Manuscript Review Report A temperature threshold to identify the driving climate forces of the respiratory process in terrestrial ecosystems
Authors: Zhang, et al

The authors reported an analysis of flux data from 152 sites to establish an empirical relationship between plant respiration (R_e) and a number of selected explanatory variables. The authors used quantile regression to study how temperature (annual mean) and other meteorological factors affect R_e. The authors hypothesized that annual mean meteorological factors will affect the maximum achievable R_e. As a result, a quantile regression model is used to examine the changes of the 99th percentile of R_e as a function of various meteorological factors.

The statistical analysis has two fatal flaws.

- When the interest is the maximum of R_e at a given temperature, 99th percentile is a poor substitute. In addition, quantile regression would be a poor choice for such extreme upper quantile. A regression model uses the general probabilistic assumption to quantify the distributional parameters. When a specific quantile is of interest, the quantile regression is appropriate, as long as the quantile is not extreme (e.g., 1 or 99%). Such extreme quantiles render conventional statistics useless as the behavior of extreme values should be modeled differently. This is a well known result originally by Gumbel (1935) (English translation, *Statistics of Extremes*, appeared in 1958). Recent interest in climate change effect further expanded the applications of extreme value statistics. When using quantile regression to model the 99th percentile, the resulting model is highly uncertain. This behavior is expected as we rarely observe extreme values. Extreme values follow a different probabilistic pattern; analyzing extreme values using conventional statistics models will likely lead to misleading results.

- Fitting R_e against MAT (or annual means of other factors) across the entire data can be misleading. When MAT ranges from -10 to 25°C, we are lumping cross globe (spatial) differences as a result of annual mean temperature. However, the inference is largely local. That is, the authors attempted to use regional differences to infer the effect of climate change at a local level. This could fall into the trap of the Simpson's paradox, which states that correlation measured in aggregated scale is not necessarily the same at individual level.

In its current form, the paper is unacceptable because of the potential misleading results. In fact, Figure 5 of the manuscript suggests that the finding of a MAT threshold of 11°C is likely an artifact. I recommend that the manuscript be declined.