Induced Seismicity Susceptibility Investigation in Kiskatinaw Area, British Columbia: A Machine Learning-Based Case Study

Ali Mehrabifard, Erik Eberhardt

Geological Engineering/EOAS, The University of British Columbia

Steve Rogers

WSP, Vancouver

Report to:

BC Oil and Gas Research and Innovation Society (BC OGRIS)

Rev 1

July 18, 2023

Acknowledgements

This work wouldn't have been possible without the support of BC OGRIS, the technical support of the BC Energy Regulator (BC-ER), and the generous donation of subsurface data by several Montney operators.

Contents

List of Figures

List of Tables

1. Introduction

A key recommendation in the independent report, Scientific Review of Hydraulic Fracturing in British Columbia (Allen et al., 2019), was the need to develop a susceptibility map of induced seismicity potential for northeastern British Columbia (NEBC). In response, Amini et al. (2021) in collaboration with Geoscience BC developed a machine learning (ML) framework for generating induced seismicity susceptibility maps, with focus placed on the Montney region of NEBC. These maps were produced relying on a combination of public domain and proprietary data, including the use of geoLOGIC systems ltd.'s geoSCOUT database.

Building on this, focus for the present study was placed on a well pad-scale case study using data provided for a seismogenic well pad in the Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) of NEBC. The key objective of this research is to find the correlation between the geological and operational features (i.e., the input parameters for ML) and the occurrence of relatively large induced seismicity events (i.e., $\geq M_W 1$). The goal in doing so is to identify the important unfavorable features that are contributing to the occurrence of these large seismic events. Furthermore, these results can be compared with the findings of earlier regional studies (e.g., Wozniakowska and Eaton, 2020; Amini et al., 2021).

1.1. Background

The advancement of multivariate statistics and machine learning techniques in analyzing large datasets is useful in dealing with the problem of induced seismicity hazards. However, there is limited experience in this area, particularly with considering the influence of both geological and operational factors. It is important to distinguish between these factors, as geological factors relate to the conditions that cannot be controlled or manipulated, while most operational factors can be manipulated to potentially mitigate induced seismicity hazards.

Several studies have recognized that the large and complex datasets generated by hydraulic fracturing (HF) activities require sophisticated analysis methods due to the inefficiency and ineffectiveness of traditional empirical and statistical analyses. These have turned to ML techniques, which provide useful tools that can detect correlations that are otherwise hidden. Pawley et al. (2018) used logistic regression to develop an induced seismicity potential map in the Duvernay play in Alberta, finding that factors such as proximity to basement, formation overpressure, minimum horizontal stress, proximity to reef margins, lithium concentrations, and natural seismicity rate are the most significant factors in inducing seismicity. Similarly, Wozniakowska and Eaton (2020) developed a Seismic Activation Potential (SAP) map in the Montney Formation using logistic regression, with the distance to the Cordilleran deformation front and injection depth being the most important features. Fox et al. (2020) used multivariate statistical analyses, testing four different machine learning models, to investigate the correlation between operational parameters and a limited set of geological parameters with induced seismicity events in the KSMMA and the North Peace Ground Motion Monitoring Area (NPGMMA) of the Montney region. The analysis also looked at both likelihood (classification of seismogenic wells) and severity (maximum magnitude regression) of induced seismicity events. The study found that the simplest model placed a relatively high negative importance on minimum horizontal stress and high positive importance on geothermal gradient, distance between wells and mean proppant per stage.

However, the more complex models placed relatively high importance on Paleozoic structure and distance to faults. In terms of magnitude regression analysis (i.e., predicting the magnitude of the induced event), most models showed a relatively high importance for the top of the Montney structure and distance to faults. Amini and Eberhardt (2021) investigated different machine-learning algorithms and found that geological features, such as pore-pressure gradient, distance to basement, distance to known faults, and azimuth of maximum horizontal stress, generally ranked higher than operational features in inducing seismicity. The completion length was the operational feature that consistently ranked as being important. Mehrabifard and Eberhardt (2021) investigated the effect of geological features on the magnitude distribution of induced seismicity events using different machine-learning algorithms and showed that the seismogenic response was most strongly influenced by the pre-injection reservoir pore pressure.

Amini et al. (2021) presented the most comprehensive ML study to date on induced seismicity, generating susceptibility maps for the Montney region. Eight different ML algorithms were tested across analyses that included both likelihood classification and severity regression. The results found the Random Forest ML model to be amongst the top performing models, and that geological features generally ranked higher in importance than operational features. Specific to susceptibility, the depth to the top of the basement was identified as the most important predictor of a well, being seismogenic. This was shown to have a negative correlation, meaning that shallower depths to the basement increases the likelihood of a well, being seismogenic. Numerical modelling was used to investigate mechanistic explanations for the results and showed that the injection depth influences the extent to which the fluid pressure invaded zone is transmitted outwards from the injection point and HF. For models simulating a deeper basement depth, the higher stresses resulted in a significantly restricted invaded zone. As the modelled distance to the top of the basement becomes shallower (i.e., lower stresses), fluid injection results in a larger volume of elevated pore pressures. Amini et al. (2021) suggested that this results in an increased likelihood of elevated pore pressures interacting with a critically stressed fault in the basement to trigger an induced seismicity event. Overall, the locations of susceptible areas in the ML models agreed with the historical location of induced seismicity events. The top predictors for event severity were identified by the ML analyses as being: b-value of events within 100 km of the well (negative correlation), standard deviation of pore pressure gradient (positive correlation), and distance of well from the Cordilleran thrust belt (negative correlation). A positive correlation between the standard deviation of the pore pressure gradient and event severity concurs with the known influence of pore pressure compartmentalization. A negative correlation between the distance of the well from the Cordilleran thrust belt and event severity corresponds with the expected increase in fault density towards the thrust belt and therefore the increasing likelihood of a HF operation encountering a critically stressed fault.

It their conclusions, Amini et al. (2021) cautioned that ML results are subject to data errors, biases, censoring and overfitting, and therefore are not a perfect predictor. They recommended that a next step in the use of ML-generated susceptibility (and severity) maps would be to consider the site-specific geology, especially where these deviate from general regional trends.

1.2. Study Area

Data was originally received from three operators (anonymized as Operator 1, 2, and 3) from within the KSMMA boundary, shown in Figure 1. The KSMMA is situated between Fort St. John in the northwest and Dawson Creek in the southeast. Operator 1 (OP1) had a full dataset that could be used for a detailed ML study. The OP1 data was selected for the case study and is reviewed in the next section.

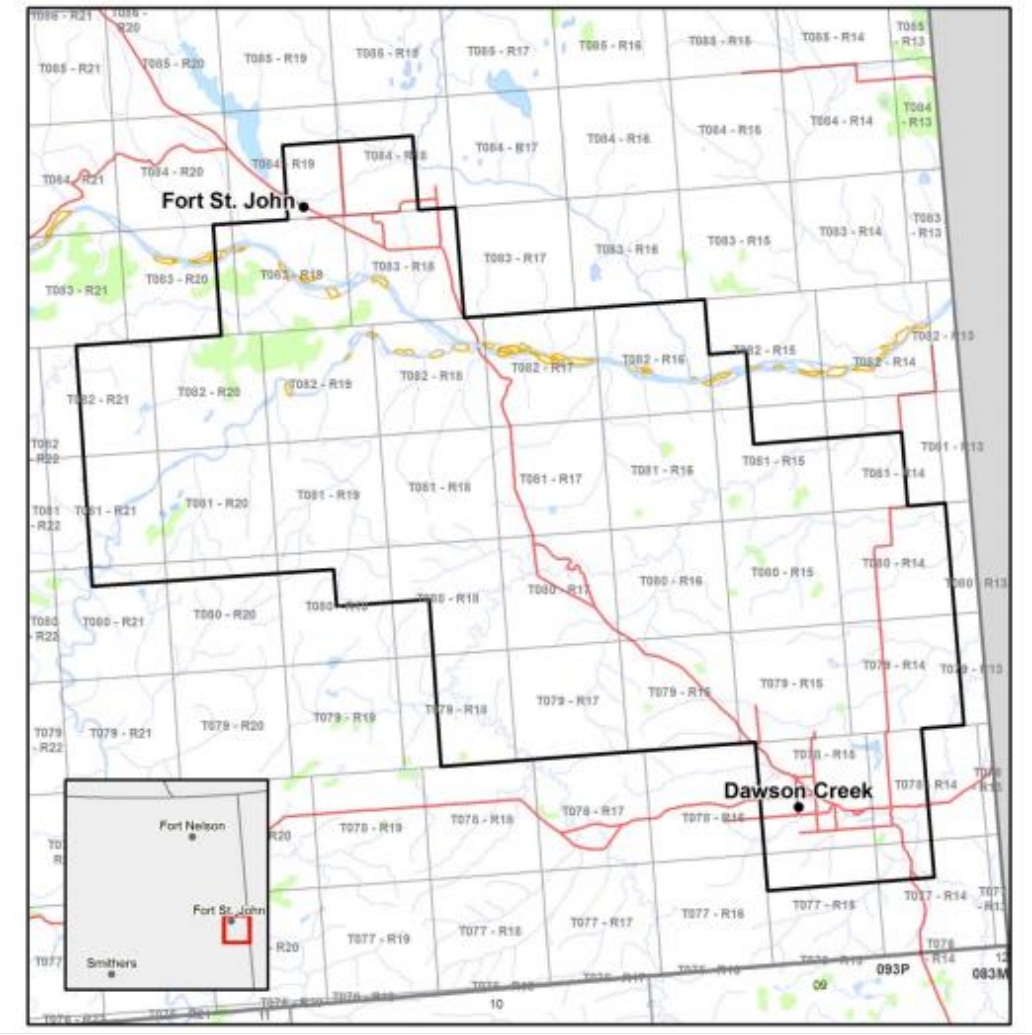


Figure 1: Map of the Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA). Source: British Columbia Energy Regulator.

2. Data

2.1. Operator 1 Data Set Summary

The data set for Operator 1 (OP1) is derived from the treatment of six east-west orientated wells, all completed in the Upper Montney using a plug and perf completion. A total of 4,434 induced seismic events were recorded by an operator-owned shallow buried array, with moment magnitudes between

-0.922 and 2.246 being observed. Induced events have been associated with the active stages from the wells using a simple temporal filter. These are shown in *Figure 2*.

The spatial distribution of events reveals strongly aligned patterns of seismic lineaments with multiple trends being present, with the main trends being 055, 025, and 115 degrees. *Figure 3* provides further details for the orientation patterns of interpreted seismic lineaments found.

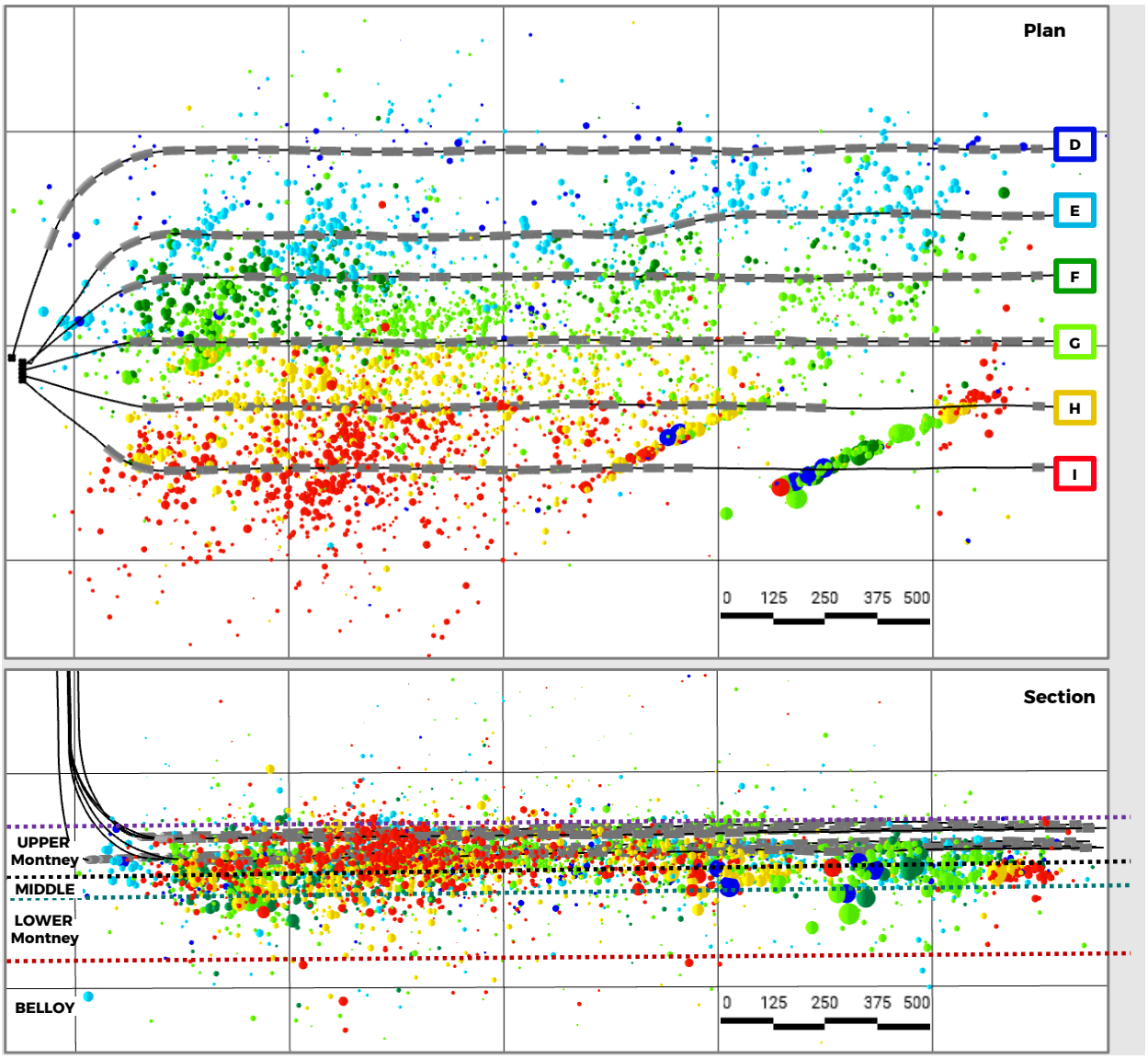


Figure 2: Plan view and sectional view of induced seismic events with size scaled by magnitude (max magnitude 2.2) and coloured by the wells they are associated with. All wells are within the Upper Montney with the Upper bench wells being E, G & I and the Lower Bench wells being D, F & H. Source: WSP 2023

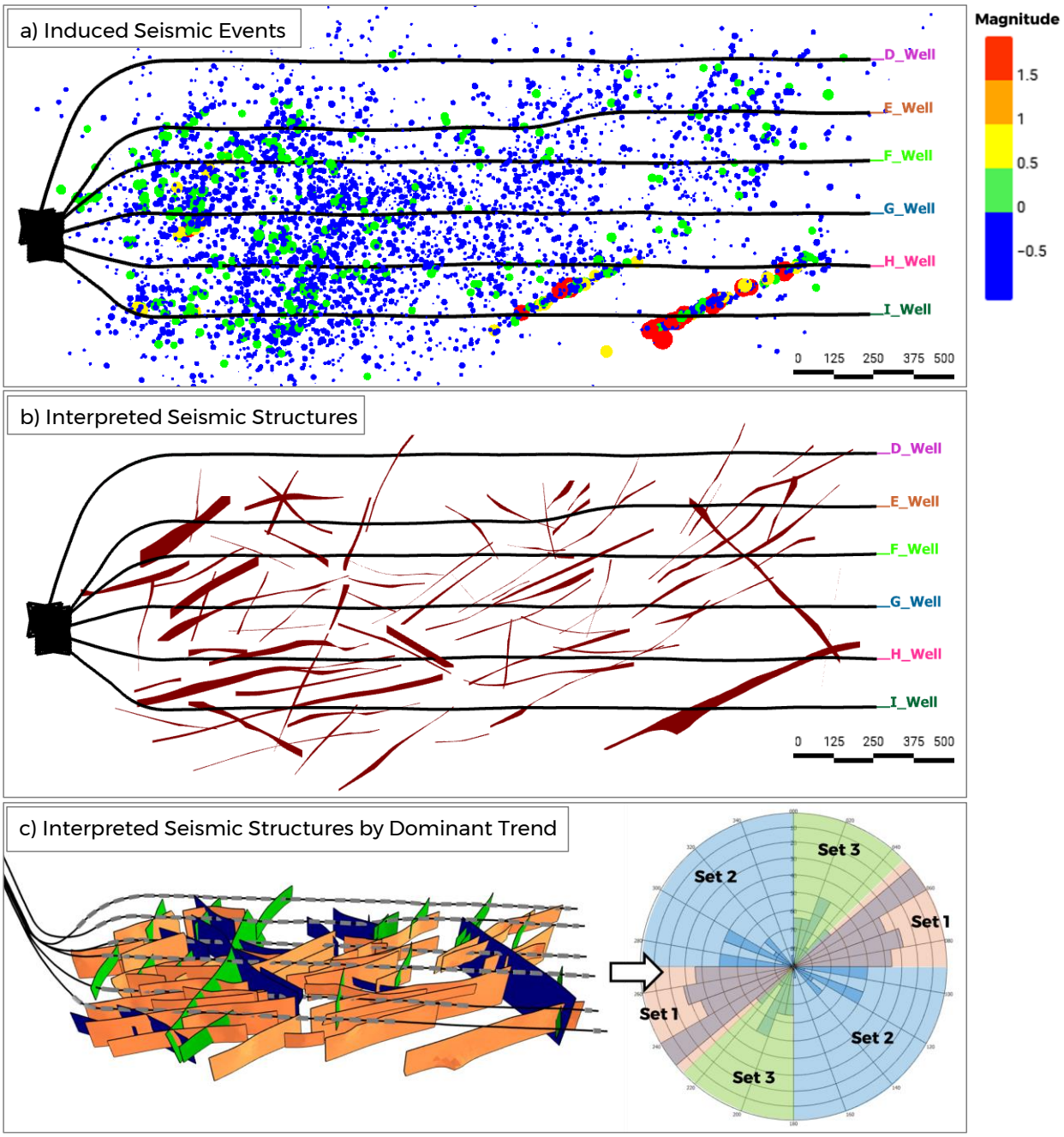
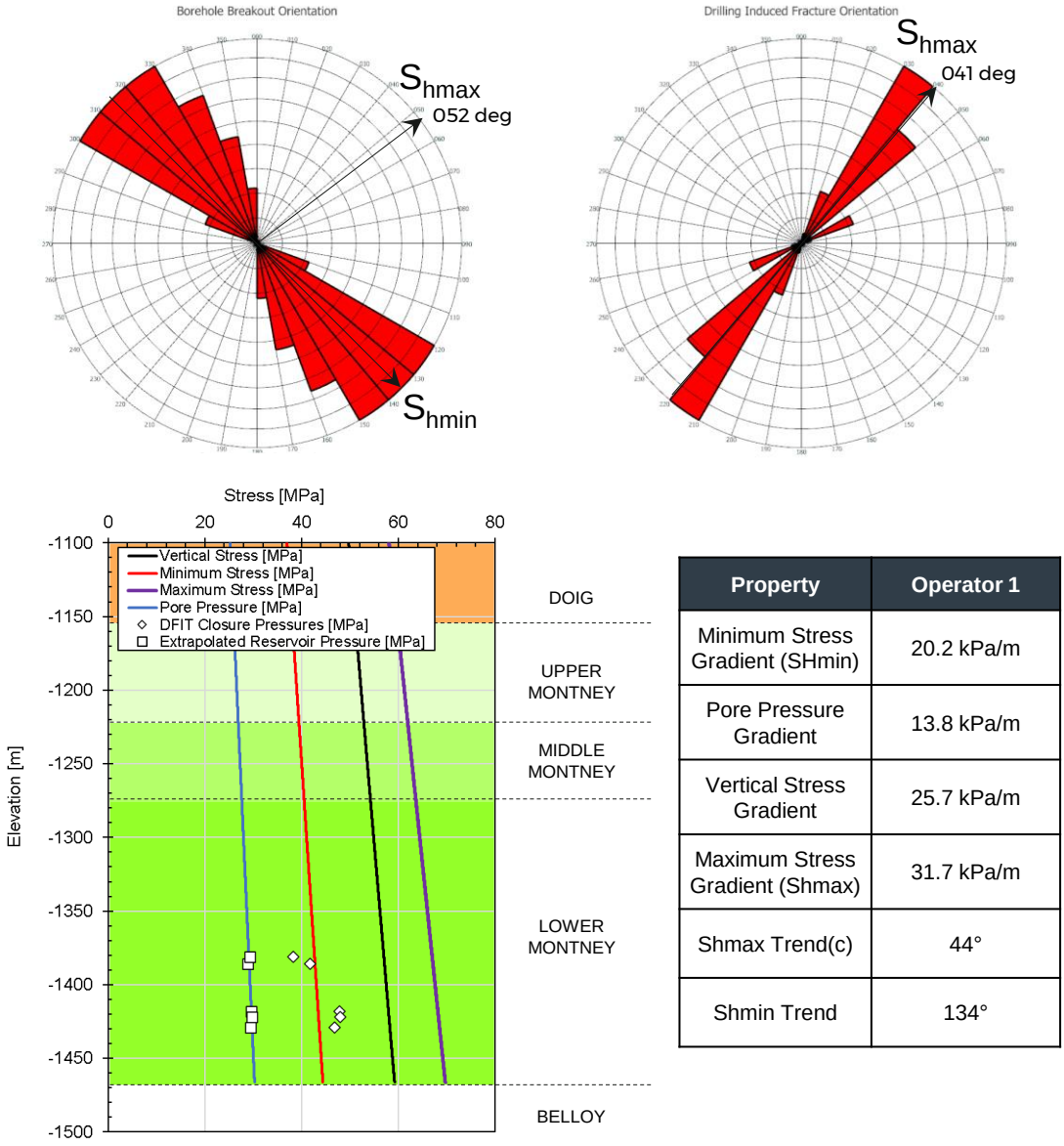


Figure 3: a) Array of induced seismic events coloured by magnitude. b) Interpreted seismic structures. c) Interpreted structures by dominant trend. Source: WSP 2023

Geomechanics data was available for this pad and stress and pore pressure gradients have been derived. An image log was available that provided the orientation of borehole breakouts and tensile drilling induced fractures, yielding an average direction of S_{Hmax} of 044 deg (Figure 4).



Property	Operator 1
Minimum Stress Gradient (SHmin)	20.2 kPa/m
Pore Pressure Gradient	13.8 kPa/m
Vertical Stress Gradient	25.7 kPa/m
Maximum Stress Gradient (Shmax)	31.7 kPa/m
Shmax Trend(c)	44°
Shmin Trend	134°

Figure 4: Top: Orientation of image log derived breakout orientation and drilling induced fracture orientations. Bottom: Summary of stress and pore pressure gradients. Source: WSP Golder 2022

2.2. Exploratory Data Analysis (EDA)

With the in-situ stress data, and the logging information available, we have performed feature engineering and assigned each induced seismic event with corresponding data. To this end, each induced event is associated with a hydraulic fracturing (HF) stage via a temporal filter (WSP 2023). All of these features are reviewed in

Table *1*.

Table 1. Description of the machine learning input data, differentiating between geological (G) and operational (O) features.

Feature	Description	Type
AverageDipDir	Average dip direction of the interpreted lineaments. Note that these are averaged because the structures are curved and not planar (see Figure 3c).	G
TraceLength_(m)	Length of each lineament measured along the longest edge.	G
number_labels	Number of interpreted lineaments connected to each HF stage. Based on the number of lineaments on which events were recorded.	G
rake_degree	Rake angle, λ, which measures the angle that the hanging wall moves during rupture relative to the fault's strike. Using Aki and Richards (2002) convention, a λ of −90° indicates normal fault slip, a λ of 0° or 180° indicates strike-slip (left- and right lateral, respectively), a λ of 90° indicates a thrust fault slip, and angles in between indicate variants of oblique slip.	G
normal_stress_MPa	Effective normal stress, σ_n', acting on each lineament, calculated based on the orientation of each lineament, the magnitude and orientation of the principal in-situ stresses and the pore pressure, and assuming a Biot's coefficient of 1.	G
Shear_stress_MPa	Shear stress, τ, calculated based on the orientation of each lineament, the magnitude and orientation of the principal in-situ stresses and the pore pressure, and assuming a Biot's coefficient of 1.	G
Slip_tendency	Ratio of shear stress to normal stress, $\tau/\sigma_n{}'$.	G
Formation_ord_enc	Ordinal encoding for formations where the events occurred, from shallowest to deepest: 0 - Doig, 1 - Upper Montney, 2 - Middle Montney, 3 - Lower Montney, 4 - Belloy.	G
DTSM [hrs/m]	Acoustic slowness of shear waves in hours per meter (hrs/m) measured from dipole shear sonic imager during well log operations. Inverse of shear wave (S-wave) velocity, V_S.	G
DTCO [hrs/m]	Acoustic slowness of compressional waves in hours per meter (hrs/m) measured from delta-T compressional logging. Inverse of compressional wave (P-wave) velocity, V_P.	G
RHOB [kg/m3]	Bulk density, ρ, from bulk density log in kilograms per cubic meter (kg/m3).	G
Prdyn	Dynamic Poisson's ratio, ν, calculated from the sonic compressional and shear wave velocities using $\nu = \frac{V_P^2 - 2V_S^2}{2(V_P^2 - V_S^2)}$, as shown by Mavko et al. (2009).	G
Ymdyn [MPa]	Dynamic Young's modulus, E, calculated from the sonic compressional and shear wave velocities and density logs using $E = \rho V_S^2 (\frac{3V_P^2 - 4V_S^2}{V_S^2 - V_P^2})$, as shown by Mavko et al. (2009).	G
Bulk_Modulus [MPa]	Dynamic bulk modulus, K, calculated from the sonic compressional and shear wave velocities using $K = \rho(V_P^2 - \frac{4}{3}V_S^2)$, as shown by Mavko et al. (2009).	G
Compressibility [1/kPa]	Inverse of bulk modulus.	G

Shear_Modulus [MPa]	Dynamic shear modulus, G, calculated from the sonic shear wave velocity and density logs using $G = \rho V_S^2$, as shown by Mavko et al. (2009).	G
Breakdown_Press_(MPa)	Formation breakdown pressure recorded during each HF stage. This represents the fracture propagation away from a wellbore (Zoback, 2007) and can be used as a proxy for formation strength. This pressure is correlated to the corresponding induced event.	G
Frac_Gradient_(kPa/m)	Fracturing gradient pressure recorded during each HF stage and correlated to the corresponding induced event.	G
ISIP_(MPa)	Instantaneous shut-in pressure (ISIP), measured after abruptly stopping flow into the well and taken as a measure of the least principal stress (Zoback, 2007).	G
Diff_m2s-1	Hydraulic diffusivity, D, calculated based on the spatio-temporal distribution of fluid-injection induced seismicity for each event (WSP 2023).	G
perm_m2	Permeability (k) calculated using $D = \frac{k}{\mu \phi C_t}$, where μ, C_t and ϕ are the viscosity of the fluid saturating the medium, the overall bulk compressibility of the saturated medium (i.e., the pore-fluid and the rock matrix and fractures), and the porosity of the medium, respectively (WSP 2023).	G
time_diff_hr	Time difference between the HF stage initiation and the corresponding events occurrence time in hours.	O-G
Avg._Rate_(m3/min)	Average injection rate in m^3/min for each stage.	O
Max._Press_(MPa)	Maximum injection pressure in MPa for each stage.	O
Avg._Press_(MPa)	Average injection pressure in MPa during each HF stage.	O
Total_Injected_(m3)	Total injected volume per HF stage. Each event is assigned to a stage via temporal filtering.	O

The value distributions of the data for each feature are shown in Figure 5. For the well logging data, since these were measured in a vertical well, we used the hypocentral depth of each event to assign the logging data value to each event. With this, we assume that the well log properties only vary in depth direction and not horizontally (i.e., the reservoir consists of transversely isotropic formation rocks).

The lineaments are assigned if the induced seismicity were used to interpret the structure. If the event has not occurred on an interpreted structure, it is assigned a negative label. This will then be ignored as a useful input data when the features are scaled.

Finally, the target class consists of two subsets: events with magnitude 1 or greater (positive class) and events with magnitude smaller than 1 (negative class).

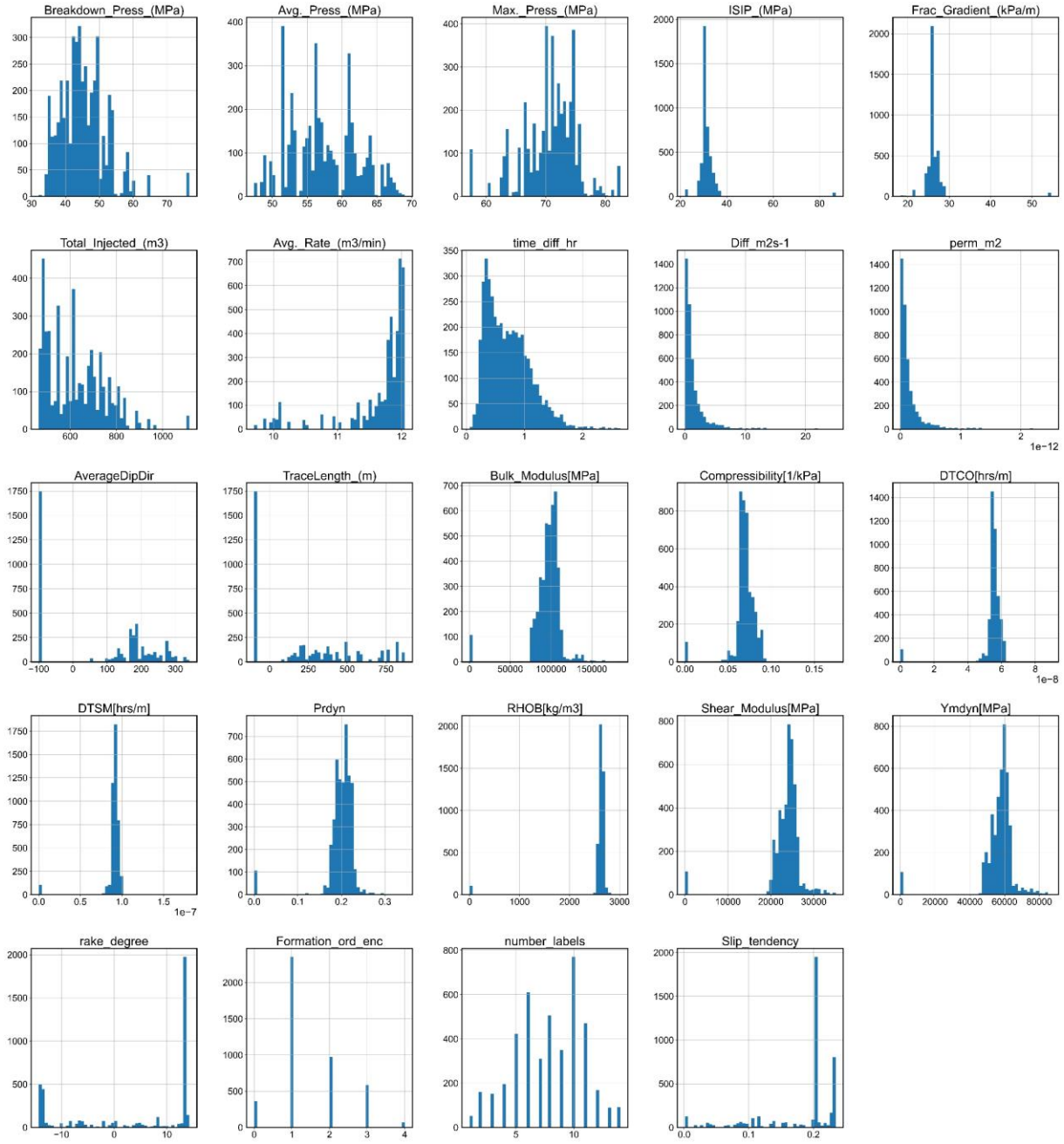


Figure 5: Distribution of values for each feature described in

Table *1*.

A correlation plot is used to visualize the linear relationship between the different features and the label. In these plots, the yellow colour indicates a strong positive correlation, and the dark blue indicates a strong negative correlation. A weak (near zero) correlation lies in the middle in green. The correlation plot for the input parameters used in this study as well as the target class is shown in Figure 6. This shows that none of the features have a high correlation with the targets, either the labels for the

classification analysis or the maximum magnitudes for the regression analysis. This again shows the complex relationship between the feature parameters and induced seismicity, and the need to use more advanced ML analyses.

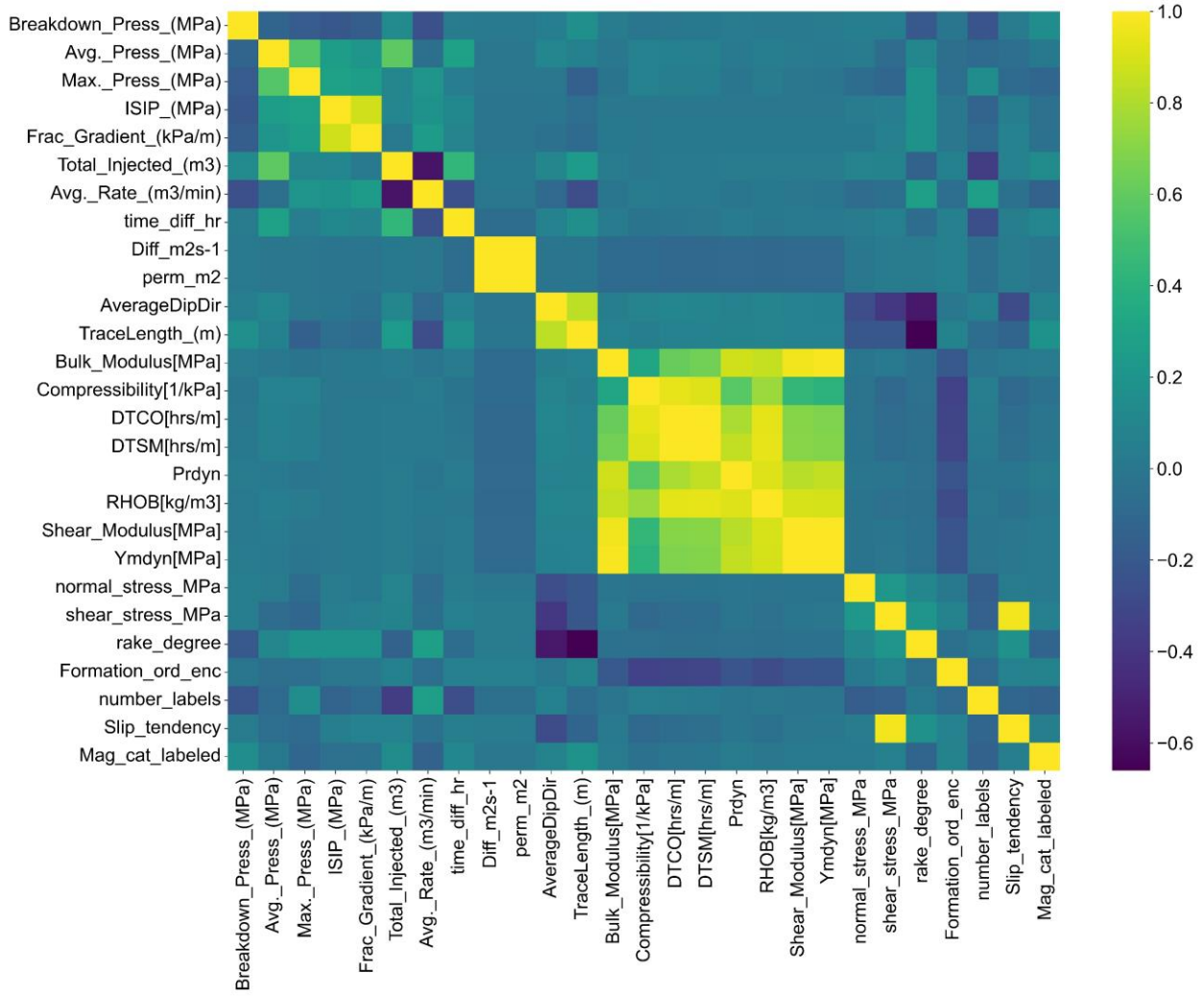


Figure 6. Correlation matrix for features used in the ML analyses, including the target value class.

3. Methodology

3.1. Machine Learning Overview

The two primary classes of ML algorithms are supervised and unsupervised techniques (Goodfellow et al., 2016). These are illustrated in Figure 77.

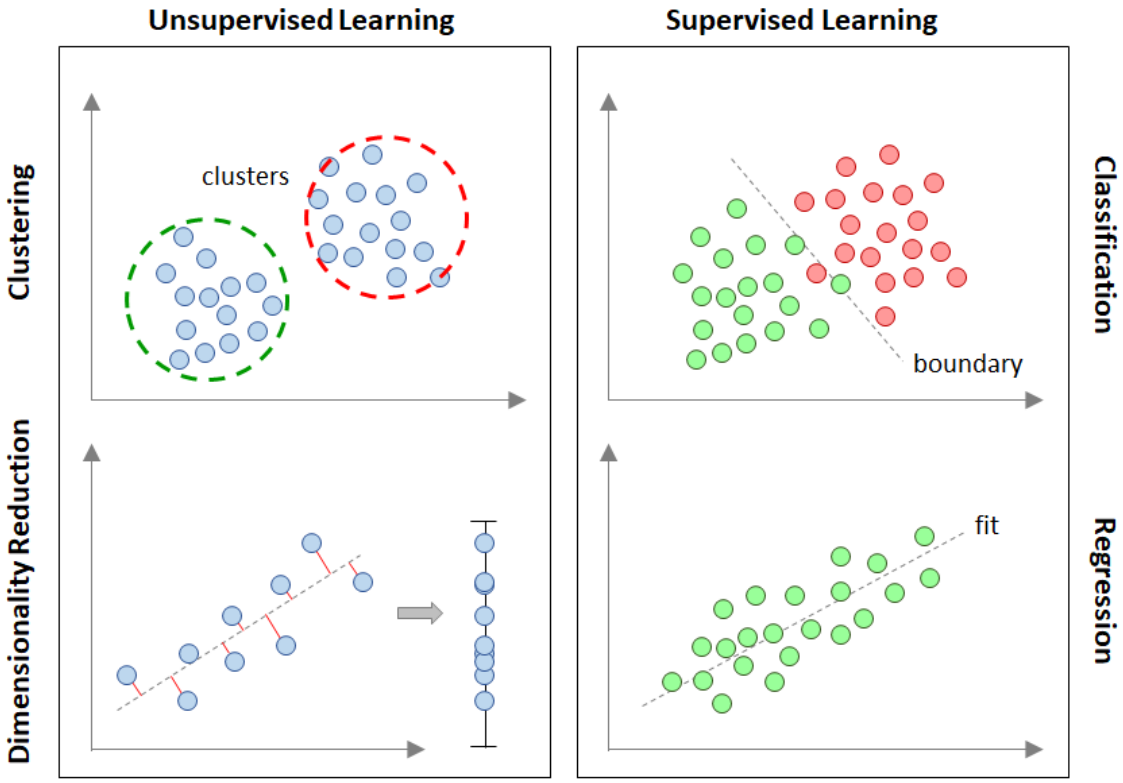


Figure 7. Types of machine learning algorithms classified as either unsupervised or supervised learning (from Amini et al., 2021).

In supervised learning, the ML algorithm "learns" or get "trained" to recognize a pattern or make general predictions using known examples. Supervised learning algorithms create a map, or model f, which relates a given set of data (or feature) in the form of a vector, x, to a corresponding label or target vector y: y = f(x). The labeled training data (data for which both the input and corresponding label (x, y) are known and given to the algorithm) are then used to optimize the model. For example, a well-trained model on different features of houses in Vancouver should be able to learn that the location of the house and its closeness to the beach matter and that these correlate positively with more expensive houses, while larger lot area may not always have a positive correlation with higher prices, for example if very large lot areas are associated with properties in more rural areas where houses may be cheaper.

Unsupervised learning methods, on the other hand, learn patterns or structure in the datasets without requiring any specific input features or labeling. In fact, these algorithms create their own features from the given data, and by detecting such features they can distinguish between different input data. A well-

known example is the feature-detection algorithm developed by Google that learned to recognize cats after being exposed to 10 million different unlabeled images (i.e., those not telling the ML if it is a cat or a tree) randomly sampled from YouTube videos (Le, 2013). Unsupervised learning is often used for exploratory data analysis or visualization in datasets for which no or few labels are available. These tend to be powerful models in some practical applications, for example self-driving cars. Although interpretability may not always be necessary in a highly accurate image recognition system, for example in the case of self-driving cars we want to make sure the algorithm distinguishes between a tree and a pedestrian but we may not care how it figured this out, it is critical when the goal is to gain physical insight into the system, which is the case for induced seismicity susceptibility (Amini et al., 2021; Bergen et al., 2019).

For this study, the objectives align with supervised learning and classification.

3.2. Machine Learning Steps

There are a number of key steps in performing a supervised classification ML data analysis. In this section, we present a general overview of these steps. The terminology used follows that commonly associated with ML analyses, including: “features” or input data (*x*-values); and “target values” or the outcome (*y*-value) that is going to be predicted by the model.

1) Divide data into training, validation and test splits.
2) Perform exploratory data analysis:
    - to detect outliers.
    - to evaluate missing data.
3) Pre-process data:
    - to clean the data.
    - to deal with missing data (e.g., by replacing them with average values).
    - to transform data into a ML-friendly database format.
    - to select and engineer features.
4) Cross-validate and tune hyperparameters. Each model has a combination of parameters that controls the learning process or how to map the input to the target value, and this step is performed to optimize such parameter configurations for the selected model features and target values through cross-validation.
5) Validate the performance of the best combination of hyperparameters using the validation dataset. If the performance is not satisfactory, the last step should be repeated until a satisfactory result is reached.
6) Using the most optimized hyperparameters, train the model on all training and validation datasets.
7) Finally, using the trained model from the previous step, investigate the model performance and feature importance ranking. The feature importance ranking shows the correlation between

each feature and its impact on the classification probability (i.e., likelihood of the outcome occurring). Note that such correlations can be either positive or negative.

It should be noted that during the data pre-processing step, we scaled all numerical inputs using a robust standard scaler (via a module in the scikit-learn Python ML tools). This not only removed the disproportionate impact of outliers (e.g., the negative values of the lineament features including TraceLength_(m)), but also ensured that the unit and the scale of the input data does not impact the performance of sensitive ML algorithms such as Logistic Regression.

Since our dataset is severely imbalanced (there are only 42 out of 4434 induced events with $\geq M_w1$), we have to use a number of metrics to evaluate the performance of the models. For likelihood classification, one of the more complete means of evaluating and visualizing model performance is the "confusion matrix". The confusion matrix breaks down the number of correct and incorrect predictions by the model, both with respect to the yes and no (positive or negative) of the binary classification problem. This is illustrated in Figure 8. In this case, the binary problem is whether a HF stage coincides with a large induced seismicity event of $\geq M_w1$ (i.e., seismogenic/positive for $\geq M_w1$ and non-seismogenic/negative for $< M_w1$). The model predictions of seismogenic and non-seismogenic are then compared to the actual field-based classification of seismogenic and non-seismogenic.

		Predicted Class		
		Seismogenic (Positive)	**Non-Seismogenic (Negative)**	
Actual Class	**Seismogenic (Positive)**	True Seismogenic (True Positive)	False Non-Seismogenic (False Negative)	**Sensitivity/Recall** $\frac{TP}{TP+FN}$
	Non-Seismogenic (Negative)	False Seismogenic (False Positive)	True Non-Seismogenic (True Negative)	**Specificity** $\frac{TN}{TN+FP}$
		Precision $\frac{TP}{TP+FP}$	**Negative Prediction** $\frac{TN}{TN+FN}$	**Accuracy** $\frac{TP+TN}{TP+TN+FP+FN}$

Figure 8. Illustration of a confusion matrix used to evaluate machine learning performance, and several advanced performance metrics calculated from the confusion matrix (from Amini et al., 2021).

The confusion matrix can then be used to calculate more advanced classification metrics, including (Figure 8):

Accuracy: Represents the number of correct predictions (both true positive and true negative) over the total number of predictions. This metric is best used for balanced datasets (both positive and negative classes have significantly different number of data occurrences).

Precision: Represents the ratio of true positive predictions to the total number of true and false positive predictions. If there are no incorrect false positive predictions, then the model has 100% precision. This metric is well suited for unbalanced datasets.

Recall: Represents the sensitivity of the model, by comparing the number of true positive predictions to the total number of true positive and false negative predictions. Thus, the recall rate is penalized whenever an incorrect false negative is predicted. This metric is also suitable for unbalanced datasets.

Precision is often used in cases where classification of true positives is a priority; i.e., it evaluates the model's success rate in correctly predicting seismogenic wells. Recall is often used in cases where classification of false negatives is a priority. In this case study, since the dataset is severely imbalanced and we are interested in finding correlations between the input features and positive class, we choose to optimize and analyze the performance of the ML models using the recall metric: out of the 42 seismogenic events how many does the model predict correctly.

3.3. Machine Learning Models Tested

Several ML algorithms were tested for this study. These are listed and described in Table 2. Note that the dummy classifier model is only used to create a baseline for the other models.

Table 2. Summary descriptions of the different machine learning models used for the classification and regression analyses.

Model Name	Abbreviation	Description
Dummy Classifier (Pedregosa et al., 2011)	Dummy Classifier	A classifier that makes predictions using simple rules. It does not use features when predicting and is useful as a simple baseline to compare with other classifiers.
Logistic Regression (Walker and Duncan, 1967)	LR	A linear model classifier. It works by fitting a logistic function to the data, which maps the input features to a probability value between 0 and 1. It predicts the class with the highest probability.
Decision Tree (Breiman, 1984)	DT	This classifier works by recursively partitioning the input space into smaller regions, based on the values of the input features. At each level of the tree, a decision is made based on the value of a single feature, with the goal of maximizing the information gain or minimizing the impurity of the resulting subsets.
Random Forest (Breiman, 2001)	RF	This ensemble classifier works by building multiple decision trees and aggregating their results to make a final prediction. Each decision tree is built on a random subset of the training data and a random subset of features. This helps

		to reduce overfitting and increase the generalization of the model.
Support Vector Classifier (Cortes and Vapnik, 1995)	SVC	This classifier works by finding a hyperplane in a high-dimensional space that maximally separates the different classes. The algorithm tries to maximize the margin between the classes and can handle non-linearly separable data by transforming the input space to a higher dimensional feature space.
XGBoost (Chen and Guestrin, 2016)	GXB	"Extreme Gradient Boosting" is another decision tree-based algorithm. It works by iteratively adding weak learners (decision trees) to the model and optimizing a loss function to minimize errors.
Light GBM (Ke et al., 2017)	LGBM	"Light Gradient Boosting Machine" is a lighter version of XGBoost classifier. It works by using a novel histogram-based approach to bin continuous features into discrete values for the decision trees.

4. Likelihood and Susceptibility Classification Analyses

4.1. Feature Importance (Geological and Operational Data)

The classification analyses focused on identifying the most important features and their correlation with respect to the likelihood of a HF stage being seismogenic or not. We performed two sets of classification modeling:

1) A feature importance analysis targeting likelihood that a HF stage produces an induced seismicity event with $\geq M_w1$. This makes use of a combined dataset of all geological and operational features at each HF stage.

2) A feature importance analysis targeting susceptibility of a well pad location being seismogenic and capable of producing events with $\geq M_w1$. This makes use of a restricted dataset limited to geological features and those operational features that apply to susceptibility.

A key distinction is made here between likelihood and susceptibility. Likelihood refers to the probability of a hazard occurring, whereas susceptibility is more specific and refers to the spatial probability of the hazard occurring (i.e., the susceptibility of a new well at a given location being seismogenic). Accordingly, the likelihood analysis makes use of all feature data, both geological and operational, whereas the susceptibility analysis excludes most operational features, as these do not apply to the pre-existing conditions that would relate to the susceptibility of a new well location to induced seismicity. Exceptions include operational features that serve as a proxy for the physical properties of the targeted formation, for example breakdown pressure and instantaneous shut-in pressure (ISIP).

Distinguishing between geological and operational features is also of interest as the geological factors relate to the conditions that are in place and cannot be controlled or manipulated (outside of avoidance), i.e., susceptibility. These differ from most operational factors (i.e., those related to HF and well completions), which can be manipulated, offering a means to potentially mitigate induced seismicity hazards for a susceptible formation.

4.1.1. Model Training and Testing Results

To begin, since the OP1 dataset involves a timeseries, the combined geological and operational features data were first sorted based on the time of event occurrences. It was then split into a training set (70%) and test set (30%). The Dummy Classifier was used as the baseline with a 99% accuracy score, which clearly indicates a severely imbalanced dataset. Thus, as explained earlier, to properly optimize the performance of the models, the calculated recall score was considered. Comparing the seven ML algorithms tested, the top performing models selected for interpretation were those with a recall of more than 50%. These are Logistic Regression (LR) with 67%, Random Forest (RF) with 67%, and Support Vector Classifier (SVC) with 78% test recall scores (see Figure 99).

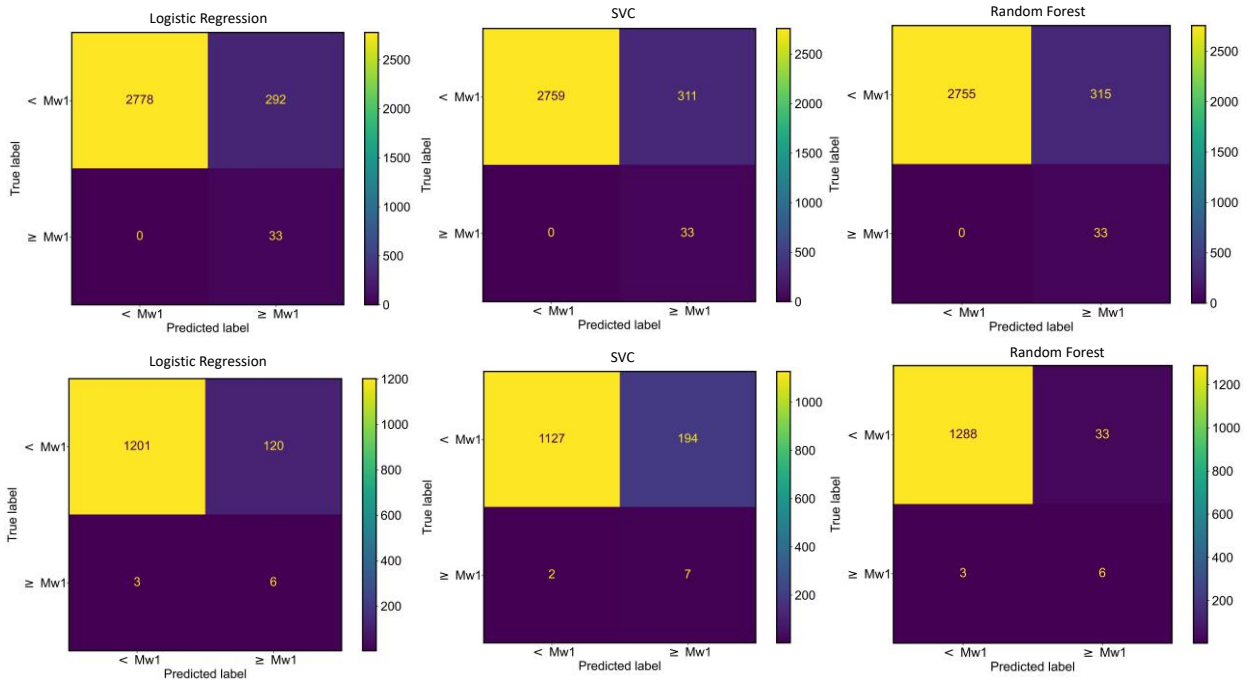


Figure 9. Confusion matrices of the training results (top) and test results (bottom), for the three top performing ML algorithm models according to recall score when considering the combined geological and operational features dataset.

4.1.2. Model Interpretation

Interpretability in machine learning refers to the ability to understand and explain how a model makes its predictions or decisions. It is an important aspect of ML because it allows us to identify correlations and understand why the model is making specific predictions. There is no mathematical definition of interpretability. However, one widely cited definition of interpretability in ML is by Miller (2017), who defines it as "the degree to which a human can understand the cause of a decision."

One of the widely used ML techniques in such interpretation is to review the strength (and the direction for linear-based models) of the influence of each feature in predicting the target classes. Thus, to thoroughly understand and interpret the performance of each model, we used three techniques, which are briefly reviewed in the following sub sections. It must be noted again that these interpretations are evaluated using a recall metric, and more importantly, correlations determined through ML outputs do not establish a cause-and-effect relationship (i.e., "correlation is not causation").

Feature Importance, Coefficients, and SHAP Interpretation

For the top three performing models (i.e., LR, SVC, RF), we first reviewed the feature coefficient and importance of the models themselves. For the linear models (LR and SVC), the "feature coefficients" were analyzed. In these models, each feature is assigned a coefficient that represents the strength and direction of its influence on the target variable. Features with larger coefficients (absolute values) are considered more important in the model's decision-making process, as they have a stronger impact on the target variable. For the tree-based model (RF), feature importance refers to the measure of the contribution of each feature in the decision-making process of the model. Features with higher importance scores are considered more important in the model's decision-making process, as they have a greater impact on the score of the model's predictions. However, it should be noted that these coefficients do not have any physical meaning and are used as relative values for comparison and not as absolute values.

Figure 10 shows the feature importance plots resulting from the LR, SVC and RF analyses for OP1.

Table 3 lists the top five important features for each of these models. These show that the lineament orientation (AverageDipDir) and its size (TraceLength_(m)) are the most important features. For the linear models (LR, SVC), the larger these values, the higher the likelihood the model will predict the stage as being seismogenic (i.e., $\geq M_W1$). Similar trends and correlations between these features and the prediction of seismogenic stages can be seen for the non-linear RF model outputs in Figure 11. Comparing the dip direction orientation of the lineaments (Figure 5) with the orientation of S_{Hmax} (*Figure **4***), these observation suggest that these sub-vertical structures will be more seismogenically active if they are striking parallel with S_{Hmax}. In addition, the findings of these models are similar to the results from Amini et al. (2021) that the geological features generally rank higher than operational features with respect to importance.

Table 3. The top five ranked features when considering all geological and operational features with respect to importance in predicting induced seismicity likelihood for the hydraulic fracturing stages of the OP1 well pad, for each of the top three performing ML classifier models. See

*Table **1** for full descriptions of the features.*

Rank	Logistic Regression	Support Vector Classifier	Random Forest
1	AverageDipDir	AverageDipDir	TraceLength_(m)
2	TraceLength_(m)	TraceLength_(m)	AverageDipDir
3	Formation_ord_enc	Formation_ord_enc	rake_degree

4	Total_Injected_(m3)	Slip_tendency	number_labels
5	number_labels	Prdyn	Total_Injected_(m3)

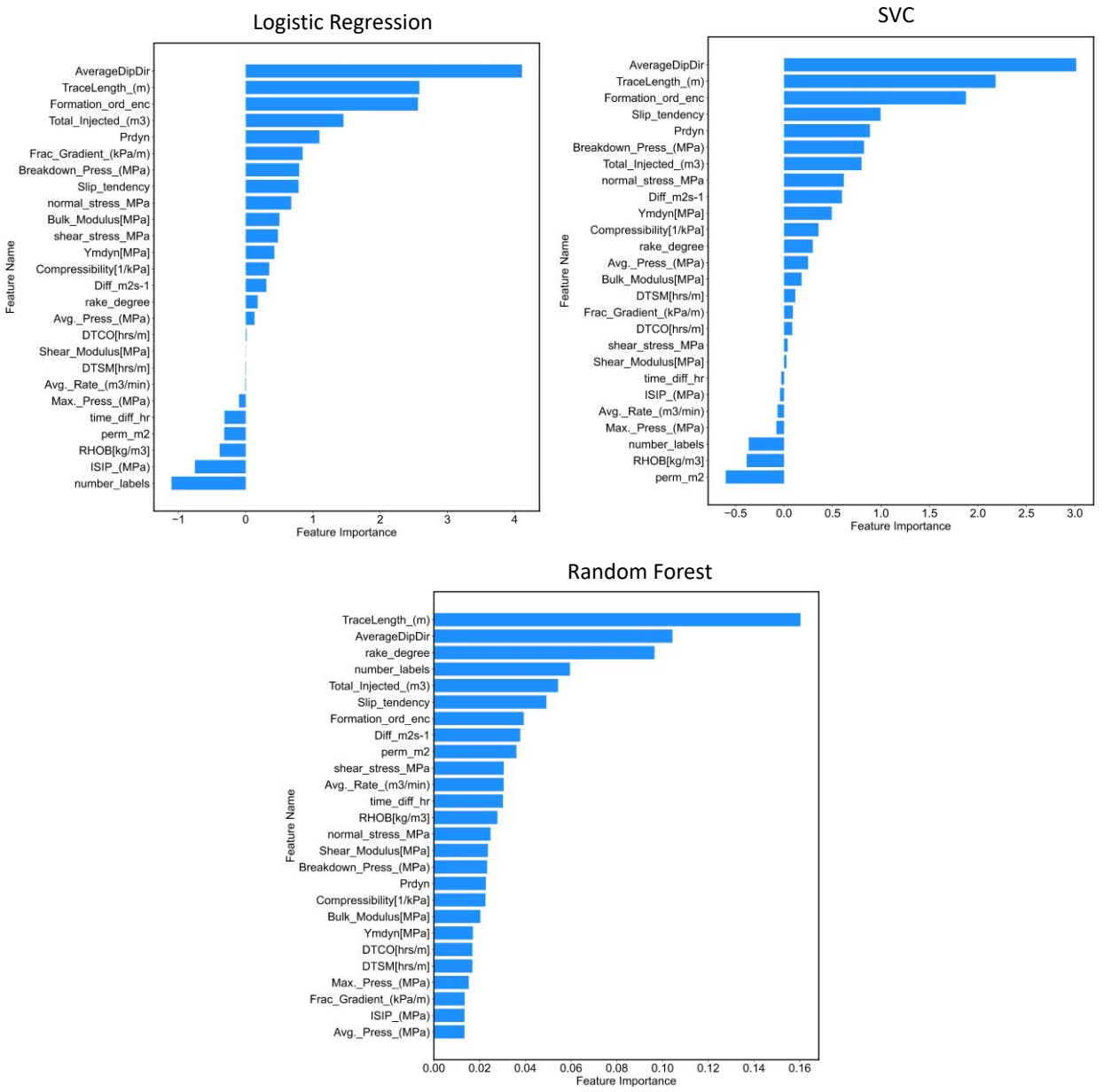


Figure 10. Importance ranking of geological and operational features for the top three performing ML classifier models applied to the OP1 dataset. Note that these coefficients do not have any physical meaning and are used as relative values for comparison and not as absolute values. See

Table **1** *for full descriptions of the features.*

Since the tree-based RF model is non-linear, to interpret the features we use SHapley Additive exPlanations (SHAP) analyses tool (Lundberg and Lee, 2017) for in-depth interpretation of the results. SHAP utilizes a game theory approach to help interpret predictions from complex model output. SHAP assigns each data point a SHAP value, which is used for the model interpretation. The SHAP value represents the estimated impact of a particular feature on the model's prediction for a given instance. A positive SHAP value indicates that the feature has a positive impact on the model's prediction (i.e., higher feature values increase the prediction), while a negative SHAP value indicates that the feature has a negative impact on the prediction (i.e., higher feature values decrease the prediction). The swarm plot consists of a scatterplot, where each point represents a single instance in the dataset. The position of the point on the x-axis represents the SHAP value for a particular feature, while the y-axis indicates the index of the instance in the dataset. By default, the points are arranged in a way that avoids overlapping, creating a "swarm" effect. The SHAP swarm summary plot for the RF model is provided in Figure 11.

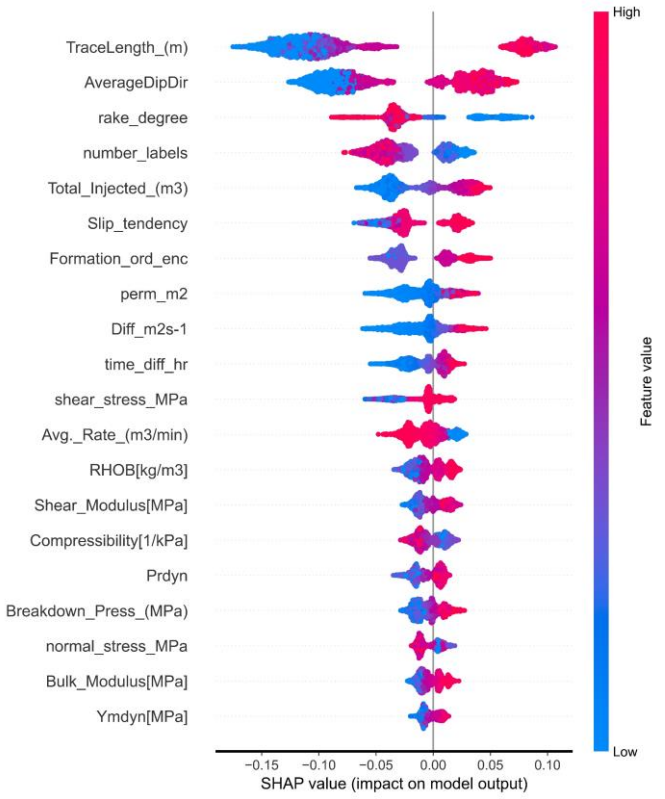


Figure 11. SHAP swarm summary plot for the Random Forest model considering both geological and operational features.

For all results (Figure 1010 and Figure 11), it can be seen that the depth of the formation correlates with the prediction of seismogenic class: strongly for the LR and SVC models, moderately for the RF model. These suggest that the deeper the lineament the larger the seismic moment release. As would be expected, with increasing depth, a higher amount of elastic strain energy would be stored along structures and released upon slip.

The number of lineaments that are connected to the HF stage (fracture intensity) is also amongst the more important geological features. Figure 10 and Figure 11 show an inverse relationship between the prediction of seismogenic class and the values of this feature. In other words, the larger the number of intersected lineaments the lower the probability of a seismogenic response. This was a key finding in the empirical and numerical results from the first part of the larger study for BC OGRIS (WSP 2023). In summary, we learned from the data that the injection into a highly fractured reservoir results in lower pressure buildup due to the higher storage capacity of the connected network of fractures. Considering the hydraulic diffusivity, fracture intensity, pressure and seismic response of the reservoir to fluid injection collectively, we concluded that there exists an inverse relationship between the hydraulic diffusivity and the number of connected structures, whereby higher-pressure buildup and larger event magnitudes are associated with higher hydraulic diffusivity values. This can also be seen in the ML results. Figure 12 isolates the SHAP results showing its correlation with fracture intensity and hydraulic diffusivity.

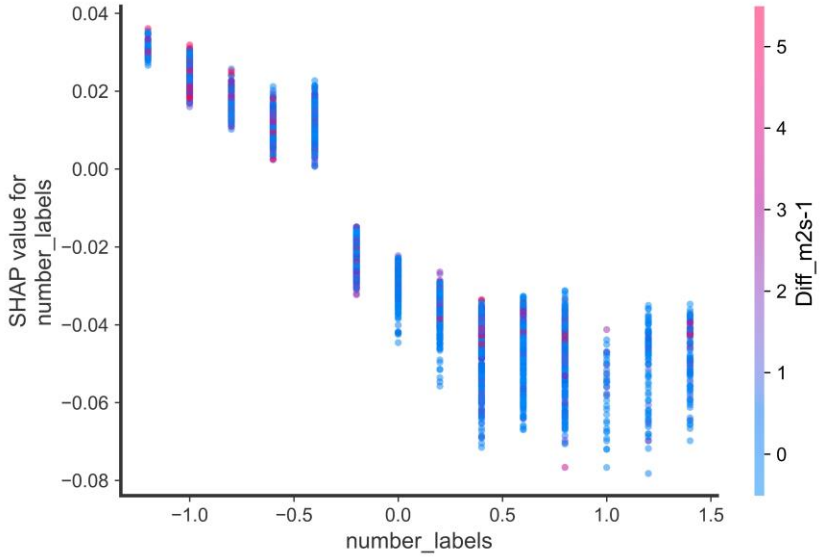


Figure 12. SHAP interpretation of the dependent correlation between fracture intensity (number_lables), hydraulic diffusivity (Diff_m2s-1), and the SHAP values.

Further analysis of Figure 10 indicates that for the linear model interpretations (LR, SVC), higher dynamic Poisson's ratio values correlate with the prediction of seismogenic class. Poisson's ratio describes how a material deforms when it is subjected to stress. It is closely related to the volume of the rock that undergoes deformation, and therefore to the amount of stored elastic strain energy. In general, rocks with high Poisson's ratios tend to deform more in response to stress, and therefore store more elastic strain energy than rocks with lower Poisson's ratios. Thus, a larger portion of stored elastic energy can be seismically released. Similarly, these models indicate that breakdown pressure as well as fracturing gradient, both proxies for formation rock strength, are among the important influential features. Again, the higher the yield strength of the rock, the larger the amount of elastic strain energy that can be stored by the rock, and hence, the increased likelihood of larger seismic moment release. It should be

noted that these features by themselves aren't lone predictors but depend on other conditional factors related to the geology and operations (e.g., the depth and stress environment).

For all three models, slip tendency is also considered an important feature. The larger this value, the higher the likelihood of shear slip and potentially larger seismic moment release. This is dependent on other features, for instance the trace length, which corresponds with rupture length potential.

Finally, the RF model results show that rake-angle is among the most top important features unlike the linear models. This again points to the importance lineament with regard to the in-situ stress. However, since the in-situ stress state is transform (i.e., $S_{Hmax} > S_V > S_{hmin}$), and the structures are vertical, the resulting rake-angles are closely generates strike-slip movement (i.e., left-lateral movement, see Aki and Richards, 2002). Mehrabifard and Eberhardt (2023) studied a global dataset of induced seismicity and concluded that the transform stress regime corresponds with a larger induced seismic hazard potential. Therefore, the presence of sub-vertical structures in the area with combination of transform stress regime lead to high seismic hazard potential, which is confirmed by observation of rake angle among the important features in this model.

The above highlights the importance of geological features. With respect to the most important operational features, the results point to the total injected fluid volume, where high injection volumes have a positive impact on model prediction of seismogenic response. This is in agreement with our regional Montney induced seismicity susceptibility study (Amini et al., 2021). The influence of injection volume is well studied and has been shown to have a positive correlation with induced seismicity (e.g., McGarr, 2014; Schultz et al., 2018).

Permutation Importance

Permutation importance analyses (Breiman, 2001) measure the increase in the prediction error of a model if the feature's values are permuted. In other words, a feature is deemed important if shuffling and reordering its values increases the model error as this indicates the model relies on the feature for a correct prediction. If shuffling and reordering the feature's values leaves the model error unchanged, then the feature did not have any significant influence on the model's ability to correctly predict and is unimportant. This procedure breaks the relationship between the feature and the target prediction, and thus any decrease in the model score is indicative of how much the model depends on the feature.

Figure 13 plots the results of the permutation importance analysis performed. The y-axis shows the features, and the x-axis shows the model score loss (i.e., increase in model error). These values do not have any physical meaning and simply provide a relative measure for comparison. The error values for each feature are reported as a box plot where the lower and upper edges of the box indicate the 25th and 75th percentiles, respectively, and the central mark indicates the median. The whiskers extend to the most extreme data points not considered to be outliers. Outliers are plotted as individual points (i.e., open circles in this case).

Overall, the results from the permutation importance analysis supports the finding from the previous technique. The results show that the lineament orientation (AverageDipDir) and its size (TraceLength_(m)) are among the most important features, the latter having the highest importance scores. Similarly, the models agree that number of connected structures (number_labels), slip tendency of the lineaments, and Poisson's ratio are also important features. The total injection volume is one of the most important operational features among both linear models (LR and SVC).

The fact that the non-linear RF model is sensitive to a smaller number of features compared to the linear SVC and LR models, suggests that this model is focusing on the most relevant and informative features in the dataset. This can be a sign of a well-generalized and interpretable model. Therefore, the findings of this analysis may indicate that Random Forest is a more reliable model between the three.

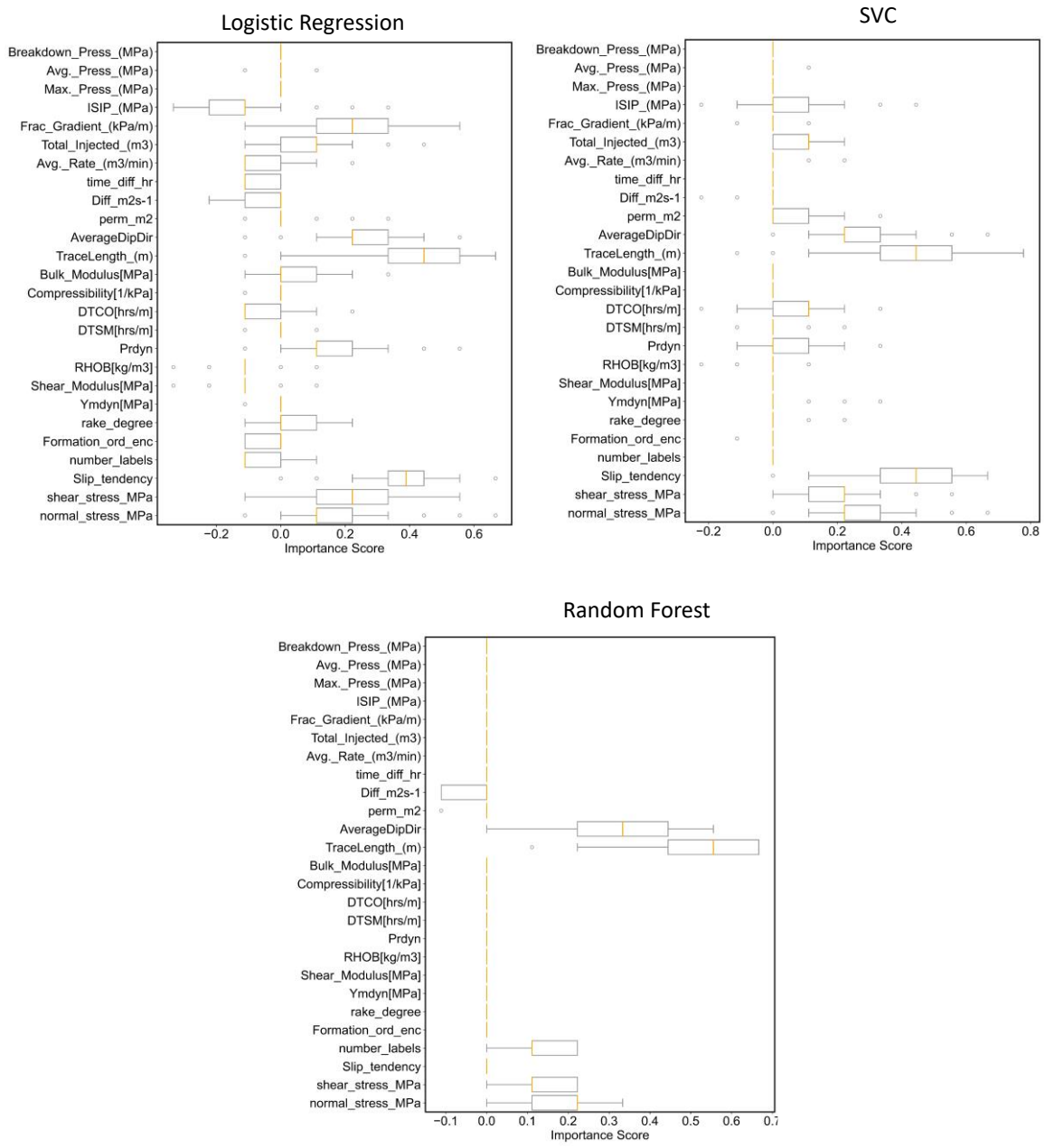


Figure 13. Permutation importance plots for the top three performing machine learning models when considering geological and operational features. See text for explanation.

Recursive Feature Elimination

Recursive feature elimination is a technique for selecting the most important features in a ML model. It works by repeatedly fitting the model with a subset of the features and eliminating the least important

feature at each iteration. The importance of each feature is measured using a model-specific metric, such as feature coefficients (for LR) or feature importance (for RF). This technique provides a ranked list of feature importance that can be used to interpret the performance of each model.

Table 4 lists the top five ranked features using this technique. It should be noted that the features with the same amount of importance are assigned the same rank. Comparing the results to Table 3, it can be seen that the lineament orientation (AverageDipDir) and its size (TraceLength_(m)) continue to be consistently ranked at the top. The two linear models (LR, SVC) also agree that the number of connected structures (number_labels) and hypocentral depth (Formation_ord_enc) are amongst the most important features, with the former also continuing to rank in the top five for the RF model. New to the top five for the LR and RF models is the slip tendency of the lineaments (Slip_tendency). The formation permeability (perm_m2) also newly appears in the top five for the RF model, ranking amongst the top important features from the recursive feature elimination analysis. It is interesting that the total injection volume is the only important operating feature that appears for the RF classifier in both the feature importance and recursive feature elimination analyses (Tables 3 and 4, respectively). Generally speaking, these results are consistent with the findings of the previous techniques.

Table 4. The top five recursive feature elimination ranked features when considering all geological and operational features with respect to importance in predicting induced seismicity likelihood for the hydraulic fracturing stages of the OP1 well-pad, for each of the top three performing ML classifier models. See

Table **1** *for full descriptions of the features. Repetitive ranking means the features are equally important.*

Logistic Regression		**Support Vector Classifier**		**Random Forest**	
Features	**Rank**	**Feature**	**Rank**	**Feature**	**Rank**
AverageDipDir	1	AverageDipDir	1	TraceLength_(m)	1
TraceLength_(m)	1	TraceLength_(m)	1	AverageDipDir	1
Formation_ord_enc	1	number_labels	1	Slip_tendency	1
number_labels	1	Formation_ord_enc	2	perm_m2	1
Slip_tendency	2	shear_stress_MPa	3	rake_degree	2
Frac_Gradient_(kPa/m)	3	Slip_tendency	4	Total_Injected_(m3)	3
ISIP_(MPa)	4	normal_stress_MPa	5	number_labels	4
Total_Injected_(m3)	5	DTCO[hrs/m]	6	time_diff_hr	5

4.2. Feature Importance (Geological Data)

Susceptibility maps are created by combining different features that contribute to a hazard, indicating where the hazard is more likely to occur spatially. However, as previously noted, not all features lend themselves to spatial susceptibility. Already in the previous section, where ML importance analyses considered both geological and operational features, the results were dominated by the importance of several geological features. The next step specified by Amini et al. (2021) in generating a ML-derived induced seismicity susceptibility map involves repeating the feature importance analysis, but with a focus on only those features that can be used to assess susceptibility. Included are all geological features

and any operational features that serve as a proxy for formation properties; these are indicated as type "G" in

Table 1).

4.2.1. Model Training and Testing Results

The same procedures as explained in Section 4.1.1. were followed in order to train and test the susceptibility-focused classification models. Again, a 70%-30% split was used for the training and testing sets, respectively. Figure 144 represents the confusions matrices for the top performing models. The testing recall scores for the LR, SVC and RF models are 55%, 67%, and 67%, respectively. Comparing the two sets of confusion matrices (Figure 9 and Figure 144), it seems that removing the operational features not related to the susceptibility, slightly reduces the recall scores.

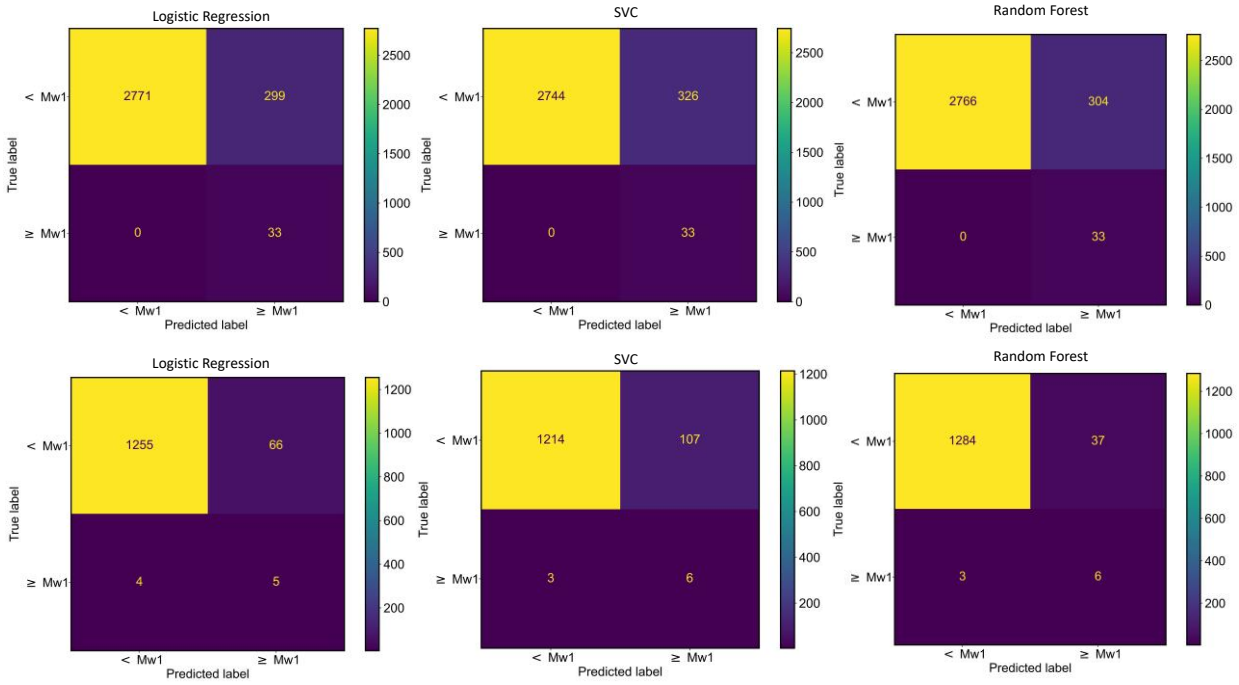


Figure 14. Confusion matrices of the training results (top) and test results (bottom), for the three top performing ML algorithm models when only considering the susceptibility-focused features dataset.

4.2.2. Model Interpretation

Feature Importance, Coefficients, and SHAP Interpretation

The training results for the three models were further analyzed to investigate their ranking of feature importance with respect to induced seismicity susceptibility. Figure 15 shows the feature importance plots resulting for these models, and Table 5 lists the top five important susceptibility features. The structural orientation and dimensions remain the most important features. These are followed by hypocentral depth, number of connected structures to the HF stage, stresses on the structures (i.e., slip-

tendency and rake angle), and Poisson's ratio. These are consistent with the findings of Section 4.1.2. Among these the inverse relation between the seismogenic class and the number of connected structures (i.e., number_labels), can be seen by the negative coefficients in Figure 155 for the linear models (LR and SVC), and the SHAP swarm summary plot in Figure 16 for the RF model (where the red dots in the plot correspond to the prediction of the non-seismogenic class; i.e., negative SHAP values).

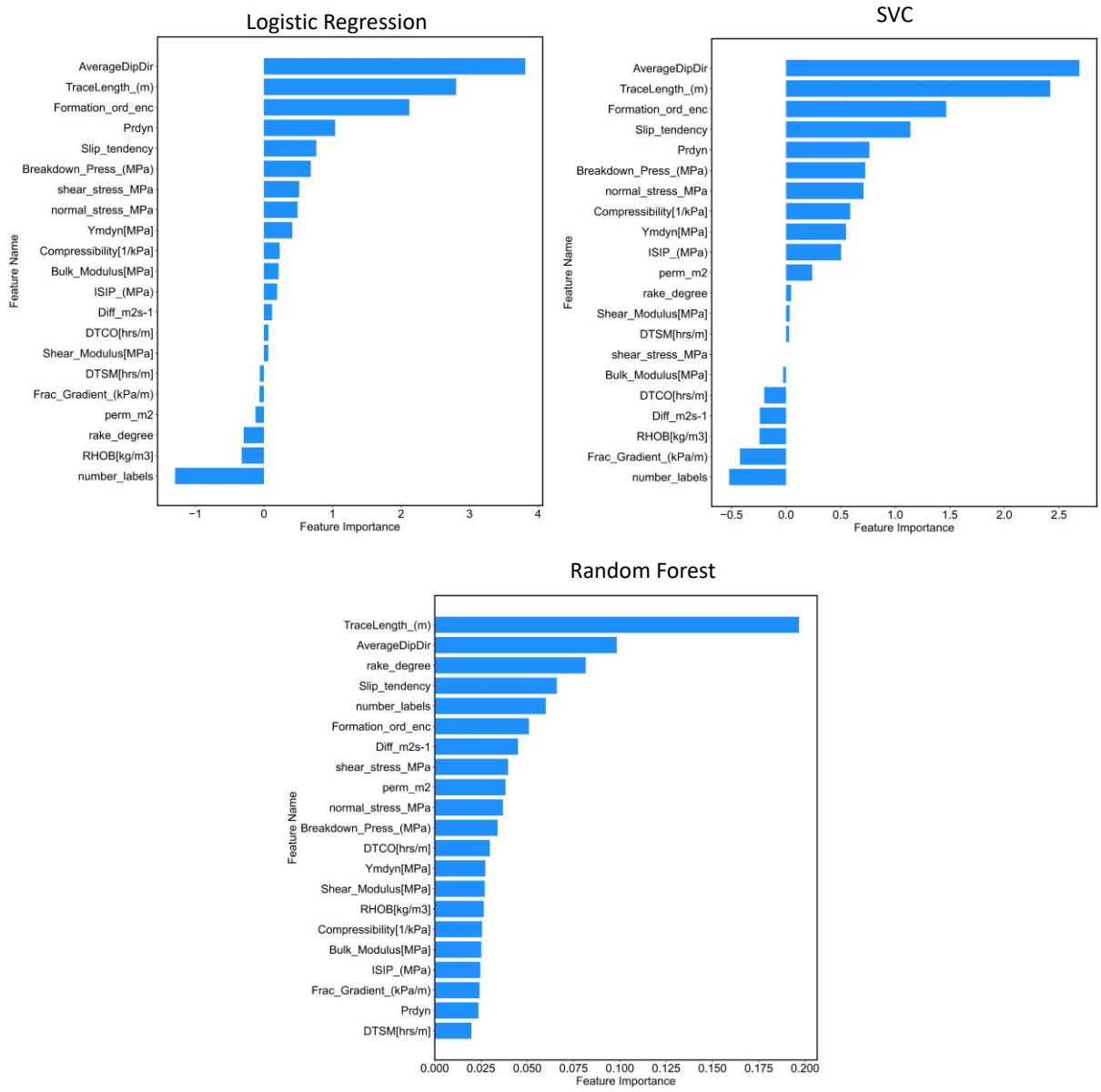


Figure 15. Importance ranking of susceptibility features for the top three performing ML models. Note that the coefficients do not have any physical meaning and are used as relative values for comparison and not as absolute values.

Table 5. The top five ranked susceptibility features with respect to importance in predicting induced seismicity likelihood for the hydraulic fracturing stages of the OP1 well-pad, for each of the top three performing ML classifier models.

Rank	Logistic Regression	Support Vector Classifier	Random Forest
1	AverageDipDir	AverageDipDir	TraceLength_(m)

2	TraceLength_(m)	TraceLength_(m)	AverageDipDir
3	Formation_ord_enc	Formation_ord_enc	rake_degree
4	number_labels	Slip_tendency	Slip_tendency
5	Prdyn	Prdyn	number_labels

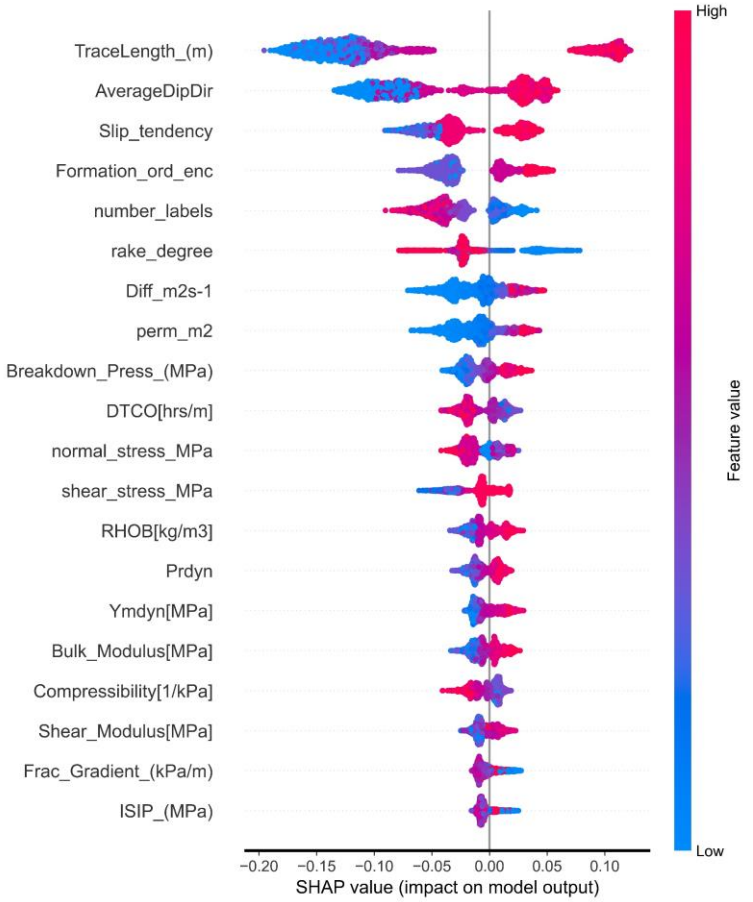


Figure 16. SHAP swarm summary plot of Random Forest model considering only the susceptibility features.

Permutation Importance

The permutation importance analyses of the susceptibility features are shown in Figure 177. Generally, the results support the findings of the feature importance analysis. The LR model is seen to be more sensitive to the slip tendency and the resultant stresses on the lineaments, whereas the RF model is most sensitive to the orientation and the dimension of the structures. The SVC model is most sensitive to the stresses on the lineaments, and their orientation and dimensions. In addition, the fact that all three models are sensitive to a similar number of features may suggest that the models are equally reliable.

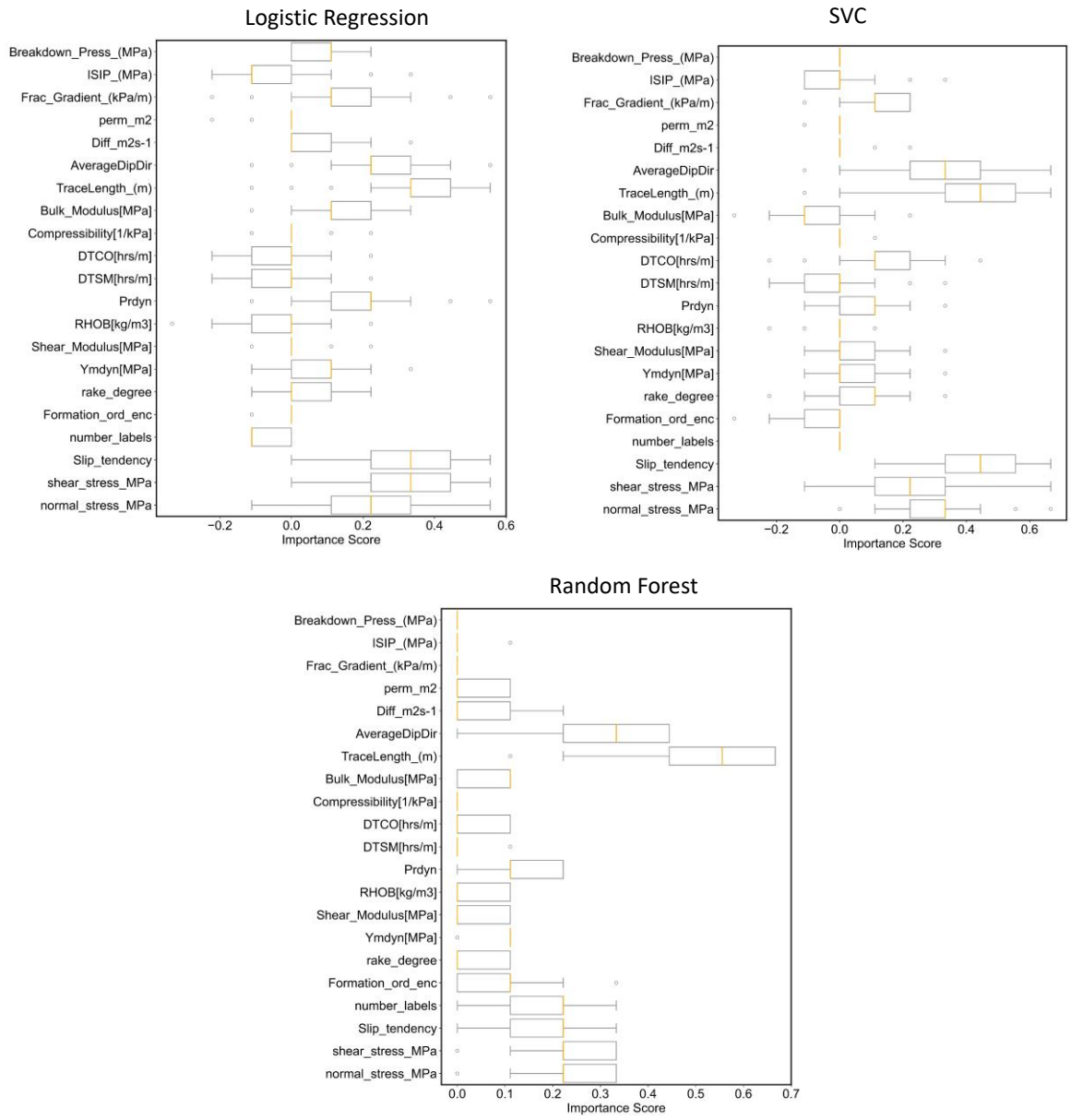


Figure 17. Permutation importance plots for the top three performing machine learning models when considering only the susceptibility features.

Recursive Feature Elimination

Table 6 lists the top five ranked susceptibility features using this technique. It should be noted that the features with similar importance have the same rank. The importance of the features agrees with the previous findings that the fracture intensity, the size and orientation of the structures, the stresses acting on these structures, and the hypocentral depth of the shear slip events are the most important features. Using this method, the RF model places a larger importance on the fluid pathway permeability, which agrees with the SHAP analysis in Figure 16 that shows the model predicting the seismogenic class for highly permeable fluid pathways. This again is consistent with the findings reported in WSP 2023.

The combination of such high permeable corridors and low fracture intensity can result in large induced seismic events in response to fluid injection (also see Figure 12).

Table 6. The top five recursive feature elimination ranked features when considering only the susceptibility features with respect to importance in predicting induced seismicity likelihood for the hydraulic fracturing stages of the OP1 well-pad, for each of the top three performing ML classifier models. See

Table **1** *for full descriptions of the features. Repetitive ranking means the features are equally important.*

Logistic Regression		**Support Vector Classifier**		**Random Forest**	
Features	**Rank**	**Feature**	**Rank**	**Feature**	**Rank**
AverageDipDir	1	AverageDipDir	1	AverageDipDir	1
TraceLength_(m)	1	TraceLength_(m)	1	TraceLength_(m)	1
number_labels	1	number_labels	1	perm_m2	1
Formation_ord_enc	2	Formation_ord_enc	2	Slip_tendency	2
Slip_tendency	3	shear_stress_MPa	3	number_labels	3
Frac_Gradient_(kPa/m)	4	Slip_tendency	4	rake_degree	4
ISIP_(MPa)	5	normal_stress_MPa	5	Diff_m2s-1	5
Prdyn	5	DTCO[hrs/m]	6	Prdyn	6

5. Induced Seismicity Susceptibility Maps and Interpretation

Machine learning results from the top two performing classification models, Linear Regression (LR) and Random Forest (RF), were used to generate induced seismicity susceptibility maps for the well pad-scale OP1 case study area in KSMMA. The Support Classifier Vector (SVC) models were not used as they are not naturally probabilistic models and do not output a prediction score between 0 and 1. (Instead, SVC models output a binary decision boundary that separates the data into two classes).

The generation of the susceptibility maps involved applying the trained models to the well-pad areal extent to calculate the probability of induced seismicity for a grid of spatial points. The models were only trained on the 70% split of the data (training set) using the features of the test dataset. It should be noted that the output of the classification model is the probability of a given point being seismogenic. In other words, the probability of observing an induced seismicity event of moment magnitude one or greater ($\geq M_W 1$). In addition, susceptibility maps only report the spatial probability of a hazard, in this case a well location being seismogenic; they do not report the temporal probability or related level of hazard (i.e., expected magnitude of the induced seismicity event).

To facilitate the comparison, both maps generated (LR and RF) are shown together in Figure 18. The network of interpreted lineaments are superimposed over these susceptibility maps. Both models predict almost identical induced seismicity hotspots with minor differences. As can be seen, where the number of connected structures to the active HF stage is low, the models predict a higher probability of seismogenic response to fluid injection.

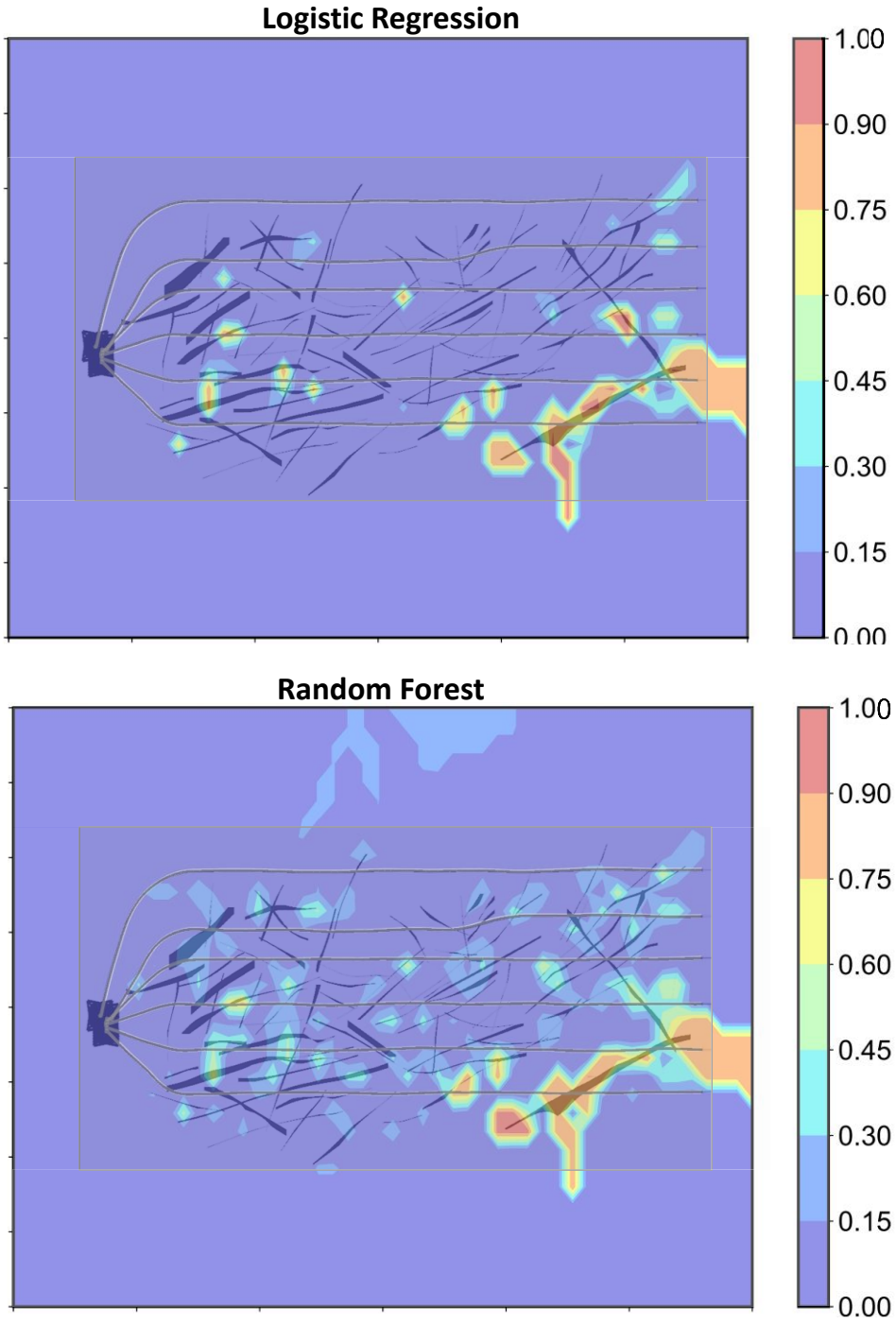


Figure 18. Induced seismicity susceptibility maps for the OP1 well-pad area in KSMMA. The color bar represents the probability of experiencing induced seismicity with ≥ M_W1. Included is the superimposed interpreted network of structures.

To evaluate the performance of the generated induced seismicity susceptibility maps, Figure *19* reproduces these with the locations of induced events with ≥M_W1 superimposed on top. This comparison shows that there is a good match between the locations of these events and areas identified by the susceptibility maps as having high probabilities of seismic potential. For both models,

there are a few small areas where the model predicts an increased potential for larger induced seismicity events for which none were recorded. This might be indicative of the parts of the reservoir where the geological conditions are unfavorable but there might be operational factors that have contributed to avoiding a seismogenic response to fluid injection.

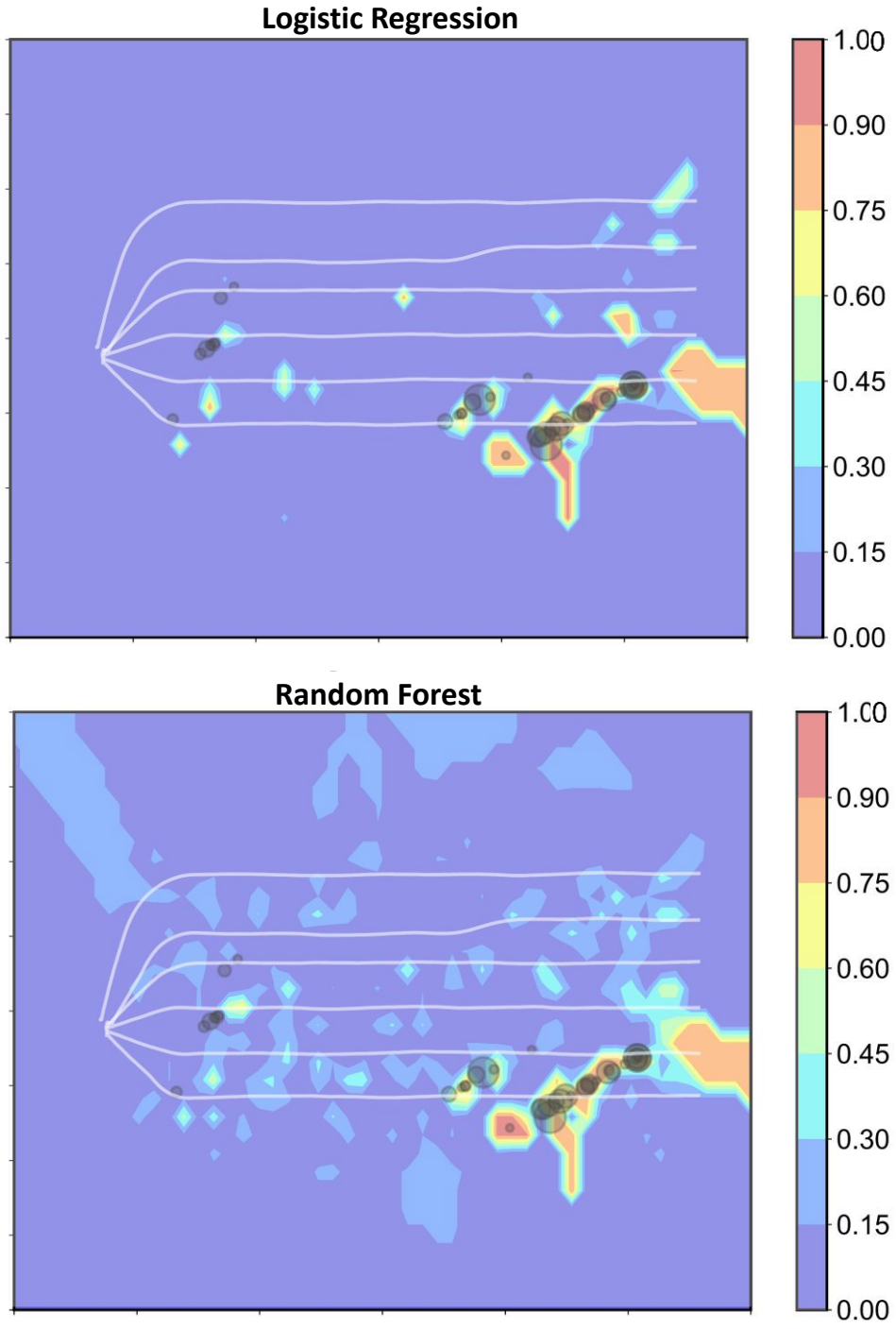


Figure 19. Locations of induced seismicity events (≥M_W1) superimposed on the susceptibility maps generated from the Logistic Regression and Random Forest classification ML models.

6. Discussion and Conclusions

Results were presented from a machine learning analysis aimed at identifying the connection between geological and operational features and the occurrence of relatively large induced seismic events ($\geq M_W1$) in the Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA). To this end, we conducted two types of classification modeling:

1) a classification analysis to identify important features by combining all geological and operational features for each hydraulic fracturing stage.
2) a classification analysis to determine spatial susceptibility by only using geological and relevant operational features.

The results of these analyses suggest that the geological features correlate more strongly with fluid injection induced seismicity. This was observed in the results of the feature importance analyses considering both geological and operational features and validated through the generated susceptibility maps superimposed with the interpreted network of structures. However, we acknowledge that this conclusion is subject to the limitations of the operational data made available and used.

Amongst all features, the dimension (trace length) and orientation of the structures were seen to impact the predictions of the machine learning models most. This is expected as the dimension of the structures (i.e., the potential shear rupture area, *A*) is linearly proportional to the seismic moment release, M_o, by means of $M_o = GA\overline{D}$ (where G is the shear modulus of the formation and $\overline{D}$ is the average shear displacement of the structure). The orientation of the fault strongly impacts the resultant shear and effective normal stresses acting on the fault and therefore its stability state relative to shear failure. The structural intensity was another important feature that consistently appeared in our machine learning results. This agrees with the findings reported in WSP 2023, which were explained mechanistically as the structural intensity influencing the storage capacity and therefore the pressure buildup and invaded zone. Lower fracture intensities were observed to be associated with larger pressure buildups that were more narrowly focused in their travel path to distal structures.

The Poisson's ratio and the formation depth were also commonly observed as being important features in the machine learning results. The higher stress environments at greater depths and increased elastic deformability of the rock (with higher Poisson's ratios) would both contribute to larger stored elastic strain energy and therefore larger magnitude events if triggered and released through shear slip.

Among the operational parameters, the injection volume was consistently seen to be an important feature. This is in agreement with our regional Montney induced seismicity susceptibility study (Amini et al., 2021). The influence of injection volume is well studied and has been shown to have a positive correlation with induced seismicity (e.g., McGarr, 2014; Schultz et al., 2018).

Using the results of the machine learning analyses, well pad-scale susceptibility maps were generated based on the probability outcomes of the Logistic Regression and Random Forest models for large induced seismicity events (i.e., $\geq M_W1$). The maps showed higher susceptibilities for an area of the reservoir footprint characterized as having low structural intensity. These findings are helpful in explaining the observation of frequent large induced seismicity in response to hydraulic fracturing operations in the Kiskatinaw area and can be useful in planning the mitigation induced seismicity hazard measures in NEBC.

Finally, we compared the approximate location of well-pad OP1 with the regional susceptibility map generated by Amini et al. (2021) for induced seismicity with magnitude 1.1 or higher. To facilitate direct comparison, focus was placed on the map versions derived using the Random Forest machine learning algorithm. The location of the well-pad is bounded by a large red circle in Figure 20 (large to maintain OP1's anonymity). This falls within an area on the regional map with a probability of 50% and higher of experiencing a Mw1.1 or larger induced seismicity event.

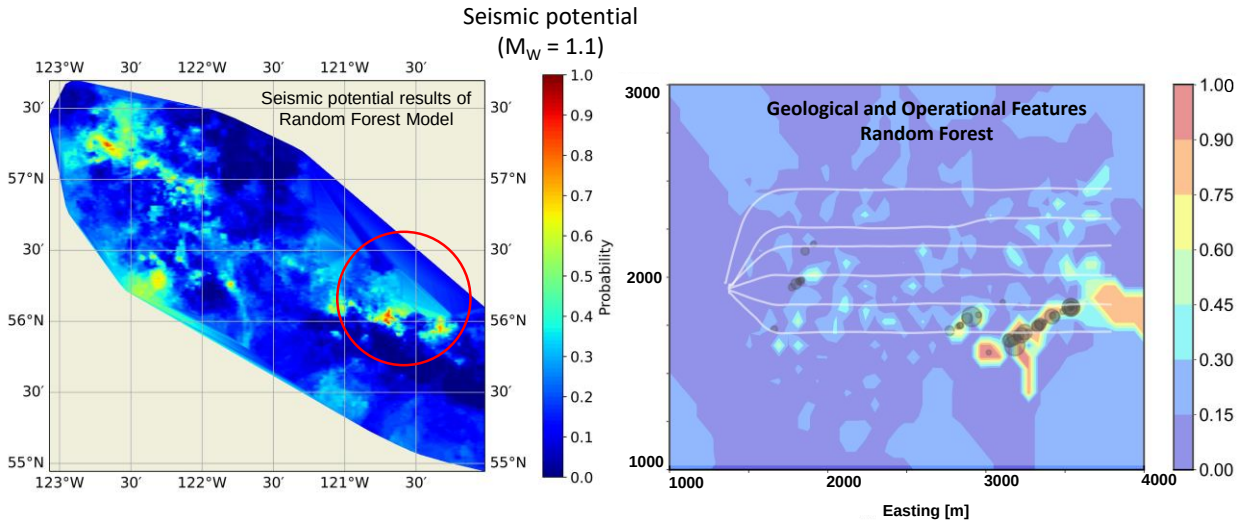


Figure 20. Induced seismicity susceptibility map developed for the Montney region by Amini et al. (2021) for events with a minimum magnitude of 1.1. This map was generated based on results from a Random Forest machine learning model. The red circle shows the approximate location of the OP1 well pad. To the right is the well pad-scale induced seismicity susceptibility map for OP1 for events with magnitude 1 or higher generated in this study.

7. Reference

Aki, K., Richards, P.G., 2002. Quantitative seismology, 2nd Ed. ed. University Science Books.

Allen, D., Eberhardt, E., Bustin, A., 2019. Scientific Review of Hydraulic Fracturing in British Columbia.

Amini, A., Mehrabifard, A., Eberhardt, E., 2021. Development of an Induced Seismicity Susceptibility Framework and Map for NEBC using an Integrated Machine Learning and Mechanistic Validation Approach by. https://doi.org/Geoscience BC Project 2019-014

Bergen, K.J., Johnson, P.A., De Hoop, M. V., Beroza, G.C., 2019. Machine learning for data-driven discovery in solid Earth geoscience. Science (80-.). 363. https://doi.org/10.1126/SCIENCE.AAU0323/ASSET/E01F2E56-540A-4836-991B-3D3E7652D547/ASSETS/GRAPHIC/363_AAU0323_F5.JPEG

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., 1984. Classification and regression trees, 1st ed, Classification and Regression Trees. CRC Press. https://doi.org/10.1201/9781315139470/CLASSIFICATION-REGRESSION-TREES-LEO-BREIMAN

Chen, T., Guestrin, C., 2016. XGBoost: A Scalable Tree Boosting System, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16. Association for Computing Machinery, New York, NY, USA, pp. 785–794. https://doi.org/10.1145/2939672.2939785

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297. https://doi.org/10.1007/BF00994018

Fox, A., McKean, S., Watson, N., 2020. Statistical Assessment of Operational Risks for Induced Seismicity from Hydraulic Fracturing in the BC Montney. Geoscience BC Project 2019-008, 12(October), 1–2.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep Learning. Cambridge, Massachusetts : The MIT Press.

Ke, G., Meng, Q., Finley, T., Wang, T., 2017. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Adv. Neural Inf. Process. Syst. 30 (NIPS 2017).

Le, Q. V., 2013. Building high-level features using large scale unsupervised learning. ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. 8595–8598. https://doi.org/10.1109/ICASSP.2013.6639343

Lundberg, S., Lee, S., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 4765–4774.

Mavko, G., Mukerji, T., Dvorkin, J., 2009. The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media, 2nd ed, The Rock Physics Handbook. Cambridge University Press. https://doi.org/10.1017/CBO9780511626753

McGarr, A., 2014. Maximum magnitude earthquakes induced by fluid injection. J. Geophys. Res. Solid Earth 119, 1008–1019. https://doi.org/10.1002/2013JB010597

Mehrabifard, A., Eberhardt, E., 2023. Reliability of earthquake-size distribution and stress regime relationships for fluid-injection-induced seismicity. Geoenergy Sci. Eng. 226, 211776. https://doi.org/10.1016/J.GEOEN.2023.211776

Mehrabifard, A., Eberhardt, E., 2021. Investigation of the Dependence of Induced Seismicity Magnitudes on Differential Stress and Pore Pressure Using Supervised Machine Learning, Northeastern British Columbia (NTS 093, 094A, B, G, H) and Globally.

Miller, T., 2017. Explanation in Artificial Intelligence: Insights from the Social Sciences. Artif. Intell. 267, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Pawley, S., Schultz, R., Playter, T., Corlett, H., Shipman, T., Lyster, S., Hauck, T., 2018. The Geological Susceptibility of Induced Earthquakes in the Duvernay Play. Geophys. Res. Lett. 45, 1786–1793. https://doi.org/10.1002/2017GL076100

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, ´Edouard, 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830. https://doi.org/10.1145/2786984.2786995

Schultz, R., Atkinson, G., Eaton, D.W., Gu, Y.J., Kao, H., 2018. Hydraulic fracturing volume is associated with induced earthquake productivity in the duvernay play. Science (80-.). https://doi.org/10.1126/science.aao0159

Walker, S.H., Duncan, D.B., 1967. Estimation of the Probability of an Event as a Function of Several Independent Variables. Biometrika 54, 167–179.

Wozniakowska, P., Eaton, D.W., 2020. Machine Learning-Based Analysis of Geological Susceptibility to Induced Seismicity in the Montney Formation, Canada. Geophys. Res. Lett. 47, e2020GL089651. https://doi.org/10.1029/2020GL089651

WSP. 2023. Modelling and analysis of KSMMA induced siesmic events to advance event mitigation strategies. Report submitted to BC OGRIS.

WSP-Golder. 2022. Review of Sub-Surface Data from Three Operators in the KSMMA Region Experiencing Elevated Induced Seismic Events. Report submitted to the BC Oil & Gas Commission

Zoback, M.D., 2007. Reservoir Geomechanics, Reservoir Geomechanics. Cambridge University Press. https://doi.org/10.1017/CBO9780511586477