

# Earth's Future



## RESEARCH ARTICLE

10.1029/2021EF002402

### Special Section:

CMIP6: Trends, Interactions, Evaluation, and Impacts

### Key Points:

- Summer temperature variance increases with warming in Northern Hemisphere midlatitudes in CMIP5 but not in CMIP6
- We develop a decomposition framework and a simple model to explain differences in variance changes across models
- Temperature sensitivity increases in both CMIP5 and CMIP6 but is offset by lower forcing variance in CMIP6

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

D. Chan,  
[duo.chan@whoi.edu](mailto:duo.chan@whoi.edu)

### Citation:

Chan, D., Rigden, A., Proctor, J., Chan, P. W., & Huybers, P. (2022). Differences in radiative forcing, not sensitivity, explain differences in summertime land temperature variance change between CMIP5 and CMIP6. *Earth's Future*, 10, e2021EF002402. <https://doi.org/10.1029/2021EF002402>

Received 2 SEP 2021  
Accepted 24 JAN 2022

### Author Contributions:

**Conceptualization:** Duo Chan, Angela Rigden, Jonathan Proctor, Pak Wah Chan, Peter Huybers

**Formal analysis:** Duo Chan

**Investigation:** Duo Chan

**Methodology:** Duo Chan, Angela Rigden, Jonathan Proctor






**Resources:** Duo Chan

**Software:** Duo Chan

© 2022 The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial License](#), which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

## Differences in Radiative Forcing, Not Sensitivity, Explain Differences in Summertime Land Temperature Variance Change Between CMIP5 and CMIP6

Duo Chan<sup>1,2</sup> , Angela Rigden<sup>2</sup> , Jonathan Proctor<sup>3</sup> , Pak Wah Chan<sup>4</sup> , and Peter Huybers<sup>2</sup> 

<sup>1</sup>Department of Physical Oceanography, Woods Hole Oceanographic Institution, Woods Hole, MA, USA, <sup>2</sup>Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA, <sup>3</sup>Center for the Environment and Data Science Initiative, Harvard University, Cambridge, MA, USA, <sup>4</sup>College of Engineering, Mathematics and Physical Sciences, University of Exeter, Exeter, UK

**Abstract** How summertime temperature variability will change with warming has important implications for climate adaptation and mitigation. CMIP5 simulations indicate a compound risk of extreme hot temperatures in western Europe from both warming and increasing temperature variance. CMIP6 simulations, however, indicate only a moderate increase in temperature variance that does not covary with warming. To explore this intergenerational discrepancy in CMIP results, we decompose changes in monthly temperature variance into those arising from changes in sensitivity to forcing and changes in forcing variance. Across models, sensitivity increases with local warming in both CMIP5 and CMIP6 at an average rate of 5.7 ([3.7, 7.9]; 95% c.i.)  $\times 10^{-3}^{\circ}\text{C per W m}^{-2}$  per  $^{\circ}\text{C}$  warming. We use a simple model of moist surface energetics to explain increased sensitivity as a consequence of greater atmospheric demand ( $\sim 70\%$ ) and drier soil ( $\sim 40\%$ ) that is partially offset by the Planck feedback ( $\sim 10\%$ ). Conversely, forcing variance is stable in CMIP5 but decreases with warming in CMIP6 at an average rate of  $-21$  ( $[-28, -15]$ ; 95% c.i.)  $\text{W}^2 \text{ m}^{-4}$  per  $^{\circ}\text{C}$  warming. We examine scaling relationships with mean cloud fraction and find that mean forcing variance decreases with decreasing cloud fraction at twice the rate in CMIP6 than CMIP5. The stability of CMIP6 temperature variance is, thus, a consequence of offsetting changes in sensitivity and forcing variance. Further work to determine which models and generations of CMIP simulations better represent changes in cloud radiative forcing is important for assessing risks associated with increased temperature variance.

**Plain Language Summary** CMIP5 models show that, in the Northern Hemisphere midlatitudes, summertime temperature variability increases as the surface warms, indicating a compound risk of extreme hot months that have important implications for climate adaptation and mitigation. CMIP6 models, however, show only a moderate increase in temperature variability that is unrelated to warming. To understand this intergenerational discrepancy in CMIP results, we develop a framework to decompose changes in temperature variability into contributions from changes in the variability of external forcing and changes in the sensitivity of temperature to that forcing. We find that both CMIP5 and CMIP6 models show consistent increases in sensitivity as the surface warms, which we demonstrate to arise mainly from warming and drying using a simple diagnostic model. Changes in forcing variability, however, differ between CMIP5 and CMIP6. Whereas, forcing variability is stable in CMIP5, it decreases substantially with warming in CMIP6 and offsets the effect of sensitivity growth. Hence, although midlatitude land surface tends to become more sensitive in all models, whether temperature variability will increase with warming remains uncertain and relies on how forcing variability changes.

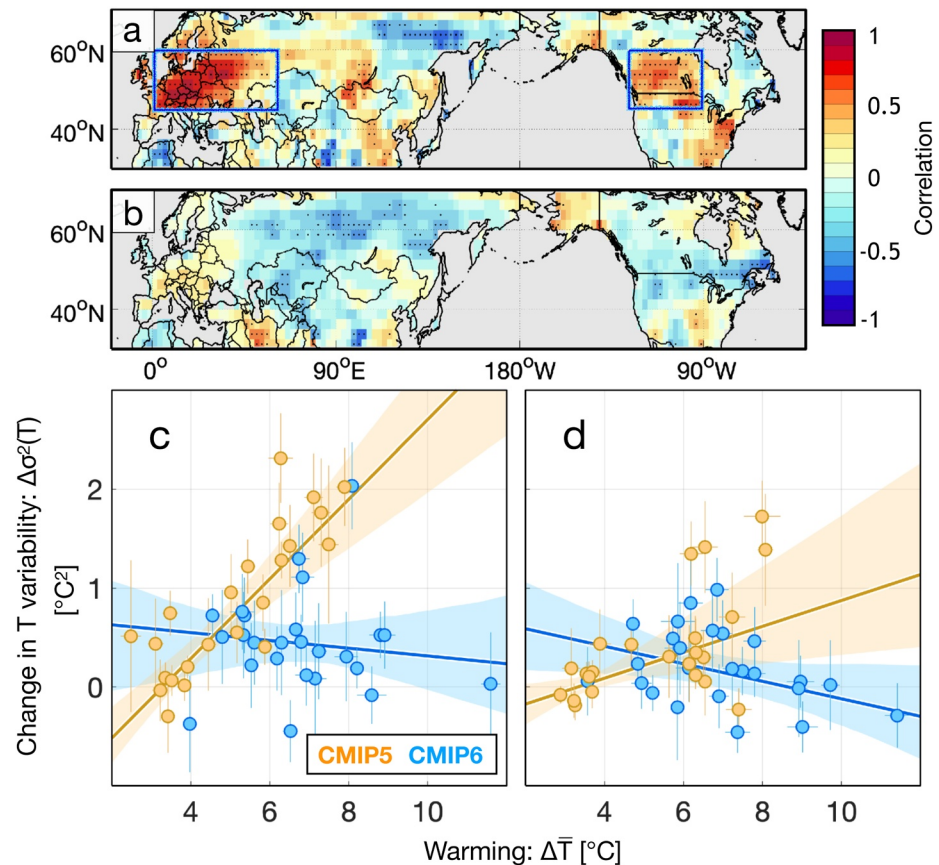
## 1. Introduction

Extreme high temperatures have significant societal and ecological implications (e.g., Deb et al., 2020; Hoag, 2014; Vautard et al., 2020), and their frequency has increased in Northern Hemisphere (NH) midlatitudes (Horton et al., 2016; IPCC, 2012). Increased frequency of extreme high temperatures is largely attributable to increases in mean temperature (Huntingford et al., 2013; Rhines & Huybers, 2013). Although significant changes in summertime temperature variance have not been observed in the historic record (McKinnon et al., 2016), climate model simulations generally indicate greater summertime temperature variance in a warming climate (e.g., Duan et al., 2020; Fischer & Schär, 2009; Gregory & Mitchell, 1995; Rowell, 2005). Increased temperature variance

**Supervision:** Peter Huybers  
**Visualization:** Duo Chan  
**Writing – original draft:** Duo Chan  
**Writing – review & editing:** Duo Chan, Angela Rigden, Jonathan Proctor, Pak Wah Chan, Peter Huybers

would further increase the frequency of extreme heat events, highlighting the importance of accurately predicting changes in summertime temperature variability.

Global climate model predictions of changes in summertime temperature variability,  $\Delta\sigma^2(T)$ , are often highly uncertain, making understanding intermodel differences important. Analyzing monthly temperatures, Chan et al. (2020) reported a positive relationship between local warming,  $\Delta\bar{T}$ , and  $\Delta\sigma^2(T)$  across projections from the Coupled Model Intercomparison Project Phase 5 (CMIP5) (Taylor et al., 2012) in NH midlatitudes (also see Figure 1a). In Europe,  $\Delta\sigma^2(T)$  increases with warming by 0.40 ([0.27, 0.51]; 95% c.i.) $^{\circ}\text{C}^2$  per  $^{\circ}\text{C}$  warming (Table 1),



**Figure 1.** Warming and changes in temperature variance. (a)–(b) Cross-model correlation between  $\Delta\bar{T}$  and  $\Delta\sigma^2(T)$  for (a) CMIP5 and (b) CMIP6. Changes, as indicated by  $\Delta$ , are the future (2071–2100) minus the historic period (1976–2005). We use historical all-forcing experiments for the historical climate and projections under high-emission scenarios for the future climate, that is, RCP8.5 for CMIP5 and SSP585 for CMIP6. For each 30-year period, the mean and linear trend are removed from each summer month, June–July–August (JJA), such that our calculated variance does not include warming signals or changes in seasonal cycles. Dots denote significance at the 95% level. As in Chan et al. (2020), the 23 CMIP5 models used are: ACCESS1-0, ACCESS1-3, CCSM4, CESM1-BGC, CESM1-CAM5, CNRM-CM5, CSIRO-Mk3-6-0, GFDL-CM3, GFDL-ESM2G, GFDL-ESM2M, GISS-E2-H, GISS-E2-H-CC, GISS-E2-R, GISS-E2-R-CC, HadGEM2-CC, HadGEM2-ES, IPSL-CM5A-MR, MIROC5, MRI-CGCM3, MRI-ESM1, NorESM1-M, NorESM1-ME, and inmcm4. The 25 CMIP6 models used are ACCESS-CM2, ACCESS-ESM1-5, AWI-CM-1-1-MR, CESM2-WACCM, CMCC-CM2-SR5, E3SM-1-1, EC-Earth3, EC-Earth3-Veg, EC-Earth3-Veg-LR, FGOALS-f3-L, FGOALS-g3, FIO-ESM-2-0, GFDL-CM4, GFDL-ESM4, INM-CM4-8, INM-CM5-0, IPSL-CM6A-LR, KACE-1-0-G, MIROC6, MPI-ESM1-2-HR, MPI-ESM1-2-LR, MRI-ESM2-0, NIESM3, NorESM2-LM, and NorESM2-MM. We use 2-m air temperatures (tas) from the r11p1/r11p1p1 member of each CMIP5/CMIP6 model. Model results are each first interpolated to a common  $0.5^{\circ}$  grid resolution and averaged to  $2.5^{\circ}$  grid resolution. As in Chan et al. (2020), models whose original resolution is lower than  $2.5^{\circ}$  in either longitude or latitude are excluded. (c)–(d)  $\Delta\sigma^2(T)$  (y-axes) versus  $\Delta\bar{T}$  (x-axes) averaged over (c) Europe and (d) Northern North America. In each panel, orange and blue markers show individual CMIP5 and CMIP6 models, respectively. Associated standard-deviation uncertainties (horizontal and vertical bars on each marker) are estimated using 2-year block bootstrapping (Chan et al., 2020). These error estimates are used as the weight in a York regression (York et al., 2004) to evaluate linear relationships between  $\Delta\sigma^2(T)$  and  $\Delta\bar{T}$  (lines). Associated 95% confidence intervals of the regression (shadings) are estimated by resampling models with replacement. More details of the York regression are in the methods section of Chan et al. (2020).

and the cross-model correlation between  $\Delta\bar{T}$  and  $\Delta\sigma^2(T)$  is 0.85 (orange markers in Figure 1c). Uncertainties are reported as 95% c.i.s if not otherwise specified. Such a positive relationship indicates a compound risk of extreme heat events due to both increases in the mean and variance of temperatures in high-emission scenarios (Chan et al., 2020).

Evaluation of CMIP6, however, shows essentially no relationship between  $\Delta\bar{T}$  and  $\Delta\sigma^2(T)$  (Figure 1b). Compared with CMIP5, the multimodel mean  $\Delta\bar{T}$  is, on average, one-third higher in CMIP6, but  $\Delta\sigma^2(T)$  is only half of that in CMIP5 (also see Figure S1 in Supporting Information S1). Across CMIP6 models,  $\Delta\sigma^2(T)$  slightly decreases with  $\Delta\bar{T}$  at a rate of  $-0.04$  ( $[-0.13, 0.09]$ )°C<sup>2</sup> per °C warming in Europe, and correlation drops to only  $-0.01$  (blue markers in Figure 1c). A positive relationship between  $\Delta\bar{T}$  and  $\Delta\sigma^2(T)$  also disappears in CMIP6 over Northern North America (Figure 1d).

This intergenerational difference in changes in temperature variance would have distinct implications for assessing risks associated with extreme high-temperature events in a warming climate. It is, therefore, important to understand why the two CMIP generations differ. Moreover, understanding differences among models and model generations provides insight into the mechanisms of summertime temperature variability and have implications for improving climate models.

We inquire into the processes responsible for distinct projections of  $\Delta\sigma^2(T)$  between CMIP5 and CMIP6, of which there are two broad categories: land-surface coupling involving evapotranspiration (Duan et al., 2020; Seneviratne et al., 2010; Vargas Zeppetello & Battisti, 2020), and atmospheric processes involving clouds (Lenderink et al., 2007), water-vapor feedback (Held & Soden, 2000; Philipona et al., 2005), and large-scale circulation (Holmes et al., 2016). Recently, Vargas Zeppetello et al. (2020) developed a simple diagnostic model, here referred to as VZ20, to understand the origin of monthly temperature variance in NH mid-latitudes. The VZ20 model captures spatial features of multimodel mean  $\Delta\sigma^2(T)$  that are attributable to land-surface coupling associated with increases in atmospheric water vapor deficit and attendant changes in evapotranspiration (Vargas Zeppetello & Battisti, 2020). On account of the complexity of full global climate models and the diversity of land-surface representations, such a simple diagnostic framework is important for understanding differences across models. The degree to which VZ20 and land-surface coupling explain inter-model and inter-generational differences in  $\Delta\sigma^2(T)$ , however, has not previously been reported.

In this study, we develop a framework that decomposes changes in temperature variance,  $\Delta\sigma^2(T)$ , into contributions from changes in forcing variance,  $\Delta\sigma^2(F)$ , and changes in temperature sensitivity to that forcing,  $\Delta\frac{\partial T}{\partial F}$ . Such a decomposition can indicate the processes responsible for divergent temperature variance predictions and provide insights for improving future model generations and projections. We show (Section 2) that both CMIP5 and CMIP6 models have consistent sensitivity growth with local warming. However, unlike CMIP5, which has stable forcing variance, CMIP6 models have substantially decreased forcing variance, which cancels sensitivity growth and stabilizes temperature variability. We show that a revised version of the VZ20 model explains the different projections of  $\Delta\sigma^2(T)$  across CMIP models (Section 3) and then use it to diagnose the specific channels through which sensitivity increases. We also show (Section 4) that clouds are partially responsible for distinct predictions of forcing variance before discussing and concluding (Section 5).

## 2. Decomposing Changes in Temperature Variance

To better understand similarities and differences between the two generations of CMIP models, we first decompose each temperature anomaly into the anomaly of radiative forcing at the land surface and the sensitivity of surface air temperature to that forcing anomaly,

$$T' = \frac{\partial T}{\partial F} F', \quad (1)$$

where,  $T'$  denotes monthly temperature anomalies relative to a 30-year monthly climatology and  $F'$  denotes forcing anomalies. Forcing comprises net shortwave radiation,  $S_n$ , and downward longwave radiation,  $L_d$ . Anomalies are measured at each 2.5° grid box during JJA after removing respective means and 30-year trends in individual months. Compared with radiative forcing, advective processes have only a weak influence on monthly

**Table 1**  
Cross-Model Relationship Between  $\Delta\sigma^2(T)$  and  $\Delta\bar{T}$

	CMIP5		CMIP6	
	$R$	York ( $^{\circ}\text{C}^2$ per $^{\circ}\text{C}$ )	$r$	York ( $^{\circ}\text{C}^2$ per $^{\circ}\text{C}$ )
EU	0.85	0.40 [0.27, 0.51]	-0.01	-0.04 [-0.13, 0.09]
NNA	0.61	0.13 [0.04, 0.26]	-0.36	-0.09 [-0.17, -0.00]
Only changing sensitivity— $\sigma_h^2(F)\Delta S$				
EU	0.90	0.43 [0.36, 0.51]	0.67	0.35 [0.21, 0.50]
NNA	0.80	0.23 [0.17, 0.32]	0.73	0.25 [0.17, 0.36]
Only changing forcing variance— $S_h\Delta\sigma^2(F)$				
EU	-0.18	-0.03 [-0.10, 0.05]	-0.55	-0.18 [-0.30, -0.11]
NNA	-0.34	-0.08 [0.17, 0.02]	-0.75	-0.22 [-0.30, -0.16]

*Note.* Shown statistics are correlations and regression slopes of  $\Delta\sigma^2(T)$  against  $\Delta\bar{T}$  over Europe (EU) and Northern North America (NNA). To account for regression dilution, we use a York regression technique (York et al., 2004) to estimate the slope. In addition to  $\Delta\sigma^2(T)$  directly calculated from temperature outputs (1<sup>st</sup>–2<sup>nd</sup> rows), also shown are the decomposition of  $\Delta\sigma^2(T)$  according to Equation 4, where only sensitivity (3<sup>rd</sup>–4<sup>th</sup> rows) or forcing variance (5<sup>th</sup>–6<sup>th</sup> rows) are allowed to change. Numbers in brackets indicate 95% confidence intervals.

summertime temperatures over noncoastal continents (Holmes et al., 2016; Vargas Zeppetello et al., 2022). For example, the month-to-month variance of the residual between radiative forcing,  $F$ , and associated response,  $R$ , is less than 10% of the variance in  $F$ .  $R$  comprises upward longwave, latent, and sensible heat fluxes, and residuals are computed between time series of  $F$  and  $R$  from individual continental grid boxes for each model. We, therefore, focus on net solar and downward longwave radiation in  $F$ .

Taking the variance of both sides of Equation 1 gives,

$$\sigma^2(T) = S\sigma^2(F), \quad (2)$$

where,  $\sigma^2$  denotes the variance of respective monthly quantities.  $S$  is the square of the sensitivity of temperature to forcing, which we calculate using the ratio of the variance of temperature to forcing,

$$S = \frac{\sigma^2(T)}{\sigma^2(F)}. \quad (3)$$

A complementary means of estimating sensitivity is to fit for a slope between temperature and forcing anomalies. Ordinary least squares, however, would suffer from regression dilution (Fuller, 2009), on account of noise in forcing estimates, for example, from neglecting advective processes (Holmes et al., 2016). One way to account for regression dilution is total least square (TLS) regression (Markovsky & Van Huffel, 2007), which minimizes the sum of distance perpendicular to the fitted slope. Because forcing and tem-

perature have different units, it is natural to standardize both variables before TLS. The solution to this normalized TLS is one because temperature and forcing anomalies are positively correlated. After inverse normalization, the TLS solution is identical to Equation 3.

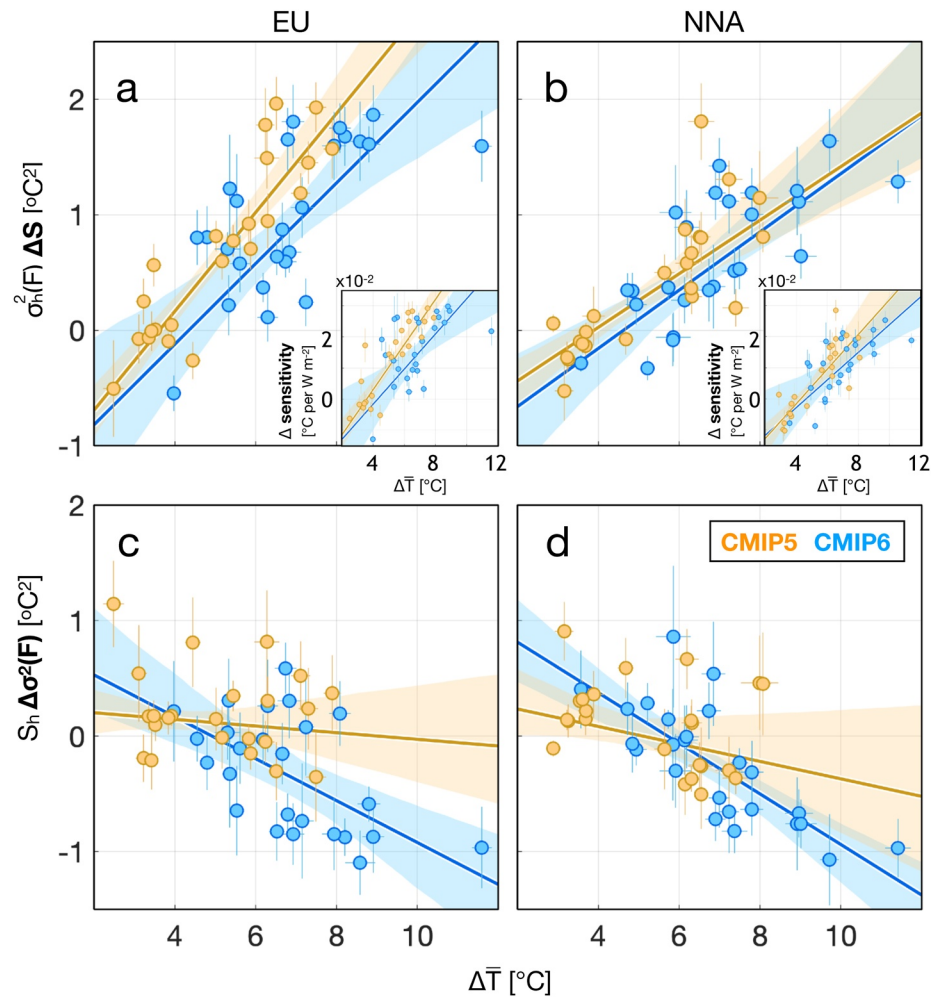
To decompose changes in temperature variance into contributions from changing sensitivity and changing forcing variance, we take the difference of Equation 2 between the future (2071–2100) and historic (1976–2005) periods,

$$\Delta\sigma^2(T) = \sigma_h^2(F)\Delta S + S_h\Delta\sigma^2(F) + O(\Delta^2), \quad (4)$$

where, a subscript  $h$  denotes historical values that we evaluate over 1976–2005. The three terms on the right-hand-side of Equation 4 denote respective contributions from changing sensitivity, changing forcing variance, and higher-order interactions, discussed below.

Applying Equation 4 to CMIP outputs, we find that sensitivity-induced changes in temperature variance,  $\sigma_h^2(F)\Delta S$ , increases with local warming in NH midlatitudes for both CMIP generations (Figures 2a and 2b). In Europe,  $\sigma_h^2(F)\Delta S$  increases with local warming at rates of 0.43 ([0.36, 0.51]) and 0.35 ([0.21, 0.50]) $^{\circ}\text{C}^2$  per  $^{\circ}\text{C}$  warming for CMIP5 and CMIP6, respectively (Figure 2a). Positive relationships are also evident in Northern North America and are consistent between CMIP generations (Figure 2b). Examining only sensitivity change confirms that changes in sensitivity, rather than intermodel differences in historic forcing variance, dominate  $\sigma_h^2(F)\Delta S$  (small panels in Figures 2a and 2b). These results show that, in both CMIP generations, land surface temperature tends to be more sensitive to forcing as it warms.

On the other hand, forcing-induced changes in temperature variance,  $S_h\Delta\sigma^2(F)$ , are distinct between CMIP5 and CMIP6 (Figures 2c and 2d). In Europe,  $S_h\Delta\sigma^2(F)$  decreases with warming at a rate of -0.03 ([-0.10, 0.05]) $^{\circ}\text{C}^2$  per  $^{\circ}\text{C}$  warming across CMIP5 models but -0.18 ([-0.30, -0.11]) $^{\circ}\text{C}^2$  per  $^{\circ}\text{C}$  warming across CMIP6 models (Figure 2c). Decreases in the European forcing variance range beyond -80  $\text{W}^2 \text{m}^{-4}$  in six of the CMIP6 models, or more than twice the largest decrease,  $\sim -40 \text{W}^2 \text{m}^{-4}$ , in the CMIP5 ensemble. Similar discrepancies in changes of forcing variance are also evident in Northern North America (Figure 2d). As a result, a decrease in the variance of radiative forcing with warming in CMIP6 models cancels out the rise in sensitivity to give near constant temperature variance.



**Figure 2.** Decomposing changes in temperature variance. (a)–(b) Regional changes in temperature variance associated with changing sensitivity of temperature to forcing (y-axes) versus local warming (x-axes) for (a) Europe and (b) Northern North America. (c)–(d) Regional changes in temperature variance associated with changing forcing variance (y-axes) versus local warming (x-axes) for (c) Europe and (d) Northern North America. The setup of individual panels is the same as panel (c) in Figure 1. Small panels in (a) and (b) show changes in sensitivity (y-axes) against warming (x-axes) in respective regions. Changes in forcing variance versus warming are in Figures 8a and 8b.

Contributions from higher-order interactions are smaller than first-order terms. Compared with a range from  $-0.5^{\circ}$  to  $2.0^{\circ}\text{C}^2$  in  $\sigma_h^2(F)\Delta S$ ,  $O(\Delta^2)$  ranges only between  $-0.2^{\circ}$  and  $0.4^{\circ}\text{C}^2$  for CMIP5 models in Europe. For CMIP6 models with high warming rates, because the effects of changes in sensitivity and forcing variance cancel each other (Figures 2a and 2c), their multiplication works in the same direction as  $S_h\Delta\sigma^2(F)$  to decrease temperature variance. For the six models whose forcing variance decreases by more than  $80\text{ W}^2\text{ m}^{-4}$  in Europe, the contribution from  $O(\Delta^2)$  is, on average, 60% of that from  $S_h\Delta\sigma^2(F)$ .

### 3. Sensitivity Growth and Suppressed Evapotranspiration

VZ20 allows for diagnosing summertime land temperature variance. In this section, we develop a revised version of VZ20, which we refer to as VZ20r, to quantify how suppressed evapotranspiration and associated processes lead to sensitivity growth in both CMIP5 and CMIP6. Compared with VZ20, in addition to diagnosing changes in temperature variance, VZ20r also allows the diagnosis of temperature sensitivity to radiative forcing, as appears in the decomposition in Equation 1. Moreover, we also show that VZ20r better reproduces differences in simulated  $\Delta\sigma^2(T)$  across CMIP models (Section 3.2).

### 3.1. Developing VZ20r From VZ20

VZ20 is derived in Vargas Zeppetello et al. (2020); here, we provide a summary of that derivation to establish nomenclature and then describe the revisions made to obtain VZ20r. A full derivation of VZ20r can be found in Appendix A.

VZ20 is premised on a quasi-equilibrium in surface energy and water budgets at monthly time scales. In the energy budget, net shortwave radiation is assigned as forcing, which is balanced by net longwave radiation, surface fluxes, and heat fluxes into the ground. Latent heat flux is parameterized as a product of atmospheric demand,  $V$ , quantified as the difference between saturated and actual specific humidity, and the saturation rate of surface soil,  $m$ . Anomalous net longwave radiation,  $L_n$ , sensible heat flux,  $H$ , and the flux into the ground,  $G$ , are parameterized as being proportional to temperature anomalies.

After solving for  $m$  from the surface water budget and substituting back to the energy budget, the VZ20 model is,

$$S'_n - \lambda\gamma P' = \left[ h + \frac{\rho_a}{r_0} \gamma \bar{m} (1 - \lambda) \frac{\partial q^*}{\partial T} \right] T', \quad (5)$$

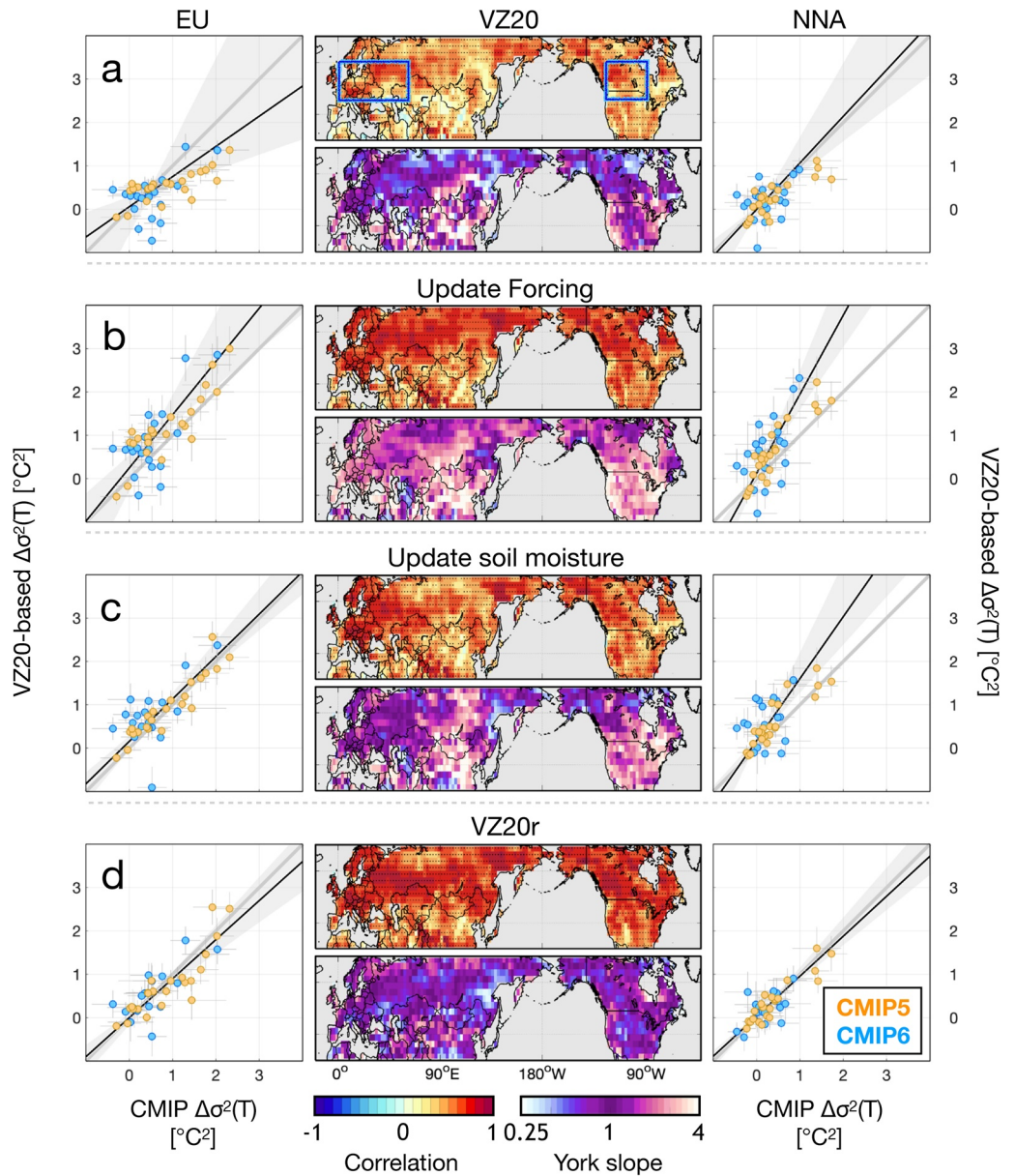
where, an over-bar denotes 30-year climatology and a prime denotes monthly anomalies.  $S'_n$  denotes anomalies in net shortwave radiation, which is the radiative forcing in VZ20. A second form of forcing in the VZ20 model comes from precipitation anomalies,  $P'$ , whose energy implications are represented using the heat of vapourization,  $\gamma$ , and a dryness index,  $\lambda$ .  $\lambda$  gives the proportion of potential water loss through evapotranspiration and ranges between zero and one. A higher  $\lambda$  is associated with a preference for evapotranspiration over runoff and indicates a drier climate (Vargas Zeppetello & Battisti, 2020). Note that  $\lambda$  and soil moisture are anticorrelated but are different conceptually. Whereas soil moisture represents water availability,  $\lambda$  reflects primarily atmospheric demand.

The right-hand-side (RHS) of Equation 5 represents longwave and sensible heat flux as well as evapotranspiration as being proportional to surface temperature anomalies,  $T'$ . The summed dependence of net-longwave, sensible heat and ground heat flux on  $T'$  is represented by  $h$ , such that  $L'_n + H' + G' = hT'$ . Other terms on the RHS denote the dependence of evapotranspiration on  $T'$ , where this dependence increases with mean soil saturation,  $\bar{m}$ , and decreases with the dryness index,  $\lambda$ . Density of air,  $\rho_a$ , and a minimum surface resistance parameter,  $r_0$ , are used in parameterizing evapotranspiration. The dependence of saturation specific humidity on temperature,  $\frac{\partial q^*}{\partial T}$ , which is evaluated at the climatological temperature, is required for evaluating anomalies in  $V$ .

As discussed in Vargas Zeppetello and Battisti (2020),  $\lambda$  is key to interpreting Equation 5. As  $\lambda$  increases and atmospheric demand grows, the surface energy balance moves from an energy-limited regime to a water-limited regime. In the energy-limited regime, positive radiative forcing anomalies increase evapotranspiration by increasing atmospheric demand. Resulting increases in latent heat flux offset forcing anomalies and damp temperature variability. Conversely, in the water-limited regime rainfall anomalies are more capable of driving anomalous latent heat flux. Because rainfall covaries negatively with insolation, decreases in latent heat anomalies tend to combine with positive anomalies in radiative forcing to increase the magnitude of temperature anomalies.

Taking the square of Equation 5 gives an expression for  $\sigma^2(T)$  that is evaluated using the variance and covariance of  $S'_n$  and  $P$  sampled from CMIP5 or CMIP6 simulations, along with sampled values for  $\bar{m}$ ,  $\bar{T}$ , and  $\bar{q}$ . Here,  $q$  denotes specific humidity, which is used together with  $\bar{T}$  to calculate  $V$  and  $\lambda$ . Although VZ20 reproduces the pattern of multimodel mean  $\Delta\sigma^2(T)$  (Vargas Zeppetello & Battisti, 2020), it has little skill in reproducing the difference in  $\Delta\sigma^2(T)$  across CMIP models, which we evaluate using Pearson's cross-correlation,  $r$ , on an individual grid-box basis across-models. On average, the  $r^2$  between CMIP and VZ20-based  $\Delta\sigma^2(T)$  over NH midlatitude (30°–70°N) continents is 0.20 for VZ20 (Figure 3a).

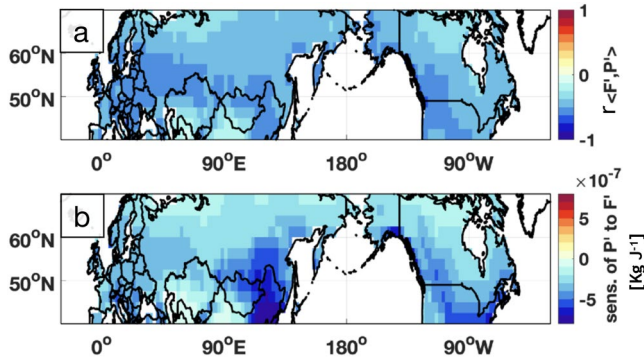
We make four changes to VZ20. First, we extend the representation of radiative forcing to include downward longwave fluxes,  $L_1$ , in addition to  $S'_n$ . This sum of  $L_1$  and  $S'_n$  gives the same  $F$  that is referred to in Equation 1. Vargas Zeppetello et al. (2019) noted that  $L_1$  involves clouds, water vapor, and air temperature and is closely related to surface temperature anomalies, and that  $L_1$  should be a response rather than forcing. Atmospheric heat convergence in the free troposphere, however, also leads to changes in  $L_1$ , making it difficult to assign causality, particularly at regional and monthly scales. As we show in Section 3.2, significant increases in model skill



**Figure 3.** Improvements from revising VZ20. Individual rows are for (a) the original VZ20, (b) assigning forcing to consist of  $S_n$  and  $L_1$  rather than only  $S_n$  (update forcing), (c) using soil moisture rather than precipitation to represent soil saturation (update soil moisture), (d) further converts precipitation anomalies into radiative forcing anomalies,  $P' = \beta F'$  (VZ20r). In each row, the two maps show, respectively, cross-model correlation (upper) and York regression slope (lower) between CMIP and VZ20-based  $\Delta\sigma^2(T)$ . Statistics are evaluated after pooling all 48 CMIP models. A higher correlation (dark red) and a slope closer to one (purple) indicate better skill. The two scatter plots show regional mean VZ20-based (y-axes) versus CMIP (x-axes)  $\Delta\sigma^2(T)$  over Europe (left) and Northern North America (right). Also shown are York regression slopes (black lines) and associated 95% confidence intervals (gray shadings).

associated with updating forcing indicate that  $L_1$ , in addition to  $S_n$ , constitutes a substantial component of effective radiative forcing.

Second, we parameterize the upward longwave radiative response using the Stefan-Boltzmann law,  $L_1 = \sigma T^4$ . Note that here,  $\sigma$  denotes the Stefan-Boltzmann constant rather than  $\sigma(\cdot)$  that we use to denote variability. An additional term,  $4\sigma\bar{T}^3 T'$ , is thus appended to the right-hand-side of Equation 5, and  $h$  is now made to only represent sensible heat and ground fluxes.



**Figure 4.** Covariability of precipitation and radiative forcing anomalies. (a) Correlation between monthly radiative forcing,  $F'$  and precipitation anomalies,  $P'$ . (b) Sensitivity of  $P'$  to  $F'$  in Northern Hemisphere midlatitudes. For both maps, statistics are first evaluated using 30-year outputs within each model and then averaged over all models and both periods.

Third, we explicitly represent the saturation of surface soil. Whereas, VZ20 approximates  $m$  using climatological precipitation divided by the maximum value over North America (Vargas Zeppetello et al., 2020), we use surface soil moisture,  $s$  (called “mrsos” in CMIP outputs). Note that CMIP models have widely variable magnitudes of  $s$ . For example, NorESM2-MM has a historical global average  $s$  of  $28 \text{ kg m}^{-2}$ , whereas FGOALS-f3-L has only  $0.22 \text{ kg m}^{-2}$ . Models having low  $s$  still show negative correlations between monthly soil moisture and latent heat anomalies in humid areas, indicating a reasonable representation of land-air coupling (Berg & Sheffield, 2018). Apparently, whereas “mrsos” is listed as total water mass in all forms in the top 10 cm of the soil layer (Taylor et al., 2012), different CMIP models integrate over different soil depths. We, therefore, normalize surface soil moisture in each model by linearly mapping values between each model's 0.1% and 95% quantile of the 1976–2005 climatology to range between 0 and  $1 \text{ kg m}^{-2}$ . Results are not qualitatively sensitive to the choice of the upper quantile of mapping, for which we also tried 90% and 99% values.

Finally, to permit a closed expression for temperature sensitivity to radiative forcing, we combine radiative forcing and precipitation anomalies on

the left-hand-side of Equation 5 by assuming that  $P' = \beta F'$ . Combining  $F'$  and  $P'$  is suggested by the fact that cloud radiative properties and rainfall generally covary negatively (Figure 4a). In Europe and northern North America, the correlation between  $F'$  and  $P'$  is, on average,  $-0.53$  and  $-0.48$ , respectively. Specifically, we assign  $\beta = -4 \times 10^{-7} \text{ kg J}^{-1}$ , a typical value diagnosed from CMIP simulations (Figure 4b). Substituting  $P'$  with  $F'$  gives VZ20r,

$$(1 - \beta\lambda\gamma)F' = \left[ 4\sigma\bar{T}^3 + h + \frac{\rho_a}{r_0}\gamma\bar{s}(1 - \lambda)\frac{\partial q^*}{\partial T} \right] T'. \quad (6)$$

VZ20r-based sensitivity of temperature to radiative forcing is,

$$\frac{\partial T}{\partial F} = \frac{1 - \beta\lambda\gamma}{4\sigma\bar{T}^3 + h + \frac{\rho_a}{r_0}\gamma\bar{s}(1 - \lambda)\frac{\partial q^*}{\partial T}}. \quad (7)$$

In evaluating Equation 7 we use CMIP climatological temperature,  $\bar{T}$ ; surface soil moisture,  $\bar{s}$ ; and specific humidity,  $\bar{q}$ . Other parameters are given in Table S1 in Supporting Information S1. VZ20r-based  $\sigma^2(T)$  is calculated by multiplying VZ20r-based sensitivity with CMIP  $\sigma^2(F)$ , calculated over 30-year periods.

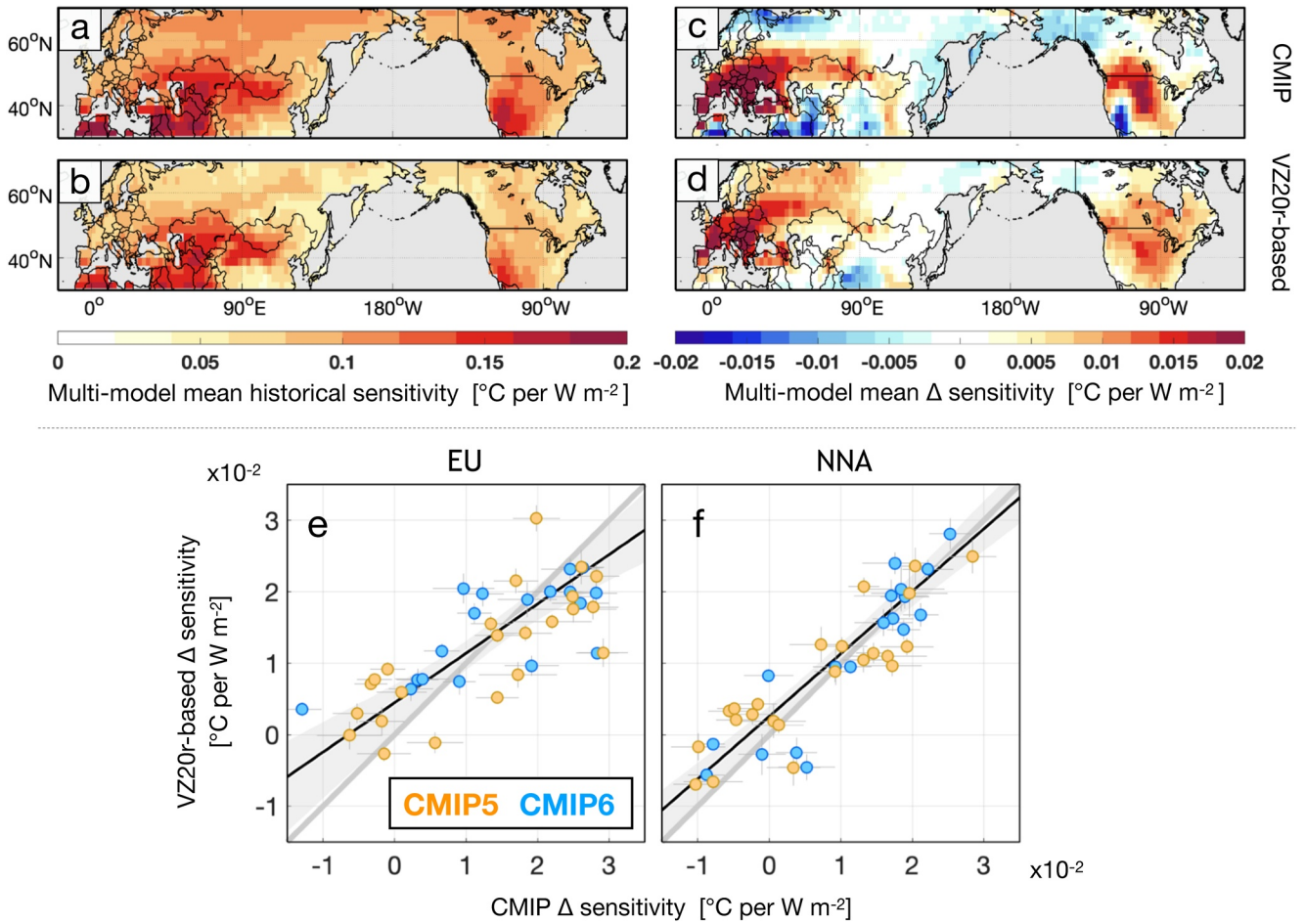
### 3.2. Skill of VZ20r

VZ20r has greater skill than VZ20 in capturing cross-model differences in  $\Delta\sigma^2(T)$  over NH midlatitudes. The averaged  $r^2$  increases from 0.20 in VZ20 to 0.43 in VZ20r (correlation maps in Figures 3a and 3d). The improvement comes mainly from including  $L_1$  in forcing, increasing  $r^2$  by 0.16 (Figure 3b), and converting precipitation anomalies into forcing anomalies, increasing  $r^2$  by another 0.11 (Figure 3d). Using normalized surface soil moisture, however, slightly decrease  $r^2$  by 0.04. Lack of improvement may reflect the fact that inter-model differences in rainfall and soil moisture variations are highly correlated. Improvements in  $r^2$  are also accompanied by improved agreement in the VZ20r predictions of regional CMIP  $\Delta\sigma^2(T)$  (scatter plots in Figure 3).

VZ20r also reproduces the multi-model mean pattern of historical CMIP sensitivity, which generally increases with aridity from  $0.08^\circ\text{C per W m}^{-2}$  in the Arctic to approximately  $0.2^\circ\text{C per W m}^{-2}$  in deserts (Figures 5a and 5b). Moreover, VZ20r reproduces the pattern of multimodel mean changes in sensitivity, which has the largest increase in Europe and central North America (Figures 5c and 5d). Most importantly, VZ20r reproduces cross-model differences in sensitivity change, with the cross-model correlation between VZ20r-based and CMIP sensitivity change being 0.75 in Europe and 0.89 in Northern North America (Figures 5e and 5f).

The skill of VZ20r is relatively low in Arctic and arid regions. An important process that is lacking in VZ20r is snow cover at high-latitudes. Another factor is that the negative  $\beta$  we use does not capture the covariance between





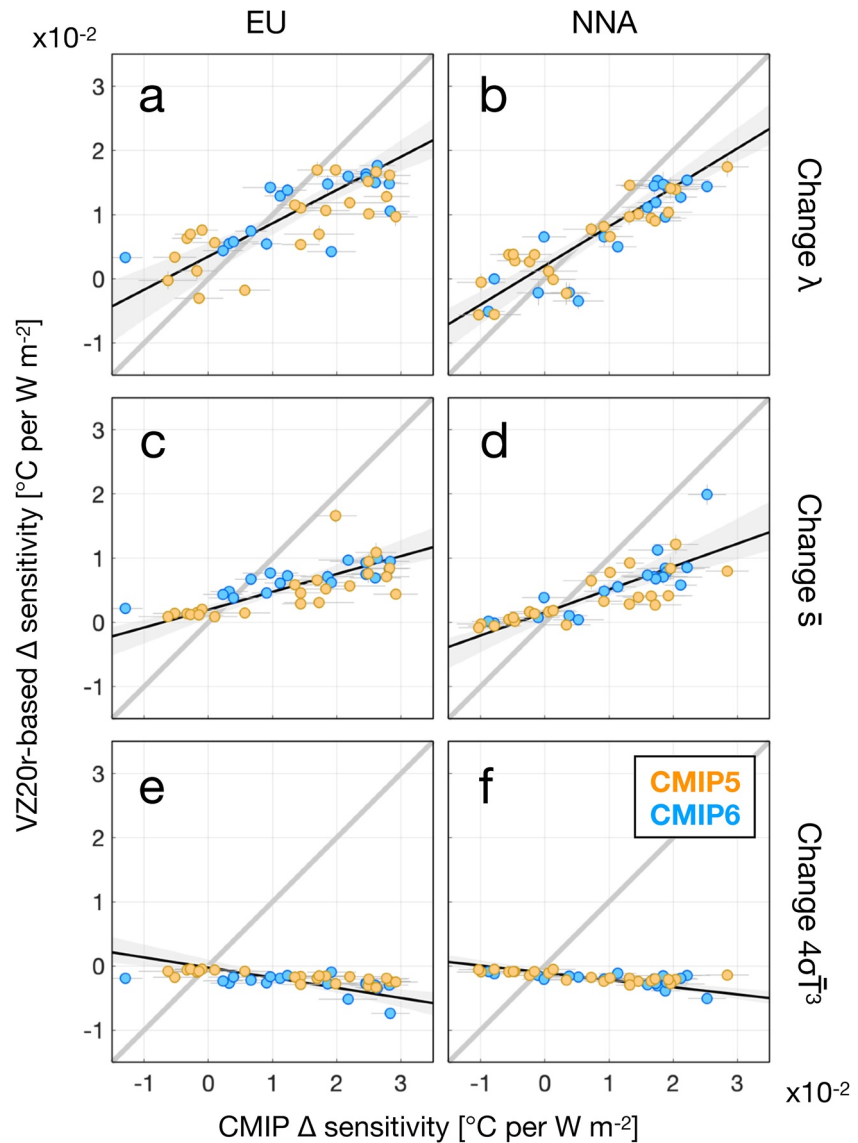
**Figure 5.** VZ20r reproduces multimodel mean and cross-model difference in sensitivity. (a) CMIP and (b) VZ20r-based multimodel mean historical sensitivity. (c)–(d) As (a)–(b) but for multimodel mean sensitivity change. (e)–(f) As scatter plots in Figure 3 but for VZ20r-based (y-axes) versus CMIP (x-axes) sensitivity change in (e) Europe and (f) Northern North America.

precipitation and radiative forcing expected in arid regions whereby, a positive rainfall anomaly increases water vapor in the air column and hence  $L_1$ . Nevertheless, VZ20r evidently captures the essential processes for understanding distinct sensitivity change over the NH midlatitudes, where the transition of evapotranspiration from a more energy-limited to a more water-limited regime takes place (Berg & Sheffield, 2018; Seneviratne et al., 2010; Vargas Zeppetello & Battisti, 2020).

### 3.3. Sensitivity Growth Is Consistently Due To Suppressed Evapotranspiration in CMIP5 and CMIP6

VZ20r indicates that suppressed evapotranspiration primarily controls sensitivity growth in both CMIP5 and CMIP6. According to Equation 7, sensitivity is a function of  $\lambda$ ,  $\bar{s}$ , and  $4\sigma\bar{T}^3$ . We decompose the contributions from each of these three factors by changing each factor individually while keeping others fixed at their historical values (Figure 6).

Both  $\lambda$ , the dryness index, and  $\bar{s}$ , surface soil moisture, are related to evapotranspiration. On average, changing  $\lambda$  explains  $\sim 70\%$  of the cross-model difference in sensitivity growth in both Europe and northern North America (Figures 6a and 6b), and changing  $\bar{s}$  contributes another  $\sim 40\%$  (Figures 6c and 6d). Furthermore, the consistent fractional contribution from  $\lambda$  and  $\bar{s}$  between CMIP5 and CMIP6 indicates that evapotranspiration processes are represented consistently between CMIP5 and CMIP6. The Planck factor,  $4\sigma\bar{T}^3$ , shows a small decrease with local warming (Figures 6e and 6f) and is consistent between CMIP5 and CMIP6.

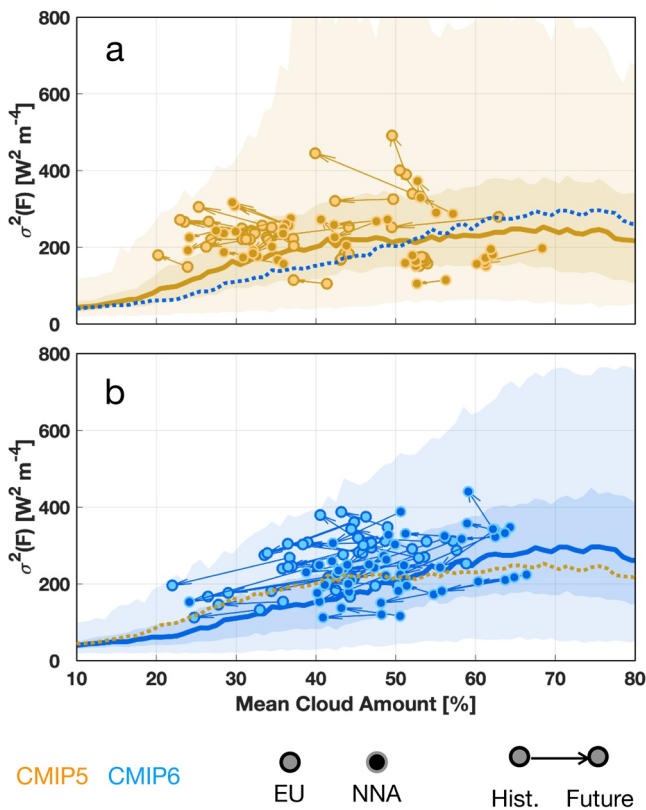


**Figure 6.** Decomposing sensitivity change. Similar to the bottom panels in Figure 5 but for decomposing VZ20r-based sensitivity change by (a–b) only changing the dryness index  $\lambda$ , (c–d) only changing surface soil moisture, and (e–f) only changing the Planck factor. The left and right columns are for Europe and Northern North America, respectively.

The fact that warm and water-limited regimes suppress evapotranspiration and amplify temperature anomalies is well recognized (e.g., Duan et al., 2020; Fischer et al., 2012; Lenderink et al., 2007; Rasmijn et al., 2018; Senviratne et al., 2010; Vargas Zeppetello & Battisti, 2020). Our contribution is to develop a revised framework to reproduce and understand differences in predictions among CMIP models. Our results indicate that processes related to land-air coupling and evapotranspiration, which will make surface temperatures more sensitive to radiative forcing under a warming and drying climate, can be captured by VZ20r and appear consistent between CMIP5 and CMIP6.

#### 4. Distinct Changes in Forcing Variability Between CMIP Generations

We showed in Section 2 that decreases in forcing variability cancel the effect of increasing sensitivity and result in a more stable temperature variance in CMIP6 models. Changes in forcing variance,  $\Delta\sigma^2(F)$ , can be understood in the context of changes in the percentage of cloud coverage. In both CMIP generations  $\sigma^2(F)$  increases with monthly mean cloud coverage until approximately 70% and then decreases (Figure 7). Stated differently, regions



**Figure 7.** Forcing variance and mean cloud amount. The distribution of radiative forcing variance versus mean cloud coverage percent for (a) CMIP5 (orange) and (b) CMIP6 (blue). Distributions are across all continental grids over both the historical (1976–2005) and the future (2071–2100) period and across all models within a CMIP generation; indicated are the median (solid line), interquartile range (dark shading), and 95% range (light showing). For comparison the median from the alternate generation is also shown (dashed line). The dependence of forcing variance on cloud percent is generally flatter in CMIP5 than CMIP6. Markers indicate averages over Western Europe (light) and Northern North America (dark) for individual models, with arrows pointing from the historical to future period.

with consistent clear-sky or fully cloudy conditions have little variability in radiative forcing. Cloud height and albedo will also influence  $\sigma^2(F)$ , but we do not undertake such a detailed analysis in the present context.

Despite the structural similarity, there are several important differences in the distribution of forcing variance conditional on mean cloud percent between CMIP5 and CMIP6. First, median  $\sigma^2(F)$  in CMIP6 has an absolute sensitivity that is higher than in CMIP5. For example, a decrease in cloud percent from 60% to 30% is associated with a median  $\sigma^2(F)$  decrease of  $150 \text{ W}^2 \text{ m}^{-4}$  in CMIP6 but only a  $75 \text{ W}^2 \text{ m}^{-4}$  decrease in CMIP5. Second, CMIP6 models have, on average, larger decreases in mean cloud amount in NH midlatitudes than CMIP5 models, which is presumably also related to their higher warming rates that decrease relative humidity and suppress cloud formation (Figure S2 in Supporting Information S1; Emanuel, 2008; Heymsfield et al., 1998; Price & Wood, 2002). Finally, the distribution of cloud variance with mean cloud amount is narrower in CMIP6 and changes in forcing variance with changes in cloud amount are more uniform. We note that such differences may reflect different cloud parameterization schemes (e.g., Jian et al., 2020; Vignesh et al., 2020; Zelinka et al., 2020), though a detailed exploration is beyond the scope of this study.

It is also helpful to decompose forcing variance into the variance of net solar radiation,  $\sigma^2(S_n)$ , downward longwave radiation,  $\sigma^2(L_\downarrow)$ , and the covariance between the two (Figure 8). We consider the contribution of each component to the scaling relationship between  $\Delta\sigma^2(F)$  and  $\Delta\bar{T}$  in the two CMIP generations. Whereas CMIP6 gives larger decreases in  $\sigma^2(F)$  than CMIP5 in Europe and northern North America, contributions from individual forcing components vary with region. Solar radiation contributions are key to differences between CMIP5 and CMIP6 in Northern North America. Specifically,  $\sigma^2(S_n)$  accounts for 7.9 out of a total difference of  $11.6 \text{ W}^2 \text{ m}^{-4}$  per  $^\circ\text{C}$  warming in the relationship between  $\Delta\sigma^2(F)$  and  $\Delta\bar{T}$  (Figures 8d, 8f, and 8h). In western Europe, however,  $\sigma^2(S_n)$ ,  $\sigma^2(L_\downarrow)$ , and their covariance contribute, respectively, 4.4, 4.0, and  $5.6 \text{ W}^2 \text{ m}^{-4}$  per  $^\circ\text{C}$  warming to the difference in the relationship between  $\Delta\sigma^2(F)$  and  $\Delta\bar{T}$  (Figures 8c, 8e, and 8g).

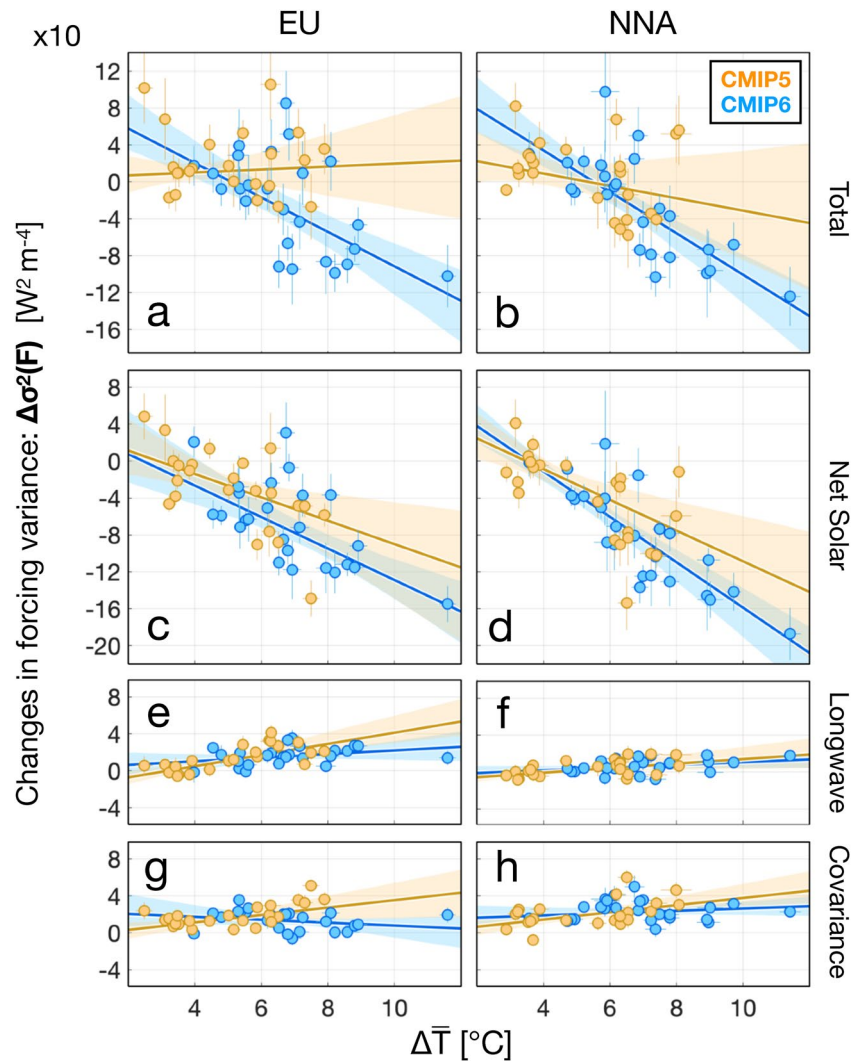
There seems to be overall consistency of median radiative forcing variance with mean cloud cover, as well as decreases in summer cloud percent, between CMIP5 and CMIP6. In CMIP5, however, changes in radiative forcing differ substantially across models for particular regions, as do the contribu-

tions to radiative forcing from solar and longwave as well as their covariation. CMIP6 shows more uniform changes in forcing variance by region and across models, though contributions from solar and longwave remain distinct. Unlike changes in the sensitivity to forcing, which uniformly increases across models and regions with overall warming and drying, we observe heterogeneous changes in radiative forcing that likely reflect multiple physical processes that are regionally and model dependent. Thus, we do not anticipate a single, cohesive explanation for changes in radiative forcing in Northern Hemisphere mid-latitudes.

## 5. Discussion and Conclusion

Our analyses show that land surfaces in semi-humid regions become more sensitive to radiative forcing in warmer and drier climates in both CMIP5 and CMIP6 generations due to suppressed surface evapotranspiration. The positive relationship between  $\Delta\bar{T}$  and  $\Delta\sigma^2(T)$  found in NH mid-latitudes in CMIP5, however, no longer holds in CMIP6 because CMIP6 simulations produce a uniform decrease in radiative forcing variance with decreasing cloud amount.

These results reveal two scenarios for changes in the distribution of midlatitude monthly summertime temperatures: compounding and canceling (Figure 9a). In the compounding scenario,  $\Delta\sigma^2(T)$  increases with warming

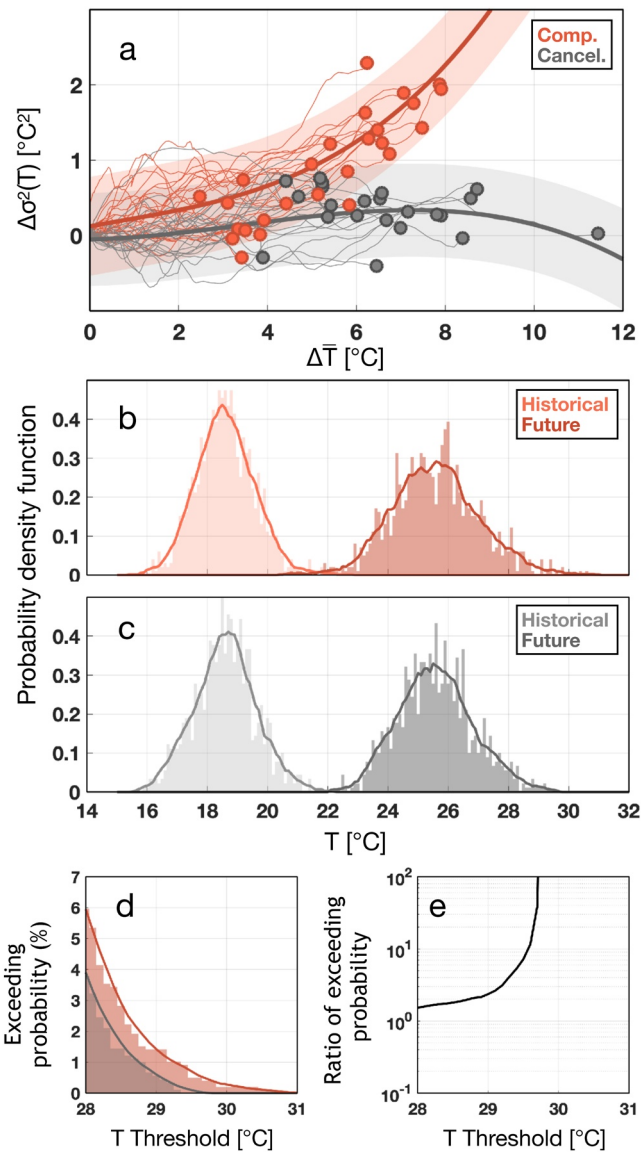


**Figure 8.** Changes in forcing variance and warming. (a)–(b) Regional  $\Delta\sigma^2(F)$  (y-axes) versus  $\Delta\bar{T}$  (x-axes) in (a) Europe and (b) Northern North America. (c)–(h) As (a)–(b) but decompose forcing variance to contributions from (c)–(d)  $\Delta\sigma^2(S_n)$ , (e)–(f)  $\Delta\sigma^2(L_l)$ , and (g)–(h) changes in the covariance between  $S_n$  and  $L_l$ .

because of increased sensitivity associated with suppressed evapotranspiration. In the canceling scenario, decreasing forcing variance cancels the effect of sensitivity growth, leading to  $\Delta\sigma^2(T)$  being almost independent of warming.

Compounding leads to more frequent extreme heat events, and we quantify this distinction for western Europe after controlling for three features of the CMIP simulations. First, although CMIP5 and CMIP6 generally respectively accord with the compounding and canceling scenarios, individual models differ. For example, three CMIP6 models—CMCC-CM2-SR5, EC-Earth3-Veg, and FIO-ESM-2-0—have  $\Delta\sigma^2(T)$  higher than  $1^\circ\text{C}^2$  in Europe (Figure 1c). We thus assign individual model simulations to the compounding and canceling scenario irrespective of CMIP generation (Figure 9a). Second, we shift the historical distribution of each model to center on the observed mean in Europe,  $18.6^\circ\text{C}$  (Harris et al., 2020), in order to control for model bias. We note that, in Europe, these biases correlate neither with  $\Delta\bar{T}$  nor  $\Delta\sigma^2(T)$  ( $r^2 < 0.1$ ). Finally, in order to control for differences in the amount of warming, we consider only models whose mean warming in 2071–2100 relative to 1976–2005 is between  $6^\circ$  and  $8.5^\circ\text{C}$ . We retain 11 models for the compounding and 10 models for the canceling scenario.

Although historical distributions of the two scenarios have similar widths, the compounding scenario has a wider distribution at the end of the 21<sup>st</sup> century (Figure 1). The probability of high-temperature extremes is 50% higher



**Figure 9.** Distinct implications for the two scenarios of  $\Delta\sigma^2(T)$ . (a) Temporal evolution (thin lines) of  $\Delta\sigma^2(T)$  and  $\Delta\bar{T}$  in Europe. Markers denotes individual models in the compounding (red) and the canceling scenario (gray). The base period is 1976–2005, and changes are calculated sliding annually from 1977–2006 to 2071–2100. See texts for more details of scenario assignment. Also shown is the central estimate (thick curve) and 95% c.i. (shading) of a cubic polynomial fitting that pools all 30-year periods from all models in each scenario. The crossmodel relationship of end-of-century changes (markers) agree with the fitted overall evolution (curve), indicating that, for specific scenario, warming differences within and across models are interchangeable. (b)–(c) Distribution of European-mean monthly temperatures in the historic (1976–2005, light) and future (2071–2100, dark) period for (b) the compounding and (c) the canceling scenario. In addition to actual histograms (bars), also shown are smoothed distributions (curves). We remove seasonal cycles and linear trends before calculating temperature anomalies, and we only use models whose European warming is between 6° and 8.5°C. (d) The probability of future temperatures exceeding certain thresholds (x-axis) in the compounding (red) and canceling scenario (gray). (e) The ratio of exceeding probability (the compounding divided by the canceling scenario) as a function of temperature threshold.

in the compounding scenario when the threshold is 28°C, and this percentage increases quickly with temperature threshold (Figures 9d and 9e). In addition to the probability of extreme events, environmental damages are often nonlinear to increasing temperatures (Schlenker & Roberts, 2009), such that overall damages increase with  $\sigma^2(T)$ . The association of greater temperature variance with low soil moisture may imply even more severe consequences when accounting for the joint roles of water availability and temperature in determining yield outcomes (Rigden et al., 2020).

Many additional analyses would be useful to better understand the distribution of simulated changes in temperatures. Our focus has been on interannual variations at monthly timescales for two regions of the NH midlatitude. A global analysis would be useful for assessing the generalizability of uniform changes in sensitivity and documenting heterogeneous changes in forcing variance. Also useful would be to investigate variations at seasonal (e.g., Santer et al., 2018) and daily timescale (e.g., Duan et al., 2020).

Scope also exists for further understanding physical processes responsible for changes in both temperature sensitivity and forcing variability. Quantifying the contribution of land-use (Christidis et al., 2013) in addition to surface warming and drying would further constrain predicted sensitivity change. Examining changes in the representation of cloud processes, including cloud characteristics and parameterization (e.g., Jian et al., 2020; Vignesh et al., 2020), appears crucial for further understanding inter-generational differences in forcing variance. Also helpful would be to document the degree to which changes in large-scale circulation (Dafka et al., 2019; Huguenin et al., 2020) and the associated free-tropospheric advection (Holmes et al., 2016) contribute to forcing variance. Other atmospheric processes, such as the strength of water-vapor feedback and atmospheric heat storage may also contribute. Decomposing and quantifying the various contributions to the variance in longwave radiation using kernel methods, which have been widely used for feedback analyses (e.g., Soden et al., 2008), could also be helpful. A related question is the degree to which VZ20r can be accurate without coupling with atmospheric processes. Our approach of separating the physics between surface-temperature sensitivity and atmospheric forcing is useful for analysis and constructing simple models. However, any explanation of temperature extremes that depends upon prescribing atmospheric forcing will be incomplete.

Finally, it would be relevant to compare simulated changes against observations. Examining whether observed changes in evapotranspiration (Pascolini-Campbell et al., 2021) have led to detectable changes in temperature sensitivity appears useful for testing model predictions. Using satellite observations and reanalyses to evaluate the relationship between forcing variance and mean cloud amount may help in evaluating which models and CMIP generation is closer to reality (e.g., Bodas-Salcedo et al., 2008). In addition, monitoring the evolution of changes in observed clouds and radiative forcing will also provide insights for constraining simulated forcing variance and uncertain predictions of temperature variability.

## Appendix A: Deriving VZ20r

For completeness, we provide the derivation of VZ20r, which generally follows that of Vargas Zeppetello et al. (2020). Where, our derivation diverges from Vargas Zeppetello et al. (2020) is explicitly noted. We begin with a surface energy balance,

$$F = L_{\uparrow} + H + \gamma E + G. \quad (\text{A1})$$

$F$  denotes radiative forcing as the summation of net shortwave and downward longwave radiation, rather than only net shortwave radiation as in VZ20. Radiative forcing is balanced by upward longwave radiation at the surface,  $L_{\uparrow}$ , sensible heat flux,  $H$ , latent heat flux,  $\gamma E$ , and heat flux into the soil column,  $G$ .

Evapotranspiration,  $E$ , is represented as,

$$E = \frac{\rho_a}{\mathcal{R}} V = \frac{\rho_a}{r_0} s V. \quad (\text{A2})$$

$V$  is the specific humidity deficit, or the difference between saturation and actual specific humidity.  $\rho_a$  is the density of air, taken as  $1.2 \text{ kg m}^{-3}$ . The bulk surface resistance,  $\mathcal{R}$ , is approximated as a minimum bulk surface resistance,  $r_0$ , divided by the saturation of surface soil,  $\frac{\rho_a}{s}$ , such that resistance increases for drier surface. Following Vargas Zeppetello et al. (2020),  $r_0$  is 75 s/m, a typical resistance for wet grassland (Jones, 1985). This approach of adjusting a minimum resistance according to variations in soil moisture is widely used to model bulk canopy conductivity (e.g., Jones, 1985). Note that in Equation A2, we ignore variations in  $\mathcal{R}$  depending on differences in vegetation, but this simple representation nevertheless reproduces sensitivity changes in Europe and Northern North America across CMIP models (see Section 3.2 in the main text).

Substituting Equation A2 into Equation A1 and taking the derivative gives

$$F' = L'_{\uparrow} + H' + \frac{\rho_a}{r_0} \gamma (\bar{s} V' + \bar{V} s') + G', \quad (\text{A3})$$

where an over-bar denotes a climatological mean, and a prime denotes monthly anomalies relative to a climatology.

Monthly anomalies in surface water balance are,

$$P' = E' + R' = \frac{\rho_a}{r_0} (\bar{s} V' + \bar{V} s') + \mu s', \quad (\text{A4})$$

where runoff anomalies,  $R'$ , are assumed to be proportional to soil moisture anomalies,  $R' = \mu s'$ . The original VZ20 model sets  $\mu$  equal to five times of the standard deviation of monthly precipitation,  $\sigma(P)$ , a quantity that varies with space. In VZ20r, we assign the median value of  $\mu$  across space and all CMIP models,  $6 \times 10^{-5} \text{ s}^{-1}$ . Rearranging terms in Equation A4 leads to a representation of soil moisture anomaly,

$$s' = \frac{P' - \frac{\rho_a}{r_0} \bar{s} V'}{\frac{\rho_a}{r_0} \bar{V} + \mu}. \quad (\text{A5})$$

Equation A5 is substituted into Equation A3 to give,

$$F' = L'_{\uparrow} + H' + \frac{\rho_a}{r_0} \gamma \bar{s} (1 - \lambda) V' + \lambda \gamma P' + G', \quad (\text{A6})$$

where  $\lambda = \frac{\frac{\rho_a}{r_0} \bar{V}}{\frac{\rho_a}{r_0} \bar{V} + \mu}$  is the proportion of potential water loss through evaporation and is interpreted as a dryness index (Vargas Zeppetello et al., 2020).

The monthly anomaly in specific humidity deficit is,

$$V' = q^{*'} - q' = \frac{\partial q^*}{\partial T} T' - \frac{\partial q}{\partial P} P', \quad (\text{A7})$$

where,  $q^*$  and  $q$  denote, respectively, saturation and actual specific humidity.  $\frac{\partial q^*}{\partial T}$  is the derivative of saturation specific humidity with respect to temperature and is evaluated at  $\bar{T}$  using the Clausius-Clapeyron equation.  $\frac{\partial q}{\partial P}$  is

the derivative of specific humidity to precipitation, but which is omitted because anomalies of  $q$  are small compared to those of  $q^*$  (Vargas Zeppetello et al., 2020).

After substituting Equation A7 into Equation A6, unlike (Vargas Zeppetello et al., 2020) that assigned anomalous longwave radiation to be proportional to  $T'$ , we apply the Stefan-Boltzmann law to represent upward longwave radiation,  $L'_l = 4\sigma\bar{T}^3T'$ , where  $\sigma$  is the Stefan-Boltzmann constant. Note that  $\sigma$  differs from  $\sigma(\cdot)$ , which we use to denote the standard deviation of a variable. Sensible heat fluxes and ground heat uptake are assumed proportional to surface temperature anomalies,  $H' + G' = hT'$ , where  $h = 4 \text{ W m}^{-2}\text{C}^{-1}$ . Applying these parameterizations to Equation A6 gives,

$$F' - \lambda\gamma P' = \left[ 4\sigma\bar{T}^3 + h + \frac{\rho_a}{r_0}\gamma\bar{s}(1 - \lambda)\frac{\partial q^*}{\partial T} \right] T'. \quad (\text{A8})$$

Finally, in order to get a closed form representation of temperature sensitivity and diverging from the approach of Vargas Zeppetello et al. (2020), we substitute  $P'$  with  $\beta F'$  to obtain VZ20r (Equation 6 in the main text),

$$(1 - \beta\lambda\gamma)F' = \left[ 4\sigma\bar{T}^3 + h + \frac{\rho_a}{r_0}\gamma\bar{s}(1 - \lambda)\frac{\partial q^*}{\partial T} \right] T'.$$

## Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

## Data Availability Statement

Data sets for this research are available in these in-text data citation reference: Taylor et al. (2012), with license of CMIP5, and Eyring et al. (2016), with license of CMIP6.

## Acknowledgments

Conversations with L. Vargas Zeppetello and D. Battisti improved the content of this manuscript. This study was supported by the Harvard Global Institute and NSF (Award 1903657). D. Chan was also supported by the Woods Hole Oceanographic Institute Weston Howland Jr. Postdoctoral Fellowship.

## References

- Berg, A., & Sheffield, J. (2018). Soil moisture–evapotranspiration coupling in CMIP5 models: Relationship with simulated climate and projections. *Journal of Climate*, 31(12), 4865–4878. <https://doi.org/10.1175/JCLI-D-17-0757.1>
- Bodas-Salcedo, A., Webb, M. J., Brooks, M. E., Ringer, M. A., Williams, K. D., Milton, S. F., & Wilson, D. R. (2008). Evaluating cloud systems in the Met Office global forecast model using simulated CloudSat radar reflectivities. *Journal of Geophysical Research*, 113(D8), D00A13. <https://doi.org/10.1029/2007JD009620>
- Chan, D., Cobb, A., Zeppetello, L. R. V., Battisti, D. S., & Huybers, P. (2020). Summertime temperature variability increases with local warming in midlatitude regions. *Geophysical Research Letters*, 47(13), e2020GL087624. <https://doi.org/10.1029/2020GL087624>
- Christidis, N., Stott, P. A., Hegerl, G. C., & Betts, R. A. (2013). The role of land use change in the recent warming of daily extreme temperatures. *Geophysical Research Letters*, 40(3), 589–594. <https://doi.org/10.1002/grl.50159>
- Dafka, S., Toreti, A., Zanis, P., Xoplaki, E., & Luterbacher, J. (2019). Twenty-first-century changes in the Eastern Mediterranean Etesians and associated midlatitude atmospheric circulation. *Journal of Geophysical Research: Atmospheres*, 124(23), 12741–12754. <https://doi.org/10.1029/2019JD031203>
- Deb, P., Moradkhani, H., Abbaszadeh, P., Kiem, A. S., Engström, J., Keellings, D., & Sharma, A. (2020). Causes of the widespread 2019–2020 Australian bushfire season. *Earth's Future*, 8(11), e2020EF001671. <https://doi.org/10.1029/2020EF001671>
- Duan, S. Q., Findell, K. L., & Wright, J. S. (2020). Three regimes of temperature distribution change over dry land, moist land, and oceanic surfaces. *Geophysical Research Letters*, 47(24), e2020GL090997. <https://doi.org/10.1029/2020GL090997>
- Emanuel, K. (2008). The hurricane-climate connection. *Bulletin of the American Meteorological Society*, 89(5), ES10–ES20. <https://doi.org/10.1175/BAMS-89-5-Emanuel>
- Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Inter-comparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, 9(5), 1937–1958. <https://doi.org/10.5194/gmd-9-1937-2016>
- Fischer, E. M., Rajczak, J., & Schär, C. (2012). Changes in European summer temperature variability revisited. *Geophysical Research Letters*, 39, L19702. <https://doi.org/10.1029/2012GL052730>
- Fischer, E. M., & Schär, C. (2009). Future changes in daily summer temperature variability: Driving processes and role for temperature extremes. *Climate Dynamics*, 33(7–8), 917–935. <https://doi.org/10.1007/s00382-008-0473-8>
- Fuller, W. A. (2009). *Measurement error models* (Vol. 305). John Wiley & Sons.
- Gregory, J. M., & Mitchell, J. (1995). Simulation of daily variability of surface temperature and precipitation over Europe in the current and 2 × CO<sub>2</sub> climates using the UKMO climate model. *Quarterly Journal of the Royal Meteorological Society*, 121(526), 1451–1476. <https://doi.org/10.1002/qj.49712152611>
- Harris, I., Osborn, T. J., Jones, P., & Lister, D. (2020). Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. *Scientific Data*, 7(1), 1–18. <https://doi.org/10.1038/s41597-020-0453-3>
- Held, I. M., & Soden, B. J. (2000). Water vapor feedback and global warming. *Annual Review of Energy and the Environment*, 25(1), 441–475. <https://doi.org/10.1146/annurev.energy.25.1.441>

- Heymsfield, A. J., Miloshevich, L. M., Twohy, C., Sachse, G., & Oltmans, S. (1998). Upper-tropospheric relative humidity observations and implications for cirrus ice nucleation. *Geophysical Research Letters*, 25(9), 1343–1346. <https://doi.org/10.1029/98GL01089>
- Hoag, H. (2014). Russian summer tops 'universal' heatwave index. *Nature*, 16(4). <https://doi.org/10.1038/nature.2014.16250>
- Holmes, C. R., Woollings, T., Hawkins, E., & De Vries, H. (2016). Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate*, 29(6), 2221–2236. <https://doi.org/10.1175/JCLI-D-14-00735.1>
- Horton, R. M., Mankin, J. S., Lesk, C., Coffel, E., & Raymond, C. (2016). A review of recent advances in research on extreme heat events. *Current Climate Change Reports*, 2(4), 242–259. <https://doi.org/10.1007/s40641-016-0042-x>
- Huguenin, M. F., Fischer, E. M., Kotlarski, S., Scherrer, S. C., Schwierz, C., & Knutti, R. (2020). Lack of change in the projected frequency and persistence of atmospheric circulation types over central Europe. *Geophysical Research Letters*, 47(9), e2019GL086132. <https://doi.org/10.1029/2019GL086132>
- Huntingford, C., Jones, P. D., Livina, V. N., Lenton, T. M., & Cox, P. M. (2013). No increase in global temperature variability despite changing regional patterns. *Nature*, 500(7462), 327–330. <https://doi.org/10.1038/nature12310>
- IPCC. (2012). C. B. Field, V. Barros, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, et al. (Eds.) *Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of working groups I and II of the intergovernmental panel on climate change*. Cambridge University Press.
- Jian, B., Li, J., Zhao, Y., He, Y., Wang, J., & Huang, J. (2020). Evaluation of the CMIP6 planetary albedo climatology using satellite observations. *Climate Dynamics*, 54(11), 5145–5161. <https://doi.org/10.1007/s00382-020-05277-4>
- Jones, M. (1985). Plant microclimate. In *Techniques in bioproductivity and photosynthesis* (pp. 26–40). Elsevier.
- Lenderink, G., Van Ulden, A., Van den Hurk, B., & Van Meijgaard, E. (2007). Summertime inter-annual temperature variability in an ensemble of regional model simulations: Analysis of the surface energy budget. *Climatic Change*, 81(1), 233–247. <https://doi.org/10.1007/s10584-006-9229-9>
- Markovsky, I., & Van Huffel, S. (2007). Overview of total least-squares methods. *Signal Processing*, 87(10), 2283–2302. <https://doi.org/10.1016/j.sigpro.2007.04.004>
- McKinnon, K. A., Rhines, A., Tingley, M. P., & Huybers, P. (2016). The changing shape of Northern Hemisphere summer temperature distributions. *Journal of Geophysical Research: Atmospheres*, 121(15), 8849–8868. <https://doi.org/10.1002/2016JD025292>
- Pascolini-Campbell, M., Reager, J. T., Chandanpurkar, H. A., & Rodell, M. (2021). A 10 percent increase in global land evapotranspiration from 2003 to 2019. *Nature*, 593(7860), 543–547. <https://doi.org/10.1038/s41586-021-03503-5>
- Philippa, R., Dürr, B., Ohmura, A., & Ruckstuhl, C. (2005). Anthropogenic greenhouse forcing and strong water vapor feedback increase temperature in Europe. *Geophysical Research Letters*, 32(19), L19809. <https://doi.org/10.1029/2005GL023624>
- Price, J. D., & Wood, R. (2002). Comparison of probability density functions for total specific humidity and saturation deficit humidity, and consequences for cloud parametrization. *Quarterly Journal of the Royal Meteorological Society*, 128(584), 2059–2072. <https://doi.org/10.1256/003590002320603539>
- Rasmijn, L., Van der Schrier, G., Bintanja, R., Barkmeijer, J., Sterl, A., & Hazeleger, W. (2018). Future equivalent of 2010 Russian heatwave intensified by weakening soil moisture constraints. *Nature Climate Change*, 8(5), 381–385. <https://doi.org/10.1038/s41558-018-0114-0>
- Rhines, A., & Huybers, P. (2013). Frequent summer temperature extremes reflect changes in the mean, not the variance. *Proceedings of the National Academy of Sciences*, 110(7), E546. <https://doi.org/10.1073/pnas.1218748110>
- Rigden, A., Mueller, N., Holbrook, N., Pillai, N., & Huybers, P. (2020). Combined influence of soil moisture and atmospheric evaporative demand is important for accurately predicting us maize yields. *Nature Food*, 1(2), 127–133. <https://doi.org/10.1038/s43016-020-0028-7>
- Rowell, D. P. (2005). A scenario of European climate change for the late twenty-first century: Seasonal means and interannual variability. *Climate Dynamics*, 25(7–8), 837–849. <https://doi.org/10.1007/s00382-005-0068-6>
- Santer, B. D., Po-Chedley, S., Zelinka, M. D., Cvijanovic, I., Bonfils, C., Durack, P. J., et al. (2018). Human influence on the seasonal cycle of tropospheric temperature. *Science*, 361(6399). <https://doi.org/10.1126/science.aas8806>
- Schlenker, W., & Roberts, M. J. (2009). Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences*, 106(37), 15594–15598. <https://doi.org/10.1073/pnas.0906865106>
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3–4), 125–161. <https://doi.org/10.1016/j.earscirev.2010.02.004>
- Soden, B. J., Held, I. M., Colman, R., Shell, K. M., Kiehl, J. T., & Shields, C. A. (2008). Quantifying climate feedbacks using radiative kernels. *Journal of Climate*, 21(14), 3504–3520. <https://doi.org/10.1175/2007JCLI1210.1>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4), 485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>
- Vargas Zeppetello, L. R., & Battisti, D. S. (2020). Projected increases in monthly midlatitude summertime temperature variance over land are driven by local thermodynamics. *Geophysical Research Letters*, 47, e2020GL090197. <https://doi.org/10.1029/2020GL090197>
- Vargas Zeppetello, L. R., Battisti, D. S., & Baker, M. B. (2020). A new look at the variance of summertime temperatures over land. *Journal of Climate*, 33(13), 5465–5477. <https://doi.org/10.1175/JCLI-D-19-0887.1>
- Vargas Zeppetello, L. R., Battisti, D. S., & Baker, M. B. (2022). The physics of heat waves: What causes extremely high summertime temperatures? *Journal of Climate*. 1–61. <https://doi.org/10.1175/JCLI-D-21-0236.1>
- Vargas Zeppetello, L. R., Donohoe, A., & Battisti, D. (2019). Does surface temperature respond to or determine downwelling longwave radiation? *Geophysical Research Letters*, 46(5), 2781–2789. <https://doi.org/10.1029/2019GL082220>
- Vautard, R., van Aalst, M., Boucher, O., Drouin, A., Hausteijn, K., Kreienkamp, F., & Wehner, M. (2020). Human contribution to the record-breaking June and July 2019 heatwaves in western Europe. *Environmental Research Letters*, 15(9), 094077. <https://doi.org/10.1088/1748-9326/aba3d4>
- Vignesh, P. P., Jiang, J. H., Kishore, P., Su, H., Smay, T., Brighton, N., & Velicogna, I. (2020). Assessment of CMIP6 cloud fraction and comparison with satellite observations. *Earth and Space Science*, 7, e2019EA000975. <https://doi.org/10.1029/2019EA000975>
- York, D., Evensen, N. M., Martinez, M. L., & De Basabe Delgado, J. (2004). Unified equations for the slope, intercept, and standard errors of the best straight line. *American Journal of Physics*, 72(3), 367–375. <https://doi.org/10.1119/1.1632486>
- Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., & Taylor, K. E. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, 47, e2019GL085782. <https://doi.org/10.1029/2019GL085782>