Dear Editor and Reviewers,

We would like to thank the reviewers for their helpful and constructive comments and have included a point-by-point response below. Reviewers' original comments are included in italic text and our responses to the reviewers are included in plain text. We have highlighted text changes that were incorporated into the manuscript in blue and included the line numbers for ease of review. The additional tracked-changes pdf version shows text that has been removed as being crossed-out, red-colored text and newly inserted text is underlined in blue.

Reviewer #1:

The authors present an analysis of seismicity clusters in the Kiskatinaw region of British Columbia. They determine vp/vs estimates for the clusters over time and observe changes that appear to be larger than the background error.

The paper is very well written and discusses a topic that is of interest within the induced seismicity community. The methodology is rigorous and is applied in a way that is logical, clear and concise. With minor revisions, this paper would be an excellent fit in this journal.

Some comments/questions:

1. Why is the term "in-situ" vp/vs used? My understanding of "in situ" refers to "in place", implying close to the source (either earthquakes or injection sites). However, given the large distances between the event clusters and stations, what is being calculated is effectively an average change in vp/vs. "In situ" may imply more resolution than there actually is.

Answer: Thanks for pointing this out. The term *"in situ"* is an established phrase by the authors who originally developed the method (see e.g., Lin 2020). To avoid confusion with readers who are not familiar to the terminology, we added the following sentence on line 62 - 64:

The term *in-situ* in this context describes the localized damaged rock volume in which closely related earthquake pairs occur that are used to resolve V_p/V_s based on P- and S-arrival-time-differences within the pairs.

2. Limitations of the Nanometrics velocity model. The authors are likely aware of the limited resolution and accuracy of a regional velocity model constructed with only 100 well logs and seismic surfaces. In my experience, thousands of well logs and significant seismic data coverage are needed to create an accurate velocity model. I'm familiar with this study area and am aware that likely what was used by Nanometrics is the only data available and as such is the best they can do at this time. However, this point should be made clear to the audience, especially given the lateral heterogeneity and complex faulting structures in this region. I would add a few sentences explaining data limitations and geologic complexity.

Answer: Thank you for pointing out the need for expanding the background on the starting model. We have modified the text to provide caveats on the limitations of the

starting model in the last sentences in Section 2 (please see the modification noted in the next response). We have also added a paragraph discussing the limitations of the Nanometrics velocity model resolution to the Discussion section as well. As detailed in the response to Comment #5, we state that two factors, namely, limited data access and existing geological heterogeneities or fault complexity, can impact the comparability of our *in-situ* estimates to the reference grid points.

3. Along the same lines as #2, I also have concerns about the estimate of 2.12% for the velocity model error. This is extremely small. In some parts of the model, it may be 2%, in others it may be 10%, there is simply no way to know given the limited velocity information in this region. I understand that for the purposes of this paper, the authors need to have some kind of baseline, which is why this is so challenging. For clarity, after line 98/99, I would add a sentence or two explaining that this is the model error and not the true velocity error, which may be significantly larger. In other words, I'm not asking for any changes with the methodology, just for more clarifying text.

Answer: Thank you for the useful comment that will help the reader avoid confusion. We added the following sentence explaining the above caveat at the end of Section 2 on line 104 - 106 to improve clarity:

We note that the assumed uncertainty of 2.12% solely reflects the model error. The ~140 sonic logs used to build the publicly available Nanometrics regional model do not enable resolving the velocity structure in high resolution or the geological structural complexity in the region.

4. I appreciate the authors rigorous approach to determining uncertainty at every step. I also appreciate the effort to "sanity-check" the results and make sure they are physically reasonable. Given the data limitations of this regional-scale analysis, I think the authors have made very wise choices about how to calculate and interpret vp/vs changes.

Answer: Thank you.

5. Perhaps I missed it, but in the Discussion do the authors also mention that some of the larger velocity changes (>2.12%) may also be reflective of regions where the velocity error may need adjusting?

Answer: The reviewers comment points out that in the initial version of this manuscript, we did not discuss that larger velocity changes might result from the uncertainties in the reference velocity model. Both the uncertainties that result from limited data access (e.g., from Nanometrics, as noted by the reviewer in Comment #3), and geological structural heterogeneities or fault complexity can generally limit the ability to compare our *in-situ* estimates to the reference grid points. We have now added a short paragraph in the Discussion section to emphasize the relative importance of considering the above factors in this kind of study. See lines 336 - 340 for text changes.

One limiting factor of our work is in the reference velocity model. While Nanometrics Inc. (2020) utilized all available data at the time to develop the velocity model, it is likely

a small fraction of a more comprehensive dataset required to resolve the geological complexity of the study area. Due to the existing resolution limit of the reference velocity model, we cannot factor out that larger changes in V_p/V_s (and hence velocity changes) are due to reference model uncertainties rather than only due to realistic changes in the earthquake cluster areas.

Reviewer #2:

The manuscript presents an application of the in-situ V_P/V_s method to earthquakes induced by hydraulic fracturing in western Canada. The authors first describe the geological setting and dataset. They then introduce the in-situ V_P/V_s method and how they adapted the method to their problem. After that, they present their results in two parts: First, they describe the overall spatiotemporal variation of the V_P/V_s of all event clusters during the entire observation period. Second, they present the detailed temporal V_P/V_s evolution of one cluster during an around two-week period near the end of the observation period. These results are further compared to the V_P/V_s computed using a two-phase poroelastic model. The authors finally discussed the significance of their results in elucidating the triggering mechanisms of the microearthquakes.

This manuscript describes one of the first applications of the in-situ V_P/V_s method to earthquakes induced by hydraulic fracturing. Given the potential of this method to image fluids in the system with unprecedented high spatial and spatial resolutions, the results of this study could be highly valuable. However, the manuscript has two major issues that render it unsuitable for publishing in its current form. The most important one is the lack of an effective discussion section. The current discussion section essentially consists of literature review and repeated mentioning of the results in previous sections. I also have concerns over the authors' quality-control procedure, which was demonstrated by a recent study to have a major impact on the final results and yet is not clearly documented in the manuscript. In addition, the manuscript also contains many typos and statements that are either ambiguous or erroneous. Based on the above impressions, I recommend a major revision to the manuscript, and I will be happy to review the revised version again. Please see my detailed review below.

Major issues:

1.) The discussion section of the manuscript is highly ineffective. The first paragraph of Section 6.1 consists largely of literature review. The second paragraph first proposes a tensile-fracture model to explain the V_P/V_s increase but then rejects it in the end. The last paragraph mentions a few seemingly irrelevant previous studies without making a clear point. Section 6.2 merely repeats the main findings presented in Section 5. The first paragraph of Section 6.3 again repeats the main findings of Section 4.1. The second paragraph is very incoherent but seems to suggest that the observed spatiotemporal variation in V_P/V_s does not provide useful information on earthquake triggering. In summary, the discussion section in its current form contains very little useful information and thus probably needs to be completely rewritten.

Answer: We have now reorganized the Discussion and edited the text of clarity. Specifically, we have trimmed Section 6.1 and merged it with Section 6.2 to be more concise and provide a clear connection the results we present here, as well as removed less directly related references. However, we respectfully disagree with the reviewer that discussion of potential mechanisms and implications of the observed Vp/Vs change are irrelevant just because in the end they do not nicely explain or fit our observation. Such "straw-man arguments" are commonly used to show consistency of results with a given interpretation. As such, our aim is to provide the readers with a list of commonly referred mechanisms and discuss how they may or may not apply to our observations. The revised text is not shown below here for brevity, but can be found on Lines 288 - 340.

2.) The quality control procedure is not well presented (last paragraph of Section). It is unclear what a "cross-correlation-based picking correction" is. The authors do not explain why they use 5.1 and 2.9 km/s as the reference P and S velocities to rule out outliers and if this step could bias the V_P/V_s estimates towards the reference value (5.1/2.9=1.76). It is also unclear how the hybrid L₁-L₂ fitting method removes outliers because in my recollection the method simply uses the L₁ norm for data points whose residues exceed a certain threshold. I suggest the authors use a flow chart to describe their quality-control procedure similar to Figure 3 of Liu et al. (2023). I emphasize the importance of quality control here because Liu et al. (2023) showed that it has a major impact on the final V_P/V_s estimates and thus deserves greater attention.

Answer: We agree that quality control criteria are crucial. We have changed the text to more extensively acknowledge the work in Liu et al., 2023, as well as added a paragraph providing a structured overview of the quality control criteria we used in this work. The new paragraph is not copied below (due to the length) but is found on lines 133 - 160 of the manuscript.

With the inclusion of the text on L. 133-160 and due to the similarity of the procedure to previously published work that has been cited in the introduction, we have chosen not to add an additional explicit workflow chart as in Liu et al. (2023).

To answer the reviewer's question regarding velocities, the values of 5.1 and 2.9 km/s follow the slope in Figure S2 and are only used to predict arrival times. We have modified the text in line 141 to improve clarity as follows: (comparable to the slope of the travel time curves in Figure S2).

We also note that in the procedure, we remove picks that deviate by a certain time as outliers following the cross-correlation calculation. We are not forcing a trend of $V_p/V_s =$ 1.76 (in fact, our estimates are lower), but merely removing incorrect phase picks (which appear as parallel lines relative to the main trend that crosses the origin in Figure 2).

Finally, we apply the same L1-L2 fitting norm as in Lin and Shearer, 2007 (the original reference). To the best of our knowledge, the regressor applies a L2 norm for values below and a L1 norm for values larger than a certain threshold.

Minor issues

Line 29: "earthquake-earthquake" is better to be changed to "inter-event"

Earthquake-earthquake interaction is a common term that in our case better describes the coseismic stress changes.

Line 32: "both" should be removed if what the authors mean is that pore-pressure increases is responsible for the earthquakes close to the well, and poroelastic stress changes are responsible for those far from the well.

We have removed "both".

Line 38: What are "shale gas plays"?

"shale gas play" is a commonly accepted term that describes shale gas accumulation in geological basins. See references in <u>Canada Energy Regulator</u> and in the literature (e.g., Schultz et al., 2015; Kim 2022; Cardott 2012).

Line 53: " V_P/V_s coupled with Poisson's ratio" is inprecise because the two are mutually dependent, i.e., there is a one-to-one relation between them.

We agree, thank you for pointing out the inconsistency. We have changed to " V_p/V_s to infer changes to Poisson's ratio"

Lines 65–66: "differential travel-time differences" is imprecise because the method uses just the differential P and S travel times without differentiating one more time.

Thanks, agreed. We have modified the text to: "...that compares differential travel times of co-located..." (now line 70)

Lines 67–68: Do "significant variations" mean variations in space, time, or both?

It refers to both. We have changed the text to clarify as follows: "significant spatiotemporal variations." (now line 72)

Lines 91–99: Was the background velocity model built using the data collected before all the injections happened? Otherwise, it will not be an appropriate reference for reflecting the V_P/V_s changes that happened after the injection began.

The background model provides a spatial average using seismic and well-log data collected up to the time the report was created. Therefore, it does not reflect a specific point in time, but rather, an average, background model to which we can compare the Vp/Vs changes. (Please see modifications implemented in response to R1 Comment #5 for additional details as well).

Line 105: What does "compensates for ray path differences" means? Does it mean the authors find a way to correct the difference in P and S ray paths?

We do not correct ray paths. We have rephrased the sentence to clarify as follows: "We apply the method of Lin and Shearer, (2007) that makes use of differential travel times per station between co-located event pairs with coincident ray paths, and by removing the need to consider their origin times." (Now on L. 111 - 113).

Line 109: "difference in origin times" should instead be "origin time errors" because if the origin times are known perfectly, this term will be zero (see Lin and Shearer (2007)).

In the derivation of the equation, Lin and Shearer describe δt_0 to be the difference in origin times between two events, $\delta t_0 = t_{02} - t_{01}$, where t_{02} is the origin time of event 2 and t_{01} is the origin time of event 1. Later in the methodology, the authors write: "The effect of the difference in origin times, δt_0 , is to shift the $(\delta T^i{}_p, \delta T^i{}_s)$ points in both coordinates by δt_0 or along a 45° line." We have elected to use the same terminology in our description in line 116 to provide consistency with the original reference.

Line 111: "eliminating all origin information" is imprecise. The demean process just removes the origin time errors without addressing the location errors.

The comment is correct, thanks. We have changed the text to the following: "requires eliminating the absolute reference to temporal origin time information" (line 118)

Line 115: I don't think the dimension of the cluster can be arbitrarily large because the difference between the P and S rays will also grow with cluster size.

Large in the context of the sentence refers to the number of events. Thank you for pointing out the ambiguity. We have modified the sentence to the following: "The V_p/V_s ratio as fitted in Equation 2 can be treated as a constant for each earthquake cluster, as long as the station-event distances are large compared to the hypocentral offsets among events in each cluster". (line 112 – 123)

Lines 119–126: The synthetic test on take-off-angle may not be relevant because the cause of the take-off-angle difference between P and S rays is the spatial variation in V_P/V_s , and the V_P/V_s variation in IASP91 may be very different from the study area.

The take-off angle test is designed to demonstrate that take-off angle is a crucial factor for our short station-event distances and shallow depths. For example, Palo et al., (2016) point to the applicability of the method and find that the takeoff-angle difference is crucial. The objective of our synthetic test is to validate the constant takeoff angle assumption by checking its variability for the station-event geometry in our study area. It demonstrates that the takeoff angle variation is negligible, and thus the validity of the assumption (stated in line 124 - 132).

Lines 165–168: Are "moderate" increase and decrease equivalent to "unsignificant" increase and decrease in Figure 3? Besides, "unsignificant" should be "insignificant".

Primarily, we want to point out that Vp/Vs changes below $\pm 1\%$ are not too small to be relevant to the interpretation, regardless of the statistical/modeled significance. Hence, we describe larger changes to be moderate. We have now reorganized the paragraph to make the above point more clear. "Figure 3 also shows 9 clusters with a relative Vp/Vs -ratio change ranging between -1% and 1%, which we interpret overall as minor changes, despite their relative lower significance. We observe a moderate increase in Vp/Vs following fluid injection for 19 out of 34 clusters, and a moderate relative decrease for the remaining 6 clusters." (line 180 – 182)

Line 172: The COVID-19 operation shut-down should be marked in Fig. 3c.

Thanks, we have added the period of seismic quiescence according to Salvage and Eaton, 2021, to the figure. We updated the caption. "The hatched, pink area shows the area of seismic quiescence due to suspension of HF operations (Salvage and Eaton, 2021) between April and August 2020."

The suspension of HF operation differs from the "lockdown phases" between 21 March (school closure on 17 March) and 5 May 2020.

Lines 178–189: Are the southeast and northwest clusters among the 34 clusters discussed in the previous section? If so, they should be marked in Fig. 1.

The length scale of individual earthquake clusters is too small for the overview in Figure 1. Alternatively, Figure 4 will provide significantly more context. However, we have now highlighted the cluster discussed in Figure 4 in Figure 3. "The example cluster highlighted in yellow is discussed in further detail in Figure 4." and "One example cluster from Figure 3a (highlighted in yellow)."

Line 187: Should "southwest" be "southeast" instead?

Yes, thank you, it is now corrected.

Lines 198–199: How the temporal windows are defined is unclear. Are there overlaps between consecutive windows?.

There are no overlaps between time windows. We divided the catalog in equally sized time segments. We have modified the sentence on I. 215-216 to make clear that the consecutive windows are non-overlapping.

"For example, Figure 5a-d shows the chronological division of 300 events in the northwestern cluster in Figure 4 (maroon box) into four equally sized groups of 67 to 68 events in non-overlapping windows"

Lines 203–205: The authors should clarify what the range in Gregory (1976) is because many readers including me may not be familiar with the study.

Thanks. We added the ranges of conditions from Gregory (1976) in line 220 - 224. The new text reads as follows: "The seemingly small absolute changes in V_p/V_s in the range of 0.06 are already significant with respect to reported values between 1.98 and 1.42 (Gregory, 1976), which were estimated for different types of consolidated sedimentary rocks with porosities ranging from 4.45% to 41.1%, water-air-saturation ratios ranging from 0% to 100%, and confining pressures ranging from 0 MPa to ~69 MPa"

Lines 207–208: The sentence in the parenthesis seems to be out of place. It is also unclear what "weakly linked event pairs" means.

We agree. We have removed the phrase "weakly linked event pairs" and rephrased the sentence at line 216 - 217 to improve clarity. It now reads: "We note that applying quality control criteria remove event pairs and hence reduces the number of actually grouped events from the original 300 to 269.".

Lines 218–220: Isn't "effective elastic moduli of fluid-filled rock" equivalent to "elastic moduli of the effective porous medium"?

We phrased this sentence to remove ambiguity. We have changed it to: "[...] such as fluid fraction, elastic modulus of each medium component, and/or fracture geometry, [...]" (line 234 – 236)

Lines 210–212: The authors should show the temporal evolution of the other three clusters at least in the supplementary materials.

We have added the temporal evolution of the other clusters in a new figure Fig. S9 (shown here below).



Line 223: "rock matrix" should be replaced by "fluid-filled rock" because "rock matrix" means the solid rock skeleton without the fluid inclusions.

We have changed the text to more clearly emphasize that the decisive parameter that we explore here is the change of the rock matrix, which subsequently changes the fluid content. The sentence has be modified to: "dependence on the rock matrix and resultant fluid content, we use" (line 239)

Lines 254–255: "It also shows that decreasing aspect ratios would be consistent with increases in fluid fraction that would lead to a decrease in Vp/Vs." is unclear to me. Please rewrite.

The revised passage has been modified to improve clarity in the explanation of the consistency argument that is being made with the four scenarios. Specifically, it is made to show that the evolution of alpha and porosity shown in Scenario #4 is consistent with HF-stimulation, and can replicate a decrease followed by an increase in Vp/Vs as is observed in the data. The new text is not copied here for brevity, but can be found on (Lines 266 - 281).

Line 258: Should "increase" be "decrease" instead? Otherwise, this sentence makes no sense to me.

The entire paragraph has now been rewritten for clarity, and the text in question no longer appears in the same form (Please see Comment #23 response above).

Line 294: Should "faults" be "," instead? Otherwise, the sentence doesn't read correctly.

The sentence was removed in the process of editing the Discussion section.

Line 300: "to be" should be deleted.

Done.

Line 332: A "to" is missing after "similar".

Corrected.

Lines 334–336: This sentence is puzzling. First, "the spatial correlation of increased pore pressure and significant V_P/V_s decreases" was never presented before. Second, "lack of significant Vp/Vs increase" contradicts the main observation shown in Fig. 3. Besides, "reduce" should be "reduced" in Line 335.

We have revised the sentence to be more concise: "On the other hand, we also observe significant V_p/V_s decreases in areas with large amounts of injected fluid (southeast end of the profile in Figure 3), suggesting that additional factors to pore

pressure increase may have an important role in activating faults here." (line 353 - 355).

Figure 1: The stations are difficult to see. The injection wells should also be shown.

We have modified the figure by adding a light background color as filling to make stations more visible. However, adding the injection wells to Figure 1 makes the figure too crowded. For this reason, we have modified Figure S1 to show the wells, as well as noted that they are depicted in S1 in the Fig. 1 caption for easy reference.

Figure 6: The color bar of (a) is saturated in the lower part of the V_P/V_s range, causing different values to be indistinguishable. This is especially problematic given that most of the observations are in this range.

The values in our model range from $V_p/V_s = 1.62$ to 6.57. For a comprehensive illustration, we decided to use colorbar limits of 1.65 (saturated) and 2.1 (not saturated for larger values). We apply those limits to focus on the relevant range of V_p/V_s values in our estimates. In general, the range of porosities can lead to highly variable ratio values. We address the above point in line 254 - 256: "The range of porosity-aspect ratio-pairs can lead to highly varying V_p/V_s estimates. For illustration purposes, Figure 6a only displays values between 1.65 and 2.1 that cover the initial V_p/V_s values observed by Gregory, (1976)."

Figure S2: What is the difference between (a) and (b)?

(a) and (b) are two events of one event pair. We have added the events' origin information in the caption to make it clear to the reader that it refers to an event pair. "a) and b) show travel time curves for P-waves (squares and blue lines) and S-waves (circles and orange lines) for an event pair. a) is a ML 2.8 on 2020-02-13T11:55:14.262, b) a co-located ML 2.7 on 2020-02-10T11:48:40.394."

Figure S7: Given that different time histories have different time windows, how the mean solution (thick black curve) is computed is unclear to me.

We resampled the individual trends to a common time vector prior to estimating the mean value. We have changed the text as follows to make the above point more clear: "The thick black line represents the interpolated mean estimate of all segments after resampling each individual trend to 11 data points."

Figure S9: The curve computed using a = 1 seems to be outside the HS+ boundary. What's the explanation?

In the following we describe our understanding from Mavko et al., (2009), Chapter 4: Effective elastic media: bounds and mixing laws. (The rock physics handbook: Tools for seismic analysis of porous media). The Hashin-Shtrikman (HS) bounds can be used to compute the upper and lower bounds for a mixture of mineral and pore

fluid. They assume each constituent is isotropic, linear, and elastic and that the rock is isotropic, linear, and elastic. HS bounds describe the narrowest possible range without specifying anything about the geometries of the constituents. Contrary to the upper Voigt "isostrain average" and lower Reuss "isostress average", if previous assumptions are not meet, the estimated bulk- and/or shear moduli violate the bounds. On the one hand, the HS-bounds do not consider pore geometry in general. The physical interpretation of HS+ (HS-) is a model with the softer (stiffer) material inside a shell of the stiffer (softer) material, respectively. On the other hand, considering the effect of pore geometry, "stiff pore shapes" cause the moduli to be higher, while softer shapes cause the value to be lower. When the aspect ratio approaches unity, we would have with a sphere the stiffest pore geometry of our model. In addition, our model contains two materials with a significant differences in individual moduli with μ = 0 (i.e., rock vs. fluid), bringing the HS bounds to its limits. Hence, we conclude that for stiff pore shapes (large α), the model is not valid. We have added following sentence to the caption to clarify: "Stiff pore geometries (i.e., α approaching unity), will cause the model to exaggerate the HS boundaries, which do not consider pore geometry."

The authors should show at least one depth cross section of the study area so that the vertical distribution of the events is clear to the readers.

We have added a new Figure S6, which describes the catalog depth distribution in the histogram, as well as the event depths relative to the well casing depth (example from Figure 4) in the inset.

