Summary

This paper presents a simple box model solving the water isotope budget in the sub-cloud layer to quantify the relative contributions of sea surface temperature, relative humidity, mid-tropospheric depletion, and the fraction of moisture from the free troposphere (r_{orig}) on the variability of δD in near-surface water vapor (δD_0). The contribution of r_{orig} is further separated into contributions of specific humidity at the surface, and the height (z_{orig}), relative humidity and temperature from which the free tropospheric air originates. Z_{orig} is found to be an important factor explaining the seasonal-spatial and daily variations of δD_0. This means that measurements of δD_0, if precise enough, can potentially be used to estimate z_{orig} and distinguish between different mixing processes in the atmosphere.

The paper is interesting and well written, and it nicely demonstrates the use of measuring water vapor isotopes on short time scales. The box model's theoretical framework is described in detail and its drawbacks are clearly identified by the authors. I only have a few comments about the methods, the rest are mainly ideas for clarifying the paper. I recommend that the paper be published after minor revisions.

General comments

1) I like the method for quantifying the contributions of different factors by linear regression. I see how this works when the contributing factors have the same units as the variable of interest, which was the case in the previous studies that used this method and are cited in this paper (Risi et al. 2010, Oueslati et al., 2016). Here the different factors all have different units, and the slope therefore depends on the units, or how much the components vary. I assume this was accounted for somehow, as the slopes in the tables are all unitless, but it is not clear from the text, and makes me a bit skeptical about the results. More explanation on that would be useful.

2) As stated in the paper, the methods rely on the assumption that the δD profile follows a Rayleigh-like line, and that there is no effect of rain evaporation. Figures 7 and 8 show that the δD profile is often closer to a mixing line than a Rayleigh line, and the large contribution of r_{orig} mainly comes from ascending regions, where clouds are most likely precipitating. It would be nice to see some quantification of how this impacts the results. A possible way to do this is to remove days/locations where the RMSE of the mixing line is smaller than the RMSE of the Rayleigh line and where there is precipitation, then repeat the analysis for these new fields and add the results in brackets in Tables 1, 2 and as dotted lines in Figures 10, 12.

3) The paper presents the new box model as an extension of the model by Benetti et al. (2015), which is technically true, but can be a bit misleading because its application is different. Rather than predicting δD_0 from z_{orig}, it predicts z_{orig} from δD_0 and therefore requires δD_0 to be known. This means it cannot be applied to initialize Rayleigh models like the model by Benetti et al. (2015), which assumes constant z_{orig}. This could be written more clearly (e.g., from the abstract it seems like the model can be used to predict $\delta D_{0,}$ which is only possible if z_{orig} is known).

4) Changing some of the colors and colormaps could make the figures easier to understand. For example, I think the contributions of different factors and how they add up in Figures 9 and 11 would be more intuitive with a perceptually uniform colormap going from light to dark colors. Also, the red and pink lines in Figures 3, 4, 7, 10, 12, 15 look very similar to each other. It would be good to use a different color for one of them.

Specific comments

P1 L13: [D]/[H] instead of [HDO]/[H_2O]
P2 L22: high bias instead of low bias?
P2 L30: Please introduce the abbreviation for LCL
P3 L4: pointed out the important role
P3 L15: "We do not call it entrained": The word entrained/entrainment still appears a few times in the text (e.g. in Fig. 2, the title of section 4.4)
P3 L23: during a field campaign, global outputs of an isotope-enabled GCM.
P3 L24: "at the global scale": Really? There are no global maps. Are the numbers in Tables 1, 2 and the lines in Figures 10, 12 from global output, or from the region shown on the maps?
P3 L28: capturing the second-order parameter d-excess
P3 L32: "MJ79 already performs quite well for d-excess": Pfahl and Wernli (2009) would probably disagree.
P5 L23: $r \rightarrow r_{orig}$
P6 L20: measurements
P6 L20: "Therefore, variations of δD_0 that are mediated by q_0 or h_0 do not interest us": But δD in the FT is prescribed as a function of q (confusing).
P8 L11: Refer to l'Hopital's rule?
P8 L21: follows ~~as~~ mixing line
P9: Fig.3: $\alpha_{eff} = \alpha_{eq}$ instead of $\alpha_{eff} = 1/\alpha_{eq}$
P11 L25: "Only profiles during the ascending phase of the balloons are considered": (Why?)
P11 L27 (title): write somewhere that these results are based on LMDZ output (not observations)
P12 Fig. 5: Describe abbreviations (LCL, EIS, SCL) in caption.
P12 L2: "if the end member is defined below 500hPa (e.g. 600hPa) results are not always reasonable": In what sense? Why?
P15 Fig. 7: What meteorological conditions do these examples represent? Would it be possible to show all (/more) simulated profiles in the background, e.g. in some transparent color, to get a better feeling for the variability? Also, I suggest adding markers to highlight where the levels are.
P15 L1: Figure 8d instead of 8c.
P15 L5: α_{eq} as a function of temperature
P16 Fig. 8: in boreal winters of all years
P17 L22: "in the cold upwelling regions": for example where?
P17 L23: probably reflects
P17 L24: "the effect of r_{orig} can be seen on the composites as a function of EIS and not as a function of ω_{500}": I don't see this, please elaborate.
P17 L30: followed by h_0 (23%), r_{orig} (16%), …

P18 Fig. 9: Are the correlations significant everywhere? Otherwise, add hatching where not significant?
P19 Fig.10: ω_{500} (hPa/d)
P20 Tab. 2: q_0 seems to be important in Fig. 12, but the slope is 0.0 here, h_0 seems to be unimportant in Fig. 12 but slope is 0.91 here. Why is that?
P20 L1: "it would translate into a lower z_{orig}.": Why?
P22 Fig.12: ω_{500} (hPa/d)
P25 L6: the cruise~~s~~ goes
P25 L8: "when considering only the 6 data points when z_{orig} < 2000m": Rationale behind this?
P25 L14: … at the seasonal-spatial and daily scale is the proportion of the water vapor in the SCL that ~~is~~ originates from above
P26 Fig. 15: r → r_{orig}
P27 L1: there → they
P27 L13: the temporal variability of α_{eff}. Is it possible to estimate the uncertainty from the spatial variability of α_{eff} as well (in the vertical, i.e. how much the δD profile differs from a Rayleigh line with constant α_{eff})?
P27 L21: estimating z_{orig} from δD_0 measurements on a daily basis (?)
P28 L2: and if we measure
P28 L3: swap trade-wind cumulus and strato-cumulus clouds
P29 L14: very precise~~d~~ estimates
P29 L18: the altitude from which the air ~~is~~ originates, and is not ~~to~~ biased by

References

Benetti, M., Aloisi, G., Reverdin, G., Risi, C., and Sèze, G.: Importance of boundary layer mixing for the isotopic composition of surface vapor over the subtropical North Atlantic Ocean, J. Geophys. Res. Atmos., 120, 2190–2209, doi:10.1002/2014JD021947, 2015.

Oueslati, B., Bony, S., Risi, C., and Dufresne, J.-L.: Interpreting the inter-model spread in regional precipitation projections in the tropics: role of surface evaporation and cloud radiative effects, Clim. Dyn., 47, 2801–2815, doi:0.1007/s00382-016-2998-6, 2016.

Pfahl, S., and H. Wernli, Lagrangian simulations of stable isotopes in water vapor: An evaluation of nonequilibrium fractionation in the Craig-Gordon model, J. Geophys. Res., 114, D20108, doi:10.1029/2009JD012054, 2009.

Risi, C., Bony, S., Vimeux, F., Frankenberg, C., and Noone, D.: Understanding the Sahelian water budget through the isotopic composition of water vapor and precipitation, J. Geophys. Res, 115, D24110, doi:10.1029/2010JD014 690, 2010.