Point-by-point response to referee's comments on Observational evidence of temperature trends at two levels in the surface layer (acp-2015-580)

We thank the referees for their insightful comments, which helped improve our manuscript. For clarity, we use italic fonts for comments from Referees, a standard font for author's point-by-point responses, and highlighted fonts in light blue with quotation marks for author's changes in manuscript.

Referee #1

Comments: The manuscript reports on analysis of near-surface lapse rate trend from 1997 to 2013 using high-quality two-height Oklahoma Mesonet observations. They show that the near-surface temperature lapse rate trend has significantly decreased, which indicates that temperatures in 9m height increased faster than 1.5m. The results provide the observational evidence of near-surface temperature changes with respect to height. Generally this manuscript is well written and illustrated. However, from global perspective, the warming is enhanced in Northern Hemisphere mid-high latitudes, especially for cold season (Enhanced cold-season warming in semi-arid regions, Atmospheric Chemistry and Physics, 12 (12), 2012). The authors may also investigate the seasonal differences of the lapse rate trend. In addition, the global warming hiatus starts around 1998, I just wonder if there are any implications by analysing the lapse rate trend for us to understand the warming hiatus?

Thanks much for the comments. First, the temperature could also have cooled more at 1.5m. The statistics do not tell us which is correct, although it appears that warming has occurred (albeit not statistically significant at the individual levels). However, a more important finding from our research, is that near the surface the magnitude of the trend depends on height.

This is a new result as there are no other long term high-quality mesonets with temperatures at two heights that we are aware of. This introduces a previously unrecognized uncertainty in the assessment of the magnitude of multi-year and multi-decadal surface temperature trends.

In addition, at this point, it is not possible to extrapolate the results for a global land effect (or to address a global hiatus in our paper) due to data from one region. In fact, this calls for similar types of measurement from other regions to address global nature of the results of this study. However, since we found that the differences in trends were largest for light winds at night, which occurs often at high latitudes in the winter season (where night can last 24 hrs per day), we suspect that our results will apply there as well.

Referee #2

General comments: (1) The authors have not done much beyond a simple land-cover classification into "grassland" and "cropland" to characterize the local environment of the stations. In particular, a decrease in surface roughness could be largely responsible for the changes seen here. Similarly, changes in irrigation practices would alter soil moisture, albedo, roughness, etc. The correlation analysis helps to get at some of this, but all of these processes are connected through nonlinear feedbacks. As a result, bivariate correlation is not able to demonstrably deconstruct the causes or interactions.

The authors would like to thank the referee for their insight and helpful comments. The major objective of this study is to determine if there is a difference in temperature trends at two heights within the surface layer, i.e., lapse rate trends in Oklahoma. It is not the intention of the authors to ignore "grassland" and "cropland" effects. We agree that changes in soil moisture, surface albedo, and surface roughness are nonlinear feedbacks thus the bivariate correlation is unable to deconstruct the interactions (as we stated in our paper 'the mechanism of decreased lapse rates and latitudinal gradients of surface lapse rate trends observed in Oklahoma warrants further study and longer observation data in the future'). Differences in lapse rates due to aerodynamic roughness certainly could be one of the influences, but we doubt it is the only one. Nonetheless, this issue of attribution is not the focus of our current study.

General comments: (2) The analysis that uses the MODIS land-cover classification doesn't address the central issue regarding land-cover. It's not the land-cover itself but any change in land cover that could be causing the differential trends. Also, "cropland" that transitions from wheat to corn during this time would have a different response from that that transitions from corn to wheat (in seasonality of the growing season, albedo, surface roughness, etc.). This part of the analysis is rather weak and doesn't add much to the conclusions.

In our initial submission, we placed this information in the supplemental information including Figures 1 to 3; then the ACP editor suggested we merge the supplemental information into the main text. We agreed with the referee that this part of the analysis is weak and doesn't add much to the conclusion. We described that 'the mechanism of decreased lapse rates and latitudinal gradients of surface lapse rate trends observed in Oklahoma from 1997 to 2013 warrants further study and longer observation data in the future' in section 3.5.

General comments: (3) Near the end of Section 3.2, the authors briefly suggest that estimating trends at a single height introduces "an uncertainty that has not yet been accounted for in the use of surface temperature trends to diagnose and monitor global warming". While this is technically correct, it's also a straw-man argument since we already know that we don't have "certainty" with regard to trends at every level in the atmosphere at all locations. What do these *results specifically say about uncertainty in monitoring global warming? For instance, is the uncertainty likely to be in one direction or the other? Based on the limited results here, the 1.5m air temperature trend would be an underestimate of the trend at slightly higher levels in the surface layer. Lumping all of this into "uncertainty" doesn't add to the understanding of the issue.*

Thank you for the comments regarding the statement about uncertainty in surface temperature warming evaluation. We agree with the referee that there is never certainty. However, our analyses show that trends diagnosed from a single level within the surface layer is not sufficient to construct a global average surface temperature trend. Whether such an average would be warmer or not is not the issue. Rather, it is yet another reason to adopt a mass weighted average of heat content to diagnose global warming as was recommended in:

Pielke Sr., R.A., 2003: Heat storage within the Earth system. *Bull. Amer. Meteor. Soc.,* 84, 331-335.

General comments: (4) Throughout the manuscript, the authors refer to the lapse rate as a response to turbulent energy exchanges. But it's also the case that the lapse rate is a determinant of turbulent energy exchange. This nonlinear interaction should lead to a more nuanced interpretation of the role of changing lapse rates in the near-surface environment.

Thank you for this insightful point. We concur that the nonlinear interaction near the surface layer is a complicated process where lapse rate changes involve interactions with turbulent energy exchanges (roughness, thermal stability, wind profiles, both air and surface water (vapor) dynamics). We changed the sentence in the section 3.2 into (Page 11, line 17):

'… the lapse rate involved interactions with stronger turbulent energy exchanges in summer and relatively weaker turbulent energy exchanges in winter in the surface boundary layer.'

In addition, authors had constructive discussions with Dr. Urs Neu at Swiss Academy of Sciences in the past for this aspect so that we acknowledged him in our acknowledgement section as,

'We also thank Dr. Urs Neu at Swiss Academy of Sciences for our constructive discussions in the past on the subject of this paper.'

Below are specific comments and authors' responses:

(1) Abstract: air temperatures at 1.5m are only used over land – much of the global average temperature is from SST measurements.

The ocean covers ~71% of the Earth's surface. The remaining 29% is still important. Moreover, the land portion has been reporting as having the larger anomaly – e.g., see http://earthobservatory.nasa.gov/GlobalMaps/view.php?d1=MOD_LSTAD_M&d2=AMSRE_SSTAn_M

(2) Abstract: "flat terrain" – please quantify.

The terrain is relatively flat. The Oklahoma topography shows that the elevations gradually increased from the east to the western panhandle (elevations of approximately ~ 80% areas are from the sea level to 300 m). There are very small portions of mountains (<5%) located in south-eastern Oklahoma.

(3) In a number of places (e.g., in the figures), the author say "surface temperature" when they should say "air temperature". Surface temperature (over land) is yet another different variable with likely different trends from near-surface air temperature (at different heights).

In the climate community we generally use surface temperatures (over land) (NCEI/NOAA) to represent air temperatures monitored at 1.5 m over land. In the text, we added that we are using surface temperature in the paper to describe air temperature at levels within the surface layer. We changed the sentence in the section 1 into (Page 3, lines 9-10):

'…surface temperature (refers to air temperature at a screen height near ground surface) increases'

(4) Section 2.3: partial derivatives are used where difference operators make more sense, as the variable is discretized and only available at two heights.

We have changed this into $-\frac{\Delta T}{\Delta z}$. Thank you.

In addition, we added one sentence here to help explain the lapse rate trend as (Page 7, lines 7 to 10),

'A negative trend in the lapse rate when the surface layer is stably stratified means that the temperature change became steeper (warmer at the higher level and/or cooler at the lower level). When the surface layer is unstably stratified, a negative trend means the temperature change with height has become less.'

(5) Section 2.3: "missing data was retained"?

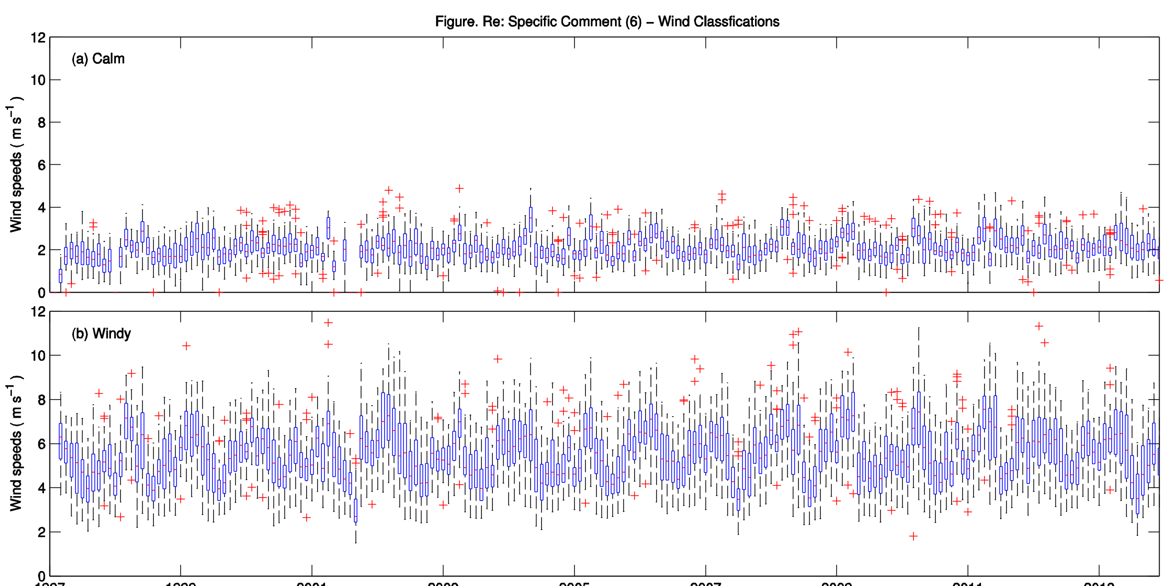
We have changed this into "missing data were not filled or interpolated" on Page 7, lines 10, 11) as:

'… missing data were not filled or interpolated by estimation …'

(6) Section 2.3: why use percentiles of wind speed to define calm and windy. The 5 windiest could be relatively calm conditions during a particularly calm month. This also results in a different definition of calm and windy in each month, which isn't that useful in this context.

We agree with the referee that this method results in a different definition of calm and wind in each month. In our manuscript, we used the same method for wind classifications used in Parker (2004; see below for this reference) by applying percentiles of daily wind speeds for each month. The specific thresholds used in our analysis are shown in Re-Fig. 1. If we use an absolute threshold determined from all months (e.g., from 17 July months) (thus, we have the same definition of calm and windy in each month) we will have unequal weighted averages of wind speeds and/or have varying numbers of missing data for some months.

Parker, D. E. 2004: Climate: large-scale warming is not urban. *Nature*, *432*(7015), 290-290.



Re-Fig 1. Boxplots showing the specific wind speed thresholds used in all 44 stations for calm conditions (a) and windy conditions (b). The blue box indicates the lower quartile, median and upper quartile values. The whiskers extend to 1.5 times the interquartile range and outliers are beyond the end of the whiskers.

(7) Section 2.3: need a reference for Eq. 5.

Thank you. We added this reference.

'Wiederhold, P. R.: *Water vapor measurement: methods and instrumentation* (Vol. 1). CRC Press, 1997'

(8) Section 2.3: Why is potential evapotranspiration referred to as "reference" evapotranspiration?

The term of potential evapotranspiration was first introduced in the late 1940s by both Penman and Thornthwaite. The potential evapotranspiration was not related to a specific crop. However, the reference evapotranspiration term was introduced in the 1970s and 1980s by assuming a fixed surface resistance and albedo as well as a uniform height of green grass under well-watered condition (following Monteith 1965). We calculated the reference evapotranspiration (ET_o).

(9) Section 2.3: not sure that that long section on adjusting the t-test is needed – just say that the effective sample size is used to correct for temporal autocorrelation (and give a reference).

We agree. We condensed this section. One sentence was modified into (Page 9, lines 6 and 7):

'…the effective sample size by correcting for the effect of temporal autocorrelation in the anomaly time series or its residual series (Santer et al., 2000; Karl et al., 2006).'

(10) Section 3: "suggesting a decrease in air humidity" – this doesn't make sense as the decrease is Td is shown. Do you mean caused by a decrease in humidity?

Yes, we wanted to say it was caused by a decrease in humidity. We changed this sentence accordingly on Page 10, line 21:

'which was caused by a decrease in air humidity'

(11) Figures 2 and 3 are virtually illegible and do not add much to this manuscript.

Thank you. Now we removed these two figures from the text in our revision. Both Figures were initially placed in our Supplemental Information (SI). As suggested by the ACP editor, all Figures in our SI should be included in the main text.

Therefore, we slightly changed the text for description of original Figures 2 and 3 in section 2.2. Also the original Figures 4 to 11 were re-ordered into Figures 2 to 9 in our revised manuscript.

(12) Figures: Should make clear that most of these are spatial averages.

All land surface temperatures (sea-surface temperatures as well) are usually calculated by area-averaged. Yes, they are the spatial averages. We described this context in section 2.3 (Page 7, lines 19 to 20) by stating 'was aggregated by using an equally weighted station average from each station when the observations were available' in our manuscript.

(13) Figure 11: why were these data smoothed? Also, trend estimates should be given even when not significant.

Authors thought the smoothing curves would provide a better visualization for possible correlations (linear) because we have a LR time series shown on the top of this figure. Authors have concluded that the presentation of statistically insignificant trends could mislead the readers. We would prefer to not give the trend estimates when they are not statistically significant.