Determination of Factors Controlling Geological Susceptibility to Induced Seismicity in the Montney Formation, Northeastern British Columbia and Northwestern Alberta, Based on a Machine-Learning Approach

P. Wozniakowska[1], **Department of Geoscience, University of Calgary, Calgary, Alberta, paulina.wozniakowska@ucalgary.ca**

D.W. Eaton, Department of Geoscience, University of Calgary, Calgary, Alberta

Wozniakowska, P. and Eaton, D.W. (2020): Determination of factors controlling geological susceptibility to induced seismicity in the Montney Formation, northeastern British Columbia and northwestern Alberta, based on a machine-learning approach; *in* Geoscience BC Summary of Activities 2019: Energy and Water, Geoscience BC, Report 2020-02, p. 19–26.

Introduction

A significant increase in the seismicity rate in western Canada in recent years has been associated with the development of unconventional oil and gas reserves, including hydraulic fracturing (Bao and Eaton, 2016) and saltwater disposal (Schultz et al., 2014). Because of incomplete understanding of the underlying spatial-temporal distribution, induced seismicity is a subject of extensive academic research (Baranova et al., 1999; Ellsworth, 2013; Guglielmi et al., 2015). Incomplete information and lack of continuous data hinder a full understanding of the distribution of seismic events, which is potentially linked to seismic hazard in the Western Canadian Sedimentary Basin (WCSB; Ghofrani et al., 2019). Therefore, understanding the mechanisms controlling the geological susceptibility to induced earthquakes is crucial for both seismic-hazard assessment and seismic-risk mitigation. Additionally, it is still not well understood why most hydraulic-fracturing and wastewater-disposal operations are not triggering higher magnitude earthquakes (Ellsworth, 2013). The goal of this project is to identify the most important factors controlling the occurrence of induced earthquakes in the Montney Formation, in northeastern British Columbia (BC) and northwestern Alberta. This will provide a source of additional information, which could help manage more effectively the seismicity induced by oil- and gas-production activities.

Monitoring of induced seismicity can significantly enhance the quality of hydraulic-fracturing stimulation, providing valuable information about the created fracture network and the reservoir mechanics (Eaton, 2018). Large volumes of data, in conjunction with the complex relations between different geological, physical and geomechanical characteristics, significantly hamper the correct interpretation of the phenomena observed in the subsurface. Due to its ability to discover hidden patterns, machine learning has proven to be a helpful tool for geoscientists and its popularity is growing. The focus of this paper was to assess the importance of properties controlling seismic activity during hydraulic-fracturing operations in the Montney Formation. Current knowledge was combined with existing technological advancements to test the common hypotheses about the nucleation of induced seismic events. An analysis was performed using decision-tree and random-forest algorithms, both of which are examples of supervised machine learning.

[1]*The lead author is a 2019 Geoscience BC Scholarship recipient.*

This publication is also available, free of charge, as colour digital files in Adobe Acrobat® PDF format from the Geoscience BC website: http://www.geosciencebc.com/updates/summary-of-activities/.

Methodology

Understanding the occurrence of induced seismic events requires a broad comprehension of the mechanisms in the subsurface and characteristics of the study area. In this study, geological, geomechanical and tectonic indicators were analyzed to estimate the influence of each characteristic on seismogenic potential for the analyzed wells. Parameters were extracted from the collection of compiled characteristics, including, among others, distance from the Rocky Mountain fold-and-thrust belt, distance from known lineaments, distribution of reservoir overpressures and vertical distance to the Debolt Formation (see below).

Overall, more than 6300 hydraulically fractured horizontal wells drilled into the Montney Formation were analyzed in this study. Wells were investigated in terms of their geological characteristics and whether seismicity occurred during or shortly after hydraulic fracturing. Ultimately, the final number of wells was determined by the quality and availability of data (see below).

The project consisted of the following steps:

1) Data collection and preparation:
   a) compilation of data from publicly available sources
   b) data preprocessing (data interpolation, incorrect data identification and removal)

c) labelling wells as seismogenic/nonseismogenic (binary classification)

2) Algorithm development

3) Feature importance analysis

Data Collection and Preparation

Output Values (Target) Determination

To implement the supervised machine-learning algorithm, it was necessary to decide on the appropriate output values. In this project, induced seismicity was considered as a binary-classification problem with respect to the observed seismic activity correlated in time and space with the coinciding hydraulic-fracturing operations. A similar methodology was introduced by Pawley et al. (2018), who investigated the potential for induced seismicity in the Duvernay Formation in Alberta caused by hydraulic fracturing and wastewater disposal. Their research revealed that seismogenic potential in the Duvernay Formation is controlled primarily by the vertical distance to the Precambrian basement, the state of stress in the formation (specifically, overpressure) and the value of the minimum horizontal stress. Due to significant differences in geological setting in the case of the Montney Formation and limited data availability, a similar study was performed using a slightly different set of input features in the analysis. Moreover, only horizontal wells were considered, as they are more likely to cause a change in the stress state and pore pressure near the faults (Atkinson et al., 2016), and are therefore more likely to result in higher magnitude induced earthquakes. Similarly, as in the study by Pawley et al. (2018), wells were flagged as seismogenic, when at least one seismic event of magnitude (M) >2.5 was located closer than 5 km from the well. This condition was valid only when the date of the injection preceded the occurrence of the seismic event and occurred no later than three months after the hydraulic-fracturing operation was completed. Wells which had no observable seismic activity near the hydraulic-fracturing operations were considered as nonseismogenic.

Input Values (Parameters) Preparation

Data used in this study were compiled from publicly available sources, which included geoSCOUT (geoLOGIC systems Ltd., 2019), BC Oil and Gas Commission (BC Oil and Gas Commission, 2019) and Alberta Energy Regulator (AER; Mossop and Shetsen, 1994) databases, as well as the Composite Alberta Seismicity Catalogue (Fereidoni and Cui, 2015). In total, 6315 oil- and gas-producing horizontal wells drilled into the Montney in BC and Alberta were analyzed. Earthquake data were sourced from the Canadian Induced Seismicity Collaboration (2019) website and included the seismic events registered before April 27, 2019. It is worth noting that the analyzed dataset does not include all the Montney horizontal wells, only those with complete (and validated) information.

Pressure Data

Previous literature examples demonstrate the relationship between the formation pressure and occurrence of seismic events (Eaton and Schultz, 2018). Here, pressure data were gathered using the geoSCOUT database and complemented with data published on the BCOGC website. Due to data sparsity and the limited number of measurements, the pressure-gradient values have been estimated using a radial basis function (RBF) interpolation method with linear kernel, which is included in the scipy interpolation Python module (Jones et al., 2001; van Rossum and Drake, 2003). A total of 2376 pressure measurements were used to calculate the pressure gradient; obtained pressure gradients ranged from 0.6 to about 20.6 kPa/m. Previous studies (e.g., Eaton and Schultz, 2018) have revealed that reservoir overpressure (areas where the pressure gradient exceeds a hydrostatic gradient value of 10 kPa/m) has the potential to influence the overall susceptibility to induced seismicity.

Regional Stress Regime

In addition to reservoir pore pressure, information about maximum horizontal-stress (S_{Hmax}) direction was another parameter characterizing the stress regime of the analyzed area. In some cases, fault orientation seems to be conformable with the regional stress field (Snee and Zoback, 2016), whereas examination of other studies suggests the potential for fault activation irrespective of their geometry (Zoback and Zoback, 1989). This implies that the regional stress direction may have a strong impact on the seismicity observed in the given area. The direction of S_{Hmax} has been investigated based on the values of S_{Hmax} azimuths from the World Stress Map (Heidbach et al., 2016), an open-access public database. These values were limited to the vicinity of the Montney Formation subcrop region and interpolated using RBF with linear kernel. Overall, 133 data points were used in the calculation, relying on stress-state variations instead of absolute values as the major factor. The difference between the S_{Hmax} direction and the average value of the S_{Hmax} azimuth observed in the WCSB (estimated to be about 45°) was used in the calculation.

Tectonic Data

Tectonically, the presence of pre-existing conductive thrust faults in the Precambrian crystalline basement (Zhang et al., 2013) as well as proximity to the Precambrian basement (Skoumal et al., 2018) are believed to influence the occurrence of natural and induced earthquakes. Within the WCSB, most of the seismic events tend to be concentrated within a band stretching between the eastern margin of the Cordilleran foreland thrust-and-fold belt and so-called undisturbed WCSB, which represent the margins of the regional seismicity. It was previously observed that the induced seismicity was most common in the vicinity of disturbed belt for two reasons: the high concentration of deep, critically stressed, pre-existing faults that are more

likely to be reactivated near the treatment wells (BC Oil and Gas Commission, 2014) and the role of the tectonic-strain rate, which is surmised to have a long-term impact on seismic potential (Kao et al., 2018). To address these hypotheses, distance from the disturbed belt was included into the analysis, together with the distance to minor lineaments indicating the potential locations of minor faults.

Stratigraphic Data

In previous studies, it was suggested that the risk of induced seismicity increased where hydraulic-fracturing operations in the Montney were performed at close distance to the crystalline basement (Skoumal et al., 2015). By also taking into consideration the Debolt Formation, which could represent a geomechanical basement, the impact of those two formations on the overall seismic-activity potential can be compared. Information about the tops of the Debolt Formation and Precambrian basement was implemented as the vertical distance between the hydraulic-fracturing operations (defined by the true vertical depth [TVD] of the well) and upper boundaries of both stratigraphic units. The Montney Formation tops were incorporated into the analysis indirectly, as one of the components of the depth factor described in detail below. Montney and Debolt formation tops were compiled from the geoSCOUT and BCOGC databases, whereas Precambrian basement-top information was sourced from the isopach and structure surface grid data collection available on the AER website. As in the case of the pressure gradient, the linear kernel variant of the RBF interpolation method was used to estimate the depths to the stratigraphic tops of the Montney and Debolt formations, and Precambrian basement.

Depth Factor

Due to the discrepancy between the levels of seismicity observed for hydraulic-fracturing operations performed in the upper, middle and lower units of the Montney Formation, establishing a possible relationship between the phenomena and the depth of injection could help further investigate any potential correlation between the induced seismicity and respective zones. In general, the lower Montney is characterized by higher induced seismic susceptibility; however, some significant seismicity is still observed in both the upper and middle units of the formation (BC Oil and Gas Commission, 2014). One possible explanation is the smaller vertical distance separating the lower unit from the underlying formations compared to the upper and middle Montney units rich in natural fractures (Nieto et al., 2018). In this study, a simple depth factor was implemented to provide a simplified way to divide the Montney into upper, middle and lower units, assuming the vertical division of the formation into three equal layers. Depth factor 'd' was calculated using the following equation:

$$d = (Z_w - Z_t) / Z_{th} \quad (1)$$

where Z_w is the TVD of the well, Z_t is the top of the Montney Formation at the location of the well, and Z_{th} is the thickness of the Montney at the location of the well.

Interpolation may carry the risk of incorrect estimation of the interpolated values. Parameter 'd' was implemented to detect the wells not matching the physical scenarios and guarantee the maximal correctness of the input data. Considering only wells drilled into the Montney interval, the depth factor should range between values 0 and 1. Values <0 indicate that the well was drilled above the top of the Montney Formation, whereas values >1 indicate that the well was drilled below the Montney. Therefore, 'd' values beyond the established range (0–1) suggest errors were made in measuring the formation top, thickness and/or well TVD. Such wells were excluded from further analysis, resulting in the final dataset consisting of 6315 wells. The depth factor was included into the models as a numerical value; however, further investigation revealed additional information about the analyzed wells.

Supposing the division of the Montney into three equal parts, it is possible to use the depth factor to differentiate between the Montney units associated with each well. According to this approach

- a 'd' value in the range of <0 to <0.34 corresponds here to the upper Montney;
- a 'd' value in the range of ≤0.34 to <0.67 corresponds here to the middle Montney; and
- a 'd' value in the range of ≤0.67 to <1 corresponds here to the lower Montney.

Table 1 presents the number of wells assigned to each specific unit of the Montney Formation using the described method.

It can be observed that the percentage of high-magnitude (M >2.5) events induced by hydraulic-fracturing operations is the highest for the wells drilled in the lower Montney, which is consistent with the hypothesis that this unit is more prone to seismic activity. At the same time, it has the fewest number of wells drilled, most of the wells having been drilled into the upper Montney unit.

Assigning Features to the Wells and Preprocessing

The value of each feature characterizing the corresponding wells was assigned using the NNjoin plugin (Tveite, 2019)

Table 1. Numbers and percentages of seismogenic and nonseismogenic wells corresponding to each of the Montney Formation units associated with M >2.5 induced seismic events.

Montney unit	Seismogenic wells		Nonseismogenic wells	
	Number	%	Number	%
upper	23	0.68	3354	99.32
middle	12	0.5	2408	99.5
lower	5	0.97	513	99.03

in QGIS, an open-source geographic information system (QGIS Development Team, 2019). The plugin allows the user to join two points from corresponding layers using the nearest neighbour method (i.e., linking points with the shortest distance). Collected features required preprocessing, which included missing-data identification and handling. For the wells that did not have data determined directly (such as formation tops, pressure and S_{Hmax} values), it was necessary to assign and use the interpolated values. In the next step, quality-control measures helped to determine only those wells that could assure the most reliable results. Due to the high number of wells with interpolated values, some of the interpolated values did not match realistic values. Data were analyzed in terms of the sequence of stratigraphic tops (Montney top – Debolt top – Precambrian top) and well TVD, with respect to the depth of the Montney top.

Algorithm Development

Determination of seismogenic potential related to the geological and geomechanical conditions in the subsurface was defined as a binary-classification problem, in which the algorithm assigns the labels according to the predicted classes (seismogenic/nonseismogenic). Two supervised learning techniques were analyzed with respect to their ability to determine the feature importance when predicting the class: decision-tree and random-forest classifiers (these are examples of tree-based methods).

In general, a decision-tree algorithm constructs the classification (or regression) tree composed of the leaves (estimated attributes) and nodes that split the data according to given conditions (Usuelli, 2014). The algorithm chooses the most important features and separates the samples into two groups until it satisfies the required conditions. Decision trees are expected to give better results than classical methods, such as linear regression, when nonlinear and complex relationships between the data are analyzed (James et al., 2013). Moreover, decision trees can be presented graphically, which makes them easier and more intuitive to interpret.

The random-forest classifier (similar to bagging and boosting machine-learning methods) is an improved and more powerful version of decision-tree algorithms (James et al., 2013) and uses sets of multiple trees to build the predictive models instead of only one. When building a random forest, small, randomly chosen collections of features are considered for splitting the nodes instead of a whole set of predictors. As a result, the feature used for splitting each node is chosen from among a limited number of available characteristics. This approach guarantees a decrease of variance between the single trees, which is desirable when building the machine-learning models.

One of the attributes of the analyzed dataset was the disproportionate size of the two classes. In other words, only a small percentage of the wells were associated with induced seismicity. To overcome this problem (which might lead to biased analysis), stratified random sampling was implemented. Stratified random sampling requires the division of the whole dataset into smaller subgroups (stratas) so that each subgroup corresponds to its class. In the next step, a random sample was drawn independently from each subgroup (Cochran, 2007).

Both models were built using Scikit-learn, a Python library for machine learning (Pedregosa et al., 2011), and trained on the number of 100 randomly shuffled stratified splits, with the training set accounting for 75% of the whole dataset and 25% for the test set. Figure 1 presents the feature importance indicated by both classification algorithms.

Discussion

Based on the decision-tree algorithm, the most important features were determined to be the distance to lineaments and the disturbed belt, the depth factor, the vertical distance to the Debolt Formation, and the vertical distance to the Precambrian basement. The results confirm the currently existing hypotheses about the influence of the tectonic setting of the wells, both in terms of distance to the faulted disturbed belt as well as known lineaments. The high importance of the depth factor potentially indicates the higher seismic risk in the lower Montney; however, this feature should be analyzed further. In contrast, the variation of the S_{Hmax} with respect to the regional stress state does not seem to influence the seismogenic potential. Surprisingly, the pressure gradient was indicated as the least important of all analyzed features, which contradicts the idea proposed by Eaton and Schultz (2018). However, this result could be biased as there were insufficient measurements of pore pressure. A different distribution of the important features might be observed for a dataset containing more exact pressure values. Comparing these results to those obtained using the second algorithm, the random-forest classifier identifies the same set of features as those with the most influence on the seismogenic potential, with similar feature-importance values. At the same time, the random-forest classifier showed a significant decrease in the error rate, which suggests a higher accuracy of the results.

It is worth noting that the machine-learning models are dependent on the quality (and quantity) of the input data. Therefore, the next step of this project is to perform extensive research and include additional parameters that may have an impact on the distribution of induced seismicity. The analysis in this study was limited to publicly available datasets only and some assumptions (e.g., the ratio between Montney intervals) were introduced. Additionally, several of the input parameters (formation tops and pressure, S_{Hmax} azimuths) were interpolated, which might mean they could

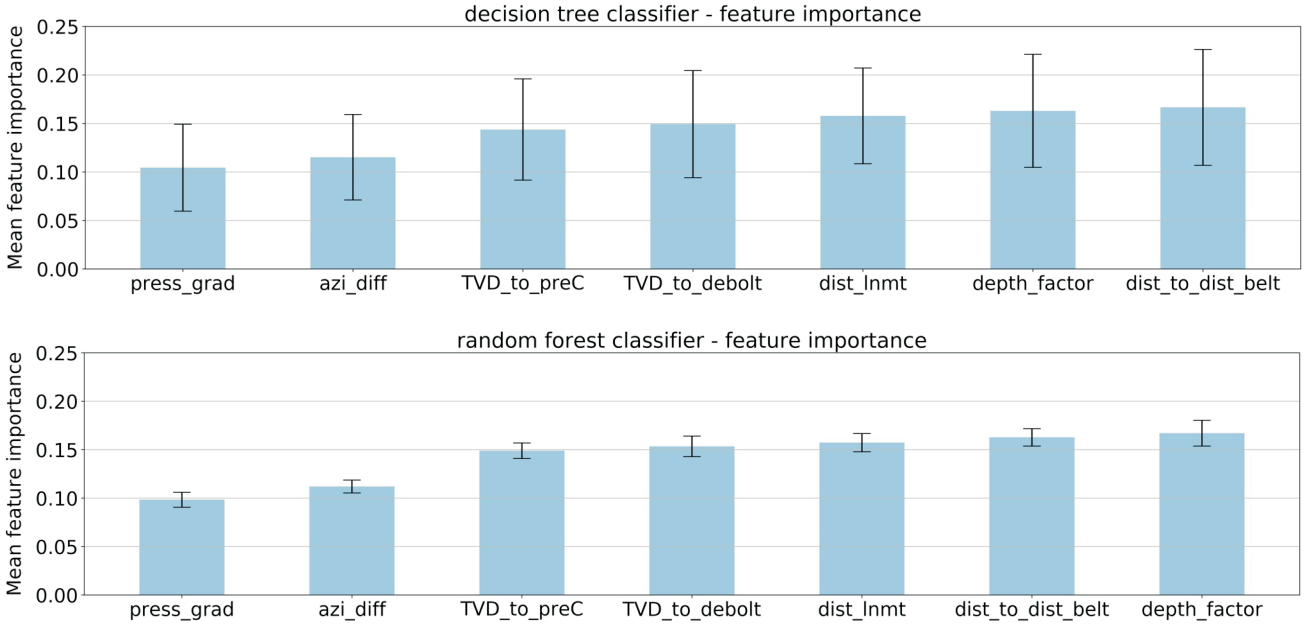

Figure 1. Feature importance calculated using decision-tree (**top**) and random-forest (**bottom**) classification algorithms. Results are presented in ascending order toward the right: pressure gradient (press_grad); maximum horizontal-stress azimuth variance (azi_diff); vertical distance from the well to the Precambrian basement and Debolt Formation (TVD_to_preC and TVD_to_debolt, respectively); distance to lineaments (dstm_lnmt); distance to the disturbed thrust-and-fold belt (dist_to_dist); and depth factor. Abbreviation: TVD, true vertical depth.

vary from the real values. Another problem is the incompleteness of data that cannot be interpolated, such as the location of unknown faults (especially strike-slip faults, which are undetectable using seismic methods). Incorporating more detailed information as well as new attributes into the algorithm will provide more accurate estimates of the geological susceptibility to induced seismicity.

Conclusions

This analysis revealed that three types of features are controlling the geological susceptibility to induced seismicity in the Montney Formation:

- the tectonic setting (specifically the distance from the disturbed belt and lineaments)
- the stratigraphy (vertical distance from the Debolt and Precambrian basement)
- the depth of the injection relative to the Montney Formation top, which correlates with the specific Montney unit (upper, middle and lower) stimulated during hydraulic fracturing.

Observations confirmed the current hypotheses about the factors controlling the induced seismicity. Pore-pressure gradient as well as variance of the S_{Hmax} from the average value were interpreted as less significant in the overall prediction of the induced seismicity for a given dataset. Moreover, a newly introduced parameter, the depth factor, confirms the current hypothesis about the higher susceptibility to induced seismicity in the lower Montney, compared to the upper and middle units. Overall, tree-based methods performed well and helped to address currently established conclusions about the factors controlling the occurrence of induced seismicity during hydraulic-fracturing operations. Including more detailed characteristics and additional features will increase the confidence level of the results and can provide information about the mechanisms responsible for the occurrence of induced seismicity.

Acknowledgments

The authors thank Geoscience BC and the Microseismic Industry Consortium for their financial support. They thank M. Hayes for his help in compiling the pressure data from the BC Oil and Gas Commission database. The authors also thank R. Weir for his comments that helped to improve this manuscript.

References

Atkinson, G.M., Eaton, D.W., Ghofrani, H., Walker, D., Cheadle, B., Schultz, R., Shcherbakov, R., Tiampo, K., Gu, J., Harrington, R.M. and Liu, Y. (2016): Hydraulic fracturing and seismicity in the western Canada sedimentary basin; Seismological Research Letters, v. 87, no. 3, p. 631–647, URL <https://pubs.geoscienceworld.org/ssa/srl/article/87/3/631/315665/Hydraulic-Fracturing-and-Seismicity-in-the-Western> [March 2016].

Bao, X. and Eaton, D. W. (2016): Fault activation by hydraulic fracturing in western Canada; Science, v. 354, no. 6318, p. 1406–1409, URL <https://science.sciencemag.org/content/354/6318/1406.full> [December 2016].

Baranova, V., Mustaqeem, A. and Bell, S. (1999): A model for induced seismicity caused by hydrocarbon production in the Western Canada Sedimentary Basin; Canadian Journal of Earth Sciences, v. 36, no. 1, p. 47–64, URL <https://www.nrcresearchpress.com/doi/10.1139/e98-080#.XaUv9fd7nV8> [January 1999].

BC Oil and Gas Commission (2014): Investigation of observed seismicity in the Montney Trend; BC Oil and Gas Commission, Technical Report, 32 p., URL <https://www.bcogc.ca/investigation-observed-seismicity-montney-trend> [December 2014].

BC Oil and Gas Commission (2019): BC Oil and Gas open data portal; data download web application, URL <https://data-bcogc.opendata.arcgis.com/> [July 2019].

Canadian Induced Seismicity Collaboration (2019), Catalogues website, URL <https://www.inducedseismicity.ca/catalogues/> [June 2019].

Cochran, W.G. (2007): Sampling Techniques (3rd edition); John Wiley and Sons, Inc., New York, New York, 442 p.

Eaton, D.W. (2018): Passive Seismic Monitoring of Induced Seismicity: Fundamental Principles and Application to Energy Technologies; Cambridge University Press, Cambridge, United Kingdom, 360 p.

Eaton, D. W. and Schultz, R. (2018): Increased likelihood of induced seismicity in highly overpressured shale formations; Geophysical Journal International, v. 214, no. 1, p. 751–757, URL <https://academic.oup.com/gji/article/214/1/751/4995198> [May 2018].

Ellsworth, W. L. (2013): Injection-induced earthquakes; Science, v. 341, no. 6142, 7 p., URL <https://science.sciencemag.org/content/341/6142/1225942.full> [July 2013].

Fereidoni, A. and Cui, L. (2015): Composite Alberta Seismicity Catalog: CASC2014-x: URL <https://www.inducedseismicity.ca/wp-content/uploads/2015/01/Composite-Alberta-Seismicity-Catalog3.pdf> [June 2019].

geoLOGIC systems ltd. (2019): geoSCOUT version 8.12; geoLOGIC systems ltd., mapping, data management and analysis software, URL <https://www.geologic.com/products/geoscout/> [August 2019].

Ghofrani, H., Atkinson, G.M., Schultz, R. and Assatourians, K. (2019): Short-term hindcasts of seismic hazard in the western Canada sedimentary basin caused by induced and natural earthquakes; Seismological Research Letters, v. 90, no. 3, p. 1420–1435, URL <https://pubs.geoscienceworld.org/ssa/srl/article/90/3/1420/530744/Short-Term-Hindcasts-of-Seismic-Hazard-in-the> [April 2019].

Guglielmi, Y., Cappa, F., Avouac, J.P., Henry, P. and Elsworth, D. (2015): Seismicity triggered by fluid injection–induced aseismic slip; Science, v. 348, no. 6240, p. 1224–1226, URL <https://science.sciencemag.org/content/348/6240/1224> [June 2015].

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.-L. and Zoback, M.D. (2018): The World Stress Map database release 2016: Crustal stress pattern across scales; Tectonophysics, v. 744, p. 484–498, URL <https://doi.org/10.1016/j.tecto.2018.07.007> [June 2019].

James, G., Witten, D., Hastie, T. and Tibshirani, R. (2013): An Introduction to Statistical Learning – with Applications in R; Springer-Verlag, New York, New York, 426 p.

Jones, E., Oliphant, T. and Peterson, P. (2016): SciPy: Open source scientific tools for Python; SciPy Developers, URL <http://www.scipy.org/> [August 2019].

Kao, H., Hyndman, R., Jiang, Y., Visser, R., Smith, B., Babaie Mahani, A., Leonard, L., Ghofrani, H. and He, J. (2018): Induced seismicity in western Canada linked to tectonic strain rate: implications for regional seismic hazard; Geophysical Research Letters, v. 45, no. 20, p. 11–104, URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2018GL079288> [October 2018].

Mossop, G.D. and Shetsen, I., compilers (1994): Precambrian structure from back model grid, elevation (m); in Geological Atlas of the Western Canada Sedimentary Basin, G.D. Mossop and I. Shetsen (comp.), Canadian Society of Petroleum Geologists and Alberta Research Council, isopach and structure surface grid data, URL <https://ags.aer.ca/publications/isopach-and-structure-surface-grid-data.htm> [July 2019].

Nieto, J., Bialowas, B., Batlai, B. and Janega, G. (2018): Managing induced seismicity in Canbriam's Altares field in the Montney Formation, N.E. British Columbia – an update; Canadian Society of Exploration Geophysicists, Recorder, v. 43, no. 7, URL <https://csegrecorder.com/articles/view/managing-induced-seismicity-in-canbriams-altares-field-in-the-montney-fm> [December 2018].

Pawley, S., Schultz, R., Playter, T., Corlett, H., Shipman, T., Lyster S. and Hauck, T. (2018): The geological susceptibility of induced earthquakes in the Duvernay play; Geophysical Research Letters, v. 45, no. 4, p. 1786–1793, URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017GL076100> [February 2018].

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V. and Vanderplas, J. (2011): Scikit-learn: machine learning in Python; Journal of Machine Learning Research, v. 12, p. 2825–2830, URL <http://www.jmlr.org/papers/v12/pedregosa11a> [August 2019].

QGIS Development Team (2019): QGIS Geographic Information System; Open Source Geospatial Foundation Project, URL <http://qgis.osgeo.org> [August 2019].

Schultz, R., Stern, V. and Gu, Y. J. (2014): An investigation of seismicity clustered near the Cordel Field, west central Alberta, and its relation to a nearby disposal well; Journal of Geophysical Research – Solid Earth, v. 119, no. 4, p. 3410–3423, URL <https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2013JB010836> [April 2014].

Skoumal, R.J., Brudzinski, M.R. and Currie, B.S. (2015): Earthquakes induced by hydraulic fracturing in Poland Township, Ohio; Seismological Society of America, Bulletin, v. 105, no. 1, p. 189–197, URL <https://pubs.geoscienceworld.org/ssa/bssa/article/105/1/189/323441/Earthquakes-Induced-by-Hydraulic-Fracturing-in> [January 2015].

Skoumal, R.J., Brudzinski, M.R. and Currie, B.S. (2018): Proximity of Precambrian basement affects the likelihood of induced seismicity in the Appalachian, Illinois, and Williston Basins, central and eastern United States; Geosphere, v. 14, no. 3, p. 1365–1379, URL <https://pubs.geoscienceworld.org/gsa/geosphere/article/14/3/1365/530435/Proximity-of-Precambrian-basement-affects-the> [April 2018].

Snee, J.E.L. and Zoback, M.D. (2016): State of stress in Texas: implications for induced seismicity; Geophysical Research Letters, v. 43, no. 19, p. 10–208, URL <https://

agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016GL070974> [December 2016].Tveite, H. (2019): NNJoin QGIS Plugin, version 1.3.3; NNJoin 3.3.3 documentation, URL <http://arken.nmbu.no/~havatv/gis/qgisplugins/NNJoin> [August 2019]

Usuelli, M. (2014): R Machine Learning Essentials; Packt Publishing Limited, Birmingham, United Kingdom, 218 p.

van Rossum, G. and Drake, F.L. (2001): Python reference manual; PythonLabs, release 2.0.1, URL <https://docs.python.org/2.0/ref/ref.html> [November 2019].

Zhang, Y., Person, M., Rupp, J., Ellett, K., Celia, M.A., Gable, C.W., Bowen, B., Evans, J., Bandilla, K., Mozley, P. and Dewers, T. (2013): Hydrogeologic controls on induced seismicity in crystalline basement rocks due to fluid injection into basal reservoirs; Groundwater, v. 51, no. 4, p. 525–538, URL <https://ngwa.onlinelibrary.wiley.com/doi/full/10.1111/gwat.12071> [June 2013].

Zoback, M.L. and Zoback, M.D. (1989): Tectonic stress field of the continental United States; *in* Geophysical Framework of the Continental United States, L.C. Pakiser and W.D. Mooney (ed.), Geological Society of America, Memoir 172, p. 523–539, URL <https://izaks.people.stanford.edu/sites/g/files/sbiybj2961/f/Zoback_Zoback_1989_Tectonic_stress_field_of_the_continental_United_States.pdf> [January 1989].