Observations and Estimates of Wet-Bulb Globe Temperature in Varied Microclimates

JORDAN CLARK[a] AND CHARLES E. KONRAD[b,c]

[a] *Nicholas Institute for Energy, Environment and Sustainability, Duke University, Durham, North Carolina*
[b] *University of North Carolina at Chapel Hill, Chapel Hill, North Carolina*
[c] *NOAA/Southeast Regional Climate Center, Chapel Hill, North Carolina*

(Manuscript received 3 May 2023, in final form 13 November 2023, accepted 20 November 2023)

ABSTRACT: Wet-bulb globe temperature (WBGT) is used to assess environmental heat stress and accounts for the influences of air temperature, humidity, wind speed, and radiation on heat stress. Measurements of WBGT are highly sensitive to slight changes in environmental conditions and can vary several degrees Celsius across small distances (tens to hundreds of meters). Relative to observations with an International Organization for Standardization (ISO)-compliant WBGT meter, this work assesses the accuracy of WBGT measurements made with a popular handheld meter (the Kestrel 5400 Heat Stress Tracker) and WBGT estimates. Measurements were made during the summers of 2019–21 in a variety of suburban and urban environments in North Carolina, including three high school campuses. WBGT can be estimated from standard weather station variables, and many of these stations report cloud cover in lieu of solar radiation. Therefore, this work also evaluates the accuracy of clear-sky radiation estimates and adjustments to those estimates based on cloud cover. WBGT estimated with the method from Liljegren et al. from a weather station were on average 0.2°C warmer than Observed WBGT, while the Kestrel 5400 WBGT was 0.7°C warmer. Large variations in WBGT were observed across surfaces and shade conditions, with differences of 0.9°C (0.3°–1.4°C) between a tennis court and a neighboring grass field. The method for estimating clear-sky radiation in Ryan and Stolzenbach was most accurate and the clear-sky radiation modified by percentage cloud cover was found to be within 75 W $m^{-2}$of observations on average.

SIGNIFICANCE STATEMENT: Wet-bulb globe temperature (WBGT) is a heat stress index that accounts for the effects of air temperature, humidity, wind, and radiation on humans. However, WBGT is not routinely measured at weather stations. This work demonstrated the accuracy of estimating WBGT with methods from Liljegren et al. (2008), finding it to be more accurate than measurements from a popular handheld meter, the Kestrel 5400 Heat Stress Tracker. Variations in WBGT that result in different danger levels were found between measurements over a tennis court and a neighboring grass field, and between sun and shade conditions. Understanding the magnitude of these differences and the biases with WBGT estimates and measurements can inform the planning of outdoor activity to robustly safeguard health.

KEYWORDS: Atmosphere; Social science; Cloud cover; Irradiance; Soil temperature; Measurements

1. Introduction

The paramount importance of environmental heat stress is evident with the fact that exposure to heat is the leading cause of weather-related death in the United States (Centers for Disease Control and Prevention 2010; National Weather Service 2020). To determine how dangerous a given set of environmental conditions are to human health, heat stress indices are utilized. While the heat index is commonly used in the United States (Hawkins et al. 2017), wet-bulb globe temperature (WBGT) is a more robust heat stress index since it accounts for the effect of air temperature, humidity, solar radiation, and wind speed on human body temperature (Budd 2008; Hondula et al. 2014). Comparatively, the heat index does not account for the effect of solar radiation or wind speed on body temperature, which is a commonly cited limitation (Budd 2008; Hondula et al. 2014).

a. WBGT

WBGT was developed in the 1950s by the U.S. military in response to numerous heat-related casualties at training camps (Budd 2008; Yaglou and Minard 1957). To calculate the WBGT, three components are utilized: dry-bulb temperature, natural wet-bulb temperature, and black globe temperature (Budd 2008).

Dry-bulb temperature refers to ambient air temperature. The natural wet-bulb temperature is similar to the dry-bulb temperature, except 1) it is not measured within a radiation shield and 2) it has a wet wick wrapped around the bulb of the thermometer (Liljegren et al. 2008). Thus, the natural wet-bulb temperature is able to account for the influence of solar radiation since it is unshielded from the radiative environment and also mimics the cooling effect of sweat evaporating off of the skin. The natural wet-bulb temperature accounts for the following environmental variables influence on human body temperature: solar radiation, wind speed, humidity, and air temperature (Budd 2008; Liljegren et al. 2008). Similarly, the black globe temperature is also sensitive to these environmental variables. However, it is a better indicator of the radiative forcings active on the body, that is, the radiant temperature (both shortwave and longwave radiation) (Budd 2008; Kopec 1977; Liljegren et al. 2008). These three components are then summed with the following weightings to derive WBGT:

Corresponding author: Jordan Clark, jordan@alumni.unc.edu

DOI: 10.1175/JAMC-D-23-0078.1

$$\mathrm{WBGT} = 0.7 \times \mathrm{NWB} + 0.2 \times \mathrm{Tg} + 0.1 \times \mathrm{Ta}, \quad (1)$$

where NWB is the natural wet-bulb temperature, Tg is the black globe temperature, and Ta is the dry-bulb temperature (air temperature).

Similar to other heat stress indices, each WBGT value has an associated level of danger, often referred to as "WBGT flag level." The values defining the thresholds for each flag level vary. For example, the threshold for a black flag is set as 33.3°C by some organizations but is most commonly 32.2°C. An example of [used by the North Carolina High School Athletic Association (NCHSAA)] common thresholds are 26.7°–29.4°C (green flag), 29.5°–31.0°C (yellow flag), 31.1°–32.1°C (red flag), and 32.2°C+ (black flag) (Casa et al. 2015; NCHSAA 2016). Flag levels correspond to increasing level of danger, with green flag requiring limited activity modifications and black flag, for example, requiring that high school football practices be halted in states such as North Carolina.

Although developed in the 1950s, WBGT is utilized in numerous applications and its use continues to rise. For example, WBGT is the recommended method for measuring heat stress by the U.S. Occupational Safety and Health Administration (OSHA 2017). Additionally, an increasing number of U.S. states require WBGT measurements to determine if athletic practices can continue safely (Grundstein et al. 2015; NCHSAA 2016), including North Carolina, South Carolina, Georgia, Florida, Minnesota, Massachusetts, New Jersey, Vermont, and Arkansas.

Despite the growing utilization of WBGT, several challenges with the index remain: First, WBGT is typically measured using a device that either directly measures or approximates each individual component of the index. Devices measuring each component directly are the most accurate but are more expensive. Furthermore, some cheaper devices use the psychrometric wet-bulb temperature instead of the natural wet-bulb temperature (Cooper et al. 2017), which is problematic since the psychrometric wet-bulb temperature can be significantly cooler, particularly at low wind speeds (Kopec 1977). However, various empirical and statistical models have been developed to estimate WBGT based on variables that weather stations widely record, including air temperature, relative humidity, wind speed, and solar irradiance. Most stations record wind speed at a height of 10 m, and therefore need to be logarithmically downscaled to 2 m to the height of humans. The developed methodologies vary in accuracy and computational complexity, with the physically based model in Liljegren et al. (2008) found to be the most accurate (Grundstein and Cooper 2018; Lemke and Kjellstrom 2012; Patel et al. 2013). While estimates of WBGT can provide useful information in evaluating the heat stress of an environment, direct measurement with a WBGT meter onsite at the time of any outdoor activity remains critical since WBGT is very sensitive to slight changes in environmental conditions, similar to the human body, such as effects from wind speed and solar radiation (Budd 2008; Grundstein et al. 2022).

b. *Microclimatic influences on heat stress*

Land cover is a large source of small-scale, local variations in the environmental variables determining thermal stress (humidity, wind speed, and radiation) (Burakowski et al. 2018; Dugord et al. 2014). Local wind speed is strongly influenced by the density, location, and height of trees and buildings in the immediate vicinity (Coutts et al. 2016; EPA 2000; Wang et al. 2015). Above buildings and trees above the ground, there are typically only small differences in the wind speed across a region (e.g., 3 km or more) (Barry and Blanken 2016; Oke 1987). However, objects extending from Earth's surface, primarily trees and buildings, result in increasing friction as height above the surface decreases, slowing wind speeds down relative to the wind higher in the atmosphere. The difference in the amount of friction, referred to as surface roughness, between locations corresponds to differences in the degree of wind deceleration at the surface (Barry and Blanken 2016; Burakowski et al. 2018; Wu et al. 2018). This, in turn, can lead to marked increases in WBGT and thermal stress on hot days. The high sensitivity of WBGT to changes in wind speed can be seen in Fig. 1, which displays an example where there can be differences of up to two flag levels when wind speeds decrease from 2 to 0.75 m s^{-1} on sunny days.

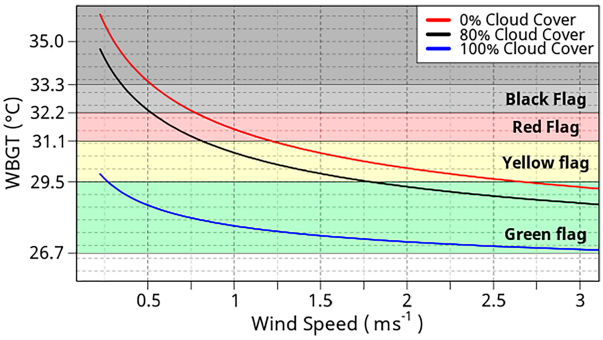

FIG. 1. Relationship between WBGT and wind speed. WBGT was estimated using the Liljegren et al. (2008) method. Relationships are depicted for three categories of cloud cover, with an air temperature of 30°C (86°F) and a dewpoint of 21.1°C (70°F).

To estimate WBGT from weather station observations, wind speed should be measured at 2 m above the ground. Since most weather stations measure wind 10 m above the surface, it must be translated into a 2-m wind speed using a vertical wind profile and logarithmic transformation. An example of this concept can be seen in Fig. 2. Utilizing wind speeds at 2 m is important since this is the general height of humans and thus the height of the wind that impacts body temperature, and since the wind speed at 10 m is consistently faster. Given the sensitivity of WBGT to wind speed, accuracy with this downscaling is critical. Common methods for downscaling wind speeds include the use of Pasquil–Gifford (PG) stability classes (Frank et al. 2020; EPA 2000) or the use of surface roughness (EPA 2000).

Variations in surface type, for example, grass versus a tennis court, produce varying degrees of thermal stress since different surfaces absorb and reflect varying amounts of radiation, which determines the amount of heat absorbed by the surface and its temperature (Barry and Blanken 2016; Burakowski et al. 2018; Oke 1987; Oke et al. 2017). This temperature, often

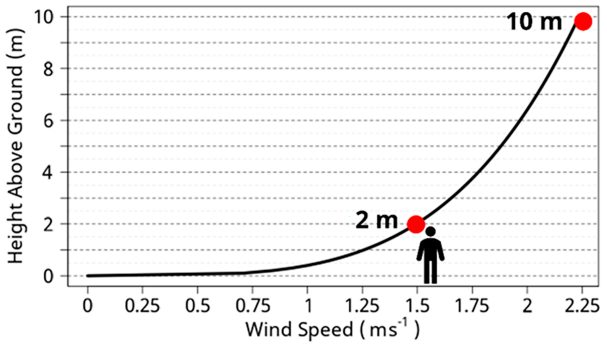

FIG. 2. Variation in wind speed as a function of height above ground. Wind speed is downscaled using PG stability classes and the "urban" exponents (see EPA 2000), with solar radiation of 500 W m^{-2} and wind speed of 2.25 m s^{-1}.

called surface temperature, is therefore proportional to the amount of longwave radiation emitted from the surface and incident on the human body (Vanos et al. 2019). Since surface temperature is largely determined by the amount of direct solar radiation incident on the surface, which is a function of cloud cover and atmospheric transmissivity, the same surface can have vastly different temperatures based on slight changes in radiation (Grundstein and Cooper 2020). The main factors influencing incident radiation, especially over small spatial and temporal scales, are land cover (shade from vegetation) and cloud cover (Burakowski et al. 2018; Coutts et al. 2016; EPA 2000). Like wind speed, changes in radiation contribute to large and sudden swings in WBGT over the course of only minutes (Grundstein and Cooper 2020; Lee et al. 2023). Thus, to reliably capture the overall heat stress, WBGT should be averaged over a set time interval, with literature recommending a variety of intervals, such as a 5-min (Thorsson et al. 2007), 30-min (Kopec 1977), or 60-min average (OSHA 2017).

The third environmental variable contributing to microclimatic variations in WBGT is humidity. Thermally, changes in humidity determine the sweat evaporation rate off human skin, and thus the rate at which the body can cool itself through evaporative cooling (Foster et al. 2022). Like wind speed and radiation, land cover is critical since nearby bodies of water, specifically the evaporation of that water, significantly contribute to local humidity levels (Barry and Blanken 2016; Oke 1987). Further, soil characteristics that vary across different soil types, such as retention capacity and drainage, result in varying amounts of water evaporating from the soil and thus local differences in humidity (Barry and Blanken 2016). Differences in soil moisture are yet another factor influencing surface temperature. Wet soil has a greater latent and smaller sensible heat flux than dry soil, meaning dry soil will be warmer than wet soil under the same amount of incident radiation (Barry and Blanken 2016; Oke 1987; Oke et al. 2017).

Last, while these environmental variables are important to consider separately as determiners of microclimatic variability, these variables act in concert as part of a feedback loop to define uniquely hot or cool microclimates, with the following example resulting in progressively less heat stress. Generally, the amount of radiation a surface receives influences surface temperature and local moisture flux from bodies of water and the soil. As the surface is continually heated, the air at the surface warms and rises. The degree of surface warming influences the amount of air being warmed and the speed at which it rises, generating wind. As this rising air is replaced by the cooler and drier air above via convective mixing, local humidity levels are changed, and together with the wind generated from the radiative heating of the surface, the overall heat stress of a particular location and microclimate is determined.

Existing research has highlighted the variations in WBGT as a function of surface type (Grundstein and Cooper 2020; Kosaka et al. 2018; Pryor et al. 2017), the wide range of variations in temperature across surfaces (Kopec 1977), and the range of impacts of the overall radiative environment to human body temperature (Hardin and Vanos 2018). Additional research has also espoused the challenges and inaccuracies associated with using WBGT and other heat stress indices estimated from a weather station far from the site of interest, emphasizing the importance of real-time, on-site measurements (Grundstein et al. 2022; Pryor et al. 2017).

The research presented here expands upon this existing literature in several ways. Relative to prior work, it examines a much larger sampling of days across microclimates in an array of athletic and suburban environments, particularly days with high thermal stress (WBGT $>$ 31.1°–32.2°C). Further, a full-size WBGT meter that conforms to standards outlined by the International Organization for Standardization is utilized to assess the accuracy of WBGT estimates and WBGT measurements made by other devices. Given the importance of accurately measuring and estimating WBGT since it is used to safeguard health, the bias of the Liljegren et al. (2008) method for estimating WBGT and one of the more commonly used and robust WBGT meters, the Kestrel 5400, are compared with ground-truth WBGT observations, with the use of Kestrel 5400 devices similar to those used by Carter et al. (2020). Additionally, WBGT measurements and surface temperature readings are made and compared over various surfaces with distinct microclimates (e.g., grass, tennis courts, forests, sun, and shade).

This research further contributes to existing literature by evaluating the accuracy of different methods for estimating clear-sky radiation and the subsequent modification of that clear-sky radiation by percentage cloud cover to attain an estimate of incident radiation at the surface. This is important since WBGT is not commonly measured at weather stations and most weather stations either have no solar radiation measurement or only provide observed cloud-cover amounts. Assessing the accuracy of solar radiation estimates and accounting for cloud-cover percentage has important implications for creating forecasts of WBGT since they typically rely on an instantaneous forecast cloud-cover value.

c. *Research questions*

The following research questions were addressed:

1) What is the overall accuracy of the Liljegren et al. (2008) method for estimating WBGT? For this method, does assuming surface temperature is equivalent to the air temperature introduce significant bias?
2) What is the accuracy of the Kestrel 5400 WBGT?
3) How does surface temperature across different surface types relate to variations in WBGT (e.g., shade vs sun; tennis court vs grass)?
4) How does the accuracy of different commonly used methods for estimating clear-sky radiation compare to one another? What is the accuracy of modifying clear-sky radiation by percentage cloud cover?

2. Data

a. Observed WBGT measurements

WBGT data were recorded at several locations, detailed below, alongside a Davis Instruments Vantage Pro 2 weather station with three different instruments: two Kestrel 5400 Heat Stress Tracker devices and a custom research unit manufactured by Kestrel to meet the specifications for a WBGT meter as outlined by the International Organization for Standardization, first used in Cooper et al. (2017) (Fig. 3). The ISO guidelines (Parsons 2006) require that the black globe have a diameter of 0.15 m and the temperature sensor measure 20°–120°C (with accuracies of ±0.5°C for 20°–50°C and ±1°C for temperatures greater than 50°C). The natural wet-bulb temperature sensor should be a cylindrical shape and measure the range 5°–40°C with an accuracy of ±0.5°C. The sensor should have a diameter of 6 ± 1 mm with a length of 30 ± 5 mm. The wick positioned over the temperature sensor must be white and made of cotton or other water-absorbent material. Last, the dry-bulb temperature sensor must be within a radiation shield and measure the range 10°–60°C with a ±1°C accuracy (Parsons 2006). As with the other instruments, the WBGT meter was designed to record data 1.5–2 m above ground level. Last, dependent upon the location and the frequency of access to the location, the WBGT meter recorded data at multiple intervals (2-, 5-, 10-, 30-, 60-, or 120-s intervals). The WBGT measured by this device will hereinafter be referred to as "Observed WBGT."

The Kestrel 5400 devices directly measure WBGT. Although the diameter of the black globe thermometer (2.5 cm) is significantly smaller than the standard black globe thermometer (0.15 m) (Fig. 3), the manufacturer adjusts the measurements of this smaller globe to account for this size difference. The devices do not directly record natural wet-bulb temperature but instead offer approximations of this component. Two Kestrel 5400 devices were purchased directly from the manufacturer in the summer of 2019, and the humidity sensor was recalibrated prior to the start of measurements in 2021. These two devices were used for concurrent measurements for comparisons of WBGT in different microclimates.

For all measurements used in this study, the Kestrel 5400s were mounted atop tripods designed for the devices with the Kestrel Vane Mount holding the device, which allows the device to rotate freely according to wind direction. In addition to the WBGT variables, the Kestrel 5400 devices also record air temperature, dewpoint temperature, relative humidity, barometric pressure, and wind speed. The Kestrel 5400s were set to record data at 1.5 m above the ground at various intervals (2-, 5-, 10-, and 30-s intervals), depending upon the location and ability to access that location regularly, detailed below.

FIG. 3. Field work instruments (WBGT meters and weather station). The WBGT meter is on the right, with a black globe thermometer, a natural wet-bulb temperature probe situated in a water reservoir, and a dry-bulb sensor in a radiation shield. A weather station (on the left) and a Kestrel 5400 (in the center) were also collocated with the WBGT meter.

b. Weather station measurements

Alongside the Kestrel 5400 devices, a weather station was utilized to collect standard meteorological variables, specifically a Davis Instruments Vantage Pro 2 Plus (Fig. 3). The station was collocated with the WBGT meter to provide data quality checks. This weather station made measurements of the following variables with a temporal resolution of 10 s: air temperature, dewpoint temperature, barometric pressure, relative humidity, solar radiation, and wind speed. The weather station instruments were mounted on a tripod to situate the station 1.5 m above ground level. Routine calibration for relative humidity was conducted by placing the temperature and humidity sensor (Sensirion SHT31) in humidity chambers, calibrating to two points, with a sodium chloride slurry and magnesium chloride slurry for calibrating to 75% and 33%, respectively.

Additionally, a temperature probe was anchored adjacent to the ground to record surface temperature measurements, also at a temporal resolution of 10 s. However, surface temperature measurements were not recorded at the high school campuses. This temperature probe was manufactured by

Davis Instruments (SKU 6475). Last, attached at the top of the weather station tripod was an IP camera positioned to take photographs of the sky conditions at an interval of 10 s (Fig. 3). The exact camera model ultimately utilized is not pictured in Fig. 3. The camera was manufactured by Amcrest (model IP8M-2496EB-V2) with a resolution of 8 megapixels (3840 × 2160) and a wide viewing angle of 125°.

c. Data collection locations

Measurements were taken at five locations throughout the summers of 2019–21 in North Carolina: 1) three high schools in the Triangle Region: Cedar Ridge High School (CRHS) in Hillsborough (36.055 94°, −79.129 684°), Green Level High School (GLHS) in Cary (35.771 784°, −78.898 320°), and Wake Forest High School (WFHS) in Wake Forest (35.985 798°, −78.516 503°), including measurements at different practice and playing fields across these campuses and tennis courts; 2) the Horace Williams Airport in Chapel Hill (35.934 270°, −79.063 597°); and 3) within suburban environments in Chapel Hill, Durham, and Shelby. Cloud-cover images were only obtained for the summer of 2021 in Durham (35.905 449°, −78.929 363°) and Shelby (35.255 419°, −81.518 499°). All instruments were checked daily to ensure accurate and consistent measurement. This was critical for the natural wet-bulb temperature to ensure that there was water in the reservoir to moisten the wick.

3. Methods

a. Observed WBGT

The Observed WBGT data (i.e., the individual measures of natural wet-bulb temperature, black globe temperature, and dry-bulb temperature) were consistently cross referenced with Station WBGT from the weather station collocated with the WBGT meter. This was done to ensure no errant data points. To ensure that the instruments were properly situated with respect to sun and shade, measurement times were cross referenced with regularly captured images of the observation area and with pyranometer measurements of solar radiation from the collocated weather station.

When the WBGT meter was put in place (both Observed WBGT and Kestrel WBGT), the first 15 min of data were discarded, since the instruments require time to equilibrate with the environment (Kestrel 2021). Routine checks were conducted to ensure the water reservoir for the natural wet-bulb temperature remained sufficiently full to keep the wick consistently moist. If the water level was insufficient to keep the wick moist, the natural wet-bulb temperature data collected between then and the last time it was checked (the prior evening) were discarded. When additional water was required, distilled water kept in the shade at ambient air temperature levels was used. After refilling the water reservoir, the observations for the next 15 min were discarded to ensure the wick had sufficient time to dampen and the temperature probe equilibrate.

Because of the rapid variability of WBGT over short time intervals and for better comparison between the three devices (Observed WBGT, Kestrel 5400, and weather station), minute-level averages were calculated from the collected data by each device, from which 15-min moving averages were derived. The 15-min average was chosen as a compromise between other recommended averaging periods for WBGT (5- and 30-min averages) and following Liljegren et al. (2008).

b. Station WBGT

WBGT was estimated from the weather station observations using the method developed in Liljegren et al. (2008), which has consistently been found to be highly accurate (Lemke and Kjellstrom 2012; Patel et al. 2013). The required methods were provided via the R package "wbgt" (Lieblich and Spector 2017).

Using the Liljegren et al. (2008) method, the equation for calculating the natural wet-bulb temperature is given as

$$T_w = T_a - \frac{\Delta H}{c_p}\frac{M_{\mathrm{H2O}}}{M_{\mathrm{Air}}}\left(\frac{\mathrm{Pr}}{\mathrm{Sc}}\right)^{\alpha}\left(\frac{e_w - e_a}{P - e_w}\right) + \frac{\Delta F_{\mathrm{net}}}{Ah}, \qquad (2)$$

where T_a is the ambient air temperature, ΔH is the heat of vaporization, C_p is the specific heat at constant pressure, M_{H2O} is the molecular weight of water vapor, M_{Air} is the molecular weight of air, Pr is the Prandtl number, Sc is the Schmidt number, α is a constant (0.56), e_w is the saturation vapor pressure of the wick, e_a is the saturation vapor pressure of the air, P is the barometric pressure, ΔF_{net} is the net radiant heat flux to the wick from the environment, A is the surface area of the wick, and h is the convective heat transfer coefficient [see Liljegren et al. (2008) for details] (Grundstein and Cooper 2018; Liljegren et al. 2008). The equation for black globe temperature is given as

$$T_g^4 = \frac{1}{2}(1 + \varepsilon_a)T_a^4 - \frac{h}{\varepsilon_g \sigma}(T_g - T_a) + \frac{S}{2\varepsilon_g \sigma}(1 - \alpha_g) \times \left\{1 + \left[\frac{1}{2\cos(\theta)} - 1\right] f_{\mathrm{dir}} + \alpha_{\mathrm{sfc}}\right\}, \qquad (3)$$

where ε_a is the emissivity of air, T_a is the ambient air temperature, h is the convective heat transfer coefficient, ε_g is the emissivity of the globe, S is the total horizontal solar irradiance, α_g is albedo of the ground, θ is the solar zenith angle, f_{dir} is the fraction of total horizontal irradiance S that is direct beam radiation, and α_{sfc} is the albedo of the surface (Grundstein and Cooper 2018; Liljegren et al. 2008). The following constants were used: $\varepsilon_g = 0.95$, $\alpha_g = 0.05$, and $\alpha_{\mathrm{sfc}} = 0.45$ (Liljegren et al. 2008).

The estimated WBGT from the weather station will be hereinafter referred to as "Station WBGT."

c. Clear-sky radiation estimates and image analysis

Estimates of clear-sky radiation from three methods were compared with measured solar radiation: 1) Kasten and Czeplak (1980) (Stein et al. 2012), 2) Ryan–Stolzenbach model (Annear and Wells 2007; Ryan and Stolzenbach 1972), and 3) Bras model (Bras 1990; Covert and Hellström 2014). These three models will hereinafter be referred to as 1) Kasten, 2) Ryan, and 3) Bras.

This analysis was conducted to quantify the accuracy of modifying clear-sky radiation by percentage cloud cover.

The Kasten and Czeplak (1980) method was utilized as it is the current method used by the NWS for their WBGT forecast (Boyer 2022). The Ryan model has input for elevation, and other parameters for all models include the location (latitude and longitude) and the date/time. Additionally, the Ryan and Bras models for estimating clear-sky radiation have customizable parameters to account for the impact of atmospheric transmissivity on the incident radiation at the surface, 0.0–1.0 for the Ryan model and an integer between 1 and 5 for the Bras model, with 5 being the most turbid. The default parameters were utilized, specifically setting the atmospheric transmissivity coefficient (TC) to 0.8 in the Ryan model and the corresponding variable in the Bras model to 3. In addition to estimates using the default parameters, these parameters were changed to determine if more accurate estimates could be made based on the observed solar radiation by the weather station. First, however, cloud-free periods and the percentage cloud cover were identified.

To estimate the percentage cloud cover for the weather station and WBGT observations, each image corresponding to the date and times of those observations was segmented and classified. The images were segmented using the simple linear iterative clustering (SLIC) method combined with the affinity propagation (AP) clustering algorithm (Zhou 2015). Superpixels using the SLIC method were constructed and then clustered based on similarity with the AP clustering algorithm (Zhu et al. 2017). This was completed in R using the packages "SuperpixelImageSegmentation" (Mouselimis 2022b; Zhou 2015) and "OpenImageR" (Achanta et al. 2010, 2012; Buchner 2013; Haghighat et al. 2015; Haghighat 2015; Sight Machine 2012; Mouselimis 2022a).

After the image segmentation and clustering, the color representing each cluster, which was determined by averaging the RGB values across all pixels in the cluster, was converted to hue, saturation, value (HSV) values in the R software. The HSV values corresponding to clouds and clear sky were determined manually from a sample of images. Once these characteristics were determined, all images were processed. To ensure adequate performance, random selections of five images per day were selected and manually reviewed to determine if the configuration was correctly segmenting and classifying clusters as clouds or clear sky. Some clusters identified in the image had very few total pixels. When these clusters were 1) unable to be classified based on the predetermined characteristics of clouds and clear sky and 2) exceeded 10% of the total pixels in the image, the cloud-cover percentage derived from that image was not included in the analysis.

Although images were captured every 10 s, only images every 30 s were utilized from 20 June 2021 to 11 September 2021, given the volume of data and the ultimate smoothing of data at this fine temporal scale, detailed below. Last, only images between the hours of 0930 and 1600 local time were used, due to issues with classifying the images when the sun angle was lower and since this was the period in which all instruments were consistently and fully exposed to the sun (or whatever the current sky conditions were). The percentage cloud cover was then calculated by dividing the total number of pixels within each cluster denoted as clouds by the total number of pictures in the entire image. Out of all images of sky conditions taken, 83 859 images were able to be used; 3036 of these images were discarded due to the percentage of pixels unable to be classified being greater than 10%.

For each of the three models, to assess the accuracy of modifying clear-sky radiation by percentage cloud cover to obtain an estimate of the solar radiation incident at the surface, the clear-sky radiation value corresponding to the time stamp of each image was modified by the calculated percentage cloud cover using

$$\text{Srad} = R_0(1 - 0.75n^{3.4}), \tag{4}$$

where n is the cloud-cover fraction (0.0–1.0) and R_0 is the clear-sky direct radiation (W m^{-2}) (detailed above) (Kasten and Czeplak 1980).

d. Analysis

The accuracy of the Liljegren WBGT estimation and the WBGT measured by the Kestrel 5400 were assessed by comparing them with the Observed WBGT. The overall accuracy and the accuracy at 0.5°–1°C WBGT increments were calculated. In addition to the general and relative biases, the accuracies of the Liljegren method and Kestrel 5400 (with values rounded to the nearest 0.5°C) were assessed using the three following accuracy verification metrics:

1) Hit rate (%): percentage of correct WBGT measurements or estimates (Jolliffe and Stephenson 2012).
2) False alarm ratio: total number of false alarms for a given WBGT divided by the total number of observations at that WBGT (Jolliffe and Stephenson 2012).
3) Bias score: measure indicating the direction of bias (positive/negative) in addition to the magnitude (ratio of the frequency of measuring/estimating a WBGT to the frequency of observations at that WBGT). Values greater than 1 correspond to positive (warm) biases. Values less than 1 correspond to negative (cool) biases (Jolliffe and Stephenson 2012; NCAR 2015).

Comparisons of the WBGT in a variety of locations and shade conditions were assessed, including 1) full sun versus full shade, 2) full sun versus shaded forest, 3) shaded field versus shaded forest, and 4) tennis court versus grassy field. These comparisons only utilize the Kestrel 5400 WBGT since there were two devices, and concurrent measurements allowed for the most robust comparisons. To assess the statistical significance of the differences in WBGT as measured by the Kestrel 5400 and estimated with the Liljegren method, and differences across the different surface types, Wilcoxon rank sum tests with continuity correction were utilized to compare means. Statistical significance was evaluated at the 95% confidence level.

The surface temperature probe was collocated with the other field work instruments to determine the impact of surface temperature on the accuracy of the estimation and Kestrel 5400 values. Since the Liljegren method assumes that the surface temperature is equivalent to the air

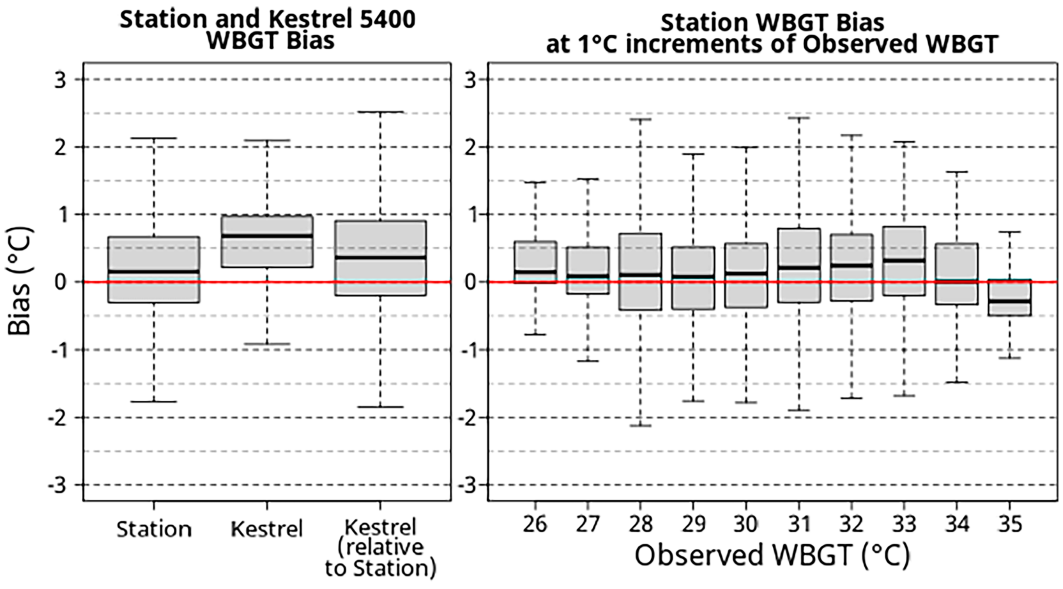

FIG. 4. WBGT biases by device and by 1°C increments of WBGT: (left) differences between the Station WBGT or Kestrel 5400 WBGT and Observed WBGT, as well as differences between the Station and Kestrel 5400 WBGT, and (right) the difference of the Station WBGT bias as a function of the Observed WBGT, incrementing at intervals of 1°C.

temperature for estimating the black globe temperature (Liljegren et al. 2008), comparisons of the bias for the black globe temperature and WBGT overall were assessed as a function of the difference in air and surface temperature.

Last, evaluating the accuracy of the three methods for estimating clear-sky radiation used here first required identifying cloud-free periods. The derived cloud-cover percentage from each image and observed solar radiation readings were averaged out to a 5-min interval to suppress differences arising from any slight mismatches in time steps of the images and observations. The percentage cloud cover at a 5-min interval was then determined to be cloud-free if it had less than 5% cloud cover. For these periods, the estimates of clear-sky radiation from each of the three methods were calculated and then directly compared with the corresponding 5-min average solar radiation observation. The accuracy assessment of modifying clear-sky radiation by percentage cloud cover [Eq. (4)] followed those same steps, except all images were used instead of only the cloud-free images (cloud-cover values 0%–100%).

4. Results

a. Biases in the station and Kestrel 5400 WBGT

Comparisons with the Observed WBGT revealed that both the Station and Kestrel 5400 WBGT display a positive (warm) median bias (Fig. 4). The Kestrel 5400 WBGT median bias was 0.4°C [95% confidence interval (CI): 0.3°–0.4°C] higher than the Station WBGT and 0.7°C (0.6°–0.7°C) higher than Observed WBGT (Fig. 4) (Table 1).

The assessment revealed a distinct pattern in the Kestrel 5400 WBGT bias (Fig. 5). As Observed WBGT increased, the Kestrel 5400 had an increasingly positive (warm) bias, with mean biases of +0.7°C (0.6°–0.7°C) at 26.7°C and +0.9°C (0.9°–0.9°C) at 32.2°C (Fig. 5) (Table 2). At WBGTs greater than 31°C, the median and the lower 75th percentile of the Kestrel 5400 WBGT bias were positive (Fig. 5). Contrasting this, the Station WBGT did not display a distinct pattern of changing bias as WBGT increased, with its median bias remaining between 0.1° and 0.4°C for all WBGTs between 26° and 35°C (Fig. 4). At extreme WBGTs of 35°C, the Station

TABLE 1. WBGT thresholds: Wilcoxon difference of means results, comparing means for 1) Station WBGT, 2) Kestrel 5400 WBGT, and 3) Kestrel 5400 WBGT relative to Station WBGT across all WBGTs, and for when WBGT was equal to or greater than 26.7°, 29.5°, 31.1°, 32.2°, and 33.3°C. 95% confidence intervals are included in parentheses.

WBGT threshold (°C)	Station vs obs	Kestrel 5400 vs obs	Kestrel vs station
All	0.2°C (0.2°–0.3°C)	0.7°C (0.6°–0.7°C)	0.4°C (0.3°–0.4°C)
26.7	0.2°C (0.2°–0.3°C)	0.7°C (0.6°–0.7°C)	0.4°C (0.3°–0.4°C)
29.5	0.3°C (0.2°–0.3°C)	0.8°C (0.7°–0.8°C)	0.4°C (0.4°–0.5°C)
31.1	0.2°C (0.2°–0.3°C)	0.8°C (0.8°–0.9°C)	0.5°C (0.5°–0.5°C)
32.2	0.2°C (0.2°–0.3°C)	0.9°C (0.9°–0.9°C)	0.6°C (0.6°–0.7°C)

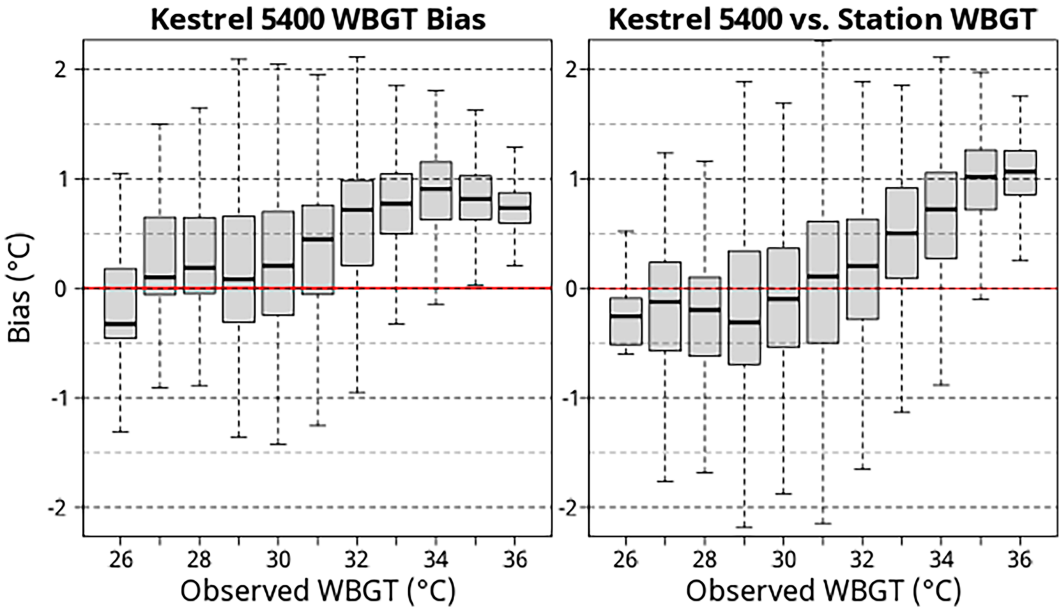

FIG. 5. Bias of the Kestrel 5400 WBGT (left) relative to Observed WBGT, and (right) relative to Station WBGT, both as a function of Observed WBGT, with the *x* axis indicating the Observed WBGT at 1°C increments.

WBGT bias had a slight negative (cool) median bias of −0.3°C (Fig. 4).

The bias of the Kestrel 5400 WBGT relative to the Station WBGT became increasingly positive as Observed WBGT increased (Fig. 5). At WBGTs less than or equal to 30°C, the Kestrel 5400 bias was slightly negative relative to the Station WBGT (Fig. 5). However, as WBGT increased, this bias became increasingly positive (warm), with the Kestrel 5400 WBGT median bias relative to the Station WBGT being +0.3° and +1.1°C, respectively, at Observed WBGTs of 31° and 35°C (Fig. 5). All comparisons of means between the Kestrel 5400 WBGT and Station WBGT were statistically significant (Table 2).

A disaggregation of the biases of the Station and Kestrel 5400 WBGT reveals that the natural wet-bulb temperature for both was positively (warm) biased (Table 2). The magnitude of bias was lowest for the Station, with a bias of 0.7°C, as compared with the Kestrel 5400, which had a bias of 1.5°C (Table 2). Contrasting this, the Station and Kestrel 5400 black globe temperatures were negatively (cool) biased. As was the case with the natural wet-bulb temperature, the station estimation of black globe temperature was more accurate than the Kestrel 5400, with biases of −1.5° and −2.2°C, respectively (Table 2).

WBGT ACCURACY METRICS

Analyses of the hit rate, false alarm ratio, and bias scores further revealed the differences between the Station WBGT and Kestrel 5400 WBGT (Fig. 6), with the Station WBGT better matching the Observed WBGT. The hit rate was comparable until WBGT exceeded 29°C, at which point the Kestrel 5400 WBGT hit rate continuously decreased, while the Station WBGT hit rate was relatively consistent from 30.5° to 34°C (Fig. 6). Likewise, the false alarm ratio was similar between the two until WBGT exceeded 29.5°C. At WBGTs greater than or equal to 32°C, the Kestrel 5400 false alarm ratio ranged from 0.9 to 0.99, while the Station WBGT false alarm ratio ranged between 0.75 and 0.82 (Fig. 6). Last, the Kestrel 5400 and Station WBGT bias scores at WBGTs less than 31°C were similar, being slightly cool-biased (Fig. 6). Overall, the Station WBGT bias scores at dangerously high WBGTs of 32.2°–34.5°C were lower than the Kestrel 5400 WBGT bias scores, although both were positively biased (Fig. 6).

b. WBGT variations: Surface types and shade

1) SURFACE TEMPERATURE

The observed surface temperatures on bare ground, alive grass, and dormant grass differed with respect to both their

TABLE 2. Wilcoxon difference of means to assess differences of the natural wet-bulb temperature and black globe temperature for the Station and the Kestrel 5400 measurements.

	Station vs obs	Kestrel 5400 vs obs	Kestrel vs station
Natural wet-bulb temperature	0.7°C (0.7°–0.7°C)	1.5°C (1.4°–1.5°C)	0.7°C (0.7°–0.8°C)
Black globe temperature	−1.5°C (from −1.6° to −1.4°C)	−2.2°C (from −2.3° to −2.1°C)	−0.9°C (from −1° to −0.8°C)

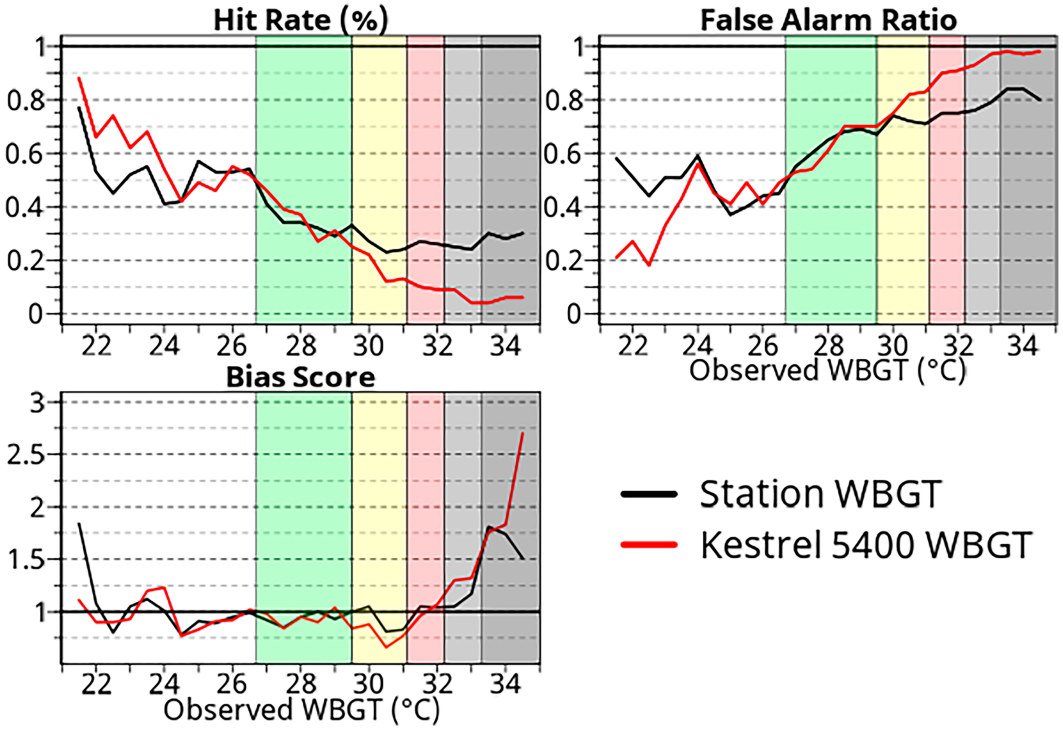

FIG. 6. WBGT accuracy metrics, showing the hit rate, false alarm ratio, and bias score for the Station and Kestrel 5400 WBGT at 0.5°C increments. Shading represents the corresponding WBGT flag level of the values (including the standards that have 33.3°C as the threshold for black flag).

mean and ranges (Fig. 7). The "alive grass" (tall fescue) was on average 7–10 cm in height. Overall, the range of observed surface temperatures was greatest for bare ground (21°–60°C), followed by dormant grass (24°–54°C), and then alive grass (24°–49°C) (Fig. 7). The surface temperature of bare ground was the highest, with a mean of 40°C, and 25th and 75th percentile values of 34° and 46°C, respectively (Fig. 7). Alive grass had a mean temperature of 34°C, while dormant grass had a mean of 36.5°C (Fig. 7). Both grass types had similar 25th percentiles of roughly 24°C. However, the 75th percentile was higher for dormant grass (43°C as compared with 38°C) (Fig. 7).

The mean bias of the Station black globe temperature [utilizing the Liljegren et al. (2008) method] was positive (warm) when the air temperature was warmer than the surface temperature, and negatively (cool) biased when surface temperature was warmer than air temperature (Fig. 8). The patterns in bias between the Station WBGT and Kestrel 5400 WBGT were similar when assessed as a function of the difference between the surface and air temperature, with both being positive (warm) when surface temperature was less than air temperature. Also, both WBGTs had a warm bias when surface temperature was greater than air temperature. When surface temperature was 6°–21°C higher than air temperature, the median bias of the Station WBGT ranged between 0° and 0.5°C, as compared with median biases for the Kestrel 5400 ranging from 0.7° to 0.85°C (Fig. 8).

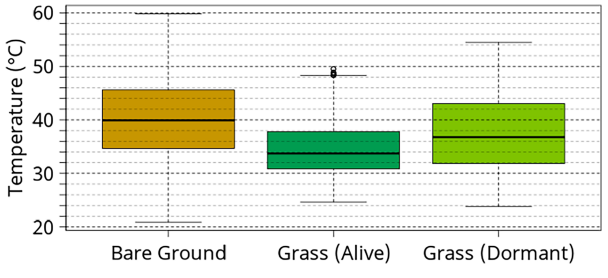

FIG. 7. Observed surface temperatures.

2) COMPARISON ACROSS SURFACES AND SHADE CONDITIONS

Statistically significant differences were identified in the means of the WBGT (measured with the Kestrel 5400) devices over a tennis court and neighboring grassy field, with the tennis court being 0.9°C (0.3°–1.4°C) warmer (Table 3).

Overall, the WBGT measured in the sun was around 4°C warmer than WBGT in the shade (excluding shaded observations in the forest) (Fig. 9). However, the difference between WBGT measured in the sun and WBGT measured in a shaded forest was greater, with a difference in WBGT of 5.4°C (5.2°–5.6°C) (Table 3). Conifers were predominant at the forest site with intermittent hard wood trees. The density of tree coverage was such that the forest floor was consistently shaded, with approximately two trees per square meter.

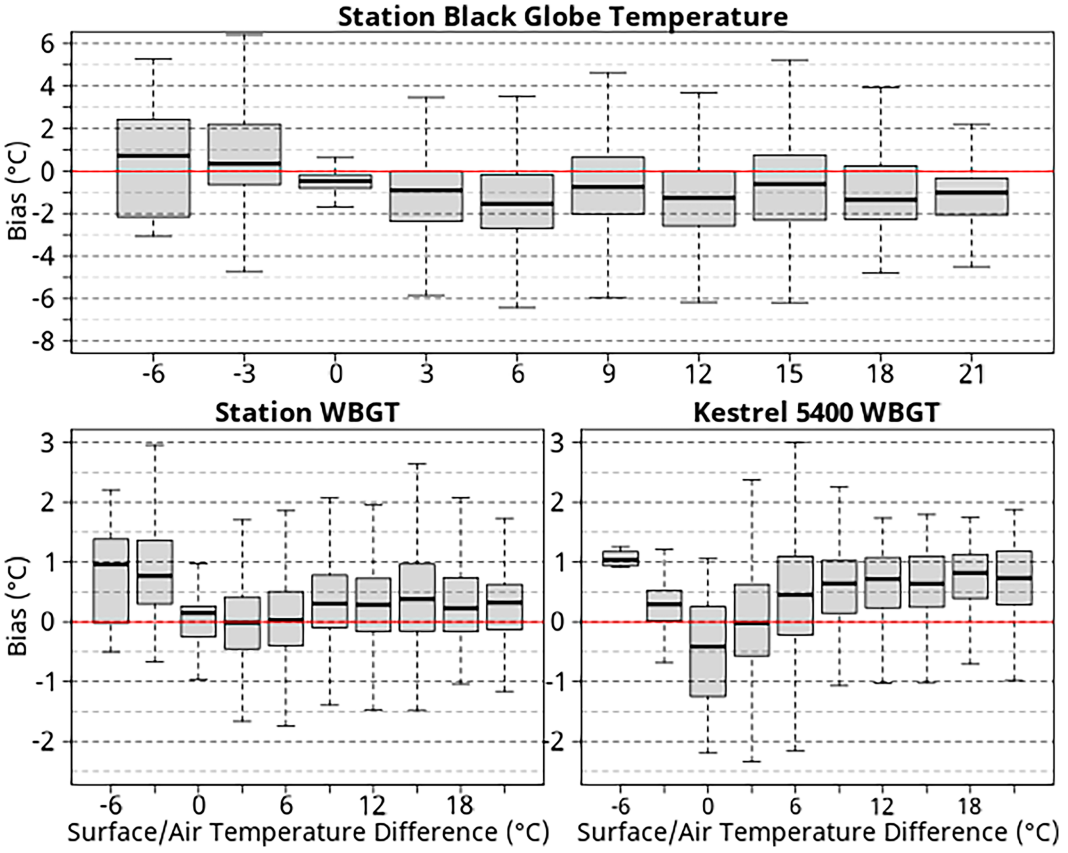

FIG. 8. Black globe temperature biases relative to difference between surface and air temperature, showing (top) a comparison of the Station vs observed black globe temperature and (bottom) the Station and Kestrel 5400 WBGT biases in comparison with Observed WBGT as a function of the difference between surface and air temperature (surface minus air temperature).

As Observed WBGT increased, the differences between the WBGT in the sun and shade steadily increased, with the median difference being +0.5°C at 25°C, +2.9°C at 29°C, and +5.1°C at 34°C (Fig. 9). A statistically significant difference between the means of the WBGT in the sun and shade were found, with a difference of 3.8°C (3.4°–4.1°C) (Table 3).

The difference in WBGT measured over 1) a grassy field that had been irradiated periodically throughout the day and 2) bare ground within a forest with a thick canopy and consistent shade was statistically significant. The WBGT measured over the shaded grassy field was 0.4°C (0.2°–0.9°C) warmer than the WBGT measured in the forested shade (Table 3).

c. *Estimating clear-sky radiation*

The methods for estimating clear-sky radiation and modifying that radiation by percentage cloud cover all performed well, but there was some variability in bias. A comparison of the estimated radiation from each of the three methods for

TABLE 3. Differences in WBGT between locations, showing the difference in means of WBGT measured concurrently by the two Kestrel 5400 devices in different locations and shade conditions. The 95% confidence interval is included in parentheses. Positive values indicate that the values at "Location 1" were higher than those at "Location 2." Sun and shade comparisons were based on measurements at all locations across three summers (multiple weeks of measurements per summer). Field and forest comparisons are based on measurements on two days at CRHS (15 and 19 Aug 2019, during 1145–1430 local time) (n = 140). The tennis court and field comparison is based on measurements on two days during 1145–1430 local time at CRHS (15 and 19 Aug 2019) and two days at GLHS (17 Jul and 7 Aug 2019) (n = 180).

Location 1	Location 2	Diff in WBGT means
Sun	Shade	3.8°C (3.4°–4.1°C)
Sunned field (n = 185)	Shaded forest (n = 33)	5.4°C (5.2°–5.6°C)
Shaded field (n = 140)	Shaded forest (n = 33)	0.4°C (0.2°–0.9°C)
Tennis court (n = 180)	Grass field (n = 180)	0.9°C (0.3°–1.4°C)

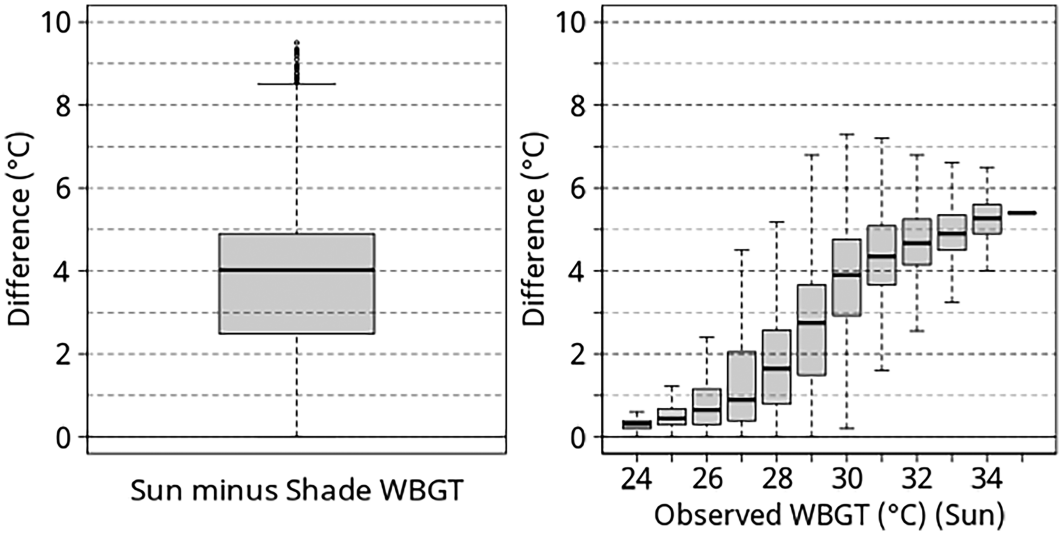

FIG. 9. WBGT sun vs WBGT shade. The difference is calculated with sunned WBGT minus shaded WBGT.

estimating clear-sky radiation, including modifying those estimates by percentage cloud cover, showed that all methods were within ±75 W m^{-2}, excluding the Ryan model with the default TC set to 0.8 (Fig. 10). Under clear sky or nearly clear sky (when cloud-cover percentage was less than or equal to 5%), the Kasten method showed the largest bias (median of +72 W m^{-2}), while the Ryan method had the lowest when TC was set to 0.65 (median of −15 W m^{-2}) (Fig. 10). Further, the Ryan method was slightly negative, while the other two methods were positively biased (Fig. 10). As with the Ryan method, the default settings in the Bras method for atmospheric turbidity resulted in estimated solar radiation being too high. Adjusting this by changing the value representing this turbidity from 3 to 4 improved results, but the Ryan method (0.65 TC) remained more accurate (Fig. 10).

The biases with all three methods for estimated solar radiation derived from modifying clear-sky radiation by percentage cloud cover are shown in Fig. 10. The pattern in bias between the methods is similar to the accuracy of the clear-sky radiation discussed above, with the Kasten method having the largest bias and the Ryan method having the smallest bias (Fig. 10). However, the mean biases of all methods were positive (too high) relative to observations, with means of 80 W m^{-2} (Kasten), 7 W m^{-2} (Ryan TC 0.65), and 65 W m^{-2} (Bras, default parameters) (Fig. 10).

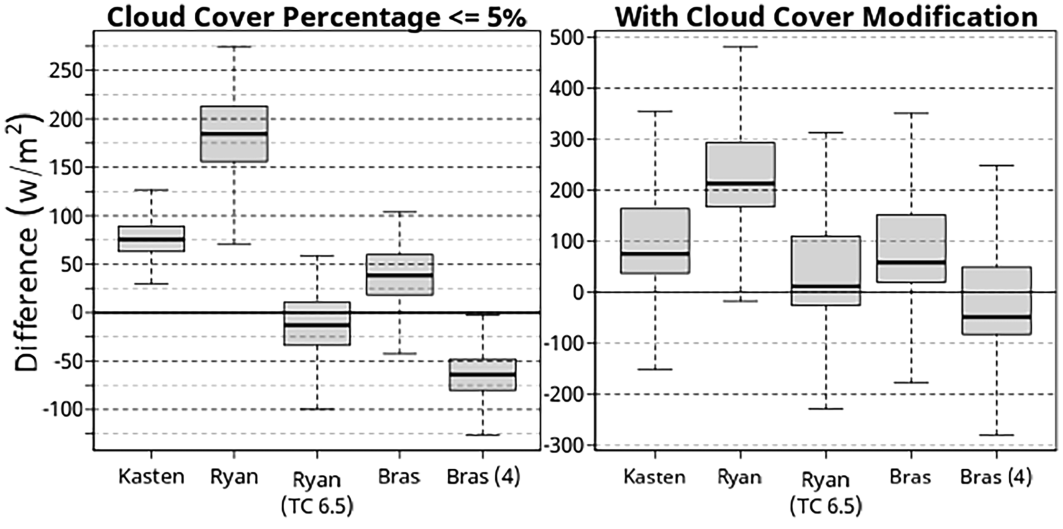

FIG. 10. Clear-sky radiation estimates: (left) Comparison of estimated clear-sky radiation and observed solar radiation from the 1) Kasten, 2) Ryan (with default TC of 0.8 and 0.65), and 3) Bras methods (with default TC setting of 3 and 4) under cloud-free sky. (right) Comparison of estimated radiation and observed solar radiation after modifying by cloud-cover percentage.

5. Discussion and conclusions

This research addressed four specific questions. The first two questions were concerned with discerning and comparing the accuracy of 1) estimated WBGT from a weather station using the Liljegren et al. (2008) method and 2) the WBGT as measured by a Kestrel 5400 WBGT meter. The Liljegren estimation method had a lower bias than the Kestrel 5400 when compared with Observed WBGT, with overall biases being +0.2° and +0.7°C warmer than Observed WBGT, respectively (Fig. 4) (Table 1). Furthermore, the bias of the Liljegren estimation does not vary significantly as WBGT increases, which was the case for the Kestrel 5400 (Fig. 4). The Station (Liljegren) WBGT and Kestrel 5400 WBGT hit rates, bias scores, and false alarm ratios were similar when WBGT was less than 29.5°C. As had already been determined, the Kestrel 5400 WBGT being more positively biased than the Observed WBGT resulted in relatively low hit rates at the most dangerous WBGTs and high numbers of false alarms.

Additionally, the Liljegren method assumes that the surface temperature is equivalent to the air temperature in the equation for estimating the black globe temperature. Given the often large differences between air and surface temperature, particularly over surfaces such as bare ground or tennis courts, this research question sought to gauge the sensitivity of the black globe estimation to this assumption and what bias it may introduce. Based on comparing the Liljegren WBGT biases when surface and air temperature were equivalent, or the surface temperature was up to 21°C warmer, little difference was revealed (Fig. 8). Thus, this assumption was not found to be a significant contributor to the overall bias of the Station WBGT, further evident given that the Station WBGT is positively biased, yet the station black globe temperature is negatively biased, as surface temperatures become increasingly warmer than the air temperature.

The third research question evaluated the differences in WBGT between measurement locations, for example, sun versus shade, and how WBGT varied as a function of surface type and the surface temperature of those surfaces. In addition to sun versus shade comparisons, measurements of WBGT from within a forested area and over tennis courts were compared. Overall, the three surfaces investigated in this work (bare ground, alive grass, and dormant grass) had markedly different surface temperatures (Fig. 7). The bare ground temperatures were the highest, largely due to the relatively lower albedo of this surface, which allowed more radiation to be absorbed. The surface temperature of dormant grass was higher than alive grass, which is similarly related to albedo (Fig. 7). However, this difference is also related to the relatively higher latent heat flux of the alive grass (with the dormant grass having a higher sensible heat flux) since the alive grass is transpiring. The grass being alive is also indicative of moister soil and increased latent heat flux.

A statistically significant difference was found in the sun versus the shade (Table 3), with the WBGT in the sun being 3.8°C (3.4°–4.1°C) warmer, and these differences increased as WBGT increased. From an average difference of around +0.5°C when WBGT was 25°C, when WBGT was equal or greater to 29° and 34°C, the WBGT in the sun was on average 2.9° and 5.1°C warmer than concurrent shaded WBGT measurements, respectively. The largest differences at the most extreme WBGTs are driven by the increased differences in the radiative components of the environment between sun and shade that influence heat stress and WBGT.

The difference between WBGT in the sun (grassy field) and WBGT in a shaded forest was statistically significant, with the sunned grassy field being +5.4°C (5.2°–5.6°C) warmer than the shaded forest WBGT. Additionally, the shaded WBGT measurements made over a grassy field versus shaded measurements from within a forest were +0.4°C (0.2°–0.9°C) warmer than the concurrent measurements in the forest, despite the inherent lower wind speeds within dense tree cover that would contribute to relatively higher WBGTs. These differences are hypothesized to be at least partially a result of the amount of radiation the respective surfaces had received throughout the day. This hypothesis would be strengthened by future research with a larger sample size. The location in the grassy field was in direct sunlight for approximately 4–5 h on average before becoming shaded by trees, after which the measurements were taken. However, the measurement location in the forest was selected based on the full coverage of shade provided by a thick tree canopy. Thus, while both sets of measurements were made concurrently in the shade, the WBGTs over the surface that had been irradiated earlier in the day were warmer than the WBGTs over the surface that had been shaded throughout the day.

Last, WBGT measurements made over a tennis court were +0.9°C (0.3°–1.4°C) warmer than concurrent measurements made over a neighboring grassy field, with roughly 200 ft between the sites. This differs from the results in Grundstein and Cooper (2020) and Kopec (1977), which showed minimal difference in WBGT between a hard surface (tennis court) and grass. This difference is hypothesized to result from differences in the weather conditions in the studies, as the mean WBGT observed in Grundstein and Cooper (2020) over all surfaces was less than 28°C while the mean WBGT of concurrent observations on a tennis court and grass was 32.2°F and 31.4°C, respectively, in this study.

The implications of these differences are important because the conditions in this study were significantly more thermally stressful, and it is at these WBGTs that health effects could be ameliorated with awareness of the differences across surfaces. Grundstein and Cooper (2020) and Kopec (1977) note that competing factors likely contributed to their finding of little difference between a tennis court and a grassy field (e.g., higher dewpoint temperatures and lower surface temperatures over the grass and the reverse over the tennis court). Thus, additional research could confirm if these competing factors are less impactful at the most extreme heat stress levels and provide further clarity on differences between tennis courts and grass overall.

The fourth question addressed in this research compared the accuracy of different methods for estimating clear-sky radiation and the accuracy of modifying that clear-sky radiation by percentage cloud cover. The three methods assessed were 1) Kasten model (Kasten and Czeplak 1980), 2) Ryan–Stolzenbach

model (Ryan and Stolzenbach 1972), and 3) Bras model (Bras 1990). Between these methods, with the default parameters (atmospheric transmissivity) in the Ryan and Bras models, the Bras model had the lowest bias in estimating clear-sky radiation. However, setting the TC to 0.65 instead of the default 0.8 in the Ryan model resulted in clear-sky radiation estimates that outperformed all other methods tested, and outperformed the resulting change in bias for the Bras model when modifying its TC (Fig. 10). Importantly, the use of varied settings of the TC should be validated further. Changing the TC setting too drastically from the default would not be recommended, especially in a different climate or season, but awareness that there are substantive differences between methods is important as it may help explain unexpected differences in the results of a given study.

The modification of clear-sky radiation by the calculated percentage cloud cover from the images of sky conditions performed well in comparison with observed solar radiation, with median differences of within ± 75 W m^{-2}. The differences between methods, in this case, mirrored the differences with estimating clear-sky radiation. While the Ryan method with an TC of 0.65 underestimated clear-sky radiation, this method with cloud cover was slightly positive (too high relative to observed radiation) by 6.9 W m^{-2} (-17.7–31.6 W m^{-2}), but overall, not different from zero to a statistically significant degree.

This study had limitations that are important to note: First, the results and comparisons made here may differ in other climates and regions, but similarities in other humid, subtropical locations would be expected. While hot and humid environments are dangerous, they are challenging to assess here due to the varying influences of humidity and wind. The suite of challenges is different for drier environments, where the influences of radiation (shortwave and longwave) would have a more profound influence. Second, the image analysis to classify and calculate percentage cloud cover could be improved in future research by more robust image classification, which is increasingly possible with the advancements in machine learning algorithms. Future research should also address the varying character of cloud types (e.g., height and density) as that certainly impacts the radiation ultimately incident at the surface. However, this would require a reformulation of modifying clear-sky radiation estimates by cloud cover, since the method used here simply allows for one value to represent the total sky conditions (4).

Overall, the research presented here further confirmed prior research highlighting the relative accuracy of the Liljegren et al. (2008) method for estimating WBGT. The Kestrel 5400 WBGT meter had a higher bias, particularly when WBGT was greater than or equal to 32°C (+0.7°–0.9°C), which is up to 0.2°C greater than the stated accuracy of the device by the manufacturer (0.7°C) (Kestrel 2020). Relatively small bias has major implications when it comes to WBGT, given that the difference between flag levels is only 1°–3°C. The bias here resulted in the device "erring on the side of caution," but if the reverse were true, the device would indicate conditions were safe for activity when conditions could be quite threatening to health. Despite these biases, the authors regard the Kestrel 5400 WBGT meter as a reliable device for its price point, although additional emphasis by the manufacturers on proper calibration of these instruments is critical. Last, it is important to reemphasize that WBGT measurements should be taken onsite and the time of activity, as estimated WBGT can vary in accuracy based on conditions, with values upon which decisions are based being averaged over time periods of 15–30 min given the rapid variability of WBGT over small time periods.

Awareness of the increasingly large difference between the sun and shade WBGT as WBGT in the sun increases provides further basis for encouraging activity to be moved to the shade on the most thermally stressful days, if possible, and emphasizes the importance of on-site measurements. This also should encourage future planning of outdoor athletic facilities to factor in shade access more heavily for the hottest times of the day. Last, the identification of a more accurate method for estimating clear-sky radiation, relative to the observations assessed here in this study, and confirmation of the accuracy of modifying that radiation by percentage cloud cover, provides useful insight for future efforts to estimate WBGT in real time from weather stations that do not measure solar radiation. This finding also emboldens efforts to forecast WBGT that rely on forecast cloud-cover values, ultimately reducing uncertainty with estimating WBGT from estimated variables themselves.

Acknowledgments. The authors acknowledge the staff at Horace Williams Airport in Chapel Hill that allowed for WBGT data collection and the North Carolina high schools that permitted collection of data across their campuses: Cedar Ridge High School in Hillsborough, Green Level High School in Cary, and Wake Forest High School in Wake Forest. The authors also thank Dr. Andrew Grundstein at the University of Georgia for permitting the use of the ISO-compliant WBGT meter.

Data availability statement. Because of the proprietary nature of the primary data used in this research and the ongoing use of some data for future publications, the datasets are currently not publicly available. A precise roadmap for replicating these results with similar data is provided in the data and methods section, with the methods for estimating WBGT based on standard meteorological variables provided in Liljegren et al. (2008) and in the R package "wbgt." The methods for creating superpixels from collecting images of cloud cover are detailed in the documentation for the R package "SuperpixelImageSegmentation."

REFERENCES

Achanta, R., A. Shaji, K. Smith, and A. Lucchi, 2010: SLIC superpixels. EPFL Tech. Rep. 149300, 15 pp., https://infoscience.epfl.ch/record/149300/files/SLIC_Superpixels_TR_2.pdf.

——, ——, ——, ——, P. Fua, and S. Süsstrunk, 2012: SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.*, **34**, 2274–2282, https://doi.org/10.1109/TPAMI.2012.120.

Annear, R. L., and S. A. Wells, 2007: A comparison of five models for estimating clear-sky solar radiation. *Water Resour. Res.*, **43**, W10415, https://doi.org/10.1029/2006WR005055.

Barry, R. G., and P. D. Blanken, 2016: *Microclimate and Local Climate*. Cambridge University Press, 316 pp.

Boyer, T., 2022: NDFD wet bulb globe temperature algorithm and software design. NWS Meteorological Development Laboratory Doc., 8 pp., https://vlab.noaa.gov/documents/6609493/7858379/NDFD+WBGT+Description+Document.pdf.

Bras, R. L., 1990: *Hydrology: An Introduction to Hydrologic Science*. Addison-Wesley, 643 pp.

Buchner, J., 2013: imagehash: A Python perceptual image hashing module. GitHub, https://github.com/JohannesBuchner/imagehash.

Budd, G. M., 2008: Wet-bulb globe temperature (WBGT)—Its history and its limitations. *J. Sci. Med. Sport*, **11**, 20–32, https://doi.org/10.1016/j.jsams.2007.07.003.

Burakowski, E., A. Tawfik, A. Ouimette, L. Lepine, K. Novick, S. Ollinger, C. Zarzycki, and G. Bonan, 2018: The role of surface roughness, albedo, and Bowen ratio on ecosystem energy balance in the eastern United States. *Agric. For. Meteor.*, **249**, 367–376, https://doi.org/10.1016/j.agrformet.2017.11.030.

Carter, A. W., B. F. Zaitchik, J. M. Gohlke, S. Wang, and M. B. Richardson, 2020: Methods for estimating wet bulb globe temperature from remote and low-cost data: A comparative study in central Alabama. *GeoHealth*, **4**, e2019GH000231, https://doi.org/10.1029/2019gh000231.

Casa, D. J., and Coauthors, 2015: National Athletic Trainers' Association position statement: Exertional heat illnesses. *J. Athl. Train.*, **50**, 986–1000, https://doi.org/10.4085/1062-6050-50.9.07.

Centers for Disease Control and Prevention, 2010: Climate change and extreme heat events. CDC Doc., 20 pp., https://www.cdc.gov/climateandhealth/pubs/climatechangeandextremeheatevents.pdf.

Cooper, E., A. Grundstein, A. Rosen, J. Miles, J. Ko, and P. Curry, 2017: An evaluation of portable wet bulb globe temperature monitor accuracy. *J. Athl.Train.*, **52**, 1161–1167, https://doi.org/10.4085/1062-6050-52.12.18.

Coutts, A. M., E. C. White, N. J. Tapper, J. Beringer, and S. J. Livesley, 2016: Temperature and human thermal comfort effects of street trees across three contrasting street canyon environments. *Theor. Appl. Climatol.*, **124**, 55–68, https://doi.org/10.1007/s00704-015-1409-y.

Covert, J., and R. Hellström, 2014: An indirect method for predicting road surface temperature in coastal areas with snowy winters. Geography Faculty Publications Paper 11, 8 pp., https://vc.bridgew.edu/cgi/viewcontent.cgi?article=1010&context=geography_fac.

Dugord, P.-A., S. Lauf, C. Schuster, and B. Kleinschmit, 2014: Land use patterns, temperature distribution, and potential heat stress risk—The case study Berlin, Germany. *Comput. Environ. Urban Syst.*, **48**, 86–98, https://doi.org/10.1016/j.compenvurbsys.2014.07.005.

EPA, 2000: Meteorological monitoring guidance for regulatory modeling applications. Tech. Rep. EPA-454/R-99-005, 171 pp., https://www.epa.gov/sites/production/files/2020-10/documents/mmgrma_0.pdf.

Foster, J., J. W. Smallcombe, S. Hodder, O. Jay, A. D. Flouris, and G. Havenith, 2022: Quantifying the impact of heat on human physical work capacity; Part II: The observed interaction of air velocity with temperature, humidity, sweat rate, and clothing is not captured by most heat stress indices. *Int. J. Biometeor.*, **66**, 507–520, https://doi.org/10.1007/s00484-021-02212-y.

Frank, C. W., B. Pospichal, S. Wahl, J. D. Keller, A. Hense, and S. Crewell, 2020: The added value of high resolution regional reanalyses for wind power applications. *Renewable Energy*, **148**, 1094–1109, https://doi.org/10.1016/j.renene.2019.09.138.

Grundstein, A., and E. Cooper, 2018: Assessment of the Australian Bureau of Meteorology wet bulb globe temperature model using weather station data. *Int. J. Biometeor.*, **62**, 2205–2213, https://doi.org/10.1007/s00484-018-1624-1.

——, and ——, 2020: Comparison of WBGTs over different surfaces within an athletic complex. *Medicina*, **56**, 313, https://doi.org/10.3390/medicina56060313.

——, C. Williams, M. Phan, and E. Cooper, 2015: Regional heat safety thresholds for athletics in the contiguous United States. *Appl. Geogr.*, **56**, 55–60, https://doi.org/10.1016/j.apgeog.2014.10.014.

——, E. R. Cooper, O. Cahill, and C. Walker, 2022: Are local weather stations a feasible substitute for on-site measurements for heat stress assessment in sports? *Int. J. Sports Sci. Coaching*, **18**, 2178–2188, https://doi.org/10.1177/17479541221117240.

Haghighat, M., 2015: gabor: Gabor feature extraction. GitHub, https://github.com/mhaghighat/gabor.

——, S. Zonouz, and M. Abdel-Mottaleb, 2015: CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Syst. Appl.*, **42**, 7905–7916, https://doi.org/10.1016/j.eswa.2015.06.025.

Hardin, A. W., and J. K. Vanos, 2018: The influence of surface type on the absorbed radiation by a human under hot, dry conditions. *Int. J. Biometeor.*, **62**, 43–56, https://doi.org/10.1007/s00484-017-1357-6.

Hawkins, M. D., V. Brown, and J. Ferrell, 2017: Assessment of NOAA National Weather Service methods to warn for extreme heat events. *Wea. Climate Soc.*, **9**, 5–13, https://doi.org/10.1175/WCAS-D-15-0037.1.

Hondula, D. M., J. K. Vanos, and S. N. Gosling, 2014: The SSC: A decade of climate–health research and future directions. *Int. J. Biometeor.*, **58**, 109–120, https://doi.org/10.1007/s00484-012-0619-6.

Jolliffe, I. T., and D. B. Stephenson, Eds., 2012: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. John Wiley and Sons, 296 pp.

Kasten, F., and G. Czeplak, 1980: Solar and terrestrial radiation dependent on the amount and type of cloud. *Sol. Energy*, **24**, 177–189, https://doi.org/10.1016/0038-092X(80)90391-6.

Kestrel, 2020: Kestrel certificate of conformity. 2 pp., https://kestrelinstruments.com/mwdownloads/download/link/id/41.

——, 2021: Users guide Kestrel 5400 heat stress tracker. 24 pp., https://kestrelinstruments.com/mwdownloads/download/link/id/14/.

Kopec, R. J., 1977: Response of the wet-bulb-globe-thermometer heat stress index to selected land use surfaces. *Southeast. Geogr.*, **17**, 133–145, https://doi.org/10.1353/sgo.1977.0009.

Kosaka, E., A. Iida, J. Vanos, A. Middel, M. Yokohari, and R. Brown, 2018: Microclimate variation and estimated heat stress of runners in the 2020 Tokyo Olympic marathon. *Atmosphere*, **9**, 192, https://doi.org/10.3390/atmos9050192.

Lee, H., H.-G. Kwon, S. Ahn, H. Yang, and C. Yi, 2023: Estimation of perceived temperature of road workers using radiation and meteorological observation data. *Remote Sens.*, **15**, 1065, https://doi.org/10.3390/rs15041065.

Lemke, B., and T. Kjellstrom, 2012: Calculating workplace WBGT from meteorological data: A tool for climate change assessment. *Ind. Health*, **50**, 267–278, https://doi.org/10.2486/indhealth.MS1352.

Lieblich, M., and J. Spector, 2017: WBGT: Wet bulb globe temperature, version 1.2. R package, https://github.com/mdljts/wbgt.

Liljegren, J. C., R. A. Carhart, P. Lawday, S. Tschopp, and R. Sharp, 2008: Modeling the wet bulb globe temperature using standard meteorological measurements. *J. Occup. Environ. Hyg.*, **5**, 645–655, https://doi.org/10.1080/15459620802310770.

Mouselimis, L., 2022a: OpenImageR: An image processing toolkit version 1.2.7. R package, https://CRAN.R-project.org/package=OpenImageR.

——, 2022b: SuperpixelImageSegmentation: Image segmentation using superpixels, affinity propagation and Kmeans clustering, version 1.0.5. R package, https://CRAN.Rproject.org/package=SuperpixelImageSegmentation.

National Weather Service, 2020: Weather related fatality and injury statistics. NOAA, accessed 20 June 2021, https://www.weather.gov/hazstat/.

NCAR, 2015: verification: Weather forecast verification utilities, version 1.42. R package, https://CRAN.R-project.org/package=verification.

NCHSAA, 2016: Health and safety. North Carolina High School Athletic Association, https://www.nchsaa.org/health-and-safety/.

Oke, T. R., 1987: *Boundary Layer Climates*. 2nd ed. Routledge, 435 pp.

——, G. Mills, A. Christen, and J. A. Voogt, 2017: *Urban Climates*. Cambridge University Press, 548 pp.

OSHA, 2017: OSHA instruction directive TED-01-00-015. OSHA Technical Manual, https://www.osha.gov/dts/osta/otm/otm_iii/otm_iii_4.html.

Parsons, K., 2006: Heat stress standard ISO 7243 and its global application. *Ind. Health*, **44**, 368–379, https://doi.org/10.2486/indhealth.44.368.

Patel, T., S. P. Mullen, and W. R. Santee, 2013: Comparison of methods for estimating wet-bulb globe temperature index from standard meteorological measurements. *Mil. Med.*, **178**, 926–933, https://doi.org/10.7205/MILMED-D-13-00117.

Pryor, J. L., R. R. Pryor, A. Grundstein, and D. J. Casa, 2017: The heat strain of various athletic surfaces: A comparison between observed and modeled wet-bulb globe temperatures. *J. Athl. Train.*, **52**, 1056–1064, https://doi.org/10.4085/1062-6050-52.11.15.

Ryan, P. J., and K. D. Stolzenbach, 1972: Environmental heat transfer. *Engineering Aspects of Heat Disposal from Power Generation*, D. R. F. Harleman, Ed., Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, Dept. of Civil Engineering, Massachusetts Institute of Technology, 1–75.

Sight Machine, 2012: SimpleCV: The open source framework for machine vision: findHOGFeatures function of the SimpleCV computer vision platform. GitHub, https://github.com/sightmachine/SimpleCV.

Stein, J. S., C. W. Hansen, and M. J. Reno, 2012: Global horizontal irradiance clear sky models: Implementation and analysis. SANDIA Tech. Rep. SAND2012-2389, 68 pp., https://www.osti.gov/servlets/purl/1039404.

Thorsson, S., F. Lindberg, I. Eliasson, and B. Holmer, 2007: Different methods for estimating the mean radiant temperature in an outdoor urban setting. *Int. J. Climatol.*, **27**, 1983–1993, https://doi.org/10.1002/joc.1537.

Vanos, J. K., E. Kosaka, A. Iida, M. Yokohari, A. Middel, I. Scott-Fleming, and R. D. Brown, 2019: Planning for spectator thermal comfort and health in the face of extreme heat: The Tokyo 2020 Olympic marathons. *Sci. Total Environ.*, **657**, 904–917, https://doi.org/10.1016/j.scitotenv.2018.11.447.

Wang, Y., F. Bakker, R. de Groot, H. Wortche, and R. Leemans, 2015: Effects of urban trees on local outdoor microclimate: Synthesizing field measurements by numerical modelling. *Urban Ecosyst.*, **18**, 1305–1331, https://doi.org/10.1007/s11252-015-0447-7.

Wu, J., J. Zha, D. Zhao, and Q. Yang, 2018: Changes in terrestrial near-surface wind speed and their possible causes: An overview. *Climate Dyn.*, **51**, 2039–2078, https://doi.org/10.1007/s00382-017-3997-y.

Yaglou, C. P., and D. Minard, 1957: Control of heat casualties at military training centers. *AMA Arch. Ind. Health*, **16**, 302–316.

Zhou, B., 2015: Image segmentation using SLIC superpixels and affinity propagation clustering. *Int. J. Sci. Res.*, **4**, 1525–1529.

Zhu, H., J. Xu, J. Hu, and J. Chen, 2017: Medical image segmentation using improved affinity propagation. *CompIMAGE 2016: Computational Modeling of Objects Presented in Images. Fundamentals, Methods, and Applications*, R. Barneva, V. Brimkov, and J. Tavares, Eds., Lecture Notes in Computer Science, Vol. 10149, Springer, 208–215, https://doi.org/10.1007/978-3-319-54609-4_15.