objectively quantify the point where wells start to interfere.

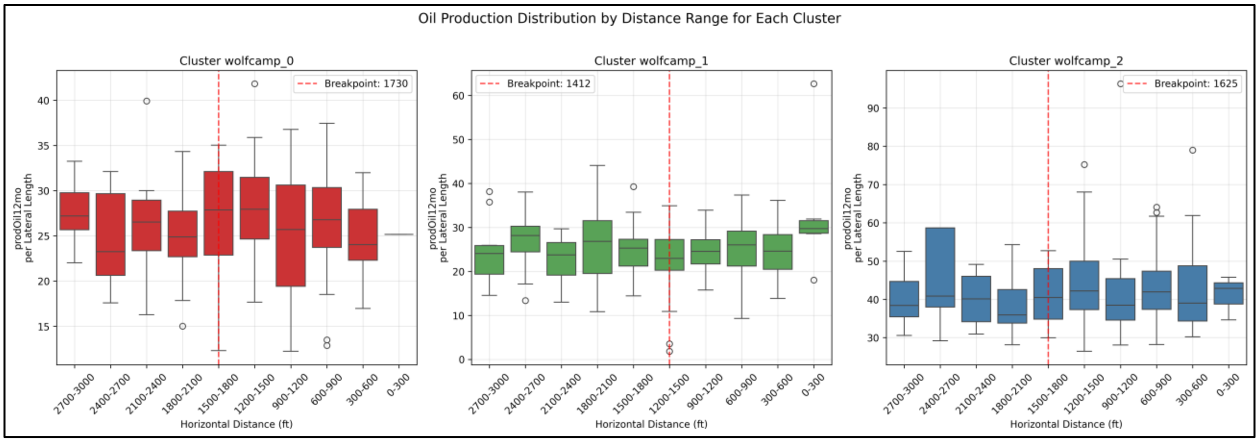

Figure 7: The production and spacing trends across the clusters, along with their breakpoint, is shown above.

These differentiated breakpoints confirm that well interference patterns are not uniform across the reservoir but rather vary significantly based on the geological and completions characteristics captured in each cluster.

Training a Spacing-Aware Model

The breakpoint analysis gives us the missing piece we need to train a spacing aware model to predict well performance. The cubic extraction ratio (CER) was used as the target variable for this model, where:

$$CER = \frac{12mon\ cum\ oil\ (bbl)}{Lateral\ Length\ (ft)\ x\ Total\ Drainage\ (ft^2)} \quad [2]$$

The feature pipeline was reconfigured to assign ellipse dimensions to each well in the training set based on that well's cluster assignment. The same parent/child/sibling logic was applied to the drainage volume for each well. This process assigns a Total Drainage value to each well. The same geological and completions features were used in this model as were used in the clustering model. The feature importances are slightly different, with TVDSS standing out as the most important feature (see Figure 8). This model had a 5-fold cross validation accuracy of 78.5%.

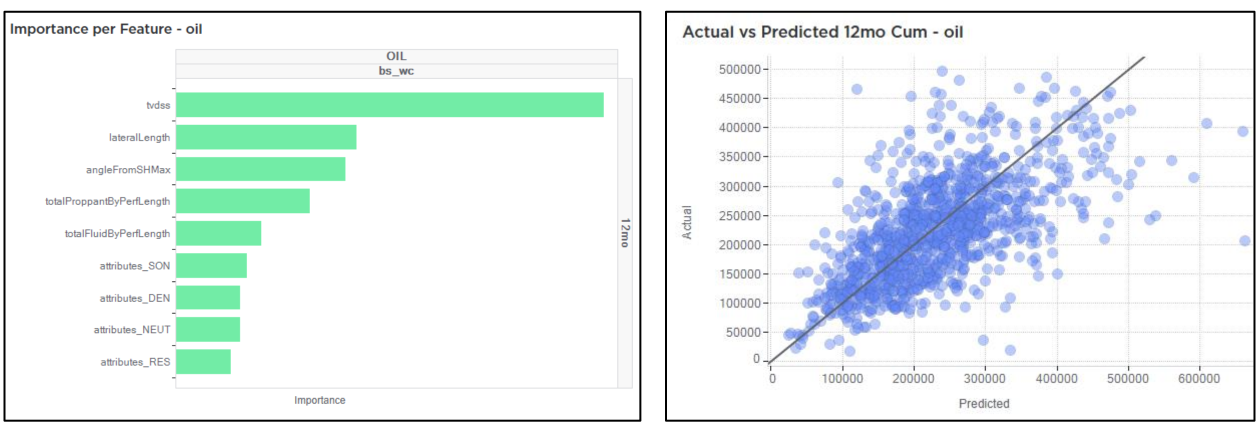

Figure 8: The feature importance and five-fold cross validation for the spacing aware model is shown above.

Spacing Sensitivity

We next select locations across the study area to run spacing sensitivities. Locations were selected in each of the three clusters and sensitivites were run from 1 to 12 wells per section (WPS). These locations are shown on the map in Figure 9 below.

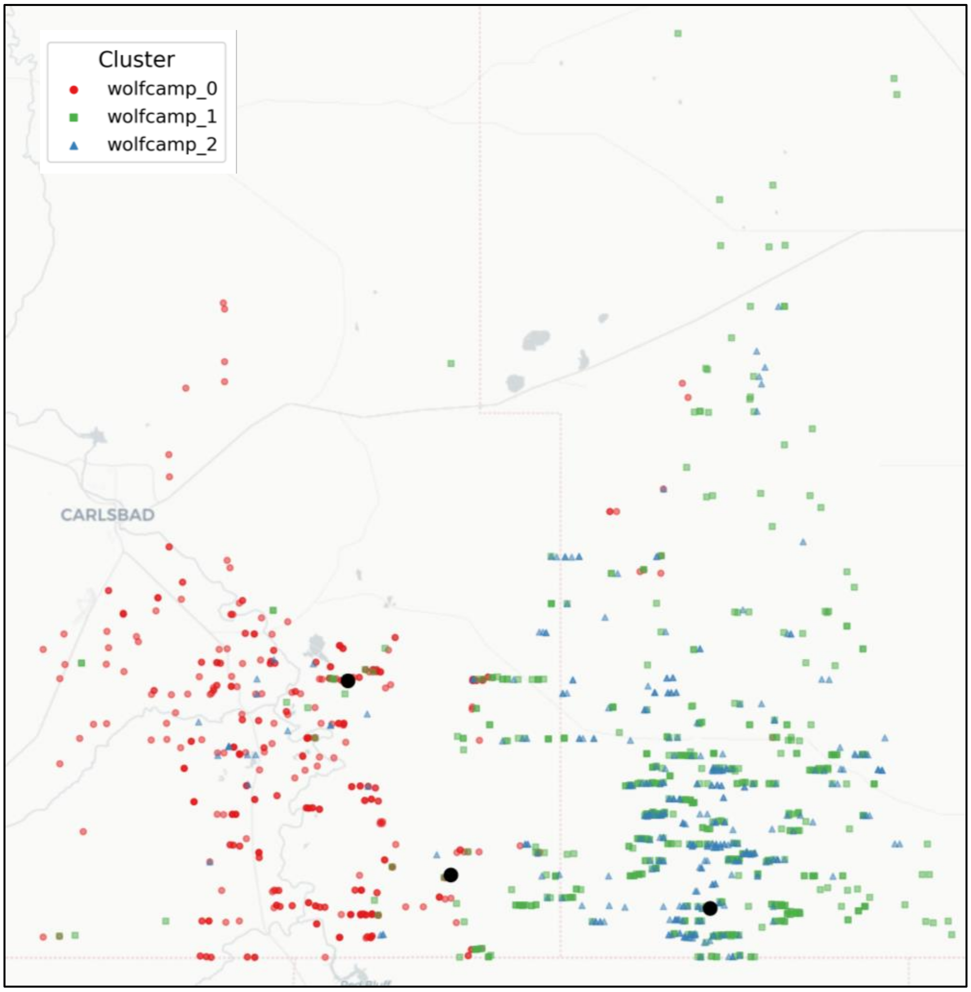

Figure 9: Three locations were selected to run spacing tests across the study area.

Each location has the dimensions of the drainage ellipse set based on our previous breakpoint analysis. These dimensions are summarized again in Table 1. These dimensions allow the model to see well interactions. The larger the dimensions, the sooner wells will start to see their neighbors. As wells get closer together, they have less volume to drain. This results in geometric degradation.

Table 1: Summary of breakpoints for each cluster.

Cluster	**Breakpoint**
wolfcamp_0	1730
wolfcamp_1	1400
wolfcamp_2	1625

The figure below shows what the volume sharing looks like at 8 WPS. Each well in the scenario gets its drainage allocation with is used to normalize the production.

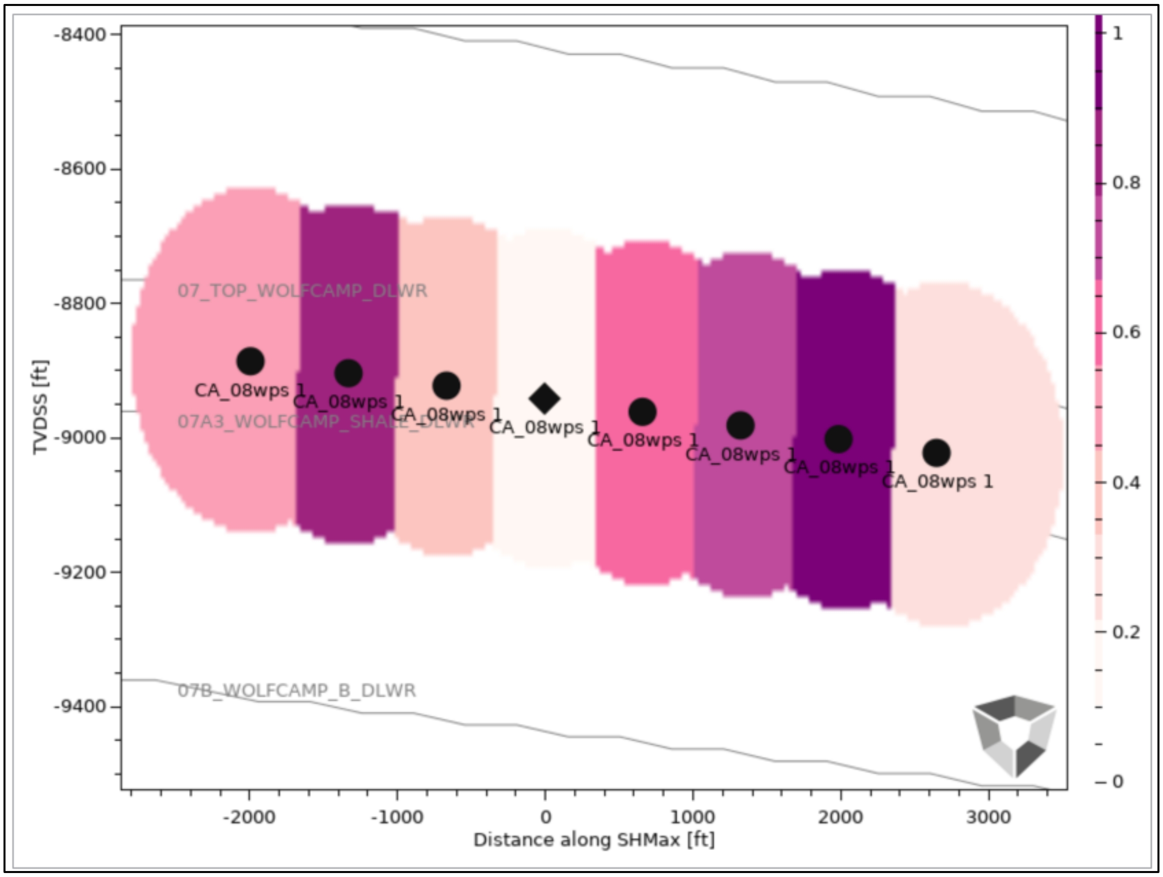

Figure 10: The system allocates available drainage area across all wells in the test. As wells move closer together, they have less volume available.

Analysis of normalized production performance (EUR Oil per ft) versus well density reveals compelling relationships between drainage geometry and well productivity. Wolfcamp_2 demonstrates superior initial productivity, achieving approximately 145 bbl/ft at low well densities (1-3 wells per section), despite having an intermediate drainage width. This is expected based on the higher productivity of wells in this cluster. The figure shows that an 8 WPS development in cluster 2 could still be as productive on a per well basis as 5 wells in clusters 0 and 1.

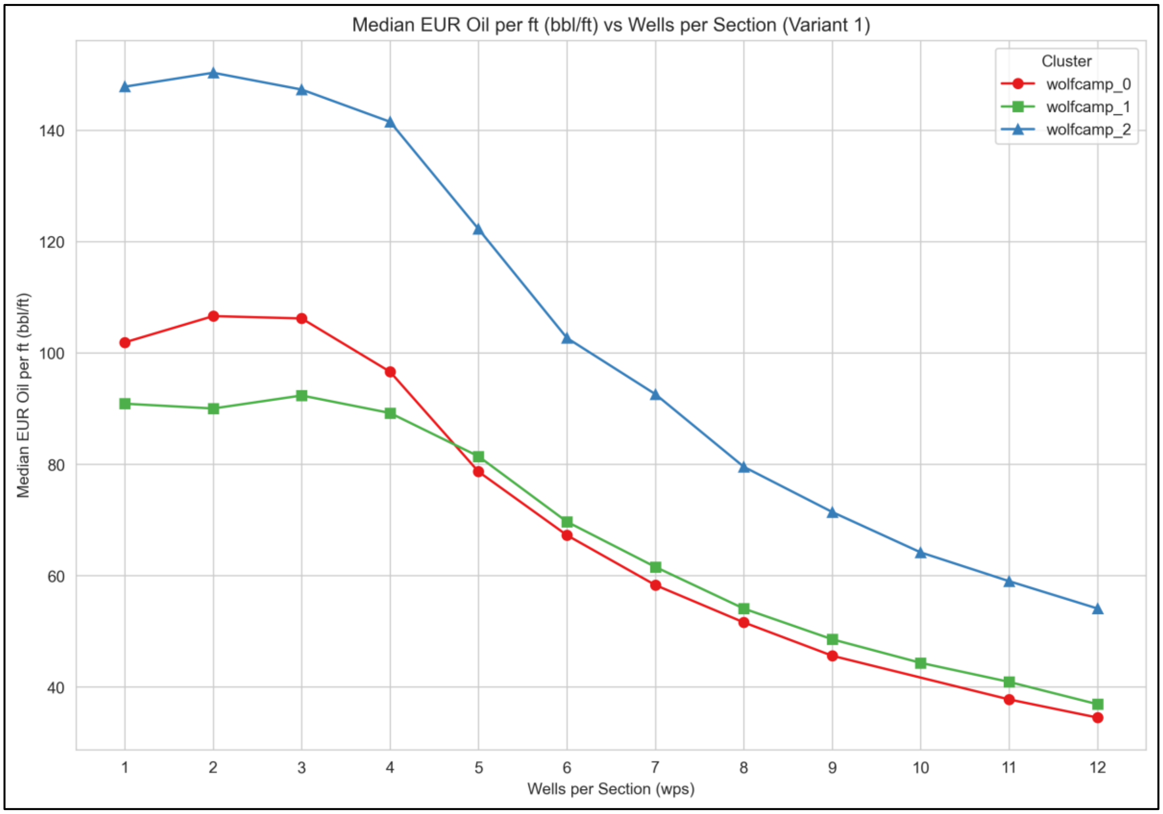

Figure 11: The EUR/ft for each of the spacing tests is shown above.

The degradation analysis, normalized to single-well performance, provides crucial insights into the impact of shared drainage volume across the clusters. Figure 12 shows the relative change at each WPS by calculating the percent difference from the baseline of a single, unbounded well.

$$Degredation = {(value - baseline)}/{baseline} \quad [3]$$

All three areas maintain relatively stable performance within ±5% variation up to 4 wells per section, indicating minimal interference effects at wider well spacing. However, beyond 4 wells per section, performance degradation accelerates non-linearly across all clusters. With its smaller drainage volume, wolfcamp_1 demonstrates the most resilient behavior to degradation not surpassing 10% degradation until 6 WPS. Wolfcamp_0 and wolfcamp_2 both pass 10% degradation at 5 WPS.

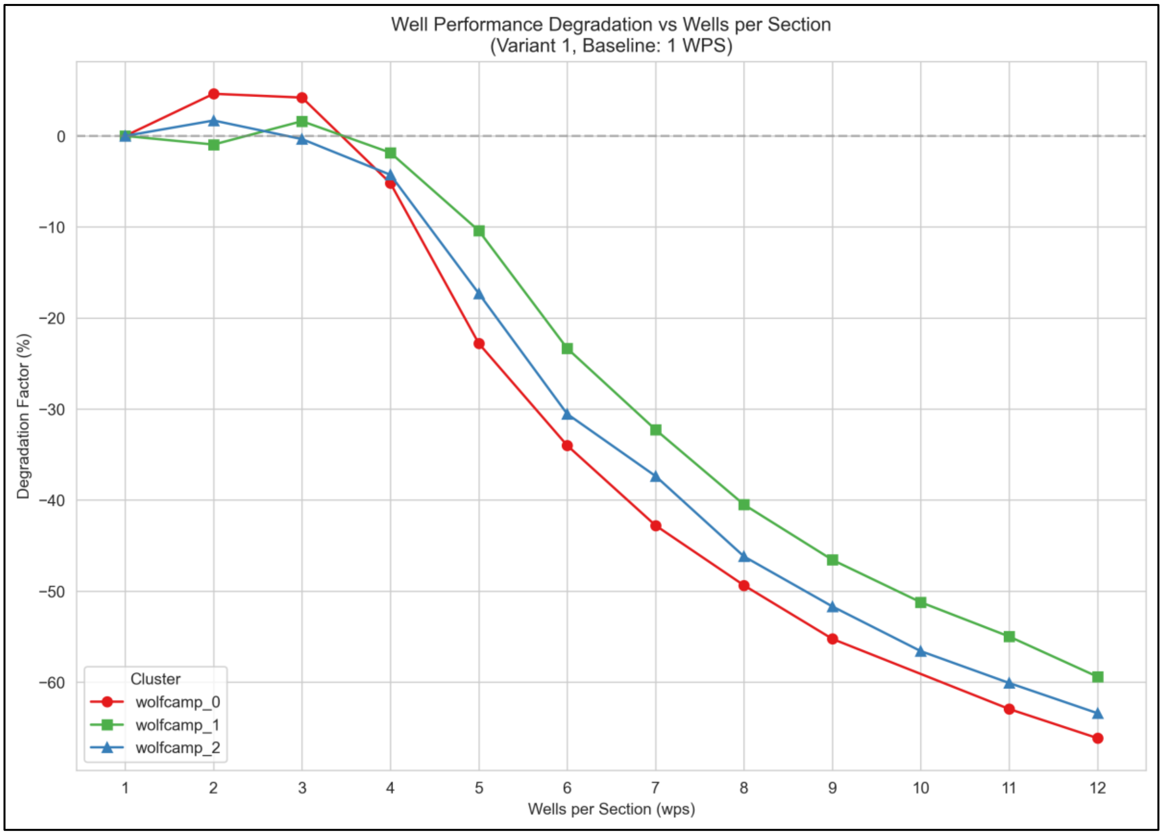

Figure 12: The degradation factor is calculated using a single well as the baseline.

The machine learning model's incorporation of drainage volume sharing through elliptical drainage areas effectively captures the physics of well interference. The observed degradation patterns validate the model's ability to represent the competition for available reservoir volume as well spacing decreases. The varying responses across clusters indicate that the model successfully captures not only the geometric aspects of drainage but also the underlying geological heterogeneity influencing production performance.

Conclusions

This study demonstrates the effectiveness of an integrated, data-driven approach to quantify well spacing effects in the geologically complex Delaware Basin. By combining advanced clustering techniques with machine learning models calibrated to real-world operational data, we have successfully isolated the impact of well spacing from other performance drivers across heterogeneous reservoir conditions.

Our analysis reveals that the northern Delaware Basin does not follow a single-spacing degradation pattern but rather exhibits distinct behavior across the three identified geological clusters. The breakpoint analysis identified critical interference thresholds at 1730 ft, 1400 ft, and 1625 ft for clusters wolfcamp_0, wolfcamp_1, and wolfcamp_2 respectively, providing concrete guidance for development planning within each region. The significant variation in these thresholds underscores the importance of treating spacing optimization as a location-specific exercise rather than applying basin-wide generalizations.

The spacing sensitivity results offer particularly valuable insights for field development. All three clusters maintain relatively stable performance with minimal degradation up to 4 wells per section, suggesting this may represent a conservative initial development density across the basin. However, the findings also suggest that spacing degradation can be even more severe in regions with superior reservoir quality. It’s not that reservoir quality prevents degradation, it’s just that in these regions wells can sustain a significant amount of degradation and still be economic.

Beyond the specific findings for the Delaware Basin, this work establishes a robust methodological framework that can be applied to other unconventional plays. By reducing reliance on anecdotal observations and interpretive biases, our approach provides a more objective basis for development decisions. The combination of geologically informed clustering with physics-based drainage modeling creates a powerful predictive tool that balances data-driven insights with reservoir engineering principles.

Future work should focus on validating these predictions with additional production history, refining the clustering approach to incorporate additional geological attributes, and extending the analysis to evaluate economic optimization metrics beyond physical well performance. Nevertheless, this study represents a significant advancement in our ability to make data-driven decisions about optimal well spacing in complex unconventional reservoirs.

References

Connell, Charles, LaMotta, Kyle, Chaudhary, Nitin, and Mickey Friedrich. "Optimizing Shale Infill Development with AI: Child Well Meta Models." Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, June 2023. doi: https://doi.org/10.15530/urtec-2023-3870467

Connell, Charles, LaMotta, Kyle, Munger, Clark, and Joe Wicker. "A Machine Learning Approach to Basin Scale Inventory Forecasting and Development Planning that Accounts for Complex Well Interactions." Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, June 2024. doi: https://doi.org/10.15530/urtec-2024-4036541

Gonzalez, K., Brusova, O., & Valenciano, A. (2023). A machine learning workflow for log data prediction at the basin scale. Extended abstracts, SEG Annual Meeting. First Break, 41(2), 73-80. https://doi.org/10.3997/1365-2397.fb2023015

Lundberg, Scott and Lee, Su-In. 2017. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17). Curran Associates Inc., Red Hook, NY, USA, 4768–4777.

Ren, Xiang (Rex), Yin, Jichao, Xiao, Feng, Miao, Sasha, Lolla, Sri, Yao, Changqing, Lonnes, Steve, Sun, Huafei, Chen, Yang, Brown, James, Garzon, Jorge, and Piyush Pankaj. "Data Driven Oil Production Prediction and Uncertainty Quantification for Unconventional Asset Development Planning Through Machine Learning." Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, June 2023. doi: https://doi.org/10.15530/urtec-2023-3865670

Zhou, Hui, and Benjamin Lascaud. "An Integrated Machine Learning Framework for Optimizing Unconventional Resources Development." Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Denver, Colorado, USA, July 2019. doi: https://doi.org/10.15530/urtec-2019-319

Zoback, Mark, Ruths, Troy, McClure, Mark, Singh, Ankush, Kohli, Arjun, Hall, Brendon, Irvin, Rohan, and Malcolm Kintzing. "Lithologically-Controlled Variations of the Least Principal Stress with Depth and Resultant *Frac Fingerprints* During Multi-Stage Hydraulic Fracturing." Paper presented at the SPE/AAPG/SEG Unconventional Resources Technology Conference, Houston, Texas, USA, June 2022. doi: https://doi.org/10.15530/urtec-2022-3722883