General comments

Köhler and Weidle present results from two temporary seismic arrays installed in permafrost soil in Svalbard. They calculate horizontal-to-vertical spectral ratios (HVSR) for each station. At some stations, they observe a peak that glides to lower frequencies during the thawing period. They convincingly show, using simple numerical models, that this could correspond to a thawing layer with low shear wave velocity in the uppermost meters, addressing also potential issues due to the Nyquist frequency limit. They then give advice for future seismic projects in permafrost regions. Finally, they discuss whether an observed seismic tremor source could be used for HVSR studies.

This article is well-written and mostly clear. I do however have a number of specific comments for the text, for the organization of the manuscript, and especially for the figures. My main concern is the relation between HVSR from ambient noise and from the tremor source, which should be discussed in more detail to improve the coherency of the paper.

After minor revisions, I think that this paper would be an interesting contribution to the special issue in environmental seismology.

Specific comments

1. I am missing a paragraph with some explicit statements about how the HVSR of the ambient noise and of the tremor complement each other. Why would you need the tremor HVSR at all, what value does it add to noise HVSR? As is, section 6 is somewhat detached from the rest of the manuscript. By relating it better to the ambient noise HVSR, this part would be better integrated into the paper.
2. Organization: Sections 5, 6 and 7 would profit from a better structuring, i.e. clearly separating results from discussion, and giving recommendations only at the very end. Personally, I found the high number of enumerated lists and sublists confusing rather than helping. Please note that the order of these sections as stated at the end of section 1 is different to the order in the abstract and in the manuscript itself.
3. Figures 2, 3, A1, A2: I appreciate that you are showing HVSR from all stations. However, I did not get why the stations are ordered the way they are, and not e.g. ascending BRA1-8, and KBS1-4 (or similar). It took me some time to find the corresponding HVSR for stations mentioned in the text.
4. p.2 l.28 I suggest to better explain jargon (e.g. GSN, BH/HH channels, trigger mode) to the potential non-seismologists in the audience
5. p.4. l.1 does "North" refer to the spectrum or the time-domain record, i.e. do you average the raw data or the spectra?
6. p.4 l.4 Do you use Konno-Ohmachi smoothing? If so, please mention this and specify your smoothing constant. If not, please explain how you smooth your spectra.
   Konno, K., & Ohmachi, T. (1998). Ground-motion characteristics estimated from spectral ratio between horizontal and vertical components of microtremor. Bulletin of the Seismological Society of America, 88(1), 228–241.
7. p.7 l.4 You assume a 1-D subsurface, "inspired" by Haldorsen and Heim (1999). Could you please explain what each layer of your model corresponds to (e.g. regarding the units in Fig. 3+4 in Haldorsen and Heim (1999))? Why do you think that the 1-D assumption is justified given the clearly dipping layers?

8. p.8 Table 1: I think a figure would be much more helpful
9. p.8 l.5 Fäh et al. (2001) and Poggi et al. (2012) divide their spectra by a factor of sqrt(2) to compare the amplitudes. From Fig. 4 it looks like this would match quite well.
Fäh, D., Kind, F., & Giardini, D. (2001). A theoretical investigation of average H/V ratios. Geophysical Journal International, 145(2), 535–549. https://doi.org/10.1046/j.0956-540x.2001.01406.x
Poggi, V., Fäh, D., Burjánek, J., & Giardini, D. (2012). The use of Rayleigh-wave ellipticity for site-specific hazard assessment and microzonation: application to the city of Lucerne, Switzerland. Geophysical Journal International, 188(3), 1154–1172. https://doi.org/10.1111/j.1365-246X.2011.05305.x
10. p.8 l.15 wind: do you think that the wind directly affects the instruments, or do you think that the wind affects the ground which then is picked up by the geophones?
11. p.8 l.17 only wind noise or other noise as well?
12. p.9 l.8 This is the first time you mention tilt of the instruments. How did the instruments look like when they were dismantled, were they still leveled? I suggest to mention this in the Data section
13. p.9 l.9 Albaric et al. (in prep) does not appear in the references. Please elaborate or remove.
14. p.10 l.13 How deep is this concrete shelter, and how far away from the active layer?
15. p.11 l.13 amplitude spectrum: please rephrase by saying that you take the Fourier transform of these two time series (amplitude spectrum is technically correct, but a bit confusing in this context)
16. p.11 l.20 I am missing an actual physical mechanism of the tremor generation. Ocean waves have a lower frequency (in the microseism band) than the observed tremor. Why do you think that the cliff would vibrate at 4-5 Hz? What exactly would vibrate? What is the role of the cave, what would be this amplification (p.16 l.31)?
p.11 l.23 "is a good explanation" I see that this phenomenon correlates with the tides, but in my opinion the source mechanism is not very clear, and should be discussed in more detail.
17. p.12 l.6 (and also in the summary p.14 l.16). Based on what test and significance criterion do you conclude that this is significant if it is within one standard deviation from the other?
18. p. 13 l.7 The horizontal and the vertical components are affected in the same way only if the source is a pure Rayleigh wave source.
19. p.15 l.15 What is the network code for KBS? Are there DOIs for the seismic datasets?
20. p.16 l.22 Only 31 tremors? In Fig. A3 I count at least 16 in one month.
21. p.17 l.23 I do not understand how the depth sensitivity plays a role, please elaborate.

Figures:

1. What is the source of the background image? Where is the borehole of Boike et al., 2018? There is a typo in Ny-Ålesund in panel b. Subtitle c is closer to panel b than panel c. It would be helpful to label the axes with North and East

2. In the top panel, it's almost impossible to distinguish red from dark red (same goes for Fig. 3 and those in the appendix)
3. see Fig. 2
4. This figure is quite busy. I suggest to make separate subplots with only the dashed lines, and subplots with only the solid lines and the same x-axis scale as d). Please also mention in the caption what the dashed lines show.
5. Why didn't you pick any peak frequencies in the end of July and in the beginning of August? I suggest to make more picks, and remove (or decrease the size of) the black dots from the figure, as you suggest that these are gliding peaks rather than discrete occurrences. Additionally, what line corresponds to which station? I recommend to plot the lines in different colors and make a legend.
7. In a), neither the "legend" nor the caption state whether dashed is summer or winter. In c), dark red and light red can hardly be distinguished. I am also missing a legend. As far as I understand, all RVSRs are from the tremor. If this is correct, please state so in the caption.

A3 It would be helpful to show the picks of your STA/LTA algorithm on this figure.

Technical corrections

p.1 l.4 (and several other places) thawn → thawed

p.2 l.15 it's → its (also p.10 l.16 and p.14 l.11)

p.2 l.20 anthropogentic → anthropogenic

p.5 l.23 extend → extent

p.6 l.5 theory,

p.7 l.2 S wave → S-wave

p.8 l.13 broadband ~~band~~

p.8 l.13 in and during beginning of the melt season → in the beginning and during the melt season?

p.9 l.13 kms^{-1} → km s^{-1}

p.9 l.14 HVSR spectrum → HVSR curve?

p.9 l.23 relative → relatively

p.9 l.26 issues → issue

p.9 l.30 peaks/amplitude increase → peaks or amplitude increases

p.9 l.31 their → they

p.10 l.4 An → A

p.10 l.6 as well?

p.10 l.16 exist → exists

p.11 l.11 Weird wording and word order. Better to say e.g. like this: However, during neap tides and low wind speeds, some tidal maxima do not have a corresponding detection.

p.14 l.2+4 However used in two consecutive sentences

p.16 l.9 are → were

p.16 l.11 mis-detections → false positives

p.17 l.22 The lack of Rayleigh wave energy in this freq. band: which component are you talking about?