

1 **An area and distance weighted analysis of the impacts of station**
2 **exposure on the U.S. Historical Climatology Network temperatures and**
3 **temperature trends**

4 **PRE-PRINT DRAFT DISCUSSION PAPER**

5 Anthony Watts
6 President, IntelliWeather, Chico, CA, USA

7
8 Evan Jones
9 IntelliWeather, Chico, CA, USA

10
11 Stephen McIntyre, Toronto, Canada

12
13 John R. Christy
14 Department of Atmospheric Science, University of Alabama, Huntsville, AL, USA

15
16
17 [plus additional co-authors that will be named at the time of submission to the journal]

18 **Abstract**

19 In Fall et al, 2011, results from the recently concluded Surface Stations Project surveying
20 the U.S. Historical Climatology Network (USHCN) were presented, using a siting
21 classification system developed by Michel Leroy for Meteofrance in 1999, and employed
22 by the National Oceanic and Atmospheric Administration (NOAA) to develop the U.S.
23 Climate Reference Network (USCRN) in 2002. In 2010, Leroy improved upon this
24 system to introduce a "maintained performance classification" which quantifies the effect
25 of heat sinks and sources within the thermometer viewshed by calculation of the area-
26 weighted and distance-weighted impact of biasing elements such as concrete, asphalt,
27 runways, tarmac, and buildings, creating a new site classification that more accurately
28 reflects the representivity of the station exposure. The new area and distance weighted
29 classification system does a more complete job of siting assessment, particularly when
30 applied retroactively to existing stations, than the original distance weighted
31 classification system described in Leroy (1999) , which performs well for new station
32 siting evaluation, but does not take into account the surface area of heat sinks and sources
33 that may encroach upon a temperature measurement station over its lifetime.

34 In Fall et al. (2011), using Leroy's 1999 classification system, it was demonstrated that
35 station exposure affects USHCNv2 temperatures, in particular the minimum
36 temperatures, but showed little difference in mean temperature trends used to assess
37 climate variability. Menne et al. (2010), and Muller et al. (2012), both of which also used
38 the older Leroy (1999) classification system, suggested there is little if any mean

39 temperature trend difference between well and poorly sited stations. Using the new Leroy
40 (2010) classification system on the older siting metadata used by Fall et al. (2011),
41 Menne et al. (2010), and Muller et al. (2012), yields dramatically different results.

42 Both raw and gridded comparisons were performed on the 30 year trends that were
43 calculated for each surveyed station, using temperature data from USHCNv2. Mean
44 temperature trend is indisputably lower for well sited stations than for poorly sited
45 stations. Minimum temperature trend shows the greatest differences between siting
46 classification while maximum temperature trend shows the smallest.

47 Well sited stations consistently show a significantly lower trend than poorly sited
48 stations, no matter which class of station is used for a baseline for comparison, and also
49 when using no baseline at all. Well sited stations, using a localized Class 4 (the most
50 common class) baseline show a trend that is 0.09°C per decade lower than poorly sited
51 stations for raw mean temperature trends. Raw mean temperature trends for well sited
52 stations are 0.145°C per decade lower than adjusted mean temperature trends for poorly
53 sited stations, and 0.145°C per decade lower than adjusted mean trend for all stations.

54 Comparisons demonstrate that NOAA adjustment processes fail to adjust poorly sited
55 stations downward to match the well sited stations, but actually adjusts the well sited
56 stations upwards to match the poorly sited stations. Well sited rural stations show a
57 warming nearly three times greater after USHCNv2 adjustments are applied.

58 It is also demonstrated that urban sites warm more rapidly than semi-urban sites, which in
59 turn warm more rapidly than rural sites. Since a disproportionate percentage of stations
60 are urban (10%) and semi-urban (25%) when compared with the actual topography of the
61 U.S., this further exaggerates mean temperature trends. Montandon et al (2011)
62 documents this large urban bias in station siting on the Global Historical Climate
63 Network.

64 These factors, combined with station siting issues, have led to a spurious doubling of U.S.
65 mean temperature trends in the 30 year data period covered by the study from 1979 -
66 2008.

67 Keywords: Surface Temperature, Historical Climate Network, U.S. Temperature Trend

68 **1. Introduction**

69 A number of recent studies have addressed the myriad of factors and biases associated
70 with temperature surface measurement in the United States. The identified biases include
71 station moves, changes in instrumentation, localized changes in instrumentation location,
72 changes in observation practices, and evolution of the local and microsite station
73 environment over time. Some of the identified changes have been addressed in previous
74 works such as where land use/cover change are considered (e.g. Asaeda et al.,(1996);
75 Baker,(1975); Karl and Williams,(1987); Karl et al.,(1988); Karl et al.,(1989); Davey and
76 Pielke,(2005); Mahmood et al.,(2006, 2010), Pielke et al.,(2007a and 2007b); Yilmaz et

77 al.,(2008); Christy et al.(2009). It has been described by these and other studies that
78 maximum and minimum temperatures measured at the station are affected in different
79 ways by the changes in the station environment. McNider et al.,(2012) shows that even
80 slight increases in the vertical mixing near the observing site (such as a local change in
81 the surface land use) can result in significant changes in the minimum temperature trend.
82 Such nearby changes in the station environment can create inhomogeneities, which in
83 turn induce artificial trends or discontinuities in long-term temperature time series and
84 can result in erroneous characterization of climate variability (Peterson et al., 1998;
85 Thorne et al., 2005). Thus, even if stations are initially placed at pristine locations, i.e.
86 “well-sited”, the station environment can change, altering the characteristics of surface
87 temperature measurements over time. As documented in surveys presented in Watts,
88 (2009), and also in Fall et al.,(2011), the USHCN has a significant portion of stations
89 affected by such changes, with approximately 10% of the USHCN remaining classified
90 as “well-sited” using the Leroy (1999) classification method.

91 There have also been a number of attempts to address these station inhomogeneities.
92 These include statistical identification methods for detecting, quantifying, and removing
93 discontinuities and various non-climatic biases that affect temperature records have been
94 employed (e.g. Karl et al., 1986; Karl and Williams, 1987; Quayle et al., 1991; Peterson
95 and Easterling, 1994; Imhoff et al., 1997; Peterson et al., 1998; Hansen et al., 2001; Vose
96 et al., 2003; Menne and Williams, 2005; Mitchell and Jones, 2005; Brohan et al., 2006;
97 DeGaetano, 2006; Runnalls and Oke 2006 Reeves et al., 2007; Menne and Williams,

98 2009; Muller et al, 2012). in order to obtain homogeneous data and create reliable long-
99 term surface temperature time series. Menne et al. (2009) for the United States Historical
100 Climatology Network, Version 2 (USHCNv2), rely exclusively on detecting changes
101 within the unadjusted surface temperature data itself to identify and correct time-varying
102 non-climatic biases. Because of the unreliability of the archived metadata, some recently
103 introduced adjustment approaches, such as that described by Menne et al. (2010), are not
104 fully comprehensive, and are a tradeoff between leaving large undocumented changes
105 uncorrected and inadvertently altering true local climate signals while also failing to
106 detect and correct for other inhomogeneities such as changes in the station siting
107 environment. An example of the incompleteness of their approach is reported, as one
108 example, in Martinez et al (2012), who reported that

109 *“Significant differences in temperature trends based on the surrounding land use were*
110 *found for minimum temperature and temperature range in the 1970–2009 period*
111 *indicating that data homogenization of the USHCN temperature data did not fully remove*
112 *this influence”*

113 The incompleteness by Menne et al. (2010) in correcting for non-climatic effects and
114 non-spatially representative trends can explain the divergence in the multi-decadal
115 temperature trend diagnosed for the surface and the lower troposphere Klotzbach et al.
116 (2009, 2010)

117 Menne et al. (2010) analyzed the 1979-2008 temperature trends of stations grouped into
118 two categories based on the quality of siting. They found that a trend bias in non-
119 compliant sites relative to compliant sites is consistent with instrumentation changes that
120 occurred in the mid- and late 1980s (conversion from Cotton Region Shelter-CRS to
121 Maximum-Minimum Temperature System-MMTS). The main conclusion of their study
122 is that there is *“no evidence that the CONUS temperature trends are inflated due to poor*
123 *station siting”*.

124 In Fall et al. (2011), it was demonstrated that station exposure affects USHCNv2
125 temperatures, in particular the minimum temperatures, but showed little difference in
126 mean temperature trends. It was noted however, that there was no century scale trend
127 observed in the diurnal temperature variation.

128 In Muller et al. (2012), there has been considerable new work done to account for known
129 inhomogeneities and obtain adjusted surface temperature datasets for climate analysis
130 using the station siting metadata from Fall et al. (2011). In Muller et al. (2012), a
131 statistical analysis identified a -0.014 ± 0.028 C per century difference between well sited
132 and poorly sited weather stations identified in the Fall et al., 2011, metadata set. Muller et
133 al.,(2012), concluded, *“The absence of a statistically significant difference indicates that*
134 *these networks of stations can reliably discern temperature trends even when individual*
135 *stations have nominally poor quality rankings.”*.

136 Independent of the recent finding in Muller et al.,(2012), the National Climatic Data
137 Center (NCDC) has long recognized the need for a climate monitoring network as free as
138 possible from non-climatic trends and discontinuities and has developed the United States
139 Climate Reference Network (USCRN) to fill this need. (NOAA/NESDIS Climate
140 Reference Network-CRN, 2002). Using the method outlined by Leroy,(1999), NOAA
141 USCRN sites were selected based on the consideration of geographic location factors
142 including their regional and spatial representivity, the suitability of each site for
143 measuring long-term climate variability, and the likelihood of preserving the integrity of
144 the site and its surroundings over a long period. The method adopted from Leroy (1999)
145 was appropriate in achieving this goal, because it attempts to quantify the impacts of
146 visible microsite issues for new climatic station sites under consideration for inclusion
147 into the USCRN. The method from Leroy (1999) relies mainly on one observed value,
148 distance from visible heat sinks and heat sources to the thermometer instrumentation, to
149 quantify the station environment as being suitable for deployment of a USCRN climate
150 monitoring site. Having no other published metric by which to gauge station siting and
151 create representative metadata, the resultant siting metadata suggested by Leroy (1999)
152 derived from the Watts (2009) survey, was utilized in Menne et al.,(2010), Fall et al.,
153 (2011), and also Muller et al.,(2012). In all cases, station siting effects on mean
154 temperature trends were observed to be small. However, this was metadata derived from
155 the Leroy (1999) siting classification system, which was designed for site pre-selection,
156 rather than retroactive siting evaluation and classification.

157 The improved Leroy (2010) siting classification system, which included a method for
158 including the surface area of heat sinks and heat sources within the viewshed of
159 thermometer was endorsed by the World Meteorological Organization Commission for
160 Instruments and Methods of Observation Fifteenth session (CIMO-XV, 2010), in
161 September 2010 stating: *“The Commission agreed that the publication of the siting
162 classification as a common WMO-ISO standard would help in assessing and improving
163 the quality of data originating from WMO-owned, cosponsored and non-WMO observing
164 networks. The Commission agreed to further develop this classification as a common
165 WMO-ISO standard.”*

166 Given that the WMO has endorsed the Leroy (2010) classification system in CIMO-XV
167 (2010) as a WMO-ISO standard, it is suitable for use in re-assessing the station quality
168 issues reported by Watts (2009)., Menne et al.,(2010), Fall et al.,(2011), and Muller et
169 al.,(2012).

170 The new siting classification system proposed in Leroy (2010) and accepted by CIMO-
171 XV is similar to the Leroy (1999) system, but adds total surface area to the distance
172 measurement as an additional metric for determining station site representivity for
173 thermometers. This resulted in a dramatic and statistically significant improvement in the
174 binning of stations quality ratings as distance alone does not quantify the amount of heat
175 emitted by a source or sink within the thermometer viewshed. As an example, in Lee
176 (1995), it was demonstrated that the design of heat sinks for electronics cooling is highly
177 dependent on the total surface area available to radiate thermal energy away from the

178 surface. The greater the surface area of the heat sink, the more efficient it is at
179 exchanging heat with the fluid medium surrounding it, and in the case of this study, that
180 is the surface layer atmosphere within the thermometer viewshed. Two physical
181 processes are involved with heat sinks and sources within the thermometer viewshed;
182 mass transfer and radiative transfer. Fourier (1822) described the process of mass
183 transfer of heat, such as between a surface and a gas. This process has been observed
184 where wind transport moves heat from nearby artificial surfaces such as asphalt, concrete,
185 and buildings to nearby thermometers, which is the basis for the distance component of
186 the Leroy (1999, 2010) rating systems: to allow adequate mixing of the boundary layer
187 atmosphere, thus minimizing the mass transfer bias before reaching the thermometer. As
188 for radiative transfer, Aseada et al. (1996) reported from measurements and analysis:

189 *“At the maximum, asphalt pavement emitted an additional 150 W m^{-2} in infrared*
190 *radiation and 200 W m^{-2} in sensible transport compared to a bare soil surface. Analyses*
191 *based on a parallel layers model of the atmosphere indicated that most of the infrared*
192 *radiation from the ground was absorbed within 200 m of the lower atmosphere, affecting*
193 *air temperature near the ground.”*

194 It follows that the total amount of infrared radiation and sensible heat released by such
195 artificial surfaces is dependent on the number of square meters of surface area within the
196 thermometer viewshed, thus making the Leroy (2010) rating system, which combines
197 surface area and distance to define the station site rating, more valuable at quantifying the

198 representivity of the station site for temperature measurements than distance alone as was
199 done in Leroy (1999) and the subsequent studies that used that rating system.

200 Many USHCNv2 stations which were previously rated with the methods employed in
201 Leroy (1999) were subsequently rated differently when the Leroy (2010) method was
202 applied in this study. This simple change in the rating system accounts for the majority of
203 differences in the data and conclusions between this study and Menne et al.,(2010), Fall
204 et al.,(2011), and Muller et al.,(2012). Effectively, the lack of accounting for the surface
205 area of heat sinks and sources using Leroy (1999) methods in Menne et al (2009), Fall et
206 al.,(2010), and Muller et al.,(2012) resulted in binning errors of trends for site
207 representivity, providing what amounted to a pseudo-randomization of the station data in
208 the context of heat sinks and sources, rendering the signal for siting issues into the noise
209 bands of the data. Once the Leroy (2010) site rating system was applied, the binning error
210 was removed, and the signal demonstrating the differences in station trends between
211 siting classes became clear.

212

213 **2. Data and methods**

214 ***2.1. USHCNv2 Climate Data***

215 The USHCNv2 monthly temperature data set is described by Menne et al. (2009). The
216 raw and unadjusted data provided by NCDC has undergone the standard quality-control

217 screening for errors in recording and transcription by NCDC as part of their normal ingest
218 process but is otherwise unaltered. The intermediate (TOB) data has been adjusted for
219 changes in time of observation such that earlier observations are consistent with current
220 observational practice at each station. The fully adjusted data has been processed by the
221 algorithm described by Menne et al. (2009) to remove apparent inhomogeneities where
222 changes in the daily temperature record at a station differs significantly from neighboring
223 stations. Unlike the unadjusted and TOB data, the adjusted data is serially complete, with
224 missing monthly averages estimated through the use of data from neighboring stations.
225 The USHCNv2 station temperature data in this study is identical to the data used in Fall
226 et al. (2011), coming from the same data set.

227 *2.2. Station Site Classification*

228 We make use of the subset of USHCNv2 metadata from stations whose sites have been
229 classified by Watts (2009), gathered by the volunteers of the surfacestations.org project
230 using the USCRN site-selection classification scheme for temperature and humidity
231 measurements (NOAA/NESDIS 2002), and originally developed by Leroy (1999). For
232 Watts (2009) and Fall et al. (2011), USHCNv2 site surveys were originally performed
233 between June 2nd, 2007 and Feb 23rd, 2010. For the purpose of this study, the original
234 site rating metadata from Fall et al (2011), also used in Muller (2012), was supplemented
235 with further refinements and additional station surveys inclusive from June 15th, 2011 to
236 July 1st, 2012, followed by application of the Leroy (2010) site survey rating system to
237 both old and new surveys (Table 1) including both a distance and an area rating

238 component. Any known changes in siting characteristics after that period are ignored. A
239 total of 1065 USHCNv2 stations were surveyed, comprising 87.4% of the 1218 station
240 USHCNv2 network. Of those 1065 stations surveyed, 779 were classified per the Leroy
241 (2010) site survey rating system (Figure 1). As a rule, LeRoy (2010) is less “strict” than
242 Leroy (1999). There is a greater number of Class 1, 2, and 3 stations, and fewer Class 4
243 stations. There are, however, a greater number of Class 5 stations, as well.

244 In our urban-rural comparisons we use the Urban, Semi-Urban, Rural classifications
245 provided by NASA. We divide the continental contiguous USA into twenty-six 6-degree
246 grid boxes so that the gridding process eliminates distribution bias.

247 Because the great majority of the station surveys occurred prior to creation the of Leroy
248 (2010) site survey rating system, site surveys previously acquired and used in Fall et al.
249 (2011) and Muller et al (2012) were retroactively resurveyed, and wherever possible, had
250 additional land and aerial photography added, so that surface area measurements required
251 for the Leroy (2010) site survey rating system could be performed. In addition to station
252 ratings, the survey provided an extensive documentation composed of station
253 photographs and detailed survey forms. Because some stations used in Fall et al. (2011)
254 and Muller et al. (2012) suffered from a lack of the necessary supporting photography
255 and/or measurement required to apply the Leroy (2010) rating system, or had undergone
256 recent station moves, there is in a smaller set of station rating metadata (779 stations)

257 than used in Fall et al (2011) and Muller et al. (2012), both of which used the data set
258 containing 1007 rated stations.

259 For each site in this study, ground and or aerial photography was obtained, distance
260 measurements of visible encroachments were made, and a calculation was done to
261 determine the percentage of area within the different radii (3m, 5m, 10m, 30m, and
262 100m) surrounding the thermometer per Leroy (2010), containing heat sinks and/or heat
263 sources. The distance and area values were applied to the final rating for each station.
264 Quality control checks were routinely done to ensure that the proper station was
265 identified, that it matched descriptions in metadata provided by NCDC, that it was
266 consistent with the latitude and longitude given for the station, and that the equipment
267 seen in photography and described in survey reports matched the equipment description
268 according to NCDC metadatabase. Where discrepancy existed, interviews were
269 conducted with the station curator when possible to resolve such discrepancy and to
270 ensure the location of the thermometer in some aerial photos that had marginal resolution.
271 Where such discrepancies could not be resolved, or it was determined from photographs,
272 metadata, or curator interviews that the station had been closed or moved after 2002, and
273 prior location could not be established, that station was excluded from consideration and
274 not included in this study. Since the site metadata is either incomplete or cannot be
275 verified for those stations that were excluded, it became impossible to bin them into their
276 siting classes for use in this study. Examples of problems that caused exclusion include
277 but are not limited to; recent station moves that made a station previously identifiable

278 now unidentifiable, obscuration of the thermometer viewshed in aerial photos preventing
279 a full distance and area measurement are; low resolution aerial photography that made it
280 impossible to identify the exact location of the thermometer for measurements, no usable
281 aerial photographic coverage at all, and inability to contact the site curator for verification
282 of details not clearly visible in aerial and ground photography.

283

284 The best sites (compliant per Leroy, 2010) consist of 160 stations classified as either
285 Class 1 (48 stations) or Class 2 (112 stations) doubling the number of compliant stations
286 used in Fall et al. 2011 (80 stations), where the Leroy (1999) site survey rating system
287 was applied. The worst (non-compliant per Leroy 2010) sites, of Classes 3, (247 stations)
288 4, (277 stations) and 5 (95 stations), comprise the majority of the USHCNv2 network
289 with 619 stations at 79.5% (Table 2). The distribution of the best and poorest sites is
290 displayed in Figure 1. Because Leroy (2010) considers both Class1 and Class 2 sites to be
291 acceptably representative for temperature measurement, with no associated measurement
292 bias, these were combined into the single “compliant” group with all others, Class,
293 3,4,and 5 as the “non-compliant” group. In contradiction to Leroy (1999) and Leroy
294 (2010) publicly available review papers for Muller et al. (2012), showed they used
295 grouping of Classes 1,2,3 as compliant sites, and Classes 4&5 as non-compliant sites. In
296 addition to the lack of class binning using surface area by applying Leroy (2010) site

297 classifications, this may also have contributed to Muller et al. (2012) finding no
298 discernible trend differences between station classes.

299 As in Fall, et al (2011), Menne (2010), and Muller (2012), only the heat source/sink
300 proximity and area ratings from Leroy 2010 and are do consider ground-level vegetation
301 or shade.

302 Shade (a cooling bias) will inevitably affect poorly sited stations more than those that are
303 well sited: The poorer sited stations are often shaded by nearby structures which result in
304 their poor rating in the first place. Therefore, if anything, not accounting for shade would
305 most likely lessen the differences between the better and poorer sites rather than increase
306 them. Ground vegetation (a warming bias), on the other hand, affects the better sites,
307 particularly stations located in rural areas, rather than the poorer and urban sites.
308 Therefore, not accounting for vegetation may well lessen the differences between good
309 and bad sites rather than increase them. Therefore we can be reasonably certain that
310 excluding these factors will not bias this study in ways that will exaggerate the
311 differences between well and poorly sited stations.

312 In any event, with the resources currently available, we are unable to rate either shade or
313 ground cover adequately. Perhaps this will be addressed in a future study (including
314 factors such as terrain and altitude). We can, however, quite accurately determine heat
315 sink coverage by use of satellite and aerial imagery and in particular, that of Google Earth
316 aerial photography and its distance measurement tool.

317 **2.3. Methods of Analysis**

318 The purpose of this study is to determine whether and to what extent regional and
319 national-scale temperatures and temperature trends estimated from poorly-sited stations
320 differ from those estimated from well-sited stations, by building on what was learned
321 from Menne et al. (2010), Fall et al. (2011, and Muller et al. (2012) and by applying the
322 new Leroy (2010) rating system against the stations surveyed by Watts (2009). The
323 analysis involves aggregating USHCNv2 monthly station data into regional and national
324 averages and comparing values obtained from different population groups of stations.

325 The process is started by computing monthly anomalies relative to a 30-year baseline
326 period, in this case 1979-2008, to be consistent for comparison with previous works of
327 Menne et al. (2010), Fall et al. (2011), and Muller et al. (2012). We then average the
328 monthly anomalies across all stations in a particular Leroy (2010) class or set of classes
329 within each of the nine NCDC-defined climate regions shown in Figure 2. In Figure 2,
330 without separating any classes of stations to provide a baseline for the CONUS, the raw
331 data average of all rated stations in each region shows a positive trend ranging from
332 0.173°C/decade in the Northwest region to 0.380 °C/decade in the Southwest region, with
333 a continental United States (CONUS) gridded value of 0.231°C/decade.

334 Further investigations include separating stations by classes, and then examining the
335 effect on the trends of the difference between classes for Tmin, Tmax, and Tmean,
336 including examinations of rural and urban stations, stations at airports versus the general

337 station population, and differences in station equipment. Finally, an overall average value
338 for the (CONUS) is computed as a gridded, area-weighted mean of the regional averages
339 for each of the station siting classes and subsets of siting classes, examining rural and
340 urban, airport and non-airport stations, and equipment differences between stations using
341 Cotton Region Shelters (CRS) and Maximum-Minimum Temperature System (MMTS)
342 electronic thermometers.

343 The multiple regional analyses presented are designed to account for the spatial variations
344 of the background climate and the variable number of stations within each region, so that
345 the national analysis is not unduly influenced by data from an unrepresentative but data-
346 rich corner of the United States. Figure 3 shows station distributions by class in the
347 CONUS.

348 Menne et al. (2010) used a gridded analysis approach for the CONUS, as in our study.
349 However, compared to the Menne et al. (2010) results, as well as the Muller (2012)
350 results, both of which found very little difference between well sited and poorly sited
351 stations in the CONUS, our gridded results based on the Leroy (2010) site ratings yields
352 national trend values for all well sited (compliant classes 1&2) stations of 0.155°C
353 /decade trend, while the poorly sited (non-compliant classes 3,4,5) stations show a
354 $0.248^{\circ}\text{C}/\text{decade}$ trend. Even greater and more significant differences are seen in the
355 regional, environmental, class, and station type specific analyses we completed.

356 The results of the analysis suggest that these differences may be due specifically to the
357 station siting characteristics or be due to other characteristics that covary with station
358 siting, such as instrument type. Siting differences directly affect temperature trends if the
359 poor siting compromises trend measurements or if changes in siting have led to artificial
360 discontinuities. In what follows, to the extent that significant differences are found
361 among classes, the well sited stations will be assumed to have more accurate
362 measurements of temperature and temperature trends than poorly sited stations.

363 **3. Results**

364 *3.1. Regional trend analysis*

365 Figure 4 shows regional decadal trends in the CONUS for 1979-2008 as calculated with
366 USHCNv2 data from all stations and all classes of stations. Clear statistically significant
367 differences between Class 1&2 (compliant) and Class 3,4,5 (non-compliant) stations are
368 indicated in the bar graphs. Without exception, in each region, compliant stations have a
369 lower decadal scale trend than non-compliant stations. In the most striking example of
370 this difference, one region, the SE, a slight negative trend exists for compliant stations of
371 $-0.02^{\circ}\text{C}/\text{decade}$ while non-compliant stations have a positive trend of $0.223^{\circ}\text{C}/\text{decade}$.
372 For the entire CONUS, the average of all regions shows the compliant Class 1&2 stations
373 have a decadal scale trend of $0.155^{\circ}\text{C}/\text{decade}$ while non-compliant Class 3,4,5 stations
374 have a $0.248^{\circ}\text{C}/\text{decade}$ trend. Fully adjusted USHCNv2 data for the entire CONUS (all
375 classes of stations) has a $0.309^{\circ}\text{C}/\text{decade}$ trend.

376 When USHCNv2 stations located at airports are considered, such differences between
377 poor and well sited stations were observed to grow even larger. Figure 5 shows that when
378 airport stations are excluded for the CONUS analysis, compliant stations have a
379 0.124°C/decade trend while non-compliant stations are almost double the trend at
380 0.246°C/decade. The difference in the SE region grew even larger with compliant stations
381 having a -0.131°C/decade trend while non-compliant stations have a 0.219 °C/decade
382 trend for a difference of 0.350°C/decade. Again, for all classes of stations, in all nine
383 regions considered, compliant stations have a lower decadal scale trend than non-
384 compliant stations. Conversely when only USHCNv2 stations sited at airports are
385 considered these differences are not as strong as seen in Figure 6. Part of the differences
386 may be attributed to the way equipment is deployed, sited, and maintained at airports.
387 May airports, due to the weather stations being placed on grassy areas in between
388 runways, are rated as “compliant” by both Leroy (1999) and Leroy (2010) rating
389 systems. However, the data from airport stations is logged with aviation monitoring
390 systems known as ASOS, from OFCM, (1994), and it has been demonstrated by error
391 reports, such as in the Senate testimony of Snowe (1998) stating “*The ASOS systems in*
392 *Maine have been very unreliable, The station in Houlton recorded more than 1400*
393 *mistakes in one year*” that the ASOS system has significant reporting problems
394 particularly with the HO-83 hygrothermometer used in the ASOS system. Problems with
395 temperature biases in the HO-83 hygrothermomter were first reported in Gall et al.
396 (1992) in connection with large errors in the Tucson ASOS station. They report that in
397 Tucson, an all-time maximum temperature record was set of 114°F, along with numerous

398 daily records during the time this ASOS station was in use, many of these records having
399 been set while no other records were broken within 1000 miles of Tucson.

400 In response to issues raised by Gall et al. (1992), ASOS hygrothermometers were
401 redesigned. In Jones and Young, 1995 they reported:

402 *“Examination of differences between the two instruments found that the original version*
403 *of the HO-83 read approximately 0.6 deg C warmer than the redesigned instrument.*
404 *Significant changes in the differences between the two instruments were noted between*
405 *winter and summer. It is suggested that for stations with climatology similar to the ones*
406 *used in this study monthly mean temperatures reported by the original version of the HO-*
407 *83 be adjusted by adding -0.4 deg C to June, July August and Sept observations and by*
408 *adding -0.7 deg C for the remainder of the year.”*

409

410 Karl et al. (1995) noted issues with the HO-83 hygrothermometer in Chicago in relation
411 to reporting temperatures during a summer heat wave. In Karl and Knight, (1996) it was
412 further discussed:

413 *“Karl et al. (1995) show that, on average, the HO-83 increased the maximum*
414 *temperature by about 0.5°C relative to the HO-63 instrument and also increased the*
415 *minimum but only by 0.1°C. Much larger effects have been noted in Tucson, for example*
416 *(Gall et al. 1992), and Jones and Young (1995) also find a consistent positive bias at*

417 *several stations they examined in the southern and central plains. This suggests that the*
418 *trends of maximum T in Chicago are biased warm not only due to increased urbanization*
419 *but by the introduction of the HO-83 instrument in 1986.”*

420 In the Snowe, 1998 testimony before the Senate, concerns over ASOS station data
421 reliability were great enough to cause this amendment to be added to the bill being
422 discussed:

423 *“The administrator of the Federal Aviation Administration shall not terminate human*
424 *weather observers for Automated Surface Observation System Stations until (1) The*
425 *secretary of transportation determines that the system provides consistent reporting of*
426 *changing meteorological conditions and notifies the Congress in writing of that*
427 *determination; and (2) 60 days have passed since the report was submitted to the*
428 *Congress.”*

429 The issues of reliability, combined with known historical problems with airport ASOS
430 station instrumentation introducing positive temperature biases into the record, taken
431 along with our findings that airport stations add a warm bias to our own siting analysis,
432 suggests that airport weather stations utilizing ASOS may produce artificially high and
433 uncorrected temperature records, and thus may not be suitable for inclusion into long
434 term climate data without detailed retroactive examinations of instrument maintenance
435 and calibration records and corrections applied to the daily data.

436 In the data shown in this study, airport stations clearly have a less definitive siting bias
437 signal. This should not be taken as a suggestion that airport stations have better siting
438 overall, as Watts 2009 demonstrated that many airport ASOS stations were near runways
439 and tarmac, but that due to errors and inconsistencies in the ASOS temperature
440 instrumentation, the temperature data may not accurately reflect the station siting bias
441 issues due to being swamped by the larger errors of ASOS instrumentation.

442 Further analysis of the USHCNv2 data, taking into account rural stations, and excluding
443 airports demonstrates even stronger bias magnitudes between compliant and non-
444 compliant stations. In figure 7, the CONUS Class 1&2 trend for rural stations without
445 airports is observed to be an even lower value at $0.108^{\circ}\text{C}/\text{decade}$, with Class 3,4,5 non-
446 compliant stations having more than double that value at $0.228^{\circ}\text{C}/\text{decade}$. The Class
447 difference in the SE region is -0.100 for compliant stations, with non-compliant stations
448 at $0.157^{\circ}\text{C}/\text{decade}$ for a difference of $0.257^{\circ}\text{C}/\text{decade}$ between compliant and non-
449 compliant stations. This is in stark contrast to figure 6, using airport stations only, where
450 the SE region negative shows a positive trend of $0.181^{\circ}\text{C}/\text{decade}$. These findings further
451 suggest that airports are not representative recorders of regional climatic trends.

452 The regional examination of classes that demonstrate the lowest decadal scale trend of all
453 subsets, that of rural MMTS stations, excluding airports, reveals some of the most
454 significant differences in siting biases between compliant and non-compliant stations.

455 Figure 8 shows that rural MMTS stations, excluding small rural airports that are
456 sometimes equipped with MMTS equipment, have the lowest decadal trends of all classes
457 and subsets of stations. The difference between compliant at $-0.207^{\circ}\text{C}/\text{decade}$ and non-
458 compliant stations at $0.113^{\circ}\text{C}/\text{decade}$, in the SE region grows to $0.310^{\circ}\text{C}/\text{decade}$, with
459 two additional regions, ENC and WNC, now showing negative decadal scale trends of -
460 $0.125^{\circ}\text{C}/\text{decade}$ and $-0.055^{\circ}\text{C}/\text{decade}$ respectively with each showing large differences
461 with their non-compliant station counterparts. The ENC region now records the largest
462 regional scale difference between compliant and non-compliant stations in the entire
463 USHCNv2 dataset at $0.365^{\circ}\text{C}/\text{decade}$.

464 The gridded average of all compliant Class 1&2 stations in the CONUS is only slightly
465 above zero at $0.032^{\circ}\text{C}/\text{decade}$, while Class 3,4,5 non-compliant stations have a trend
466 value of $0.212^{\circ}\text{C}/\text{decade}$, a value nearly seven times larger. NOAA adjusted data, for all
467 classes of rural non-airport stations has a value of $0.300^{\circ}\text{C}/\text{decade}$ nearly ten times larger
468 than raw data from the compliant stations.

469 These large differences demonstrated between regional and CONUS trends accomplished
470 by removal of airports and choosing the rural subset of stations to remove any potential
471 urbanization effects suggests that rural MMTS stations not situated at airports may have
472 the best representivity of all stations in the USHCNv2.

473 ***3.2. Temperature bias analysis by site classification and equipment type.***

474 Significant decadal trend differences were observed between compliant CRS stations and
475 compliant MMTS stations, with MMTS stations generally being cooler, confirming what
476 was observed in Menne et al (2010). But, this effect is swamped by the larger effect of
477 siting bias in the non-compliant stations, particularly in the trends of the Tmin,
478 suggesting a sensitivity to heat sinks within the thermometer viewshed, which is the basis
479 of the Leroy classification system. In Watts 2009 it was observed that with the
480 introduction of the MMTS electronic thermometers in the NOAA COOP network starting
481 in 1983, difficulties in trenching past obstacles (sidewalks, driveways, roadways, etc.)
482 due to cabling, placed MMTS thermometers closer to offices and domiciles of the COOP
483 observers. Our findings confirm this to have a real effect across all classes, with non-
484 compliant MMTS stations having warmer trends. Additionally, it was observed that the
485 Tmax trends of compliant CRS stations was significantly higher, suggesting that
486 maintenance issues, such as paint deterioration over time and differences as discussed in
487 Watts (2009), and seen in figure 9 darkened the wood, and lowered the surface albedo of
488 the CRS equipped stations, making them more susceptible to solar insolation effects near
489 the time of Tmax.

490 *3.2.1 Comparison by site classifications*

491 For the CONUS, we compare the average temperature of each Leroy (2010) class with
492 the average of each of the other classes within each grid. This results in these baseline
493 comparisons.

494 Figure10: Class 4 Comparisons (with each other Class ratings within each grid, then all
495 results are averaged). Figure11: Class 3 Comparisons Figure12: Class 1 & 2
496 Comparisons.

497 The results are listed in order of robustness: There are more Class 4 stations than any
498 other rating, so the Class 4 comparisons are examined first, followed by Class 2, then
499 Class 1&2 stations. There is insufficient CONUS grid box coverage of Class 5 stations to
500 use them as a baseline for a gridded comparison.

501

502 In figure 10, the columns represent the following measurements:

503 Class 1&2 compared with Class 4 within each grid box. The resulting differences for
504 each Class 1&2 station are then averaged. Class 3 is compared with Class 4,
505 Class 4 compared with Class 4 (the baseline, so the result will be 0.), Class 5 compared
506 with Class 4 and all lower classes.

507

508 Note that the well sited stations (Class 1 & 2) show a substantial difference in the Tmean
509 trend compared with poorly sited stations. As reported in Fall et al. (2010), the difference
510 is most significant in terms of Tmin. Tmax shows a very similar pattern to Tmin,
511 although the differences are smaller.

512

513 Note also that while all classes of stations higher than Class 1&2 demonstrate higher
514 trends than nearby Class 3&4 stations, Class 5 stations appear to be overwhelmed with

515 waste heat which appears to be masking the trend. Note also that for Class 5 stations we
516 observe a reversal of the decadal trend for Tmax and Tmin compared to all other classes
517 of stations. We posit that this reflects the thermal latency of nearby heat sinks and
518 sources for Class 5 stations that are applying a dampening effect on the surface layer
519 thermometer, limiting its sensitivity to the surface layer atmosphere diurnal range. Such
520 an effect would be demonstrated by a reversal of trends as heat sinks in the immediate
521 proximity of Class 5 stations, such as concrete, asphalt, and buildings, dump stored heat
522 from daytime solar insolation into the nighttime Tmin period, buffering the minimum
523 temperature. Conversely, during the day, a large area of nearby heat sinks can act as solar
524 radiation absorbers, buffering the ability of the local surface atmosphere to reach a
525 representative Tmax compared to nearby stations. The overall result would be higher
526 absolute temperatures, but, at the same time, lower temperature trends.

527

528 Both of these observations

- 529 1.) Poorly sited stations show greater trend results than well sited stations.
- 530 2.) Class 5 stations show smaller increases in trend results, which effect is possibly
531 due to overwhelming by waste heat.

532 Will be either supported or disputed by the many various comparisons which follow.

533 In Figure 11, gridded with a Class 3 baseline, we see the same pattern as in Figure 10
534 observing that Tmean trend is indisputably higher for well sited stations than for poorly
535 sited stations. Tmin shows the greatest differences between station classes, while Tmax
536 shows the smallest.

537

538 In figure 12, Class 1 & 2 gridded comparisons, we observe that all remaining non-
539 compliant classes of stations, and the non-compliant grouped class 3\4\5 have higher
540 decadal trends than the compliant stations of Class 1&2. As in figures 10 and 11, Tmin
541 shows the greatest differences between station classes, while Tmax shows the smallest.

542

543

544 *3.2.2 Equipment Comparisons*

545

546 We next examine whether these differences are an artifact of equipment or whether they
547 hold true for both MMTS and CRS stations.

548

549 The USHCNv2 ground level photographic survey of Watts (2009), plus subsequent re-
550 surveys and new station surveys for the purpose of this study reveal that the great
551 majority of USHCNv2 stations consists of either CRS or MMTS equipped stations. There
552 is a smaller number of airport based ASOS/AWOS stations and a very small population
553 of non-standard equipment, such as consumer grade weather stations approved for use at
554 a private station by the local National Weather Service Office COOP manager. The
555 population of USHCN stations equipped with consumer grade instrumentation is too
556 small to provide a statistically meaningful comparison and is ignored for the purposes of
557 this study.

558

559 For purposes of evaluation in this study, we are classifying as MMTS any station which
560 converted to MMTS prior to 1995, (and/or has been MMTS for a plurality of the study
561 period), and the same applies for ASOS/AWOS. We classify as CRS any station which
562 converted to MMTS (or other non-CRS equipment) in 1995 or later.

563

564 Comparing equipment alone, we observe in figure 13 that ASOS stations equipped with
565 electronic hygrometers, such as the problematic HO-83, have the highest raw
566 (ungridded) Tmean trends at 0.277 °C/decade, followed by CRS equipped stations at
567 0.265 °C/decade, and MMTS equipped stations at 0.192 °C/Decade. MMTS equipped
568 stations are observed to have significantly lower Tmean trends than the two other
569 equipment types.

570

571 This is of particular importance, considering that ASOS/AWOS systems are by far the
572 better sited systems. 57% of rated ASOS/AWOS systems are Class 1&2, as opposed to
573 23% of CRS stations and a mere 14% of MMTS.

574

575 In order to demonstrate that these differences are a result of equipment bias and not
576 actually a sign that poorly sited stations tend to show a smaller Tmean warming trend, we
577 examine Class 1&2, plus Class 3, 4, and 5 stations for MMTS and CRS equipped
578 stations. There is not a significant enough population of ASOS equipped USHCN stations
579 for a statistically significant gridded comparison and that comparison is not done for that
580 reason.

581

582 The following set of figures shows the gridded comparisons of each, calculated using the
583 same method as for figures 10 through 12.

584

585 In figures 14 and 15 above, showing gridded comparison of CRS and MMTS equipped
586 stations, respectively, we observe the same basic Tmean pattern for both sets of
587 equipment. The only difference is that Class 5 CRS stations have a lower comparative
588 difference than do Class 5 MMTS stations. In the case of both MMTS and CRS, well
589 sited (Class 1&2) stations show a significantly smaller trend compared with poorly sited
590 (Class 3,4,5) stations. Furthermore, in the case of MMTS stations (the most prevalent
591 station type), the difference is about twice as great as for the CRS stations.

592

593 Another question that arises is whether microsite differences are masked by mesosite
594 considerations of rural vs. urban environment.

595

596 To examine this question, we first look at overall mesosite trends for all stations, and then
597 for Class 1&2, Class 3, Class 4, and Class 5 stations. For purposes of mesosite
598 classification, we use the terms provided by NASA Goddard Institute for Space Studies
599 for their GISTEMP database: Urban, Semi-Urban, Rural. Shown in Figure 16 is a six
600 panel comparison showing comparisons for Urban, Semi-Urban, Rural stations with raw
601 and adjusted data for all stations, raw and adjusted data for Class 1&2 stations, raw and

602 adjusted data for Class 3,4,5 stations, raw and adjusted data for Class 3 stations, raw and
603 adjusted data for Class 4 stations, and raw and adjusted data for Class 5 stations.

604

605 We observe that for the Tmin value identified in Fall et al. (2011) as being the most
606 affected by siting issues, significant differences exist in Tmin raw data between urban
607 and rural compliant Class 1&2 stations and between urban and rural non-compliant Class
608 3,4,5 stations. Rural Class 1&2 stations have a Tmin raw trend of $0.127^{\circ}\text{C}/\text{decade}$ while
609 urban stations have a Tmin raw trend of $0.278^{\circ}\text{C}/\text{decade}$. Rural Class 3,4,5 stations have
610 a Tmin raw trend of $0.278^{\circ}\text{C}/\text{decade}$, while urban Class 3,4,5 stations have a Tmin raw
611 trend of $0.420^{\circ}\text{C}/\text{decade}$, the highest in the dataset. This suggests that no matter what the
612 microsite level issues of siting, urban sited stations are proportionately more affected in
613 the Tmin by the mesoscale heat sinks and sources that make up urbanity. When looking
614 at the Tmin USHCNv2 adjusted data for rural stations, we observe that it is adjusted
615 higher in value, from $0.127^{\circ}\text{C}/\text{decade}$ to $0.249^{\circ}\text{C}/\text{decade}$, effectively doubling the trend,
616 and with that adjustment very nearly matches the rural Class 3,4,5 Tmin adjusted value of
617 $0.265^{\circ}\text{C}/\text{decade}$. This suggests that USHCNv2 data homogenization methods are
618 erroneously adjusting pristine Tmin data from rural Class 1&2 stations to be similar to
619 that of rural Class 3,4,5 stations, effectively eliminating the preferred station
620 representivity defined by Leroy (2010) .

621

622 In order to demonstrate that microsite considerations prevail, regardless of mesosite
623 condition, we examine Class 1&2, Class 3, Class 4, and Class 5 averages (ungridded) for
624 Rural, Semi-urban, and Urban environments in figure 17

625

626

627 This confirms that the microsite conditions we are seeing remain consistent in Rural and
628 Semi-urban settings. In urban settings (10% of all stations), the differences are somewhat
629 masked, especially in the case of Class 5 stations.

630

631 This is consistent with the hypothesis that artificially heated areas tend to overwhelm
632 microsite considerations after a certain point. Note that urban Class 5 stations have the
633 lowest trend, and that rural Class 4 stations have a lower trend than urban Class 4 stations
634 as they are beginning to be overwhelmed by heat sink/source effects as well. This is
635 further supported by the observation that the behavior of Class 5 stations in non-urban
636 settings parallels the behavior of Class 4 stations in urban settings.

637

638

639 ***3.2.3 Discussion of Adjustments***

640 Finally, just to confirm our overall findings, we present in figure 18 the USHCNv2 raw
641 and adjusted gridded average for all stations. We do this by simply averaging each Class
642 of station within each of our 26 grid boxes seen in figure 19 and then we average all the

643 boxes for each Class of station. This removes the distribution bias, and is standard
644 procedure for calculating temperature trends.

645

646 For all stations, T_{mean} trends are adjusted upwards from 0.23 °C per decade to 0.31°C
647 per decade, an increase of 35%.

648

649 One will note that the adjusted T_{mean} trends “correct” the inequities caused by microsite
650 quality – not by adjusting the poorly sited station trends down, to match the well sited
651 stations, but by adjusting the well sited station trends upward by 92% to match the poorly
652 sited stations. The poorly sited stations are adjusted warmer by 23%, as well.

653

654 After these adjustments, T_{mean} trends from poorly and well sited stations match almost
655 exactly. This suggests that much of the representivity for well sited stations defined by
656 Leroy (2010) are being discarded in adjustment processes.

657

658 In figure 20, the differences in regional and gridded CONUS decadal scale trends
659 between all compliant, all non-compliant, and final NOAA USHCNv2 adjusted data for
660 the CONUS are illustrated. The compliant thermometers (Class 1&2) have a trend value
661 of 0.155°C/decade, the non-compliant thermometers (Class 3,4,5) have a trend value of
662 0.248°C/decade, and the NOAA final adjusted USHCNv2 data have a trend value of
663 0.309°C/decade, nearly double that of all compliant thermometers in the CONUS.

664 This disparity suggests that a combination of siting issues and adjustments are creating a
665 spurious doubling of the U.S. surface temperature record for the 30 year period of this
666 study. When rural, non-airport stations are considered, the CONUS trend is almost one
667 third that of the NOAA adjusted record.

668

669 ***3.2.4 Statistical Significance Testing***

670 In order to separately assess the effects of ratings, urbanization, equipment, max-min and
671 region, a random effects model was constructed using the R-package lme4 by Pinheiro et
672 al., (2012) as follows:

673 (1) $\text{trend} \sim (1|\text{ratings}) + (1|\text{Type}) + (1|\text{equipment}) + (1|\text{max}) + (1|\text{Grid})$

674 where ratings is a factor with two classes: “compliant”= Class 1-2 and “non-compliant”=
675 Class 3-5; Type is a factor for urbanization with three classes: R(ural); S(mall); U(rban);
676 equipment is a factor with three classes: MMTS, CRS, and ASOS, max is a factor with
677 three classes: max, min, mean; and Grid is a factor with 26 classes each representing a
678 geographic region.

679

680 The base model considered the network of 779 stations with valid metadata (as defined
681 above), less four stations with “other” equipment, reducing the base network slightly to
682 775 stations. Trends were calculated using “raw” USHCN v2 data.

683 The base model was compared to random effects models leaving each random effect out
684 one by one using an anova test. Each random effect was highly significant as summarized
685 in table 3.

686 The difference between trends for “compliant” and “non-compliant” stations was
687 $0.105^{\circ}\text{C}/\text{decade}$; between rural and urban stations was $0.066^{\circ}\text{C}/\text{decade}$; between min and
688 max measurements was $0.090^{\circ}\text{C}/\text{decade}$ (max has lower trend) and between non-MMTS
689 and MMTS approximately $0.06^{\circ}\text{C}/\text{decade}$ (MMTS cooler), as shown in the figure 21

690 When a similar analysis was carried out on USHCN v2 adjusted data, the random effects
691 for rating urbanization and equipment were completely eliminated; none were statistically
692 significant, as seen in Figure 22. The sign of the max-min random effects was reversed.
693 The fixed effect for adjusted data was $0.31^{\circ}\text{C}/\text{decade}$ (as compared to $0.25^{\circ}\text{C}/\text{decade}$ for
694 raw data.)

695

696 Our interpretation of these results is that the USCHNv2 adjustment method from Menne
697 et al (2009) is over-homogenizing the data and, in the process, removing statistically
698 significant and important information. Because of the interaction between max-min,
699 urbanization and rating, the following variation of the above was used to illustrate the
700 interaction (see Figure 23):

701

702 (1) trend~(1|max:ratings:Type)+(1|equipment)+(1|Grid)

703

704 The left panel shows trends for Class 1-2 (“compliant”) stations by urbanization class for
705 trends for max, min and mean temperatures. The right panel shows the same information
706 for Class 3-5 (non-compliant) stations. The trends for maximum and minimum
707 temperatures for compliant stations are virtually identical for each of the three
708 urbanization classes, with a difference of about 0.092°C/decade between rural and urban
709 stations. In contrast, non-compliant stations show a dramatic difference between trends
710 of maximum and minimum temperatures of approximately 0.14°C/decade, in accordance
711 with previous observations.

712

713 ***3.2.5 Reconciliation to NOAA and NASA GISS***

714 The trend (mean temperatures) for “compliant” rural stations is 0.16°C/decade,
715 substantially less than corresponding trends for the corresponding periods for the
716 continental U.S. as calculated by NOAA (0.31°C/decade) and by GISS (0.31°C/decade).
717 These values are identical to the fixed effect using adjusted USHCN data, also
718 0.31°C/decade as noted above. Both NOAA and GISS indices use adjusted USHCN data
719 in their calculations. Both NOAA and GISS estimates more or less correspond to trends
720 from non-compliant stations. Berkeley (BEST), Muller et al. (2012) adjustment

721 methodology is substantially similar to USHCN adjustment methodology and
722 accordingly yields almost identical results to NOAA.

723

724 GISS formerly (prior to the present version) ran noticeably cooler in the continental U.S.
725 than NOAA (or CRU). This was because their prior methodology did not use USHCN
726 adjusted data; GISS instead established trends from a network of “rural” stations (as
727 defined by nightlights) using less processed USHCN data. This method (as noted in
728 online discussions at the time) yielded trends more similar to that from “compliant”
729 stations in the surface stations study. GISS’ adoption of USHCN adjusted data therefore
730 appears to be a retrogression in their analysis.

731

732 Within “compliant” stations, the effect of urbanization is as expected and ranges from
733 0.11 to 0.14°C/decade. Similarly, the effect of ratings on rural stations is directionally as
734 expected at the outset of the surface stations project but with a marked interaction with
735 max-min: the effect of ratings is much stronger with minimum temperatures
736 (0.15°C/decade) than for maximum temperatures (only 0.03°C/decade), in line with the
737 emphasis of Christy et al (2008) on maximum temperatures as an indicator.

738

739 By way of comparison, the University of Alabama Huntsville (UAH) Lower Troposphere
740 CONUS trend over this period is 0.25°C/decade and Remote Sensing Systems (RSS) has
741 0.23°C/decade, the average being 0.24°C/decade. This provides an upper bound for the
742 surface temperature since the upper air is supposed to have larger trends than the surface
743 (e.g. see Klotzbach et al (2011)). Therefore, the surface temperatures should display some
744 fraction of that 0.24°C/decade trend. Depending on the amplification factor used, which
745 for some models ranges from 1.1 to 1.4, the surface trend would calculate to be in the
746 range of 0.17 to 0.22, which is close to the 0.155°C/decade trend seen in the compliant
747 Class 1&2 stations.

748 **4. Discussion and Conclusions**

749 The analysis demonstrates clearly that siting quality matters. Well sited stations
750 consistently show a significantly cooler trend than poorly sited stations, no matter which
751 class of station is used for a baseline, and also when using no baseline at all.

752 Statistically significant differences between compliant and non-compliant stations exist,
753 as well as urban and rural stations. We have demonstrated evidence that USCHNv2
754 adjustments are over-homogenizing the data and, in the process, removing statistically
755 significant and important information.

756 It is demonstrated that stations with poor microsite (Class 3, 4, 5) ratings have
757 significantly higher warming trends than well sited stations (Class 1, 2): This is true for,

758 in all nine geographical areas of all five data samples. The odds of this result having
759 occurred randomly are quite small.

760 It is demonstrated that stations with poor mesosite (airports and urbanized areas) show an
761 increase in temperature trends of both well and poorly microsited stations, alike. Over a
762 third of all stations are located in a poor mesosite environment. This is extremely
763 unrepresentative of the topography the stations purport to represent. Poor mesosite has
764 its greatest effect on Class 1, 2 stations (over 40% spurious exaggeration of trend), as so
765 many of them are located in airports.

766

767 Well sited stations, using a localized Class 4 (the most common class) baseline show a
768 trend of 0.09°C per decade lower than poorly sited stations for raw Tmean trends. The
769 Raw Tmean trend for well sited stations is 0.14°C per decade lower than adjusted Tmean
770 trend for poorly sited stations.

771 Not only does the NOAA USCHNv2 adjustment process fail to adjust poorly sited
772 stations downward to match the well sited stations, but actually adjusts the well sited
773 stations upwards to match the poorly sited stations.

774 In addition to this, it is demonstrated that urban sites warm more rapidly than semi-urban
775 sites, which in turn warm more rapidly than rural sites. Since a disproportionate

776 percentage of stations are urban (10%) and semi-urban (25%) when compared with the
777 actual topography of the U.S., this further exaggerates Tmean trends.

778 NOAA adjustments procedure fails to address these issues. Instead, poorly sited station
779 trends are adjusted sharply upward (not downward), and well sited stations are adjusted
780 upward to match the already-adjusted poor stations. Well sited rural stations show a
781 warming nearly three times greater after NOAA adjustment is applied. We have shown
782 that the site-classification value is a clear factor in the calculation of the trend magnitude.
783 We are investigating other factors such as Time-Of-Observation changes which for the
784 adjusted USHCNv2 is the dominant adjustment factor during 1979-2008.

785 Future investigations could test to see if the siting issue is broader. Given that USHCN
786 stations overlap and are a part of the GHCN, the siting issue should be examined for all
787 of the GHCN and BEST sites used in Muller (2012).

788 Class 5 sites, even more so than Class 3 and 4, have a multitude of major non-climatic
789 effects and local microclimate which result making it difficult, if not impossible, to
790 explain the behavior of its trend signal. This includes shading from buildings and trees,
791 cooling of dry bulb temperatures by evaporation from grasses around the site in otherwise
792 dry vegetation areas, their location on roof tops with more wind ventilation, etc. There is
793 also the likelihood of more evaporation of water vapor into the air such as from water
794 treatment plants and non-representative nearby vegetation such as lawns and shrubs.

795

796 In future analyses, the assessment of moist enthalpy trends could provide more insight.
797 As shown in Pielke et al (2004), Davey et al (2006), Fall et al (2010), and Peterson et al
798 (2011) concurrent long term trends in the absolute humidity of the surface air make the
799 interpretation of the dry bulb temperature trend more difficult. However, it is the
800 combined effect of dry bulb temperature and absolute humidity that are the true measure
801 of heating and cooling.

802

803 As shown in Figure 11 in Pielke et al (2007), for example, the hottest time of the day in
804 the dry bulb temperature is not the hottest in the physics unit of heat (i.e. Joules per kg of
805 air). It could be that in the urban area the added water vapor from those sites could be
806 resulting in really warm conditions in terms of Joules per kg, but the dry bulb temperature
807 is suppressed. This certainly could be true around sites at water treatment plants, of which
808 a significant population exists in the USHCN.

809 There is the further issue of equipment inhomogeneity. Modern MMTS sensors show a
810 significantly lower warming trend than the obsolete CRS shelters. Yet rather than
811 lowering the trends of CRS stations, the trends of MMTS stations are sharply adjusted
812 upwards. It is difficult, however, to be certain of the true effect thanks to the relatively
813 small number of Class 1,2, rural, non-airport stations.

814 Taken *in toto*, these factors identified in this study have led to a spurious doubling of U.S.
815 Tmean trends from 1979 - 2008.

816

817 **Acknowledgments**

818 The authors wish to acknowledge the many cooperative observers that unselfishly carry
819 out COOP observations which are the backbone of climate monitoring. We also
820 acknowledge the many volunteers that made the surfacestations.org project possible with
821 their personal time and efforts in gathering the nationwide survey. Special thanks are
822 given to these prominent volunteers who expended special efforts and expertise in
823 metadata collection and collation: Gary Boden, Kristen Byrnes, Don and Liz Healy, Eric
824 Gamberg, John Goetz, Don Kostuch, Steven McIntyre, John Slayton, Ted Semon, Russell
825 and Ellen Steele, and Barry Wise. Acknowledgement is given to former California State
826 Climatologist James Goodridge, who was inspirational with surveys he made of
827 California COOP stations during his tenure.

828 Special thanks are given to Dr. Roger Pielke Sr. for inspiration, advice, and technical
829 proofreading of this study.

830

831 **References**

832 Asaeda T., Ca V. T., Wake A., 1996; Heat storage of pavement and its effect on the lower
833 atmosphere. *Atmospheric environment*. CUTEST '92 : conference on the urban thermal
834 environmental studies in Tohwa, Fukuoka, JAPON (07/09/1992) 1996, vol. 30, n° 3 (185
835 p.) (30 ref.), pp. 413-427

836 Baker, D. G., 1975: Effect of observation time on mean temperature estimation. *Journal*
837 *of Applied Meteorology*, **14**, 471-476.

838 Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones, 2006: Uncertainty
839 estimates in regional and global observed temperature changes: A new dataset from 1850,
840 *Journal of Geophysical Research*, **111**, D12106, doi:10.1029/2005JD006548.

841 Christy, J. R., 2002: When was the hottest summer? A State Climatologist struggles for
842 an answer. *Bulletin of American Meteorological Society*, **83**, 723-734.

843 Christy, J. R., W. Norris, K. Redmond, and K. Gallo, 2006: Methodology and results of
844 calculating central California surface temperature trends: Evidence of a human induced
845 climate change. *Journal of Climate*, **19**, 548-563.

846 Christy, J.R., W.B. Norris, and R.T. McNider, 2009: Surface Temperature Variations in
847 East Africa and Possible Causes. *Journal of Climate*, **22**, 3342–3356.

848 Davey, C.A., and R.A. Pielke Sr., 2005: Microclimate exposures of surface-based
849 weather stations - implications for the assessment of long-term temperature trends.
850 *Bulletin of American Meteorological Society*, **86**, 497-504.

851 Davey, C.A., R.A. Pielke Sr., and K.P. Gallo, 2006: Differences between near-surface
852 equivalent temperature and temperature trends for the eastern United States - Equivalent
853 temperature as an alternative measure of heat content. *Global and Planetary Change*, **54**,
854 19–32.

855 DeGaetano, A.T., 2006: Attributes of several methods for detecting discontinuities in
856 mean temperature series. *Journal of Climate*, **19**, 838-853.

857 Easterling, D. R., T. R. Karl, E.H. Mason, P. Y. Hughes, and D. P. Bowman, 1996:
858 *United States Historical Climatology Network (U.S. HCN) Monthly Temperature and*
859 *Precipitation Data*. ORNL/CDIAC-87, NDP-019/R3. Carbon Dioxide Information

860 Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak
861 Ridge, Tennessee.

862 Easterling, D. R., B. Horton, P. D. Jones, T. C. Peterson, T. R. Karl, D. E. Parker, M. J.
863 Salinger, V. Razuvayev, N. Plummer, P. Jamason, C. K. Folland, 1997: Maximum and
864 minimum temperature trends for the globe. *Science*, **277**, 364-367.

865 Fall, S., D. Niyogi, R. A. Pielke Sr., A. Gluhovsky, E. Kalnay and G. Rochon, 2009:
866 Impacts of land use land cover on temperature trends over the continental United States:
867 assessment using the North American Regional Reanalysis, *International Journal of*
868 *Climatology*, DOI: 10.1002/joc.1996.

869 Fall, S., N. Diffenbaugh, D. Niyogi, R.A. Pielke Sr., and G. Rochon, 2010: Temperature
870 and equivalent temperature over the United States (1979 – 2005). *Int. J. Climatol.*, DOI:
871 10.1002/joc.2094.

872 Fall, S., Watts, A., Nielsen-Gammon, J. Jones, E. Niyogi, D. Christy, J. and Pielke, R.A.
873 Sr., 2011, Analysis of the impacts of station exposure on the U.S. Historical Climatology
874 Network temperatures and temperature trends, *Journal of Geophysical Research*, **116**,
875 D14120, doi:10.1029/2010JD015146, 2011

876 Fourier, J. B., 1822, *Theorie analytique de la chaleur*, Paris; Freeman, A., 1955,
877 translation, Dover Publications, Inc, NY. Available online at:
878 http://www.math.umass.edu/~lr7q/ps_files/finalfourier.pdf

879 Gall, R, K. Young, R. Schotland, and J. Schmitz , 1992. *The Recent Maximum*
880 *Temperature Anomalies in Tueson: Are They Real or an Instrumental Problem?* *Journal*
881 *of Climate* Volume 5, Issue 6 (June 1992) pp. 657-665

882 Hansen, J. E., R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson,
883 and T. Karl, 2001: A closer look at United States and global surface temperature change,
884 *Journal of Geophysical Research*, **106**, 23947–23963.

885 Hubbard K.G., and X. Lin, 2006. Reexamination of instrument change effects in the U.S.
886 Historical Climatology Network. *Geophysical Research Letters* **33**: L15710, DOI:
887 10.1029/2006GL027069.

888 Imhoff, M.L., W. T. Lawrence, D. C. Stutzer, and C. D. Elvidge, 1997. A Technique for
889 Using Composite DM SP/OLS "City Lights" Satellite Data to Map Urban Area. *Remote*
890 *Sensing of Environment* **61**, 361-370.

891 Jones, Cyrus G. and Young, Kenneth C., An Investigation of Temperature Discontinuities
892 Introduced by the Installation of the HO-83 Thermometer *Journal of Climate* Volume 8,
893 Issue 5 (May 1995) pp. 139-140

894 Kalnay, E. and M. Cai, 2003: Impact of urbanization and land-use change on climate.
895 *Nature* **423**: 528– 531.

896 Kalnay, E., M. Cai, H. Li, and J. Tobin, 2006. Estimation of the impact of land-surface
897 forcings on temperature trends in eastern United States. *Journal of Geophysical Research*
898 **111**, (D06106) 1-13.

899 Kalnay, E., M. Cai, M. Nunez, and Y. Lim, 2008. Impacts of urbanization and land
900 surface changes on climate trends. *International Association for Urban Climate* **27**, 5–9.

901 Karl, T. R., G. Kukla, and J. Gavin, 1984: Decreasing diurnal temperature range in the
902 United States and Canada from 1941 through 1980. *J. Climate Appl. Meteor.*, **23**, 1878-
903 1886.

904 Karl, T.R., C.N. Williams, Jr., P.J. Young, and W.M. Wendland, 1986: A model to
905 estimate the time of observation bias associated with monthly mean maximum,

906 minimum, and mean temperature for the United States, *Journal of Climate and Applied*
907 *Meteorology*, **25**, 145-160.

908 Karl, T.R., and C.N. Williams Jr., 1987: An approach to adjusting climatological time
909 series for discontinuous inhomogeneities. *Journal of Climate and Applied Meteorology*,
910 **26**, 1744-1763.

911 Karl, T.R., H.F. Diaz, and G. Kukla, 1988: Urbanization: its detection and effect in the
912 United States climate record, *Journal of Climate*, **1**, 1099-1123.

913 Karl, T. R., J. D. Tarpley, R. G. Quayle, H. F. Diaz, D. A. Robinson, and R. S. Bradley,
914 1989: The recent climate record: What it can and cannot tell us, *Rev. Geophys.*, **27**, 405–
915 430, doi:10.1029/RG027i003p00405

916 Karl, T.R., C.N. Williams, Jr., F.T. Quinlan, and T.A. Boden, 1990: *United States*
917 *Historical Climatology Network (HCN) Serial Temperature and Precipitation Data*,
918 Environmental Science Division, Publication No. 3404, Carbon Dioxide Information and
919 Analysis Center, Oak Ridge National Laboratory, Oak Ridge, TN, 389 pp.

920 Karl, T. R., and Coauthors, 1993: Asymmetric trends of daily maximum and minimum
921 temperature. *Bull. Amer. Meteor. Soc.*, **74**, 1007-1023.

922 Karl, T.R., and Coauthors, 1995: Critical issues for long-term climate monitoring.
923 *Climate Change*, **31**, 185-221.

924 Karl, Thomas R., Knight, Richard W. 1997: The 1995 Chicago Heat Wave: How Likely
925 Is a Recurrence?. *Bull. Amer. Meteor. Soc.*, **78**, 1107–1119.

926 Klotzbach, P.J., R.A. Pielke Sr., R.A. Pielke Jr., J.R. Christy, and R.T. McNider, 2009:
927 An alternative explanation for differential temperature trends at the surface and in the
928 lower troposphere. *J. Geophys. Res.*, 114, D21102, doi:10.1029/2009JD011841.

929 Klotzbach, P.J., R.A. Pielke Sr., R.A. Pielke Jr., J.R. Christy, and R.T. McNider, 2010:
930 Correction to: "An alternative explanation for differential temperature trends at the
931 surface and in the lower troposphere. *J. Geophys. Res.*, 114, D21102,
932 doi:10.1029/2009JD011841", *J. Geophys. Res.*, 115, D1, doi:10.1029/2009JD013655.

933 Leroy, M., 1999: *Classification d'un site*. Note Technique no. 35. Direction des Systèmes
934 d'Observation, Météo-France, 12 pp.

935 Leroy, M., 2010: Siting Classification for Surface Observing Stations on Land, Climate,
936 and Upper-air Observations *JMA/WMO Workshop on Quality Management in Surface*,
937 Tokyo, Japan 27-30 July 2010

938 Lee, Seri, 1995 : Optimum design and selection of heat sinks, *Semiconductor Thermal*
939 *Measurement and Management Symposium*, 1995. SEMI-THERM XI., Eleventh Annual
940 IEEE, 7-9 Feb 1995

941 Lin, X., R.A. Pielke Sr., K.G. Hubbard, K.C. Crawford, M. A. Shafer, and T. Matsui,
942 2007: An examination of 1997-2007 surface layer temperature trends at two heights in
943 Oklahoma. *Geophysical Research Letters*, 34, L24705, doi:10.1029/2007GL031652.

944 Mahmood, R. L., S. A. Foster, and D. Logan, 2006: The GeoProfile metadata, exposure
945 of instruments, and measurement bias in climatic record revisited. *International Journal*
946 *of Climatology*, **26**, 1091-1124.

947 Mahmood, R., R.A. Pielke Sr., K.G. Hubbard, D. Niyogi, G. Bonan, P. Lawrence, B.
948 Baker, R. McNider, C. McAlpine, A. Etter, S. Gameda, B. Qian, A. Carleton, A. Beltran-
949 Przekurat, T. Chase, A.I. Quintanar, J.O. Adegoke, S. Vezhapparambu, G. Conner, S.
950 Asefi, E. Sertel, D.R. Legates, Y. Wu, R. Hale, O.W. Frauenfeld, A. Watts, M. Shepherd,

951 C. Mitra, V.G. Anantharaj, S. Fall, R. Lund, A. Nordfelt, P. Blanken, J. Du, H.-I. Chang,
952 R. Leeper, U.S. Nair, S. Dobler, R. Deo, and J. Syktus, 2010: Impacts of land use land

953 cover change on climate and future research priorities. *Bulletin of the American*
954 *Meteorological Society*, 91, 37–46, DOI: 10.1175/2009BAMS2769.1

955 Martinez, C.J., Maleski, J.J., Miller, M.F, 2012: Trends in precipitation and temperature
956 in Florida, USA. *Journal of Hydrology*. volume 452-453, issue , year 2012, pp. 259 – 281

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

961 Montandon, L.M., S. Fall, R.A. Pielke Sr., and D. Niyogi, 2011: Distribution of
962 landscape types in the Global Historical Climatology Network. *Earth Interactions*, 15:6,
963 doi: 10.1175/2010EI371

964 Menne, M.J., and C.N. Williams, Jr., 2005: Detection of undocumented changepoints
965 using multiple test statistics and composite reference series. *Journal of Climate*, **18**,
966 4271- 4286.

967 Menne, M.J., C.N. Williams, Jr., and R.S. Vose, 2009: The United States Historical
968 Climatology Network monthly temperature data – Version 2. *Bulletin of American*
969 *Meteorological Society*, **90**, 993-1007, doi: 10.1175/2008BAMS2613.1.

970 Menne, M.J., and C.N. Williams, Jr., 2009: Homogenization of temperature series via
971 pairwise comparisons. *Journal of Climate*, **22**, 1700-1717.

972 Menne, M. J., C. N. Williams Jr., and M. A. Palecki, 2010: On the reliability of the U.S.
973 surface temperature record, *J. Geophys. Res.*, **115**, D11108, doi:10.1029/2009JD013094

974 Mitchell, T.D. and P.D. Jones, 2005: An improved method of constructing a database of
975 monthly climate observations and associated high-resolution grids. *International Journal*
976 *of Climatology* **25**, 693-712

977 Muller, R.A., Curry, J., Groom, D. Jacobsen, R., Perlmutter, S. Rohde, R. Rosenfeld, A.,
978 Wickham, C., Wurtele, J., 2012: Earth Atmospheric Land Surface Temperature and
979 Station Quality in the United States. *JGR Special Publication The Third Santa Fe*
980 *Conference on Global and Regional Climate Change manuscript number 12JD018146*
981 <http://berkeleyearth.org/pdf/berkeley-earth-station-quality.pdf>

982 NOAA/NESDIS, 2002: *Climate Reference Network Site Information Handbook*. National
983 Oceanic and Atmospheric Administration, NOAA-CRN/OSD-2002-0002R0UD0. 19 pp.
984 Available online at:
985 [ftp://ftp.ncdc.noaa.gov/pub/data/uscrn/documentation/program/X030FullDocumentD0.pd](ftp://ftp.ncdc.noaa.gov/pub/data/uscrn/documentation/program/X030FullDocumentD0.pdf)
986 [f](ftp://ftp.ncdc.noaa.gov/pub/data/uscrn/documentation/program/X030FullDocumentD0.pdf)

987 OFCM, 1994: Office of the Federal Coordinator for Meteorology, *Federal Standard for*
988 *Siting Meteorological Sensors at Airports*, FCM-S4-1994, Washington, D.C. August 1994
989 Chap. 2- Sensor Exposure, Section 2.6, Temperature and Dew Point Sensors, Available
990 online at: <http://www.ofcm.gov/siting/text/a-cover.htm>

991 Peterson, T.C., and D.R. Easterling, 1994: Creation of homogeneous composite
992 climatological reference series, *International Journal of Climatology*, **14**, 671–679.

993 Peterson TC, Easterling DR, Karl TR, Groisman P, Nicholls N, Plummer N, Torok S,
994 Auer I, Bohm R, Gullett D, Vincent L, Heino R, Tuomenvirta H, Mestre O, Szentimrey
995 T, Salinger J, Folland EJ, Hanssen-Bauer I, Alexandersson H, Jones P, Parker, D. 1998.
996 Homogeneity adjustments of in situ atmospheric climate data: a review. *International*
997 *Journal of Climatology* **18**: 1493-1517

998 Peterson, T.C., 2003: Assessment of urban versus rural in situ surface temperatures in the
999 contiguous United States: No difference found. *Journal of Climate*, **16**, 2941–2959.

1000 Peterson, T.C., 2006: Examination of potential biases in air temperature caused by poor
1001 station locations. *Bull. Amer. Meteor. Soc.*, **87**, 1073-1080.

1002 Peterson, T. C., K. M. Willett, and P. W. Thorne (2011), Observed changes in surface
1003 atmospheric energy over land, *Geophys. Res. Lett.*, **38**, L16707,
1004 doi:10.1029/2011GL048442

1005 Pielke, R.A., Sr., T. Stohlgren, L. Schell, W. Parton, N. Doesken, K. Redmond, J. Money,
1006 T. McKee, and T.G.F. Kittel., 2002: Problems in evaluating regional and local trends in
1007 temperature: An example from eastern Colorado, USA. *International Journal of*
1008 *Climatology*, **22**, 421–434.

1009 Pielke, R.A. Sr., K. Wolter, O. Bliss, N. Doesken, and B. McNoldy, 2007: The July 2005
1010 Denver heat wave: How unusual was it? *Nat. Wea. Dig.*, **31**, 24-35.

1011

1012 Pielke Sr., R.A., C. Davey, and J. Morgan, 2004: Assessing "global warming" with
1013 surface heat content. *Eos*, **85**, No. 21, 210-211.

1014 Pielke Sr., R.A. J. Nielsen-Gammon, C. Davey, J. Angel, O. Bliss, N. Doesken, M. Cai,
1015 S. Fall, D. Niyogi, K. Gallo, R. Hale, K.G. Hubbard, X. Lin, H. Li, and S. Raman,
1016 2007a: Documentation of uncertainties and biases associated with surface temperature
1017 measurement sites for climate change assessment. *Bulletin of American Meteorological*
1018 *Society*, **88**, 913-928.

1019 Pielke Sr., R.A., C. Davey, D. Niyogi, S. Fall, J. Steinweg-Woods, K. Hubbard, X. Lin,
1020 M. Cai, Y.-K. Lim, H. Li, J. Nielsen-Gammon, K. Gallo, R. Hale, R. Mahmood, S.
1021 Foster, R.T. McNider, and P. Blanken, 2007b: Unresolved issues with the assessment of
1022 multi-decadal global land surface temperature trends. *Journal of Geophysical Research*,
1023 **112**, D24S08, doi:10.1029/2006JD008229.

1024 Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar and the R Development
1025 Core Team (2012). nlme: Linear and Nonlinear Mixed Effects Models. R package
1026 version 3.1-104. Available online at: <http://cran.r-project.org/web/packages/nlme/>

1027 Quayle, R.G., D.R. Easterling, T.R. Karl, and P.Y. Hughes, 1991, Effects of recent
1028 thermometer changes in the Cooperative Station Network, *Bulletin of American*
1029 *Meteorological Society*, **72**, 1718– 1723.

1030 Reeves, J., J. Chen, X.L. Wang, R. Lund, and Q.Q. Lu, 2007: A review and comparison
1031 of changepoint detection techniques for climate data. *Journal of Applied Meteorology*
1032 *and Climatology*, **46**, 900-914.

1033 Runnalls, K.E. and Oke, T.R. 2006. A technique to detect microclimatic inhomogeneities
1034 in historical records of screen-level air temperature. *Journal of Climate*, **19**, 959-978.

1035 Snowe, Olympia J., Senator from Maine, Testimony before the U.S. Senate on
1036 Automated Surface Observations System Stations, Congressional Record, V. 144, Pt. 15,
1037 September 22, 1998 to September 26, 1998, page 158, Accessible online at :
1038 http://books.google.com/books?id=MfrwH4TB93QC&lpg=PA158&ots=kEEsz2_h9k&dq=ASOS%20errors%20Congressional%20record&pg=PA158#v=onepage&q&f=false
1039

1040 Stone, D. A. and A. J. Weaver , 2003: Factors contributing to diurnal temperature range
1041 trends in the twentieth and twenty first century simulations of the CCama coupled
1042 model, *Climate Dynamics*, **20**, 435-445.

1043 Stott P. and P. Thorne; 2010: Proposals for surface-temperature data bank now open for
1044 scrutiny *Nature*, **466**, 1040-1040

1045 Thorne, P.W., D.E. Parker, J.R. Christy, and C.A. Mears, 2005: Uncertainties in climate
1046 trends: Lessons from upper-air temperature records, *Bulletin of American Meteorological*
1047 *Society*, **86**, 1437-1442.

1048 Vose, R.S., C.N. Williams Jr., T.C. Peterson, T.R. Karl, and D.R. Easterling, 2003: An
1049 evaluation of the time of observation bias adjustment in the U.S. Historical Climatology
1050 Network, *Geophysical Research Letters*, **30**, 2046, doi:10.1029/2003GL018111.

1051 Vose, R. S., D. R. Easterling, T. R. Karl, and M. Helfert, 2005a: Comments on
1052 “Microclimate exposures of surface-based weather stations.” *Bulletin of American*
1053 *Meteorological Society*, **86**, 504–506.

1054 Vose, R. S., D. R. Easterling, and B. Gleason, 2005b: Maximum and minimum
1055 temperature trends for the globe: An update through 2004. *Geophys. Res. Lett.*, **32**,
1056 L23822, doi:10.1029/2005GL024379.

1057 Walters, J. T., R. T. McNider, X. Shi, and W. B. Norris, 2007: Positive surface
1058 temperature feedback in the stable nocturnal boundary layer. *Geophysical Research*
1059 *Letters*, doi:10.1029/2007GL029505.

1060 Watts, A., 2009: Is the U.S. surface temperature record reliable? The Heartland Institute,
1061 Chicago, IL. 28 pp.

1062 Willmot, C.J., S.M. Robeson, and J.J. Feddema, 1991: Influence of spatially variable
1063 instrument networks on climatic averages, *Geophysical Research Letters*, **18**, 2249–
1064 2251.

1065 World Meteorological Organization Commission for Instruments and Methods of
1066 Observation, Fifteenth session, (CIMO-XV, 2010) WMO publication Number 1064,
1067 available online at: [http://www.wmo.int/pages/prog/www/CIMO/CIMO15-
1068 WMO1064/1064_en.pdf](http://www.wmo.int/pages/prog/www/CIMO/CIMO15-WMO1064/1064_en.pdf)

1069 Yilmaz, H., Toy, S., Irmak, M.A., Yilmaz, S., Bulit, Y., 2008: Determination of
1070 temperature differences between asphalt concrete, soil, and grass surfaces of the City of
1071 Ezurum, Turkey. *Atmosfera*, **21**, 135-146