

Urban heat islands during heat waves: a comparative study between Boston and Phoenix

Liang Wang¹, Dan Li^{1*}

1: Department of Earth and Environment, Boston University, MA, 02215, USA.

* corresponding author: lidan@bu.edu

Abstract In this study, we simulate the magnitude of urban heat islands (UHIs) during heat wave (HWs) in two cities with contrasting climates (Boston and Phoenix) using the Weather Research and Forecasting (WRF) model and quantify their drivers with a newly developed attribution method. During the daytime, a surface UHI (SUHI) is found in Boston mainly caused by the higher urban surface resistance (r_s) that reduces the latent heat flux, and the higher urban aerodynamic resistance (r_a) that inhibits convective heat transfer between the urban surface and the lower atmosphere. In contrast, a surface urban cool island (SUCI) is found in Phoenix mainly due to the lower urban r_a that facilitates convective heat transfer. In terms of near-surface air UHI (AUHI), there is almost no daytime AUHI in either city. At night, a SUHI and an AUHI are identified in Boston due to the stronger release of heat storage in urban areas. In comparison, the lower urban r_a in Phoenix enhances convective heat transfer from the atmosphere to the urban surface at night, leading to a positive SUHI but no AUHI. Our study highlights that the magnitude of UHIs or UCIs is strongly controlled by urban-rural differences in terms of aerodynamic features, vegetation and moisture conditions, and heat storage, which show contrasting characteristics in different regions.

Keywords Aerodynamic resistance • Heat waves • Surface resistance • Urban heat islands • WRF

1 Introduction

Despite having no universal definitions, heat waves (HWs) usually refer to a sustained period (typically more than 2 days) when the temperatures (e.g., daily maximum, mean, or minimum temperatures) exceed a certain threshold (Robinson 2001). Recent years have witnessed numerous disastrous HWs worldwide, such as the 2003 HW in Europe (García-Herrera et al. 2010), the 2010 Russian HW (Dole et al. 2011) and the 2013 HW in eastern China (Xia et al. 2016), incurring substantial socioeconomic costs (Xia et al. 2018) and raising concerns about human health (Campbell et al. 2018; Mora et al. 2017; Petkova et al. 2014), wildfires (Parente et al. 2018), crop failures (Fontana et al. 2015; Wreford and Adger 2010; Zampieri et al. 2017), and infrastructure damage (Rübelke and Vögele 2011). What is worse is that the intensity, duration and frequency of HWs have been increasing and will likely continue to increase in many parts of the world in a warming climate (Brown et al. 2008; Donat et al. 2013; Fischer and Schär 2010; Lau and Nath 2012, 2014; Meehl and Tebaldi 2004; Schär et al. 2004; Tebaldi et al. 2006).

With the high thermal risks imposed by HWs in mind (Patz et al. 2005; Tan et al. 2010), the urban population, which comprises more than half of the world's population and is projected to reach 68% by 2050 (Grimm et al. 2008; United Nations 2019), usually experiences hotter conditions than the rural counterpart due to the well-known urban heat island (UHI) effects (Arnfield 2003; Oke 1982). The UHI effects have important impacts on the atmospheric boundary layer flow, the dispersion of pollutants, the energy and water consumption in cities, and so on (Han et al. 2014; Hidalgo et al. 2009; Miao et al. 2009; Pal et al. 2012; Zhang et al. 2014). The UHI intensity, which characterizes the magnitude of the UHI effect, is typically defined as the temperature difference between the urban and

the surrounding rural areas, which can be based on either near-surface air temperature (i.e., air UHI or AUHI) or surface temperature (i.e., surface UHI or SUHI) (Stewart 2011). The negative AUHI and SUHI are called air urban cool island (AUCI) and surface urban cool island (SUCI), respectively, in this study.

Broadly, UHIs can be explained by the surface energy balance equation. For an imaginary control volume with an arbitrary horizontal scale that extends from the ground to the roof level and has no net horizontal advection of heat through the sides of the control volume, the energy balance equation can be expressed as (Oke et al. 2017):

$$AF + (1 - \alpha)SW_{in} + \varepsilon LW_{in} = H + LE + G + \varepsilon \sigma T_s^4 \quad (1)$$

where AF is the anthropogenic heat flux (W m^{-2}), α is the surface albedo, SW_{in} is the incoming shortwave radiation (W m^{-2}), ε is the surface emissivity, LW_{in} is the incoming longwave radiation (W m^{-2}), H is the sensible heat flux (W m^{-2}), LE is the latent heat flux (W m^{-2}), G is the ground heat flux (also called heat storage, W m^{-2}), σ is the Stefan-Boltzmann constant ($\text{W m}^{-2} \text{K}^{-4}$), and T_s is the land surface temperature (K). The terms on the left-hand side of Eq. (1) represent the energy input to the control volume either from anthropogenic emissions or from the radiation (after subtracting the reflected radiation), which includes the solar insolation and the longwave radiation from the ambient atmosphere. The input energy can be transferred away from the control volume by either convection into the lower atmosphere (H), conduction into the ground (G), or radiation in the longwave band ($\varepsilon \sigma T_s^4$). Part of the input energy is consumed by the evapotranspiration process in order to transform the water from the liquid phase to the vapor phase, namely, the latent heat flux (LE). UHIs are mainly caused by the unique urban characteristics such as less evapotranspiration associated with limited green space and low surface moisture,

lower albedo owing to radiative trapping, larger heat release at night due to the higher thermal admittance of built materials as well as the larger anthropogenic emissions (Arnfield 2003; Grimm et al. 2008; Grimmond 2007; Oke 1982; Ramamurthy et al. 2014; Taha 1997; Zhao et al. 2014). These unique urban characteristics either increase the energy input on the left-hand side of Eq. (1) or inhibit the efficiency of heat transfer on the right-hand side of the equation, leading to hotter urban surfaces and near-surface conditions.

Although the causes of UHIs are generally well understood, it remains a challenge to quantify and compare the contribution of each individual process to the magnitude of UHIs, which strongly depends on the nature of the urban and rural environments, human activities, and meteorological conditions (Grimmond 2007). Moreover, the spatial variations of UHIs across cities and background climates and their key controlling factors are still under debate (Li et al. 2019; Manoli et al. 2019; Zhao et al. 2014). To fill this research gap, the present study simulates the magnitude of UHIs during about 20 HW events in the period of 2007-2016 over two cities in the United States (i.e., Boston and Phoenix) with the Weather Research and Forecasting (WRF) model. The objective of this study is to quantify and compare the underlying drivers of UHIs during HWs in these two cities with contrasting background climates. Specifically, Boston has a humid continental climate while Phoenix has a hot desert climate. The annual mean temperature is 9.8 °C in Boston while 21.5 °C in Phoenix, and the annual mean rainfall is 1122 mm in Boston while only 211 mm in Phoenix. Therefore, conducting the same analyses over these two cities can provide insights into the spatial variability of UHIs and their controlling factors.

The paper is organized as follows: sect. 2 describes the data and methodology, sect. 3 presents the main results, sect. 4 discusses the implications and limitations of our work, and sect. 5 concludes the paper.

2 Data and Methodology

2.1 Observational data

To identify HWs, we use the 2-m air temperatures from the Integrated Surface Database (ISD) provided by the National Oceanic and Atmospheric Administration (NOAA, [ncdc.noaa.gov/isd/data-access](https://www.ncdc.noaa.gov/isd/data-access)) and measured at the Boston Logan International Airport and the Phoenix Sky Harbor International Airport. These data are used because of their long-term span, which is required for HW identification. However, for validation of WRF simulation results, we use the 2-m air temperatures from the Meteorological Assimilation Data Ingest System (MADIS, madis-data.ncep.noaa.gov/madisPublic1/data/archive) due to their larger spatial coverage (see blue triangles on Figure 1a & c). We only use the MADIS stations whose data availability in the simulation period is more than 90%.

To evaluate the WRF-simulated boundary-layer structure, we use commercial aircraft data provided by the Aircraft Communications Addressing and Reporting System (ACARS; madis-data.cprk.ncep.noaa.gov/madisPublic1/data/archive). The ACARS data are proved to be as good as the radiosonde data in terms of quality in the lower atmosphere, but with much higher temporal and vertical resolutions (Petersen 2016; Zhang et al. 2019). Specifically, the root-mean-square error (RMSE) of ACARS against radiosonde below 850 hPa is approximately 1.3 K, while the mean bias error (MBE) falls between -0.16 K and -0.32 K (Zhang et al. 2019). In this study, we use the temperature profiles collected by airlines when they take off or land at the Boston Logan International Airport and the

Phoenix Sky Harbor International Airport (see blue circles on Figure 1b & d), which have been processed and interpolated onto regular height levels at the hourly scale (Zhang et al. 2019, 2020).

To validate the WRF-simulated land surface temperatures, we use the land surface temperature data from the 1-km MOD11A1 Version 6 product provided by the Moderate Resolution Imaging Spectroradiometer (MODIS; lpdaac.usgs.gov/products/mod11a1v006).

2.2 HW identification

The HW definitions used in previous studies vary in terms of the temperature variable, the threshold of its magnitude, and the duration/extent (Anderson and Bell 2011; Chen and Zhai 2017; Lau and Nath 2014; Liao et al. 2018a; Luo and Lau 2016; Meehl and Tebaldi 2004; Peng et al. 2011; Perkins 2015; Sillmann et al. 2013; Yang et al. 2017). In our research, we identify a HW as a period of at least 2 consecutive days whose daily mean temperature exceeds the 95 percentile of the local climatology (Anderson and Bell 2009, 2011; Hajat et al. 2006; Zhang et al. 2020), which is defined based on daily mean temperatures during the warm seasons (May 1 to September 30) of 2007-2016. The thresholds of two days and 95 percentile are used to ensure enough HW days selected for our analysis and also distinct thermal conditions between HW and non-HW days. We further corroborate these selected HWs using the ACARS temperature profiles (i.e., we manually check if the start and the end of selected HWs coincide with obvious changes in the ACARS temperature profiles, see Zhang et al. 2020). Overall, 41 HWs are selected during 2007-2016, with 20 HWs in Boston and 21 HWs in Phoenix (Table 1).

2.3 WRF simulations

133 The WRF model has been widely used to study urban climate (Chen et al. 2011; Georgescu
134 et al. 2011; Heaviside et al. 2015; Li and Bou-Zeid 2014; Li et al. 2018; Meir et al. 2013;
135 Ramamurthy et al. 2017; Ramamurthy and Bou-Zeid 2016; Tewari et al. 2019). In this
136 study, the WRF version 4.0 is used to simulate each HW event, with 5 pre-HW days and 5
137 post-HW days. Three nested model domains with spatial resolutions of 9, 3 and 1 km,
138 respectively, are used (Figure 1). For Boston, these three domains have 149×149 , $150 \times$
139 150 and 150×150 grid cells, respectively. The 9-km domain covers most of the
140 northeastern United States, and the innermost domain covers Boston as well as the
141 surrounding landscapes which are mostly forests. For Phoenix, these three domains have
142 299×299 , 300×300 and 180×180 grid cells, respectively. The 9-km domain covers most
143 of the southwestern United States as well as a sizeable portion of Mexico, and the innermost
144 domain covers Phoenix and the surrounding landscapes which are mostly scrublands. All
145 model domains have 55 vertical levels, and the model top is set as 100 hPa. The North
146 American Regional Reanalysis (NARR) data with a spatial resolution of about 32 km and
147 a temporal resolution of 3 hours are used for the initial and boundary conditions. National
148 Land Cover Database 2011 (NLCD 2011) is used as land use input to the model (Homer et
149 al. 2015). When used in WRF, this land cover dataset has a 9-second spatial resolution.
150 The urban land is classified into three categories (commercial urban, high-intensity
151 residential urban, low-intensity residential urban). Thus, not all urban grid cells have the
152 same surface properties. The physical parameterizations for the WRF simulations follow
153 closely the studies by Li and Bou-Zeid (2014) and Wang and Li (2019), including the
154 Dudhia scheme for shortwave radiation (Dudhia 1988), the rapid radiative transfer model
155 (RRTM) scheme for longwave radiation (Mlawer et al. 1997), the single-moment 6-Class

(WSM6) microphysics scheme (Hong and Lim 2006), and the Noah land surface model (Ek et al. 2003) coupled with the single-layer urban canopy model (Kusaka et al. 2001; Kusaka and Kimura 2004). Following Wang and Li (2019), we test three planetary boundary layer (PBL) schemes: the asymmetric convective model, version 2 (ACM2) scheme (Pleim 2007), the Yonsei University (YSU) scheme (Hong et al. 2006), and the Mellor-Yamada Janjic (MYJ) scheme (Mellor and Yamada 1974). We focus on a sensitivity test to the PBL scheme as turbulent heat transfer is an extremely important process that regulates the magnitude of UHIs (Li and Bou-Zeid 2014). For model validation, we select 6 HW events with more complete observational records (denoted with an asterisk in Table 1). We will use the validation results to select a PBL scheme with the best consistency with the observational data, which will be used for simulating the remaining HW events.

2.4 The attribution method

2.4.1 Attribution of SUHI

The attribution of SUHI is based on the Two-Resistance Mechanism (TRM) model used in a number of recent studies (Chen et al. 2020; Li et al. 2019; Li and Wang 2019; Liao et al. 2018b; Moon et al. 2020; Rigden and Li 2017; Wang et al. 2019, 2020). The TRM model starts from the surface energy balance equation as Eq. (1), with H and LE parameterized using the concepts of aerodynamic resistance and surface resistance (Brutsaert 1982, 2005; Monteith and Unsworth 2007), as follows:

$$H = \frac{\rho c_p}{r_a} (T_s - T_a) \quad (2)$$

$$LE = \frac{\rho L_v}{r_a + r_s} (q^*(T_s) - q_a) \quad (3)$$

where ρ is the air density (kg m^{-3}), c_p is the specific heat of air at constant pressure (J kg^{-1}
 K^{-1}), r_a is the aerodynamic resistance between the surface and the lower atmosphere (s m^{-1}), T_a is the potential temperature of the lower atmosphere (K, not the near-surface or 2-m
 air temperature as discussed below), L_v is the latent heat of vaporization (J kg^{-1}), r_s is the
 surface resistance (s m^{-1}), q^* is the saturated specific humidity at T_s (kg kg^{-1}), and q_a is the
 specific humidity of the lower atmosphere (kg kg^{-1}). The aerodynamic resistance (r_a)
 represents the efficiency with which the land surface convects sensible heat to the lower
 atmosphere (Brutsaert 1982, 2005; Garratt 1992). It is controlled mostly by the capacity of
 atmospheric turbulence in transporting sensible heat but also molecular diffusion at the
 interface between the land and the atmosphere (Brutsaert 1982, 2005; Garratt 1992). From
 Eq. (2) one can see that a higher aerodynamic resistance results in a lower sensible heat
 flux with a given temperature gradient. On the other hand, the surface resistance (r_s)
 represents the efficiency with which water is extracted from the saturated zone to the
 surface or from the vegetation inside to the leaf surface, which is strongly dependent on
 soil moisture and vegetation stresses (Brutsaert 1982, 2005; Garratt 1992). From Eq. (3)
 one can see that moisture transfer from the land to the atmosphere also experiences the
 aerodynamic resistance (r_a), in addition to the surface resistance (r_s). With everything else
 being equal, a higher surface resistance leads to a smaller latent heat flux.

Substituting Eqs. (2-3) into Eq. (1) yields a non-linear equation for T_s , which is further
 linearized by applying first-order Taylor series expansion to the outgoing longwave
 radiation and the saturated specific humidity terms, so that an analytical expression for T_s
 can be obtained:

$$T_s = \frac{AF + SW_{in}(1 - \alpha) + \varepsilon LW_{in} - \varepsilon \sigma T_a^4 - G - \frac{\rho L_v}{r_a + r_s}(q^*(T_a) - q_a)}{4\varepsilon \sigma T_a^3 + \frac{\rho c_p}{r_a} + \frac{\rho L_v}{(r_a + r_s)} \frac{\partial q^*}{\partial T} |_{T_a}} + T_a \quad (4)$$

Based on this solution, the difference in the surface temperature between urban and rural surfaces can be further expressed as the sum of the contributions from various factors again using first-order Taylor series expansion, as follows:

$$\Delta T_s = \left(\frac{\partial T_s}{\partial AF} \right) \Delta AF + \left(\frac{\partial T_s}{\partial \varepsilon} \right) \Delta \varepsilon + \left(\frac{\partial T_s}{\partial \alpha} \right) \Delta \alpha + \left(\frac{\partial T_s}{\partial G} \right) \Delta G + \left(\frac{\partial T_s}{\partial r_a} \right) \Delta r_a + \left(\frac{\partial T_s}{\partial r_s} \right) \Delta r_s + \dots \\ + \left(\frac{\partial T_s}{\partial SW_{in}} \right) \Delta SW_{in} + \left(\frac{\partial T_s}{\partial LW_{in}} \right) \Delta LW_{in} + \left(\frac{\partial T_s}{\partial T_a} \right) \Delta T_a + \left(\frac{\partial T_s}{\partial q_a} \right) \Delta q_a + \left(\frac{\partial T_s}{\partial P} \right) \Delta P \quad (5)$$

where P is the pressure (Pa), and Δ indicates the urban-rural difference in each factor (i.e., urban minus rural values). In Eq. (5), the partial derivative represents the sensitivity of the surface temperature to the change in each factor, whose analytical formulation can be obtained using Eq. (4). Throughout the paper, each term on the right-hand side of Eq. (5) will be called a contribution. Each contribution is the product of the sensitivity (Table 3), which represents how sensitive the surface temperature is to a change in each factor, and the difference (Table 5), which represents the urban-rural contrast of that factor (i.e., urban minus rural values).

Compared with previous attribution methods (such as Li et al. 2019; Zhao et al. 2014), Eq. (5) not only considers the influence of differences in land surface biophysical parameters (such as albedo) on SUHIs, but also the influence of differences in atmospheric conditions (such as incoming shortwave radiation) on SUHIs. This is important for our study because unlike previous studies assuming that urban and rural lands share the same atmospheric conditions due to their use of sub-grid outputs from global climate or earth system models (Li et al. 2019; Zhao et al. 2014), the SUHIs in our study are the surface

temperature differences between urban and rural grid cells, which have different overlying atmospheric conditions.

2.4.2 Attribution of AUHI

The original TRM method was designed for the attribution of surface temperature differences (i.e., SUHIs). In this study, we further develop the TRM method to attribute the near-surface air temperature differences (i.e., AUHIs). The mathematical link between the surface temperature and near-surface air temperature is provided by the constant heat flux assumption in the atmospheric surface layer. This assumption is the basis for deriving the so-called 2-m air temperature in models like WRF. With this assumption, the sensible heat flux derived from the temperature difference between the land surface and the lower atmosphere (i.e., Eq. 2) should be equal to the sensible heat flux derived from the temperature difference between the air at 2 meters above the displacement height and the lower atmosphere. Thus the 2-m air temperature (T_2) can be related to the surface temperature (T_s) and the potential temperature of the lower atmosphere (T_a), as follows:

$$T_2 = \frac{r'_a}{r_a} (T_s - T_a) + T_a \quad (6)$$

where r'_a is the aerodynamic resistance between the reference height (i.e., 2 meters above the displacement height) and the lower atmosphere (s m^{-1}). Hence, r'_a is different from r_a , which is the aerodynamic resistance between the surface and the lower atmosphere.

Thus, the urban-rural difference in the 2-m air temperature can be expressed as follows:

$$\begin{aligned}
\Delta T_2 = & f \left(\frac{\partial T_s}{\partial AF} \right) \Delta AF + f \left(\frac{\partial T_s}{\partial \varepsilon} \right) \Delta \varepsilon + f \left(\frac{\partial T_s}{\partial \alpha} \right) \Delta \alpha + f \left(\frac{\partial T_s}{\partial G} \right) \Delta G + \dots \\
& + \left(f \left(\frac{\partial T_s}{\partial r_a} \right) - \frac{r'_a}{r_a^2} (T_s - T_a) \right) \Delta r_a + \left(\frac{T_s - T_a}{r_a} \right) \Delta r'_a + f \left(\frac{\partial T_s}{\partial r_s} \right) \Delta r_s + \dots \\
& + f \left(\frac{\partial T_s}{\partial SW_{in}} \right) \Delta SW_{in} + f \left(\frac{\partial T_s}{\partial LW_{in}} \right) \Delta LW_{in} + \left(f \left(\frac{\partial T_s}{\partial T_a} \right) - f + 1 \right) \Delta T_a + \dots \\
& + f \left(\frac{\partial T_s}{\partial q_a} \right) \Delta q_a + f \left(\frac{\partial T_s}{\partial P} \right) \Delta P
\end{aligned} \tag{7}$$

where $f = \frac{r'_a}{r_a}$. Again, Δ indicates the urban-rural difference in each factor (Table 5) while the remainder of each term represents the sensitivity of T_2 to each factor (Table 4).

2.5 Application of the attribution method to the WRF outputs

The attribution analysis in general follows previous research by Liao et al. (2018b) and Li et al. (2019). Most variables needed in Eqs. (5 & 7) are direct outputs of the WRF model including the surface temperature, 2-m air temperature, emissivity, albedo, ground heat flux, incoming shortwave radiation, incoming longwave radiation, pressure, and the potential temperature at the lowest level of the atmospheric model (which is about 30 m in our simulations). The specific humidity is calculated from the water vapor mixing ratio ($QVAPOR$) through $q_a = QVAPOR / (QVAPOR + 1)$. Since our WRF simulations do not include anthropogenic heat fluxes in the form of both sensible and latent heat, AF is set to zero in the attribution. The implication of neglecting anthropogenic heat fluxes will be discussed in section 4.2.

The aerodynamic resistances (r_a and r'_a) and the surface resistance (r_s) are not direct outputs from the WRF model. To infer them at the grid-cell scale, we use the WRF simulated sensible and latent heat fluxes, as well as the surface temperature, 2-m air temperature, and the potential temperature and specific humidity at the lowest level of the atmospheric model following Eqs. (2, 3, and 6). It should be stressed that we use the

potential temperature and specific humidity at the lowest level of the atmospheric model (~30 m), instead of the 2-m air temperature and specific humidity, to represent T_a and q_a in Eqs. (2, 3, and 6), because the 2-m air temperature and specific humidity are not prognostic variables in the model. We also stress that r_a is different from r'_a , arising from the fundamental dependence of aerodynamic resistance on the height z in the bulk formulations for surface fluxes (Garratt 1992): r_a represents the resistance to convective heat transfer between the surface and the lowest level of the atmospheric model while r'_a represents the resistance to convective heat transfer between the reference height (i.e., 2 m above the displacement height) and the lowest level of the atmospheric model.

The aerodynamic resistances and surface resistance are inferred using Eqs. (2, 3, and 6) at the hourly scale for grid cells in the innermost domain. In this study, we only consider grid cells with the dominant urban land cover (i.e., high-intensity residential urban for both cities) and the dominant rural land cover (i.e., evergreen forest for Boston and shrub for Phoenix). In total, there are 1372 (1452) urban grid cells and 1798 (25502) rural grid cells in Boston (Phoenix). We consider only one urban type because each urban type corresponds to a unique set of urban parameters and thus considering only one urban type simplifies the comparison between the two cities and reveals more information about the role played by the background climate. We use the default values of urban parameters in WRF (Chen et al. 2011). The high-intensity residential urban corresponds to the average building height of 7.5 m, the average roof and road widths of 9.4 m, and the impervious surface fraction of 0.9. Other thermodynamic parameters such as albedo, emissivity, heat capacity, and thermal conductivity can be found in Chen et al. (2011). When the resistances are estimated at the hourly scale, they may be of negative values, especially for

aerodynamic resistances. This issue is fundamentally linked to the fact that the TRM model parameterizes turbulent heat fluxes using the bulk transfer relations (i.e., Eqs. 2-3), which are local turbulence closures and assume uniform distributions of heat sources at the surface (i.e., with a single surface temperature). In the WRF model, the heat sources are not uniformly distributed (e.g., the single-layer urban canopy model considers the roof, the wall, and the ground separately). In this situation, the inferred aerodynamic resistances (e.g., based on Eq. 2) might appear to be negative. Since the negative resistances are physically meaningless, the grid cells with negative aerodynamic or surface resistances are filtered out following previous work (Li et al. 2019; Liao et al. 2018b). The numbers of the filtered urban/rural grid cells vary from hour to hour, and are on average about 6% (1%) for Boston (Phoenix) during the daytime, and about 56% (3%) for Boston (Phoenix) during the nighttime. To further reduce the uncertainties of the inferred resistances, we remove the grid cells in which the magnitudes or absolute values of sensible and latent heat fluxes are smaller than 15 W m^{-2} in the daytime (10am-4pm, local standard time) and 0.1 W m^{-2} in the nighttime (10pm-4am, local standard time). These thresholds are needed because the fluxes appear in the denominator when inferring the resistances. The exact values of these thresholds are chosen as a compromise between ensuring that the uncertainty of the inferred resistance is sufficiently small but in the meantime still maintaining a reasonably large sample of grid cells. After applying these data filtering strategies, there remain on average about 1300 and 350 urban grid cells in Boston during the daytime and nighttime, respectively, and 1500 and 1000 rural grid cells in Boston during the daytime and nighttime, respectively. For Phoenix, there are on average about 200 (10000) and 100 (20000) urban (rural) grid cells during the daytime and nighttime, respectively. We note

that these data filtering strategies are only applied when the attribution analysis is conducted (i.e., in section 3.3). The consistency between the average diurnal cycle results in section 3.2 and the attribution results in section 3.3, as shall be seen later, implies that these data filtering strategies do not alter our key findings.

Then we spatially average each variable across the dominant urban and rural land cover types, which results in one urban value and one rural value for each variable in each hour. After performing the spatial average, we further average the data over daytime and nighttime for each HW day and then conduct the attribution analysis at the daytime average and nighttime average scales (hereafter the daily scale). In doing so, the consistency between the TRM modelled UHIs and the WRF simulated UHIs is higher compared to performing the attribution at the hourly scale and then aggregating the results to the daily scale. This is similar to what previous studies found (Li et al. 2019; Liao et al. 2018b).

3 Results

3.1 Assessment of the WRF simulations with different PBL schemes

This section aims to assess the WRF-simulated results, with a focus on comparing the performances of WRF with different PBL schemes. To do so, 6 out of the 41 HWs are selected (noted with an asterisk in Table 1), with 3 in Boston and 3 in Phoenix.

First, we validate the WRF-simulated temperature profiles against the ACARS observations at the airports. Figure 2 shows an example (HW case 1, June 25-28, 2007) in Boston. We find that the WRF model simulates the “heat dome” over HW days (June 25-28) as observed by the ACARS data. A significant warming and a sharp cooling are captured by WRF in the lowest 4 km of the atmosphere on the days right before (June 24) and after (June 29) the HW, respectively (Figure 2b-d). We do not observe obvious

differences among the WRF simulations with different PBL schemes (Figure 2b-d), which is also reflected by the similar RMSEs between the WRF simulated results with the three PBL schemes and the ACARS data (Table 2). When the validation against the ACARS data is further applied to the other 5 HW cases, the weighted average RMSEs in Table 2 show that ACM2 is slightly superior to the other two PBL schemes but the differences between the PBL schemes are rather small.

Second, we validate the WRF-simulated 2-m air temperatures against the weather station data from MADIS. Figure 3 shows the results from the same HW event as in Figure 2 at two weather stations (i.e., the urban site KBED and the rural site KMQE in Boston marked by yellow triangles in Figure 1b). Although some biases exist, the WRF-simulated results show overall good agreement with the observations and can reasonably reflect the onset and demise of the HW. Different PBL schemes show contrasting performances. For example, at the urban site KBED (Figure 3a, c & e), the WRF results using MYJ deviate from the observations more strongly than those using the other two PBL schemes. However, at the rural site KMQE (Figure 3b, d & f), the WRF results with MYJ are closer to the observations than those with the other two PBL schemes. The weighted average RMSEs based upon all available weather station data (Table 2) show that ACM2 is slightly better than YSU, with MYJ showing the worst performance, which is consistent with the previous findings using ACARS.

Third, we compare the land surface temperatures simulated by WRF and remotely sensed by MODIS. Figure 4 shows two snapshots of the simulated and observed land surface temperature patterns at around 11 am and 9 pm (local standard time) on June 27, 2007 in Boston. Unlike ACARS and MADIS that have continuous observations over the

course of the day, MODIS only provides land surface temperature measurements at its overpassing time. Moreover, the availability of MODIS data is often reduced by the presence of clouds. For this particular HW day, the WRF simulations are able to capture the broad urban-rural land surface temperature contrast, with some underestimation of SUHI during the daytime and overestimation during the nighttime. When the temperatures are averaged over urban areas, the WRF simulations yield 311.0 K (300.5 K), 311.2 K (300.2 K), 310.4 K (299.6 K) for ACM2, YSU and MYJ, respectively; and the MODIS gives 312.3 K (298.4 K) during the daytime (nighttime). For the rural average land surface temperatures, the WRF simulations yield 304.7 K (296.2 K), 304.8 K (295.2 K), 305.0 K (294.9 K) for ACM2, YSU and MYJ, respectively, and the MODIS gives 305.7 K (296.3 K) during the daytime (nighttime).

The performance of the WRF model presented here is broadly consistent with previous studies (Kalverla et al. 2016; Li and Bou-Zeid 2014; Liao et al. 2014; Meir et al. 2013; Ramamurthy et al. 2017; Salamanca et al. 2018; Wang and Li 2019). Based on the findings from Figures 2 to 4 and Table 2, we choose the ACM2 scheme to perform simulations for the remaining HW cases.

3.2 WRF-simulated urban-rural differences in temperatures and fluxes

In this section, we investigate the WRF-simulated urban-rural differences in surface and 2-m air temperatures and fluxes using average diurnal cycles over HW days. The average diurnal cycles present a general picture of urban-rural contrasts during HWs and how they differ between day and night and between Boston and Phoenix.

During the daytime (10am-4pm, local standard time), Phoenix is much hotter than Boston (Figure 5a & c), as expected from its lower latitude and desert climate. When the

urban and rural surface temperatures are compared, Boston shows a daytime SUHI, while Phoenix presents a daytime SUCI (Figure 5b). During the nighttime (10pm-4am, local standard time), both cities show a SUHI (Figure 5b). Compared to the surface temperature, the 2-m air temperature shows smaller differences between cities (Figure 5c) and also smaller urban-rural differences for each city (Figure 5d). As far as the daytime and nighttime average results are concerned, Phoenix shows almost no AUHI during both daytime and nighttime; Boston shows no AUHI during the daytime but a nighttime AUHI (Figure 5d). However, we also point out that there are some variations within the daytime/nighttime periods. For example, the urban-rural differences of 2-m air temperature in both cities drop in the early morning, which is possibly due to the rapid growth of the convective boundary layer as shown by Theeuwes et al. (2015).

The radiative fluxes including incoming shortwave and longwave radiation, and outgoing shortwave and longwave radiation are examined in Figure 6. During the daytime, the incoming shortwave radiation of Phoenix is on average larger than that of Boston due to its lower latitude (Figure 6a). Within each city, the incoming shortwave radiation is nearly identical for both urban and rural surfaces (Figure 6a). Note that we did not modify the aerosol profiles in WRF, and thus the urban-rural difference in terms of aerosol loading is not considered here. In contrast, the urban-rural difference in outgoing shortwave radiation is large (Figure 6c), which reflects the large urban-rural contrast of albedo considering that the urban-rural difference in incoming shortwave radiation is rather small. Specifically, the urban land of Boston has a higher surface albedo than the rural land of Boston (forests) and thus reflects more shortwave radiation. Conversely, the urban land of

Phoenix has a smaller surface albedo than the rural land of Phoenix (shrubs) and thus reflects less shortwave radiation.

The urban-rural contrast in the incoming longwave radiation is positive for both cities (Figure 6b), which is consistent with the findings from previous studies using flux tower measurements in other metropolitan regions (Li et al. 2015; Ao et al. 2019). The incoming longwave radiation is dependent on the temperature and humidity profiles in the whole atmospheric column (Brutsaert 2005) and thus it is difficult to identify exactly which factors cause such differences. In terms of urban-rural difference in outgoing longwave radiation (Figure 6d), the results in both cities are consistent with the expectation that the outgoing longwave radiation difference should be of the same sign as the surface temperature difference (Figure 5a; i.e., a hotter surface tends to have larger outgoing longwave radiation).

Now we examine the turbulent fluxes. During the daytime, the urban sensible heat flux is much larger than the rural one in both cities (Figure 7a), with a significant time lag in Boston that has been also observed by other studies (Oke et al. 2017; Ramamurthy et al. 2014). To explain this daytime urban-rural difference in sensible heat flux, we first note that the urban-rural difference in T_a is very small for both cities (less than 0.1 K) (Figure 7c). This is again because T_a in our study refers to the potential temperature at the lowest level of the atmospheric model (~30 m). The dynamics of T_a are controlled by horizontal advection of heat and mixing in the boundary layer, which are complicated and simulated by the atmospheric component of the WRF model. However, it is not surprising that the urban-rural differences of T_a are smaller than the surface and 2-m air temperature differences due to the mixing power of surface-layer turbulence. With this in mind, we

further note that the urban-rural difference in sensible heat flux is strongly modulated by
 aerodynamic resistances (i.e., r_a & r'_a) according to Eqs. (2) and (6). During the daytime,
 the urban-rural difference in r_a is positive in Boston but negative in Phoenix (Figure 8a &
 b). This implies that the urban land is more efficient in convecting heat from the surface to
 the lower atmosphere than the rural land in Phoenix, but the opposite is true in Boston.
 This is broadly consistent with the fact that the rural land of Phoenix is short shrubs and
 thus has lower surface roughness, while the rural land of Boston is tall forests and thus has
 higher surface roughness. Typically, larger obstacles generate more turbulence with which
 the heat can be transferred more efficiently from the surface to the overlying air, although
 we stress that convective heat transfer is fundamentally different from momentum transfer
 (Brutsaert 1982; Garratt 1992). Here the aerodynamic resistance refers to the resistance to
 convective heat transfer, not to momentum transfer. In comparison, the urban-rural
 difference in r'_a is smaller for both Boston and Phoenix (Figure 8c & d). This highlights the
 important difference between r_a and r'_a and demonstrates the height-dependence of
 aerodynamic resistance as alluded to earlier. During the nighttime, sensible heat flux
 becomes negative (Figure 7a), indicating that the lower atmosphere in turn heats the surface.
 The magnitude of the urban sensible heat flux is smaller than the rural sensible heat flux in
 Boston but larger in Phoenix (Figure 7a), which indicates that the urban land in Boston
 (Phoenix) has a weaker (stronger) capacity to transfer heat from the overlying air to the
 surface than the rural land at night. Combined with the daytime results, this reflects that
 the urban land is less (more) efficient than the rural land in convecting heat between the
 land surface and the lowest model level in Boston (Phoenix).

During the daytime, the urban latent heat flux is much lower than the rural latent heat flux in Boston, while in Phoenix both urban and rural latent heat fluxes are extremely small (Figure 7b). In both cities, the urban-rural differences in latent heat flux are not caused by the differences in the specific humidity at the lowest level of the atmospheric model, which are very small (Figure 7d). According to Eq. (3), the latent heat flux is strongly modulated by surface resistance in addition to aerodynamic resistance. We use Eq. (3) to infer the surface resistance and find that the urban surface resistance is larger than the rural surface resistance in Boston during the daytime (Figure 8e). This explains why the urban latent heat flux is much smaller than the rural counterpart in Boston during the daytime. Both urban and rural surface resistances in Phoenix inferred from Eq. (3) are exceptionally large due to the desert climate (Figure 8f). During the nighttime, there are rather small urban-rural differences in latent heat flux for both cities as urban and rural latent heat fluxes tend to be close to zero (Figure 7b).

Last but not the least, we examine the ground heat flux. Between the two cities, the magnitude of rural ground heat flux differs strongly: the forests in Boston have much smaller ground heat fluxes than the shrubs in Phoenix (Figure 9). This has been observed before and is due to the difference in the canopy height: a taller canopy tends to reduce the shortwave radiation reaching the ground and thus has less ground heat flux (Garratt 1992). When the urban and rural ground heat fluxes are compared, we find that in both cities, the urban ground heat flux is on average larger than the rural ground heat flux during the daytime (Figure 9), but is more negative during the nighttime. The results indicate that in both cities the urban land stores more energy into the ground than the rural land during the

daytime. As a result, the urban surface and near-surface air are heated more at night due to the heat storage release.

3.3 Attribution results

While the results shown in the previous section are informative, they do not directly quantify the contributions of different factors to the magnitude of UHIs. To do so, we apply the TRM attribution method to the WRF-simulated UHIs and UCIs in terms of surface temperature T_s (section 3.3.1) and 2-m air temperature T_2 (section 3.3.2). The comparison between Boston and Phoenix results is presented in section 3.3.3.

The credibility of the TRM method is supported by the consistency between the TRM-modelled (Figures 10 & 11, red bars) and the WRF-simulated (Figures 10 & 11, orange bars) urban-rural temperature differences. Based on the TRM-modelled results, we find a SUHI ($\Delta T_s = 6.1$ K) in Boston but a SUCI ($\Delta T_s = -3.3$ K) in Phoenix during the daytime (Figure 10). At night, we find a SUHI ($\Delta T_s = 3.3$ K) in Boston and also a SUHI ($\Delta T_s = 3.7$ K) in Phoenix. In terms of the AUHI, we find almost no daytime AUHIs in both cities (Figure 11). At night, we find an AUHI ($\Delta T_2 = 2.4$ K) in Boston and no AUHI in Phoenix. In the following, we discuss the individual contributions from different factors to SUHIs and AUHIs (Figures 10 & 11, blue bars). The individual sensitivities and urban-rural differences are presented in Tables 3 to 5. Note that the values reported in Tables 3 to 5 are the median values of the results over all HW days.

3.3.1 SUHI attribution results

3.3.1.1 Boston

During the daytime, the urban-rural difference in the surface temperature (ΔT_s) in Boston is mainly contributed by r_s (+73.3%), r_a (+69.9%), G (-30.2%), and α (-15.9%) (see

Figure 10a). Hence the hotter urban surface in Boston during the daytime, when compared to the rural surface (forests), is mainly caused by the drier nature (i.e., a larger urban r_s) and the lower heat transfer efficiency between the surface and the lower atmosphere (i.e., a larger urban r_a). The larger G and the larger α tend to lower the urban surface temperature, but do not overcome the positive contributions from the resistances. Specifically, the sensitivity of T_s to r_s is positive (Table 3), implying that a drier surface tends to be hotter. Therefore, the r_s makes a large positive contribution to the ΔT_s due to the larger urban r_s than the rural r_s in Boston (Table 5). The sensitivity of T_s to r_a is also positive (Table 3), indicating that a surface with a larger r_a tends to be hotter during the daytime. This is because with a larger r_a it is more difficult to convect the heat to the atmosphere. As the urban r_a is larger than the rural r_a in Boston (Table 5), the r_a also makes a large positive contribution to the ΔT_s . The sensitivity of T_s to α is negative (Table 3), showing that a darker surface tends to be hotter. Thus, the α makes a negative contribution to the ΔT_s as the urban surface has a larger α than the rural surface in Boston (Table 5). The sensitivity of T_s to G is also negative (Table 3), which means that a larger ground heat flux going into the soil or built materials tends to cause a cooler surface during the daytime. Therefore, the larger urban G (Table 5) results in a small negative contribution to the ΔT_s .

The nighttime ΔT_s in Boston is primarily contributed by G (see Figure 10b). That is, the hotter urban surface mostly results from the larger heat storage release at night. The sensitivity of T_s to G is negative (Table 3), which indicates that the surface becomes hotter when there is a larger release of the heat storage (i.e., a more negative G). Thus, when the

urban G is smaller (i.e., more negative) than the rural G (Table 5 and Figure 9a), the G makes a positive contribution to the ΔT_s .

3.3.1.2 Phoenix

In Phoenix, the daytime ΔT_s is mainly contributed by r_a (-159.7%) and α (+67.9%) (see Figure 10c). Hence the cooler urban surface in Phoenix during the daytime is mainly because the urban areas have a stronger convective heat transfer efficiency compared to the rural areas. Although the urban surface is darker in Phoenix, it does not overcome the negative contribution from the aerodynamic resistance.

The nighttime ΔT_s in Phoenix is mainly contributed by r_a (+56.9%), LW_{in} (+33.0%), and G (+22.1%). Namely, the hotter urban surface is caused by its lower urban aerodynamic resistance, the larger urban incoming longwave radiation, and the stronger heat storage release. While the latter two factors are straightforward to understand, the importance of r_a requires some explanation. The sensitivity of T_s to r_a is negative (Table 3), which indicates that a smaller r_a leads to a higher T_s . This is because more heat can be transferred downward to warm the surface at night. Therefore, the smaller urban r_a in Phoenix (Table 5) contributes positively to the ΔT_s . Here we highlight that the sensitivity of T_s to r_a changes its sign compared to the daytime counterpart. During the daytime, a surface with a larger r_a tends to be hotter due to the reduced efficiency to transfer sensible heat to the atmosphere. However, during the nighttime a surface with a smaller r_a tends to be hotter because it becomes easier for the atmosphere to transfer heat towards the surface. This diurnal asymmetry has been well documented in studies examining the impacts of deforestation on land surface temperature (Burakowski et al. 2018; Lee et al. 2011; Liao et al. 2018b; Schultz et al. 2017).

3.3.2 AUHI attribution results

Close inspection of the AUHI attribution results reveals that most factors work in a similar way as in the SUHI attribution results but with a smaller magnitude (Figures 10 & 11). This is understandable as the 2-m air temperature is effectively derived by interpolating between the surface temperature and the potential temperature at the lowest level of the atmospheric model (see Eq. 6). Hence most factors affecting SUHI also impact AUHI, but the magnitude of such impact is damped due to turbulent mixing between the surface and the 2-m level, which manifests as the difference between r_a and r'_a (i.e., the f in Eq. 7). The major differences between the AUHI and SUHI attribution results lie in the new contributions from r'_a and the different sensitivities of T_s and T_2 to r_a , which are of opposite signs as shown in Tables 3 and 4. We will focus on these differences in this section.

3.3.2.1 Boston

During the daytime, the urban-rural difference in the 2-m air temperature (ΔT_2) is mainly controlled by r_s and r'_a in Boston (see Figure 11a). The new contribution from r'_a is negative (Figure 11a). The sensitivity of T_2 to r'_a is three orders of magnitude larger than the sensitivity of T_2 to r_a (Table 4), which indicates that the 2-m air temperature is more controlled by r'_a than r_a . Moreover, the sensitivity of T_2 to r'_a is positive, which means that the 2-m air temperature tends to be smaller when the heat at the 2-m level can be more easily transferred upward during the daytime (i.e., with a smaller r'_a). The urban r'_a is slightly lower than the rural r'_a in Boston (Table 5), resulting in a negative contribution to the ΔT_2 .

During the nighttime, the ΔT_2 is predominately contributed by G in Boston (see Figure 11b). Nonetheless, the r'_a makes a negative contribution (Figure 11b). The sensitivity of T_2

to r'_a is negative (Table 4), which has an opposite sign compared to the daytime counterpart because the heat is transferred upward in the day but downward at night. As the r'_a becomes larger, it is more difficult for the heat to be transferred downward to warm the near-surface air at 2 m at night, leading to a lower T_2 . Therefore, when r'_a is larger in urban areas than in rural areas at night (Table 5), the r'_a makes a negative contribution to the ΔT_2 . The magnitude of the sensitivity of T_2 to r'_a is also 3.5 times that of the sensitivity of T_2 to r_a , explaining why the contribution from r_a is much smaller than that from r'_a .

3.3.2.2 Phoenix

In Phoenix, the daytime ΔT_2 is more controlled by r'_a than r_a , with also some positive contribution from α (Figure 11c). This is because the sensitivity of T_2 to r'_a is positive and two orders of magnitude larger than the sensitivity of T_2 to r_a (Table 4). This is similar to the Boston results and again indicates that a larger r'_a results in a higher T_2 because it is more difficult for heat at the 2-m level to be transferred to the atmosphere above during the daytime. Therefore, the larger urban r'_a (Table 5) leads to a positive contribution to the ΔT_2 .

During the nighttime, both r'_a and r_a make important contributions to ΔT_2 (Figure 11d), but they play opposite roles and roughly cancel each other. The sensitivity of T_2 to r'_a is negative at night (Table 4), which indicates that a lower r'_a raises T_2 as the near-surface air can gain heat more easily from the atmosphere above. Thus, the smaller urban r'_a makes a positive contribution to the ΔT_2 . On the other hand, the r_a makes a negative contribution. The contribution from r_a changes its sign when compared to the SUHI results because the sensitivity of T_2 to r_a always has the opposite sign as the sensitivity of T_s to r_a (Tables 3 & 4). Physically this means that convection either raises the surface temperature at the expense of reducing the near-surface air temperature, or raises the near-surface air

temperature at the expense of reducing the surface temperature. In this case, the smaller urban r_a causes more heat to be transferred from the overlying air to the surface, which positively contributes to SUHI (Figure 10) but negatively contributes to AUHI (Figure 11).

3.3.3 Differences in the attribution results between Boston and Phoenix

Comparing the attribution results between Boston and Phoenix, we find that the major contributors work in an opposite way during the daytime due to the opposite sign of the urban-rural difference in these factors (Table 5). First, compared to the urban land, the rural land in Boston has a lower r_a but the rural land in Phoenix has a higher r_a . Thus, the urban-rural difference in the aerodynamic resistance between the surface and the lowest atmospheric model level (Δr_a) is positive in Boston but negative in Phoenix. Therefore, the r_a contribution to ΔT_s is positive in Boston, but negative in Phoenix (Figures 10 & 11). Second, the urban-rural difference in the aerodynamic resistance between the 2-m level and the lowest atmospheric model level ($\Delta r'_a$) is negative in Boston but positive in Phoenix during the daytime (Table 5). Therefore, the contribution of r'_a to ΔT_2 is negative in Boston but positive in Phoenix (Figure 11). Third, compared to urban areas, rural areas (forests) in Boston have lower albedo (α) while rural areas (shrubs) in Phoenix have higher α (Table 5). Thus, the contributions of α to ΔT_s and ΔT_2 are negative in Boston, but positive in Phoenix (Figures 10 & 11). Fourth, the urban-rural difference in surface resistance (Δr_s) is positive in Boston (Table 5), which leads to positive contributions to ΔT_s and ΔT_2 . But the urban-rural difference in surface resistance (Δr_s) has almost no effects on ΔT_s and ΔT_2 in Phoenix (Figures 10 & 11) due to the much lower sensitivities ($\partial T_s / \partial r_s$ & $\partial T_2 / \partial r_s$) (Tables 3 & 4). This indicates that the surface temperature is more sensitive to the surface moisture

availability in Boston than in Phoenix, possibly due to the already very dry conditions in Phoenix.

The comparison of the nighttime results is simpler. The nighttime SUHI and AUHI in Boston are predominantly caused by the stronger urban heat storage release. For Phoenix, the contributions from r_a and r'_a are large, but the heat storage release still plays a role especially for SUHI.

4 Discussions

4.1 Comparison to previous studies

Our SUHI attribution results can shed important insights on the recent debate about the spatial variability of daytime SUHIs over North America. Cities with more annual mean rainfall, such as Boston in our case, tend to have stronger daytime SUHIs than those with less rainfall, such as Phoenix in our case, as shown by a number of previous studies (Clinton and Gong 2013; Imhoff et al. 2010; Li et al. 2019; Peng et al. 2012; Zhao et al. 2014). Zhao et al. (2014) found that aerodynamic resistance is the dominant factor controlling the daytime SUHI spatial variability across cities in North America at the annual mean scale. On the other hand, Li et al. (2019) found that it is surface resistance that more strongly controls the daytime SUHI spatial variability at the annual mean scale, although the attribution results in Li et al. (2019) also indicated an important role of aerodynamic resistance. Our results show that both aerodynamic resistance and surface resistance contribute significantly to the differences between Boston and Phoenix results in the daytime, which is consistent with the findings of Li et al. (2019) and Manoli et al. (2019) albeit at vastly different temporal scales. The importance of aerodynamic resistance for the SUHI spatial variability results from the fact that the urban-rural difference in

aerodynamic resistance shows opposite signs in the two cities. Our results show that in Boston the aerodynamic resistance to convective heat transfer between the surface and the lower atmosphere is larger in urban areas than in rural areas, while the opposite is true in Phoenix. On the other hand, the importance of surface resistance for the daytime SUHI variability stems from the fact that the urban-rural difference in surface resistance contributes strongly to the SUHI in Boston but not in Phoenix.

Our nighttime results in Boston are consistent with the traditional paradigm that the stronger release of ground heat flux in urban areas predominantly causes the nighttime SUHIs (Oke et al. 2017). However, the results in Phoenix suggest that the role of aerodynamic resistance cannot be neglected. When the land surface becomes cooler than the atmosphere at night, the sensible heat goes from the atmosphere to the surface. This transfer of heat is more efficient over surfaces with a lower r_a (the urban surface in Phoenix's case) and hence positively contributes to the nighttime SUHI. While this effect is less documented in observational studies on SUHIs, it has been demonstrated in studies on the impacts of deforestation on land surface temperature (Burakowski et al. 2018; Lee et al. 2011; Liao et al. 2018b; Schultz et al. 2017).

4.2 Limitations of our study

There are a few limitations of our study that are important to appreciate. First, we do not consider anthropogenic heat flux (AF) in our simulations and attributions. Within the confines of our attribution method (see Eq. 4), the sensitivity of T_s and T_2 to AF should be identical to that to ground heat flux (G) but with an opposite sign, which is on the order of 10^{-2} to 10^{-1} K m² W⁻¹ for T_s and on the order of 10^{-3} to 10^{-2} K m² W⁻¹ for T_2 . Using the fixed diurnal profile for AF prescribed in the single-layer urban canopy model (Kusaka et

al. 2001; Kusaka and Kimura 2004), the urban-rural difference in AF is estimated to be on the order of 75 W m^{-2} (including both sensible and latent anthropogenic heat fluxes). Thus, the contributions of AF are estimated to be around 0.75 to 7.5 K for the SUHI and 0.075 to 0.75 K for the AUHI. We point out that the simple estimate presented here might suffer from the uncertainties associated with the magnitude of AF (Allen et al. 2011; Salamanca et al. 2014). Moreover, the simulated sensitivities of T_s and T_2 to AF depend on how exactly AF is incorporated in the energy balance computation in WRF, and might be different from the values shown here. While these complications are left for future investigations, we note that the signs of the sensitivities of T_s and T_2 to AF and the urban-rural contrast of AF are expected to be positive. Hence AF always tends to make positive contributions to the UHIs.

Second, our analysis only focuses on HW days and thus the findings might not be extrapolated to other weather conditions. We do not investigate the responses of UHIs to HWs by comparing the results in pre-HW, HW, and post-HW days. Many studies have shown that there exist synergies between UHIs and HWs, namely, the magnitude of UHIs is larger under HWs (e.g., Ao et al. 2019; Founda et al. 2015; Founda and Santamouris 2017; Li and Bou-Zeid 2013; Li et al. 2015; Ramamurthy et al. 2017; Schatz and Kucharik 2015). However, there are also studies reporting that the magnitude of UHIs is reduced under HW conditions (e.g., Rogers et al. 2019; Scott et al. 2018). Our attribution method may shed new insights into this debate but this is left for future research.

Third, as with most studies using models like WRF, the validation is centered on land surface temperature patterns, 2-m air temperatures at a number of weather stations, as well as the boundary layer temperature profiles at the two airports due to data availability. Given

the lack of observations, we could not directly validate the fluxes from our simulations. More thorough validation, especially on parameters and variables related to near-surface heat transfer, is recommended.

5. Conclusions

In the present study, we investigate the controlling factors of UHIs and UCIs during HWs in Boston and Phoenix using WRF simulations. During the daytime, we find a SUHI in Boston mainly caused by the higher urban r_s that reduces the latent heat flux, and the higher urban r_a which inhibits convective heat transfer from the urban surface to the atmosphere. In contrast, we find a daytime SUCI in Phoenix mainly due to the lower urban r_a that facilitates convective heat transfer. At night, we identify a SUHI and an AUHI in Boston due to the stronger release of urban heat storage. In comparison, the lower urban r_a in Phoenix facilitates convective heat transfer from the atmosphere to the urban surface at night, leading to a SUHI but no AUHI. Our study highlights that the magnitude of UHIs or UCIs is strongly controlled by urban-rural differences in terms of aerodynamic features, vegetation and moisture conditions, and heat storage, which show contrasting characteristics in different regions. Further investigations on the roles of anthropogenic heat flux and weather conditions, and more thorough validation of the simulated results are recommended.

Acknowledgments: This material is based upon work supported by the U.S. National Science Foundation (Grant No. ICER-1854706) and the U.S. Army Research Office (Grant No. W911NF-18-1-0360). We acknowledge the high-performance computing support from Boston University Shared Computing Cluster and from Cheyenne

690 (doi:10.5065/D6RX99HX) provided by NCAR's Computational and Information Systems
691 Laboratory, sponsored by the National Science Foundation. We also thank the three
692 reviewers whose comments led to significant improvement of the quality of our paper.

References

- Allen, L., Lindberg, F., Grimmond, C.S.B., 2011: Global to city scale urban anthropogenic heat flux: model and variability. *International Journal of Climatology*, **31**, 1990–2005, [10.1002/joc.2210](https://doi.org/10.1002/joc.2210).
- Anderson, G.B., Bell, M.L., 2009: Weather-related mortality: how heat, cold, and heat waves affect mortality in the United States. *Epidemiology*, **20**, 205–213, <https://doi.org/10.1097/EDE.0b013e318190ee08>.
- Anderson, G.B., Bell, M.L., 2011: Heat waves in the United States: mortality risk during heat waves and effect modification by heat wave characteristics in 43 U.S. communities. *Environ Health Perspect*, **119**, 210–218, <https://doi.org/10.1289/ehp.1002313>.
- Ao, X., Wang, L., Zhi, X., Gu, W., Yang, H., Li, D., 2019: Observed synergies between urban heat islands and heat waves and their controlling factors in shanghai, China. *J. Appl. Meteor. Climatol*, **58**, 1955–1972, <https://doi.org/10.1175/JAMC-D-19-0073.1>.
- Arnfield, A.J., 2003: Two decades of urban climate research: a review of turbulence, exchanges of energy and water, and the urban heat island. *International Journal of Climatology*, **23**, 1–26, <https://doi.org/10.1002/joc.859>.
- Brown, S.J., Caesar, J., Ferro, C.A.T., 2008: Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research: Atmospheres*, **113**, D05115, <https://doi.org/10.1029/2006JD008091>.
- Brutsaert, W., 1982: *Evaporation into the Atmosphere: Theory, History and Applications*. Springer Netherlands, 302 pp.
- Brutsaert, W., 2005: *Hydrology: An Introduction*. Cambridge University Press, 605 pp.

715 Burakowski, E., Tawfik, A., Ouimette, A., Lepine, L., Novick, K., Ollinger, S., Zarzycki,
 716 C., Bonan, G., 2018: The role of surface roughness, albedo, and Bowen ratio on ecosystem
 717 energy balance in the Eastern United States. *Agricultural and Forest Meteorology*, **249**,
 718 367–376, <https://doi.org/10.1016/j.agrformet.2017.11.030>.
 719 Campbell, S., Remenyi, T.A., White, C.J., Johnston, F.H., 2018: Heatwave and health
 720 impact research: A global review. *Health & Place*, **53**, 210–218,
 721 <https://doi.org/10.1016/j.healthplace.2018.08.017>.
 722 Chen, C., Wang, L., Ranga, M., Li, D., 2020: Attribution of land-use/land-cover change
 723 induced surface temperature anomaly: How accurate is the first-order Taylor series
 724 expansion? *Journal of Geophysical Research: Biogeosciences*, **125**, e2020JG005787,
 725 <https://doi.org/10.1029/2020JG005787>.
 726 Chen, F., and Coauthors, 2011: The integrated WRF/urban modelling system:
 727 Development, evaluation, and applications to urban environmental problems. *International*
 728 *Journal of Climatology*, **31**, 273–288, <https://doi.org/10.1002/joc.2158>.
 729 Chen, Y., Zhai, P., 2017: Revisiting summertime hot extremes in China during 1961–2015:
 730 Overlooked compound extremes and significant changes. *Geophysical Research Letters*,
 731 **44**, 5096–5103, <https://doi.org/10.1002/2016GL072281>.
 732 Clinton, N., Gong, P., 2013: MODIS detected surface urban heat islands and sinks: Global
 733 locations and controls. *Remote Sensing of Environment*, **134**, 294–304,
 734 <https://doi.org/10.1016/j.rse.2013.03.008>.
 735 Dole, R., and Coauthors, 2011: Was there a basis for anticipating the 2010 Russian heat
 736 wave? *Geophysical Research Letters*, **38**, L06702,
 737 <https://doi.org/10.1029/2010GL046582>.

738 Donat, M.G., and Coauthors, 2013: Updated analyses of temperature and precipitation
 739 extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal*
 740 *of Geophysical Research: Atmospheres*, **118**, 2098–2118,
 741 <https://doi.org/10.1002/jgrd.50150>.

742 Dudhia, J., 1988: Numerical study of convection observed during the winter monsoon
 743 experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.*, **46**, 3077–3107,
 744 [https://doi.org/10.1175/1520-0469\(1989\)046<3077:NSOCOD>2.0.CO;2](https://doi.org/10.1175/1520-0469(1989)046<3077:NSOCOD>2.0.CO;2).

745 Ek, M. B., Mitchell, K.E., Lin, Y., Rogers, E., Grunmann, P., Koren, V., Gayno, G.,
 746 Tarpley, J.D., 2003: Implementation of Noah land surface model advances in the National
 747 Centers for Environmental Prediction operational mesoscale Eta model. *J. Geophys. Res.*,
 748 **108**, 8851, <https://doi.org/10.1029/2002JD003296>.

749 Fischer, E.M., Schär, C., 2010: Consistent geographical patterns of changes in high-impact
 750 European heatwaves. *Nature Geoscience*, **3**, 398–403, <https://doi.org/10.1038/ngeo866>.

751 Fontana, G., Toreti, A., Ceglar, A., De Sanctis, G., 2015: Early heat waves over Italy and
 752 their impacts on durum wheat yields. *Natural Hazards and Earth System Sciences*, **15**,
 753 1631–1637, <https://doi.org/10.5194/nhess-15-1631-2015>.

754 Founda, D., Pierros, F., Petrakis, M., Zerefos, C., 2015: Interdecadal variations and trends
 755 of the Urban Heat Island in Athens (Greece) and its response to heat waves. *Atmospheric*
 756 *Research*, **161–162**, 1–13, <https://doi.org/10.1016/j.atmosres.2015.03.016>.

757 Founda, D., Santamouris, M., 2017: Synergies between urban heat island and heat waves
 758 in Athens (Greece), during an extremely hot summer (2012). *Scientific Reports*, **7**, 10973,
 759 <https://doi.org/10.1038/s41598-017-11407-6>.

760 García-Herrera, R., Díaz, J., Trigo, R.M., Luterbacher, J., Fischer, E.M., 2010: A review
 761 of the European summer heat wave of 2003. *Critical Reviews in Environmental Science*
 762 *and Technology*, **40**, 267–306, <https://doi.org/10.1080/10643380802238137>.

763 Garratt, J.R., 1992: *The Atmospheric Boundary Layer*. Cambridge University Press, 316
 764 pp.

765 Georgescu, M., Moustauoui, M., Mahalov, A., Dudhia, J., 2011: An alternative explanation
 766 of the semiarid urban area “oasis effect”. *J. Geophys. Res.*, **116**, D24113,
 767 <https://doi.org/10.1029/2011JD016720>.

768 Grimm, N.B., Faeth, S.H., Golubiewski, N.E., Redman, C.L., Wu, J., Bai, X., Briggs, J.M.,
 769 2008: Global change and the ecology of cities. *Science*, **319**, 756–760,
 770 <https://doi.org/10.1126/science.1150195>.

771 Grimmond, S., 2007: Urbanization and global environmental change: Local effects of
 772 urban warming. *The Geographical Journal*, **173**, 83–88, [https://doi.org/10.1111/j.1475-](https://doi.org/10.1111/j.1475-4959.2007.232_3.x)
 773 [4959.2007.232_3.x](https://doi.org/10.1111/j.1475-4959.2007.232_3.x).

774 Hajat, S., Armstrong, B., Baccini, M., Biggeri, A., Bisanti, L., Russo, A., Paldy, A., Menne,
 775 B., Kosatsky, T., 2006: Impact of high temperatures on mortality: is there an added heat
 776 wave effect? *Epidemiology*, **17**, 632–638.

777 Han, J.Y., Baik, J.J., Lee, H., 2014: Urban impacts on precipitation. *Asia-Pac. J.*
 778 *Atmospheric Sci*, **50**, 17–30, <https://doi.org/10.1007/s13143-014-0016-7>.

779 Heaviside, C., Cai, X.M., Vardoulakis, S., 2015: The effects of horizontal advection on the
 780 urban heat island in Birmingham and the West Midlands, United Kingdom during a
 781 heatwave. *Q. J. R. Meteorol. Soc.*, **141**, 1429–1441, <https://doi.org/10.1002/qj.2452>.

782 Hidalgo, J., Masson, V., Gimeno, L., 2009: Scaling the daytime urban heat island and
 783 urban-breeze circulation. *J. Appl. Meteorol. Climatol.*, **49**, 889–901,
 784 <https://doi.org/10.1175/2009JAMC2195.1>.
 785 Homer, C., and Coauthors, 2015: Completion of the 2011 National Land Cover Database
 786 for the conterminous United States – Representing a decade of land cover change
 787 information. *Photogrammetric Engineering and Remote Sensing*, **81**, 345–354.
 788 Hong, S.Y., Lim, J.O.J., 2006: The WRF single-moment 6-class microphysics scheme
 789 (WSM6). *J. Korean Meteor. Soc.*, **42**, 129–151.
 790 Hong, S.Y., Noh, Y., Dudhia, J., 2006: A new vertical diffusion package with an explicit
 791 treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341,
 792 <https://doi.org/10.1175/MWR3199.1>.
 793 Imhoff, M.L., Zhang, P., Wolfe, R.E., Bounoua, L., 2010: Remote sensing of the urban
 794 heat island effect across biomes in the continental USA. *Remote Sensing of Environment*,
 795 **114**, 504–513, <https://doi.org/10.1016/j.rse.2009.10.008>.
 796 Kalverla, P.C., Duine, G.J., Steeneveld, G.J., Hedde, T., 2016: Evaluation of the Weather
 797 Research and Forecasting Model in the Durance Valley complex terrain during the
 798 KASCADE Field Campaign. *J. Appl. Meteorol. Climatol.*, **55**, 861–882,
 799 <https://doi.org/10.1175/JAMC-D-15-0258.1>.
 800 Kusaka, H., Kimura, F., 2004: Coupling a single-layer urban canopy model with a simple
 801 atmospheric model: Impact on urban heat island simulation for an idealized case. *Journal*
 802 *of the Meteorological Society of Japan*, **Ser. II 82**, 67–80,
 803 <https://doi.org/10.2151/jmsj.82.67>.

804 Kusaka, H., Kondo, H., Kikegawa, Y., Kimura, F., 2001: A simple single-layer urban
805 canopy model for atmospheric models: Comparison with multi-layer and slab models.
806 *Boundary-Layer Meteorology*, **101**, 329–358, <https://doi.org/10.1023/A:1019207923078>.

807 Lau, N.C., Nath, M.J., 2012: A model study of heat waves over North America:
808 Meteorological aspects and projections for the Twenty-First Century. *J. Climate*, **25**, 4761–
809 4784, <https://doi.org/10.1175/JCLI-D-11-00575.1>.

810 Lau, N.C., Nath, M.J., 2014: Model simulation and projection of European heat waves in
811 present-day and future climates. *J. Climate*, **27**, 3713–3730, [https://doi.org/10.1175/JCLI-](https://doi.org/10.1175/JCLI-D-13-00284.1)
812 [D-13-00284.1](https://doi.org/10.1175/JCLI-D-13-00284.1).

813 Lee, X., and Coauthors, 2011: Observed increase in local cooling effect of deforestation at
814 higher latitudes. *Nature*, **479**, 384–387, <https://doi.org/10.1038/nature10588>.

815 Li, D., Bou-Zeid, E., 2013: Synergistic interactions between urban heat islands and heat
816 waves: The impact in cities is larger than the sum of its parts. *J. Appl. Meteor. Climatol.*,
817 **52**, 2051–2064, <https://doi.org/10.1175/JAMC-D-13-02.1>.

818 Li, D., Bou-Zeid, E., 2014: Quality and sensitivity of high-resolution numerical simulation
819 of urban heat islands. *Environ. Res. Lett.*, **9**, 055001, [https://doi.org/10.1088/1748-](https://doi.org/10.1088/1748-9326/9/5/055001)
820 [9326/9/5/055001](https://doi.org/10.1088/1748-9326/9/5/055001).

821 Li, D., Liao, W., Rigden, A.J., Liu, X., Wang, D., Malyshev, S., Shevliakova, E., 2019:
822 Urban heat island: Aerodynamics or imperviousness? *Science Advances*, **5**, eaau4299,
823 <https://doi.org/10.1126/sciadv.aau4299>.

824 Li, D., Sun, T., Liu, M., Yang, L., Wang, L., Gao, Z., 2015: Contrasting responses of urban
825 and rural surface energy budgets to heat waves explain synergies between urban heat

826 islands and heat waves. *Environ. Res. Lett.*, **10**, 054009, [https://doi.org/10.1088/1748-](https://doi.org/10.1088/1748-9326/10/5/054009)
827 9326/10/5/054009.

828 Li, D., Wang, L., 2019: Sensitivity of surface temperature to land use and land cover
829 change-induced biophysical changes: The scale issue. *Geophysical Research Letters*, **46**,
830 9678–9689, <https://doi.org/10.1029/2019GL084861>.

831 Li, H., Zhou, Y., Wang, X., Zhou, X., Zhang, H., Sodoudi, S., 2018: Quantifying urban
832 heat island intensity and its physical mechanism using WRF/UCM. *Sci. Total Environ.*,
833 **650**, 3110–3119, <https://doi.org/10.1016/j.scitotenv.2018.10.025>.

834 Liao, J., Wang, T., Wang, X., Xie, M., Jiang, Z., Huang, X., Zhu, J., 2014: Impacts of
835 different urban canopy schemes in WRF/Chem on regional climate and air quality in
836 Yangtze River Delta, China. *Atmospheric Res.*, **145–146**, 226–243,
837 <https://doi.org/10.1016/j.atmosres.2014.04.005>.

838 Liao, W., and Coauthors, 2018a: Stronger contributions of urbanization to heat wave trends
839 in wet climates. *Geophysical Research Letters*, **45**, 11,310–11,317,
840 <https://doi.org/10.1029/2018GL079679>.

841 Liao, W., Rigden, A.J., Li, D., 2018b: Attribution of local temperature response to
842 deforestation. *Journal of Geophysical Research: Biogeosciences*, **123**, 1572–1587,
843 <https://doi.org/10.1029/2018JG004401>.

844 Luo, M., Lau, N.C., 2016: Heat waves in southern China: Synoptic behavior, long-term
845 change, and urbanization effects. *J. Climate*, **30**, 703–720, [https://doi.org/10.1175/JCLI-](https://doi.org/10.1175/JCLI-D-16-0269.1)
846 D-16-0269.1.

847 Manoli, G., and Coauthors, 2019: Magnitude of urban heat islands largely explained by
848 climate and population. *Nature*, **573**, 55–60, <https://doi.org/10.1038/s41586-019-1512-9>.

849 Meehl, G.A., Tebaldi, C., 2004: More intense, more frequent, and longer lasting heat waves
850 in the 21st Century. *Science*, **305**, 994–997, <https://doi.org/10.1126/science.1098704>.

851 Meir, T., Orton, P.M., Pullen, J., Holt, T., Thompson, W.T., Arend, M.F., 2013:
852 Forecasting the New York City urban heat island and sea breeze during extreme heat
853 events. *Weather Forecast*, **28**, 1460–1477, <https://doi.org/10.1175/WAF-D-13-00012.1>.

854 Mellor, G.L., Yamada, T., 1974: A hierarchy of turbulence closure models for planetary
855 boundary layers. *J. Atmos. Sci.*, **31**, 1791–1806, [https://doi.org/10.1175/1520-0469\(1974\)031<1791:AHOTCM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<1791:AHOTCM>2.0.CO;2).

857 Miao, S., Chen, F., LeMone, M.A., Tewari, M., Li, Q., Wang, Y., 2009: An observational
858 and modeling study of characteristics of urban heat island and boundary layer structures in
859 Beijing. *J. Appl. Meteorol. Climatol.*, **48**, 484–501,
860 <https://doi.org/10.1175/2008JAMC1909.1>.

861 Mlawer, E.J., Taubman, S.J., Brown, P.D., Iacono, M.J., Clough, S.A., 1997: Radiative
862 transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the
863 longwave. *Journal of Geophysical Research*, **102**, 16, <https://doi.org/10.1029/97JD00237>.

864 Monteith, J., Unsworth, M., 2007: *Principles of Environmental Physics*. Academic Press,
865 440 pp.

866 Moon, M., Li, D., Liao, W., Rigden, A.J., Friedl, M.A., 2020: Modification of surface
867 energy balance during springtime: The relative importance of biophysical and
868 meteorological changes. *Agricultural and Forest Meteorology*, **284**, 107905,
869 <https://doi.org/10.1016/j.agrformet.2020.107905>.

870 Mora, C., and Coauthors, 2017: Global risk of deadly heat. *Nature Climate Change*, **7**,
871 501–506, <https://doi.org/10.1038/nclimate3322>.

872 Oke, T.R., 1982: The energetic basis of the urban heat island. *Quarterly Journal of the*
873 *Royal Meteorological Society*, **108**, 1–24, <https://doi.org/10.1002/qj.49710845502>.

874 Oke, T.R., Mills, G., Christen, A., Voogt, J.A., 2017: *Urban Climate*. Cambridge
875 University Press, 546 pp.

876 Pal, S., and Coauthors, 2012: Spatio-temporal variability of the atmospheric boundary layer
877 depth over the Paris agglomeration: An assessment of the impact of the urban heat island
878 intensity. *Atmos. Environ.*, **63**, 261–275, <https://doi.org/10.1016/j.atmosenv.2012.09.046>.

879 Parente, J., Pereira, M.G., Amraoui, M., Fischer, E.M., 2018: Heat waves in Portugal:
880 Current regime, changes in future climate and impacts on extreme wildfires. *Sci. Total*
881 *Environ.*, **631–632**, 534–549, <https://doi.org/10.1016/j.scitotenv.2018.03.044>.

882 Patz, J.A., Campbell-Lendrum, D., Holloway, T., Foley, J.A., 2005: Impact of regional
883 climate change on human health. *Nature*, **438**, 310–317,
884 <https://doi.org/10.1038/nature04188>.

885 Peng, R.D., Bobb, J.F., Tebaldi, C., McDaniel, L., Bell, M.L., Dominici, F., 2011: Toward
886 a quantitative estimate of future heat wave mortality under global climate change. *Environ.*
887 *Health Perspect.*, **119**, 701–706, <https://doi.org/10.1289/ehp.1002430>.

888 Peng, S., and Coauthors, 2012: Surface urban heat island across 419 global big cities.
889 *Environ. Sci. Technol.*, **46**, 696–703, <https://doi.org/10.1021/es2030438>.

890 Perkins, S.E., 2015: A review on the scientific understanding of heatwaves—Their
891 measurement, driving mechanisms, and changes at the global scale. *Atmospheric Research*,
892 **164–165**, 242–267, <https://doi.org/10.1016/j.atmosres.2015.05.014>.

893 Petersen, R.A., 2016. On the impact and benefits of AMDAR observations in operational
894 forecasting—Part I: A review of the impact of automated aircraft wind and temperature

895 Reports. *Bull. Am. Meteorol. Soc.*, **97**, 585–602, <https://doi.org/10.1175/BAMS-D-14->
896 00055.1.

897 Petkova, E.P., Morita, H., Kinney, P.L., 2014: Health impacts of heat in a changing climate:
898 how can emerging science inform urban adaptation planning? *Curr Epidemiol Rep*, **1**, 67–
899 74, <https://doi.org/10.1007/s40471-014-0009-1>.

900 Pleim, J.E., 2007: A combined local and nonlocal closure model for the atmospheric
901 boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–
902 1395, <https://doi.org/10.1175/JAM2539.1>.

903 Ramamurthy, P., Bou-Zeid, E., 2016: Heatwaves and urban heat islands: A comparative
904 analysis of multiple cities. *Journal of Geophysical Research: Atmospheres*, **122**, 168–178,
905 <https://doi.org/10.1002/2016JD025357>.

906 Ramamurthy, P., Bou-Zeid, E., Smith, J.A., Wang, Z., Baeck, M.L., Saliendra, N.Z., Hom,
907 J.L., Welty, C., 2014: Influence of subfacet heterogeneity and material properties on the
908 urban surface energy budget. *J. Appl. Meteor. Climatol.*, **53**, 2114–2129,
909 <https://doi.org/10.1175/JAMC-D-13-0286.1>.

910 Ramamurthy, P., Li, D., Bou-Zeid, E., 2017: High-resolution simulation of heatwave
911 events in New York City. *Theor Appl Climatol*, **128**, 89–102,
912 <https://doi.org/10.1007/s00704-015-1703-8>.

913 Rigden, A.J., Li, D., 2017: Attribution of surface temperature anomalies induced by land
914 use and land cover changes. *Geophysical Research Letters*, **44**, 6814–6822,
915 <https://doi.org/10.1002/2017GL073811>.

916 Robinson, P.J., 2001: On the definition of a heat wave. *J. Appl. Meteor.*, **40**, 762–775,
917 [https://doi.org/10.1175/1520-0450\(2001\)040<0762:OTDOAH>2.0.CO;2](https://doi.org/10.1175/1520-0450(2001)040<0762:OTDOAH>2.0.CO;2).

918 Rogers, C.D.W., Gallant, A.J.E., Tapper, N.J., 2019: Is the urban heat island exacerbated
919 during heatwaves in southern Australian cities? *Theoretical and Applied Climatology*, **137**,
920 441-457, <https://doi.org/10.1007/s00704-018-2599-x>.

921 Rübbelke, D.T.G., Vögele, S., 2011: Impacts of climate change on European critical
922 infrastructures: The case of the power sector. *Environmental Science & Policy*, **14**, 53–63,
923 <https://doi.org/10.1016/j.envsci.2010.10.007>.

924 Salamanca, F., Georgescu, M., Mahalov, A., Moustauoi, M., Wang, M., 2014:
925 Anthropogenic heating of the urban environment due to air conditioning. *Journal of*
926 *Geophysical Research: Atmospheres*, **119**, 5949-5965,
927 <https://doi.org/10.1002/2013JD021225>.

928 Salamanca, F., Zhang, Y., Barlage, M., Chen, F., Mahalov, A., Miao, S., 2018: Evaluation
929 of the WRF-urban modeling system coupled to Noah and Noah-MP land surface models
930 over a semiarid urban environment. *Journal of Geophysical Research: Atmospheres*, **123**,
931 2387-2408, <https://doi.org/10.1002/2018JD028377>.

932 Schär, C., Vidale, P.L., Lüthi, D., Frei, C., Häberli, C., Liniger, M.A., Appenzeller, C.,
933 2004: The role of increasing temperature variability in European summer heatwaves.
934 *Nature*, **427**, 332, <https://doi.org/10.1038/nature02300>.

935 Schatz, J., Kucharik, C.J., 2015: Urban climate effects on extreme temperatures in
936 Madison, Wisconsin, USA. *Environ. Res. Lett.*, **10**, 094024, [https://doi.org/10.1088/1748-](https://doi.org/10.1088/1748-9326/10/9/094024)
937 [9326/10/9/094024](https://doi.org/10.1088/1748-9326/10/9/094024).

938 Schultz, N.M., Lawrence, P.J., Lee, X., 2017: Global satellite data highlights the diurnal
939 asymmetry of the surface temperature response to deforestation. *Journal of Geophysical*
940 *Research: Biogeosciences*, **122**, 903–917, <https://doi.org/10.1002/2016JG003653>.

941 Scott, A.A., Waugh, D.W., Zaitchik, B.F., 2018: Reduced Urban Heat Island intensity
 942 under warmer conditions. *Environ. Res. Lett.*, **13**, 064003, [https://doi.org/10.1088/1748-](https://doi.org/10.1088/1748-9326/aabd6c)
 943 9326/aabd6c.

944 Sillmann, J., Kharin, V.V., Zwiers, F.W., Zhang, X., Bronaugh, D., 2013: Climate extremes
 945 indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of*
 946 *Geophysical Research: Atmospheres*, **118**, 2473–2493, <https://doi.org/10.1002/jgrd.50188>.

947 Stewart, I.D., 2011. A systematic review and scientific critique of methodology in modern
 948 urban heat island literature. *Int. J. Climatol.*, **31**, 200–217,
 949 <https://doi.org/10.1002/joc.2141>.

950 Taha, H., 1997: Urban climates and heat islands: albedo, evapotranspiration, and
 951 anthropogenic heat. *Energy and Buildings*, **25**, 99–103, [https://doi.org/10.1016/S0378-](https://doi.org/10.1016/S0378-7788(96)00999-1)
 952 7788(96)00999-1.

953 Tan, J., and Coauthors, 2010: The urban heat island and its impact on heat waves and
 954 human health in Shanghai. *Int J Biometeorol*, **54**, 75–84, [https://doi.org/10.1007/s00484-](https://doi.org/10.1007/s00484-009-0256-x)
 955 009-0256-x.

956 Tebaldi, C., Hayhoe, K., Arblaster, J.M., Meehl, G.A., 2006: Going to the extremes: An
 957 intercomparison of model-simulated historical and future changes in extreme events.
 958 *Climatic Change*, **79**, 185–211, <https://doi.org/10.1007/s10584-006-9051-4>.

959 Tewari, M., Yang, J., Kusaka, H., Salamanca, F., Watson, C., Treinish, L., 2019:
 960 Interaction of urban heat islands and heat waves under current and future climate conditions
 961 and their mitigation using green and cool roofs in New York City and Phoenix, Arizona.
 962 *Environ. Res. Lett.*, **14**, 034002, <https://doi.org/10.1088/1748-9326/aaf431>.

963 Theeuwes, N.E., Steeneveld, G.J., Ronda, R.J., Rotach, M.W., Holtslag, A.A.M., 2015:
 964 Cool city mornings by urban heat. *Environ. Res. Lett.*, **10**, 114022,
 965 <https://doi.org/10.1088/1748-9326/10/11/114022>.
 966 United Nations, 2019: *World Urbanization Prospects: The 2018 Revision*. United Nations,
 967 103 pp.
 968 Wang, L., Li, D., 2019: Modulation of the urban boundary-layer heat budget by a heatwave.
 969 *Quarterly Journal of the Royal Meteorological Society*, **145**, 1814-1831,
 970 <https://doi.org/10.1002/qj.3526>.
 971 Wang, L., Li, D., Zhang, N., Sun, J., Guo, W., 2020: Surface urban heat and cool islands
 972 and their drivers: an observational study in Nanjing, China. *Journal of Applied*
 973 *Meteorology and Climatology*, **59**, 1987-2000, <https://doi.org/10.1175/JAMC-D-20->
 974 0089.1.
 975 Wang, P., Li, D., Liao, W., Rigden, A., Wang, W., 2019: Contrasting evaporative responses
 976 of ecosystems to heatwaves traced to the opposing roles of vapor pressure deficit and
 977 surface resistance. *Water Resources Research*, **55**, 4550–4563,
 978 <https://doi.org/10.1029/2019WR024771>.
 979 Wreford, A., Adger, W.N., 2010: Adaptation in agriculture: Historic effects of heat waves
 980 and droughts on UK agriculture. *International Journal of Agricultural Sustainability*, **8**,
 981 278–289, <https://doi.org/10.3763/ijas.2010.0482>.
 982 Xia, J., Tu, K., Yan, Z., Qi, Y., 2016: The super-heat wave in eastern China during July–
 983 August 2013: A perspective of climate change. *International Journal of Climatology*, **36**,
 984 1291–1298, <https://doi.org/10.1002/joc.4424>.

985 Xia, Y., and Coauthors, 2018: Assessment of the economic impacts of heat waves: A case
 986 study of Nanjing, China. *Journal of Cleaner Production*, **171**, 811–819,
 987 <https://doi.org/10.1016/j.jclepro.2017.10.069>.
 988 Yang, X., Leung, L.R., Zhao, N., Zhao, C., Qian, Y., Hu, K., Liu, X., Chen, B., 2017:
 989 Contribution of urbanization to the increase of extreme heat events in an urban
 990 agglomeration in east China. *Geophysical Research Letters*, **44**, 6940–6950,
 991 <https://doi.org/10.1002/2017GL074084>.
 992 Zampieri, M., Ceglar, A., Dentener, F., Toreti, A., 2017: Wheat yield loss attributable to
 993 heat waves, drought and water excess at the global, national and subnational scales.
 994 *Environ. Res. Lett.*, **12**, 064008, <https://doi.org/10.1088/1748-9326/aa723b>.
 995 Zhang, N., Wang, X., Peng, Z., 2014: Large-eddy simulation of mesoscale circulations
 996 forced by inhomogeneous urban heat island. *Bound. Layer Meteorol.*, **151**, 179–194,
 997 <https://doi.org/10.1007/s10546-013-9879-x>.
 998 Zhang, Y., Li, D., Lin, Z., Santanello, J.A., Gao, Z., 2019: Development and evaluation of
 999 a long-term data record of planetary boundary layer profiles from aircraft meteorological
 1000 reports. *Journal of Geophysical Research: Atmospheres*, **124**, 2008–2030,
 1001 <https://doi.org/10.1029/2018JD029529>.
 1002 Zhang, Y., Wang, L., Santanello, J.A., Pan, Z., Gao, Z., Li, D., 2020: Aircraft observed
 1003 diurnal variations of the planetary boundary layer under heat waves. *Atmospheric*
 1004 *Research*, **235**, 104801, <https://doi.org/10.1016/j.atmosres.2019.104801>.
 1005 Zhao, L., Lee, X., Smith, R.B., Oleson, K., 2014: Strong contributions of local background
 1006 climate to urban heat islands. *Nature*, **511**, 216–219, <https://doi.org/10.1038/nature13462>.
 1007

1008

Table 1 Identified heat waves (HWs) during 2007-2016 in Boston and Phoenix

Boston			Phoenix		
Case No.	Date	HW days	Case No.	Date	HW days
1*	20070625 – 20070628	4	21*	20070705 – 20070706	2
2	20070802 – 20070804	3	22	20070717 – 20070719	3
3	20090817 – 20090819	3	23	20070812 – 20070813	2
4	20100703 – 20100706	4	24	20080616 – 20080617	2
5	20100716 – 20100718	3	25	20090710 – 20090714	5
6	20100728 – 20100729	2	26	20090716 – 20090719	4
7*	20100829 – 20100902	5	27	20090727 – 20090728	2
8	20110717 – 20110718	2	28	20090803 – 20090806	4
9	20110721 – 20110723	3	29	20100714 – 20100716	3
10	20120620 – 20120622	3	30*	20110701 – 20110703	3
11	20120629 – 20120701	3	31	20110823 – 20110826	4
12	20120715 – 20120717	3	32	20110829 – 20110901	4
13	20130530 – 20130602	4	33	20120707 – 20120711	5
14	20130623 – 20130625	3	34	20120806 – 20120814	9
15	20130703 – 20130707	5	35	20130628 – 20130630	3
16*	20130714 – 20130720	7	36	20130816 – 20130817	2
17	20160712 – 20160716	5	37	20140722 – 20140724	3
18	20160721 – 20160723	3	38	20150617 – 20150619	3
19	20160811 – 20160812	2	39	20150814 – 20150816	3
20	20160814 – 20160815	2	40*	20160619 – 20160620	2
			41	20160726 – 20160728	3

Note. The asterisk indicates that the HW is selected for validation.

1009

1010

1011

1012

1013

1014

Table 2 Averaged root-mean-square errors (RMSEs) between the model simulated results and the observations (MADIS and ACARS) over 6 heat waves

ACARS						MADIS					
Case No.	ACM2	YSU	MYJ	Availability (%)	Sample size	ACM2	YSU	MYJ	Availability (%)	Sample size	
1	1.91	1.90	1.89	64.48	21946	2.16	2.13	2.50	99.06	10015	
7	2.21	2.21	2.17	60.42	22031	2.04	2.05	2.44	96.91	10495	
16	2.31	2.45	2.47	60.18	24861	2.35	2.27	2.43	98.79	10505	
21	2.32	2.26	2.45	46.65	13618	3.05	3.19	3.60	98.93	6290	
30	2.47	2.45	2.73	65.62	20743	3.11	3.09	3.34	98.49	6782	
40	2.76	2.78	2.79	88.81	25922	2.91	3.03	3.49	99.87	4618	
Weighted mean	2.33	2.34	2.42			2.49	2.50	2.82			

Note. Availability refers to the percentage of available data for the whole HW period, including 5 pre-HW days, all HW days and 5 post-HW days. The weighted mean RMSE is calculated based upon the weight of the sample size.

1015

1016

1017

1018

Table 3 Sensitivities of surface temperature to changes in various factors

	$\frac{\partial T_s}{\partial \varepsilon}$ (K)	$\frac{\partial T_s}{\partial \alpha}$ (K)	$\frac{\partial T_s}{\partial r_a}$ (K m/s)	$\frac{\partial T_s}{\partial r_s}$ (K m/s)	$\frac{\partial T_s}{\partial G}$ (K m ² /W)	$\frac{\partial T_s}{\partial SW_{in}}$ (K m ² /W)	$\frac{\partial T_s}{\partial LW_{in}}$ (K m ² /W)	$\frac{\partial T_s}{\partial T_a}$ (-)	$\frac{\partial T_s}{\partial q_a}$ (K)	$\frac{\partial T_s}{\partial P}$ (K/Pa)
DB	-3×10^0	-2×10	2×10^{-1}	7×10^{-3}	-2×10^{-2}	2×10^{-2}	2×10^{-2}	7×10^{-1}	1×10^2	-5×10^{-5}
DP	-5×10^0	-2×10	3×10^{-1}	5×10^{-5}	-2×10^{-2}	2×10^{-2}	2×10^{-2}	8×10^{-1}	6×10^0	-8×10^{-5}
NB	-3×10^0	0	-4×10^{-3}	6×10^{-6}	-7×10^{-2}	6×10^{-2}	7×10^{-2}	6×10^{-1}	3×10	1×10^{-5}
NP	-1×10	0	-1×10^{-3}	4×10^{-7}	-1×10^{-1}	1×10^{-1}	1×10^{-1}	2×10^{-1}	4×10^0	1×10^{-5}

Note. DB = daytime Boston; DP = daytime Phoenix; NB = nighttime Boston; NP = nighttime Phoenix. The value is the median of the results at the daily scale.

Table 4 Sensitivities of 2-m air temperature to changes in various factors

	$\frac{\partial T_2}{\partial \varepsilon}$ (K)	$\frac{\partial T_2}{\partial \alpha}$ (K)	$\frac{\partial T_2}{\partial r_a}$ (K m/s)	$\frac{\partial T_2}{\partial r'_a}$ (K m/s)	$\frac{\partial T_2}{\partial r_s}$ (K m/s)	$\frac{\partial T_2}{\partial G}$ (K m ² /W)	$\frac{\partial T_2}{\partial SW_{in}}$ (K m ² /W)	$\frac{\partial T_2}{\partial LW_{in}}$ (K m ² /W)	$\frac{\partial T_2}{\partial T_a}$ (-)	$\frac{\partial T_2}{\partial q_a}$ (K)	$\frac{\partial T_2}{\partial P}$ (K/Pa)
DB	-3×10^{-1}	-2×10^0	-1×10^{-4}	2×10^{-1}	1×10^{-3}	-2×10^{-3}	2×10^{-3}	2×10^{-3}	1×10^0	2×10	-5×10^{-6}
DP	-5×10^{-1}	-2×10^0	-2×10^{-3}	3×10^{-1}	4×10^{-6}	-2×10^{-3}	1×10^{-3}	2×10^{-3}	1×10^0	5×10^{-1}	-6×10^{-6}
NB	-2×10^0	0	2×10^{-3}	-7×10^{-3}	5×10^{-6}	-6×10^{-2}	5×10^{-2}	5×10^{-2}	6×10^{-1}	3×10	2×10^{-5}
NP	-3×10^0	0	1×10^{-3}	-6×10^{-3}	1×10^{-7}	-3×10^{-2}	3×10^{-2}	3×10^{-2}	8×10^{-1}	1×10^0	2×10^{-6}

Note. DB = daytime Boston; DP = daytime Phoenix; NB = nighttime Boston; NP = nighttime Phoenix. The value is the median of the results at the daily scale.

Table 5 Urban-rural differences in terms of various factors (Δ : Urban minus rural values)

	$\Delta \varepsilon$ (-)	$\Delta \alpha$ (-)	Δr_a (s/m)	$\Delta r'_a$ (s/m)	Δr_s (s/m)	ΔG (W/m ²)	ΔSW_{in} (W/m ²)	ΔLW_{in} (W/m ²)	ΔT_a (K)	Δq_a (kg/kg)	ΔP (Pa)
DB	2×10^{-2}	5×10^{-2}	2×10	-3×10^0	6×10^2	8×10	6×10^{-1}	4×10^0	4×10^{-1}	-4×10^{-4}	6×10^2
DP	5×10^{-2}	-1×10^{-1}	-2×10	2×10^0	-1×10^3	-5×10^0	-3×10^0	2×10	-1×10^{-1}	-4×10^{-5}	3×10^3
NB	2×10^{-2}	5×10^{-2}	-2×10	1×10	2×10^4	-4×10	0	5×10^0	5×10^{-1}	2×10^{-5}	5×10^2
NP	5×10^{-2}	-1×10^{-1}	-2×10^3	-4×10^2	4×10^5	-6×10^0	0	9×10^0	-6×10^{-1}	-2×10^{-4}	2×10^3

Note. DB = daytime Boston; DP = daytime Phoenix; NB = nighttime Boston; NP = nighttime Phoenix. The value is the median of the results at the daily scale.

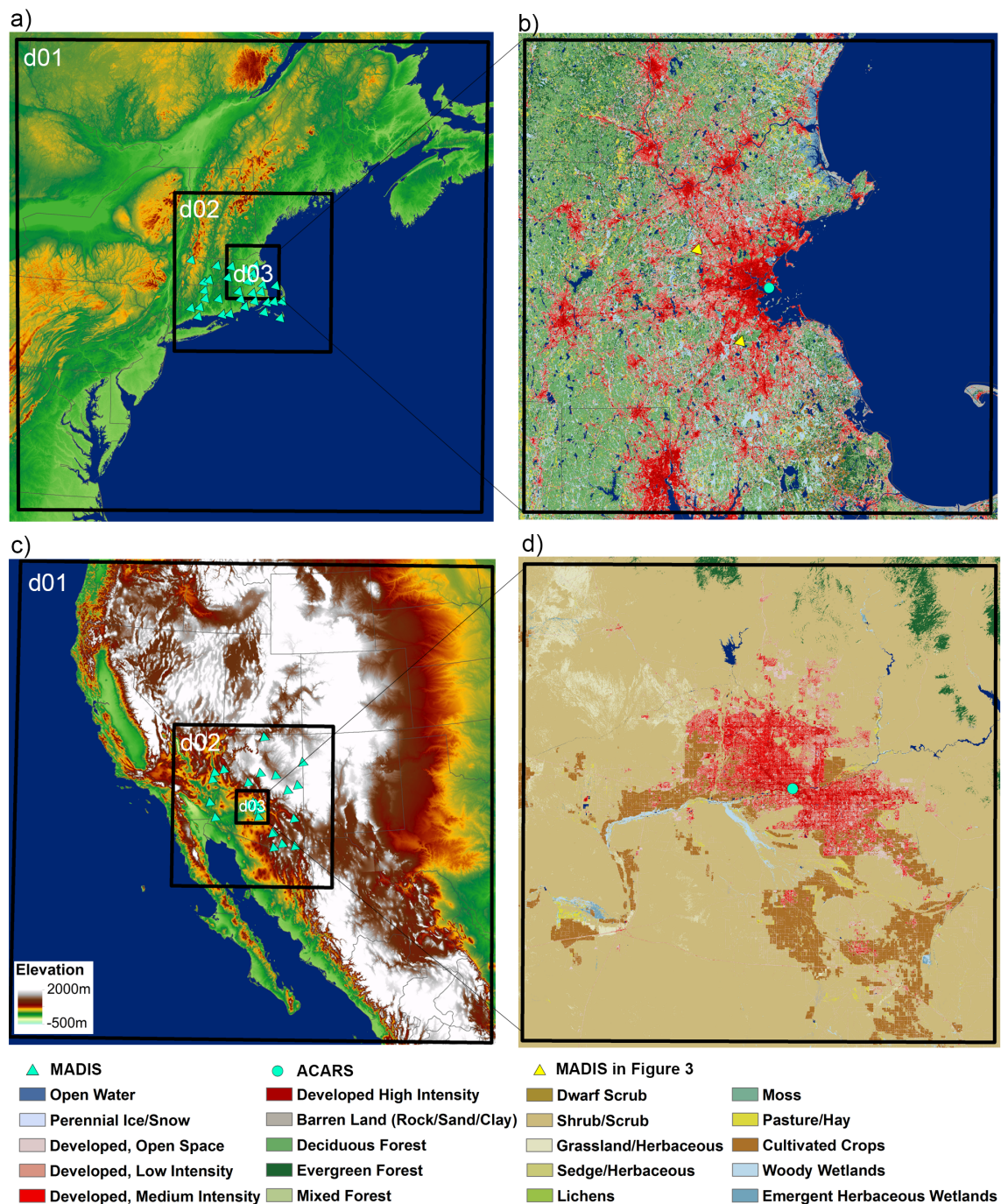


FIGURE 1 The WRF domain configuration (left), terrain height (left), and land use map (right) over Boston (a, b) and Phoenix (c, d). The yellow triangles mark the locations of the two MADIS sites used in Figure 3. The northern one is an urban site “KBED” at 42.47°N, 71.28°W, and the southern one is a rural site “KMQE” at 42.21°N, 71.11°W. The blue circles mark the locations of the Boston and Phoenix airports.

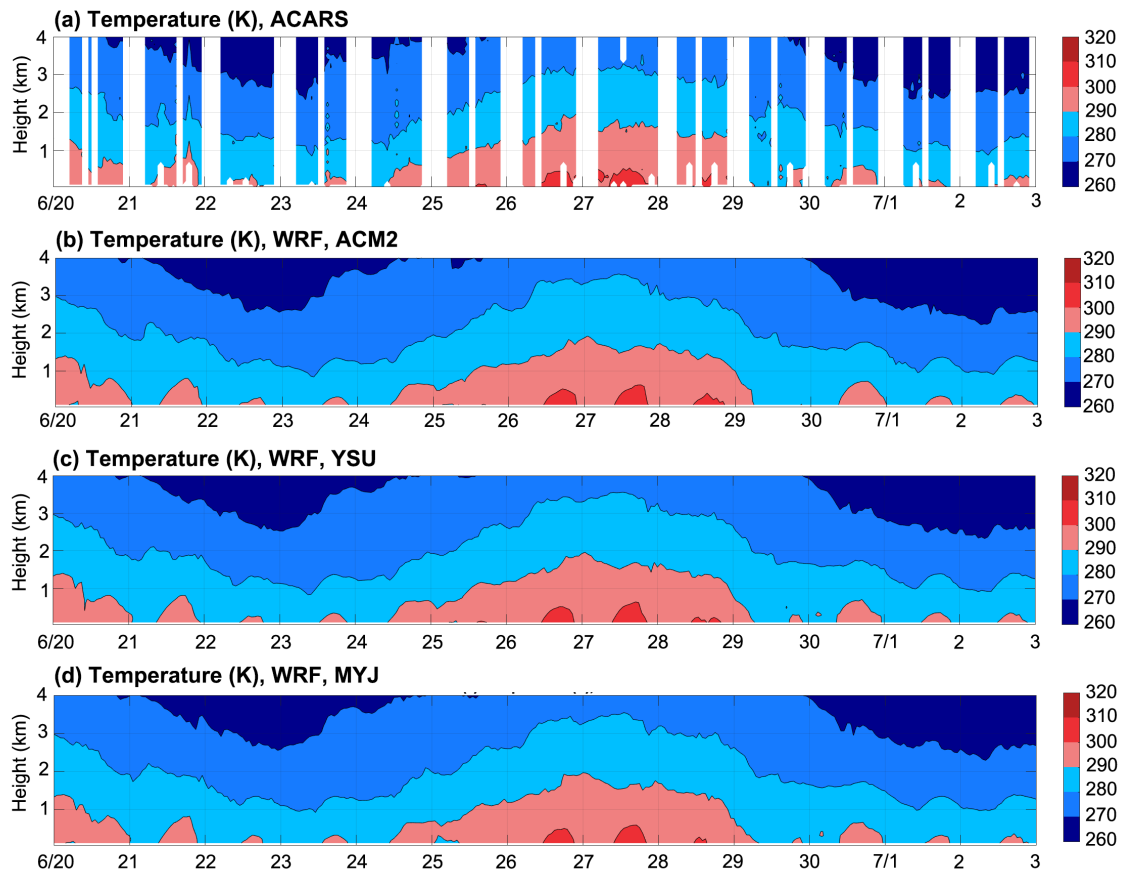


FIGURE 2 The evolution of air temperature profile in the lowest 4 km at the Boston Logan International Airport from June 20 to July 3, 2007. June 25 to 28 is the heat wave period. The panels are (a) ACARS data and WRF simulation results with the (b) ACM2, (c) YSU and (d) MYJ schemes. The height is above the ground level. The blank areas are due to the lack of ACARS data, which typically occur from late night to early morning.

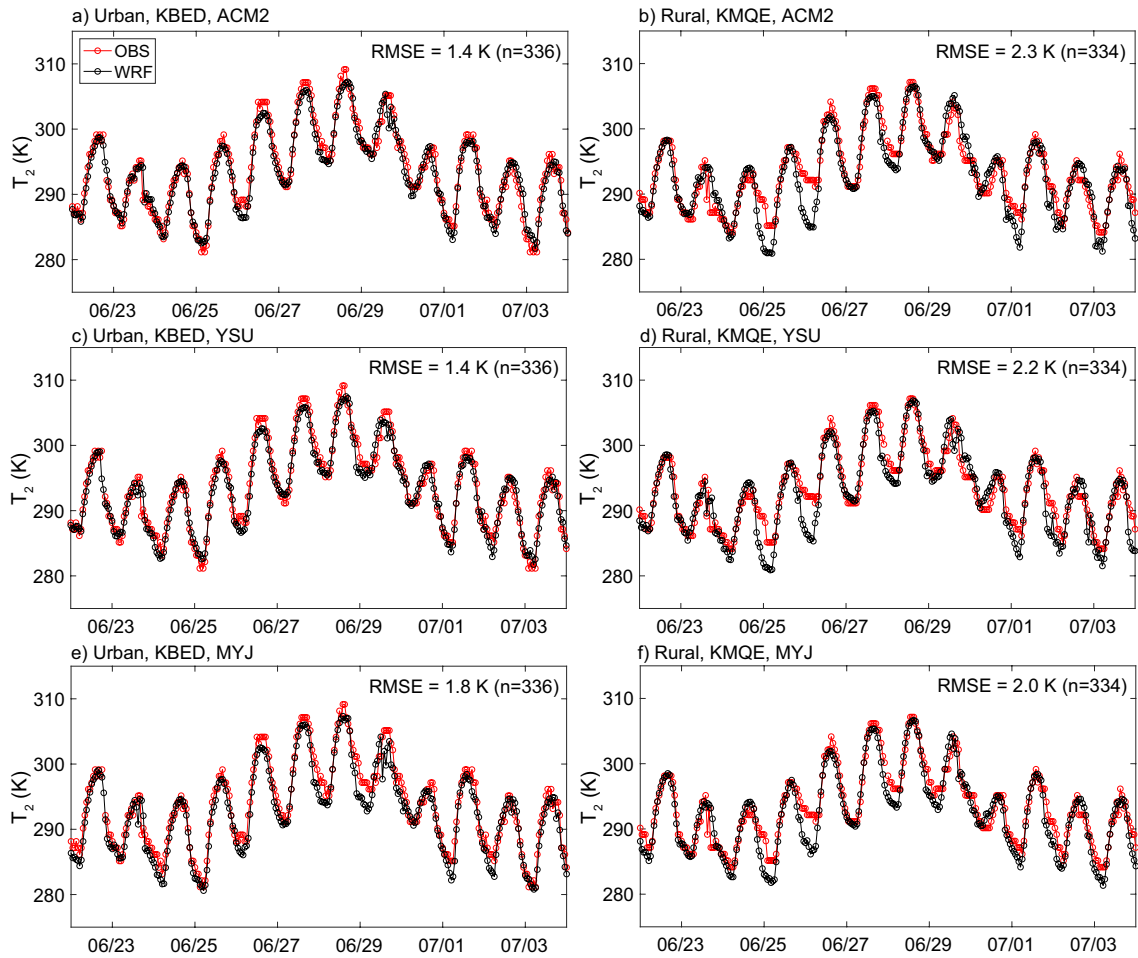


FIGURE 3 Comparisons between WRF-simulated (black circle) and MADIS weather-station measured (red circle) 2-m air temperature from June 20 to July 3, 2007 in Boston. June 25 to 28 is the heat wave period. The left panels are comparisons at an urban site (KBED) for WRF simulations with (a) ACM2, (c) YSU and (e) MYJ. The right panels are comparisons at a rural site (KMQE) for WRF simulations with (b) ACM2, (d) YSU and (f) MYJ. Please see the yellow triangles in Figure 1 for the locations of KBED and KMQE. The root-mean-square error (RMSE) and sample size (n) are denoted in the upper right corner of each panel.

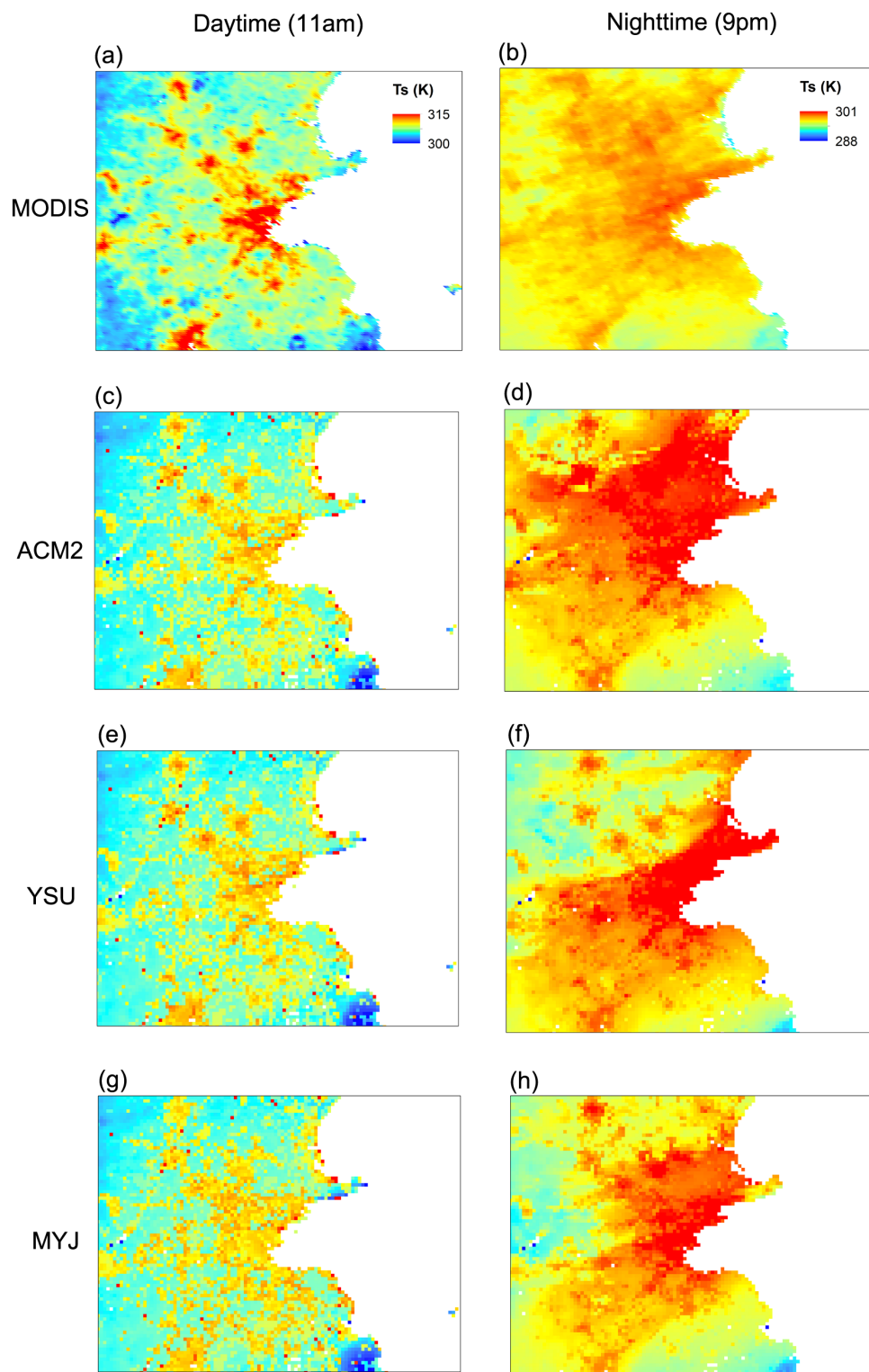


FIGURE 4 Land surface temperatures from (a, b) MODIS and WRF simulations using (c, d) ACM2, (e, f) YSU and (g, h) MYJ schemes at (a, c, e, g) 11 am and (b, d, f, h) 9 pm on June 27, 2007 in Boston.

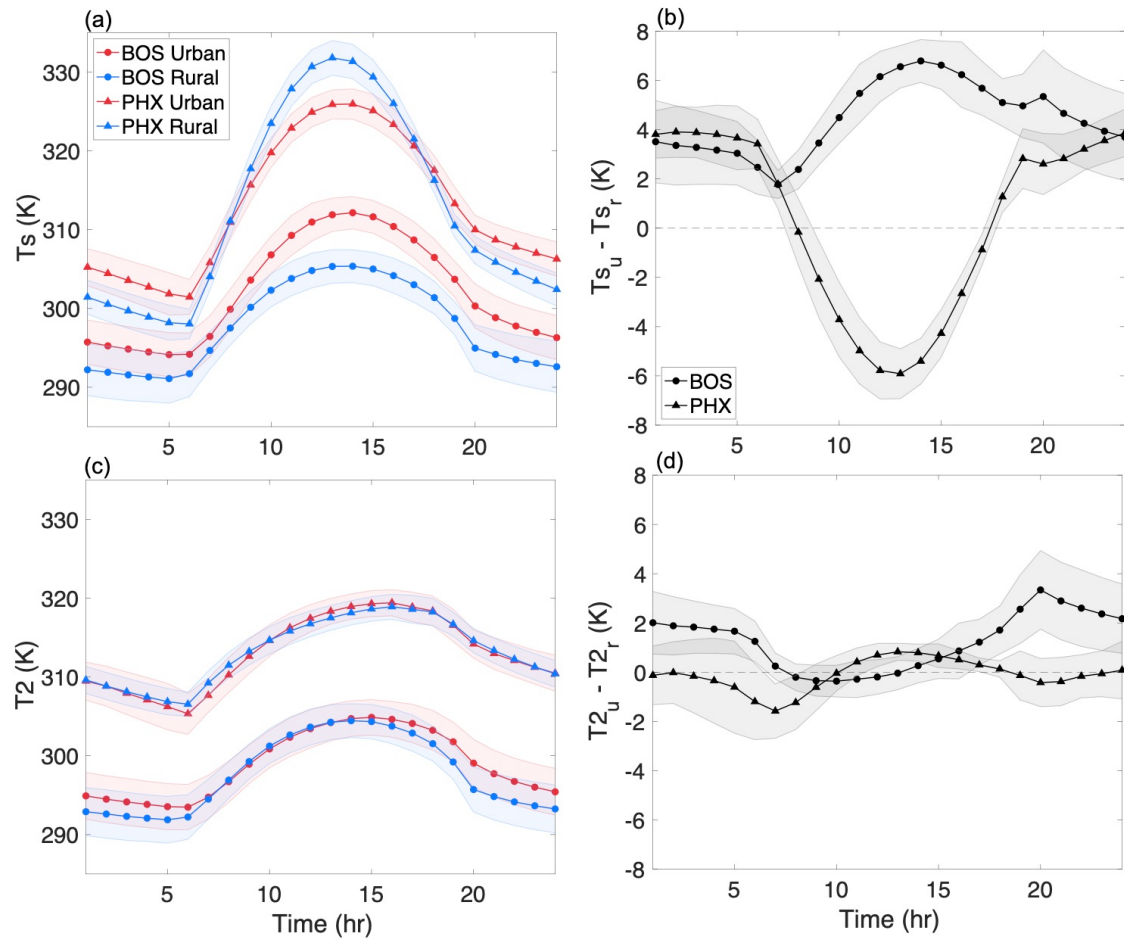


FIGURE 5 Average diurnal cycles of (a) surface temperatures (T_s), and (c) 2-m air temperature (T_2), and the urban-rural (urban minus rural temperature) differences in (b) surface temperature and (d) 2-m air temperature from WRF simulations. The shading denotes standard deviations. BOS = Boston; PHX = Phoenix. Time represents the local standard time.

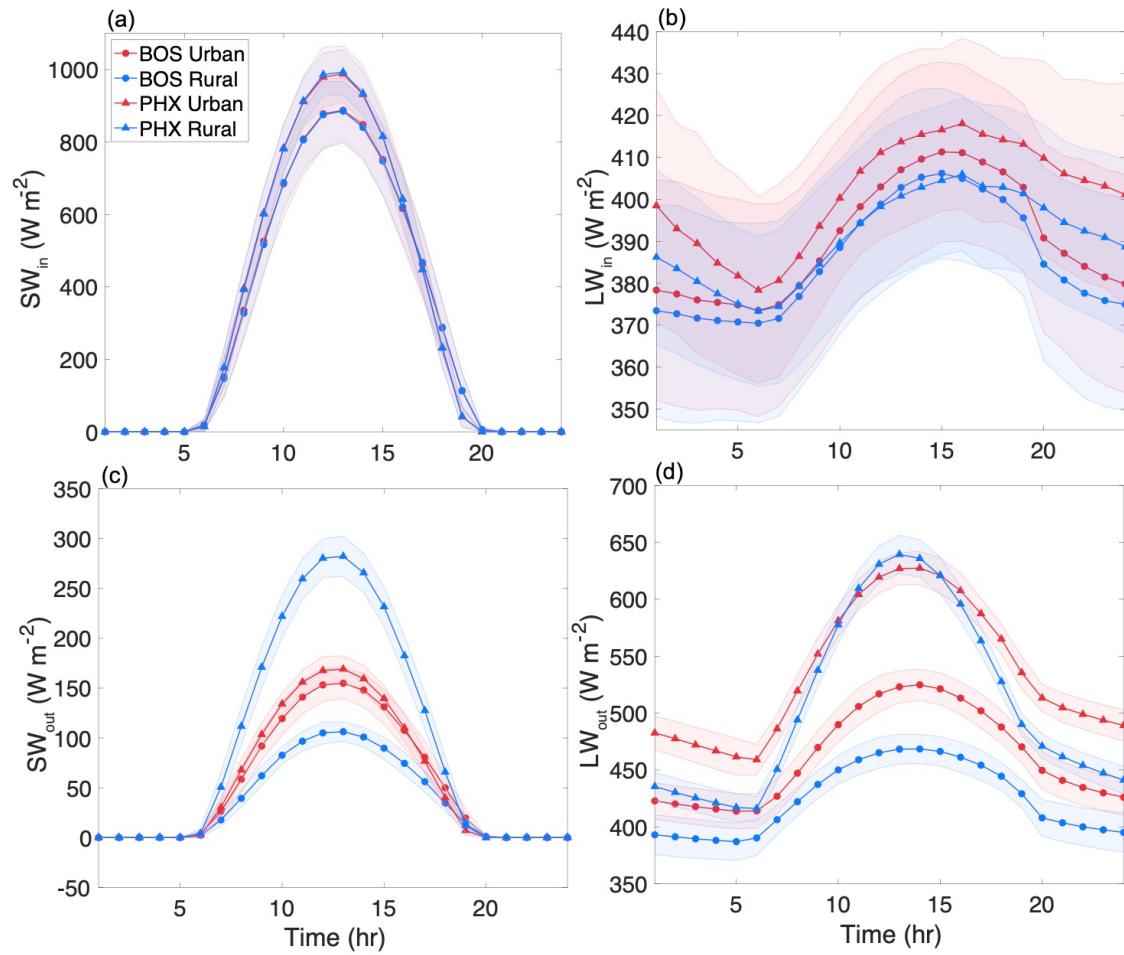


FIGURE 6 Average diurnal cycles of (a) incoming shortwave radiation (SW_{in}), (b) incoming longwave radiation (LW_{in}), (c) outgoing shortwave radiation (SW_{out}), and (d) outgoing longwave radiation (LW_{out}) from WRF simulations. The shading denotes standard deviations. BOS = Boston; PHX = Phoenix. Time represents the local standard time.

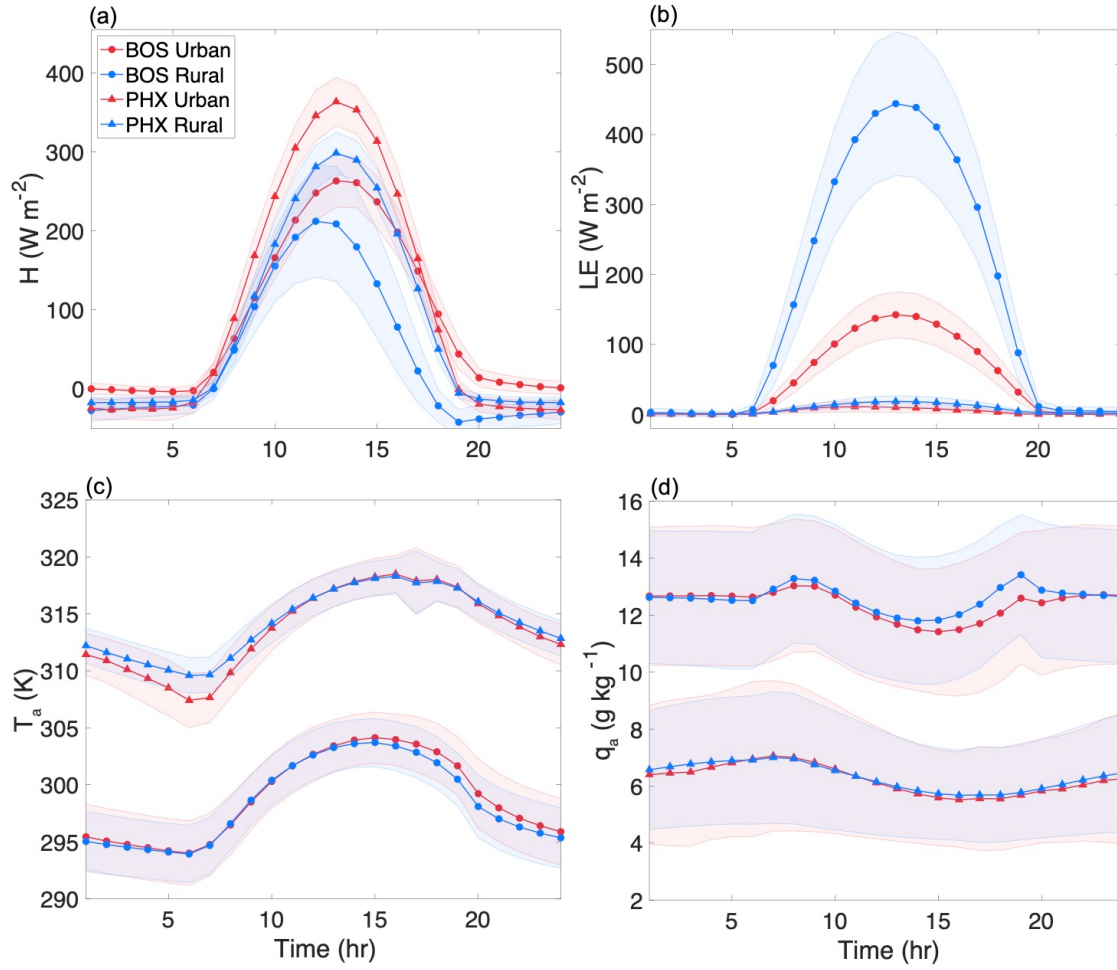


FIGURE 7 Average diurnal cycles of (a) sensible heat flux (H), (b) latent heat flux (LE), (c) potential temperature (T_a) at the lowest level of the atmospheric model, and (d) specific humidity (q_a) at the lowest level of the atmospheric model from WRF simulations. The shading denotes standard deviations. BOS = Boston; PHX = Phoenix. Time represents the local standard time.

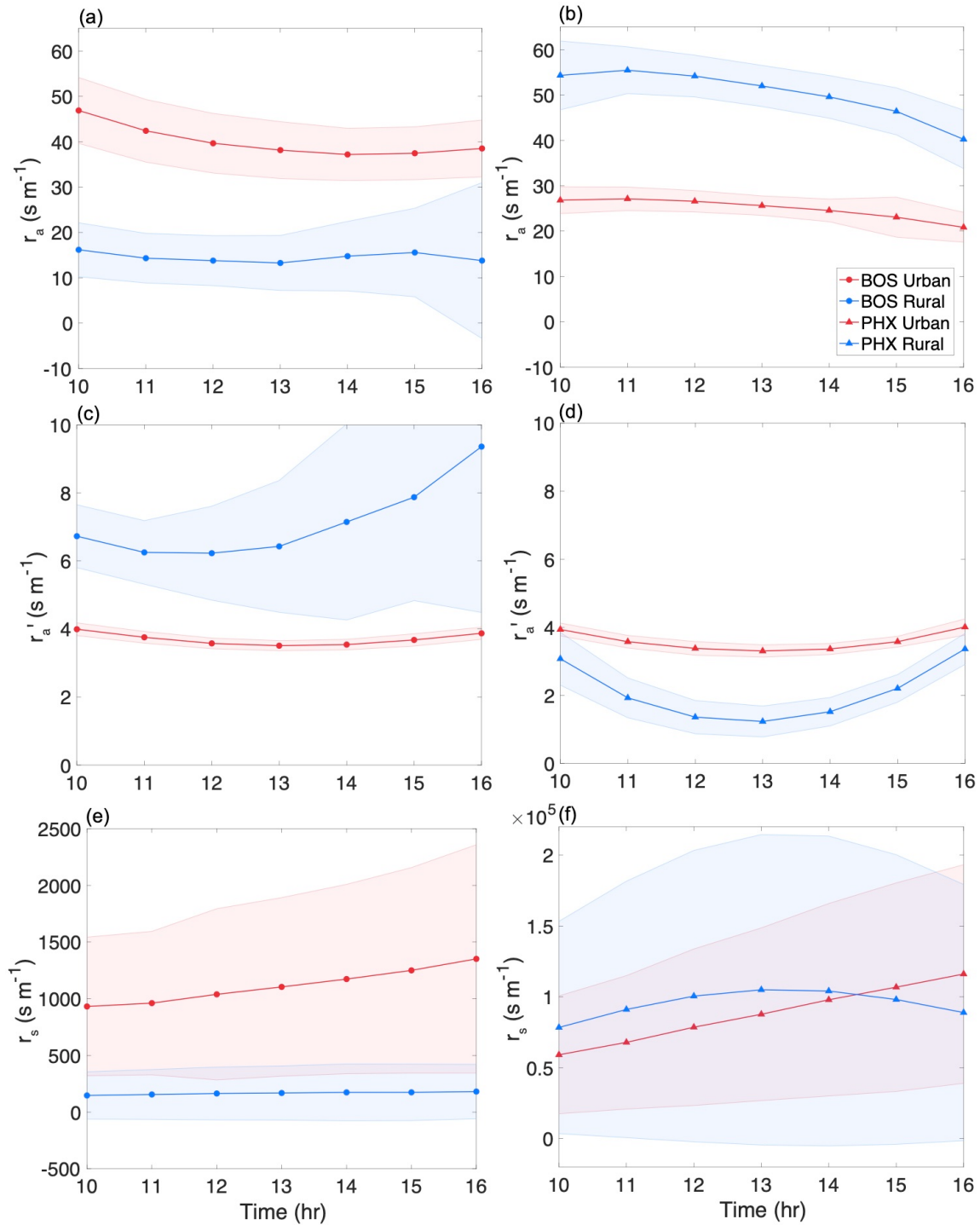


FIGURE 8 The inferred resistances from WRF simulations, including (a, b) aerodynamic resistance between the surface and the lowest atmospheric model level (r_a), (c, d) aerodynamic resistance between the 2-m level and the lowest atmospheric model level (r'_a), and (e, f) surface resistance (r_s). The shading denotes standard

1076 deviations. BOS = Boston; PHX = Phoenix. Time represents the local standard time. Only the daytime (10am
1077 to 4pm local standard time) results are shown.

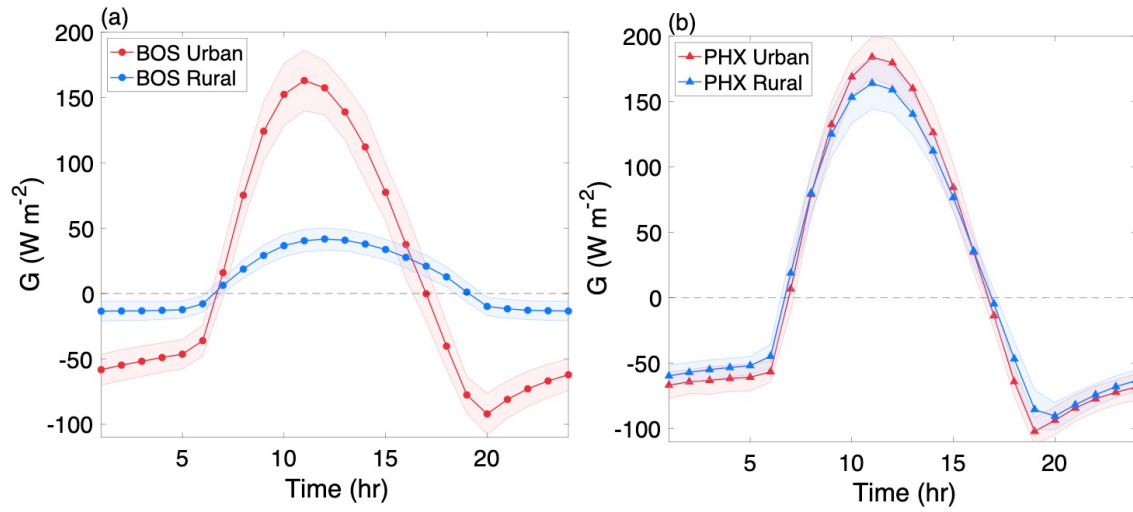


FIGURE 9 Average diurnal cycles of ground heat flux (G) from WRF simulations in (a) Boston and (b) Phoenix. The shading denotes standard deviations. BOS = Boston; PHX = Phoenix. Time represents the local standard time.

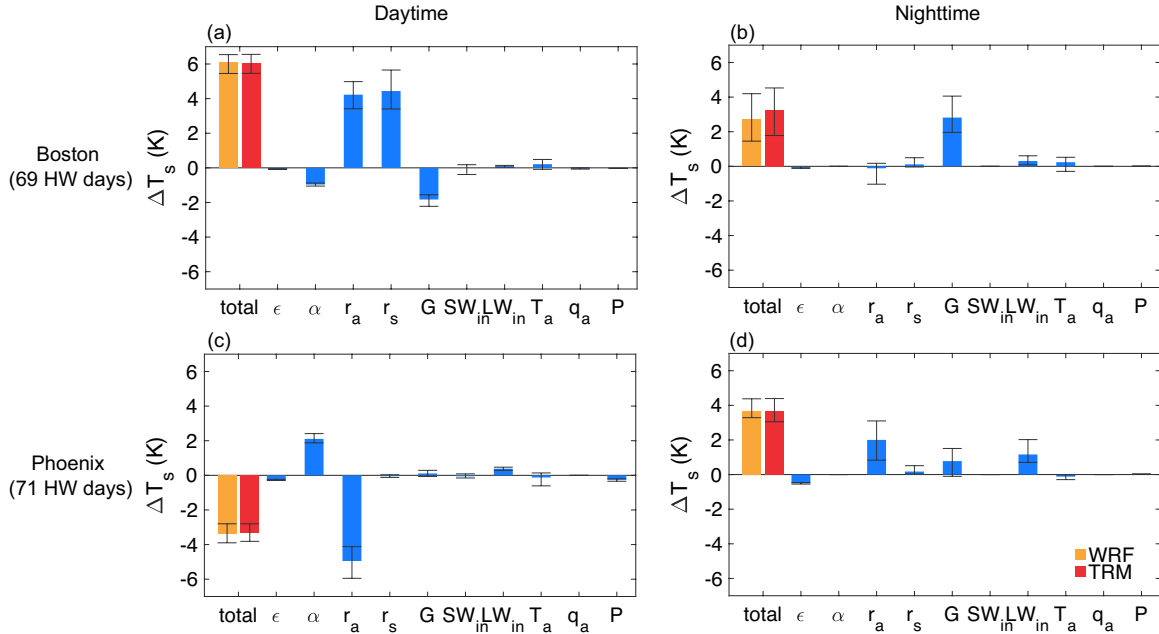


FIGURE 10 Attribution of surface urban heat islands (ΔT_s , urban minus rural surface temperature) using the Two Resistance Mechanism (TRM) model during (a, c) daytime (10am-4pm) and (b, d) nighttime (10pm-4am) in (a, b) Boston and (c, d) Phoenix. The orange and red bars represent WRF-simulated and TRM-modelled ΔT_s , respectively. The blue bars represent contributions from different factors, including emissivity (ϵ), albedo (α), aerodynamic resistance between the surface and the lowest atmospheric model level (r_a), surface resistance (r_s), ground heat flux (G), incoming shortwave radiation (SW_{in}), incoming longwave radiation (LW_{in}), potential temperature (T_a) at the lowest atmospheric model level, specific humidity (q_a) at the lowest atmospheric model level, and pressure (P). The sample size is noted in the bracket on the left. HW = heat wave. The column indicates the median of the attribution results at the daily scale. The error bars are the 80th and 20th percentiles of the results, respectively, representing the day-to-day variability of the attribution results.

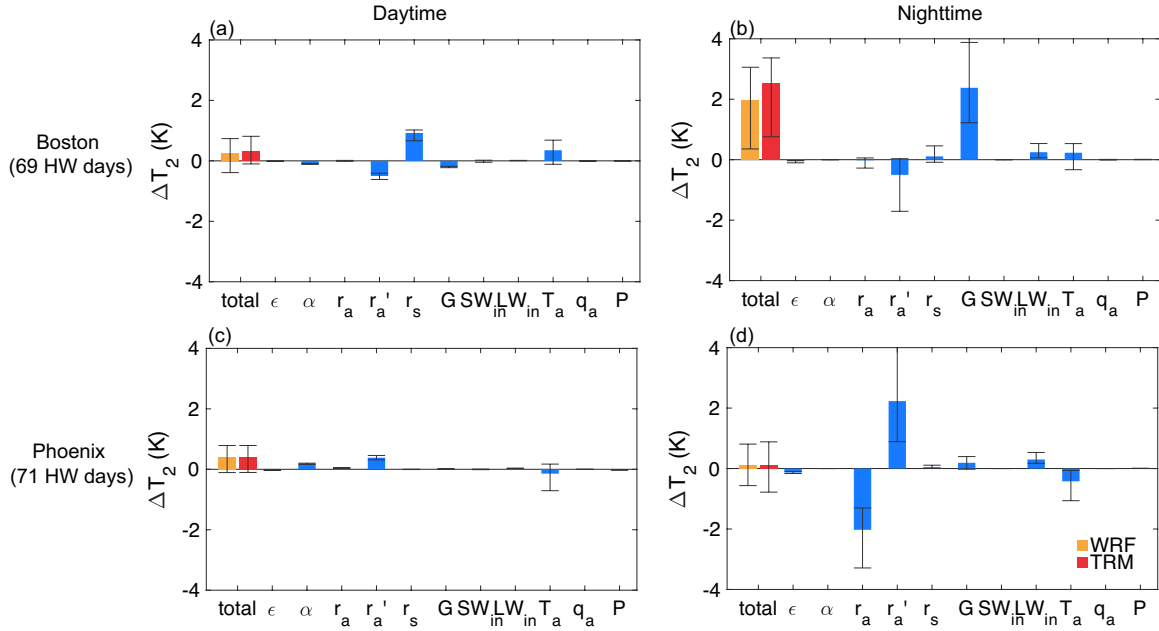


FIGURE 11 Attribution of near-surface air urban heat islands (ΔT_2 , urban minus rural 2-m air potential temperature) using the Two Resistance Mechanism (TRM) model during (a, c) daytime (10am-4pm) and (b, d) nighttime (10pm-4am) in (a, b) Boston and (c, d) Phoenix. The orange and red bars represent WRF-simulated and TRM-modelled ΔT_2 , respectively. The blue bars represent contributions from different factors, including emissivity (ϵ), albedo (α), aerodynamic resistance between the surface and the lowest atmospheric model level (r_a), aerodynamic resistance between the 2-m level and the lowest atmospheric model level (r'_a), surface resistance (r_s), ground heat flux (G), incoming shortwave radiation (SW_{in}), incoming longwave radiation (LW_{in}), potential temperature (T_a) at the lowest atmospheric model level, specific humidity (q_a) at the lowest atmospheric model level, and pressure (P). The sample size is noted in the bracket on the left. HW = heat wave. The column indicates the median of the attribution results at the daily scale. The error bars are the 80th and 20th percentiles of the results, respectively, representing the day-to-day variability of the attribution results.