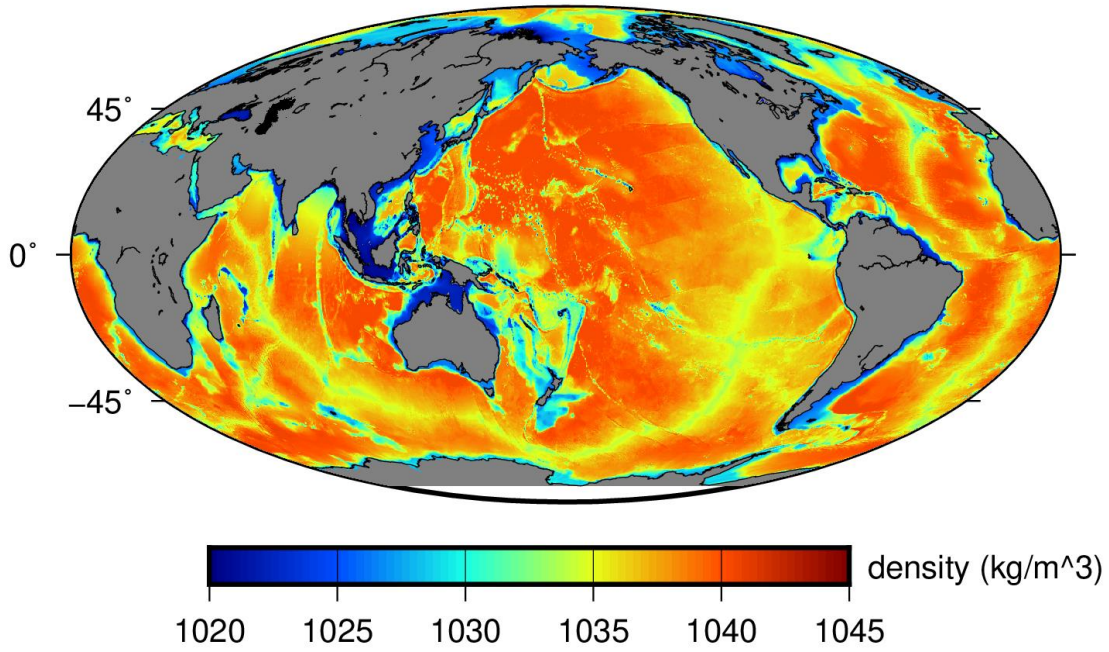
Thank you for the review and comments. We provide below our responses to the points raised, including details on the modifications made to the manuscript.

1. I want to bring up a possible systematic error which, if not addressed here, needs to be addressed at some point by this group as they continue to do these kinds of studies. I think it could cause errors of a few percent, which could be significant. Specifically, I'm beginning to think these very precise geodetic applications need to account for the variable density of seawater – that rho in Eqns (1-2) should stay inside the integral sign.

Thank you for pointing this out. The rho (density of seawater) stays inside the integral sign in Eqns (1-2), and the latest version of SPOTL accounts for the same variable density of surface seawater as before. In order to examine the effect of the variable density of the whole column seawater, we have computed the mean seawater density for each column on a 0.25x0.25° grid using the World Ocean Atlas (Boyer et al., 2013). The result is shown in the figure below.



We have verified that this corresponds closely to Figure 6a of Ray et al. (2013) if $1/(g\rho_{mean})$ is plotted. For the loading computations we needed to extend some grid cells towards the coast to cover all water areas. The CARGA program was used (Bos, M S., and Baker, T. F.: An estimate of errors in gravity ocean tide loading computations, J. Geod., 79, 50-63, https://doi.org/10.1007/s00190-005-0442-5, 2005) to perform some sensitivity tests. First, a constant sea water density of 1030 kg m^{-3} was used (label=1030 in Table R1.1). Second, the spatially-varying mean sea water density was used in the OTL computations (label=spatial). Finally, we corrected the mean sea water density also for compressibility using the formula given in Ray (2013), with label 'compr'. Some statistics for the changes arising in M2 height displacement amplitude and phase-lag for the 102 GPS station locations in this study are presented in Table R1.1. The mean difference for the amplitude is around 0.11 mm if compressibility is also taken into account, and this is smaller than the 0.2-0.3 mm uncertainty of the GPS observations. The maximum difference reaches 0.37 mm, which is starting to be noticeable and in future investigations this should be treated more carefully. Nevertheless, this is still much smaller than the observed discrepancies of over 1.5 mm and we have added a paragraph to the end of

Section 5 which summarises these computations and comparisons.

Table R1.1 Influence of spatially-varying seawater density on M_2 vertical OTL displacement at the 102 GPS sites

	ΔAmplitude (mm)			ΔPhase-lag (°)		
Comparison	Min	Mean	Max	min	mean	max
spatial-1030	-0.13	0.03	0.16	-0.59	-0.22	0.38
compr-1030	-0.17	0.10	0.37	-0.99	-0.39	0.65

2. Abstract, line 13, recommend inserting "the regional model" before NAO99Jb, because many readers, even in the tide community, may see NAO99 and think it refers to Matsumoto's global altimeter-based model.

We have inserted "the regional model" as suggested.

3. Next line: "the most accurate." Well, this is risky because who knows if someone has developed another regional model here. I'd say "an accurate" – but it's up to the authors.

We have changed the part of the Abstract to now state, "By comparison with tide gauge observations, we establish that from nine selected ocean tide models (DTU10, EOT11a, FES2014b, GOT4.10c, HAMTIDE11a, NAO99b, NAO99Jb, OSU12, TPXO9-Atlas), the regional model NAO99Jb is the most accurate in this region".

4. In Abstract, and also page 15, bottom, authors quote 1.5 mm and 0.8 mm. I don't understand this. From Table 4, this looks to be "comparing apples and oranges." One is a maximum error and the other is RMS. Seems misleading, unless I'm just not following where they get these numbers.

0.8 mm was not in fact the RMS but the typical maximum error arising on applying the anelastic Green's functions, which we agree should have been made clear and used statistics consistent with Table 4. We have modified Table 4 to now incorporate minimum, maximum, 90th percentile and RMS values, and we now refer to these explicitly in the Abstract, the discussion in Section 5 and in Section 6 (Conclusions).

5. Page 13, line 5: Dahlen & Tromp is a massive book. I and many readers would appreciate your quoting the page number or even the Eqn number you're using when you cite this book.

We have inserted the equation number (9.66) to the text where Dahlen and Tromp (1998) is cited.

6. Page 2, line 36, where GPS is assigned an error of 0.3 mm. I don't accept this, because surely the errors in GPS are dependent on the length of the time series.

Penna et al. (2015) showed that with 2.5 years of GPS data, a semi-diurnal harmonic displacement could be estimated with an accuracy of around 0.2-0.4 mm. This led to our choice of 0.3 mm for the GPS observational error. To clarify, we have amended the text to state "the GPS observational error is assigned a STD of 0.3 mm following Penna et al. (2015), and which assumes that at least 2.5 years of continuous GPS data will be available".