Dr. Alex Guenther
Editor, Geoscientific Model Development
Laboratory Fellow
Pacific Northwest National Laboratory
Richland WA USA
Date 24/11/14

Dr. Guenther:

Please consider our revised manuscript, "Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5". We thank the reviewers for their helpful, constructive comments to strengthen our manuscript.

> Editor's guidance:
>
> The response to the Referees shall be structured in a clear and easy to follow sequence: (1) comments from Referees, (2) author's response, (3) author's changes in manuscript. In addition, please provide a marked-up manuscript version showing the changes made

Both reviewers had similar comments about our manuscript, and we summarize key points below. Both reviewers stated the value of this manuscript:

> Reviewer 1: *This implementation could be **very useful** and so the effort is applauded, because it is **not possible to reconstruct peak (or average) heat stress from standard model outputs** such as peak temperature or from daily-mean meteorological quantities.*
>
> Reviewer 2: *This effort is **very valuable and potentially useful for a large scientific community** running this climate model or analysing its output.*

However, each reviewer also noted that the manuscript was 'wordy', 'lengthy', and 'poorly organized'. Additionally, the reviewers both stated they recommend redoing the figures, that the box-whisker plots (Fig. 4-7) are difficult to read, the maps are complicated (Fig. 2-3), and Figure 1 is unnecessary. The reviewers specifically wanted to see the advantages of online versus offline calculation:

> Reviewer 1: *What I really want to see in this paper is a **demonstration of the value added** by what the authors have done. How do online, time-step level metrics improve on what we could have done with archived model output?*
>
> Reviewer 2: *How do the results in the **online calculation differ from the indices calculated based on the model output**? To what extent does a calculation based on daily average output miss the extreme heat stress values?*

To this end, we have rewritten significant portions of the document to enhance clarity, removed ~6 pages to reduce the length, and replaced all of the figures.

- Our new Figure 1 is an assessment of:

    1. The minimal differences between new and legacy vapor pressure calculations on heat stress over the effective range of CLM4.5 (Fig. 1a).
    2. The importance of accurate moist thermodynamic calculations (Fig. 1b).
    3. The large systematic differences of online versus offline computation of these metrics at low and high frequency time fields (Fig. 1c,d).
- Our new Figure 2 describes a *T-Q* expected rank classification that may only be accomplished with model time step calculations because they arise from instantaneous covariances. This reduces the number of panels required to show how extreme *T* and extreme *Q* effect each heat stress metric (reducing the number of panels per figure by 1/3), thus, enhancing clarity.
- Our new Figures 3-6 apply the *T-Q* expected rank value to describe regimes of extreme moisture and extreme temperature. These results require online model calculations.
    1. We show that partitioning of *T* and *Q* begins in low latitudes at the 75th percentile, and expands as the heat stress metrics approach the 99th percentile (all novel results).
    2. We also show that the *T-Q* regimes are different between metrics (also a novel result) demonstrating the added value of online calculations of multiple metrics at the same time.

Both reviewers stated they wanted to see a model validation:

> Reviewer 1: *Or, another interesting question would be how well does the model do at reproducing observed values of these metrics?*
>
> Reviewer 2: *Furthermore, it would be desirable to evaluate the model's performance at some places with in-situ measurements or gridded data in order to see whether potential model biases in temperature and humidity add up or cancel out.*

CLM4.5, in this study, is driven by CRUNCEP, an observation corrected reanalysis product, so there would be little practical value in such further analysis. We believe validation at each weather station is outside the scope of our manuscript. An added value of the HumanIndexMod is enabling researchers to be able to execute effective model validation studies with commonly used heat stress metrics, and accurate, efficient moist thermodynamic quantities. These studies may use the HumanIndexMod to 1) calculate heat stress with their datasets, and 2) compare their results with CLM4.5 (and hopefully other land surface models).

Below, we address each reviewer comment point-by-point and indicate where we have altered the content of the manuscript. Reviewer's comments are italicized in blue. Our edited manuscript appears in red. Key recommendations or responses are in bold. We look forward to hearing from you.

Sincerely yours,

Jonathan R. Buzan, Keith Oleson, and Matthew Huber

Response to Reviewer 1

This paper describes an implementation of several heat-stress metrics into the land component of a climate model. This implementation could be very useful and so the effort is applauded, because it is not possible to reconstruct peak (or average) heat stress from standard model outputs such as peak temperature or from daily-mean meteorological quantities.

We thank the reviewer for the detailed response to our manuscript. We believe the comments have enhanced the readability and fine-tuned our objectives for the manuscript. As suggested by the reviewer, we rewrote the entire document. Below are the changes to the manuscript. We believe we have addressed all of the major and minor comments.

*The description and implementation is detailed which is useful, although the paper is wordy at times and its **50-page length should be shortened if possible without losing key information**. One thing that would help is to get rid of the Appendix (see point 42 below), and I recommend **eliminating Figure 1** (point 44) and **redoing or reconsidering nearly all the other figures**.*

We have:

- Rewritten the document to enhance clarity.
- Removed ~6 pages to reduce length, removing irrelevant information.
- Eliminated Figure 1.
- Redesigned all figures.

We believe we have addressed all issues raised by the reviewer. Below are detailed responses to each comment by the reviewer.

MAJOR COMMENTS:
I find that there are many problems with the manuscript which will require major revisions before it i s acceptable for publication. The principle problems are:

***Major Comment 1:** The figures are **unintelligible** and **require better explanations in the captions and text**. Ideally one should be able to understand the figures in a paper just based on reading the captions, but that would certainly not be possible here. Even after reading the text I was often unable to figure out what exactly they were showing. The description of the statistics computed is inadequate and/or incoherent.*

We have rewritten the figure captions. For example:

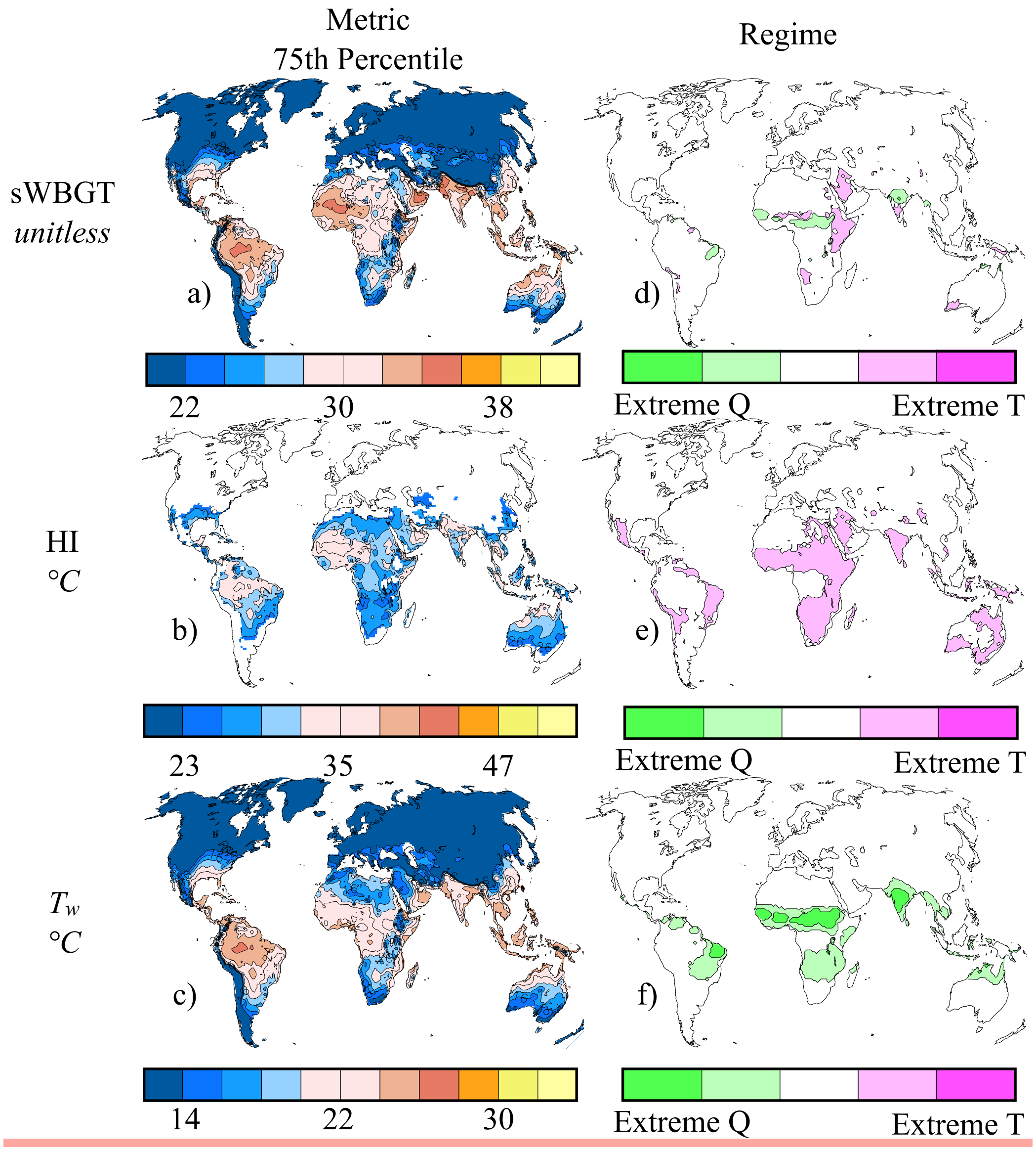


Figure 3. 75th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

Major Comment 2, Part 1: *A bigger point around the figures and indeed the results in general, is that I question whether the authors have chosen the right things to show.* ***Their main point*** *(if I have understood correctly what they have done, which might not be the case) seems to be t****hat in humid regions, heat stress thresholds are reached at lower temperatures and heat stress is less variable than in arid regions. These points are obvious*** *(and have been quantified in several of the papers cited), and* ***you don't need a GCM****, let alone online diagnostics, to show it. A secondary finding is that formulas used in GCMs for things like saturation water vapor pressure* *are not perfectly accurate, but i do not believe this point is significant (see 6 below).*

We apologize we were not clear. As noted by both reviewers, our figures did not adequately illustrate our main point of this section. The results, shown below, show that a GCM is required to determine what causes extreme high heat stress events.

An open question is what drives extreme high heat stress events, which are, by definition, rare events. For example, we cannot determine from the mean climate state or from theory, in a warm and humid climate, if abnormally high temperature, abnormally high moisture, or a combination of the two, caused a heat stress event. This is a question of the covariance of perturbations of temperature and humidity, not a statement of mean conditions, and there is no theory to explain these situations. Our new Figure 2 shows a new classification system, *T-Q* regimes. The new Figures 3-6 are applications of this new classification system, demonstrating that heat stress thresholds are not intuitively obvious.

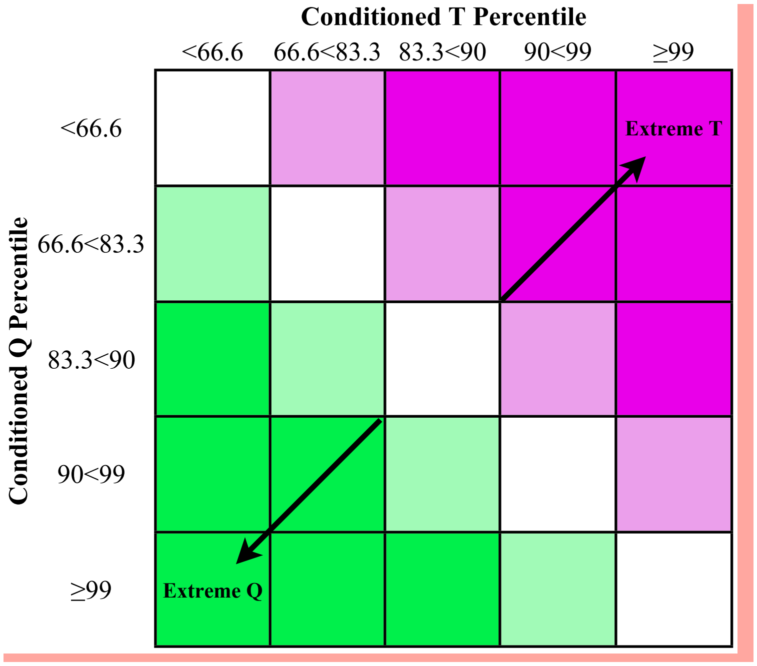


Figure 2. Expected value rank. *T* and *Q* conditioned upon exceedance value of a heat stress or moist thermodynamic metric. The *T* and *Q* values are compared to their respective time series as a percentile. These *T* and *Q* percentiles are binned and are compared to each other. Extreme *Q* are greens and extreme *T* are magentas.

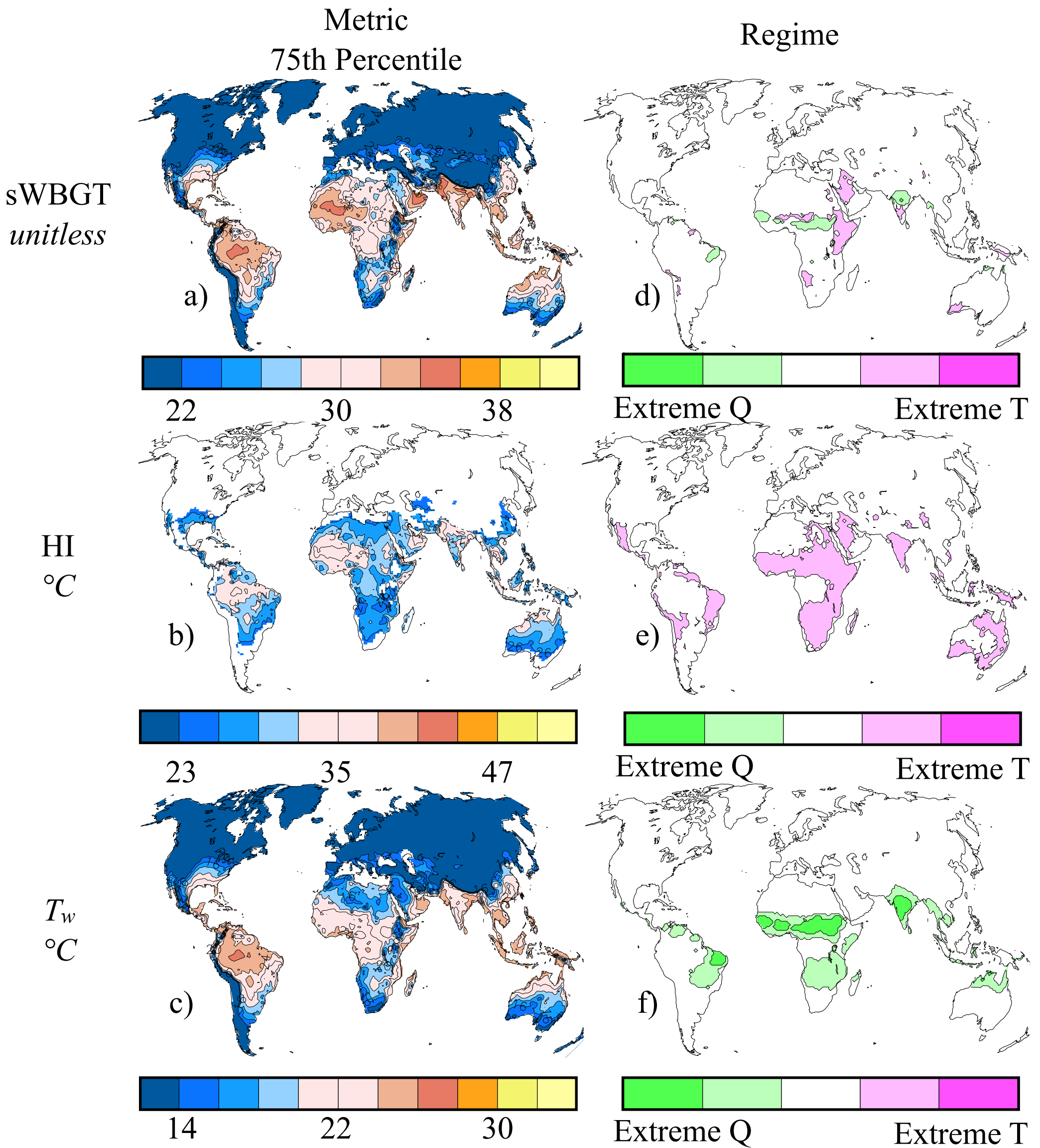


Figure 3. 75th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

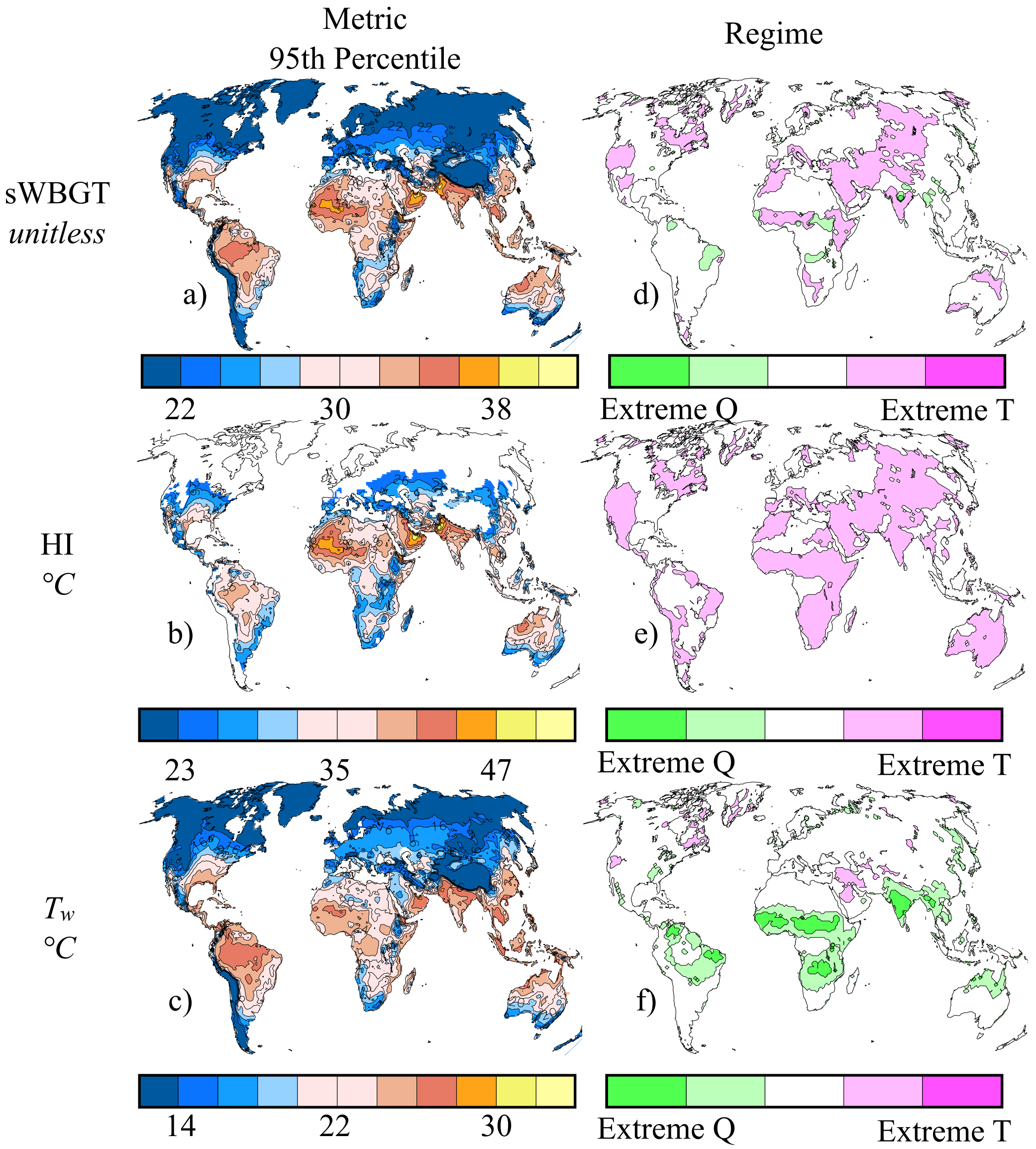


Figure 4. 95th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

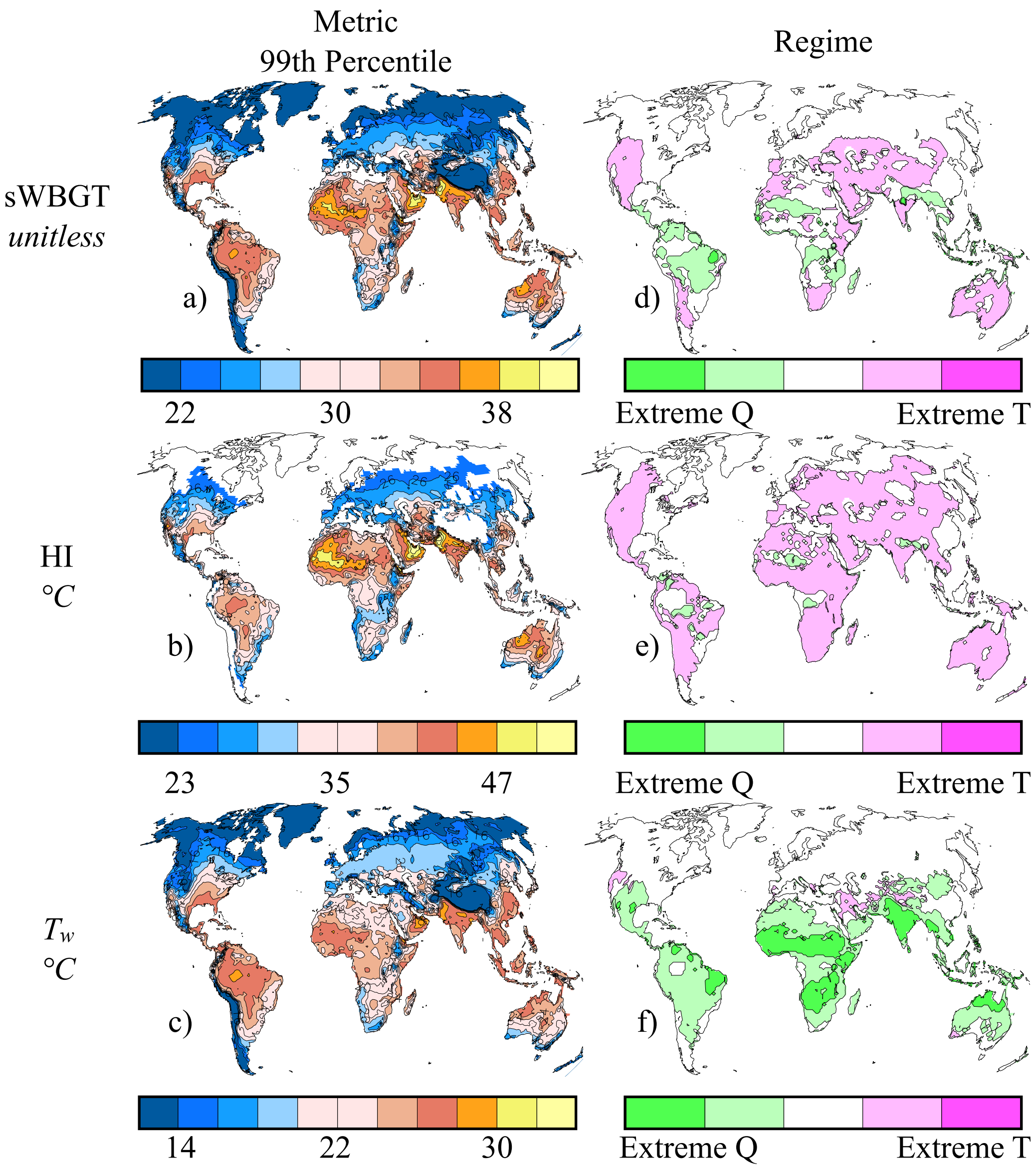


Figure 5. 99th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

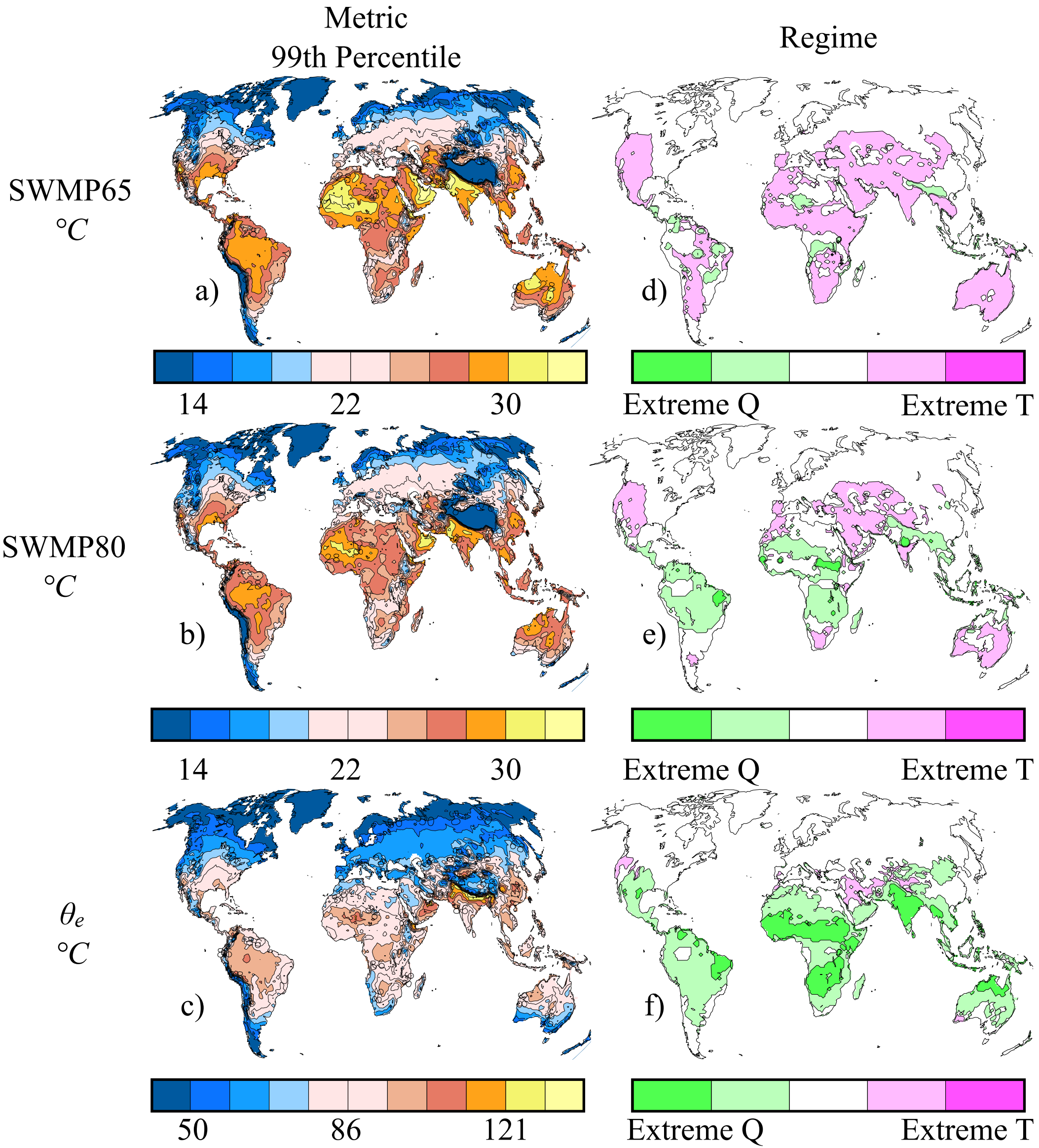


Figure 6. 99th percentile exceedance value of 3 metrics for a) SWMP65, b) SWMP80, and c) θ_E (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

Major Comment 2, Part 2: *What I really want to see in this paper is* ***a demonstration of the value added*** *by what the authors have done.* ***How do online, time-step level metrics improve on what we could have done with archived model output?*** *If we had daily or 3-hourly model output would that be enough to get the same results? Are the results they have shown, also evident in daily or even monthly mean fields? How independent are the different metrics, are there some that can be discarded as having no additional information content beyond others (some information is given on this but with very little discussion)?*

We realize that we were not clear what the value added to CLM4.5 was. We have replaced Figure 1 with our evaluation of the value added to CLM4.5. We have eliminated sections 2.1-2.2, incorporated elements into Section 3 to shorten the manuscript, and succinctly describe the value added by having the metrics calculated online.

revised portion of:

Section 3.2

Moist thermodynamic water vapor quantities in CLM4.5 are calculated within QSatMod. We use the outputs from QSatMod as the inputs to the HumanIndexMod. Within the HumanIndexMod, we created a subroutine, QSat_2, that has all the same functionalities as QSatMod. This subroutine uses the August-Roche-Magnus (ARM) equation (Eq. A.13), rather than the Flatau et. al. (1992) polynomial equations for vapor pressure in QSatMod. The log derivative of ARM (Eq. A.15) is a critical component of the calculation of T_w, and is not available in QSatMod. Additionally, QSat_2 calculates $f(\theta_E)$ (Eq. A.18) with respect to the input temperature, and the subsequent derivatives. These are required to calculate T_w (Eq. A.22) using Davies-Jones (2008), and cannot be accomplished using QSatMod. We show acceptable differences between the Stull version of wet bulb temperature (T_{wS}) calculated using either QSatMod and QSat_2 (Fig. 1a). **The new subroutines improve CLM4.5 by calculating previously unused thermodynamic quantities.** Additionally, these routines are useful moist thermodynamic routines for other datasets for researchers to use, thus expanding the capacity of the HumanIndexMod.

revised portion of:

4 Results

We present a snap shot of the many metrics calculated. First, we present results of our evaluation the improved moist thermodynamic calculations and the implementation these metrics into CLM4.5 (Fig. 1). Second, we show an example of the possible global applications for these metrics. This approach characterizes heat stress within CLM4.5 in response to one observation reanalysis product, the CRUNCEP.

4.1 Evaluation of improved moist thermodynamic quantities

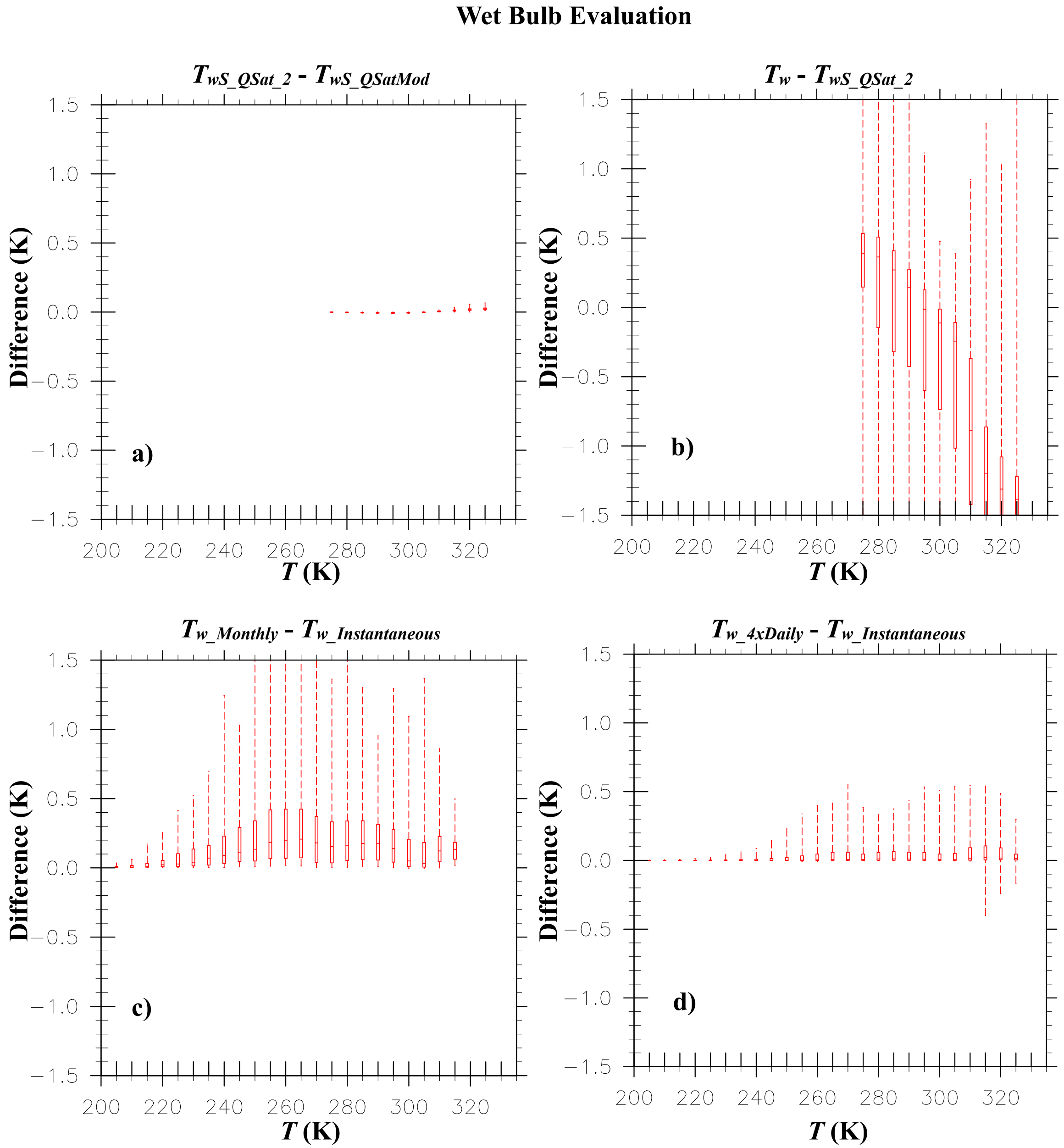
We present a series of box and whisker plots demonstrating the value added of implementing 1) accurate and efficient moist thermodynamic quantities, and 2) online calculation of the heat stress metrics is an improvement over calculating these metrics using monthly or 4x daily model output (Figure 1). **Figure 1a shows the difference in the Stull (2011) wet bulb temperature calculated using the saturated vapor pressure from Davies-Jones (2008) (QSat_2) and Flatau et al. (1992) (QSatMod).** The differences are minimal. However, our point is that the Davies-Jones (2008) method for wet bulb temperature is preferred. **We show the difference between wet bulb temperatures using Stull (2011) calculated with QSat_2, and Davies-Jones (2008) (which requires QSat_2) (Fig. 1b). Differences are greater than 1K between Stull (2011) and Davies-Jones (2008) methods, and they are temperature dependent (Fig. 1b).** Lastly, we **show the difference between calculating Davies-Jones (2008) T_w using monthly and 4x daily averaged model data vs the model instantaneous calculations (Fig. 1c**

and 1d, respectively). Using model averaged data instead of the instantaneous data systematically overestimates T_w by more than 1K for monthly and 0.5K for 4x daily output.

revised portion of Section 5:

5 Discussion

We designed the HumanIndexMod to calculate diagnostic heat stress and moist thermodynamics systematically. There are many approaches to evaluating heat stress. Monthly and seasonal temperature and moisture averages were used for general applications (Dunne et al., 2013), **however these averages overestimate the potential severity of heat stress** (Fig. 1c,d). **Even using daily or sub-daily averages** (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) **potentially overestimates heat stress. This is due to the non-linear covariance of T and Q, and averages miss these extremes.** Ultimately, capturing the diurnal cycle is crucial for quantifying heat stress extremes (Oleson et al., 2013b). Heat stress related illness is exacerbated by high heat stress nights as well as daytime. **To accurately calculate these extremes, one needs either high temporal resolution data, or by directly computing them at each time step within climate models.** We discuss the results from the implementation separately: moist thermodynamics and heat stress.



Wet Bulb Evaluation

$T_{wS\_QSat\_2}$ - $T_{wS\_QSatMod}$

T_w - $T_{wS\_QSat\_2}$

$T_{w\_Monthly}$ - $T_{w\_Instantaneous}$

$T_{w\_4xDaily}$ - $T_{w\_Instantaneous}$

Figure 1. Evaluation of wet bulb temperatures. The boxes represent the 90% confidence interval. The upper and lower tails represent the 100% confidence interval. The horizontal line in each box is the median value. a) difference between T_{wS} using QSat_2 saturated vapor pressure and QSatMod saturated vapor pressure over the valid range for T_{wS}. b) difference between T_w (Davies-Jones, 2008) and T_{wS} (Stull, 2011) (both using QSat_2 saturated vapor pressure calculation) over the valid range for T_{wS}. c) is the difference between using model monthly averaged input fields and model instantaneous fields to calculate monthly T_w. d) difference between using model 4x Daily averaged input fields and model instantaneous fields to calculate 4x Daily T_w. For a), b), and d) the inputs of T, P, and Q are derived from model 4x Daily fields from the years 2001-2010. For c) the inputs of T, P, and Q are derived from model Monthly fields from the years 2001-2010.

Major Comment 2, Part 3: *Or, another interesting question would be how well does the model do at reproducing observed values of these metrics? This could be contrasted with its skill in quantities more likely to be looked at such as peak temperature, daily and monthly means. Do these diagnostics reveal previously hidden problems?*

We drove CLM4.5 with CRUNCEP, an observationally corrected reanalysis product. Our new figures analyze the partitioning of T and Q from this data product. We believe that CLM4.5 station to model validation is outside the scope of this manuscript. We feel that the addition of the HumanIndexMod enables researchers to run model validation experiments of extreme events within CLM4.5.

Major Comment 3: *I think the abstract is too long, has too many details (e.g. names of parts of the model code) and does not adequately explain the motivation and objectives of the study. The Summary section, 6, is a better abstract (in some ways anyway) than what the authors have written for the abstract.*

We have rewritten the abstract. We split the abstract into two components:

1. What we have added to CLM4.5 and why.
2. What is the added value of these metrics within a GCM.

Abstract

We implement and analyse 13 different metrics (4 moist thermodynamic quantities and 9 heat stress metrics) in the Community Land Model (CLM4.5), the land surface component of the Community Earth System Model (CESM). We call these routines the HumanIndexMod. We limit the algorithms of the HumanIndexMod to meteorological inputs of temperature, moisture, and pressure for their calculation. All metrics assume no direct sunlight exposure. **The goal of this project is to implement a common framework for calculating operationally used heat stress metrics, in climate models, offline output, and locally sourced weather datasets, with the intent that the HumanIndexMod may be used with the broadest of applications.** The **thermodynamic quantities use the latest accurate and efficient algorithms available,** which in turn are used as inputs to the heat stress metrics. There are **three advantages** of adding these metrics to CLM4.5 **1) improved moist thermodynamic quantities, 2) quantifying heat stress in every available environment within CLM4.5, and 3) these metrics may be used with human, animal, and industrial applications.**

We demonstrate the capabilities of the HumanIndexMod in a default configuration simulation using CLM4.5. We output 4x daily temporal resolution globally. **We show that the advantage of implementing these routines into CLM4.5 is capturing the nonlinearity of the covariation of temperature and moisture conditions.** For example, **we show that there are systematic biases of 1.5°C between monthly and ±0.5°C between 4x daily offline calculations and the online instantaneous calculation, respectively.** Additionally, **we show**

that the differences between an inaccurate wet bulb calculation and the improved wet bulb calculation are ±1.5°C. These **differences are important due to human responses to heat stress being non-linear.** Furthermore, **we show heat stress has unique regional characteristics. Some metrics have a strong dependency on regionally extreme moisture, while others have a strong dependency on regionally extreme temperature.**

Major Comment 4: *I am discouraged that the authors (apparently) did not incorporate heat-stress metrics that account for wind and radiation variations, since this seems to be one place where they can truly innovate by assessing how overall patterns of heat stress are (or not) altered by taking these into account. Currently every heat stress study has to basically hand-wave that these factors aren't important—this study could make a real contribution by finally tackling the issue, since all the needed inputs are right there in the model.*

We understand the reviewer's sentiment about the current state of global heat stress analysis, and we agree that radiation and thermal transfer are hand-waved. However, there are numerous datasets that do not include terms for radiation in their data. One of the goals of these algorithms is to unify the research, such that all datasets and climate models are using the same equations. This allows model to dataset comparisons, but model to model comparisons are possible, as well. Radiation and physiological variables are under future research and are outside the scope of this manuscript.

We make it clear in the new abstract (see major comment 3) that we do not use radiation. Additionally, we rewrote the entirety of section 2, and added a statement about radiation and winds.

From **Abstract**:
All metrics assume no direct sunlight exposure.

2 Heat stress modeling

2.1 Background

The primary focus of this paper is on atmospheric variable based heat stress metrics that we introduce into the HumanIndexMod. The models for determining heat stress for humans vary greatly; ranging from simple indices to complex prognostic physiology modeling (Table 1). Prognostic thermal models are beyond the scope of this paper, as they require more than atmospheric inputs. **Additionally, metrics that include radiation and wind (with one exception, Apparent Temperature) are also beyond the scope of this paper.** Each index that we chose uses a combination of atmospheric variables: temperature (*T*), humidity (*Q*), and pressure (*P*). **We chose these metrics because they are in operational use globally by industry, governments, and weather services.** Furthermore, **these metrics may be applied to the broadest range of uses: climate and weather forecasting models, archive datasets, and local weather stations.**

Major Comment 5: *There are numerous instances, which will appear in the detailed comments, where the authors have used terminology inaccurately or confusingly.*

We agree that the vocabulary used to describe the metrics and results were misapplied, and we are addressed this issue. Specifically, we are replacing all mentions of 'joint distributions' with 'conditional distributions'.

As an example:
From our new Section 3.4:
Every 6 hour period that exceeds the percentiles were located within the time series, and we calculated the **conditional distributions**.

Major Comment 6: *The paper implies in a very misleading calculation (lower panels of Fig. 1) that there are significant problems with the way saturation vapor pressure is calculated in GCMs, when the errors are tiny and well-documented in the existing literature.*

We have eliminated all instances of this from the manuscript (Sections 2.1-2.2, and Section 3). Additionally, we have removed the original Figure 1. Our new Figure 1a demonstrates that the differences between the two vapor calculations are minimal (see Major Comment 2).

Major Comment 7: *The writing of the paper is quite poor, with badly organized paragraph structure, and many unclear sentences, which makes it hard to read. I have not tried to point all of these out individually but would request that the authors read carefully and try to find ways of rearranging or rewriting text that is unclear or jumps from one topic to another within a paragraph. A good rule of thumb is that everything said in a paragraph should usually relate back to and expand on the opening sentence of the paragraph.*

We thank the reviewer for pointing out the structural problems with the manuscript, and the numerous comments to help us address these issues. We have rewritten the entire manuscript to address all major and minor comments.

MINOR COMMENTS

1. 5198,11 etc. The word "covariance" is misused; this specifically means <uv> averaged over an ensemble of joint measurements {u,v}. I think "covariation" is what they are after.

We have revised our use of covariance. Our analysis technique is to isolate the turbulent events (extremes) within CLM4.5 which is the covariance of T and Q.

from Section 3.4:
An open question is what drives extreme high heat stress events, which are, by definition, rare events. For example, we cannot determine from the mean climate state or from theory, in a warm and humid climate, if abnormally high temperature, abnormally high moisture, or a

combination of the two, caused a heat stress event. **This is a question of the covariance of perturbations of temperature and humidity, not a statement of mean conditions, and there is no theory to explain these situations**. For example, we may apply Reynolds averaging to the NWS Heat Index equation (Eq. 3):

$$\overline{HI} = a + \overline{b\overline{T}} + \overline{c\overline{RH}} + \overline{d\overline{T}\,\overline{RH}} + \overline{e\overline{T}^2} + \overline{f\overline{RH}^2} + \overline{g\overline{T}^2\overline{RH}} + \overline{h\overline{T}\,\overline{RH}^2} + \overline{\imath\overline{T}^2\overline{RH}^2} + \left[\overline{dRH'T'} + \overline{eT'^2} + \overline{fRH'^2} + \overline{gT'^2RH'} + \overline{hT'RH'^2} + \overline{\imath T'^2RH'^2}\right] \quad (13)$$

Also from Section 3.4:
Every 6 hour period that exceeds the percentiles were located within the time series, and we calculated the conditional distributions. For example, the 99th percentile exceedance of HI isolated the top 1606 hottest time steps for each grid cell. After isolating these time steps, we use this distribution as a mask to isolate all other quantities (e.g., temperature and humidity), allowing cross comparison between all metrics and HI. **The goal is to develop an analysis technique comparing all covariances of the metrics within CLM4.5.**

from Section 5:

5 Discussion

We designed the HumanIndexMod to calculate diagnostic heat stress and moist thermodynamics systematically. There are many approaches to evaluating heat stress. Monthly and seasonal temperature and moisture averages were used for general applications (Dunne et al., 2013), however these averages overestimate the potential severity of heat stress (Fig. 1c,d). Even using daily or sub-daily averages (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) potentially overestimates heat stress. **This is due to the non-linear covariance of *T* and *Q*, and averages miss these extremes.** Ultimately, capturing the diurnal cycle is crucial for quantifying heat stress extremes (Oleson et al., 2013b). Heat stress related illness is exacerbated by high heat stress nights as well as daytimes. To accurately calculate these extremes, one needs either high temporal resolution data, or directly computing them at each time step within climate models. We discuss the results from the implementation separately: moist thermodynamics and heat stress.

2. 15-16. This statement is too vague; in principle one can calculate T_W from r and T which are prognostic variables in the model.

We have replaced the language to show the differences between inaccurate and the improved wet bulb temperature calculations. The changes are reflected in major comment 3 (the Abstract).

from the Abstract:
Additionally, we show that the differences between an **inaccurate wet bulb calculation and the improved wet bulb calculation are ±1.5°C**.

from Section 4.1

We show the difference between wet bulb temperatures using Stull (2011) calculated with QSat_2, and Davies-Jones (2008) (which requires QSat_2) (Fig. 1b). **Differences are greater than 1K between Stull (2011) and Davies-Jones (2008) methods, and they are temperature dependent (Fig. 1b)**.

3. 28-end: quite confusing to first assert two regimes (arid and non-arid) and then begin talking about others "strong convection". Also confusing to have "strong convection" and "equatorial" as separate regimes (isn't there strong convection near the equator?)

We have eliminated this phrasing from the manuscript to reduce the length and enhance clarity. We eliminated the original box and whisker plots that used this language (Fig. 4-7), and eliminated Table 5, which also used this language.

4. 5199,2 "Heat death" -> "Heat related conditions" ?

We have replaced this language.

1 Introduction

Heat related conditions are the number one cause of death from natural disaster in the United States; more than tornadoes, flooding, and hurricanes combined (NOAAWatch, 2014).

5. 5200,7. It is implied that these studies do not consider the diurnal cycle but in some cases they did look at daytime high and nighttime values, even if not characterizing the full diurnal cycle.

We have reviewed the 7 citations that are in this sentence, and at best these papers used daily values as the highest temporal resolution.

6. 12: The text should be clearer about what type of "inaccuracies" are being considered here. Are the authors claiming that inaccurate formulas for well-defined quantities such as dewpoint or wet-bulb temperature have been carelessly applied? Or are they referring to heat stress measures such as sWBGT that are widely-used approximations to other measures? It is not clear from the papers listed as offenders here; for example Dunne et al. used a standard WBGT approximation, while Sherwood and Huber 2010 considered only T_W, for which they used the same Davies-Jones formulas used in this paper unless I am missing something (since Huber is an author on both papers this at least should be cleared up). This distinction matters because even if sWBGT departs significantly from WBGT, it is still a well-defined index that is no more or less relevant a priori than most of the other indices (AT, HI etc.). WBGT itself is based on a very approximate physical analog of the human body, and is not a standard of truth.

We realize that this paragraph was too brief in its descriptions of what inaccuracies were being considered. We have added a detailed descriptions of which moist thermodynamic quantities and heat stress metrics in which papers that have been inaccurate or erroneously applied.

From Section 1:

There are a limited number of studies validating, exploring, or using heat stress metrics on a global scale (Kjellstrom et al., 2009b; Hyatt et al., 2010; Sherwood and Huber, 2010; Fischer and Schar, 2010; Fischer et al., 2012; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Algorithms for measuring heat stress and labor capacity are based upon sub-daily rates of exposure to heat stress (Parsons, 2006). Most of these studies do not capture the diurnal cycle of heat stress (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013), thus not representing both night time highs, and daytime extremes. Only one study includes solar radiation as a component in heat stress (Kjellstrom et al., 2013). Different metrics are used between each study, and only one study attempts to compare more than two metrics (Oleson, et al., 2013b).

Various forms of moist thermodynamic calculations (Buck, 1981; Davies-Jones, 2008; Stull, 2011) and heat stress metrics are criticized for their inaccuracies (Budd, 2008; Alfano et al., 2010; Davies-Jones, 2008). **Buck (1981) moist thermodynamic calculations are not as accurate as Bolton (1980), yet are used in a recent study (Kjellstrom et al., 2013). Wet bulb temperature calculations are different between multiple studies (Hyatt et al., 2010; Sherwood and Huber, 2010; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Hyatt et al. (2010) and Kjellstrom et al. (2013) use natural wet bulb temperature—a calculation, that due to non-linear components of its equation, may have multiple steady state solutions (Alfano et al., 2010). Oleson et. al. (2013b) uses a recent formulation of wet bulb temperature that is limited in effective range of input temperatures (Stull, 2011) (we go into further detail on this equation in section 2).** Sherwood and Huber (2010) and Dunne et al. (2013) **both use Davies-Jones (2008) as their source paper for their calculation of wet bulb temperature. However, Sherwood and Huber's (2010) wet bulb temperature calculations use Bolton's (1980) equivalent potential temperature Eq. (38), rather than the more accurate Eq. (39)** (Bolton, 1980; Davies-Jones, 2008; Davies-Jones, 2009). Furthermore, **their calculation is only valid for wet bulb temperatures above 10°C.** Dunne et al. (2013), on **the other hand, uses wet bulb potential temperature Eq. (3.4)** in Davies-Jones (2008), **yet the recommended equations for wet bulb potential temperature are Eq. (3.5-3.7, and 3.8)** (Davies-Jones, 2008) for the temperature ranges used in their paper. Dunne et al. (2013) a**lso uses Bolton's (1980) equivalent potential temperature Eq. (40), rather than the more accurate Eq. (39)** (Bolton, 1980). Additionally**, Dunne et al. (2013) uses a variation of WBGT that is heavily criticized, the indoorWBGT (Budd, 2008).**

Occasionally, results using heat stress limits are misinterpreted. One study confuses wet bulb temperature thresholds with dry bulb temperature thresholds (Benestad, 2011). **This has misleading consequences as their results do not include moisture metrics, yet the author cites Sherwood and Huber (2010)'s wet bulb threshold (35°C) as the threshold value for their temperature analysis.** The **wet bulb temperature at 35°C is a theoretical limit where**

humans would die from heat stress after 6 hours of exposure. Benestad (2011)'s **misapplication implies that most humans should die every year, because a great portion of the world reaches temperatures of 35°C for more than a 6 hour period.**

7. 23-end: The goals of the study are not clear (problem also in Abstract). If the goal is to compare a large suite of metrics, why do we need a land surface model? Why not just use station observations? If the objective is to make a bunch of new metrics available in a GCM then that should be more clearly stated, and the reasons why this is a good thing should be more clearly explained.

We have rewritten this paragraph to make the objectives of this manuscript clear.

from Section 1:

Our goal here is to improve the situation by creating a module that calculates a **large suite of metrics**, using the **most accurate and efficient algorithms available**, that may be **used with as many applications as possible: climate models, offline archive data, model validation studies, and weather station datasets**. We call this module the HumanIndexMod. The module calculates 4 moist thermodynamic quantities and 9 heat stress metrics. These heat stress metrics are in operational use worldwide, and cover a wide range of assumptions.

8. 5201,7: Do you need a citation here? Also, is all the discussion of biomes etc. necessary for interpreting the results of your study? It seems tangential.

We have added the citation for the release date of CLM4.5. We agree that the descriptions of the biomes is tangential to the aims of the paper, and have reduced the paragraph to the most relevant items for this manuscript. Additionally, we have removed this section and combined it with section 3 to reduce the length and eliminate redundancy.

from Section 3

3.1 The structure of Community Land Model version 4.5

We use CLM version 4.5 which was released in June, 2013 (Oleson et al., 2013a). Boundary conditions for CLM4.5 consist of land cover and atmospheric weather conditions. Each grid cell in CLM4.5 can include vegetation, lakes, wetlands, glacier, and urban. There are new parameterizations and models for snow cover, lakes, crops, a new biogeochemical cycles model, and new urban classifications (Oleson et al., 2013a). The urban biome, a single-layer canyon model, is designed to represent the 'heat island', where temperatures are amplified by urban environments (Oleson et al., 2008a,b; Oleson et al., 2010a,b,c). The 'heat island' effect can increase the likelihood of complications from human heat stress (Oleson, 2012).

> *9. 5202,6: Subtitle is strange—these are not water calculations but moist thermodynamic quantities. We need to know how much water is in the air already*

before we can do these calculations. Also, I think it would make more sense to move this discussion together with the current section 3.1 where all is revisited and much is restated.

10. 10-19: This text is confusing two issues. One is the accurate calculation of the saturation humidity, and the other is how to calculate T_W efficiently. Please write more clearly. GCMs may use imperfect approximations to do the former; do they calculate the latter at all? I do not believe there is a case for updating calculations of saturation, unless you can show that the errors are physically significant. People have been well aware of this situation for years and compared to other modeling uncertainties like boundary layer and cloud parameterization, the small errors in approximate formulas for e_s or r_s seem trivial.

11. 20-21. This sentence is confusing because Bolton (1980) did not present a T_W calculation. The relation between T_W and theta_e needs to be explained.

12. 5203,5: by "calculating" do you mean "implementing"?

13. 12: unrealistic compared to what?

For comments 9-13, we have eliminated sections 2.1, 2.2, and 3.1 to reduce the length of the manuscript, reduce redundancy, and clarify our statements. As mentioned previously, we replaced Figure 1 to clarify the statements regarding efficient calculations of T_w.

From section 3.2

Moist thermodynamic water vapor quantities in CLM4.5 are calculated within QSatMod. We use the outputs from QSatMod as the inputs to the HumanIndexMod. Within the HumanIndexMod, we created a subroutine, QSat_2, that has all the same functionalities as QSatMod. This subroutine uses the August-Roche-Magnus (ARM) equation (Eq. A.13), rather than the Flatau et. al. (1992) polynomial equations for vapor pressure in QSatMod. The log derivative of ARM (Eq. A.15) is a critical component of the calculation of T_w, and is not available in QSatMod. Additionally, QSat_2 calculates $f(\theta_E)$ (Eq. A.18) with respect to the input temperature, and the subsequent derivatives. These are required to calculate T_w (Eq. A.22) using Davies-Jones (2008), and cannot be accomplished using QSatMod. We show acceptable differences between the Stull version of wet bulb temperature (T_{wS}) calculated using both QSatMod and QSat_2 (Fig. 1a). The new subroutines improve CLM4.5 by calculating previously unused thermodynamic quantities. Additionally, these routines are useful moist thermodynamic routines for other datasets for researchers to use, thus expanding the capacity of the HumanIndexMod.

We implement specific thermodynamic routines developed by Davies-Jones (2008) to accurately calculate T_w (see Appendix A). Equation (A.4) is the most accurate and efficient θ_E calculation available (Bolton, 1980; Davies-Jones, 2009). Calculating Eq. (A.4) required implementing T_L and θ_{DL} (Eq. A.2 and A.3, respectively) into the HumanIndexMod. T, P, and Q from CLM4.5 are used to calculate θ_E and T_E (Eq. A.5). T_E, a quantity used in a previous heat stress study (Fischer and Knutti, 2012), is an input into QSat_2 for calculating the initial guess of T_w, and subsequently followed by the accelerated Newton-Raphson method (Eq. A.9-A.22). We found it advantageous to split the heat stress quantities into their own subroutines, allowing the user to choose what quantities to be calculated. The minimum requirements to execute the entire module are T (K), P (Pa), RH (%), Q (g/kg), e (Pa), and u_{10m} (m/s). Table 4 shows the subroutines, input requirements, and outputs in HumanIndexMod.

14. 22: Don't some of them use radiation and wind speed? Ok, a few lines later you mention this, but the wording implies that you won't look at the ones that use wind speed. Why not? Isn't that a strength of your approach that you can do this?

We have rewritten the paragraphs to clarify these statements about radiation and wind and why we do not use them.

from the new Section 2

2 Heat stress modeling

2.1 Background

The primary focus of this paper is on atmospheric variable based heat stress metrics that we introduce into the HumanIndexMod. The models for determining heat stress for humans vary greatly; ranging from simple indices to complex prognostic physiology modeling (Table 1). Prognostic thermal models are beyond the scope of this paper, as they require more than atmospheric inputs. **Additionally, metrics that include radiation and wind (with one exception, Apparent Temperature) are also beyond the scope of this paper.** Each index that we chose uses a combination of atmospheric variables: temperature (T), humidity (Q), and pressure (P). **We chose these metrics because they are in operational use globally by industry, governments, and weather services. Furthermore, these metrics may be applied to the broadest range of uses: climate and weather forecasting models, archive datasets, and local weather stations.**

15. 5204, 4: I do not understand how winds can be implicitly included. Either the metric explicitly includes wind speed, or is calculated assuming some average or typical wind speed, or inherently does not depend on wind speed (e.g. T_W). Implicitly including it would mean that it is included indirectly because it correlates with some other variable that is included.

We have removed this language to reduce the length of the manuscript.

16. 5205, eq (1-2). Why e_RH and e_sPa? What does RH stand for? The normal notation is to simply use e for the vapor pressure and e_s or e for the saturation vapor pressure (I see e_s in Table 2, is e_sPA a misprint?). I don't see any other e so there is no need for a subscript except to distinguish actual and saturation.*

We break from the normal notation for the various vapor pressures because we are using the notation in the HumanIndexMod. Additionally, there are units differences between e_s and e_sPa (e_s is in mb and e_sPa is in Pa). We use e_RH because we want the reader to remember that we calculate vapor pressure using relative humidity.

From section 2.2:
where the vapor pressure (e_{RH}) is in Pascals and is calculated from the relative humidity (RH in %), and saturated vapor pressure (e_{sPa}, also in Pascals). **We use this notation because e_s (Table 2) is in millibars. These variable names are the explicit names of the variables in the HumanIndexMod**.

17. eq (3): citation needed (not clear if this polynomial fit is from Steadman but as written it implies not). Since Fahrenheit was a person, the subscript should be F not f (consistent with C for Celcius)

We attached the reference for the Heat Index formula, and also adjust notation to reflect the capitalization.

Heat Index (HI) was developed using a similar process as AT. The United States National Weather Service (NWS) required a heat stress early warning system, and the index was created as a polynomial fit to Steadman's (1979a) comfort model **(Rothfusz, 1990)**.

$$\begin{aligned} HI = {} & -42.379 + 2.04901523T_F + 10.14333127RH + -0.22475541T_F RH \\ & + -6.83783\times10^{-3}T_F^2 + -5.481717\times10^{-2}RH^2 + 1.22874\times10^{-3}T_F^2 RH \\ & + 8.5282\times10^{-4}T_F RH^2 + -1.99\times10^{-6}T_F^2 RH^2 \end{aligned}$$

18. 5206,13: What warning system? Citation needed.

We replaced the 'warning system' with index values, and added the citation.

HUMIDEX is unitless because the authors recognized that the index is a measure of heat load. **The index has a series of thresholds: 30 is some discomfort, 46 is dangerous, and 54 is imminent heat stroke (Masterson and Richardson, 1979).**

19. 17: why do you use the calibration for pigs? Does it make much difference?

We have clarified the sentence.

The Temperature Humidity Index for Comfort (THIC) is a modification of the Temperature Humidity Index (THI) (Ingram, 1965). **Comfort was quantified for livestock through THIC (NWSCR, 1976). We use the original calibration, which is for pigs (Ingram, 1965).** The index is unitless:

where wet bulb temperature (T_w) is in units of Celsius. The index is used to describe behavioral changes in large animals due to discomfort (seeking shade, submerging in mud, etc.). The index is in active use by the livestock industry for local heat stress and future climate considerations (Lucas et al., 2000; Renaudeau et al., 2012). The index describes qualitative threat levels for animals: 75 is alert, 79-83 is dangerous, and 84+ is very dangerous. There are different approaches to the development of THIC, including considerations of physiology of large animals.

> *20. 5207,12: waht is "radiation temperature"? And why wasn't wind incorporated as an input?*
> *21. 13: in what sense are these required? You just said the inputs were T, T_W, and "radiation temperature" so aren't those what would be required to calculate the index value? Do you mean sweat rates etc. were used to develop the index?*
>
> *22. 14: In the introduction you implied that the failure to account for wind and radiation was an important shortcoming of past studies, but now you are saying you aren't interested in those either. More discussion is needed, in the context of the rationale for the study. I think to be honest, if you don't intend to deal with metrics that account for these factors, then in the introduction where you raise this issue you should say up front that this study also will avoid them (currently one will guess the opposite). Finally, this paragraph should state more clearly that the authors are *not* using UTCI. This is a big decision, since the UTCI is arguably the most sophisticated index and was designed to be incorporated into models such as this one (as I understand it).*
>
> *23. 2508,14: This curious statement requires elaboration. You mean there are three different ways to construct a natural wet bulb that yield different temperatures? That a natural wet bulb exhibits hysteresis and doesn't have a unique equilibrium temperature? That there are three different equations for predicting the natural wet bulb temperature and we don't know which one to use? What does it mean for a metric to have "multiple end members"? Confused.*
>
> *24. 5209,7: you mean its accuracy in reproducing WBGT may be questionable. Not just may be, it is guaranteed not to except in particular conditions, since it ignores factors that affect Tg. But it may be OK for diagnosing the effect of a change in T or humidity on human comfort (other things not changing much).*

Minor comments 20-24: we have cut parts of this section to reduce manuscript length and only discuss the metrics we implemented into the HumanIndexMod.

from Section 2.4

2.4 Empirical algorithms

The last category of metrics are derived from first principle thermo-physiology models, or changes in worker productivity, etc., and then reduced by empirical fit. The first metric we present is widely used modification of an industry labor standard, the Simplified Wet Bulb Globe Temperature (sWBGT):

sWBGT was designed for estimating heat stress in sports medicine, adopted by the Australian Bureau of Meteorology, and is acknowledged that its accuracy of representing the original labour industry index may be questionable (ABOM, 2010; ACSM, 1984; ACSM, 1987). We chose, however, to implement sWBGT due to its wide use. sWBGT is unitless, and its threat levels are: 26.7-29.3 is green or be alert, 29.4-31.0 is yellow or caution, 31.1-32.1 is red or potentially dangerous, and ≥32.2 is black or dangerous conditions (US Army, 2003).

25. 5211,10-19: Here the authors give some important information on objectives and rationale, that should have been given much earlier. One thing they should make clear is why they are putting these metrics into the land model, rather than the atmosphere model (I know the answer, but readers who are not familiar with GCM construction may be puzzled).

We have added a part to the introduction (section 1) explaining the goals of implementing the HumanIndexMod into CLM4.5.

from Section 1:
As an example of numerous applications, **we implement the HumanIndexMod into the Community Land Model (CLM4.5)**, a component model of the Community Earth System Model (CESM), maintained by the National Center for Atmospheric Research (NCAR) (Hurrell et al., 2013). **The metrics are directly calculated at the sub-grid scale, capturing heat stress in every environment: urban, lakes, vegetation, and bare ground.** We show examples of the advantages of calculating these metrics at the model time step as compared to lower temporal resolution, and the importance of using accurate moist thermodynamic calculations. We also show that having all metrics calculated at the same time allows for comparison of metrics between each other, and allows for unique analysis of conditional distributions of the inputs. Finally, we show that the metrics may also be used as model diagnostics.

26. 11: you mean the joint distribution of T, P, and Q conditional on high value of metric X?

27. 12: "hottest" means "hottest according to metric X" (not T)?

28. 18: by "of the percentiles" you mean, of the extremes in metric X?

29. 20: "median of the joint distribution" is a non-sequiter, joint distributions do not have medians. What do you actually mean here? Do you mean the median value of X for all points in the top 99%?

> *30. 5215,10: These maps do not present joint distributions - a joint distribution is the probability density of a multidimensional state vector. All that is presented here is a single statistic of the distribution (at each location). We are not told what statistic, so I do not know what these maps actually show. I also don't know what "metric B given metric A" means - does this mean the value of B conditioned on a globally fixed value of A according to a fit (say, multivariate linear or Gaussian) to the sample joint distribution, or a subsampling of all values within some tolerance of A? If so then what value of A is used? There is way too little information here, and what information is provided doesn't make sense.*

Minor comments 26-30: we have rewritten the results section (see major comment 2), renaming the joint distributions to conditional distributions, as well as removing all language to 'a' given 'b', etc.

from Section 3.4:

Every 6 hour period that exceeds the percentiles were located within the time series, and we calculated the **conditional distributions**. For example, the 99th percentile exceedance of HI isolated the top 1606 hottest time steps for each grid cell. After isolating these time steps, we **use this distribution as a mask to isolate all other quantities** (e.g., temperature and humidity), allowing cross comparison between all metrics and HI. The goal is to develop an analysis technique comparing all covariances of the metrics within CLM4.5.

After the **conditional distributions are calculated, we, again, compute the statistical dispersion (mean, variance, exceedance, etc.) of the percentiles**. We display this analysis with maps in two ways. 1) we show the exceedance value of a metric, and 2) we show *T-Q* regime plots of that same metric. We calculate the *T-Q* regimes through expected rank values (Fig. 2). This required a series of steps. 1) We take the **conditional distribution** of *T* and *Q* that represent exceedance percentile of the source heat stress or moist thermodynamic metric. 2) **We take the expected value (median) of the conditional distributions of *T* and *Q* and determine what percentile they come from in their respective time series.** 3) We **condition these values on each other to create the expected rank values (Fig. 2).**

4.2 Exceedance values and regime maps

We **show exceedance and *T-Q* regime maps for the 75th, 95th percentiles of 3 metrics, and 99th percentiles of 6 metrics.** The maps show spatial patterns of heat stress and characteristics. Equatorial and monsoonal regions show moderate levels of heat stress in the 75th percentile (Figure 3a-c). sWBGT shows values exceeding minimum metric warning levels (e.g. China, Northern Africa), whereas HI does not have necessarily the same warning. The 95th percentile shows that moderate levels of heat stress have expanded into higher latitudes (Figure 4a-c). At equatorial and monsoonal regions, heat stress labor reductions should be in effect as it is not safe to work outside, and in some cases (West Africa, the Arabian Peninsula, and the Himalayan Wall), no work at all. At the 99th percentile, severe heat stress is experienced in the monsoonal regions (Figure 5a-c). These maxima correlate with maxima in T_w (Figure 5c).

The *T-Q* regime maps show that partitioning of heat stress into *T* and *Q* begins in regional locations at the 75th percentile (Fig. 3d-f). The partitioning occurs in low latitudes, and is not consistent between metrics. At the 95th percentile, the partitioning expands into higher latitudes, however, many areas (continental interiors) remain equally dependent on *T* and *Q* (Fig. 4d-f). Tw is largely driven by extreme moisture (Fig. 4f) and in some locations (monsoonal Africa, Indian sub-continent, and equatorial South America) very extreme moisture. HI is driven by *T* (Fig. 4e), and sWBGT is mixed between extreme *Q* and extreme *T* (Fig. 4d). All three metrics agree with *T* in the Western United States and Middle East. At the 99th percentile, HI, although dominated by *T* worldwide, shows sign reversals in very small locations (Fig. 5e). Extreme *Q* expands for Tw, and all of the low latitudes experience moisture dependence except for Western United States and Middle East (Fig. 5f). sWBGT has some reversal of *T* to *Q* dominated heat stress (Western Africa). *Q* largely expands worldwide. In all instances, except for HI, high latitudes are equally dependent on *Q* and *T* for heat stress.
Our final maps show SWMP65, SWMP80, and θ_E at the 99th percentile. Maxima for θ_E are spatially the same as T_w (Figure 5c and 6c). Additionally, θ_E partitions towards *Q*, just as T_w (Figure 5f and 6f). Spatial patterns between SWMP65 and HI are similar (Figure 5b and 6a), and their regime maps show similar partitioning toward *T* globally, except for select locations of strong monsoonal locations that show *Q* dependency (Figure 5e and 6d). Lastly, SWMP80 and sWBGT share similar spatial patterns (Figure 5a and 6b). As with the other paired metrics, their *T-Q* regime maps share the same characteristics (Figure 5d and 6e). Low latitudes show strong *Q* dependence, and higher latitudes switch to a *T* dependence.

31. 16: the plural of maximum is maxima

We appreciate this vocabulary definition, and have incorporated its use in our results.

from Section 4.2
These **maxima** correlate with **maxima** in T_w (Figure 5c).

Maxima for θ_E are spatially the same as T_w (Figure 5c and 6c).

32. 5216,19: These four categorizations are not on all fours. Equatorial regions are convective, some arid regions are found in mid-latitudes. I cannot make sense out of this classification.

We have removed Table 5 and removed the box and whisker plots.

*33. Figure captions. The captions of Figures 2-7 each repeat the same unimportant information "1901–2010 CLM4.5 forced by CRUNCEP" - this only needs to be stated in the text (once). But the captions do *not* tell us what is plotted except "joint distribution" (which is incorrect). Please tell us enough so we can figure out how to read the plot. Also, these plots are a bit small and hard to read.*

We have created short, succinct, clear figure captions (see Major Comment 1).

34. Figs. 4-7: These are getting closer to being actual joint distributions but not quite - they are conditional distributions of T for various values of each metric X and "regional association" (try to be consistent between the terminology "regime", "category" and "regional association"). They are not joint distributions because they don't show the distribution of X (except its limits), only the conditional distributions of T given X. Why are there many points on the bottom axis?

We have adopted the conditional distribution vocabulary, see minor comment 30.

35. 5218,4-5: Shouldn't the criterion for saying these are unreasonable be because they disagree with observations (how do we know a priori what T_W should be there)? Why don't you look at some station data or HadCRUH to see what observed humidity is there? Also, does this problem originate in unrealistic CRUNCEP fields or poor behavior of the land model? The first "reason" given for the error does not make sense, since over bare ground any flux at the surface will match that at 2m as long as you are averaging over more than a few minutes (little water vapor or heat can accumulate in 2 meters of air). The second reason should be checked by looking at the distribution of soil moisture values; on its face, it also seems an unlikely explanation since in arid regions there should rarely be rainfall so this problem presumably would not occur very often? And there is no "sand parameterization", there is a soil parameterization that assumes some physical characteristics for sand.

In order to reduce the length of the manuscript, we have removed this section.

> *36. 5219,1: which features?*
>
> *37. 2-3: This paper has not shown that implementing their metrics reduces uncertainty in anything, let alone justifying such a sweeping statement.*

Minor comments 36 and 37: we have rewritten this paragraph.

The **spatial distributions of high heat stress** are **robust between CLM model versions** (Oleson et al., 2011; Fischer et al., 2012; Oleson et al., 2013b). Due to the conservation of energy and entropy, calculating moist thermodynamic variables shows that climate models and reanalysis fall along constant lines of T_E (Eq. A.5), even out to the 99th percentile of daily values (Fischer and Knutti, 2012). **The spread between models is small as compared to the spread in T, thus using heat stress metrics in Earth system modeling may reduce the uncertainties of climate change** (Fischer and Knutti, 2012).

38. 4-6: I don't understand what this sentence means, and don't recall T_E being defined.

We apologize for leaving out that T_E is calculated in the Appendix, and will add this to the discussion, and implementation sections.

from Section 3.2
We implement specific thermodynamic routines developed by Davies-Jones (2008) to accurately calculate T_w (see Appendix A). Equation (A.4) is the most accurate and efficient θ_E calculation available (Bolton, 1980; Davies-Jones, 2009). Calculating Eq. (A.4) required implementing T_L and θ_{DL} (Eq. A.2 and A.3, respectively) into the HumanIndexMod. **T, P, and Q from CLM4.5 are used to calculate θ_E and T_E (Eq. A.5). T_E, a quantity used in a previous heat stress study** (Fischer and Knutti, 2012), is an input into QSat_2 for calculating the initial guess of T_w, and subsequently followed by the accelerated Newton-Raphson method (Eq. A.9-A.22). We found it advantageous to split the heat stress quantities into their own subroutines, allowing the user to choose what quantities to be calculated. The minimum requirements to execute the entire module are T (K), P (Pa), RH (%), Q (g/kg), e (Pa), and u_{10m} (m/s). Table 4 shows the subroutines, input requirements, and outputs in HumanIndexMod.

from Section 5.1
The spatial distributions of high heat stress are robust between CLM model versions (Oleson et al., 2011; Fischer et al., 2012; Oleson et al., 2013b). Due to the conservation of energy and entropy, calculating moist thermodynamic variables shows that climate models and reanalysis fall along **constant lines of T_E (Eq. A.5)**, even out to the 99th percentile of daily values (Fischer and Knutti, 2012). The spread between models is small as compared to the spread in T, thus using heat stress metrics in Earth system modeling may reduce the uncertainties of climate change (Fischer and Knutti, 2012).

39. 18-22: these statements require qualification if they apply to the top 1% of events. Indices that are similar at this extreme mibht be different at lower temperatures.

We have replaced the figures and expanded the exceedance to include the 75th and 95th (Figure 3 and 4) (See Major Comment 2). Thus we have also rewritten the discussion.

40. 22-30: the paper swerves into reopening the discussion of assumptions in the metrics. That should all be done earlier, unless these assumptions are key to interpreting the results or future uses of the software package presented (if so that isn't clear at all from what is written).

In order to reduce manuscript length, we have removed all discussions of development of metrics in section 2. However, in the 5.2 section, we have rewritten the statements about calibrations/assumptions of heat stress metrics to clarify this.

The assumptions/calibrations that derived the heat stress metrics in the HumanIndexMod are another avenue of research that may be explored using a global model. For example, the original equation that sWBGT was derived from was calibrated using US Marine Corps Marines during basic training (Minard et al., 1957), who are in top physical condition. HI was calibrated for an ‘average’ American male (Steadman, 1979a; Rothfusz, 1990). Calculating these heat stress metrics, and the many others in the HumanIndexMod, at every time step within climate models were previously intractable due to insufficient data storage capabilities for high temporal resolution variables. We show that SWMP65 and SWMP80 diverge in their values (Fig. 6a,b and 6d,e). Yet, SWMP80 and sWBGT are similar in spatial patterns and regimes, while HI and SWMP65 have similar patterns and regimes. What links SWMP65 and SWMP80 together is T_w. Swamp coolers are evaporators, and as their efficiency approaches 100%, their solutions approach T_w. Figures 5 and 6 are similar to a circuit resistor, or stomatal resistance (Oke, 1987), which is measure of efficiency. The ‘average’ person (HI) may be acting as a stronger resistor to evaporation than one that is acclimatized (sWBGT). The HumanIndexMod may explore the effects of acclimatization, and its impact on efficiency of evaporative cooling through climate modeling. This type of research may ultimately reduce the number metrics required for computing heat stress.

41. 5220,1-8: A limitation of the approach the authors use is that many heat stress impacts seem to depend on multi-day exposure duration. Come to think of it, is the authors’ plan to write out all their indices multiple times per day? If so then they don’t seem to offer any advantage over just writing out T, r, and p and calculating from output. If not then what summary quantities would one output? Peak heat stress during the month, for monthly mean output, for example? Or number of days above a few heat-stress thresholds? There are more issues that must be confronted before these metrics will have practical value, it seems.

Monthly averages are the default setting for total, rural, and urban environments (13 metrics, 39 fields total). We agree that some sort of average monthly min, max, and mean would be useful, however, users may not want these fields as defaults, because that would add 117 fields altogether. The current setup in CLM4.5 allows the user to change the frequency of output through a name list adjustment.

from Section 6

Furthermore, we have implemented the HumanIndexMod into the latest public release version of CLM4.5. **Archival is flexible**, as the **user may choose to turn on high frequency output**, and the default is monthly averages. **Additionally, monthly urban and rural output of the metrics is default**. We show that the module may be used to explore new avenues of research: characterization of human heat stress, model diagnostics, and intercomparisons of heat stress metrics. Our results show that there are two regimes of heat stress, extreme moisture and extreme temperature, yet all of the most extreme heat stress events are tied to maximum moisture.

42. The Appendix, as far as I can tell, just goes through the contents of the Davies-Jones (2008) paper. What is the point of this? Why not just cite Davies-Jones. If there are key formulas that need to be invoked in the text, put those where they are invoked.

We agree that we should have stated why the Davies-Jones equations are listed in the Appendix. We list these equations because we want every single equation that is used in the HumanIndexMod to also be available in this manuscript. As mentioned in minor point 6, different equations for the moist thermodynamic quantities and heat stress metrics are common in this field of science. From our new Figure 1 (see major point 2), we show that these different calculations may impact their results. We have modified the first paragraph of the Appendix to inform the reader why we have the Appendix.

Davies-Jones (2008) shows multiple methods of computing T_w, and we implemented the most accurate equations, described below. We introduce terminology to describe the Davies-Jones (2008) calculation. All temperature subscripts that are capitalized are in Kelvin, while lower case are in Celsius. κ_d is the Poisson constant for dry air (0.2854), and λ is the inverse (3.504). Many of the following equations are scaled using non-dimensional pressure (also known as the Exner function), π:

43. Table 3. This caption needs more information - what does "modern" and "future" mean? Are you just describing past work here or does this refer to your calculations? Some of those studies may have used more than what is listed (e.g. Sherwood and Huber used reanalysis data not just CCSM3).

We expanded the description of table 3 to be more detailed, and thank the reviewer for catching the error on Sherwood and Huber (2010).

Table 3. List of previous heat stress studies. Studies using datasets, reanalysis, and/or model output that range from ~1900 until ~2010 are labeled 'Modern' and from ~2005 to ~2100 are labeled Future. Some studies do not analyze heat stress quantitatively (Assessment).

44. Figure 1. The variable "q" should I guess be "Q", the specific humidity? The plot for q is un- useful and indeed misleading because it implies a very large error in q which is not true. What is actually happening is that you the computation is being done at fixed total pressure p, but as e approaches p this becomes impossible and implies a vanishing (and then negative) dry air pressure. This is not sensible. If the calculation is done at fixed dry air pressure (more sensible since this is what would actually happen with a fixed mass of dry air and g), the curve for q will look similar to that for e. I recommend deleting the figure entirely and dropping all claims or innuendo in the paper about the inaccuracy of saturation algorithms—you are beating a dead horse, these small errors are already documented in the literature, and there is no way that errors of no more than 2% that don't begin to appear until temperatures are 30C higher than any on Earth today are of any significance.

We have eliminated Figure 1, and replaced it with a new Figure 1 (see Major comment 2).

Response to Reviewer 2

The manuscript by Buzan et al. documents the implementation of an online calculation of heat stress indices in one of the leading Earth System Models. ***This effort is very valuable and potentially useful for a large scientific community running this climate model or analysing its output****. As far as I can judge without having access to the source code,* ***the approach used and its implementation are sound and the set of indices is comprehensive and justified****.*

We thank the reviewer for their response to our manuscript. We agree that this is potentially a useful set of tools, and modifications to GCMs, for the larger scientific community.

However, the reviewer said that:

> *However, in my opinion the manuscript is not ready for publication. The manuscript is* ***poorly organized, in places lengthy, and generally hard to read****. There is a* ***high level of detail in the first part*** *including the* ***discussion of indices that are not even implemented****, or* ***lengthy discussions on aspects that are not directly relevant for the manuscript****, whereas in the* ***second part it is very hard to find all the relevant information documenting the figures and results****. The* ***first half of the manuscript*** *is fine but* ***should be substantially shortened*** *to the essential information documenting the indices and their implementation, whereas the* ***second part requires major revisions*** *including* ***revisiting the selection of results presented.***
> ***In my opinion the complete analysis of the model experiment output needs to revisited, the corresponding sections need to be rewritten and the figures need to be redone with the goal of demonstrating the added value of the great effort done in this project.***

As described further below, we have addressed the reviewer's concerns for our manuscript.
We have:

- Rewritten the document, enhancing clarity and organization.
- Removed ~6 pages; eliminating lengthy discussions and metrics not implemented.
- Redesigned all figures, revising the selection of results.
- Demonstrated added value of the metrics within CLM4.5.

We believe we have addressed all issues raised by the reviewer. Below are detailed responses to each comment by the reviewer. We have reordered the general questions the reviewer asked, linking similar topics together.

Major Comment 1: *The authors argue that "The three advantages of adding these metrics to CLM4.5 are (1) improved thermodynamic calculations within climate models, (2) quantifying human heat stress, and (3) that these 20 metrics may be applied to other animals as well as* *industrial applications.” Given the results presented* ***I am not yet convinced that (2) and (3) would require an online calculation of the indices…***

…I do not see much evidence that (1) makes much of a difference for the temperature range relevant in the troposphere.

How do the results in the online calculation differ from the indices calculated based on the model output? To what extent does a calculation based on daily average output miss the extreme heat stress values?

The redesign of our figures shows the importance of online calculation of the indices, and that approximations of moist thermodynamic quantities may have a large impact on the results.
Our new Figure 1 is an assessment of:

1. As the reviewer stated, the vapor pressure calculations have minimal differences between new and legacy calculations of wet bulb temperatures over the effective environmental range of CLM4.5 (Fig. 1a).
2. The importance of accurate moist thermodynamic calculations (Fig. 1b).
3. The large systematic differences of online versus offline computation of these metrics at low and high frequency time fields (Fig. 1c,d).

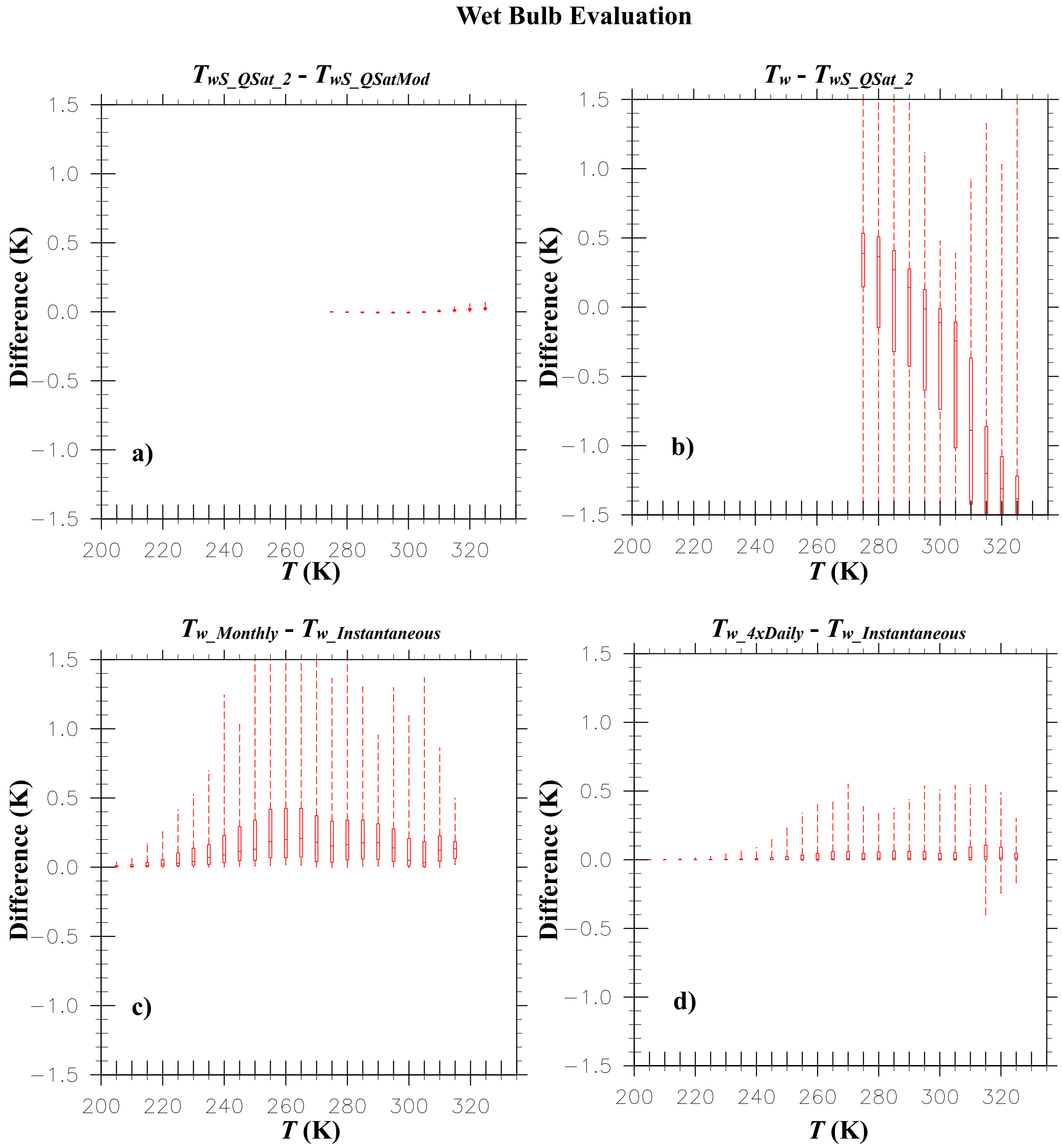


Figure 1. Evaluation of wet bulb temperatures. The boxes represent the 90% confidence interval. The upper and lower tails represent the 100% confidence interval. The horizontal line in each box is the median value. a) difference between T_{wS} using QSat_2 saturated vapor pressure and QSatMod saturated vapor pressure over the valid range for T_{wS}. b) difference between T_w (Davies-Jones, 2008) and T_{wS} (Stull, 2011) (both using QSat_2 saturated vapor pressure calculation) over the valid range for T_{wS}. c) is the difference between using model monthly averaged input fields and model instantaneous fields to calculate monthly T_w. d) difference between using model 4x Daily averaged input fields and model instantaneous fields to calculate 4x Daily T_w. For a), b), and d) the inputs of T, P, and Q are derived from model 4x Daily fields from the years 2001-2010. For c) the inputs of T, P, and Q are derived from model Monthly fields from the years 2001-2010.

We revised our results section to reflect the new figure:

4.1 Evaluation of improved moist thermodynamic quantities

We present a series of box and whisker plots demonstrating the value added of implementing 1) accurate and efficient moist thermodynamic quantities, and 2) online calculation of the heat stress metrics is an improvement over calculating these metrics using monthly or 4x daily model output (Fig. 1). Figure 1a shows the difference in the Stull (2011) wet bulb temperature calculated using the saturated vapor pressure from Davies-Jones (2008) (QSat_2) and Flatau et al. (1992) (QSatMod). The **differences are minimal**. However, our point is that the **Davies-Jones (2008) method for wet bulb temperature is preferred**. We show the difference between wet bulb temperatures using Stull (2011) calculated with QSat_2, and Davies-Jones (2008) (which requires QSat_2) (Fig. 1b). **Differences are greater than 1K between Stull (2011) and Davies-Jones (2008) methods, and they are temperature dependent (Fig. 1b).** Lastly, we show the **difference between calculating** Davies-Jones (2008) T_w **using monthly and 4x daily averaged model data vs the model instantaneous calculations** (Fig. 1c and 1d, respectively). **Using model averaged data** instead of the instantaneous data **systematically overestimates T_w by more than 1K for monthly and 0.5K for 4x daily output**.

We revised our discussion section to reflect the new figure:

5 Discussion

We designed the HumanIndexMod to calculate diagnostic heat stress and moist thermodynamics systematically. There are many approaches to evaluating heat stress. **Monthly and seasonal temperature and moisture averages were used for general applications** (Dunne et al., 2013), **however these averages overestimate the potential severity of heat stress (Fig. 1c,d). Even using daily or sub-daily averages** (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) **potentially overestimates heat stress. This is due to the non-linear covariance of T and Q, and averages miss these extremes.** Ultimately, capturing the diurnal cycle is crucial for quantifying heat stress extremes (Oleson et al., 2013b). Heat stress related illness is exacerbated by high heat stress nights as well as daytime. To accurately calculate these extremes, one needs either high temporal resolution data, or by directly computing them at each time step within climate models. We discuss the results from the implementation separately: moist thermodynamics and heat stress.

Major Comment 2: *Based on a first application of the new code the authors conclude that some indices are more sensitive to temperatures and others more to humidity, and that "arid regions consistently have higher temperatures and lower humidities than the non-arid areas."* ***I am afraid I do not see why an online calculation would be needed to draw these pretty obvious conclusions.***

Instead of presenting these obvious findings, *I would expect the results and discussion section* ***to demonstrate the added value of the implemented online calculation of heat stress indices…*** *Thus, there is* ***no need for ground breaking research and highly innovative new findings*** *but at least I would expect the results and discussion section* ***to demonstrate the accuracy and relevance of the online calculation of heat stress indices over a post processing of the daily output…***

In this regard, I wonder ***why the authors decided to output the 6h average values rather than the maximum and minimum values at any time step in the 6 hour or daily interval.***

We believe that we have confused the reviewer, and our revised manuscript rectifies this issue.

Our analysis technique, and our revised results, demonstrate that:

1. Our results are derived from extreme events.
2. There is no intuitive or theory to analytically derive these results because they arise from instantaneous covariances.
3. A GCM is required for capturing non-linear covariances.
4. To calculate the covariances, we need the heat stress events to correspond to the T and Q combinations, thus output is 6 hour averages.
5. The results are novel.

We have rewritten portions of the methods, section 3.4:

An open question is what drives extreme high heat stress events, which are, by definition, **rare events**. For example, **we cannot determine from the mean climate state or from theory**, **in a warm and humid climate, if abnormally high temperature, abnormally high moisture, or a combination of the two, caused a heat stress event**. This is a question of the **covariance of perturbations of temperature and humidity**, not a statement of mean conditions, and there is **no theory to explain these situations**. For example, we may apply Reynolds averaging to the NWS Heat Index equation (Eq. 3):

$$\overline{HI} = a + \overline{b\,\overline{\overline{T}}} + \overline{c\,\overline{\overline{RH}}} + \overline{d\,\overline{\overline{T}}\,\overline{\overline{RH}}} + \overline{e\,\overline{\overline{T^2}}} + \overline{f\,\overline{\overline{RH^2}}} + \overline{g\,\overline{\overline{T^2}}\,\overline{\overline{RH}}} + \overline{h\,\overline{\overline{T}}\,\overline{\overline{RH^2}}} + \overline{\imath\,\overline{\overline{T^2}}\,\overline{\overline{RH^2}}} +$$

$$\left[\overline{dRH'T'} + \overline{eT'^2} + \overline{fRH'^2} + \overline{gT'^2RH'} + \overline{hT'RH'^2} + \overline{\imath T'^2RH'^2}\right] \qquad (13)$$

where a, b, c, d, e, f, g, h, and i are constants in the polynomial. *RH* and *T* are relative humidity and temperature, respectively. **We are not concerned with the terms outside the brackets, as they are the means**. The **terms within the bracket** are representative of **turbulent effects** on the Heat Index, which we are discussing. It is these **turbulent states where a GCM is able to determine these individual factors**, by **calculating** the heat stress metrics and thermodynamic quantities at **every model time step**. Furthermore, each heat stress metric has different assumptions (such as body size, or physical fitness, etc.) that weigh temperature and humidity differently. A high heat stress event indicated by one metric does not necessarily transfer onto another metric.

To present our results in a clearer fashion, we condensed our 2 extreme *T* and extreme *Q* maps per heat stress metric, into 1 map. To do this, we created an expected rank system, and mapped *T*-*Q* regimes:

From section 3.4:
After the conditional distributions are calculated, we, again, compute the statistical dispersion (mean, variance, exceedance, etc.) of the percentiles. We display this analysis with maps in two ways. 1) we show the exceedance value of a metric, and 2) we show *T*-*Q* regime plots of that same metric. We calculate the *T*-*Q* regimes through expected rank values (Fig. 2). This required a series of steps. 1) We take the conditional distribution of *T* and *Q* that represent exceedance percentile of the source heat stress or moist thermodynamic metric. 2) We take the expected value (median) of the conditional distributions of *T* and *Q* and determine what percentile they come from in their respective time series. 3) We condition these values on each other to create the expected rank values (Fig. 2).

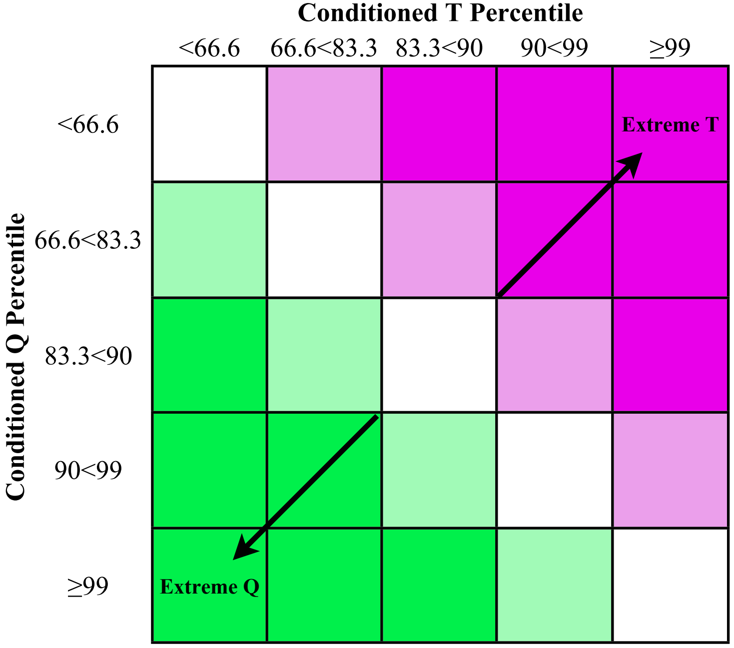


Figure 2. Expected value rank. *T* and *Q* conditioned upon exceedance value of a heat stress or moist thermodynamic metric. The *T* and *Q* values are compared to their respective time series as a percentile. These *T* and *Q* percentiles are binned and are compared to each other. **Extreme *Q* are green and extreme *T* are magentas.**

Our novel analysis technique shows that extremes in *T* and *Q* partition at low latitudes at the 75th percentile heat stress events, and expand to higher latitudes at 95th and 99th percentile heat stress events. Additionally, our novel technique shows that heat stress metrics vary in their response to extremes in *T* and *Q*. These metrics may have opposite signs when compared to other metrics, and opposite signs within their distributions:

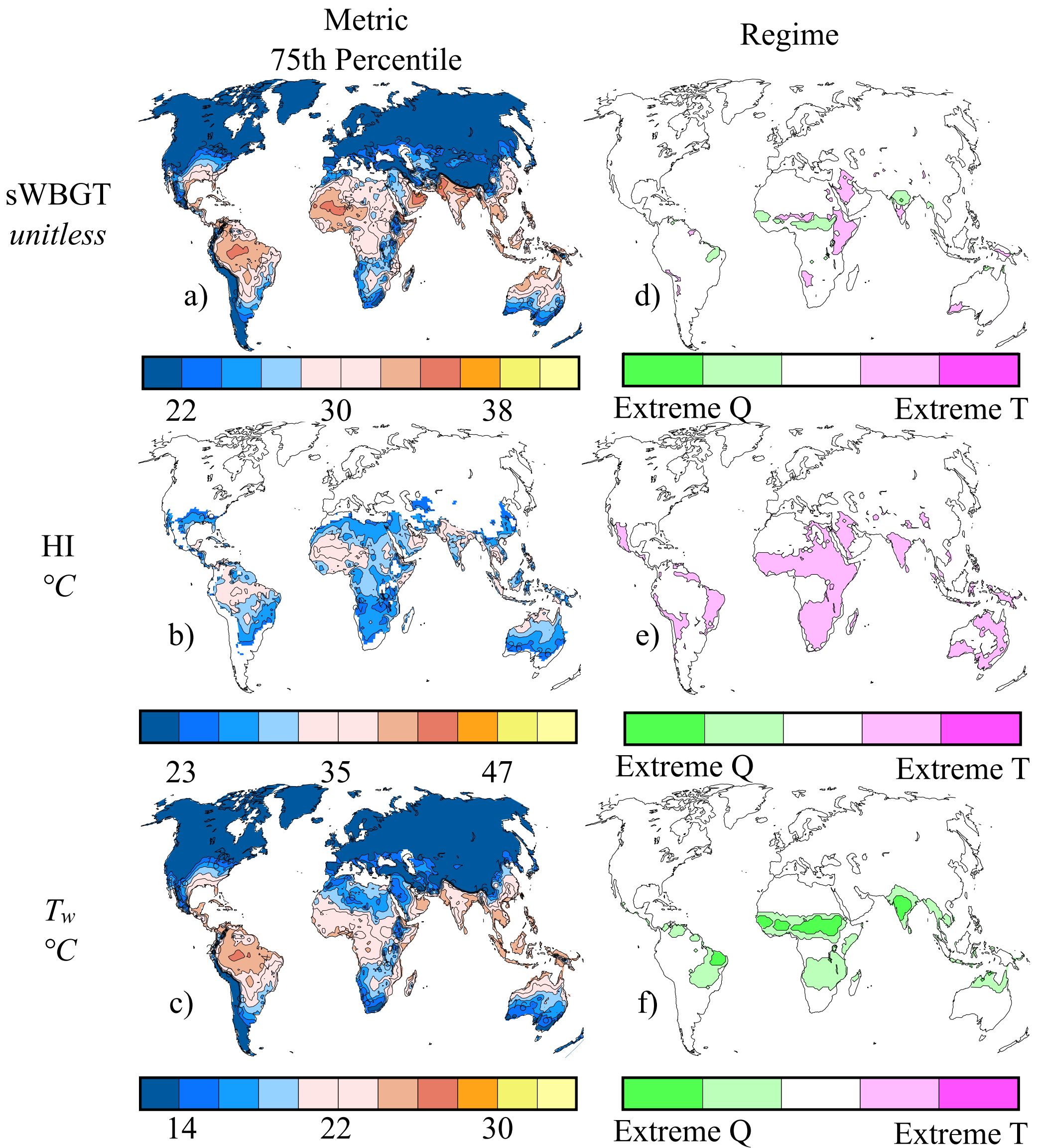


Figure 3. 75th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

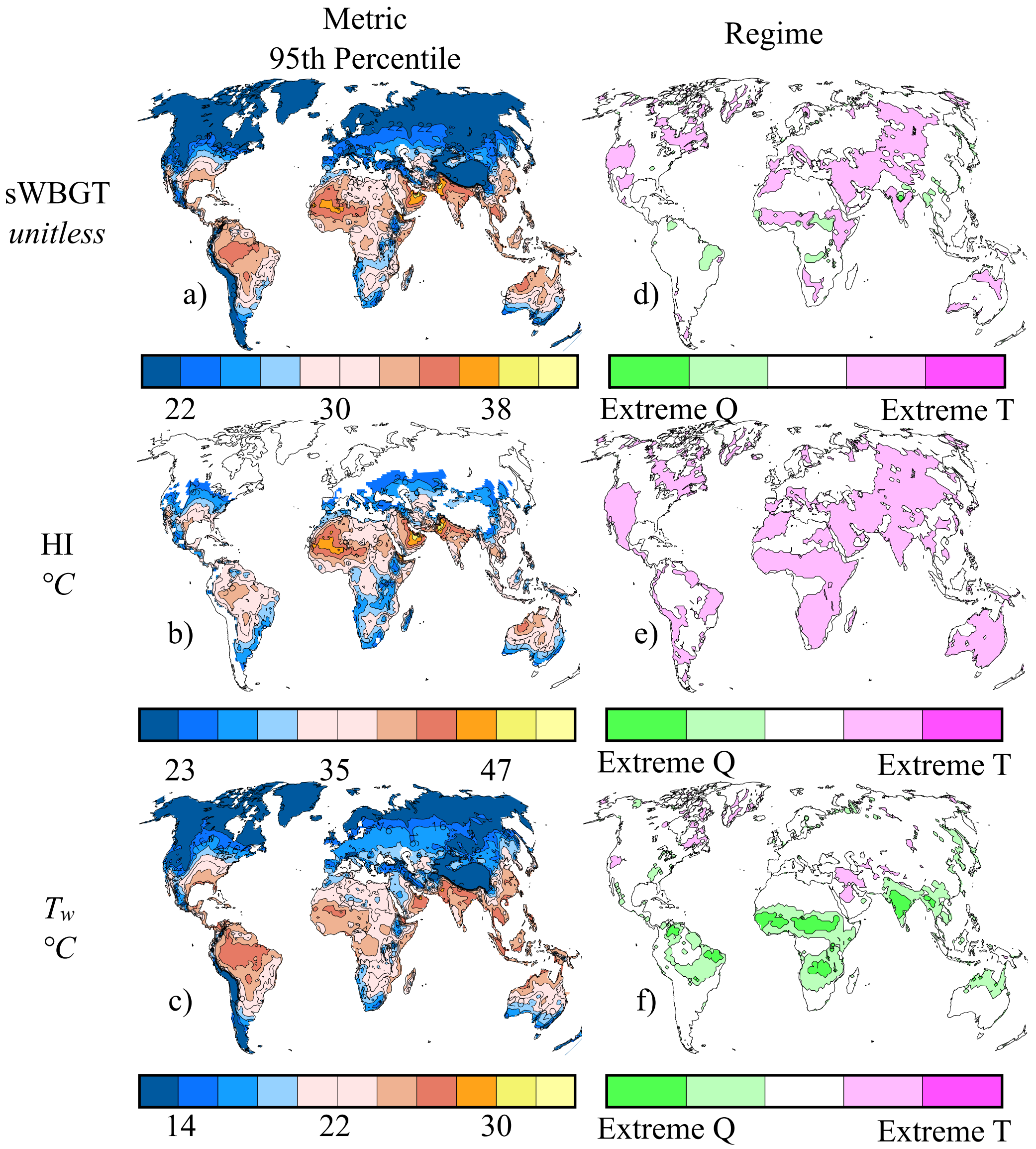


Figure 4. 95th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

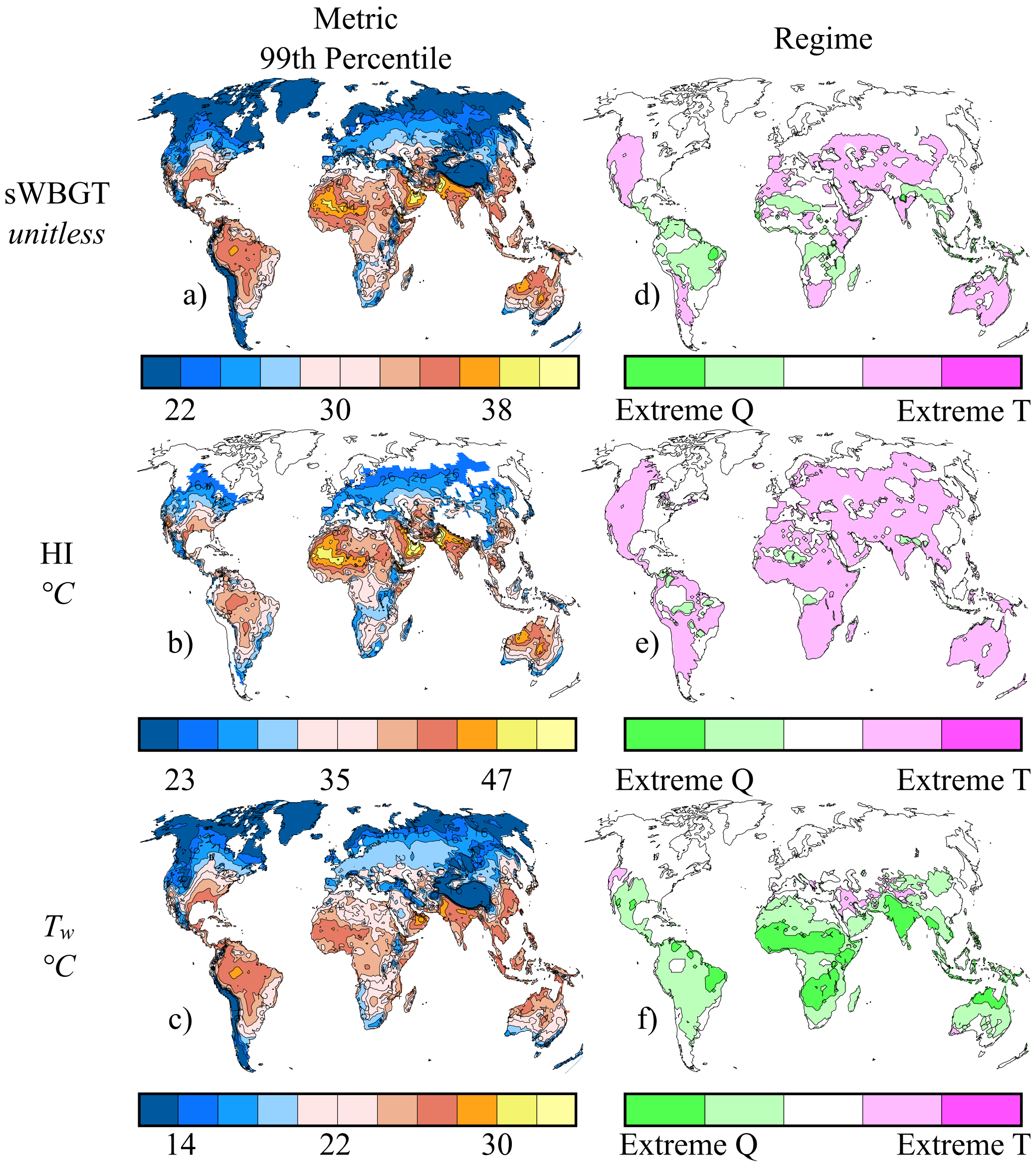


Figure 5. 99th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

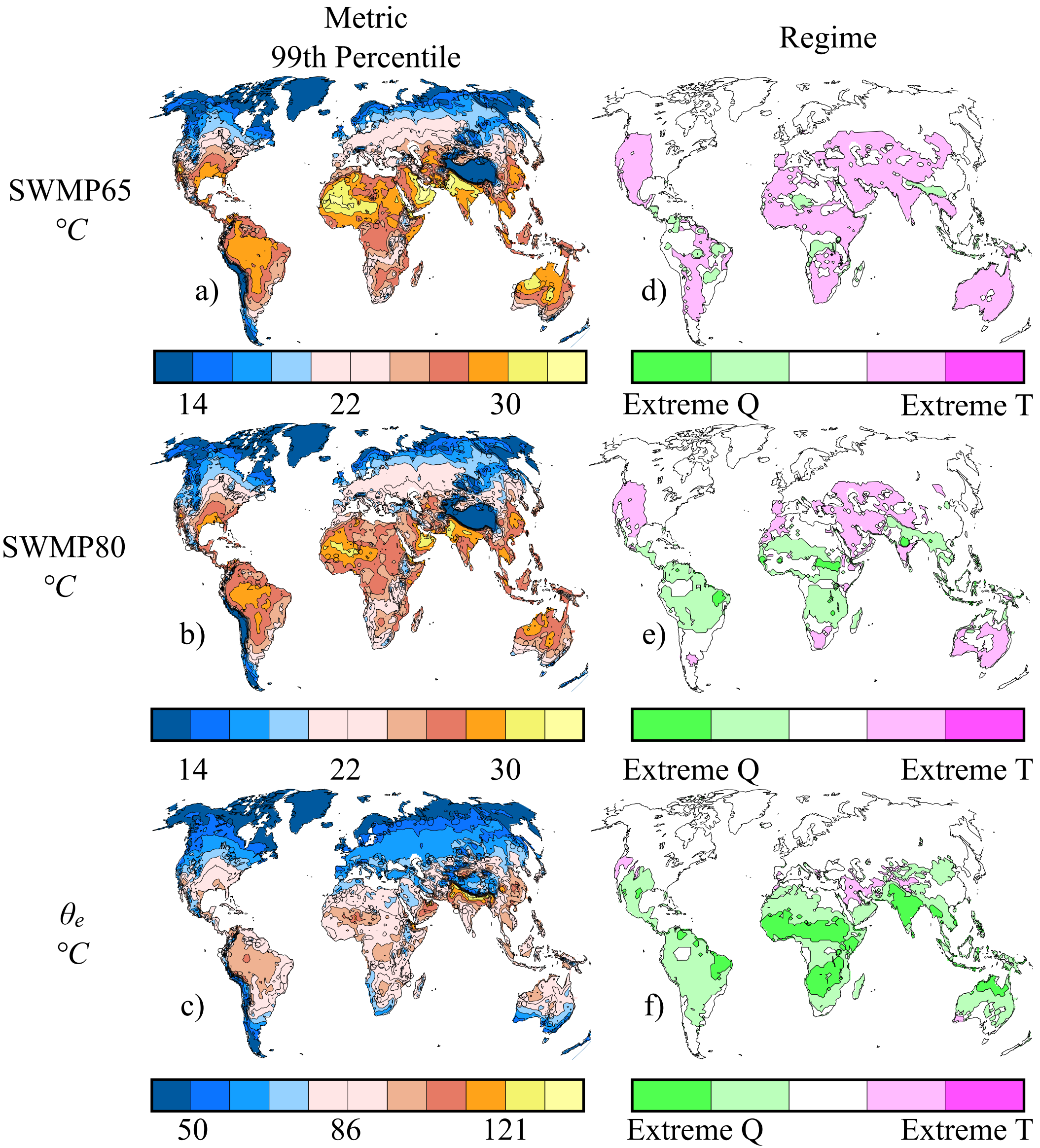


Figure 6. 99th percentile exceedance value of 3 metrics for a) SWMP65, b) SWMP80, and c) θ_E (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Fig. 2.

We have rewritten our results sections (section 4) to reflect the revised figures:

The *T-Q* regime maps show that partitioning of heat stress into *T* and *Q* begins in regional locations at the 75th percentile (Fig. 3d-f). The partitioning occurs in low latitudes, and is not consistent between metrics. At the 95th percentile, the partitioning expands into higher latitudes, however, many areas (continental interiors) remain equally dependent on *T* and *Q* (Fig. 4d-f). Tw is largely driven by extreme moisture (Fig. 4f) and in some locations (monsoonal Africa, Indian sub-continent, and equatorial South America) very extreme moisture. HI is driven by *T* (Fig. 4e), and sWBGT is mixed between extreme *Q* and extreme *T* (Fig. 4d). All three metrics agree with *T* in the Western United States and Middle East. At the 99th percentile, HI, although dominated by *T* worldwide, shows sign reversals in very small locations (Fig. 5e). Extreme *Q* expands for Tw, and all of the low latitudes experience moisture dependence except for Western United States and Middle East (Fig. 5f). sWBGT has some reversal of *T* to *Q* dominated heat stress (Western Africa). *Q* largely expands worldwide. In all instances, except for HI, high latitudes are equally dependent on *Q* and *T* for heat stress.

Our final maps show SWMP65, SWMP80, and θ_E at the 99th percentile. Maxima for θ_E are spatially the same as T_w (Fig. 5c and 6c). Additionally, θ_E partitions towards *Q*, just as T_w (Fig. 5f and 6f). Spatial patterns between SWMP65 and HI are similar (Fig. 5b and 6a), and their regime maps show similar partitioning toward *T* globally, except for select locations of strong monsoonal locations that show *Q* dependency (Fig. 5e and 6d). Lastly, SWMP80 and sWBGT share similar spatial patterns (Fig. 5a and 6b). As with the other paired metrics, their *T*-*Q* regime maps share the same characteristics (Fig. 5d and 6e). Low latitudes show strong *Q* dependence, and higher latitudes switch to a *T* dependence.

The novel results are reflected within our rewritten discussion section:

5 Discussion

We designed the HumanIndexMod to calculate diagnostic heat stress and moist thermodynamics systematically. There are many approaches to evaluating heat stress. **Monthly and seasonal temperature** and moisture averages were used for general applications (Dunne et al., 2013), **however these averages overestimate the potential severity of heat stress (Fig. 1c,d)**. Even **using daily or sub-daily averages** (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) **potentially overestimates heat stress. This is due to the non-linear covariance of *T* and *Q*, and averages miss these extremes**. **Ultimately, capturing the diurnal cycle is crucial for quantifying heat stress extremes** (Oleson et al., 2013b). Heat stress related illness is exacerbated by high heat stress nights as well as daytime. **To accurately calculate these extremes, one needs either high temporal resolution data, or by directly computing them at each time step within climate models.** We discuss the results from the implementation separately: moist thermodynamics and heat stress.

5.2 Heat stress

We show that there are **two regimes of heat stress globally in agreement between metrics** in the CRUNCEP CLM4.5 simulation, *T* (Western United States and Middle East) and *Q* (monsoonal regions). Western United States and Middle East regions consistently have higher temperatures and lower humidities than the monsoonal areas. However, **we show that maximum heat stress is partitioned between *T* and *Q* globally**. **Characterizing arid regions versus non-arid regions may require different heat stress metrics** (e.g. Oleson el al., 2013b, specifically the comparison between Phoenix and Houston). The **HumanIndexMod provides this capability**.

The assumptions/calibrations that derived the heat stress metrics in the HumanIndexMod are another avenue of research that may be explored using a global model. For example, the original equation that sWBGT was derived from was calibrated using US Marine Corps Marines during basic training (Minard et al., 1957), who are in top physical condition. HI was calibrated for an 'average' American male (Steadman, 1979a; Rothfusz, 1990). Calculating these heat stress metrics, and the many others in the HumanIndexMod, at every time step within climate models were previously intractable due to insufficient data storage capabilities for high temporal resolution variables. **We show that SWMP65 and SWMP80 diverge in their values** (Fig. 6a,b and 6d,e). **Yet, SWMP80 and sWBGT are similar in spatial patterns and regimes**, while **HI and SWMP65 have similar patterns and regimes**. What links SWMP65 and SWMP80 together is T_w. Swamp coolers are evaporators, and as their efficiency approaches 100%, their solutions approach T_w. **Figures 5 and 6 are similar to a circuit resistor, or stomatal resistance (Oke, 1987), which is measure of efficiency.** The 'average' person (HI) may be acting as a stronger resistor to evaporation than one that is acclimatized (sWBGT). The HumanIndexMod may explore the effects of acclimatization, and its impact on efficiency of evaporative cooling through climate modeling. **This type of research may ultimately reduce the number metrics required for computing heat stress.**

Major Comment 3: *Furthermore, it would be desirable to evaluate the model's performance at some places with in-situ measurements or gridded data in order to see whether potential model biases in temperature and humidity add up or cancel out.*

CLM4.5, in this study, is driven by CRUNCEP, an observationally corrected reanalysis product. Our new Figures 3-6 show partitioning of T and Q in the CLM4.5 simulation. We believe that validation of CLM4.5 with station datasets is outside the scope of this manuscript. The HumanIndexMod, however, enables researchers to carryout validation experiments of extreme events. We feel that the addition of these tools opens new avenues of research. One of the goals of the project is to enable as many different disciplines as possible to conduct their research. We have clarified this in our abstract, and throughout the manuscript.

From the Abstract:
The goal of this project is to implement a common framework for calculating operationally used heat stress metrics, in climate models, offline output, and locally sourced weather datasets, with the intent that the HumanIndexMod may be used with the broadest of applications.

From the Section 1:
Our goal here is to improve the situation by creating a module that calculates a large suite of metrics, using the most accurate and efficient algorithms available, that may be used with as many applications as possible: climate models, offline archive data, model validation studies, and weather station datasets.

From the Section 2.1:
We chose these metrics because they are in operational use globally by industry, governments, and weather services. Furthermore, these metrics may be applied to the broadest range of uses: climate and weather forecasting models, archive datasets, and local weather stations.

Minor Comments:

1. The author emphasize the implementation of more accurate moist thermodynamic calculations. However, based on the results shown here I am not yet convinced that this is particularly relevant. At least Fig.1 suggest that there is hardly any difference in the typical range of tropospheric temperatures. Please quantify the effect for the application here.

We agree with the reviewer that Figure 1 does not present the advantages of the new calculations. We replaced Figure 1 with a new Figure 1 as seen in Major Comment 1. We are removing the old Figure 1, because the figure was meant to demonstrate that the differences between water vapor calculations are minor, which our new Figure 1a demonstrates. However, the improved water vapor thermodynamic calculations for T_w, T_E, and θ_E, are a substantial improvement to CLM4.5. There are no T_E and θ_E calculations in CLM4.5, and we show with our new Figure 1b that the accurate T_w calculation is substantially different from the legacy approximation.

2. The discussion of the existing literature on page 5200, line 4-22 is misleading. It suggest that there are major issues in the existing literature on heat stress, pointing to the inaccurate moist thermodynamic calculations. However, from Fig.1 it seems that this effect is either small or even completely irrelevant for the findings. I am also surprised about the claim that there is an error in Benestad (2010). I do not know the study but please provide more detail about and verify with authors before making such a claim in a side remark. Also the criticism of the studies looking at monthly values is not justified as for instance Dunne et al. are carefully motivating why they choose the monthly time scale.

We expanded our section reviewing previous work, as Reviewer #1 also wanted clarification in that section. Benestad (2010) does not look at moisture metrics in their analysis, yet cites Sherwood and Huber (2010) for the T_w 35°C threshold limits for the temperature analysis they perform. Dunne et al. (2013) use labor work capacities algorithms that are designed for sub-hourly extremes, yet, Dunne et al. (2013) does not show why the sub-daily effects of heat stress may be ignored in their analysis.

We have revised the Introduction:
We restructured the first part:
There are a limited number of studies validating, exploring, or using heat stress metrics on a global scale (Kjellstrom et al., 2009b; Hyatt et al., 2010; Sherwood and Huber, 2010; Fischer and Schar, 2010; Fischer et al., 2012; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Algorithms for measuring heat stress and labor capacity are based upon sub-daily rates of exposure to heat stress (Parsons, 2006). Most of these studies do not capture the diurnal cycle of heat stress (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013), thus not representing both night time highs, and daytime extremes. Only one study includes solar radiation as a component in heat stress (Kjellstrom et al., 2013). Different metrics are used between each study, and only one study attempts to compare more than two metrics (Oleson, et al., 2013b).

Expanded the section detailing criticisms:

Various forms of moist thermodynamic calculations (Buck, 1981; Davies-Jones, 2008; Stull, 2011) and heat stress metrics are criticized for their inaccuracies (Budd, 2008; Alfano et al., 2010; Davies-Jones, 2008). **Buck (1981) moist thermodynamic calculations are not as accurate as Bolton (1980), yet are used in a recent study (Kjellstrom et al., 2013). Wet bulb temperature calculations are different between multiple studies** (Hyatt et al., 2010; Sherwood and Huber, 2010; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Hyatt et al. (2010) and Kjellstrom et al. (2013) use **natural wet bulb temperature—a calculation, that due to non-linear components of its equation, may have multiple steady state solutions** (Alfano et al., 2010). **Oleson et. al. (2013b) uses a recent formulation of wet bulb temperature that is limited in effective range of input temperatures** (Stull, 2011) (we go into further detail on this equation in section 2) **Sherwood and Huber (2010) and Dunne et al. (2013) both use Davies-Jones (2008) as their source paper for their calculation of wet bulb temperature. However, Sherwood and Huber's (2010) wet bulb temperature calculations use Bolton's (1980) equivalent potential temperature Eq. (38), rather than the more accurate Eq. (39) (Bolton, 1980; Davies-Jones, 2008; Davies-Jones, 2009).** Furthermore, **their calculation is only valid for wet bulb temperatures above 10°C. Dunne et al. (2013), on the other hand, uses wet bulb potential temperature Eq. (3.4)** in Davies-Jones (2008), **yet the recommended equations for wet bulb potential temperature are Eq. (3.5-3.7, and 3.8)** (Davies-Jones, 2008) for the temperature ranges used in their paper. **Dunne et al. (2013) also uses Bolton's (1980) equivalent potential temperature Eq. (40), rather than the more accurate Eq. (39) (Bolton, 1980). Additionally, Dunne et al. (2013) uses a variation of WBGT that is heavily criticized, the indoorWBGT (Budd, 2008).**

And went into detail about how heat stress results may be misinterpreted:

Occasionally, results using heat stress limits are misinterpreted. One study confuses wet bulb temperature thresholds with dry bulb temperature thresholds (Benestad, 2011). This has misleading consequences as their results do not include moisture metrics, **yet the author cites Sherwood and Huber (2010)'s wet bulb threshold (35°C) as the threshold value for their temperature analysis**. The **wet bulb temperature at 35°C is a theoretical limit where humans would die from heat stress after 6 hours of exposure. Benestad (2011)'s misapplication implies that most humans should die every year, because a great portion of the world reaches temperatures of 35°C for more than a 6 hour period.**

3. I do not see why the paper defines and discusses indices such as the indoor WBGT when in the end it is not even implement in the code. The entire first half of the manuscript should be substantially shortened to the essential parts. I recommend defining the indices in a table or list to enhance readability and give only a short discussion of their strength and weaknesses in the text.

Both reviewers stated that the manuscript is lengthy, and recommend shortening. We reduce the entire metric background section.

4. The ***joint distribution analysis is not really convincing****, it is not clear what research question it addresses and thus it does not add any novel understanding.* ***Do you want to understand which indices give more weight to temperature or to humidity?*** *If so you could basically do that in an xy-plot showing temperature on one axis and relative humidity on the other axis, and then add the isolines for the individual indices. The slopes would then tell you which indices give more weight to temperature or to humidity. Another approach would be to produce a QQ-plot of temperature and humidity versus each index or correlate their time series. If the emphasis is more on the* ***spatial pattern I would like to see a more quantitative analysis*** *like a pattern correlation of the contributing variables and the indices. But again,* ***emphasis in the results section should be put on demonstrating the added value*** *of the new code implementation.*

We thank the reviewer for pointing out that our joint/conditional distribution plots were not convincing. We have removed all of the box and whisker plots (Figures 4-7). Previous work has explored the differences between heat stress metrics (Epstein and Moran, 2006). However, as there is no theory or analytical approach to dealing with turbulent (i.e. extreme) events, we cannot rely on linear relationships to describe our results. Our new Figure 2, and subsequently its applications in Figures 3-6, show that these differences in T and Q show up as completely different responses in extreme heat stress events within metrics, and between metrics (see Major Comment 2).

5. The figure captions are highly cryptic and the readability of the figures is poor. Please spell out the abbreviations in the caption and clearly describe what is shown. The caption should allow a reader to understand the key message of a figure without having to read the whole manuscript.

We have replaced all figure captions with clear, concise explanations.

Ex.

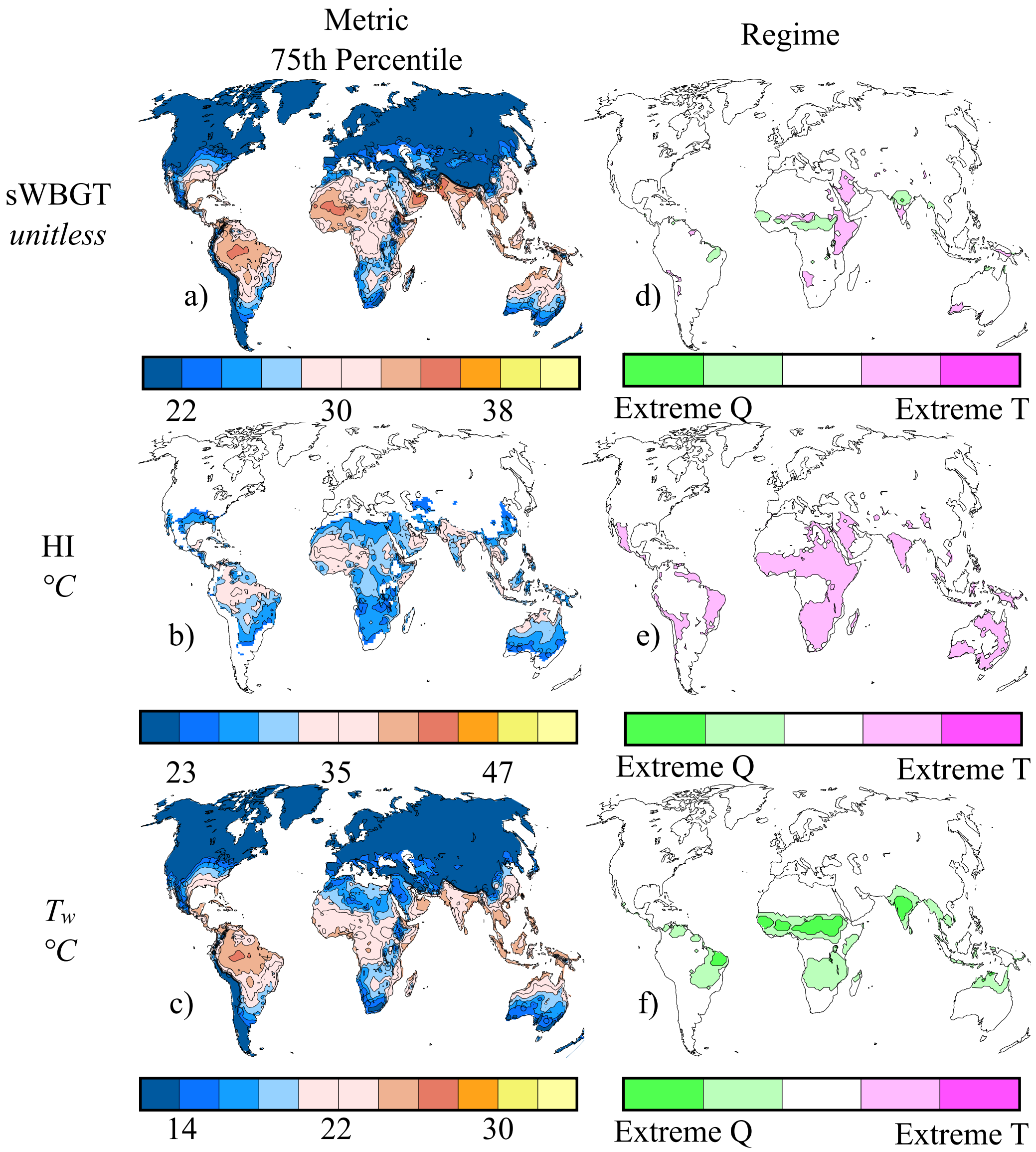


Figure 3. 75th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

6. Abstract: the abstract is too long and does not provide a concise summary of what is new in this manuscript

We have rewritten the abstract reduce its length make it concise.

We implement and analyze 13 different metrics (4 moist thermodynamic quantities and 9 heat stress metrics) in the Community Land Model (CLM4.5), the land surface component of the Community Earth System Model (CESM). We call these routines the HumanIndexMod. **We limit** the algorithms of the HumanIndexMod **to meteorological inputs of temperature, moisture, and pressure** for their calculation. **All metrics assume no direct sunlight exposure. The goal of this project is to implement a common framework for calculating operationally**

used heat stress metrics, in climate models, offline output, and locally sourced weather datasets, with the intent that the HumanIndexMod may be used with the broadest of applications. The **thermodynamic quantities use the latest accurate and efficient algorithms** available, which in turn are used as inputs to the heat stress metrics. There are three advantages of adding these metrics to CLM4.5 **1) improved moist thermodynamic quantities, 2) quantifying heat stress in every available environment within CLM4.5, and 3) these metrics may be used with human, animal, and industrial applications.**

We demonstrate the capabilities of the HumanIndexMod in a default configuration simulation using CLM4.5. We output 4x daily temporal resolution globally. **We show that the advantage of implementing these routines into CLM4.5 is capturing the nonlinearity of the covariation of temperature and moisture conditions.** For example, **we show that there are systematic biases of 1.5°C between monthly and 4x daily offline calculations and the online instantaneous calculation, respectively.** Additionally, **we show that the differences between an inaccurate wet bulb calculation and the improved wet bulb calculation are ±1.5°C.** These differences are important due to human responses to heat stress being non-linear. Furthermore, **we show heat stress has unique regional characteristics. Some metrics have a strong dependency on regionally extreme moisture, while others have a strong dependency on regionally extreme temperature.**

7. 5199, line 6: replace "heat death" by "heat-related mortality" or similar 5199, 12: note that there are very large uncertainties in these numbers, see e.g. Robine, J., S. et al. (2008), Death toll exceeded 70,000 in Europe during the summer of 2003, C. R. Biol., 331(2), 171–178, doi:10.1016/j.crvi.2007.12.001. Instead of giving a number that may be very inaccurate, provide a range of values given in the literature or just refer to "tens of thousands".

We have rewritten the sentence and incorporated your reference.

Although there is high uncertainty in the number of deaths, the 2003 European heat wave killed 40,000 people during a couple weeks in August (Garcia-Herrera et al., 2010), **and tens of thousands more altogether for the entire summer (Robine et al., 2008).**

8. 5199, 18: Statement "Another study shows that the trend in hot extremes has continued despite the warming hiatus". The statement "even without El Niño" is confusing. I would argue if there hadn't been an El Niño in 1998, the trend would have been even more pronounced and I think this is also what the paper says.

We have reworded the sentence:

Another study shows that even with including the global warming 'hiatus', there is an increasing occurrence of extreme temperatures (Seneviratne et al., 2014).

9. 5202, 11: "has" -> "have"

This section is being combined with beginning of Section 3 to reduce redundancy and enhance readability.

10. 5207, 14-15: Motivate why you do not use wind and radiation. I assume there are good reasons why we do not have too high confidence in these variables in coarse resolution models. Anyway, the limitations may not be obvious for non-experts and should be briefly discussed.

The wind, radiation, etc., are outside the scope of this paper, as we are focusing on atmospheric constrained variables. We mention address this in multiple sections.

From the abstract:
All metrics assume no direct sunlight exposure.

From the Introduction:
Our goal here is to improve the situation by creating a module that calculates a large suite of metrics, using the most accurate and efficient algorithms available, that may be used with as many applications as possible: climate models, offline archive data, model validation studies, and weather station datasets.

From the Section 2:

2 Heat stress modeling

2.1 Background

The primary focus of this paper is on atmospheric variable based heat stress metrics that we introduce into the HumanIndexMod. The models for determining heat stress for humans vary greatly; ranging from simple indices to complex prognostic physiology modeling (Table 1). Prognostic thermal models are beyond the scope of this paper, as they require more than atmospheric inputs. **Additionally, metrics that include radiation and wind (with one exception, Apparent Temperature) are also beyond the scope of this paper.** Each index that we chose uses a combination of **atmospheric variables: temperature (*T*), humidity (*Q*), and pressure (*P*). We chose these metrics because they are in operational use globally by industry, governments, and weather services.** Furthermore, **these metrics may be applied to the broadest range of uses: climate and weather forecasting models, archive datasets, and local weather stations.**

11. 5212, 24: Quantify how much difference it makes.

Bolton (1980) and Davies-Jones (2009) make an extensive effort to evaluate their numerous forms of approximation for θ_E. There are multiple tables and figures that map out the errors and evaluate the differences between equations. Our objectives are not to prove that one metric is better than another within Bolton's or Davies-Jones's work, only to use their recommended equations.

12. 5214, 2-7: Why don't you output minimum and maximum values across the 6 hour intervals?

For the conditional analysis that we show, we need to use 4x Daily values for comparability. One of the advantages to having the HumanIndexMod is the capability to compare each metric with another using the same inputs at the model time step resolution. Our new figures reflect this and demonstrates that the same inputs to all metrics do not produce the same 99th percentiles. This type of analysis cannot be accomplished with minimum and maximum values within the 6 hour intervals (See Major Comment 2).

13. 5214, 10-11: "within the 99th percentiles", I assume you mean exceeding the 99th percentile.

We have removed this language.

14. 5214, 13-14: Rephrase sentence, it is unclear. What do you mean by "the time domain"?

We have rephrased this section to enhance readability.
From Section 3.4:
Every 6 hour period that exceeds the percentiles were located within the time series, and we calculated the conditional distributions. For example, the 99th percentile exceedance of HI isolated the top 1606 hottest time steps for each grid cell. **After isolating these time steps, we use this distribution as a mask to isolate all other quantities** (e.g., temperature and humidity), allowing cross comparison between all metrics and HI. The goal is to develop an analysis technique comparing all covariances of the metrics within CLM4.5.

15. 5214, 24-26: Sentence unclear.
16. 5214, 26-28: What figures are you referring to?

We have removed the box and whisker plots (Figure 4-7), and redesigned our analysis. We have removed any mention to these figures.

17. 5215, 18: "This is unsurprising given their underlying similarity as a buoy- ancy measure." I am not sure I understand.

θ_E is a function of T_w, and they are both measurements of buoyancy. However, to reduce length of the manuscript, we have removed this sentence.

18. 5216, 18 and Table 5: These categories seem unconventional. Either you group the locations by latitudes or by regime but then I would expect an objective criterion with no overlap. E.g. I do not see why the tropics are not considered as "moist convective".

We have removed Figures 4-7, and removed the reference table for those Figures, Table 5, to reduce the size of the manuscript and enhance clarity.

19. 5217, 4-5: Extremely long tail in Fig. 6. It is not clear what you are referring to.

We have removed Figure 6.

20. 5217, 19: I think Diffenbaugh et al. used daily data and Dunne et al. were very clear that they never intended to quantify peak heat stress but rather to explore sustained stress causing work inefficiency.

We have corrected the reference to Diffenbaugh et al., and thank the reviewer for pointing out this error. Dunne et al. (2013) use equations that are designed for hourly measurements of heat stress in the workplace. Additionally, as stated in Major Comment 1, the monthly data over estimates the heat stress values, overestimating the threat to humans.

There are many approaches to evaluating heat stress. **Monthly and seasonal temperature** and moisture averages were used for general applications **(Dunne et al., 2013)**, **however these averages overestimate the potential severity of heat stress** (Figure 1c,d). Even using daily or sub-daily averages (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) potentially overestimates heat stress. This is due to the non-linear covariance of T and Q, and averages miss these extremes.

21. 5218, 7-19: This is a very detailed feature and its discussion seems odd in this part of the manuscript. If the authors think this bias is crucial and needs to be documented, it would need to be illustrated with a figure or introduced and discussed with more background so that it is understandable for a non-expert knowing all the details of this parameterization.

In order to reduce the length of the manuscript, we have removed this section.

References

Epstein, Y., & Moran, D. S. Thermal comfort and the heat stress indices. Industrial Health, 44(3), 388-398, 2006.

Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5

J. R. Buzan^{1,2}, K. Oleson^{3}, and M. Huber^{1,2}

[1]{Department of Earth Sciences, University of New Hampshire, Durham New Hampshire, USA}

[2]{Earth Systems Research Center, Institute for the Study of Earth, Ocean, and Space, University of New Hampshire, Durham New Hampshire, USA}

[3]{National Center for Atmospheric Research, Boulder Colorado, USA}

Correspondence to: J. R. Buzan (jonathan.buzan@unh.edu)

Abstract

We implement and analyse 13 different metrics (4 moist thermodynamic quantities and 9 heat stress metrics) in the Community Land Model (CLM4.5), the land surface component of the Community Earth System Model (CESM). We call these routines the HumanIndexMod. We limit the algorithms of the HumanIndexMod to meteorological inputs of temperature, moisture, and pressure for their calculation. All metrics assume no direct sunlight exposure. The goal of this project is to implement a common framework for calculating operationally used heat stress metrics, in climate models, offline output, and locally sourced weather datasets, with the intent that the HumanIndexMod may be used with the broadest of applications. The thermodynamic quantities use the latest accurate and efficient algorithms available, which in turn are used as inputs to the heat stress metrics. There are three advantages of adding these metrics to CLM4.5 1) improved moist thermodynamic quantities, 2) quantifying heat stress in every available environment within CLM4.5, and 3) these metrics may be used with human, animal, and industrial applications.

We demonstrate the capabilities of the HumanIndexMod in a default configuration simulation using CLM4.5. We output 4x daily temporal resolution globally. We show that the advantage of implementing these routines into CLM4.5 is capturing the nonlinearity of the covariation of temperature and moisture conditions. For example, we show that there are

JRBuzan 12/2/14 3:51 PM
Deleted: Sciences

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: analyze

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: These heat stress metrics embody three philosophical approaches: comfort, physiology, and empirically based

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: . The metrics are directly ... [1]

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: are

JRBuzan 12/2/14 3:51 PM
Formatted ... [2]

JRBuzan 12/2/14 3:51 PM
Deleted: calculations within climate models

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: human

JRBuzan 12/2/14 3:51 PM
Formatted ... [3]

JRBuzan 12/2/14 3:51 PM
Deleted: that

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: applied to other animals as well as

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: Additionally, an offline ver ... [4]

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: , we analyze the top 1% of h ... [5]

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: . We cross compare these e ... [6]

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted: heat stress may be divided i ... [7]

systematic biases of up to 1.5°C between monthly and ±0.5°C between 4x daily offline calculations and the online instantaneous calculation, respectively. Additionally, we show that the differences between an inaccurate wet bulb calculation and the improved wet bulb calculation are ±1.5°C. These differences are important due to human responses to heat stress being non-linear. Furthermore, we show heat stress has unique regional characteristics. Some metrics have a strong dependency on regionally extreme moisture, while others have a strong dependency on regionally extreme temperature.

1 Introduction

Heat related conditions are the number one cause of death from natural disaster in the United States; more than tornadoes, flooding, and hurricanes combined (NOAAWatch, 2014). Short-term duration (hours) of exposure to heat while working may increase the incidence of heat exhaustion and heat stroke (Liang et al., 2011). However, long-term exposure (heat waves or seasonally high heat), even without working, may drastically increase morbidity and mortality (Kjellstrom et al., 2009). Although there is high uncertainty in the number of deaths, the 2003 European heat wave killed 40,000 people during a couple weeks in August (Garcia-Herrera et al., 2010), and tens of thousands more altogether for the entire summer (Robine et al., 2008). The 2010 Russian heat wave, the worst recorded heat wave, killed 55,000 people over the midsummer (Barriopedro et al., 2011).

A growing literature is concerned with the frequency and duration of heat waves (Seneviratne et al., 2012 and references therein). One study concluded that intensification of 500-hPa height anomalies will produce more severe heat waves over Europe and North America in the future (Meehl and Tebaldi, 2004). Another study shows that even with including the global warming 'hiatus', there is an increasing occurrence of extreme temperatures (Seneviratne et al., 2014). Multiple studies associate lack of precipitation and/or low soil moisture to contributing to high temperatures (Fischer et al., 2007; Mueller and Seneviratne, 2012; Miralles et al., 2014).

Regarding humans, however, temperature differences are not the primary method for heat dissipation. Evaporation of sweat is crucial to maintaining homeostasis, and none of the before mentioned studies incorporate atmospheric moisture to measure heat stress. Many diagnostic and prognostic methods were developed to diagnose heat stress (over a 100 year history, Table 1), such as the Wet Bulb Globe Temperature (WBGT), the Discomfort Index

Deleted: convection (±30° latitude). ... [9]

JRBuzan 12/2/14 3:51 PM
Formatted ... [8]

JRBuzan 12/2/14 3:51 PM
Deleted: heat stress as compared to t ... [11]

JRBuzan 12/2/14 3:51 PM
Formatted ... [10]

JRBuzan 12/2/14 3:51 PM
Formatted ... [12]

JRBuzan 12/2/14 3:51 PM
Formatted ... [13]

JRBuzan 12/2/14 3:51 PM
Deleted: death is

JRBuzan 12/2/14 3:51 PM
Formatted ... [14]

JRBuzan 12/2/14 3:51 PM
Deleted: tornados

JRBuzan 12/2/14 3:51 PM
Formatted ... [15]

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Formatted ... [16]

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Formatted ... [17]

JRBuzan 12/2/14 3:51 PM
Deleted: The

JRBuzan 12/2/14 3:51 PM
Formatted ... [18]

JRBuzan 12/2/14 3:51 PM
Deleted: García

JRBuzan 12/2/14 3:51 PM
Formatted ... [19]

JRBuzan 12/2/14 3:51 PM
Formatted ... [20]

JRBuzan 12/2/14 3:51 PM
Deleted: in

JRBuzan 12/2/14 3:51 PM
Formatted ... [21]

JRBuzan 12/2/14 3:51 PM
Deleted: absence of El Ninos since 1997

JRBuzan 12/2/14 3:51 PM
Formatted ... [22]

JRBuzan 12/2/14 3:51 PM
Deleted: ,

JRBuzan 12/2/14 3:51 PM
Formatted ... [23]

JRBuzan 12/2/14 3:51 PM
Formatted ... [24]

JRBuzan 12/2/14 3:51 PM
Deleted: different heat

JRBuzan 12/2/14 3:51 PM
Formatted ... [25]

JRBuzan 12/2/14 3:51 PM
Deleted: indices

JRBuzan 12/2/14 3:51 PM
Formatted ... [26]

(DI), or Heat Index (HI), and policy makers have decided to incorporate these indices in weather warning systems (Epstein and Moran, 2006; Parsons, 2006; Parsons, 2013; Rothfusz, 1990; Fiala et al., 2011).

There are a limited number of studies validating, exploring, or using heat stress metrics on a global scale (Kjellstrom et al., 2009b; Hyatt et al., 2010; Sherwood and Huber, 2010; Fischer and Schar, 2010; Fischer et al., 2012; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Algorithms for measuring heat stress and labor capacity are based upon sub-daily rates of exposure to heat stress (Parsons, 2006). Most of these studies do not capture the diurnal cycle of heat stress (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013), thus not representing both nighttime highs, and daytime extremes. Only one study includes solar radiation as a component in heat stress (Kjellstrom et al., 2013). Different metrics are used between each study, and only one study attempts to compare more than two metrics (Oleson, et al., 2013b).

Various forms of moist thermodynamic calculations (Buck, 1981; Davies-Jones, 2008; Stull, 2011) and heat stress metrics are criticized for their inaccuracies (Budd, 2008; Alfano et al., 2010; Davies-Jones, 2008). Buck (1981) moist thermodynamic calculations are not as accurate as Bolton (1980), yet are used in a recent study (Kjellstrom et al., 2013). Wet bulb temperature calculations are different between multiple studies (Hyatt et al., 2010; Sherwood and Huber, 2010; Dunne et al., 2013; Kjellstrom et al., 2013; Oleson et al., 2013). Hyatt et al. (2010) and Kjellstrom et al. (2013) use natural wet bulb temperature—a calculation, that due to non-linear components of its equation, may have multiple steady state solutions (Alfano et al., 2010). Oleson et al. (2013b) uses a recent formulation of wet bulb temperature that is limited in effective range of input temperatures (Stull, 2011) (we go into further detail on this equation in section 2). Sherwood and Huber (2010) and Dunne et al. (2013) both use Davies-Jones (2008) as their source paper for their calculation of wet bulb temperature. However, Sherwood and Huber's (2010) wet bulb temperature calculations use Bolton's (1980) equivalent potential temperature Eq. (38), rather than the more accurate Eq. (39) (Bolton, 1980; Davies-Jones, 2008; Davies-Jones, 2009). Furthermore, their calculation is only valid for wet bulb temperatures above 10°C. Dunne et al. (2013), on the other hand, uses wet bulb potential temperature Eq. (3.4) in Davies-Jones (2008), yet the recommended equations for

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic, English (UK)

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic, English (UK)

JRBuzan 12/2/14 3:51 PM
Formatted: English (UK)

JRBuzan 12/2/14 3:51 PM
Formatted: Font color: Black, English (UK)

JRBuzan 12/2/14 3:51 PM
Deleted:).

JRBuzan 12/2/14 3:51 PM
Moved down [1]: are criticized for their inaccuracies (Budd, 2008; Alfano et al., 2010; Davies-Jones, 2008).

JRBuzan 12/2/14 3:51 PM
Deleted: These calculations, however, are widely used (Sherwood and Huber, 2010; Hyatt et al., 2010; Fischer et al., 2012; Willett and Sherwood, 2012; Dunne et al., 2013; Kjellstrom et al., 2013).

JRBuzan 12/2/14 3:51 PM
Moved down [2]: Occasionally, results using heat stress limits are misinterpreted.

JRBuzan 12/2/14 3:51 PM
Moved down [3]: 2010).

JRBuzan 12/2/14 3:51 PM
Deleted: night time highs, and daytime extremes. For example, one study used monthly averages to calculate extreme heat stress (Dunne et al., 2013). Various forms of moist thermodynamic calculations (Buck, 1981; Davies-Jones, 2008; Stull, 2011) and heat stress metrics (ACSM, 1984)

JRBuzan 12/2/14 3:51 PM
Deleted: One study confuses wet bulb temperature thresholds with dry bulb temperature thresholds, which has misleading consequences (Benestad,

JRBuzan 12/2/14 3:51 PM
Deleted: no

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Formatted: Font color: Auto

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [1]

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [3]

wet bulb potential temperature are Eq. (3.5-3.7, and 3.8) (Davies-Jones, 2008) for the temperature ranges used in their paper. Dunne et al. (2013) also uses Bolton's (1980) equivalent potential temperature Eq. (40), rather than the more accurate Eq. (39) (Bolton, 1980). Additionally, Dunne et al. (2013) uses a variation of WBGT that is heavily criticized, the indoorWBGT (Budd, 2008).

Occasionally, results using heat stress limits are misinterpreted. One study confuses wet bulb temperature thresholds with dry bulb temperature thresholds (Benestad, 2011). This has misleading consequences as their results do not include moisture metrics, yet the author cites Sherwood and Huber (2010)'s wet bulb threshold (35°C) as the threshold value for their temperature analysis. The wet bulb temperature at 35°C is a theoretical limit where humans would die from heat stress after 6 hours of exposure. Benestad (2011)'s misapplication implies that most humans should die every year, because a great portion of the world reaches temperatures of 35°C for more than a 6 hour period.

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [2]

Our goal here is to improve the situation by creating a module that calculates a large suite of metrics, using the most accurate and efficient algorithms available, that may be used with as many applications as possible: climate models, offline archive data, model validation studies, and weather station datasets. We call this module the HumanIndexMod. The module calculates 4 moist thermodynamic quantities and 9 heat stress metrics. These heat stress metrics are in operational use worldwide, and cover a wide range of assumptions.

JRBuzan 12/2/14 3:51 PM
Deleted: implementing and comparing

JRBuzan 12/2/14 3:51 PM
Deleted: into a commonly used open source

JRBuzan 12/2/14 3:51 PM
Deleted: , using the highest temporal variability and accurate moist thermodynamics we can bring to bear currently. This paper implements 13 different metrics

As an example of numerous applications, we implement the HumanIndexMod into the Community Land Model (CLM4.5), a component model of the Community Earth System Model (CESM), maintained by the National Center for Atmospheric Research (NCAR) (Hurrell et al., 2013). The metrics are directly calculated at the sub-grid scale, capturing heat stress in every environment: urban, lakes, vegetation, and bare ground. We show examples of the advantages of calculating these metrics at the model time step as compared to lower temporal resolution, and the importance of using accurate moist thermodynamic calculations. We also show that having all metrics calculated at the same time allows for comparison of metrics between each other, and allows for unique analysis of conditional distributions of the inputs. Finally, we show that the metrics may also be used as model diagnostics.

The outline of the paper is as follows: section 2 (Heat stress modeling) focuses on the development, calculation, and use of these 13 metrics. Section 3 (Methods) describes the implementation and model setup. Section 4 (Results) presents the results of a model

JRBuzan 12/2/14 3:51 PM
Deleted: Background

simulation using these metrics. Section 5 (Discussion) discusses the implications of the research, and section 6 (Summary) presents the conclusions of the paper.

2 Heat stress modeling

2.1 Background

The primary focus of this paper is on atmospheric variable based heat stress metrics that we introduce into the HumanIndexMod. The models for determining heat stress for humans vary greatly; ranging from simple indices to complex prognostic physiology modelling (Table 1). Prognostic thermal models are beyond the scope of this paper, as they require more than atmospheric inputs. Additionally, metrics that include radiation and wind (with one exception, Apparent Temperature) are also beyond the scope of this paper. Each index that we chose uses a combination of atmospheric variables: temperature (*T*), humidity (*Q*), and pressure (*P*). We chose these metrics because they are in operational use globally by industry, governments, and weather services. Furthermore, these metrics may be applied to the broadest range of uses: climate and weather forecasting models, archive datasets, and local weather stations.

Sections 2.2-2.4 describe the metrics that we have chosen to implement in the HumanIndexMod (see variables defined in Table 2). Most of the metrics have units of temperature, which may be misleading. The metrics have temperature scales for comparative purposes only, as the metrics are an index, not a true thermodynamic quantity. We break these metrics into three categories, based upon design philosophies: comfort, physiological response, and empirical fit. Comfort based algorithms are a quantification of behavioural or ‘feels like’ reactions to heat in both animals and humans. Physiological indices quantify the physical response mechanisms within a human or animal, such as changes in heart rate or core temperatures. The empirical indices quantify relationships between weather conditions and a non-physical or comfort related attribute. For example, an empirical algorithm’s result may determine how much work may be completed per hour per weather condition.

2.2 Comfort algorithms

We use Apparent Temperature, Heat Index, Humidex, and Temperature Humidity Index for Comfort to account for comfort level. These metrics were either tailored to the global

Formatted ... [27]
JRBuzan 12/2/14 3:51 PM
Moved (insertion) [4] ... [28]
JRBuzan 12/2/14 3:51 PM
Formatted ... [29]
JRBuzan 12/2/14 3:51 PM
Formatted ... [30]
JRBuzan 12/2/14 3:51 PM
Deleted: structure
JRBuzan 12/2/14 3:51 PM
Formatted ... [31]
JRBuzan 12/2/14 3:51 PM
Moved down [5]: . Each grid cell ... [34]
JRBuzan 12/2/14 3:51 PM
Moved down [6]: urban classificat ... [36]
JRBuzan 12/2/14 3:51 PM
Deleted: Community Land Mo ... [32]
JRBuzan 12/2/14 3:51 PM
Formatted ... [33]
JRBuzan 12/2/14 3:51 PM
Deleted: flexible
JRBuzan 12/2/14 3:51 PM
Deleted: CLM4.5 includes a carbon ... [35]
JRBuzan 12/2/14 3:51 PM
Deleted: . Radiation is absorbed wit ... [37]
JRBuzan 12/2/14 3:51 PM
Formatted ... [38]
JRBuzan 12/2/14 3:51 PM
Deleted: average) between the urban ... [39]
JRBuzan 12/2/14 3:51 PM
Formatted ... [40]
JRBuzan 12/2/14 3:51 PM
Deleted: . Specifically, the potential ... [41]
JRBuzan 12/2/14 3:51 PM
Formatted ... [42]
JRBuzan 12/2/14 3:51 PM
Deleted: are used for determining he ... [43]
JRBuzan 12/2/14 3:51 PM
Moved down [8]: 1).
JRBuzan 12/2/14 3:51 PM
Deleted: Both ARM and Flatau’s ... [44]
JRBuzan 12/2/14 3:51 PM
Moved up [4]: . ... [45]
JRBuzan 12/2/14 3:51 PM
Formatted ... [46]
JRBuzan 12/2/14 3:51 PM
Deleted: modeling
JRBuzan 12/2/14 3:51 PM
Formatted ... [47]
JRBuzan 12/2/14 3:51 PM
Deleted: human
JRBuzan 12/2/14 3:51 PM
Formatted ... [48]
JRBuzan 12/2/14 3:51 PM
Deleted: and instead, we are focusin ... [49]
JRBuzan 12/2/14 3:51 PM
Formatted ... [50]
JRBuzan 12/2/14 3:51 PM
Deleted: The different atmospheric ... [51]
JRBuzan 12/2/14 3:51 PM
Formatted ... [52]
JRBuzan 12/2/14 3:51 PM

JRBuzan 12/2/14 3:51 PM
Formatted ... [53]
JRBuzan 12/2/14 3:51 PM

JRBuzan 12/2/14 3:51 PM
Formatted ... [54]
JRBuzan 12/2/14 3:51 PM

JRBuzan 12/2/14 3:51 PM
Formatted ... [55]
JRBuzan 12/2/14 3:51 PM
... [56]
JRBuzan 12/2/14 3:51 PM
Formatted ... [57]

locations where they were developed, or streamlined for ease of use from physiology models. The underlying philosophical approach to deriving comfort metrics is representing behavioral reactions to levels of comfort (Masterson and Richardson, 1979; Steadman, 1979a). The goal of these equations of comfort is to match the levels of discomfort to appropriate warnings for laborers (Gagge et al., 1972) and livestock (Renaudeau et al., 2012). Discomfort in humans sets in much earlier than physiological responses, i.e. the human body provides an early warning to the mind that continuing the activity may lead to disastrous consequences. For example, when heat exhaustion sets in, the body is sweating profusely, and often there are symptoms of dizziness. However, the actual core temperature for heat exhaustion is defined at 38.5°C, which is considerably lower than heat stroke (42°C). We describe the 4 comfort based algorithms below.

Apparent Temperature (AT) was developed using a combination of wind, radiation, and heat transfer to measure thermal comfort and thermal responses in humans (Steadman, 1994). AT is used by the Australian Bureau of Meteorology, and was developed for climates in Australia (ABM, 2014). The metric is an approximation of a prognostic thermal model of human comfort (Steadman 1979a,b; Steadman, 1984).

$$AT = T_c + \frac{3.3e_{RH}}{1000} - 0.7u_{10m} - 4 \qquad (1)$$

$$e_{RH} = \left({}^{RH}/_{100}\right)e_{sPa} \qquad (2)$$

where the vapor pressure (e_{RH}) is in Pascals and is calculated from the relative humidity (RH in %), and saturated vapor pressure (e_{sPa}, also in Pascals). We use this notation because e_s (Table 2) is in millibars. These variable names are the explicit names of the variables in the HumanIndexMod. AT uses the wind velocity (m/s) measured at the 10m height (u_{10m}). Air temperature (T_C) and AT are in units of degrees Celsius. AT is the only metric in the HumanIndexMod that includes an explicit calculation for wind velocity; the other metrics assume a reference wind. We included this metric due to a previously used legacy version within CLM4.5 (Oleson et al., 2013b). An assumption made by AT is that the subject is outside, but not exposed to direct sunlight. AT has no explicit thresholds; rather, the index shows an amplification of temperatures. Previous work, however, has used temperature percentiles to describe AT (Oleson et al., 2013b).

JRBuzan 12/2/14 3:51 PM
Deleted: actual

JRBuzan 12/2/14 3:51 PM
Deleted: Apparent Temperature, Heat Index, Humidex, and Temperature Humidity Index account for the comfort level, and they were tailored to the world locations where they were developed in, or streamlined for ease of use, as described further

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted:), and is as follows:

JRBuzan 12/2/14 3:51 PM
Deleted: vapour

JRBuzan 12/2/14 3:51 PM
Deleted: vapour pressures (e_{RH} and

JRBuzan 12/2/14 3:51 PM
Deleted: respectively),

JRBuzan 12/2/14 3:51 PM
Deleted: in Pascals. u_{10m} (m/s) is

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: RH (%) is the relative humidity. Of the metrics we implement, this

JRBuzan 12/2/14 3:51 PM
Deleted: explicitly

JRBuzan 12/2/14 3:51 PM
Deleted: others

JRBuzan 12/2/14 3:51 PM
Deleted: for

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: ,

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

Heat Index (HI) was developed using a similar process as AT. The United States National Weather Service (NWS) required a heat stress early warning system, and the index was created as a polynomial fit to Steadman's (1979a) comfort model (Rothfusz, 1990).

$$HI = -42.379 + 2.04901523T_F + 10.14333127RH + -0.22475541T_F RH + -6.83783\times 10^{-3}T_F^2 + -5.481717\times 10^{-2}RH^2 + 1.22874\times 10^{-3}T_F^2 RH + 8.5282\times 10^{-4}T_F RH^2 + -1.99\times 10^{-6}T_F^2 RH^2 \qquad (3)$$

Here, air temperature (T_F) and HI are in Fahrenheit. HI has a number of assumptions. The equation assumes a walking person in shorts and T-shirt, who is male and weighs ~147lbs (Rothfusz, 1990). Additionally, this subject is not in direct sunlight. As with AT, HI represents a 'feels like' temperature, based upon levels of discomfort. HI uses a scale for determining heat stress: 27-32°C is caution, 33-39°C is extreme caution, 40-51°C is danger, and ≥52°C is extreme danger.

Humidex (HUMIDEX) was developed for the Meteorological Service of Canada, and describes the 'feels like' temperature for humans (Masterson and Richardson, 1979). The original equation used dew point temperature, rather than specific humidity. The equation was modified to use vapor pressure, instead:

$$HUMIDEX = T_C + \frac{5}{9}\left(\frac{e_{RH}}{100} - 10\right) \qquad (4)$$

HUMIDEX is unitless because the authors recognized that the index is a measure of heat load. The index has a series of thresholds: 30 is some discomfort, 46 is dangerous, and 54 is imminent heat stroke (Masterson and Richardson, 1979).

The Temperature Humidity Index for Comfort (THIC) is a modification of the Temperature Humidity Index (THI) (Ingram, 1965). Comfort was quantified for livestock through THIC (NWSCR, 1976). We use the original calibration, which is for pigs (Ingram, 1965). The index is unitless:

$$THIC = 0.72T_w + 0.72T_c + 40.6 \qquad (5)$$

where wet bulb temperature (T_w) is in units of Celsius. The index is used to describe behavioral changes in large animals due to discomfort (seeking shade, submerging in mud, etc.). The index is in active use by the livestock industry for local heat stress and future climate considerations (Lucas et al., 2000; Renaudeau et al., 2012). The index describes

JRBuzan 12/2/14 3:51 PM
Formatted ... [64]

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Deleted: $T_f T_F + 10.14333127R$ ... [65]

JRBuzan 12/2/14 3:51 PM
Deleted: $T_f^2 T_F^2 + -5.481717\times$ ... [66]

JRBuzan 12/2/14 3:51 PM
Deleted: T_f^2

JRBuzan 12/2/14 3:51 PM
Deleted: T_f

JRBuzan 12/2/14 3:51 PM
Deleted: , and RH is as described previously

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted ... [67]

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: vapour

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: where air temperature (T_c) and vapor pressure (e_{RH}) are as described previously, and

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: ,…because the authors reco ... [68]

JRBuzan 12/2/14 3:51 PM
Formatted ... [69]

JRBuzan 12/2/14 3:51 PM
Deleted: both humans and

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted:), however, we…. We use ... [70]

qualitative threat levels for animals: 75 is alert, 79-83 is dangerous, and 84+ is very dangerous. There are different approaches to the development of THIC, including considerations of physiology of large animals.

2.3 Physiology algorithms

Numerous metrics are based upon direct physiological responses within humans and animals; however, almost all of them are complicated algorithms (e.g. Moran et al., 2001; Berglund and Yokota, 2005; Gribox et al., 2008; Maloney and Forbes, 2011; Havenith et al., 2011; Gonzalez et al., 2012; Chan et al., 2012). Most metrics require radiation measurements, or heart rates, and/or even sweat rates. The available metrics that are calibrated for physiological responses using only meteorological inputs, though, are limited, such as the Temperature Humidity Index for Physiology (THIP; Ingram, 1965):

$$THIP = 0.63T_w + 1.17T_c + 32 \qquad (6)$$

THIP and THIC are modifications of the Temperature Humidity Index (THI). Additionally, THIC and THIP have applications beyond heat stress. THIP and THIC threshold levels are computed from both indoor and outdoor atmospheric variables. The differences between outdoor and indoor values are used to evaluate evaporative cooling mechanisms, e.g. swamp coolers (Gates et al., 1991a,b).

2.4 Empirical algorithms

The last category of metrics are derived from first principle thermo-physiology models, or changes in worker productivity, etc., and then reduced by empirical fit. The first metric we present is widely used, modification of an industry labor standard, the Simplified Wet Bulb Globe Temperature (sWBGT):

$$sWBGT = 0.56T_C + \frac{0.393e_{RH}}{100} + 3.94 \qquad (7)$$

sWBGT was designed for estimating heat stress in sports medicine, adopted by the Australian Bureau of Meteorology, and is acknowledged that its accuracy of representing the original labor industry index may be questionable (ABOM, 2010; ACSM, 1984; ACSM, 1987). We chose, however, to implement sWBGT due to its wide use. sWBGT is unitless, and its threat levels are: 26.7-29.3 is green or be alert, 29.4-31.0 is yellow or caution, 31.1-32.1 is red or potentially dangerous, and ≥32.2 is black or dangerous conditions (US Army, 2003).

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: .2

JRBuzan 12/2/14 3:51 PM
Deleted: ,… however, almost all of t ... [71]

JRBuzan 12/2/14 3:51 PM
Deleted: The Temperature Humidity Index for Physiology (*THIP*) is one such metric:

JRBuzan 12/2/14 3:51 PM
Deleted: where the temperature inputs are the same as in Eq. (5) (Ingram, 1965).

JRBuzan 12/2/14 3:51 PM
Formatted ... [72]

JRBuzan 12/2/14 3:51 PM
Deleted: generalized by the format: ... [73]

JRBuzan 12/2/14 3:51 PM
Formatted ... [74]

JRBuzan 12/2/14 3:51 PM
Deleted: 3.3

JRBuzan 12/2/14 3:51 PM
Deleted: third…ast category of algo ... [75]

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted:) and Indoor Wet Bulb Globe Temperature (*indoorWBGT*). *sWBGT* is based upon the Wet Bulb Globe Temperature (*WBGT*) that was developed as a decision making tool for the United States Marine Corps to mitigate heat stress casualties during training (Minard et al., 1957). The *WBGT* uses a combination of wet bulb and dry bulb temperatures as well as a globe thermometer (T_g).

JRBuzan 12/2/14 3:51 PM
Deleted: $WBGT = 0.7T_w + 0.2T_g + 0.1T_c$ (8) ... [76]

JRBuzan 12/2/14 3:51 PM
Deleted: $T_c + \frac{0.393e_{RH}}{100} T \quad c \quad +$ ... [77]

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: We did not implement *WBGT*, nor *indoorWBGT*. *WBGT* requires radiation, and is outside the scope of this work. *indoorWBGT* is criticized explicitly for no recalibration due to removing radiation (Budd, 2008).

JRBuzan 12/2/14 3:51 PM
Formatted ... [78]

Discomfort Index (DI) was developed in the 1950s as a calibration for air conditioners (Thom, 1959). It was adapted by the Israeli Defense Force as a decision making tool regarding heat stress (Epstein and Moran, 2006). DI requires T_w and T_C. The computation of T_w in the past was computationally expensive, and the DI equations often used approximations (e.g., Oleson et al., 2013b):

$$T_{wS} = T_C \arctan\left(0.151977\sqrt{RH + 8.313659}\right) + \arctan(T_C + RH) - \arctan(RH - 1.676331) + 0.00391838 RH^{3/2} \arctan(0.023101 RH) - 4.68035 \tag{8}$$

where T_{wS} is the wet bulb temperature in Celsius (Stull, 2011). Stull's function has limited range of effective accuracy.

$$\begin{matrix} -20 < T_C < 50 \\ -2.27T_C + 27.7 < RH < 99 \end{matrix} \tag{9}$$

We compute DI with both T_{wS} and T_w calculated using our implementation of Davies-Jones (2008) (Eq. A.22). We keep the legacy version (Stull, 2011) for comparative purposes. DI is calculated from these inputs:

$$DI = 0.5T_w + 0.5T_c \tag{10}$$

where the DI is unitless and the values are an indicator of threats to the populations: 21-24 is <50% of population in discomfort, 24-27 >50% of population in discomfort, 27-29 most of the population in discomfort, 29-32 severe stress, and >32 is state of emergency (Giles et al., 1990).

The last index we present is a measurement of the capacity of evaporative cooling mechanisms. Often, these are referred to as swamp coolers. Large-scale swamp coolers generally work by spraying a 'mist' into the air, or blowing air through a wet mesh. This mist then comes in contact with the skin, and subsequently evaporated, thus cooling down the subject. In dry environments, they can be an effective mass cooling mechanism. Unfortunately, swamp coolers raise the local humidity considerably, reducing the effectiveness of direct evaporation from the skin. Swamp coolers are measured by their efficiency:

$$\eta = \frac{T_c - T_t}{T_c - T_w} 100\% \tag{11}$$

where η (%) is the efficiency, and T_t is the target temperature for the room to be cooled towards in Celsius (Koca et al., 1991). Rearranging Eq. (11) and solving for T_t:

JRBuzan 12/2/14 3:51 PM
Deleted: has had a similar developm ... [80]

JRBuzan 12/2/14 3:51 PM
Formatted ... [79]

JRBuzan 12/2/14 3:51 PM
Deleted:), it

JRBuzan 12/2/14 3:51 PM
Formatted ... [81]

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: difficult

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: $(T_c + RH)$

JRBuzan 12/2/14 3:51 PM
Deleted: 11

JRBuzan 12/2/14 3:51 PM
Deleted: $\begin{matrix} -20<T_c<50 \\ -2.27T_c+27.7<RH<99 \end{matrix}$

JRBuzan 12/2/14 3:51 PM
Deleted: 12

JRBuzan 12/2/14 3:51 PM
Deleted: where, not only is the funct ... [82]

JRBuzan 12/2/14 3:51 PM
Moved down [9]: DI is calculated ... [83]

JRBuzan 12/2/14 3:51 PM
Formatted ... [84]

JRBuzan 12/2/14 3:51 PM
Deleted: 13) ... [85]

JRBuzan 12/2/14 3:51 PM
Formatted ... [86]

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [9] ... [87]

JRBuzan 12/2/14 3:51 PM
Formatted ... [88]

JRBuzan 12/2/14 3:51 PM
Deleted: The index

JRBuzan 12/2/14 3:51 PM
Formatted ... [89]

JRBuzan 12/2/14 3:51 PM
Deleted: ,

JRBuzan 12/2/14 3:51 PM
Deleted: warning levels

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Deleted: 14

JRBuzan 12/2/14 3:51 PM
Deleted: 14

$$T_t = T_c - \frac{\eta}{100}(T_c - T_w) \qquad (12)$$

where T_t is now the predicted temperature based upon environmental variables. The maximum efficiency of typical swamp coolers is 80%, and a typical value of a sub-standard mechanism is 65% (Koca et al., 1991). Thus, we calculate T_t with two different efficiencies: SWMP80, for η at 80%, and SWMP65 for η at 65%. With the mist-injected air cooled to T_t, T_t is approximately equal to a new local T_w. Humid environments or environments that are hot and have an above average *RH* relative to their normally high *T*, severely limit the cooling potential of swamp coolers. The livestock industry uses evaporative cooling mechanisms for cooling, and often in conjunction with THIP and THIC, as mentioned previously (Gates et al., 1991a,b). Due to their low cost, swamp coolers are used throughout the world as a method of cooling buildings and houses. No one has implemented SWMP65 and SWMP80 in global models, and we believe that this will provide many uses to industry by its inclusion in CLM4.5. Table 2 shows what metrics are discussed in this paper.

3 Methods

Our approach is to choose a subset of heat stress metrics that are in common use operationally by governments and/or used extensively in prior climate modeling studies (Table 3). We do this in order to provide a framework to allow comparisons of metrics across studies, and we designate the algorithms the HumanIndexMod. Section 3.1 describes CLM4.5. Section 3.2 discusses the implementation of the HumanIndexMod into CLM4.5. Section 3.3 describes our simulation setup that we use to demonstrate the capabilities of the HumanIndexMod. The simulation is for showcasing the HumanIndexMod, not as an experiment for describing real climate or climate change. Section 3.4 describes a unique application method for analyzing heat stress.

3.1 The structure of Community Land Model version 4.5

We use CLM version 4.5, which was released in June, 2013 (Oleson et al., 2013a). Boundary conditions for CLM4.5 consist of land cover and atmospheric weather conditions. Each grid cell in CLM4.5 can include vegetation, lakes, wetlands, glacier, and urban. There are new parameterizations and models for snow cover, lakes, crops, a new biogeochemical cycles model, and new urban classifications (Oleson et al., 2013a). The urban biome, a single-layer

JRBuzan 12/2/14 3:51 PM
Deleted: 15

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: <#>Heat stress modeling

JRBuzan 12/2/14 3:51 PM
Deleted: 2

JRBuzan 12/2/14 3:51 PM
Deleted: 3

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [5]

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [6]

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [7]

canyon model, is designed to represent the ‘heat island’, where temperatures are amplified by urban environments (Oleson et al., 2008a,b; Oleson et al., 2010a,b,c). The ‘heat island’ effect can increase the likelihood of complications from human heat stress (Oleson, 2012).

3.2 HumanIndexMod design and implementation

There are two philosophical aspects to the design of the HumanIndexMod. 1) Accurate and efficient moist thermodynamic algorithms, and 2) a modular format to increase use through both narrowly focused applications and up to broad based studies. The module is in an open source format, and is incorporated into the CLM4.5 developer branch (the module itself is available from the corresponding author’s website). The modular format encourages adapting the code to specific needs; whether that focus is on moist thermodynamics or heat stress. The inclusion of heat stress metrics covering comfort, physiology, and empirical philosophies encourages the use of HumanIndexMod for many applications.

We directly implemented the code into the CLM4.5 architecture through seven modules. Four of these modules—BareGroundFuxesMod, CanopyFluxesMod, SlakeFluxesMod, and UrbanMod—call the HumanIndexMod. The HumanIndexMod is calculated for every surface type in CLM4.5. The design of CLM4.5 allows the urban and rural components, where the rural component represents the natural vegetation surface, to be archived separately for intercomparison. The HumanIndexMod uses the 2-meter calculations of water vapor, temperature, and pressure, as well as 10-meter winds. Three other modules are modified with the implementation process. These modules—clmtype, clmtypeInitMod, and histFldsMod—are used for initializing memory and outputting variable history files.

Moist thermodynamic water vapor quantities in CLM4.5 are calculated within QSatMod. We use the outputs from QSatMod as the inputs to the HumanIndexMod. Within the HumanIndexMod, we created a subroutine, QSat_2, which has all the same functionalities as QSatMod. This subroutine uses the August-Roche-Magnus (ARM) equation (Eq. A.13), rather than the Flatau et al. (1992) polynomial equations for vapor pressure in QSatMod. The log derivative of ARM (Eq. A.15) is a critical component of the calculation of T_w, and is not available in QSatMod. Additionally, QSat_2 calculates $f(\theta_E)$ (Eq. A.18) with respect to the input temperature, and the subsequent derivatives. These are required to calculate T_w (Eq. A.22) using Davies-Jones (2008), and cannot be accomplished using QSatMod. We show

JRBuzan 12/2/14 3:51 PM
Deleted: the improvement of

JRBuzan 12/2/14 3:51 PM
Deleted: quantities regarding water

JRBuzan 12/2/14 3:51 PM
Deleted: author

JRBuzan 12/2/14 3:51 PM
Deleted: improving water vapor calculations

JRBuzan 12/2/14 3:51 PM
Deleted: As previously mentioned in Section 2.2, moist

JRBuzan 12/2/14 3:51 PM
Deleted: that

JRBuzan 12/2/14 3:51 PM
Deleted: These algorithms calculate T_w using Davies-Jones (2008). We show acceptable differences between QSatMod and QSat_2, previously mentioned in section 2.2 (Fig. 1).

JRBuzan 12/2/14 3:51 PM
Deleted: ,

acceptable differences between the Stull version of wet bulb temperature (T_{wS}) calculated using both QSatMod and QSat_2 (Fig. 1a). The new subroutines improve CLM4.5 by calculating previously unused thermodynamic quantities. Additionally, these routines are useful moist thermodynamic routines for other datasets for researchers to use, thus expanding the capacity of the HumanIndexMod.

We implement specific thermodynamic routines developed by Davies-Jones (2008) to accurately calculate T_w (see Appendix A). Equation (A.4) is the most accurate and efficient θ_E calculation available (Bolton, 1980; Davies-Jones, 2009). Calculating Eq. (A.4) required implementing T_L and θ_{DL} (Eq. A.2 and A.3, respectively) into the HumanIndexMod. T, P, and Q from CLM4.5 are used to calculate θ_E and T_E (Eq. A.5). T_E, a quantity used in a previous heat stress study (Fischer and Knutti, 2012), is an input into QSat_2 for calculating the initial guess of T_w, and subsequently followed by the accelerated Newton-Raphson method (Eq. A.9-A.22). We found it advantageous to split the heat stress quantities into their own subroutines, allowing the user to choose what quantities to be calculated. The minimum requirements to execute the entire module are T (K), P (Pa), RH (%), Q (g/kg), e (Pa), and u_{10m} (m/s). Table 4 shows the subroutines, input requirements, and outputs in HumanIndexMod.

JRBuzan 12/2/14 3:51 PM
Deleted: uncalculated

JRBuzan 12/2/14 3:51 PM
Deleted: quantities, thus creating new opportunities

JRBuzan 12/2/14 3:51 PM
Deleted: future

JRBuzan 12/2/14 3:51 PM
Deleted: replace QSatMod with QSat_2

JRBuzan 12/2/14 3:51 PM
Deleted: all of the

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Deleted: the

3.3 CLM4.5 experimental setup

CLM4.5 may be executed independently of the other models in CESM, called an I-Compset. To do so, CLM4.5 requires atmospheric boundary conditions. We use the default dataset for CLM4.5—CRUNCEP. CRUNCEP is the NCEP/NCAR reanalysis product (Kalnay et al., 1996) corrected and downscaled by the Climatic Research Unit (CRU) gridded observations dataset from the University of East Anglia (Mitchell and Jones, 2005). The time period is 4x daily from 1901-2010, and is on a regular grid of ~0.5°x0.5°. The combination of CRU and NCEP products was to correct for biases in the reanalysis product, and improve overall resolution (Casado et al., 2013). To drive CLM4.5 we used surface solar radiation, surface precipitation rate, temperature, specific humidity, zonal and meridional winds, and surface pressure.

Our simulation has the carbon and nitrogen cycling on (biogeophysics 'CN'). The simulation was initialized at year 1850, on a finite volume grid of 1°x1°, using boundary conditions provided from NCAR (Sam Levis, personal communication). The simulation spun up while

JRBuzan 12/2/14 3:51 PM
Deleted: CLM4.5 was released in June, 2013, and the model has substantial improvements over previous versions—including improved urban canyon components, as well as new biogeochemical cycles (Oleson et al., 2013a).

JRBuzan 12/2/14 3:51 PM
Deleted: The simulation has the HumanIndexMod included.

cycling 3 times over CRUNCEP 1901-1920 forcings. Once completed, our experiment used the spun up land conditions, and ran the entirety of 1901-2010.

3.4 Heat stress indices analysis

An open question is what drives extreme high heat stress events, which are, by definition, rare events. For example, we cannot determine from the mean climate state or from theory, in a warm and humid climate, if abnormally high temperature, abnormally high moisture, or a combination of the two, caused a heat stress event. This is a question of the covariance of perturbations of temperature and humidity, not a statement of mean conditions, and there is no theory to explain these situations. For example, we may apply Reynolds averaging to the NWS Heat Index equation (Eq. 3):

$$\overline{HI} = a + \overline{b\overline{T}} + \overline{c\overline{RH}} + \overline{d\overline{T}\,\overline{RH}} + \overline{e\overline{T}^2} + \overline{f\overline{RH}^2} + \overline{g\overline{T}^2\overline{RH}} + \overline{h\overline{T}\,\overline{RH}^2} + \overline{\iota\overline{T}^2\overline{RH}^2} + \left[\overline{dRH'T'} + \overline{eT'^2} + \overline{fRH'^2} + \overline{gT'^2RH'} + \overline{hT'RH'^2} + \overline{\iota T'^2RH'^2}\right] \qquad (13)$$

where *a*, *b*, *c*, *d*, *e*, *f*, *g*, *h*, and *i* are constants in the polynomial. *RH* and *T* are relative humidity and temperature, respectively. We are not concerned with the terms outside the brackets, as they are the means. The terms within the bracket are representative of turbulent effects on the Heat Index, which we are discussing. It is these turbulent states where a GCM is able to determine these individual factors, by calculating the heat stress metrics and thermodynamic quantities at every model time step. Furthermore, each heat stress metric has different assumptions (such as body size, or physical fitness, etc.) that weigh temperature and humidity differently. A high heat stress event indicated by one metric does not necessarily transfer onto another metric.

Thus, we outputted 4x daily averages of the heat stress metrics and the corresponding surface pressure (P), 2-meter temperature (T), 10-meter winds (u_{10m}), and 2-meter humidity (Q) fields. We computed statistics for the time series (mean, variance, exceedance, etc.). We focus primarily on the 99th percentiles (hottest 1606 six hour intervals, ~402 days), but also show some of the robust features with the 75th (hottest 40,150 six hour intervals, ~10,038 days) and 95th percentiles (hottest 8030 six hour intervals, ~2008 days).

Every 6-hour period that exceeds the percentiles was located within the time series, and we calculated the conditional distributions. For example, the 99th percentile exceedance of HI isolated the top 1606 hottest time steps for each grid cell. After isolating these time steps, we

JRBuzan 12/2/14 3:51 PM
Deleted: We

JRBuzan 12/2/14 3:51 PM
Deleted: The 4x daily files are compiled into yearly files with 1460 time steps, and a concatenated file from 1901-2010.

JRBuzan 12/2/14 3:51 PM
Deleted: The 99th percentile between different heat stress metrics may not involve the same *T*, *P*, and *Q* combinations. To quantify these differences requires analyzing the original inputs used to calculate the heat stress metric.

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Deleted: within

JRBuzan 12/2/14 3:51 PM
Deleted: 99th

JRBuzan 12/2/14 3:51 PM
Deleted: were

JRBuzan 12/2/14 3:51 PM
Deleted: joint distribution.

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: in

JRBuzan 12/2/14 3:51 PM
Deleted: latitude by longitude. After calculating the joint distribution

use this distribution as a mask to isolate all other quantities (e.g., temperature and humidity), allowing cross comparison between all metrics and HI. The goal is to develop an analysis technique comparing all covariances of the metrics within CLM4.5.

After the conditional distributions are calculated, we, again, compute the statistical dispersion (mean, variance, exceedance, etc.) of the percentiles. We display this analysis with maps in two ways. 1) We show the exceedance value of a metric, and 2) we show *T*-*Q* regime plots of that same metric. We calculate the *T*-*Q* regimes through expected rank values (Fig. 2). This required a series of steps. 1) We take the conditional distribution of *T* and *Q* that represent exceedance percentile of the source heat stress or moist thermodynamic metric. 2) We take the expected value (median) of the conditional distributions of *T* and *Q* and determine what percentile they come from in their respective time series. 3) We condition these values on each other to create the expected rank values (Fig. 2).

4 Results

We present a snap shot of the many metrics calculated. First, we present results of our evaluation the improved moist thermodynamic calculations and the implementation these metrics into CLM4.5 (Fig. 1). Second, we show an example of the possible global applications for these metrics (Fig. 3-6). This approach characterizes heat stress within CLM4.5 in response to one observation reanalysis product, the CRUNCEP.

4.1 Evaluation of improved moist thermodynamic quantities

We present a series of box and whisker plots demonstrating the value added of implementing 1) accurate and efficient moist thermodynamic quantities, and 2) online calculation of the heat stress metrics is an improvement over calculating these metrics using monthly or 4x daily model output (Fig. 1). Figure 1a shows the difference in the Stull (2011) wet bulb temperature calculated using the saturated vapor pressure from Davies-Jones (2008) (QSat_2) and Flatau et al. (1992) (QSatMod). The differences are minimal. However, our point is that the Davies-Jones (2008) method for wet bulb temperature is preferred. We show the difference between wet bulb temperatures using Stull (2011) calculated with QSat_2, and Davies-Jones (2008) (which requires QSat_2) (Fig. 1b). Differences are greater than 1K between Stull (2011) and Davies-Jones (2008) methods, and they are temperature dependent (Fig. 1b). Lastly, we show the difference between calculating Davies-Jones (2008) T_w using

JRBuzan 12/2/14 3:51 PM
Deleted: the time domain

JRBuzan 12/2/14 3:51 PM
Deleted: ,

JRBuzan 12/2/14 3:51 PM
Deleted: was

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: heat stress

JRBuzan 12/2/14 3:51 PM
Deleted: joint

JRBuzan 12/2/14 3:51 PM
Deleted: 99th

JRBuzan 12/2/14 3:51 PM
Deleted: developed two methods of displaying the output for visual

JRBuzan 12/2/14 3:51 PM
Deleted: point comparisons. These maps are the 99th percentiles of the metrics and the medians

JRBuzan 12/2/14 3:51 PM
Deleted: their joint distribution. We selected a variety of regional city localities around the world to demonstrate latitudinal and regional influences on heat stress (Table 5). These locations were chosen due to high concentrations of people, or unique environments (i.e. deserts, coastal, monsoons, etc.). Due to the variance that an individual grid cell may have, we averaged the statistical dispersion (mean, variance, exceedance, etc.) information of all 8 nearest local grid cells together.

JRBuzan 12/2/14 3:51 PM
Deleted: plotted the joint distributions as box and whisker diagrams. The 25th and 75th percentiles are the box edges, the

JRBuzan 12/2/14 3:51 PM
Deleted: as the horizontal bar, the lower whisker as 5th, and the 90th as the upper whisker (the upper tail is discussed in section 5). The following section displays some

JRBuzan 12/2/14 3:51 PM
Deleted: results and their characterization within CLM4.5.

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [8]

JRBuzan 12/2/14 3:51 PM
Deleted: We

JRBuzan 12/2/14 3:51 PM
Deleted: . Additionally, we break down the analysis into two sections: joint distribution maps and box plots.

JRBuzan 12/2/14 3:51 PM
Deleted: <#>Joint Distribution ... [90]

monthly and 4x daily averaged model data versus the model instantaneous calculations (Fig. 1c and 1d, respectively). Using model-averaged data instead of the instantaneous data systematically overestimates T_w by more than 1K for monthly and 0.5K for 4x daily output.

4.2 Exceedance values and regime maps

We show exceedance and T-Q regime maps for the 75th and 95th percentiles of 3 metrics, and 99th percentiles of 6 metrics. The maps show spatial patterns of heat stress and characteristics. Equatorial and monsoonal regions show moderate levels of heat stress in the 75th percentile (Fig. 3a-c). sWBGT shows values exceeding minimum metric warning levels (e.g. China, Northern Africa), whereas HI does not have necessarily the same warning. The 95th percentile shows that moderate levels of heat stress have expanded into higher latitudes (Fig. 4a-c). At equatorial and monsoonal regions, heat stress labor reductions should be in effect as it is not safe to work outside, and in some cases (West Africa, the Arabian Peninsula, and the Himalayan Wall), no work at all. At the 99th percentile, severe heat stress is experienced in the monsoonal regions (Fig. 5a-c). These maxima correlate with maxima in T_w (Fig. 5c).

The T-Q regime maps show that partitioning of heat stress into T and Q begins in regional locations at the 75th percentile (Fig. 3d-f). The partitioning occurs in low latitudes, and is not consistent between metrics. At the 95th percentile, the partitioning expands into higher latitudes, however, many areas (continental interiors) remain equally dependent on T and Q (Fig. 4d-f). T_w is largely driven by extreme moisture (Fig. 4f) and in some locations (monsoonal Africa, Indian sub-continent, and equatorial South America) very extreme moisture. HI is driven by T (Fig. 4e), and sWBGT is mixed between extreme Q and extreme T (Fig. 4d). All three metrics agree with T in the Western United States and Middle East. At the 99th percentile, HI, although dominated by T worldwide, shows sign reversals in very small locations (Fig. 5e). Extreme Q expands for T_w, and all of the low latitudes experience moisture dependence except for Western United States and Middle East (Fig. 5f). sWBGT has some reversal of T to Q dominated heat stress (Western Africa). Q largely expands worldwide. In all instances, except for HI, high latitudes are equally dependent on Q and T for heat stress.

Our final maps show SWMP65, SWMP80, and θ_E at the 99th percentile. Maxima for θ_E are spatially the same as T_w (Fig. 5c and 6c). Additionally, θ_E partitions towards Q, just as T_w

shows (Fig. 5f and 6f). Spatial patterns between SWMP65 and HI are similar (Fig. 5b and 6a), and their regime maps show similar partitioning toward T globally, except for select locations of strong monsoonal locations that show Q dependency (Fig. 5e and 6d). Lastly, SWMP80 and sWBGT share similar spatial patterns (Fig. 5a and 6b). As with the other paired metrics, their T-Q regime maps share the same characteristics (Fig. 5d and 6e). Low latitudes show strong Q dependence, and higher latitudes switch to a T dependence.

5 Discussion

We designed the HumanIndexMod to calculate diagnostic heat stress and moist thermodynamics systematically. There are many approaches to evaluating heat stress. Monthly and seasonal temperature and moisture averages were used for general applications (Dunne et al., 2013), however these averages overestimate the potential severity of heat stress (Fig. 1c,d). Even using daily or sub-daily averages (Kjellstrom et al., 2009b; Hyatt et al., 2010; Fischer and Schar, 2010; Fischer and Knutti, 2012; Willett and Sherwood, 2012; Kjellstrom et al., 2013) potentially overestimates heat stress. This is due to the non-linear covariance of T and Q, and averages miss these extremes. Ultimately, capturing the diurnal cycle is crucial for quantifying heat stress extremes (Oleson et al., 2013b). Heat stress related illness is exacerbated by high heat stress nights as well as daytimes. To accurately calculate these extremes, one needs either high temporal resolution data, or directly computing them at each time step within climate models. We discuss the results from the implementation separately: moist thermodynamics and heat stress.

5.1 Moist thermodynamics

The spatial distributions of high heat stress are robust between CLM model versions (Oleson et al., 2011; Fischer et al., 2012; Oleson et al., 2013b). Due to the conservation of energy and entropy, calculating moist thermodynamic variables shows that climate models and reanalysis fall along constant lines of T_E (Eq. A.5), even out to the 99th percentile of daily values (Fischer and Knutti, 2012). The spread between models is small as compared to the spread in T, thus using heat stress metrics in Earth system modeling may reduce the uncertainties of climate change (Fischer and Knutti, 2012).

Previous modeling studies have demonstrated that urban equatorial regions transition to a nearly permanent high heat stress environment when considering global warming (Fischer et

JRBuzan 12/2/14 3:51 PM
Deleted: Diffenbaugh et al., 2007;

JRBuzan 12/2/14 3:51 PM
Deleted: underestimate

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [10]

JRBuzan 12/2/14 3:51 PM
Formatted: Font color: Auto

JRBuzan 12/2/14 3:51 PM
Deleted: not shown).

JRBuzan 12/2/14 3:51 PM
Deleted: Recent research used daily values (Kjellstrom et al.,

JRBuzan 12/2/14 3:51 PM
Moved up [10]: 2009b; Hyatt et al.,

JRBuzan 12/2/14 3:51 PM
Deleted: monthly

JRBuzan 12/2/14 3:51 PM
Deleted: the

JRBuzan 12/2/14 3:51 PM
Formatted: Font color: Auto

JRBuzan 12/2/14 3:51 PM
Deleted: 2010), but ultimately

JRBuzan 12/2/14 3:51 PM
Deleted: daytime

JRBuzan 12/2/14 3:51 PM
Deleted: by

JRBuzan 12/2/14 3:51 PM
Deleted: HumanIndexMod has applications beyond just human heat stress. For example, these metrics maybe used as a model diagnostic. The high Q values in the arid regions appear to be unreasonable because their absolute maximum T_w are equivalent to the values found in the centers of hurricanes (Zhang et al., 2002; Smith and Montgomery, 2012) (we omit those CLM4.5 values in our results with the top whisker limited to the 90th percentile). Upon further analysis, we believe there are two reasons. 1) the BareGroundFuxesMod fluxes are calculated at the surface, not at 2-meter height as with the CanopyFluxesMod, SlakeFluxesMod, and UrbanMod modules. This could cause high anomalous water quantities to be interpolated to the 2m height. 2) the sand ... [91]

JRBuzan 12/2/14 3:51 PM
Moved down [11]: Characterizin ... [92]

JRBuzan 12/2/14 3:51 PM
Deleted: The features

JRBuzan 12/2/14 3:51 PM
Deleted: Using heat stress metrics ... [93]

JRBuzan 12/2/14 3:51 PM
Deleted: heat stress

JRBuzan 12/2/14 3:51 PM
Deleted: ,

al., 2012; Oleson et al., 2013b). The convective regions are areas with the highest heat stress maximums and are often near coastal locations. Many of these metropolitan areas are in monsoonal regions, which have strong yearly moisture variability, yet the partitioning of heat stress is towards Q, not T, in these regions (Fig. 5d-f and 6d-f). Heat stress in both equatorial and monsoonal regions is expected to increase dramatically when considering global warming (Kjellstrom et al., 2009b; Fischer and Knutti, 2012; Dunne et al., 2013; Oleson et al., 2013b). Accurate moist thermodynamic calculations from the HumanIndexMod will aid future characterizations of heat stress.

5.2 Heat stress

We show that there are two regimes of heat stress globally in agreement between metrics in the CRUNCEP CLM4.5 simulation, T (Western United States and Middle East) and Q (monsoonal regions). Western United States and Middle East regions consistently have higher temperatures and lower humidities than the monsoonal areas. However, we show that maximum heat stress is partitioned between T and Q globally. Characterizing arid regions versus non-arid regions may require different heat stress metrics (e.g. Oleson el al., 2013b, specifically the comparison between Phoenix and Houston). The HumanIndexMod provides this capability.

The assumptions/calibrations that derived the heat stress metrics in the HumanIndexMod are another avenue of research that may be explored using a global model. For example, the original equation that sWBGT was derived from was calibrated using US Marine Corps Marines during basic training (Minard et al., 1957), who are in top physical condition. HI was calibrated for an 'average' American male (Steadman, 1979a; Rothfusz, 1990). Calculating these heat stress metrics, and the many others in the HumanIndexMod, at every time step within climate models were previously intractable due to insufficient data storage capabilities for high temporal resolution variables. We show that SWMP65 and SWMP80 diverge in their values (Fig. 6a,b and 6d,e). Yet, SWMP80 and sWBGT are similar in spatial patterns and regimes, while HI and SWMP65 have similar patterns and regimes. What links SWMP65 and SWMP80 together is T_w. Swamp coolers are evaporators, and as their efficiency approaches 100%, their solutions approach T_w. Figures 5 and 6 are similar to a circuit resistor, or stomatal resistance (Oke, 1987), which is measure of efficiency. The 'average' person (HI) may be acting as a stronger resistor to evaporation than one that is acclimatized (sWBGT). The HumanIndexMod may explore the effects of acclimatization,

JRBuzan 12/2/14 3:51 PM
Deleted: (Fig. 4, red)

JRBuzan 12/2/14 3:51 PM
Deleted: .

JRBuzan 12/2/14 3:51 PM
Deleted: Calculating multiple

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [11]

JRBuzan 12/2/14 3:51 PM
Deleted: at every time step within climate models opens up avenues

JRBuzan 12/2/14 3:51 PM
Deleted: were previously intractable due to insufficient data storage capabilities for high temporal resolution variables.

JRBuzan 12/2/14 3:51 PM
Moved down [12]: We show that SWMP65 and SWMP80 diverge in their values (Fig.

JRBuzan 12/2/14 3:51 PM
Deleted: 5d, 7d and 5f, 7f, respectively). *SWMP80* and *sWBGT* are similar in patterns with T_w while *HI* and *SWMP65* have similar patterns. These relationships

JRBuzan 12/2/14 3:51 PM
Deleted: related to the assumptions that were used to derive *sWBGT* and *HI*. *WBGT*, where

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Deleted: ,

JRBuzan 12/2/14 3:51 PM
Deleted: for

JRBuzan 12/2/14 3:51 PM
Formatted: Font:Not Italic

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [12]

JRBuzan 12/2/14 3:51 PM
Deleted: This is

JRBuzan 12/2/14 3:51 PM
Deleted: . An avenue of research that may be explored through climate modeling using the

JRBuzan 12/2/14 3:51 PM
Deleted: is

and its impact on efficiency of evaporative cooling through climate modeling. This type of research may ultimately reduce the number metrics required for computing heat stress.

Exposure to high moist temperatures, ultimately, threatens humans physically, and long-term exposure may lead to death. Extreme moist temperatures are projected to increase in the future, and potentially may reach deadly extremes, permanently in some regions (Sherwood and Huber, 2010). Heat stress indices have the ability to diagnose instantaneous exposure. Diagnostic models, however, cannot measure or evaluate the potential impacts of long-term exposure to heat stress accurately. Prognostic thermal physiological models can be used to predict the complexities of heat stress on humans.

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Deleted: Evaluating

JRBuzan 12/2/14 3:51 PM
Deleted:

JRBuzan 12/2/14 3:51 PM
Deleted: , however, cannot be measured

JRBuzan 12/2/14 3:51 PM
Deleted: by diagnostic models

Prognostic thermal physiology considers wind, ambient temperature, and moisture from the environment, as well as internal processes, such as blood flow and sweat. There are numerous different forms of prognostic models (Table 1). Some of them are quite complicated, using hundreds of grid cells to represent all parts of the body (Fiala et al., 1999). Less complicated models represent the human body as a single cylinder with multiple layers (Kraning and Gonzalez, 1997). Neither computational method is currently coupled to Earth system models, and this is a significant gap in determining future heat stress impacts that the HumanIndexMod may not be able to fulfill. To make progress towards representing the effects of heat stress on the human body prognostically, we recommend, as a first step, incorporating mean radiant temperature of humans. Radiation is a major component of human energy balance, and implementing this also allows incorporating more accurate diagnostics, such as Wet Bulb Globe Temperature (Minard et al., 1957) and the Universal Thermal Climate Index (Havenith et al., 2012).

JRBuzan 12/2/14 3:51 PM
Deleted: *WBGT* and *UTCI*.

6 Summary

We present the HumanIndexMod that calculates 9 heat stress metrics and 4 moist thermodynamical quantities. The moist thermodynamic variables use the latest accurate and efficient algorithms available. The heat stress metrics cover three developmental philosophies: comfort, physiological, and empirically based algorithms. The code is designed, with minimal effort, to be implemented into general circulation, land surface, and weather forecasting models. Additionally, this code may be used with archived data formats and local weather stations.

Furthermore, we have implemented the HumanIndexMod into the latest public release version of CLM4.5. Archival is flexible, as the user may choose to turn on high frequency output, and the default is monthly averages. Additionally, monthly urban and rural output of the metrics is default. We show that the module may be used to explore new avenues of research: characterization of human heat stress, model diagnostics, and intercomparisons of heat stress metrics. Our results show that there are two regimes of heat stress, extreme moisture and extreme temperature, yet all of the most extreme heat stress events are tied to maximum moisture.

JRBuzan 12/2/14 3:51 PM
Deleted: arid

JRBuzan 12/2/14 3:51 PM
Deleted: non-arid

JRBuzan 12/2/14 3:51 PM
Deleted: is

JRBuzan 12/2/14 3:51 PM
Deleted: maximum temperatures with

Our approach has limitations. None of the metrics in the HumanIndexMod include the effects of solar and thermal radiation. Radiation is a non-negligible component of heat stress. As a consequence, the heat stress metrics presented always assume that the subject is not in direct solar exposure. Additionally, the indices represent a diagnostic environment for heat stress. These metrics do not incorporate prognostic components or complex physiology of the human thermal system.

Overall, the HumanIndexMod provides a systematic way for implementing an aspect of thermo-animal physiology into an Earth system modeling framework. Incorporating the HumanIndexMod into a variety of different models would provide a baseline for model-model comparisons of heat stress, such as the Coupled Model Intercomparison Project (CMIP) (Taylor et al., 2012) and other collaborative modeling frameworks. We encourage researchers to incorporate the HumanIndexMod within their research environments.

Code Accessibility

We will make the HumanIndexMod available at the University of New Hampshire Data Discover Center New Hampshire Climate section. The NSF-funded New Hampshire EPSCoR Ecosystem and Society Project manage this data archive. Additionally, we will upload the HumanIndexMod to Data.gov, a free repository for data, metrics, and results for public use. The United States Government manages this repository.

Appendix A: Moist Thermodynamics

Davies-Jones (2008) shows multiple methods of computing T_w, and we implemented the most accurate equations, described below. We introduce terminology to describe the Davies-Jones

(2008) calculation. All temperature subscripts that are capitalized are in Kelvin, while lower case are in Celsius. κ_d is the Poisson constant for dry air (0.2854), and λ is the inverse (3.504). Many of the following equations are scaled using non-dimensional pressure (also known as the Exner function), π:

$$\pi = (p/p_0)^{1/\lambda} \tag{A.1}$$

where p is the pressure (mb), and p_0 is a reference pressure (1000mb).

To define T_w (the wet bulb temperature), we solve for the equivalent potential temperature, θ_E. Determining θ_E is a three-step process. First, we solve for the lifting condensation temperature (T_L):

JRBuzan 12/2/14 3:51 PM
Deleted:

$$T_L = \frac{1}{\frac{1}{T-55} - \frac{\ln(RH/100)}{2840}} + 55 \tag{A.2}$$

where T is the parcel temperature (Kelvin). For example, we use the 2 m air temperature in CLM4.5. RH (%) is taken at the same height as T. T_L (Eq. A.2), from Eq. (22) Bolton (1980), is the temperature at which a parcel that is lifted, following a dry adiabatic lapse rate, begins to condense. Second, as the air rises further, the parcel now follows a moist potential temperature, θ_{DL}:

$$\theta_{DL} = T\left(\frac{p_0}{p-e}\right)^{\kappa_d}\left(\frac{T}{T_L}\right)^{0.00028r} \tag{A.3}$$

where e is the parcel vapor pressure (mb) (using CLM4.5, this is the 2 m vapor pressure), and r is the mixing ratio (g/kg) (this is converted from the 2 m height Q to r in CLM4.5). Third, the parcel is raised to a great height where all latent heat is transferred to the air parcel, and the water is rained out, giving the solution to θ_E. There are many methods for representing this process. The analytical solution (Holton, 1972) is computationally prohibitive in atmospheric and land surface models. There are various approximations of different aspects of potential and saturated temperatures to calculate θ_E (Betts and Dugan, 1973; Simpson, 1978), however, many of them have large errors. These errors are compared in Bolton (1980), and Eq. (39) (Bolton's formulation) is up to an order of magnitude more accurate:

$$\theta_E = \theta_{DL}\exp\left[\left(\frac{3.036}{T_L} - 0.001788\right)r(1 + 0.000448r)\right] \tag{A.4}$$

Equivalent temperature, T_E, is θ_E scaled by π:

$$T_E = \theta_E \pi \tag{A.5}$$

The initial guess for T_w is based upon regions where the second order derivative of θ_E reaches a linear relationship with variations in T_w and λ. Two coefficients are derived (Davies-Jones, 2008):

$$k1 = -38.5\pi^2 + 137.81\pi - 53.737 \quad \text{(A.6)}$$

$$k2 = -4.392\pi^2 + 56.831\pi - 0.384 \quad \text{(A.7)}$$

The initial guess of T_w for coldest temperatures:

$$T_w = T_E - C - \frac{Ar_s(T_E,\pi)}{1 + Ar_s(T_E,\pi)\frac{\partial \ln(e_s)}{\partial T_E}} \quad \text{(A.8)}$$

where C is freezing temperature, A is a constant (2675), and r_s is the saturated mixing ratio. The evaluation of errors at a various pressures necessitated that Davies-Jones develop a regression line on colder regions of the initial guess:

$$\left({}^{C}/_{T_E}\right)^{\lambda} > D(\pi);\ D = \left(0.1859\frac{p}{p_0} + 0.6512\right)^{-1} \quad \text{(A.9)}$$

where D is calculating transition points between quadratic fits to the second order derivatives of θ_E. T_w for all other temperature regimes is governed by:

$$T_w = k1(\pi) - 1.21cold - 1.45hot - (k2(\pi) - 1.21cold)\left({}^{C}/_{T_E}\right)^{\lambda} + \left(\frac{0.58}{\left({}^{C}/_{T_E}\right)^{\lambda}}\right)hot \quad \text{(A.10)}$$

$$cold\begin{cases} = 0: 1 \leq \left({}^{C}/_{T_E}\right)^{\lambda} \leq D(\pi) \\ \quad = 1 \end{cases} \quad \text{(A.11)}$$

$$hot\begin{cases} = 1: T_E > 355.15 \\ \quad = 0 \end{cases} \quad \text{(A.12)}$$

where the combination of equations' initial guesses are valid from 1050mb down to 100mb. Following the initial guess, up to two iterations using the Newton-Raphson method are required to reach the true wet bulb temperature. Using T_W, saturation vapor pressure is solved by the August-Roche-Magnus formulation of the Clausius-Clayperon equation (Bolton, 1980; Lawrence, 2005):

$$e_s(T_W) = 6.112\exp\left(\frac{a(T_W - C)}{T_W - C + b}\right) \quad \text{(A.13)}$$

where e_s is in mb, a and b are constants. The saturation mixing ratio, r_s, is dependent on e_s:

$$r_s(T_W) = \frac{\varepsilon e_s(T_W)}{\left(p_0\pi^\lambda - e_s(T_W)\right)} \tag{A.14}$$

where ε is a constant (~0.622). Following Davies-Jones, we use the derivative of ARM equation for calculating the derivative of r_s:

$$\frac{\partial \ln(e_s)}{\partial T_W} = \frac{ab}{(T_W - C + b)^2} \tag{A.15}$$

$$\frac{\partial e_s}{\partial T_W} = e_s \frac{\partial \ln(e_s)}{\partial T_W} \tag{A.16}$$

$$\left(\frac{\partial r_s}{\partial T_W}\right)_\pi = \frac{\varepsilon p}{\left(p - e_s(T_W)\right)^2}\frac{\partial e_s}{\partial T_W} \tag{A.17}$$

Now, we return to θ_E, and substitute T_W for T_L:

$$f(T_W;\pi) = \left(C/T_W\right)^\lambda \left[1 - \frac{e_s}{p_0\pi^\lambda}\right]^{\kappa_d\lambda} \exp\left(-\lambda G(T_W;\pi)\right) \tag{A.18}$$

where:

$$G(T_W;\pi) = \left(\frac{3036}{T_W} - 1.78\right)\left[r_s(T_W;\pi) + 0.448 r_s^2(T_W;\pi)\right] \tag{A.19}$$

The derivative of the function Eq. (A.18) is required for the Newton-Raphson method:

$$f'(T_W;\pi) = -\lambda\left[\frac{1}{T_W} + \frac{\kappa_d}{\left(p - e_s(T_W)\right)}\frac{\partial e_s}{\partial T_W} + \left(\frac{\partial G}{\partial T_W}\right)_\pi\right] \tag{A.20}$$

where the derivative of $G(T_W;\pi)$:

$$\left(\frac{\partial G}{\partial T_W}\right)_\pi = -\frac{3036\left(r_s(T_W) + 0.448 r_s^2(T_W)\right)}{T_W^2}\left(\frac{3036}{T_W} - 1.78\right)\left(1 + 2\left(0.448 r_s(T_W)\right)\right)\left(\frac{\partial r_s}{\partial T_W}\right)_\pi \tag{A.21}$$

and, due to the linear relationship of the second order derivative of Eq. (A.18), we may accelerate the Newton-Raphson method using the initially calculated T_W and T_E:

$$T_w = T_w - \frac{f(T_W;\pi) - \left(C/T_E\right)^\lambda}{f'(T_W;\pi)} \tag{A.22}$$

Acknowledgements

Computing resources from Information Technology at Purdue University supported this research. JRB thanks the reviewers for helpful and constructive comments. JRB thanks Aaron Goldner, Jacob Carley, and Nick Herold for helpful comments and support. He also thanks his friends and family for their continued, unwavering, support. K.W. Oleson acknowledges support from the NCAR WCIASP and from NASA grant NNX10AK79G (the

JRBuzan 12/2/14 3:51 PM
Deleted: This research was supported by computing

JRBuzan 12/2/14 3:51 PM
Deleted: Finally, JRB thanks his family.

SIMMER project). NCAR is sponsored by the National Science Foundation. MH acknowledges support from the NSF-funded New Hampshire EPSCoR Ecosystem and Society Project.

References

Adam-Poupart, A., LaBreche, F., Smargiassi, A., Duguay, P., Busque, M., Gagne, C., Rintamaki, H., Kjellstrom, T., and Zayed, J. Climate change and occupational health and safety in a temperate climate: potential impacts and research priorities in Quebec, Canada. Industrial Health, 51(1), 68-78, 2013.

Australian Bureau of Meteorology. About the WBGT and Apparent Temperature indices– Australian Bureau of Meteorology. http://www.bom.gov.au/info/thermal_stress/, 2014.

American College of Sports Medicine. Position stand on the prevention of thermal injuries during distance running. Medical Journal of Australia, 141, 876–879, 1984.

American College of Sports Medicine. Position stand on the prevention of thermal injuries during distance running. Medicine and Science in Sports and Exercise, 19(5), 529-533, 1987.

Alfano, F. R. D. A., Palella, B. I., and Riccio, G. On the Problems Related to Natural Wet Bulb Temperature Indirect Evaluation for the Assessment of Hot Thermal Environments by Means of WBGT. Annals of occupational hygiene, 56(9), 1063-1079, 2012.

Barriopedro, D., Fischer, E., Luterbacher, J., Trigo, R., García-Herrera, R. The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe. Science 332(6026) pp. 220-224, 2011.

Belding, H. S. and Hatch, T. F. Index for evaluating heat stress in terms of resulting physiological strain. Heating, piping and air conditioning, 27(8), pp. 129, 1955.

Benestad. A New Global Set of Downscaled Temperature Scenarios. Journal of Climate. vol. 24 (8) pp. 2080-2098, 2011.

Berglund, L. G., & Yokota, M. Comparison of human responses to prototype and standard uniforms using three different human simulation models: HSDA, Scenario_J and Simulink2NM (No. USARIEM-T05-08). Army Research Inst of Environmental Medicine Natick MA Biophysics and Biomedical Modeling Div, 2005.

Betts, A. K. and Dugan, F. J. Empirical formula for saturation pseudoadiabats and saturation equivalent potential temperature. Journal of Applied Meteorology, 12(4), 731-732, 1973.

Bolton, D. The computation of equivalent potential temperature. Monthly weather review, 108(7), 1046-1053, 1980.

Bonan, G. B., Levis, S., Kergoat, L., & Oleson, K. W. Landscapes as patches of plant functional types: An integrating concept for climate and ecosystem models. Global Biogeochemical Cycles, 16(2), 1021, 2002.

Brake, D. J. Calculation of the natural (unventilated) wet bulb temperature, psychrometric dry bulb temperature and wet bulb globe temperature from standard psychrometric measurements. J Mine Vent Soc S Afr, 54(108), 12, 2001.

Breckenridge, J. R., & Goldman, R. F. Solar heat load in man. Journal of Applied Physiology, 31(5), 659-663, 1971.

Bröde, P., Krüger, E. L., Rossi, F. A., & Fiala, D. Predicting urban outdoor thermal comfort by the Universal Thermal Climate Index UTCI—a case study in Southern Brazil. International journal of biometeorology, 56(3), 471-480, 2012.

Bröde, P., Blazejczyk, K., Fiala, D., Havenith, G., Holmer, I., Jendritzky, G., Kuklane, K., and Kampmann, B. The Universal Thermal Climate Index UTCI compared to ergonomics standards for assessing the thermal environment. Industrial Health, 51(1), 16-24, 2013.

Budd, G. M. Wet-bulb globe temperature (WBGT)—its history and its limitations. Journal of Science and Medicine in Sport, 11(1), 20-32, 2008.

Casado, M., Ortega, P., Masson-Delmotte, V., Risi, C., Swingedouw, D., Daux, V., Genty, D., Maignam, F. Solomina, O., Vinther, B., Viovy, N., & Yiou, P. Impact of precipitation intermittency on NAO-temperature signals in proxy records. Climate of the Past, 9(2), 871-886, 2013.

JRBuzan 12/2/14 3:51 PM
Deleted: Brooks,

JRBuzan 12/2/14 3:51 PM
Moved down [13]: R.

JRBuzan 12/2/14 3:51 PM
Deleted: H., and Corey, T. Hydraulic properties of porous media. Hydrology Papers, Colorado State University, Fort Colins, Colorado, 3, 1-30, 1964.
... [94]

Chan, A. P., Yam, M. C., Chung, J. W., & Yi, W. Developing a heat stress model for construction workers. Journal of Facilities Management, 10(1), 59-74, 2012.

Davies-Jones, R. An efficient and accurate method for computing the wet-bulb temperature along pseudoadiabats. Monthly Weather Review, 136(7), 2764-2785, 2008.

Davies-Jones, R. On formulas for equivalent potential temperature. Monthly Weather Review, 137(9), 3137-3148, 2009.

Diffenbaugh, N. S., Pal, J. S., Giorgi, F., & Gao, X. Heat stress intensification in the Mediterranean climate change hotspot. Geophysical Research Letters, 34(11), 2007.

Dufton, A. M. The eupatheostat. Journal of scientific instruments, 6(8), 249, 1929.

Dunne, J. P., Stouffer, R. J., & John, J. G. Reductions in labour capacity from heat stress under climate warming. Nature Climate Change, 2013.

Epstein, Y., & Moran, D. S. Thermal comfort and the heat stress indices. Industrial Health, 44(3), 388-398, 2006.

Fiala, D., Lomas, K. J., & Stohrer, M. A computer model of human thermoregulation for a wide range of environmental conditions: the passive system. Journal of Applied Physiology, 87(5), 1957-1972, 1999.

Fiala, D., Lomas, K. J., & Stohrer, M. Computer prediction of human thermoregulatory and temperature responses to a wide range of environmental conditions. International Journal of Biometeorology, 45(3), 143-159, 2001.

Fiala, D., Psikuta, A., Jendritzky, G., Paulke, S., Nelson, D., D. van Marken Lichtenbelt, W., Frijns., W. Physiological modeling for technical, clinical and research applications. Front Biosci S, 2, 939-968, 2010.

JRBuzan 12/2/14 3:51 PM
Deleted: ... [95]

Fiala, D., Havenith, G., Brode, P., Kampmann, B., and Jendritzky, G. UTCI-Fiala multi-node model of human heat transfer and temperature regulation. Int J Biometeorol, pp. 1-13, 2011.

Fischer, E. M., Seneviratne, S., Luthi, D., and Schar, C. Contribution of land-atmosphere coupling to recent European summer heat waves. Geophys. Res. Lett., 34(6), L06707, 2007.

Fischer, E. M., Oleson, K. W., & Lawrence, D. M. Contrasting urban and rural heat stress responses to climate change. Geophysical research letters, 39(3), 2012.

Fischer, E. M., & Knutti, R. Robust projections of combined humidity and temperature extremes. Nature Climate Change, 2012.

Fischer, E. M., and Schar, C. Consistent geographical patterns of changes in high-impact European heatwaves. Nature Geoscience, 3(6), 398, 2010.

Flatau, P. J., Walko, R. L., & Cotton, W. R. Polynomial fits to saturation vapor pressure. Journal of Applied Meteorology, 31, 1507-1507, 1992.

Gagge, A. P. An effective temperature scale based on a simple model of human physiological regulatory response. ASHRAE Trans., 77, 247-262, 1972.

García-Herrera, R., Diaz, J., Trigo, R. M., Luterbacher, J., and Fischer, E. M. A Review of the European Summer Heat Wave of 2003. Critical Reviews in Environmental Science and Technology, 40(4), 267-306, 2010.

Gates, R. S., Timmons, M. B., & Bottcher, R. W. Numerical optimization of evaporative misting systems. Transactions of the ASAE, 34, 1991.

Gates, R. S., Usry, J. L., Nienaber, J. A., Turner, L. W., & Bridges, T. C. An optimal misting method for cooling livestock housing. Transactions of the ASAE, 34, 1991.

Giles, B. D., Balafoutis, C., and Maheras, P. Too hot for comfort: the heatwaves of Greece in 1987 and 1988. International Journal of Biometeorology, 34(2), 98-104, 1990.

Gonzalez, R. R. SCENARIO revisited: comparisons of operational and rational models in predicting human responses to the environment. Journal of Thermal Biology, 29(7), 515-527, 2004.

Gonzalez, R. R., Cheuvront, S. N., Ely, B. R., Moran, D. S., Hadid, A., Endrusick, T. L., & Sawka, M. N. Sweat rate prediction equations for outdoor exercise with transient solar radiation. Journal of Applied Physiology, 112(8), 1300-1310, 2012.

Haslam, R. and Parsons, K. Using computer-based models for predicting human thermal responses to hot and cold environments. TERG, 37(3), 399-416, 1994.

Havenith, G., Fiala, D., Błazejczyk, K., Richards, M., Bröde, P., Holmér, I., Rintamaki, H., Benshabat, Y., & Jendritzky, G. The UTCI-clothing model. International journal of biometeorology, 56(3), 461-470, 2012.

Heat stress control and heat casualty management. Technical Bulletin Medical 507/Air Force Pamphlet, 48-152, US Army, 2003.

Höppe, P. The physiological equivalent temperature—a universal index for the biometeorological assessment of the thermal environment. International Journal of Biometeorology, 43(2), 71-75, 1999.

Houghton, F., Yaglou, C. Determining equal comfort lines. J Am Soc Heat Vent Eng., 29, 165–176, 1923.

Hurrell, J. W., Holland, M. M., Gent, P. R., Ghan, S., Kay, J. E., Kushner, P. J., Lamarque, J. F., Large, W. G., Lawrence, D., Lindsay, K., Lipscomb, W. H., Long, M. C., Mahowald, N., Marsh, D. R., Neale, R. B., Rasch, P., Vavrus, S., Vertenstein, M., Bader, D., Collins, W. D., Hack, J. J., Kiehl, J., and Marshall, S.: The Community Earth System Model: A Framework for Collaborative Research, B Am Meteorol Soc, 94, 1339-1360, doi:10.1175/bams-d-12-00121.1, 2013.

Hyatt, O. M., Lemke, B., & Kjellstrom, T. Regional maps of occupational heat exposure: past, present, and potential future. Global health action, 3, 2010.

Ingram, D. L. Evaporative cooling in the pig. Nature, (207), 415-416, 1965.

SREX IPCC. Summary for Policymakers. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, 1-19, 2012.

Jendritzky, G., & Tinz, B. The thermal environment of the human being on the global scale. Global Health Action, 2, 2009.

Jendritzky, G., Havenith, G., Weihs, P., & Batchvarova, E. Towards a Universal Thermal Climate Index UTCI for assessing the thermal environment of the human being. Final Report COST Action, 730, 2009.

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Leetmaa, A., Reynolds, R., Jenne, R., & Joseph, D. The NCEP/NCAR 40-year reanalysis project. Bulletin of the American meteorological Society, 77(3), 437-471, 1996.

Khan, Z. A., Maniyan, S., Mokhtar, M., Quadir, G. A., & Seetharamu, K. N. A Generalised Transient Thermal Model for Human Body. Jurnal Mekanikal, (18), 78-97, 2004.

Kjellstrom, T., Gabrysch, S., Lemke, B., and Dear, K. The ‘Hothaps’ programme for assessing climate change impacts on occupational health and productivity: an invitation to carry out field studies. Global Health Action, 2, 2009a.

Kjellstrom, T., Kovats, R. S., Lloyd, S. J., Holt, T., & Tol, R. S. The direct impact of climate change on regional labor productivity. Archives of Environmental & Occupational Health, 64(4), 217-227, 2009b.

Kjellstrom, T., Holmer, I., and Lemke, B. Workplace heat stress, health and productivity–an increasing challenge for low and middle-income countries during climate change. Global Health Action, 2, 2009c.

Kjellstrom, T., Lemke, B., and Otto, M. Mapping occupational heat exposure and effects in South-East Asia: ongoing time trends 1980-2011 and future estimates to 2050. Industrial Health, 51(1), 56-67, 2013.

Koca, R. W., Hughes, W. C., & Christianson, L. L. Evaporative cooling pads: test procedure and evaluation. Applied Engineering in Agriculture, 7, 1991.

Kraning, K. K., & Gonzalez, R. R. A mechanistic computer simulation of human work in heat that accounts for physical and physiological effects of clothing, aerobic fitness, and progressive dehydration. Journal of thermal biology, 22(4), 331-342, 1997.

Keuhn, L. A., Stubbs, R. A., & Weaver, R. S. Theory of the Globe Thermometer (No. DRET-RP-745). Defense Research Establishment Toronto Downsview (Ontario), 1970.

Lawrence, M. G. The relationship between relative humidity and the dewpoint temperature in moist air: A simple conversion and applications. Bulletin of the American Meteorological Society, 86(2), 225-233, 2005.

Lawrence, D. M., Oleson, K. W., Flanner, M. G., Thornton, P. E., Swenson, S. C., Lawrence, P. J., Zeng, X., Yang, Z.-L., Levis, S., Sakaguchi, K., Bonan, G. B., & Slater, A. G. Parameterization improvements and functional and structural advances in version 4 of the Community Land Model. Journal of Advances in Modeling Earth Systems, 3(1), 2011.

Liang, C., Zheng, G., Zhu, N., Tian, Z., Lu, S., Chen, Y. A new environmental heat stress index for indoor hot and humid environments based on Cox regression. Building and Environment, 46(12), 2472-2479, 2011.

Liljegren, J. C., Carhart, R. A., Lawday, P., Tschopp, S., and Sharp, R. Modeling the wet bulb globe temperature using standard meteorological measurements. Journal of Occupational and Environmental Hygiene, 5(10), 645-655, 2008.

Lucas, E. M., Randall, J. M., & Meneses, J. F. Potential for evaporative cooling during heat stress periods in pig production in Portugal (Alentejo). Journal of Agricultural Engineering Research, 76(4), 363-371, 2000.

Maloney, S. K., & Forbes, C. F. What effect will a few degrees of climate change have on human heat balance? Implications for human activity. International journal of biometeorology, 55(2), 147-160, 2011.

Masterson, J.M., Richardson, F.A. Humidex, a method of quantifying human discomfort due to excessive heat and humidity. Environment Canada, Atmospheric Environment Service, Downsview, Ontario, CLI 1-79, 1979.

Meehl and Tebaldi. More intense, more frequent, and longer lasting heat waves in the 21st century. Science, 305(5686), 994-7, 2004.

Minard, D., Belding, H. S., & Kingston, J. R. Prevention of heat casualties. Journal of the American Medical Association, 165(14), 1813-1818, 1957.

Miralles, D., Teuling, A., Van Heerwaarden, C., Arellano, J. Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. Nature Geosci, 7(5), 345-349, 2014.

Mitchell, T. D., & Jones, P. D. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. International journal of climatology, 25(6), 693-712, 2005.

Moran, D. S., Pandolf, K. B., Shapiro, Y., Heled, Y., Shani, Y., Mathew, W. T., & Gonzalez, R. R. An environmental stress index (ESI) as a substitute for the wet bulb globe temperature

JRBuzan 12/2/14 3:51 PM

Deleted: Meehl, G. A., Arblaster, J. M., Lawrence, D. M., Seth, A., Schneider, E. K., Kirtman, B. P., & Min, D. Monsoon regimes in the CCSM3. Journal of climate, 19(11), 2482-2495, 2006. ... [96]

(WBGT). Journal of thermal biology, 26(4), 427-431, 2001.

Mueller, B. and Seneviratne, S. Hot days induced by precipitation deficits at the global scale. Proceedings of the National Academy of Sciences, 109(31), 12398-12403, 2012.

Nag, P., Dutta, P., and Nag, A. Critical body temperature profile as indicator of heat stress vulnerability. Industrial Health, 51(1), 113-122, 2013.

Nilsson, M. and Kjellstrom, T. Climate change impacts on working people: how to develop prevention policies. Global health action, 3, 2010.

NOAAWatch: http://www.noaawatch.gov/themes/heat.php, last access 12 March, 2014.

Oke, T. R. Boundary Layer Climates, 2nd Edition. London. Methuen and Co. Chapter 4, 1987.

Oleson, K. W., Niu, G. Y., Yang, Z. L., Lawrence, D. M., Thornton, P. E., Lawrence, P. J., Stockli, R., Dickinson, R. E., Bonan, G. B., Levis, S., Dai, A., & Qian, T. Improvements to the Community Land Model and their impact on the hydrological cycle. Journal of Geophysical Research: Biogeosciences (2005–2012), 113(G1), 2008a.

Oleson, K. W., Bonan, G. B., Feddema, J., Vertenstein, M., & Grimmond, C. S. B. An urban parameterization for a global climate model. Part I: Formulation and evaluation for two cities. Journal of Applied Meteorology and Climatology, 47(4), 1038-1060, 2008b.

Oleson, K. W., Bonan, G. B., Feddema, J., & Vertenstein, M. An urban parameterization for a global climate model. Part II: Sensitivity to input parameters and the simulated urban heat island in offline Simulations. Journal of Applied Meteorology and Climatology, 47(4), 1061-1076, 2008c.

Oleson, K.W., D.M. Lawrence, G.B. Bonan, M.G. Flanner, E. Kluzek, P.J. Lawrence, S. Levis, S.C. Swenson, P.E. Thornton, A. Dai, M. Decker, R. Dickinson, J. Feddema, C.L. Heald, F. Hoffman, J.-F. Lamarque, N. Mahowald, G.-Y. Niu, T. Qian, J. Randerson, S.

Running, K. Sakaguchi, A. Slater, R. Stockli, A. Wang, Z.-L. Yang, Xi. Zeng, and Xu. Zeng. Technical Description of version 4.0 of the Community Land Model (CLM). NCAR Technical Note NCAR/TN-478+STR, National Center for Atmospheric Research, Boulder, CO, 257 pp, 2010a.

Oleson, K.W., G.B. Bonan, J. Feddema, M. Vertenstein, and Kluzek, E Technical Description of an Urban Parameterization for the Community Land Model (CLMU). NCAR Technical Note NCAR/TN-480+STR, DOI: 10.5065/D6K35RM9, 2010b.

Oleson, K. W., Bonan, G. B., Feddema, J., & Jackson, T. An examination of urban heat island characteristics in a global climate model. International Journal of Climatology, 31(12), 1848-1865, 2011.

Oleson, K. Contrasts between urban and rural climate in CCSM4 CMIP5 climate change scenarios. Journal of Climate, 25(5), 1390-1412, 2012.

Oleson, K.W., D.M. Lawrence, G.B. Bonan, B. Drewniak, M. Huang, C.D. Koven, S. Levis, F. Li, W.J. Riley, Z.M. Subin, S.C. Swenson, P.E. Thornton, A. Bozbiyik, R. Fisher, E. Kluzek, J.-F. Lamarque, P.J. Lawrence, L.R. Leung, W. Lipscomb, S. Muszala, D.M. Ricciuto, W. Sacks, Y. Sun, J. Tang, Yang, Z.-L. Technical Description of version 4.5 of the Community Land Model (CLM). Ncar Technical Note NCAR/TN-503+STR, National Center for Atmospheric Research, Boulder, CO, 422 pp, DOI: 10.5065/D6RR1W7M, 2013a.

Oleson, K. W., Monaghan, A., Wilhelmi, O., Barlage, M., Brunsell, N., Feddema, J., Hu, L., & Steinhoff, D. F. Interactions between urbanization, heat stress, and climate change. Climatic Change, 1-17, 2013b.

Parsons, K. Heat stress standard ISO 7243 and its global application. Industrial Health, 44(3), 368-379, 2006.

Parsons, K. Occupational health impacts of climate change: current and future ISO standards for the assessment of heat stress. Industrial Health, 51(1), 86-100, 2013.

Pradhan, B., Shrestha, S., Shrestha, R., Pradhanang, S., Kayastha, B., and Pradhan, P. Assessing climate change and heat stress responses in the Tarai region of Nepal. Industrial Health, 51(1), 101-12, 2013.

Renaudeau, D., Collin, A., Yahav, S., De Basilio, V., Gourdine, J. L., & Collier, R. J. Adaptation to hot climate and strategies to alleviate heat stress in livestock production. Animal, 6(05), 707-728, 2012.

Robine, J-M., Cheung, S. L. K., Roy, S. L., Oyen, H.Van, Griffiths, C., Michel, J-P, Herrmann, F. R. Death toll exceeded 70,000 in Europe during the summer of 2003. C. R. Biologies, 331,171–178, 2008.

Rothfusz, L. P., & Headquarters, N. S. R. The heat index equation (or, more than you ever wanted to know about heat index). Fort Worth, Texas: National Oceanic and Atmospheric Administration, National Weather Service, Office of Meteorology, 90-23, 1990.

Scholander, P., Hock, R., Walters, V., and Irving, L. Adaptation to cold in arctic and tropical mammals and birds in relation to body temperature, insulation, and basal metabolic rate. The Biological Bulletin, 99(2), 259-271, 1950.

Seneviratne, S.I., N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, and X. Zhang. Changes in climate extremes and their impacts on the natural physical environment. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press, Cambridge, UK, and New York, NY, USA, 109-230, 2012.

Seneviratne, S., Donat, M., Mueller, B., and Alexander, L. No pause in the increase of hot temperature extremes. Nature Climate Change. 4,161-163, 2014.

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [13]

JRBuzan 12/2/14 3:51 PM
Moved (insertion) [14]

Sheffield, P. E., Herrera, J. G. R., Lemke, B., Kjellstrom, T., & Romero, L. E. B. Current and Future Heat Stress in Nicaraguan Work Places under a Changing Climate. Industrial health, 51(1), 123-127, 2013.

Sherwood, S. C., & Huber, M. An adaptability limit to climate change due to heat stress. Proceedings of the National Academy of Sciences, 107(21), 9552-9555, 2010.

Simpson, R. H. On the computation of equivalent potential temperature. Monthly Weather Review, 106(1), 124-130, 1978.

Steadman, R. G. The assessment of sultriness. Part I: A temperature-humidity index based on human physiology and clothing science. Journal of Applied Meteorology, 18(7), 861-873, 1979a.

Steadman, R. G. The Assessment of Sultriness. Part II: Efiects of Wind, Extra Radiation and Barometric Pressure on Apparent Temperature. Journal of Applied Meteorology, 18, 874-884, 1979b.

Steadman, R. G. A universal scale of apparent temperature. Journal of Climate and Applied Meteorology, 23(12), 1674-1687, 1984.

Steadman, R. G. Norms of apparent temperature in Australia. Aust. Met. Mag, 43, 1-16, 1994.

Stoecklin-Marois, M., Hennessy-Burt, T., Mitchell, D., and Schenker, M. Heat-related illness knowledge and practices among California hired farm workers in the MICASA study. Industrial Health, 51(1), 47-55, 2013.

Stolwijk, J. A mathematical model of physiological temperature regulation in man. Report CR-1855 NASA, Washington, D.C, 1971.

Stolwijk, J. Mathematical models of thermal regulation. Annals of the New York Academy of Sciences 335(1) pp. 98-106, 1980.

JRBuzan 12/2/14 3:51 PM
Deleted: Smith,

JRBuzan 12/2/14 3:51 PM
Moved up [14]: R.

JRBuzan 12/2/14 3:51 PM
Deleted: K. and Montgomery, M. T. How important is the isothermal expansion effect in elevating equivalent potential temperature in the hurricane inner core?. Quarterly Journal of the Royal Meteorological Society, 201 ... [97]

Stolwijk, J. and Hardy, J. Temperature regulation in man-A theoretical study. Ptlilgers Archiv, 129, 129-162, 1966.

Stull, R. Wet-bulb temperature from relative humidity and air temperature. Journal of Applied Meteorology and Climatology, 50(11), 2267-2269, 2011.

Tawatsupa, B., Yiengrugsawan, V., Kjellstrom, T., Berecki-Gisolf, J., Seubsman, S., and Sleigh, A. Association between heat stress and occupational injury among Thai workers: Findings of the Thai Cohort Study. Industrial Health, 51(1), 34-46, 2013.

Taylor, K. E., Stouffer, R. J., & Meehl, G. A. An overview of CMIP5 and the experiment design. Bulletin of the American Meteorological Society, 93(4), 485-498, 2012.

Thom, E. C. The discomfort index. Weatherwise, 12(2), 57-61, 1959.

Wexler, A. Vapor pressure formulation for water in range 0 to 100 Degrees C. A revision. Journal of Research of the National Bureau of Standards–A. Physics and Chemistry, 80A(5), 1976.

Wexler, A. Vapor pressure formulation for ice. Journal of Research of the National Bureau of Standards–A. Physics and Chemistry, 81(1), 5-19, 1977.

Willett, K. M., & Sherwood, S. Exceedance of heat index thresholds for 15 regions under a warming climate using the wet‐bulb globe temperature. International Journal of Climatology, 32(2), 161-177, 2012.

Wyndham, C. and Atkins, A. A physiological scheme and mathematical model of temperature regulation in man. Pflügers Archiv. 303(1) pp. 14-30, 1968.

Yokota, M., Berglund, L., Cheuvront, S., Santee, W., Latzka, W., Montain, S., Kolka, M., and Moran, D. Thermoregulatory model to predict physiological status from ambient environment and heart rate. Computers in biology and medicine, 38(11), 1187-1193, 2008.

JRBuzan 12/2/14 3:51 PM
Deleted: van Genuchten, M. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Science Society of America Journal, 44(5), 892-898, 1980. ... [98]

JRBuzan 12/2/14 3:51 PM
Formatted: Space Before: 0 pt, No widow/orphan control, Don't adjust space between Latin and Asian text, Don't adjust space between Asian text and numbers, Tabs: 0.39", Left + 0.78", Left + 1.17", Left + 1.56", Left + 1.94", Left + 2.33", Left + 2.72", Left + 3.11", Left + 3.5", Left + 3.89", Left + 4.28", Left + 4.67", Left

JRBuzan 12/2/14 3:51 PM
Deleted: &

JRBuzan 12/2/14 3:51 PM
Deleted: ... [99]

Table 1. Heat stress diagnostics and prognostic models.

Metric	Type	Ref.
Wet Bulb Temperature	Temperature	Haldane (1905)
Effective Temperature	Index	Houghton and Yaglou (1923)
Equivalent Temperature	Temperature	Dufton (1929)
Heat Stress Index	Index	Belding and Hatch (1955)
Wet Bulb Globe Temperature	Index	Yaglou and Minard (1957)
Discomfort Index	Index	Thom (1959)
Temperature Humidity Index	Index	Ingram (1965)
Temp. Regulation in Man	Prognostic	Stolwijk and Hardy (1966)
Physiological Mathematical Model	Prognostic	Wyndham and Atkins (1968)
Solar Heat in Man	Index	Breckenridge and Goldman (1971)
Mathematical Model Temperature in Man	Prognostic	Stolwijk (1971)
New Effective Temperature	Index	Gagge (1972)
Humidex	Index	Masterson and Richardson (1979)
Sultriness Index	Index	Steadman (1979a)
Mathematical Model Thermal Regulation	Prognostic	Stolwijk (1980)
Apparent Temperature	Index	Steadman (1984)
Heat Index	Index	Rothfusz (1990)
Computer Based Thermal Response	Prognostic	Haslam and Parsons (1994)
SCENARIO	Prognostic	Kraning and Gonzalez (1997)

Computer Model Human Thermo-Regulation	Prognostic	Fiala et al. (1999); Fiala et al. (2001)
PET	Index	Höppe (1999)
Environmental Stress Index	Index	Moran et al. (2001)
SCENARIO Monte Carlo	Prognostic	Gonzalez (2004)
Generalized Transient Thermal Model	Prognostic	Khan et al. (2004)
ISO 7243 WBGT	Index	Parsons (2006)
IDCA	Prognostic	Yokota et al. (2008)
Physiological Equivalent Temperature	Index	Jendritzky et al. (2009)
UTCI	Index	Fiala et al. (2010)
UTCI-Fiala Model	Index-Prognostic	Fiala et al. (2011)
Index of Equivalent Temperature	Index	Liang et al. (2011)

Table 2. Moist temperature variables and heat stress metrics.

Metric	Variable	Equation #
Temperature (Kelvin)	T	N/A
Temperature (Celsius)	T_C	N/A
Temperature (Fahrenheit)	T_F	N/A
Pressure	P	N/A
Relative humidity	RH	N/A
Specific humidity	Q	N/A
10 m Winds	u_{10m}	N/A
Vapor Pressure (mb)	e_{RH}	2
Vapor Pressure (Pa)	e_{sPa}	N/A
Saturated vapor pressure (mb)	e_s	A.13
Derivative saturated vapor pressure	de_s/dT	A.16
Log derivative saturated vapor pressure	$d(ln(e_s))/dT$	A.15

JRBuzan 12/2/14 3:51 PM
Deleted: Output

JRBuzan 12/2/14 3:51 PM
Deleted: Calculated

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [101]

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [102]

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [100]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [103]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [104]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [105]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [106]

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [107]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [108]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

Mixing ratio	r_s	A.14
Derivative mixing ratio	dr_s/dT	A.17
Function of equivalent potential temperature	$f(\theta_E)$	A.18
Derivative of function of equivalent potential temperature	$f'(\theta_E)$	A.20
Wet Bulb Temperature	T_w	A.22
Wet Bulb Temperature, Stull	T_{wS}	8-9
Lifting condensation temperature	T_L	A.2
Moist potential temperature	θ_{DL}	A.3
Equivalent potential temperature	θ_E	A.4
Equivalent temperature	T_E	A.5
Heat Index	HI	3

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: 11-12

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

Apparent Temperature	*AT*	1
Humidex	*HUMIDEX*	4
Simplified WBGT	*sWBGT*	7
Discomfort Index	*DI*	10
Temperature Humidity Index for Comfort	*THIC*	5
Temperature Humidity Index for Physiology	*THIP*	6
Swamp cooler efficiency 65%	*SWMP65*	11
Swamp cooler efficiency 80%	*SWMP80*	11

Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: Wet Bulb Globe Tempera ... [109]

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [110]

JRBuzan 12/2/14 3:51 PM
Deleted: 10

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [111]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [112]

JRBuzan 12/2/14 3:51 PM
Deleted: Universal Thermal Climat ... [113]

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [114]

JRBuzan 12/2/14 3:51 PM
Deleted: 13

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [115]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [116]

JRBuzan 12/2/14 3:51 PM
Deleted: Temperature Humidity In ... [117]

JRBuzan 12/2/14 3:51 PM
Formatted Table ... [118]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [119]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted Cells ... [120]

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: 15

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM
Deleted: 15

JRBuzan 12/2/14 3:51 PM
Deleted: X

JRBuzan 12/2/14 3:51 PM

JRBuzan 12/2/14 3:51 PM
... [121]

Table 3. List of previous heat stress studies. Studies using datasets, reanalysis, and/or model output that range from ~1900 until ~2010 are labeled 'Modern' and from ~2005 to ~2100 are labeled Future. Some studies do not analyze heat stress quantitatively (Assessment).

Location	Metric	Time	Model	Ref.
Mediterranean Sea	*HI*	Modern and Future	RegCM3	Diffenbaugh et al. (2007)
Delhi	*WBGT*	Modern	NOAA	Kjellstrom et al. (2009a)
World	*sWBGT*	Future	HadCM3	Kjellstrom et al. (2009b)
World Cities	*WBGT, T*	Modern and Future	NOAA/Various Models	Kjellstrom et al. (2009c)
Global	*PET variation*	Future	ECHAM4	Jendritzky and Tinz (2009)
Global	T_w	Modern and Future	CCSM3/ERA Interim	Sherwood and Huber (2010)
Europe	*HI, HUMIDEX*	Future	ENSEMBLES	Fischer and Schar (2010)
Global	*indoorWBGT*	Modern and Future	NOAA	Hyatt et al. (2010)
Global	—	Modern	Assessment	Nilsson and Kjellstrom (2010)
Southern Brazil	*UTCI*	Modern	Direct Measurement	Bröde et al. (2012)
Global	*sWBGT*	Modern and Future	CLM4	Fischer et al. (2012)
Global	*sWBGT*	Modern and Future	HadCRUH/ISD-NCDC	Willett and Sherwood

				(2012)
Global	*T*	Modern	Various Datasets	SREX IPCC (2012)
Western India	*WBGT, T*	Modern	Direct Measurement	Nag et al. (2013)
California Farms	—	Modern	Assessment	Stoecklin-Marois et al. (2013)
Thailand	—	Modern	Assessment	Tawatsupa et al. (2013)
Nepal	*sWBGT, HI, HUMIDEX*	Modern	Direct Measurement	Pradhan et al. (2013)
South East Asia	*WBGT*	Modern and Future	GSOD/CRU/BCM2	Kjellstrom et al. (2013)
Quebec	*T*	Future	Assessment	Adam-Poupart et al. (2013)
Global	*indoorWBGT*	Modern and Future	ESM2M/NCEP-NCAR	Dunne et al. (2013)
United States	*sWBGT, DI, HI, HUMIDEX, AT*	Modern and Future	CLM4/CLMU/WRF	Oleson et al. (2013b)

Table 4. The HumanIndexMod: subroutine names, required inputs, and variables calculated.

Name	Subroutine	Input	Calculates
Moist Thermodynamics	*Wet_Bulb*	T, e_{RH}, P, RH, Q	T_E, θ_E, T_w
Wet Bulb Temperature, Stull	*Wet_BulbS*	T_C, RH	T_{wS}
Heat Index	*HeatIndex*	T_C, RH	HI
Apparent Temperature	*AppTemp*	T_C, e_{RH}, $Wind$	AT
Simplified WBGT	*swbgt*	T_C, e_{RH}	$sWBGT$
Humidex	*hmdex*	T_C, e_{RH}	$HUMIDEX$
Discomfort Index	*dis_coi*	T_C, T_w	DI
Discomfort Index w/Stull	*dis_coiS*	T_C, T_{wS}	DI
Temperature Humidity Index	*THIndex*	T_C, T_w	$THIC$, $THIP$
Swamp Cooler Efficiency	*SwampCoolEff*	T_C, T_w	$SWMP65$, $SWMP80$
Kelvin to Celsius	*KtoC*	T	T_C
Vapor Pressure	*VaporPres*	RH, e_s	e_{RH}
Saturated Vapor Pressure	*QSat_2*	T, P	e_s, de_s/dT, $d(ln(e_s))/dT$, r_s, dr_s/dT, $f(\theta_E)$, $f'(\theta_E)$

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: T_c

JRBuzan 12/2/14 3:51 PM
Deleted: Table 5. Regional city location and regime.

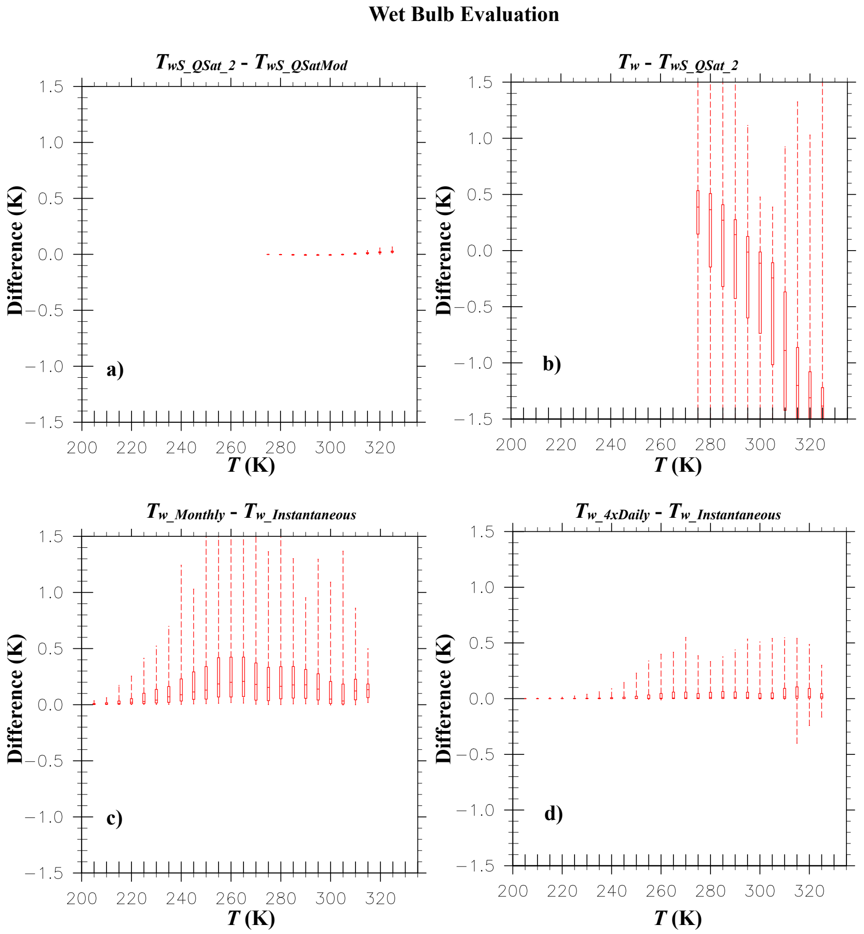


Figure 1. Evaluation of wet bulb temperatures. The boxes represent the 90% confidence interval. The upper and lower tails represent the 100% confidence interval. The horizontal line in each box is the median value. a) difference between T_{wS} using QSat_2 saturated vapor pressure and QSatMod saturated vapor pressure over the valid range for T_{wS}. b) difference between T_w (Davies-Jones, 2008) and T_{wS} (Stull, 2011) (both using QSat_2 saturated vapor pressure calculation) over the valid range for T_{wS}. c) is the difference between using model monthly averaged input fields and model instantaneous fields to calculate monthly T_w. d) difference between using model 4x Daily averaged input fields and model instantaneous fields to calculate 4x Daily T_w. For a), b), and d) the inputs of T, P, and Q are derived from model

4x Daily fields from the years 2001-2010. For c) the inputs of *T*, *P*, and *Q* are derived from model Monthly fields from the years 2001-2010.

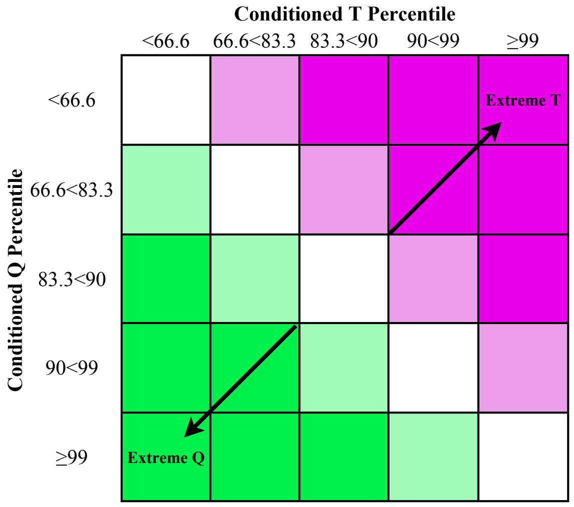


Figure 2. Expected value ranking. *T* and *Q* conditioned upon exceedance value of a heat stress or moist thermodynamic metric. The *T* and *Q* values are compared to their respective time series as a percentile. These *T* and *Q* percentiles are binned and are compared to each other. Extreme *Q* are greens and extreme *T* are magentas.

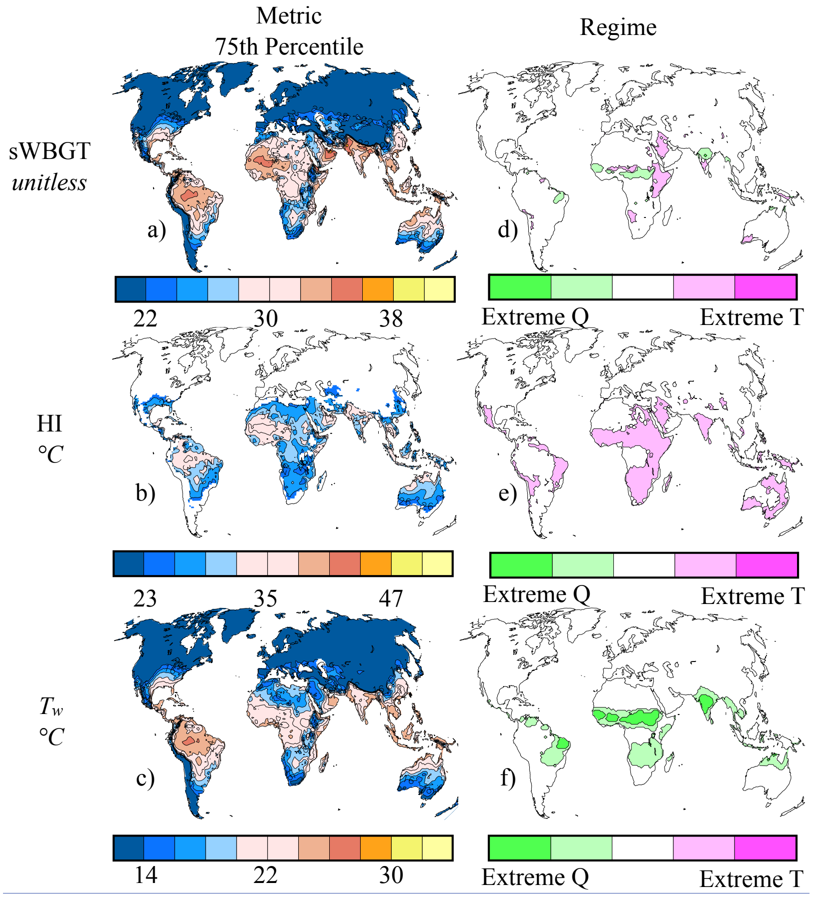


Figure 3. 75th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value T-Q regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

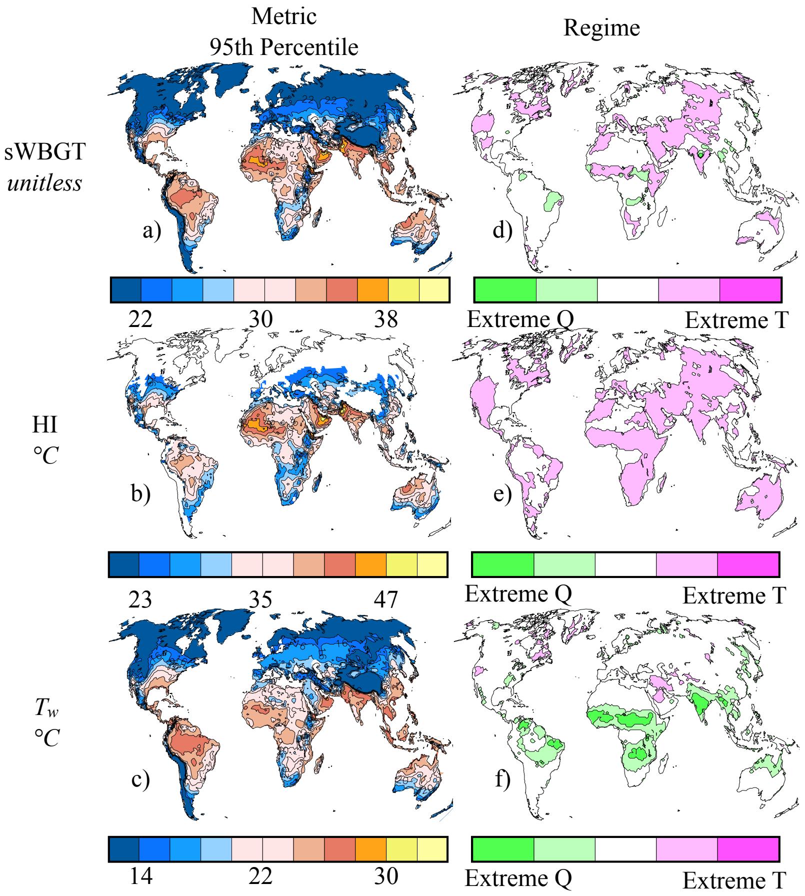


Figure 4. 95th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value *T*-*Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

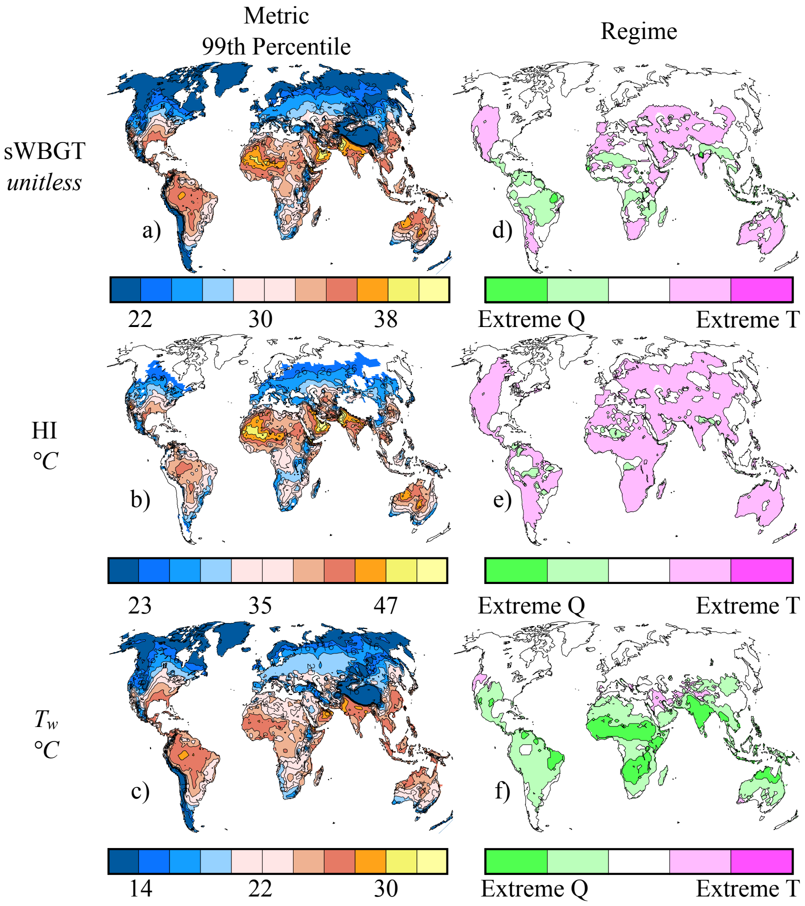


Figure 5. 99th percentile exceedance value of 3 metrics for a) sWBGT, b) HI, and c) T_w (left). Expected rank value T-Q regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.

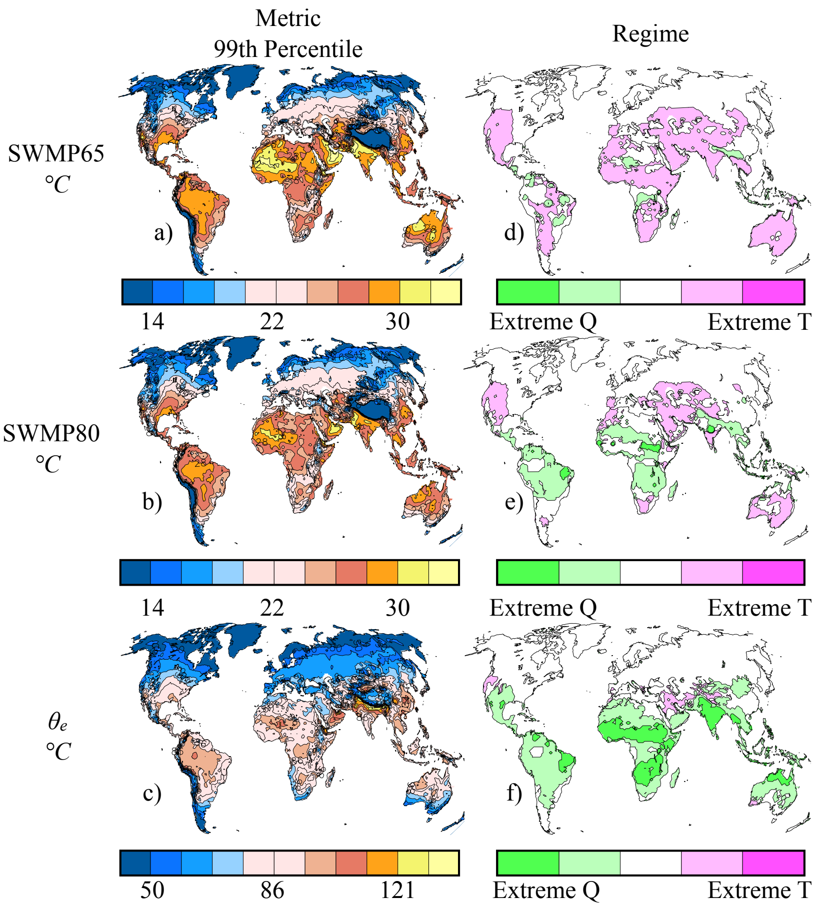


Figure 6. 99th percentile exceedance value of 3 metrics for a) SWMP65, b) SWMP80, and c) θ_E (left). Expected rank value *T-Q* regime maps d), e), and f) (right) conditioned by a), b) and c), respectively. Rank values for d)-f) are described in Figure 2.