Review of “Seismic full-waveform inversion constrained by weighted correlation envelope with learnable envelope power”

Dear authors:

The motivation of the manuscript is to mitigate the notorious cycle-skipping issue by utilizing the data envelope, meanwhile, improve the inversion accuracy by neural network (NN). From my perspective, the manuscript is overall well-structured and the writing is clear. However, in my opinion, there are still some key points that are not fully explained and validated, please find my comments and questions below.

The core content (novelty) of the manuscript is the objective function design based on envelope correlation with a learnable envelope power parameter. Mitigating the cycle-skipping problem by incorporating the envelope information is not much a novel idea since it is early proposed by *Chi et al., (2013).* Joint inversion by combination of waveform and waveform envelope has also proposed by Liu and Zhang (2017.). The objective functions between the proposed FWI-NN-WECP and previous WECI (*Song et al., 2023*) differs mainly in the envelope power parameter selection. Therefore, I think a key point of this paper is to explain the role and influence of the variable envelope power in the inversion. However, regarding this issue, on the one hand, in the methodology section, the authors' analysis based on Figure 3 and Figure 4 in the manuscript is not convincing enough. From my perspective, the data of different envelope power parameters does not show much difference, especially for deep reflections. In other words, what impact will this data difference have on the inversion? On the other hand, in numerical example part, for comparison, the author should explain why the envelope power parameter in the WECI methods is set equal to 1, because I noticed that in the study of *Song et al. (2023)*, inversion tests with envelope power of 1.5 obtained better results. I understand that fair comparisons help to better reveal the difference the variable envelope power parameters bring in the inversion.

Another innovative point of the manuscript is that the proposed FWI-NN-WECP method utilized a neural network to reparametrize the velocity model, incorporating FWI-WECP loss function that features a learnable power of envelope. The results in the manuscript show that the inversion results of the FWI method with the neural network (FWI-NN) are generally improved. Can you further explain the reason behind this phenomenon? By the way, I believe it is better to include more description about the conventional and neural network-based FWI workflows. In addition, for different experimental tests, can you further explain the computational efficiency of these methods? For example, how long does it take for 1500 iterations in the sigsbee model experiment? If possible, how does it compare with the traditional adjoint state method based FWI? (This is a crucial aspect in practical 3D FWI applications).

References:
-Benxin Chi, Liangguo Dong, Yuzhu Liu, 2013, Full waveform inversion based on envelope objective function,73th Annual International Conference and Exhibition, EAGE, Extended Abstracts, London, UK.
-Liu Z Y, Zhang J. 2017. Joint traveltime, waveform, and waveform envelope inversion for near-surface imaging. Geophysics, 82(4), 235–244.
-Song, C., Y. Wang, A. Richardson, and C. Liu, 2023, Weighted envelope correlation-based waveform inversion using automatic differentiation: IEEE Transactions on Geoscience and Remote Sensing, 61, 1–11, doi: 10.1109/TGRS.2023.3300127.

The literature review is generally fluent since the authors are focused on utilizing seismic envelope data and incorporating neural networks to mitigate the cycle-skipping issue. However, since this is a very fundamental challenge in FWI, I think it would be better to include a wider range of related studies, for example, seismic traveltime inversion (*Ma and Hale, 2013; Wang et al., 2023*), phase inversion (*Choi et al., 2015*), adaptive waveform inversion (*Warner et al., 2016; Yong et al., 2022, 2023*) and so on. In addition, in recent years, FWI combined with neural networks has developed to alleviate the cycle-skipping. These highly related studies should be included in the discussion to some extent.

References:
-Choi, Y., & Alkhalifah, T. (2015). Unwrapped phase inversion with an exponential damping. Geophysics, 80(5), R251-R264.
-Ma, Y., and D. Hale, 2013, Wave-equation reflection traveltime inversion with dynamic warping and full-waveform inversion: Geophysics, 78, no. 6, R223 R233.
-Wang, Y., L. Dong, J. Wang, and J. Zhang, 2023, Characteristic reflector-based wave equation reflection traveltime inversion: Geophysics, 88, no. 3, R323-R337.
-Warner, M., and L. Guasch, 2016, Adaptive waveform inversion: Theory: Geophysics, 81, no. 6, R429-R445.
-Yong, P., R. Brossier, L. Métivier, and J. Virieux, 2022, Localized adaptive waveform inversion: theory and numerical verification: Geophysical Journal International, 233, no. 2, 1055-1080.
-Yong, P., R. Brossier, L. Métivier, and J. Virieux, 2023, Localized adaptive waveform inversion: regularizations for Gabor deconvolution and 3-D field data application: Geophysical Journal International, 235, no. 1, 448-467.

Some minors:

1. In equation 3 on page 4, *p* (represents the power of envelope) is a subscript? Please check.

$$J_{ECI} = 1 - \frac{1}{N_s N_r} \sum_s \sum_r \frac{1}{\left\| e_p(x_s, x_r, t) \right\| \left\| e_d(x_s, x_r, t) \right\|} \int_t e_p(x_s, x_r, t) e_d(x_s, x_r, t) \mathrm{d}t, . \tag{3}$$

2. In the two synthetic examples, the authors provided 6 inversion results with different FWI methods, but does not show more inversion details. I believe that more quality control of the inversion would be helpful to readers, such as the matching of data before and after inversion.

3. Check the literature citation on page 17, line 10: "we present their corresponding reverse time migration (RTM) images in Figure 8 (*Richardson, 2023*). "

4. On page 20, line 2, "…ranging from 2000 km/s to 3800 km/s,", the units should be " m/s ".

5. In the inversion results of the field data (Figure 15), we can see that the incorporation of the neural networks has significantly improved the inversion results on the right side of the model (where data coverage may be insufficient), which is also the area where the well-log is located (Figure 15). Can you further explain why the neural network brings about improvements here?

6. Which PDE is used to predict the data in Figure 16, acoustic or elastic wave equation?

I wish my comments are fairly straightforward and helpful. Again, I hope to see an updated version of the manuscript. I wish you the best of luck.

Sincerely,
Reviewer