Reviewer # 1 comments:

This is an interesting and timely manuscript which has the merit to apply seismic tomography methods to a problem of societal relevance: the security of nuclear power plants.

The main point seems to be that the comparison of Vp and Vs tomography results suggests the presence of fluid in the region of the Iwaki earthquake. It is indeed a good idea to combine Vp and Vs results to get information on Earth structure other than elastic parameters. Some of the difficulties involved by this endeavor are, however, not discussed. In particular, how was regularization selected for the Vp and Vs inversion? In principle, if the results are to be compared, the regularization scheme in the two cases should be the "same", but the problem is that the Vp and Vs inverse problems presumably have different properties (coverage, signal-to-noise ratio of the P and S data...) so that the same numerical values of the regularization parameters would not result in an equivalent level of regularization. I believe that clarifying this point would make the paper more convincing.

A related issue is that of the metric used to compare different models: what exactly is the "structural similarity index", or SSIM? please give at least a reference of a publication where this quantity is clearly defined. Is SSIM, like correlation, just sensitive to the pattern of heterogeneity, or does it also compare amplitude?

Note that a general issue with tomography is that, while the geographic pattern of structural heterogeneity is reasonably easy to constrain, it is much harder to determine robustly its amplitude. As the authors' inferences rely on the ratio of Vp to Vs, rather than just on the distribution of lateral variations in either quantity, amplitude is here very important. How stable are values of amplitude with respect to variations in the regularization scheme? How does amplitude of the solution models respond, in synthetic tests, to the level of random noise added to the synthetics?

I am sure that it will not be difficult for the authors to address the points I mentioned, and believe a very minor revision will be sufficient to make this nice manuscript fully acceptable.

Response:

We are very grateful to Prof. L. Boschi for his thoughtful comments on our manuscript.

We agree with the reviewer that the Vp and Vs inverse problems have different properties. In most cases, the inversion depends on choosing a suitable regularization parameter. In this study, we selected the optimal regularization parameters (including the damping and smoothing parameters) after detailed analyses of the trade-off between the data variance reduction and model smoothness for the Vp and Vs inversions, respectively. Based on such analyses, the damping and smoothing parameters are chosen to be 5.0 and 0.1 for the Vp inversion, and they are 6.0 and 0.1 for the Vs inversion. We found that the optimal damping and smoothing parameters for the ray tomography are almost the same as those for the finite-frequency tomography. Therefore we used the same regularization parameters for the finite-frequency and ray tomographic inversions to make a valid comparison between the two tomographic methods. We have added related discussions on the determination of the regularization parameters in the revised manuscript. Please see Lines 110-116 on Pages 5-6.

The structural similarity index (SSIM) was defined in our previous publication (Tong et al., 2011, GJI), and it was used in the present work. Here we quote its definition as

The SSIM index between two (velocity or any positive physical parameter) structures A and B can be defined as

$$SSIM\,(A,B) = \frac{2\mu_A \mu_B \sigma_{AB}}{(\mu_A^2 + \mu_B^2)(\sigma_A^2 + \sigma_B^2)} + 0.5$$

where μ_A *is the average of A,* μ_B *is the average of B,* σ_A^2 *is the variance of A,* σ_B^2 *is the variance of B, and* σ_{AB} *is the covariance of A and B. The resultant SSIM index is a decimal value between 0.0 and 1.0, and it is 1.0 only when A and B are identical.*

This index can show quantitatively the degree of similarity between two models (velocity models in this study). Since the averages and variances of two models are involved in the definition of the SSIM index, this index not only reflects the pattern of tomography but also compare the amplitudes of velocity anomalies. In the revised manuscript, we have explicitly stated that the SSIM index was defined by Tong, Zhao, and Yang (2011) GJI. Please see Line 121 in Page 6.

As the reviewer mentioned, it is hard to determine robustly the amplitudes of velocity anomalies with a regularization scheme. In this work, we have used nearly the same amount of P and S wave data (199,363 P-wave and 184,919 S-wave arrival times from 6506 earthquakes) in the tomographic inversions. The same large number of P and S wave data result in dense and uniform ray-path coverage in the study area, thus both the pattern and amplitudes of Vp and Vs anomalies are well recovered. This is verified by the resolution tests. The similar results generated by the two different tomographic methods also suggest that both Vp and Vs models are reliably determined.

In the synthetic tests, random errors with a standard deviation of 0.1 s were added to the synthetic arrival times calculated for the checkerboard model to account for the picking errors existing in the real data. The amplitudes of velocity anomalies can be well recovered as shown in Figs. S1-S8. However, if we increase the standard deviation of the random errors from 0.1 s to 0.3 s, the structural similarity indices between the inversion results and checkerboard model will decrease. But if the standard deviation of the random errors is less than 0.1 s, reflecting the actual picking error, the structural similarity indices have no significant variations. This suggests that the data set used in this study is good enough to recover the Vp, Vs and the Poisson's ratio structures. In the revised manuscript, we have added related discussions on the determination of the Poisson's ratio image. Please see Lines 102-107 in Page 5.