Title Page

Observational evidence of temperature trends at two levels in the surface layer

X. Lin[1*], R.A. Pielke, Sr.[2], R. Mahmood[3], C.A. Fiebrich[4], and R. Aiken[5]

[1]Department of Agronomy, Kansas Climate Center, Kansas State University, Manhattan, KS

[2]CIRES and ATOC, University of Colorado, Boulder, CO

[3]Department of Geography and Geology and Kentucky Climate Center, Western Kentucky University, Bowling, KY

[4]University of Oklahoma, Oklahoma Climatological Survey, Norman, OK

[5]Kansas State University, Northwest Research Center, Colby, KS

* *Correspondence to:* X. Lin (xlin@ksu.edu)

Abstract

Long-term surface air temperatures at 1.5 m screen level over land are used in calculating a global average surface temperature trend. This global trend is used by the IPCC and others to monitor, assess, and describe global warming or warming hiatus. Current knowledge of near-surface temperature trends with respect to height, however, is limited and inadequately understood because surface temperature observations at different heights in the surface layer of the world are rare especially from a high-quality and long-term climate monitoring network. Here we use high-quality two-height Oklahoma Mesonet observations, synchronized in time, fixed in height, and situated in relatively flat terrain, to assess temperature trends and differentiating temperature trends with respect to heights (i.e., near-surface lapse rate trend) over the period 1997 to 2013. We show that the near-surface lapse rate has significantly decreased with a trend of $-0.18 \pm 0.03\ ^{o}\mathrm{C}\ (10\ \mathrm{m})^{-1}$ per decade indicating that the 9 m height temperatures increased faster than temperatures at the 1.5 m screen level and/or conditions at the 1.5 m height cooled faster than at the 9 m height. However, neither of the two individual height temperature trends by themselves were statistically significant. The magnitude of lapse rate trend is greatest under lighter winds at night. Nighttime lapse rate trends were significantly more negative than daytime lapse rate trends and the average lapse rate trend was three times more negative under calm conditions than under windy conditions. Our results provide the first observational evidence of near-surface temperature changes with respect to height that could enhance the assessment of climate model predictions.

1 Introduction

Physical properties of the atmosphere and dynamic processes mix heat vertically and horizontally, yielding the highest temperatures, on average, at the surface with marked seasonal and spatial variations (IPCC, 2013; Karl, et al., 2006). The thermal structure near the surface is affected by various surface forcings (e.g., radiation absorbed and emitted, turbulent mixing, and vegetation interaction) which result in the near-surface lapse rate varying considerably with location and season as well as with atmospheric humidity (Stone and Carlson, 1979; Karl et al., 2006; Mahrt, 2006; Pielke et al., 2007). In the entire troposphere, climate models indicate a distinct height-dependent temperature response to surface temperature increases (refers to air temperature at a screen height near ground surface) (Gaffen et al., 2000; Santer et al., 2005; Karl et al., 2006; Thorne et al., 2011; Seidel et al., 2012; Mitchell et al., 2013). Most of these height-dependent temperature studies focused on tropospheric temperature trends by using radiosonde and satellite observations and climate models (Thorne et al., 2011), however, the near-surface temperature lapse rate has rarely been studied in the surface layer of the atmosphere.

Natural internal climate variability and noise in the data make the detectability of long-term temperature trends in the surface boundary layer difficult. One reason is that the boundary layer typically changes from a convective turbulent regime, with a gain of sensible heat (daytime), to a thermodynamically stable, long-wave radiationally cooled regime (nighttime) with a loss of sensible heat (Pielke et al., 2007; Baldocchi and Ma, 2013). The high-quality two-height surface observations in the Oklahoma Mesonet (Shafer et al., 2000; Lin et al., 2007), however, provide for the first time, an accurate observational network to extract the temperature trend signal at two heights in the surface layer. The temperature observations are synchronized in time, fixed in height, and situated in relatively flat terrain, thus providing a unique opportunity to evaluate near-surface temperature trends and thus the lapse rate trends.

This study is the first observational investigation of two-height, near-surface temperatures to examine lapse rate trends and variability over more than a decade period, a 17-year timescale from 1997 to 2013, which substantially increases the signal-to-noise ratio for trend analysis (Santer et al., 2011) compared to a decade observation (Lin et al., 2007). In this study, our objective is to provide observational evidence for near-surface lapse rate and temperature trends over 1997 to 2013 in Oklahoma.

2 Climate stations and data analysis methods

2.1 Climate stations and data

We selected stations from the Oklahoma Mesonet, which is a world-class network of environmental monitoring stations as reported in 2009 the National Research Council (NRC) recommended the Oklahoma Mesonet as the "gold standard" for statewide weather and climate networks (https://www.mesonet.org/, accessed on 4 May 2015). For two-height temperatures, quality-controlled hourly observations from the Oklahoma Mesonet were used. They include air temperatures at 1.5 m and 9.0 m, relative humidity at 1.5 m, wind speeds (WS) at 2 m and 10 m, global incoming solar radiation (SR), and precipitation. The uncertainties in observations prior to 1997 in the Oklahoma Mesonet were due to an incomplete thermometer's processing algorithm (a delay time required in HMP35C temperature sensors for temperature and humidity measurements) (Shafer et al., 2000) so our study period was from January 1997 to December 2013. Stations that experienced relocation and missing high-level (i.e., 9 m) temperature measurements were excluded leaving a total of 44 Oklahoma Mesonet stations selected (Fig. 1) from 104 stations initially commissioned in January 1994 in Oklahoma.

The US Historical Climatology Network (USHCN, version 2.5) consists of 44 high-quality stations in Oklahoma and the data quality of monthly average temperatures has been rigorously examined (Menne et al., 2009) (Fig. 1). These 44 USHCN stations have long been commonly selected for use in evaluating climate changes on the global, regional, and state scales and thus the USHCN temperature is considered as a reference temperature change when evaluating climate change. It was assumed that both the 44 USHCN stations and the 44 Oklahoma Mesonet stations are representative of the Oklahoma state region in this study.

2.2 Homogeneity tests of temperature time series in the Oklahoma Mesonet

In the USHCN dataset, the instrument change adjustments in a climate series "is a regional average" (Quayle et al., 1991; Hubbard and Lin, 2002; Hubbard and Lin, 2006). The exact effect at individual stations may vary depending on local environmental or climate factors such as the direction of sunlight and wind speeds around the radiation shields. Temperature data used in the study from the Oklahoma Mesonet are quality controlled and thermometers used in the network have been calibrated every 24 to 60 months. The air temperature at 9 m height was measured by a thermistor in a naturally ventilated radiation shield from 1997 to 2013. Air temperature instruments at 1.5 m height were changed from a naturally ventilated radiation shield into an aspirated radiation shield in late 2008. Therefore, homogeneity tests of monthly temperatures for individual Mesonet stations in both $T_{1.5m}$ (temperatures at the 1.5 m height) and T_{9m} (temperatures at the 9 m height) series over 1997 to 2013 were evaluated using two methods:

standard normal homogeneity test (SNHT) (Alexandersson and Moberg, 1997; Peterson et al., 1998) and multiple linear regression (MLR) (Vincent, 1998; Reeves et al., 2007). Note that a time series was classified as homogeneous only if the null hypothesis of homogeneity was not rejected at the 95% level, using both methods to evaluate the single-most-probable discontinuities (or step changes). The reference series (R_i) was formed by using five nearest stations weighted by squared correlation coefficients ($\rho_{i,j}$). In their simplest form, the SNHT and MLR are written as,

$$Q_i = (y_i - \overline{y_i}) - \frac{\sum_{j=1}^{5} \rho_{i,j}{}^2 (x_{i,j} - \overline{x_j})}{\sum_{j=1}^{5} \rho_{i,j}{}^2} \tag{1}$$

$$y_i = a_2 + b_2\, I_{(i \geq c)} + c_2 R_i + e_{i,2} \tag{2}$$

The second part of SNHT's Equation (1) is the reference series. The x_j is a surrounding station series and y_i is the candidate station series to be tested. In the MLR's Equation (2), the I variable is a binary variable which is zero prior to the change point (c) and one after the occurrence of that change point (c). The e_i in Equation (2) is the regression residual term. Note that for R_i in Equation (2), the reference series is the same as the second part of Equation (1).

The 44 $T_{1.5m}$ candidate series were tested against the nearest five USHCN stations, creating the reference series. Three documented change points and five undocumented change points were detected in the $T_{1.5m}$ temperature series. Three documented change points were adjusted in this study. For the 44 T_{9m} candidate series, the instruments have been consistently operated by naturally ventilated radiation shields from 1997 to 2013. Larger ambient wind speeds at the 9 m height relative to the 1.5 m, reduce radiative errors for T_{9m} temperatures (Hubbard and Lin, 2002). When the 44 T_{9m} series were tested by using a reference series created from the five nearest Oklahoma Mesonet stations at the 9 m height, only two change points were found which were undocumented. These undocumented changes were not adjusted in our T_{9m} temperature series.

2.3 Data and trend analysis

The lapse rate is defined as $-\frac{\Delta T}{\Delta z}$ by using the hourly temperatures observed at 1.5 m and 9.0 m in units of ^{o}C $(10\ m)^{-1}$. 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. All missing data were not filled or interpolated by estimation and no outlier screening was implemented in the study. When there were three hourly temperatures missing, the daily lapse rate was excluded. The monthly data were excluded when more than 5 days were missing in a month. The air temperatures at two heights for daytime and nighttime were calculated based on the sunrise and sunset hours (rounded into an integral hour) during any calendar day. The mean wind speed of 2 m and 10 m heights was used to classify wind regimes as windy (87% percentile or above, i.e., 5 windiest days in a month) or calm (17% percentile or below, i.e., 5 calmest days in a month) conditions on a monthly basis.

Monthly anomalies for lapse rates, temperatures, and other climatic variables were departures from monthly climatology for the period from January 1997 to December 2013. The regional time series was aggregated by using an equally weighted station average from each station when the observations were available.

The computation of complementary variables shown in this study is briefly described here. The total energy content of a unit parcel of air (per kg) is provided by the sum of the kinetic energy, latent heat, enthalpy, and gravitational potential energy (Peterson et al., 2011). Without considering the gravitational potential energy and kinetic energy, the air heat content (H) was then calculated by (Pielke et al., 2004; Peterson et al., 2011)

$$H = C_p T + Lq \tag{3}$$

Where T is the Kelvin temperature (K) and q is the specific humidity (kg kg^{-1}). Both the specific heat of air at constant pressure C_p (J K^{-1} kg^{-1}) and the latent heat of evaporation L (J kg^{-1}) are calculated by a function of ambient humidity and temperature (Stull, 1988).

The water vapor pressure deficit (*VPD*) was calculated using,

$$VPD = e_s - e_a \tag{4}$$

e_s and e_a are the equilibrium (or saturated) vapor pressure and actual vapor pressure with respect to water obtained from (Wiederhold, 1997),

$$e_w = (1.0007 + 3.46\times10^{-6}P)6.1121e^{[\frac{17.502T}{240.97+T})} \tag{5}$$

where P is the atmospheric pressure (mb) and e_w is the equilibrium vapor pressure (mb); for e_s (mb), T is the ambient temperature (^{o}C); for e_a (mb), T is the dew point temperature (^{o}C). The dew point was calculated from ambient temperature and relative humidity observed in the Oklahoma Mesonet. The pressure P was estimated based on the station elevation. The calculation of reference evapotranspiration (ET_o) used the Penman-Monteith equation (Allen, 2000). All variables in the ET_o calculation are either directly available at the stations or were estimated from empirical equations (Allen, 2000).

For the trend analysis, the adjusted standard error and adjusted degrees of freedom method was used for evaluating the statistical significance of regional temporal trends and individual station trends at the 95% or otherwise specified confidence levels (Santer et al., 2000; Karl et al., 2006). This approach is a modification of the ordinary least squares linear regression to substitute 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)

3 Results

3.1 Surface temperature-related trends at individual levels

Here we present the first observational investigation of two-height, near-surface temperatures to examine lapse rate trends and variability over more than a decade period. For the period of 1997 to 2013, when trends of surface temperature anomalies are evaluated by individual surface temperatures at 1.5 m ($T_{1.5m}$) and 9.0 m ($T_{9.0m}$) from Oklahoma Mesonet stations, statistically non-significant trends of $+0.065\pm0.59$ ^{o}C per decade and $+0.281\pm0.58$ ^{o}C per decade, respectively were documented (Fig. 2a and b). However, trends could not be confirmed for either of these two individual surface temperatures over Oklahoma (derived from $T_{1.5m}$ and $T_{9.0m}$) when adjusting the statistical analysis for first order autocorrelation effects as shown by the adjusted p values in the trend analysis. When we used the USHCN data, the surface temperatures (T_{USHCN}) again showed a statistically non-significant trend of 0.079 ± 0.58 ^{o}C per decade (Fig. 2c) over 1997 to 2013.

In terms of month-to-month variability of these three time series (Figs. 2a to 2c), the standard deviations over the period studied were 1.63, 1.64, and 1.65 ^{o}C for $T_{1.5m}$, T_{9m}, and T_{USHCN}, respectively, without any statistical differences. To further examine the change over 1997 to 2013 at a single height, the surface air heat content (H) (Pielke et al., 2004; Peterson et al., 2011) was evaluated. Again, we were unable to confirm a statistically significant trend in H although the H showed an apparent ‘cooling’ trend (i.e. -0.737±1.08 kJ/kg per decade) (Fig. 2d), which was caused by a decrease in air humidity (Fig. 2e).

The air heat content variability was very similar to the air temperature’s month-to-month variability although it was coupled with air humidity (Fig. 2d and e). The temperature difference between measurements at 1.5 m of the Oklahoma Mesonet and USHCN ($T_{1.5m} - T_{USHCN}$) had an overall standard deviation of 0.17°C where less variation occurred during the first 10 years, relative to the subsequent 7 years. A slightly positive $T_{1.5m} - T_{USHCN}$ difference, observed during the last three years, cannot be attributed to the thermometer’s exposure changes in the Oklahoma Mesonet because the aspirated thermometers could have a cool-bias compared to non-aspirated thermometers at observing stations. Nonetheless, the overall 0.17°C standard deviation of $T_{1.5m} - T_{USHCN}$ is of the order of uncertainties associated with any current thermometer used in climate monitoring networks (Hubbard and Lin, 2002; Lin et al., 2005).

3.2 Surface lapse rate trends and seasonality

Figure 3 shows the lapse rate changes and changes of its monthly anomalies for daily, daytime, and nighttime conditions. The lapse rate is defined as $-\frac{(T_{9m}-T_{1.5m})}{7.5\text{m}}$ values plotted in Figures in units of ^{o}C (10 m)$^{-1}$. There was a substantial and clear seasonality signal in the daily lapse rate time series (Figs. 3a and 4a). The lapse rates in summertime were larger than in the wintertime which indicated that 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 (Figs. 3a and 4a).

The statistically significant trend of the daily lapse rate was -0.18±0.03 ^{o}C (10 m)$^{-1}$ per decade, and this daily lapse rate trend is the average of daytime (-0.16 ^{o}C (10 m)$^{-1}$ per decade) and nighttime (-0.20 ^{o}C (10 m)$^{-1}$ per decade) lapse rate trends as expected; all at the 99.9% confidence levels. The nighttime lapse rate not only showed a larger trend than in the daytime but also varied significantly more (Fig. 3b and d).

In Fig. 3b, the metadata inventory of thermometer changes suggests that there could be systematic biases which might compromise trend analysis. In addition to the routine quality-control and instrument calibrations of $T_{1.5m}$ and T_{9m}, we conducted multiple lines of data examination in $T_{1.5m}$ and T_{9m} for their fidelity: (1) data homogeneity tests were conducted and documented change points were adjusted; (2) the $T_{1.5m} – T_{USHCN}$ time series (Fig. 2f) showed that no systematic biases existed in $T_{1.5m}$ due to instrument changes late in 2008; and (3) the relatively flat lapse rate anomalies from 2009 to 2013 did not support a systematic bias caused by changes from naturally ventilated thermometers to aspirated thermometers (Fig. 3b). Therefore, it is unlikely that changes from naturally ventilated to aspirated thermometers in 2008 and 2009 contribute to the lapse rate trends.

That the lapse rate trend is statistically significant is initially surprising, since the individual two-height temperatures have no significant trends (Fig. 3a and b). We explained how this can occur in the Appendix A (see Figs. A1 and A2). Results in Fig. 3 indicated that the temperature difference between $T_{9.0m}$ and $T_{1.5m}$ had a statistically significant increasing trend. Considering the statistically non-significant trends in $T_{1.5m}$ and $T_{9.0m}$ (Fig. 3a and b), we infer that the near-surface vertical temperatures at 9 m were warming faster than temperatures at the screen level (1.5 m) in the surface boundary layer. However, it is possible that cooling (which is within the range of statistical uncertainty) at the 1.5 m level could account for the increased temperature difference ($T_{9m} - T_{1.5m}$). The -0.18°C $(10\ m)^{-1}$ per decade of lapse rate trend with a 7.5 m height difference is equivalent to a warming trend +0.135°C per decade of the (T_{9m} minus $T_{1.5m}$) temperatures. Regardless of whether the individual levels can be shown to have statistically significant positive or negative trends over time, our results in Oklahoma present clear evidence for changes of near-surface vertical temperature profiles over the period 1997 to 2013. This means that measurements of trends at a single height introduce an uncertainty that has not yet been accounted for in the use of surface temperature trends to diagnose and monitor global warming.

The seasonality shown in the daytime lapse rate was clearer than in the nighttime lapse rate (Figs. 3 and 4), suggesting that strong turbulent mixing controlled the daytime mixing layer but as expected, there was stabilized surface air (weak turbulence) in the nocturnal boundary layer (Stone and Carlson, 1979; Stull, 1988; Karl et al., 2006; McNider et al., 2012). Thus, the nighttime lapse rate clearly consistently varied much more than the daytime lapse rate over 1997 to 2013 (Figs. 3a and 4a). Figure 4 indicates that part of the daytime was unstably stratified in the surface boundary layer, however, for most of the time over a 24 hour period, the lapse rates show a stable surface boundary layer for all months in the Oklahoma region. During the spring season, the daytime lapse rates were relatively suppressed while nighttime lapse rates were suppressed during the fall season in Oklahoma (Fig. 4a). All daytime, daily, and nighttime lapse rates showed a change between the averages of the first ten years and the last ten years (Fig. 4b).

3.3 Spatial distributions of station's lapse rate trends

To examine spatial aspects of lapse rate changes, the lapse rate trends in 44 individual stations are shown in Figure 5 for daily, daytime, and nighttime lapse rates. All but one station lapse rate trend showed a decrease irrespective of whether they were the daily, daytime, or nighttime analyses. About 16%, 36%, and 23% of all stations showed statistically non-significant trends for the daily, daytime, and nighttime time series, respectively. The majority of stations showed significant decreasing trends, especially for daily lapse rates (Fig. 5). The histogram of individual trends for nighttime indicated trends were more negative relative to daily and daytime lapse trends (meaning the higher level temperature increased more (or decreased less) than the lower lever temperature). Across Oklahoma, the lower latitude region showed more negative lapse rate trends. When dividing all of Oklahoma into northern and southern areas by a 35.4^o N line in latitude, the average lapse rate trends in the southern area were significantly more negative than the average trends in the northern area at 98% confidence levels for daily lapse rates ($Trend_N$ = -0.14 *vs.* $Trend_S$ = -0.24 ^oC $(10\ m)^{-1}$ per decade), daytime lapse rates ($Trend_N$ = -0.12 *vs.* $Trend_S$ = -0.22 ^oC $(10\ m)^{-1}$ per decade), and nighttime lapse rates ($Trend_N$ = -0.16 *vs.* $Trend_S$ = -0.25 ^oC $(10\ m)^{-1}$ per decade) (Fig. 5).

3.4 Wind influences on lapse rate trends

Daytime and nighttime lapse rate trends demonstrate different properties largely due to the diurnal solar cycle, wind speed and its interaction with the land surface (Pepin, 2001; Karl et al., 2006; Mahrt, 2006; McNider et al., 2012). Wind strongly influences turbulent mixing and surface boundary layer depth (Stull, 1988; Pepin, 2001). Figure 6 shows the lapse rate trend and variations under windy and calm conditions. There was no significant lapse rate trend observed under windy daytime conditions (Fig. 6c). The most negative lapse rate trend, -0.40±0.03 ^{o}C (10 m)$^{-1}$ per decade, was found under calm nighttime conditions (Fig. 6f). Both the trend magnitude and variation of lapse rate under calm nighttime conditions were the largest among all classified lapse rates as shown in Figure 6. Since the stable nocturnal boundary layer is very sensitive to local radiative effects from atmospheric CO_2 and water vapor, wind speed, surface roughness, and soil heat capacity (Pielke et al., 2007; McNider et al., 2012), slight changes to the surface layer structure from these local effects could explain part of the observed trends. The observed slight increase in wind speed could have resulted in the 9 m level being above the nocturnal cool level more often later in the observational period, thus a positive temperature trend would be seen in the data due to this effect.

3.5 Trends of related climate variables in Oklahoma over 1997 to 2013

The MODIS Land Cover product (MOD12Q1) was used for the year 2005 (Friedle et al., 2010) to classify all 44 Oklahoma Mesonet stations into 34 grassland stations and 10 cropland stations to examine possible effects of land use and land cover on lapse rates (Fig. 1). Figures 7 and 8 showed that there were no statistical differences among respective lapse rate trends between grassland and cropland stations.

Due to the complexity of the surface vertical temperature profile variations (Stone and Carlson, 1979; Pepin, 2001; Mahrt, 2014), here we simply presented a monthly smoothed anomaly time series of climatic variables including solar radiation (SR, W m^{-2}), water vapor pressure deficit (VPD, kPa), mean wind speed (WS, m s^{-1}), precipitation (mm), and reference evapotranspiration (ET_o, mm $month^{-1}$), and their correlations with the monthly lapse rate time series (Fig. 9). Only mean wind speed and reference evapotranspiration showed significant trends; both of which were increasing (Fig. 9d and f). In terms of correlation with lapse rates, solar radiation, reference evapotranspiration, and vapor pressure deficit showed significant correlations with values of -0.55, -0.46, -0.18, and 0.35, respectively.

In summary, for related climate variables, it is understandable that solar radiation is the most correlated due to its strong role on turbulent sensible heat flux from the ground surface associated with vertical temperature gradients and stability. The wind speed did play a role for lapse rate changes in the surface boundary layer (Pepin, 2001; Pielke et al., 2007; McNider et al., 2012; Baldocchi and Ma, 2013). Precipitation changes can provide information about soil moisture changes and its effect on variations of the daytime surface energy budget and heating of atmospheric temperatures (McNider et al., 2012; Baldocchi and Ma, 2013). Nevertheless, 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.

4 Summary and concluding remarks

Our study has the following major findings. First, using the lapse rate (defined as the difference in temperature at two levels) trends can be diagnosed with more statistical confidence than considering temperature trends from each level separately. Second, trends of surface temperature depend on the height at which the measurements are made. A greater warming at the 9 m level, or larger cooling at the 1.5 m screen level would explain such an observation. This is important as the surface temperature is used to diagnose and model global warming (IPCC, 2013). Using just the 1.5 m level trends would provide a different magnitude of trend than if obtained from the temperatures at 9 m [at least in Oklahoma and this may be true elsewhere]. Third, the near-surface lapse rate trends were altered by wind speed. Fourth, lapse rate trends in southern Oklahoma were significantly more negative than further north in the state. Our study suggests a positive temperature trend at 9 m could be due in part, to a change in wind speed during the time period such that the 9 m level more often remains above the nocturnal cool layer later during the observing period.

Finally, since land surface temperatures are often not taken at the same height above the ground, if the magnitude of long-term trends depends on the height of the measurement, it further complicates the ability to accurately quantify global warming using a global average surface temperature trend from a single height of observation at each location used in the construction of the global assessment (IPCC, 2013). This research should provide impetus for building additional or vertical expansion of current in-situ observational infrastructure for a more robust understanding of climate change.

5 Appendix A: How do two individual heights show no statistically significant trends, but the difference or the lapse rate does?

One might question how measurements from two individual heights can show no significant trends but the difference does. To evaluate this, we first generated two monthly temperature anomaly series, representing measurements at 9 m height (m_1) and 1.5 m height (m_2) with a length of 360-month values (i.e., 30 years). The correlation coefficient between m_1 and m_2 was preset at 0.97, which was a typical value for the monthly T_{9m} and $T_{1.5m}$ series in this study. The simulated m_1 and m_2 were generated by introducing fields of random month-to-month temperatures that were normally distributed with a mean of zero and a variance of one. Secondly, the initial trends and noise values in m_1 and m_2 were added to produce the s_1 and s_2 series as,

$$s_1 = m_1 + trend_1 + n_1 \tag{A1}$$

$$s_2 = m_2 + trend_2 + n_2 \tag{A2}$$

where $trend_1$ and $trend_2$ are initial trends imposed on the series, which have four combinations of a non-trended series and a linear trended series. These four trend combinations were [0.00 0.00], [0.00 0.12], [0.12 0.00], and [0.12 0.12] ^{o}C per decade. The n_1 and n_2 are normally distributed noise and n_2's power level was set four times larger than the power level in n_1 because it was assumed that surface temperatures at $T_{1.5m}$ may have larger non-climatic and local-climatic noise than T_{9m}. In terms of noise level, the normally distributed noise n_1 had a zero mean and 0.2 of standard deviation.

The third step was to run simulations 1000 times to generate 1000 pairs of s_1 and s_2 series for the four trend combinations individually, resulting in 1000 difference series of s_1-s_2 for each set of trend conditions. Figure A1 illustrates an example result out of running 1000 simulations when $trend_1$ and $trend_2$ were [0.12 0.00] ^{o}C per decade. This example shows that two individual temperatures (s_1 and s_2) can show no statistically significant trends but the difference (s_1-s_2) does (Fig. A1).

Finally, trend analyses were conducted for the s_1, s_2, and s_1-s_2 series. The results indicate that there were about 600 chances out of 1000 simulations, where two trends of s_1 and s_2 were not significant but the s_1-s_2 trend was significant, that is the [001] status shown in Figures A2b and A2c, under the combination of trends imposed by [0.00 0.12] and [0.12 0.00] °C per decade. When both trends were zero or both trends were 0.12 °C per decade, there was a rare chance to have a significant s_1-s_2 trend (Fig. A2a and A2d).

In summary, a differential process ($s_1 - s_2$) is able to robustly suppress noise common to the s_1 and s_2 series relative to the difference signal (s_1-s_2). Therefore, an improved signal-to-noise ratio series of $s_1 - s_2$ could show a statistically significant trend, but two individual s_1 and s_2 series do not show statistically significant trends.

Acknowledgements

This work was partially supported by the United States Geological Survey (USGS), National Science Foundation (NSF), the U.S. Climate Program Office (CPO), and the USDA Ogallala Initiative Program. Authors thank Monica Deming at the Oklahoma Climatological Survey for permission to use Oklahoma Mesonet data and for valuable discussions. Authors benefitted from insightful comments and discussion from Oklahoma Mesonet supporting staff. We also thank Dr. Urs Neu at Swiss Academy of Sciences for our constructive discussions in the past on the subject of this paper.

References

Alexandersson, H., and Moberg, A.: Homogenization of Swedish temperature data. Part I: Homogeneity test for linear trends. *Int. J. Climatol.* 17, 25-34, 1997.

Allen, R. G.: Using the FAO-56 dual crop coefficient method over an irrigated region as part of an evapotranspiration intercomparison study. *J. Hydrology* 229, 27-41, 2000.

Baldocchi, D, and Ma, S.: How will land use affect air temperature in the surface boundary layer? Lessons learned from a comparative study on the energy balance of an oak savanna and annual grassland in California, USA. *Tellus B* 65, 19994, doi:10.3402/tellusb.v65i0, 2013.

Friedl, M. A., Sullan-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., and Huang, X.: MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sens. Environ*. 114, 168-182, 2010.

Gaffen, D. J., Santer, B. D., Boyle, J. S., Christy, J. R., Graham, N. E., and Ross, R. J.: Multidecadal changes in the vertical temperature structure of the tropical troposphere. *Science* 287, 1242–1245, 2000.

Hubbard, K. G., and Lin, X.: Realtime data filtering models for air temperature measurements. *Geophys. Res. Lett.* 29(67), 1-4, doi:10.1029/2001GL013191, 2002.

Hubbard, K. G., and Lin, X.: Reexamination of the effects of instrument change in the U.S Historical Climatology Network. *Geophys. Res. Lett.* 33, L15710, doi:10.1029/2006GL027069, 2006.

IPCC: Summary for Policymakers. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, UK and New York, NY, USA, 2013.

Karl, T. R., Arguez, A., Huang, B., Lawrimore, J. H., McMahon, J. R., Menne M. J., Peterson, T. C., Vose, R. S., and Zhang, H. M.: Possible artifacts of data biases in the recent global surface warming hiatus. *Science*, 348, 1469-1472, 2015.

Karl, T. R., Hassol, S. J., Miller, C. D., and Murray W. L.: Temperature trends in the lower atmosphere: Steps for understanding and reconciling differences*,* Synth. Assess. Prod. 1.1*, U.S. Climate Change Sci. Program,* Washington, DC, 2006.

Lin, X, Hubbard, K. G., and Baker, C. B.: Surface air temperature records biased by snow covered surface. *Int. J. Climatol*. 25, 1223-1236, doi: 10.1002/joc.1184, 2005.

Lin, X., Pielke Sr., R. A., Hubbard, K. G., Crawford, K. C., Shafer, M. A., and Matsui, T.: An examination of 1997–2007 surface layer temperature trends at two heights in Oklahoma. *Geophys. Res. Lett.* 34, L24705, doi: 10.1029/2007GL031652, 2007.

Mahrt, L.: Variation of surface air temperature in complex terrain. *J. Appl. Meteorol. Climatol.* 45, 1481–1493, 2006.

McNider, R. T., Steeneveld, G. J., Holtslag, A. A. M., Pielke Sr., R. A., Mackaro, S., Pour-Biazar, A., Walters, J., Nair, U., and Christy, J.: Response and sensitivity of the nocturnal boundary layer over land to added longwave radiative forcing. *J. Geophys. Res*. 117, D14106, doi: 10.1029/2012JD017578, 2012.

Menne, M. J., Williams Jr., C. N., and Russell, S. V.: The United States Historical Climatology Network Monthly Temperature Data - Version 2. *Bull. Amer. Meteor. Soc*. 90**,** 993-1107, doi:10.1175/2008BAMS2613.1, 2009.

Mitchell, D. M., Thorne, P. W., Stott, P. A., and Gray, L. J.: Revisiting the controversial issue of tropical tropospheric temperature trends. *Geophys. Res. Lett.* 40, 1-6, doi:10.1002/grl.50465, 2013.

Pepin, N.: Lapse rate changes in northern England. *Theor. Appl. Climatol*. 68, 1-16, 2001.

Peterson, T., C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholl, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R. Tuomenvirfa, H., Mestre, O., Szentimrey, T., Salinger, J., Forland, E. J., Hanssend-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: A review. *Int. J. Climatol*. 18:1493–1517, 1998.

Peterson, T. C., Willett, K. M., and Thorne, P. W.: Observed changes in surface atmospheric energy over land. *Geophys. Res. Lett.,* 38, L16707, doi: 10.1029/2011GL048442, 2011.

Pielke Sr., R. A., Davey, C. A., Niyogi, D., Fall, S., Steinweg-Woods, J., Hubbard, K. G., Lin, X., Cai, M., Lim, Y. K., Li, H., Nielsen-Gammon, J., Gallo, K., Hale, R., Mahmood, R., Foster, S., McNider, R. T., and Blanken, P.: Unresolved issues with the assessment of multi-decadal global land temperature trends. *J. Geophys. Res.* 112, D24S08, doi: 10.1029/2006JD008229, 2007.

Pielke Sr., R. A., Davey, C., and Morgan, J.: Assessing “global warming” with surface heat content, *Eos Trans. AGU* 85(21), doi: 10.1029/2004EO210004, 2004.

Quayle, R. G., Easterling, D. R., Karl, T. R., Hughes, P. Y.: Effects of recent thermometer changes in the cooperative station network. *Bull. Amer. Meteor. Soc*. 72, 1718–1723, 1991.

Reeves, J., Chen, J., Wang, X. L., Lund, R., and Lu, Q.: A review and comparison of changepoint detection techniques for climate data. *J. Appl. Meteor. Climatol.* 46, 900-915, 2007.

Santer, B. D., Wigley, T. M. L., Boyle, S., Gaffen, G. J., Hnilo, J. J., Nychka, D., Parker, D. E., and Taylor K. E.: Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *J. Geophys. Res*. 105, 7337–7356, doi: 10.1029/1999JD901105, 2000.

Santer, B. D., Wigley, T. M. L., Mears, C., Wentz, F. J., Klein, S. A., Seidel, D. J., Taylor, K. E., Thorne, P. W., Wehner, M. F., Gleckler, P. J., Boyle, J. S., Collins, W. D., Dixon, K. W., Doutriaux, C., Free, M., Fu, Q., Hansen, J. E., Jones,G. S., Ruedy, R., Karl, T. R., Lanzante, J. R., Meehl, G. A., Ramaswamy, V., Russell, G., and Schmidt, G.: Amplification of surface temperature trends and variability in the tropical atmosphere. *Science* 309, 1551–1556, 2005.

Santer, B. D., Mears, C., Doutriaux, C., Caldwell, P., Gleckler, P.J., Wigley, T. M. L., Solomon, S., Gillett, N. P., Ivanova, D., Karl, T. R., Lanzante, J. R., Meehl, G. A., Stott, P. A., Taylor, K. E., Thorne, P. W., Wehner, M. F., and Wentz, F. J.: Separating signal and noise in atmospheric temperature changes: The importance of timescale. *J. Geophys. Res.,* 116, D22105. doi:10.1029/2011JD016263, 2011.

Seidel, D. J., and Lanzante, J. R.: An assessment of three alternatives to linear trends for characterizing global atmospheric temperature changes. *J. Geophys. Res.* 109, D14108, doi:10.1029/2003JD004414, 2004.

Seidel, D. J., Free, M., and Wang J. S.: Reexamining the warming in the tropical upper troposphere: Models versus radiosonde observations. *Geophys. Res. Lett.* 39, L22701, doi: 10.1029/2012GL053850, 2012.

Shafer, M. A., Fiebrich, C. A., Arndt, D. S., Fredrickson, S. E., and Hughes, T. W.: Quality assurance procedures in the Oklahoma Mesonet. *J. Atmos. Oceanic Technol.* 17, 474–494, 2000.

Stone P. H. and Carlson, J. H.: Atmospheric lapse rate regimes and their parameterization. *J. Atmos. Sci.* 36, 415–423, 1979.

Stull, R. B.: *An Introduction to Boundary Layer Meteorology*. Kluwer Acad., Norwell, MA, 1988.

Thorne, P. W., Lanzante, J. R., Peterson, T. C., Seidel, D. J., and Shine, K. P.: Tropospheric temperature trends: history of an ongoing controversy. WIREs Climate Change, 2, 66–88, doi: 10.1002/wcc.80, 2011.

Vincent, L. A.: A technique for the identification of inhomogeneities in Canadian temperature series. J. Climate. 11, 1094–1104, 1998.

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

Figure Legends

Figure 1. 44 Oklahoma Mesonet stations (filled circles) and 44 USHCN stations (open circles), in which the Oklahoma Mesonet stations include 34 grassland stations (green) and 10 cropland stations (black circles). The MODIS Land Cover product (MOD12Q1) (Friedle et al., 2010) in 2005 was used to classify the Oklahoma Mesonet stations into grassland (34 stations) and cropland stations (10 stations). The thin lines indicate the borders of nine climate divisions in Oklahoma.

Figure 2. Monthly time series of (**a**) surface temperature at 1.5 m ($T_{1.5m}$), (**b**) surface temperature at 9 m ($T_{9.0m}$), (**c**) USHCN surface temperature (T_{USHCN}), (**d**) near-surface air heat content (H) at 1.5 m, (**e**) surface dew point temperature (T_d), and (**f**) surface temperature difference between $T_{1.5m}$ and T_{USHCN}. The ± values define the 95% confidence intervals for specific trends shown in each panel. The first two months of H and T_d and the first month of $T_{1.5m}$ were missed due to fewer available observations.

Figure 3. Changes of monthly lapse rate (LR) ($^oC\ (10\ m)^{-1}$) in Oklahoma over 1997-2013: (**a**) absolute daily (blue), daytime (red), and nighttime (green) lapse rates, (**b**) daytime anomaly lapse rates, (**c**) daily anomaly lapse rates derived from 24-hour averaged over two heights, and (**d**) nighttime anomaly lapse rates. The straight lines are least squares trends ($^oC\ (10\ m)^{-1}$ per decade) with adjusted p-values shown. The ± values define the 95% confidence intervals for trends. The shaded region around lapse rate anomalies shows the standard deviation of 44 Oklahoma Mesonet stations. The metadata for dates of thermometer status are shown at the bottom of (**b**) for changes of thermometer radiation shields at 9 m and 1.5 m.

Figure 4. (**a**) Lapse rate seasonality averaged over 1997 to 2013 for daily (blue), daytime (red), and nighttime (green) periods; standard deviations are represented by shaded areas. The significant trends in daily, daytime, and nighttime lapse rates (Figs. 3b, 3c and 3d) were removed before calculating the standard deviation. (**b**) The same as (**a**) but averaged over the first ten years (1997-2006; grey areas with black lines) and averaged over the last ten years (2004-2013; blue for daily, red for daytime, and green for nighttime).

Figure 5. Individual station trends of monthly lapse rate anomalies (^{o}C (10 m) $^{-1}$ per decade) for (**a**) daily lapse rate, (**b**) daytime lapse rate, and (**c**) nighttime lapse rate. The histogram of individual station trends is presented to the right side of each panel. The x axis is the color bar of trends and the y axis represents the number of stations having that trend. Pink squares indicate non-significant trends, otherwise, the stations are significant at 95% confidence levels accounting for serial autocorrelation. The dotted orange line is the line of 35.4^{o} in latitude to divide Oklahoma into northern and southern areas, both of which have 22 stations for evaluating averaged lapse rate trends of southern ($Trend_S$, unit is ^{o}C (10 m) $^{-1}$ per decade) and northern ($Trend_N$, ^{o}C (10 m) $^{-1}$ per decade) areas. The p-value shown is from the two-sample t-test.

Figure 6. Trends and variations of monthly lapse rate (LR) (y-axis units are ^{o}C (10 m) $^{-1}$) classified by windy and calm conditions for daily (**a**) and (**b**), daytime (**c**) and (**d**), and nighttime (**e**) and (**f**) from 1997-2013. Windy conditions were the days that mean wind speeds were above the 87% percentile in a month (i.e., 5 windiest days in each month), and calm conditions were the days where the mean wind speed was below the 17% percentile (i.e., 5 calmest days for each

month). The straight lines are least squares trends (^{o}C (10 m) $^{-1}$ per decade) with adjusted p-values shown. The ± values define the 95% confidence intervals for trends. The shaded region shows the standard deviation of 44 stations.

Figure 7. Changes of monthly lapse rates (LR) (^{o}C (10 m) $^{-1}$) averaged only by grassland stations in Oklahoma over 1997-2013: (**a**) daily anomaly derived from 24-hour averaged over two heights, (**b**) daytime anomaly, and (**c**) nighttime anomaly. The straight lines are least squares trends (^{o}C (10 m) $^{-1}$ per decade) with adjusted p-values shown. The ± values define the 95% confidence intervals for trends. The shade region shows the standard deviation from 34 grassland stations.

Figure 8. The same as Fig. 7 but only for 10 cropland station averages.

Figure 9. Monthly smoothed anomaly time series over 1997-2013 of (**a**) lapse rate (LR, ^{o}C (10 m)$^{-1}$), (**b**) solar radiation (SR, W m^{-2}), (**c**) water vapor pressure deficit (VPD, kPa), (**d**) mean wind speed (WS, m s^{-1}), (**e**) precipitation (mm), and (**f**) reference evapotranspiration (ETo, mm month^{-1}). The p-value$_{adj}$ is p values in the trend analysis, p_r values are p values given in the correlation analysis, and r correlation coefficients with the lapse rate time series. All time series are 7-month moving averages (used as a smoother) of the original monthly data, which were expressed as a departure from the 1997 to 2013 average. The significant trends or correlations are indicated by specified p values but non-significant trends were evaluated at the 95% confidence levels.

Figure A1. An example of two non-significant trends in s_1 (**a**) and s_2 (**b**) temperature time series individually but differentiating them, s_1-s_2 temperature series (**c**) shows a significant trend. This is one realization example taken from the simulations; s_1 and s_2 were constructed with trend values of 0.12 and 0.00 ^{o}C decade $^{-1}$, respectively.

Figure A2. These figures illustrate the frequency of outcomes (shown as the 'y'-axis) for four combinations of initial trends for series s_1 and s_2. The eight possible combinations (shown as the 'x'-axis) are represented by 3-bit binary numbers: the first bit represents the s_1 trend status; the second bit represents the s_2 trend status; and the last bit represents the s_1-s_2 trend status. Each trend status has two possibilities of either a non-significant trend (0) or a significant trend (1). For example, the 001 in the x-axis stands for a combination of a non-significant trend (0) in s_1, non-significant trend (0) in s_2, and significant trend in s_1-s_2 (1). Initial trends of s_1 and s_2 were imposed as (**a**) 0.00 and 0.00; (**b**) 0.00 and 0.12; (**c**) 0.12 and 0.00; and (**d**) 0.12 and 0.12 for each corresponding set of 1000 realizations. The trend units are ^{o}C per decade. The y-axis represents the number out of 1000 simulations for eight combinations of the s_1, s_2, and s_1-s_2 trend status.