RESPONSE TO REVIEWER #3

We thank the reviewer for the constructive feedback, which has helped us to improve the manuscript. We have addressed each of the reviewer comments below.

Major remarks:

The paper is well-written and provides interesting new results on historic trends in evapotranspiration (ET). ET links the land water balance and the energy balance and is a key variable of the hydroclimatological system. Previous assessments on trends in ET show diverse results. This study gives further insights on ET-trends for the 1960-1999 period for a wide-range of different river basins, following a comprehensive approach. The authors also attribute detected changes to variations in meteorological or land surface-related properties. However, the authors possibly should address the following comments to improve the manuscript:

1) Please provide a brief description of the variance partitioning method. Probably many potential readers of the study are not used to it and as you draw many conclusions based on this method, it would be nice to know at least basics without buying the referenced textbook. I guess, 2 to 3 additional sentences would be enough... or maybe add a small appendix.

   We have added additional information on the variance partitioning method in the revised manuscript (section 2.5).

2) You make very strong statements based on figure 3, like e.g. 'Wind speed significantly affected ET variability in some small coastal basins (see the Sacramento basin in Fig. 3).' (5749 I. 6-7). A strong relationship between precipitation and ET is obvious, but a relationship between ET and e.g. wind is not. Please support these statements with calculating at least a correlation (and its corresponding significance). Also, please comment shortly on causality. Of course, time series of ET and e.g. wind for Sacramento basin look maybe quite similar, but unfortunately, this tells you nothing about a causal relationship between them. I would guess its more probable to find a link between wind and precipitation, because when you have a lot of wind, you probably have a lot of precip or vice versa.

   We do not attempt to show causality of ET variability, as that would involve methods more complicated than those employed here. The wording of the section has been revised as not to imply causality is shown and correlation coefficients are now shown in a revised Figure 3. We have explained the choice of each predictor variable in the Introduction and provided a short explanation as to why those variables may influence ET trends and variability.

3) Provide some information on how you've tested on significance. You give some p- Values in your table, but it is not clear how they are computed. Also provide significance information on each basin in figure 4. Equal trends in tropical and subpolar basins will definitely have different significance levels.., so maybe you could highlight basins having significant trends and leave out those with insignificant trends.

The table captions state the slopes and p-values were derived from a multiple regression analysis but we have also clarified this in the Analysis section of the revised manuscript (section 2.5). The significant basins are now highlighted by using shading in Figure 4.

4) Why do you use just two precip datasets. Your analysis would definitely improve from using more datasets.

In response to reviewer comments, we also analysed (CRU-based) WATCH precipitation forcing dataset (Weedon et al., 2011), which is corrected for gauge undercatch. However, we found little effect on our results (discussed in more detail in our response to Reviewer Comment #2).

Specific comments:

1) 57431. 15: Do you really need to fill gaps? You are able to calculate trends and other basic statistics without interpolating. However, if you want to interpolate, it is definitely better to choose another approach, as your technique is probably creating inconsistencies. Imagine a dry year with a missing month within two wet years, creating you a wet month in an overall dry year.

It was necessary to fill gaps in order to look at interannual variability, as it wouldn't have been possible to calculate annual totals with missing months present. We filled ≤10 missing months per basin over a 39-year period (max. 2% of the time series) so we do not believe the gap filling or our chosen method has created significant artifacts in our analysis.

2) 57431. 25-26: In fact, you do not need to introduce the Budyko framework here, as you just use the aridity index to determine 'dry' and 'wet' areas. Also, please motivate the choice of an aridity index of '1.5' for classification. Why don't you use an aridity index of '2', following UNEP's definition of semiarid and semihumid regions (see Middleton et al., 1997)? The distribution of wet to dry basins in figure 1 seems also a little bit random. So, if you decide not to mention the Budyko framework here, also remove it from the abstract. If you decide to mention it here, introduce it briefly.

We believe it is worth mentioning the Budyko framework in this section as the concept of energy/water limitation is derived from the framework and is the basis for why we have chosen to classify our basins into 'wet' and 'dry'. We have added additional references for the framework.

The threshold of 1.5 for aridity index (A) was chosen because basin behaviour seems to change around A = 1.5, where basins with A > 1.5 showed converging relationship with precipitation (ET vs. precipitation R^2 values were consistently high after crossing this threshold). The threshold of 1.5 is also used in a recent paper by Zhang et al. (2012).

3) 57451. 5: This datasets has a pretty coarse resolution. Could you please comment on this and emphasize corresponding limitations to your analysis.

We acknowledge that the coarse resolution is a limitation (despite interpolation to a higher resolution of 0.5°), and this may particularly affect

small basins prone to large spatial variations in wind speed.

4) 57451.19: Could you please provide some insight on this kind of temporal interpolation and if it is appropriate or not?

We acknowledge that the temporal resolution of the original dataset is limited and we have pointed out in the manuscript that land use data have inherent uncertainties in the absence of high temporal/spatial resolution satellite records for our study period. However, we think it is acceptable to assume that land use changes are gradual at the basin scale and should show limited interannual variability, and thus can be interpolated to annual time steps without introducing artefacts to the data.

5) 57481. 17: If you like to draw conclusions from the Budyko framework, please introduce it briefly (like mentioned above) or give maybe the original reference: 'Budyko, M. I. Climate and life (Academic Press, New York, 1974)'.

We have added Budyko (1974) as a reference.

6) Regarding the figures: All the text in figures 2,4 and the tickmarks in figure 3 are quite small.

We have improved the figures.

References:

Budyko, M. I.: Climate and life, International Physics Series, Academic, New York, 1974.

Weedon, G. P., Gomes, S., Viterbo, P., Shuttleworth, W. J., Blyth, E., Österle, H., Adam, J. C., Bellouin, N., Boucher, O., and Best, M.: Creation of the WATCH Forcing Data and its use to assess global and regional reference crop evaporation over land during the twentieth century, J. Hydrometeorol., 12, 823-848, doi:10.1175/2011JHM1369.1, 2011.

Zhang, Y., Leuning, R., Chiew, F. H. S., Wang, E., Zhang, L., Liu, C., Sun, F., Peel, M. C., Shen, Y., and Jung, M.: Decadal trends in evaporation from global energy and water balances, J. Hydrometeorol., 13, 379-391, doi:10.1175/JHM-D-11-012.1, 2012.