How Do Different Processes Shape Temperature Probability Distributions? A Percentile-Averaged Temperature Tendency Decomposition

HENG QUAN,[a] BOER ZHANG,[b] STEPHEN BOURGUET,[c] MARIANNA LINZ,[b,c] AND GANG CHEN[d]

[a] *Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing, China*
[b] *School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts*
[c] *Department of Earth and Planetary Sciences, Harvard University, Cambridge, Massachusetts*
[d] *Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California*

(Manuscript received 28 July 2022, in final form 22 February 2023, accepted 13 April 2023)

ABSTRACT: Studying temperature probability distributions and the physical processes that shape them is important for understanding extreme temperature events. Previous work has used a conditional mean temperature framework to reveal whether horizontal temperature advection drives temperature to extreme or median values at a specific location as a method to dynamically interpret temperature probability distributions. In this paper, we generalize this method to study how other processes shape temperature probability distributions and explore the diverse effects of horizontal temperature advection on temperature probability distributions at different locations and different temperature percentiles. We apply this generalized method to several representative regions to demonstrate its use. We find that temperature advection drives temperatures toward more extreme values over most land in the midlatitudes (i.e., cold air advection occurs during cold anomalies and warm air advection occurs during warm anomalies). In contrast, we find that horizontal temperature advection dampens temperature anomalies in some coastal summer monsoon regions, where extreme temperatures result from other processes, such as horizontal humidity advection and vertical temperature advection. By calculating the mean of processes conditioned on the temperature percentile, this method enables composite analysis of processes that contribute to events for all percentiles and a range of processes. We show examples of composites at different percentiles for certain processes and regions to illustrate the conditional mean analysis. This general approach may benefit future studies related to temperature probability distributions and extreme events.

KEYWORDS: Atmospheric circulation; Extreme events; Monsoons; Statistics

1. Introduction

Modern climate change not only increases the mean temperature but also affects higher-order statistical moments (like variance and skewness) and thus the overall shape of temperature probability distributions (Hoskins and Woollings 2015; McKinnon et al. 2016; Rhines et al. 2017; Tamarin-Brodsky et al. 2019). Extreme events like heat waves are closely related to the shape of temperature probability distributions (Ruff and Neelin 2012; Huybers et al. 2014; Loikith and Neelin 2015; Perkins 2015). Therefore, it is important to determine the physical processes that influence temperature probability distributions and how they will change in a warming climate, so that we can better predict the likelihood, impact, and trend of extreme events from a dynamical perspective. We are still far from a thorough understanding of the physical processes that shape temperature probability distributions (Hoskins and Woollings 2015), though many approaches have been proposed recently to interpret observed temperature probability distributions (Grotjahn et al. 2015; Garfinkel and Harnik 2017; Linz et al. 2018; Tamarin-Brodsky et al. 2019, 2020; Catalano et al. 2021).

One recent approach calculated the mean of horizontal temperature advection at each temperature percentile at a given location to determine whether horizontal temperature advection drives temperature toward extreme values or back toward the median at that location (Linz et al. 2020). This analysis was done on data generated by an aquaplanet model, and the authors concluded that horizontal temperature advection drives temperature toward extreme values at most places on Earth outside the tropics, which could be interpreted by a simple theory. However, aquaplanet models do not capture many real-world phenomena due to their lack of realistic topography, land–sea contrast, and land–atmosphere interactions, so these results could not be directly applied to Earth's climate. Zhang et al. (2022) applied this method to reanalysis data and found that the role of horizontal temperature advection is spatially heterogeneous: it drives temperature toward extreme values in most places, but it also drives temperature back toward median values in a few regions (especially coastal monsoon regions). Both Linz et al. (2020) and Zhang et al. (2022) found that the role of horizontal temperature advection can vary with temperature percentiles at a given location. For example, it can drive temperature toward extreme values when temperature is high and toward median values when temperature is low in some places. The mechanisms controlling such diverse roles of horizontal temperature advection

Denotes content that is immediately available upon publication as open access.

Quan's current affiliation: Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey.

Corresponding author: Heng Quan, hengquan@princeton.edu

DOI: 10.1175/JCLI-D-22-0556.1

are likely also quite diverse. In addition to the horizontal temperature advection, Zhang et al. (2022) studied how vertical processes (like vertical temperature advection) affect local temperature probability distributions, but they did not consider the roles of other processes (like those related to humidity) that can affect local temperature.

This article aims to generalize the original conditional mean framework of temperature in Linz et al. (2020) in order to study the roles of other processes. We use the framework to address two questions: First, why does horizontal temperature advection play different roles at different locations and at different temperature percentiles? Second, in addition to horizontal temperature advection, how do other processes shape temperature probability distributions?

The rest of the paper is organized as follows. In section 2 we review the original conditional mean temperature framework of Linz et al. (2020) and generalize it in order to study how different processes shape temperature probability distributions. We also provide an explanation for interpreting this more general treatment. In section 3 we present three case studies to demonstrate how to apply the generalized method, and we explain the diverse impacts of horizontal temperature advection using composite analyses. We also explained the roles of other processes. Finally, in section 4, we present conclusions, discussions, and potential extensions.

2. Method: The generalized conditional mean temperature framework

Global warming is simultaneously increasing the global mean temperature and changing the shape of local temperature probability distributions. Many studies have examined extreme events in particular regions and how these events have changed or are likely to change (Ruff and Neelin 2012; Huybers et al. 2014; Loikith and Neelin 2015; Perkins 2015). Analysis of composites of weather patterns that precede or follow an "event" (defined however is most relevant to the phenomenon of interest) can provide dynamical insights into how those events come to be. This work and related studies (Linz et al. 2020; Zhang et al. 2022) are motivated by the desire to take this local detective work and make it more widely applicable both geographically and with respect to what constitutes an event. Accordingly, we examine all percentiles of temperature at each location in our study (though we focus on land), and we take the mean of the temperature tendencies due to different processes at each of these percentiles (T^e, where e is the percentile). The conditional mean temperature framework was first proposed in Linz et al. (2020) and then used in Zhang et al. (2022) [a related framework was presented for precipitation in Chen et al. (2019), Norris et al. (2019a,b), and Ma et al. (2020)]. We encourage the interested reader to refer to Linz et al. (2020) for details of the conditional mean framework used to examine the role of horizontal temperature advection in driving extremes. Here, we present an explanation of how to interpret these results and then focus on temperature advection before generalizing the method to explore how other processes shape temperature probability distributions.

a. *A simplified presentation of the conditional mean framework*

Consider a case where temperature is only affected by two processes, call them A and B. The temperature budget equation is

$$\frac{\partial T}{\partial t} = A + B. \tag{1}$$

This balance applies instantaneously, and it also applies to the average at a given percentile [see Linz et al. (2020)], so if we use the notation $\langle * \rangle_{T^e}$ to represent the mean at a given percentile e of temperature, we can also write the following balance:

$$\left\langle \frac{\partial T}{\partial t} \right\rangle_{T^e} = \langle A \rangle_{T^e} + \langle B \rangle_{T^e}. \tag{2}$$

This can be rewritten to express the left-hand side as the rate of change of the temperature at a given percentile:

$$\frac{\partial T^e}{\partial t} = \langle A \rangle_{T^e} + \langle B \rangle_{T^e}. \tag{3}$$

Since we are considering the temperature distributions to be stationary in time, there is no time tendency at a particular percentile. At a given percentile, then, there is a balance between A and B:

$$0 = \langle A \rangle_{T^e} + \langle B \rangle_{T^e}. \tag{4}$$

How can we make any inferences about causality in this case? Each percentile of temperature has a balance of terms—which of these is "driving" temperature anomalies toward more extreme values and which is "damping" them toward the median? To think about this, consider a particular event that falls into, say, the 75th percentile T^{75}, where it is getting warmer. That means that right now

$$\frac{\partial T}{\partial t} = A + B > 0. \tag{5}$$

A and B were balanced, but in order to get warmer, now A has to get larger or B does. Which one is causing the temperatures to get hotter? We return to the statistical perspective and see what the balance is, statistically, when it is warmer. If the conditional means of A and B look like the example shown in Fig. 1a, $\langle A \rangle_{T^{75}} < \langle A \rangle_{T^{76}}$, so on average the contribution of A is even more positive at the 76th percentile than the 75th. At the 76th percentile, B is more negative, so B cannot be doing the work of making the temperatures more extreme. Therefore A drives the temperature from the 75th to the 76th percentile. A more complicated example is shown in Fig. 1b. Now if at a particular time, the temperature is at the 75th percentile and it is getting warmer, $\langle A \rangle_{T^{75}} > \langle A \rangle_{T^{76}}$, but $\langle B \rangle_{T^{75}} < \langle B \rangle_{T^{76}}$, so now B is driving the warm extreme values (even though the tendency due to B is actually a negative tendency). The behavior changes over the distribution in this example also, so A drives cold extremes, but B drives warm ones.

The simplest way to understand the role of different processes in causing the shape of the temperature distribution heuristically is to remember that a term with a positive linear

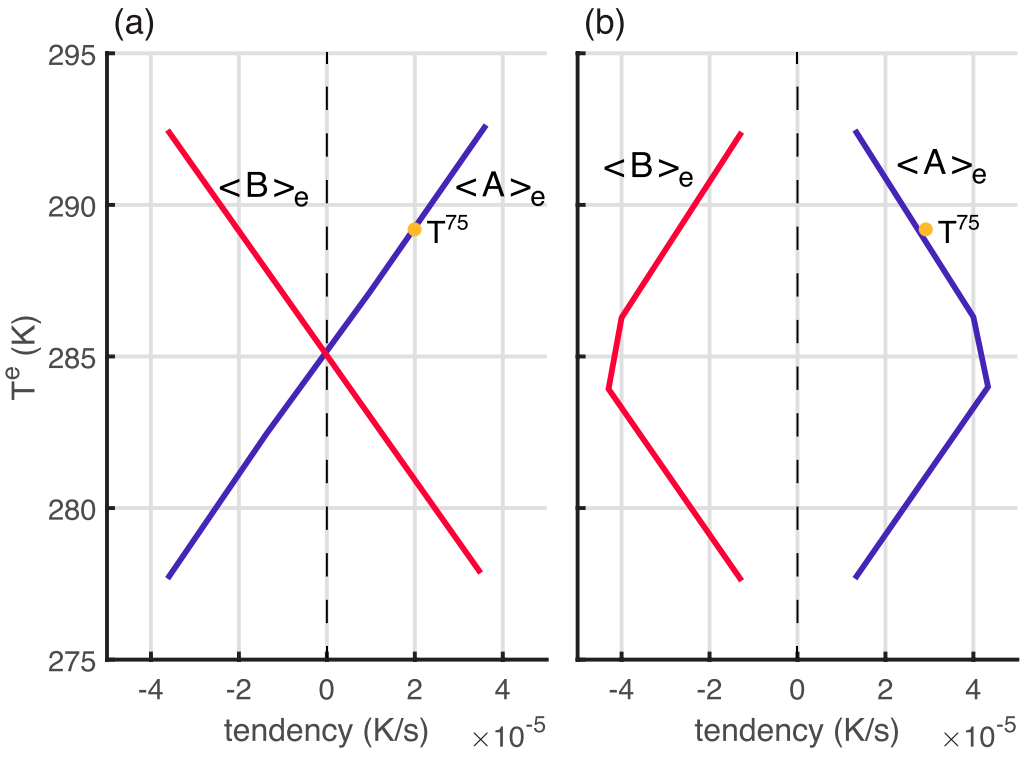

FIG. 1. Two examples of the balance described in Eq. (4). (a) The two physical processes act consistently between hot and cold percentiles. (b) A less intuitive example where the process that always acts as a negative tendency (B) drives warmer extremes and the process that always acts as a positive tendency (A) drives cold extremes.

slope (in T^e versus tendency) for a certain range of percentiles is driving temperatures toward more extreme values in that range. One with a negative slope is acting to damp anomalies. With that in mind, we will consider temperature advection alone before moving on to other processes.

b. *The role of horizontal temperature advection in shaping temperature distributions*

We start by looking at the horizontal temperature advection ($-\mathbf{v}\cdot\nabla T$), because this is the easiest term to understand. At each grid point, we first calculate the conditional mean temperature values T^e, where T^e represents temperature at the eth percentile. Then we calculate the conditional mean horizontal temperature advection $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ when temperature is at T^e. Notice that $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ represents the conditional mean horizontal temperature advection at the eth temperature percentile, rather than the conditional mean value at the advection percentile $\langle -\mathbf{v}\cdot\nabla T\rangle_{(-\mathbf{v}\cdot\nabla T)^e}$, which in general are very different [see Fig. 1 in Zhang et al. (2022)]. The relationship between T^e and $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ shows how horizontal temperature advection shapes the temperature probability distribution at that grid point.

Following Zhang et al. (2022), in this article we calculate the conditional mean values on 49 different percentiles ($e = 2, 4, \ldots, 98$), and each conditional mean value is the average over temperature percentile range $[e - 0.5, e + 0.5]$. All calculations are based on 6-hourly 850-hPa ERA5 reanalysis data in boreal summer (JJA). We only focus on land, though the approach is also applicable over ocean. Since we use data on the 850-hPa pressure level, regions where atmospheric pressure does not reach 850 hPa are ignored.

We then calculate the Pearson correlation between T^e and $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$, $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$, to explore the relationship between horizontal temperature advection and the local temperature probability distribution. Figure 2a shows the global distribution of $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$. The implication of $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$ can be explained by idealized examples in Fig. 2b: if $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) = +1$ (the red line), there is warm (cold) advection when temperature is high (low). Therefore, horizontal temperature advection drives temperature to extreme values by making hot (cold) days hotter (colder). By contrast, the horizontal temperature advection drives temperature back to the median value by making hot (cold) days colder (hotter) in areas where $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) = -1$ (the blue line). Most places in Fig. 2a have a significant positive correlation, while a few small regions have a negative one. This is consistent with the result in Zhang et al. (2022) (also based on ERA5 reanalysis data), but slightly different from Linz et al. (2020) (based on data generated by an idealized aquaplanet model) due to the lack of land–sea contrast in the aquaplanet model.

Note that the sign of $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$ does not tell the entire story, especially in the tropics and subtropics. For example, at the Arabian Peninsula, the fact that $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx +1$

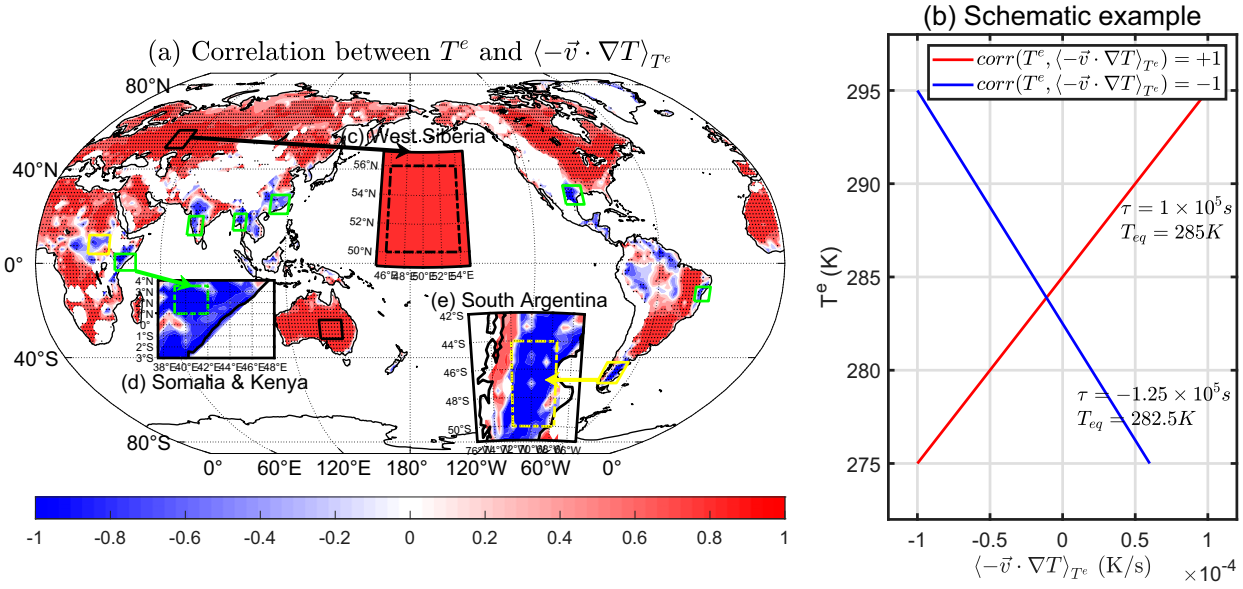

FIG. 2. (a) The distribution of the Pearson correlation between 850-hPa JJA conditional mean horizontal temperature advection and temperature [corr(T^e, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$)] on land. Regions where the 850-hPa pressure level is under the surface are ignored. Grid points whose correlation value meets the 5% confidence level in a shuffling test are shaded. In this map we mark two black squares, six green squares, and two yellow squares. We will choose one representative region from each color to do a case study in section 3. (b) Idealized examples when conditional mean horizontal temperature advection and temperature have a perfect linear relationship $T^e = \tau\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e} + T_{\text{eq}}$. (c)–(e) The corr($T^e$, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$) in three representative regions marked by black, green, and yellow squares, respectively. The dashed squares indicate subregions in which we will compute regional averages. We will do case studies in these three regions in section 3.

in Fig. 2a tells us horizontal temperature advection drives temperature to extreme values, but we cannot say the hot (cold) extreme temperatures result from warm (cold) advections, because the Arabian Peninsula is among the hottest places on Earth in JJA and hot temperatures there do not advect from other hotter places. The conditional mean horizontal temperature advection $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ is negative at every T^e at the Arabian Peninsula, and more negative at low temperatures while close to zero at high temperatures. As shown in Fig. 1b, "horizontal temperature advection drives temperature to extreme values" means "horizontal temperature advection increases with temperature" but not necessarily means "warm (cold) advection occurs at hot (cold) temperatures." Therefore, the correlation between conditional mean horizontal temperature advection and temperature percentiles is thus a quick way to check the general behavior in a region, but a more detailed examination is sometimes necessary.

The fact that corr(T^e, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$) $\approx +1$ in many places can be interpreted to mean that all other processes act as a linear damping. We can take the Taylor approximation of the total derivative of temperature,

$$\dot{T} = a_0 + a_1 T + a_2 T^2 + \cdots \tag{6}$$

and rewrite the equation to the first order as a Newtonian relaxation to an equilibrium temperature T_{eq},

$$\dot{T} \approx -\frac{T - T_{\text{eq}}}{\tau}. \tag{7}$$

Therefore, $T_{\text{eq}} = -a_0/a_1$, with a positive relaxation time scale $\tau = -1/a_1$ (τ is similar to a relaxation time scale when positive and is physically uninterpretable when negative).

Neglecting the vertical temperature advection $-\omega(\partial T/\partial p)$ (considering it to be part of the relaxation term), we can rewrite the local temperature tendency equation as

$$\frac{\partial T}{\partial t} = -\mathbf{v}\cdot\nabla T - \frac{T - T_{\text{eq}}}{\tau}, \tag{8}$$

and then calculate the conditional mean value of both sides of the equation. The conditional mean local temperature tendency can be neglected, as it is nonzero with global warming but nevertheless far smaller than the other terms in the equation (see Zhang et al. 2022). We then get

$$T^e = \tau\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e} + T_{\text{eq}}, \tag{9}$$

which tells us that when the approximation in Eq. (7) is valid, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ will be positively and linearly related to T^e. Therefore, horizontal temperature advection drives temperatures to extreme values in the areas in Fig. 2 where corr(T^e, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$) $\approx +1$, while other processes together serve as a relaxation process driving temperatures back to the median value.

Although τ and T_{eq} have been written this way in analogy to a Newtonian relaxation, they cannot be understood in the standard way. A positive τ gives a sense of the time scale of persistence of the temperature anomalies due to advection.

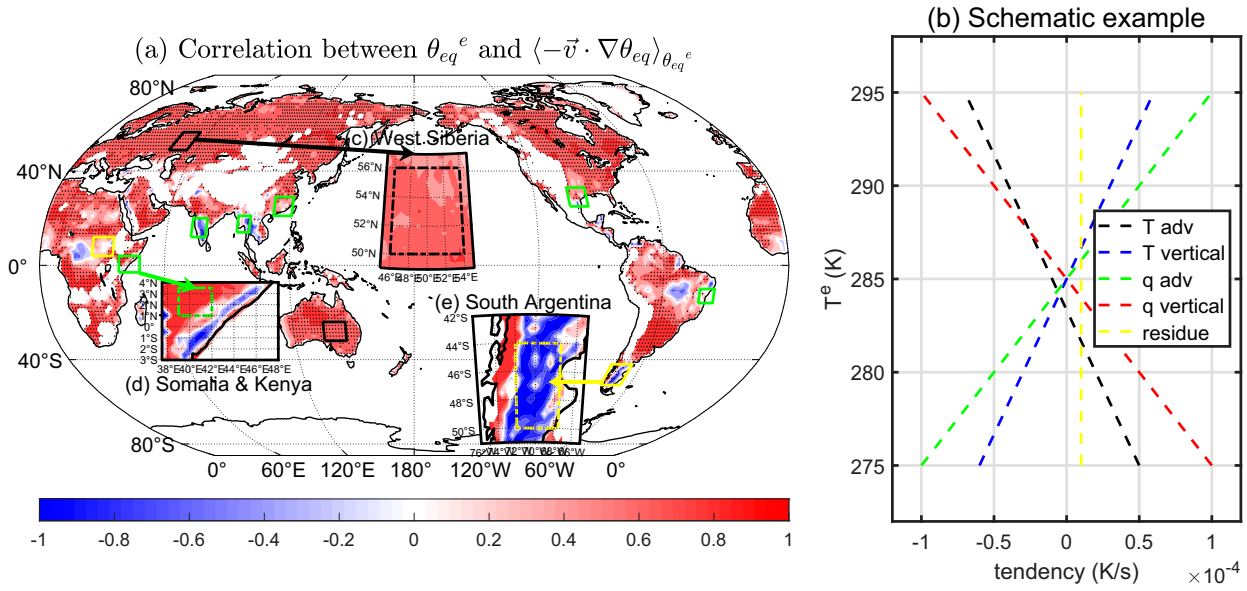

FIG. 3. (a) The distribution of the Pearson correlation between 850-hPa JJA conditional mean horizontal θ_{eq} advection and θ_{eq} on land. Regions where the 850-hPa pressure level is under the surface are ignored. Grid points whose correlation value meets the 5% confidence level in a shuffling test are shaded. The squares are same as those in Fig. 2a. (b) Iidealized scatterplots when all terms have a perfect linear relationship with T^e. The sum of all terms is zero according to Eq. (13). (c)–(e) the corr(θ_{eq}^e, $\langle -\mathbf{v}\cdot\nabla\theta_{eq}\rangle_{\theta_{eq}^e}$) in three representative regions. The dashed squares indicate subregions in which we will compute regional average.

Notice that $T_{eq} \neq T^{50}$, and T_{eq} is useful only in so much as the sign of $T_{eq} - T^{50}$ will give a sense of the role of other processes at T^{50} [if T_{eq} is larger/smaller than T^{50}, "other processes" $-(T^{50} - T_{eq})/\tau$ will be positive/negative and have a warming/cooling effect at T'^{50}]. The $\tau < 0$ is physically uninterpretable as this simplified picture is not relevant when temperature advection is acting to damp anomalies (applying τ and T_{eq} fields derived here in a dynamical core would not be meaningful and would indeed be unstable). The way to use these values is as a sanity check and, where $\tau > 0$, as a comparison of the strength of other processes relative to temperature advection. Alternatively, consider them to just be coefficients from a Taylor expansion that have been rewritten to have simpler units. They are determined by doing a linear regression, and we label them in Fig. 2b.

c. *The roles of other processes in shaping temperature distributions*

In the regions where corr(T^e, $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$) ≈ -1 (highlighted by blue color in Fig. 2a), processes other than the horizontal temperature advection drive temperature to extreme values, while the horizontal temperature advection drives temperature back to the median value. Most of these regions are coastal summer monsoon regions (such as the six green squares in Fig. 2a; please refer to appendix A for the definition of monsoon region and Fig. A1 for a map), so we expect humidity to play an important role in shaping their local temperature distributions.

To test this, we can replace temperature T by equivalent potential temperature θ_{eq} and see how conditional mean horizontal θ_{eq} advection (calculated at different θ_{eq} percentiles) is related to θ_{eq}. The θ_{eq} takes both temperature and humidity into consideration, so differences between T and θ_{eq} can tell us aboout the role of humidity in shaping temperature distributions. Equivalent potential temperature is calculated by the following simplified formula (Stull 1988):

$$\theta_{eq} = \theta + \frac{L_v}{c_p} q = T\left(\frac{p_0}{p}\right)^{R_d/c_p} + \frac{L_v}{c_p} q, \tag{10}$$

where T is temperature, θ is potential temperature, q is specific humidity, L_v is the latent heat of vaporisation of water, c_p is the heat capacity of air, $p_0 = 1000$ hPa, $p = 850$ hPa, and R_d is the specific gas constant.

Figure 3a shows the global distribution of corr(θ_{eq}^e, $\langle -\mathbf{v}\cdot\nabla\theta_{eq}\rangle_{\theta_{eq}^e}$), which is calculated using the same data as Fig. 2a; therefore, a comparison of the correlations displayed in these figures can reveal the role of humidity in setting local temperature distributions. In particular, regions with positive correlations in Fig. 2a retain their positive correlation in Fig. 3a (like west Siberia in Fig. 2c), while many regions with negative correlations in Fig. 2a (especially coastal monsoon regions marked by green squares like Somalia and Kenya in Fig. 2d) are positive in Fig. 3a, indicating a switch from a damping role for temperature advection to amplifying one for θ_{eq} advection. In those regions horizontal temperature and moisture advection together cause θ_{eq} values to be more extreme.

To show the roles of humidity related processes more explicitly, next we calculate the total derivative of θ_{eq},

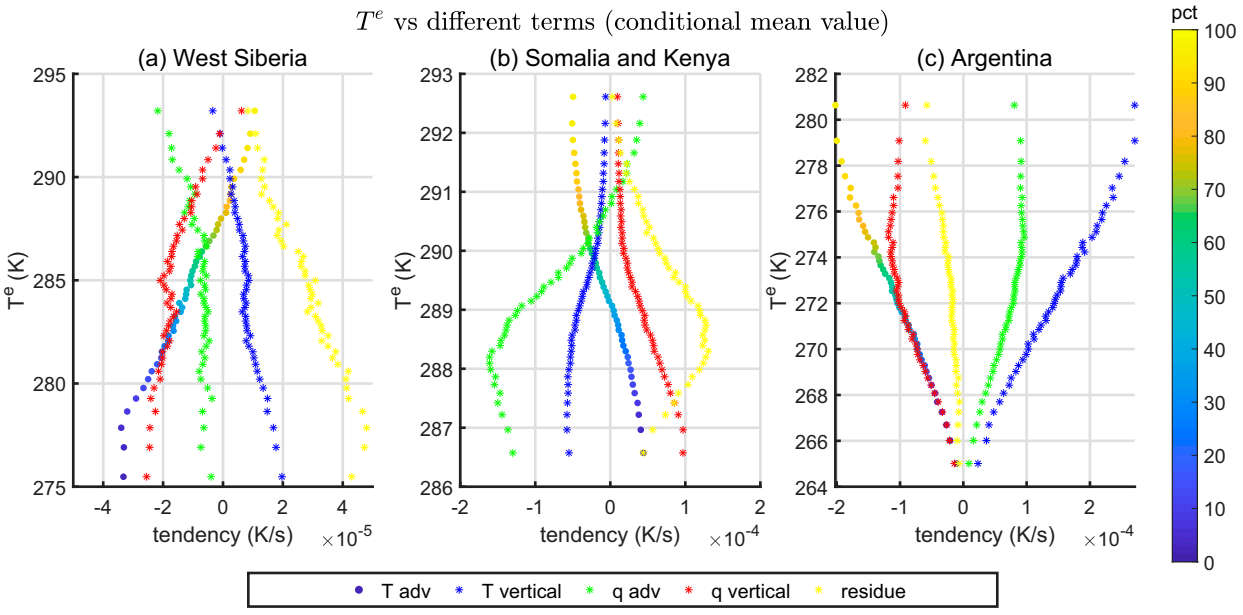

FIG. 4. The relationships between conditional mean temperature T^e (averaged in the dashed squares in Figs. 2c–e) vs conditional mean horizontal temperature advection $(p_0/p)^{R_d/c_p}\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ (points with gradient color indicating percentile e), vertical temperature advection $(p_0/p)^{R_d/c_p}\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ (blue points), horizontal humidity advection $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e}$ (green points), vertical humidity advection $(L_v/c_p)\langle -\omega(\partial q/\partial p)\rangle_{T^e}$ (red points), and residual term R_{T^e}(yellow points) averaged in those dashed squares in Figs. 2c–e. (a) West Siberia; (b) Somalia and Kenya, east coast of equatorial Africa; and (c) south corner of Argentina. Note we have used the opposite signs from Zhang et al. (2022, their Fig. 13).

$$\frac{d\theta_{eq}}{dt} = \frac{\partial\theta_{eq}}{\partial t} + \mathbf{v}\cdot\nabla\theta_{eq} + \omega\frac{\partial\theta_{eq}}{\partial p}, \tag{11}$$

and expand this equation to separate temperature and humidity,

$$\frac{\partial\theta_{eq}}{\partial t} + \left(\frac{p_0}{p}\right)^{R_d/c_p}(\mathbf{v}\cdot\nabla T) + \frac{L_v}{c_p}(\mathbf{v}\cdot\nabla q) + \omega\frac{\partial}{\partial p}\left[\left(\frac{p_0}{p}\right)^{R_d/c_p}T\right] + \omega\frac{L_v}{c_p}\frac{\partial q}{\partial p} = R, \tag{12}$$

where the residual term $R = d\theta_{eq}/dt$ incorporates other processes that can change θ_{eq}, like radiation and sensible heat flux. Note that R does not include latent heat flux, since θ_{eq} is conserved in moist adiabatic processes.

Then we calculate the conditional mean value of both sides in Eq. (12) at different temperature percentiles. Since the conditional mean local θ_{eq} tendency can be ignored (the largest trend is on the order of 0.1 K yr^{-1}, much smaller than advection terms), we get

TABLE 1. We apply the generalized conditional mean framework to 10 representative regions marked by squares in Fig. 2a. Based on results in the table, they can be roughly divided into three classes (with different square colors). The first column shows the locations of representative regions. The second column shows the signs of $r(T) = \mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$. The third column shows the signs of $r(\theta_{eq}) = \mathrm{corr}(\theta^e_{eq}, \langle -\mathbf{v}\cdot\nabla\theta_{eq}\rangle_{\theta^e_{eq}})$. The fourth column shows the classes that the regions belong to. The fifth column shows the mechanisms of $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx -1$ (why the horizontal temperature advection drives temperature to median value) in summer (JJA), while $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx +1$ is common hence not explained. The last column lists terms (e.g., T_{adv} means horizontal temperature advection) that drive temperature to extreme values, and the term with the asterisk is the dominant one with larger magnitude.

Region	$r(T)$	$r(\theta_{eq})$	Class	Why $r(T) < 0$	Driving term
Central Australia	+	+	1	—	T_{adv}*, $q_{vertical}$
West Siberia	+	+	1	—	T_{adv}*, $q_{vertical}$
Somalia, Kenya	−	+	2	Monsoon	q_{adv}*, $T_{vertical}$
Brazil east coast	−	+	2	Monsoon	q_{adv}*, $T_{vertical}$
India west coast	−	+	2	Monsoon	q_{adv}*, R
China south coast	−	+	2	Monsoon	q_{adv}*, $T_{vertical}$
Myanmar	−	+	2	Monsoon	q_{adv}*, R
U.S. Texas coast	−	+	2	Unknown	q_{adv}*
South Argentina	−	−	3	Topography	$T_{vertical}$*, q_{adv}
Central Africa	−	0	3	Topography	q_{adv}*, $T_{vertical}$

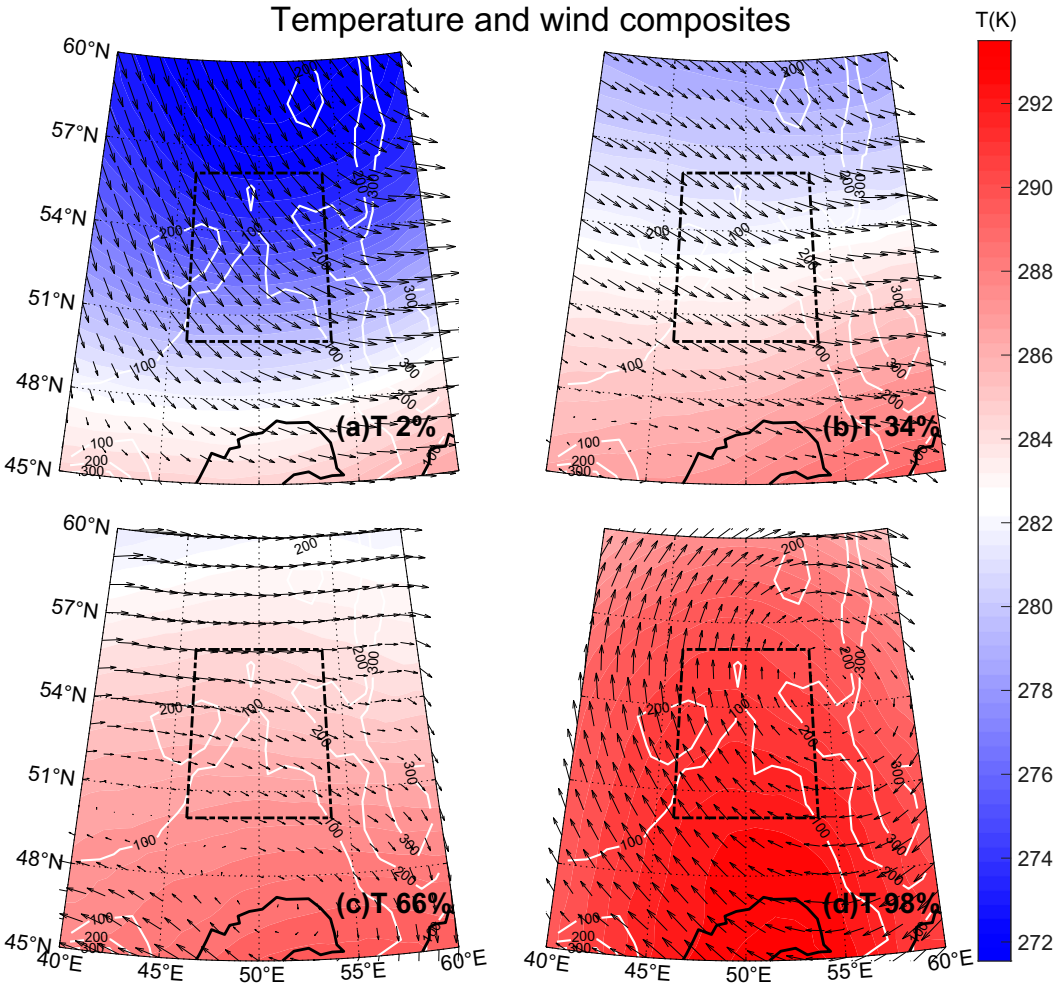

FIG. 5. The composite maps in west Siberia, Russia (Fig. 2c), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The white lines are elevation contours, and the dashed black line marks the subregion where we compute regional averages. The black arrows represent conditional mean wind vectors, and the color indicates temperature.

$$\underbrace{\left(\frac{p_0}{p}\right)^{R_d/c_p}\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}}_{1}+\underbrace{\frac{L_v}{c_p}\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e}}_{2}+\underbrace{\left(\frac{p_0}{p}\right)^{R_d/c_p}\left\langle -\omega\frac{\partial T}{\partial p}+\omega\frac{\kappa T}{p}\right\rangle_{T^e}}_{3}+\underbrace{\frac{L_v}{c_p}\left\langle -\omega\frac{\partial q}{\partial p}\right\rangle_{T^e}}_{4}+\underbrace{R_{T^e}}_{5}=0, \tag{13}$$

where each term is averaged at a temperature percentile T^e. We will call the five terms in Eq. (13) conditional mean horizontal temperature advection, horizontal humidity advection, vertical temperature advection, vertical humidity advection, and a residual term, respectively. If one starts from the primitive T and q equations and derives the moist static energy (MSE) budget, one will get the same result [please refer to Eq. (10.3) in Schneider and Sobel (2007)]. Later when we do composite analysis, we will omit the factors $(p_0/p)^{R_d/c_p}$ and L_v/c_p for simplicity.

We can study the roles of these five terms by studying each one's relationship with conditional mean temperature T^e. At grid points where the result behaves like the idealized example in Fig. 3b, the horizontal humidity advection and vertical temperature advection drive temperature to extreme values, while the horizontal temperature advection and the vertical humidity advection drive temperature back to the median value. The residual term has a negligible effect. Zhang et al. (2022) did a similar diagnostic analysis, but they mainly focused on vertical temperature advection $\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ and did not take humidity into consideration.

So far, we have generalized the original conditional mean framework in Linz et al. (2020) and Zhang et al. (2022) to study how processes other than the horizontal temperature advection shape temperature probability distributions. Next,

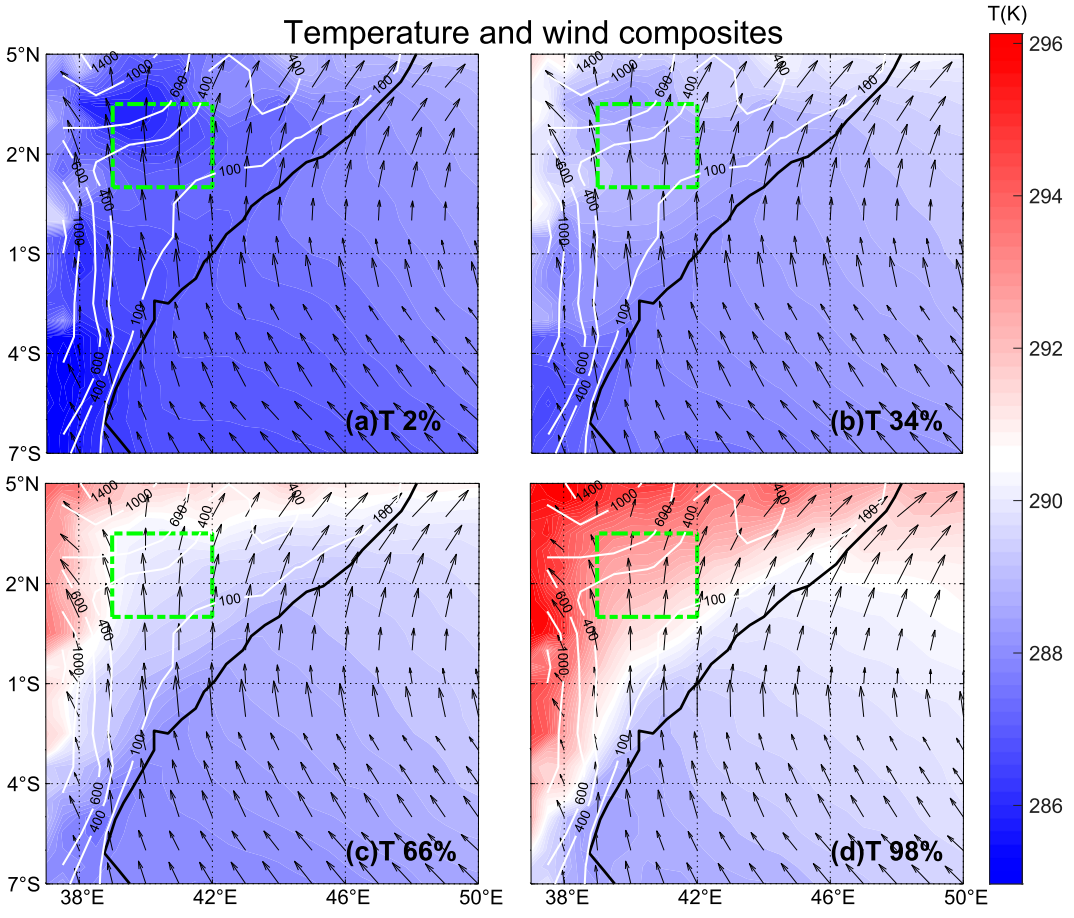

FIG. 6. The composite maps in Somalia and Kenya, the east coast of equatorial Africa (Fig. 2d), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, the white lines are elevation contours, and the dashed green line marks the subregion where we compute regional averages. The black arrows represent conditional mean wind vectors, and the color indicates temperature.

we will apply the method to the three representative regions shown in Figs. 2c–e as case studies.

3. Results

Figure 4 shows scatterplots of T^e versus the five terms (conditional mean value) in Eq. (13) in the three representative regions (they represent three different classes) in Figs. 2c–e. In this section we focus on these three representative regions to study how horizontal temperature advection and other processes shape local temperature probability distributions and explain the diverse roles of horizontal temperature advection in shaping local temperature probability distributions by composite analysis. Results of other regions are summarized in Table 1.

a. The role of horizontal temperature advection

First, we will examine the role of horizontal temperature advection ($-\mathbf{v}\cdot\nabla T$) in setting the temperature distribution, and explain its diverse roles in different regions by composite analysis.

The first example is west Siberia, an inland plain in Fig. 2c, which represents most regions where $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx +1$. [See Zhang et al. (2022) for a comparison of regions.] The points with color gradient show horizontal temperature advection in Fig. 4a with the color indicating the percentile. This region has a cold (warm) horizontal temperature advection when temperature is low (high), so the horizontal temperature advection drives temperature to extreme values. To study the mechanism connecting horizontal temperature advection to local temperature distribution, we create and analyze composite maps at different temperature percentiles.

Composite maps in Fig. 5 show how the horizontal temperature advection ($-\mathbf{v}\cdot\nabla T$) drives temperature to the extreme values: the direction of the temperature gradient ∇T does not depend on temperature percentile T^e (the south side is hotter than the north side in each of the four maps, though the magnitude of temperature gradient depends on temperature. See Fig B1 in appendix B for details), but the direction of wind vector $\mathbf{v}$ depends on T^e. When temperature is low (Fig. 5a, on cold days), the wind heads from north (cooler) to south (warmer), creating cold temperature advection in the dashed

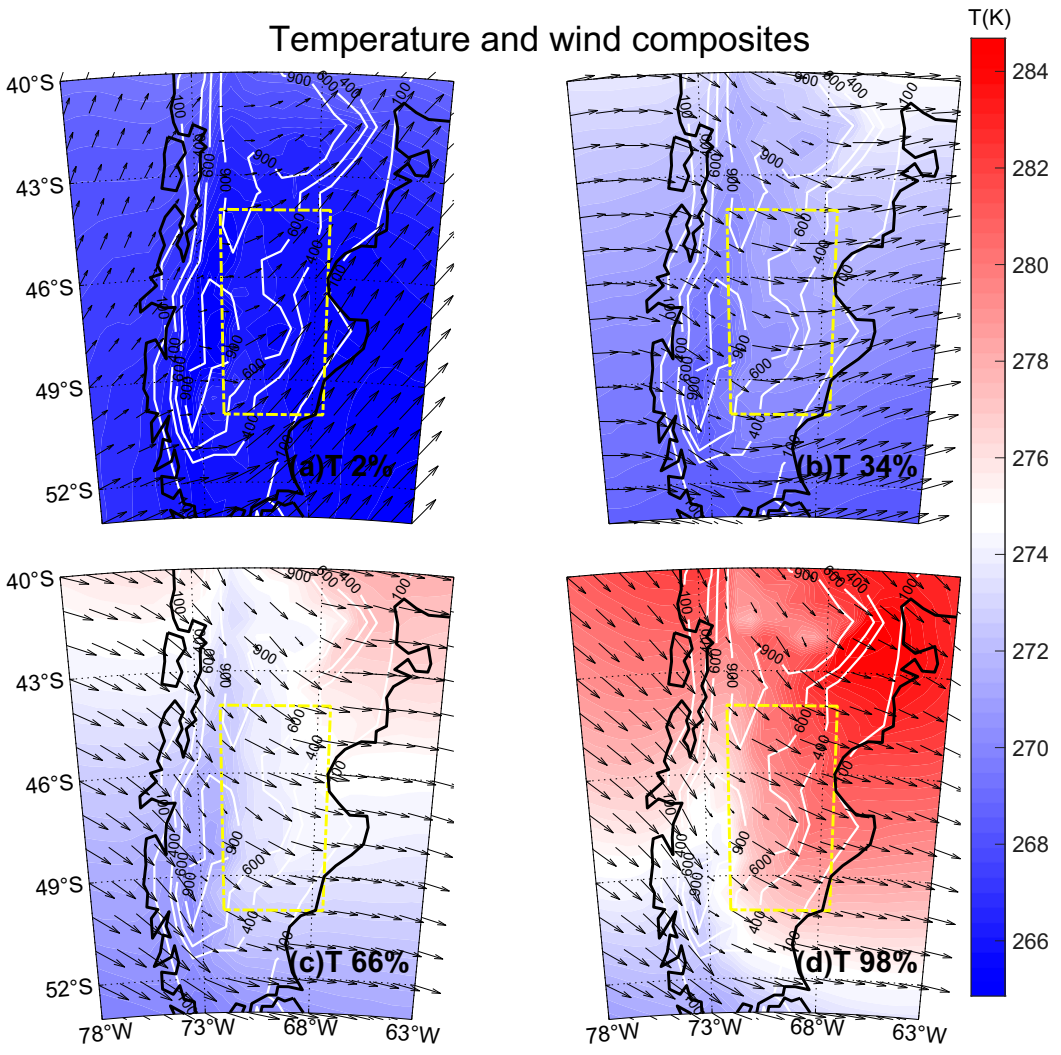

FIG. 7. The composite maps in the south corner of Argentina (Fig. 2e), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, the white lines are elevation contours, and the dashed yellow line marks the subregion where we compute regional averages. The black arrows represent conditional mean wind vectors, and the color indicates temperature.

black square ($-\mathbf{v}\cdot\nabla T<0$). By contrast, when temperature is high at these points (Fig. 5d, on hot days), southerly winds drive warm temperature advection ($-\mathbf{v}\cdot\nabla T>0$). The other two composite maps (Figs. 5b,c) are simply transitional states between the extremes. Therefore, the horizontal temperature advection ($-\mathbf{v}\cdot\nabla T$) tends to make hot days hotter and cold days colder, thereby driving temperature to extreme values. Figure 4a also shows that the vertical humidity advection has a positive slope for the upper half of the distribution, so it also contributes to warm extremes. This example is a simple one that nevertheless corresponds to many regions across the globe.

Our second case study is Somalia and Kenya, on the east coast of equatorial Africa in Fig. 2d. This region and many other tropical regions where $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})\approx -1$ (marked by green squares in Fig. 2a) are typical coastal summer monsoon regions (Li and Zeng 2000; An et al. 2015), which suggests that the coastal summer monsoon influences the role of horizontal temperature advection in these regions. [We use a monsoon index (Li and Zeng 2000) to define monsoon regions; details are in appendix A]. This analysis is only done for boreal summer, so we cannot comment on $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$ in these regions during other nonmonsoonal seasons.

From the points with color gradient (which represents horizontal temperature advection) in Fig. 4b we know that this region has warm (cold) horizontal temperature advection when temperature is low (high). Composite maps in Fig. 6 show that the coastal summer monsoon contributes to the negative $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e})$ in Somalia and Kenya. The direction of the wind vector $\mathbf{v}$ does not depend on temperature percentile T^e ($\mathbf{v}$ has the same directions in four maps in Fig. 6), because

the prevailing background monsoon wind is from the ocean to the land. Conversely, the direction of temperature gradient ∇T depends on T^e. When temperature is low (Fig. 6a, in cold days), the ocean is warmer than the land (because of the ocean's higher heat capacity), so the sea breeze transports warm air from the ocean to the land, resulting in the warm advection in the dashed green square ($-\mathbf{v} \cdot \nabla T > 0$). When temperature is high (Fig. 6d, in hot days), the ocean is cooler than the land and cold air is advected onto land by the prevailing winds ($-\mathbf{v} \cdot \nabla T < 0$). The other two composite maps (Figs. 6b,c) are transitional states between the extremes. Therefore, the horizontal temperature advection ($-\mathbf{v} \cdot \nabla T$) associated with the coastal summer monsoon tends to make hot days cooler and cold days warmer, thereby driving temperatures back to the median values.

Our third case study is the south corner of Argentina (Patagonia) in Fig. 2e, where $\text{corr}(T^e, \langle -\mathbf{v} \cdot \nabla T \rangle_{T^e}) \approx -1$ but there is not a coastal monsoon. The other region of this class is in central Africa (as denoted by a yellow square in Fig. 2a). Both of these are distinguished by substantial local topography, which should play an important role in setting the shape of temperature distributions (Lutsko et al. 2019). Unlike the coastal monsoon regions, the horizontal temperature advection $\langle -\mathbf{v} \cdot \nabla T \rangle_{T^e}$ (points with color gradient) in Fig. 4c always has a negative value: when temperature is high (low), the cold advection is stronger (weaker). Therefore, the horizontal temperature advection ($-\mathbf{v} \cdot \nabla T$) drives temperature back to median values, and it has a cooling effect at both high and low temperatures.

The composite maps at different temperature percentiles in Fig. 7 differ from those in Somalia and Kenya (Fig. 6). Here, the wind direction depends on temperature percentile T^e, and the temperature distribution does not have land–sea contrast. Figure 7 shows that the horizontal temperature advection is actually a negative contribution at every temperature percentile, with a stronger negative tendency at the highest percentiles. This example shows that the negative correlation does not show the entire story, as at the coldest temperatures, horizontal temperature advection still has a negative tendency, thereby has a cooling effect. As we mentioned earlier, the correlation between conditional mean horizontal temperature advection and temperature percentiles is a quick way to check the general behavior in a region, but a more detailed examination is sometimes necessary.

In Fig. 8 we explain the diverse roles of horizontal temperature advection in shaping local temperature probability distribution at different places. In most inland regions (e.g., west Siberia in the Northern Hemisphere Fig. 2c), the south is always warmer than the north. Northerly wind at low temperatures cools the black box in Fig. 8a to lower temperatures, while southerly wind at high temperatures warms the black box in Fig. 8b to higher temperatures, so horizontal temperature advection drives temperature to extreme values (see Fig. 5). By contrast, in most coastal monsoon regions (e.g., Somalia and Kenya at the east coast of equatoral Africa in Fig. 2d), there is always sea breeze (because we are looking at the conditional mean wind that filters out the local land–sea

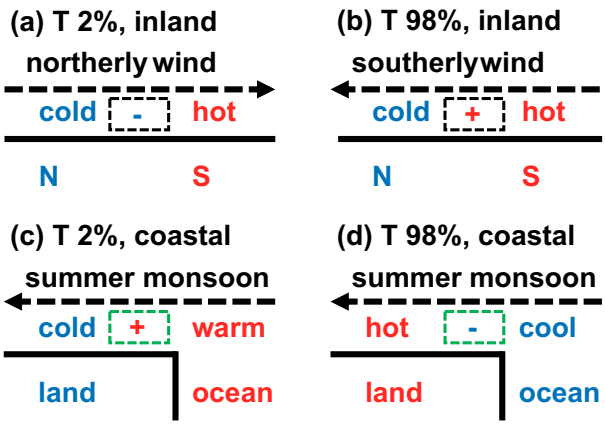

FIG. 8. The schematic figure explaining the diverse roles of horizontal temperature advection in shaping local temperature probability distribution at different places: (a) an inland region (in the Northern Hemisphere), when temperature is low; (b) an inland region, when temperature is high; (c) a monsoon region, when temperature is low; and (d) a coastal monsoon region, when temperature is high. The minus sign means cold advection, while the plus sign means warm advection.

circulation) in boreal summer. When temperature is low (high), the ocean is warmer (cooler) than the land because of higher heat capacity, resulting in warm (cold) advection in the green box in Fig. 8c (Fig. 8d), so the horizontal temperature advection drives temperature back to median values (see Fig. 6). Nevertheless, there are also a few regions with special topography (like south Argentina in Fig. 2e) where the negative $\text{corr}(T^e, \langle -\mathbf{v} \cdot \nabla T \rangle_{T^e})$ is not caused by summer monsoon. There are some regions where the slope of horizontal temperature advection changes sign for different percentiles, but generally, temperature advection is well behaved so that the sign of the correlation coefficient provides information over almost all of the distribution.

b. *The roles of other processes*

Now we look back to Figs. 3 and 4 to study the roles of other processes in shaping local temperature probability distribution.

Considering θ_{eq} instead of T, the correlation turns from negative to positive in Somalia and Kenya (Fig. 3d), but remains approximately the same in west Siberia and south Argentina (Figs. 3c,e). Such contrast shows that humidity plays an important role in determining the sign of the correlation in the monsoon (Somalia and Kenya) region, while it is not important in the inland (west Siberia) or nonmonsoon (Argentina) region. This is consistent with the results in Figs. 2a and 3a, where we see that most coastal monsoon regions with green squares have negative $\text{corr}(T^e, \langle -\mathbf{v} \cdot \nabla T \rangle_{T^e})$ but positive $\text{corr}(\theta^e_{\text{eq}}, \langle -\mathbf{v} \cdot \nabla \theta_{\text{eq}} \rangle_{\theta^e_{\text{eq}}})$.

We study the roles of different processes more explicitly by the budget analysis in Fig. 4. In west Siberia (Fig. 4a), horizontal temperature advection $\langle -\mathbf{v} \cdot \nabla T \rangle_{T^e}$ (points with color gradient), and, to a lesser extent, vertical humidity advection (L_v/c_p)

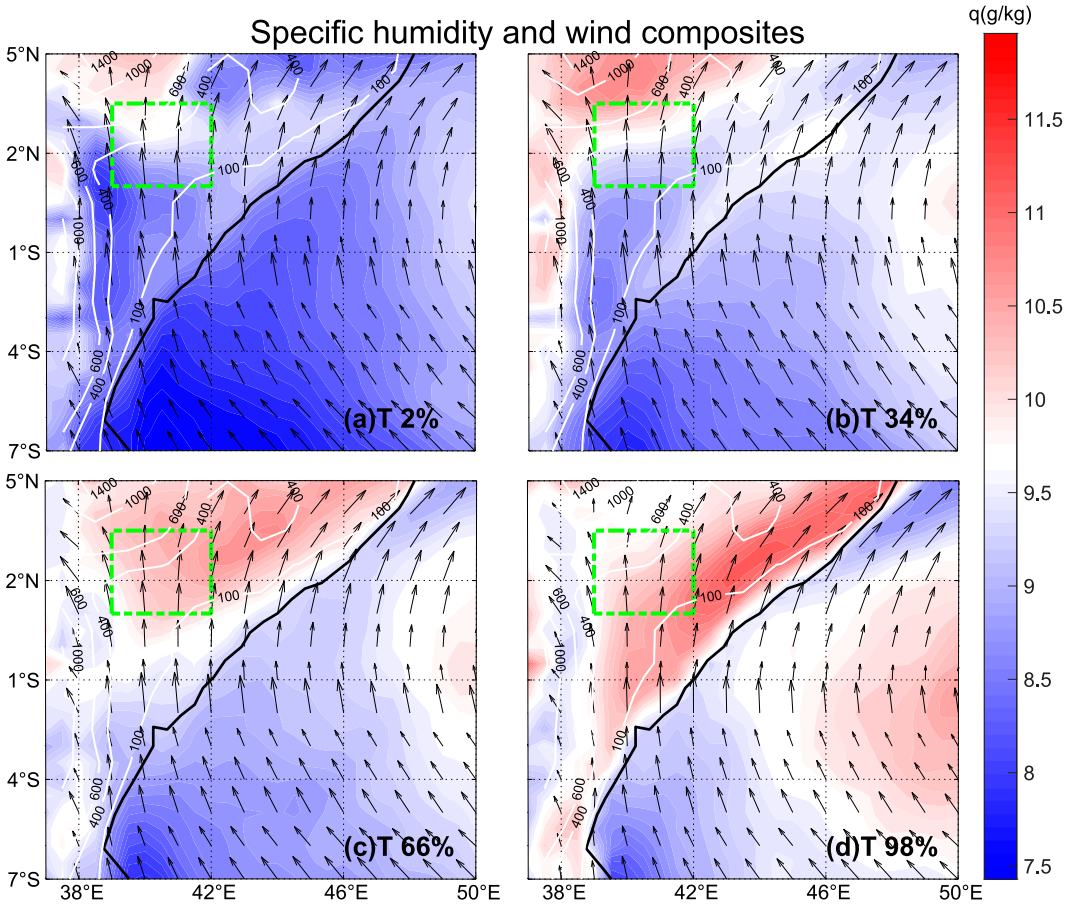

FIG. 9. The specific humidity and wind composite maps in Somalia and Kenya, the east coast of equatorial Africa (Fig. 2d), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, the white lines are elevation contours, and the dashed green line marks the subregion where we compute regional averages. The black arrows represent conditional mean wind vectors, and the color indicates specific humidity.

$\langle -\omega(\partial q/\partial p)\rangle_{T^e}$ (red points) drive temperature to extreme values. In Somalia and Kenya (Fig. 4b), horizontal humidity advection $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e}$ (green points), and, to a lesser extent, vertical temperature advection $\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ (blue points) drive temperature to extreme values. In south corner of Argentina (Fig. 4c), vertical temperature advection $(p_0/p)^{R_d/c_p}\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ (blue points), and, to a lesser extent, horizontal humidity advection $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e}$ (green points) drive temperature to extreme values. A summary of the same diagnostic analysis in other regions (marked by squares in Fig. 2a) is shown in Table 1. Now we analyze the roles of other processes by case studies.

1) THE ROLE OF HORIZONTAL HUMIDITY ADVECTION

We find that in lots of coastal monsoon regions (marked by green squares in Fig. 2a), horizontal humidity advection, instead of horizontal temperature advection, is the dominant term that drives temperature to extreme values. Therefore, we take Somalia and Kenya in Fig. 2d as an example to explain the role of horizontal humidity advection (it drives temperature to extreme values, see the green points in Fig. 4b) in detail by looking at humidity and wind composites at different temperature percentiles.

First, we focus on specific humidity and wind composites. Similar to Fig. 6, in Fig. 9 the direction of the wind vector $\mathbf{v}$ does not depend on temperature percentile T^e, because the prevailing background wind is always from the ocean to the land. Conversely, the direction of specific humidity gradient ∇q depends on T^e. When temperature is low, (Fig. 9a, in cold days), the sea breeze transports dry air near the coast to wetter inland regions, so $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e} < 0$ in the dashed green square; when temperature is high (Fig. 9d, in hot days), the sea breeze transports wet air near the coast to drier inland regions, so $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e} > 0$ in the dashed green square. The other two composite maps (Figs. 9b,c) are transitional states between the extremes. Therefore, horizontal humidity advection might drive temperature to extreme values by affecting latent heat release from condensation: At low temperatures dry advection reduces latent heat release and has a cooling effect, while at high temperatures moist advection increases latent heat release and has a warming effect.

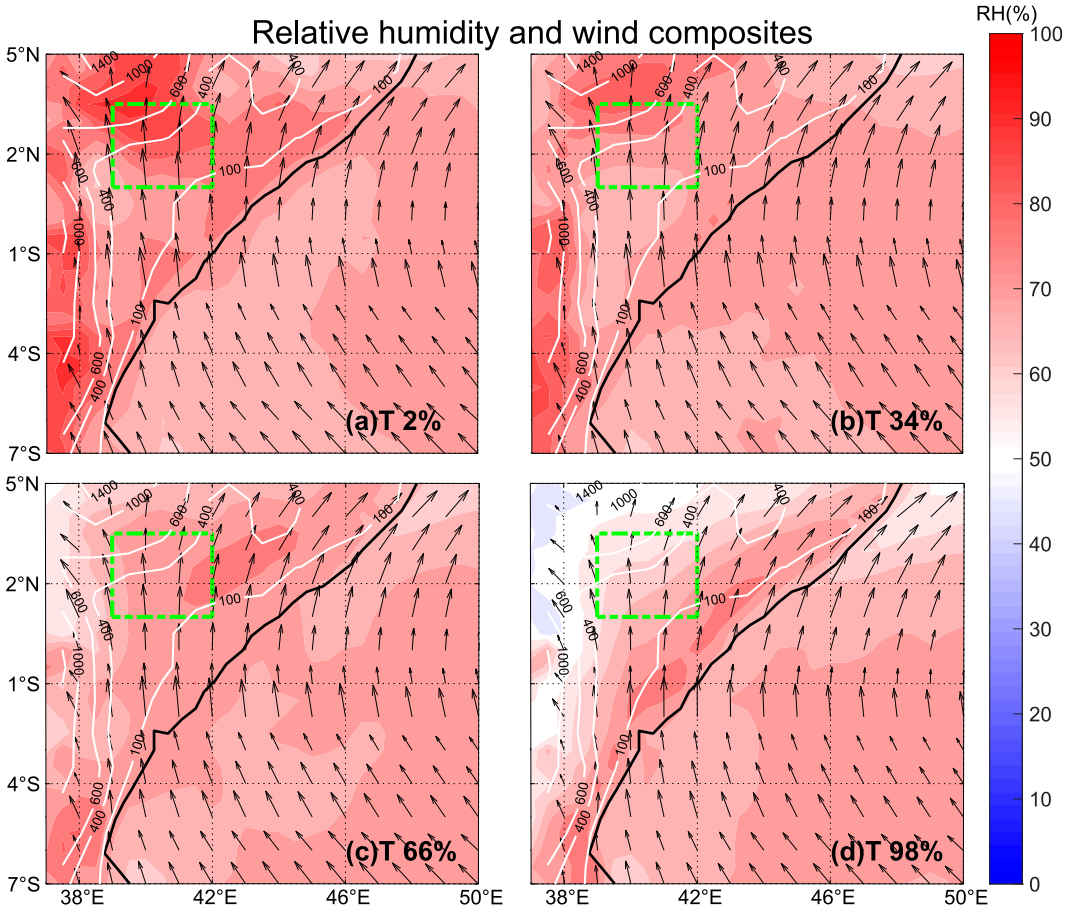

FIG. 10. The relative humidity and wind composite maps in Somalia and Kenya, the east coast of equatorial Africa (Fig. 2d), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, the white lines are elevation contours, and the dashed green line marks the subregion where we compute regional averages. The black arrows represent conditional mean wind vectors, and the color indicates relative humidity.

To see whether horizontal humidity advection drives temperature to extreme values via latent heat release from condensation, we not only look at specific humidity (Fig. 9 above), but also look at relative humidity (RH, Fig. 10 below). At low temperatures (Fig. 10a), the highest RH in the subregion enclosed by the green square is already 100%, so condensation releases latent heat. The dry advection (Figs. 9a and 10a) replaces the saturated moist air by dry air and reduces latent heat release, so it has a cooling effect. However, at high temperatures (Fig. 10d), RH in the green square or the source of the advection are both far from 100%. Although moist advection (Figs. 9d and 10d) makes the green square wetter, the air is not saturated, so the moist advection does not increase latent heat release. Therefore, horizontal humidity advection has a cooling effect at low temperatures by reducing latent heat release, but has a warming effect at high temperatures through other mechanisms at Somalia and Kenya. We are not able to explain the latter right now, and we leave this for future research.

Similar humidity and wind composite analysis can be done in other representative regions. We find that in most coastal monsoon regions (green squares in Fig. 2a), horizontal humidity advection is the dominant factor driving temperature to extreme values (see Table 1 for details), although the underlying mechanisms might be different. That is why $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) < 0$ (Fig. 2d) but $\mathrm{corr}(\theta^e_{\mathrm{eq}}, \langle -\mathbf{v}\cdot\nabla\theta_{\mathrm{eq}}\rangle_{\theta^e_{\mathrm{eq}}}) > 0$ (Fig. 3d) in coastal monsoon regions.

2) THE ROLE OF THE RESIDUAL TERM (DIABATIC HEATING)

Somalia and Kenya in Fig. 2d is also a good example to study the role of the residual term. In Fig. 4, the residual term R_{T^e} (yellow points) is presumed to be from diabatic heating, which is dominated by radiation absorption at the 850-hPa pressure level. Figure 4b tells us that at Somalia and Kenya the residual term R_{T^e} is large and positive at low temperatures, while close to zero at high temperatures. Therefore, numerically we would expect that radiation absorption related to cloud cover has a significant warming effect at low extremes but does not play an important role at high extremes.

Figure 11 shows the cloud cover fraction in Somalia and Kenya at different temperature percentiles. At low temperatures

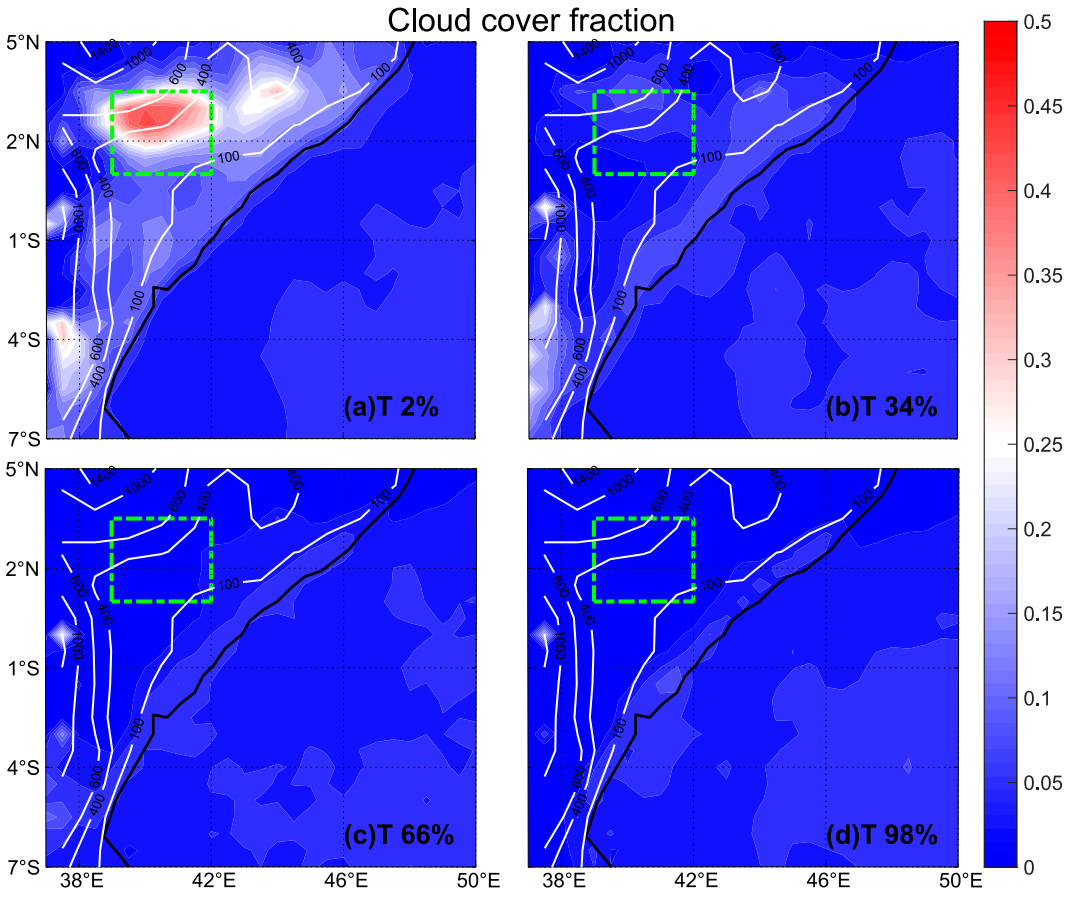

FIG. 11. The cloud cover fraction at Somalia and Kenya, the east coast of equatorial Africa (Fig. 2d), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, the white lines are elevation contours, and the dashed green line marks the subregion where we compute regional averages. The color indicates cloud cover fraction.

(Fig. 11a), the cloud cover in the green square is anomalously high. Low cloud at 850 hPa absorbs solar radiation and results in a positive heating rate, so radiation absorption related to cloud cover has a warming effect at low temperatures. At high temperatures (Fig. 11d), the cloud cover in the green square is close to zero, so the heating rate is very small and radiation absorption does not affect temperature. Results from Fig. 11 are consistent with our expectation above based on Fig. 4b.

3) THE ROLE OF VERTICAL TEMPERATURE ADVECTION

Subsidence during summer is known to be important for warm extremes. Now we use south Argentina as the example to study the role of vertical temperature advection. Figure 4c tells us that vertical temperature advection (blue points) has a positive slope, so numerically it drives temperature to extreme values, and seems to have a large warming effect at high temperatures at the south corner of Argentina.

Figure 12 shows the horizontal and vertical velocity composites in south Argentina at different temperature percentiles. At low temperatures (Fig. 12a), the vertical velocity in the yellow square is relatively small, and close to zero on average. As a result, vertical temperature advection is small at low temperatures. At high temperatures (Fig. 12d), the yellow square is dominated by subsidence [vertical velocity $\omega = dP/dt > 0$]. Given that $-(\partial T/\partial p) + (\kappa T/p) > 0$ in this region (figure not shown), vertical temperature advection $\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ is positive and has a significant warming effect. Because the horizontal advection comes from the Andes, we can conclude that adiabatic warming from downslope winds off the Andes is a major contributor to warm extremes in this region.

4. Conclusions and discussion

In this article we present a more general conditional mean temperature framework to study how various processes shape local temperature probability distributions. We also explore the mechanisms underlying the diverse effects that horizontal temperature advection can have on temperature probability distributions at different locations and different temperature percentiles. We then apply the generalized conditional mean temperature framework and perform composite analyses in several representative regions (marked by squares in Fig. 2a)

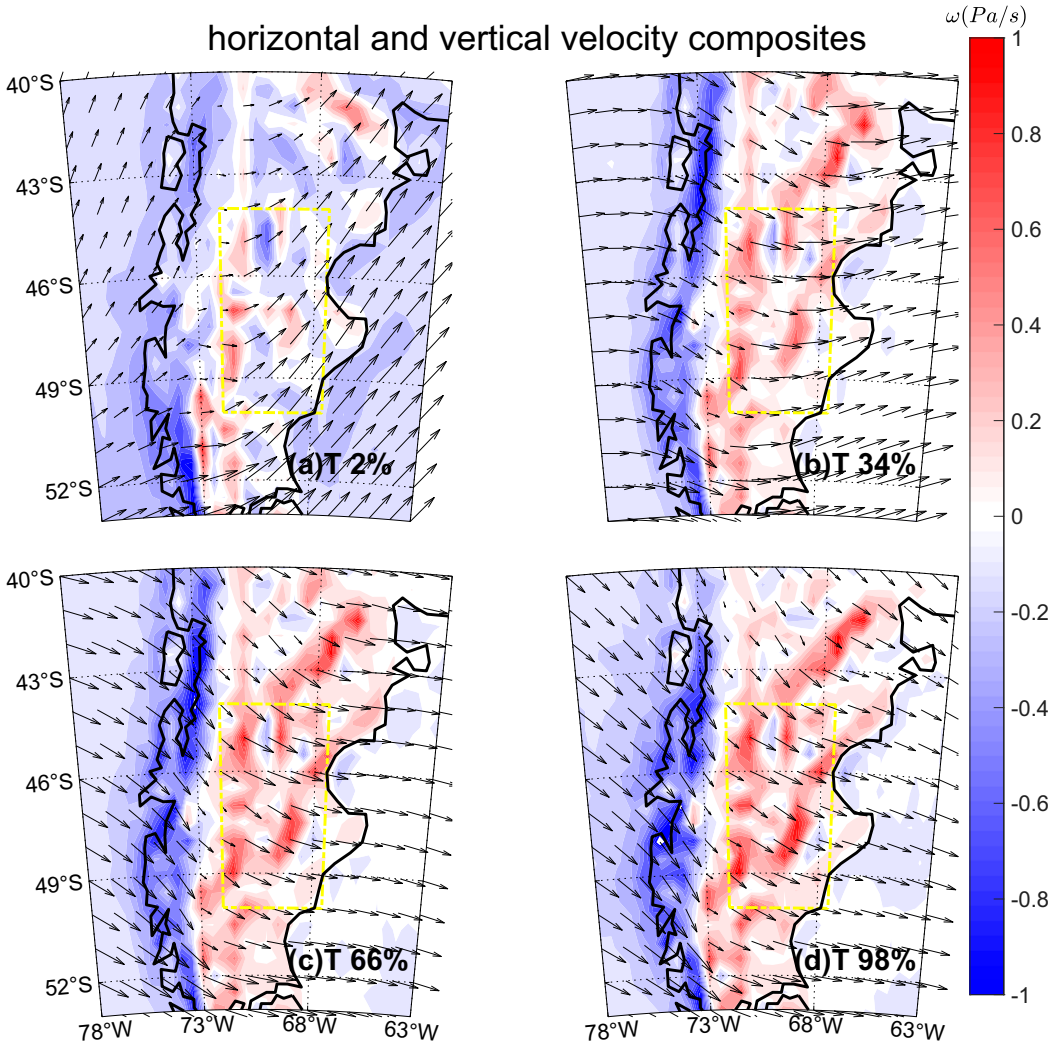

FIG. 12. The horizontal and vertical velocity $\omega = dP/dt$ composite maps in the south corner of Argentina (Fig. 2e), at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The solid black line is the coastline, and the dashed yellow line marks the subregion where we compute regional averages. The black arrows represent horizontal wind vectors, and the color indicates vertical velocity $\omega = dP/dt$ (positive means subsidence).

and present results of three case studies in section 3. We summarize the results of all 10 case studies in Table 1.

Based on the results in Table 1, the 10 representative regions can be roughly divided into three classes. We now compare our results to the results from k-means clustering analysis for all land area based on the relationship between conditional mean temperature T^e and horizontal temperature advection $\langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}$ (Zhang et al. 2022).

Class 1 in Table 1 features $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx +1$, and it belongs to cluster 1, the most common cluster in Fig. 9 in Zhang et al. (2022). This class represents regions on land where $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx +1$, including central Australia and west Siberia. In those regions the horizontal temperature advection $(-\mathbf{v}\cdot\nabla T)$ is the dominant term that drives temperature to extreme values. In Linz et al.'s (2020) analysis of temperature advection in an aquaplanet model, almost all places outside the tropics behave like those in class 1.

Class 2 and class 3 in Table 1 feature $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) \approx -1$, and they both belong to cluster 3 in Fig. 9 in Zhang et al. (2022). We divide this cluster into two classes here because horizontal temperature advection plays the same role in shaping the temperature distribution in each region, but the mechanism underlying this relationship differs bewteen the two. Class 2 represents coastal summer monsoon regions, where the horizontal temperature advection $(-\mathbf{v}\cdot\nabla T)$ drives temperature to median values. Humidity is important in coastal monsoon regions, so $\mathrm{corr}(\theta^e_{\mathrm{eq}}, -\langle \mathbf{v}\cdot\nabla\theta_{\mathrm{eq}}\rangle_{\theta^e_{\mathrm{eq}}}) \approx +1$ for this class. The conditional mean vertical temperature advection $\langle -\omega(\partial T/\partial p) + \omega(\kappa T/p)\rangle_{T^e}$ and horizontal humidity advection $(L_v/c_p)\langle -\mathbf{v}\cdot\nabla q\rangle_{T^e}$ drive temperature toward extreme values

in class 2 regions. Class 3 regions are not coastal summer monsoon regions, but they are affected by substantial local topography, where horizontal temperature advection ($-\mathbf{v} \cdot \nabla T$) drives temperature to median values and other processes (such as subsidence) drive temperature to extreme values.

The generalized conditional mean framework developed here can be used in future studies to explore how different processes shape temperature probability distributions. We have used it in a series of case studies to examine diverse roles of horizontal temperature advection in shaping temperature probability distributions at different locations and different temperature percentiles. In particular, this framework can be useful for studying the temperature probability distributions for extreme events. This method can be thought of as a budget analysis or composite analysis for each temperature percentile.

Furthermore, this conditional mean framework can be applied to other tracers besides temperature, like precipitation (Chen et al. 2019; Norris et al. 2019a,b; Ma et al. 2020) or trace gas concentration. In fact, it can be applied to any tracer A that follows the tendency equation:

$$\frac{\partial A}{\partial t} = -v \cdot \nabla A + \dot{A}. \tag{14}$$

Thus, one could apply this to medium- to long-lived trace gases to see the relative role of chemistry or transport over a range of pollution conditions.

Our application of the conditional mean framework to temperature shows that the correlation between local temperature and horizontal temperature advection can reveal where temperature advection amplifies or dampens temperature anomalies. In regions where the correlation is approximately +1, the amplifying behavior is consistent with previous studies that found skewness in the temperature distribution resulting from eddy advection of temperature (Linz et al. 2018; Tamarin-Brodsky et al. 2019; Garfinkel and Harnik 2017). We spend much of this paper focused on a more detailed application of the framework to examine the relationship between local temperature and horizontal temperature advection in regions that are less straightforward. In the coastal monsoon regions, land–sea temperature contrasts and background monsoon cause horizontal temperature advection to dampen extreme values. In other regions, the role of horizontal temperature advection is more complicated. In this study we focus on the season where less agreement is found (JJA), but of course this analysis could be expanded to examine DJF. This will not generally be applicable to shoulder seasons (MAM and SON), because we explicitly assume a stationary temperature distribution, which will obviously not be the case for those times. Overall, this method is an attempt to generalize the insights gained from regional extreme event case studies to global understanding of the full distribution. Last, we want to clarify that our approach statistically implies causality rather than examines the evolution of a particular event, and there is a Lagrangian back trajectory method (Catalano et al. 2021) that solves the latter problem better.

We have made the conditional mean budget terms available globally, and we hope this will be useful for understanding temperature distributions and how they might change as the underlying processes shift with global warming.

Acknowledgments. Heng Quan thanks Dr. Da Nian, Dr. Yu Huang, and Prof. Zuntao Fu at Peking University for helpful discussions.

Data availability statement. The data used in this article are the ERA5 reanalysis (Hersbach et al. 2020) from the European Center for Medium-Range Weather Forecasts (ECMWF). We use temperature, horizontal wind, specific humidity, and vertical velocity ($\omega = dp/dt$) at the 850-hPa pressure level during 1979–2018. Temperature and specific humidity at 825 and 875 hPa are also used to calculate vertical processes. The temporal resolution is 6 h, while the spatial resolution is 0.5° × 0.5°. Instantaneous data are used unless specified otherwise. The dataset used in this study can be accessed through Copernicus Climate Change Service (C3S) (DOI: 10.24381/cds.bd0915c6). Postponed data are available upon request.

APPENDIX A

The Global Monsoon Index

The monsoon can be defined in various ways depending on an observer's perspective (Trenberth et al. 2000; Qian 2000; Chang et al. 2000; Wang and Ding 2006, 2008; An et al. 2015). One of the most straightforward definitions is the seasonality of the wind, which is an objective index of the intensity and location of the monsoon (Li and Zeng 2000, 2002, 2003). We use this definition here both for its ease of calculation and because it fits this study well: the wind direction contributes to negative corr(T^e, $\langle -\mathbf{v} \cdot \nabla T \rangle_{T^e}$) in coastal monsoon regions (see section 3).

Li and Zeng (2000) proposed a global monsoon index based on the seasonality of the wind called static normalized seasonality (SNS):

$$\delta = 2\frac{|\mathbf{v}_1 - \mathbf{v}_7|}{|\mathbf{v}_1 + \mathbf{v}_7|} - 2, \tag{A1}$$

where $\mathbf{v}_1$ and $\mathbf{v}_7$ represent climatological wind in January and July respectively.

The SNS monsoon index δ has two features [as proved in Li and Zeng (2000)]:

1) If we fix the magnitude of $\mathbf{v}_1$ and $\mathbf{v}_7$, δ will be a strictly increasing function of the angle α between $\mathbf{v}_1$ and $\mathbf{v}_7$.

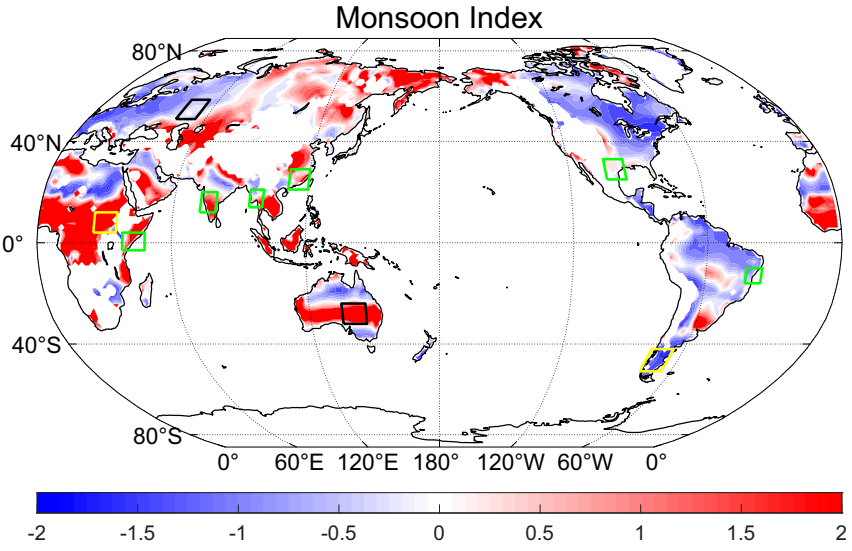

FIG. A1. The distribution of the global monsoon SNS index δ, based on monthly average ERA5 horizontal wind data at the 850-hPa pressure level. Regions where $\delta > 0$ are considered monsoon regions. Here we copy the 10 squares in Fig. 2a. The coastal regions marked by the six green squares are monsoon regions ($\delta > 0$). Most regions where $\mathrm{corr}(T^e, \langle -\mathbf{v}\cdot\nabla T\rangle_{T^e}) < 0$ (colored in blue in Fig. 2a) are monsoon regions (colored in red in this figure).

2) When $\alpha = 90°$, $\delta = 0$; when $\alpha > 90°$, $\delta > 0$; when $\alpha < 90°$, $\delta < 0$.

Li and Zeng (2000) defined the monsoon regions globally as regions where the SNS index $\delta > 0$ (where $\mathbf{v}_1$ and $\mathbf{v}_7$ have an angle $\alpha > 90°$). The threshold $\delta > 0$ (or $\alpha > 90°$) is subjective, but if we use $\alpha > 120°$ instead of $\alpha > 90°$, our results only change slightly [see Fig. 1 in Li and Zeng (2000)]; hence the definition is robust.

We repeat Li's work and plot the global distribution of the SNS monsoon index in Fig. A1, based on monthly average ERA5 horizontal wind data on the 850-hPa pressure level. According to Li's definition, regions colored in red ($\delta > 0$) in Fig. A1 are considered monsoon regions.

APPENDIX B

Temperature Gradients at Different Temperature Percentiles

In the west Siberia case study, we found that the direction of temperature gradient does not depend on temperature, but the magnitude of temperature gradient might depend on temperature. Here we briefly study the temperature dependence of the magnitude of the horizontal temperature gradient $|\nabla T|$, at west Siberia.

From Fig. B1 we learn that $|\nabla T|$ depends on temperature. The $|\nabla T|$ is larger at low temperatures but smaller at high temperatures.

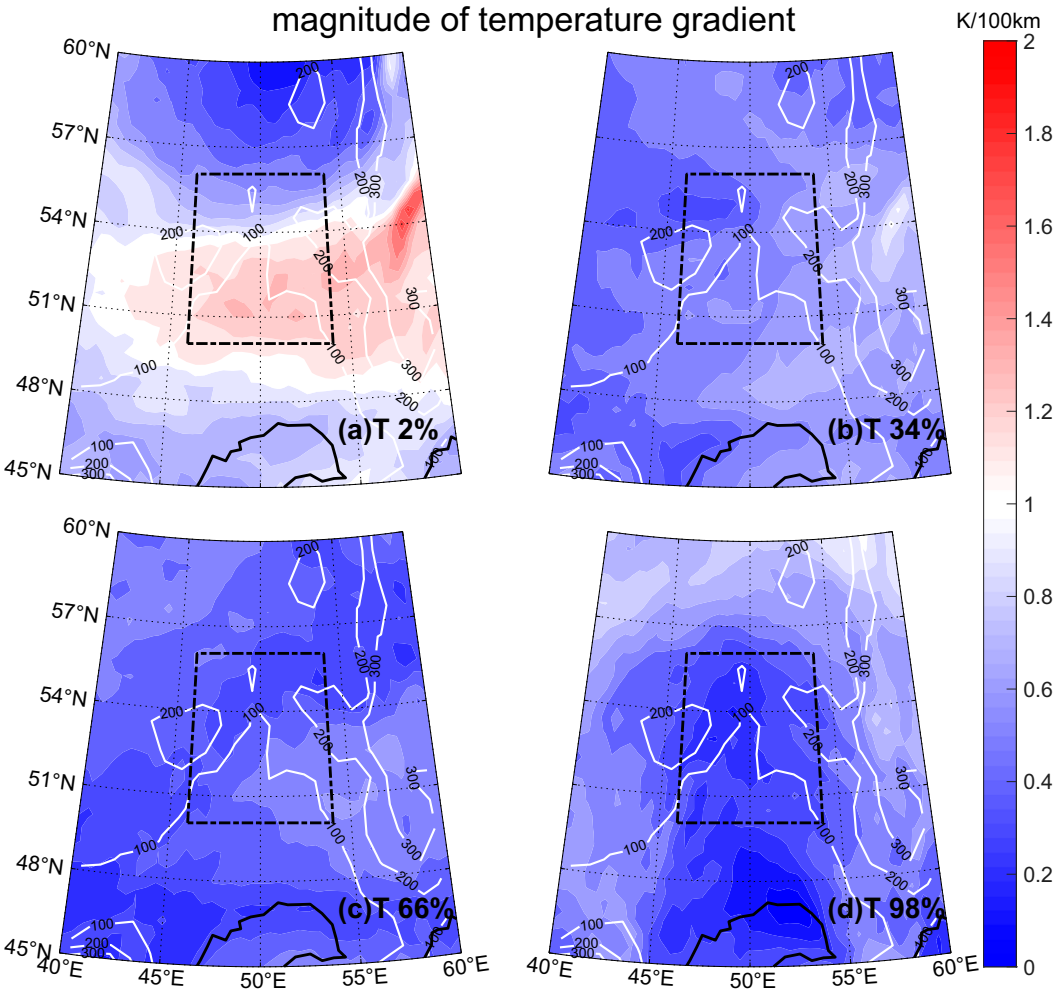

FIG. B1. The magnitude of horizontal temperature gradient $|\nabla T|$ in west Siberia, Russia (Fig. 2c) at different temperature percentiles: (a) 2%, (b) 34%, (c) 66%, and (d) 98%. The white lines are elevation contours, and the dashed black line marks the subregion where we compute regional averages.

REFERENCES

An, Z., and Coauthors, 2015: Global monsoon dynamics and climate change. *Annu. Rev. Earth Planet. Sci.*, **43**, 29–77, https://doi.org/10.1146/annurev-earth-060313-054623.

Catalano, A., P. Loikith, and J. Neelin, 2021: Diagnosing non-Gaussian temperature distribution tails using back-trajectory analysis. *J. Geophys. Res. Atmos.*, **126**, e2020JD033726, https://doi.org/10.1029/2020JD033726.

Chang, C.-P., Y. Zhang, and T. Li, 2000: Interannual and interdecadal variations of the East Asian summer monsoon and tropical Pacific SSTs. Part I: Roles of the subtropical ridge. *J. Climate*, **13**, 4310–4325, https://doi.org/10.1175/1520-0442(2000)013<4310:IAIVOT>2.0.CO;2.

Chen, G., J. Norris, J. D. Neelin, J. Lu, L. R. Leung, and K. Sakaguchi, 2019: Thermodynamic and dynamic mechanisms for hydrological cycle intensification over the full probability distribution of precipitation events. *J. Atmos. Sci.*, **76**, 497–516, https://doi.org/10.1175/JAS-D-18-0067.1.

Garfinkel, C. I., and N. Harnik, 2017: The non-Gaussianity and spatial asymmetry of temperature extremes relative to the storm track: The role of horizontal advection. *J. Climate*, **30**, 445–464, https://doi.org/10.1175/JCLI-D-15-0806.1.

Grotjahn, R., and Coauthors, 2015: North American extreme temperature events and related large scale meteorological patterns: A review of statistical methods, dynamics, modeling, and trends. *Climate Dyn.*, **46**, 1151–1184, https://doi.org/10.1007/s00382-015-2638-6.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hoskins, B., and T. Woollings, 2015: Persistent extratropical regimes and climate extremes. *Curr. Climate Change Rep.*, **1**, 115–124, https://doi.org/10.1007/s40641-015-0020-8.

Huybers, P., K. A. McKinnon, A. Rhines, and M. Tingley, 2014: U.S. daily temperatures: The meaning of extremes in the context of nonnormality. *J. Climate*, **27**, 7368–7384, https://doi.org/10.1175/JCLI-D-14-00216.1.

Li, J., and Q. Zeng, 2000: Significance of the normalized seasonality of wind field and its rationality for characterizing the monsoon. *Sci. China Earth Sci.*, **43**, 646–653, https://doi.org/10.1007/BF02879509.

——, and ——, 2002: A unified monsoon index. *Geophys. Res. Lett.*, **29**, 1274, https://doi.org/10.1029/2001GL013874.

——, and ——, 2003: A new monsoon index and the geographical distribution of the global monsoons. *Adv. Atmos. Sci.*, **20**, 299–302, https://doi.org/10.1007/s00376-003-0016-5.

Linz, M., G. Chen, and Z. Hu, 2018: Large-scale atmospheric control on non-Gaussian tails of midlatitude temperature distributions. *Geophys. Res. Lett.*, **45**, 9141–9149, https://doi.org/10.1029/2018GL079324.

——, ——, B. Zhang, and P. Zhang, 2020: A framework for understanding how dynamics shape temperature distributions. *Geophys. Res. Lett.*, **47**, e2019GL085684, https://doi.org/10.1029/2019GL085684.

Loikith, P. C., and J. D. Neelin, 2015: Short-tailed temperature distributions over North America and implications for future changes in extremes. *Geophys. Res. Lett.*, **42**, 8577–8585, https://doi.org/10.1002/2015GL065602.

Lutsko, N. J., J. W. Baldwin, and T. W. Cronin, 2019: The impact of large-scale orography on Northern Hemisphere winter synoptic temperature variability. *J. Climate*, **32**, 5799–5814, https://doi.org/10.1175/JCLI-D-19-0129.1.

Ma, W., J. Norris, and G. Chen, 2020: Projected changes to extreme precipitation along North American west coast from the CESM large ensemble. *Geophys. Res. Lett.*, **47**, e2019GL086038, https://doi.org/10.1029/2019GL086038.

McKinnon, K. A., A. Rhines, M. P. Tingley, and P. Huybers, 2016: The changing shape of Northern Hemisphere summer temperature distributions. *J. Geophys. Res. Atmos.*, **121**, 8849–8868, https://doi.org/10.1002/2016JD025292.

Norris, J., G. Chen, and J. D. Neelin, 2019a: Changes in frequency of large precipitation accumulations over land in a warming climate from the CESM large ensemble: The roles of moisture, circulation, and duration. *J. Climate*, **32**, 5397–5416, https://doi.org/10.1175/JCLI-D-18-0600.1.

——, ——, and ——, 2019b: Thermodynamic versus dynamic controls on extreme precipitation in a warming climate from the Community Earth System Model large ensemble. *J. Climate*, **32**, 1025–1045, https://doi.org/10.1175/JCLI-D-18-0302.1.

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

Qian, W., 2000: Dry/wet alternation and global monsoon. *Geophys. Res. Lett.*, **27**, 3679–3682, https://doi.org/10.1029/1999GL011255.

Rhines, A., K. A. McKinnon, M. P. Tingley, and P. Huybers, 2017: Seasonally resolved distributional trends of North American temperatures show contraction of winter variability. *J. Climate*, **30**, 1139–1157, https://doi.org/10.1175/JCLI-D-16-0363.1.

Ruff, T. W., and J. D. Neelin, 2012: Long tails in regional surface temperature probability distributions with implications for extremes under global warming. *Geophys. Res. Lett.*, **39**, L04704, https://doi.org/10.1029/2011GL050610.

Schneider, T., and A. H. Sobel, 2007: *The Global Circulation of the Atmosphere*. Princeton University Press, 400 pp.

Stull, R. B., 1988: *An Introduction to Boundary Layer Meteorology*. Springer, 670 pp.

Tamarin-Brodsky, T., K. Hodges, B. J. Hoskins, and T. G. Shepherd, 2019: A dynamical perspective on atmospheric temperature variability and its response to climate change. *J. Climate*, **32**, 1707–1724, https://doi.org/10.1175/JCLI-D-18-0462.1.

——, ——, ——, and ——, 2020: Changes in Northern Hemisphere temperature variability shaped by regional warming patterns. *Nat. Geosci.*, **13**, 414–421, https://doi.org/10.1038/s41561-020-0576-3.

Trenberth, K. E., D. P. Stepaniak, and J. M. Caron, 2000: The global monsoon as seen through the divergent atmospheric circulation. *J. Climate*, **13**, 3969–3993, https://doi.org/10.1175/1520-0442(2000)013<3969:TGMAST>2.0.CO;2.

Wang, B., and Q. Ding, 2006: Changes in global monsoon precipitation over the past 56 years. *Geophys. Res. Lett.*, **33**, L06711, https://doi.org/10.1029/2005GL025347.

——, and ——, 2008: Global monsoon: Dominant mode of annual variation in the tropics. *Dyn. Atmos. Oceans*, **44**, 165–183, https://doi.org/10.1016/j.dynatmoce.2007.05.002.

Zhang, B., M. Linz, and G. Chen, 2022: Interpreting observed temperature probability distributions using a relationship between temperature and temperature advection. *J. Climate*, **35**, 705–724, https://doi.org/10.1175/JCLI-D-20-0920.1.