The authors added a new soil N_2O emissions module to the dynamic global land model LM3V-N, and tested its sensitivity to soil moisture regime, as well as its responses to elevated CO_2 and temperature. However, I am not sure what the main objective of the paper is – whether this was mainly a model development paper or whether they wanted to conduct different sensitivity analyses. As noted by the other two reviewers, I think this paper needs major revisions before it can be published. In my opinion, the most important is to: 1) include more analyses instead of the speculations presented in its current status; and 2) highlight the original contributions in this paper, specifically illustrating what is different from what already has been published in Xu-Ri et al., 2012.

I first list some major concerns, followed by minor comments.

Major points:

The authors argue in the abstract ln. 7-9 on p. 3102 that "[t]he model was capable of reproducing the average of cross-site observed annual mean emissions, although differences remained across individual sites if stand-level measurements were representative of gridcell emissions." It is not obvious how they concluded that the model was indeed capable of reproducing the observed emissions. From the Figure 3, it is also not clear if the model is capable or not. I would expect to see more rigorous model-obs comparisons, if this is a model development/validation paper. As the second reviewer suggested, I would also like to see hourly/monthly comparisons at multiple sites, and I find it odd that the model is run at "an annual time step" as they state on ln. 8-9, p. 3113. Why don't they get the annual average from their half-hourly simulation? Also, I recommend that they at least add their modeled values in Table B1 as well, so that the reader can directly compare their modeled values to the observations.

With regard to soil moisture, why does Figure 3 use different methods for the different data sets? I understand that there are three methods that the authors used for each of the three different data sets but it does not make much sense to do a model-obs comparison in a panel, using method 3 for part a and method 2 for parts b and c. Why not use one of the methods for all parts? If the authors agree that soil moisture values larger than 0.6 are not reasonable, what about the validity of the maximum water method that leads to a global mean WFPS higher than 0.6 (Figure 4)?

I also found that there are some statements in this paper that should be better justified. First, on ln. 5, p. 3115 authors state that "[t]he negative impacts (reduced N_2O flux), which are also reported from manipulative experiments, are likely from increased plant N and immobilization demand under CO_2 fertilization, reducing N availability for nitrifiers and denitrifiers" but is this what they see in the model? I believe they can also draw a similar graph, illustrating plant N and immobilization rate in time-series to see if this is indeed the response they are seeing in the model. The same goes for the positive impacts. I think it is important to see if the litter production and soil moisture have been increased, as well as stomatal conductance and leaf transpiration reduction, as they imply in the paper.

I'm not sure I understand the reasoning behind the statement on ln. 14-16, p. 3118: "Patterns of seasonality, and the correlates between N_2O emissions vs. temperature and soil moisture suggest that moisture is the dominant driver of N_2O emission in tropical regions and soil temperature critical elsewhere." What does "dominant" mean in this case? I think that in order to make such a statement, one needs to show the impact of different variables that are important and how that affects their N_2O emissions.

The authors write on p. 3104 that "[s]imulations with LPJ-DyN and O–CN demonstrated a positive response of N_2O emissions to historical warming and a negative response to historical CO_2 increase, globally. This negative CO_2 response seems to be in disagreement with one meta-analysis of manipulative field experiments showing an increase in N_2O emissions at elevated levels of CO_2 (Zaehle et al., 2011; Xu et al., 2012; van Groenigen et al., 2011). The discrepancy in response to global change factors needs to be addressed both in models and in the interpretation of manipulative field experiments." It seems that authors are misinterpreting the work of Xu-Ri et al. (2012) (which authors write as Xu et al, (2012)). Xu-Ri et al. (2012) states that "[i]ncreasing CO_2 generally enhanced the N_2O emission in tropical and temperate moist forests, whilst reducing the N_2O emission in some other regions (Fig. 6)," which is essentially the same as the argument made in the current paper. I think it would be helpful if the authors could clarify what it is that they are arguing that is different from the conclusions in the Xu-Ri et al. (2012), as this was not obvious to me.

Minor comments:

I am a bit confused about the Figure 1. The MEI values (http://www.esrl.noaa.gov/psd/enso/mei/table.html) are higher than 0.6 on several occasions between 1975 and 1980 (1976 Jun-Oct, 1977 Jun-1978 Mar, 1979 Jul-1980 Jul) and yet, this figure is only showing a one gray zone during that period. Also, it is unclear which WFPS method was used for this calculation. It would be helpful if they showed the range in interannual emissions, based on the 3 different methods and datasets they used. The same goes for Figure 2.

Are $RNOx:N_2O$ and $RN_2:N_2O$ values calculated at every time step for every grid cell? Or how does it work?

L. 4, P. 3102 – typo "reponses"
L. 19, p. 3109 – typo "equalibrium"
l. 1, p. 3117 – typo "exsit"
l. 5, p. 3117 – "constraint" to "constrain"
l.29, p. 3117 – typo "oringinal"
l. 8, p. 3119 – typo "aboitic"
l. 13, p. 3119 – typo "unstand"
l. 9, p. 3120, typo – "speicies"