Sutanto et al. compare global scale water vapor δD patterns across several different satellite datasets and the ECHAM4 GCM. The primary contributions of this work are to demonstrate the improvements in TES v5 and to examine issues in making model-data comparisons, relating to the application of averaging kernels to model fields in the presence of possible humidity biases, both of which are worthwhile.

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 TESv5, it is therefore not clear what advance is being made by this study in terms of using these signals as benchmarks.

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.

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

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?

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.

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
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?

P9099 L29: change retrieve to retrieved

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.

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?

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.

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.

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

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.
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

P9112 L11: change greatly increased in to increased
P9112 L17: change datasets to data points
P9112 L21: change significantly to significant

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

P9116 L10: in this section please discuss previous studies that have considered Rayliegh 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.

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

P9118 L23: split these references up according to each factor

P9120 L3: change region to regions