(1) Line 16-17: “Our results do not provide any conclusive evidence for upwelling mantle plume rooted in the CMB beneath the Emeishan LIP” - given that the model region extends down to 800/900 km depth and the mantle layer is ~3000 km thick, there is no prospect of conclusive evidence for a plume rooted in the CMB, irrespective of what the results show.

This have been rewritten, please see below:
Our results do not provide any conclusive evidence that upwelling mantle from the lower mantle which exclude upwelling mantle plume rooted in the core-mantle boundary (CMB) beneath the Emeishan LIP.

(2) In general, the written English is ok, but there are a number of places where the sentences don’t quite make sense (e.g. Lines 62-65).

This have been revised, please see below:
Previous studies suggested the Emeishan flood basalts generated by mantle plume impingement at the base of the lithosphere and caused large-scale regional up-doming prior to volcanism (Shellnutt et al., 2012; Shellnutt, 2013; Li et al., 2002; Gao et al., 1999; Liu et al., 2008) and led to a short eruption of less than 1 Ma (Song et al. 2004).

(3) Lines 74-76: If the LIP formed a quarter of a billion years ago, present day mantle dynamics are unlikely to be able to help shed light on its origin.

The tectonic framework of Emeishan LIP is characterized by the Longmenshan thrust fault in the northwest and the Ailaoshan-Red River strike slip fault in the southwest. It is possible that the assembly of Yangtze block with another crustal block in the Late Permian and Early Triassic, which is the largest tectonic event in the Emeishan area, might have led to crustal thickening and large-scale delamination of the lower crust and (or) mantle lithosphere. The delamination resulted in the mantle upwelling and generation of crustal melts that triggered plume-like upwelling, eventually, resulting in the Emeishan LIP formation.

Generally, high and low velocity relics generated by subduction slab or crustal and mantle lithospheric delamination and upwelling mantle in the asthenospheric mantle can be retained for over tens of millions of years (Cook et al. 1999; Balling 2000; Svenningsen et al., 2007; Zhai et al., 2007; He et al., 2015). These low and high velocity structure can be image by tomography (D. Zhao et al., 1992, 1994; L. Zhao et al., 2016).

(4) Line 91: “velocity structure or mantle dynamics” - if you apply traveltime tomography, then there is no choice but to recover velocity structure. You can only make inferences about the mantle dynamics from these results.

Thank you, this have been revised, please see below:
The target of this study is with a view to construct the velocity structure and investigate mantle dynamics of the upper mantle beneath the Emeishan area, and further discuss the Emeishan LIP formation.

(5) Lines 102-103: What 1D velocity model is used, and what is the “fast raytracing technique”? I know that references are provided, but a one sentence summary would be useful.

Thank you, it is revised, please below:
The location of the seismic ray crosses through the boundary of the study region was determined by a 1D (or 1D IASP91) velocity model.

(6) Lines 104-107: Is this a regular grid in spherical coordinates? And linear interpolation is not possible in 3-D when the velocity field is a function of 8 surrounding control points (in general, gradV will not be constant inside a cell). I think most people refer to it as pseudo-linear interpolation. This parameterization is fairly standard, so why are so many papers by Zhao cited in lieu of a definition?

Thank you, it has been revised:

Theoretical travel time and seismic ray paths are obtained by the fast raytracing technique (or pseudobending technique) (Um and Thurber, 1987; Zhao et al., 1992). 3D grids are employed to express the velocity perturbation values, and any point in the model space can be calculated from values of the surrounding eight nodes by pseudo-linear interpolation (Zhao et al., 1992, 1994).

(7) Lines 111-112, and Figure 2 caption: It is stated that events between 30-85 degrees angular distance are used, but the plot shows events out to ~100 degrees. Azimuthal coverage is very varied too, with most events from the south and east. In order to improve coverage from other quadrants, is there any prospect of employing other global phases, such as PP, SKP, Pdiff etc?

Epicenter distance range from 30 degree to 85 degree for each station-event pair rather than for the center of study region-event pair. In this study, we only used P-travel-time arrival, based on our assessment and comparing with other similar studies in this area, our results should be accepted.

(8) Line 115: Does capping the maximum residual mean that larger residuals are due to noise or an inability to model signal? In practice, this difference is perhaps not worth dwelling upon, but I always find it interesting when influential data (large residuals demand more significant model perturbations) are ignored.

Generally, we suggest the larger residual are due to noise or an inability signal.

(9) Line 118: But at least in Figure 6 (for example), the model seems to extend in depth to 900 km, not 800 km.

Thank you, it is revised.

(10) Line 119: The contribution of the crust to the pattern of teleseismic arrival time residuals can be quite significant, particularly in regions with large changes in elevation. More information about how this correction was made and which model was used would be useful. Looking at the results (e.g. Figure 9), shallow high velocity zones tend to be associated with regions of low elevation, where one would expect thinner crust, and hence a negative contribution to the arrival time residual compared to where there is thicker crust. Hence, if the crustal correction doesn't adequately take this into account, it may result in artefacts in the upper mantle velocity structure.

In this version, we further explain the crustal correction, please see below:

In teleseismic tomography, rays do not crisscross well in the crust and the uppermost mantle beneath the study region. Therefore, the effect of crustal heterogeneity need to be removed through correcting the relative travel-time residuals, which is called crustal correction (Zhao et al., 2006; Jiang et al., 2009, 2015). In this work, the CRUST1.0 model (Laske et al., 2012) is used to make the crustal correction to the relative travel-time residuals following the scheme of Jiang et al. (2015). Here,
we are calculating the crustal correction for the upper 50 km of the earth.

(11) Lines 122-126: Why only quantify (and illustrate) the damping vs. data fit trade-off if smoothing is also applied? Also, is the tomography iterative non-linear, or just linear (rays only traced through reference model)?

In this code, teh smoothing parameter is fixed, so we only quantity the damping vs data fit trade-off.
In this study, the LSQR algorithm (Paige and Saunders, 1982) was used to solve the large and sparse system of observation equations with damping and smoothing regularizations (Zhao, 2004).

(12) Paragraph starting line 128: "...tracing the actual rays through a synthetic structure..." What is meant by the "actual rays"? Are these the ray geometries from the initial or final model of the observational study? In other words, is the checkerboard test a purely linear inverse problem? At the risk of blowing my own trumpet, I suggest reading the paper:

N. Rawlinson, W. Spakman; On the use of sensitivity tests in seismic tomography. Geophys J Int 2016; 205 (2): 1221-1243. doi: 10.1093/gji/ggw084 which outlines some of the pitfalls of using a synthetic recovery test with a relatively tight pattern of uniformly-sized anomalies. I think in this case the quality of the recovery overstates what can actually be achieved with the real data. Apart from data noise, which I doubt is accounted for here, this kind of test is strongly preconditioned to produce favourable results. For instance, there is very little evidence of smearing, yet the use of teleseismic body waves generally results in some near-vertical smearing. Finally, lines 135-136 simply repeat lines 117-118 for no good reason.

The repeat lines have been removed.

Thank you, this part have been improved, please see below:

For evaluating the resolution of the 3-D velocity structure, we carried out checkerboard resolution tests (CRTs) (Zhao et al., 1994; Zhao, 2001; Rawlinson and Spakman, 2016) and assigned positive and negative velocity perturbations of ±5% to all the 3D grid nodes. Synthetic travel times are calculated for the checkerboard model. The locations for stations/events are the same in the synthetics as in the real data. We then inverted the synthetic data with the same algorithm as that for the real data. Although the CRTs have a number of potential drawbacks (Rawlinson and Spakman, 2016), however, it bascally reflects the resolution of the tomography and become method of rotine checking. (13) Lines 154-155: Given that this is a teleseismic tomography study based on relative arrival time residuals, velocity perturbations can only be discussed in a relative rather than absolute sense, unless constraints from elsewhere are applied.

(14) Line 157-159: The recent study by Huang et al (2015) is frequently referred to, and indeed one lingering question is what the new teleseismic tomography study brings to the table that the 2015 study does not. A quick comparison of the arrays used show that Huang has a much denser station network, but in this study the array extends further north and east. The implications of these differences should be discussed somewhere. Also, Huang jointly inverts a large database of local earthquakes along with the teleseisms in order to constrain crustal structure, which ostensibly is an advance on what is done in the current study. Therefore, some discussion and perhaps justification of the current study relative to those that precede it is probably warranted.

Thank you, in this version, we have improved these, please see below:

Huang et al. (2015) carried out a tomographic study using 411 temporary stations within 20-33° N and 95-110° E and obtained the velocity structure of the crust and upper mantle in Chuandian area. Here, we carry out an extended study in the region northward and eastward within 20-35° N and 97-111° E so as to cover all the regions of the

Emeishan LIP. Although we used almost same amount of data as that of Huang et al. (2015), our teleseismic data are of higher quality and were recorded by 228 permanent stations. Our target is to construct the velocity structure and investigate the mantle dynamics beneath the Emeishan area, based on which we evaluate the geodynamics of Emeishan LIP formation.

(15) Line 197: "...with crust and (or) lithospheric delamination." - It would be more correct to say "mantle lithosphere" rather than just "lithosphere".

Thank you, this has been revised.

(16) Line 225: How does a receiver function study demonstrate convective circulation in the mantle? A few more details of this study would be useful to include.

Thank you, we have improved this part, please see below:
A receiver function study revealed a felsic lower crust in the Emeishan area, suggesting crustal delamination (He et al., 2014). Simultaneously, the MTZ beneath this region shows a cold domain (He et al., 2014), which might suggest that the delaminated cold material dropped down into the upper MTZ. Generally, crustal delamination can induce mantle upwelling (Elkins-Tanton and Hager, 2000; Elkins-Tanton, 2005), which might have eventually resulted in a convective circulation system between the lower crust and the MTZ beneath the Emeishan area (He et al., 2014). This further confirms the present location of the Emeishan LIP.

(17) Line 248: Following from the above, how do the receiver functions identify felsic lower crust?

The Vp/Vs ratio can be obtained by H-K stacking of receiver function technique, generally, the crustal delamination resulted in felsic lower crust with low Vp/Vs ratio (<1.75).

(18) Lines 260-264: Getting back to an earlier point I made, I would be very careful about associating what you see at 500-600 km depth with a LIP that formed a quarter of a billion years ago. I'm not saying that there cannot be an association, but it would be good to see some independent evidence e.g. from geodynamic modelling. If the delaminated lithosphere is distributed as inferred by this study, how long would it take to go from initial separation to this state, and given plate motion over time, is it located where you would expect? And when might the volcanism occur? While undertaking modelling of this type is clearly beyond the scope of this study, I still think that the readers of the journal will need a bit of convincing that what is proposed is actually possible.

1. Previous receiver function have demonstrated the Emeishan LIP formation may be associated with the delamination, since then, there aren't any evidence demonstrated

the location of the Emeishan LIP change (He et al., 2014).

2. The vestige of the delamination (High velocity perturbation) and mantle upwelling (low velocity perturbation) can be retained over tens of millions years. (Cook et al. 1999; Balling 2000; Svenningsen et al., 2007; Zhai et al., 2007; He et al., 2015). These low and high velocity structure can be image by tomography (D. Zhao et al., 1992, 1994; L. Zhao et al., 2016).

(19) The Conclusions are rather brief, and would benefit from being fleshed out a bit.you

Thank you, this have been revised, please see below.

The tectonic framework of Emeishan LIP is characterized by the Longmenshan thrust fault in the northwest and the Ailaoshan-Red River strike slip fault in the southwest. It is possible that the assembly of Yangtze block with another crustal block in the Late Permian and Early Triassic, which is the largest tectonic event in the Emeishan area, leading to crustal thickening and large-scale delamination of the lower crust and (or) mantle lithosphere. The delamination resulted in the mantle upwelling and generation of crustal melts that triggered plume-like upwelling, eventually, resulting in the Emeishan LIP formation.

Our results show that there are no low velocity perturbation rooted in the lower mantle beneath the Emeishan LIP, suggesting the absence of any vestiges of a mantle plume rising from the CMB beneath the Emeishan area.

Our results also further confirm the model of vetical convective circulation in the mantle as the major trigger for the Emeishan LIP.