Understanding and Reducing Warm and Dry Summer Biases in the Central United States: Analytical Modeling to Identify the Mechanisms for CMIP Ensemble Error Spread

CHAO SUN[a,b] AND XIN-ZHONG LIANG[a,b]

[a] *Department of Atmospheric and Oceanic Science, University of Maryland, College Park, College Park, Maryland*
[b] *Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*

(Manuscript received 12 April 2022, in final form 24 October 2022)

ABSTRACT: Most climate models in phase 6 of the Coupled Model Intercomparison Project (CMIP6) still suffer pronounced warm and dry summer biases in the central United States (CUS), even in high-resolution simulations. We found that the cloud base definition in the cumulus parameterization was the dominant factor determining the spread of the biases among models and those defining cloud base at the lifting condensation level (LCL) performed the best. To identify the underlying mechanisms, we developed a physically based analytical bias model (ABM) to capture the key feedback processes of land–atmosphere coupling. The ABM has significant explanatory power, capturing 80% variance of temperature and precipitation biases among all models. Our ABM analysis via counterfactual experiments indicated that the biases are attributed mostly by surface downwelling longwave radiation errors and second by surface net shortwave radiation errors, with the former 2–5 times larger. The effective radiative forcing from these two errors as weighted by their relative contributions induces runaway temperature and precipitation feedbacks, which collaborate to cause CUS summer warm and dry biases. The LCL cumulus reduces the biases through two key mechanisms: it produces more clouds and less precipitable water, which reduce radiative energy input for both surface heating and evapotranspiration to cause a cooler and wetter soil; it produces more rainfall and wetter soil conditions, which suppress the positive evapotranspiration–precipitation feedback to damp the warm and dry bias coupling. Most models using non-LCL schemes underestimate both precipitation and cloud amounts, which amplify the positive feedback to cause significant biases.

KEYWORDS: Feedback; Bias; Numerical analysis/modeling; Convective parameterization; Cumulus clouds

1. Introduction

Most climate models have long-lasting biases in the central United States (CUS); their simulated summer surface conditions are significantly warmer and drier than observations (Klein et al. 2006; Cheruy et al. 2014; Lin et al. 2017; Zhang et al. 2018; Ma et al. 2018; Morcrette et al. 2018; Van Weverberg et al. 2018). Given nonlinear interactions and feedbacks in the Earth system, these biases diminish the reliability of climate predictions and scenario projections (Liang et al. 2008; Lin et al. 2017; Palmer and Stevens 2019; Jiang et al. 2021). Consequently, the warm and dry biases disrupt the nonlinear relationships among surface heat fluxes and water storage and thus reduce the accuracy in predicting the critical feedback processes (Rupp et al. 2017). They decrease model ability in predicting extreme heat or rainfall events and severe weather outbreaks (Kunkel et al. 2010; Pendergrass et al. 2020; Sun and Liang 2020a), which have profound socioeconomic impacts (Smith and Matthews 2015; NOAA 2021; Kotz et al. 2022; Liang 2022), including extensive damages on agricultural productivity (Liang et al. 2017; Y. Li et al. 2019; Ortiz-Bobea et al. 2021). Therefore, they deteriorate the prediction reliability on climate–crop interactions in the CUS—the heartland of U.S. agriculture with summer as the primary growing season (Mueller et al. 2016). To boost the confident use of climate predictions or projections in decision-making, significant model improvements must be made to correct these essential biases.

The modeling community has made tremendous efforts to reduce these biases. Among these are various empirical bias-correction techniques (Bellprat et al. 2016; Ardilouze et al. 2019; Chang et al. 2019) to provide improved climate data necessary for impact assessments. Such posterior corrections, while useful, do not actually improve the modeling system representation of nonlinear physical relationships between atmosphere and land processes, rendering model predictions hard to explain, and thus may mislead decision-making that requires integrated knowledge (Ehret et al. 2012). Given that the failure to capture convective scale processes may cause the warm and dry biases (Liang et al. 2007; Lin et al. 2017; Sun and Liang 2020a), many efforts turn to high-resolution modeling to replace physics parameterizations. Liu et al. (2017) demonstrated that even using a grid spacing of 4 km (convection-permitting), severe biases (up to 6°C) in the CUS still exist. S. Li et al. (2019) tuned model parameters through a multiphase refinement approach, which reduced the warm biases to about 3°C. Cheruy et al. (2020) improved atmospheric and land surface physics, whereas more than 2°C warm biases remain. Barlage et al. (2021) incorporated groundwater processes in a convection-permitting model to reduce the warm biases from 5°–6°C to still 2°–3°C. Sun and Liang (2020b) attempted to infer the physical causes of the

Supplemental information related to this paper is available at the Journals Online website: https://doi.org/10.1175/JCLI-D-22-0255.s1.

Corresponding author: Xin-Zhong Liang, xliang@umd.edu

DOI: 10.1175/JCLI-D-22-0255.1

dry biases through the structural equation model analysis. Morcrette et al. (2018), Van Weverberg et al. (2018), Ma et al. (2018), and Zhang et al. (2018) established a special project named CAUSES (Clouds Above the United States and Errors at the Surface) to understand the physical mechanisms leading to CUS warm biases based on surface radiation budget decomposition analysis. They found that surface evaporative fraction underestimation and solar radiation overestimation were the two leading contributors to the warm biases. However, all previous studies have not fully disentangled the complex feedback mechanisms among radiation, convection, precipitation, and surface processes that result in the coupled warm and dry biases. These endeavors revealed the challenges in resolving the CUS warm and dry biases.

Our companion paper Sun and Liang (2022, hereafter S&L) took a solid step to understand and reduce CUS warm and dry summer biases in the regional Climate–Weather Research and Forecasting (CWRF) Model (Liang et al. 2012) by improving its ensemble cumulus parameterization (ECP; Qiao and Liang 2015, 2016, 2017). The improvements included a trigger for mesoscale convective systems to occur in unfavorable environmental conditions, a constraint to suppress convections from shallow boundary layers, and a cloud-to-rainwater conversion to favor anvil formation. In addition, the trigger defines the cloud base at the lifting condensation level (LCL) rather than the level of free convection (LFC), which lowered the cloud base and increased the cloud depth. Together, the new ECP produced more low- and high-level clouds to reduce surface shortwave radiation and increase outgoing longwave radiation, and hence eliminated the warm and dry biases in a physically consistent manner. S&L demonstrated the importance of identifying the physical processes and understanding the underlying mechanisms in order to actually reduce or eliminate such biases in a specific model.

The CAUSES project has been working on three competing hypotheses of the key causes for the CUS warm biases: underestimation of precipitation (Lin et al. 2017), cloud (Cheruy et al. 2014), and evapotranspiration (Mueller and Seneviratne 2014). To this end, S&L explained how improving cumulus parameterization can increase both low-level cloud and precipitation amounts to consistently eliminate the warm and dry biases with a realistic atmospheric energy balance. We did that from the perspective of atmospheric forcing since all changes started from altering cumulus parameterization that led to surface responses. As such, we did not address the specific role of evapotranspiration in the loop, although its interannual differences between the new and original ECP contained about 20%–70% (depending on regions) of total variance in surface air temperature and precipitation biases. Presumably, there exist strong land–atmosphere coupling mechanisms that link underestimations of low cloud, precipitation, and evapotranspiration to the warm and dry biases. However, such links are complex, strongly depending on climate regimes and model configurations, and it is particularly challenging to determine which mechanism dominates and whether the atmosphere or surface controls the pathways (Findell and Eltahir 2003a,b; Koster et al. 2004; Ferguson and Wood 2011; Taylor et al. 2012; Santanello et al. 2018; Wei and Dirmeyer 2012; Zhou et al. 2019).

Land–atmosphere interactions involve soil moisture and temperature, evapotranspiration and sensible heat, solar absorption and infrared emission, and cloud, convection, and precipitation through various positive and negative (local or nonlocal) feedbacks (Santanello et al. 2018). Numerous studies offered diagnostic methods to dissect soil moisture–precipitation coupling mechanisms containing all above variables and processes (Brubaker and Entekhabi 1995, 1996; Schär et al. 1999; Salvucci et al. 2002; Findell and Eltahir 2003a; Dirmeyer and Brubaker 2007; Wei and Dirmeyer 2012; Wulfmeyer et al. 2018; Hu et al. 2021). Most of these methods focused on analyzing the feedbacks in individual models, which may reach varying conclusions specific to land component deficiencies or different parameterizations used in the coupled land–atmosphere system (Williams et al. 2016; Wei et al. 2021). They may not be as effective for application to identify the key mechanisms that govern the spread of biases in all related variables among a large suite of models. They also require high-frequency data that are typically not available from all models in comparison.

This study aims to identify the model deficiencies that may likely explain common CUS warm and dry biases among recent climate models and develop analytical modeling to quantify the major contributing factors and associated feedbacks. The identification and attribution were approached from the dominant behavior of a large multimodel ensemble rather than a single model like CWRF. However, the rigorous process understanding based on CWRF sensitivity experiments in S&L offered a unique guidance for the ensemble quantification. Section 2 describes the ensemble simulations from the Coupled Model Intercomparison Project phase 5 (CMIP5) and phase 6 (CMIP6) and the High-Resolution Model Intercomparison Project (HighResMIP, including only models with grid spacing < 30 km) as well as observational data used. Section 3 presents the warm and dry biases common to these climate models and conducts composite, decomposition, and regression analyses to distinguish their strong sensitivity to cumulus parameterization and statistical relationships among key variables. In particular, the analyses confirmed that defining the cloud base at LCL rather than LFC is a major contributor to reducing the biases. Section 4 develops the analytical bias model (ABM) to understand the responsible physical processes and quantify the key feedback mechanisms. Section 5 conducts counterfactual experiments with ABM to disentangle the feedback loops and identify the key mechanisms that dominate the systematic biases and spreads among the latest CMIP6 plus HighResMIP models. Especially, we found thresholds for varying evapotranspiration–precipitation feedbacks to occur and thus cause systematic warm and dry biases in CUS. Section 6 gives the conclusions.

2. CMIP simulations and observational data

This study analyzed all present-day simulations currently available from 61 CMIP6 and 11 HighResMIP models, where different resolutions of a same model are counted as different models (see Tables S1 and S2 in the online supplemental material for model resolution and data availability). Most models have multiple simulations (or variants) that were realized

TABLE 1. Observations, resolutions, available years, and their references.

Meteorology variable	Temporal/spatial resolution and mapping method	Years	Source
Precipitation	Daily at 1-km grid (conservative interpolated)	1980–present	Daymet: daily surface weather data (Thornton et al. 2016)
Temperature at 2 m	Daily at 1-km grid (linear interpolated)		
Shortwave and longwave radiation at surface and top of atmosphere (clear/all)	Monthly at 1° × 1° latitude–longitude grid (linear interpolated)	2000–present	Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) Edition 4.1 (Loeb et al. 2018)
Sensible heat and latent heat flux	Monthly at 0.5° × 0.5° latitude–longitude grid	2001–13	The FLUXCOM ensemble of global land–atmosphere energy fluxes (Jung et al. 2019)
Cloud fraction Cloud top pressure Cloud optical depth	Eight-day 1° × 1° equal-angle grid (linear interpolated)	2002–present	MODIS Cloud Optical Properties: The Collection 6/6.1 Level-2 (Platnick et al. 2015)
Height, zonal/meridional wind, omega, temperature, and specific humidity on pressure levels	Monthly 32-km grid with 29 pressure levels (linear interpolated)	1979–present	NCEP North American Regional Reanalysis (NARR) (Mesinger et al. 2006)

with small perturbations in initial states and physical parameters. In total, there are 592 simulations, with the actual number of variants available ranging from 592 to 402 among the 15 variables examined (Table 1). The analyses below were mainly based on the ensemble mean of all available variants for a specific variable in each model together with these individual variants depicting the uncertainty. To identify feedbacks and mechanisms in which the physical consistency is essential, we chose the first member (rather than ensemble mean) of each model with the variant label "r1i1p1f1" (with only a few exceptions due to naming differences), where the letters r, i, p, and f denote respectively the realization, initialization, physics, and forcing indices of the simulation. We also evaluated 49 CMIP5 models with a total of 212 variants (Table S3) to see if any reductions in the warm and dry biases were made over the years. For all these model simulations, our analyses focused on summer (June–August) averages during the common period from 1 January 1980 through 31 December 2014.

Table 1 summarizes the observational data used in this study as the reference to calculate model biases. All data were available at varying temporal and spatial resolutions and mapped onto CWRF 30-km grid by conservative or linear interpolation.

3. Common climate model biases and sensitivity to cumulus parameterization

Figure 1 compares the geographic distributions of summer temperature and precipitation biases from observations over the contiguous United States among the models separately in the CMIP5, CMIP6, and HighResMIP-high (grid spacing < 30 km) ensembles. To avoid the result statistics being dominated by models with large realizations, we first calculated the mean of all the variants from each model and then averaged the mean biases over all models in an ensemble. The CMIP5 ensemble showed average warm biases of 1°–3°C in CUS. The CMIP6 ensemble produced similar warm biases but in more extensive areas, expanded from CUS into the central-southern Great Plains. The HighResMIP-high ensemble simulated even larger warm biases over broader areas, with centers of greater than 3°C. Correspondingly, all ensembles underestimated precipitation in CUS, with the magnitudes and coverages of the dry biases (over 0.5 mm day^{-1}) both increased from CMIP5 to CMIP6 to HighResMIP-high.

Thus, CMIP6 still simulated large warm and dry biases in CUS over broader areas than CMIP5, albeit incorporating more advanced physical parameterizations and higher resolutions (Eyring et al. 2019). Even HighResMIP-high with grid sizes reaching 25 km could not resolve the warm and dry bias (see Fig. S1 in the supplemental material for a more balanced comparison between high- versus low-resolution simulations from the same models). Simply increasing model resolution cannot eliminate the biases, while refining physics representation such as for mesoscale convective systems is essential (S&L). CMIP5 had other biases, including wet biases in the northern states between the Rocky Mountains and Great Plains and along the Appalachian Mountains as well as dry biases in the coastal areas along the Gulf of Mexico. These precipitation biases exhibited no apparent link with those of temperature and were reduced in CMIP6 and more significantly diminished in HighResMIP-high. This reduction could partly result from model resolution increase, a topic warranting a separate study.

S&L clearly demonstrated the importance of cumulus parameterization to the CWRF summer warm and dry biases in CUS among other major regions. Given the corresponding physical mechanisms identified earlier, here we focused on the key characteristic differences in cumulus parameterization that may explain the warm and dry biases common to current global climate models. One of key findings in S&L was that the cloud base definition in cumulus schemes plays a critical role. Below we separated all 72 models from CMIP6 plus

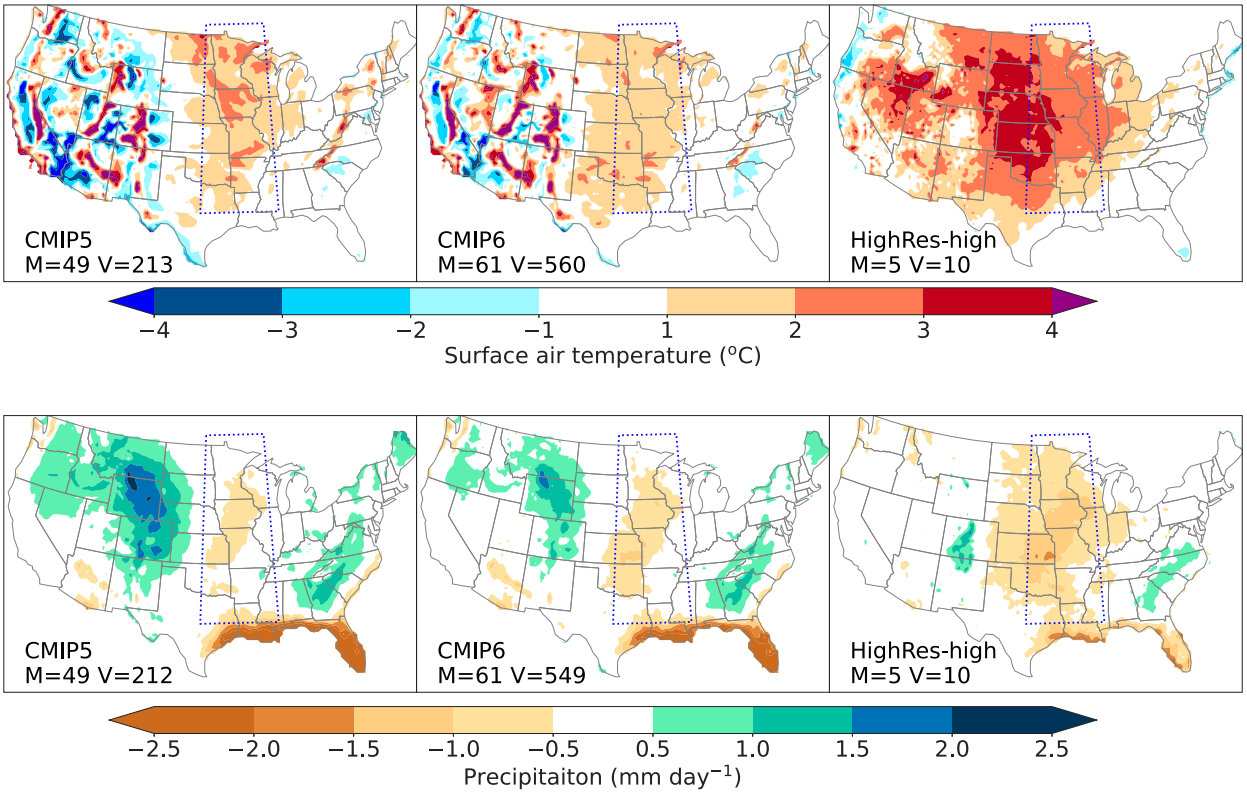

FIG. 1. Summer surface air temperature (°C) and precipitation (mm day^{-1}) biases during 1980–2014 averaged among CMIP5, CMIP6, and HighResMIP-high (grid spacing < 30 km) models. The "M" stands for the total number of models, and the "V" is the total number of variations used in calculating the ensemble mean. The blue dot box (31°–52°N, 89°–98°E) outlines the boundary of CUS.

HighResMIP (hereafter referred to as CMIP6X for brevity) into nine groups by the cloud base definition in their cumulus scheme as summarized in Table S4, which includes the respective abbreviation, main features, and references. The models were first categorized by cumulus schemes, including BMJ (Betts–Miller–Janjić; Betts and Miller 1986), AS (Arakawa and Schubert 1974), GR (Gregory and Rowntree 1990), PCMT (Prognostic Condensates Microphysics and Transport; Lopez 2002; Piriou et al. 2007; Geleyn et al. 2008; Guérémy 2011), LMDZ (Hourdin et al. 2020), ZM (Zhang and McFarlane 1995), GY (Del Genio and Yao 1993), and TDK (Tiedtke 1989). Since BMJ does not use cloud mass flux explicitly, the first-order relationships among downdraft, cloud depth, and others are irrelevant. AS and GR define the cloud base at LFC, PCMT at LCL, and LMDZ at 40 hPa above LCL, while ZM and GY define it at the particle lifting level (PLL). In contrast, TDK adopts either the LCL or LFC definition and hence is subdivided into TDK_LCL or TDK_LFC.

Figure 2 compares the ensemble mean temperature biases among the nine groups. Each group contains all the models adopting a same cumulus scheme, and each model is also shown with its own multirealization mean biases averaged in CUS. In general, the groups defining the cloud base at LCL (TDK, PCMT) simulated smaller warm biases than LFC (TDK, AS, GR). The most direct comparison was made with TDK, in which using LCL rather than LFC produced systematically cooler air in most of the eastern United States, essentially eliminating the large warm biases in CUS. This TDK_LCL resembled our new ECP in CWRF improved by S&L as they used the same cumulus trigger. In contrast, LMDZ still had warm biases in CUS and the Northeast, likely because it elevated the cloud base by 40 hPa above LCL. On the other hand, ZM estimated PLL as the maximum moist static energy level. Since this PLL is usually lower than LCL (Wu 2012), a parcel rising from PLL to LCL may experience negative buoyancy (Zhang and McFarlane 1991), reducing the total cloud-work function. The net effect of defining the cloud base at PLL is equivalent to reducing the effective cloud depth. This may partially explain why ZM using PLL generated large warm biases like other schemes using LFC. A seeming contradiction appeared with GY, which also used PLL but yielded cold biases (see below for further discussion).

Figure 3 compares the ensemble mean precipitation biases among the nine groups. Clearly, TDK using LCL (rather than LFC) largely reduced the dry biases in CUS. Meanwhile, the cumulus schemes using LFC (TDK, GR) and PLL (ZM) simulated notable dry biases. These were consistent with the reduction of the warm biases from LCL versus LFC. On the other hand, the dry biases in CUS (excluding the southern region) were not evident in AS using LFC but obvious in PCMT and LMDZ using LCL. In addition, GY even overestimated precipitation, accompanying a cold bias. This could result from the precipitable water increase by allowing convective rain

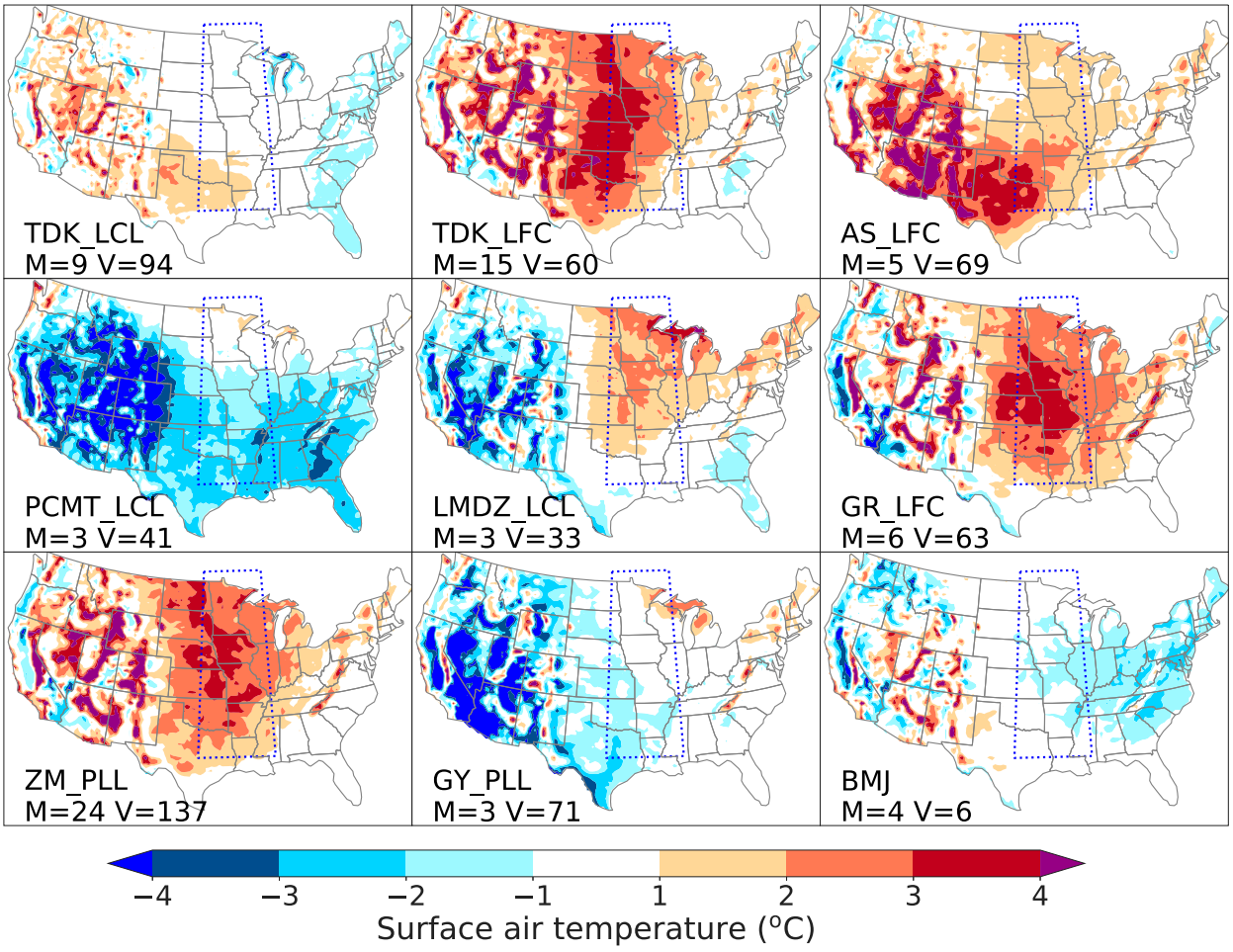

FIG. 2. Summer surface air temperature (°C) biases, grouped 72 CMIP6X (CMIP6 plus HighResMIP) models into nine types by cumulus schemes and cloud base definitions. See Fig. 1 for other information.

evaporation above the cloud base (Kim et al. 2012), which was deactivated in the latest model (Rind et al. 2020). It is difficult to interpret these mixed results since no clean experiment was available for these cumulus schemes directly separating LCL versus LFC as in TDK. Nevertheless, the choice of the cumulus parameterization scheme plays a significant, if not dominant, role in the contrast in precipitation biases among the models. While parameterized (convective) and resolved (large-scale) precipitations affect cloud–radiation interactions very differently (Lin et al. 2013), their relative contributions to model biases or even extreme events are not fully understood (Liang et al. 2019; Sun and Liang 2020b). However, our analysis of convective precipitation (Fig. S2) and its ratio to total amount (Fig. S3) showed no significant correspondence with CUS warm and dry biases (Fig. S4).

Figure 4 compares CUS regional mean temperature, precipitation, total cloud cover, and precipitable water biases among 72 CMIP6X models as grouped by cumulus schemes and cloud base definitions. In general, the models using the cumulus schemes defining the cloud base at LCL (PMCT, TDK, LMDZ) simulated the least dry and warm biases associated with more cloud cover and less precipitable water, while those defining at LFC (AS, TDK, GR) or PLL (GY, ZM) produced larger biases with less cloud and more precipitable water (see Figs. S5 and S6 for their spatial distributions). Since high- and low-level clouds have generally opposite radiative warm and cooling effects (partially explained LMDZ's total cloud underestimate), a more physically consistent picture would be drawn if cloud data at different levels were available for direct comparison. As discussed earlier, GY_PLL is an exception, having cold and wet biases. On the other hand, BMJ behaved differently from all others, producing cold but dry biases, indicating inconsistent physics representation (see below for further discussion). TDK_LFC had the largest bias spread; a closer scrutiny (Table S5) showed that the outliers (with largest high warm and dry biases) were from two high-resolution simulations (MPI-ESM1-2-XR and INM-CM-5).

Figure 5 compares CUS regional mean biases in temperature, precipitation, cloud albedo [shortwave reflection; see the supplemental material for calculation following Betts (2007)], and surface energy fluxes among 72 CMIP6X models as grouped by cumulus cloud base definitions, which include LCL (15),

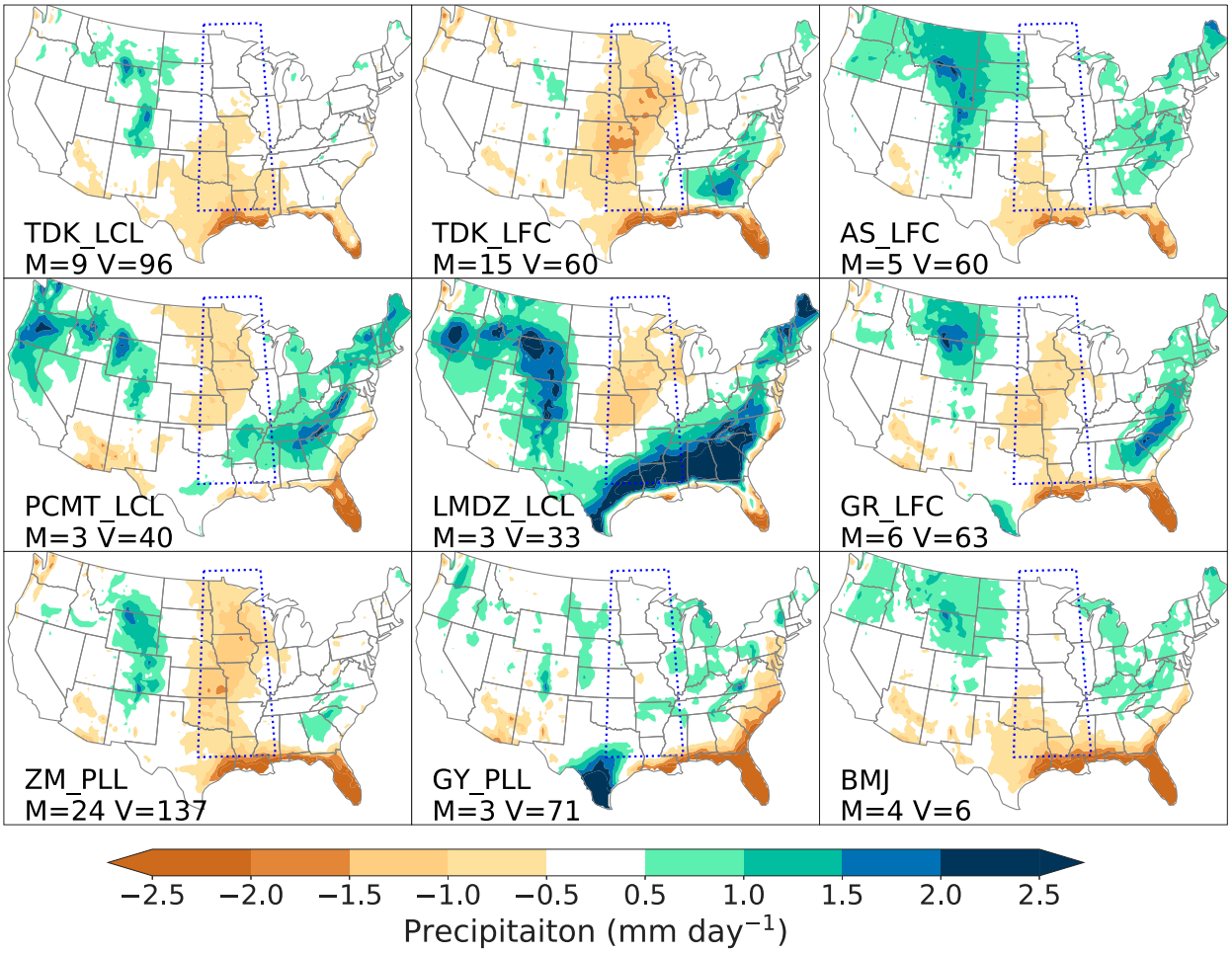

FIG. 3. As in Fig. 2, but for precipitation biases (mm day^{-1}).

LFC (26), PLL (27), and BMJ (4). Apparently, BMJ is an outlier with large cold but relatively small dry biases (−0.6°C, −0.3 mm day^{-1}). It substantially overestimated both shortwave surface reflection and cloud attenuation (17.6 and 5.7 W m^{-2}), causing large underestimates of surface net shortwave and hence total (shortwave plus longwave) radiation fluxes (−10.6 and −7.0 W m^{-2}). Corresponding to its cold bias, BMJ simulated surface net (shortwave, longwave, total) radiation and sensible heat biases totally opposite to those by other cloud base definitions, indicating inconsistent treatment for surface reflection and cloud effect.

Clearly, LCL had the smallest temperature and precipitation biases (0.2°C, −0.3 mm day^{-1}), whereas LFC yielded much larger biases (2.3°C, −0.6 mm day^{-1}). Consistently, LCL simulated the most realistic surface shortwave cloud radiative effect, net shortwave, longwave, and total radiation, and sensible heat with small biases (−0.4, 7.4, −3.0, 4.4, and 4.0 W m^{-2}). As demonstrated in S&L, LCL reduced excessive convective mass flux and thus cumulus drying and warming so to generate sufficient low-level clouds and attenuate more downwelling shortwave radiation. On the other hand, LFC largely underestimated cloud amount so to systematically overestimate cloud radiative effect and surface fluxes (17.3, 16.4, −8.3, 8.1, and 16.9 W m^{-2}), all of which were much larger in magnitude than the corresponding values of LCL. PLL resembled LFC with not only close temperature and precipitation biases (2.2°C, −0.7 mm day^{-1}) but also similar biases in respective cloud radiative effect and surface fluxes (6.1, 13.8, −7.3, 6.5, and 15.2 W m^{-2}). The main PLL difference from LFC was in reducing the cloud radiative effect by 65%. LCL also simulated a much smaller surface latent heat bias than LFC and PLL (−10.4 vs −18.8 and −19.6 W m^{-2}). This was well reflected in the contrast among their precipitation biases.

In summary, the CUS summer warm and dry biases and their spread among the latest CMIP6X models strongly depend on cumulus parameterization, in which the cloud base definition is a critical factor separating systematic model differences. In particular, defining the cumulus cloud base at LCL results in the least warm and dry biases with more realistic surface energy partitioning among sensible and latent heat, and radiation components.

Figure 6 shows the statistical relationships among CUS summer biases in cloud albedo, surface radiation fluxes, and surface heat fluxes simulated by 72 CMIP6X models. For each model and each field, all its variants were used to

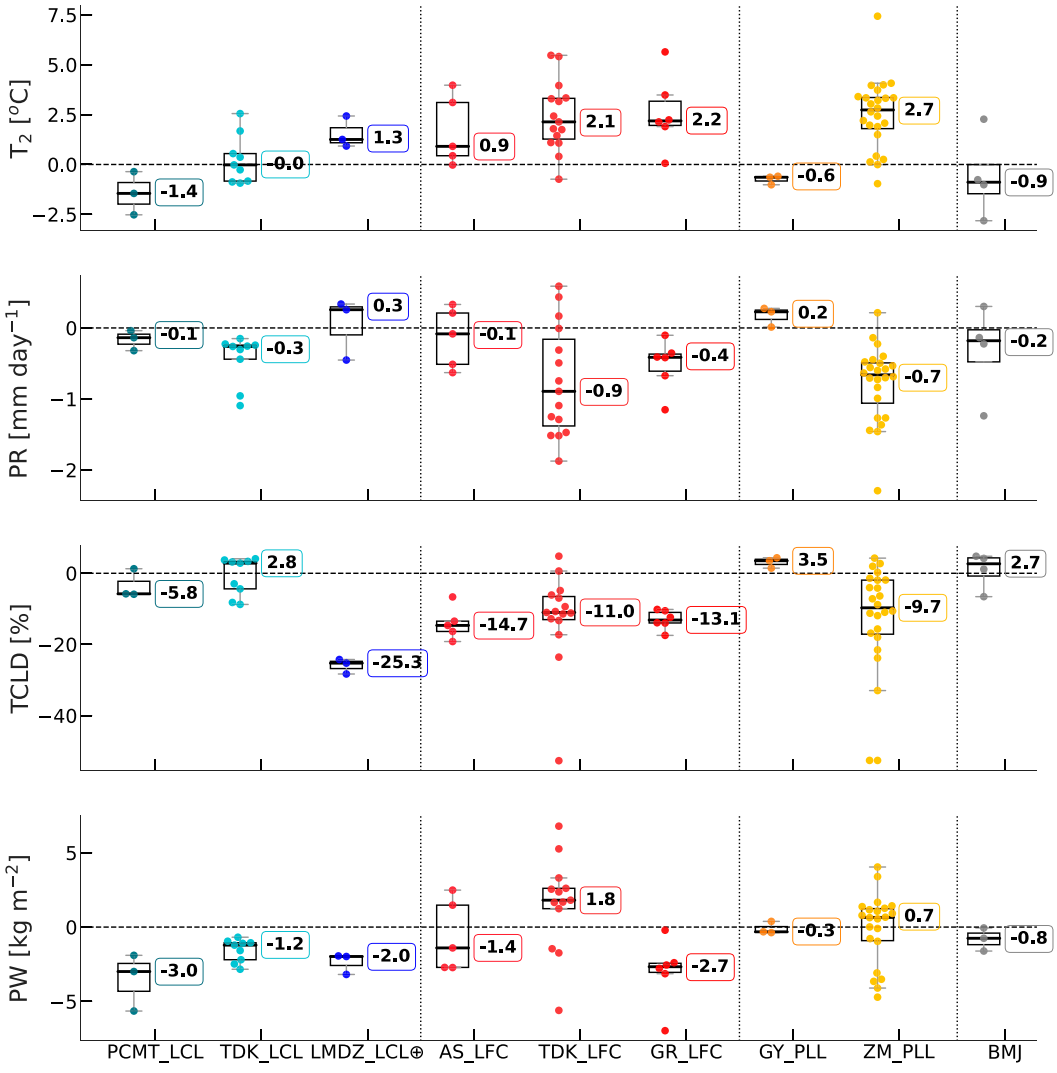

FIG. 4. Summer surface air temperature (T_2; °C), precipitation (PR; mm day^{-1}), total cloud cover (TCLD; %), and precipitable water (PW; kg m^{-2}) biases during 1980–2014 in CUS grouped 72 CMIP6X models into nine types by cumulus schemes and cloud base definitions.

calculate the mean and standard deviation, with the latter depicting uncertainty. Linear regressions were then made on these means averaged in CUS among these fields, with the respective one deviation depicting their uncertainty range. For the surface energy budget averaged over all models, biases (in W m^{-2}) in radiation components were the largest for SW$_d$ (23.9), reduced by about half for SW$_u$ (11.8) and LW$_u$ (10.5), and the smallest for LW$_d$ (4.3). The net total radiation (R_n) was overestimated by 5.9 W m^{-2}, which was compensated by smaller LH (−16.6) and larger SH (12.1). This resulted in a net energy surplus of 10.4 W m^{-2}, which was balanced by larger heat flux into the ground.

For the two radiation components directly affected by the atmosphere and as the energy inputs to the surface, the variance explained is 89% for SW$_d$ by cloud albedo alone and 64% for LW$_d$ by precipitable water. While the two inputs are independent (with 0 covariance), their partitioning into other surface energy components is our main concern. The variance percentage explained for SW$_u$, LW$_u$, SH, and LH was respectively 22, 37, 27, and 4 by SW$_d$ and 2, 46, 16, and 8 by LW$_d$. It is surprising to notice that biases in LH (hence evapotranspiration or ET) and SW$_d$ had almost no correspondence among the models. Little correspondence was seen between SW$_u$ and LW$_d$ biases. The model spread in SW$_d$ biases was distributed mostly into those in LW$_u$, SH, and SW$_u$, while the LW$_d$ spread was partitioned mainly among those in LW$_u$, SH, and LH. It is important to note that the spread in net radiation gain (R_n) explained 95% variance of moist enthalpy flux (SH + LH), whereas the residual (ground heat flux) barely contributed any to temperature (0.07) or precipitation bias (0.05). The sum of LW$_d$ and SH explained 83% variance of LW$_u$, much more than each individually (46% and 69%; not shown).

Figure 7 illustrates the statistical relationships among CUS summer biases in precipitation and its decomposed diabatic

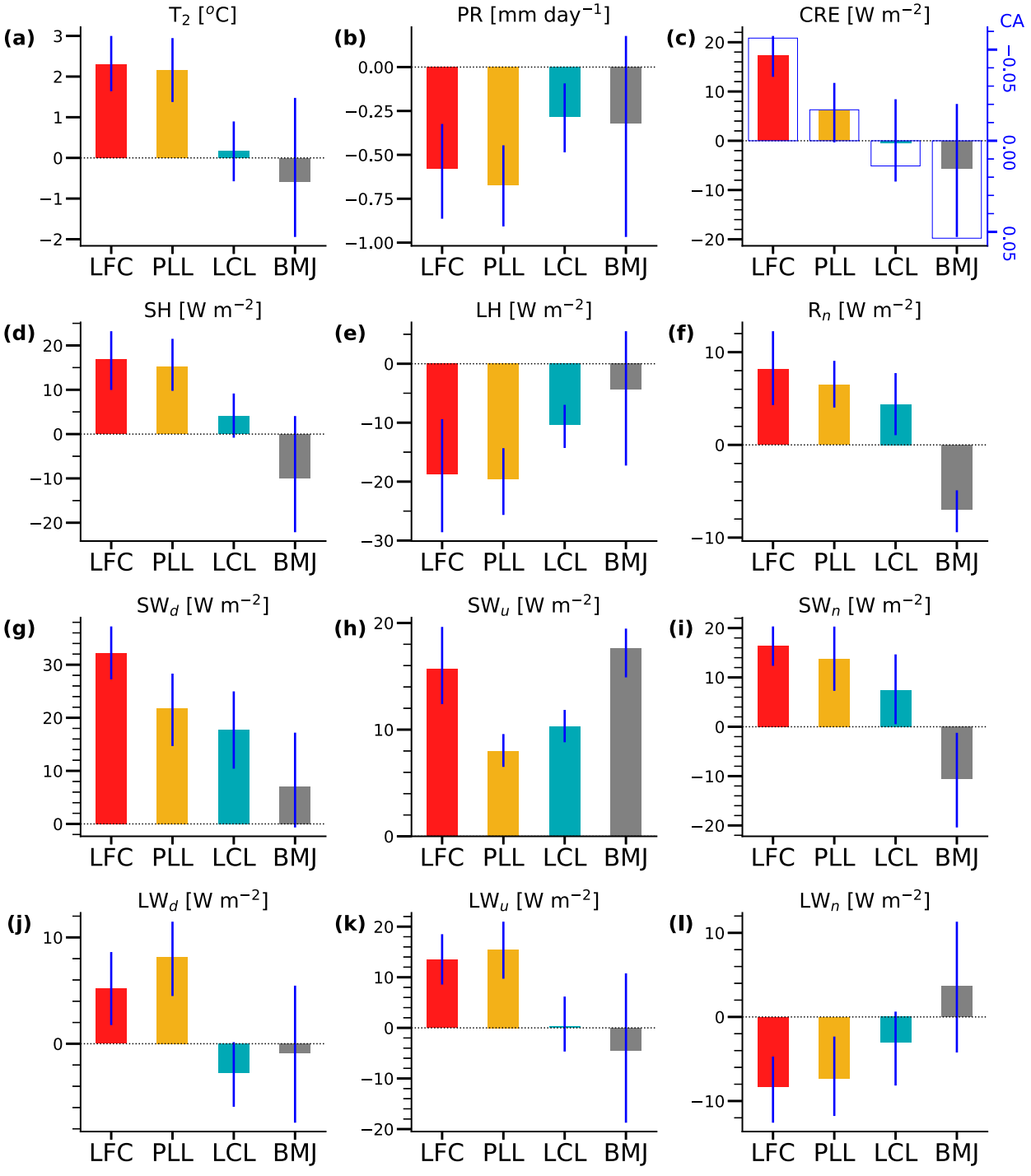

FIG. 5. Characteristics of summer CUS bias among multiple physical variables from four broad types (CMIP6X total 72 models) during 1980–2014. They are (a) near-surface temperature (T_2), (b) daily precipitation amount (PR), (c) cloud radiation effect (CRE) and total cloud albedo (CA), (d) surface sensible heat (SH), (e) surface latent heat (LH), (f) net radiation at the surface (R_n), (g) downwelling shortwave radiation at the surface (SW_d), (h) upwelling shortwave radiation at the surface (SW_u), (i) net shortwave radiation at the surface (SW_n), (j) downwelling longwave radiation at the surface (LW_d), (k) surface longwave emission (LW_u), and (l) net longwave radiation at the surface (LW_n). Blue lines represent 95% confidence interval estimated by the bootstrapping method of Linnet (2000) with 1000 resamples.

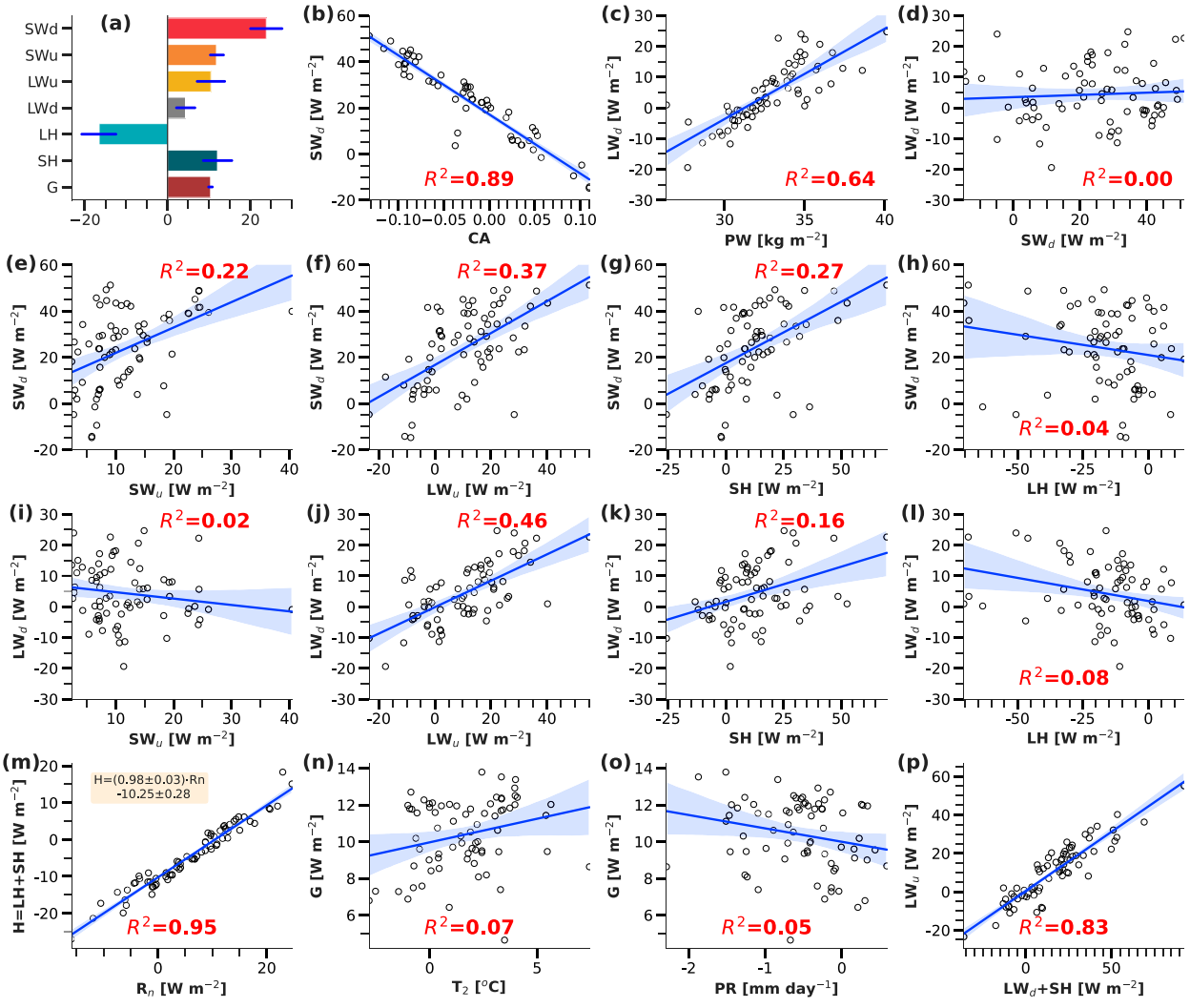

FIG. 6. Association of biases in downwelling shortwave/longwave radiation at the surface (SW$_d$, LW$_d$,) with upwelling shortwave (SW$_u$) and emission longwave radiation (LW$_u$) at the surface, sensible/latent heat (SH, LH), and the cloud albedo (CA) or precipitable water (PW). (a) Averaged bias over all models, with blue lines representing the 95% confidence interval estimated by the bootstrapping method of Linnet (2000) with 1000 resamples. (b)–(l) The other scatter diagrams show the linear regression line with the 95% confidence interval (shading) and explained variance (red text). Regional mean (during 1980–2014 in CUS) biases from individual models (CMIP6X total 72 models) are denoted with empty black circles. Also shown are the associations (m) between the moist enthalpy flux (H) and the net radiation at surface (R_n), (n),(o) between the ground heat flux (G) and the temperature/precipitation bias, and (p) between LW$_u$ and the summation of LW$_d$ and SH.

cooling components simulated by the 72 models. The model spread in precipitation biases was mainly determined by those in SH and LW$_u$, which explained respectively 64% and 55% variance. The remaining radiation components, LW$_d$, SW$_u$, SW$_d$, TOA shortwave reflection (SW$_u^{\mathrm{TOA}}$), and outgoing longwave radiation (OLR), contributed much less to the diabatic cooling bias spread, explaining only 20%, 1%, 6%, 10%, and 4% variance, respectively; their total contribution was 20%. The last component LH was mostly compensated by SH (negatively correlated, explaining 78% covariance), canceling each other with their sum explained only 9% variance of precipitation biases. As such, precipitation biases were positively correlated with LH biases, which explained 74% variance, somewhat larger than SH. This result agrees with the widely held view that summer CUS has the wet-soil advantage, where wetter soil more favorably triggers convection (Findell and Eltahir 2003a,b; Koster et al. 2004; Ferguson and Wood 2011; Santanello et al. 2018). The energetic negative association between SH and PR reflected the underlying mechanism, that is, less LH reduced PR and left more energy for SH.

4. Developing the analytical bias model (ABM) to explain CMIP ensemble errors

The results presented in section 3 with Figs. 1–7 were all based on statistical analyses of CMIP model biases, showing a

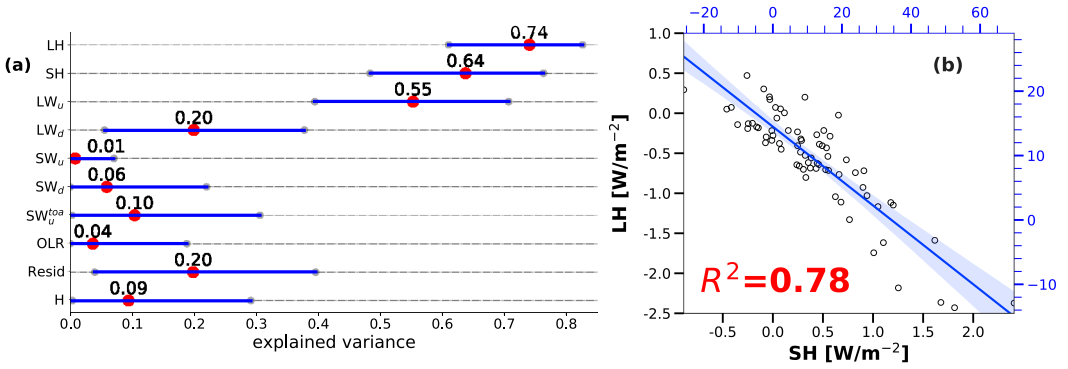

FIG. 7. (a) The variance of PR biases explained by biases in LH, SH, radiation fluxes (see Table S6 for the notations), the residual of diabatic cooling (Resid = $LW_d + SW_u + SW_d + SW_u^{TOA} + OLR$), and moist enthalpy flux (LH + SH). Red circles with text above represent the explained variance, while blue segments with gray dots represent its 95% confidence interval estimated by the bootstrapping method of Steck and Jaakkola (2003) with 1000 resamples. (b) The association of SH and LH biases. The scatter diagrams show the linear regression line with the 95% confidence interval (shading) and explained variance (red text). Regional mean (during 1980–2014 in CUS) biases from individual CMIP6X models are denoted with empty black circles. Shown are in both diabatic cooling (W m^{-2}) (blue axes), and the equivalent precipitation (mm day^{-2}) units (black axes).

strong sensitivity to cumulus parameterization and significant relationships among key variables. These statistics cannot reveal nonlinear feedbacks essential to the coupled climate system nor reflect the causality of the actual physical system (Pearl 2009). Therefore, below we developed the ABM that captures the physical mechanisms underlying the CUS warm and dry biases. The LCL-based schemes simulated more low-level cloud and less total precipitable water, which reduced net shortwave and longwave downwelling at the surface (Figs. 4–6; S&L). Thus, the purpose of the ABM is to link the coupled (T_2, PR) biases to these surface radiation (LW_d, SW_n) differences. Our focus is on the key mechanisms and feedbacks that dominate the systematic biases and spreads among the latest CMIP6X models. Table S4 summarizes the important notations used in this study. For radiation fluxes, unless specifically denoted with a superscript TOA for the top of atmosphere, they are all at the surface, where the subscripts d, u, and n represent downwelling, upwelling, and net (d minus u) fluxes, respectively. Table 2 lists the key parameters used in the ABM, where the subscript 0 depicts a prescribed value, and the overbar denoted CUS summer averaging over all analysis years (35) and across all CMIP6X models (72).

As illustrated in Fig. 8, the ABM approximates the interactions among the three major physical processes that link 1) PR to R_n, SH, and ground temperature (T_g) (the blue hub), 2) LH to PR and R_n (the green hub), and 3) LW_u and T_g to LW_d and SH (the orange hub). The term T_2 is the consequence of these interactions, coupling T_g, LH, LW_d, and SW_n with various feedbacks. While all these variables can be directly acquired from individual CMIP6X models, our strategy is to keep some as intermediate or latent variables (rather than inputs) in the ABM by parameterizing the physical relationships to activate the essential feedbacks. The derivation of this ABM starts with the approximations for the atmosphere,

TABLE 2. Model parameters and their reference values.

Symbol	Description	Numerical value
A_{B0}	Ball parameter	0.2
A_{pr0}	$A_{pr0} \equiv \alpha_{pr0}\overline{T}_{g0}/4 + \mathcal{E}_0$	4.5 W m^{-2}
α_{H0}	Slope for warm advection effects on for dry-static energy divergence parameterization	3.8 W m^{-2} K^{-1}
α_{pr0}	$\alpha_{pr0} \equiv \alpha_{H0} - C_{g0}/G_0$	0.37 W m^{-2} K^{-1}
C_{g0}	$4\sigma T_{g0}^3$	5.9 W m^{-2} K^{-1}
C_{h0}	Multiplication of air density, near-surface temperature, heat capacity of air, and drag coefficient	31 W m^{-2} K^{-1}
ε_0	Combined terms (the shortwave reflectance at TOA, reference value for dry-static energy divergence and OLR with other residual terms)	−24 W m^{-2}
G_0	Subsurface energy flux	10 W m^{-2}
$\mathcal{G}_0$	Greenhouse factor	2.0
$\mathcal{L}$	Precipitation to energy unit conversion	29 mm day^{-1} W^{-1} m^2
$\overline{T}_{g0}$	Reference temperature in CUS summer (ensemble mean of CMIP6X)	298 K
γ_{eb0}	Ratio of the pressure at the emission height over the pressure depth of the boundary layer	0.8
ω_{ap0}	Ratio of additional water amount for evapotranspiration to total precipitation	0.24

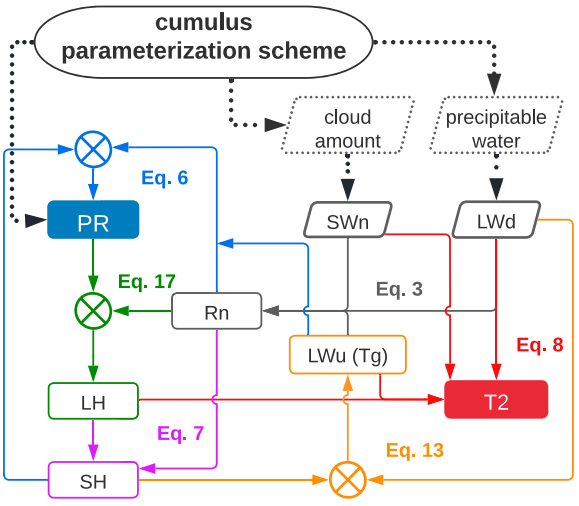

FIG. 8. The conceptual diagram summarizing the main components and linkages of the analytical bias model (ABM). A parallelogram lists a main input (SW_n, LW_d), while a rectangle lists a key intermediate variable (LH, SH, LW_u, R_n) and a final output (T_2, PR) with color shading. An arrow links a major equation with its respective in/out variables using the same color for text and lines. A dashed parallelogram lists a latent variable (cloud amount, precipitable water) not directly used in the ABM. A circled X depicts an essential hub linking the associated physical processes. Different cumulus parameterization schemes define different cloud bases and hence produce variations in low-level cloud and precipitable water amounts that drive the ABM.

land surface, and boundary layer energy budgets ($E_{\rm atm}$, $E_{\rm sfc}$, $E_{\rm bl}$) averaged in the CUS summer,

$$\frac{\partial E_{\rm atm}}{\partial t} = {\rm PR} \times \mathscr{L} - Q - H_{\rm div}, \tag{1a}$$

$$\frac{\partial E_{\rm sfc}}{\partial t} \approx R_n - G_0 - ({\rm SH} + {\rm LH}), \quad \text{and} \tag{1b}$$

$$\frac{\partial E_{\rm bl}}{\partial t} \approx \left(\frac{\partial E_{\rm bl}}{\partial t}\right)_{\rm rad} + \left(\frac{\partial E_{\rm bl}}{\partial t}\right)_{\rm sfc} + \left(\frac{\partial E_{\rm bl}}{\partial t}\right)_{\rm ent}, \tag{1c}$$

where the subscripts div, rad, sfc, and ent depict the energy tendencies due to atmospheric horizontal divergence, surface radiation, surface turbulence, and boundary entrainment, respectively.

Following Muller and O'Gorman (2011), Eq. (1a) decomposes the latent heat into total diabatic cooling (Q) and dry static energy divergence ($H_{\rm div}$) in the atmosphere. Assuming the long-term mean in an *atmospheric energy balance* state $\partial E_{\rm atm}/\partial t \approx 0$, we can estimate precipitation as

$${\rm PR} \times \mathscr{L} = Q + H_{\rm div} \equiv (R_n + {\rm OLR} + {\rm SW}_n^{\rm TOA} - {\rm SH}) + H_{\rm div}, \tag{2}$$

where surface net radiation is

$$R_n \equiv {\rm SW}_n + {\rm LW}_n \equiv ({\rm SW}_d - {\rm SW}_u) + ({\rm LW}_d - {\rm LW}_u) \tag{3}$$

and $\mathscr{L} \equiv L_v/\rho_w = 29\,({\rm mm^{-1}\,day\,W\,m^{-2}})$ is the unit conversion from precipitation to energy. Adopting the conventional definition of the coupled surface–atmosphere Earth system emission temperature $T_e \equiv ({\rm OLR}/\sigma)^{1/4}$ and greenhouse factor $\mathscr{G} \equiv (T_g/T_e)^4$, where T_g is the ground (radiative) temperature $T_g \equiv ({\rm LW}_u/\sigma)^{1/4}$, we can approximate OLR in Eq. (2) using its Taylor expansion with respect to $T_{g0} \equiv \overline{T}_g = 298\,{\rm K}$,

$$\begin{aligned}{\rm OLR} \equiv \sigma T_e^4 = \frac{\sigma T_g^4}{\mathscr{G}} &= \frac{\sigma}{\mathscr{G}}\left[T_{g0}^4 + \frac{\partial {\rm OLR}}{\partial T_g} dT_g + \mathscr{O}^2(T_g)\right] \\ &\approx \frac{\sigma}{\mathscr{G}_0}[T_{g0}^4 + 4T_{g0}^3(T_g - T_{g0})] \equiv \frac{C_{g0}T_{g0}}{4\mathscr{G}_0} + \frac{C_{g0}}{\mathscr{G}_0}T_g',\end{aligned} \tag{4}$$

where $C_{g0} \equiv 4\sigma T_{g0}^3 \approx 5.9\,{\rm W\,m^{-2}\,K^{-1}}$, and $T_g' \equiv T_g - T_{g0}$. Our analysis showed that the warm and dry bias is insensitive to $\mathscr{G}$, which is thus approximated as $\mathscr{G}_0 \equiv \sigma T_g^4/{\rm OLR} \approx 1.78$ (Table 2).

We further decompose $H_{\rm div}$ in Eq. (2) into a reference state and its thermal plus dynamic perturbations—the biases in this study. We consider that in midlatitudes the dynamic contribution is much larger (than the thermal term) and linearly depends on vertical velocity at 500 hPa (Muller and O'Gorman 2011). Then we apply the omega equation to link vertical velocity to the near-surface temperature latitudinal gradient. Taken together [see derivation details in the online supplemental material for Eq. (5)]:

$$H_{\rm div} \approx \mathscr{E}_H - \alpha_H T_g', \tag{5}$$

where all residual terms are merged into $\mathscr{E}_H$ (Table 2). Equation (5) implies that a warmer CUS is associated with a smaller temperature latitudinal gradient, which reduces the warm advection from the Gulf states, suppresses the rising motion in CUS, and thus contributes negatively to local precipitation through the flux divergence (α_H). The advection contributes to the bias through the dry-static energy, rather than the moisture transport through the Great Plains low-level jet (LLJ). Although the moisture transport was considered as a dominant factor determining summer CUS precipitation and extremes (Li et al. 2011; Gimeno et al. 2016), our analysis based on the indices of LLJ (Liang et al. 2004) or Bermuda high (Zhu and Liang 2013) showed no statistically significant relationship with precipitation or temperature biases (Fig. S7).

Combining Eqs. (2)–(5) yields (see the corresponding processes of the blue hub in Fig. 8)

$${\rm PR} \times \mathscr{L} \approx R_n - {\rm SH} - (\alpha_{H0} - C_{g0}/\mathscr{G}_0) \times T_g' + \mathscr{E}_0. \tag{6}$$

Our derivation of the α_{H0} term (dry-static energy divergence dependence on temperature) offers a new insight into the conceptually similar energy transport coefficient in the Sellers-Budyko model. By fitting CMIP6X data, we obtained $\alpha_{H0} = 3.8\,{\rm W\,m^{-2}\,K^{-1}}$, which is the same as the transport coefficient used in McGuffie and Henderson-Sellers (2014). Similarly, we obtained the residual term $\mathscr{E}_0 = {\rm SW}_n^{\rm TOA} + [C_{g0}T_{g0}/(4\mathscr{G}_0)] + \mathscr{E}_H = -24\,({\rm W\,m^{-2}})$ (see the online supplemental material for details). For brevity we define $\alpha_{\rm pr0} \equiv \alpha_{H0} - C_{g0}/\mathscr{G}_0 = 0.37\,{\rm W\,m^{-2}K^{-1}}$ (Table 2).

Assuming the long-term mean in a *surface energy balance* state $\partial E_{\rm sfc}/\partial t \approx 0$, Eq. (1b) gives

$$\mathrm{SH} = R_n - G_0 - \mathrm{LH}. \tag{7}$$

Our analysis showed that the subsurface flux G_0 has no coherent relationship with the warm and dry bias in CUS summer among the CMIP6X models (Fig. 6). Thus, this term can be approximated as $G_0 \approx (\overline{\mathrm{SW}}_n + \overline{\mathrm{LW}}_n) - (\overline{\mathrm{LH}} + \overline{\mathrm{SH}})$. In CUS summer, G_0 is about 10 W m^{-2} as calculated from CMIP6X simulations (Fig. 6), but the observed value (merged energy flux measurements from eddy covariance towers with remote sensing and meteorological data; Jung et al. 2019) is much smaller (0.45 W m^{-2}). This may indicate systematic energy leaking from land surface components in the CMIP6X models.

Letting SH $\approx C_{h0}(T_g - T_2)$ and taking Eq. (7) for the departure from the long-term mean of all CMIP6X models yield a diagnostic approximation (see Table 2 for C_{h0} and C_{g0} values) of

$$T_2' \approx T_g'\left(1 + \frac{C_{g0}}{C_{h0}}\right) + \frac{\mathrm{LH}' - \mathrm{LW}_d' - \mathrm{SW}_n'}{C_{h0}}, \tag{8}$$

where $F' \equiv F - \overline{F}_0$ for any variable F such as T_g, LH, LW_d, or SW_n.

To estimate LW_u [and hence T_g in Eq. (8)], we assume a *boundary layer energy balance* state $\partial E_{\rm bl}/\partial t \approx 0$ in Eq. (1c). We neglect the phase change effect such as due to raindrop evaporation or condensation in this layer (Brubaker and Entekhabi 1995, 1996; Otterman 1990), horizontal advection terms (Otterman 1990; Betts 2000), and shortwave absorption (Otterman 1990; Gentine et al. 2018). We also assume equilibrium potential temperature in the boundary layer is vertically invariant (Brubaker and Entekhabi 1995, 1996), which can be derived from entropy maximization (Bohren and Albrecht 2000), and is proportional to T_2. Under these conditions, we have

$$\frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm LW} + \frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm SH} + \frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm BE} \approx 0. \tag{9}$$

Equation (9) indicates that the T_2 tendency is dominated by surface longwave plus sensible heating as well as entrainment cooling at the top of the boundary layer.

To estimate the longwave cooling $(\partial T_2/\partial t)_{\rm LW}$, we assume the local thermodynamic equilibrium and follow Coakley and Yang (2014) for the longwave emission and its pressure (p) gradient at the surface:

$$\begin{cases} \mathrm{LW}_n = -\pi\mathscr{B}[\![p_t]\!]\exp\left(-\dfrac{p_s}{p_e}\right) - \pi\mathscr{B}[\![p_t]\!]\left[1 - \exp\left(\dfrac{p_t - p_s}{p_e}\right)\right] \\ \left.\dfrac{d\mathrm{LW}_n}{dp}\right|_{p\to p_s} = \left\{\pi\mathscr{B}[\![p_t]\!]\exp\left(-\dfrac{p_s}{p_e}\right) + \pi\mathscr{B}[\![p_t]\!]\left[1 - \exp\left(\dfrac{p_t - p_s}{p_e}\right)\right]\right\}\Big/p_e = -\mathrm{LW}_n/p_e, \end{cases} \tag{10}$$

where a double bracket denotes for a function of the enclosed variables, $\mathscr{B}$ is the Plank function, and p_s, p_t, and p_e are the pressures at the surface, tropopause, and emission height. This leads to an approximation:

$$\frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm LW} \approx -\left.\frac{d\mathrm{LW}_n}{dp}\right|_{p\to p_s} = \frac{\mathrm{LW}_n}{p_e} \equiv \frac{\mathrm{LW}_d - \mathrm{LW}_u}{p_e}. \tag{11}$$

The PBL top entrainment cooling can be parameterized as a fraction (A_{B0}; Table 2) of the surface sensible heating (Stull 1988), which in turn is determined by the pressure gradient of SH (Nicholls and Smith 1982). Thus, the tendency due to these two factors can be jointly estimated as

$$\frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm SH} + \frac{C_p}{g}\left(\frac{\partial T_2}{\partial t}\right)_{\rm BE} \approx (1 - A_{B0})\frac{\partial \mathrm{SH}}{\partial p} \approx (1 - A_{B0})\frac{\mathrm{SH}}{p_s - p_b}, \tag{12}$$

where A_{B0} is the fractional Ball parameter and p_b the pressure at the PBL top.

Combining Eqs. (9), (11), and (12) leads to the estimation of LW_u from LW_d and SH (see the corresponding processes of the orange hub in Fig. 8):

$$\mathrm{LW}_u - \mathrm{LW}_d \approx \frac{p_e}{p_s - p_b}(1 - A_{B0}) \times \mathrm{SH} \equiv \gamma_{\rm eb} \times \mathrm{SH}, \tag{13}$$

where $\gamma_{\rm eb}$ is a dimensionless parameter defined as the product of the emission to PBL depth pressure ratio $p_e/(p_s - p_b)$ and the Ball fraction residual $(1 - A_{B0})$. A more stable surface layer, which depresses the PBL depth, may also be associated with a lower cloud layer, which lowers the emission height; both conditions increase $\gamma_{\rm eb}$. Reversely, an unstable surface layer along with a higher cloud layer decrease $\gamma_{\rm eb}$. Their disassociation can change $\gamma_{\rm eb}$ in either direction. Thus, $\gamma_{\rm eb}$ depicts the partitioning of surface net longwave cooling between sensible heat loss and cloud warming feedback. Based on ERA5 daily mean analysis data, the CUS summer long-term average $\gamma_{\rm eb0} \approx 0.8$ (see the supplemental material for calculation details). As limited by data availability (e.g., the daily PBL depth was missing in many models), we explored the ABM sensitivity to $\gamma_{\rm eb}$ variations around $\gamma_{\rm eb0}$ in counterfactual experiments (Fig. S8). We found that temperature is more sensitive to $\gamma_{\rm eb}$ than precipitation and a larger $\gamma_{\rm eb}$ leads to a larger temperature bias.

To estimate the remaining unknown term (LH) in Eq. (8), we adopt the Budyko (1974) curve. Using the conventional definition of the Bowen ratio BR ≡ SH/LH and the aridity

index AI ≡ E_p/PR, which depicts soil water availability, we can estimate surface potential ET (E_p) in the energy form:

$$E_p = (R_n - G_0) \times \mathscr{L}^{-1}. \tag{14}$$

Considering actual ET ≡ LH × $\mathscr{L}^{-1}$ and dividing Eq. (7) by PR leads to

$$\mathrm{AI} \equiv \frac{E_p}{\mathrm{PR}} = \frac{\mathrm{ET}}{\mathrm{PR}} + \frac{\mathrm{SH} \times \mathscr{L}^{-1}}{\mathrm{PR}} = \frac{\mathrm{ET}}{\mathrm{PR}} + \frac{\mathrm{SH}}{\mathrm{LH}} \times \frac{\mathrm{ET}}{\mathrm{PR}} = \frac{\mathrm{ET}}{\mathrm{PR}}(1 + \mathrm{BR}) \Rightarrow \mathrm{ET} = \frac{\mathrm{AI} \times \mathrm{PR}}{1 + \mathrm{BR}}. \tag{15}$$

Budyko (1974) presented his curve: $\mathrm{BR} \approx \mathscr{C}_b \mathrm{AI} = \mathrm{AI}/\sqrt{\mathrm{AI} \times \tanh[\![\mathrm{AI}^{-1}]\!](1 - e^{-\mathrm{AI}})} - 1$. The curve assumes that soil water storage and ground heat flux are negligible (Gerrits et al. 2009). This assumption is reasonable for the long-term annual average, but not valid for CUS summer where soil water storage is an important source (Ting et al. 2021). Considering that soil moisture provides additional water for ET, which is equivalent to increasing precipitation, we may define an effective aridity index $\mathrm{AI}_{\mathrm{eff}} \equiv \mathrm{AI}(1 + \omega_{\mathrm{ap0}})^{-1}$ to approximate the storage effect,

$$\mathrm{BR} \approx \mathscr{C}_b \mathrm{AI}_{\mathrm{eff}} = \mathscr{C}_b \mathrm{AI}[\![(1 + \omega_{\mathrm{ap0}})^{-1}]\!], \tag{16}$$

where the dimensionless ω_{ap} is defined as the ratio of additional water (WA) for ET to total precipitation ($\omega_{\mathrm{ap}} \equiv$ WA/PR). Following Hartmann (2015), ignoring minor surface condensation, summer average soil water storage change rate WS ≈ PR − ET − RO, where RO is runoff. That is, PR − ET ≈ RO + WS. We can see that any reduction from (RO + WS) would be equivalent to the increasing water WA for ET. Thus, $\omega_{\mathrm{ap}} \approx -$(RO + WS)/PR. Our ABM sensitivity analysis (Fig. S8) showed the warm and dry bias is insensitive to ω_{ap} in the range from 0.1 to 0.5, which may cause insignificant uncertainty in temperature bias ≤ 0.1 K and precipitation bias ≤ 0.1 mm day^{-1}. Since soil moisture and runoff data were not available for many models, we used ERA5 to derive the approximation as $\omega_{\mathrm{ap0}} \approx 0.24$ (Table 2). Conceptually, ω_{ap} is like the recycling rate, and our estimate is close to the value from the diagnosis of Dominguez et al. (2006).

Combining Eqs. (14)–(16) links LH with PR and R_n as a function of the modified Budyko curve (see the corresponding processes of the green hub in Fig. 8):

$$\begin{cases} \mathrm{LH} \equiv \mathscr{L} \times \mathrm{ET} \approx \dfrac{\mathscr{L} \times \mathrm{PR} \times \mathrm{AI}}{1 + \mathscr{C}_b[\![\mathrm{AI}_{\mathrm{eff}}]\!]} \equiv \mathscr{L} \times \mathrm{PR} \times f_b[\![\mathrm{AI}_{\mathrm{eff}}]\!] \\ f_b[\![\mathrm{AI}_{\mathrm{eff}}]\!] \equiv (1 + \omega_{\mathrm{ap0}})\sqrt{\mathrm{AI}_{\mathrm{eff}} \times \tanh[\![\mathrm{AI}_{\mathrm{eff}}^{-1}]\!] \times (1 - e^{-\mathrm{AI}_{\mathrm{eff}}})}\ . \\ \mathrm{AI}_{\mathrm{eff}}[\![R_n, \mathrm{PR}]\!] \equiv \dfrac{R_n - G_0}{\mathscr{L} \times \mathrm{PR} \times (1 + \omega_{\mathrm{ap0}})} \end{cases} \tag{17}$$

The complete ABM consists of Eqs. (3), (6)–(8), (13), and (17). It requires specific input (LW_d, SW_n) as long-term (1980–2010) summer averages from individual CMIP6X models, the prescribed parameters from the ensemble mean of all CMIP6X models (α_{pr0}, C_{h0}, $\mathscr{E}_0$, G_0, T_{g0}) and the ERA5 reanalysis (ω_{ep0}, γ_{eb0}), and some constants (Table 2). Through recursive iterations, the ABM solves the coupled (T_2, PR) responses to the biases and spreads of (LW_d, SW_n) among the CMIP6X models, where LH, SH, R_n, LW_u, and E_p are intermediate dependent variables. However, the ABM is a nonlinear balance system that can have multiple solutions, depending most sensitively on the starting PR value (PR_s). This mainly results from the feedback processes that depend on the Budyko curve. Dry conditions may lead to a stronger positive soil moisture–precipitation feedback and greater sensitivity than wet anomalies (Hohenegger et al. 2009; Koster et al. 2004; Zaitchik et al. 2013). Therefore, to obtain a realistic solution, we start the ABM iterations with PR_s set to the long-term averaged summer mean PR from each of the CMIP6X models.

5. ABM simulations and stability analyses for CMIP6X bias mechanisms

Figure 9 compares the (T_2, PR) biases simulated by the ABM and CMIP6X models. The ABM can explain 82% temperature and 81% precipitation variance of the biases among all CMIP6X models. These percentages or coefficients of determination (R^2) were based on the one-to-one relationship between the ABM and CMIP6X biases, rather than the linear regression fitting them. This avoids overlooking the confidence of the result with a high correlation but incorrect amplitude (Legates and McCabe 1999). The regression (slope, intercept) of CMIP6X with ABM are (0.81° ± 0.04°, 0.32° ± 0.10°C) for temperature and (1.19 ± 0.04, −0.12 ± 0.02 mm day^{-1} for precipitation. As compared with the existing CAUSES outcome (R^2 = 0.61), our result represents a significant improvement for both temperature and precipitation, having higher R^2 scores, slopes closer to 1, and intercepts closer to 0. The first two measures indicate that the ABM well captures the spread of biases among all CMIP6X models, while the third measure depicts a high certainty in estimating the system sensitivity as derived in Eq. (S8). Therefore, the ABM can represent the principal relationships among surface climate variables in summer CUS, rendering a significant confidence for application to study the physical mechanisms of biases.

Below we conducted ABM *counterfactual* experiments to identify the key mechanisms for the warm and dry bias. In each experiment, for a specific input factor (LW_d, SW_n, PR_s), we replaced its values from all individual CMIP6X models with the mean of the subset models using the LCL-based cumulus parameterizations to quantify its contribution to the bias (see the supplemental material for calculation details). This experimental approach enables us to consider interactions among feedback processes in the entire system, which is critical but often overlooked in previous studies on bias mechanisms (Stephens 2005). Additional experiments were conducted to examine the ABM result sensitivity to the prescribed key parameters (γ_{eb0}, ω_{ap0}). This approach may not test uncertainties associated with all physical assumptions and numerical approximations used in building the ABM. However, our focus is to seek a useful *alternative* physical proxy that enables a better understanding of the complex nonlinear feedbacks underlying the

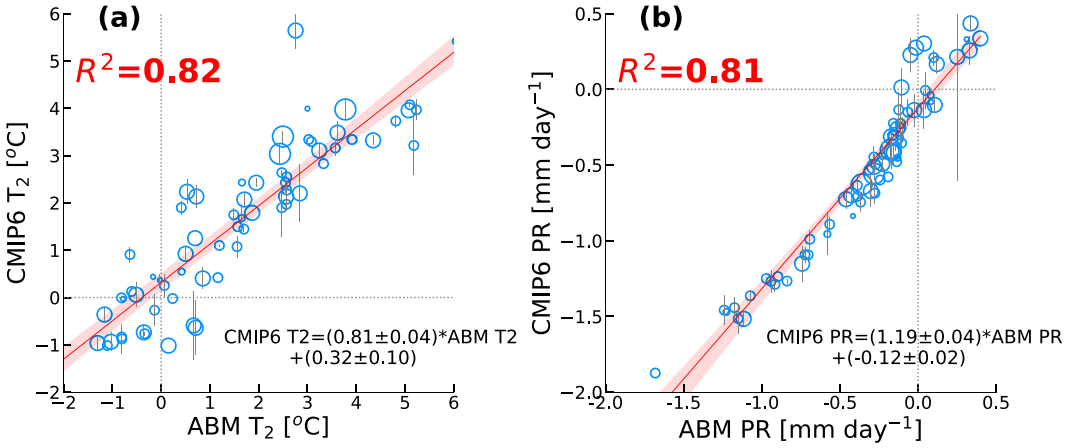

FIG. 9. The analytical bias model (ABM) estimated vs the CMIP6X simulated CUS summer (1980–2014) mean biases in temperature (T_2; °C) and precipitation (PR; mm day^{-1}). The red line depicts the linear regression, the red shading shows the 95% confidence intervals, and the red label gives the explained variance (R^2 score based on one-to-one bias correspondences rather than their regression fitting). The blue circle denotes the bias from each of the 72 CMIP6X models averaged over all its realizations, the light vertical line for one standard deviation among the realizations, and the radius for the model resolution.

fully coupled CMIP6X simulations and hence offers an explicit explanation for the spread of biases among models. This alternative approach is recommended for feedback analysis (Stephens 2005).

Figure 10 compares the attributions to the CUS summer warm and dry bias. When adopting a common LW_d value as the mean of the subset models using the LCL-based parameterization, other cumulus schemes (LFC, PLL, BMJ) would significantly reduce the warm bias by (1.7°, 2.1°, 0.7°C) and increase precipitation by (0.6, 0.6, 0.4 mm day^{-1}). The BMJ models contain significantly larger uncertainty than other groups. Correspondingly, when adopting a common SW_n

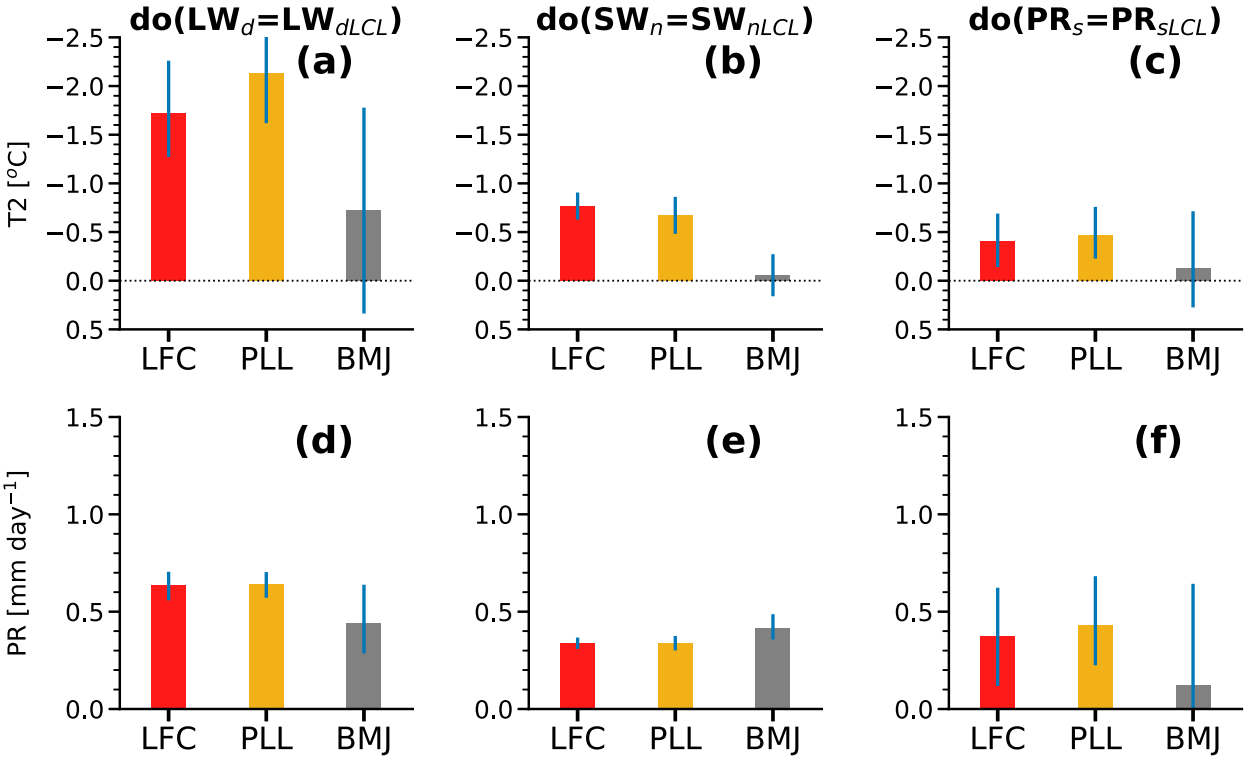

FIG. 10. The ABM estimated (a)–(c) temperature and (d)–(f) precipitation effects if assigning the result of the LCL cumulus scheme to the respective values in all models for (left) LW_d, (center) SW_n, and (right) PR_s. Different cumulus schemes are separated by colors.

value from the mean of the LCL models, other cumulus schemes would reduce the warm bias by 0.7°, 0.7°, and 0.0°C and increase precipitation by 0.3, 0.3, and 0.4 mm day^{-1}, respectively. These warm and dry bias reductions induced by SW_n are much smaller than those by LW_d, with the former accounting for only 20%–50% of the latter. See appendix for the explanation why CUS summer is more sensitivity to LW_d. The uncertainties in estimating the contributions decrease from LW_d to SW_n. In addition, when adopting a common PR_s value from the mean of the LCL models, other cumulus schemes would reduce the warm bias by 0.4°, 0.5°, and 0.1°C and increase precipitation by 0.4, 0.4, and 0.1 mm day^{-1}, respectively. These bias reductions are comparable with those induced by SW_n, although the uncertainties of the estimation increase.

The above ABM experiments allow us to attribute the major cause for the warm and dry bias to the excessive radiative energy reaching the surface together with insufficient rainfall. An important question is why most of the excessive energy flows into the sensible rather than latent heat and what mechanisms amplify initial perturbations toward significant biases—a potential runaway problem. To address this question, we conducted a feedback analysis following Roe (2009) and showed that the system bias behavior depends on the f_b range (see appendix for derivations). The f_b range can be divided by a threshold at 0.6, which corresponds approximately to a mean precipitation amount of 8.8 mm day^{-1}. Since summer CUS mean precipitation is only 3.3 mm day^{-1} in observations and less than 3.9 mm day^{-1} in all CMIP6X models, the combined radiative forcing $SW'_n \times \gamma_{eb} + LW'_d \times (1 + \gamma_{eb})$ in Eq. (A5) induces a runaway feedback for both temperature and precipitation. Larger surface radiative forcing leads to warmer T_g and so larger LW_u. Warmer T_g produces larger SH and so higher T_2, which in turn suppresses warm advection toward CUS and hence reduces vertical motion. Meanwhile, larger LW_u entering the atmosphere reduces the demand for latent heat release from precipitation. Both lead to smaller PR [Eq. (6)] and so drier soil for less ET and LH [Eq. (17)], which is balanced by larger SH [Eq. (7)]. Given the energy balance in the surface boundary layer [Eq. (1c)], larger SH leads to higher LW_u. This positive feedback evolves into a runaway problem in most CMIP6X models using non-LCL cumulus schemes, causing significant warm and dry biases.

6. Summary and conclusions

This study showed that substantial warm and dry biases persisted in summer CUS among most climate models despite remarkable improvements in the latest CMIP6 (Eyring et al. 2019) and even in high-resolution simulations. Meanwhile, S&L demonstrated that CWRF defining the cloud base at LCL in cumulus parameterization can significantly reduce the biases. Following this idea, we grouped all 72 CMIP6X models by four cloud base definitions in their cumulus schemes. Our composite analysis indicated that the models using the LCL-based schemes simulated systematically smaller warm and dry biases than others. A more comprehensive analysis of precipitation decompositions, radiative fluxes, and surface budgets confirmed that the cloud base definition is the dominant factor determining the spread of the biases among the CMIP6X models and those using the LCL definition performed the best.

The statistical analyses cannot determine the causality or physical mechanisms underlying the relationships among model biases. To identify these feedback mechanisms and quantify the key contributions to the spread of the biases among the CMIP6X models, we developed a physically based ABM that captures the principal energy balances of the coupled surface–atmosphere system. The ABM consists of three major physical parameterizations that link 1) precipitation with surface net radiation minus sensible heat flux and ground temperature departure from a reference state; 2) surface latent heat flux (and hence evapotranspiration) with precipitation in proportion to a generalized Budyko curve of an effective aridity index; and 3) surface net longwave radiation with a negative proportion of sensible heat flux. With these new parameterizations, we introduced two key dimensionless numbers, γ_{eb} and ω_{pe}, to respectively characterize the land–atmosphere coupling strength through energy and water exchanges. The ABM solves recursively the coupled (T_2, PR) responses to the driving biases and spreads of (LW_d, SW_n, PR_s) among the CMIP6X models. This is analogous to the surface forcing and response framework (Andrews et al. 2009), the ABM enables us to gain physical insight into the bias causes. Using the prescribed parameters from observational data or multimodel ensemble means, the ABM has significant explanatory power for CUS summer warm and dry biases, capturing 80% variance of temperature and precipitation biases among all CMIP6X models.

We conducted ABM counterfactual experiments to attribute the key factors for the warm and dry bias and found that the LCL cumulus parameterization reduces the warm and dry bias through two principal mechanisms. First, it produces more low clouds and smaller total precipitable water, which reduce respectively the downwelling (and net) shortwave and longwave radiation reaching the surface and thus provide less energy available for surface heating and evapotranspiration, causing a cooler and wetter soil. Second, it produces more rainfall and wetter soil conditions, which prevent the land surface runaway toward the hot-drought state as driven by the strong positive ET–PR feedback and hence eliminate the warm and dry bias. Further analysis revealed that, under relative dry conditions in most CMIP6X models, primary surface downwelling longwave errors combined with secondary net shortwave flux errors drive the climate system into the runaway stage, causing a lock in a dry and warm bias loop.

This study offered a systematic approach to determine common deficiencies, identify sensitive factors, quantify relative contributions, and understand dominant mechanisms for CUS summer warm and dry biases and their spreads across the CMIP climate models. The developed ABM, with large explanatory power for CUS summer biases, can be applied to other regions with similar climate characteristics. This analytical modeling and physical understanding laid a solid foundation for not only climate model improvement but also more interpretable and reliable climate prediction in CUS. Our ABM analysis supports the projected coexistence of severe heat stress with more devastating drought (Zhou et al. 2019; Ting et al. 2021) and the observational interpretation of the potential feedback processes in CUS summer (Taylor et al.

2012). It also explains how underestimation of precipitation (Lin et al. 2017), cloud (Cheruy et al. 2014), and evapotranspiration (Mueller and Seneviratne 2014) can jointly cause the warm and dry bias and quantifies their relative contributions. Our ABM analysis indicated that CUS summer warm and dry biases among CMIP6X models are attributed mostly by surface downwelling longwave radiation errors and second by surface net shortwave radiation errors, with the former 2–5 times larger. These two errors as weighted by their relative contributions form an effective radiative forcing to induce runaway temperature and precipitation feedbacks, which collaborate to cause CUS summer warm and dry biases. Note that the ABM circumvents explicit representation of the large uncertainty associated with clouds by taking input of surface radiation fluxes from CMIP6X models. This implicit treatment could underestimate the buffering effect from cloud-induced negative feedbacks (Stephens and Webster 1981). The consequence of this limitation warrants further investigation.

Acknowledgments. The research was supported by the U.S. National Science Foundation Innovations at the Nexus of Food, Energy and Water Systems under Grant EAR1639327 and the U.S. Department of Agriculture–National Institute of Food and Agriculture under Grant 20206801231674 for developing the Dashboard for Agricultural Water use and Nutrient management at the University of Maryland and Grant 2015-34263-24070 for the UV-B Monitoring and Research Program at Colorado State University. Additional support came from the U.S. Department of Commerce Educational Partnership Program–National Oceanic and Atmospheric Administration Center for Atmosphere and Meteorology (NCAS-M) under Grant NA16SEC4810006. The simulations and analyses were conducted on supercomputers supported by the National Center for Atmospheric Research Computational and Information Systems Lab, the Maryland Advanced Research Computing Center, and the Atmospheric River Program funded by the California Department of Water Resources and the Forecast Informed Reservoir Operations Program funded by the U.S. Army Corps of Engineers Engineer Research and Development Center. We acknowledge the World Climate Research Programme with its Working Group on Coupled Modeling for coordinating the CMIP5/6 efforts. We thank all the climate modeling groups for producing and making available their CMIP5/6 outputs, the Earth System Grid Federation (ESGF) for archiving this data and providing access, and the multiple funding agencies who support CMIP5/6 and ESGF.

Data availability statement. Data analyzed in this study were a re-analysis of existing data, which are openly available at locations cited in the reference section. Further documentation about data processing is available at the supplemental material.

APPENDIX

To linearize the ABM system, we fix Budyko function $f_b \approx f_{b0}$, approximate $(1 + x)^{1/4} \approx 1 + x/4$, expand $\mathrm{LW}_u = \mathrm{LW}_{u0} + \mathrm{LW}'_u$, and use $\mathrm{LW}_{u0} \equiv \sigma T_{g0}^4$. This leads to

$$T_g = \left(\frac{\mathrm{LW}_{u0} + \mathrm{LW}'_u}{\sigma}\right)^{1/4} \approx \left(\frac{\mathrm{LW}_{u0}}{\sigma}\right)^{1/4}\left(1 + \frac{\mathrm{LW}'_u}{4\mathrm{LW}_{u0}}\right)$$
$$= T_{g0}\left(\frac{3}{4} + \frac{\mathrm{LW}_u}{4\mathrm{LW}_{u0}}\right) = \frac{3T_{g0}}{4} + \frac{\mathrm{LW}_u}{C_{g0}}. \quad \text{(A1)}$$

The ABM system can now be simplified and linearized as

$$\begin{cases} \mathrm{LH} = \mathscr{L} f_{b0}\mathrm{PR} \\ \mathrm{SH} = R_n - \mathrm{LH} - G_0 \\ \mathrm{PR} = \mathscr{L}^{-1}(R_n + A_{\mathrm{pr0}} - \mathrm{SH} - \alpha_{\mathrm{pr0}}C_{g0}^{-1}\mathrm{LW}_u), \\ \mathrm{LW}_u = \gamma_{\mathrm{eb}}\mathrm{SH} + \mathrm{LW}_d \\ R_n = \mathrm{LW}_d - \mathrm{LW}_u + \mathrm{SW}_n \end{cases} \quad \text{(A2)}$$

where $A_{\mathrm{pr0}} \equiv \alpha_{\mathrm{pr0}}\overline{T}_{g0}/4 + \mathscr{E}_0 = 4.5\,(\mathrm{W\,m^{-2}})$. Since T_2 is a diagnostic variable, we proxy its feedback effect through T_g as they explain each other 95% of covariance (Fig. S9) and then substitute T_g by LW_u as in Eq. (A1).

From Eq. (A2),

$$\begin{cases} \mathrm{LW}_u = \underbrace{(f_{b0}n - \gamma_{\mathrm{eb}})}_{\mathrm{F_{LW_u}}}\mathrm{LW}_u + \mathscr{R}_0 \\ \mathrm{PR} = \underbrace{f_{b0}n}_{\mathrm{F_{PR}}}\mathrm{PR} + \mathscr{L}^{-1}\mathscr{R}_1 \\ n \equiv 1 + \gamma_{\mathrm{eb}}(1 + C_{g0}{}^{-1}\alpha_{\mathrm{pr0}}) \end{cases} \Rightarrow_{\frac{\partial f_b}{\partial \mathrm{PR}}<0} \begin{cases} \dfrac{\partial F_{\mathrm{LW}_u}}{\partial \mathrm{PR}} < 0 \\ \dfrac{\partial F_{\mathrm{PR}}}{\partial \mathrm{PR}} < 0 \end{cases}, \quad \text{(A3)}$$

where F_{LW_u} and F_{PR} are the feedback factors, and $\mathscr{R}_0$ and $\mathscr{R}_1$ are the radiative forcings:

$$\begin{cases} \mathscr{R}_0 \equiv (1 - f_{b0})\mathscr{R}_{\mathrm{e}} + c_0 \\ \mathscr{R}_1 \equiv -m\mathscr{R}_{\mathrm{e}}\gamma_{\mathrm{eb}}\alpha_{\mathrm{pr0}}C_{g0}{}^{-1} + c_1 \\ \mathscr{R}_{\mathrm{e}} \equiv \mathrm{SW}_n\gamma_{\mathrm{eb}} + \mathrm{LW}_d(1 + \gamma_{\mathrm{eb}}) \\ m \equiv \dfrac{1 - f_{b0}n}{1 - f_{b0}n + \gamma_{\mathrm{eb}}} \\ c_0 \equiv -A_{\mathrm{pr0}}f_{b0}\gamma_{\mathrm{eb}} - G_0\gamma_{\mathrm{eb}} \\ c_1 \equiv mA_{\mathrm{pr0}}(1 + \gamma_{\mathrm{eb}}) + mG_0[\gamma_{\mathrm{eb}}\alpha_{\mathrm{pr0}}C_{g0}{}^{-1} + (1 + \gamma_{\mathrm{eb}})] \end{cases} \quad \text{(A4)}$$

Taking c_0 and c_1 as fixed values, the perturbed forcings (biases) for the system are

$$\begin{cases} \mathscr{R}'_0 = (1 - f_{b0})\mathscr{R}'_e \\ \mathscr{R}'_1 = -m\mathscr{R}'_e\gamma_{\mathrm{eb}}\alpha_{\mathrm{pr0}}C_{g0}^{-1} \\ \mathscr{R}'_2 = \mathrm{SW}'_n\gamma_{\mathrm{eb}} + \mathrm{LW}'_d(1 + \gamma_{\mathrm{eb}}) \end{cases} \quad \text{(A5)}$$

As shared between $\mathscr{R}'_{0,1}$, the effective forcing $\mathscr{R}'_e$ combines surface radiative perturbations SW'_n and LW'_d as weighted by γ_{eb} and $(1 + \gamma_{\mathrm{eb}})$. The fraction of the weights depicts the relative contribution of SW'_n and LW'_d to the bias of T_g ($\propto T_2$) or PR. As such, the LW_d forcing plays a dominant role.

Equation (A3) indicates that drier conditions exacerbate the feedback. We can derive two thresholds to determine whether the feedback is in a negative or positive or runaway state:

$$\begin{cases} \kappa_{\mathrm{lw0}} \equiv n^{-1}\gamma_{\mathrm{eb}} \\ \kappa_{\mathrm{lw1}} \equiv n^{-1}(\gamma_{\mathrm{eb}} + 1) \end{cases} \Rightarrow \begin{cases} f_{b0} < \kappa_{\mathrm{lw0}} \Rightarrow F_{\mathrm{LW}_u} < 0 \\ \kappa_{\mathrm{lw0}} < f_{b0} < \kappa_{\mathrm{lw1}} \Rightarrow 0 < F_{\mathrm{LW}_u} < 1 \\ \kappa_{\mathrm{lw1}} < f_{b0} \Rightarrow F_{\mathrm{LW}_u} > 1 \end{cases}. \tag{A6}$$

For LW_u and so T_g, it is a negative feedback when $f_{b0} < \kappa_{\mathrm{lw0}} \approx 0.4$, but a positive feedback when $\kappa_{\mathrm{lw0}} < f_{b0} < \kappa_{\mathrm{lw1}} \approx 0.97$. Since the summer CUS perturbation of surface net shortwave plus downwelling longwave radiative is positive ($\mathrm{LW}'_d > 0, \mathrm{SW}'_n > 0$), the forcing is positive $\mathscr{R}'_0 > 0$. The positive feedback along with a positive forcing leads to warmer T_g. Once f_{b0} exceeds κ_{lw1}, the feedback enters a runaway stage, the feedback-factor analysis is no longer valid (Roe 2009). However, we can prove that the runaway effect always enhances warm and dry biases (see the online supplemental material for derivation details).

For precipitation it is a positive feedback when $f_{b0} < n^{-1} \approx 0.6$. Given $m \equiv (1 - f_{b0}n)/(1 - f_{b0}n + \gamma_{\mathrm{eb}}) > 0$, and for $\mathrm{LW}'_d > 0, \mathrm{SW}'_n > 0$, we have $\mathscr{R}'_1 > 0$. Thus, the positive feedback with a negative forcing reduces precipitation. Drier conditions increase AI and f_b, which strengths the feedback [Eq. (A3)] and thus drive precipitation to enter a runaway condition.

REFERENCES

Andrews, T., P. M. Forster, and J. M. Gregory, 2009: A surface energy perspective on climate change. *J. Climate*, **22**, 2557–2570, https://doi.org/10.1175/2008JCLI2759.1.

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. *J. Atmos. Sci.*, **31**, 674–701, https://doi.org/10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2.

Ardilouze, C., L. Batté, B. Decharme, and M. Déqué, 2019: On the link between summer dry bias over the U.S. Great Plains and seasonal temperature prediction skill in a dynamical forecast system. *Wea. Forecasting*, **34**, 1161–1172, https://doi.org/10.1175/WAF-D-19-0023.1.

Barlage, M., F. Chen, R. Rasmussen, Z. Zhang, and G. Miguez-Macho, 2021: The importance of scale-dependent groundwater processes in land–atmosphere interactions over the central United States. *Geophys. Res. Lett.*, **48**, e2020GL092171, https://doi.org/10.1029/2020GL092171.

Bellprat, O., S. Kotlarski, D. Lüthi, R. De Elía, A. Frigon, R. Laprise, and C. Schär, 2016: Objective calibration of regional climate models: Application over Europe and North America. *J. Climate*, **29**, 819–838, https://doi.org/10.1175/JCLI-D-15-0302.1.

Betts, A. K., 2000: Idealized model for equilibrium boundary layer over land. *J. Hydrometeor.*, **1**, 507–523, https://doi.org/10.1175/1525-7541(2000)001<0507:IMFEBL>2.0.CO;2.

——, and M. J. Miller, 1986: A new convective adjustment scheme. Part II: Single column tests using GATE wave, BOMEX, ATEX and Arctic air-mass data sets. *Quart. J. Roy. Meteor. Soc.*, **112**, 693–709, https://doi.org/10.1002/qj.49711247308.

——, 2007: Coupling of water vapor convergence, clouds, precipitation, and land–surface processes. *J. Geophys. Res.*, **112**, D10108, https://doi.org/10.1029/2006JD008191.

Bohren, C. F., and B. A. Albrecht, 2000: Atmospheric thermodynamics. *Amer. J. Phys.*, **68**, 1159, https://doi.org/10.1119/1.1313524.

Brubaker, K. L., and D. Entekhabi, 1995: An analytic approach to modeling land–atmosphere interaction: 1. Construct and equilibrium behavior. *Water Resour. Res.*, **31**, 619–632, https://doi.org/10.1029/94WR01772.

——, and ——, 1996: Analysis of feedback mechanisms in land–atmosphere interaction. *Water Resour. Res.*, **32**, 1343–1357, https://doi.org/10.1029/96WR00005.

Budyko, M. I., 1974: *Climate and Life*. Vol. 18. Academic Press, 508 pp.

Chang, Y., S. D. Schubert, R. D. Koster, A. M. Molod, and H. Wang, 2019: Tendency bias correction in coupled and uncoupled global climate models with a focus on impacts over North America. *J. Climate*, **32**, 639–661, https://doi.org/10.1175/JCLI-D-18-0598.1.

Cheruy, F., J. L. Dufresne, F. Hourdin, and A. Ducharne, 2014: Role of clouds and land–atmosphere coupling in midlatitude continental summer warm biases and climate change amplification in CMIP5 simulations. *Geophys. Res. Lett.*, **41**, 6493–6500, https://doi.org/10.1002/2014GL061145.

——, and Coauthors, 2020: Improved near-surface continental climate in IPSL–CM6A–LR by combined evolutions of atmospheric and land surface physics. *J. Adv. Model. Earth Syst.*, **12**, e2019MS002005, https://doi.org/10.1029/2019MS002005.

Coakley, J. A., Jr., and P. Yang, 2014: *Atmospheric Radiation: A Primer with Illustrative Solutions*. John Wiley & Sons, 256 pp.

DelGenio, A. D., and M. S. Yao, 1993: Efficient cumulus parameterization for long-term climate studies: The GISS scheme. *The Representation of Cumulus Convection in Numerical Models, Meteor. Monogr.*, No. 46, Amer. Meteor. Soc., 181–184.

Dirmeyer, P. A., and K. L. Brubaker, 2007: Characterization of the global hydrologic cycle from a back-trajectory analysis of atmospheric water vapor. *J. Hydrometeor.*, **8**, 20–37, https://doi.org/10.1175/JHM557.1.

Dominguez, F., P. Kumar, X.-Z. Liang, and M. Ting, 2006: Impact of atmospheric moisture storage on precipitation recycling. *J. Climate*, **19**, 1513–1530, https://doi.org/10.1175/JCLI3691.1.

Ehret, U., E. Zehe, V. Wulfmeyer, K. Warrach-Sagi, and J. Liebert, 2012: HESS opinions "Should we apply bias correction to global and regional climate model data?" *Hydrol. Earth Syst. Sci.*, **16**, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012.

Eyring, V., and Coauthors, 2019: Taking climate model evaluation to the next level. *Nat. Climate Change*, **9**, 102–110, https://doi.org/10.1038/s41558-018-0355-y.

Ferguson, C. R., and E. F. Wood, 2011: Observed land–atmosphere coupling from satellite remote sensing and reanalysis. *J. Hydrometeor.*, **12**, 1221–1254, https://doi.org/10.1175/2011JHM1380.1.

Findell, K. L., and E. A. B. Eltahir, 2003a: Atmospheric controls on soil moisture–boundary layer interactions. Part I: Framework development. *J. Hydrometeor.*, **4**, 552–569, https://doi.org/10.1175/1525-7541(2003)004<0552:ACOSML>2.0.CO;2.

——, and ——, 2003b: Atmospheric controls on soil moisture–boundary layer interactions. Part II: Feedbacks within the continental United States. *J. Hydrometeor.*, **4**, 570–583, https://doi.org/10.1175/1525-7541(2003)004<0570:ACOSML>2.0.CO;2.

Geleyn, J.-F., B. Catry, Y. Bouteloup, and R. Brozkova, 2008: A statistical approach for sedimentation inside a microphysical precipitation scheme. *Tellus*, **60**, 649–662, https://doi.org/10.1111/j.1600-0870.2007.00323.x.

Gentine, P., G.-J. Steeneveld, B. G. Heusinkveld, and A. A. M. Holtslag, 2018: Coupling between radiative flux divergence

and turbulence near the surface. *Quart. J. Roy. Meteor. Soc.*, **144**, 2491–2507, https://doi.org/10.1002/qj.3333.

Gerrits, A. M. J., H. H. G. Savenije, E. J. M. Veling, and L. Pfister, 2009: Analytical derivation of the Budyko curve based on rainfall characteristics and a simple evaporation model. *Water Resour. Res.*, **45**, W04403, https://doi.org/10.1029/2008WR007308.

Gimeno, L., and Coauthors, 2016: Major mechanisms of atmospheric moisture transport and their role in extreme precipitation events. *Annu. Rev. Environ. Resour.*, **41**, 117–141, https://doi.org/10.1146/annurev-environ-110615-085558.

Gregory, D., and P. R. Rowntree, 1990: A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure. *Mon. Wea. Rev.*, **118**, 1483–1506, https://doi.org/10.1175/1520-0493(1990)118%3c1483:AMFCSW%3e2.0.CO;2.

Guérémy, J. F., 2011: A continuous buoyancy based convection scheme: One- and three-dimensional validation. *Tellus*, **63A**, 687–706, https://doi.org/10.1111/j.1600-0870.2011.00521.x.

Hartmann, D. L., 2015: *Global Physical Climatology*. 2nd ed. Elsevier, 481 pp.

Hohenegger, C., P. Brockhaus, C. S. Bretherton, and C. Schär, 2009: The soil moisture–precipitation feedback in simulations with explicit and parameterized convection. *J. Climate*, **22**, 5003–5020, https://doi.org/10.1175/2009JCLI2604.1.

Hourdin, F., and Coauthors, 2020: LMDZ6A: The atmospheric component of the IPSL climate model with improved and better tuned physics. *J. Adv. Model. Earth Syst.*, **12**, e2019MS001892, https://doi.org/10.1029/2019MS001892.

Hu, H., L. R. Leung, and Z. Feng, 2021: Early warm-season mesoscale convective systems dominate soil moisture–precipitation feedback for summer rainfall in central United States. *Proc. Natl. Acad. Sci. USA*, **118**, e2105260118, https://doi.org/10.1073/pnas.2105260118.

Jiang, R., L. Sun, C. Sun, and X.-Z. Liang, 2021: CWRF downscaling and understanding of China precipitation projections. *Climate Dyn.*, **57**, 1079–1096, https://doi.org/10.1007/s00382-021-05759-z.

Jung, M., and Coauthors, 2019: The FLUXCOM ensemble of global land–atmosphere energy fluxes. *Sci. Data*, **6**, 74, https://doi.org/10.1038/s41597-019-0076-8.

Kim, D., A. H. Sobel, A. D. Del Genio, Y. Chen, S. J. Camargo, M. S. Yao, M. Kelley, and L. Nazarenko, 2012: The tropical subseasonal variability simulated in the NASA GISS general circulation model. *J. Climate*, **25**, 4641–4659, https://doi.org/10.1175/JCLI-D-11-00447.1.

Klein, S. A., X. Jiang, J. Boyle, S. Malyshev, and S. Xie, 2006: Diagnosis of the summertime warm and dry bias over the US southern Great Plains in the GFDL climate model using a weather forecasting approach. *Geophys. Res. Lett.*, **33**, L18805, https://doi.org/10.1029/2006GL027567.

Koster, R. D., and Coauthors, 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, https://doi.org/10.1126/science.1100217.

Kotz, M., A. Levermann, and L. Wenz, 2022: The effect of rainfall changes on economic production. *Nature*, **601**, 223–227, https://doi.org/10.1038/s41586-021-04283-8.

Kunkel, K. E., X.-Z. Liang, and J. Zhu, 2010: Regional climate model projections and uncertainties of U.S. summer heat waves. *J. Climate*, **23**, 4447–4458, https://doi.org/10.1175/2010JCLI3349.1.

Legates, D. R., and G. J. McCabe Jr., 1999: Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.*, **35**, 233–241, https://doi.org/10.1029/1998WR900018.

Li, S., and Coauthors, 2019: Reducing climate model biases by exploring parameter space with large ensembles of climate model simulations and statistical emulation. *Geosci. Model Dev.*, **12**, 3017–3043, https://doi.org/10.5194/gmd-12-3017-2019.

Li, W., L. Li, R. Fu, Y. Deng, and H. Wang, 2011: Changes to the North Atlantic subtropical high and its role in the intensification of summer rainfall variability in the southeastern United States. *J. Climate*, **24**, 1499–1506, https://doi.org/10.1175/2010JCLI3829.1.

Li, Y., K. Guan, G. D. Schnitkey, E. DeLucia, and B. Peng, 2019: Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Global Change Biol.*, **25**, 2325–2337, https://doi.org/10.1111/gcb.14628.

Liang, X.-Z., 2022: Extreme rainfall slows the global economy. *Nature*, **601**, 193–194, https://doi.org/10.1038/d41586-021-03783-x.

——, L. Li, A. Dai, and K. E. Kunkel, 2004: Regional climate model simulation of summer precipitation diurnal cycle over the United States. *Geophys. Res. Lett.*, **31**, L24208, https://doi.org/10.1029/2004GL021054.

——, M. Xu, K. E. Kunkel, G. A. Grell, and J. S. Kain, 2007: Regional climate model simulation of U.S.–Mexico summer precipitation using the optimal ensemble of two cumulus parameterizations. *J. Climate*, **20**, 5201–5207, https://doi.org/10.1175/JCLI4306.1.

——, K. E. Kunkel, G. A. Meehl, R. G. Jones, and J. X. L. Wang, 2008: Regional climate models downscaling analysis of general circulation models present climate biases propagation into future change projections. *Geophys. Res. Lett.*, **35**, L08709, https://doi.org/10.1029/2007GL032849.

——, and Coauthors, 2012: Regional climate–weather research and forecasting model. *Bull. Amer. Meteor. Soc.*, **93**, 1363–1387, https://doi.org/10.1175/BAMS-D-11-00180.1.

——, and Coauthors, 2017: Determining climate effects on US total agricultural productivity. *Proc. Natl. Acad. Sci. USA*, **114**, E2285–E2292, https://doi.org/10.1073/pnas.1615922114.

——, Q. Li, H. Mei, and M. Zeng, 2019: Multi-grid nesting ability to represent convections across the gray zone. *J. Adv. Model. Earth Syst.*, **11**, 4352–4376, https://doi.org/10.1029/2019MS001741.

Lin, Y., M. Zhao, Y. Ming, J. C. Golaz, L. J. Donner, S. A. Klein, V. Ramaswamy, and S. Xie, 2013: Precipitation partitioning, tropical clouds, and intraseasonal variability in GFDL AM2. *J. Climate*, **26**, 5453–5466, https://doi.org/10.1175/JCLI-D-12-00442.1.

——, W. Dong, M. Zhang, Y. Xie, W. Xue, J. Huang, and Y. Luo, 2017: Causes of model dry and warm bias over central U.S. and impact on climate projections. *Nat. Commun.*, **8**, 881, https://doi.org/10.1038/s41467-017-01040-2.

Linnet, K., 2000: Nonparametric estimation of reference intervals by simple and bootstrap-based procedures. *Clin. Chem.*, **46**, 867–869, https://doi.org/10.1093/clinchem/46.6.867.

Liu, C., and Coauthors, 2017: Continental-scale convection-permitting modeling of the current and future climate of North America. *Climate Dyn.*, **49**, 71–95, https://doi.org/10.1007/s00382-016-3327-9.

Loeb, N. G., and Coauthors, 2018: Clouds and the Earth's Radiant Energy System (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product. *J. Climate*, **31**, 895–918, https://doi.org/10.1175/JCLI-D-17-0208.1.

Lopez, P., 2002: Implementation and validation of a new prognostic large-scale cloud and precipitation scheme for climate and

data-assimilation purposes. *Quart. J. Roy. Meteor. Soc.*, **128**, 229–257, https://doi.org/10.1256/00359000260498879.

Ma, H.-Y., and Coauthors, 2018: CAUSES: On the role of surface energy budget errors to the warm surface air temperature error over the central United States. *J. Geophys. Res. Atmos.*, **123**, 2888–2909, https://doi.org/10.1002/2017JD027194.

McGuffie, K., and A. Henderson-Sellers, 2014: *The Climate Modelling Primer*. 4th ed. John Wiley and Sons, 456 pp.

Mesinger, F., and Coauthors, 2006: North American Regional Reanalysis. *Bull. Amer. Meteor. Soc.*, **87**, 343–360, https://doi.org/10.1175/BAMS-87-3-343.

Morcrette, C. J., and Coauthors, 2018: Introduction to CAUSES: Description of weather and climate models and their near-surface temperature errors in 5 day hindcasts near the southern Great Plains. *J. Geophys. Res. Atmos.*, **123**, 2655–2683, https://doi.org/10.1002/2017JD027199.

Mueller, B., and S. I. Seneviratne, 2014: Systematic land climate and evapotranspiration biases in CMIP5 simulations. *Geophys. Res. Lett.*, **41**, 128–134, https://doi.org/10.1002/2013GL058055.

Mueller, N. D., E. E. Butler, K. A. McKinnon, A. Rhines, M. Tingley, N. M. Holbrook, and P. Huybers, 2016: Cooling of US Midwest summer temperature extremes from cropland intensification. *Nat. Climate Change*, **6**, 317–322, https://doi.org/10.1038/nclimate2825.

Muller, C. J., and P. A. O'Gorman, 2011: An energetic perspective on the regional response of precipitation to climate change. *Nat. Climate Change*, **1**, 266–271, https://doi.org/10.1038/nclimate1169.

Nicholls, S., and F. B. Smith, 1982: On the definition of the flux of sensible heat. *Bound.-Layer Meteor.*, **24**, 121–127, https://doi.org/10.1007/BF00121804.

NOAA, 2021: NOAA National Centers for Environmental Information (NCEI) U.S. billion-dollar weather and climate disasters. Accessed December 2021, https://www.ncdc.noaa.gov/billions/.

Ortiz-Bobea, A., T. R. Ault, C. M. Carrillo, R. G. Chambers, and D. B. Lobell, 2021: Anthropogenic climate change has slowed global agricultural productivity growth. *Nat. Climate Change*, **11**, 306–312, https://doi.org/10.1038/s41558-021-01000-1.

Otterman, J., 1990: A simple two-system-parameter model for surface-effected warming of the planetary boundary layer. *Bound.-Layer Meteor.*, **51**, 213–227, https://doi.org/10.1007/BF00122138.

Palmer, T., and B. Stevens, 2019: The scientific challenge of understanding and estimating climate change. *Proc. Natl. Acad. Sci. USA*, **116**, 24 390–24 395, https://doi.org/10.1073/pnas.1906691116.

Pearl, J., 2009: *Causality*. 2nd ed. Cambridge University Press, 464 pp.

Pendergrass, A. G., and Coauthors, 2020: Flash droughts present a new challenge for subseasonal-to-seasonal prediction. *Nat. Climate Change*, **10**, 191–199, https://doi.org/10.1038/s41558-020-0709-0.

Piriou, J.-M., J.-L. Redelsperger, J.-F. Geleyn, J.-P. Lafore, and F. Guichard, 2007: An approach for convective parameterization with memory: Separating microphysics and transport in grid-scale equations. *J. Atmos. Sci.*, **64**, 4127–4139, https://doi.org/10.1175/2007JAS2144.1.

Platnick, S., and Coauthors, 2015: MODIS cloud optical properties: User guide for the collection 6 level-2 MOD06/MYD06 product and associated level-3 datasets. Version 1, 145 pp., https://modis-images.gsfc.nasa.gov/_docs/C6MOD06OPUserGuide.pdf.

Qiao, F., and X.-Z. Liang, 2015: Effects of cumulus parameterizations on predictions of summer flood in the central United States. *Climate Dyn.*, **45**, 727–744, https://doi.org/10.1007/s00382-014-2301-7.

——, and ——, 2016: Effects of cumulus parameterization closures on simulations of summer precipitation over the United States coastal oceans. *J. Adv. Model. Earth Syst.*, **8**, 764–785, https://doi.org/10.1002/2015MS000621.

——, and ——, 2017: Effects of cumulus parameterization closures on simulations of summer precipitation over the continental United States. *Climate Dyn.*, **49**, 225–247, https://doi.org/10.1007/s00382-016-3338-6.

Rind, D., and Coauthors, 2020: GISS Model E2.2: A climate model optimized for the middle atmosphere—Model structure, climatology, variability, and climate sensitivity. *J. Geophys. Res. Atmos.*, **125**, e2019JD032204, https://doi.org/10.1029/2019JD032204.

Roe, G., 2009: Feedbacks, timescales, and seeing red. *Annu. Rev. Earth Planet. Sci.*, **37**, 93–115, https://doi.org/10.1146/annurev.earth.061008.134734.

Rupp, D. E., S. Li, P. W. Mote, K. M. Shell, N. Massey, S. N. Sparrow, D. C. Wallom, and M. R. Allen, 2017: Seasonal spatial patterns of projected anthropogenic warming in complex terrain: A modeling study of the western US. *Climate Dyn.*, **48**, 2191–2213, https://doi.org/10.1007/s00382-016-3200-x.

Salvucci, G. D., J. A. Saleem, and R. Kaufmann, 2002: Investigating soil moisture feedbacks on precipitation with tests of Granger causality. *Adv. Water Resour.*, **25**, 1305–1312, https://doi.org/10.1016/S0309-1708(02)00057-X.

Santanello, J. A., Jr., and Coauthors, 2018: Land–atmosphere interactions: The LoCo perspective. *Bull. Amer. Meteor. Soc.*, **99**, 1253–1272, https://doi.org/10.1175/BAMS-D-17-0001.1.

Schär, C., D. Lüthi, U. Beyerle, and E. Heise, 1999: The soil–precipitation feedback: A process study with a regional climate model. *J. Climate*, **12**, 722–741, https://doi.org/10.1175/1520-0442(1999)012<0722:TSPFAP>2.0.CO;2.

Smith, A. B., and J. L. Matthews, 2015: Quantifying uncertainty and variable sensitivity within the US billion-dollar weather and climate disaster cost estimates. *Nat. Hazards*, **77**, 1829–1851, https://doi.org/10.1007/s11069-015-1678-x.

Steck, H., and T. Jaakkola, 2003: Bias-corrected bootstrap and model uncertainty. *NIPS'03: Proc. 16th Int. Conf. on Neural Information Processing Systems*, Whistler, BC, Canada, https://papers.nips.cc/paper/2003/hash/2aaaddf27344ee54058548dc081c6541-Abstract.html.

Stephens, G. L., 2005: Cloud feedbacks in the climate system: A critical review. *J. Climate*, **18**, 237–273, https://doi.org/10.1175/JCLI-3243.1.

——, and P. J. Webster, 1981: Clouds and climate: Sensitivity of simple systems. *J. Atmos. Sci.*, **38**, 235–247, https://doi.org/10.1175/1520-0469(1981)038<0235:CACSOS>2.0.CO;2.

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

Sun, C., and X.-Z. Liang, 2020a: Improving U.S. extreme precipitation simulation: Sensitivity to physics parameterizations. *Climate Dyn.*, **54**, 4891–4918, https://doi.org/10.1007/s00382-020-05267-6.

——, and ——, 2020b: Improving US extreme precipitation simulation: Dependence on cumulus parameterization and underlying mechanism. *Climate Dyn.*, **55**, 1325–1352, https://doi.org/10.1007/s00382-020-05328-w.

——, and ——, 2022: Understanding and reducing warm and dry summer biases in the central United States: Improving

cumulus parameterization. *J. Climate*, **36**, 2015–2034, https://doi.org/10.1175/JCLI-D-22-0254.1.

Taylor, C. M., R. A. de Jeu, F. Guichard, P. P. Harris, and W. A. Dorigo, 2012: Afternoon rain more likely over drier soils. *Nature*, **489**, 423–426, https://doi.org/10.1038/nature11377.

Thornton, P. E., M. M. Thornton, B. W. Mayer, Y. Wei, R. Devarakonda, R. S. Vose, and R. B. Cook, 2016: Daymet: Daily surface weather data on a 1-km grid for North America, version 3. ORNL DAAC, accessed December 2021, https://doi.org/10.3334/ORNLDAAC/2129.

Tiedtke, M, 1989: A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.*, **117**, 1779–1800, https://doi.org/10.1175/1520-0493(1989)117<1779:ACMFSF>2.0.CO;2.

Ting, M., R. Seager, C. Li, H. Liu, and N. Henderson, 2021: Future summer drying in the U.S. Corn Belt and the role of midlatitude storm tracks. *J. Climate*, **34**, 9043–9056, https://doi.org/10.1175/JCLI-D-20-1004.1.

Van Weverberg, K., and Coauthors, 2018: CAUSES: Attribution of surface radiation biases in NWP and climate models near the U.S. Southern Great Plains. *J. Geophys. Res. Atmos.*, **123**, 3612–3644, https://doi.org/10.1002/2017JD027188.

Wei, J., and P. A. Dirmeyer, 2012: Dissecting soil moisture–precipitation coupling. *Geophys. Res. Lett.*, **39**, L19711, https://doi.org/10.1029/2012GL053038.

——, J. Zhao, H. Chen, and X.-Z. Liang, 2021: Coupling between land surface fluxes and lifting condensation level: Mechanisms and sensitivity to model physics parameterizations. *J. Geophys. Res. Atmos.*, **126**, e2020JD034313, https://doi.org/10.1029/2020JD034313.

Williams, I. N., Y. Lu, L. M. Kueppers, W. J. Riley, S. C. Biraud, J. E. Bagley, and M. S. Torn, 2016: Land–atmosphere coupling and climate prediction over the U.S. southern Great Plains. *J. Geophys. Res. Atmos.*, **121**, 12125–12144, https://doi.org/10.1002/2016JD025223.

Wu, T., 2012: A mass-flux cumulus parameterization scheme for large-scale models: Description and test with observations. *Climate Dyn.*, **38**, 725–744, https://doi.org/10.1007/s00382-011-0995-3.

Wulfmeyer, V., and Coauthors, 2018: A new research approach for observing and characterizing land–atmosphere feedback. *Bull. Amer. Meteor. Soc.*, **99**, 1639–1667, https://doi.org/10.1175/BAMS-D-17-0009.1.

Zaitchik, B. F., J. A. Santanello, S. V. Kumar, and C. D. Peters-Lidard, 2013: Representation of soil moisture feedbacks during drought in NASA Unified WRF (NU-WRF). *J. Hydrometeor.*, **14**, 360–367, https://doi.org/10.1175/JHM-D-12-069.1.

Zhang, C., S. Xie, S. A. Klein, H.-Y. Ma, S. Tang, K. Van Weverberg, C. J. Morcrette, and J. Petch, 2018: CAUSES: Diagnosis of the summertime warm bias in CMIP5 climate models at the ARM Southern Great Plains site. *J. Geophys. Res. Atmos.*, **123**, 2968–2992, https://doi.org/10.1002/2017JD027200.

Zhang, G. J., and N. A. McFarlane, 1991: Convective stabilization in midlatitudes. *Mon. Wea. Rev.*, **119**, 1915–1928, https://doi.org/10.1175/1520-0493(1991)119<1915:CSIM>2.0.CO;2.

——, and ——, 1995: Sensitivity of climate simulations to the parameterization of cumulus convection in the Canadian Climate Centre general circulation model. *Atmos.–Ocean*, **33**, 407–446, https://doi.org/10.1080/07055900.1995.9649539.

Zhou, S., and Coauthors, 2019: Land–atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proc. Natl. Acad. Sci. USA*, **116**, 18 848–18 853, https://doi.org/10.1073/pnas.1904955116.

Zhu, J., and X.-Z. Liang, 2013: Impacts of the Bermuda high on regional climate and ozone over the United States. *J. Climate*, **26**, 1018–1032, https://doi.org/10.1175/JCLI-D-12-00168.1.