Dear Dr. Muir,

We hereby submit a revised version of my manuscript “Deep learning detects uncataloged low-frequency earthquakes across regions” to Seismica. Thank you for the handling of the manuscript.

In our revision, we considered all the helpful and constructive points raised by the reviewers. Please find attached a point-by-point response (written in blue) to the comments of the reviewers (written in black) with the revised article in two versions, one with changes tracked and one without. The line numbers in the response letter refer to the version with highlighted changes.

We hope the manuscript now meets the high standards associated with Seismica.

Best regards,

Jannes Münchmeyer for all authors

Editors minor comments:

- It would be good to specify that the Phase-Net architecture is 3 component when describing it (otherwise 1D could be misconstrued)

> We’ve incorporated the suggestion. (Line 117)

- It would be useful to expand the section on spurious detection slightly, in its current terse form it is a little difficult to follow the argument being made (i.e. I'm assuming you then re-associate after scrambling to see what catalogs are formed but this is only implicit in the text as written).

> Indeed, we re-associate the scrambled picks and compare the resulting catalog to the original. We’ve added this clarification to the manuscript. (Lines 234f)

- It would be helpful to provide some geological context for the gaps in Nankai.

> Thank you for the suggestion. However, we’ve decided not to add a discussion on this, as the geological conditions leading to slow earthquakes are still debated and many aspects remain unclear. See, for example, Nishikawa et al. (2023), Section 5.2.

Reviewer A:

This manuscript describes the development of an automatic Low-Frequency-Earthquake (LFE) detection & location workflow based on deep learning, that is applied to data from 3 subduction zones (Cascadia, Nankai, Guerrero) and the San Andreas Fault. The neural network is trained using a superposition of noise sequences (known not to contain LFEs) and LFE template waveforms from the different LFE regions. Results are assessed based on further synthetic tests, and through the assembly of new catalogs for the 4 LFE regions permitting comparison with original template-matching-based catalogs. with generally good results. In particular, regional tremor epicentral migration patterns are well recovered and evidence is provided that the deep learning approach recovers substantial numbers of previously undetected LFEs.

> We thank the reviewer for their encouraging words.

The motivation for this work is certainly worthy, namely, improved tracking of LFEs will lead to better characterization of the details of slow slip with potential implications for improved understanding of how slow-slip and earthquake rupture interplay. My main criticism involves the discussion and comparison with relevant previous work (aside from template matching studies) which I expand upon below.

1) My main criticism is that, aside from template-matching, the current work fails to acknowledge the range of approaches that has been previously applied to detect and locate tremor, and the advantages (and/or disadvantages?) that the deep learning approach might hold over them. I list some pertinent references below for Cascadia. Principal among these are the "cross-station" approaches which include the original work by Armbruster, Rubin and colleagues. Two specific tests are listed in 2) and 3)

> We thank the reviewer for this suggestion and the wealth of references provided. We've added a paragraph discussing methods for tremor detection. However, we'd like to point out that tremor detection and LFE detection are distinct tasks with different challenges. Tremors are usually several minutes long, making them easier to detect than LFEs that are short signals, with waveforms lasting at most a few seconds. Therefore, tremor detection methods can not be directly compared to LFE detection methods. As for the work of Savard et al, we now discuss their method in the related work and provide a comparison to their results (see the answer to the next point). (Lines 79ff)

2) One of references listed below (Savard et al) applied a cross-station approach for regional catalogue assembly with results much like those presented here, at least visually. In addition, they presented kernel density plots of their detections (see their Figure 7). I think it would be useful if the authors considered preparing similar plots, at least for Vancouver Island, for comparison purposes. Tremor on Vancouver island is unevenly distributed with some particularly active hotspots separated by regions which radiate little or no energy. The degree to which the deep learning framework could improve on the delineation of these hotspots in Cascadia over those in Savard would provide testament to its superiority as a regional detector/locator. See also studies by Armbruster, Rubin and colleagues regarding the high precision definition of these hotspots.

> Thank you for the suggestion. We've added density plots to the supplementary material (Figure S5). We compared the figure to the KDE plots from Savard and Bostock (2015). Indeed, the results show good alignment with the hotspot regions identified in their study. This behaviour is particularly clear for the patches A and C

highlighted in Figure 7 of Savard and Bostock. Our catalog differs in the substantially higher number of events underneath the tip and just offshore Vancouver Island. This is consistent with our later findings regarding the diversity of events. We added this discussion to the manuscript. (Lines 193ff)

3) Another question that would be useful to address is how effective is the deep learning detector/locator in less favourable station-separation circumstances? In much of Cascadia, the regular station sampling is more like ~30 km. The Wech/PNSN tremor algorithm manages to detect and map averaged tremor epicenters reasonably well with an accuracy probably on the scale of +/- ~ 10 km along the entire Cascadia margin. It would be useful to demonstrate how favourably the deep learning detector/locator would compare to the Wech scheme in this 30 km station spacing scenario. I imagine it should be easy to set up a comparison with the Wech catalog for a given ETS event using a common set of stations(?) Note that the online PNSN catalog is at https://pnsn.org/tremor/

Thank you for the suggestion. We've added a comparison to the PNSN tremor catalog for a 31 day period in 2021, using exclusively waveforms from the CN network with its ~30 km station spacing (Figure S6). The study region for this comparison focuses on Vancouver Island but takes into account a larger part of it than our comparisons with the Bostock et al catalog. The results show a clear agreement between our LFEs and the PNSN tremors, even though the LFE catalog shows higher scatter. Nonetheless, we'd like to reiterate our statement from the reply to comment 1), that LFE and tremor detection are fundamentally different, even though the underlying processes are closely related. (Lines 202f)

4) Re line 219, are the uncataloged (new) detections for Cascadia distributed through the mapped tremor region in Figure 2? or are they concentrated at the southern and northern borders?

Additional detections are more prevalent at the borders but occur also in significant amounts within the center of the region, where the coverage in the Bostock catalog is good. We've clarified this in the manuscript now. (Lines 243f)

5) To add value to the current study, the authors should be encouraged to supply their novel "uncataloged" detection catalogs for the 4 LFE regions in a data repository or as supplemental materials, for the benefit of future tremor research.

Thank you for the suggestion. For now, we refrain from making our catalogs openly available, as these catalogs only cover short time periods and are not manually vetted. However, we make our model available in an easily accessible format (and with an associated DOI) to enable interested scientists to build their own catalogs.

6) On line 188 it is stated that "LFEs show a continuous decrease in energy from low frequencies onward". This doesn't seem to be true for Cascadia, ie the left panel in Figure 4, or am I missing something?

Your observation is correct. We rephrased this part to be more precise. (Lines 216ff)

Reviewer C:

This study applies deep-learning-based phase picking to detect low frequency earthquakes (LFEs) in multiple geographic regions. As the authors emphasize, this is an important yet challenging task, with potential to identify new, previously unrecognized areas or times of LFE activity, providing important constraints on the associated fault slip behaviors. The authors train and benchmark their analyses using template-based LFE catalogs (Cascadia, San Andreas, Guerrero) as well as the Nankai LFE catalog obtained by traditional phase picking. The authors do a very nice job describing the strengths and weaknesses of template versus deep-learning methods, including how they can be complementary to each other. This is an impressive study – careful, comprehensive, clearly written, with clear figures. I think it will be a model for future efforts. In my opinion, it can be published after very minor revisions considering the comments below.

> Dear Dr. Shelly, thanks for your positive comments about our study. Please find the responses to your detailed comments below.

Specific comments:

Lines 59-60: "such a workflow is not applicable to LFE detection..." This is mostly true, but the analyzed JMA catalog from the Nankai region does provide one counter example. It might be worth mentioning here – presumably that catalog is enabled by extremely high-quality seismic data and manual effort. Nevertheless, as noted later in the manuscript, this catalog contains relatively few LFEs.

> Thanks for pointing this out. We've now corrected the statement, explicitly including the case of the JMA catalog and explaining why this workflow is applicable here. (Lines 61ff)

Line 72-73 (also Line 238-240): It might be worth clarifying the location part here – although many template-matching studies don't attempt to individually locate events because of low SNR (making individual locations worse than just assuming that detections are co-located with a template event), that doesn't mean that a location step cannot follow a detection step for events detected via template matching. In fact, this is commonly done for earthquakes identified via matched-filtering, and occasionally done for LFEs (here's one example for LFEs: https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2008GL036367) - but it's obviously difficult because of low SNR. Regardless of which technique (deep learning or matched filtering) is used for detection, a location step can be applied.

> We have now clarified this point and added a reference to the suggested paper as an example for locating individual LFEs identified through template matching. (Lines 274ff)

Line 150 (and Figure S4): It's striking that there is such a huge range among the areas in terms of how many events are in the "overlap" category (detected by both LFE and earthquake models). Can you elaborate on this? It seems that perhaps this should be emphasized discussed more, because of the potential pitfall of detecting "typical" earthquakes instead of LFEs with an LFE model. This study wisely took the extra step of applying an earthquake-trained deep-learning model, but others considering similar work might miss the importance of that step unless it is further emphasized.

Thanks for pointing this out. We've now added a discussion on this overlap in detections. While we don't know the reason for this overlap, and especially why it manifests differently in the different regions, we now point it out more clearly and provide hypotheses regarding the cause. (Lines 164 ff)

Line 206: "The" -> "Therefore"?? (though therefore is also used later in the sentence?)

Thanks, we fixed this.

References

Nishikawa, Tomoaki, Satoshi Ide, and Takuya Nishimura. 2023. "A Review on Slow Earthquakes in the Japan Trench." Progress in Earth and Planetary Science 10(1):1. doi: 10.1186/s40645-022-00528-w.