Review Round 1

Reviewer 1

General Comments

Overall, this is a concise, well-written article that applies an established approach to a known case of induced seismicity in the Dallas–Fort Worth, Texas area. While the work is promising, several general and specific issues should be addressed before publication.

For readers unfamiliar with induced seismicity in Texas, the authors should replace qualitative distance descriptions (e.g., “not far,” “far”) with quantitative measures. This will improve clarity and reproducibility.

I am concerned that the conclusion—stating no significant link between saltwater disposal (SWD) and the Venus, Azle, or Cleburne earthquake sequences—may reflect limitations of the modeling approach rather than the actual geologic and operational reality. It is unclear whether the authors differentiate between deep SWD (below the producing layer) and shallow SWD (above the producing layer). If they do not, they should explain why. Numerous studies have shown that deep injection can influence seismicity over long distances. The Ellenburger Formation, referenced in the paper, qualifies as deep injection. Although it has low porosity and permeability, fluid migration over considerable distances along fractures has been documented.

Given this, the models should account for a larger potential influence radius for deep wells and perhaps a smaller one for shallow SWDs—or the analysis could be limited to deep wells. While modeling these processes is challenging, greater attention is needed to ensure that model assumptions align with observed realities in the region. As currently presented, the lack of distinction between SWD practices and geologic settings in the study area reduces confidence in the stated conclusions.

Specific Comments

Abstract: “..large D values.” D is not explained in abstract. Need to either remove and/or refer to other process

Abstract: “Notably, no significant link with SWD is observed for the Venus, Azle or Cleburne earthquakes.” I do not think you can conclude this. This likely means your model is wrong somewhere, and needs to be adjusted.

Line 115: “the causative fault is very far”- How far is very far? Put number.

Line 130: “Five SWD wells were within 5 km the Venus cluster..“ should be “Five SWD wells were within 5 km of the Venus cluster”

Lines 227-229: explain the d and w parameters. Choosing values makes no sense without the context of knowing what the parameters represent.

Line 283: SWD may impact not just neighbor blocks! Could be km away. In some areas of Texas, that distance could well over 10 km, the assumption you use. Can this be explored? I would say for deep wells, this number should be at least 25 km, based on many authors work.

Lines 287-290: With deep injection, the diffusivity will be different than shallow injection. Would exploring these differences and the end results.

Lines 328-333: This sort of addresses my previous comment, but separating the wells (deep/shallow) is critical for this to be a robust result.

Lines 335-337: You are stating that many SWD wells are within 15 km of seismicity, then why are you using 10 km for you parametrization? I think you are getting a result based on the input to your model, instead of inputting parameters that reflect would the consensus for deep injection.

Figure 2: Would like to see the differentiation of deep (below producing layer) and shallow wells.

Reviewer 2

This paper applies a statistical method to assess the probable causality of seismicity in the Dallas-Fort Worth area between 2008 and 2023. The method was previously applied successfully to induced seismicity in Oklahoma as well as the Delaware Basin and the Eagle Ford play in Texas, so this paper extending its use to the DFW area is a welcome addition. I found the presentation of the history of seismicity in the region and the method to be clear and interesting (with a few minor exceptions detailed in the line-by-line comments below). However, the presentation and analysis of the results was somewhat confusing. First, in the presentation of the HF results, many of the results were based on the HFR method, even though in the methods section this was described as a secondary method prone to coincidental findings. However, the overall conclusion of this section that most of the

seismicity was not due to HF except one unnamed cluster was well-supported. The SWD results showed a lower correlation than expected from previous work in this area, and I think this paper could do more explore why that might be the case. While some possible explanations were offered, they weren't really evaluated against each other in terms of likelihood, and some additional work could help distinguish them. On the possibility that the data completeness is insufficient for this type of analysis, it might be helpful to examine a subset of more recent data that would have better completeness. Examining a shorter time period might also help with concerns that the seismic response is not consistent across time. If this isn't possible, you should explain why. On the possibility that the diffusivity is high enough to permit complete spreading of SWD to all the blocks, some additional discussion of how plausible various D values are would be helpful. In short, this paper either suggests that a higher quality catalogue is necessary to apply this method, or there is something different about the FWB that causes it to react differently than the previously studied areas, or neither SWD or HF is likely responsible for most of the seismicity in the FWB. In any case it would be an interesting result, and the paper would be significantly strengthened by a more thorough investigation into which it is (even if it's still not possible to come to a definitive conclusion).

Line-by-line suggestions follow:

16 "either SWD and hydraulic fracturing (HF) are potential causal factors" should be "either SWD or hydraulic fracturing (HF) are potential causal factors"

21 You mention but don't define "D values" (I assume it's diffusivity)

22 "That said, the method assumes a temporally consistent seismic response within each grid-cell, which does not appear to be the case in the basin. Rather, the basin is dominated by isolated sequences."

Obviously you don't want to go into too much detail in the abstract, but it's not clear what you mean by a non- "temporally consistent seismic response" why that is linked to a lack of widespread fault reactivation, and this is not elaborated on in the paper either.

65 "In the past 15 years, there have been six main earthquakes sequences in the basin (Figure 1). The" should be "In the past 15 years, there have been six main earthquakes sequences in the basin (Figure 1): the"

68 You discuss all the other sequences in more detail, but not the Lake Lewisville sequence.

91 “Since the shut-in of the nearest SWD well, the sequence continued into at least 2015 slowly migrating northeast along the 6 km causative fault (Ogwari et al., 2018).”

When were the SWD wells shut in? What happened to the earthquake sequence after 2015?

103 “widely felt” should be “most widely felt”

131 “later” I think should be “latter” or more clearly “hypocenters”

142 “Notably, ever decreasing disposal rates from their peak levels in 2014 did not appear to coincide with lower rates of seismicity across the basin at the time of publication (Figure 2b).”

The rate of seismicity shown in Fig 2b does look like it’s decreased significantly since the peak in 2014/2015. Possibly plotting yearly EQs would give a smoother graph that would be easier to interpret.

Figure 1. It would be very helpful to add polygons demarcating the different sequences described in the paper, as I had a fair amount of difficulty figuring out which earthquakes were being described in the discussion of the different sequences.

192 “Notably this area that was considered aseismic until 2008.” Remove “that”

208 “Thus, we opted for a magnitude of completeness (Mc) value of 2 for our investigation period.” It might be a good idea to estimate the pre-2013 Mc for your catalogue to justify this choice, since many of the earthquake sequences began before then. I know you said your method doesn’t strongly depend on Mc, but is that still true if you pick too low of an Mc, and the Mc also changes drastically over time?

245 Can you include a little more detail about how the injection data was reshuffled? Was is reshuffled just in space or also in time? You present results with diffused injection, how was that handled with the shuffling?

276 “G-R” should be “Gutenberg–Richter”

296 You say you employed this method only as a sanity check, but the results are presented prominently, I suggest either changing this framing or reducing the prominence of the results.

316 “Finally, the largest events linked to HF according to the HFR was M 3, and it occurred in July 17 2011, less than 1km away from the 2015 Venus Mw 4 event. The block hosting the M 3 event has however high pHF value, indicating that the spatio-temporal link with HF might be coincidental.”

You seem to be treating the HFR result here more seriously than seems justified from the description in the methods. With the large number of HF wells, some coincidental results are expected, and I would expect that this event is likely coincidental, not just might be coincidental. It would be interesting to know the largest magnitude event in a $p<.05$ block to get an idea of the size more likely HF caused events.

319 I think this results section could use a little more detail. There is a $p<.05$ block near Cleburne in the map, but this isn’t discussed. The two methods seem to mostly agree in terms of producing low, moderate, or high potential HF induced seismicity except the Azle sequence, which has a high p-value but moderate HFR. Is there an explanation for that?

Figure 3. The HFR and p-value maps are flipped from the order you discuss them in the results section, I suggest either flipping the order of discussion or the order of the maps.

330 I know you show the D=1.0 results in the appendix, but neither 0.3 or 2.0 are very close to 0.7, so the 1.0 results seem the most likely to be relevant and it would be good to show them in the actual paper. If you think the D=2.0 results are plausible, it would be good to include some discussion of why the 0.7 estimate might be too low.

344 “Perhaps a lower magnitude cut-off is needed for that sequence.” Did you try a lower magnitude cutoff? Why would you expect that to help? This might be better off moved to a new section where you discuss in more depth the possible reasons for the high p-values you observed.

347 “We should note that our framework assumes a consistent seismic response within each block across time, which does not appear to be the case in the FWB, which is dominated by short-lived isolated sequences, with the exception of the Venus sequence (Quinones et al. 2019).”

Can you explain this in more detail? What do you mean by saying the FWB doesn’t have a consistent seismic response within each block over time? If the sequences are short-lived, but the triggers are also short-lived, couldn’t that still be consistent over time with your model? And if the Venus sequence isn’t short-lived, why does it also show no correlation with SWD?

352 “since the completeness issues in the available earthquake catalogs are significant”

Have you tried running your model on just the more recent data with a lower Mc? This wouldn’t work for the sequences that ended before 2017, but could help you disentangle whether the low correlation is due to data quality, or if it’s a real result and there really is a low correlation with SWD.

371 “Our findings regarding SWD are inconclusive since they are very sensitive to the chosen large-scale diffusivity value.”

Is there any work that could be done to narrow down the possible D-values? I’m not suggesting you actually do it, just mention it as future work that could be done to help determine if D=2.0 is plausible or likely.

“That said, our framework assumes a consistent seismic response within each block across time, which does not appear to be the case in the FWB.”

How long of a period does your method require to work? Can you analyze a smaller time period where the response might be more consistent?

379 “We conclude that the Fort Worth basin is not prone to widespread fault reactivation, allowing for effective mitigation measures.”

As I mentioned with the abstract, it’s not clear to me how your results support this conclusion. If anything, since your results suggest a high diffusivity is necessary to explain SWD triggering of the observed seismicity, it seems like mitigating on a well-by-well basis would not be effective, since it’s the overall volume of injection that would be triggering the seismicity, not the near-field effect of a particular well.

Appendix

Figure S3b. It would be helpful to see the diffused results for the other values of D.

Round 2

Reviewer 1

General Comments

Overall, this is a concise and well-written article that effectively applies a previously developed approach to a well-documented case of induced seismicity in the Dallas–Fort Worth region of Texas. The seismology is outstanding in this paper, and the modeling is state of the art. Having reviewed an earlier version, I believe the manuscript is worthy of publication, pending revision to the conclusions. My main disagreement is not on the results, but how the results are being framed.

This is the response to my comment on:

My comment: “I am concerned that the conclusion—stating no significant link between saltwater disposal (SWD) and the Venus, Azle, or Cleburne earthquake sequences—may reflect limitations of the modeling approach rather than the actual geologic and operational reality…”

Response: “…Next, we should note that our methodology is fully capable of accounting for persistent far-field effects of deep SWD (see section A4 in the Appendix, Figure S2b). Even for very large D values (2 m2/s), we could not link the Venus, Azle, or Cleburne sequences with SWD. The issue is thus not far-field triggering, but rather the limitation that “our framework assumes a consistent seismic response within each block across time, which does not appear to be the case in the FWB, which is dominated by short-lived isolated sequences”. We mention this limitation in section 4.2, in the conclusions and in the abstract. This likely explains why Azle and Cleburne were not flagged. Moving on to the Venus sequence, which was not particularly short-lived, we still do not find a high likelihood of a causal link. This is not very surprising given how different our approach is the forward deterministic modeling approaches adopted by the previous studies. These studies demonstrated that a causal link is physically possible given reasonable parametrization. The uniqueness and statistical significance of this association remained unquantified. We believe there is value in addressing the same question using different scientific frameworks and we encourage further research into the causal factor of the Venus sequence. Lack of consensus is often the first step towards deeper understanding of such chaotic processes. We stand by our key conclusion was that “the FWB is not prone to widespread fault reactivation at the basin-wide scale”. “

I am fine with the conclusion that the FWB is not prone to widespread fault reactivation (a solid conclusion based on their results), but again, stating that there is no connection between saltwater disposal (SWD) and the earthquakes in the Venus, Azle, or Cleburne areas is misleading, even though it may be consistent with your model. Given that reductions in injection volumes were followed by reduced seismic activity, a causal link appears supported by observations. Clarifying this point in the paper (abstract, discussion, and

conclusion) would strengthen the manuscript and ensure the conclusions are accurately represented.

Reviewer 2

I thank the authors for their detailed responses to my comments. The new additions to the introduction help clarify the location and timing of the events described. I also appreciate the reorganization of the methods section, and the significantly increased clarity in the results section. However there are still some areas that I think could be clearer, listed below (line numbers refer to my original comments, not the revised paper). Specifically, there were some helpful clarifications you made in the response to reviewer that you didn't include in the paper. Finally, I am still somewhat confused by your final concluding paragraph, as some of the conclusions drawn there don't seem to obviously follow from the work done in the paper.

142 I suggest you clarify directly in the paper that you are comparing the pre-2014 level of earthquakes to the post-2014 level, as you do in the response. I still suggest the figure would be significantly easier to interpret if you plotted yearly earthquakes instead of monthly, given the low and irregular values involved.

330 I wasn't trying to suggest that you should rerun your results with D=0.7, just that you should put your D=1.0 results in the main paper instead of the appendix, since that's the closest value to 0.7. If you can only put two figures in the main paper for space reasons, I suggest that the D=0.3 could be moved to the appendix, to more clearly show that even a D value slightly greater than the 0.7 you suggest is most likely doesn't produce strong correlations with SWD over most of your study area.

371 I suggest you directly include this response in the paper somewhere.

379 I still don't understand how this statement is justified as a conclusion of the work done in this paper. The spatial correlation of SWD wells with seismicity was already known, and does suggest that mitigation can be done on a well-by-well basis. The novel results from this work, as I understand them, are the HF results and that SWD can be strongly linked to the Irving-Dallas and Lake Lewisville sequences for high D values. You also suggest that this method may not be fully applicable to this region due to the short-lived sequences and inconsistencies in the catalog over time. These are all interesting results, and I appreciate the work that went into producing them. But I don't see how they contribute to this conclusion as stated. You also say in your response "Furthermore, even for the 2 far-field SWD clusters, we did not observe an expanding seismicity front or cascading fault reactivation, justifying the first part of our sentence," but I don't see where this is discussed in the paper. My recommendation is to either further clarify the link between your novel results and this conclusion, or to move the statement to a different part of the paper so that it doesn't appear to be your main conclusion.

Reviewer 3

Grigoratos and colleagues conduct a statistical proximity analyses of earthquake occurrence and locations of both hydraulic fracturing and waste water injection in the Dallas-Fort Worth region. They have employed a similar technique in the Permian and Eagle Ford shale gas exploitation prior to this study of a smaller and possibly more heterogeneous earthquake data set. I think that they have adequately addressed comments from the first review, but I saw areas where the context could be clarified and the paper made more accessible to a broader audience. That is, details of interest to those conducting independent modelling studies are referred to as 'geology', 'basement'. The inclusion of a stratigraphic column in SM, or a few sentences in background regarding depth to basement, orientation of active faults are important to readers comparing different types of analyses, or with other shale gas basin regions.

A second concern is the implicit assumption that the locations of pre-existing fault zones and their orientation with respect to the present-day stress field are not key factors, whereas as other models of, for example, the Timpson swarm clearly show that these are factors. Readers will want to see this explicitly stated, and I certainly hope that the authors move forward with new techniques in future analyses of TexNet data sets.

A third is the assumption of mainshock - aftershock sequences in declustering. Are there clear main shocks, or are these truly swarms? Is largest by 0.1 significant, given magnitude estimation uncertainties? Swarms in extensional areas are commonly groups of similar magnitude events and declustering needs clarification - e.g., Zaliapin and Ben-Zion, 2023 . More should be said here or in SM.

Minor comments

68-69: why not shorten to "Venus sequence in northeast...which hosted the May 2015 Mw 4.0 earthquake?

126 - difficult to drill because of what aspect of the geology? Overpressure?

Fig 2 top - labels are too small.

207-208 - What orientation of normal faults and how far below the injection horizons? Again, a strat column in SM or more information in the text are needed.

301-303 - 2D, not 3D comparison, correct?

334-335 - Too late in text to learn that Ellenburger Fm is highly fractured - see comments above.

In Summary, I recommend publication with further clarification of assumptions, and hope that the authors can move forward with tools that consider blocks with favorably oriented faults to enable comparisons with forward modelling papers.

Round 3

Reviewer 1

I appreciate the authors efforts in addressing my comments. I am fine with the manuscriot as it is now, and beleive it should be published.

Reviewer 3

The reviewers have done a good job clarifying points in the text, and to make the paper of broader interest. There are still parts that ask the readers to do background reading, and that's OK - I just see that a paper with material in main and SM leads to more citations. I am particularly happy to see that the authors found that declustering/no declustering makes little to no difference.

The added comments that the FWB shows heterogeneous responses and is therefore not prone to widespread fault reactivation, and can be handled on a well-by-well basis were much needed. It is bold, but it's also fine to say in terms of the study results. I am used to dynamic systems where everything is time-varying - activation of a basement fault could change that, which is something I am sure the authors know.

I now have a broader perspective on induced seismicity and benefited from reading the comments of the other reviewers.