Thank you for the review and comments. We provide below our responses to the points raised, including details on the modifications made to the manuscript.

1. To make a general point at the outset, using RMS as the sole measure of disagreement is hazardous. For example, in Figure 2, how do we know that the large RMS on the coast of China is not just one badly discrepant model rather than a Gaussian-like scatter? Or, alternatively, perhaps this RMS is large because the many global models used do not agree well: what I would want to know is how well the three most modern ones (FES2014b, TPXO9-Atlas, and GOT4.10c) agree.

We use the inter-model standard deviation (as shown in Figure 2) and then in Table 2 list the RMS agreement *per model* with respect to tide gauges in each defined sub-area. For the eastern China sub-area, the per model RMS values in Table 2 indicate that there are many models contributing to the sub-area's large scatter shown in Figure 2a, and that only NAO99Jb and FES2014b are accurate there, with RMS differences of 10-12 cm, whereas TPXO9-Atlas and GOT4.10c have RMS differences of 30-35 cm. To further confirm this, in Figure R3.1 we show, per pair, plots of the phasor differences among the three most modern models, namely FES2014b-GOT4.10c, FES2014b-(TPXO9-Atlas) and GOT4.10c-(TPXO9-Atlas). It can be seen that for all three pairs very similar patterns are obtained as for the standard deviations for all nine models shown in Figure 2, and we have added a couple of sentences to Section 2 (paragraph 3) to inform the reader. The only noticeable difference is that the older model differences tend to tail off slightly less rapidly on moving from the coast. These plots, together with the RMS differences with respect to the tide gauges listed in Table 2, do not suggest that the three most modern models should be chosen over the others in this region.

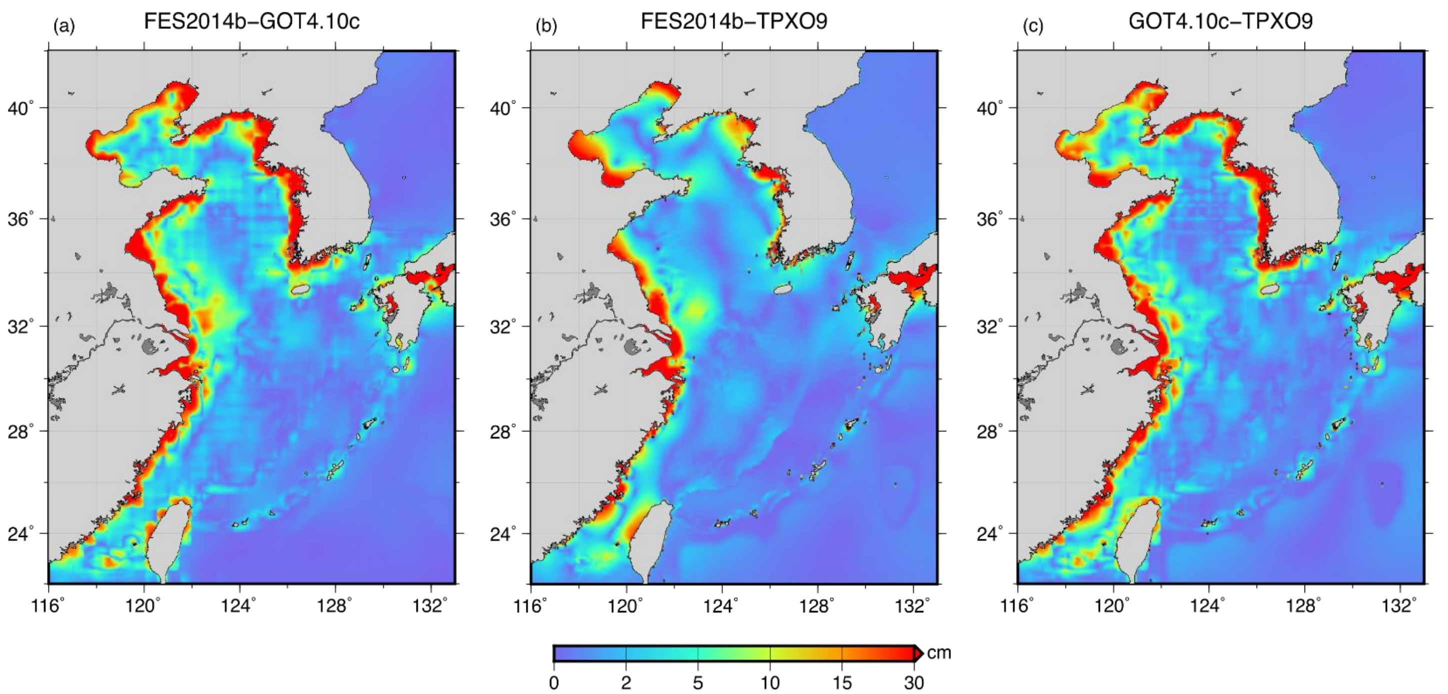

Figure R3.1 The M_2 phasor differences between each of the three most modern ocean tide models: FES2014b, GOT4.10c and TPXO9-Atlas.

2. Another general point is that the "East China Sea" in the title is misleading: the authors use data from the many GPS stations on Kyushu, a smaller number (but still quite a few) from the Ryukyu Islands, three in Korea, two in Taiwan, and one on the Chinese mainland. For tide gauges the same distribution is similar, except that there are six stations on the Chinese mainland and none on Korea. Any results, particularly any RMS values, will therefore be only about the first two areas, and especially Kyushu: for the GPS, the Pacific is likely to be as or more important than the East China Sea in producing almost all of the loads. I appreciate that the authors want to use as many stations as they can, but I think the paper would be much better if the few non-Japanese stations were omitted. This would also avoid a problem with Figures 4 and 5, which is that where most of the data is, it is impossible to see the results in any detail. Even if the authors do keep the few other stations, they should use a set of more focused maps, perhaps with the Kyushu-Ryukyu stations shown using an oblique Mercator.

We have changed the manuscript title to now state "around the East China Sea", not "in the East China Sea". To improve the presentation of Figures 4 and 5 (as also suggested by Reviewer 2), we now include an enlargement of Kyushu and most of the Ryukyu Islands on oblique Mercator plots. We consider the non-Japanese sites to still provide useful information as they are further from the local loads and hence provide a control on deeper mantle behaviour. Regarding the load contribution from the Pacific Ocean, we agree that a significant (but not dominant) proportion of the OTL is caused by this (as shown by Figure R3.2 below and in Table 3); but again the non-Japanese sites are useful in widening the aperture of our array to allow its effects to be distinguished from local loads. The Rest of NAO99Jb phasors in Figure R3.2 include the part of the Pacific Ocean contributing most to the loading, but inter-model variations are small, as they also are for the Central ECS contribution. The biggest impact on the loading and errors at the three sites considered comes from the very local sea areas, and where the importance of using NAO99Jb is shown.

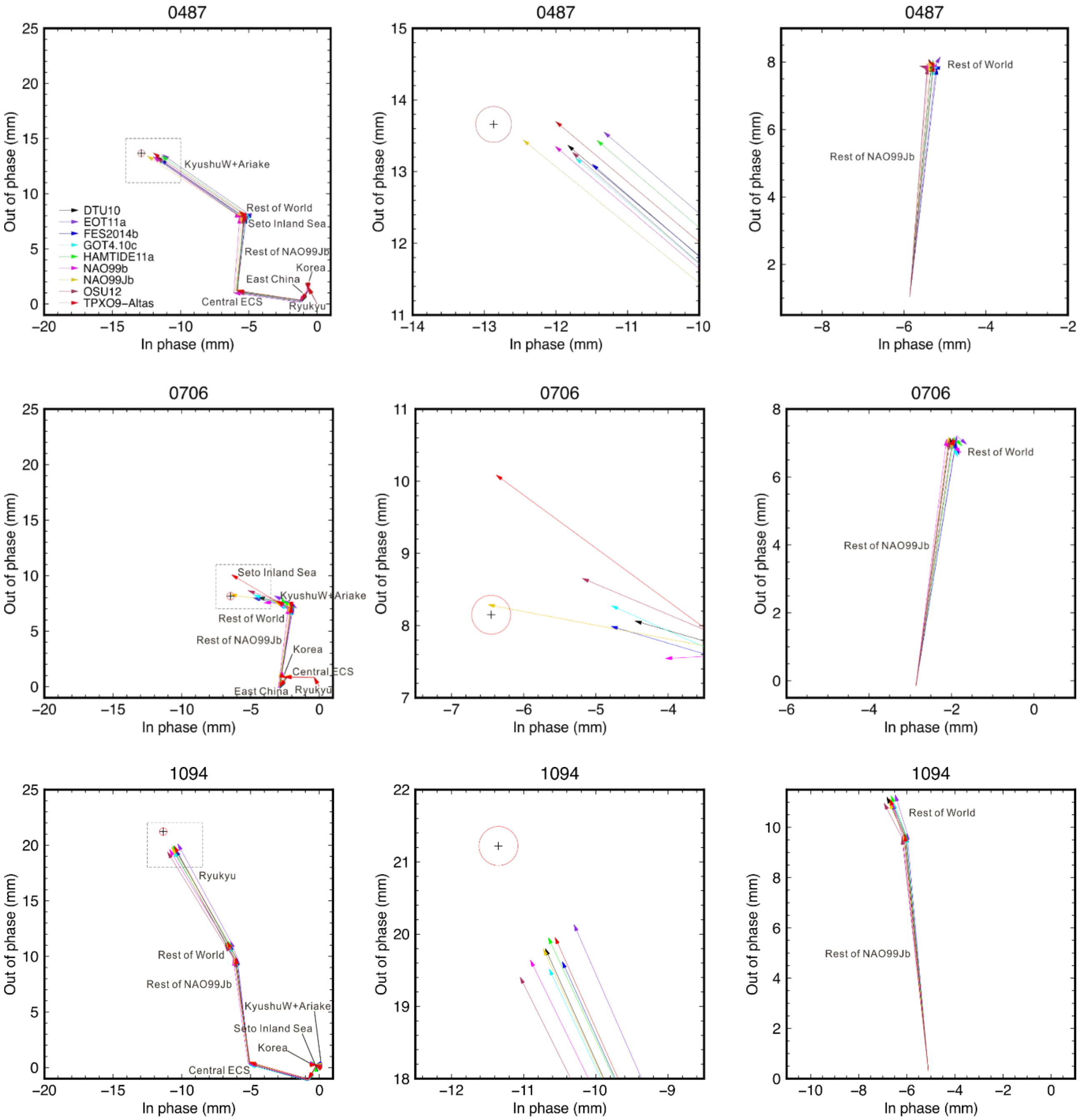

Figure R3.2 Phasor plot of the M_2 vertical OTL displacement contributions from the water sub-areas Eastern China, Korea, Central ECS, Seto Inland Sea (including the Kanmen Straits), Ryukyu Islands, Rest of NAO99Jb (comprises all the NAO99Jb coverage except the aforementioned sub-areas) and Rest of World.

3. This geographic imbalance leads to another problem, namely the authors' conclusion that the NAO99Jb model should be used, despite its age, because of its lower RMS compared to the tide gauges. But the authors' own Table 2 shows that for the most modern high-resolution global tide models (again, FES2014b, TPXO9-Atlas, and GOT4.10c) this lower RMS is confined to nearly-enclosed seas: for these NAO99Jb does much better. As the authors note, this is hardly surprising. The question is, how important are these enclosed seas in computing the loads?
I computed loads in two ways. A was to use all of the NAO99Jb model, and TPXO7.2atlas for the remaining global parts: close to the authors' procedure. B was to use the NAO99Jb model only inside the polygons and TPXO7.2atlas everywhere else. Figure 2 shows the results, as contours of the ratio of the M2 amplitude in vertical displacement for B, divided by the same thing for A. Two features of this plot are notable. First, the ratio is spatially smooth, which means that these enclosed seas only contribute to the estimated load for very nearby stations, so that NAO99Jb needs to be used only in these limited areas. The other is that there is, clearly, a systematic difference between loads that used NAO99Jb regionally and those that used it locally: this systematic difference might well make a difference in the authors' comparisons and conclusions. So I'd like to see the authors compute the loads using NAO99Jb only for limited areas, and more modern models (the three I've mentioned) for everywhere else.
We disagree that the lower RMS of NAO99Jb over FES2014b, GOT4.10c and TPXO9-Atlas (compared with tide gauge observations) is confined to the nearly-enclosed Ariake and Seto Inland seas: this is only the case for FES2014b. As shown in Table 2, the NAO99Jb M_2 RMS error with respect to tide gauges for the eastern China sub-area is 11.7 cm, whereas for GOT4.10c and TPXO9-Atlas the error is much larger at 30.1 cm and 34.5 cm respectively. Then for the Ryukyu Islands sub-area, TPXO9-Atlas has an RMS error of 11.0 cm compared with 2.4 cm for NAO99Jb, 3.1 cm for FES2014b and 3.9 cm for GOT4.10c. For the open sea areas where there are no tide gauges, Figure 2 suggests that all of FES2014b, GOT4.10c, TPXO9-Atlas and NAO99Jb agree to within 1-2 cm inter-model station deviation and so the choice of model here is immaterial.

To evaluate how important the regional improvements of NAO99Jb are in computing the load, we computed six sets of M_2 vertical OTL displacement for all 102 GPS sites, using two runs for each of the three most recent global models:

1. FES2014b:
   a. FES2014b augmented with all of NAO99Jb
   b. FES2014b augmented with NAO99Jb only for the Ariake and Seto Inland seas
2. GOT4.10c:
   a. GOT4.10c augmented with all of NAO99Jb
   b. GOT4.10c augmented with NAO99Jb only for the Ariake and Seto Inland seas
3. TPXO9-Atlas:
   a. TPXO9-Atlas augmented with all of NAO99Jb
   b. TPXO9-Atlas augmented with NAO99Jb only for the Ariake and Seto Inland seas

The PREM elastic Green's function was used and the minimum, maximum and RMS of the model M_2 vertical phasor residuals with respect to the GPS observations are shown in Table R3.1 for each ocean tide model combination. It can be seen that for all three models, the residuals when using all of NAO99Jb are smaller than when using it only for the Ariake and Seto Inland seas. The differences obtained among the global models when augmented with all of NAO99Jb are indistinguishable, which confirms that our use of FES2014b (which was based on the accuracy tests with respect to the available tide gauges in the East China Sea region) for the Green's function comparisons is valid. This indistinguishability is also consistent with the small contributions and close model agreements for the areas outside the NAO99Jb extents shown in Figure R3.2. Whilst TPXO9-Atlas has a slightly lower RMS for the NAO99Jb augmentation with only the Ariake and Seto Inland seas compared with FES2014b and GOT4.10c, this is likely a result of improvements in areas with no tide gauges, but using all of NAO99Jb still gives lower RMS and maximum residual values.

Table R3.1 Phasor differences (in mm) between M_2 vertical OTL displacement and GPS observations at the 102 GPS sites using three different global ocean tide models and different augmentations of the regional NAO99Jb model, all with the PREM elastic Green's function

Model	(a) Use of all NAO99Jb			(b) NAO99Jb for Ariake and Seto Inland seas only		
	Min	Max	RMS	Min	Max	RMS
1. FES2014b	0.08	1.59	0.53	0.06	1.80	0.69
2. GOT4.10c	0.06	1.70	0.53	0.01	1.82	0.68
3. TPXO9-Atlas	0.09	1.64	0.50	0.02	1.76	0.58

4. Another major problem is that the conclusion about determining Earth structure seems inadequately supported by the evidence. Table 4 shows that once we adjust for anelastic attenuation, PREM gives RMS values that are basically indistinguishable from those for the regional model (which the authors more or less admit). Changing the model can reduce the RMS a bit more, but there is no demonstration that the reduction is significant given the added degrees of freedom: certainly the conclusion about asthenosphere depth (p. 13 lines 18-19) is not at all warranted.

The paper has first (to end of Section 4) demonstrated that the systematic M_2 residuals of about 1.3 mm amplitude arise from deficiencies in the elastic PREM Earth model. Our original intention with Section 5 was to explain these deficiencies by first obtaining the optimal model for the region. However, as the reviewer points out (and we had already noted), differences between anelastic global PREM, anelastic regional S362ANI and our modified regional S362ANI are relatively small. We have rewritten and reordered Section 5 to reflect this, whereby we first consider if the original elastic S362ANI regional model results in reductions in the residuals over elastic global PREM, and the effect is small (RMS for the whole region and the Ryukyu Islands alone both reduced by 0.08 mm). We then describe how accounting for anelasticity at the M_2 frequency for both PREM (globally) and S362ANI (in the ECS region) reduces the residuals (S362ANI_M2 results in slightly smaller RMS values than PREM_M2), but still residuals at the ~0.7 mm level for the Ryukyu Islands remain. Therefore we can only test optimality of the Green's function by computing a range of Green's functions based on different asthenosphere depths and values of Q. However, we did not find any significant improvement over S362ANI_M2 on doing this, and have therefore taken care to avoid any suggestion or claim that our observations *require* any change in the asthenosphere depth from S362ANI for this region. Instead, we have described our search tests to ascertain optimality, but that they result in similar reductions as S362ANI_M2 for the ECS region, and also how, for this region, using the global PREM_M2 model leads to residuals of almost comparable size. As well as rewriting Section 5, we have modified both the Abstract and the Conclusions to reflect this. For Figure 5, we now show the S362ANI_M2 residuals, not those of mod_S362ANI_M2.

5. I have grave doubts about this method of finding errors in the loading computation. It depends, as the authors note, on the terms in the sum being uncorrelated, and that they certainly are not. So I am dubious about all subsequent invocations of errors in the loads.

In this same vein, Figure 3 shows standard deviations much larger than the RMS values of the loads from different models: this suggests that the computed errors are much too large.

The method presented is intended, to first order, to enable a further indication to be obtained that the large, systematic 1.3 mm M_2 residuals seen across the Ryukyu Islands and also on Kyushu are not caused by errors in the most accurate NAO99Jb model (as ascertained in Section 2). The inter-model standard deviations shown in Figure 3a for all nine ocean tide models are only around 0.3 mm across the Ryukyu Islands, suggesting that ocean tide model error contributions do not explain the discrepancies. However, around parts of Kyushu, namely the Seto Inland Sea and the Ariake Sea, the standard deviations increase to 2.5 mm, but in these regions we showed in Section 2 that the global models are erroneous. In terms of evaluating the contributions of the errors in each of the sub-area polygons to the total error at a specific location, in practice correlations are likely to exist among the polygons. For neighbouring regions, these correlations are likely to be positive, so equation 4 provides a very conservative (upper) bound on the expected level of model error. We have modified the text of Section 3 to emphasise this. For the Ryukyu Islands, the errors derived from equation 4 are about 0.5 mm, still much smaller than the 1.3 mm GPS observational residuals, and similarly across Kyushu, they are around 0.3 mm, indicating that the 2.5 mm inter-model standard deviations shown in Figure 3a are caused by models other than NAO99Jb. The errors within our defined eastern China sub-area polygon are around 1-2 mm and larger than the inter-model standard deviations, but this is because we have defined our eastern China polygon as considerably larger than the discrepant inter-model areas very close to the coast for both the ocean tides (Figure 2a) and OTL displacements (Figure 3a). Hence a conservative NAO99Jb error of 11.7 cm across all of the eastern China polygon has been applied, but which has been computed using eight tide gauges within the polygon, most of which are within the near coastal zone of much larger inter-model agreement than arises for much of the eastern China polygon. As well as mentioning in Section 3 that these large errors arise because of the fairly large 11.7 cm RMS error for NAO99Jb used, we have added an explanation that this is likely very conservative and results in errors which are too large for much of the area.

6. I hope the final version of the paper will include a supplement with text files giving the authors' M2 estimates (GPS and tide gauges) as well as the Green functions.

The GPS-estimated total vertical M_2 amplitudes and Greenwich phase lags and M_2 tide gauge estimates are included as a text file supplement. The Green's functions have also been moved from the appendices to this text file.