Anonymous Referee #1

We would like to thank the reviewer for the comments and valuable suggestions.

Major concerns:

1. The goal of the paper is stated at P9100 L27: "The focus of this paper is to establish to what degree the different datasets reproduce several well-established isotope signals". Many of these isotope signals have been seen in previous satellite and model studies, some referenced, but consideration of these previous studies is uneven or absent. Beyond the improvements in TES_{V5}, it is therefore not clear what advance is being made by this study in terms of using these signals as benchmarks.

Our impression from the existing literature is that a systematic evaluation of the available data sets is missing. We investigate three data sets benchmarking them at least qualitatively and comparing the pros and cons of different remote sensing systems. The transition between use of the TES V4 and V5 data in scientific studies occurred without detailed investigation of the differences and improvements between V4 and V5. Here we present the implications of the improved sensitivity of TES V5 data in terms of the isotope effects. For example, we report that the isotope amount effect in the tropics is well documented in the TES V5 dataset and not in the V4 dataset. This is relevant since a number of earlier studies used V4 TES data in the tropics, where we could show that considerable improvements were made in the V5 dataset. In addition, we demonstrate the issue of humidity biases in the models and discuss methods to avoid these biases.

2. In addition, for a largely methodological paper, it is unclear how some significant methodological issues were addressed, namely that a single model year (2001) is being compared to satellite data for different periods (2003-2005 for SCIAMACHY, 2006 for TES). This is not sufficiently discussed as a possible source of difference between the satellite and model datasets.

We added a discussion about this possible source of difference in section 4.1. However, the isotope effects discussed here are dominating each year's isotope pattern in a very similar way. In general, they do not depend strongly on meteorological interannual variability. The issue of using mean "climatological" AKs instead of AK from the specific date is discussed in the following.

3. Related to this, a detailed description is given in section 2.4.2 of the rationale for and approach to averaging kernel application. At P9110 some subtle issues relating to model-data comparisons and humidity biases are examined, which are important and I suspect not recognized by those making model-data comparisons. But it is not clear how:

- The averaging kernels were selected for each model profile in the first place, knowing that their vertical structure can be different under different meteorological conditions (e.g. described by Lee et al. (2011)) especially considering that the model simulation year was outside of the TES observation period.

We use monthly averaged AKs interpolated on the horizontal grid of the model while in the vertical, the model was interpolated on the TES vertical layers (P9104L16). We mention the non-overlapping time periods of model simulations and satellite observations as a possible source of error in chapter 4.1.

- Whether the model was sampled at points coincident with the orbital paths of the satellites. Werner, Langebroek et al. (2011, JGR), for example, weighted their mean model δD according to the regionally-varying SCIAMACHY sampling frequency. Risi et al. (2012) sample the model output at points coincident with the satellite measurements.

We couldn't sample the model output exactly synchronously with satellite measurements since model results were not available from the specific dates. We therefore use both the monthly model and the monthly TES V5 LITE datasets. This LITE dataset is in any case a monthly data product (P9102L1519, TES technical document). Thus, the HDO/H_2O data from TES V5 is itself an averaged value of many profiles taken at numerous orbital positions of the satellite. Furthermore, we interpolate the TES data on the horizontal grid of the model. We added this information in chapter 2.4.

- How TES DOF and quality filtering was accounted for in treating the model data.

The TES DOF and SRQ filtering were carried out during the interpolation of TES data into horizontal model grid. Thus, we only use the TES data which have DOF > 0.5 and SRQ =1 for both version 4 and version 5.

4. Furthermore, Werner et al. 2011 compared the newer ECHAM5 isotopic GCM to SCIAMACHY, but this is not discussed in the paper. Has ECHAM5 agreement with SCIAMACHY changed?

The difference between ECHAM5 and SCIAMACHY is ~20 ‰ in the tropics and ~25 ‰ at the higher latitudes (60 N-S, Werner et al. 2011). From our results, the differences between ECHAM4 and SCIAMACHY are ~10 ‰ and ~20 ‰ in the tropics and at the higher latitudes (60 N-S), respectively (Fig. 2). We added information on ECHAM5 results in our discussion in section 4.1.

5. Lastly, in Section 4.3, the data and model output are examined according to how well they follow a simple Rayleigh distillation model. But would anyone currently argue that it could? Worden et al. (2007), Brown et al. (2008), Risi et al. (2012), etc. discuss non Rayleigh-like behavior in satellite measurements and models. I appreciate that the Rayleigh distillation model is a useful conceptual framework, but the point of, and advance made by this discussion, needs to be made more clear in the context of previous studies that consider Rayleigh vs. non-Rayleigh like distributions.

Point taken. The Rayleigh distillation model is for us rather a benchmark to evaluate to what degree the atmosphere in various regions is affected by mixing or cloud processes, such as ice crystal lifting, which disturb the simple distillation behavior. One of the interesting results is that the Rayleigh model fits better when signals in the vertical coordinate are compared and does not fit well when large-scale transport is considered. We clarified this approach in chapter 4.3.

Specific comments:

P9097 L4: omit 'only'

P9097 L13: change kernel to kernels

P9097 L15: change extent to extents

P9097 L25: separate these references across the different applications mentioned.

P9098 L18: change has to have

P9098 L19: remove 'the' from 'the ground based'

P9099 L14: 'Adequate tools' – this is an awkward way to describe the models

We will change the sentences accordingly.

P9099 L28: in this paragraph, rather than just listing these comparison studies, summarize the most important points across them. Werner et al. 2011 compared the newer ECHAM5 isotopic GCM to SCIAMACHY, but is absent from this list. How do ECHAM4 and ECHAM5 each compare to SCIAMACHY and what are possible sources of any differences?

The comparisons between ECHAM4 to SCIAMACHY and ECHAM5 to SCIAMACHY yield similar results (see above). SCIAMACHY results look very similar to Werner et al. 2011. We added this information into the text.

P9099 L29: change retrieve to retrieved.

Has been changed

P9102 L20: Worden et al. (2012) used a DOF threshold of 1.0 for the two TES versions. Please explain why this threshold was lowered and what effect it potentially had on the results.

The reason using the DOF>0.5 for both TES versions is to have more data available at higher latitudes e.g. from mid latitudes ~60N-S to Polar Regions. Our impression is that with such a modified threshold at least TESV5 version gives more valuable information at mid-latitudes.

P9105 L15: At P9102 L20, the TES data filtering procedures are described. How is this accounted for before applying the averaging kernels? i.e. is the model data filtered in any way to account for possible systematic biases introduced by the TES data filtering?

The TES data filtering for DOF>0.5 and SRQ=1 was carried out before we apply the AK to the model outputs. With this filtering process, we select all TES data (AK, H_2O and HDO), which have passed the filtering criteria. Thus we used these data for AK application. We do not filter the model data based on DOF and SRQ criteria; however, the bias resulting from the AK application is corrected using eq. 8 for $ECHAM_{AK4}$ by 3%.

P9107 L7: Section 3 needs to include a summary of the extent to which the isotope signals have or have not been seen in previous satellite-based studies (eg. Werner et al., (2011), Risi et al. (2012)) and what advance you are trying to make with this analysis.

We see (section 3) a number of interesting improvements of TES V5 related to the isotope effects: a more realistic latitudinal effect especially at higher latitudes, a clear continental effect, a sharp amount effect above the West Pacific ocean and the seasonally varying isotope amount effect due to the movement of ITCZ (Fig.4&5). The latter is a feature of the water isotopologues that has not been documented in previous studies using satellite remote sensing instruments. We added a paragraph describing the progress made here compared to former studies.

P9109 L11: These latitudinal gradients need to be compared to those presented for Werner et al. (2011) for ECHAM5 and SCIAMACHY and Risi et al. (2012) for LMDz, TESv4 and SCIAMACHY. The differences for TES v5 are interesting, but it is not clear what we are learning from them compared to these previous studies.

It was not intended here to make a full comparison between all available isotope models and all available satellite data sets. However we agree that this is a worthwhile comparison to make. The data from Risi et al (2012) were given at a different vertical resolution, which is why a straightforward comparison was not possible. TES V4 measures δD around -180 ‰ at latitude 60N while SCIAMACHY measures δD around -220 (also in Werner et al 2011). This -40 ‰ difference between TES_{V4} and SCIAMACHY is not seen in TES_{V5}, which measures δD around -220 ‰ similar to SCIAMACHY (Fig.2). The difference between V4 and V5 at higher latitudes is due to the fact that V4 is more strongly influenced by the a priory field than V5. The same is true when the models are evaluated with the respective Aks. This is one of the improvements of the TES_{V5} dataset that has not been illustrated in the previous studies.

P9110 L4: are Risi et al., 2012b, 2013 discussing ECHAM4 (as part of a model comparison), or another model?

Risi et al. 2012b used ECHAM4 as part of a model comparison while Risi et al. 2013 used only LMDZ.

L9110 L24: Please compare these results for the tropical continents to those from Werner et al. (2011) for ECHAM5 and SCIAMACHY , Risi et al. (2012a) for SCIAMACHY and TES v4, and, especially, Brown, Worden and Noone (2008, JGR) for TES v4.

We added a discussion on a comparison between our and their results.

P9111 L24: As far as TES v4 not picking up the depleted δD over the Pacific Warm Pool, and for any differences between v5 and v4 generally, how much is due to the different vertical regions over which the data are averaged, rather than v5 retrieval improvements? I understand that v5 is more sensitive over a wider vertical range, but this does not preclude trimming that vertical range to that of v4 for the purposes of a more meaningful comparison.

The amount of δD difference between V4 and V5 at West Pacific is up to 22 ‰. We added this point in the text.

P9111 L26: please discuss how the zonal mean δD compare to Risi et al., 2012a whose comparison included TES v4 and SCIAMACHY, and Werner et al. (2011) whose comparison included SCIAMACHY and ECHAM5.

The zonal mean δD from our study is in a good agreement with Werner et al. 2011. Both ECHAM4 and ECHAM5 show similar results with δD values at 60N around -200 ‰ and for SCIAMACHY around -220 ‰.

P9112 L11: change greatly increased in to increased .

P9112 L17: change datasets to data points.

P9112 L21: change significantly to significant.

We will change the sentences accordingly.

P9113 L1: change section name to 'Seasonal isotope distribution'. Also, this section needs generally to address whether the different simulation and TES/SCIA observation periods contribute to differences in isotopic distributions.

We have changed the sub title and added a discussion on the issue of the different simulation period.

P9113 L17-L22: Is the SCIAMACHY data filter and clear sky bias incorporated into the model comparison?

For SCIAMACHY, the same data filtering procedures as in Frankenberg et al. 2009 were applied. We do not apply any correction to the model for comparison with SCIAMACHY.

P9116 L10: in this section please discuss previous studies that have considered Rayleigh vs. non-Rayleigh isotopic distributions (e.g. Worden et al. (2007), Brown et al. (2008), Risi et al. (2012)), and compare your results to those.

We added a discussion on spatial and vertical Rayleigh distillation in section 4.3. The general conclusion that condensation processes alone cannot account for the observed isotope signals in the atmospheric water vapor is in agreement with studies from Worden et al. (2007) and Brown et al. (2008). We added a discussion about former studies in this section.

P9117 L4-L11: here, please also discuss the contribution of the differences in time periods to differences between the datasets.

The discussion in P9117 L4-11 discusses the missing data in the polar/higher latitudes regions from satellites, which contribute significantly to flatter slopes of ln(TPW) vs ln(HDO/H_2O) in the satellite datasets compared to model results. We feel that this is a major difference between the datasets and not the difference in time periods.

P9118 L23: split these references up according to each factor.

P9120 L3: change region to regions.

Points taken.