
Geoscience BC Report 2021-11

Development of an Induced Seismicity Susceptibility Framework and Map for NEBC using an Integrated Machine Learning and Mechanistic Validation Approach

(Geoscience BC Project 2019-014)

by

Afshin Amini, Ali Mehrabifard and Erik Eberhardt

University of British Columbia

Prepared for Geoscience BC

September 1, 2021

Executive Summary

This report presents the results of a study to develop a machine learning framework for generating induced seismicity susceptibility maps, with focus placed on the Montney region of northeastern British Columbia. Data was compiled from several publicly available and proprietary data sources, and was comprised of earthquake data, geological data, and hydraulic fracturing operational and survey data. Eight different machine learning algorithms were tested across four sets of machine learning analyses.

The machine learning analyses conducted, with reference to the Montney region, include:

- 1) a classification feature importance analysis, based on a combined dataset of all geological and operational features at each hydraulic fracturing well;
- 2) a classification susceptibility analysis, based on a restricted dataset limited to geological features and those operational features that apply to susceptibility;
- 3) a classification susceptibility mapping analysis, based on the restricted dataset of geological and susceptibility-relevant features interpolated to a 2.5 km grid spacing; and
- 4) a regression severity analysis, based on a combined dataset of geological and operational features that included the maximum magnitude of the induced seismicity events associated with each well.

The classification analyses focussed on identifying the most important features and their correlation with respect to the likelihood of a well being seismogenic or not; the regression analyses focussed on identifying the most important features and their correlation with respect to the maximum magnitude of the induced seismicity associated with each seismogenic well (i.e., severity). The machine learning workflow included training and iteratively validating and fine tuning each model, and then testing it against unseen data. The top performing models were then selected for interpretation. The interpretation of the machine learning results was further supported by a series of advanced numerical models using the 3-D hydraulic fracturing simulator XSite™. These were used to provide mechanistically-based cause and effect insights into key machine learning correlations that were unclear or counterintuitive.

The results for the classification and regression analyses found the Light GBM and Random Forest machine learning models to be amongst the top performing models. The results for the feature importance analysis found that, in general, geological features ranked higher in importance than operational features. This lends itself to susceptibility mapping as most operational features do not apply to spatial susceptibility. 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. The numerical modelling results support this finding, showing that the injection depth influences the extent to which the

fluid pressure invaded zone is transmitted outwards from the injection point and hydraulic fracture. 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. This in turn results in an increased likelihood of elevated pore pressures interacting with a critically stressed fault in the basement to trigger an induced seismicity event.

The machine learning results were extended using the Light GBM and Random Forest models to generate induced seismicity susceptibility maps based on a 2.5 km grid resolution for the Montney region. Two sets of maps were generated for each machine learning model, one relative to a minimum moment magnitude (M_w) of 1.1 and one for a minimum M_w of 2.0. Overall, the locations of susceptible areas in both models agree with the historical location of induced seismicity events. It is cautioned in using these maps that machine learning results are subject to data errors, biases, censoring and overfitting, and therefore are not a perfect predictor. Use must also consider the site specific geology, especially where these deviate from general regional trends; machine learning analyses only report data relationships and correlations, but do not connect these to cause and effect mechanisms. Susceptibility is also distinctly different from hazard, with the latter being influenced by operational factors and considering key questions like maximum magnitude potential.

A preliminary basis for developing induced seismicity hazard maps is presented in the form of a regression analyses of maximum event magnitudes associated with seismogenic wells. The top predictors for event severity were identified by the machine learning models 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 hydraulic fracturing operation encountering a critically stressed fault.

The numerical modelling proved to be especially useful for explaining other correlations that were less clear or counterintuitive. The b-value was found to have a positive correlation between higher values and higher event magnitudes, whereas more typically, lower b-values suggest a higher likelihood of a large magnitude event. The numerical simulation results showed that faults in stiffer rock masses are more susceptible to larger magnitude events, and similarly, those within a strike-slip far-field stress regime are more likely to produce large magnitude events ($M_w \geq 4$). The latter was seen to relate to how the far-field stresses transform onto the fault surface (relative to the fault orientation) and how much of the resulting fault surface area is critically stressed. The simulations show that, for the same injection depth, adversely oriented faults under a strike-slip far-field stress regime can be more critically

stressed over a larger area and therefore are more sensitive to injected fluid pressures causing slip over a large area beyond the area of the fluid pressure perturbation. In contrast, a fault slip event might be triggered locally where the fluid injection pressures interact with the fault, but if less of the fault is in a critical state (either due to the fault orientation or the far-field stress state condition), the slip area and therefore the event magnitude might have a smaller upper limit. This was seen to be the case for faults modelled assuming a reverse-fault far-field stress regime for the same injection depth. These results were specific to the far-field stresses and not the focal mechanism. The modelling results showed that in the case of fluid-injection induced seismicity it is possible to generate a large magnitude event with a reverse fault focal mechanism (locally) but in a strike-slip far-field stress regime (regionally). This depends on the orientation of the fault experiencing the pore pressure increase, relative to the orientation of the far-field stress regime.

Together, the machine learning and numerical modelling results suggest that if operating within a strike-slip far-field stress regime in stiffer rocks and there is a lack of early, significant seismicity (of $M_w > 1$ or 2) with increasing injection time, then this might be a sign that strain energy is accumulating towards a potentially large event ($M_w \geq 4$). Similarly, in the context of a Traffic Light Protocol, the detection of a $M_w 2$ event might be more concerning when the formation targeted by the injection, or adjacent to it, is stiffer and if operating within a strike-slip far-field stress regime.

Table of Contents

Executive Summary.....	2
Table of Contents.....	5
1. Introduction.....	7
1.1. Previous Work	7
1.2. Study Area	9
1.3. Supplementary Materials.....	10
2. Data Compilation and Preparation.....	11
2.1. Seismic Event Catalogue	11
2.2. Geological Database.....	11
2.3. Well and Completions Database	12
2.4. Data Preparation	12
2.5. Data and Analysis Uncertainties	16
3. Methodology	18
3.1. Machine Learning Overview	18
3.2. Model Generation.....	21
3.2.1. Workflow.....	21
3.2.2. Machine Learning Models Tested.....	23
4. Exploratory Data Analysis (EDA).....	25
5. Susceptibility Analysis.....	29
5.1. Feature Importance (Geological and Operational Data)	29
5.1.1. Model Training, Tuning and Comparison Testing.....	29
5.1.2. Model Interpretation	30
5.2. Induced Seismicity Susceptibility (Geological Data)	40
5.2.1. Model Training, Tuning and Comparison Testing.....	40
5.2.2. Model Interpretation	43
6. Induced Seismicity Susceptibility Maps.....	51
6.1. Cut-Off Magnitude 1.1	51
6.2. Cut-Off Magnitude 2.0	55
6.3. Use and Limitations.....	62
7. Severity Analysis	63
7.1. Model Training, Tuning and Comparison Testing	63
7.2. Model Interpretation	64
8. Mechanistic Analysis	69
8.1. Numerical Method Overview	69
8.2. Model Setup and Inputs	71
8.3. Procedure for Calculating Induced Seismicity Moment Magnitude.....	77
8.4. Simulation Results.....	78
8.4.1. Importance of Depth to Top of Basement.....	78

8.4.2. Influence of Formation Stiffness.....	80
8.4.3. Influence of Fault Orientation	89
9. Conclusions, Recommendations and Future Work.....	98
9.1. Conclusions.....	98
9.2. Recommendations/Guidelines.....	101
9.3. Future Works.....	103
Acknowledgments.....	104
References	105
Appendix A - Feature Definitions.....	112
Appendix B - Databases and Codes	116

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). Cited was a similar map developed for the Duvernay Play in Alberta using machine learning (Pawley et al., 2018). However, the map produced by Pawley et al. was limited to the data relationships determined by computer algorithms without connecting these to cause and effect mechanisms. Refining the machine learning output through mechanistic validation would add considerable value in terms of improved reliability for decision makers.

This report presents the results of research undertaken to develop an induced seismicity susceptibility map for NEBC, with focus placed on the Montney region. Of interest was the development of an initial susceptibility map, but also a framework and workflow to guide practitioners in using machine learning to prepare induced seismicity susceptibility maps for any region, or to update these as more data becomes available over time. Three key objectives were targeted:

1. To compile data for NEBC, with focus on the Montney region, that integrates induced seismicity data with geological data and well activity data (related to hydraulic fracturing).
2. To investigate and test different machine learning algorithms for their robustness in identifying the most influential factors controlling induced seismicity for the Montney region, and use this to develop an induced seismicity susceptibility map together with a framework and guidelines for developing susceptibility maps using machine learning.
3. To validate and refine the machine learning results using 3-D numerical simulations to investigate and improve understanding of the cause and effect relationships influencing induced seismicity behaviour.

Reported are the results for each of these objectives, including the methodologies developed and applied, and the corresponding results and key findings. Delivered are outcomes and products in the form of a robust susceptibility map that can be used to help decision makers with their planning of hydraulic fracturing activities and induced seismicity hazard management planning. The results presented also help to identify areas requiring additional focused research.

1.1. Previous Work

The rapid development of multivariate statistics and machine learning techniques to analyze large data sets is especially attractive and conducive to the induced seismicity hazard problem (e.g., likelihood, severity, etc.). However, experience in doing so is limited, especially in considering the influence of both geological and operational factors together. Distinguishing between these is 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 hydraulic fracturing and completions), which can be manipulated offering a means to potentially mitigate induced seismicity hazards for a susceptible formation.

Several studies have recognized that a significant amount of geological, operational, and seismic data is being collected from hydraulic fracturing activities, and the size and complexity of these datasets have made traditional empirical and statistical analyses inefficient and ineffective. A study by Pawley et al. (2018) combined tectonic, geomechanical and hydrological data with induced seismicity data related to hydraulic fracturing operations in the Duvernay play in Alberta to train a logistic regression algorithm to develop an induced seismicity potential map. Their results suggest that the proximity to basement, formation overpressure, minimum horizontal stress, proximity to reef margins, lithium concentrations, and natural seismicity rate are the dominant controlling factors to triggering induced seismicity within the study area. Wozniakowska and Eaton (2020) used logistic regression to investigate the influence of geological features on the distribution of induced seismicity in the Montney Formation, and developed a susceptibility map of Seismic Activation Potential (SAP) in the region. Their study indicated that distance to the Cordilleran deformation front and injection depth are 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 Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) and the North Peace Ground Motion Monitoring Area (NPGMMA) of the Montney region. Their analyses examined both likelihood and severity of induced seismicity events. The results showed 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, while the more complex models placed a relatively high importance on Paleozoic structure and distance to faults. In the magnitude regression analysis, most models have shown a relatively high importance for the top of the Montney structure and distance to faults. Amini and Eberhardt (2021) investigated the application of different machine-learning algorithms to determine the relative importance of an expanded set of geological and operational parameters in relation to the triggering of induced seismicity. They tested different algorithms using data related to hydraulic fracturing activities in the Montney region. Their analyses showed that geological features generally ranked higher than operational features with respect to correlation with wells associated with induced seismicity. In all models, pore-pressure gradient ranked highly as having a significant influence. Other important geological features included distance to basement, distance to known faults and the azimuth of the maximum horizontal stress. The operational feature that most consistently ranked as being important was the completion length.

This present study builds upon these previous studies, expanding upon the number of geological features compiled and assessed, while also improving on the analysis and interpretation methodologies developed.

1.2. Study Area

The Lower Triassic Montney Formation is aerially extensive, covering approximately 130,000 km² from central Alberta to NEBC (National Energy Board, 2013). It is also thick, typically ranging from 100 to 300 m, but thinning to zero at its eastern and northeastern margins while increasing to over 300 m on its western margin before it begins outcropping in the Rocky Mountains. The Montney Formation unconformably overlies Carboniferous or Permian strata and consists of variable amounts of interbedded shale, siltstone, and sandstone. These strata developed during the first of three major transgression-regression cycles (Edwards et al., 1994; Gibson & Barclay, 1989). Based on the lithostratigraphy, Davies et al. (2018) subdivides the Montney Formation into three members: the Upper, Middle, and Lower Montney. These members are separated by a basin-wide unconformity that developed due to tectonic uplift of the basin margin (Dixon, 2009).

The area of research for this project encompasses the Montney region of NEBC (Figure 1).

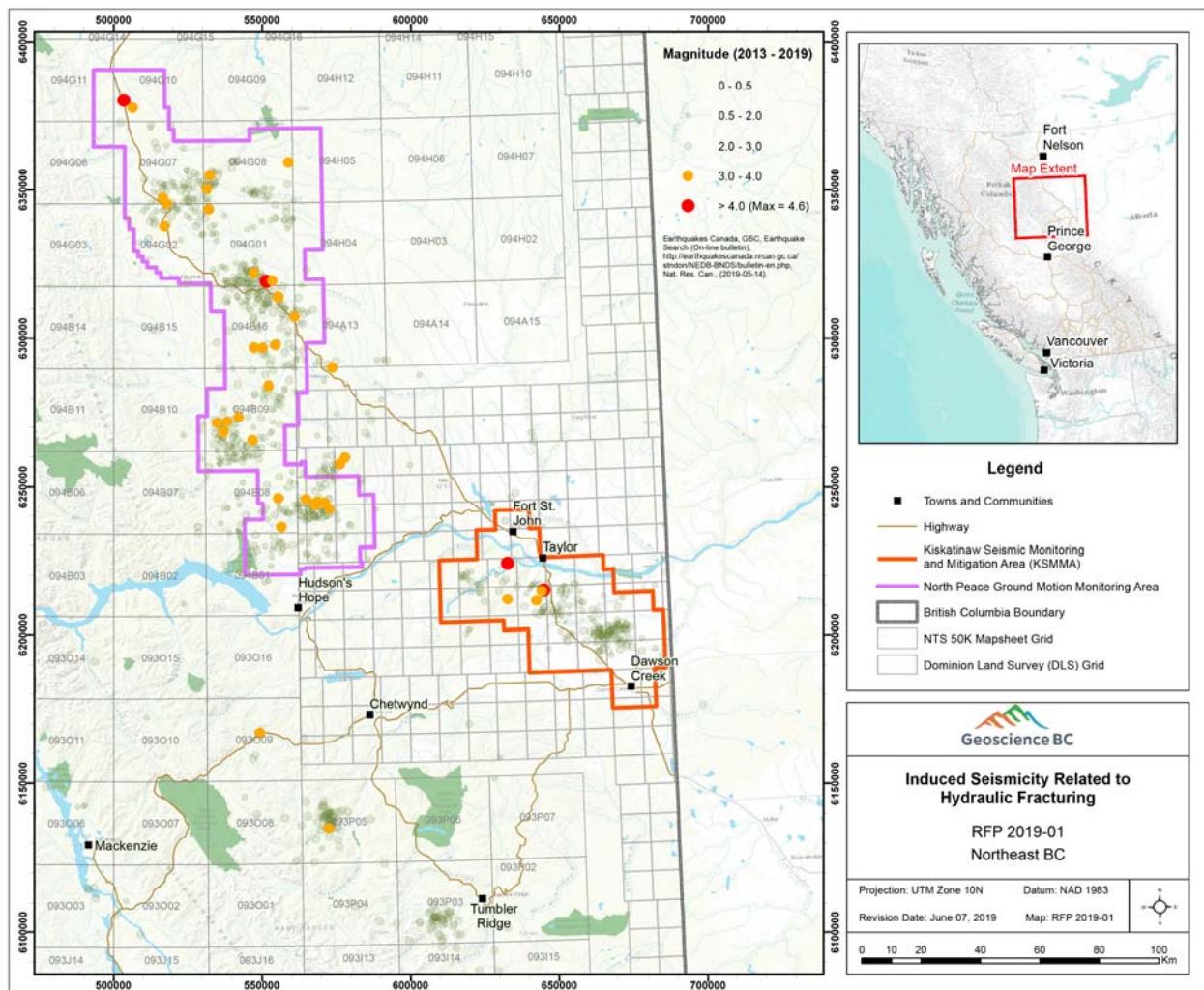


Figure 1 – Map of the study area. Modified after Geoscience BC RFP 2019-01.

1.3. Supplementary Materials

The following appendices are included with this report:

Appendix A: Feature Definitions – Detailed descriptions of each feature included in one or more of the machine learning analyses.

Appendix B: Database and Codes – Included as separate folders and files to this report.

2. Data Compilation and Preparation

Several datasets were compiled and integrated for this study using both public domain and proprietary sources. These include seismic event catalogues, geological databases, and well and hydraulic fracturing data. Each is discussed separately in the following subsections. Identified from these were a number of geological and operational “features” that form the basis for the machine learning analysis. Definitions of these features and details regarding their data quality and completeness are provided in Appendix A.

2.1. Seismic Event Catalogue

The seismic event catalogue we compiled made use of several existing public-domain earthquake catalogues. It should be noted that these catalogues include all recorded seismic events without specifying whether they are natural earthquakes or induced seismicity events. The necessary filtering for this is discussed in Section 2.4. The primary source was the catalogue compiled by Enlighten Geosciences (Fox et al., 2020). This combines:

- i) the Composite Alberta Seismicity Catalog by Fereidoni and Cui (2015);
- ii) the Geological Survey of Canada (GSC) catalogue by Visser et al. (2017); and
- iii) the BC Oil and Gas Commission (BCOGC) and GSC catalogue by (Babaie Mahani et al., 2020).

Duplicate events were removed by Enlighten Geosciences resulting in a catalogue of 9843 seismic events recorded between January 1, 2000 and April 16, 2019. To this we added an update to the GSC catalogue that included another 109 events recorded in the Montney region between April 16, 2019 and February 17, 2020. This resulted in a final catalogue of 9949 events used in this study, with moment magnitudes (M_w) ranging from 0.3 to 4.6 and a magnitude of completeness (M_c) of 1.1 (see Fox et al., 2020).

The seismic event catalog provides two targets for our analysis. The first is the seismogenic target, which indicates whether a well is associated with seismicity or not. The second is the maximum magnitude of the events associated with each well. For the regression analysis, wells without associated seismicity are excluded.

2.2. Geological Database

A high level of focus was placed on compiling data of different geological indicators (features), as these represent the conditions in place that influence the susceptibility and seismogenic potential of a given well location. For this analysis, 27 geological features were extracted from the various sources consulted. These include:

- i) a regional fault model by Petrel Robertson Consulting Ltd.;
- ii) an isopach map of the top of the Precambrian basement by Enlighten Geosciences Ltd.;

- iii) geophysical logs accessed through geoLOGIC systems Ltd.'s geoSCOUT database, which was made available via a research license agreement between geoLOGIC and UBC; and
- iv) the World Stress Map database (Heidbach et al., 2018).

Some geological features are used as proxies where more direct measurements are unavailable. For example, gamma ray readings from geophysical logs are used as a proxy for clay concentration. Recent examples of using features as proxies include Wozniakowska and Eaton (2020) where they used distance to the Cordilleran deformation front as a proxy for strain rate, and Pawley et al. (2018) where they used dolomite occurrences as a proxy for deep subsurface faults. The compiled geological datasets are identified and described in detail in Appendix A.

2.3. Well and Completions Database

Hydraulic fracturing data for the Montney region was sourced from geoLOGIC systems Ltd.'s Well Completions & Frac Database (geoLOGIC) through their gDCweb software. This was made available through a license agreement between geoLOGIC systems Ltd. and Geoscience BC. The data was compiled according to two groupings: a well dataset used to access hydraulic fracturing information such as well completion length, fluid injection volume and number of stages; and a stage dataset used to calculate parameters such as average injection rate, maximum instantaneous shut-in pressure (ISIP) and maximum breakdown pressure for each well. In total, data from 6,530 horizontal wells incorporating 95,807 hydraulic fracturing stages was used. The data was structured according to wells and not well pads (although both were identified) to maximize the number of data points for the machine learning analyses. A complete listing of the well and completion parameters used is provided in Appendix A.

2.4. Data Preparation

The first step in data preparation required the events in the induced seismicity catalogue to be spatially and temporally correlated with the database of well locations and hydraulic fracturing activities, respectively. For this, the spatio-temporal filter developed by Amini and Eberhardt (2019) was used. A spatial filter was applied to search for all event locations that were within a 5 km radius of an active hydraulic fracturing well. The 5 km radius represents the uncertainty in the event location accuracy reported for the earthquake catalogue. Next, a one-month temporal filter was applied relative to each well's activity. Thus, if a seismic event was recorded as occurring within 5 km of an active well and within 30 days from the start date of hydraulic fracturing operations, it was considered to be an induced seismicity event. The corresponding well was then labelled as being seismogenic. It should be noted that similar filtering techniques have been used in other studies and have proven to be effective. However, it is also important to recognize the limitations associated with their use. First is the fixed spatial filtering value. In this case, the filter dismisses any induced events that are farther than 5 km away from the well in question. Next is the duration of the temporal window. Verdon and Bommer (2020) investigated the temporal evolution of hydraulic fracturing-induced seismicity events from case

studies around the world. Their study showed the largest time-lag observed between the end of hydraulic fracturing and an increase in seismic magnitudes was 23 days. Note that the 30-day temporal filter used in this case is with respect to the start of hydraulic fracturing. Finally, multiple wells can be associated with a single induced seismic event if they each meet the spatial and temporal limits. Likewise, natural earthquakes can be mistakenly identified as induced seismicity events if they meet the filtering criteria. Both are limitations that arises from the uncertainty in calculating seismic event locations and can be improved over time as more seismic stations are installed and the accuracy of event locations improve.

Use of the 5 km radius, 30 day spatio-temporal filter resulted in a subset of 937 events being identified as induced seismicity events from the 9949 events contained in the seismic event catalogue. From this, each well was classified as either having experienced induced seismicity or not, assuming a minimum magnitude cut-off of 1.1 (Figure 2).

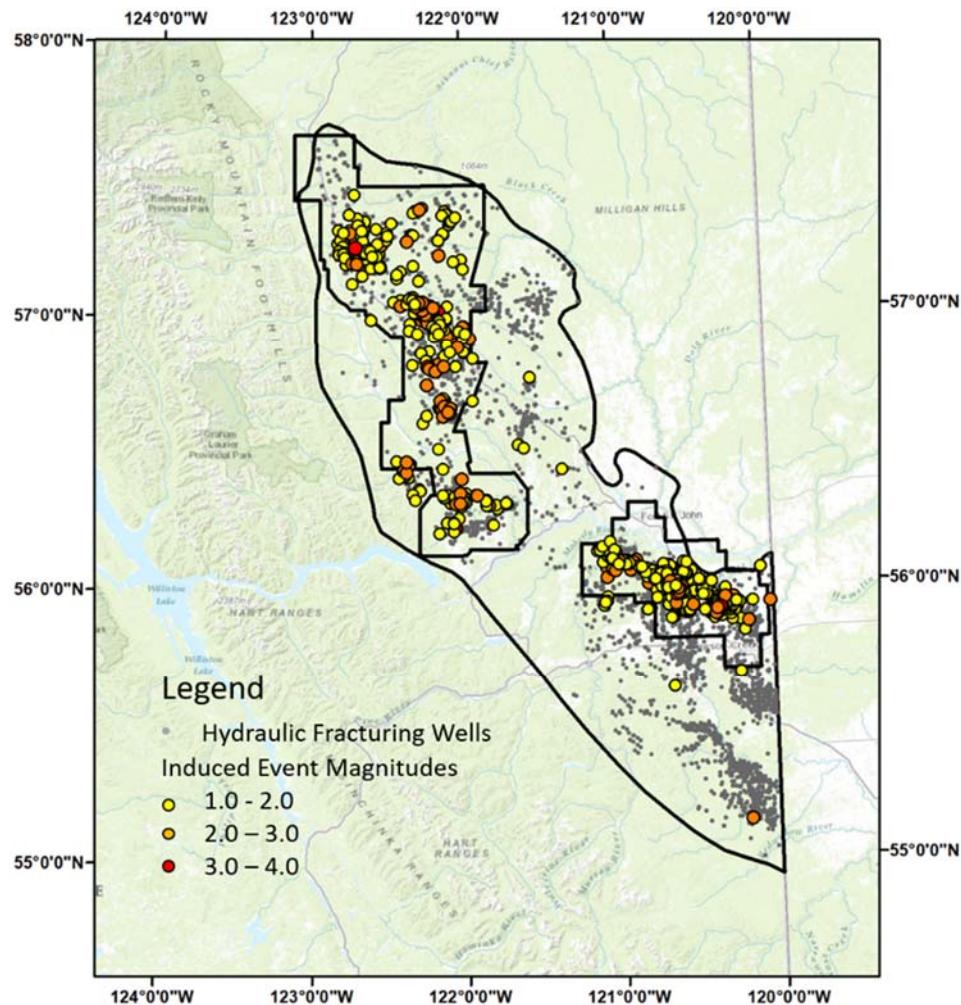


Figure 2 – Location map of induced seismicity events and their magnitudes determined using a spatio-temporal filter relative to well locations and activities.

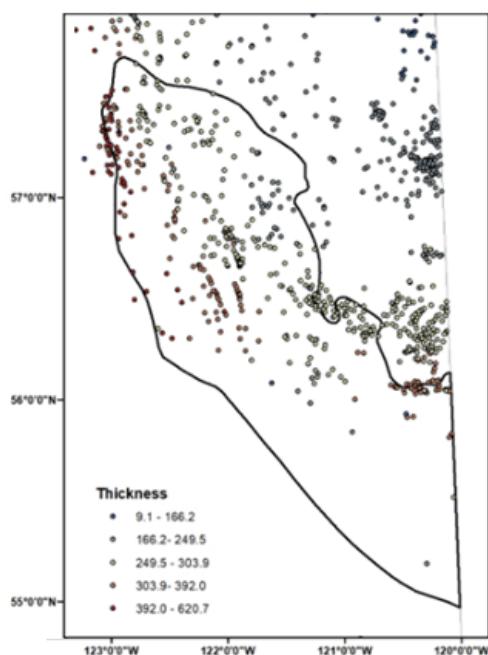
The geological data was next prepared by interpolating the available data for each feature to obtain a distribution of values for the entire study area (the Montney region). This was done using the Inverse Distance Weighted (IDW) interpolation tool in ArcGIS, which uses the 12 closest data points to interpolate a value for each designated point. The designated points were defined according to two different needs (Figure 3). The first was to assign values for each geological feature to each well in the database. This dataset was used to train and test the performance of the machine learning algorithms with respect to identifying which factors have the most influence on the occurrence of induced seismicity in the Montney region. The second need was to assign values for each geological feature to a 2.5 km grid spacing for the entire Montney region. This dataset was used as the input to the trained machine learning algorithms to estimate the seismogenic activation potential across the Montney region, as required to develop an induced seismicity susceptibility map.

The final data preparation step was to integrate the geological data, well completion data, deviation surveys, and the combined seismic event catalogue. These were loaded and processed using the open source programming language Python, and entailed standardizing the Unique Well Identifiers (UWIs) and geospatial projections. The stage-by-stage completion data for each horizontal well was aggregated, since the analysis was conducted on an individual well basis. Wells were analyzed individually, meaning that wells on the same pad were not aggregated but treated separately. Geospatial layers and deviation surveys were projected to NAD83 UTM Zone 10N prior to distance calculations. The wells were then associated with the combined seismic event catalogue and the data combined into a single data set.

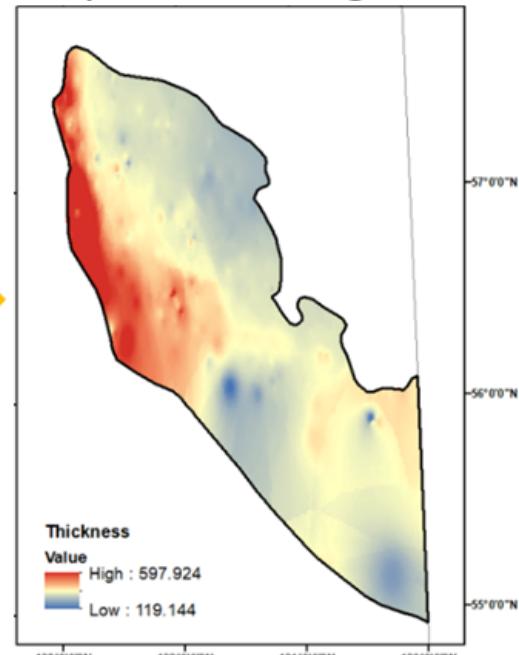
Based on these data preparation steps, four separate datasets were generated from the compiled database to be used as the input data for the different machine learning analyses:

- i) Well Features Dataset: A dataset containing hydraulic fracturing well IDs, latitudes and longitudes, the value of each geological and operational feature at each well location, and whether the well was associated with seismic activity or not. This dataset was used to train, test and validate the machine learning classification models used for a feature importance analysis (Section 5.1).
- ii) Well Susceptibility Dataset: A limited version of the Well Features Dataset that focusses on the geological features and those operational features that are relevant to spatial susceptibility mapping. Note that this excludes most well completions 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). Another exception is the well density (within 1, 3, 5, and 10 km of the observed well), which accounts for any interference or preconditioning related to previous activities that may influence the susceptibility of a new well in a given location. This dataset was used to train, test and validate the machine learning models used for the susceptibility analyses (Section 5.2).

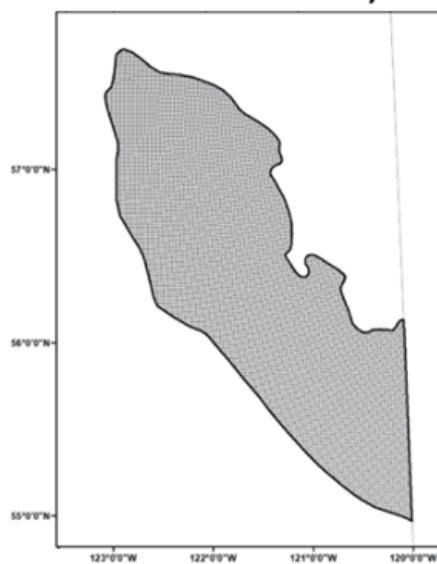
Input Geological Data



Interpolated Geological Data



Extract at Montney Grid



Extract at HF wells

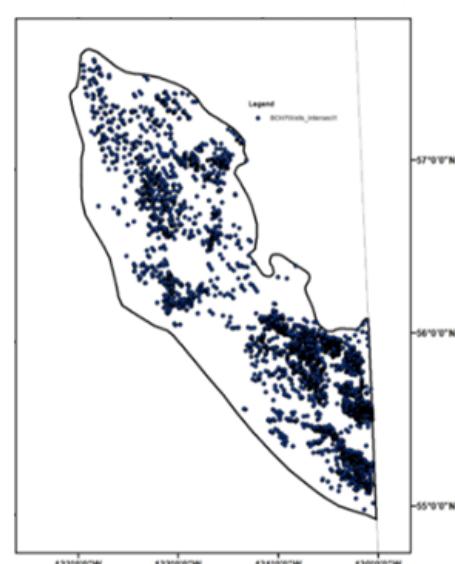


Figure 3 – Procedure used to interpolate values for the geological features and assign to each well to train and test the machine learning algorithms used, and to assign to a 2.5 km grid spacing for development of an induced seismicity susceptibility map. The example shown is for thickness of the Montney formation.

- iii) Montney Grid Dataset: A dataset containing latitudes and longitudes for a 2.5 km grid spacing across the entire Montney region and the interpolated values of the Well Susceptibility dataset at these locations. This dataset was used to generate induced seismicity susceptibility maps using the best performing machine learning models (Section 6).
- iv) Induced Seismicity Severity Dataset: This dataset is similar to the Well Features Dataset but with two differences: it only includes the wells that are seismogenic, and it includes the maximum magnitude of the induced seismicity events associated with each well. This dataset was used for the regression analysis to study the relationship between the feature values and the severity of induced seismicity events (Section 7).

2.5. Data and Analysis Uncertainties

Considerable effort was taken to reduce errors and omissions related to the data, and subsequently, the machine learning and numerical modelling performed. After compiling the data, in some cases from several sources, it was manually inspected to complete the record for each feature (via imputing, interpolation, etc.). Effort was also taken to manage uncertainty, recognizing that uncertainty can take different forms (e.g., aleatoric, epistemic, etc.). These are briefly described below:

Data Uncertainty: This includes the amount of data available for each feature analyzed, the degree of variability in the observations (e.g., noise), and any errors in measuring, recording or transcribing the data. Data coverage may be incomplete, irregular or require assumptions in its calculation. For example, the b-value for each well is calculated from the induced seismicity events that occur within a certain distance from the well, which is arbitrarily selected. In this study, the distance used was 100 km, which was arrived at after some sensitivity testing. In addition, because the distribution of induced seismicity events across the Montney varies, the number of events available to base the calculation on for each well varies significantly when comparing wells in different parts of the Montney (see geospatial uncertainty below). Another relevant example is the seismogenic index, which includes the injection volume in its calculation relative to the seismogenic response recorded. However, the injection volumes are taken from the well and completions database, which is specific to each well, whereas the seismogenic response is associated to a well based on a spatial-temporal filter (as described in Section 2.4). The calculation of the seismogenic index therefore combines factual data with interpolated data.

Geospatial Uncertainty: This recognizes that the geospatial distribution of data points across the Montney is not uniform. There are areas with high concentrations of data focussed around areas of past and present unconventional gas development, and areas with little or no data in areas not yet developed. This distribution of data influences the quality of any required interpolation. Another form of geospatial uncertainty arises from the focus placed on using data from unconventional gas wells (i.e., those with hydraulic fracturing

completions); data from wastewater disposal wells or older conventional wells were not used. This means that any local stress changes or other influences resulting from wastewater injection activities or older conventional gas production are not recognized and accounted for in the interpretation of the machine learning results. Furthermore, the omission of wastewater injection wells from consideration also means that any induced seismicity events triggered from these activities that meet the spatial-temporal filter criteria, relative to a nearby hydraulic fracturing operation, would be associated with the hydraulic fracturing well.

Model Uncertainty: All statistical models are an estimation of reality and therefore are imperfect. More specifically, machine learning models are subject to uncertainty resulting from the procedures used to prepare the model, the choice and preparation of data, the choice of hyperparameters used in training, and the interpretation of model output. As a result, the machine learning model may have high bias where the model is underfitted (i.e., the data is poorly mapped), or high variance where the model has been overfitted (i.e., over trained) and it has essentially memorized the given data relations without learning the true data correlations. To minimize this, several steps are considered such as testing and comparing multiple models and evaluating each model's performance.

3. Methodology

3.1. Machine Learning Overview

Machine learning algorithms can be broadly categorized as unsupervised or supervised (Figure 4). Unsupervised learning searches for previously unknown patterns within a data set containing no pre-existing labels or known outcomes. Supervised learning uses labelled data for which there is a known outcome and seeks to learn the patterns between the input features and known outcome (Goodfellow et al., 2016).

For this study, each hydraulic fracturing well in the database is labelled with respect to a known outcome; i.e., whether the well has been associated with one or more induced seismicity events or not, and if so, the event magnitude(s). Accordingly, supervised learning was used to fit the machine learning models to two analyses: likelihood classification (i.e. susceptibility) and severity regression. Associated with each of these is feature importance. Likelihood classification involves supervised learning using binary classification (Figure 4): a given well in the database has either been associated with a past induced seismicity event or not (either yes or no). Severity regression targets the maximum magnitude of the induced events associated with each well. Although the likelihood classification and severity regression analyses require different machine learning models, the initial data input used is the same.

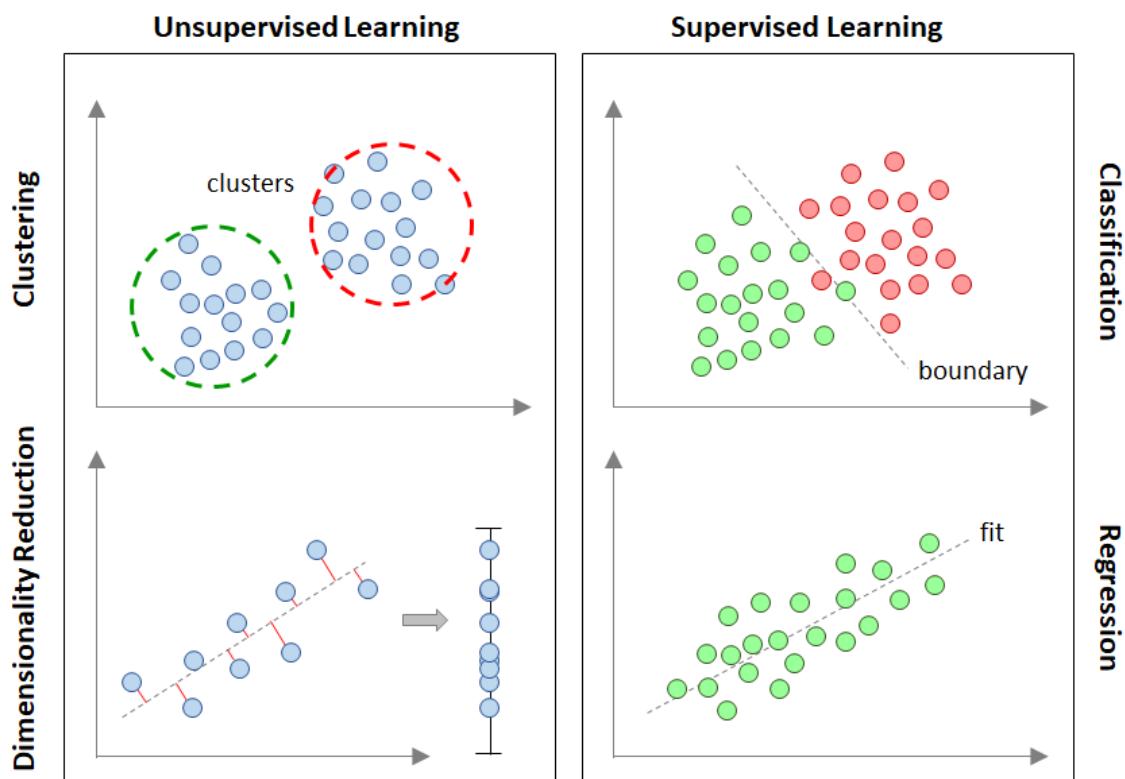


Figure 4 – Types of machine learning algorithms, classified as either unsupervised or supervised learning.

In building a machine learning model, an important initial step is to first divide the available data into a training split and test split (Figure 5). For this study, we used a standard 75%-25% split, respectively. The test split is a sub-sample of the available data (e.g., 25% of all available data) that is held back until the end of the model building and tuning process to provide an unbiased evaluation of the final model performance on unseen data. This allows the prediction error of the model to be calculated and for competing models to be compared to one another. The training split is used to fit the machine learning model; the model learns from this data.

Model training involves further subdividing the training split into a training set and validation set. The training set is used to train the model and the validation set is introduced as new data to evaluate the fit of the model. Because this is an intermediary evaluation of the model fit, separate from the final test analysis, the procedure can be carried out iteratively to improve and fine-tune the model performance. In this study, we use the k -fold cross-validation technique to iteratively tune the model. This procedure repeats the division of the training split into training and validation sets so that all data in the training split serves at least one time as the validation dataset, while being used to train the model during other iterations (Figure 5).

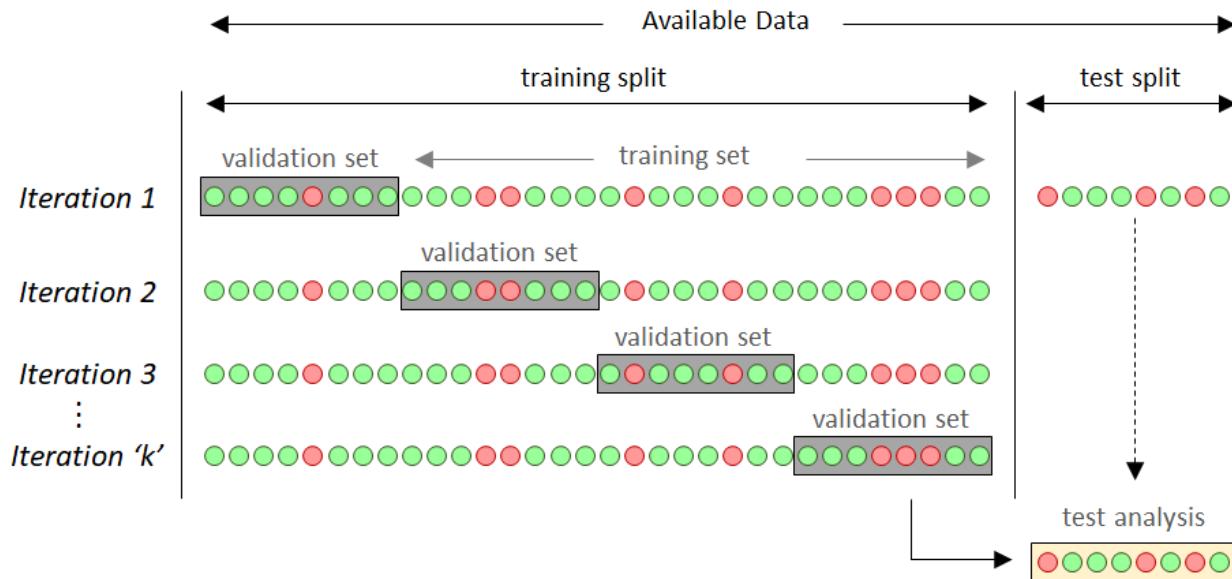


Figure 5 – Illustration of the machine learning analysis procedure involving first dividing all available data into a training and test split, and then sub-dividing the training split into a training and validation data set. Shown is the k -fold cross validation procedure, which iteratively repeats the subdivision of the training split to allow evaluation and fine-tuning of the machine learning model performance.

The cross-validation tuning of the model is also used to prevent overfitting errors. Overfitting is an error that occurs when a model result is too closely aligned to a limited set of data points (training set) and may therefore fail to fit additional data (validation set) or predict future observations reliably. Different machine learning algorithms employ a set of parameters known as hyperparameters whose values are used to control the learning process. These hyperparameters are used to tune the model to avoid overfitting. For this study, a Bayesian search algorithm was used. This method is considered to be more advanced than other algorithms such as grid search or random search. Models were trained using a 50-fold cross validation (i.e., 50 iterations; Figure 5).

Evaluation of a model's performance, whether against the validation data for intermediary tuning of the model or test data for final evaluation of the model, can be carried out using a number of metrics. 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 6. In this case, the binary problem is whether a hydraulic fracturing well has experienced induced seismicity (seismogenic; positive) or not (non-seismogenic; negative). The model predictions of seismogenic and non-seismogenic are then compared to the actual 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 6 – Illustration of a confusion matrix used to evaluate machine learning performance, and several advanced performance metrics calculated from the confusion matrix.

The confusion matrix can then be used to calculate more advanced classification metrics (Figure 6). Three of the more important of these are:

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; the model has a high success rate in correctly predicting seismogenic wells. Recall is often used in cases where classification of false negatives is a priority. Ideally, both are desired in a good machine learning classifier model. The F1-score is a metric which takes into account both precision and recall and is defined as follows:

$$F1\text{-score} = \frac{2 * Recall * Precision}{Recall + Precision}$$

For classification analyses, the F1-score is used to evaluate and compare the performance of competing machine learning models. This is done by measuring precision and recall at the same time and using a harmonic mean in place of the arithmetic mean to better account for extreme values. This makes two models with low precision and high recall or vice versa, comparable.

3.2. Model Generation

3.2.1. Workflow

We compared the performance of eight different machine learning algorithms, first with respect to the classification problem of induced seismicity occurrence (i.e., likelihood), and then the regression analysis of the maximum magnitude of the induced seismicity events associated with each well (i.e. severity). Figure 7 illustrates the workflow and the components of the analysis, together with the metric used to evaluate the model performance.

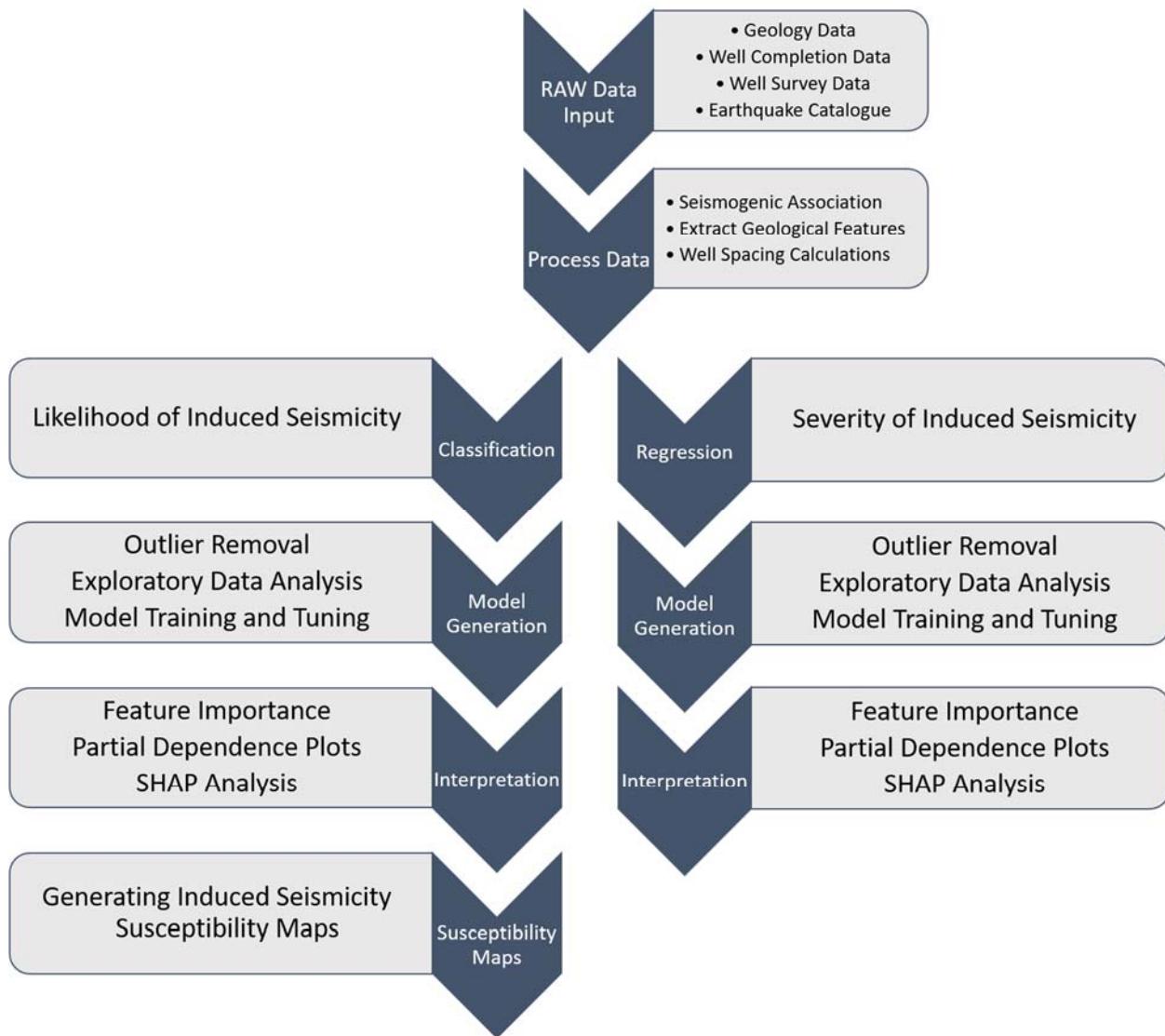


Figure 7 – Machine learning analysis workflow, examining both the classification problem of induced seismicity likelihood and the regression problem of induced seismicity severity, and the corresponding steps of model generation and interpretation.

For this analysis, the machine learning models were generated using SciKit-learn (Pedregosa et al., 2011), which is a machine learning library for the Python programming language. The workflow to generate each model is shown in Figure 8. The first step addresses data quality-assurance and quality-control (QA/QC). The data within each database was manually inspected to make sure the values were physically correct. For example, the depth to the top of the lower Montney was checked to make sure it was not above the depth of the top of the Montney. This step recognizes that transcription and recording errors are common in large datasets compiled

from different sources (e.g., wells involving different operators and service providers). Next outliers were removed after vetting for validity and manually checking whether the values were well outside accepted ranges (e.g., anomalously high injection rates relative to those cited by operators). This was followed by imputation to make sure there were no missing values for each feature for each well. The data was then scaled to have zero-mean and unit-variance. This step ensures that the scale of the data does not impact the performance of models that are sensitive to this (e.g., logistic regression).

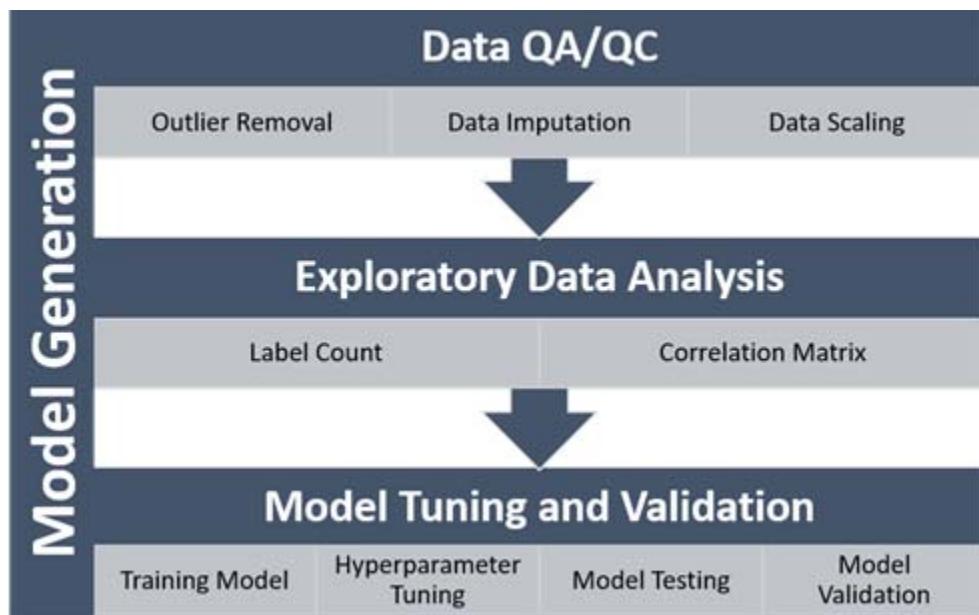


Figure 8 – Workflow showing key steps taken in generating each machine learning model, including data quality-assurance and –control (QA/QC), exploratory data analysis (EDA), and model tuning and validation.

The second step involves performing an Exploratory Data Analysis (EDA) to summarize the main characteristics of the data and determine which evaluation metric should be used when comparing different model's performances. The final step involves model tuning and validation as described above (e.g., Figures 5 and 6). These two steps are described with the reporting of the machine learning results.

3.2.2. Machine Learning Models Tested

The eight different models generated and tested for this study are listed and described in Table 1. Note that the dummy classifier model is only used to create a baseline for the other models.

Table 1 – 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 & Duncan, 1967)	LR	A linear model used for both classification and regression. In this model, the probabilities describing the possible outcomes of a single trial are modelled using a logistic function.
Decision Tree (Boehmke & Greenwell, 2019)	DT	A non-parametric supervised learning method used for classification and regression. The goal is to create a model that predicts the value of a target variable by learning simple decision rules inferred from the data features.
Random Forest (Breiman, 2001)	RF	A random forest is an estimator that fits a number of decision tree classifiers on various sub-samples of the dataset and uses averaging to improve the predictive accuracy and control overfitting.
Support Vector Classifier (Cortes & Vapnik, 1995)	SVC	Support vector machine classifiers are a set of supervised learning methods used for classification, regression and outlier detection.
K-Nearest Neighbour (Boehmke & Greenwell, 2019)	KNN	A non-parametric supervised/unsupervised learning method used for classification and regression.
XGBoost (Chen & Guestrin, 2016)	GXB	“Extreme Gradient Boosting”. It provides a parallel tree boosting to solve supervised learning problems.
Light GBM (Ke et al., 2017)	LGBM	“Light Gradient Boosting Machine”. It is a gradient boosting framework for machine learning based on decision tree algorithms and used for classification, regression and other machine learning tasks.

4. Exploratory Data Analysis (EDA)

An Exploratory Data Analysis (EDA) was performed to summarize the main characteristics of the data and determine which metric should be used when evaluating model performances. First, we look at label count (number of seismogenic vs. non-seismogenic wells) in the dataset. Table 2 shows that 20% of the wells are labelled as seismogenic and 80% as non-seismogenic.

Table 2 – Well label percentages.

Category	Percentage
Seismogenic well	20%
Non-seismogenic well	80%

This difference in population is an indication of class imbalance (i.e., the majority of the wells are non-seismogenic). Having an imbalanced dataset impacts a model in two ways. First, it means there are fewer samples of seismogenic wells for the machine learning models to train on. To overcome this, more weight was assigned to the seismogenic wells during training and hyperparameter tuning. Second, when performing classification, the ‘accuracy’ metric can no longer be used to evaluate the performance of the models.

A pair plot of several of the key features can be seen in Figure 9. The plot shows the pairwise relationship between dataset features, which are colour-coded by their labels (i.e., seismogenic or non-seismogenic). The diagonals include the univariate distribution plots that show the marginal distribution of the data in each column. This shows that the distributions of features are overlapping between seismogenic (orange) and non-seismogenic (blue) wells, which points to the complex intrinsic relationships between the data and label. The overlap also demonstrates why machine learning is required to analyze and determine trends in the data.

The correlation plot of seismogenic well classification is shown in Figure 10, and that for the maximum magnitude regression is shown in Figure 11. The 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 negative strong correlation. A weak (near zero) correlation lies in the middle in green. Based on this analysis, three groups are identified as having high positive correlations:

- The depths to the Top of the Montney, Top of the Middle Montney and Top of the Lower Montney.
- Distance from the injection point to the Top of the Montney and the Lower Formations (either Debolt or Belloy).
- Well density at 3, 5 and 10 km radius.

Another observation from the correlation plots is 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 machine learning analyses.

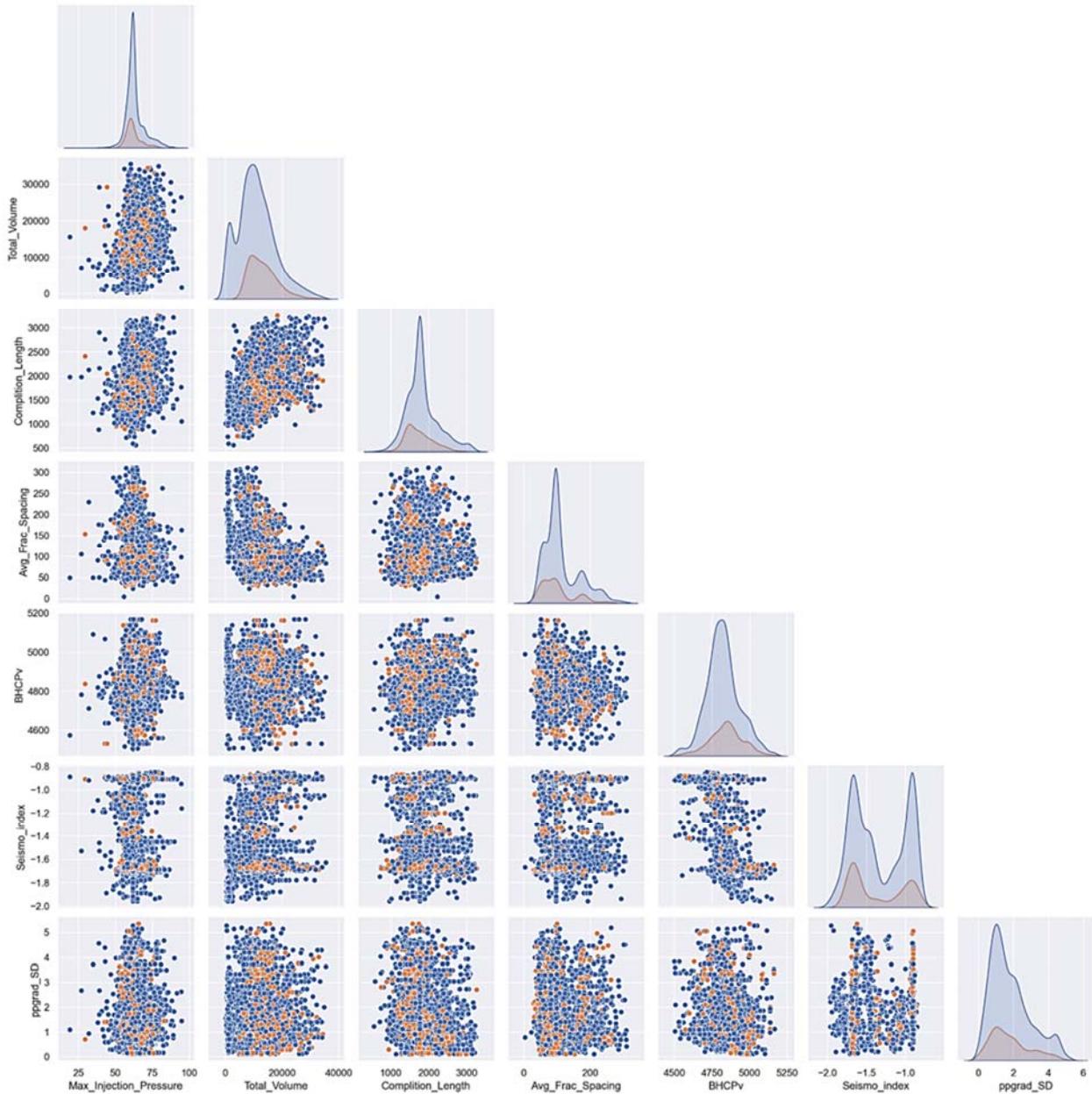


Figure 9 – Pair plot of sub selection of features. The seismogenic well cases and distributions are coloured orange, and the non-seismogenic well cases and distributions are coloured blue.

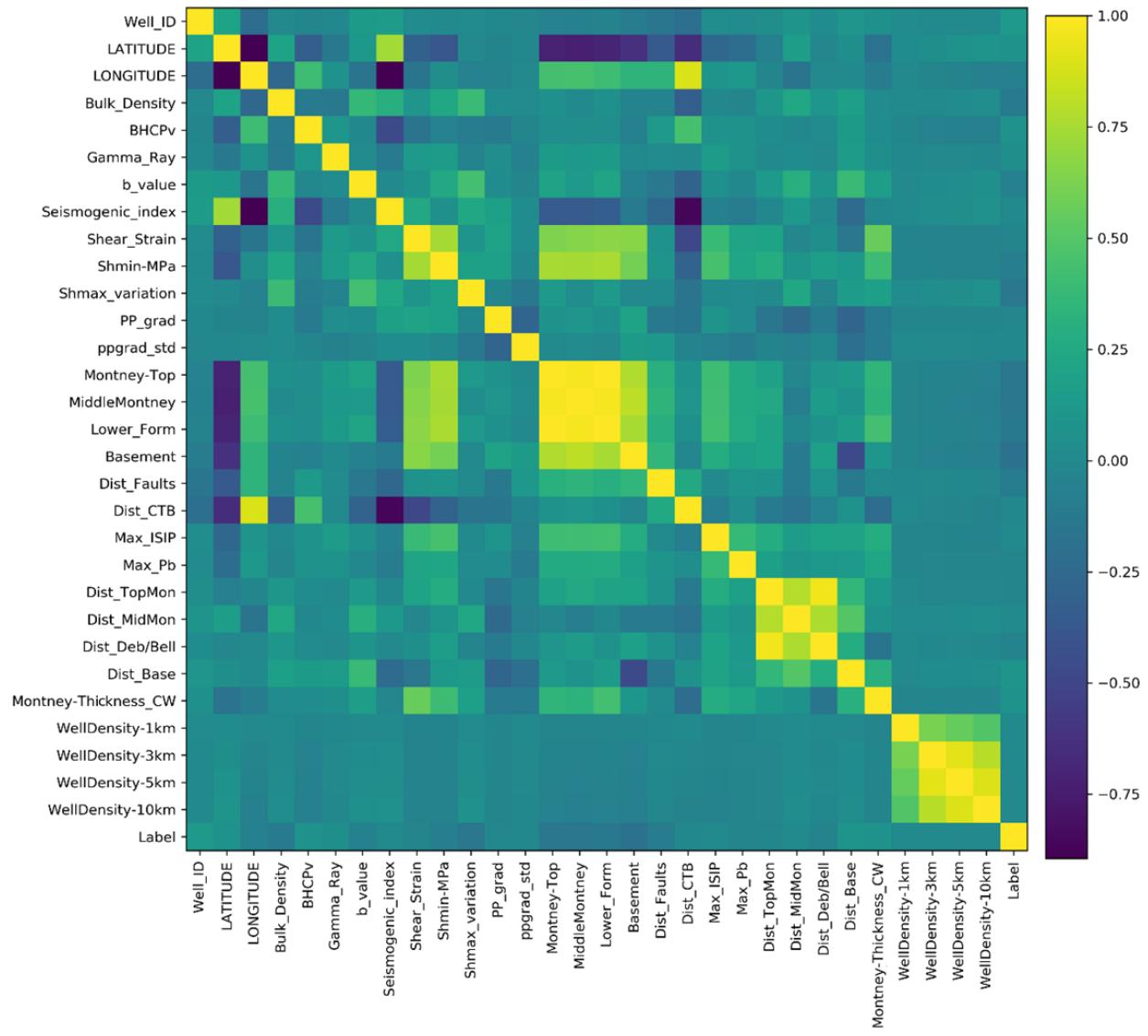


Figure 10 – Correlation matrix for the classification analysis of induced seismicity likelihood.

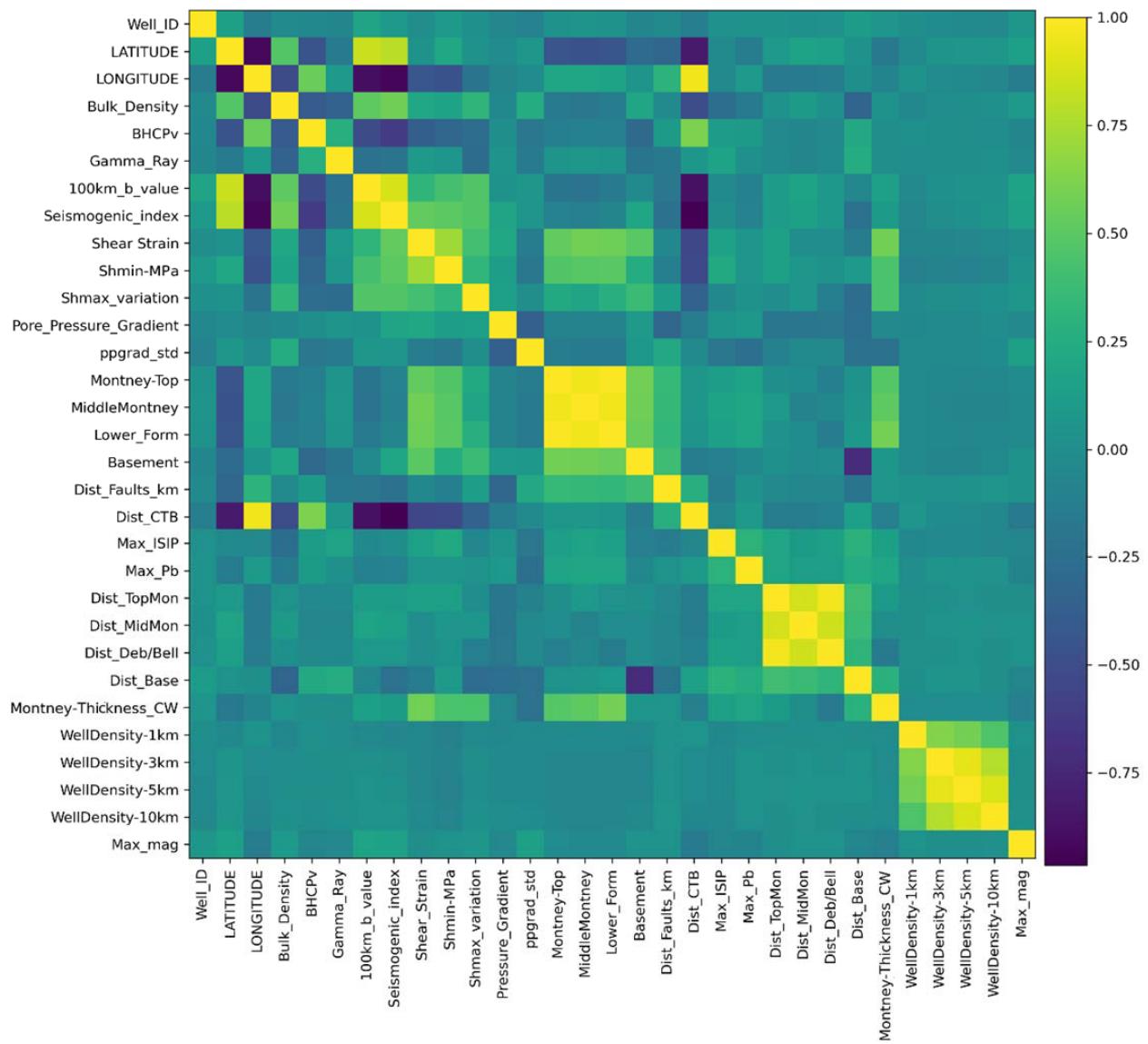


Figure 11 – Correlation matrix for the regression analysis of event magnitude severity.

5. Susceptibility Analysis

5.1. Feature Importance (Geological and Operational Data)

The first analysis conducted was a classification analysis of feature importance. The classification model provides a likelihood, and the feature importance technique uses machine learning to assign a score to input features based on how useful they are at predicting a target variable, in this case the likelihood of a well being seismogenic. For features, both the geological and operational parameter datasets (i.e., Well Features Dataset; Section 2.4) were used to train, test and validate the different machine learning models tested. (This differs from the induced seismicity susceptibility analysis that follows in the next section where a number of operational features are removed from consideration as they do not lend themselves to the concept of susceptibility).

5.1.1. Model Training, Tuning and Comparison Testing

To begin, the integrated Well Features Dataset was split into a training set (75%) and test set (25%). Table 3 shows the training results after tuning using a 50-fold cross validation (i.e., 50 iterations). The Dummy Classifier model was used to establish a baseline with 80% accuracy. This model does not take features into consideration and predicts based on the most frequent population. Since 80% of the wells in the dataset are non-seismogenic, the Dummy Classifier predicts everything as non-seismogenic resulting in 80% accuracy. Accordingly, because no wells are predicted to be seismogenic, the Precision and Recall (and therefore F1-score) for this model are zero.

Table 3 – Cross validation results of the feature-importance classification models, which considers all geological and operational features in the compiled Well Features Dataset. Values given are the average over 50-fold iterations; values in parentheses are standard deviations.

Model	Validation Accuracy	Validation Recall	Validation Precision	Validation F1-score
Dummy Classifier	0.789 (+/- 0.009)	0.000 (+/- 0.000)	0.000 (+/- 0.000)	0.000 (+/- 0.000)
Logistic Regression	0.679 (+/- 0.071)	0.707 (+/- 0.134)	0.372 (+/- 0.080)	0.484 (+/- 0.089)
Decision Tree	0.847 (+/- 0.053)	0.599 (+/- 0.150)	0.656 (+/- 0.139)	0.620 (+/- 0.130)
Random Forest	0.895 (+/- 0.034)	0.589 (+/- 0.142)	0.881 (+/- 0.102)	0.696 (+/- 0.116)
SVC	0.891 (+/- 0.042)	0.656 (+/- 0.122)	0.799 (+/- 0.124)	0.715 (+/- 0.110)
KNN	0.856 (+/- 0.036)	0.546 (+/- 0.126)	0.719 (+/- 0.127)	0.611 (+/- 0.102)
XGBoost	0.905 (+/- 0.038)	0.684 (+/- 0.133)	0.843 (+/- 0.116)	0.748 (+/- 0.107)
Light GBM	0.919 (+/- 0.039)	0.724 (+/- 0.124)	0.877 (+/- 0.111)	0.788 (+/- 0.105)

Based on this analysis, the top four performing models (i.e., highest F1-scores) were selected to be applied to the unseen test dataset. These were the Light GBM, XGBoost, Support Vector Classifier (SVC) and Random Forest models. Figure 12 shows the confusion matrix results for these four models, which returned the following F1-scores: Light GBM (77.4%), XGBoost (70.1%), SVC (69.2%) and Random Forest (69.0%). These test scores were similar to the validation F1-scores for these models, meaning that they perform well on unseen data (Figure 12). Thus, these four models were selected for detailed interpretation.

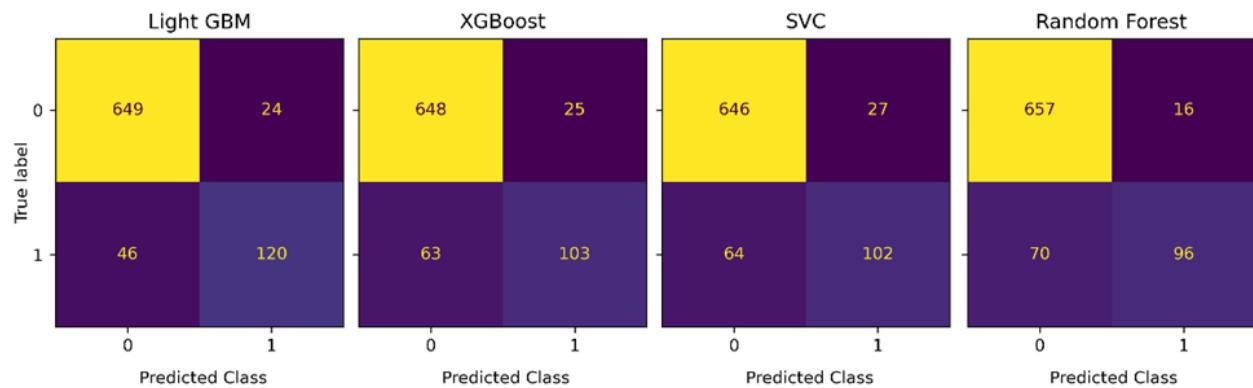


Figure 12 – Confusion matrices for the trained model test results against unseen data for the top four performing machine learning models, where label = 0 is a non-seismogenic well and label = 1 is a seismogenic well. See Figure 6 for reference.

5.1.2. Model Interpretation

There is no mathematical definition of interpretability. A (non-mathematical) definition by Miller (2017) is that interpretability is the degree to which the cause of a decision can be understood. The higher the interpretability of a machine learning model, the easier it is to understand the predictions being made. Several interpretation tools/methods are introduced here and then used again in later sections. These include feature importance plots, SHAP summary plots, and permutation importance plots.

Feature Importance Plots

Figures 13 to 15 show the feature importance plots resulting from the Light GBM, XGBoost and Random Forest models, respectively. (Note that the feature importance plot for the Support Vector Classifier model isn't included here due to the high computational times it requires). The importance of each feature in these plots is indicated as a coefficient; these coefficients do not have any physical meaning and are used as relative values for comparison and not as absolute values. Table 4 reports the top five important features for each of these models.

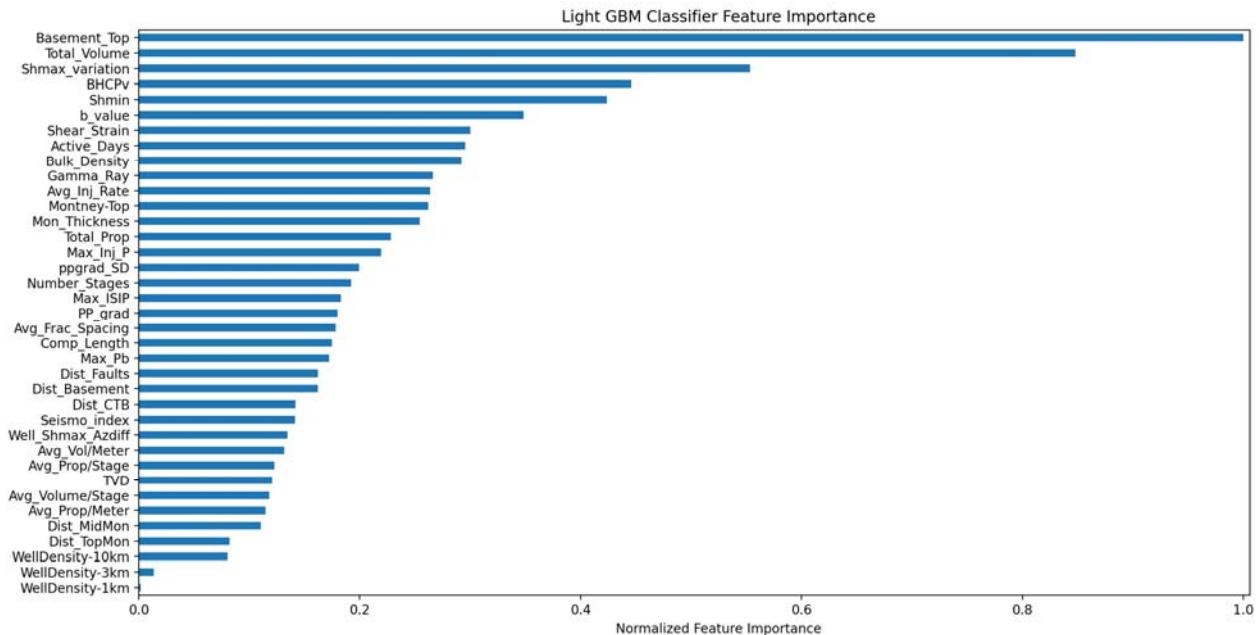


Figure 13 – Light GBM classifier results ranking feature importance for induced seismicity likelihood.

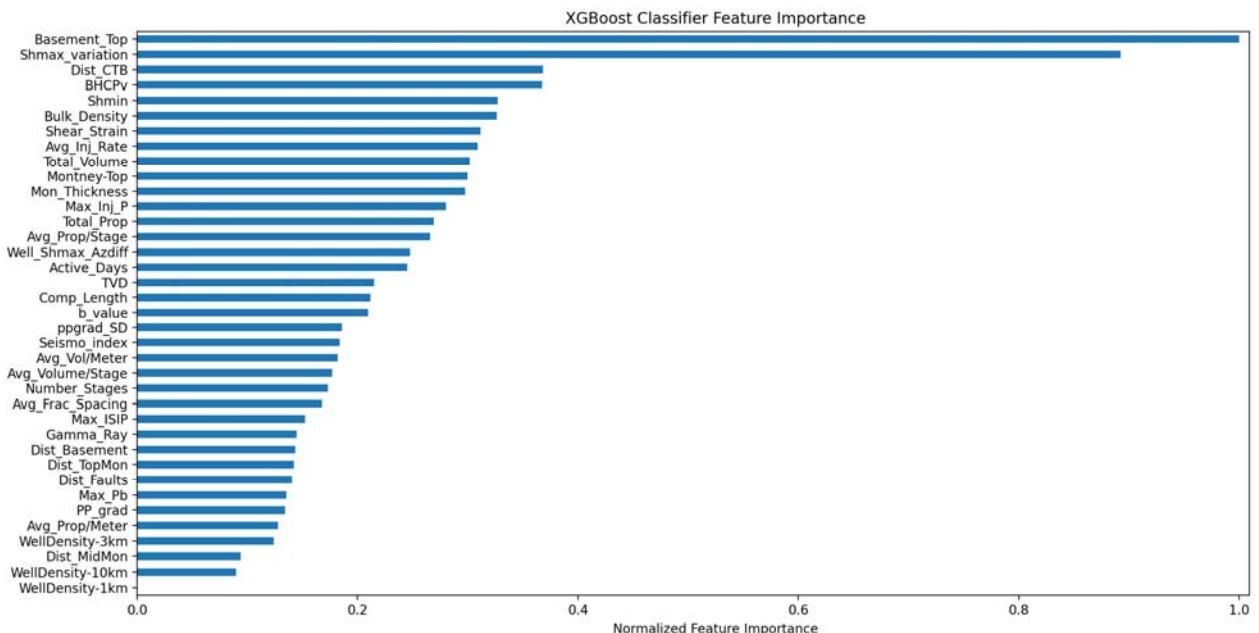


Figure 14 – XGBoost classifier results ranking feature importance for induced seismicity likelihood.

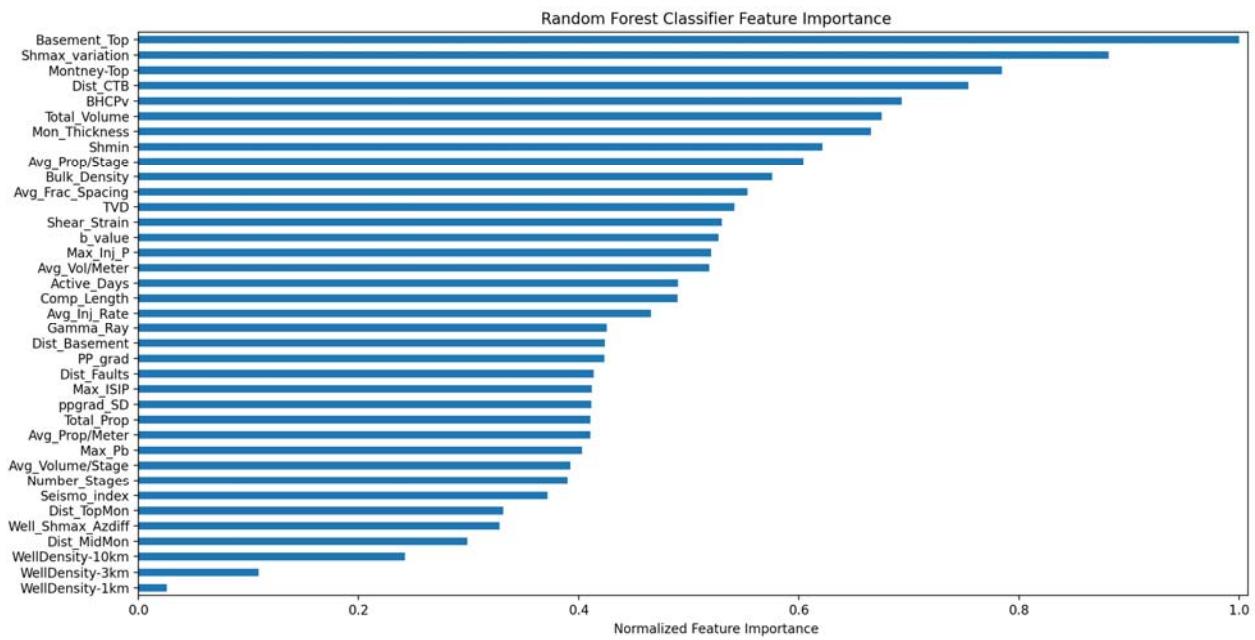


Figure 15 – Random Forest classifier results ranking feature importance for induced seismicity likelihood.

Table 4 – Top five ranked features with respect to importance in predicting induced seismicity likelihood for a hydraulic fracturing well, for each of the top performing classifier models.

Rank	Light GBM	XGBoost	Random Forest
1	Depth to the top of basement	Depth to the top of basement	Depth to the top of basement
2	Total injection volume	Variation of local Shmax from regional orientation	Variation of local Shmax from regional orientation
3	Variation of local Shmax from regional orientation	Distance of well from Cordilleran thrust belt	Depth to the top of the Montney
4	P-wave velocity of Montney at well location	P-wave velocity of Montney at well location	Distance of well from Cordilleran thrust belt
5	Minimum horizontal stress magnitude	Minimum horizontal stress magnitude	P-wave velocity of Montney at well location

As shown in Table 4, all three top-performing classifier models rank the depth to the top of the Paleozoic basement at the well location as the most important feature for correctly predicting the likelihood of a well being seismogenic. There is also agreement regarding the importance of the deviation of the orientation of the maximum horizontal stress at the well location relative to bearing 045 degrees, which is the general regional orientation. This was the second ranked important feature for the XGboost and Random Forest models, and the third ranked for the Light GBM model. The other feature that ranked in the top five for all three models was the average P-wave velocity for the Montney formation at the well location calculated from the borehole sonic logs. Features that ranked in the top five for two of the three models include the minimum horizontal stress magnitude and distance of the well from the Cordilleran thrust belt. Other important features rounding out the top five for the models include the injection volume and depth to the top of the Montney formation. Overall, the models generally ranked geological features higher than operational features.

SHAP Summary Plots

SHapley Additive exPlanations (SHAP) analyses utilize a game theory approach to help interpret predictions from complex model output. SHAP assigns each feature an importance value for a particular prediction and shows there is a unique solution for each class of additive feature importance that adheres to desirable properties (Lundberg & Lee, 2017). The SHAP Tree Explainer tool is a subcategory of SHAP that is specifically built for interpreting tree-based models such as decision trees, random forests and gradient boosting methods. The SHAP value plot shows both the positive and negative relationships of the predictors for a target variable.

Figures 16 to 18 present the summary plots from the SHAP analyses for the top performing classification models that are tree-based models (Light GBM, XGBoost and Random Forest). The plots combine feature importance with feature impact. To read these plots, each feature is ordered according to its importance (starting with the most important). Next, points are plotted to show the distribution of the SHAP values using color to represent the feature value.

Overlapping points are stacked in the vertical direction to give a sense of the distribution of the SHAP values. The impact of the stacked points (both positive and negative) is shown along the horizontal axis. In doing so, the effect of a feature value (i.e., high value or low value) is shown relative to its association with either a higher or lower prediction. For example, the SHAP results for the Light GBM model (Figure 16) show that lower values of depth to the top of the basement (blue points) generally have a higher positive correlation with a well being seismogenic. The same correlation is seen in the other SHAP plots, as well as exceptions where some higher values of depth to the top of the basement (red points) also have high positive correlations with wells being seismogenic. This indicates that the depth of the basement by itself isn't a lone predictor but depends on other conditional factors related to the geology and operations. Correlations determined through machine learning outputs do not establish a cause-and-effect relationship ("correlation is not causation").

The most important operational feature in these plots is the total injected fluid volume, where high injection volumes have a positive impact on model correlation. The influence of injection volume is well studied and has been shown to have a positive correlation with induced seismicity (Schultz et al., 2018; Fasola et al., 2019; Amini, 2020). The SHAP plots also show the reverse, that a smaller injection volume has a high negative correlation. Hydraulic fracturing wells with smaller injected fluid volumes have a significantly lower likelihood of triggering an induced seismicity event.

The ranking of the total injected fluid volume as one of the top five important features in the Light GBM, XGBoost and Random Forest SHAP analyses, differs from the feature importance analyses where it only ranked highly in the Light GBM model. However, it should be noted that the ranking order of feature importance in a SHAP analysis might differ slightly from those reported in the feature importance plots. This is because the SHAP plots are calculated for one instance of each model's validation, whereas the ranking in the feature importance plots is based on an averaging of 50 cross validation runs. Still, features such as depth to the top of the basement, azimuth variation between the local and regional maximum horizontal stress, and minimum horizontal stress, ranked highly in both the SHAP and feature importance analyses. The SHAP analysis results then add further transparency and interpretability, showing the relative impact of each feature on either a higher or lower prediction and whether it correlates positively or negatively with the likelihood of a well being seismogenic.

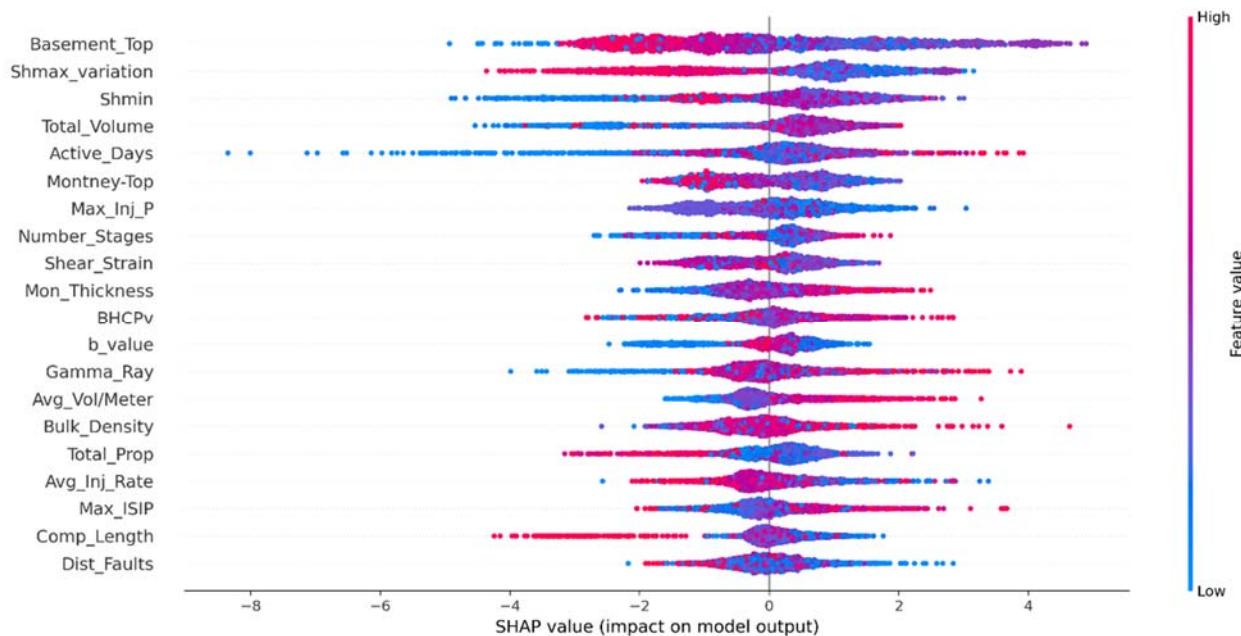


Figure 16 – Light GBM model SHAP summary plot. See text for explanation.

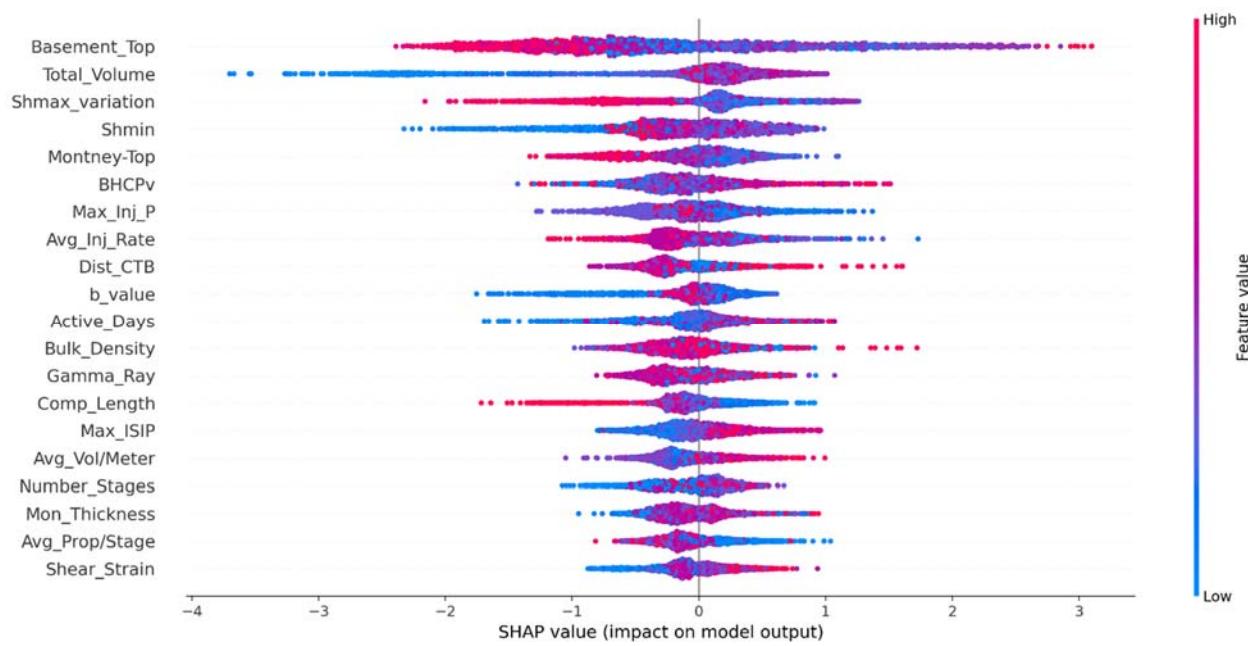


Figure 17 – XGBoost model SHAP summary plot. See text for explanation.

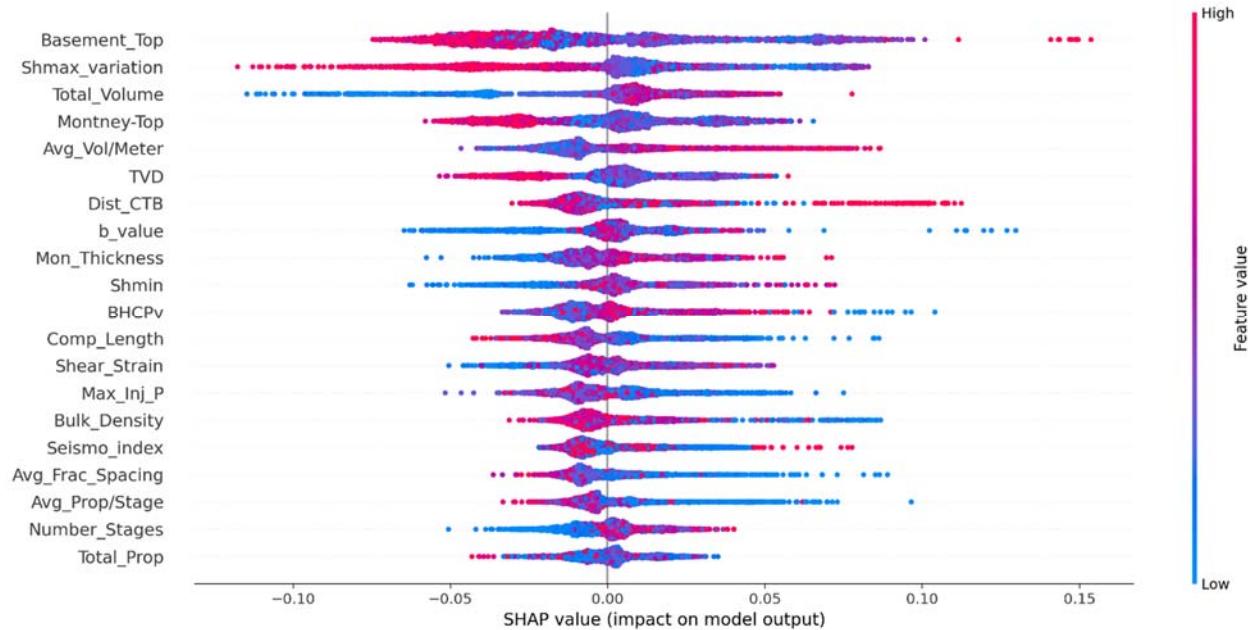


Figure 18 – Random Forest model SHAP summary plot. See text for explanation.

Figure 19 provides the SHAP analysis results for the other top performing model, the Support Vector Classifier (SVC) model. As previously noted, the SHAP Tree Explainer tool used for the other classifier models is specifically built for interpreting tree-based models. The SVC model is not a tree classifier and requires use of a kernel explanation tool instead. This procedure is computationally very heavy and time demanding. For the analysis shown in Figure 19, a subset of 300 data samples was used instead of the full dataset. Therefore, this analysis is less accurate compared to the SHAP plots shown for the other models. This is reflected in the order of feature importance, which is somewhat different from the previous models (note that a feature importance plot was not possible for the SVC model for further comparison). Features such as the thickness of the Montney, the breakdown pressure, and number of hydraulic fracturing stages appear in the top five with respect to importance. Still some similarities can be seen such as the importance of the depth to the top of basement and the azimuth variation between the local and regional maximum horizontal stress being among the top five important features.

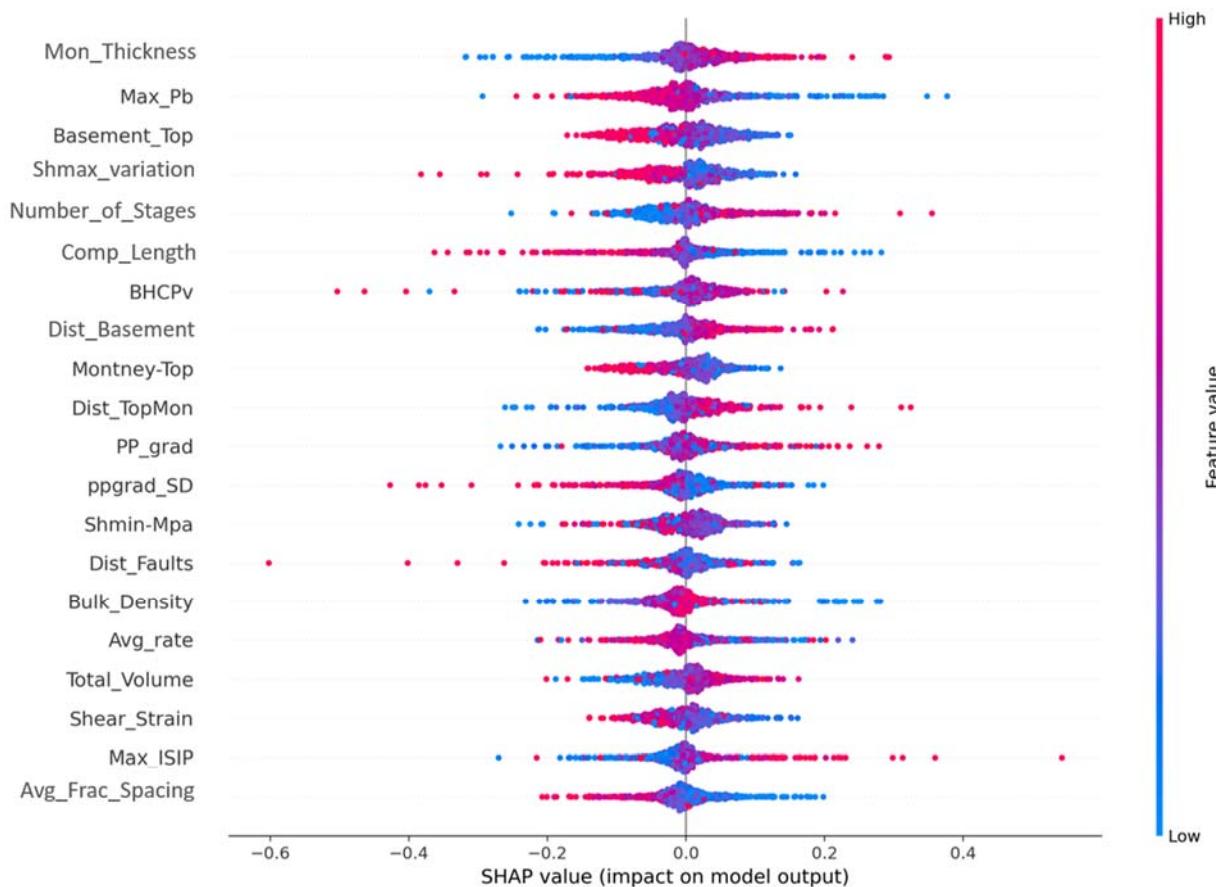


Figure 19 – Support Vector Classifier (SVC) model SHAP summary plot. See text for explanation.

Permutation Importance Plots

Permutation importance analyses (Breiman, 2001) measure the increase in the prediction error of a model if a 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 relied 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 the decrease in the model score is indicative of how much the model depends on the feature.

The permutation importance plots for the four top performing machine learning models are provided in Figures 20 to 23. The y-axis of these plots shows the features in order of their importance with the most important feature being listed at the top. 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).

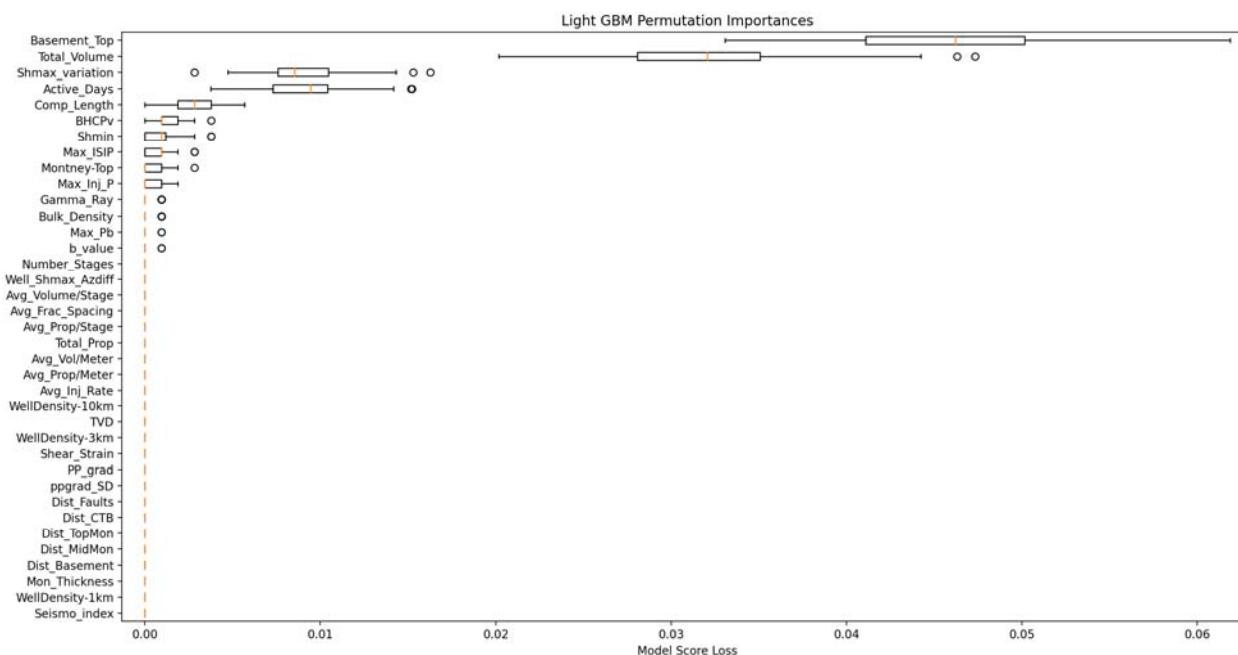


Figure 20 – Light GBM model permutation importance plot. See text for explanation.

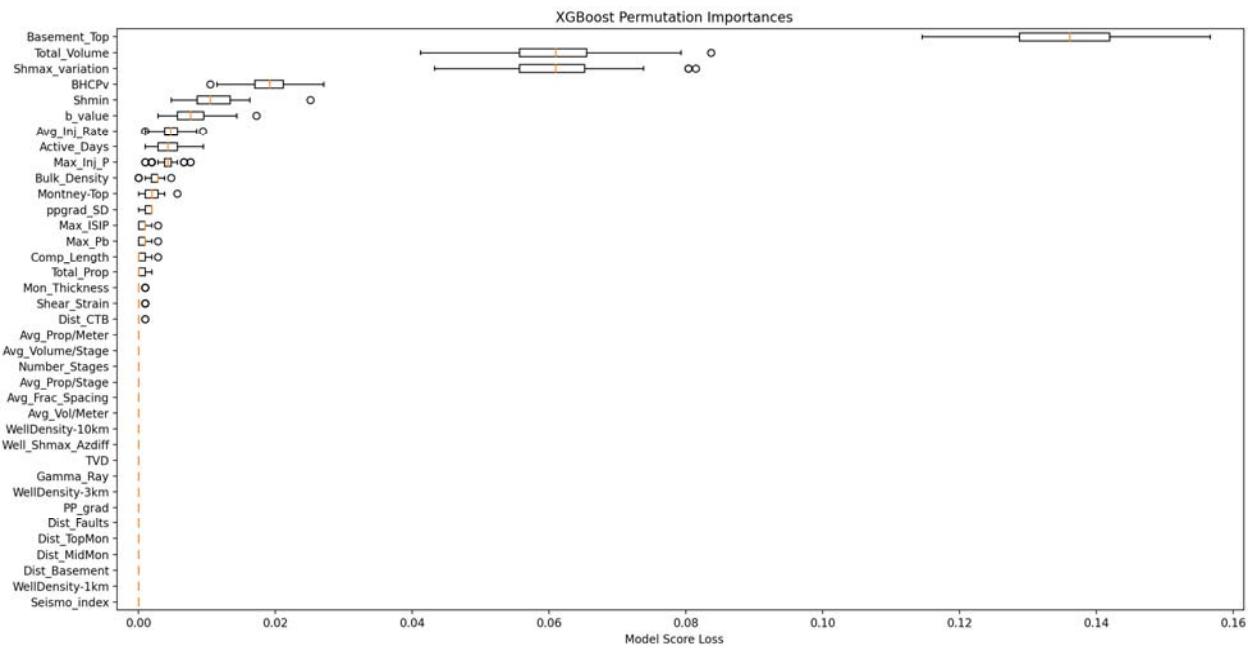


Figure 21 – XGBoost model permutation importance plot. See text for explanation.

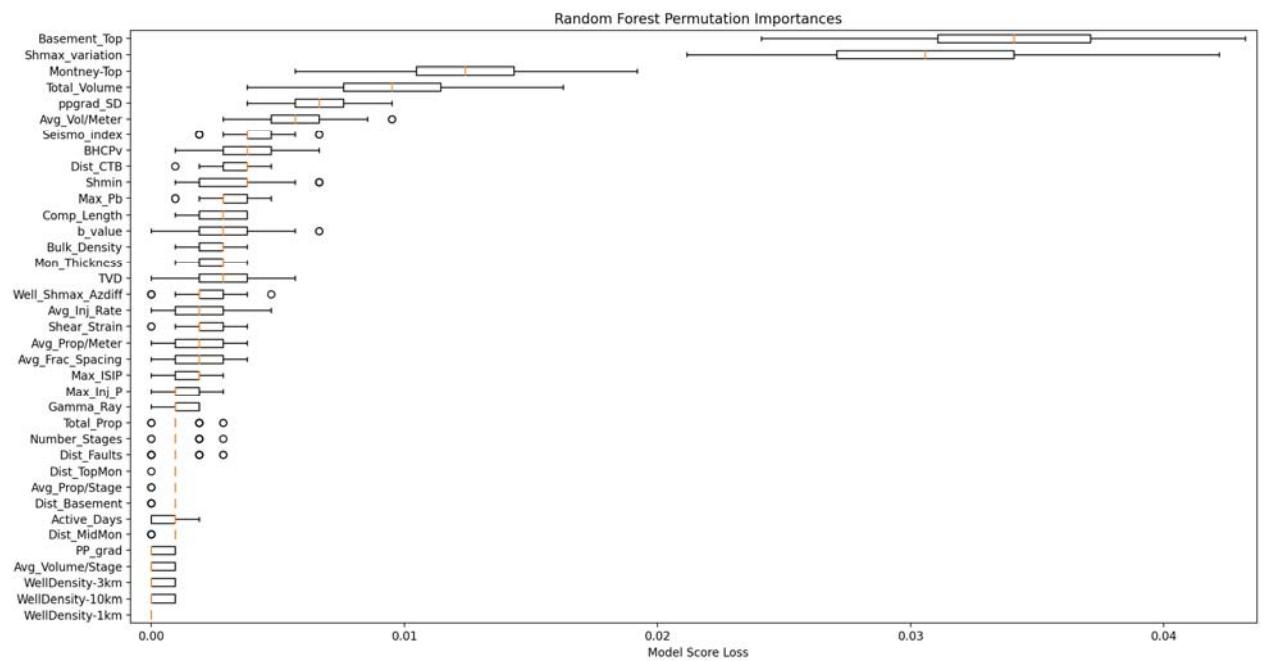


Figure 22 – Random Forest models permutation importance plot. See text for explanation.

The results from the permutation importance analyses for the tree-based classifier models (Light GBM, XGBoost and Random Forest) support the previous results derived from the SHAP analyses. Again, these show that the depth to the top of basement and azimuth variation between the local and regional maximum horizontal stress are highly important as geological factors, while the total volume of fluids injected is one of the most important operational factors. More specifically, the permutation importance plots show that these three features contribute significantly more to correct predictions than the others. For the other features, as their importance decreases, the model's sensitivity to these features also decreases.

The SVC permutation importance results in Figure 23 show a different response. The most important features in this case are the bulk density and pore pressure gradient, but the results also show a high sensitivity to a much wider range of features. This sensitivity to all features for the SVC model is reflective of the need to use a nonlinear kernel, in contrast to the tree-based classifier models. The difficult interpretability of the SVC model results cautions whether it is a suitable model for this specific purpose (i.e., induced seismicity susceptibility analysis).

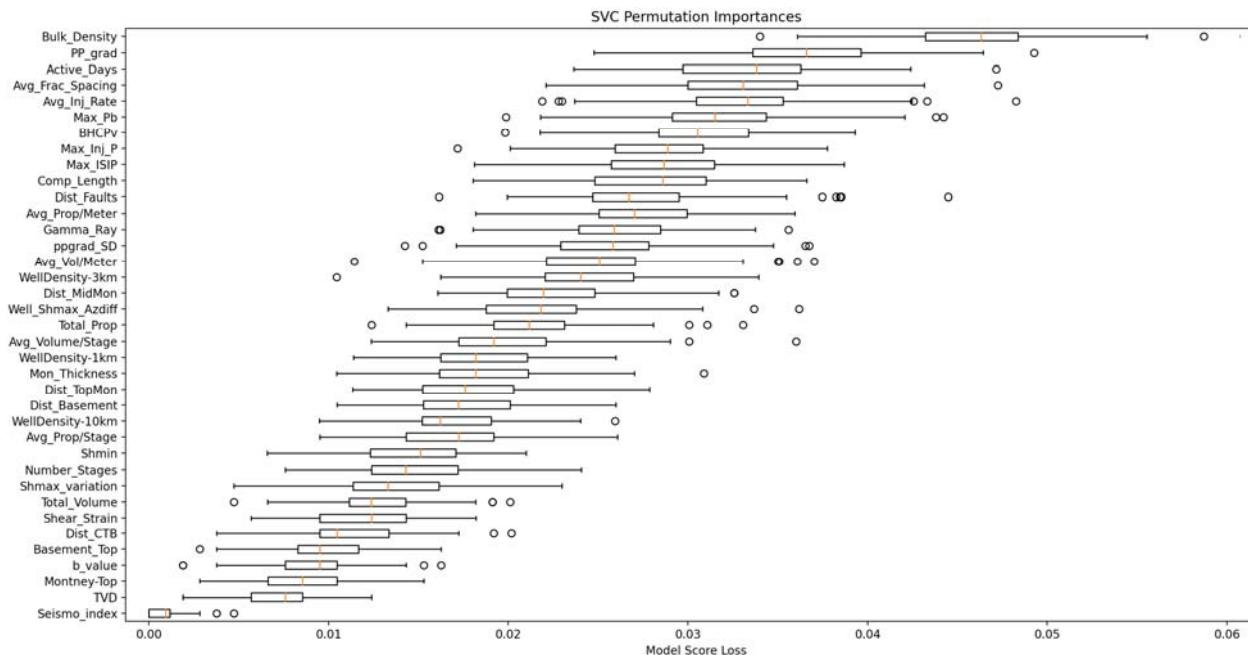


Figure 23 – SVC model permutation importance plot. See text for explanation.

5.2. Induced Seismicity Susceptibility (Geological Data)

The objective of developing an induced seismicity susceptibility map builds on the feature importance analysis presented in the previous section, but does so recognizing that not all features lend themselves to susceptibility. Susceptibility maps combine different features that contribute to a hazard to give an indication of where the hazard is more likely to occur; it refers to the spatial likelihood or probability of the hazard occurring but not the temporal. In the context of this study, data is being utilized from existing hydraulic fracturing wells to predict the susceptibility of a new well at a given location to being seismogenic. This differs from hazard maps, which include reference to a time frame to also consider the hazard frequency (temporal reoccurrence) and magnitude of the hazard event. The terms susceptibility and hazard are often erroneously used as synonyms.

Thus, the next step was to repeat the feature importance analysis but using the Well Susceptibility Dataset (Section 2.4), which limits the models to only those features that lend themselves to susceptibility. This includes the geological features plus any operational features that act as a proxy for formation properties (e.g., breakdown pressure, instantaneous shut-in pressure) or account for possible preconditioning of a given area related to previous hydraulic fracturing activities (e.g. well density).

With respect to severity, the susceptibility analysis only speaks to the magnitude distribution of the induced seismicity events making up the seismic event catalogue used. This had a magnitude of completeness (M_c) of 1.1, meaning that reference to susceptibility in the analysis that follows is with respect to the likelihood of generating an induced seismicity event with a minimum magnitude of 1.1.

5.2.1. Model Training, Tuning and Comparison Testing

The same procedures as explained in Section 5.1.1. were followed in order to train and test the susceptibility-focussed classification models. Again, a 75%-25% split was used for the training and testing sets, respectively. Table 5 presents the final cross-validation results of the model training in which the analysis was restricted to the reduced susceptibility-focussed data set. As before, model training and tuning was carried out using a 50-fold cross validation (i.e., 50 iterations) and the Dummy Classifier model was used to establish a baseline with 80% accuracy.

Examining the cross validation results (Table 5), the top three performing models (i.e. highest F1-scores) were the Light GBM, XGBoost and Random Forest models. These were the same top performing tree-based models as determined for the feature importance analysis, and were subsequently selected to be applied to the unseen test dataset. Figure 24 shows the confusion matrix results for these three models, which returned the following F1-scores: Light GBM (73.4%), Random Forest (70.1%) and XGBoost (61.6%). The F1-scores of the Light GBM and Random Forest models were similar to their validation training F1-scores in Table 5, meaning they perform well on unseen data. In contrast, there is a 10% difference between the XGBoost model's training and testing F1-scores (72.6% and 61.6%, respectively).

Table 5 – Cross validation results of the susceptibility-focussed classification models, which considers all geological features but only those operational features that serve as a proxy for formation properties. Values given are the average over 50-fold iterations; values in parentheses are standard deviations.

Model	Validation Accuracy	Validation Recall	Validation Precision	Validation F1-score
Dummy Classifier	0.801 (+/- 0.004)	0.000 (+/- 0.000)	0.000 (+/- 0.000)	0.000 (+/- 0.000)
Logistic Regression	0.639 (+/- 0.066)	0.591 (+/- 0.118)	0.301 (+/- 0.061)	0.396 (+/- 0.071)
Decision Tree	0.861 (+/- 0.036)	0.659 (+/- 0.109)	0.658 (+/- 0.111)	0.652 (+/- 0.086)
Random Forest	0.904 (+/- 0.030)	0.641 (+/- 0.125)	0.850 (+/- 0.102)	0.722 (+/- 0.099)
SVC	0.858 (+/- 0.041)	0.447 (+/- 0.154)	0.741 (+/- 0.192)	0.546 (+/- 0.149)
KNN	0.844 (+/- 0.034)	0.416 (+/- 0.127)	0.682 (+/- 0.168)	0.508 (+/- 0.129)
XGBoost	0.904 (+/- 0.032)	0.662 (+/- 0.137)	0.823 (+/- 0.100)	0.726 (+/- 0.107)
Light GBM	0.905 (+/- 0.035)	0.660 (+/- 0.137)	0.833 (+/- 0.113)	0.729 (+/- 0.111)

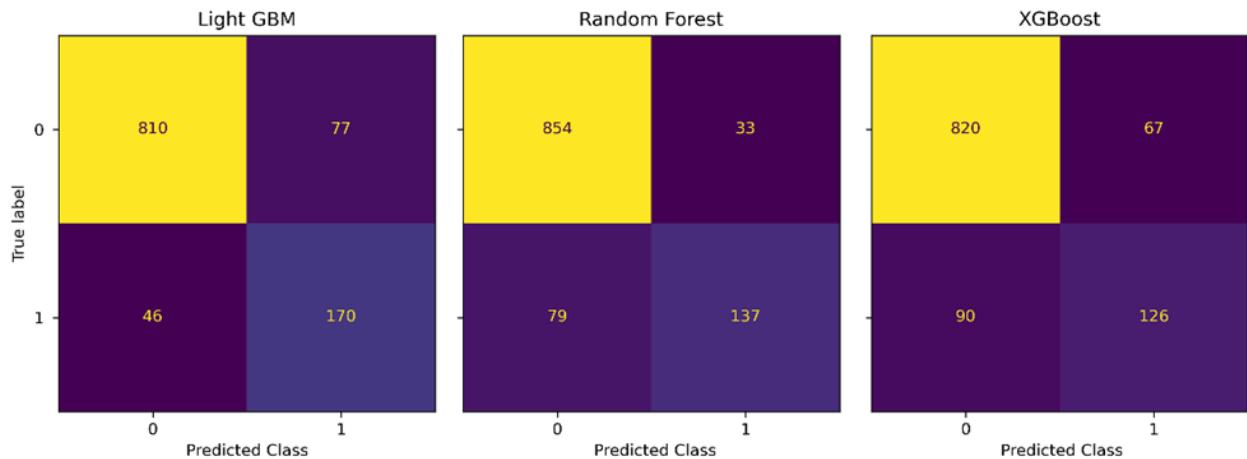


Figure 24 – Confusion matrices for the trained model test results against unseen data for the top three performing susceptibility-focussed classifier models, where label = 0 is a non-seismogenic well and label = 1 is a seismogenic well. See Figure 6 for reference.

To investigate the performance of the XGBoost model further, Figure 25a plots the receiver operating characteristic (ROC) curves for the top three performing models being compared. A ROC curve depicts the diagnostic ability of a binary classifier system as its discrimination threshold is varied. It plots the true positive rate (Recall) against the false positive rate ($1 - \text{Specificity}$; see Figure 6) at various threshold settings. The best possible prediction method would yield a point in the upper left corner of the ROC plot (True Positive Rate = 1, False

Positive Rate = 0), representing 100% sensitivity (no false negatives) and 100% specificity (no false positives). A random guess would plot along a diagonal line (line of no-discrimination) from the bottom left to the top right corners. Thus, a better performing model would generate a ROC curve that is closer to the upper left corner of ROC space. As shown in Figure 25, both the Light GBM and Random Forest models are performing similarly and better than the XGBoost classifier. This is in agreement with the results of F1-score calculations. Figure 25b plots the ROC curves for all models that were tested, for comparison.

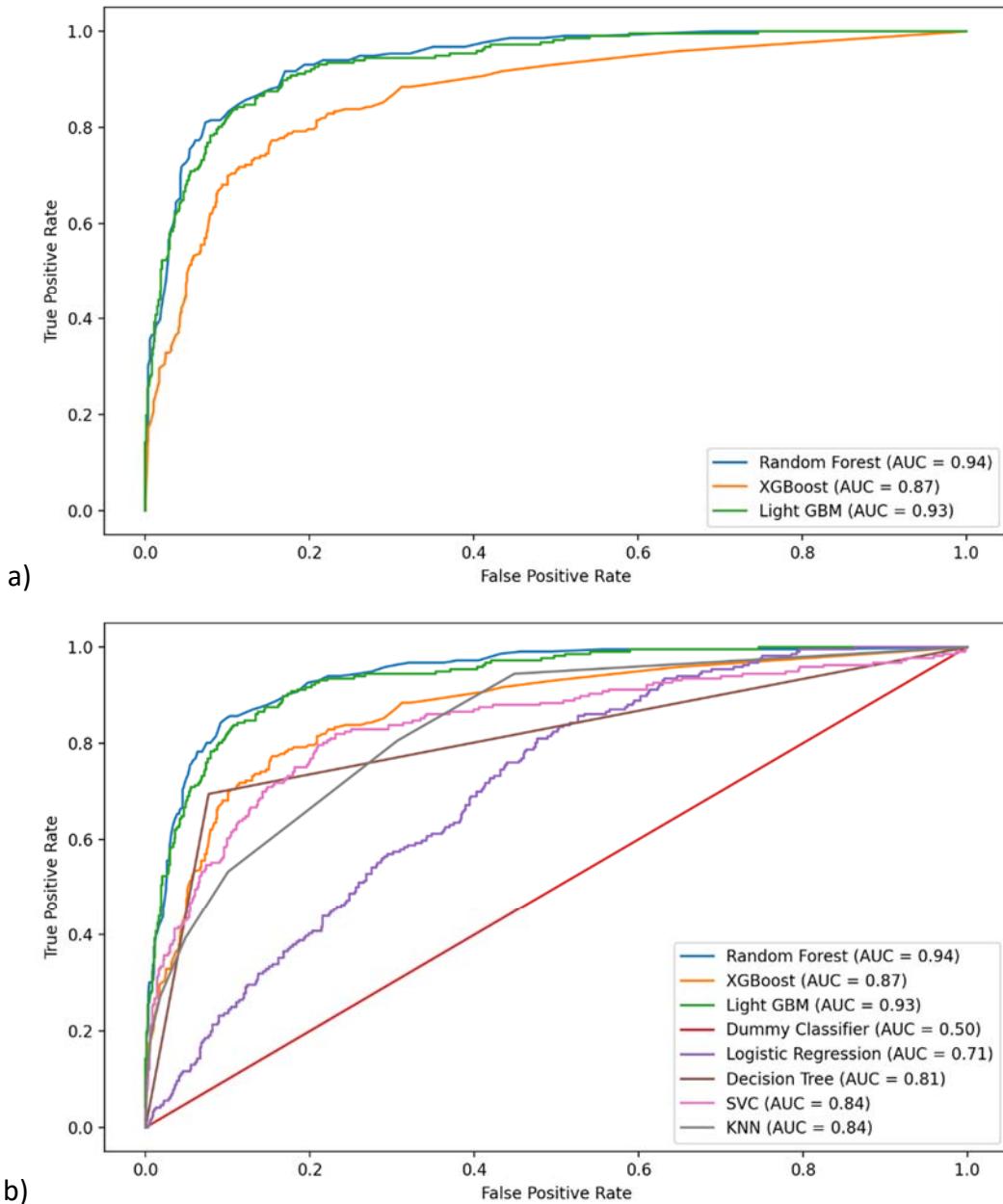


Figure 25 – Receiver operating characteristic (ROC) curves, comparing those for: a) the top three classifiers, and b) all classifier models.

Based on the model validation and test results, the Light GBM and Random Forest models were selected as the best performing models and were used for model interpretation and generating induced seismicity susceptibility maps.

5.2.2. Model Interpretation

Feature Importance Plots

The training results for the Light GBM and Random Forest models were further analyzed to investigate their ranking of feature importance with respect to induced seismicity susceptibility. Figures 26 and 27 show the feature importance plots computed. Table 6 reports the top five important features for each of these models.

As would be expected, the results show a similar ranking of the top five important features for the susceptibility analysis as the previous feature importance analysis, minus the operational factors that were excluded from the susceptibility analysis. This was limited to total injection volume, which appeared in the top five for the Light GBM model and was replaced by the depth to the top of the Montney formation. However, without consideration and weighting given to the operational features not included, the order of the individual rankings also slightly changed for each model. In the case of the Random Forest model, the minimum horizontal stress magnitude replaced the average P-wave velocity of the Montney at the well location in the top five. Overall, both models show similar results with respect to the ranking of feature importance specific to induced seismicity susceptibility.

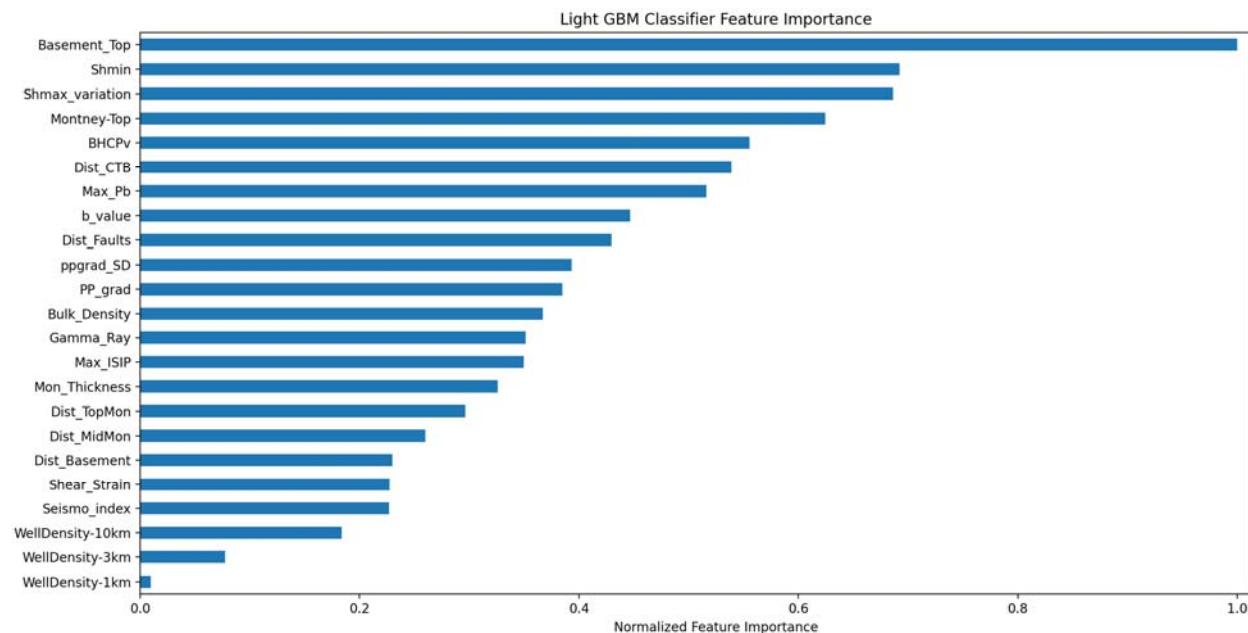


Figure 26 – Feature importance results for the Light GBM classifier model with respect to induced seismicity susceptibility.

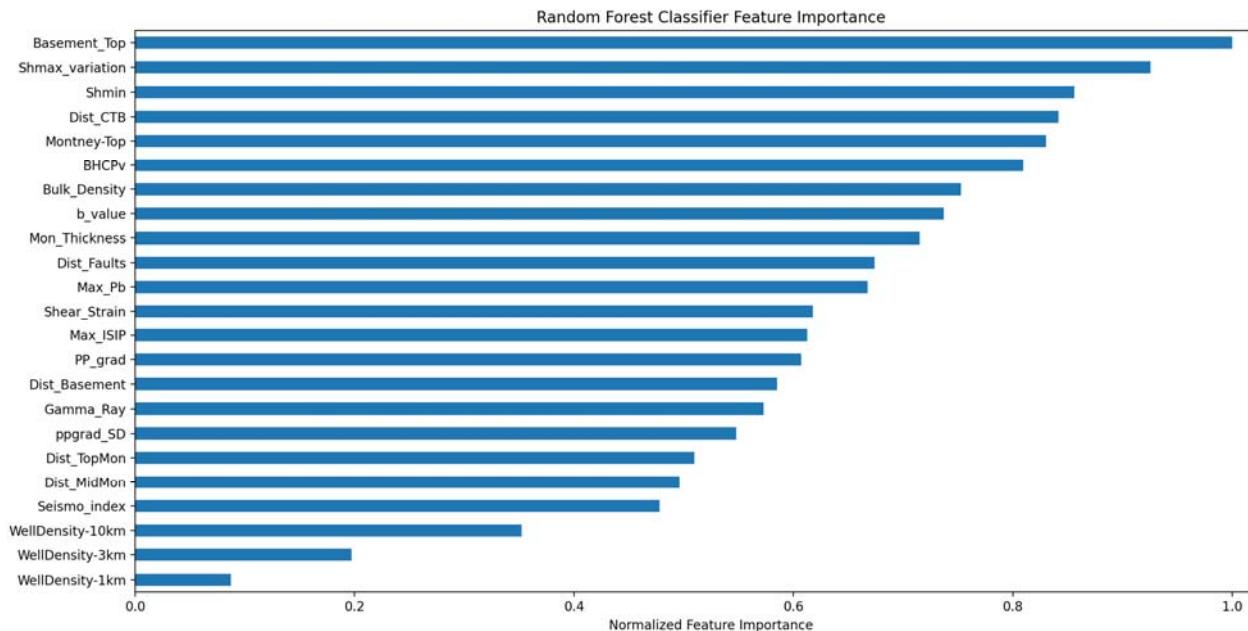


Figure 27 – Feature importance results for the Random Forest classifier model with respect to induced seismicity susceptibility.

Table 6 – Top five ranked features with respect to importance to induced seismicity susceptibility, for each of the top performing classifier models.

Rank	Light GBM	Random Forest
1	Depth to the top of basement	Depth to the top of basement
2	Minimum horizontal stress magnitude	Variation of local Shmax from regional orientation
3	Variation of local Shmax from regional orientation	Minimum horizontal stress magnitude
4	Depth to the top of the Montney	Distance of well from Cordilleran thrust belt
5	P-wave velocity of Montney at well location	Depth to the top of the Montney

Partial Dependence Plots

Partial dependence plots (PDP) show the marginal effect a feature of interest has on the predicted outcome of a machine learning model (Friedman, 2001). The analysis works by marginalizing (i.e., integrating) the model output over the distribution of the other features, so that the resulting relationship only depends on the feature of interest. The plot will then show whether the relationship is monotonic, linear or more complex across the feature's range of values in the dataset.

Figures 28 and 29 show the PDP for several features of interest for the Light GBM and Random Forest classifier models, respectively. Note that the x-axes are plotted in zero-mean and unit variance. The black tick marks at the bottom of the plot represent the deciles of the distribution of feature values. The results for the influence of distance from known faults provides a first example. The PDP for this feature show a relatively monotonic (i.e., constant) relationship, indicating that the distance from known faults has a similar impact on the model's predictions across its range of values in the dataset. There is a small, notable step in the trend indicating that lower values (i.e., wells closer to known faults) have slightly more impact than higher values (i.e., wells farther from known wells). However, this shift is relatively small compared to those seen for other features. For example, both sets of results (Light GBM and Random Forest) show that shallower depths to the top of the basement have a relatively high impact on the model predictions up to their mean value; values greater than the mean depth see a sharp linear decrease in the impact the feature has on the predicted outcome.

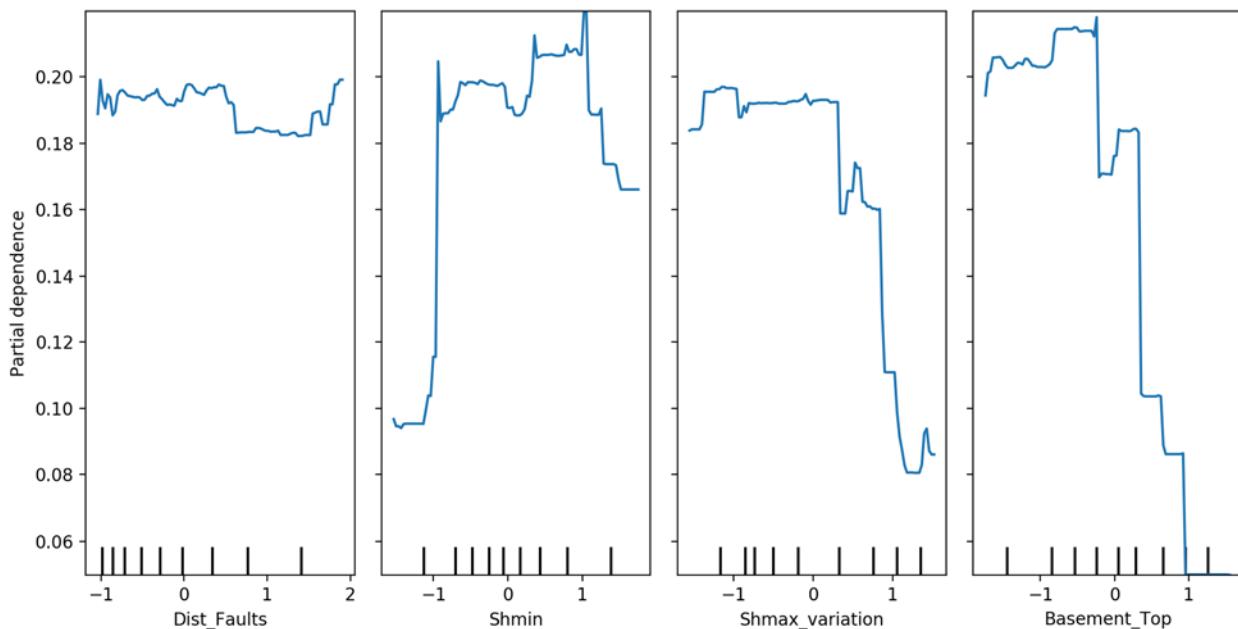


Figure 28 – Partial dependence plot (PDP) for the Light GBM classifier model with respect to induced seismicity susceptibility. See text for explanation. Note that the x-axes are plotted in zero-mean and unit-variance.

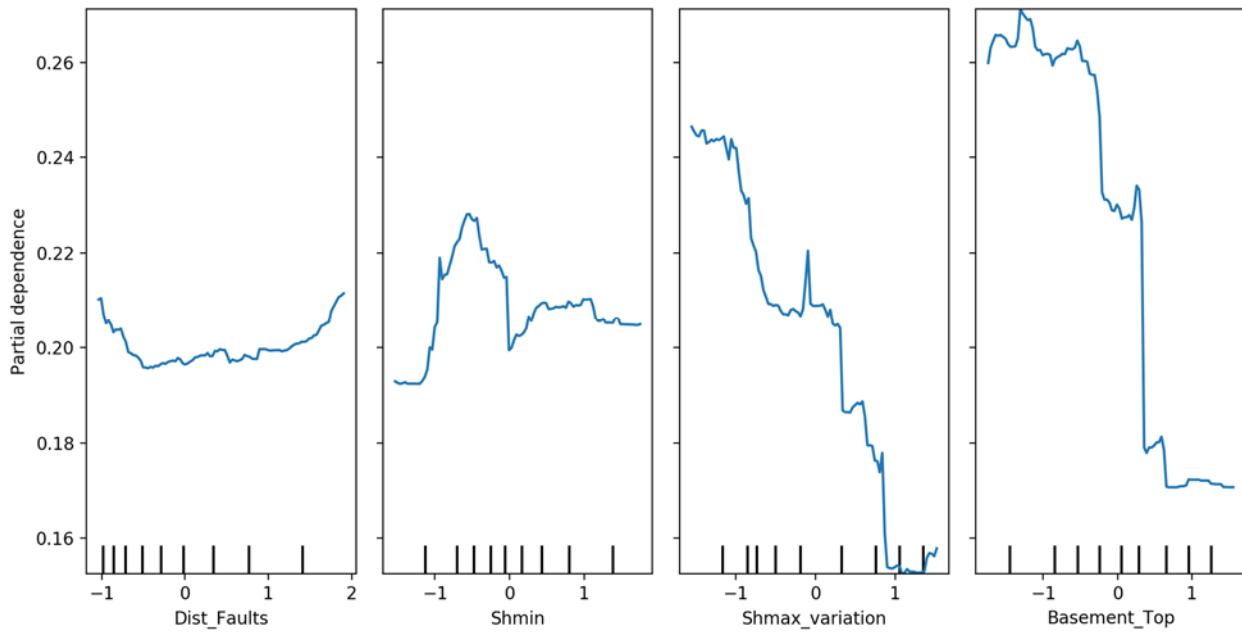


Figure 29 – Partial dependence plot (PDP) for the Random Forest classifier model with respect to induced seismicity susceptibility. See text for explanation. Note that the x-axes are plotted in zero-mean and unit-variance.

Individual Conditional Expectation Plots

Individual conditional expectation (ICE) plots provide an extension to the partial dependence plots (PDP). Whereas PDPs show the average effect of a feature's value range on the predicted response, ICE plots disaggregate these averages to show the relationship for each instance (Goldstein et al., 2015). Thus, the plots are comprised of a series of trend lines with each representing the effect of varying the feature's value for a particular instance on the model's prediction, while the values for all other features remain constant. This allows the ICE plots to be used to highlight variations in the fitted model across the value range of a feature, and to suggest where and to what extent heterogeneities (i.e., interactions) might exist.

Figures 30 and 31 show the ICE plots for several features of interest for the Light GBM and Random Forest classifier models, respectively. The average dependence trend line, coinciding with that from the PDP, is shown as a thick blue line. In this case, the y-axes are scaled to show the full partial dependence range (0 to 1). As before, the x-axes are plotted in zero-mean and unit variance, and the black tick marks at the bottom of the plot represent the deciles of the distribution of feature values. Using the same example as before to start, the ICE plot for influence of distance from known faults shows that the partial dependence of this feature on average is relatively constant (as already indicated based on the PDP). However, when looking at the partial dependence trends of the individual instances, it can be seen that there are cases with a higher variation across the range of values for the feature. This suggests that for some wells, the sensitivity to faults being nearby is greater with respect to induced seismicity.

susceptibility, but in general the partial dependence decreases with distances being farther away from known faults. Again, for distances from the top of the basement, most instances show a sharp decrease in sensitivity for those deeper than the mean value, and this trend is fairly consistent for the different instances.

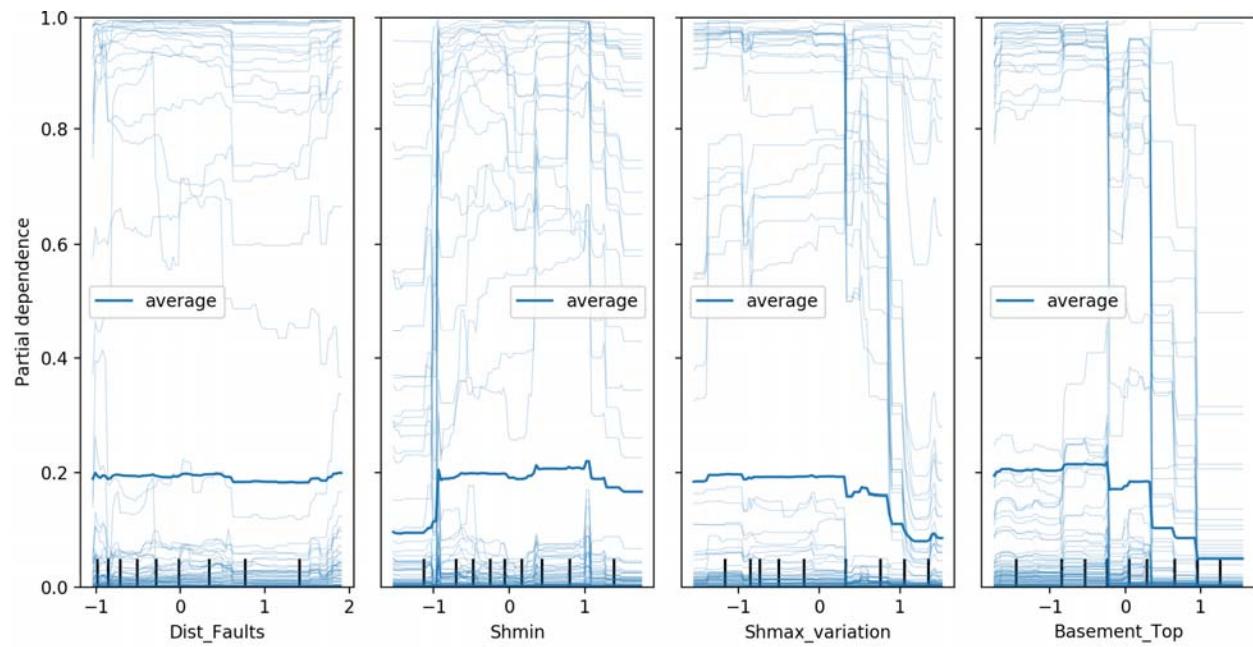


Figure 30 – Individual conditional expectation (ICE) plot for the Light GBM classifier model with respect to induced seismicity susceptibility. Note that the x-axes are plotted in zero-mean and unit-variance.

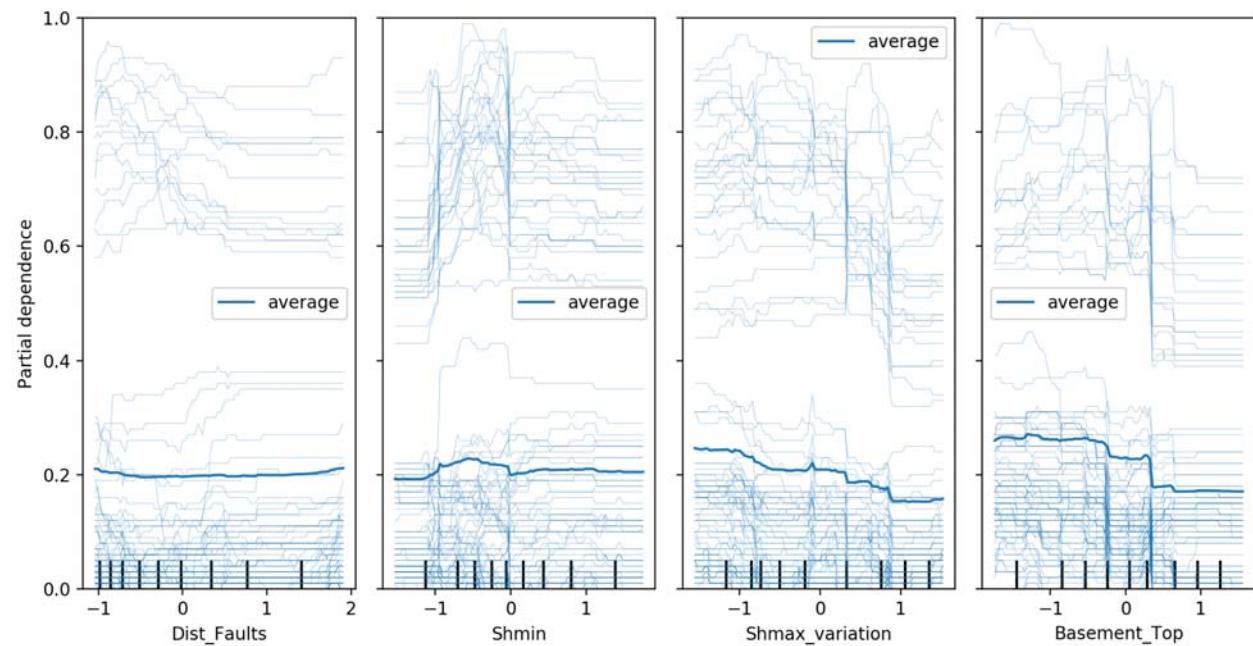


Figure 31 – Individual conditional expectation (ICE) plot for the Random Forest classifier with respect to induced seismicity susceptibility. Note that the x-axes are plotted in zero-mean and unit-variance.

SHAP Summary Plots

Figures 32 and 33 present the summary plots from the SHAP analysis for the Light GBM and Random Forest models, respectively, with regards to induced seismicity susceptibility. These results reflect those seen in the feature importance plots, and indicate whether the correlation between the feature values and induced seismicity susceptibility is positive or negative. Features such as depth to top of basement, azimuth variation between local and regional maximum horizontal stress, and distance to the Cordilleran thrust belt have a negative correlation with induced seismicity susceptibility. For example, wells closer to the Cordilleran thrust belt have a higher susceptibility to being seismogenic. In contrast, features like pore pressure gradient and seismogenic index have a positive correlation. The influence of pore pressures in the Montney has been studied by Eaton and Schultz (2018), who demonstrated a positive relationship between over-pressured areas and induced seismicity. The positive correlation determined for the seismogenic index is also logical as by definition, higher seismogenic indexes indicate a higher sensitivity to fluid injection-induced seismicity (independent of the injection parameters).

Another observation from the SHAP summary plots is that although several features have a clear correlation with induced seismicity susceptibility, a number of features do not. For example, smaller values of the minimum horizontal stress have a negative correlation, but many small values also have a positive correlation. Similarly, larger values of the minimum horizontal stress do not have a strictly positive impact and therefore a clear distinction cannot be made.

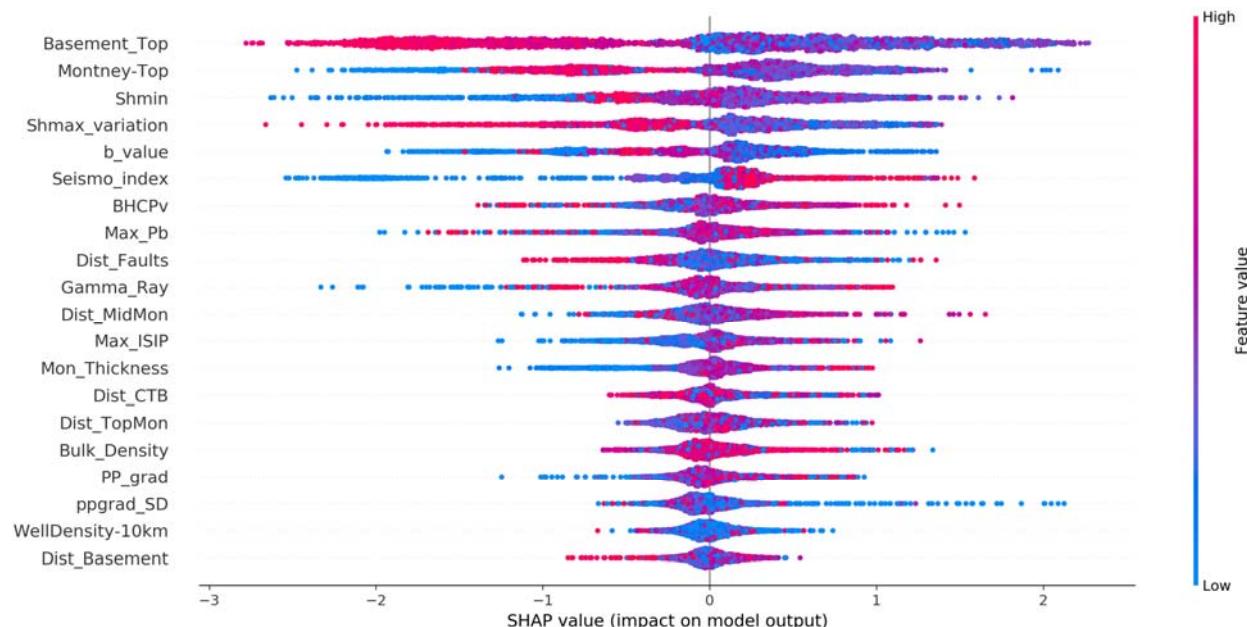


Figure 32 – SHAP summary plot for the Light GBM classifier model with respect to induced seismicity susceptibility.

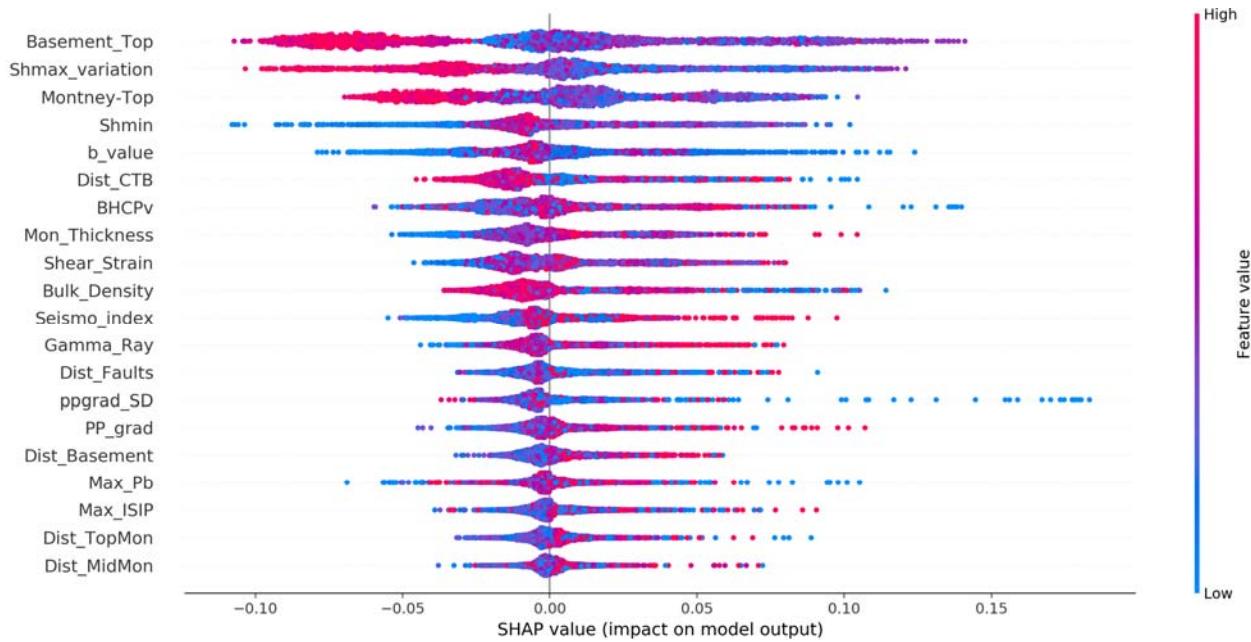


Figure 33 – SHAP summary plot for the Random Forest classifier model with respect to induced seismicity susceptibility.

Permutation Importance

Figures 34 and 35 show the permutation importance plots for the Light GBM and Random Forest models, respectively, with regards to induced seismicity susceptibility. Comparing these, both models are most sensitive to the depth to the top of the Basement.

Otherwise, the Light GBM model is most sensitive to the top five important features, which includes: depth to top of basement, azimuth variation between local and regional maximum horizontal stress, the depth to the top of the Montney, the b-value and minimum horizontal stress magnitude. The Random Forest model is most sensitive to only the top three features: depth to top of basement, azimuth variation between local and regional maximum horizontal stress, and the depth to the top of the Montney.

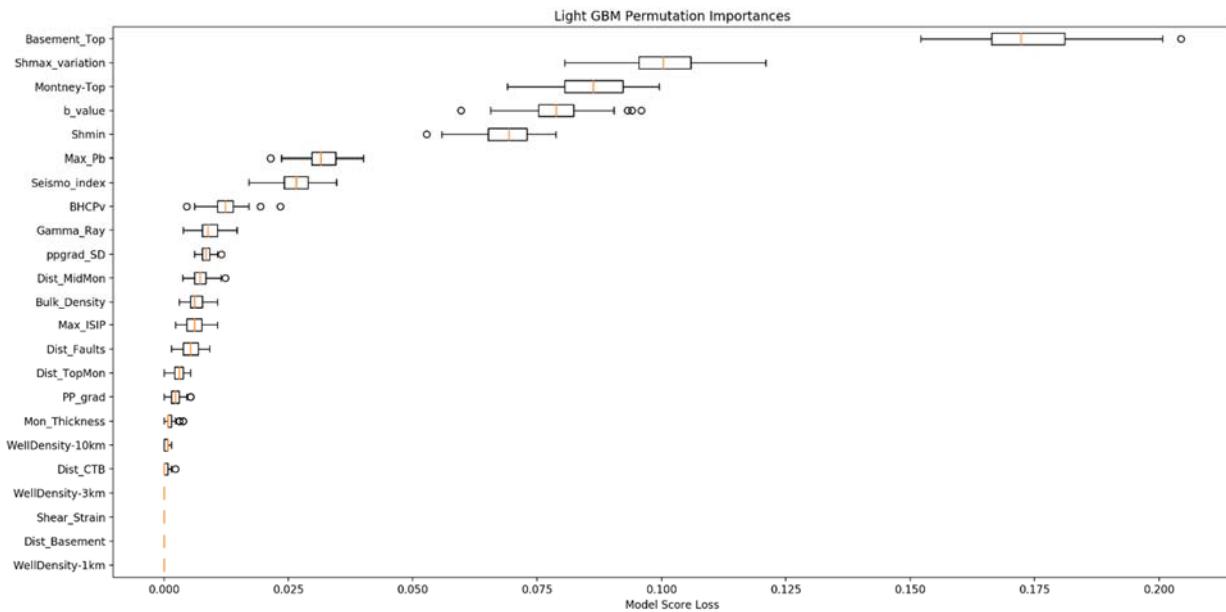


Figure 34 – Permutation importance plot for the Light GBM model with respect to induced seismicity susceptibility.

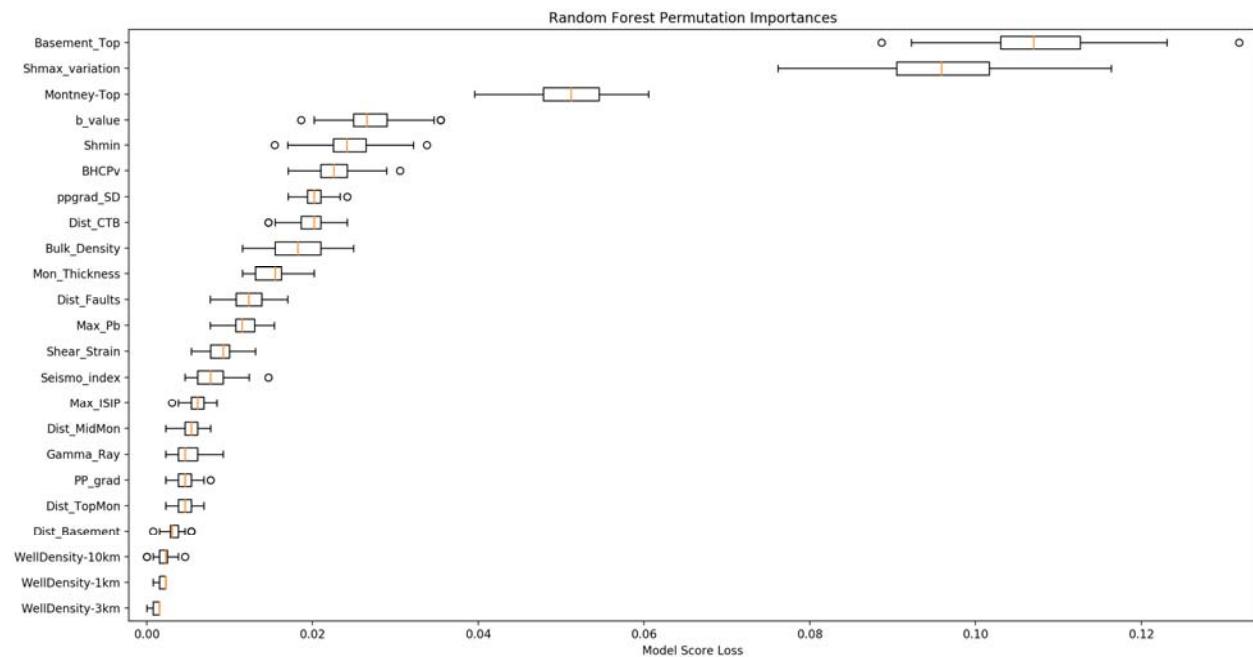


Figure 35 – Permutation importance plot for the Random Forest model with respect to induced seismicity susceptibility.

6. Induced Seismicity Susceptibility Maps

6.1. Cut-Off Magnitude 1.1

Machine learning results from the top two performing classification models, Light GBM and Random Forest, were used to generate induced seismicity susceptibility maps for the Montney region. This involved applying the tuned and tested models to the Montney Grid Dataset to calculate the probability of induced seismicity for each grid point entry in the dataset. This approach differs from the recent work of Wozniakowska and Eaton (2020) where probabilities were calculated at the well locations and then interpolated throughout the Montney region. As described in Section 2.4, the Montney Grid Dataset contains latitudes and longitudes on a 2.5 km grid spacing, for which the corresponding feature values are interpolated from the Well Susceptibility Dataset (Figure 3). The Montney Grid Dataset represents unseen data by the models.

It should be noted that the output of the classification model is the probability of a given grid point being seismogenic. No error is associated with the calculated probabilities, although this can be inferred from the standard deviations calculated for the cross-validation results. When the probabilities are calculated, the initial classification of seismogenic or non-seismogenic requires a user-defined probability threshold. For this study, we used the standard default probability threshold of 50%. This means that the model algorithm classifies a sample as being positive (i.e., seismogenic) if the probability of that sample being positive is greater than 50%, and classifies a sample as being negative if the probability of that sample being positive is less than 50%.

Figures 36 and 37 present the susceptibility maps generated from the Light GBM and Random Forest models, respectively. Both maps show the same induced seismicity hotspots, but differ in that areas with probabilities greater than 10% are more concentrated around these hotspots for the Light GBM susceptibility map and are more dispersed for the Random Forest susceptibility map. As previously discussed in Section 5.2, 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). Instead, information on induced seismicity severity derived from these maps is limited to that of a minimum threshold defined by the magnitude of completeness of the seismic catalogue used to train the machine learning models.

Accordingly, the susceptibility maps presented in Figures 36 and 37 show the probability of a location generating an induced seismicity event with a magnitude greater than 1.1.

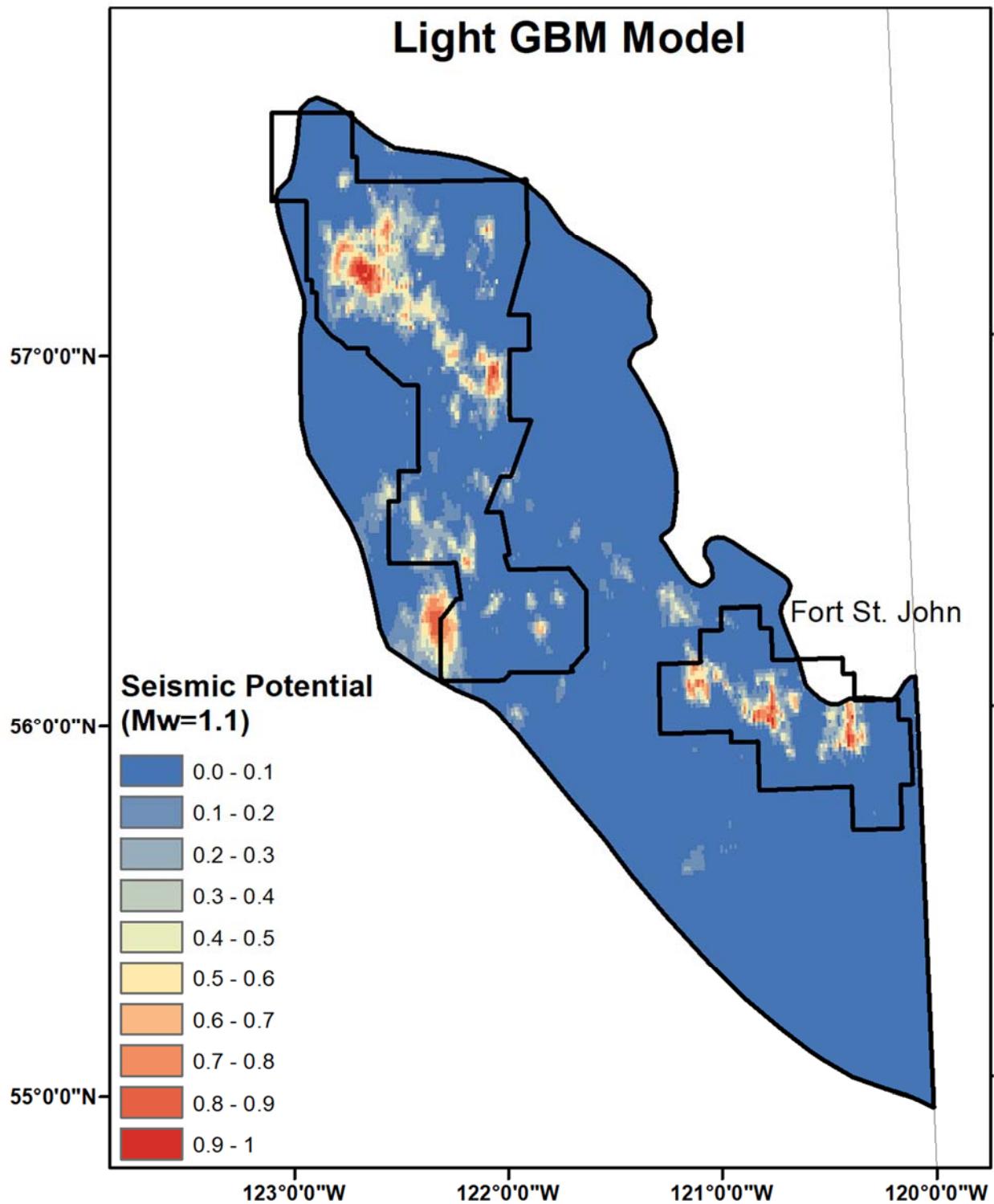


Figure 36 – Induced seismicity susceptibility map for the Montney region for events with a minimum magnitude of 1.1 based on results from the Light GBM machine learning model. Included are outlines of the North Peace Ground Motion Monitoring Area (NPGMMA) to the northwest, and Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) to the southeast.

Random Forest Model

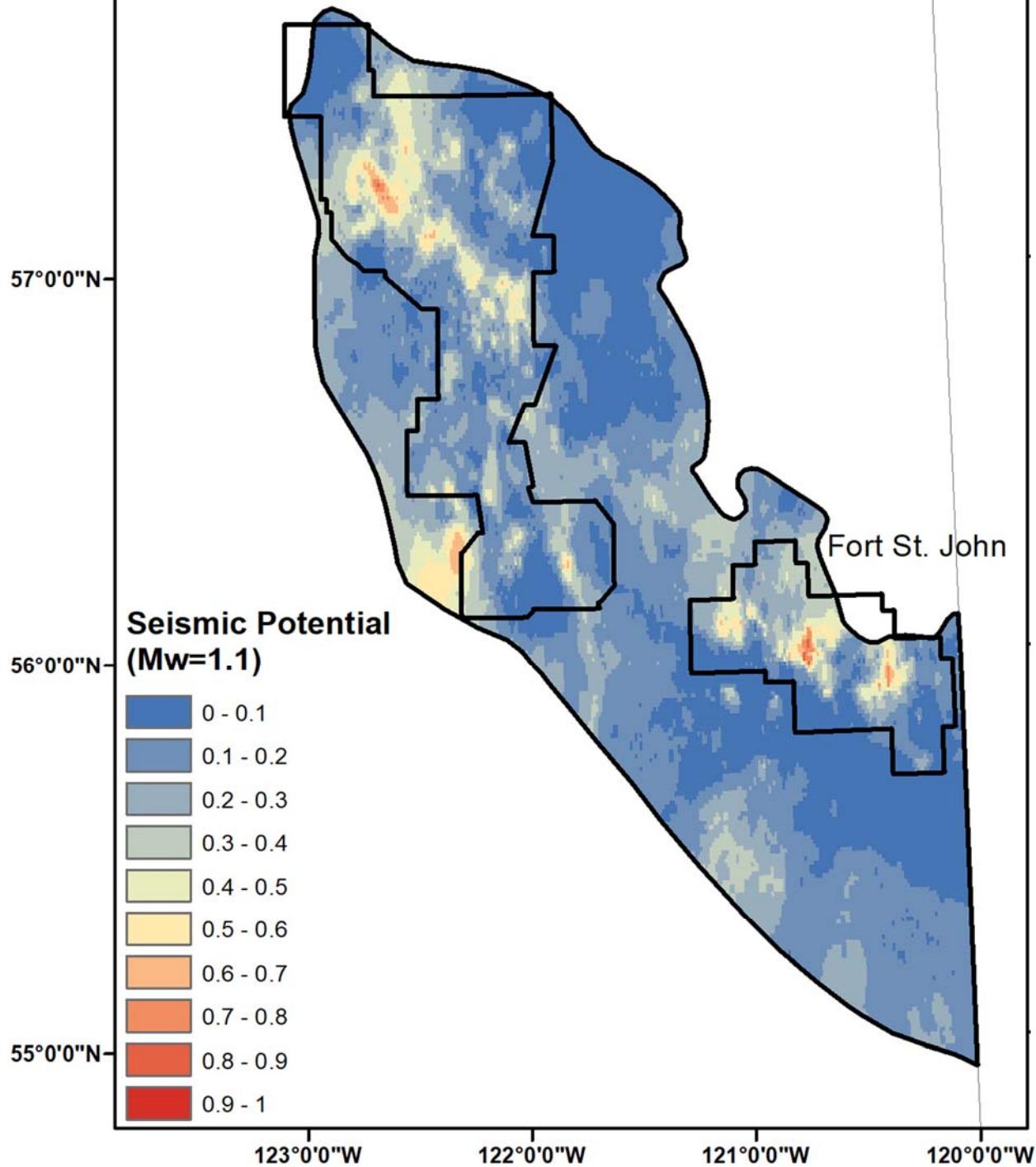


Figure 37 – Induced seismicity susceptibility map for the Montney region for events with a minimum magnitude of 1.1 based on results from the Random Forest machine learning model. Included are outlines of the North Peace Ground Motion Monitoring Area (NPGMMA) to the northwest, and Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) to the southeast.

To evaluate the accuracy of the generated induced seismicity susceptibility maps, Figure 38 reproduces these with the locations of historic events superimposed on top. This comparison shows that there is a good match between the locations of historic induced seismicity events and areas identified by the susceptibility maps as having high probabilities of seismic potential. For the Light GBM susceptibility map, there are a small number of isolated events that are located within areas identified as having low probabilities (0-10%). In contrast, because the probabilities in the Random Forest susceptibility map are more distributed, it shows these isolated events as falling within areas with higher probabilities.

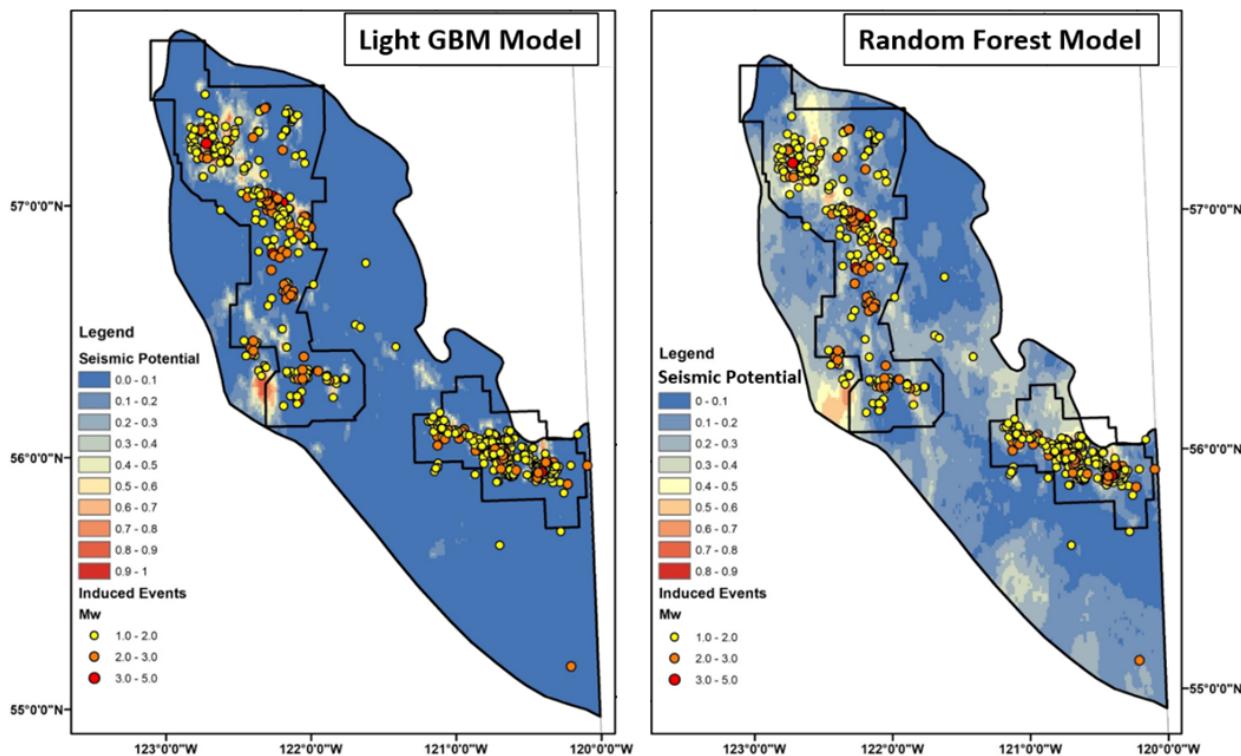


Figure 38 – Locations of historic induced seismicity events (greater than magnitude 1.1) superimposed on top of the susceptibility maps generated from the Light GBM and Random Forest classification models.

Figure 39 shows a second comparison, but this time superimposing the locations of all hydraulic fracturing wells in the study area. Two areas are highlighted, circled in yellow and orange, where there are a large number of wells but no or very few recorded induced seismicity events. Both models show these areas as having low probabilities of seismic potential. This agreement demonstrates that the models perform well with respect to both areas with low and high seismic potential probabilities.

Lastly, circled in red is an area that both models predict as having a high susceptibility to induced seismicity, which has seen a lower concentration of wells compared to other

concentrated areas of well development activity. If development begins to extend into this area, these can be used to further test the predictive performance of the machine learning models and corresponding induced seismicity susceptibility maps.

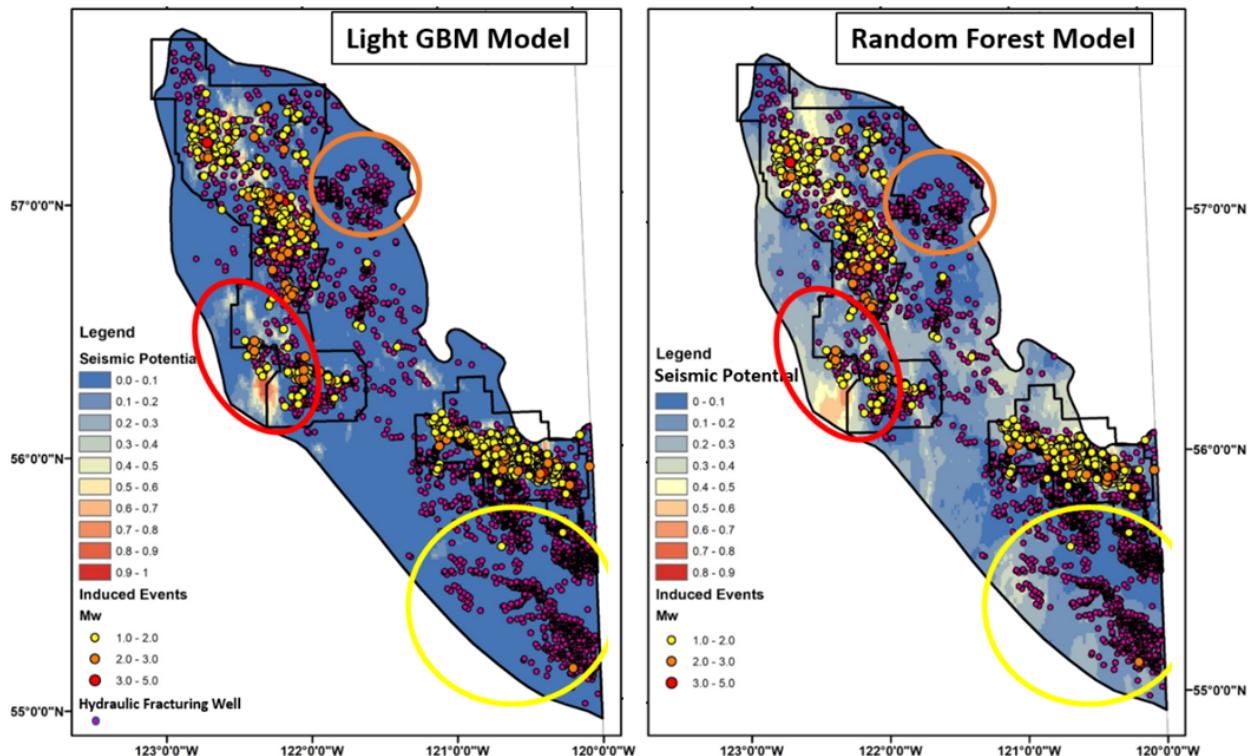


Figure 39 – Locations of existing hydraulic fracturing wells superimposed on top of the susceptibility maps generated from the Light GBM and Random Forest machine learning models. Areas highlighted by the yellow and orange circles are those with a high population of wells but few recorded induced seismicity events (with a magnitude greater than 1.1). The area circled in red is one that both models show to be susceptible to induced seismicity but where there are fewer drilled wells.

6.2. Cut-Off Magnitude 2.0

The induced seismicity susceptibility maps presented in the previous section were developed to maximize the data available. In general, machine learning models improve their predictive performance with increasing dataset size. By maximizing the data available, specifically that in the seismic event catalogue, the resulting maps indicate the seismic potential relative to the magnitude of completeness of the seismic event catalogue used to train the models. In this case, the magnitude of completeness of the full dataset was 1.1, and therefore the maps show the spatial probability of an induced seismicity event with a minimum moment magnitude (M_w) of 1.1.

It is also recognized that there might be other minimum event thresholds that are of interest. For example, focus might be placed on large events with a minimum magnitude of 4.0, as this is the threshold used in the B.C. Drilling and Production Regulation (DPR, 2021; section 21.1) when hydraulic fracturing and fluid injection activities must be suspended. However, this results in a dataset that is much too small to conduct a machine learning analysis. Alternatively, a minimum threshold of 2.0 is used by the Alberta Energy Regulator (AER, 2015) as part of their traffic light protocol for the Duvernay, where requirements switch from a green light and normal operations to a yellow light and the need to report and mitigate. This offers a possible compromise.

The key impact that increasing the cut-off magnitude has on the machine learning model is that the number of seismogenic wells in the dataset decreases, thereby resulting in the data becoming more imbalanced. Otherwise, the steps to prepare and process the data are exactly the same as before, as are the steps to train, validate, and test the models. Figure 40 plots the receiver operating characteristic (ROC) curves for the different machine learning classification models being compared, relative to their performance using the susceptibility-focussed dataset (see Section 5.2). The results indicate that the Random Forest, XGBoost and Light GBM models performed similarly, and significantly better than the other models.

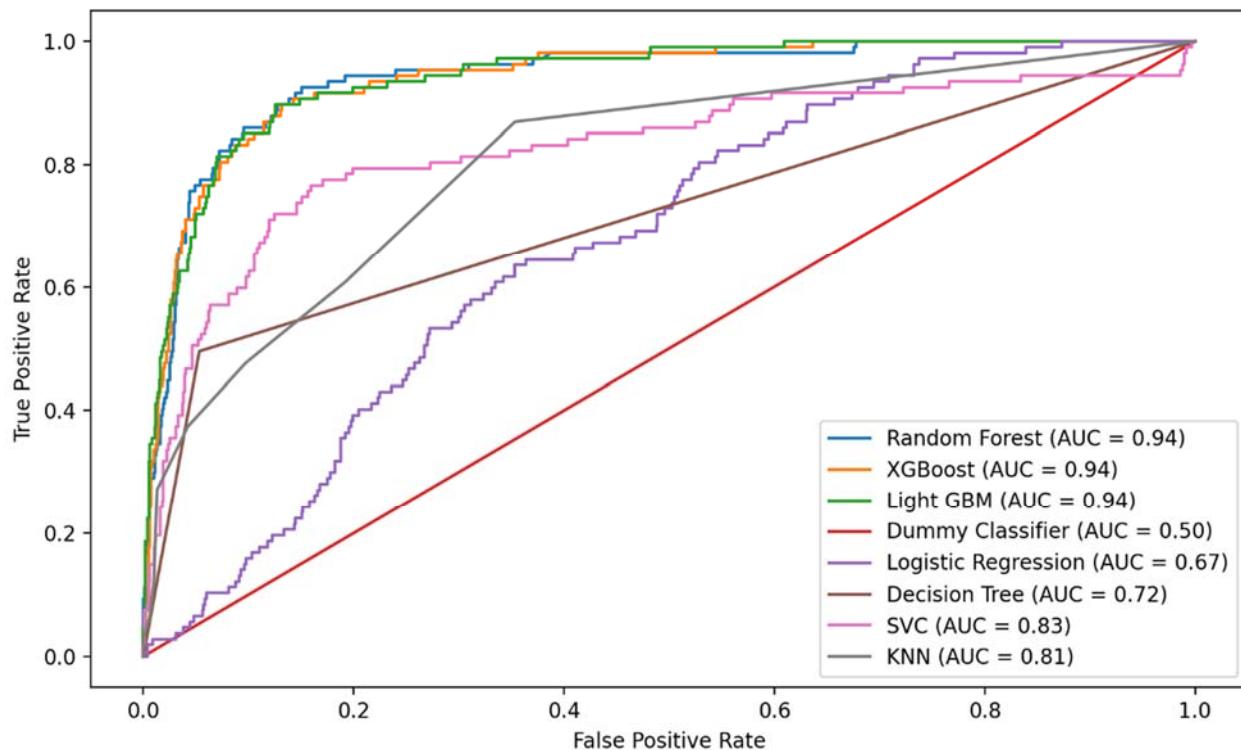


Figure 40 – Receiver operating characteristic (ROC) curves for the classifier models tested against the susceptibility-focussed dataset with a minimum event cut-off of magnitude 2.0.

In order to allow comparison with the previous susceptibility results (with a cut-off magnitude of 1.1), the Light GBM and Random Forest models were carried forward and applied to the Montney grid dataset. Table 7 shows the top five important features identified by these models, together with those for the magnitude 1.1 cut-off for comparison. The results show that by filtering out induced seismicity events with magnitudes less than 2.0, the models' sensitivity to different features has changed. Although features like depth to the top of the basement are still ranked as being highly important, other features like the distance of the well from the Cordilleran thrust belt increase in importance. Other features like azimuth variation between the local and regional maximum horizontal stress were seen to drop from the top five, whereas other features like the regional shear strain, standard deviation of the pore pressure gradient and average bulk density of the Montney at the well location, rise into the top five important features. The importance of distance from the Cordilleran thrust belt and regional shear strains possibly points to their association with stored tectonic strain energy accumulated, and therefore energy release potential and larger event magnitudes. The importance of standard deviation of the pore pressure gradient possibly points to zones of overpressure contributing to larger magnitude events. The importance of features relative to induced seismicity event magnitude is investigated more explicitly through a machine learning regression analysis in Section 7.

Table 7 – Top five ranked features with respect to importance in predicting induced seismicity susceptibility, for datasets limited to a minimum moment magnitude (M_w) of 1.1 and 2.0, for each of the top two performing classifier models.

Rank	Light GBM ($M_w = 2.0$)	Light GBM ($M_w = 1.1$)	Random Forest ($M_w = 2.0$)	Random Forest ($M_w = 1.1$)
1	Depth to the top of basement	Depth to the top of basement	Distance of well from Cordilleran thrust belt	Depth to the top of basement
2	Minimum horizontal stress magnitude	Minimum horizontal stress magnitude	Depth to the top of basement	Variation of local Shmax from regional orientation
3	Distance of well from Cordilleran thrust belt	Variation of local Shmax from regional orientation	Minimum horizontal stress magnitude	Minimum horizontal stress magnitude
4	Regional shear strain field	Depth to the top of the Montney	Depth to the top of the Montney	Distance of well from Cordilleran thrust belt
5	Standard deviation of pore pressure gradient	Distance of well from Cordilleran thrust belt	Average bulk density of Montney at well location	Depth to the top of the Montney

Figures 41 and 42 present the susceptibility maps of induced seismicity potential for event magnitudes 2.0 and greater as generated from the Light GBM and Random Forest models, respectively. Figure 43 provides a side-by-side comparison of the maps for both models against the corresponding maps derived from the magnitude 1.1 cut-off dataset analyses.

As would be expected, increasing the cut-off magnitude to 2.0 reduces the extent and probability magnitudes of seismic potential across both maps. Comparing the two Light GBM susceptibility maps, similar hotspots are seen (compare Figure 43 a with b). These have similarly high probabilities (up to 80-90% in hotspot locations), but the extent of the higher probability areas surrounding them are smaller for the magnitude 2.0 cut-off map. Another key difference is a new hotspot area that appears in the southwestern part of the Montney for the magnitude 2.0 cut-off map (circled in red in Figure 43b). The magnitude 1.1 cut-off map had previously identified this area as having a low probability of seismic potential. However, the training of the Light GBM model against the magnitude 2.0 cut-off dataset results in slightly different key features being emphasized as important, for example distance from the Cordilleran thrust belt and regional shear strain field. These then combine to indicate an additional area of high seismic potential with probabilities of up to 70% of generating an induced seismicity event of magnitude 2.0 or greater.

The Random Forest susceptibility maps likewise show that areas indicating a higher probability of seismic potential are reduced in extent for the magnitude 2.0 cut-off map compared to the magnitude 1.1 cut-off map (compare Figure 43 d with c). In addition, the probability magnitudes for the hotspot locations are lower. The highest probabilities in the hot spot locations for the magnitude 1.1 cut-off analysis were in the range of 80-100% (Figure 43c) compared to those in the magnitude 2.0 cut-off analysis, which are in the range of 50-70% (Figure 43d).

Overall, these results demonstrate the sensitivity of the induced seismicity susceptibility maps generated to the machine learning algorithms used, as well as to the minimum event magnitude cut-off threshold used. However, they also demonstrate the robustness of the map predictions as both sets of models identify similar hotspots with similar probability magnitude distribution trends; the differences seen are generally small and localized relative to the accuracy and precision that can be expected from a model and its output. The results point to the value of using two or three top performing models in parallel, as well as best practices in which the different induced seismicity susceptibility maps generated are considered together to inform decision making. The performance of these models and map products need to be periodically updated and tested as hydraulic fracturing operations in the Montney continue.

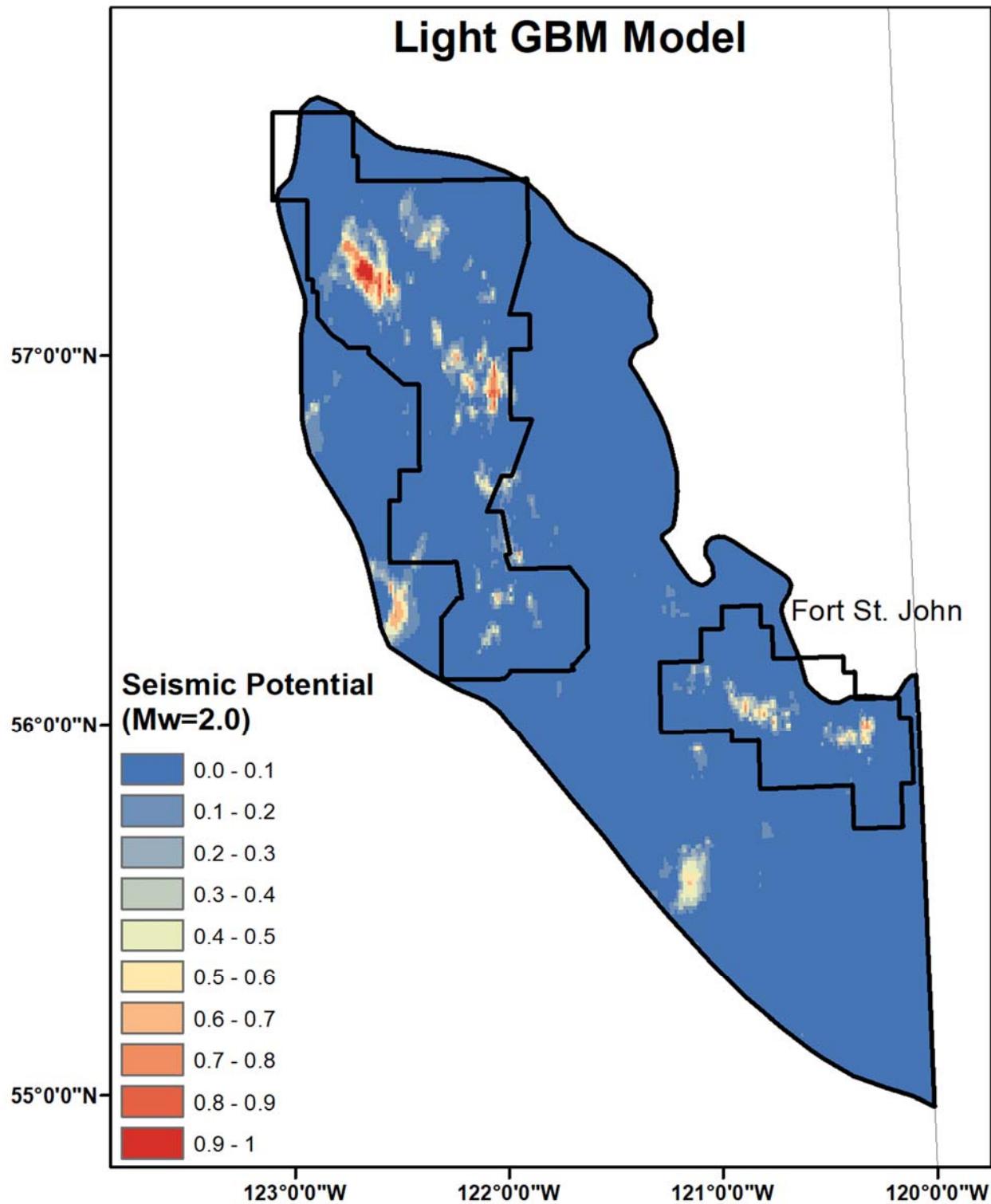


Figure 41 – Induced seismicity susceptibility map for the Montney region for events with a minimum magnitude of 2.0 based on results from the Light GBM machine learning model. Included are outlines of the North Peace Ground Motion Monitoring Area (NPGMMA) to the northwest, and Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) to the southeast.

Random Forest Model

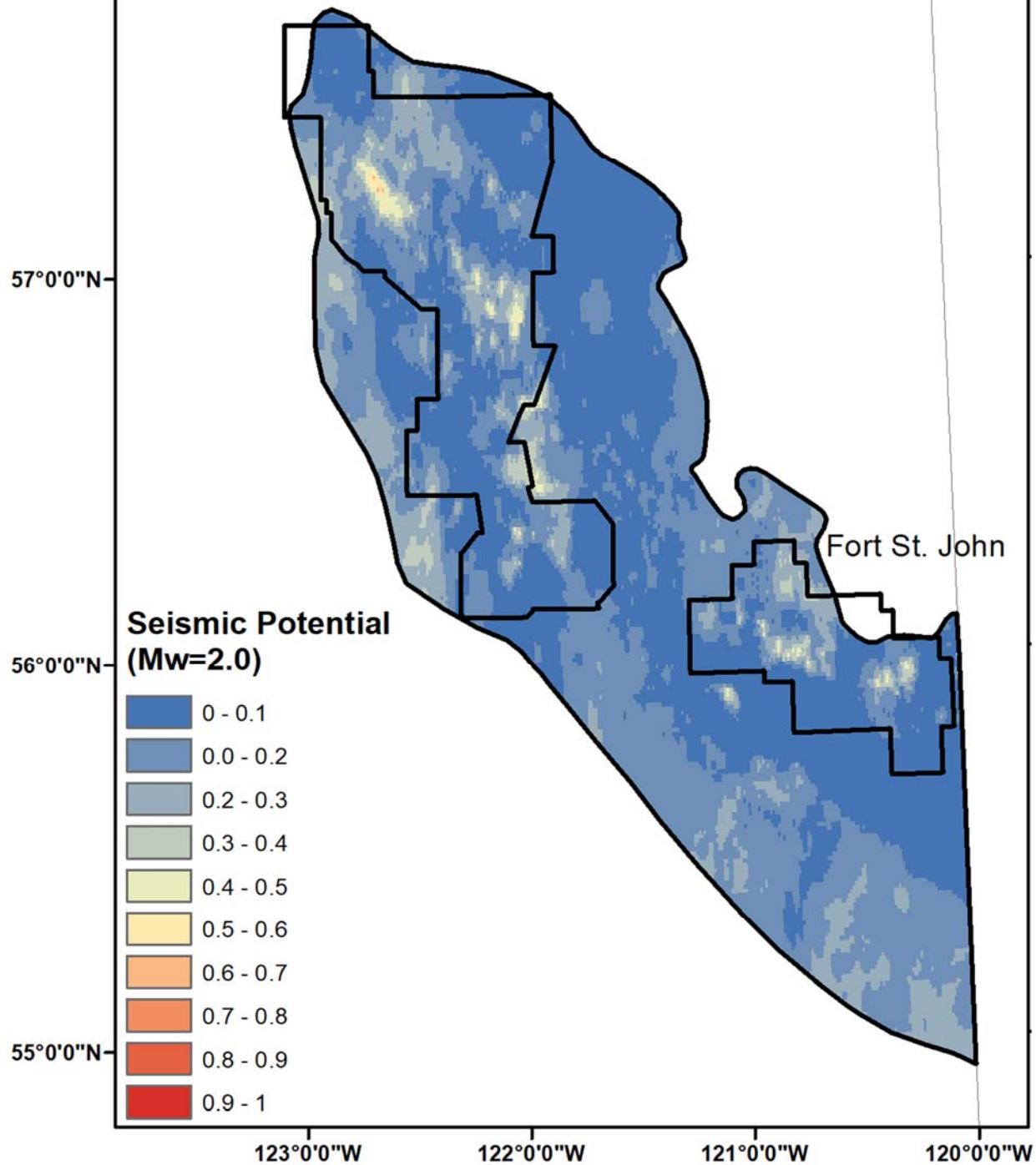


Figure 42 – Induced seismicity susceptibility map for the Montney region for events with a minimum magnitude of 2.0 based on results from the Random Forest machine learning model. Included are outlines of the North Peace Ground Motion Monitoring Area (NPGMMA) to the northwest, and Kiskatinaw Seismic Monitoring and Mitigation Area (KSMMA) to the southeast.

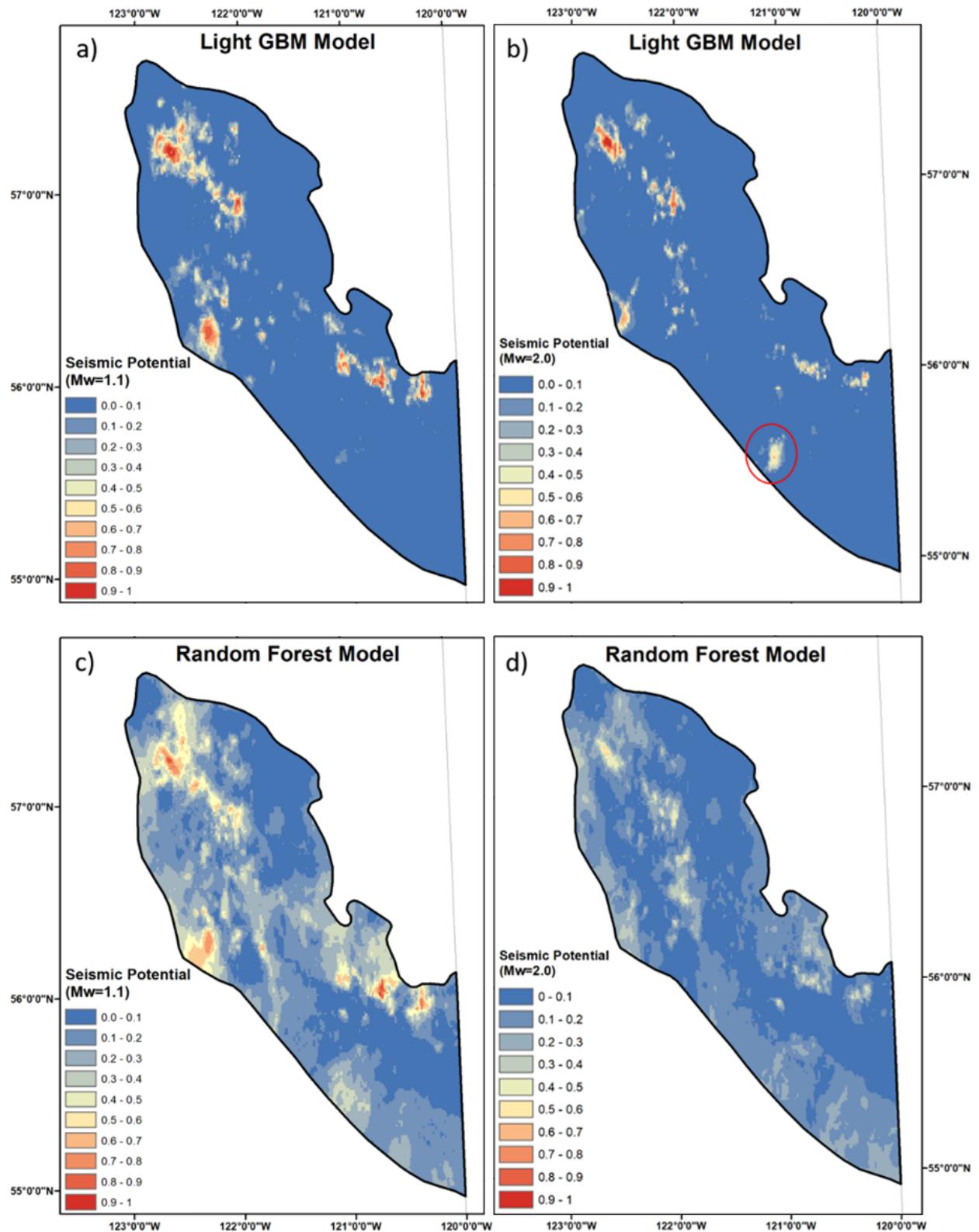


Figure 43 – Comparison of induced seismicity susceptibility maps based on results from the Light GBM model for a minimum magnitude cut-off of: a) 1.1 and b) 2.0, and from the Random Forest model for a minimum magnitude cut-off of: c) 1.1 and d) 2.0.

6.3. Use and Limitations

Below are several key points regarding the use of the induced seismicity susceptibility maps generated, including considerations that affect their accuracy and/or confidence in their use.

- The generated maps are raster maps with a 2.5 km grid spacing covering the Montney region. Each raster pixel reports the calculated spatial probability of a well at that location triggering an induced seismicity event greater than the cut-off magnitude (1.1 or 2.0, as specified). This information can be used towards planning, hazard and risk analyses, and future decision making, as pixels with higher probabilities indicate areas with a higher likelihood of triggering induced seismicity.
- The resolution of the generated raster maps (i.e., 2.5 km) should be taken into consideration when using these maps. Any interpolation of the results at a smaller resolution is beyond the accuracy of this work and should not be relied on.
- The induced seismicity susceptibility maps are intended to serve as a screening tool and should not be relied upon in the absence of other information for absolute prediction. Indications of relatively lower or higher susceptibility should be used as one piece of information to help support judgement and decision making as to the level of effort needed for further work, and not as the sole basis for a decision about activities.
- The influence of other related activities (e.g., wastewater injection) have not been considered in developing the induced seismicity susceptibility maps. The effects of such activities would require further site-specific analysis.
- The generated maps are only as reliable as the data used. As more data with higher quality is collected over time, these maps should be updated.

7. Severity Analysis

Building on the previous classification-based susceptibility analyses, the question of induced seismicity severity potential was further investigated employing machine learning regression analysis techniques (e.g., Figure 4). The workflow used, relative to that for the susceptibility analysis, is provided in Figure 7. Several machine learning models were trained, tested and their performances evaluated with respect to estimating the maximum magnitude of the induced seismicity events associated with each hydraulic fracturing well. For this analysis, the Induced Seismicity Severity Dataset was used (see Section 2.4). The Mean Square Error (MSE) and Mean Absolute Percent Error (MAPE) were used as the key metrics for evaluating model performance and hyperparameter tuning of the regression models.

7.1. Model Training, Tuning and Comparison Testing

Table 8 reports the cross validation results for the maximum magnitude regression models after training and tuning. Between the eight regression models tested, the Light GBM and Random Forest models were seen to perform the best based on having the lowest MSE and MAPE validation scores, while having high R^2 values. The MSE and MAPE are calculated from the difference between the actual and predicted values, and therefore lower values indicate a better model performance. Table 8 shows that the XGBoost model also has low MAE and MAPE values, but the zero values during training and r^2 of 1.000 warns that the model is overfitting.

Table 8 – Cross validation results of the maximum magnitude regression models. Values given are the average over 50-fold iterations; values in parentheses are standard deviations.

Model	Training MSE	Validation MSE	Training r^2	Validation r^2	Training MAPE	Validation MAPE
Dummy Regressor	0.187 (+/- 0.002)	0.188 (+/- 0.091)	0.000 (+/- 0.000)	-0.140 (+/- 0.207)	0.168 (+/- 0.001)	0.168 (+/- 0.035)
Linear Regression	0.184 (+/- 0.093)	0.203 (+/- 0.130)	0.016 (+/- 0.494)	-0.330 (+/- 0.828)	0.165 (+/- 0.034)	0.173 (+/- 0.053)
Decision Tree	0.069 (+/- 0.005)	0.122 (+/- 0.099)	0.630 (+/- 0.029)	0.188 (+/- 0.713)	0.094 (+/- 0.003)	0.118 (+/- 0.046)
Random Forest	0.013 (+/- 0.001)	0.090 (+/- 0.061)	0.933 (+/- 0.003)	0.445 (+/- 0.345)	0.037 (+/- 0.001)	0.099 (+/- 0.035)
SVR	0.008 (+/- 0.000)	0.155 (+/- 0.087)	0.955 (+/- 0.000)	0.068 (+/- 0.282)	0.045 (+/- 0.000)	0.150 (+/- 0.037)
KNN	0.103 (+/- 0.001)	0.146 (+/- 0.081)	0.450 (+/- 0.007)	0.100 (+/- 0.383)	0.116 (+/- 0.001)	0.140 (+/- 0.039)
XGBoost	0.000 (+/- 0.000)	0.134 (+/- 0.120)	1.000 (+/- 0.000)	0.124 (+/- 0.813)	0.000 (+/- 0.000)	0.112 (+/- 0.053)
Light GBM	0.002 (+/- 0.000)	0.087 (+/- 0.078)	0.988 (+/- 0.001)	0.458 (+/- 0.450)	0.013 (+/- 0.001)	0.097 (+/- 0.038)

Based on the cross-validation results, the performance of the Light GBM and Random Forest models were next tested against the test data split, which represents unseen data (e.g., Figure 5). Table 9 reports the test results and indicates that both models performed similarly. The test scores were also similar to the cross-validation scores for these models (Table 8), and thus both were selected for detailed interpretation.

Table 9 – Test results comparing the top two performing regression models against the unseen test data.

Model	Test MSE	Test r^2	Test MAPE
Light GBM	0.11	0.53	0.118
Random Forest	0.18	0.25	0.137

7.2. Model Interpretation

Feature Importance Plots

The features considered in the maximum magnitude regression analysis included all geological and operational features, similar to the feature importance classification analysis presented in Section 5.1. The key difference is that the Induced Seismicity Severity Dataset used for the regression analysis (the Well Features Dataset was used previously), only includes the wells that were classified as being seismogenic. For each of these, the maximum magnitude of the associated induced seismicity events is ascribed (Section 2.4). Figures 44 and 45 show the resulting feature importance plots with respect to maximum induced seismicity event magnitude for the Light GBM and Random Forest regression algorithms, respectively. Table 10 reports the top five important features for each of these models.

Both models show agreement that the two most important features with respect to predicting maximum event magnitude are the b-value of induced seismicity events (within 100 km of the well) and the standard deviation of the pore pressure gradient. The relationship between b-value and induced seismicity magnitude has been investigated by Amini and Eberhardt (2019), who showed that the tectonic stress regime influences the magnitude distribution of induced seismicity events and therefore the b-value. Interestingly, the distance of a well from the Cordilleran thrust belt ranks in the top five important features for both models, as do other stress regime indicators like azimuth difference between the local and regional maximum horizontal stress (for the Light GBM regression model) and minimum horizontal stress (for the Random Forest regression model). The importance of pore pressure gradient is possibly related to pore pressure compartmentalization observed in the Montney Formation (e.g., Fox & Watson, 2019). However, its influence on event magnitude still needs to be investigated. With respect to operational features, total volume is no longer the most important feature as was the case in the classification analysis, and instead the number of active days for completions

and maximum injection pressure were seen to be the most important operational features with respect to induced seismicity severity.

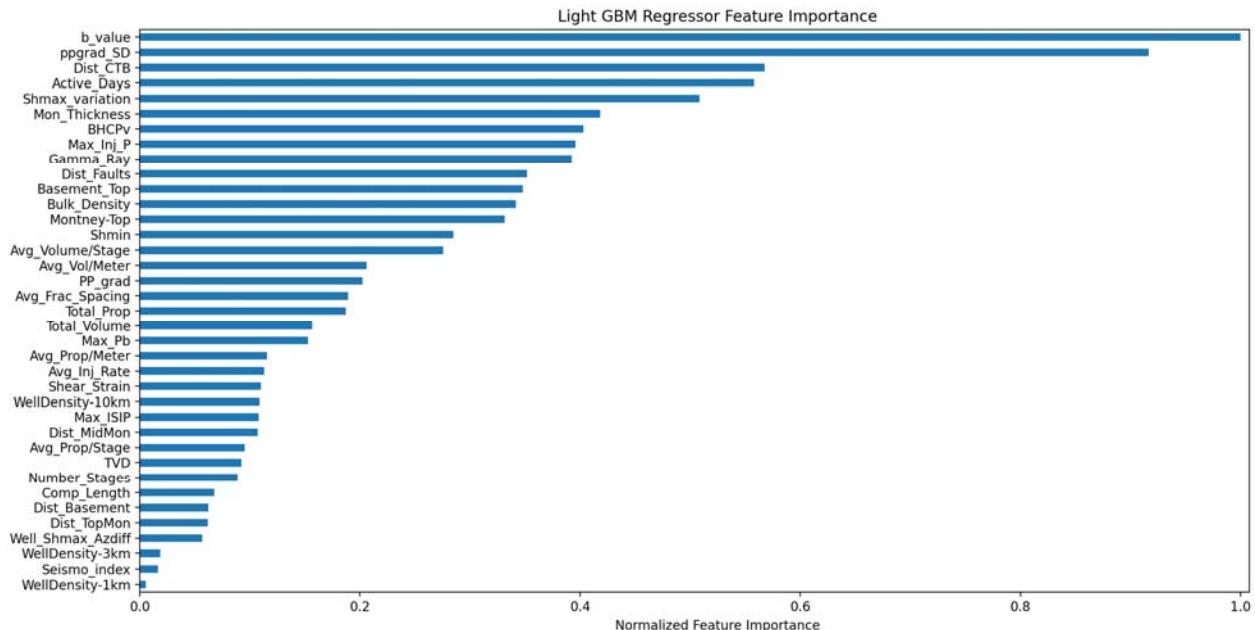


Figure 44 – Feature importance results for the Light GBM regression model with respect to maximum induced seismicity event magnitude.

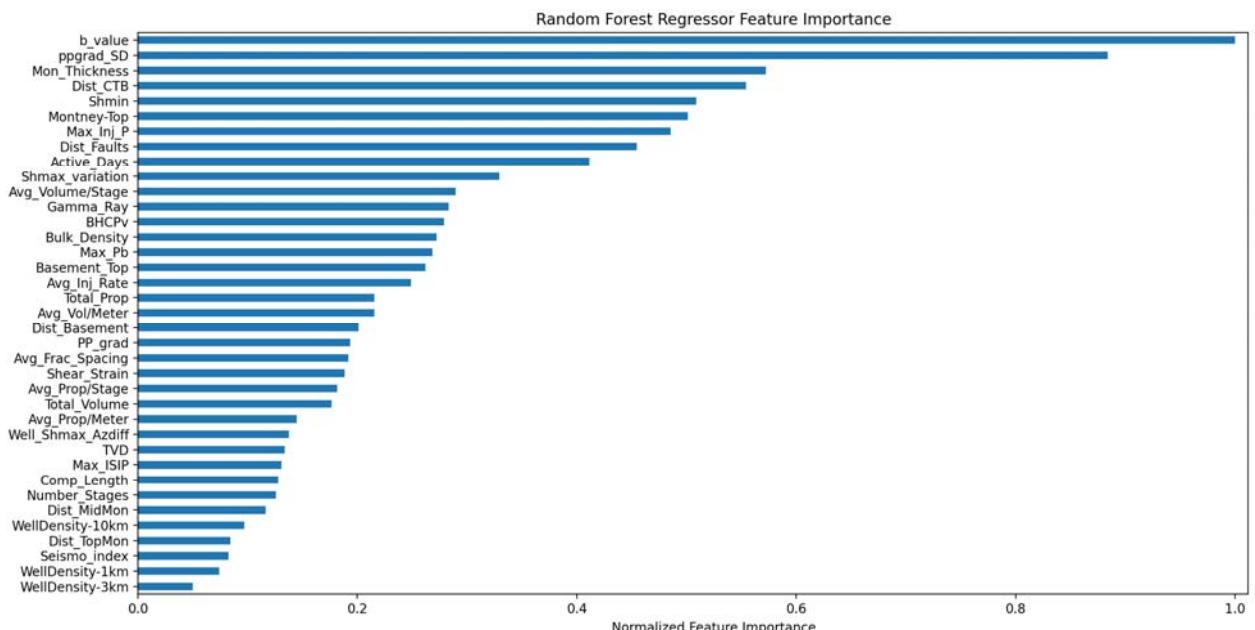


Figure 45 – Feature importance results for the Random Forest regression model with respect to maximum induced seismicity event magnitude.

Table 10 – Top five ranked features with respect to importance to maximum induced seismicity event magnitude, for each of the top performing regression models.

Rank	Light GBM	Random Forest
1	b-value of events (within 100 km of well)	b-value of events (within 100 km of well)
2	Standard deviation of pore pressure gradient	Standard deviation of pore pressure gradient
3	Distance of well from Cordilleran thrust belt	Thickness of Montney formation
4	Number of active completion days	Distance of well from Cordilleran thrust belt
5	Variation of local Shmax from regional orientation	Minimum horizontal stress magnitude

SHAP Summary Plots

Figures 46 and 47 present the summary plots from the SHAP analysis for the two top performing regression models. Based on these plots, we can see that the b-value ranks highest in terms of importance and has a positive correlation with maximum event magnitude. The higher values in this case correspond to a strike-slip stress regime, suggesting that larger magnitude induced seismicity events correlate more strongly with strike-slip stress conditions. (Note that the lower values for this feature correspond with a reverse fault stress regime, and these have less impact on the regression model predictions of maximum event magnitude). Similarly, the azimuth difference between the local and regional maximum horizontal stress orientation for the Light GBM regression model is ranked as important and shows a positive correlation. This does not necessarily mean that higher variations correlate directly with higher magnitude events since the maximum range of variation in the dataset is only 10 degrees. However, together, both features indicate the importance of the stress regime and that depending on the relative fault orientation, some faults become more critically stressed than others. This finding is explored in more detail in the next section (Section 8).

Another important feature ranked highly by both regression models is the standard deviation of pore pressure gradient. This is seen to have a positive correlation, although the results are somewhat more ambiguous for the Light GBM regression model, which shows high values for this feature having both positive and negative impacts on the model's output. This response might be an indication of the dependency of fault slip (and large versus small magnitude events) on the effective stress state and its heterogeneity, in combination with its sensitivity to the fault orientation (i.e., how critically stressed the fault might be or not). A similar bi-modal

response is seen for the distance of the well from the Cordilleran thrust belt; smaller values have both a negative and positive impact on the model predictions of severity. Again, this is specifically with respect to event magnitudes. For wells that are closer to the Cordilleran thrust belt, the results indicate a higher induced seismicity maximum magnitude, perhaps related to more faults being present within the deformation limits of the Cordillera. However, this feature does not specify the size or orientation of the faults in proximity to the well, and therefore both small and large magnitude events can be expected depending on the orientation of the fault(s) the injected fluids are interacting with.

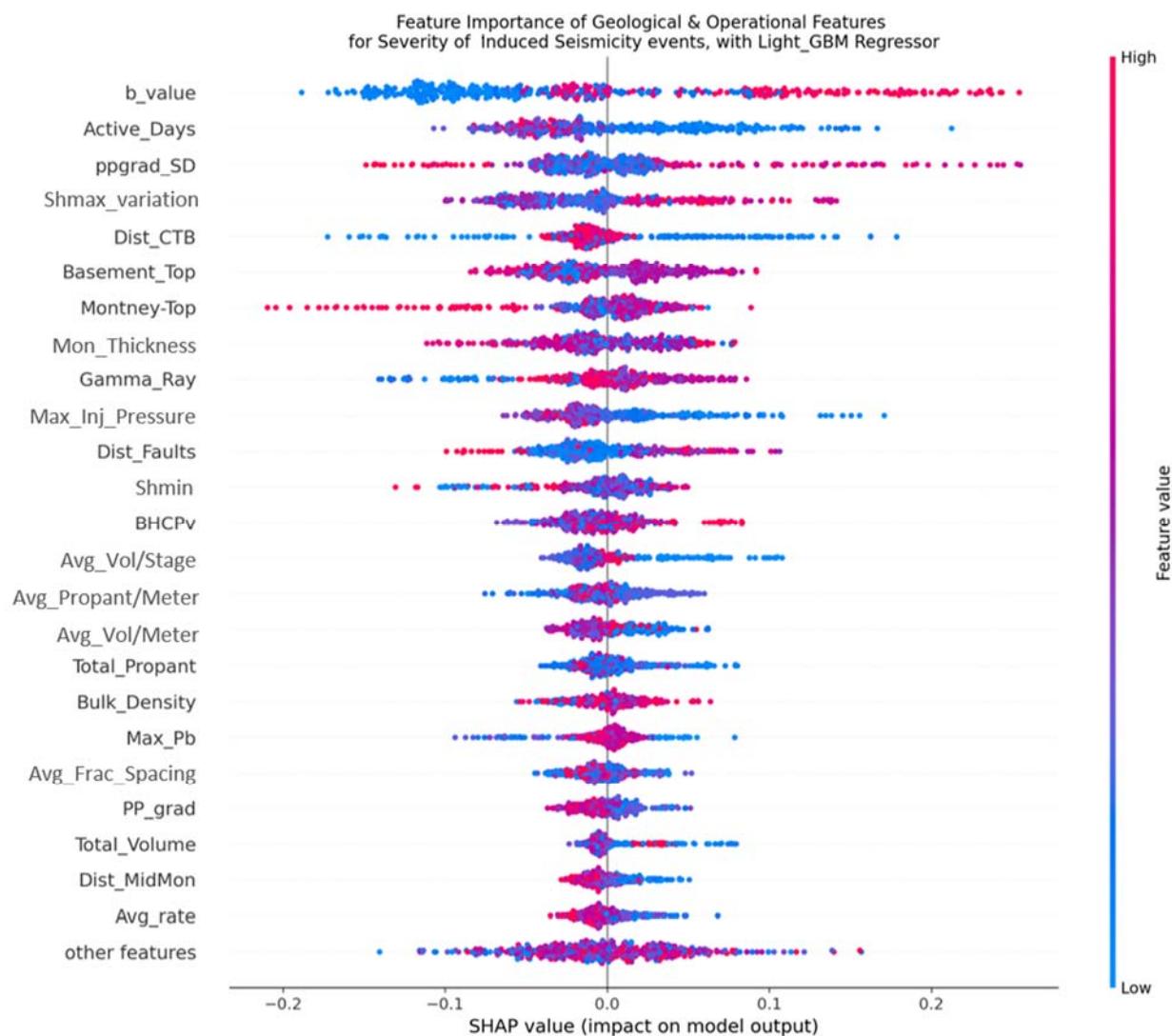


Figure 46 – SHAP summary plot for the Light GBM regression model with respect maximum induced seismicity event magnitude.

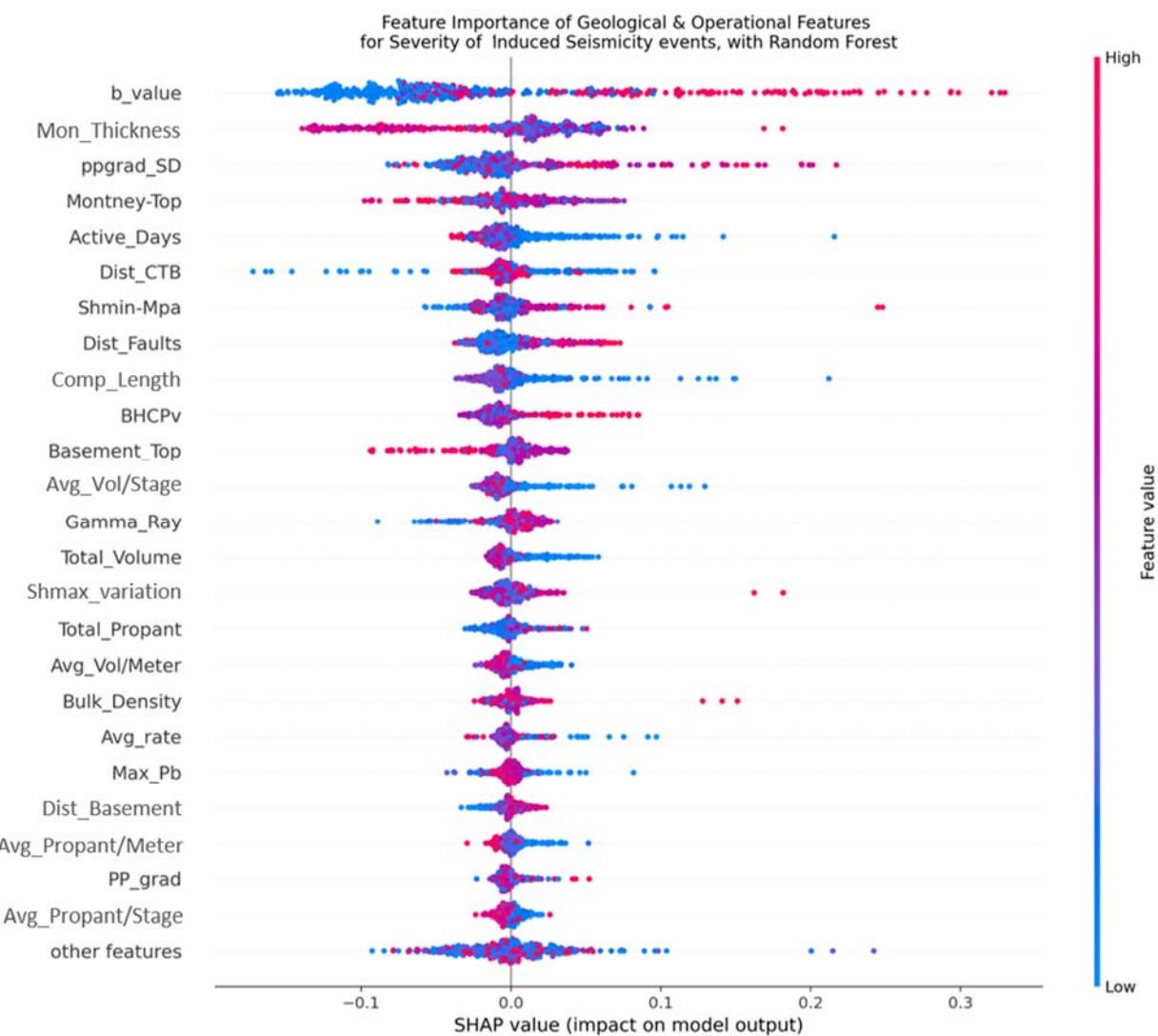


Figure 47 – SHAP summary plot for the Random Forest regression model with respect maximum induced seismicity event magnitude.

8. Mechanistic Analysis

To support the machine learning results, a series of numerical simulations were performed to demonstrate their value in investigating the mechanistic cause and effect relationships behind the patterns and correlations uncovered. As an investigative tool, numerical modelling allows the effects of different parameters to be tested for their sensitivities and interactions with one another. The Itasca Consulting Group's 3-D hydraulic fracture simulator XSite™ (version 3.0) was used for this purpose.

The models generated are conceptual, but were guided and constrained by reported reservoir characteristics for the Montney in British Columbia. They are designed to be generic to study the mechanistic cause of induced seismicity due to fluid injection and are not meant to be predictive for any specific site location or set of conditions. By adopting a simpler representation involving a single planar fault, focus was placed on capturing the physics of fluid pressure perturbation and stress transfer as triggering mechanisms.

8.1. Numerical Method Overview

XSite™ is a 3-D hydraulic fracturing numerical simulation program that employs a fully-coupled hydro-mechanical bonded-lattice modelling method (Damjanac et al., 2016). This numerical approach is based on the Synthetic Rock Mass (SRM) concept by Pierce et al. (2007). The SRM has been developed as a more realistic means to model the mechanical behavior of a fractured rock mass, and is comprised of two components: i) the bonded particle model (BPM), which simulates the deformation and fracturing of intact rock (Potyondy & Cundall, 2004); and ii) the smooth joint model (SJM), which simulates slip and opening behavior along discontinuities (Damjanac et al., 2020). The SRM and its components are shown in Figure 48.

The 3-D lattice model utilizes bonded particle (Potyondy & Cundall, 2004) and discrete element (Cundall & Strack, 1979) principles, but simplifies these to gain computational efficiencies to deliver faster calculation speeds. The lattice is made up of a quasi-random 3-D assembly of nodes with mass connected by two sets of springs: one representing the shear and the other the normal contact stiffness (Figure 49). These add up to represent the elasticity of the rock mass. Microcracks are simulated in the model through the failure of springs; these cluster to represent the growth of larger 3-D nonplanar fractures. A fluid element is assigned to each newly generated microcrack and a fluid channel will connect adjacent fluid elements. This allows non-steady fluid flow and pressure to be implemented in both the pre-defined joints, for example that specified through a discrete fracture network (DFN), and any stress-induced fractures generated during time-stepping. The fracture fluid flow is fully coupled (i.e., two-way interaction) with mechanical deformation. The pore pressures in the joints act to load the surrounding solid model, whereas the deformation of the solid model causes pore pressure and aperture (i.e., permeability) changes in the discontinuities.

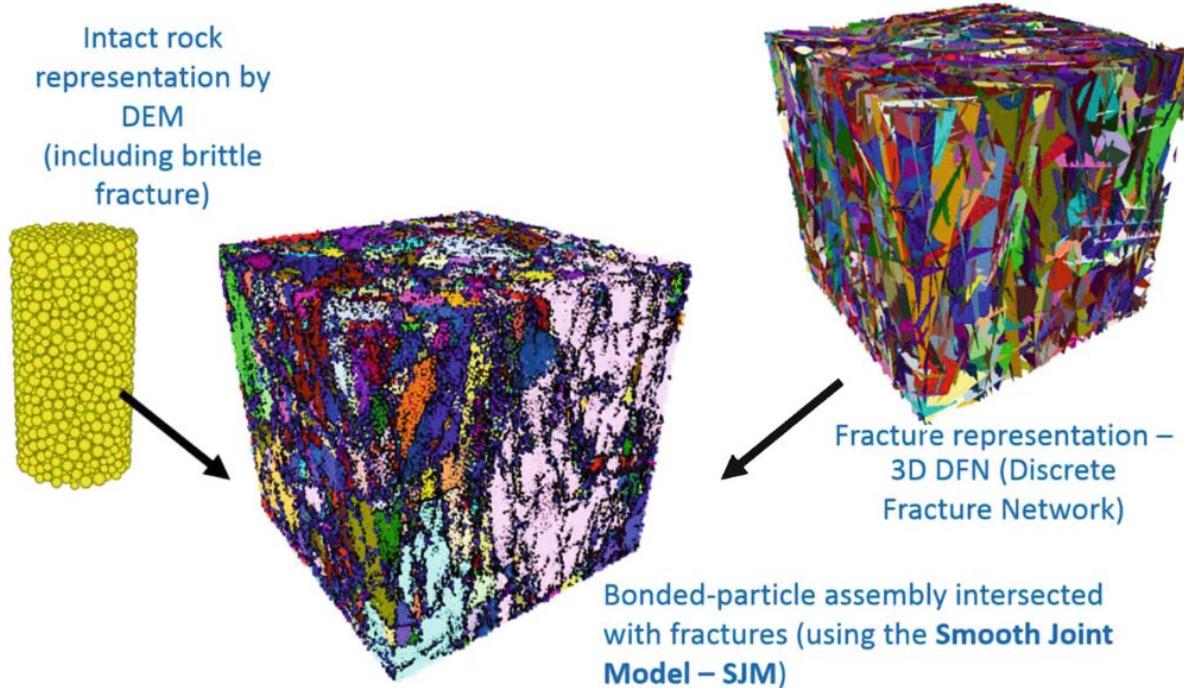


Figure 48 – Synthetic rock mass (SRM) method, which combines a bonded particle model (BPM) to simulate intact rock behavior, including the generation of stress or fluid pressure induced brittle fracturing, and a smooth joint model (SJM) to simulate discontinuity behaviour, including shear slip.

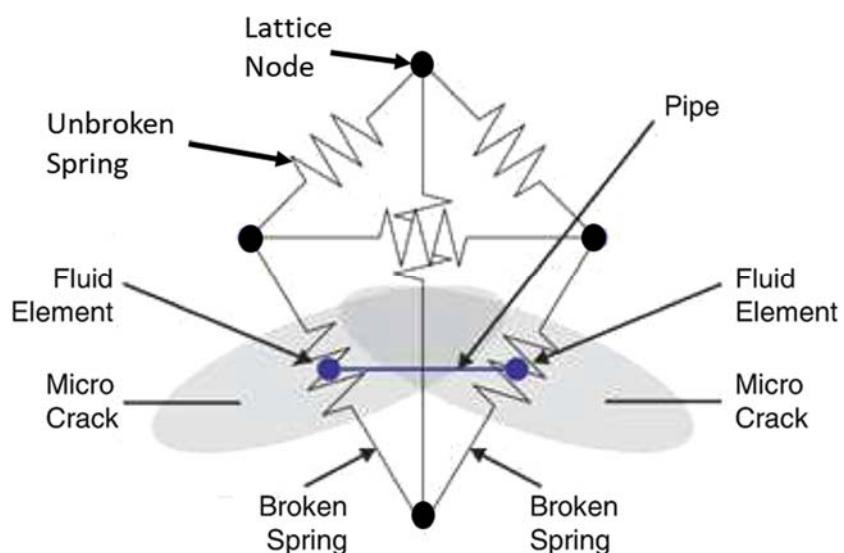


Figure 49 – Correlation between the 3-D lattice model and corresponding pipe network formed by the generation of microcracks (after Damjanac et al., 2016).

8.2. Model Setup and Inputs

The XSite™ model was setup to consider the general geology and typical reservoir depths present in the Montney region. Consideration was also given to studies suggesting that some of the largest induced seismicity events recorded for the Montney have occurred in the basement (e.g., Peña Castro et al., 2020). At the same time, it was noted that the vertical component of source location for these events suffers from poor detection resolution as well as use of a simple seismic velocity model, with updated seismic velocity models showing one of these events (the M_w 4.2 event on November 30, 2018) as occurring at the reservoir depth (personal communication, BCOGC). The model was thus developed to be adaptable. For this, a 3-D block domain was generated with dimensions of $1.6 \text{ km} \times 1.6 \text{ km} \times 1.5 \text{ km}$ (in the X, Y and Z directions, respectively). This model dimension permits the accommodation of a circular fault with a radius of 650 m (Figure 50). By applying different vertical stresses to the top of the model, different fault depths could be modelled.

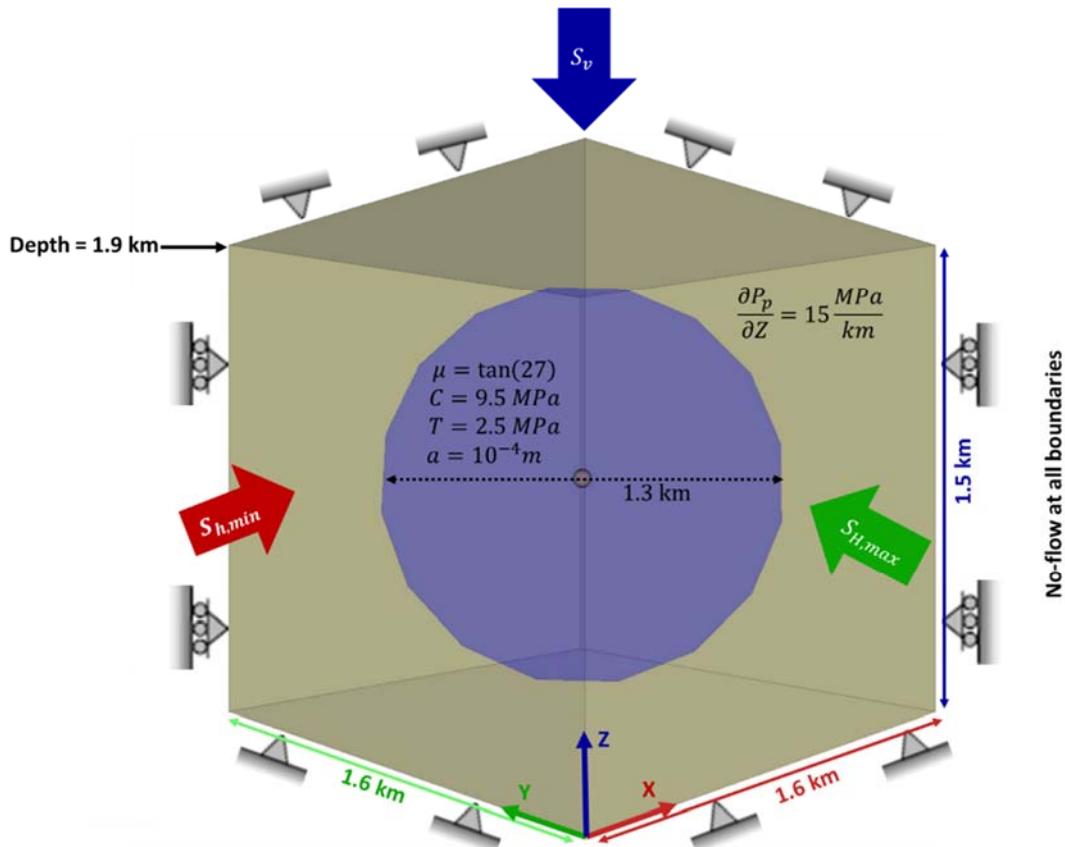


Figure 50 – Schematic of the base XSite™ model geometry and boundary conditions. The directions of the principal stresses are shown (note that each stress is parallel to the Cartesian axes of the model as indicated by the matching arrow and axes colours). The geometry, initial aperture (a), and frictional strength properties of the central circular fault are shown. The far-field displacement boundary conditions are fixed at the top and bottom faces, and free to move in the Z-direction along the side faces. All the boundaries are impermeable.

The fault orientation (i.e., dip and dip direction) was not fixed relative to a specific scenario but was varied in relation to the applied stress conditions as part of a parametric analysis. This was in recognition of the complex tectonic history of NEBC, which has contributed to the development of normal, strike-slip and reverse faults (i.e., faults with different dip angles and dip directions). For example, the main tectonic structure in the Montney region is the Fort St. John Graben (FSJG), where a significant number of the large magnitude events have been observed (Mahani et al., 2017). Although grabens are associated with extensional stress regimes and therefore normal faulting, the development of the FSJG was coupled with strike-slip motion and related compressional and rotational movement (Barclay et al., 1990). Evidence of fault orientations based on seismic reflection data is generally not available in the public domain, while the mapping of sub-vertical faults in reflection seismic surveys is challenging. Typically, these unmapped faults are more seismogenically problematic (personal communication with BCOGC).

With respect to present-day tectonics, stress inversions of induced seismicity focal mechanisms suggest that both reverse and strike-slip faulting stress regimes are present in the Montney (Mahani et al., 2017, 2019; Eaton et al., 2019; Peña Castro et al., 2020). For example, Eaton et al. (2019) analyzed the focal mechanisms of the 2018/11/30 earthquake sequence near Fort St. John, which include a M_w 4.6 event. Figure 51 shows the epicenters of events with $M > 1.5$ (both moment magnitude, M_w , and local magnitude, M_L), as well as the focal mechanism diagrams (beach balls) for the 12 largest events. The beach balls are coloured according to distinct clusters of events (i.e., events that occurred spatio-temporally close together). The primary event is displayed in green and represents a predominantly reverse fault mechanism, interpreted as a reverse reactivation of a normal fault that forms the southern boundary of the FSJG. Other events exhibit a diverse range of focal mechanisms, including strike-slip and low-angle thrust faulting. This diversity of induced seismicity source type further supports the presence of a complex fault system.

It should also be noted that for induced seismicity, the focal mechanism is not necessarily the same as the far-field stress regime; the focal mechanism is representative of the near-field stresses acting locally on the fault and the corresponding slip direction. As will be discussed later with the simulation results, this might be different from the far-field stress regime depending on the orientation of the fault experiencing the pressure perturbation from the fluid injection. Accordingly, reference to the modelled stress scenarios here differentiates between the far-field stress regime applied as a boundary condition to the model (simulating the tectonic condition) and the transformed stresses acting locally on the fault which relate to the focal mechanism when a fault slip event is triggered.

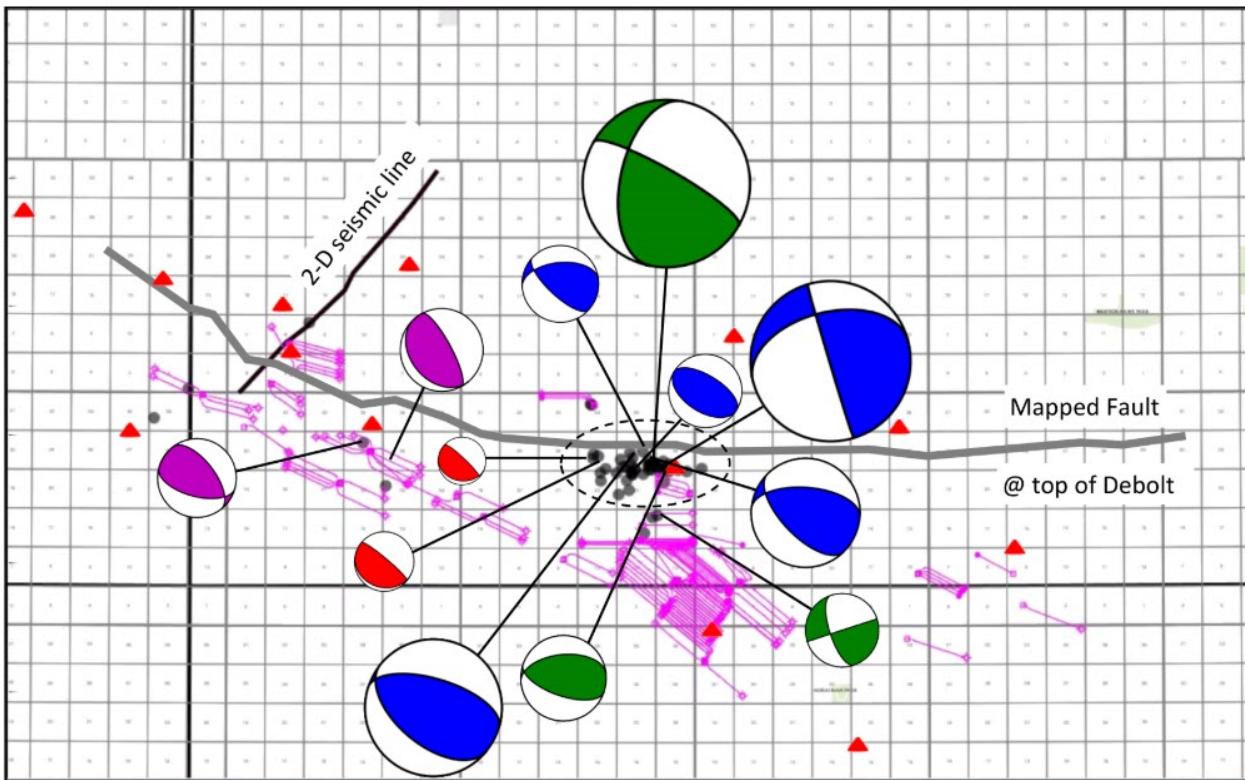


Figure 51 – Map of the 2018/11/30 earthquake sequence near Fort St. John, BC. The magenta lines represent the horizontal wells and laterals. The gray dots show the events recorded by the local seismograph array (red triangles). The beach ball diagrams show the focal mechanisms scaled by magnitude and coloured by event clusters (i.e., events that occurred spatiotemporally close together). Location of the southern bounding fault of the Fort St. John Graben (FSJG) is also shown in the map. The map grid spacing is 1.6 x 1.6 km. After Eaton et al. (2019).

Accordingly, the XSite™ model was set up to simulate both strike-slip and reverse fault far-field stress regimes as part of the parametric analyses performed. The estimation of the in-situ stresses was further guided by published well logging analyses (e.g., Rogers & Patrick, 2014; Rogers et al., 2015; Fox et al., 2019). Figure 52 shows the stress and fluid pressure profiles implemented as a function of depth. The three principal stresses are orthogonal. The vertical stress, fluid pressure, and the maximum horizontal stress gradients were kept the same for both stress regime models; only the minimum horizontal stress gradient was changed. By using the same vertical stress and fluid pressure gradients, the injection depths for both models remain the same, thus allowing results to be directly compared. The magnitude of the vertical stresses was easily determined using density log data together with the average densities for the overlying formations. The vertical stress was the minimum principal stress for the reverse far-field stress regime scenario ($S_H > S_h > S_V$). The magnitude and orientation of the minimum horizontal stress was determined primarily from minifrac test results. The minimum horizontal stress was the minimum principal stress for the strike-slip far-field stress regime scenario ($S_H > S_V > S_h$). The

maximum horizontal stress was determined by finding a value that would cause stress-induced borehole failure (or lack thereof) that matches observed failure in borehole image log data (e.g., FMI, CBIL, etc.). The orientation of the observed failure provides the azimuth of S_H (see Zoback, 2007; Fox et al., 2019). The gradients for each stress component are provided in Table 11.

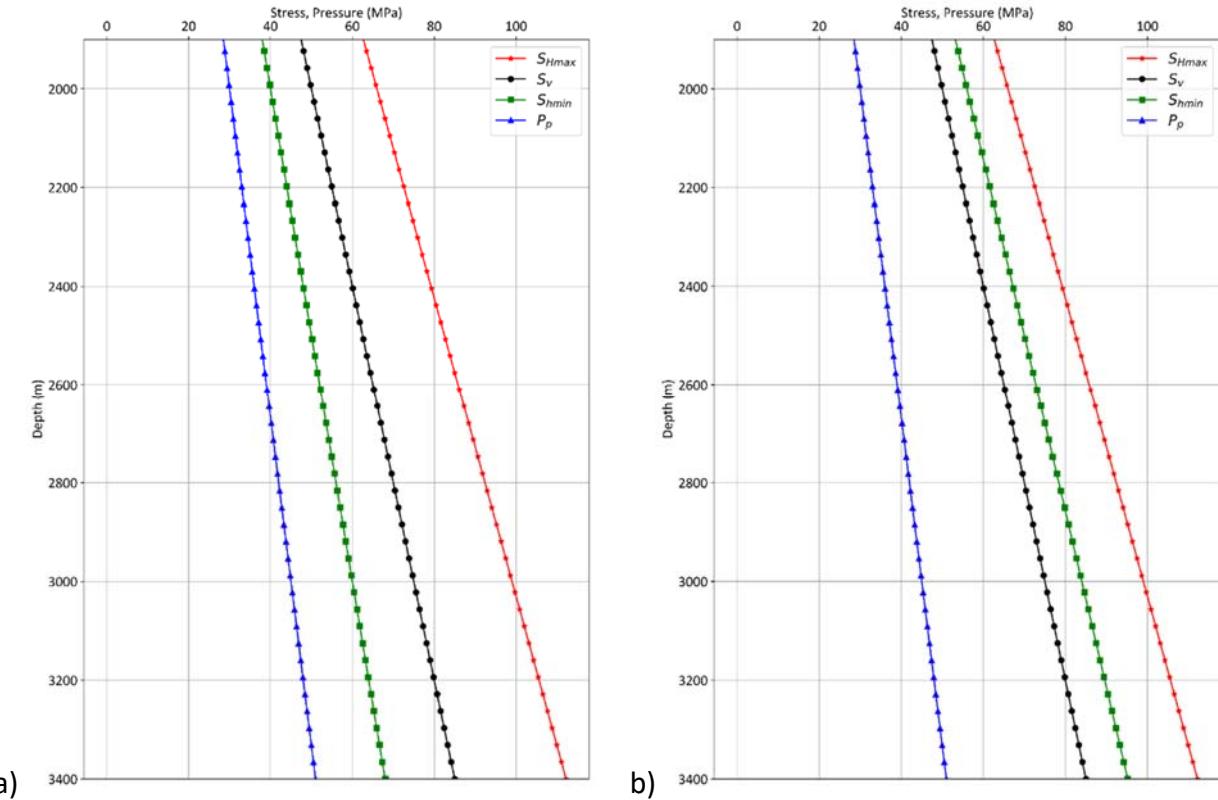


Figure 52 – In-situ stress and pore pressure profiles modelled, for: a) strike-slip, and b) reverse fault far-field stress regimes applied as model boundary conditions. Note that the injection depth (vertical stress) is the same for both scenarios.

Table 11 – XSite™ model inputs, and sources, for fluid injection and induced seismicity simulation, including in-situ stress conditions and rock mass and fault properties.

Property	Description
Model geometry	1.6 km × 1.6 km × 1.5 km
Fault radius, initial aperture	$r = 650 \text{ m}$, $a = 0.0001 \text{ m}$
Fault strength properties	$\mu = \tan(27)^{1,2}$, $c = 9.5 \text{ MPa}^3$, $T = 2.5 \text{ MPa}^2$
Injection rate at point of intersection with fault ⁴	$q_{\text{inj}} = 0.2 \text{ m}^3/\text{s}$ ⁵
S_v (Z-axis)	Strike-slip: $48 \text{ MPa} \pm 25 \text{ MPa/km}$, Reverse: $48 \text{ MPa} \pm 25 \text{ MPa/km}$ ^{1,6}
$S_{H\max}$ (Y-axis)	Strike-slip: $63 \text{ MPa} \pm 33 \text{ MPa/km}$, Reverse: $63 \text{ MPa} \pm 33 \text{ MPa/km}$ ^{1,6}
$S_{h\min}$ (X-axis)	Strike-slip: $38 \text{ MPa} \pm 20 \text{ MPa/km}$, Reverse: $53 \text{ MPa} \pm 28 \text{ MPa/km}$ ^{1,6}
Pore pressure gradient	15 MPa/km ^{1,6}
Depth at top of the model	1.9 km
Formation density	2600 kg/m^3 ^{1,2,7}
Elastic properties	$E = 32 \text{ GPa}$, $v = 0.22$ ^{1,2,7}
Matrix internal friction angle	47° ^{1,7}
Unconfined Compression Strength (UCS)	120 MPa ⁷
Fracture toughness	$1.25 \text{ MPa.m}^{0.5}$ ^{8,9}
Tensile strength	15 MPa ⁹
Matrix permeability and porosity	$k = 50 \text{ nD}$ ^{1,10} , $\phi = 5\%$ ^{10,11}

¹ McLellan (2016); Rogers & Patrick (2014).

² Vaisblat et al. (2019).

³ Determined based on sensitivity analysis. Note that the cohesion adds more brittle behavior to the fault and thus a more sudden release of energy. This value is selected to simulate a conservative scenario.

⁴ The model is designed to examine the fluid injection rates, pressures and volumes as experienced by the fault. The implementation of the fluid injection represents a conservative scenario where the borehole or the hydraulic fracture intersects with the DFN connected with the fault (see Figure 53).

⁵ Zoback & Kohli (2019).

⁶ Fox et al. (2019); the gradients implemented simulate injection into an over-pressured reservoir. This represents a conservative scenario. The vertical stress gradient is calculated from the average density of the formations from the ground surface to the reservoir depth.

⁷ Trican Geological Solutions (2013).

⁸ Chen & Wong (2018).

⁹ Abdelaziz et al. (2019).

¹⁰ Javaheri et al. (2017).

¹¹ Vishkai & Gates (2019).

Table 11 also lists the rock mass and fault properties used for model input. These were compiled from reports on laboratory testing results for Montney rock specimens retrieved from reservoir depths (Trican Geological Solutions report, 2013; McLellan, 2016; Vaisblat et al., 2019). It was assumed that the rock mass matrix is homogeneous and isotropic, and that the fault properties were uniform along the planar fault. A Coulomb shear slip criterion was used to model the fault strength. It is recognized that fault properties are more likely to be strongly heterogeneous and encompass damage zones adjacent to the fault; this consideration is the topic of planned future research, but was partly accounted for here by assigning cohesive and tensile strength values to the fault to account for asperities, non-planarity and intact rock bridges.

The injected fluid was assumed to be incompressible relative to the rock mass compressibility (Damjanac et al., 2020). This assumption simplifies the interpretation of the simulation results. As shown in Figure 50, the initial fluid pressures in the matrix and fault were set to have the same gradient of 15 MPa/km. The matrix is otherwise assumed to be impermeable, reflective of the ultra-low permeability rocks present in the Montney. All model boundaries are likewise treated as impermeable (i.e., no flow is permitted in or out of the modelled domain). Fluid injection was simulated directly into the fault to gain computational advantages in simulation computing times. This simplified assumption is representative of a scenario where either the borehole, hydraulic fracture or a discrete fracture network (DFN) intersects the fault and transmits fluid pressure radially into the more permeable fault (Figure 53). As the injection continues, the zone of interaction (or the injection pressure front) transmitted to the fault gets larger. The injection flow rate used was $0.2 \text{ m}^3/\text{s}$ (i.e., $\sim 75.5 \text{ bbl/min}$), which represents an average value for well stimulations lasting two hours (Zoback & Kohli, 2019).

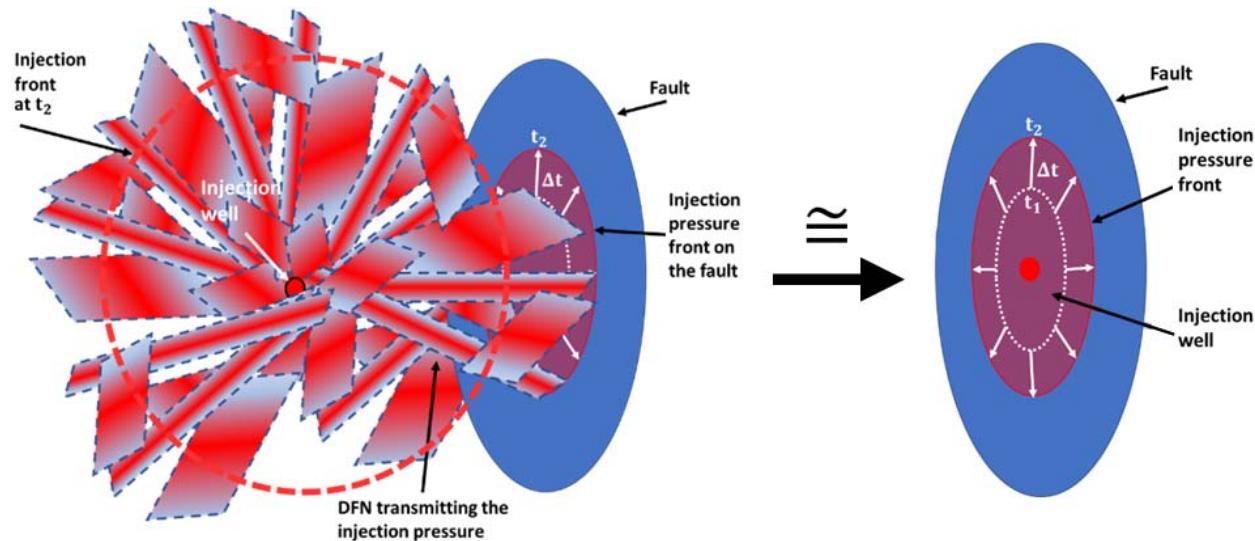


Figure 53 – Fluid injection method used for the XSite™ modelling. In this scenario, a fracture network intersects with a fault, transmitting fluid pressure as a cone-shaped pressure front that radiates outwards from the injection point and intersects and perturbs an area of the fault, which increases from t_1 (white dotted circle) to t_2 (solid red circle) with increasing injection time.

8.3. Procedure for Calculating Induced Seismicity Moment Magnitude

Figure 54 represents the relationships between earthquake moment magnitude (moment release), rupture radius, stress drop and shear displacement, derived using procedures reported in Stein and Wysession (2003) and Scholz (2019).

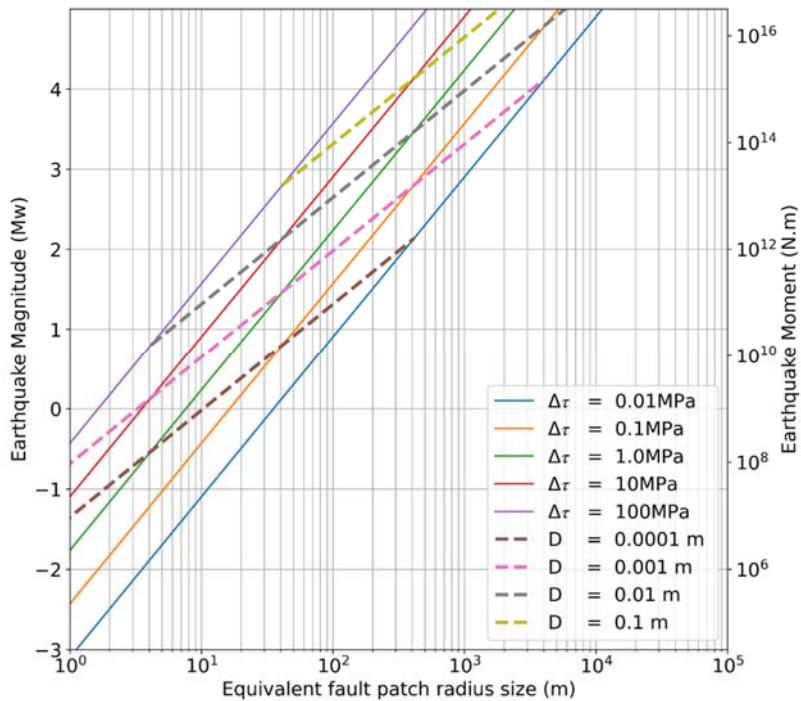


Figure 54 – Relationships between different scaling parameters for earthquake moment magnitude derived using procedures reported in Stein and Wysession (2003) and Scholz (2019). A shear modulus of 30 GPa was assumed for the rock mass hosting the fault. Larger earthquake magnitudes result from larger fault areas (x-axis) and slip displacements (dashed lines), as well as stress drops (solid lines).

From Figure 54, depending on the displacement magnitude and stress drop, the modelled fault size in this case (650 m radius) is capable of producing earthquakes of $M_w \geq 3.0$. The relationship between the shear displacement in an earthquake, its source dimension, and its seismic moment release is closely tied to the magnitude of the stress released by the earthquake, or the stress drop ($\Delta\tau$). The earthquake releases the strain energy stored in the fault due to the applied far-field stress. The change in the strain energy ($\Delta\varepsilon_{shear}$) for an earthquake whose rupture radius is R and its average slip dimension is \bar{D} , could be expressed as:

$$\Delta\varepsilon_{shear} \approx \frac{\bar{D}}{R} \quad (1)$$

Hence the stress drop averaged over the fault would be:

$$\Delta\tau \approx \frac{G\bar{D}}{R} \quad (2)$$

where, G is the shear modulus. We can calculate the average stress drop over a circular slip area using the moment release (M_o) as following:

$$\Delta\tau = \frac{7}{16} \frac{M_o}{R^3} \quad (3)$$

where $M_o = G\bar{D}\pi R^2$ (Scholz, 2019).

From this relationship, we are able to calculate the moment magnitude for a modelled fault slip event in XSite™, as both the average stress drop and shear slip magnitude distribution along the fault. The strain energy released can also be calculated as $W = \Delta\tau \times \pi R^2 \times \bar{D}$ (i.e., work equals force times distance).

8.4. Simulation Results

8.4.1. Importance of Depth to Top of Basement

The first set of model results investigated the influence of depth to the top of the basement at the well location on induced seismicity susceptibility. This feature was repeatedly ranked as the most important by the machine learning classification models for the susceptibility analysis. The XSite™ model in this case was only used in a limited sense as the full model presented in Section 8.2 is designed more specifically to model induced seismicity event severity (i.e., maximum moment magnitude). Note that the depth to the basement top did not rank in the top five important features in the machine learning regression models for the severity analysis.

According to the machine learning classification model results, the distance to the top of the basement showed a negative correlation with increasing likelihood of a well being seismogenic (e.g., see Figures 32 and 33). In other words, wells located where the depth to the top of the basement is shallower are more likely to be seismogenic. This correlation does not specify what the corresponding injection depth is, but it can be postulated that as the distance between the injection point and the top of the basement increases, it is less likely that the injection pressure front interacts with any basement faults. At the same time, there is the added consideration that as the basement depth increases, deeper faults will be under higher stresses and therefore potentially in a more critical state that is susceptible for generating seismicity.

To test this interaction, the modelled simulation made use of a realistic discrete fracture network (DFN) realization developed for the Montney formation (courtesy of Dr. Steve Rogers, Golder). The in-situ stress and pore pressure gradients initialized in the models were the same as the strike-slip stress regime provided in Figure 52a. However, to simulate a low and high stress environments, the top of the model was selected to be at a depth of 1.9 and 2.5 km, respectively. These depths were selected in accordance with the Well Features Dataset compiled for the machine learning analyses. In both scenarios, the depth of the injection point relative to the

depth of the basement (and any basement faults) is the same. This allows the influence of the stress magnitude (as a proxy for depth) to be singled out in the models.

The model results are presented in Figure 55. These show the extent of the fluid pressure invaded zone after 30 minutes of injection in the form of the modelled shear-slip distributions along the individual fractures making up the DFN. These indicate that for the deeper basement and injection scenario (Figure 55b), the corresponding higher stresses act to limit the extent of the invaded zone transmitted outwards from the injection point and hydraulic fracture. Conversely, the results for the lower basement depth (Figure 55a) show that the lower stress environment results in a significantly larger invaded zone (i.e., larger volume of rock mass subjected to increased pore pressures). Thus, as the distance to the top of the basement becomes smaller, fluid injection results in a larger volume of elevated pore pressures, which in turn results in an increased likelihood of interacting with a critically stressed fault in the basement to trigger an induced seismicity event.

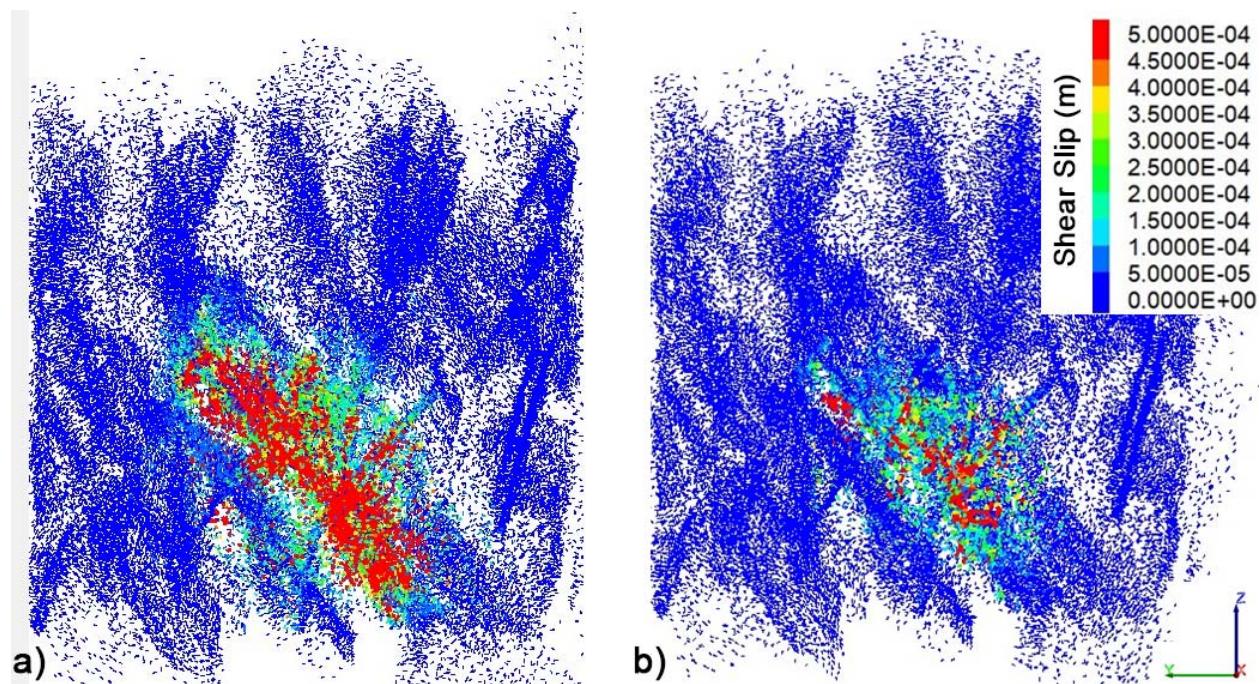


Figure 55 – Shear displacements generated along faults within a discrete fracture network at the end of 30 minutes of injection as a rate of $0.2 \text{ m}^3/\text{s}$ for two different depth scenarios: a) injecting at the depth of 2.3 km (top of the model is at 1.9 km depth), b) injection at the depth of 2.9 km (top of the model is at 2.5 km depth). The strike-slip far-field stress regime boundary condition and pore pressure profiles are shown in Figure 52a.

8.4.2. Influence of Formation Stiffness

Modelled Parameter Range

The next series of XSite™ models investigated the influence of elastic stiffness for the rock mass hosting a critically stressed fault. Rock mass stiffness is known to be a key influencing factor in fault-slip rockburst hazards encountered in deep mining. Figure 56 reports the range of laboratory-measured Young's Modulus (E) values for cores obtained from the Montney formation at reservoir depths between 2.3 and 2.6 km, as reported by Trican Geological Solutions (2013). These show values ranging from 16 to 50 GPa. McKean and Priest (2019) report an even higher range (42 to 72 GPa), albeit for samples obtained from Montney surface outcrops. To ensure a thorough analysis of the parameter space, we selected five different Young's modulus values covering the reported ranges. As previously noted, the rock properties were assumed to be homogeneous and isotropic in each model representing a different stiffness of the rock mass domain.

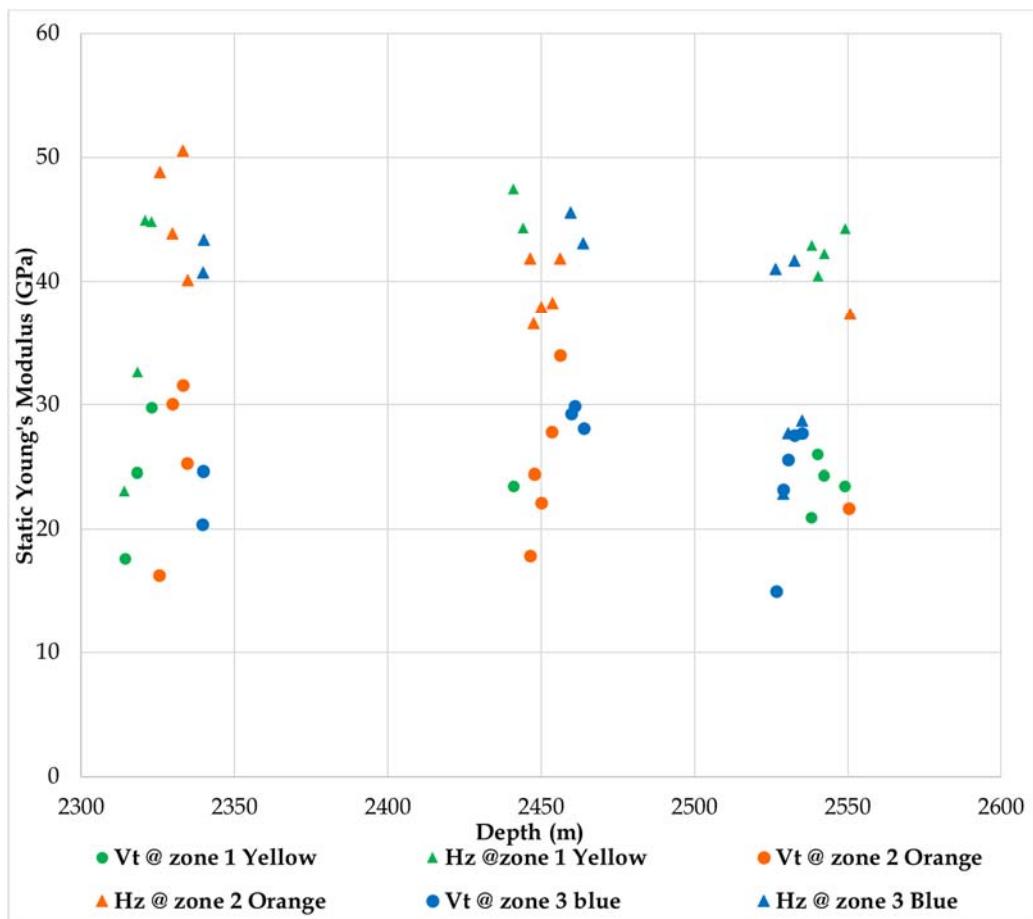


Figure 56 – Laboratory measured Young's modulus (E) values for rock cores retrieved from the Montney formation from the same well, as reported by Trican Geological Solutions (2013). The coloured circles and triangles correspond with values for samples tested vertically and horizontally, respectively, relative to the core axis. The colours represent three different zones designated as having different rock facies.

As the parameter of interest was the formation stiffness, the model geometry used assumed the same fault orientation for each model scenario (dip direction of 45°, clockwise from the maximum horizontal stress, and dip angle of 80°). Modelling for this fault orientation considered both the strike-slip and reverse fault stress regimes. Once the in-situ stresses and pore pressures are initialized and fluid injection begins, the fluid pressures acting on the fault begin to increase. As the fluid pressure increases, the effective shear strength of the fault decreases until the shear slip criterion is satisfied (i.e., $\tau - [(\sigma_n - P_f) \tan(27^\circ) + 9.5] \geq 0$) and shear slip occurs.

It was observed that as the rock mass stiffness increases, fluid pressure in the fault builds up more as there is more resistance to rock mass deformation (i.e., $\varepsilon_v = \frac{\sigma}{K}$, where ε_v is the volumetric strain and K is the bulk modulus), which restricts the ability for the fault aperture to increase. This results in a higher fluid pressure gradient (ΔP_f) at the fault, resulting in a larger perturbed area in a shorter injection time ($q \propto \Delta P_f$, and $q \cdot \Delta t_{inj} = \Delta V = a\pi R^2$, where Δt_{inj} is the injection duration). At the same time, the stiffer rock mass adjacent to the fault provides more confinement (i.e., higher normal stresses), and therefore resistance to slip, resulting in more elastic strain energy being stored. Figure 57 illustrates the increasing amount of strain energy stored for a given amount of shear strain, ε_0 , as a function of increasing rock mass shear stiffness, G (note that the stored elastic strain energy is proportional to the area under the shear stress versus shear strain plot for the rock). Thus, for the same slip area (πR^2) and displacement (D), it is expected that stiffer rock will result in a larger moment release (recall that $M_o = GD\pi R^2$).

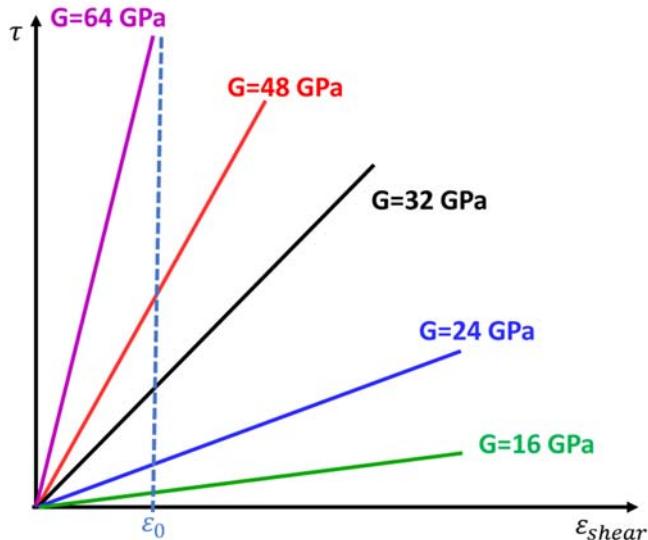


Figure 57 – Schematic illustration of the stored elastic strain energy for different rock stiffnesses. The stored elastic strain energy is proportional to the area under the stress-strain plot for the rock. Thus, as rock shear stiffness, G , increases, so does the strain energy stored in the rock mass for the same amount of deformation, ε_0 .

Reverse Fault Far-Field Stress Regime

Figure 58 plots the moment release as a function of injection time for the reverse fault far-field stress regime boundary condition. The solid lines are for the base-case maximum horizontal stress gradient (i.e., 33 MPa/km), as described for Figure 52b, whereas the dashed lines are for a higher horizontal stress gradient of 38 MPa/km. The latter represents an adjusted gradient that conforms to the observed horizontal stress ratio (i.e., $1.1 \leq S_H/S_h \leq 1.4$; see Rogers & Patrick, 2014), but also results in a differential stress (i.e., S_1-S_3) that is the same as for the strike-slip stress regime scenario (Figure 52a). This facilitates comparison to results for the strike-slip stress regime scenario presented later. Another important consideration with respect to interpreting this plot, is that XSite™ is a quasi-static program, while fault slip rupture events are dynamic. To address this limitation, the maximum moment magnitude for a simulation can be calculated by summing the displacement events over the total fluid injection time, thus treating these as a cluster of cascading events resulting in a single moment magnitude, even though the fault slip response in the model occurs over an extended injection time. Figure 58 shows the calculated maximum moment magnitudes (M_w) with arrows pointing to the end of the injection time.

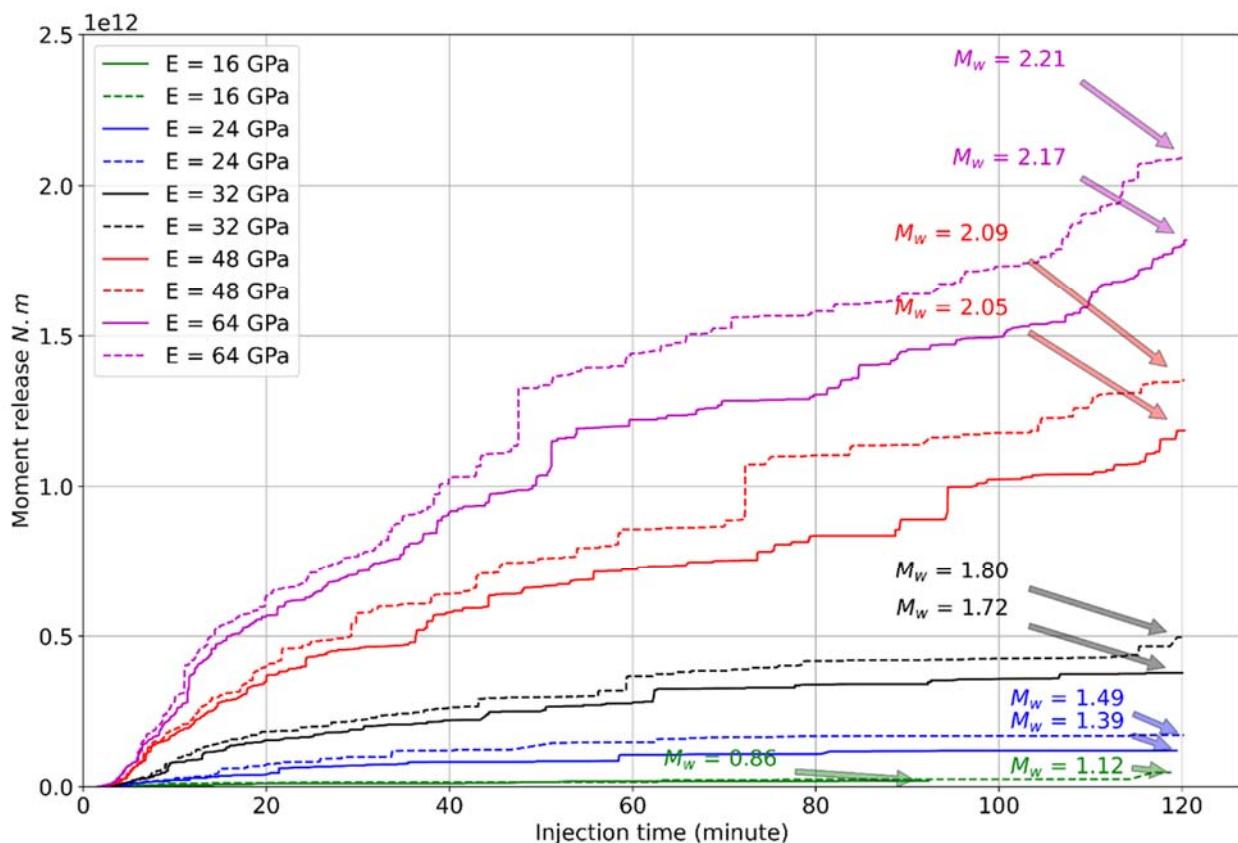


Figure 58 – Moment release with fluid injection for models with different formation stiffnesses. The modelled scenario is for a fault with a dip of 80° and dip direction of 45° relative to a reverse fault far-field stress regime, assuming either a base-case (solid lines) or maximum (dashed lines) stress gradient. Reported for each is the quasi-static maximum moment magnitude (M_w), as described in the text.

As expected, the overall trend in Figure 58 indicates that higher rock mass stiffnesses result in the potential for increased moment release, and thus, larger magnitude events. Comparing the solid and dashed lines, the results also suggest that as the differential stress increases, the moment release and maximum event magnitudes increase further. The latter can be explained mechanistically by comparing the shear to normal stress ratio acting on the fault prior to injection, which for the larger differential stress is slightly higher (i.e., $\tau/\sigma_n = 0.30$ for the maximum stress gradient compared to $\tau/\sigma_n = 0.18$ for the base-case stress gradient; Figure 59). Consequently, a higher differential stress results in a more critically stressed fault with a larger potential slip area, resulting in larger magnitude events. Viewed another way, a larger differential stress results in a larger stress drop potential and therefore larger magnitude events. This is shown in Figure 60, which compares the stress drop with fault slip for the larger differential stress case (blue symbols) to that for the smaller differential stress case (green symbols).

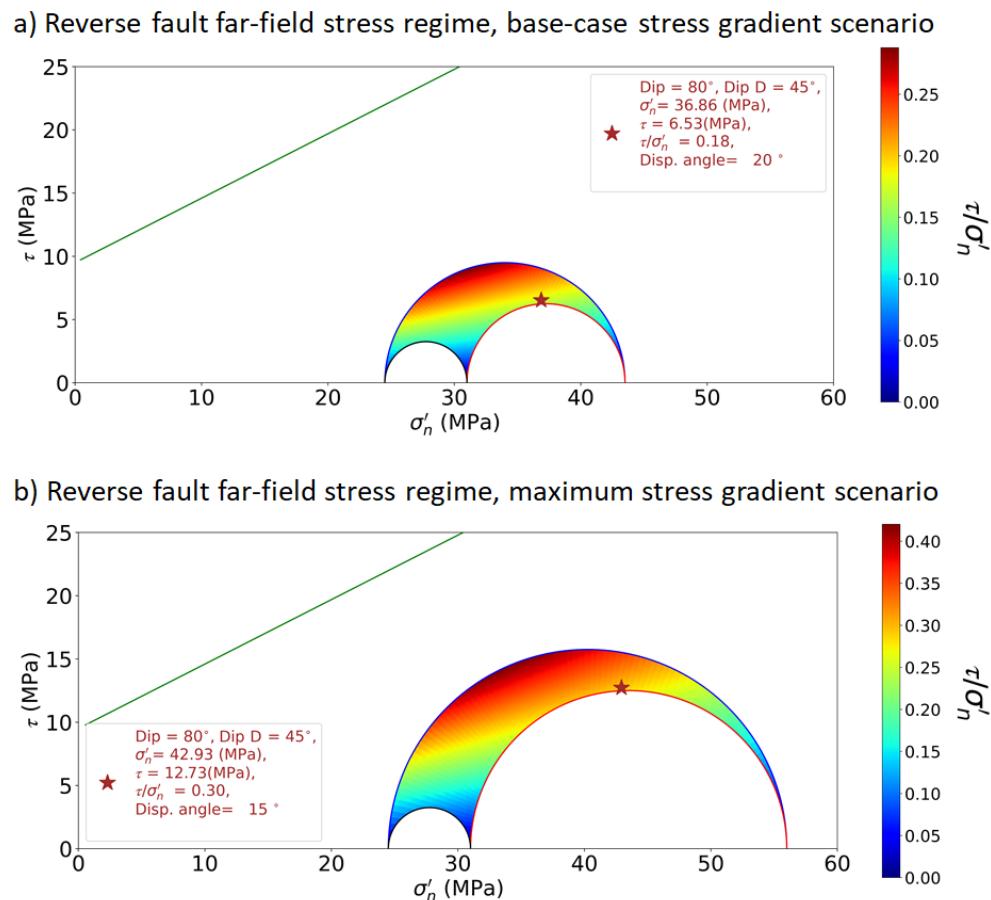


Figure 59 – 3-D Mohr-circle representations of the pre-injection effective normal and shear stress ratios calculated for a depth of 2.4 km under a reverse fault far-field stress regime for the: a) base-case, and b) maximum stress gradient scenarios. The red star indicates the stress state for a fault dipping at 80 degrees with a dip direction of 45° relative to the reverse fault far-field stresses. Also reported are the corresponding effective normal stress (σ'_n), shear stress (τ), τ/σ'_n ratio, and displacement angle (i.e., the clockwise angle of the fault rake from the strike).

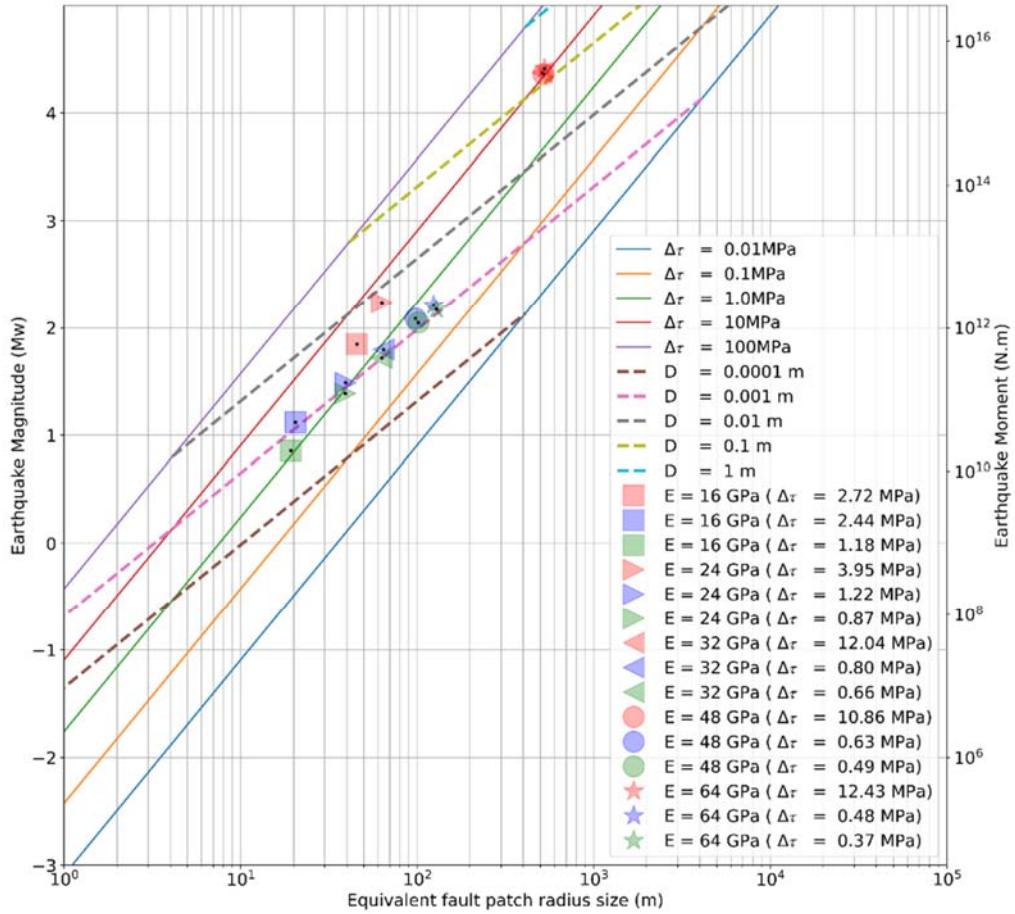


Figure 60 – Average stress drop ($\Delta\tau$) and shear displacement (D), and corresponding event magnitude, derived from the simulation results as a function of formation stiffness. The modelled scenario is for a fault dipping at 80° with dip direction 45° relative to a reverse fault far-field stress regime and base-case (green symbols) and maximum (blue symbols) stress gradient, and strike-slip far-field stress regime (red symbols). These are based on the maximum induced seismicity event after 120 minutes of fluid injection.

Strike-Slip Far-Field Stress Regime

Figure 60 also includes for comparison the average stress drops modelled for the strike-slip far-field stress regime. These are seen to be significantly larger for the same rock mass stiffness than the stress drops for the reverse fault far-field stress regime, including the maximum stress gradient case. This is further shown in Figure 61, which plots the moment release as a function of simulated injection time for each rock mass stiffness. Comparing Figure 61 to the same plot for the reverse fault far-field stress boundary condition (Figure 58), the moment release and event magnitudes are significantly higher for the strike-slip far-field stress boundary condition. This relates to the τ/σ_n ratio for the strike-slip far-field stress regime ($\tau/\sigma_n = 0.56$; see Figure 62) being double that for the reverse fault far-field stress regime (compare to Figure 59). This shows that the strike-slip far-field stress regime results in the fault, for the assumed fault orientation, being more critically stressed, thus requiring a smaller fluid pressure perturbation to cause slip over a larger area resulting in much larger events.

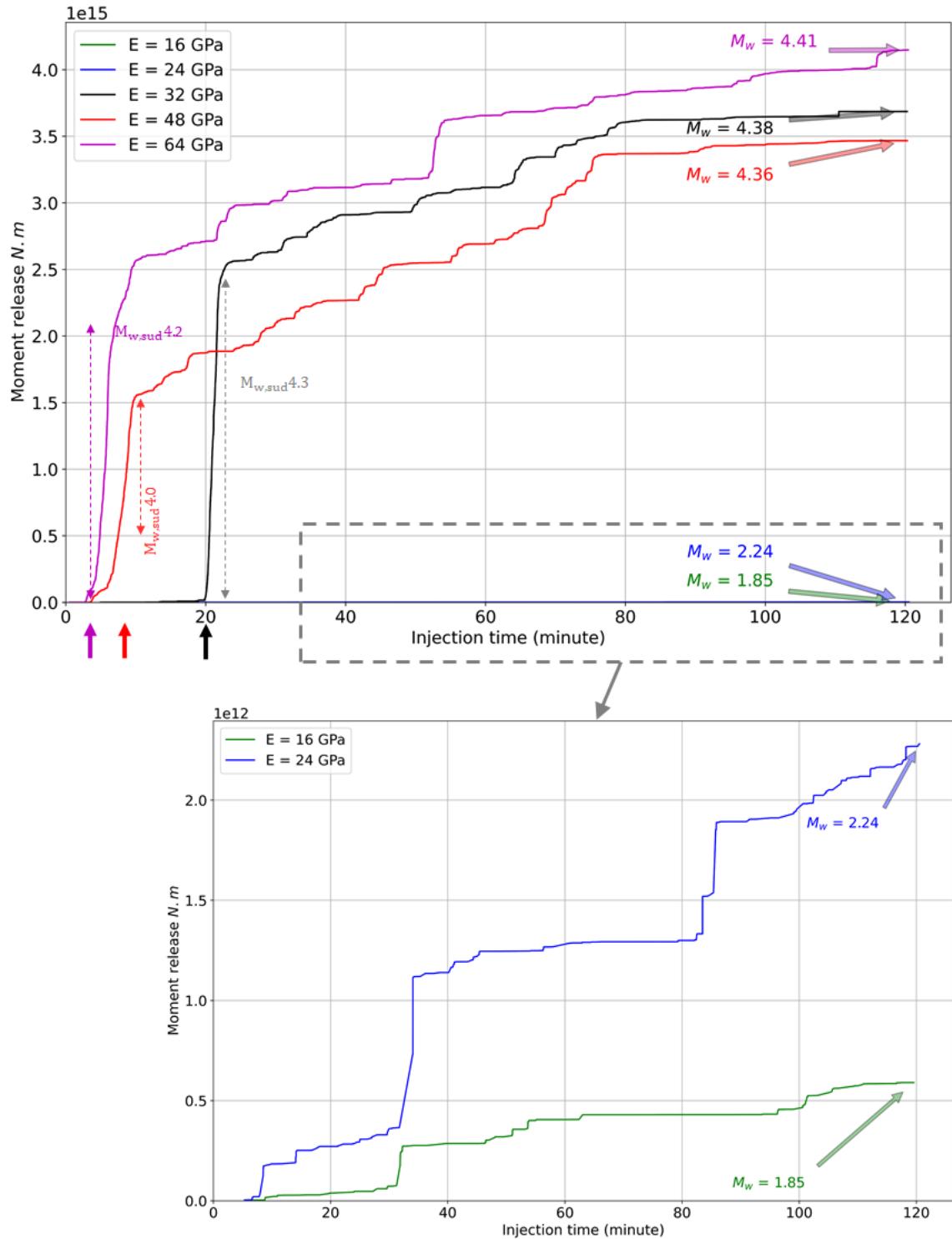


Figure 61 – Moment release with fluid injection for models with different formation stiffnesses, for a fault dipping at 80° with dip direction 45° relative to a strike-slip far-field stress regime. Reported for each case is the quasi-static maximum moment magnitude (M_w), as described in the text. The dashed arrows represent the sudden energy release and corresponding moment magnitude ($M_{w,sud}$). Shown below the main plot is one where the Moment Release scale bar is three orders of magnitude smaller to show the details of the response for the lower rock stiffness cases.

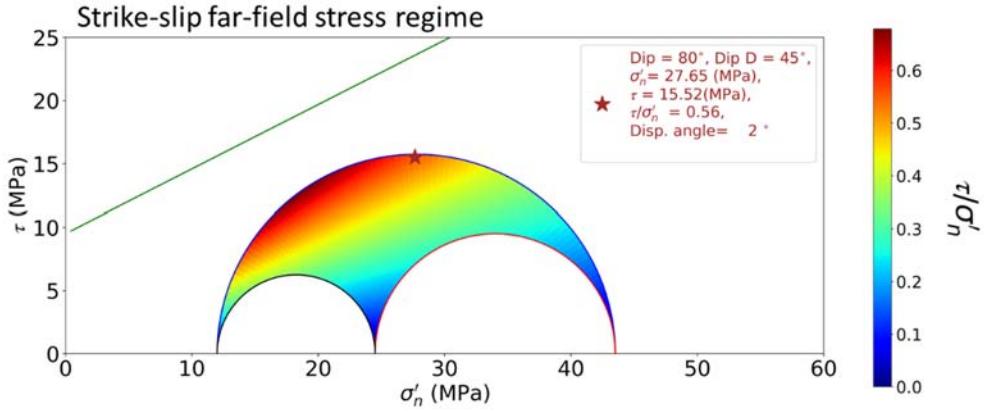


Figure 62 – 3-D Mohr-circle representation of the pre-injection effective normal and shear stress ratios calculated for a depth of 2.4 km under a strike-slip far-field stress regime. The red star indicates the stress state for a fault dipping at 80 degrees with a dip direction of 45° relative to the strike-slip far-field stresses. Also reported are the corresponding effective normal stress (σ'_n), shear stress (τ), τ/σ'_n ratio, and displacement angle (i.e., the clockwise angle of the fault rake from the strike).

Figure 61 also shows an interesting deviation from the general trend of increasing moment release with increasing rock mass stiffness for the intermediate case of $E = 32$ GPa. The results for $E = 32$ GPa show a higher moment release than the stiffer $E = 48$ GPa case. Exploring this further, it can also be seen that there is a difference related to the injection time when the first sudden moment release ($M_{w,sud}$) is experienced. For the two higher stiffness cases ($E = 64$ and 48 GPa), this occurs after an injection time of 10 minutes, whereas for the intermediate case ($E = 32$ GPa) it is 20 minutes, and for the two lower stiffness cases ($E = 24$ and 16 GPa) it is 30 minutes. As previously discussed, for the same fluid injection rate, the corresponding fluid pressures acting on the fault increase as the stiffness of the surrounding rock mass increases. Figure 63 shows the bottom-hole fluid pressures during the simulated injection (dashed lines), which confirms these are higher for the two stiffer formations during the first ten minutes of injection. What is seen in the models is that the critical pressure ($P_{f,crit.}$) to initiate a fault slip event is reached with smaller injection volumes for stiffer rocks. At the same time, the smaller injection volumes mean that the areal extent of the increased pore pressures acting on the fault surface is also less. Therefore, two factors are at play in the XSite™ models with respect to the magnitude of the induced seismicity event: i) how critically stressed the fault is in relation to the in-situ stress condition, and the corresponding stress drop; and ii) the areal extent along the fault experiencing the pore pressure increase resulting from the fluid injection. The latter introduces a temporal component related to the injection rate.

This is illustrated in Figure 64. As the formation stiffness decreases, it takes a higher fluid pressure acting on the fault to trigger a slip event, which requires a longer injection time and thus a larger injection volume is able to radiate further into the fault resulting in a larger affected fault area. Comparing $E = 64$ GPa to $E = 48$ GPa, the models are comparable in their sensitivity to the fluid pressure and therefore they slip at similar times, with the $E = 64$ case generating a larger slip area

(solid lines in Figure 64) and stress drop as the fault is more critically stressed. In contrast, the $E = 24$ and 16 GPa models are less critically stressed, and thus require longer injection times (larger injection volumes) for pressures acting on the fault to be sufficient, relative to the corresponding aerial extent of fault weakening, to trigger slip.

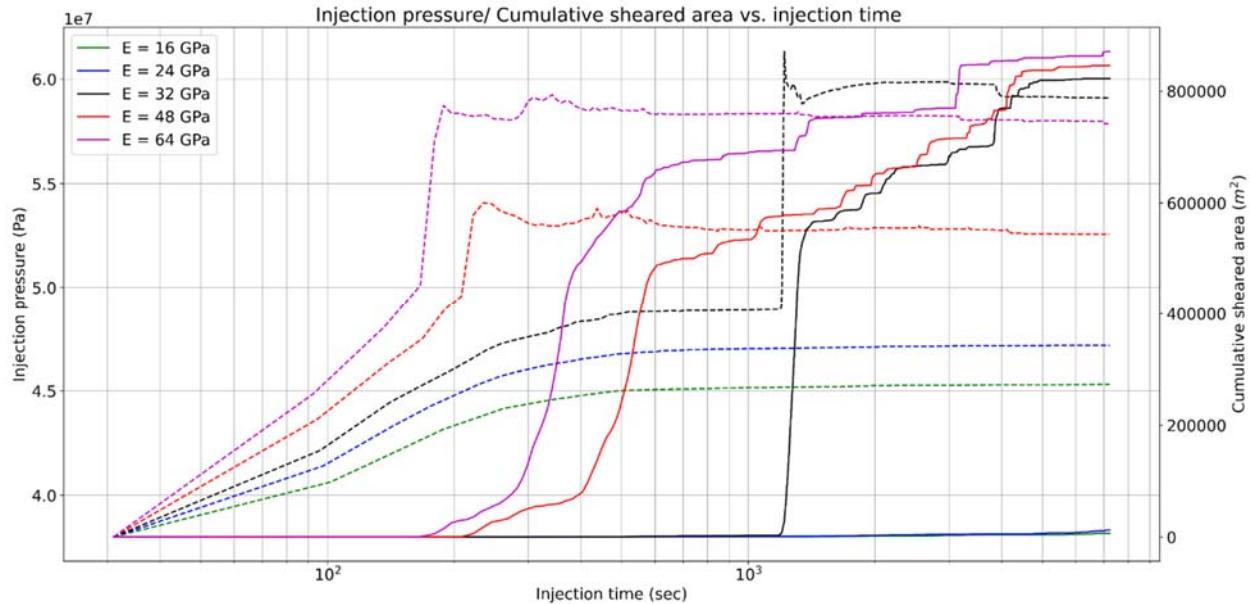


Figure 63 – Bottom-hole injection pressures (dashed lines) and sheared slipped area on the fault (solid lines) as a function of injection time, for different formations. The modelled scenario is for a fault with dip direction = 45° and dip = 80° in a strike-slip stress regime.

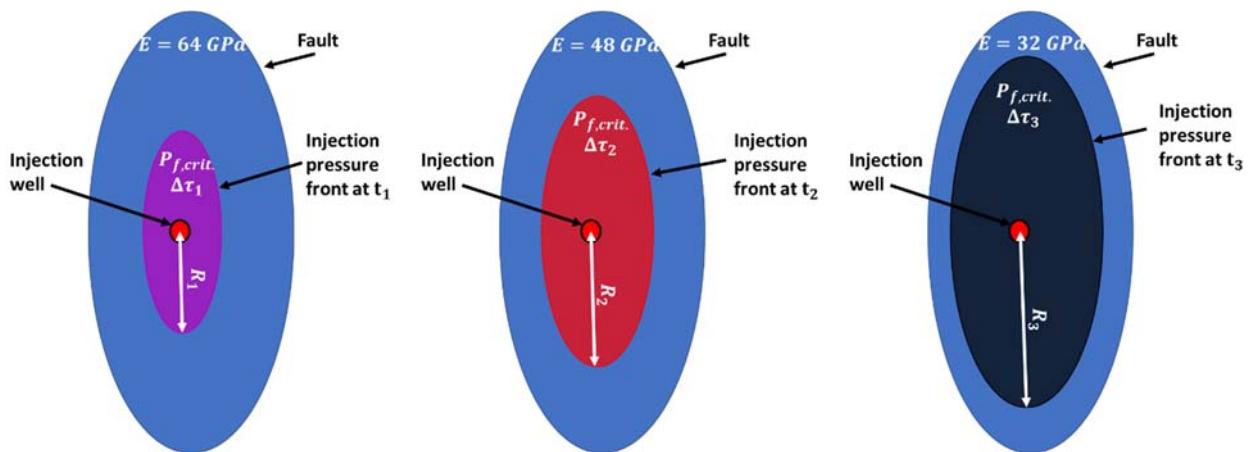


Figure 64 – Schematic representation of the fluid injection front at the time of the first seismic moment release, as observed in Figure 61. The colours of the injection pressure fronts above match those for the rock stiffnesses in Figure 61. These show the area of fluid pressure acting on the fault increasing with injection time. The corresponding stress drops are, from left to right, $\Delta\tau_1 \geq \Delta\tau_2 > \Delta\tau_3$.

It should be noted that although the area on the fault affected by the pore pressure increase is larger for the lower rock mass stiffness cases, this does not translate into larger induced seismicity event magnitudes. Although the influenced area is larger, the area satisfying the slip condition ($\tau - [(\sigma_n - P_f) \tan(27^\circ) + 9.5] \geq 0$) is smaller, resulting in a smaller seismic moment.

Thus, for the $E = 32$ GPa case, an intermediate condition exists. Because the fault in this case is less critically stressed compared to the higher stiffness cases, a longer injection time is possible allowing more fault area to see an increase in fluid pressure (and decrease in effective stress). At the same time, the rock is still stiff enough to store considerable strain energy. The combined result is that when the induced seismicity event is triggered, there is more seismic energy released suddenly for the $E = 32$ GPa case compared to the $E = 48$ GPa case.

Comparison to Machine Learning Results

The machine learning results for the regression modelling of induced seismicity severity indicate that an increase in P-wave velocity (“BHC Pv”) has a positive correlation with induced seismicity event magnitude (see the SHAP plots in Figures 46 and 47). At the same time, there is a well established positive correlation between seismic wave velocity and dynamic Young’s Modulus, where higher P-wave velocities are associated with higher rock stiffnesses. Thus, the numerical modelling results explain that the correlation between higher P-wave values and larger magnitude induced seismicity events is through their association with higher stiffness rock.

The machine learning severity analyses also indicates that the number of active completion days (“Active_Days”) is among the most important features and shows a negative correlation with induced seismicity magnitude. In other words, a shorter activity period correlates with larger events. This means that there is a higher likelihood of seeing a large magnitude event in the early days of an injection operation. The numerical modelling results would suggest that these are possibly wells that are in proximity to faults in higher stiffness rock in a strike-slip far-field stress regime (see Figure 61). Although the XSite™ modelled scenario includes numerous simplifications and generalizations, it is interesting to note that the average formation stiffness for the Montney is approximately 32 GPa (Figure 56), and that the basement rocks would be even stiffer. Together, the machine learning and numerical modelling results suggest that if operating in a strike-slip far-field stress regime in stiffer rocks and there are fewer early events of $M_w > 1$ or 2 than expected, that this might be a sign that strain energy is accumulating towards a large $M_w > 4$ event. Similarly, in the context of a Traffic Light Protocol (e.g., Walters et al., 2015) and seismic hazard, the detection of a $M_w 2$ event might be more concerning when the formation targeted by the injection, or adjacent to it, is stiffer (e.g., $E > 32$ GPa). It is acknowledged that these results assume a good hydraulic communication between the injection wells and the fault causing observation of early seismicity (i.e., causing a strong correlation between injection and the observation of induced seismicity). It is also noted that the results do not consider activity after shut-in and large induced seismicity events are possible after well shut-in (e.g., Häring et al., 2008).

8.4.3. Influence of Fault Orientation

Modelled Parameter Range

Given the complex tectonic history of NEBC, the development and presence of faults with different orientations can be expected. Similarly, their orientation relative to the present-day far-field stress regime, or any localized stress heterogeneity, is also likely varied. To investigate this further, the XSite™ model was used to examine the seismogenic response relative to different fault orientations for the reverse fault and strike-slip far-field stress boundary conditions.

For the reverse fault far-field stress condition (Figure 52b), Andersonian faulting theory (Anderson, 1951) suggests that when $S_H > S_h > S_V$, gently dipping faults ($\sim 30^\circ$) with a dip direction subparallel to S_H will be the most likely to activate. Based on this, the XSite™ model was run assuming a fault dip angle of 30° and varying the dip direction from 0° (parallel to S_H) to 90° (parallel to S_h) in 15° increments. This analysis was then repeated for a steeply dipping fault (80°) to consider a fault with an orientation that is not optimally aligned with the Andersonian reverse fault stress assumption.

For the strike-slip far-field stress condition (Figure 52a), Andersonian faulting theory predicts that when $S_H > S_V > S_h$, faults that are subvertical ($\sim 80\text{--}90^\circ$) and with a dip direction of 60° from S_H are the most critical. Accordingly, a model series was run with a fault dip angle of 80° and varying the dip direction from 0° (parallel to S_H) to 90° (parallel to S_h) in 15° increments. This follows the same orientations as modelled for the reverse fault far-field stress condition to aid comparison. A second set of models was run with a dip direction that is slightly outside the optimum range at 45° to S_H .

Reverse Fault Far-Field Stress Regime

Figure 65a shows the results for a fault dipping at 30° in a reverse fault far-field stress regime with varying dip orientation. As predicted by Andersonian fault mechanics, the fault dipping in the direction of S_H is the most seismogenic, producing the highest magnitude event. Overall, the maximum potential magnitudes for the range of dip directions modelled range from 1.29 to 1.74. Deviating from the most critical fault dip angle, Figure 65b provides the results for a subvertical fault dipping at 80° . A similar range in maximum potential magnitudes was obtained (1.51 to 1.77), but in this case faults with dip directions between 30° and 60° to S_H had slightly larger moment magnitudes. Nevertheless, even when extending these models by applying injection times of up to 8 hours, no large magnitude events ($M_w \geq 3$) were produced.

Mechanistically, the models show that for the reverse fault far-field stress condition, the pre-injection effective normal and shear stresses acting on the fault at any orientation are in a stable state and far from failure (see example of Mohr stress circles and failure envelope in Figure 59). The result is that when the fault slip event is triggered by the injected fluid pressures, the slip area is limited in how far it extends beyond the area of fault experiencing the pore pressure increase. When the slip event propagates beyond this area, it is met by fault contacts where the strength is in excess of the initial shear stresses and therefore have capacity to resist the dynamic

stress change. (This will be shown to be different for the strike-slip far-field stress regime where more of the fault area is in a critically stressed state). For the range of parameters tested for the reverse fault far-field stress scenario, it was not possible to generate events with $M_w \geq 3$.

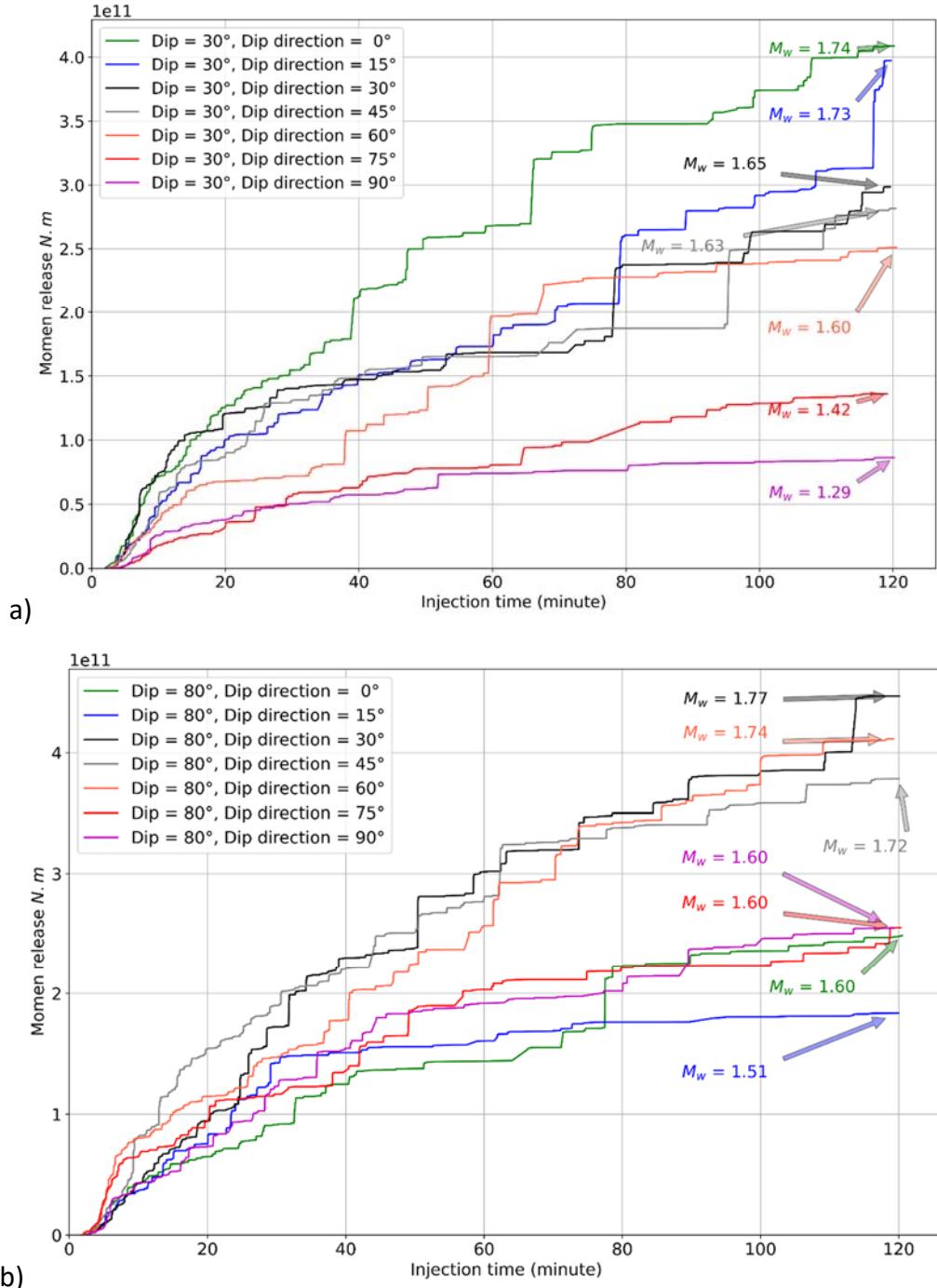


Figure 65 – Moment release with fluid injection for models assuming a reverse fault far-field stress regime and varying fault dip direction, for faults dipping at: a) 30°, and b) 80°. Reported for each case is the quasi-static maximum moment magnitude (M_w), as described in the text.

The other interesting result from these analyses is that the events generated included both strike-slip and reverse faulting focal mechanisms. In other words, even though the far-field in-situ stress boundary condition imposed was specifically that for a reverse fault stress regime, depending on the fault orientation relative to the stress field, both reverse and strike-slip fault slip events could be generated locally in response to the pore pressure increase interacting with the fault.

Strike-Slip Stress Regime

Figure 66 shows the results for a fault dipping at 80° in a strike-slip far-field stress regime with varying dip orientation. As predicted by Andersonian fault mechanics, the fault dipping in the direction of 60° from S_H is the most seismogenic, producing the highest magnitude event. In contrast to the reverse fault far-field stress regime results, where the maximum magnitudes generated were less than M_w 2, a sizeable M_w 4.4 event was generated. Overall, the maximum potential magnitudes for the range of dip directions modelled range from 1.67 to 4.44. Examining the effective normal and shear stresses acting on the fault for these models, it could be seen that for the dip directions between 60° and 90° from S_H (i.e., striking sub-parallel to S_H), the entire fault surface area is in a critically stressed state. The result is that when the injected fluid pressure triggers fault slip, the rupture propagates beyond the pressurized fault interval and it is met by fault contacts already near their limit state resulting in the cascading growth of the slip area, thereby generating a much larger magnitude event, in this case exceeding M_w 4.

Deviating from the most critical fault dip direction, Figure 67 provides the results for varying dip angles with a dip direction of 45° from S_H . A similar broad range in maximum potential magnitudes was obtained (1.70 to 4.41), with the larger moment magnitudes coinciding with sub-vertical fault dips compared to those that were sub-horizontal. This again is in agreement with Andersonian faulting theory, which defines sub-vertical faults as being the most active in a strike-slip far-field stress regime (Anderson, 1951). As before, the pre-injection effective normal and shear stresses acting on faults dipping from 75° to 90° results in these faults being critically stressed over most of the length of the fault (i.e., closer to the shear failure envelope; see for example, compare Figure 62 with Figure 59). Thus, with a small pressure perturbation, a larger area of the fault becomes unstable and slips, resulting in a large seismic moment release.

Figure 68 illustrates the stress drop for the strike-slip far-field stress boundary condition analyses. This shows that the stress drop for the more critically oriented faults, relative to the same fluid injection rate and time (i.e., same injection volume) are ten times larger than the other fault orientations. The stress drop is indicative of the criticality of the fault. Likewise, the shear displacements and slip areas for the critical faults are almost 100 times larger than for the less critical faults. Hence, under a strike-slip far-field stress regime, the seismic energy release ($W = \Delta\tau \times \pi R^2 \times \bar{D}$) for a critically oriented faults is approximately 36,000 times larger than for less critically oriented faults, or as well, faults of any orientation in a reverse fault far-field stress regime.

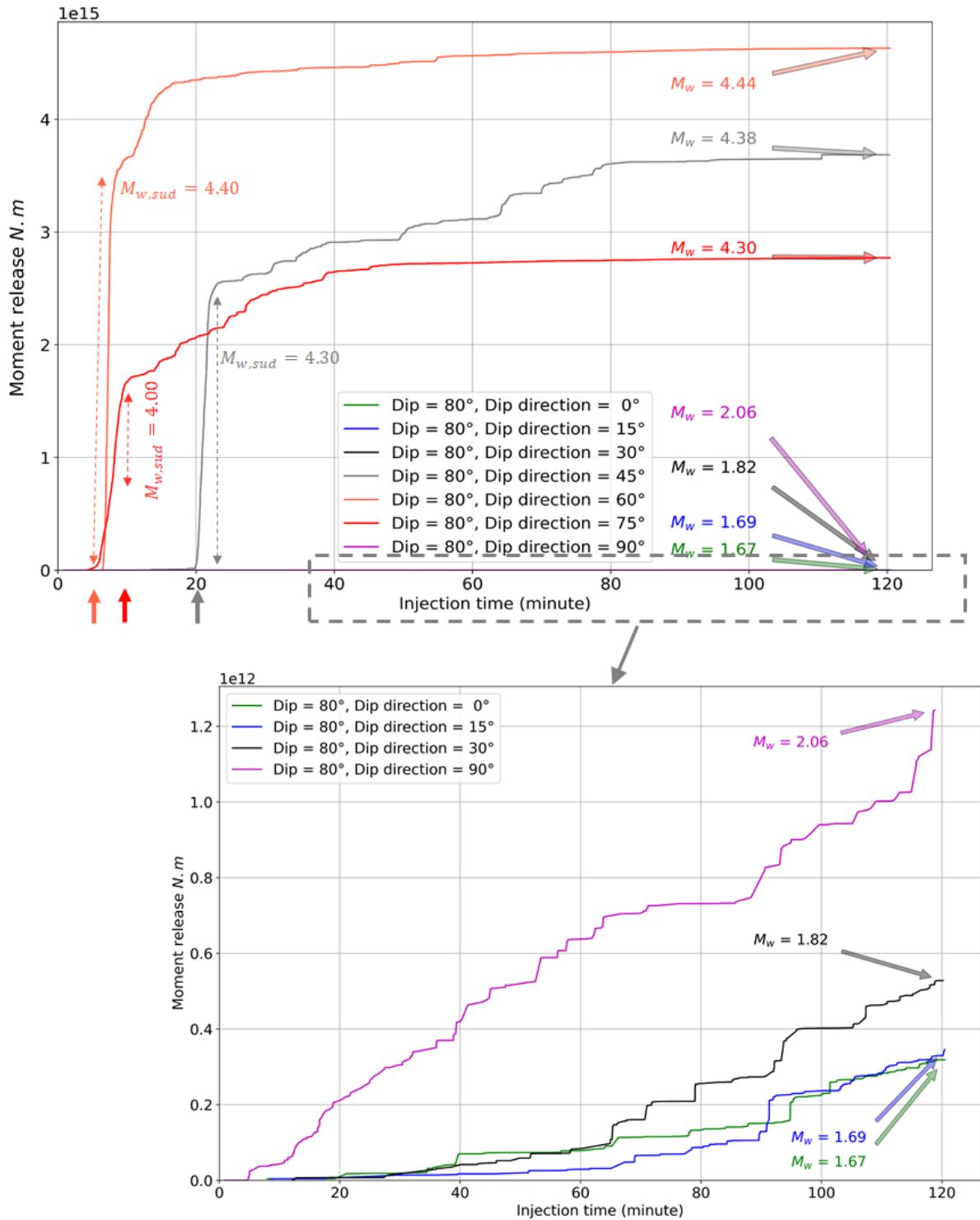


Figure 66 – Moment release with fluid injection for models with a fault dipping at 80° but with different dip directions, assuming a strike-slip far-field stress regime. For reference, a dip direction of 0° is in the direction of S_H (i.e., 90° dips parallel to S_H). Reported for each case is the quasi-static maximum moment magnitude (M_w), as described in the text. The dashed arrows represent the sudden energy release and corresponding moment magnitude ($M_{w,sud}$). Shown below the main plot is one where the Moment Release scale bar is three orders of magnitude smaller to show the details of the smaller events.

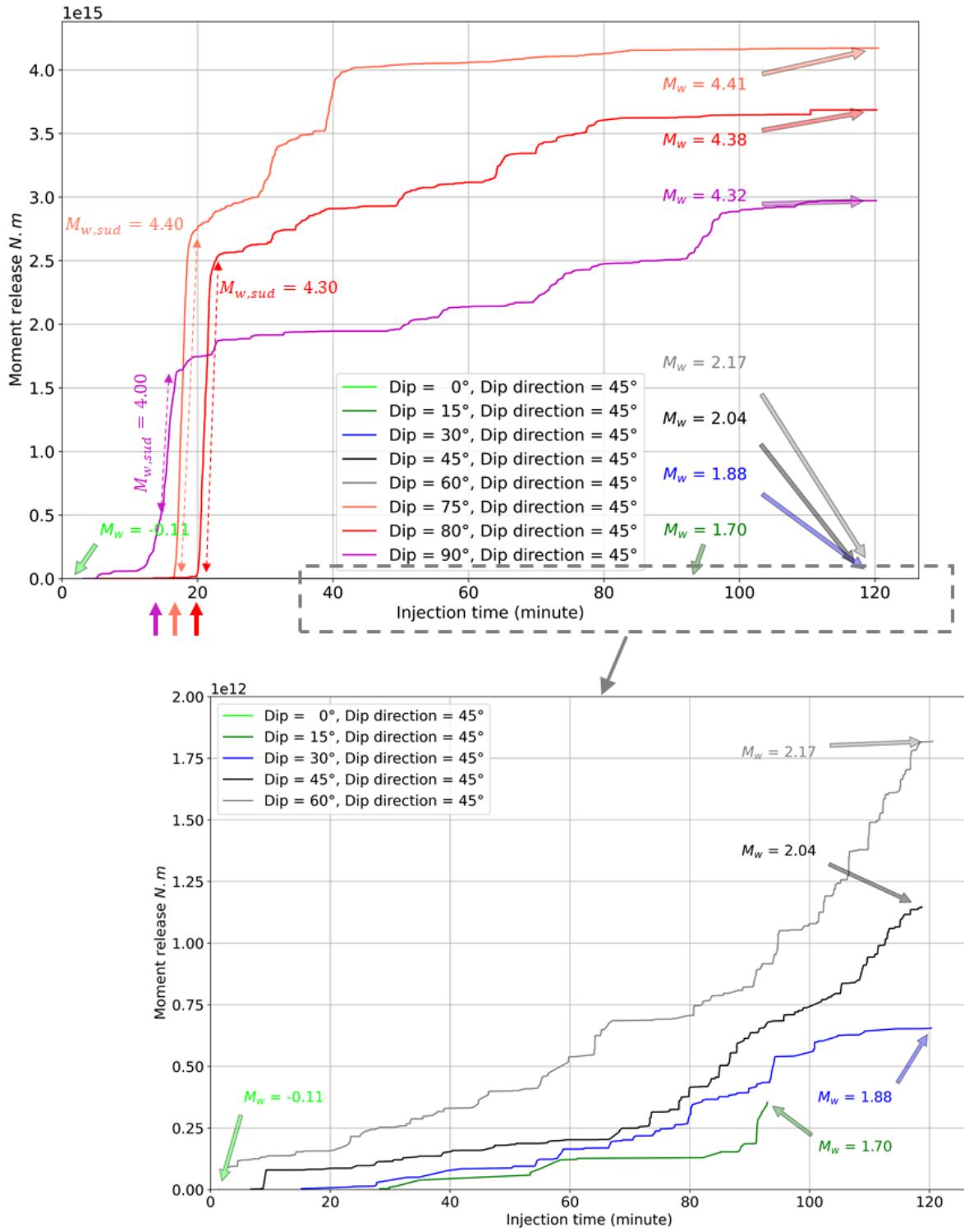


Figure 67 – Moment release with fluid injection for models with a fault dipping at different angles with a dip direction of 45° to S_H , assuming a strike-slip far-field stress regime. Reported for each case is the quasi-static maximum moment magnitude (M_w), as described in the text. The dashed arrows represent the sudden energy release and corresponding moment magnitude ($M_{w,sud}$). Shown below the main plot is one where the Moment Release scale bar is three orders of magnitude smaller to show the details of the smaller events.

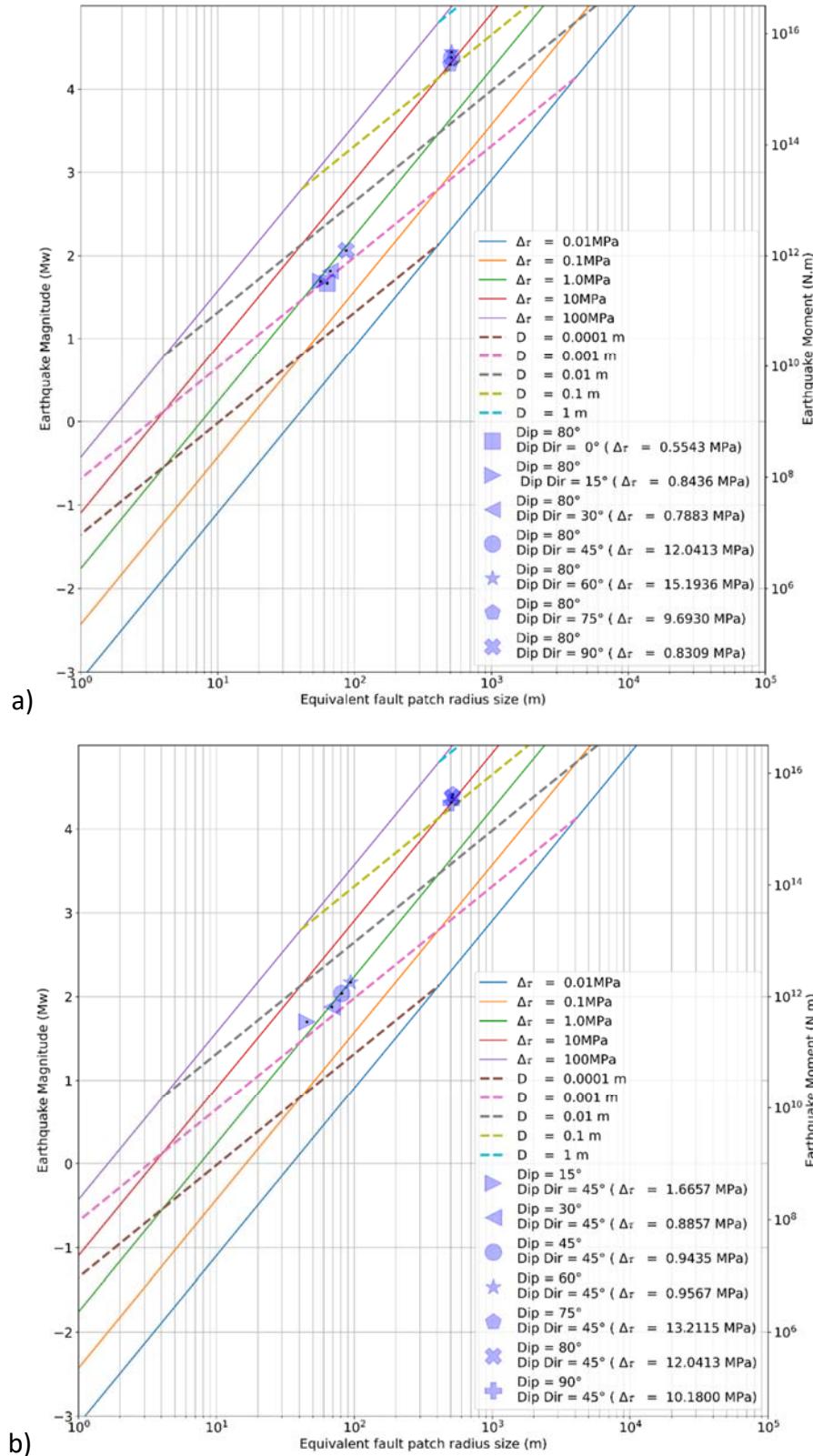


Figure 68 – Average stress drop ($\Delta\tau$) and shear displacement (D), and corresponding event magnitude, derived from simulations with varying fault orientations within a strike-slip far-field stress regime. These are based on the maximum induced seismicity event after 120 minutes of fluid injection.

It is also noteworthy that moderately dipping faults (i.e., dipping 15° to 60°) were seen to be capable of producing M_w 2 events. According to Andersonian fault mechanics, moderately dipping faults in a strike-slip stress regime would not be considered capable of generating earthquakes, at least in the context of natural earthquakes (i.e., those driven by regional tectonics and strain rates). However, in the context of induced seismicity, the modelling results show that a moderately dipping fault in a strike-slip far-field stress regime can generate seismicity if it locally experiences a pore pressure increase sufficiently large enough to trigger slip. Similar to the reverse fault far-field stress regime model results, it was found that both reverse and strike-slip focal mechanisms could be generated locally, even though the far-field stress boundary condition was strike-slip. This again was observed to depend on the fault orientation relative to the stress field and pore pressure increase interacting with the fault. The results also suggest that not all faults that trigger a M_w 2 event might be capable of triggering an even larger event with continued injection. The orientation of the fault relative to the far-field stress condition may impose an upper limit on the maximum magnitude possible. This could have significant implications with respect to developing and implementing a traffic light protocol.

Comparison to Machine Learning Results

The XSITE™ simulation results indicate that critically oriented faults in a strike-slip far-field stress regime are more likely to generate large magnitude events ($M_w \geq 4$), relative to less critically oriented faults or those involving any orientation in a reverse fault far-field stress regime. Mechanistically, the key consideration is how much of the fault is near its limit equilibrium condition and therefore more sensitive to injected fluid pressures causing slip over a large area beyond the area of the fluid pressure perturbation. A fault slip event might be triggered locally where the fluid injection pressures directly interact with a fault, but if less of the fault is in a critical state (either due to the fault orientation or far-field stress regime), the slip area and therefore the event magnitude might have a smaller upper limit.

A point of comparison to the machine learning results is with respect to the relationship between the b-value and assumed far-field stress conditions. We first acknowledge that established relationships between b-values and far-field stresses are based on focal mechanisms for natural earthquakes where the distribution of events used to calculate the b-value are derived from a large number of faults within the same tectonic region, each being critically stressed by the same tectonic strain rate under the same far-field stress regime. By extension, these faults are assumed to be similarly oriented with respect to the orientations most critical for the given far-field stress regime per Andersonian fault mechanics (i.e., strike-slip faults for a strike-slip stress regime, reverse faults for a reverse far-field stress regime, etc.), and thus the focal mechanisms are considered to be representative of the far-field stresses. Under these conditions, studies show that the b-value decreases, meaning a higher likelihood of larger magnitude events, when comparing a normal ($S_V > S_H > S_h$) to a strike-slip ($S_H > S_V > S_h$) to a reverse ($S_H > S_h > S_V$) far-field stress regime (Schorlemmer et al., 2005; Scholz, 2015; Mehrabifard & Eberhardt, 2021). For the case of induced seismicity, a clear relationship is not expected between b-value and stress regime when the b-value is calculated from events recorded within a small area, for example the

response from hydraulic fracturing operations for a single well pad. At this scale, the recorded seismicity would be restricted to that derived from a single dominant fault structure, or several structures in close proximity, that may or may not result in a focal mechanism representative of the far-field stress conditions; as previously shown, the focal mechanism can differ locally from the far-field stresses depending on the orientation of the fault being triggered during fluid injection. For the machine learning analyses, the b-values used represent an intermediate case, as they are calculated from induced seismicity events occurring within a 100 km radius of each well. By calculating the b-value over a large area, it is expected that the distribution of event magnitudes encompasses a large number of faults spanning a significant part of the Montney, albeit with possibly varying orientations. Hence, it is reasonable that the b-values calculated in this fashion could follow the same logic as the b-values for natural earthquakes with respect to serving as an indicator of the far-field stress regime.

From the machine learning regression results, the b-value of events within 100 km radius of a well was identified to be the top predictor of induced seismicity event severity (Figures 44 and 45). The SHAP plots further showed that this has a positive correlation, meaning that wells that have experienced larger magnitude events have higher calculated b-values (Figures 46 and 47). This might seem contradictory at first as higher b-values typically suggest a lower likelihood of large magnitude events (e.g., van der Elst, 2021). One explanation could be as follows. Using the b-value as an indicator for the far-field stress regime (Schorlemmer et al., 2005), the higher b-values would correspond with a strike-slip stress regime, which the numerical simulation results suggest has a greater potential to generate large magnitude events (for critically oriented faults). In other words, the smaller population of wells that have experienced more severe induced seismicity events are surrounded by a larger population of wells (within a 100 km radius) that have more generally experienced events with a magnitude distribution (i.e., b-value) characteristic of a strike-slip far-field stress regime. This is independent of the measured focal mechanism for these large magnitude events, which the numerical results show can be a reverse slip mechanism locally within a strike-slip far-field stress regime.

A second feature from the machine analysis related to the stress regime was the variation of the local orientation of S_H (observed at the well) relative to the regional trend ("Shmax_variation"). This was seen to be a top five predictor for the classification susceptibility analyses (Figures 26 and 27). The SHAP plots for these (Figures 32 and 33) show a negative correlation, meaning the likelihood of generating seismicity decreases when values for this feature increase; i.e., a higher likelihood of seismicity can be expected where the local S_H orientation is similar to the regional. However, this does not speak to the orientation of the population of faults in proximity to the wellbore and is inherently biased towards reverse faults. It is also noted that there are high values in the SHAP plots (red dots) that also correlate with higher susceptibilities (possibly attributed to strike-slip stress conditions). With respect to the regression severity analyses, Shmax_variation was again seen to be an important feature (Figures 44 and 45), but in this case, the SHAP plots show a positive correlation (Figures 46 and 47). This suggests an increased likelihood of larger magnitude events when the local orientation of S_H deviates from the regional. One interpretation

is that as the orientation of S_H deviates from the regional orientation, and assuming the population of faults are randomly oriented and not necessarily aligned with the present day tectonic stresses, the potential for strike-slip faulting increases. The numerical simulation results suggest that sub-vertical faults with dip directions deviating from S_H become more critically stressed resulting in a higher potential for generating a larger magnitude event.

Together, the machine learning and numerical results (acknowledging the numerous model assumptions and simplifications related to the representation of the fault and geology), suggest that in a real-time seismic risk evaluation using a Traffic Light Protocol, the detection of a M_w 2 event in a strike-slip far-field stress regime should be treated differently than a M_w 2 event in a reverse fault far-field stress regime. The susceptibility of generating induced seismicity might actually be higher in a reverse fault far-field stress regime, but more limited with respect to the maximum magnitude potential. Whereas in a strike-slip far-field stress regime, the generation of smaller M_w 2 events have a higher likelihood of serving as a pre-cursor (i.e., early warning) of a possible larger M_w 4 event. Again, this finding is specific to the far-field stress regime and not the focal mechanism, which in the case of fluid-injection induced seismicity, as observed in the numerical simulation results, can include a reverse fault slip event occurring locally in a strike-slip far-field stress regime.

9. Conclusions, Recommendations and Future Work

9.1. Conclusions

Machine Learning

The application of machine learning was investigated for the purpose of generating induced seismicity susceptibility maps and ranking the influence of relevant geological and operational features on the likelihood (classification of wells being seismogenic or not) and severity (maximum magnitude regression) of induced seismicity events. For this purpose, a comprehensive database was compiled using both publically available and proprietary data for the Montney region of northeastern British Columbia. From this database, datasets were generated specific to the analyses performed: 1) a classification feature importance analysis, based on a combined dataset of all geological and operational features at each hydraulic fracturing well; 2) a classification susceptibility analysis, based on a restricted dataset limited to geological features and those operational features that apply to susceptibility; 3) a classification susceptibility mapping analysis, based on the restricted dataset of geological and susceptibility-relevant features interpolated to a 2.5 km grid spacing; and 4) a regression severity analysis, based on a combined dataset of geological and operational features that included the maximum magnitude of the induced seismicity events associated with each well. Eight different machine learning algorithms were tested for these analyses. The machine learning workflow included training and tuning (i.e., calibrating) the models through iteration against 75% of the total, and then testing the model performance against the remaining 25% of the data unseen by the trained models.

For the classification analyses (i.e., likelihood of wells being seismogenic), the Light GBM, XGBoost and Random Forest machine learning models were observed to be the top performers for both the feature importance and susceptibility analyses. The results of the feature importance analysis, which considered all features, ranked the depth to the top of the basement as being the most important predictors 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. The most important operational feature was shown to be the total injection volume, which had a positive correlation; higher injection volumes increases the likelihood of a well being seismogenic. Overall, the geological features were found to generally rank higher in importance than the operational features. For the susceptibility analysis, which considered the geological features and only operational features relevant to susceptibility, the results showed the most important predictors to be the depth to the top of the basement (negative correlation), magnitude of the minimum horizontal stress (positive correlation), azimuth difference between the local and regional orientations of the maximum horizontal stress (negative correlation), and depth to the top of the Montney (negative correlation).

These findings informed the selection of the best performing classification models (Light GBM and Random Forest), which were then applied to the susceptibility-relevant feature data interpolated to the Montney grid to generate induced-seismicity maps. Two sets of maps were generated for each model, one relative to a minimum moment magnitude (M_w) of 1.1 and one for a minimum M_w of 2.0. Overall, the locations of susceptible areas in both models agree with the location of induced seismicity events. The Light GBM susceptibility map was seen to show areas with probabilities greater than 10% as being more concentrated around identified hotspots, whereas the Random Forest susceptibility map showed more dispersed probabilities.

For the regression analyses (i.e., severity potential with respect to maximum magnitude of the induced seismicity events associated with a hydraulic fracturing well), the top performing machine learning models were found to be Light GBM and Random Forest. The feature importance results ranked the b-value of events (within 100 km of the well) as being the most important predictor of severity, having a positive correlation. Other important predictors of severity were found to include standard deviation of pore pressure gradient (positive correlation), distance of well from the Cordilleran thrust belt (negative correlation), number of active completion days (negative correlation) and the thickness of the Montney formation (negative correlation).

Investigating the correlations found in the likelihood and severity analyses further, some of these appear to be straight forward. A positive correlation between the standard deviation of the pore pressure gradient and increased likelihood and severity of induced seismicity events, 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 expectation of fault density increasing towards the Cordilleran thrust belt and therefore the closer a well is to the thrust belt the higher the likelihood of a hydraulic fracturing operation encountering a critically stressed fault.

Other correlations are less clear, including for some of the top predictors. For example, the negative correlation between depth to the top of the basement and induced seismicity susceptibility suggests that shallower depths to the basement increases the likelihood of a well being seismogenic. This is possibly explained as the distance between the injection well and basement top increasing, resulting in a reduced likelihood of the fluid injection pressure front interacting with basement faults. However, the correlation does not speak to the depth of the injection wells or the influence of higher stresses for deeper wells on susceptibility. The positive correlation between higher b-values and higher event magnitudes is counterintuitive. Typically, lower b-values suggest a higher likelihood of a large magnitude event.

Overall, it is recognized that machine learning analyses only find data relationships and correlations, but does not connect these to cause and effect mechanisms. Models must be properly vetted for data errors, biases, censoring or overfitting, and results considered with caution. It was found that nonlinear machine learning algorithms far outperformed the linear algorithms, suggesting a nonlinear nature of the key parameters influencing induced seismicity.

The Light GBM model consistently performed better than the other models tested. This was attributed to the Light GBM decision tree algorithm splitting its tree leaf-wise following the best fit, whereas other boosting algorithms like XGBoost split the tree depth-wise or level-wise. In doing so, the leaf-wise algorithm reduces more loss than the level-wise algorithm, and hence results in better accuracy. It was also surprisingly fast, hence the reference to ‘Light’

Advanced Numerical Modelling

To improve the interpretability of the machine learning results, advanced 3-D numerical modelling was performed to investigate its utility in providing mechanistic cause and effect insights into correlations that are unclear or counterintuitive. Numerical modelling was performed using the 3-D hydraulic fracturing simulator XSsite™. For the negative correlation between basement depth and induced seismicity susceptibility, the modelling results show that the extent of the injected fluid pressure invaded zone is stress (i.e., depth) sensitive, where for deeper basement and injection scenarios, the corresponding higher stresses act to limit the extent of the invaded zone transmitted outwards from the injection point and hydraulic fracture. For shallower basement depths, the lower stress environment results in a significantly larger invaded zone, and thus, as the distance to the top of the basement becomes smaller (i.e., lower stresses), fluid injection results in a larger volume of elevated pore pressures, which in turn results in an increased likelihood of interacting with a critically stressed fault in the basement to trigger an induced seismicity event.

With respect to severity, the numerical modelling results showed that faults in stiffer rock masses are able to store more strain energy and therefore are more likely to produce larger magnitude events. This can be connected to machine learning features like P-wave velocity, which showed a positive correlation with event magnitude. Higher P-wave velocities correspond with stronger and stiffer rock. The same numerical simulations also suggest that induced seismicity will be observed earlier for fault slips in stiffer rock. Similarly, the machine learning outcomes revealed that larger magnitude events are associated with completions involving a smaller number of active days. Larger magnitude events were also observed for simulations assuming a strike-slip far-field stress regime compared to a reverse fault far-field stress regime. Collectively, these results demonstrate the value of numerical modelling in providing insights when multiple factors contribute to the mechanistic response, in this case the dependence of induced seismicity magnitude on rock mass stiffness, in-situ stress regime and fluid injection time. For example, a transitional condition was observed for an intermediary case of rock mass stiffness ($E = 32 \text{ GPa}$), where the fault was seen to be less critically stressed compared to the higher stiffness cases, allowing a longer fluid injection time. This in turn allowed more fault area to see an increase in fluid pressure (and decrease in effective stress), which when combined with the intermediary stiffness still allowing for considerable strain energy to be stored, resulted in a larger magnitude event than was observed for the stiffer rock mass case of $E = 48 \text{ GPa}$.

Numerical modelling results were also presented with respect to the influence of fault orientation. These results showed that critically oriented faults in a strike-slip far-field stress regime are more likely to produce large magnitude events ($M_w \geq 4$), whereas those in a reverse fault far-field stress regime may be more limited with respect to the maximum magnitude possible. This was seen to relate to how the far-field stresses transform onto the fault surface and how much of the resulting fault surface area is critically stressed. The simulation results showed that for adversely oriented faults in a strike-slip far-field stress regime, more of the fault surface is initially near its limit equilibrium condition and therefore these faults are more sensitive to injected fluid pressures causing slip over a large area beyond the area of the fluid pressure perturbation. Simulation results show for other scenarios that a fault slip event might be triggered locally where the fluid injection pressures directly interact with the fault, but if less of the fault is in a critical state (either due to the fault orientation or far-field stress regime), the slip area and therefore the event magnitude might have a smaller upper limit. This was seen to be the case for faults modelled assuming a reverse fault far-field stress regime.

These findings helped to explain the machine learning results that counterintuitively correlated higher b-values with increased induced seismicity event magnitudes. Expected b-value trends are related to natural earthquakes where the key driver is the tectonic environment and only faults that are critically stressed are seismogenic. In such an environment, both strike-slip and reverse faults are critically stressed over large areas (this is dependent on depth, with natural reverse fault earthquakes generally occurring at much greater depths than natural strike-slip earthquakes). In the case of induced seismicity, if sampling event magnitudes over a large area (e.g., 100 km radius of a well), the b-value calculated might likewise be indicative of the far-field in-situ stress regime with higher b-values corresponding with a strike-slip far-field stress regime and lower b-values with a reverse fault far-field stress regime. Within these distributions, the modelling results show that it is possible to generate a large magnitude induced seismicity event with a reverse fault focal mechanism (locally) but in a strike-slip far-field stress regime (regionally). This depends on the orientation of the fault experiencing the pore pressure increase, relative to the orientation of the far-field stress regime.

Together, the machine learning and numerical modelling results suggest that if operating in a strike-slip stress regime in stiffer rocks and there is a lack of early, significant seismicity (of $M_w > 1$ or 2) with increasing injection time, then this might be a sign that strain energy is accumulating towards a potentially large event ($M_w \geq 4$). Similarly, in the context of a Traffic Light Protocol, the detection of a $M_w 2$ event might be more concerning when the formation targeted by the injection, or adjacent to it, is stiffer (e.g., $E > 32$ GPa) and if operating in a strike-slip far-field stress regime.

9.2. Recommendations/Guidelines

It was demonstrated that the quantity and quality of data is of central importance in any empirical analysis, especially in the training and interpretation of machine learning models. As

unconventional gas development in northeastern British Columbia and the Montney continues, the continuous collection of geology, stress regime and operational data will create opportunities to further improve the machine learning results and interpretations presented in this report, leading to updating of the induced seismicity susceptibility maps generated.

Based on the experiences gained from this study, and limitations encountered with respect to data quantity and quality, the following recommendations are made as guidelines with respect to using machine learning to generate induced seismicity susceptibility maps:

- *Data Quality Assurance and Control (QA/QC):* Both publicly available and proprietary datasets contain missing data and/or outliers. Care must be taken to remove outliers and fill in missing data (e.g. through extrapolation) before starting a machine learning analysis.
- *Exploratory Data Analysis:* At the start of an analysis, it is recommended that an exploratory data analysis be performed in order to obtain an initial understanding of the data and plan the next steps accordingly. This includes conducting a correlation analysis between the features and target, between the features themselves, and for the label count when performing a classification likelihood analysis.
- *Model Performance Metric:* In a binary classification analysis, the number of data points in each classification group (i.e., labels) can result in the dataset being either balanced (similar ratio between data points in each class) or unbalanced (majority of data points fall into one class). Based on the class label count appropriate model evaluation metrics should be considered. Particularly when data is imbalanced, the accuracy metric is not an appropriate option and other metrics such as sensitivity (recall), precision or F1-score are more suitable. In this analysis we used F1-score as it considers both sensitivity and precision.
- *Algorithm Selection:* There is no single machine learning algorithm (i.e., model) that is always suitable for either likelihood classification or severity regression. Several different models should be tested and evaluated as the performance of models depend on many factors such as data quantity, quality and type. In this study, for each analysis eight different algorithms were tested and the best performing models were selected for validation testing and interpreting the results.
- *Validation Testing:* Validating the algorithms and testing the model performance against unseen data is an essential step in machine learning analyses. As was demonstrated in this report for the susceptibility likelihood analysis, the XGBoost model was seen to perform well on the training data but not the unseen test data and therefore was omitted from the rest of this specific analysis.
- *Nonlinear Machine Learning Algorithms:* The Support Vector Classifier showed good performance when investigating induced seismicity likelihood for the combined

geological and operational dataset. However, because this model uses a nonlinear kernel, the interpretation of its results is time consuming, limiting its use to that for a subset of the data. Using this model is not recommended unless there is a specific need that requires it.

- *Susceptibility Maps:* When developing induced seismicity susceptibility maps, it is recommended to use more than one top performing algorithm. Different algorithms are sensitive to different types of data and/or reduce model error by reducing either bias or variance. Therefore, comparing the output from different models adds more confidence to the generated maps. It should also be understood that susceptibility is a spatial probability term and is different from hazard. Induced seismicity susceptibility maps derived from machine learning analyses therefore communicate the spatial likelihood of a well at a specific location being seismogenic. Seismogenic is relative to a minimum moment magnitude, which is based on the magnitude of completeness (i.e., lower cut off) of the data used in the analysis. As was shown in this study, using all data available resulted in susceptibility maps for events greater than M_w 1.1; but the data was also filtered to a higher magnitude threshold to generate susceptibility maps for events greater than M_w 2.0.

9.3. Future Works

Several opportunities for future work have been identified, to address limitations or to extend the analyses presented in this report. These include:

- Proprietary databases like geoLOGICS's geoSCOUT potentially contain additional yet to be investigated data that might prove to be significant. For example, through access we were granted by geologic to their geoSCOUT database, we were able to extract data from several key geophysical well logs. This data represents a direct measure of the reservoir characteristics and conditions, and to our knowledge, this is the first time such data have been included in a machine learning analysis. However, our use was not complete as the data available is substantial and extracting it time consuming. A thorough vetting of such sources for additional features to be considered would help to extend this work with respect to improving confidence and reliability in the susceptibility maps generated.
- It is recognized that without fluid injection there would not be any induced seismicity. The features included in the analyses presented were limited to those that are readily available and considered most important (e.g., injection volume, rate and pressure). Further work should more thoroughly investigate the influence of different injection scenarios on induced seismicity susceptibility.

- An extension of the numerical modelling work presented would be to develop and apply a mechanistically-informed, physics-based feature selection process in the development of improved induced seismicity susceptibility maps.
- The work presented and maps generated focussed specifically on spatial susceptibility of induced seismicity likelihood. The analyses related to severity were limited to feature importance and correlations. A clear next step in this work would be to extend the analysis to hazard and risk mapping. For example, a hazard map could be generated by linking the influence of operational features to susceptibility and severity to develop a hazard map that speaks to the maximum magnitude potential.

Acknowledgments

The authors wish to thank Geoscience BC for their funding of this work. We also acknowledge and thank geoLOGIC systems Ltd. for their significant contribution of data and software used in this study. All geoLOGIC systems Ltd. data and software is ©2021.

In addition, the authors would like to thank Enlighten Geosciences Ltd. for providing the Phanerozoic Isopach data as well as their earthquake database and assistance, Petrel Robertson Consulting Ltd. for providing the regional mapped fault data, and the B.C. Oil and Gas Commission who provided access to well survey data and data on the depths of key formations. Valuable insights regarding different datasets were gained through discussions with Dr. Mostafa Gorjani, Dr. Pat McLellan, Mr. Stuart Venables, Dr. Honn Ko and Dr. Paulina Woźniakowska. We would especially like to thank Mr. Scott McKean for providing his insights and experience related to machine learning.

The authors also wish to acknowledge and thank the Itasca Consulting Group Inc. for providing access to their XSite™ hydraulic fracturing software, and especially Dr. Branko Damjanac and Maurilio Torres for providing technical support. We would also like to thank Golder Associates for providing access to their FracMan™ software, and Dr. Steve Rogers for his technical support and insights.

References

- Abdelaziz, A., Ha, J., Khair, H. A., Adams, M., Tan, C. P., Musa, I. H. & Grasselli, G. (2019). Unconventional shale hydraulic fracturing under true triaxial laboratory conditions, the value of understanding your reservoir. In *Proceedings, SPE Annual Technical Conference and Exhibition, Calgary*. Society of Petroleum Engineers, SPE-196026-MS. <https://doi.org/10.2118/196026-ms>
- AER (2015). Subsurface Order No. 2. Alberta Energy Regulator. Issued February 19, 2015. <https://static.aer.ca/prd/documents/orders/subsurface-orders/SO2.pdf>
- Allen, D., Eberhardt, E. & Bustin, A. (2019). Scientific Review of Hydraulic Fracturing in British Columbia. Report by the Scientific Hydraulic Fracturing Review Panel to the B.C. Minister of Energy, Mines and Petroleum Resources, 236 pp. https://www2.gov.bc.ca/assets/gov/farming-natural-resources-and-industry/natural-gas-oil/responsible-oil-gas-development/scientific_hydraulic_fracturing_review_panel_final_report.pdf
- Amini, A. (2020). Investigation of Induced Seismicity Mechanisms and Magnitude Distributions Under Different Stress Regimes, Geomechanical Factors, and Fluid Injection Parameters. PhD Thesis, Geological Engineering, The University of British Columbia. <https://open.library.ubc.ca/cIRcle/collections/ubctheses/24/items/1.0394963>
- Amini, A. & Eberhardt, E. (2019). Influence of tectonic stress regime on the magnitude distribution of induced seismicity events related to hydraulic fracturing. *Journal of Petroleum Science and Engineering*, 182. <https://doi.org/10.1016/j.petrol.2019.106284>
- Amini, A. & Eberhardt, E. (2021). Machine-learning analysis of factors influencing induced seismicity susceptibility in the Montney play area, northeastern British Columbia (NTS 093P, 094A, B, G, H). In *Geoscience BC Summary of Activities 2020: Energy and Water*. Geoscience BC, Report 2021-02, pp. 45–56. http://www.geosciencebc.com/i/pdf/SummaryofActivities2020/EW/Project%202019-014_EWSOA2020.pdf
- Anderson, E.M. (1951). The dynamics of faulting and dyke formation: with applications to Britain. *Nature*, 149: 651-652. <https://doi.org/10.1038/149651b0>
- Babaie Mahani, A., Esfahani, F., Kao, H., Gaucher, M., Hayes, M., Visser, R. & Venables, S. (2020). A systematic study of earthquake source mechanism and regional stress field in the southern Montney unconventional play of northeast British Columbia, Canada. *Seismological Research Letters*, 91(1): 195–206. <https://doi.org/10.1785/0220190230>
- Boehmke, B. & Greenwell, B. (2019). Hands-On Machine Learning with R. Chapman and Hall/CRC. <https://doi.org/10.1201/9780367816377>
- Breiman, L. (2001). Random Forests. *Machine Learning*, 45(1): 5–32. <https://doi.org/10.1023/A:1010933404324>

Chen, Q. & Wong, R. C. K. (2018). Mode I and Mode II Fracture Toughness of Montney Shale. In *52nd U.S. Rock Mechanics/Geomechanics Symposium, Seattle*, ARMA-2018-127.

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*, pp. 785–794. <https://doi.org/10.1145/2939672.2939785>

Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3): 273–297. <https://doi.org/10.1007/BF00994018>

Davies, G.R., Watson, N., Moslow, T.F. & MacEachern, J.A. (2018). Regional subdivisions, sequences, correlations and facies relationships of the Lower Triassic Montney Formation, west-central Alberta to northeastern British Columbia, Canada — with emphasis on role of paleostructure. *Bulletin of Canadian Petroleum Geology*, 66(1): 23–92.

Damjanac, B., Detournay, C. & Cundall, P. A. (2016). Application of particle and lattice codes to simulation of hydraulic fracturing. *Computational Particle Mechanics*, 3(2): 249–261. <https://doi.org/10.1007/s40571-015-0085-0>

Damjanac, B., Detournay, C. & Cundall, P. (2020). Numerical simulation of hydraulically driven fractures. In *Modelling Rock Fracturing Processes*, Springer International Publishing, pp. 531–561. https://doi.org/10.1007/978-3-030-35525-8_20

Dixon, J. (2009). The Lower Triassic Shale member of the Montney Formation in the subsurface of northeast British Columbia. Geological Survey of Canada, Open File 6274. <https://doi.org/10.4095/248223>

DPR (2021). Drilling and Production Regulation. British Columbia Oil and Gas Activities Act, Regulation 282/2010. https://www.bclaws.gov.bc.ca/civix/document/id/complete/statreg/282_2010

Eaton, D.W., Mahani, A., Salvage, R., Kao, H. & Van Der Baan, M. (2019). A tale of three earthquakes: new insights into fault activation in light of recent occurrences in western Canada. In *New Directions in Geosciences for Unconventional Resources, Banff*, pp. 1–11.

Eaton, D.W. & Schultz, R. (2018). Increased likelihood of induced seismicity in highly overpressured shale formations. *Geophysical Journal International*, 214(1): 751–757. <https://doi.org/10.1093/gji/ggy167>

Edwards, D.E., Barclay, J.E., Gibson, D.W., Kvill, G.E. & Halton, E. (1994). Triassic Strata of the Western Canada Sedimentary Basin. In *Geological Atlas of the Western Canada Sedimentary Basin*, Canadian Society of Petroleum Geologists and Alberta Research Council. https://ags.aer.ca/document/Atlas/chapter_16.pdf

Fasola, S.L., Brudzinski, M.R., Skoumal, R.J., Langenkamp, T., Currie, B.S. & Smart, K.J. (2019). Hydraulic fracture injection strategy influences the probability of earthquakes in the Eagle Ford

shale play of South Texas. *Geophysical Research Letters*, 46, 12958-12967.
<https://doi.org/10.1029/2019GL085167>

Fereidoni, A. & Cui, L. (2015). Composite Alberta Seismicity Catalog.

Fox, A., McKean, S.H. & Watson, N. (2020). Statistical Assessment of Operational Risks for Induced Seismicity from Hydraulic Fracturing in the BC Montney. Geoscience BC Project 2019-008. <http://www.geosciencebc.com/projects/2019-008/>

Fox, A.D. & Watson, N.D. (2019). Induced seismicity study in the Kiskatinaw Seismic Monitoring and Mitigation Area, British Columbia. Technical Report, BC Oil and Gas Commission.
<https://www.bcofc.ca/files/reports/Technical-Reports/final-report-enlighten-geoscience-kssma-phase-1-study-2019w-appendices-links.pdf>

Friedman, J.H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29(5): 1189-1232.

Gibson, D.W. & Barclay, J.E. (1989). Middle Absaroka sequence, the Triassic stable craton. In Ricketts (ed.), *Western Canada Sedimentary Basin: A Case History*. Canadian Society of Petroleum Geologists, pp. 219-231.

Goldstein, A., Kapelner, A., Bleich, J. & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1): 44-65. <https://doi.org/10.1080/10618600.2014.907095>

Goodfellow, I., Bengio, Y. & Courville, A. (2016). Deep Learning. MIT Press.

Grasby, S.E., Allen, D.M., Bell, S., Chen, Z., Ferguson, G., Jessop, A., Kelman, M., Ko, M., Majorowicz, J., Moore, M., Raymond, J. & Therrien, R. (2011). Geothermal energy resource potential of Canada. Geological Survey of Canada, Open File 6914.
<https://doi.org/10.4095/288745>

Häring, M., Schanz, U., Ladner, F. & Dyer, B. (2008). Characterisation of the Basel 1 enhanced geothermal system. *Geothermics*, 37: 469-495.
<https://doi.org/10.1016/j.geothermics.2008.06.002>

Hayes, B.J., Anderson, J.H., Cooper, M., McLellan, P.J., Rostron, B. & Clarke, J. (2021). Wastewater disposal in the maturing Montney play fairway, northeastern British Columbia (NTS 093P, 094A, B, G, H). In *Geoscience BC Summary of Activities 2020: Energy and Water*, Geoscience BC, Report 2021-02, pp. 91-102.
http://www.geosciencebc.com/i/pdf/SummaryofActivities2020/EW/Project%202019-004_EWSOA2020.pdf

Heidbach, O., Rajabi, M., Cui, X., Fuchs, K., Müller, B., Reinecker, J., Reiter, K., Tingay, M., Wenzel, F., Xie, F., Ziegler, M. O., Zoback, M. Lou & Zoback, M. (2018). The World Stress Map

database release 2016: Crustal stress pattern across scales. *Tectonophysics*, 744: 484–498.

<https://doi.org/10.1016/j.tecto.2018.07.007>

Javaheri, A., Dehghanpour, H. & Wood, J.M. (2017). Tight rock wettability and its relationship to other petrophysical properties: A Montney case study. *Journal of Earth Science*, 28(2): 381–390. <https://doi.org/10.1007/s12583-017-0725-9>

Kao, H., Hyndman, R., Jiang, Y., Visser, R., Smith, B., Babaie Mahani, A., Leonard, L., Ghofrani, H. & He, J. (2018). Induced seismicity in western Canada linked to tectonic strain rate: implications for regional seismic hazard. *Geophysical Research Letters*, 45(20): 11,104–11,115.

<https://doi.org/10.1029/2018GL079288>

Ke, G., Meng, Q., Finley, T. & Wang, T. (2017). LightGBM: A Highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems (NIPS 2017)*, 30: 3149–3157.

<https://dl.acm.org/doi/10.5555/3294996.3295074>

Lundberg, S.M & Lee, S.-I. (2017). A unified approach to interpreting model predictions.

Advances in Neural Information Processing Systems (NIPS 2017), 30: 4768–4777.

<https://dl.acm.org/doi/10.5555/3295222.3295230>

Mahani, A.B., Kao, H., Atkinson, G.M., Assatourians, K., Addo, K. & Liu, Y. (2019). Ground-motion characteristics of the 30 November 2018 injection-induced earthquake sequence in northeast British Columbia, Canada. *Seismological Research Letters*, 90(4): 1457–1467.

<https://doi.org/10.1785/0220190040>

Mahani, A.B., Schultz, R., Kao, H., Walker, D., Johnson, J. & Salas, C. (2017). Fluid injection and seismic activity in the Northern Montney Play, British Columbia, Canada, with special reference to the 17 August 2015 Mw 4.6 induced earthquake. *Bulletin of the Seismological Society of America*, 107(2): 542–552. <https://doi.org/10.1785/0120160175>

McGarr, A. (2014). Maximum magnitude earthquakes induced by fluid injection. *Journal of Geophysical Research: Solid Earth*, 119(2): 1008–1019. <https://doi.org/10.1002/2013JB010597>

McKean, S.H. & Priest, J.A. (2019). Multiple failure state triaxial testing of the Montney Formation. *Journal of Petroleum Science and Engineering*, 173: 122–135.

<https://doi.org/10.1016/J.PETROL.2018.09.004>

McLellan, P. (2016). Geomechanical characterization of the Farrell Creek Montney Reservoir, Northeast British Columbia. In *SPE Workshop: The Montney Play, Calgary*.

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. In *Geoscience BC Summary of Activities 2020: Energy and Water*, Geoscience BC, Report 2021-02, pp. 57–66.

http://www.geosciencebc.com/i/pdf/SummaryofActivities2020/EW/Sch_Mehrabifard_EWSOA_2020.pdf

Mossop, G.D. & Shetsen, I., comp. (1994). Geological Atlas of the Western Canada Sedimentary Basin. Canadian Society of Petroleum Geologists and Alberta Research Council.

<https://ags.aer.ca/reports/atlas-western-canada-sedimentary-basin>

National Energy Board (2013). The Ultimate Potential for Unconventional Petroleum from the Montney Formation of British Columbia and Alberta - Energy Briefing Note. <https://www.cer-rec.gc.ca/en/data-analysis/energy-commodities/natural-gas/report/ultimate-potential-montney-formation/the-ultimate-potential-unconventional-petroleum-from-montney-formation-british-columbia-alberta-energy-briefing-note.pdf>

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. *Geophysical Research Letters*, 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, E. (2011). Scikit-learn: machine learning in Python. *Journal of Machine Learning Research*, 12: 2825–2830. <https://doi.org/10.1145/2786984.2786995>

Peña Castro, A.F., Roth, M.P., Verdecchia, A., Onwuemeka, J., Liu, Y., Harrington, R.M., Zhang, Y. & Kao, H. (2020). Stress chatter via fluid flow and fault slip in a hydraulic fracturing-induced earthquake sequence in the Montney Formation, British Columbia. *Geophysical Research Letters*, 47(14): e2020GL087254. <https://doi.org/10.1029/2020GL087254>

Pierce, M., Cundall, P. & Potyondy, D. (2007). A synthetic rock mass model for jointed rock. In Proceedings of the 1st Canada–U.S. Rock Mechanics Symposium, Vancouver, pp. 341–349.

Potyondy, D. O., & Cundall, P. A. (2004). A bonded-particle model for rock. *International Journal of Rock Mechanics and Mining Sciences*, 41(8): 1329–1364.
<https://doi.org/10.1016/j.ijrmms.2004.09.011>

Rogers, S., McLellan, P. & Webb, G. (2015). Discrete fracture network modelling of hydraulic fracturing in a structurally controlled area of the Montney Formation, BC. In *GeoConvention 2015: New Horizons*, 1: 1–5.

Rogers, S. & Patrick, M. (2014). Investigation of the effects of natural fractures and faults on hydraulic fracturing in the Montney Formation, Farrell Creek Gas Field, British Columbia. In *International Discrete Fracture Network Engineering Conference, Vancouver*, pp. 1–13.

Scholz, C.H. (2015). On the stress dependence of the earthquake b value. *Geophysical Research Letters*, 42(5): 1399–1402. <https://doi.org/10.1002/2014GL062863>

Scholz, C.H. (2019). Mechanics of Earthquakes. Cambridge University Press.

Schorlemmer, D., Wiemer, S. & Wyss, M. (2005). Variations in earthquake-size distribution across different stress regimes. *Nature*, 437(7058): 539–542.

<https://doi.org/10.1038/nature04094>

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*, 359(6373): 304–308. <https://doi.org/10.1126/science.aao0159>

Shapiro, S.A. & Dinske, C. (2021). Stress drop, seismogenic index and fault cohesion of fluid-induced earthquakes. *Rock Mechanics and Rock Engineering*. <https://doi.org/10.1007/s00603-021-02420-3>

Stein, S. & Wysession, M. (2003). An Introduction to Seismology, Earthquakes, and Earth Structure. Blackwell Publishing.

Trican Geological Solutions, Ltd. (2013). Geomechanical Testing of Talisman Hz Altares A1-32-83-25W6. Prepared for Talisman Energy Inc. by Trican Geological Solutions Ltd.

Vaisblat, N., Shokri, A.R., Ayrancı, K., Harris, N. & Chalaturnyk, R.J. (2019). Significance of rock compositional control on geomechanical properties and hydraulic fracturing of the Montney formation, Western Canadian Basin. In *SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference 2019 (APUR 2019)*. <https://doi.org/10.15530/ap-urtec-2019-198199>

van der Elst, N. J. (2021). B-positive: a robust estimator of aftershock magnitude distribution in transiently incomplete catalogs. *Journal of Geophysical Research: Solid Earth*, 126(2), e2020JB021027. <https://doi.org/10.1029/2020JB021027>

Verdon, J.P. & Bommer, J.J. (2021). Green, yellow, red, or out of the blue? An assessment of traffic light schemes to mitigate the impact of hydraulic fracturing-induced seismicity. *Journal of Seismology*, 25: 301–326. <https://doi.org/10.1007/s10950-020-09966-9>

Vishkai, M. & Gates, I. (2019). On multistage hydraulic fracturing in tight gas reservoirs: Montney Formation, Alberta, Canada. *Journal of Petroleum Science and Engineering*, 174: 1127–1141. <https://doi.org/10.1016/j.petrol.2018.12.020>

Visser, R., Smith, B., Kao, H., Babaie Mahani, A., Hutchinson, J. & McKay, J.E. (2017). A comprehensive earthquake catalogue for northeastern British Columbia and western Alberta, 2014–2016. Geological Survey of Canada, Open File, 8335. <https://doi.org/10.4095/306292>

Walker, S.H. & Duncan, D.B. (1967). Estimation of the probability of an event as a function of several independent variables. *Biometrika*, 54(1/2): 167–179.
<http://www.jstor.org/stable/2333860>

Walters, R.J., Zoback, M.D., Baker, J.W. & Beroza, G.C. (2015). Characterizing and responding to seismic risk associated with earthquakes potentially triggered by fluid disposal and hydraulic

fracturing. *Seismological Research Letters*, 86(4): 1110–1118.

<https://doi.org/10.1785/0220150048>

Wozniakowska, P. & Eaton, D.W. (2020). Machine learning-based analysis of geological susceptibility to induced seismicity in the Montney Formation, Canada. *Geophysical Research Letters*, 47(22): e2020GL089651. <https://doi.org/10.1029/2020GL089651>

Zoback, M.D. (2007). Reservoir Geomechanics. Cambridge University Press.

<https://doi.org/10.1017/CBO9780511586477>

Zoback, M. & Kohli, A. (2019). Unconventional Reservoir Geomechanics: Shale gas, Tight oil, and Induced Seismicity. Cambridge University Press.

Appendix A - Feature Definitions

The following describes all features included in one or more of the different machine learning analyses. Each feature is listed with its abbreviation and a short description, together with other relevant information such as the rationale for its consideration, the number of missing features, and relevant references. The number of missing features where reported are relative to a total of 6530 wells and 95,807 stages in the database compiled. For easy reference, the features are listed in alphabetical order according to the abbreviations used in the figures in the report.

Abbreviation	Feature Name (units)	Description
Active_Days	Well - number of active days (d)	Total number of active completion days. (See Verdon and Bommer, 2020).
Avg_Frac_Spacing	Well - avg frac spacing (m)	Average spacing between stages at each well. Data for 2609 wells (of 6530) are missing. (See Fox et al., 2020).
Avg_Inj_Rate	Well – avg injection rate (m^3/min)	The average injection rate for the well. Calculated by averaging the fluid injection rates for each stage. Data for 10,514 stages (of 95,807) are missing. (See Schultz et al., 2018).
Avg_prop/Meter	Well - avg proppant placed per meter (t/m)	The average proppant injected in the well normalized to the well's completion length. (See Fox et al., 2020).
Avg_prop/stage	Well - avg proppant placed per stage (t/stage)	The average proppant injected per stage. Calculated from dividing the total proppant per well by the number of stages. Data for 1074 wells (of 6530) are missing. (See Fox et al., 2020).
Avg_Vol/Meter	Well - avg fluid pumped per meter (m^3/m)	The average volume injected in the well normalized to the well's completion length. (See Fox et al., 2020).
Avg_Vol/Stage	Well - avg fluid pumped per stage (m^3/stage)	The average fluid volume injected per stage. Calculated from dividing the total injected volume per well by the number of stages. Data for 848 wells (of 6530) are missing. (See Fox et al., 2020).
b_value	100km_b_value	The b-value of induced seismicity events located within a 100 km radius of each well. The 100 km radius is selected to ensure a minimum of 300 events for calculating the b-value. (See Amini and Eberhardt, 2019).

Basement_Top	Precambrian_Basement_Top_TVD (m)	True vertical depth to the top of the Precambrian basement. (See Pawley et al., 2018).
BHCPv	BHCPv (m/sec)	P-wave velocity of the Montney formation at the well location. Calculated as the average value from the borehole compensated sonic logs over the Montney's depth interval. This data was accessed through geoSCOUT.
Bulk_Density	Bulk_Density (kg/m ³)	Bulk density of the Montney formation at the well location. Calculated as the average value from the borehole density logs over the Montney's depth interval. This data was accessed through geoSCOUT.
Comp_length	Well – completed length (m)	Total completed length of well, calculated from the beginning of the first stage until the end of the last stage. Data for 2436 wells (of 6530) are missing.
Dist_Base	Dist_Base (m)	Vertical distance from the injection point to the top of the Precambrian basement. (See Pawley et al., 2018).
Dist_CTB	Distance_CTB (km)	Horizontal distance from the well head to the Cordilleran thrust belt. (See Mossop and Shetsen, 1994).
Dist_Faults	Distance_Faults (km)	Shortest distance between well heads and mapped faults. (See Hayes et al., 2021).
Dist_LowerForm	Dist_Deb/Bell (m)	Vertical distance from the injection point to the top of the formation below the Montney. The formation below the Montney varies between the Debolt and Belloy. At each well, the formation that is closer to the Montney was selected as the formation below the Montney. (See Wozniakowska and Eaton, 2020).
Dist_MidMon	Dist_MidMon (m)	Vertical distance from the injection point to the top of the Middle Montney Formation. (See Wozniakowska and Eaton, 2020).
Dist_TopMon	Dist_TopMon (m)	Vertical distance from the injection point to top of the Montney Formation. (see Wozniakowska and Eaton, 2020).
Gamma_Ray	Gamma_Ray (GAPI)	Gamma ray value of the Montney formation at the well location. Calculated as the average value from the borehole gamma ray logs over the Montney's depth interval. This feature serves as a proxy for clay content in Montney. This data was accessed through geoSCOUT.

Lower_Form	Lower_Foreation_TVD (m)	True vertical depth to the top of the formation below the Montney. The formation below the Montney varies between the Debolt and Belloy. At each well, the formation that is closer to the Montney was selected as the formation below the Montney. See Wozniakowska and Eaton, 2020).
Max_Inj_P	Well – maximum injection pressure (MPa)	The maximum injection pressure for the well. Calculated by selecting the highest fluid injection pressure across all stages for the well. Data for 19,675 stages Data for 10,514 stages (of 95,807) are missing.
Max_ISIP	Max_ISIP (MPa)	Maximum instantaneous shut-in pressure for the well.
Max_Pb	Max_Pb (MPa)	Maximum breakdown pressure for the well.
Mon_Thickness	Montney_Thickness (m)	Thickness of the Montney formation calculated from the top of the Montney to the top of formation below. For this feature, only wells that all three top formations (Montney, Belloy and Debolt) were identified are considered. Next, either the Debolt or Belloy was used depending on which was closer to the Montney. Data for 5576 wells (of 6530) are missing.
Montney-Mid	Middle_Montney_TVD (m)	True vertical depth to the top of the Middle Montney.
Montney-Top	Upper_Montney_TVD (m)	True vertical depth to the top of the Upper Montney.
Number_Stages	Well – stages actual	Number of actual completed stages per well.
pp_grad	Pore Pressure Gradient (kPa/m)	Reservoir pore pressure to depth ratio. (See Pawley et al., 2018).
ppgrad_SD	Standard Deviation of Pore Pressure Gradient	Standard deviation of the reservoir pressure to depth ratio calculated between each well and the five closest pore pressure gradient values.
Seismo_index	Seismogenic_index	The seismogenic index is a statistical characteristic of the fluid-induced seismicity. This quantity is independent of injection time and of any other injection characteristics. It is completely defined by seismotectonic features of a given location. The larger the index value, the larger the probability of a significant magnitude event.
Shear_Strain	Shear_Strain	Regional shear strain field calculated from horizontal motions measured at Global

		Navigational Satellite System (GNSS) sites throughout western Canada. (See Kao et al., 2018).
Shmax_variation	Shmax Variation (degree)	Variation between local maximum horizontal stress orientation (derived from World Stress Map and Grasby et al., 2012) and regional maximum horizontal stress orientation (assumed to be N45°E).
Shmin	Shmin (MPa)	The minimum horizontal stress. This data was accessed through geoSCOUT. (See Pawley et al., 2018).
Total_Prop	Well - Total Proppant Placed (t)	The total proppant injected for the well. Data for 1074 wells (of 6530) are missing. (See Fox et al., 2020).
Total_Volume	Well - Total Fluid Pumped (m³)	The total fluid injected in the well. Calculated by summing the fluid injected for each stage. Data for 719 wells (of 6530) are missing. (See Schultz et al., 2018).
WellDensity	Well density	The number of horizontal wells within 1, 3, 5, and 10 km of the observed well. This feature serves as a proxy for well interference and the density of field development. (See Fox et al., 2020).

Appendix B - Databases and Codes

This appendix contains the list of all databases used in this report including the data description, format, file size(s) and metadata which will be provided separately from the report.

[Hydraulic Fracturing Completion Data](#)

Hydraulic fracturing data for the Montney region was sourced from geoLOGIC systems ltd.'s Well Completions & Frac Database (geoLOGIC) through their gDCweb software. This was made available through a license agreement between geoLOGIC systems ltd. and Geoscience BC. (Not publishable)

[Earthquake Catalogues](#)

Data	Description	Format	Size
Earthquake Catalogue from 2000 to Feb-17-2020	Combined from Enlighten's report and NRCan	CSV file	530 KB

[Geological Features](#)

Data	Description	Format	Size
Geological features excluding proprietary data (Depth to top of Basement, Shear Strain, Mapped faults, Bulk Density, P-wave velocity and clay content)	Raw Geological data plus interpolated data at hydraulic fracturing wells locations and at 2.5 km grid across Montney	CSV, Shape file and Raster files	500 MB

[Final Databases](#)

Data	Description	Format	Size
Well Features Database. Excluding proprietary data (Depth to top of	Interpolated geological data at hydraulic fracturing wells used for	CSV file	5 MB

Basement, Shear Strain, Mapped faults, Bulk Density, P-wave velocity and clay content)	training/testing machine learning models.		
Montney Grid Database. Excluding proprietary data (Depth to top of Basement, Shear Strain, Mapped faults, Bulk Density, P-wave velocity and clay content)	Interpolated geological data at 2.5 km grid used for generating susceptibility maps	CSV file	2 MB

Codes

Data	Description	Format	Size
Codes used in this analysis	Python codes in Jupyter notebook environment, developed for this analysis	IPYNB files	10-15 MB

High-Resolution Figures

Data	Description	Format	Size
Full size figures	Full size images used in the final report	PNG files	100 MB