The representation of soil moisture-atmosphere feedbacks across the Tibetan Plateau in CMIP6

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1007/s00376-023-2296-2

Publisher: Springer

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

• Original Paper •

The Representation of Soil Moisture–Atmosphere Feedbacks across the Tibetan Plateau in CMIP6※

Joshua TALIB[*1], Omar V. MÜLLER[2,3], Emma J. BARTON[1], Christopher M. TAYLOR[1,4], and Pier Luigi VIDALE[2,5]

[1]*U.K. Centre of Ecology and Hydrology*, *Wallingford*, *OX*10 8*BB, U.K.*

[2]*National Centre for Atmospheric Science*, *Reading*, *RG*6 6*BB, U.K.*

[3]*Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)*, *Centro de Estudios de Variabilidad y Cambio Climático (CEVARCAM)*, *Facultad de Ingeniería y Ciencias Hídricas*, *Universidad Nacional del Litoral*, *Santa Fe*, *Argentina*

[4]*National Centre for Earth Observation*, *Wallingford*, *OX*10 8*BB, U.K.*

[5]*Department of Meteorology*, *University of Reading*, *Reading, RG*6 6*BB, U.K.*

ABSTRACT

Thermal processes on the Tibetan Plateau (TP) influence atmospheric conditions on regional and global scales. Given this, previous work has shown that soil moisture–driven surface flux variations feed back onto the atmosphere. Whilst soil moisture is a source of atmospheric predictability, no study has evaluated soil moisture–atmosphere coupling on the TP in general circulation models (GCMs). In this study, we use several analysis techniques to assess soil moisture–atmosphere coupling in CMIP6 simulations including: instantaneous coupling indices; analysis of flux and atmospheric behaviour during dry spells; and a quantification of the preference for convection over drier soils. Through these metrics we partition feedbacks into their atmospheric and terrestrial components.

Consistent with previous global studies, we conclude substantial inter-model differences in the representation of soil moisture–atmosphere coupling, and that most models underestimate such feedbacks. Focusing on dry spell analysis, most models underestimate increased sensible heat during periods of rainfall deficiency. For example, the model-mean bias in anomalous sensible heat flux is 10 W m^{-2} ($\approx$25%) smaller compared to observations. Deficient dry-spell sensible heat fluxes lead to a weaker atmospheric response. We also find that most GCMs fail to capture the negative feedback between soil moisture and deep convection. The poor simulation of feedbacks in CMIP6 experiments suggests that forecast models also struggle to exploit soil moisture–driven predictability. To improve the representation of land–atmosphere feedbacks requires developments in not only atmospheric modelling, but also surface processes, as we find weak relationships between rainfall biases and coupling indexes.

Key words: model evaluation, land–atmosphere feedbacks, Tibetan Plateau, precipitation, surface energy balance

Article Highlights:

- Substantial inter-model differences in the representation of soil moisture–atmosphere feedbacks on the Tibetan Plateau.
- All models underestimate surface flux-atmospheric coupling during three-day dry spells.
- CMIP6 models are typically in contrast with observations and tend to favour deep convection over locally wetter soils.

---

1. Introduction

With an area greater than 2.5 million km^2 and an average elevation of approximately 4500 m, the Tibetan Plateau (TP) is the largest and highest plateau in the world. Thermal processes on the TP influence regional and global characteristics

of the atmospheric circulation (Duan and Wu, 2005; Jiang et al., 2008), whilst glacial melt from the plateau is a key source of several major Asian rivers (Immerzeel et al., 2010). Typical of a semi-arid environment, the partitioning of surface turbulent fluxes across the TP is partly controlled by soil moisture variability (Fan et al., 2019; Cui et al., 2020; Talib et al., 2021). For example, evaporative fraction decreases when soils are anomalously dry. Soil moisture–driven variations in surface turbulent fluxes not only influence low-level atmospheric humidity, but also partly control boundary-layer temperatures (Fan et al., 2019; Talib et al., 2021), the formation of deep convection (Barton et al., 2021; Zhao et al., 2022), and the regional atmospheric circulation (Chow et al., 2008; Wan et al., 2017; Talib et al., 2021). Across the TP and in other semi-arid regions, the atmospheric response to soil moisture variability is a crucial source of atmospheric predictability (Koster et al., 2010; Dirmeyer et al., 2018b). However, evaluations of the representation of soil moisture–atmosphere feedbacks in general circulation models (GCMs) are fairly limited, especially those focusing on processes across the TP. In this study, we assess to what extent the latest state-of-the-art GCMs from CMIP6 correctly represent soil moisture–atmosphere feedbacks across the TP.

Evaluating the representation of soil moisture–atmosphere feedbacks requires a consideration of two processes: (1) the surface flux response to soil moisture fluctuations (terrestrial); and (2) the sensitivity of atmospheric conditions to surface flux variations (atmospheric). Focusing on the terrestrial component, most models agree on the location of semi-arid regions where surface fluxes strongly respond to soil moisture variations (Koster et al., 2006; Dirmeyer, 2011; Schwingshackl et al., 2017). However, a positive evapotranspiration bias in GCMs (Mueller and Seneviratne, 2014) leads to an amplification of terrestrial coupling (Dirmeyer et al., 2018a). As well as this, the magnitude of terrestrial coupling substantially varies amongst models (Dirmeyer, 2011; Schwingshackl et al., 2017; Gallego-Elvira et al., 2019). These differences still remain when driving surface models with identical atmospheric forcing, indicating that differences in simulated terrestrial coupling is partly driven by the representation of surface characteristics and evaporative dynamics (Gevaert et al., 2018).

Alongside evaluating the terrestrial component of soil moisture–atmosphere feedbacks, studies have also assessed surface flux–atmospheric coupling. Given that the majority of GCMs overestimate soil moisture–driven surface flux variations (Dirmeyer et al., 2018a; Gallego-Elvira et al., 2019), it is unsurprising that most models amplify the occurrence and persistence of high temperatures associated with negative evapotranspiration anomalies (Ukkola et al., 2016). With regards to the influence of soil moisture on precipitation characteristics, the differing impacts of surface moisture on local atmospheric conditions makes it challenging to simulate the correct feedback. For example, when the surface is dry and moisture sourced from the surface decreases, low-level temperature and instability increase. The fine interplay between decreased moisture and increased instability can trigger and suppress local precipitation depending on the prevailing conditions. In general, GCMs tend to agree that rainfall is influenced by soil moisture across semi-arid regions (Koster et al., 2004; Dirmeyer et al., 2006; Müller et al., 2021b). However, the sign and magnitude of simulated soil moisture-precipitation feedbacks is sensitive to the representation of convection, horizontal resolution and choice of circulation model (Hohenegger et al., 2009; Taylor et al., 2013). GCMs with a horizontal resolution typical of an Earth system model (ESM; 100 to 250 km) commonly simulate a positive feedback. Meanwhile, a negative feedback is favoured when increasing horizontal resolution and using an explicit representation of convection, which is more consistent with observations (Taylor et al., 2012; Guillod et al., 2015; Barton et al., 2021).

The favouring of sensible heat over evapotranspiration across semi-arid regions of the TP when soils are dry (Guo et al., 2017; Cui et al., 2020) leads to a deepening of the planetary boundary layer and the development of a heat low circulation (Wan et al., 2017; Talib et al., 2021). Not only does this surface-driven heat low circulation affect local atmospheric conditions, but circulation characteristics beyond the TP are influenced by the development of an upper-level Rossby wave. In addition to controlling regional-scale circulation characteristics, Barton et al. (2021) show that soil moisture–driven surface flux variations partly control the initiation of deep convection on the TP. Deep convection is favoured over dry soils, close to wet-dry boundaries, due to the development of daytime mesoscale circulations induced by differential heating across surface soil moisture gradients (Pielke Sr, 2001). For the TP specifically, the sensitivity of deep convection to soil moisture gradients decreases with increased local topographic complexity (Barton et al., 2021), as orographic lifting can trigger deep convection irrespective of the surface state (Imamovic et al., 2017). Whilst previous studies have highlighted that soil moisture on the TP is a key source of atmospheric predictability (Wang et al., 2008; Talib et al., 2021; Barton et al., 2021), an evaluation of simulated soil moisture–atmosphere coupling remains to be performed.

To investigate the impact of anthropogenic climate change on environmental processes across the TP, a large number of studies use GCMs that simulate historical and future climates under various emission scenarios (i.e. Immerzeel et al., 2010). Whilst climate models have significantly developed over the past few decades, errors still exist partly due to our current level of understanding of the complex climate system. CMIP6 is the latest release of state-of-the-art simulations from different institutions around the world (Eyring et al., 2016). Most CMIP6 simulations have a cold near-surface temperature bias and overestimate total precipitation accumulations across the TP (Zhu and Yang, 2020). Given the sensitivity of atmospheric conditions to soil moisture on the TP (Wang et al., 2008; Talib et al., 2021; Barton et al., 2021), we assess the representation of soil moisture–atmosphere feedbacks in CMIP6 simulations. In

the subsequent section we discuss model data (section 2.1), observations (section 2.2) and analysis techniques used in this work (section 2.3). In section 3, we then document inter-model differences in previously-used soil moisture–atmosphere coupling metrics. Following this, in section 4, we assess land–atmosphere feedbacks using observational metrics specifically designed for the TP. This section is partitioned into two components including an evaluation of surface fluxes and atmospheric conditions during three-day dry spells (section 4.1), and an assessment of the sensitivity of deep convection to anomalous soil moisture (section 4.2). Finally, section 5 closes the paper with a discussion and conclusions.

2.   Methodology

2.1.   *Model data*

In this study, we examine soil moisture–atmosphere feedbacks in 18 historical CMIP6 simulations (Eyring et al., 2016). Table 1. summarises the details of each simulation, including atmospheric resolution and land surface model. CMIP6 experiments were selected based on the availability of sub-daily data of precipitation, surface-layer soil moisture and surface energy balance components. The CMIP6 model outputs used in this study are resolved at two temporal resolutions: precipitation, convective precipitation and surface fluxes are outputted as three-hourly means; whilst surface-layer soil moisture and surface air pressure are diagnosed

instantaneously every three hours. Unless stated, we only analyse model outputs from a single ensemble member (r1i1p1f1 or lowest available), between years 1980 to 2014, and during boreal summer months (June to August; JJA).

2.2.   *Observations*

Building on the analysis performed by Talib et al. (2021), in section 4.1 we assess the behaviour of simulated surface fluxes during dry spells. To evaluate the simulated surface energy balance, we approximate real world radiative and turbulent fluxes through amalgamating weather station measurements and satellite-based observations. Here we provide a brief overview of our methodology to derive surface energy balance components, however more detail can be found in Talib et al. (2021).

Six-hourly data from 49 weather stations above 3000 m from the China Meteorological Administration (CMA), locations later shown in Fig. 2u, is used to approximate surface sensible heat flux (SHF, W m^{-2}) and upward longwave radiation (LW$_{up}$, W m^{-2}). Using measurements of surface temperature (T_s, K), near-surface air temperature (T_a, K) and 10 m wind speed (v_{10m}, m s^{-1}), we estimate the surface SHF using a bulk formula:

$$\text{SHF} = \rho C_p C_{DH} v_{10m} (T_s - T_a), \qquad (1)$$

where C_p is the specific heat capacity of dry air at constant pressure (1005 J kg^{-1} K^{-1}); ρ is density (kg m^{-3}) and decreases exponentially with height; and C_{DH} is the drag coefficient for heat [assumed to be 4.0×10^{-3} for all stations fol-

Table 1.   CMIP6 models used in this study. Third and fourth columns show the horizontal and vertical resolution of the model's atmospheric component. We follow the typical convention of the modelling institution in stating the model resolution. "T" and "TL" denote spectral models with a triangular truncation with an "L" signifying models with a linear Gaussian grid. "C" refers to a cubed-sphere finite volume model, whilst an "N" prefix is used before stating the total number of two-gridpoint zonal waves that can be represented. Following the grid specification, the dimensions of the model output on a Gaussian longitude/latitude grid is given alongside the stated nominal resolution from Taylor et al. (2017).

Model name	Institution	Horizontal resolution	Vertical resolution	Land surface model
ACCESS-CM2	CSIRO-ARCCSS	N96; 192 × 144; 250 km	85 levels to 85 km	CABLE 2.5
ACCESS-ESM1-5	CSIRO	N96; 192 × 144; 250 km	85 levels to 85 km	CABLE 2.4 with biogeochemistry
BCC-CSM2-MR	BCC	T206; 320 × 160; 100 km	46 levels to 1.46 hPa	BCC-AVIM2
CNRM-CM6-1	CNRM-CERFACS	T127; 384 × 192; 150 km	91 levels to 0.01 hPa	ISBA-CTRIP
CNRM-ESM2-1	CNRM-CERFACS	T127; 384 × 192; 150 km	91 levels to 0.01 hPa	Surfex 8.0c
EC-Earth3-Veg	EC-Earth consortium	TL255; 512 × 256; 100 km	91 levels to 0.01 hPa	HTESSEL
GFDL-CM4	NOAA-GFDL	C96; 360 × 180; 100 km	33 levels to 1 hPa	GFDL-LM 4.0.1
GISS-E2-1-G	NASA-GISS	C48; 144 × 90; 250 km	40 levels to 0.1 hPa	GISS LSM
HadGEM3-GC31-LL	MOHC	N96; 192 × 144; 250 km	85 levels to 85 km	JULES-HadGEM3-GL7.1
HadGEM3-GC31-MM	MOHC	N216; 432 × 324; 100 km	85 levels to 85 km	JULES-HadGEM3-GL7.1
HadGEM3-GC31-HM	MOHC	N512; 1024 × 768; 50 km	85 levels to 85 km	JULES-HadGEM3-GL7.1
IPSL-CM6A-LR	IPSL	N96; 192 × 144; 250 km	79 levels to 40 km	ORCHIDEE
KACE-1-0-G	NIMS-KMA	N96; 192 × 144; 250 km	85 levels to 85 km	JULES-HadGEM3-GL7.1
MIROC6	MIROC	T85; 256 × 128; 250 km	81 levels to 0.004 hPa	MATSIRO6.0
MPI-ESM1-2-HAM	HAMMOZ consortium	T63; 192 × 96; 250 km	95 levels to 0.01 hPa	JSBACH 3.20
MPI-ESM1-2-HR	MPI-M	T127; 384 × 192; 100 km	95 levels to 0.01 hPa	JSBACH 3.20
MPI-ESM1-2-LR	MPI-M	T63; 192 × 96; 250 km	47 levels to 0.01 hPa	JSBACH 3.20
SAM0-UNICON	SNU	C96; 288 × 192; 100 km	30 levels to ≈ 2 hPa	CLM 4.0

lowing Duan and Wu (2008)]. We compute the outgoing surface LW_{up} using the Stefan-Boltzmann equation:

$$LW_{up} = \epsilon \sigma T_s^4 \,, \qquad (2)$$

where ϵ is the surface emissivity (assumed here to be fixed at 0.95) and σ is the Stefan-Boltzmann constant (5.67 × 10^{-8} W m^{-2} K^{-4}). We then combine computed surface SHF (Eqn. 1) and LW_{up} (Eqn. 2) with radiative surface fluxes derived from the Clouds and the Earth's Radiant Energy System (CERES; Loeb et al., 2003) to partition the surface energy balance. Radiative surface fluxes are outputted on a 1° latitude × 1° longitude grid. For each station, fluxes are selected from the nearest CERES grid point. The following equation is formulated after partitioning the surface energy balance into land surface forcings (left hand side) and surface fluxes that depend on land surface characteristics (right hand side):

$$SW_{net} + LW_{down} = LW_{up} + SHF + LHF + G \,, \qquad (3)$$

where SW_{net} denotes the net-downward shortwave radiation (W m^{-2}); LW_{down} denotes the downward longwave radiation (W m^{-2}); LHF denotes the surface latent heat flux (W m^{-2}); and G denotes the ground heat flux (W m^{-2}). To minimise errors associated with the spatial misalignment between in situ observations and gridded satellite products, we only analyse station-mean anomalies relative to monthly climatologies. If we assume that surface albedo changes are minimal, only components on the right-hand side of (Eq. 3) depend on changes in surface characteristics. Upon subtracting SHF and LW_{up} from surface radiation ($SW_{net} + LW_{down}$), the remainder is assumed to be a combination of LHF and G.

Approximated observed surface fluxes are calculated instantaneously every six hours, whilst simulated fluxes are outputted as three-hourly means (i.e. 0900−1200, 1200−1500 UTC etc.; section 2.1). To enable a suitable comparison between observed and simulated surface fluxes during dry spells, we perform a temporal cubic interpolation of simulated surface fluxes. To do so we assume, for example, that the three-hourly mean between 0600 and 0900 UTC is an approximation of the instantaneous value at 0730 UTC. To then estimate the value at 0600 UTC, we perform a cubic interpolation of three-hourly mean surface fluxes centered at 0430 and 0730 UTC. With regards to referencing the local time of day, we conclude it is inappropriate to use Beijing time (BT) as a reference for local solar conditions on the TP as it covers a large longitudinal range. Instead, we define local solar time (LST) as six hours ahead of UTC as the eastern TP is situated at approximately 90° longitude.

2.3. *Analysis techniques*

2.3.1. *Coupling metrics*

In this study, soil moisture-precipitation feedbacks are defined by the complete pathway with which soil moisture variations lead to precipitation changes through fluctuations in the partitioning of surface turbulent fluxes. Components of soil moisture-precipitation feedbacks are quantified through comparing covariances of evaporative fraction (EF, dimensionless), surface soil moisture (SM, m^3 m^{-3}) and precipitation (P, mm d^{-1}). This follows multiple studies which have quantified components of land−atmosphere feedbacks through comparing covariances between surface and atmospheric fields (i.e. Dirmeyer, 2011; Dirmeyer et al., 2014; Müller et al., 2021b). Our coupling metrics are derived from Dirmeyer et al. (2014) and are the same as in Müller et al. (2021b), except for the use of EF instead of LHF. The use of EF removes the dependence of surface turbulent fluxes on radiation, and instead focuses on the partitioning of turbulent fluxes.

The terrestrial leg of soil moisture-precipitation feedbacks, which identifies areas where anomalous soil moisture drives surface flux variability, is quantified by a terrestrial coupling index (TCI, dimensionless):

$$TCI = \frac{cov(SM, EF)}{\sigma(SM)} \,. \qquad (4)$$

where $cov(x, y)$ and $\sigma(x)$ denote the covariance between two variables and the temporal standard deviation of single variable respectively. Based on this definition, strong soil moisture-surface flux coupling is quantified in regions where soil moisture conditions drive evapotranspiration dynamics. Following this, the atmospheric component of soil moisture-precipitation feedbacks, which highlights regions where surface flux changes lead to a precipitation response, is defined by an atmospheric coupling index (ACI, mm d^{-1}):

$$ACI = \frac{cov(EF, P)}{\sigma(EF)} \,. \qquad (5)$$

ACI complements TCI through highlighting areas where the partitioning of surface turbulent fluxes impacts precipitation. Whilst previous studies have used more intermediate atmospheric variables to compute coupling indices, such as the lifting condensation level (Dirmeyer et al., 2014), we use precipitation as it ensures that the full cycle of land−atmosphere coupling is analysed given the direct feedback between precipitation and surface conditions (Müller et al., 2021b). Even though using precipitation will likely result in a weaker concluded coupling between surface fluxes and atmospheric conditions, our definition of ACI represents the full cycle in surface flux-precipitation coupling, which is not guaranteed when using intermediate atmospheric variables. Precipitation is also one of only suitable atmospheric diagnostics regularly outputted at a sub-daily temporal resolution.

Regions with strong soil moisture-precipitation feedbacks are identified using a two-legged coupling index (TLCI, mm d^{-1}):

$$TLCI = \frac{cov(SM, EF)cov(EF, P)}{\sigma(SM)\sigma^2(EF)} \,. \qquad (6)$$

TLCI quantifies the anomalous precipitation influenced

by moisture–driven surface flux variability and is derived through combining TCI and ACI. Hot spots with substantial soil moisture-precipitation coupling are regions where both TCI and ACI are large. In these regions we expect a feedback from the atmosphere to the land, completing the mechanistic loop (Guo et al., 2006). Partitioning TLCI into its two components highlights locations where the relationship between soil moisture and evaporative fraction is strong (TCI), and where anomalous evaporative fraction influences precipitation (ACI).

For calculating all three indices we use daily time series of anomalies relative to a monthly climatology. To ensure that we sample the impact of surface conditions on daytime precipitation, we compute coupling metrics using three-hourly means of surface SM and EF between 06 and 09 LST. We also only analyse precipitation accumulations between 09 and 18 LST. In principle, one could compare simulated indices with those computed using reanalysis data, however in practice, due to the use of complex land surface models and an inadequate representation of orographic precipitation (Tong et al., 2014; Hu and Yuan, 2021; Müller et al., 2021a), it is unreliable to evaluate simulated metrics with those calculated using reanalysis data. It is also challenging to obtain reliable surface flux observations across the whole of the TP. Therefore, we do not compute coupling metrics using observations.

2.3.2. *The sensitivity of daytime convective precipitation to soil moisture*

To evaluate the simulated feedback between soil moisture and daytime rainfall, we use a metric derived by Taylor et al. (2012). For the rest of this study this metric is referred to as "T12" and denoted by δ_e. T12 quantifies soil moisture-rainfall coupling by assessing anomalous antecedent soil moisture differences between locations with daytime precipitation maxima and minima. The metric was originally developed for global applications and has been used to diagnose soil moisture-rainfall coupling in both observations and models (Taylor et al., 2012, 2013). T12 is computed using:

$$\delta_e = \overline{\Delta S_e} - \overline{\Delta S_c}, \tag{7}$$

where $\overline{\Delta S_e}$ is the composite-mean difference in pre-rainfall soil moisture anomalies between locations of maximum and minimum rainfall, and $\overline{\Delta S_c}$ is a control sample of typical soil moisture anomaly mean differences between those locations. Locations of maximum and minimum rainfall are identified from accumulated convective precipitation between 0900 to 1800 LST. The inclusion of rainfall between 0900 to 1200 LST accounts for the early diurnal onset bias in simulated precipitation (Christopoulos and Schneider, 2021). A 3 × 3 pixel box is centered on an afternoon convective precipitation event with rainfall exceeding 2 mm. The minimum is located within the 3 × 3 pixel box. We decided to use a lower precipitation threshold than Taylor et al. (2012), 2 mm compared to 3 mm, due to low convective precipitation rates in several models. Where two boxes overlap, the box

containing the more intense maxima is retained. If there is more than one minima within a 3 × 3 pixel box, the average soil moisture anomaly is taken. Soil moisture anomalies are sampled at 0600 to 0900 LST. If the total precipitation for sampled pixels exceeds 0.1 mm during this or the preceding time-step (0300 to 0900 LST), the event is excluded. This ensures that only pre-event soil moisture is sampled. Daily soil moisture anomalies are generated with respect to a 35-year (1980–2014) monthly climatology. For each event, soil moisture anomalies for the control sample are taken from the same day of year in non-event years.

δ_e is expressed as a percentile of typical δ values derived from random reassignment of ΔS values. Percentile values less than 10 denote a negative feedback at a 10% significance level or lower, whilst those greater than 90 indicate a positive feedback. Strong negative and positive feedbacks at a 1% signficance level or lower are indicated by percentile values less than 1 and greater than 99 respectively. Whilst a combination of high topographic complexity and poor quality remotely-sensed soil moisture data over the TP means that we cannot directly compare simulated T12 metrics with observations, we can compare the sign of the simulated metric with more detailed analysis from Barton et al. (2021).

3. Simulated soil moisture–atmosphere coupling metrics

To understand the variety of model behaviours in the representation of soil moisture–atmosphere coupling, we use metrics which aim to diagnose the intensity of such feedbacks (section 2.3.1). Figure 1 shows the TCI, ACI and TLCI for each CMIP6 simulation, alongside the ensemble-mean and ensemble-coefficient of variation across all GCMs analysed. The ensemble-coefficient of variation is calculated by dividing the standard deviation amongst GCMs by the ensemble-mean. We also show the ensemble-mean for low- (≈ 250 km) and medium-resolution (⩽ 100 km) models. Focusing on the control of soil moisture on surface turbulent fluxes, the ensemble-mean of the TCI highlights strong coupling in central and western regions of the TP (Fig. 1v). Across the east of the TP, evaporative fraction is relatively insensitive to soil moisture variability. However, whilst most models simulate a strong west-to-east gradient in soil moisture-surface flux coupling (in particular those in the MPI and HadGEM3 model families), associated with surface aridity variations, some models (e.g., BCC-CSM2-MR and MIROC6) simulate minimal surface coupling across the whole of the TP. This model variability is illustrated by large values in the ensemble-coefficient of variation, in particular, across edges of the TP (Fig. 1u1).

As surface conditions are sensitive to precipitation characteristics, we might expect that simulated rainfall differences affect the intensity of soil moisture-surface flux coupling. For example, greater precipitation across the arid surface in the west may moisten soils such that surface moisture has a strong control on the evaporative fraction. To understand

whether simulated precipitation differences affect the representation of soil moisture-surface flux coupling, Fig. 2

shows boreal summer seasonal rainfall biases in each CMIP6 simulation relative to absolute values from the Inte-
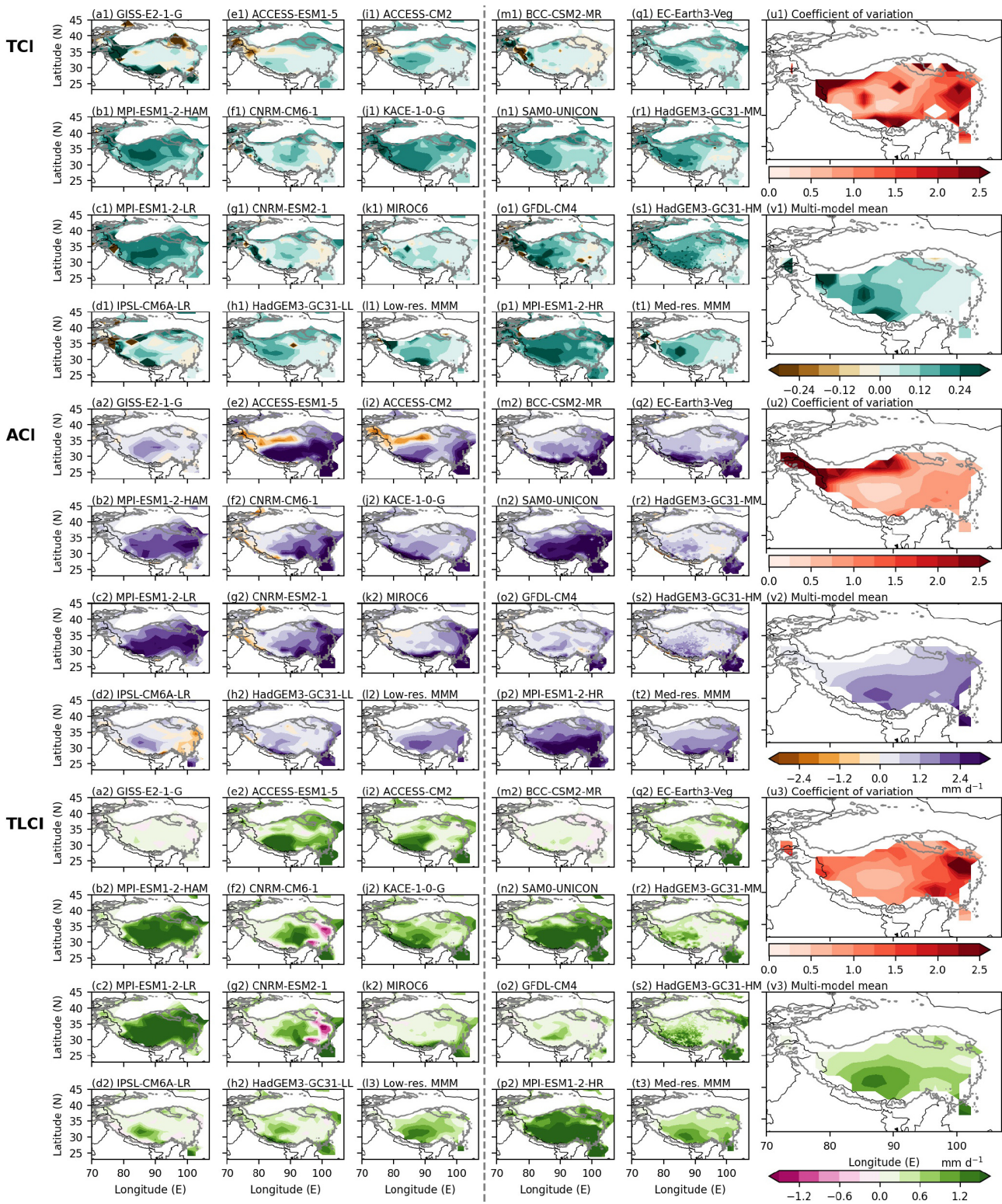

Fig. 1. (a–s, excluding l) TCI (top, suffix 1, dimensionless), ACI (middle, suffix 2, mm d⁻¹) and TLCI (bottom, suffix 3, mm d⁻¹) in each CMIP6 simulation along with the (u) ensemble-coefficient of variation and (v) multi-model mean. We also show the multi-model mean for low- (≈ 250 km) and medium-resolution (≤ 100 km) models in panels (l) and (t) respectively. For all maps we only show grid points with an elevation above 1500 m. Models are ordered based on horizontal resolution with a vertical dashed line between low- and medium-resolution models. The grey outline of the TP denotes an elevation of 3000 m.
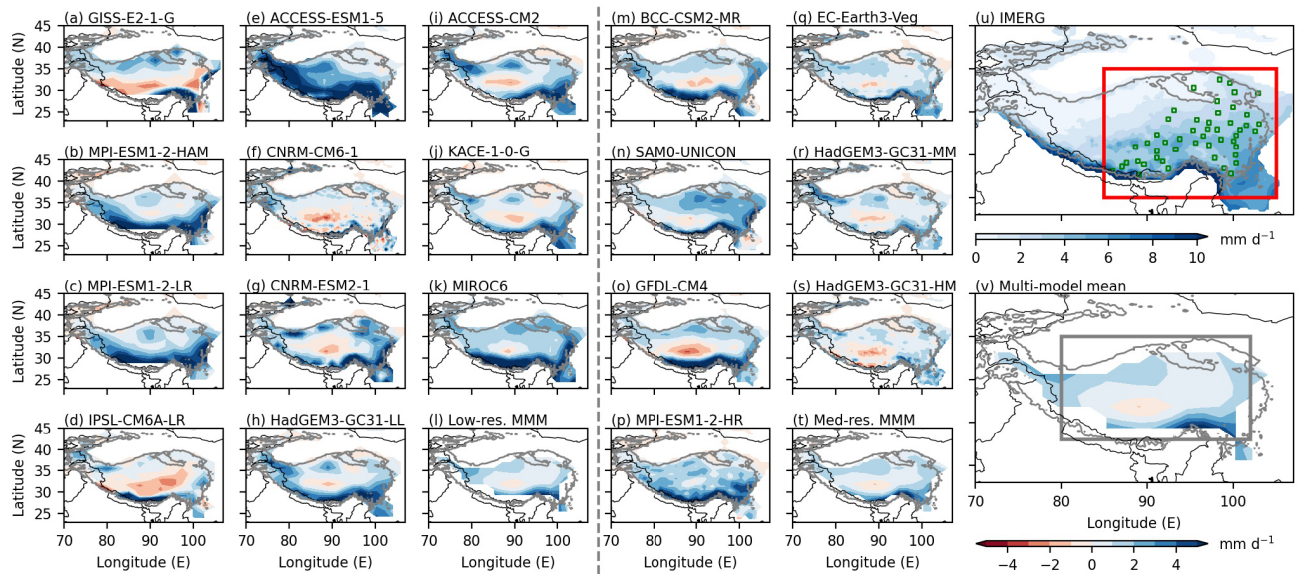
Fig. 2. (a–s, excluding l) Boreal summer-mean rainfall biases (mm d^{-1}) in each CMIP6 simulation relative to (u) IMERG precipitation totals (mm d^{-1}). Multi-model mean rainfall biases in low-resolution, medium-resolution and all models are shown in panels (l), (t) and (v) respectively. For all maps we only show grid points with an elevation above 1500 m. The vertical grey dashed line between panels partitions low- and medium-resolution models. Green squares in panel (u) denote the location of CMA weather stations, whilst the large red square denotes the region used for dry spells in CMIP6 experiments. The large grey box in panel (v) denotes the region used to composite convective events. The grey outline of the TP denotes an elevation of 3000 m.

grated Multi-satellitE Retrievals for the Global Precipitation Measurement mission (GPM IMERG) version 6B (Fig. 2u), which utilises satellite-based passive microwave and geosynchronous infrared measurements (Huffman et al., 2019). We use retrievals from GPM IMERG as previous studies have shown that this is one of the most reliable rainfall products on the TP (Zhang et al., 2018). As data from GPM IMERG begins in 2000, we only analyse simulated precipitation from boreal summer seasons between 2000 and 2014. In general, and consistent with Zhu and Yang (2020), too much rainfall is simulated across the TP (Fig. 2v). Large biases are observed across the southern edge of the plateau, associated with intense orographic uplift, along with small negative biases in central regions. When comparing rainfall biases with simulated values of TCI (Figs. 1a–v1 and 2), we find no distinctive relationship between simulated precipitation and the magnitude of soil moisture-surface flux coupling. For example, models with a strong northwest to southeast gradient in soil moisture-surface flux coupling, such as SAM0-UNICON and KACE-1-0-G, show similar rainfall biases to models with minimal coupling across the whole of TP, i.e. ACCESS-ESM1.5 and CNRM-ESM2-1. Figure A1 in the Appendix a shows the relationship between mean rainfall and simulated TCI on grid points with an elevation greater than 1500 m. Whilst a significant relationship between the two variables is concluded when using a relaxed confidence level of 10%, regional-scale model differences in simulated TCI are not solely driven by precipitation biases. Therefore we conclude that inter-model differences in surface dynamics are partly responsible for inter-model variability in soil moisture-surface flux coupling.

Positive values of ACI across southern and eastern

parts of the TP (Fig. 1v2) indicate that simulated rainfall is partly controlled by surface flux variations and favoured when EF is high. Across the north and west of the plateau, values of ACI are minimal due to little rainfall (Fig. 2u) and small precipitation variability (Wang et al., 2017). Similar to the TCI, we find large intermodel variability in the ACI with some models, i.e. GFDL-CM4 and HadGEM3-GC31-MM, simulating minimal surface flux-precipitation coupling across a large area of the TP. The ensemble-coefficient of ACI highlights large model differences across the north and west of the TP (Fig. 1u2). This is largely influenced by experiments from the ACCESS model family simulating negative ACI values, whilst others simulate positive values. Bringing together the dependence of turbulent fluxes to soil moisture fluctuations (TCI) with the sensitivity of precipitation to evaporative fraction variations (ACI), the ensemble-mean of the TLCI illustrates strong positive soil moisture-precipitation coupling across the south-west and south-east corner of the TP (Fig. 1v3). We find weak soil moisture-precipitation coupling across northern parts of the TP, associated with weak correlations between evaporative fraction and precipitation (i.e., small values of ACI). Whilst most models simulate strong soil moisture-precipitation coupling across the southwest, where precipitation biases (Fig. 2v) and the coefficient of variation are low (Fig. 1u3), across the rest of the TP, simulated differences in ACI and TCI leads to high intermodel variability in TLCI. For example, BCC-CSM2-MR and experiments from the CNRM model family simulate negative values of TLCI across eastern regions of the TP, whilst all other models have positive values. This different behaviour seen in CNRM experiments and BCC-CSM2-MR can be asso-

ciated with relatively strong values of ACI and weak values of TCI. Large inter-model differences in soil moisture-precipitation coupling across most of the TP is consistent with simulated differences in soil moisture-precipitation coupling over the Sahel (Taylor et al., 2013) and variability amongst GCMs in the surface flux response to dry spells (Gallego-Elvira et al., 2019).

4. Assessing components of land–atmosphere feedbacks

Coupling metrics illustrate large inter-model differences in simulated soil moisture-precipitation feedbacks across the TP (section 3.). Whilst these metrics provide a good overview of the coupling strength between soil moisture and precipitation, they can be influenced by atmospheric or rainfall persistence (Guillod et al., 2015). It is also challenging to understand the processes responsible for different coupling characteristics. In light of this, for the rest of this study we gain insight from examining observational metrics specifically designed for the TP. In the following subsection we evaluate the surface flux and atmospheric response to regional-scale dry spells. After this we investigate the sensitivity of convective precipitation to soil moisture heterogeneity (section 4.2).

4.1. *Simulated surface flux and atmospheric response to dry spells*

To assess simulated surface fluxes during three-day dry spells, we first show the behaviour of the surface energy balance in the real world. In this study, we define a regional dry spell as a period of three consecutive days when the regional-mean precipitation rate is below the non-zero 20th percentile boreal summer daily rainfall accumulation. To identify real world regional-scale dry spells we use station-mean daily precipitation accumulations at 1200 UTC from CMA weather stations (Fig. 2u). Due to the time span of weather station data (section 2.2), we only composite real world dry spells between 2000 and 2014. Figure 3a shows surface flux variations across the TP during observed three-day dry spells. Day 0.0 is defined as the start of a three-day regional dry spell, whilst anomalies are only shown at 1200 LST as this is the time of day with the largest surface flux response. Unsurprisingly, a dry spell across the TP increases downward radiation into the surface due to reduced cloud cover. Surface drying during a dry spell changes the partitioning of this enhanced radiation with LHF decreasing by approximately 60 W m^{-2} and SHF increasing by approximately 40 W m^{-2} between days 0 to 2. We also observe increased LW$_{up}$ by approximately 30 W m^{-2} associated with increased surface temperatures.

We next compare this observed surface flux behaviour to that exhibited in CMIP6 experiments. To identify simulated dry spells, we use regional-mean daily precipitation accumulations on grid points above 3000 m between 25° to 40° latitude and 85° to 105° longitude, region denoted by a red rectangle in Fig. 2u. Whilst for observations we were only able to composite dry spells between 2000 and 2014, for CMIP6 experiments we use data from 1980 to 2014 to increase the number of composited dry spells. Even though each GCM simulates a reasonable number of three-day dry spells when using its own 20th percentile precipitation rate, all of the simulated precipitation thresholds are larger compared to observations, which is unsurprising given positive precipitation biases (Fig. 2). When selecting simulated dry spells using the observed precipitation threshold, only eleven out the eighteen GCMs have a substantial number of dry spells ($\geqslant$ 20). Due to the smaller number of individual models with a substantial number of dry spells when using the observed threshold, we focus on the anomalous surface energy balance during dry spells that are defined using simulated 20th percentile precipitation rates. Given that simulated dry-spell precipitation rates are greater than observations, we expect simulated surface flux variations to be dampened.

Figures 3b to 3e highlight the variety of model behaviours in the CMIP6 ensemble by focusing on BCC-CSM2-MR, HadGEM3-GC31-HM, MIROC6 and MPI-ESM1-2-HR. For CMIP6 experiments we are able to composite simulated latent heat fluxes, whilst for observations, we approximate the sum of latent and ground heat fluxes. Whilst all four chosen models simulate increased downward surface radiation, associated with clear skies, they all have a different surface energy balance response. All four chosen models simulate increased SHF and LW$_{up}$, however, changes in these surface energy balance components are typically underestimated. Only BCC-CSM2-MR well represents changes in SHF, whilst anomalies in HadGEM3-GC31-HM, for example, are 50% smaller compared to observations. We also find that three of our chosen models exhibit small latent heat flux changes during a dry spell compared to observations, indicating that the surface dries more rapidly in the real world. Whilst we do see a latent heat flux decrease in MPI-ESM1-2-HR of a similar magnitude to observations, the reduction in evapotranspiration takes several days longer.

Following on from focusing on four chosen models, Figs. 4a to 4c show the average anomalous surface SHF, LW$_{up}$, and radiation reaching the surface during observed and simulated three-day dry spells. Consistent with the subset of models analysed in Fig. 3., the majority of models underestimate increases in SHF and LW$_{up}$ (Figs. 4a and 4b). For example, by day 2 of a three-day dry spell, the anomalous model-mean bias in SHF is approximately 10 W m^{-2} smaller than observations, whilst anomalous LW$_{up}$ is underestimated by approximately 20 W m^{-2}. CMIP6 simulations better represent the amplitude of surface radiation anomalies across the TP during dry spells (Fig. 4c). This indicates that errors in cloud representation are less of a concern compared to surface flux dynamics. Given that it may be the case that changes in surface SHF and LW$_{up}$ are poorly simulated due to underestimated anomalous radiation, we compute the fraction of radiation inputted into the surface that is re-emitted as SHF or LW$_{up}$. In this study, this term is referred to as the "fraction of downwelling radiation":
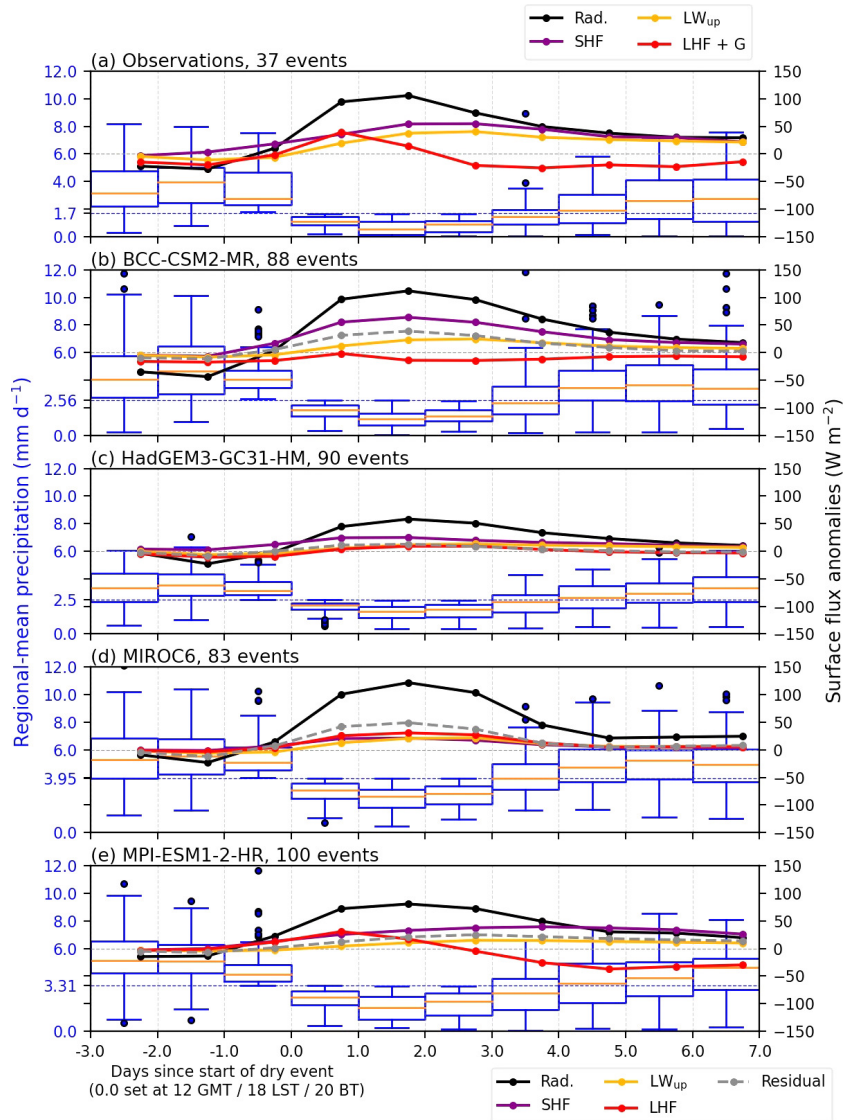
Fig. 3. Anomalous surface fluxes (W m^{-2}) and daily precipitation accumulations (mm d^{-1}) preceding, during and after three-day regional dry spells in (a) observations, (b) BCC-CSM2-MR, (c) HadGEM3-GC31-HM, (d) MIROC6 and (e) MPI-ESM1-2-HR. A three-day regional dry spell is defined when the precipitation accumulation is smaller than the twentieth percentile of boreal summer daily precipitation, denoted by the blue dashed horizontal line, for three consecutive days. We show the following components of the surface energy balance: (orange) upward longwave radiation; (purple) sensible heat flux, (black) and sum of net-downward shortwave and longwave downward radiation. In panel (a) the red line denotes the sum of latent and ground heat fluxes, whilst for panels (b) to (e) it denotes the latent heat flux only. Panels (b) to (e) also include a (dashed grey) "residual" term which is the remainder when subtracting sensible and latent heat fluxes from net-downward radiation. Box-and-whisker plots show station-mean or regional-mean daily precipitation accumulations. The orange line within each box denotes the median; the top and bottom of the box denotes the upper and lower quartiles; and the blue whiskers denote the 10th and 90th percentiles. Filled blue circles denote outliers in precipitation rates.

$$\text{Fraction of downwelling radiation} = \frac{\text{SHF} + \text{LW}_{\text{up}}}{\text{SW}_{\text{net}} + \text{LW}_{\text{down}}} \,. \quad (8)$$

Figure 4d shows the change in the anomalous fraction of downwelling radiation during a dry spell in both observa-

tions and CMIP6 simulations. The increased fraction of downwelling radiation during a dry spell in observations illustrates that in the real world changes in surface characteristics lead to anomalous sensible heating and surface longwave emission. However, all CMIP6 simulations underestimate the
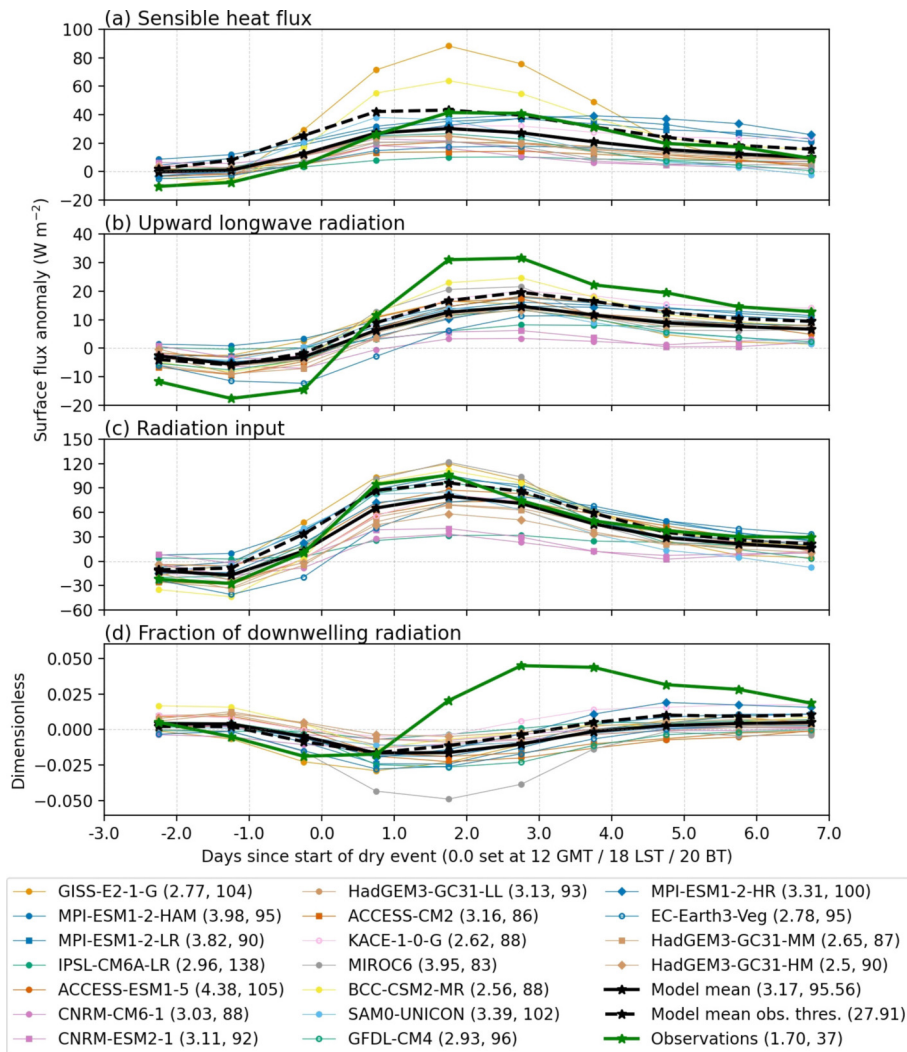
Fig. 4. Anomalous surface (a) sensible heat flux, (b) upward longwave radiation, (c) radiation inputted into the surface, and (d) fraction of downwelling radiation that is re-emitted as sensible heat and upward longwave radiation, preceding, during and after a three-day dry event. All models from the same model family are denoted by the same line colour with individual configurations distinguished by marker style. Observations are denoted by a green line whilst the model mean is denoted by a black line. The model mean for simulated dry spells with the observed precipitation threshold is denoted by a dashed black line. The values to the right of each model name in the legend include the regional-mean 20th percentile boreal summer rainfall rate and the number of three-day dry spells identified.

change in partitioning of incoming radiation towards SHF and LW_{up}. This difference between observations and CMIP6 simulations highlights that surface dynamics are poorly represented on the TP during a dry spell, and that errors in the surface energy balance are not solely due to biases in atmospheric radiation.

In comparison with observations, all CMIP6 simulations poorly represent the favouring of SHF over evapotranspiration during a dry spell. However, it may be the case that this weak surface response during simulated dry spells is associated with high dry-spell precipitation rates. For example, Fig. 3 shows larger anomalies in SHF and LW_{up} in models with smaller dry-spell precipitation rates (BCC-CSM2-MR and HadGEM3-GC31-HM). To investigate the hypothesis

that anomalous surface fluxes are small in CMIP6 experiments due to large dry-spell precipitation rates, Fig. 5a compares the average change in SHF and LW_{up} during dry spells with the prescribed precipitation threshold. We also calculate the linear least-squares regression between simulated values and take note of the Pearson correlation coefficient and p-value for a single-sided t-test assuming a negative relationship. Whilst one would expect excessive dry spell rainfall to suppress the surface flux response, the slope of the linear regression is not significantly negative. This provides evidence that an improved representation of anomalous surface fluxes during a dry spell requires more than just a better representation of dry-spell precipitation intensities. To analyse our hypothesis further, Fig. A2 shows composited surface
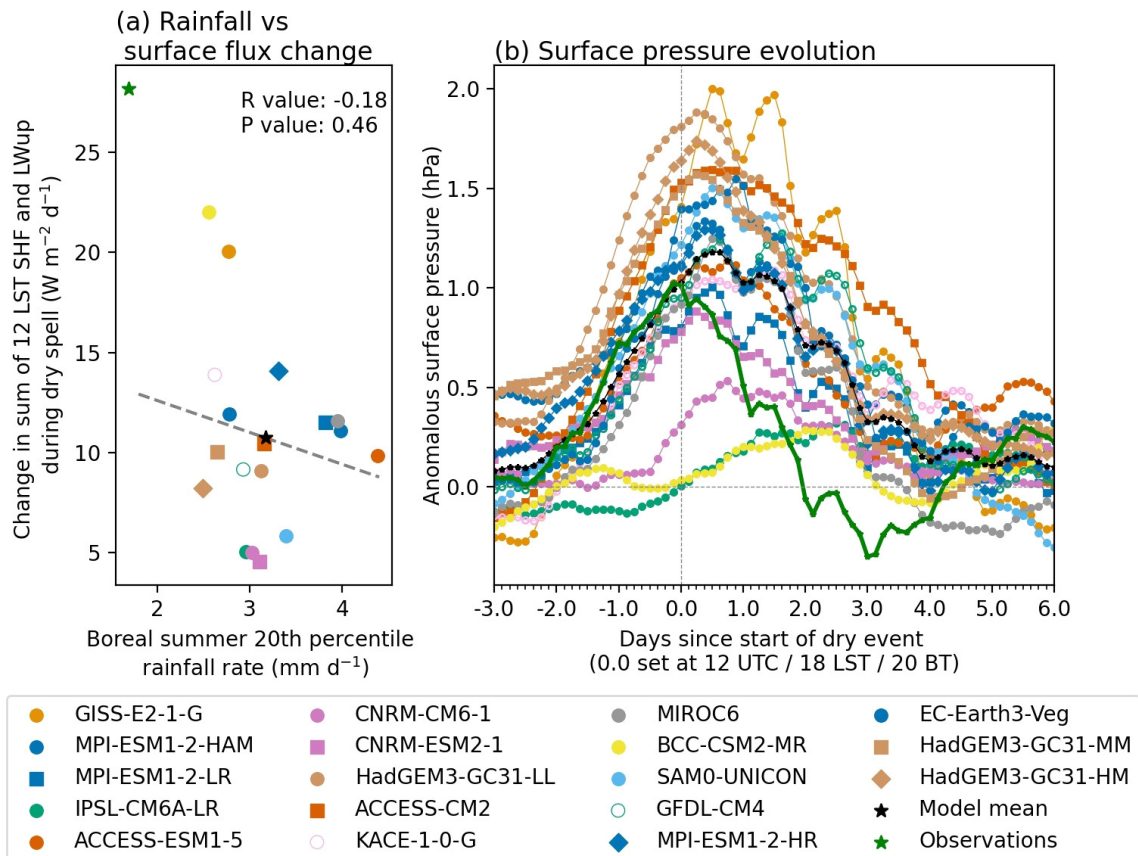
Fig. 5. (a) Comparison of boreal summer 20th percentile precipitation rate (mm d^{-1}) against the change in surface sensible heat flux and upward longwave radiation (W m^{-2}) at 12 LST between days -0.25 and 2.75 of a regional dry spell. The grey dashed line in panel (a) denotes the linear least-squared fit between simulated values. The line's Pearson correlation coefficient value (R) and p-value for a single-sided t-test assuming a negative relationship (P) is shown in the top right hand corner. (b) Anomalous surface pressure (hPa) between $25°$ to $40°$ latitude and $85°$ to $105°$ longitude preceding, during and after a three-day dry spell. Dashed grey horizontal and vertical lines in panel (b) denote the zeroth value. All models from the same model family are denoted by the same line colour with individual configurations distinguished by marker style. Observations are denoted by green circular markers, whilst the model mean is denoted by black stars.

flux anomalies during simulated dry spells using the observed precipitation threshold. As discussed previously, only a selection of models simulate a substantial number of dry spells when using the observed precipitation threshold. The model-mean surface flux response during these dry spells is also denoted in Fig. 4 by black dashed lines. Whilst Figs. 4 and A2 illustrate that using the observed precipitation threshold leads to a better simulation of anomalous SHF and LW$_{up}$, models still do not adequately capture the strong change in surface flux partitioning. This indicates that land surface schemes in GCMs are unable to represent soil moisture–driven short-term ($\approx$ few days) variability in evapotranspiration on the TP.

The inter-model variability in the behaviour of surface fluxes during dry spells is consistent with differences in TCI values (Figs. 1 and 4). Models which simulate unrealistic large dry-spell precipitation rates, such as MIROC6 and ACCESS-ESM1-5 (Fig. 5a), simulate relatively weak TCI values. In these simulations it is likely that the dry-spell precipitation rate is greater than potential evapotranspiration. This

may lead to an unrealistic representation of the land surface as it rarely dries out and uses all additional radiative energy to increase evapotranspiration. We also find that models with a relatively large change in the fraction of downwelling radiation during dry spells, such as MPI-ESM2-2-HR (Fig. 4d), are typically those with high TCI values. Whilst our analysis of surface fluxes during three-day dry spells cannot fully explain simulated TCI differences due to influence of variability on longer timescales, we find a good agreement between the magnitude of TCI values simulated and the response of surface fluxes during a dry spell.

In the real world surface drying on the TP favours sensible heat, a deeper planetary boundary layer, and a negative near-surface pressure tendency (Wan et al., 2017; Talib et al., 2021). Due to the lack of a surface energy balance response to regional dry spells in the majority of GCMs (Fig. 4), we predict that models underestimate the surface pressure response to surface drying. To illustrate the favouring of a heat low circulation across the TP during dry spells in observations, Fig. 5b shows regional-mean anomalous sur-

face pressure from the European Centre for Medium-Range Weather Forecasts (ECWMF) Reanalysis version 5 (ERA5; Hersbach et al., 2020) at a 1° resolution between 25° to 40° latitude and 85° to 105° longitude. The negative pressure tendency associated with surface drying maximises after sunset which is indicative of a heat low circulation as negative pressure tendencies are limited until a stable boundary layer has formed. Fig. 5b also shows the anomalous surface pressure during simulated dry spells in each GCM. Whilst most models correctly simulate diurnal fluctuations in anomalous surface pressure, the magnitude of pressure tendencies during a dry spell are smaller compared to observations. For example, the model-mean surface pressure anomaly decreases by approximately 0.7 hPa from days 0.0 to 3.0, compared to 1.3 hPa in observations. We also observe distinct pressure anomalies in BCC-CSM2-MR and IPSL-CM6A-MR, which we associate with synoptic forcing dominating any effects from surface heating. The simulation of weaker pressure tendencies in CMIP6 experiments is consistent with underestimated changes in surface energy balance components. The weak surface pressure response is likely to limit the impact of soil moisture-atmospheric coupling on large-scale circulation anomalies (Wan et al., 2017; Talib et al., 2021).

4.2. *Simulated feedback between soil moisture and convective initiation*

In the previous subsection we show that the dampened behaviour of surface fluxes during dry spells in CMIP6 simulations leads to a misrepresentation of surface-induced atmospheric pressure fluctuations. In this subsection we investigate whether CMIP6 models correctly favour deep convection

initiation over dry soils, as observed by Barton et al. (2021). To do this, we first analyse simulated pre-rainfall soil moisture anomalies before a strong convective precipitation event (section 2.3.2). We pool all events that occur within 80° to 102° longitude and 28° to 40° latitude, as denoted by the grey box in Fig. 2v, where elevation exceeds 2500 m. The T12 metric for 12 out of the 18 GCMs is shown in Fig. 6, with the order of GCMs determined by horizontal resolution (increasing from left to right). For the remaining six CMIP6 simulations, an insufficient number of convective precipitation events (< 100) were identified due to either persistent early-morning rainfall (GISS-E2-1-G, IPSL-CM6A-LR, CNRM-CM6-1 and CNRM-ESM2-1) or minimal precipitation rates (GFDL-CM4 and BCC-CSM2-MR). These models are also those with relatively low values of deep convective precipitation (Fig. A3). Almost all remaining GCMs fail to capture the observed negative feedback between soil moisture and deep convection found in Barton et al. (2021). Only ACCESS-CM2 simulates a significant strong negative feedback, whilst six GCMs show a significant strong positive feedback.

Given that the majority of rainfall in CMIP6 experiments with a substantial number of events is associated with deep convection (Fig. A3) and most experiments simulate positive values of TCI and ACI (Fig. 1), it is unsurprising that a positive soil moisture-deep convection feedback is seen in most models (Fig. 6). The lack of a simulated negative feedback is similar in other semi-arid regions (Taylor et al., 2012) and can be explained by a typically strong dependence of convective parameterisations on low-level humidity, which is favoured across wet soils. In reality, convective initiation
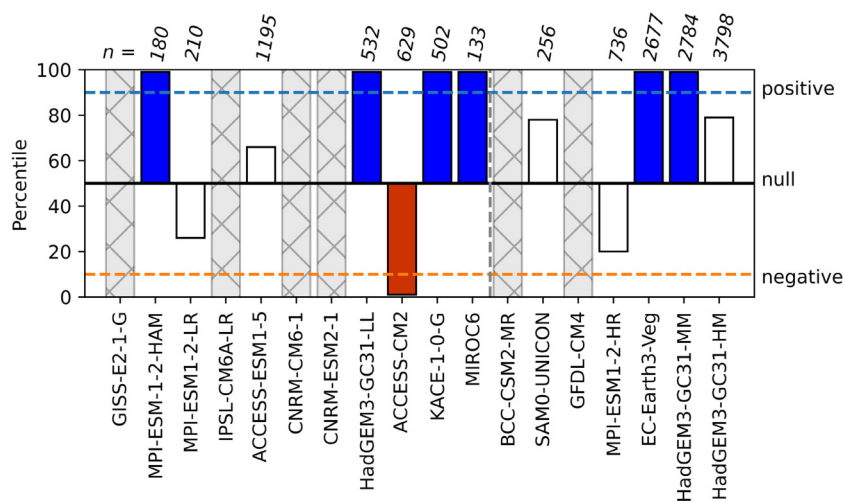

Fig. 6. T12 metric (δ_e, percentile) for events within 80° to 102° longitude and 28° to 40° latitude for CMIP6 simulations. For clarity, bars are plotted as a deviation from 50 where values larger and smaller than 50 denote positive and negative feedbacks respectively. Blue and red filled bars denote a preference for afternoon convection over wet and dry soils respectively with a significance level below 10%. Grey hatched bars denote models with fewer than 100 events. Blue and orange horizontal dashed lines denote the 10% significance level for positive and negative feedbacks respectively, whilst the grey dashed vertical line partitions low- and medium-resolution models. Italicised values above each bar denote the number of afternoon convective events used to calculate the metric.

occurs later in the day than in climate models (Christopoulos and Schneider, 2021), which favours stronger daytime mesoscale circulations and more intense heavy precipitation. We might expect that increasing the horizontal resolution of a climate model improves the model's ability at simulating the formation of realistic circulations in response to soil moisture heterogeneity. However, if the model's convection scheme is typically triggered before these circulations can develop, then a positive feedback may persist, irrespective of resolution, as illustrated in Taylor et al. (2013).

Considering all twelve GCMs for which we were able to compute the T12 metric, there is no obvious improvement with resolution (Fig. 6, left to right). However, if we compare different resolutions within the same model family, i.e. MPI-ESM1-2-HR with MPI-ESM1-2-LR and HadGEM3-GC31-HM with HadGEM3-GC31-MM, increased resolution decreases the value of the T12 metric and weakens the positive feedback. To examine this behaviour in more detail, we focus on HadGEM3-GC31 simulations for which we have a sufficient number of events at all three resolutions to subdivide the TP into four $11°$ longitude $\times 6°$ latitude quadrants (regions shown in Fig. 7). For the low (HadGEM3-GC31-

LL) and medium (HadGEM3-GC31-MM) resolutions, a significant positive feedback is simulated in all four quadrants (Figs. 7a and 7b). HadGEM3-GC31-HM on the other hand, simulates varying behaviour across the TP with a negative/positive feedback in the south-east/north-east quadrant (Fig. 7c). Comparing simulated feedbacks with resolved topography (Figs. 7d to 7f) gives some indication that increased topographic complexity influences the sign of the feedback. To investigate whether topographic complexity in the model influences the feedback between soil moisture and deep convection in more detail, we partitioned all events in HadGEM3-GC31-HM into two groups based on the grid-scale topographic complexity. The grid-scale topographic complexity is calculated as the standard deviation in altitude across a 3×3 pixel which is centered on a rainfall event. As shown in Table 2, events centered where topographic complexity is low have a weak negative feedback, whilst where topographic complexity is high, a strong positive feedback is concluded. This indicates that when increasing resolution, and hence improving the representation of orographic convection, we begin to favour negative soil moisture-convection feedbacks across regions with low topo-
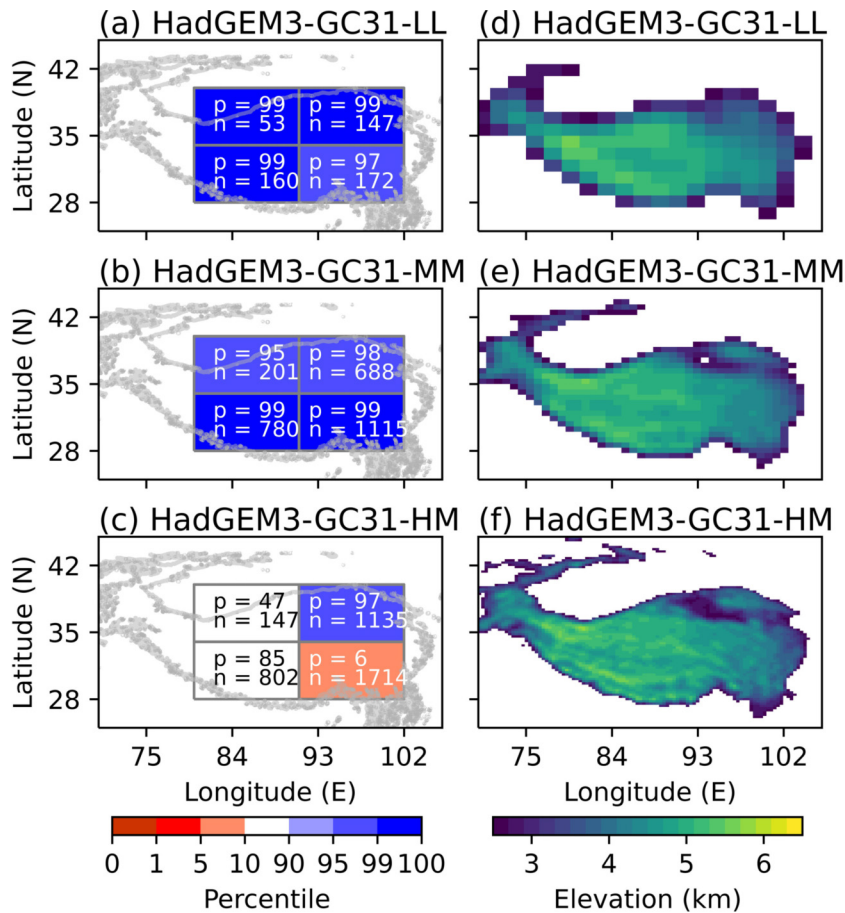

Fig. 7. (a–c) T12 metric (δ_e, percentile) over $11°$ longitude $\times 6°$ latitude quadrants and (d–f) resolved topography (km) for (a, d) HadGEM3-GC31-LL, (b, e) HadGEM3-GC31-MM and (c, f) HadGEM3-GC31-HM. In panels (a) to (c) blue and red shading denotes a preference for afternoon convection over wet and dry soils respectively. Coloured shading is only applied in quadrants with a significance level below 10%. The grey outline of the TP denotes an elevation of 2500 m.

Table 2. Evaluation of topographic dependence on soil moisture-deep convection coupling over TP in HadGEM3-GC31-HM. Regions with low and high topographic complexity have a standard deviation in altitude below or above 100 m, respectively, over a 3 × 3 pixel box. Blue and red shading denotes a preference for afternoon convection over wet and dry soils respectively.

Topographic complexity	Number of events	δ_e
Low	1842	10
High	1956	98

graphic complexity. It is known that convection-permitting resolutions are needed to fully capture soil moisture-convection feedbacks (Hohenegger et al., 2009; Taylor et al., 2013), but these configurations are currently too expensive to run globally across climate relevant time scales. The fact that HadGEM3-GC31-HM, a current medium-resolution global climate model, can begin to overcome a significant feedback bias on the TP is promising for future generations of ESMs.

5. Discussion and conclusions

In this study, we use three analysis techniques to assess the representation of soil moisture–atmosphere coupling on the TP. These techniques include: daily coupling metrics which partition the terrestrial and atmospheric components of soil moisture–atmosphere feedbacks (section 3.; Dirmeyer, 2011; Dirmeyer et al., 2014); analysis of the surface flux and atmospheric behaviour during three-day dry spells (section 4.1; Talib et al., 2021); and an index which quantifies the favouring of convective precipitation over dry soils (section 4.2; Taylor et al., 2012, 2013). Whilst previous studies have used similar techniques to perform global model assessments (Dirmeyer, 2011; Dirmeyer et al., 2014; Gallego-Elvira et al., 2019), this is the first study to focus on evaluating such feedbacks across the TP.

We find substantial inter-model differences in simulated soil moisture–atmosphere feedbacks across the TP, consistent with studies focusing on other semi-arid regions (Taylor et al., 2013; Gallego-Elvira et al., 2019). Partitioning feedbacks into their terrestrial and atmospheric segments highlights substantial model variability in both feedback components. GCMs typically underestimate the feedback of surface flux dynamics on atmospheric conditions during three-day dry spells. We note that whilst GCMs overestimate the limitation of evapotranspiration by soil moisture deficiency over relatively long periods ($\geqslant$ 10 days) (Ukkola et al., 2016; Dirmeyer et al., 2018a; Gallego-Elvira et al., 2019), they can still underestimate soil moisture–driven surface flux variations on shorter timescales ($\leqslant$ 3 days). Such behaviour suggests that simulated evapotranspiration fluctuations are too restrained by root-zone soil moisture and insufficiently sensitive to rapid variations in near-surface moisture. The high dependence of evapotranspiration on near-surface soil moisture has been highlighted by surface flux observations across semi-arid environments on the TP (Cui et al., 2020). Given that observations show that anomalous near-surface

soil moisture on the TP is source of atmospheric predictability (Talib et al., 2021; Barton et al., 2021), it is likely that inhibited soil moisture–atmosphere coupling in CMIP6 models is also present in dynamical forecast models, which reduces daily to seasonal predictive skill.

We conclude that to improve the representation of soil moisture–atmosphere feedbacks on the TP requires a better representation of both rainfall and surface dynamics. Focusing on precipitation, the positive rainfall bias in the majority of CMIP6 simulations is likely to change the representation of soil moisture-surface flux coupling. In semi-arid environments for example, greater precipitation may dampen soils such that the partitioning of surface turbulent fluxes is predominately controlled by radiation instead of near-surface soil moisture. However, whilst we hypothesise that correcting the simulation of rainfall ought to improve the simulation of soil moisture–atmosphere feedbacks, weak correlations between rainfall biases and coupling strengths suggest that feedback errors are not solely due to precipitation errors. Alongside improving the representation of precipitation, a more realistic simulation of surface dynamics is required. This is consistent with Gevaert et al. (2018), who found large model variability in soil moisture-surface flux coupling when driving several surface models with the same atmospheric data.

Alongside highlighting that an improved representation of soil moisture-surface flux coupling requires a better simulation of both surface dynamics and rainfall variability, our work illustrates that deep convection on the TP occurs too frequently over wet soils for the majority of CMIP6 models. Our analysis is consistent with previous studies which argue that deep convection parameterisation schemes are too sensitive to low-level humidity and therefore favour positive soil moisture-deep convection feedbacks (Hohenegger et al., 2009; Taylor et al., 2013). In the real world, the favouring of deep convection over dry soils occurs due to influence of soil moisture gradients on the formation of mesoscale circulations (Taylor et al., 2011; Barton et al., 2021). However in coarse-resolution ESMs, the influence of soil moisture on the development of mesoscale circulations is limited due to the inability to resolve mesoscale circulations and the early onset of daytime rainfall (Christopoulos and Schneider, 2021). Only ACCESS-CM2 correctly simulates a negative soil moisture-deep convection feedback, which we associate with large convective rainfall totals and negative values of ACI across the north-west of the TP. We also find that increasing horizontal resolution improves simulated soil moisture–atmosphere coupling for the HadGEM3 family. We hypothesise that increasing horizontal resolution improves the representation of mesoscale flows driven by orography or soil moisture variability, and their impact on convective initiation. For example, the highest resolution of HadGEM3-GC31 in CMIP6 simulates a negative soil moisture-deep convection feedback across regions with low topographic complexity. These results suggest that future modelling developments will improve the simulation of soil moisture-deep convection feedbacks.

Alongside efforts in model development, improved observations of the soil moisture and surface flux response to precipitation variability will support our understanding of land–atmosphere interactions. As well as this, an improved quantification of spatial variability in surface characteristics on the TP will develop our ability at parameterising surface processes in course-resolution GCMs. Not only can our knowledge of real world soil moisture–atmosphere feedbacks be developed by increasing the number of stations that monitor surface fluxes and weather conditions on the TP, but developing reliable analysis techniques of satellite retrievals across a topographically-complex region will also support our understanding. Given the influence of TP surface characteristics on atmospheric conditions across East Asia and beyond (Wan et al., 2017; Talib et al., 2021), an improved understanding and representation of surface–atmosphere feedbacks will improve atmospheric predictability beyond the plateau itself. In addition to improving daily to seasonal atmospheric predictability, a greater understanding of surface–atmosphere feedbacks on the TP will improve our understanding of the mechanisms responsible for amplified climate change-induced warming across the TP, lead to a better attribution of anthropogenic climate change on observed environmental changes, and reduce model uncertainties in future predictions of hydrological and atmospheric changes (You et al., 2020).
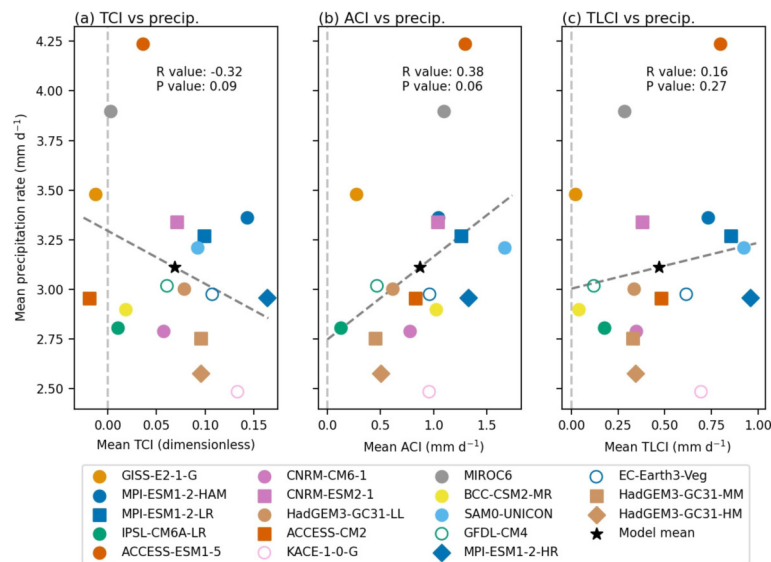
Fig. A1. Comparison of average boreal summer precipitation across the TP with simulated values of (a) TCI (dimensionless), (b) ACI (mm d^{-1}), and (c) TLCI (mm d^{-1}). The grey dashed line in each panel denotes the linear least-squared fit between simulated values. The line's Pearson correlation coefficient value (R) and p-value for a single-sided t-test assuming a negative relationship (P) is shown in the top right hand corner. Light grey dashed vertical lines denote the zeroth value. All models from the same model family are denoted by the same colour with individual configurations distinguished by marker style. The model mean is denoted by a black star.
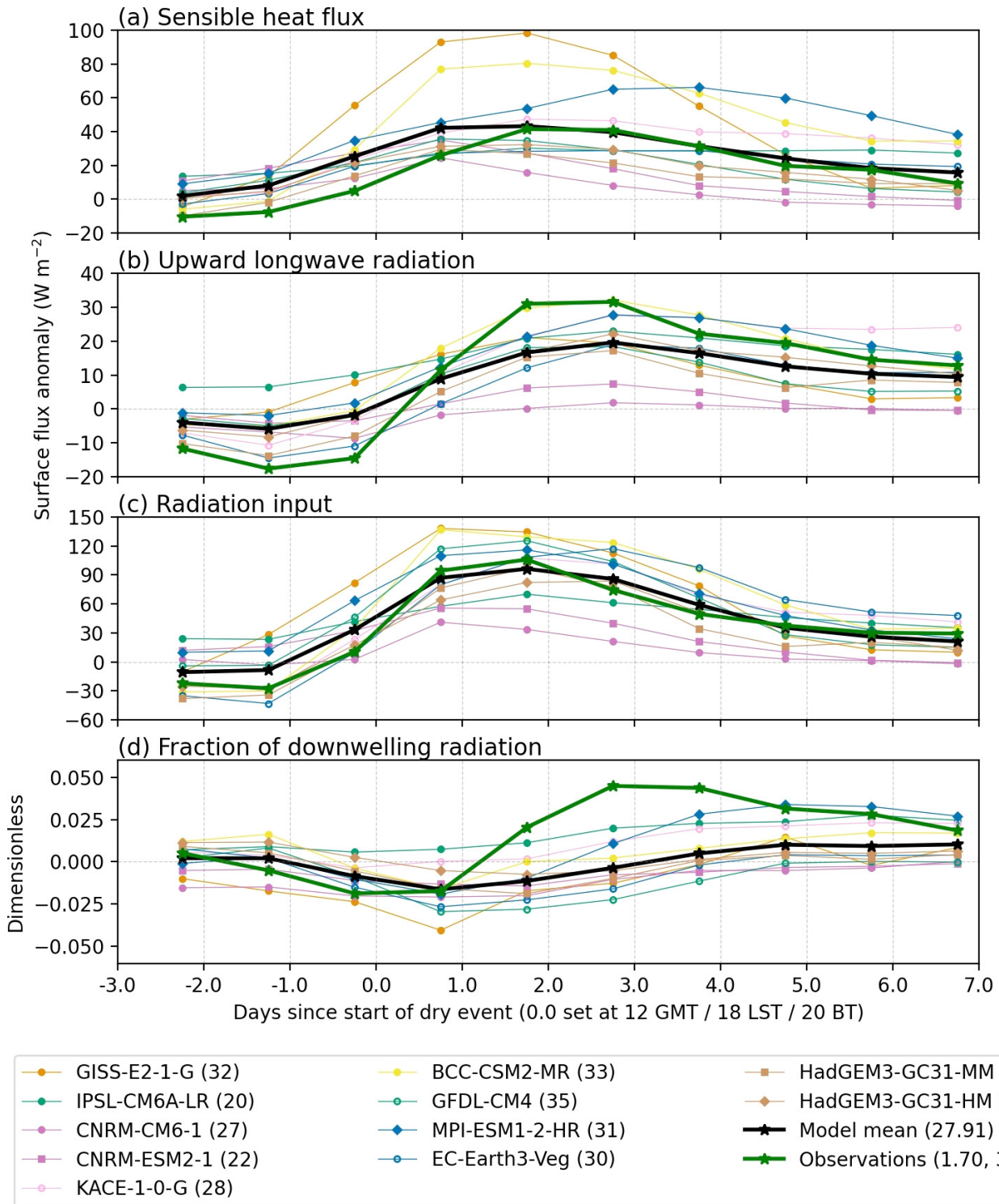
Fig. A2. Anomalous surface (a) sensible heat flux, (b) upward longwave radiation, (c) radiation inputted into the surface, and (d) fraction of downwelling radiation that is re-emitted as sensible heat and upward longwave radiation, preceding, during and after a three-day dry event. For this figure all dry events are defined using the observed dry-spell precipitation rate (1.70 mm d^{-1}). All models from the same model family are denoted by the same line colour with individual configurations distinguished by marker style. Observations and the model-mean are denoted by green and black lines respectively. The values to the right of each model name in the legend are the number of three-day dry spells identified. We also note the observed dry spell precipitation threshold next to the observations label.
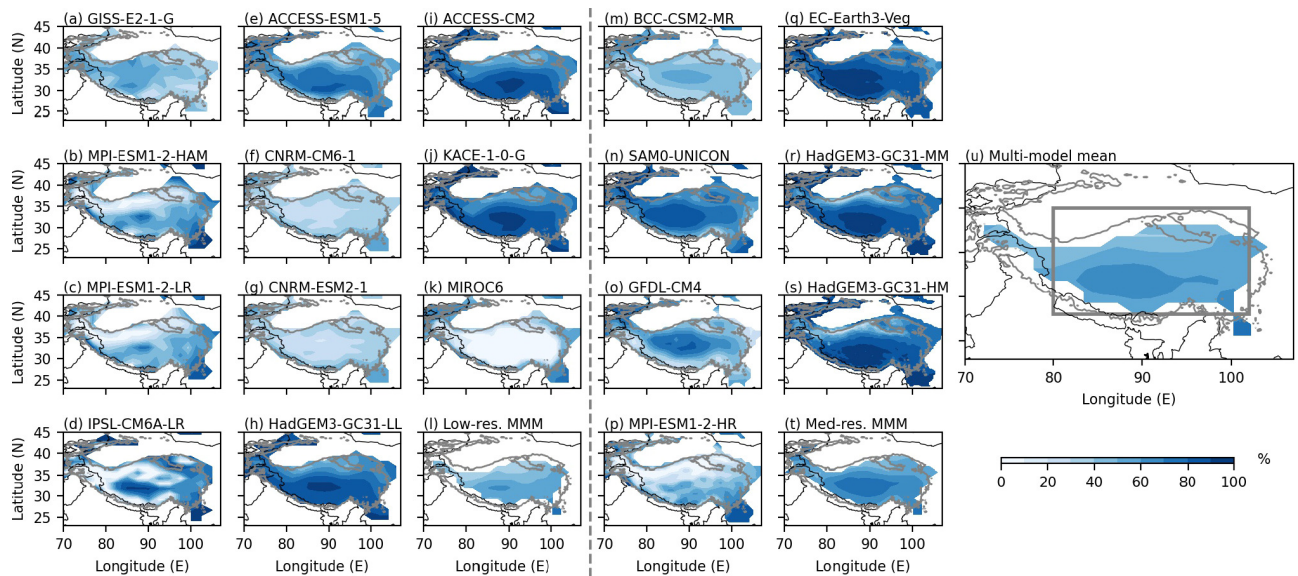
Fig. A3. (a–s, excluding l) Percentage of daytime (9–18 LST) boreal summer precipitation attributable to the deep convective parameterisation scheme in each CMIP6 simulation. Multi-model averages in low-resolution, medium-resolution and all models are shown in panels (l), (t) and (u) respectively. For all maps we only show grid points with an elevation above 1500 m. The vertical grey dashed line between panels partitions low- and medium-resolution models.

REFERENCES

Barton, E. J., C. M. Taylor, C. Klein, P. P. Harris, and X. Meng, 2021: Observed soil moisture impact on strong convection over mountainous Tibetan Plateau. *Journal of Hydrometeorology*, **22**, 561–572, https://doi.org/10.1175/JHM-D-20-0129.1.

Chow, K. C., J. C. L. Chan, X. L. Shi, Y. M. Liu, and Y. H. Ding, 2008: Time-lagged effects of spring Tibetan Plateau soil moisture on the monsoon over China in early summer. *International Journal of Climatology*, **28**, 55–67, https://doi.org/10.1002/joc.1511.

Christopoulos, C., and T. Schneider, 2021: Assessing biases and climate implications of the diurnal precipitation cycle in climate models. *Geophys. Res. Lett.*, **48**, e2021GL093017, https://doi.org/10.1029/2021GL093017.

Cui, J. P., L. D. Tian, Z. W. Wei, C. Huntingford, P. Wang, Z. Y. Cai, N. Ma, and L. X. Wang, 2020: Quantifying the controls on evapotranspiration partitioning in the highest alpine meadow ecosystem. *Water Resour. Res.*, **56**, e2019WR024815, https://doi.org/10.1029/2019WR024815.

Dirmeyer, P. A., 2011: The terrestrial segment of soil moisture–climate coupling. *Geophys. Res. Lett.*, **38**, L16702, https://doi.org/10.1029/2011GL048268.

Dirmeyer, P. A., and Coauthors, 2018a: Verification of land–atmosphere coupling in forecast models, reanalyses, and land surface models using flux site observations. *Journal of Hydrometeorology*, **19**, 375–392, https://doi.org/10.1175/JHM-D-17-0152.1.

Dirmeyer, P. A., R. D. Koster, and Z. C. Guo, 2006: Do global models properly represent the feedback between land and atmosphere. *Journal of Hydrometeorology*, **7**, 1177–1198, https://doi.org/10.1175/JHM532.1.

Dirmeyer, P. A., Z. Y. Wang, M. J. Mbuh, and H. E. Norton, 2014: Intensified land surface control on boundary layer growth in a changing climate. *Geophys. Res. Lett.*, **41**, 1290–1294, https://doi.org/10.1002/2013GL058826.

Dirmeyer, P. A., S. Halder, and R. Bombardi, 2018b: On the harvest of predictability from land states in a global forecast model. *J. Geophys. Res.*, **123**, 13 111–13 127, https://doi.org/10.1029/2018JD029103.

Duan, A. M., and G. X. Wu, 2005: Role of the Tibetan Plateau thermal forcing in the summer climate patterns over subtropical Asia. *Climate Dyn.*, **24**, 793–807, https://doi.org/10.1007/s00382-004-0488-8.

Duan, A. M., and G. X. Wu, 2008: Weakening trend in the atmospheric heat source over the Tibetan Plateau during recent decades. Part I: Observations. *J. Climate*, **21**, 3149–3164, https://doi.org/10.1175/2007JCLI1912.1.

Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016: Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016.

Fan, K. K., Q. Zhang, V. P. Singh, P. Sun, C. Q. Song, X. D. Zhu, H. Q. Yu, and Z. X. Shen, 2019: Spatiotemporal impact of soil moisture on air temperature across the Tibet Plateau. *Science of the Total Environment*, **649**, 1338–1348, https://doi.org/10.1016/j.scitotenv.2018.08.399.

Gallego-Elvira, B., C. M. Taylor, P. P. Harris, and D. Ghent, 2019: Evaluation of regional-scale soil moisture-surface flux dynamics in Earth system models based on satellite observations of land surface temperature. *Geophys. Res. Lett.*, **46**, 5480–5488, https://doi.org/10.1029/2019GL082962.

Gevaert, A. I., D. G. Miralles, R. A. M. de Jeu, J. Schellekens, and A. J. Dolman, 2018: Soil moisture-temperature coupling in a set of land surface models. *J. Geophys. Res.*, **123**, 1481–1498, https://doi.org/10.1002/2017JD027346.

Guillod, B. P., B. Orlowsky, D. G. Miralles, A. J. Teuling, and S. I. Seneviratne, 2015: Reconciling spatial and temporal soil moisture effects on afternoon rainfall. *Nat. Commun.*, **6**, 6443, https://doi.org/10.1038/ncomms7443.

Guo, X. Y., L. D. Tian, L. Wang, W. S. Yu, and D. M. Qu, 2017:

River recharge sources and the partitioning of catchment evapotranspiration fluxes as revealed by stable isotope signals in a typical high-elevation arid catchment. *J. Hydrol.*, **549**, 616–630, https://doi.org/10.1016/j.jhydrol.2017.04.037.

Guo, Z. C., and Coauthors, 2006: GLACE: The global land–atmosphere coupling experiment. Part II: Analysis. *Journal of Hydrometeorology*, **7**, 611–625, https://doi.org/10.1175/JHM511.1.

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

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.

Hu, X. L., and W. H. Yuan, 2021: Evaluation of ERA5 precipitation over the eastern periphery of the Tibetan plateau from the perspective of regional rainfall events. *International Journal of Climatology*, **41**, 2625–2637, https://doi.org/10.1002/joc.6980.

Huffman, G., E. Stocker, D. Bolvin, E. Nelkin, and T. Jackson, 2019: GPM IMERG Final Precipitation L3 1 Day 0.1 Degree x 0.1 Degree v06 (GPM_3IMERGDF). Available from https://doi.org/10.5067/GPM/IMERGDF/DAY/06.

Imamovic, A., L. Schlemmer, and C. Schär, 2017: Collective impacts of orography and soil moisture on the soil moisture-precipitation feedback. *Geophys. Res. Lett.*, **44**, 11 682–11 691, https://doi.org/10.1002/2017GL075657.

Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens, 2010: Climate change will affect the Asian water towers. *Science*, **328**, 1382–1385, https://doi.org/10.1126/science.1183188.

Jiang, D. B., Z. L. Ding, H. Drange, and Y. Q. Gao, 2008: Sensitivity of East Asian climate to the progressive uplift and expansion of the Tibetan Plateau under the mid-Pliocene boundary conditions. *Adv. Atmos. Sci.*, **25**, 709–722, https://doi.org/10.1007/s00376-008-0709-x.

Koster, R. D., and Coauthors, 2006: GLACE: The global land–atmosphere coupling experiment. Part I: Overview. *Journal of Hydrometeorology*, **7**, 590–610, https://doi.org/10.1175/JHM510.1.

Koster, R. D., and Coauthors, 2010: Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophys. Res. Lett.*, **37**, L02402, https://doi.org/10.1029/2009GL041677.

Loeb, N. G., N. Manalo-Smith, S. Kato, W. F. Miller, S. K. Gupta, P. Minnis, and B. A. Wielicki, 2003: Angular distribution models for top-of-atmosphere radiative flux estimation from the Clouds and the Earth's Radiant Energy System instrument on the Tropical Rainfall Measuring Mission satellite. Part I: Methodology. *J. Appl. Meteorol. Climatol.*, **42**, 240–265, https://doi.org/10.1175/1520-0450(2003)042<0240:ADMFTO>2.0.CO;2.

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.

Müller, O. V., P. L. Vidale, B. Vannière, R. Schiemann, and P. C. McGuire, 2021a: Does the HadGEM3-GC3.1 GCM overestimate land precipitation at high resolution? A constraint based on observed river discharge *J. Hydrometeorol.*, **22**, 2131–2151.

Müller, O. V., P. L. Vidale, B. Vannière, R. Schiemann, R. Senan, R. J. Haarsma, and J. H. Jungclaus, 2021b: Land–atmosphere coupling sensitivity to GCMs resolution: A multimodel assessment of local and remote processes in the Sahel hot spot. *J. Climate*, **34**, 967–985, https://doi.org/10.1175/JCLI-D-20-0303.1.

Koster, R. D., P. A. Dirmeyer, Z. Guo, G. Bonan, E. Chan, P. Cox, C. Gordon, S. Kanae, E. Kowalczyk, D. Lawrence, et al., 2004: Regions of strong coupling between soil moisture and precipitation. *Science*, **305**, 1138–1140, https://doi.org/10.1126/science.1100217.

Pielke R. A. Sr, 2001: Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. *Rev. Geophys.*, **39**, 151–177, https://doi.org/10.1029/1999RG000072.

Schwingshackl, C., M. Hirschi, and S. I. Seneviratne, 2017: Quantifying spatiotemporal variations of soil moisture control on surface energy balance and near-surface air temperature. *J. Climate*, **30**, 7105–7124, https://doi.org/10.1175/JCLI-D-16-0727.1.

Talib, J., C. M. Taylor, A. M. Duan, and A. G. Turner, 2021: Intraseasonal soil moisture–atmosphere feedbacks on the Tibetan Plateau circulation. *J. Climate*, **34**, 1789–1807, https://doi.org/10.1175/JCLI-D-20-0377.1.

Taylor, C. M., A. Gounou, F. Guichard, P. P. Harris, R. J. Ellis, F. Couvreux, and M. De Kauwe, 2011: Frequency of Sahelian storm initiation enhanced over mesoscale soil-moisture patterns. *Nature Geoscience*, **4**, 430–433, https://doi.org/10.1038/ngeo1173.

Taylor, C. M., R. A. M. 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.

Taylor, C. M., C. E. Birch, D. J. Parker, N. Dixon, F. Guichard, G. Nikulin, and G. M. S. Lister, 2013: Modeling soil moisture-precipitation feedback in the Sahel: Importance of spatial scale versus convective parameterization. *Geophys. Res. Lett.*, **40**, 6213–6218, https://doi.org/10.1002/2013GL058511.

Taylor, K. E., and Coauthors, 2017: CMIP6 Global Attributes, DRS, Filenames, Directory Structure, and CV's. Tech. Rep. v6.2.6., Program for Climate Model Diagnosis and Intercomparison.

Tong, K., F. G. Su, D. Q. Yang, L. L. Zhang, and Z. C. Hao, 2014: Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals. *International Journal of Climatology*, **34**, 265–285, https://doi.org/10.1002/joc.3682.

Ukkola, A. M., M. G. De Kauwe, A. J. Pitman, M. J. Best, G. Abramowitz, V. Haverd, M. Decker, and N. Haughton, 2016: Land surface models systematically overestimate the intensity, duration and magnitude of seasonal-scale evaporative droughts. *Environmental Research Letters*, **11**, 104012, https://doi.org/10.1088/1748-9326/11/10/104012.

Wan, B. C., Z. Q. Gao, F. Chen, and C. G. Lu, 2017: Impact of Tibetan Plateau surface heating on persistent extreme precipitation events in southeastern China. *Mon. Wea. Rev.*, **145**, 3485–3505, https://doi.org/10.1175/MWR-D-17-0061.1.

Wang, B., Q. Bao, B. Hoskins, G. X. Wu, and Y. M. Liu, 2008: Tibetan Plateau warming and precipitation changes in East Asia. *Geophys. Res. Lett.*, **35**, L14702, https://doi.org/10.1029/2008GL034330.

Wang, Z. Q., A. M. Duan, S. Yang, and K. Ullah, 2017: Atmo-

spheric moisture budget and its regulation on the variability of summer precipitation over the Tibetan Plateau. *J. Geophys. Res.*, **122**, 614–630, https://doi.org/10.1002/2016JD025515.

You, Q. L., F. Y. Wu, L. C. Shen, N. Pepin, Z. H. Jiang, and S. C. Kang, 2020: Tibetan Plateau amplification of climate extremes under global warming of 1.5°C, 2°C and 3°C. *Global and Planetary Change*, **192**, 103261, https://doi.org/10.1016/j.gloplacha.2020.103261.

Zhang, S. J., D. H. Wang, Z. K. Qin, Y. Y. Zheng, and J. P. Guo, 2018: Assessment of the GPM and TRMM precipitation products using the rain gauge network over the Tibetan Plateau. *Journal of Meteorological Research*, **32**, 324–336, https://doi.org/10.1007/s13351-018-7067-0.

Zhao, C. L., X. H. Meng, Y. Q. Li, S. Lyu, J. P. Guo, and H. Z. Liu, 2022: Impact of soil moisture on afternoon convection triggering over the Tibetan Plateau based on 1-D boundary layer model. *J. Geophys. Res.*, **127**, e2021JD035591, https://doi.org/10.1029/2021JD035591.

Zhu, Y.-Y. and S. N. Yang, 2020: Evaluation of CMIP6 for historical temperature and precipitation over the Tibetan Plateau and its comparison with CMIP5. *Advances in Climate Change Research*, **11**, 239–251, https://doi.org/10.1016/j.accre.2020.08.001.