A systematic evaluation of high-cloud controlling factors

Sarah Wilson Kemsley[1*], Paulo Ceppi[2], Hendrik Andersen[3,4], Jan Cermak[3,4], Philip Stier[5], Peer Nowack[3,6]

[1] Climatic Research Unit, University of East Anglia, UK
[2] Department of Physics, Imperial College London, UK
[3] Institute of Meteorology and Climate Research, Karlsruhe Institute of Technology (KIT), Germany
[4] Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Germany
[5] Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, UK
[6] Institute of Theoretical Informatics, Karlsruhe Institute of Technology (KIT), Germany

* *Correspondence to*: s.wilson-kemsley@uea.ac.uk

Formatted: Superscript

Abstract

Clouds strongly modulate the top-of-the-atmosphere energy budget and are a major source of uncertainty in climate projections. "Cloud Controlling Factor" (CCF) analysis derives relationships between large-scale meteorological drivers and cloud-radiative anomalies, which can be used to constrain cloud feedback. However, the choice of meteorological CCFs is crucial for a meaningful constraint. While there is rich literature investigating ideal CCF setups for low-level clouds, there is a lack of analogous research explicitly targeting high clouds. Here, we use ridge regression to systematically evaluate the addition of five candidate CCFs to previously established core CCFs within large spatial domains to predict longwave high-cloud radiative anomalies: upper-tropospheric static stability (S_{UT} ~~S_{UT}~~), sub-cloud moist static energy, convective available potential energy, convective inhibition, and upper-tropospheric wind shear (ΔU_{300}). ~~All combinations of tested CCFs predict historical, monthly variability well for most locations at grid-cell scales. Differences between configurations for predicting globally-aggregated radiative anomalies are more pronounced, where configurations including SUT outperform others.~~ We identify an optimal configuration for predicting high-cloud radiative anomalies that includes S_{UT} and ΔU_{300}, and show that ~~for predicting local, historical anomalies,~~ spatial domain size is more important than the selection of CCFs for predictive skill. We also find~~,~~ ~~finding~~ an important discrepancy between the optimal domain sizes required for predicting local and globally-aggregated radiative anomalies. Finally, we scientifically interpret the ridge regression coefficients, where we show that S_{UT} ~~S_{UT}~~ captures physical drivers of known high-cloud feedbacks, and ~~thus~~ deduce that inclusion of S_{UT} ~~S_{UT}~~ into observational constraint frameworks may reduce uncertainty associated with changes in anvil cloud amount as a function of climate change. Therefore, we highlight S_{UT} ~~S_{UT}~~ as an important CCF for high clouds and longwave cloud feedback.

1 Introduction

Changes in clouds are the primary source of uncertainty in the quantification of equilibrium climate sensitivity (ECS) – the long-term global warming response to a doubling of atmospheric carbon dioxide (Sherwood et al., 2020; Zelinka et al., 2022). Cloud-induced radiative anomalies (R) at the top-of-the-atmosphere (TOA) refer to changes in the balance of incoming and outgoing radiation caused by interaction with clouds. While most evidence suggests that the change in R at the TOA as a function of global warming likely has a positive effect on Earth's energy balance and thus amplifies warming (e.g., Ceppi and Nowack, 2021), the magnitude of this global cloud feedback remains highly uncertain (Ceppi et al., 2017; Sherwood et al., 2020; Zelinka et al., 2022).

Motivated by the role of clouds as a key uncertainty factor, much progress has been made towards understanding the mechanisms that drive changes in R, considering different cloud types under both natural unforced variability and long-term climate change. In particular, such work includes theoretical understanding of cloud feedback processes (e.g., Zelinka and Hartmann, 2010; Rieck, Nuijens and Stevens, 2012; Bony *et al.*, 2016); idealized regional modelling studies (Siebesma et al., 2003; Bretherton, 2015); convection-permitting global climate models (Rio et al., 2019); and climate model evaluation studies (Zelinka et al., 2022).

Here, we aim to systematically advance an alternative approach widely used for understanding and constraining uncertainties in cloud variability and trends in the form of Cloud Controlling Factor (CCF) analysis. Exploiting observed relationships between large-scale satellite cloud observations and meteorological predictor variables, CCF analyses have, for example, been used to derive observational constraints on cloud-related uncertainty estimates (Myers and Norris, 2016; Andersen et al., 2017, 2022; Fuchs, Cermak and Andersen, 2018; Ceppi and Nowack, 2021; Myers et al., 2021). In particular, meteorological CCFs for low marine and boundary-layer clouds have been widely assessed (Qu et al., 2015; Brient and Schneider, 2016; Klein et al., 2017; Scott et al., 2020; Andersen et al., 2022), with typical frameworks including CCFs such as surface temperature (T_{sfc}~~sfc~~), temperature advection, estimated boundary layer inversion strength (EIS~~EIS~~), vertical velocity, 700 hPa relative humidity (RH_{700}) and near-surface wind speed. However, comparatively less research has specifically targeted the CCFs for high clouds, despite their significant – and highly uncertain – contributions towards the total estimated feedback (Sherwood et al., 2020). A systematic comparison of CCF candidates for high clouds within a range of spatial domains will therefore be the main subject of this paper.

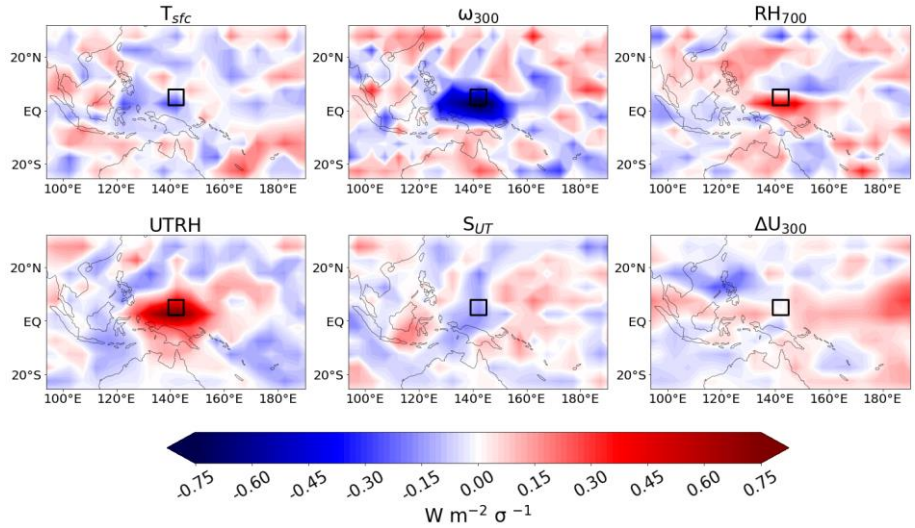

Figure 1. CMIP multi-model mean longwave cloud-radiative sensitivities for a sample 5◦× 5◦ target grid box (center 2.5° N, 142.5° E, indicated by the black box) to surface temperature (T_{sfc} ~~T_{sfc}~~), vertical velocity at 300 hPa (ω_{300}), relative humidity at 700 hPa and in the upper troposphere (RH_{700} and UTRH, respectively), wind shear at 300 hPa (ΔU_{300}), and upper-tropospheric static stability (S_{UT}) using a 21x11 domain of grid-boxes around the target (corresponding to 110° longitude x 55° latitude area, centered on the grid-box.). Radiative anomalies are normalized for a one-SD (σ) anomaly in the controlling factors, based on monthly variability.

Our work builds on a modification to a previous CCF approach, which was introduced by Ceppi and Nowack (2021, hereafter CN21). CN21 used ridge regression for their analyses, which allowed them to consider large spatial domains of CCF predictor patterns around target grid points in which cloud-radiative anomalies were predicted, with an example shown in Fig. 1. This approach contrasts with previous CCF analyses using standard multiple linear regression, which are constrained to a small number of predictors (typically < 10). This allowed their analysis to be extended beyond specific cloud regimes. As shown in CN21, the consideration of larger-scale CCF patterns led to improvements in predictive skill for both shortwave (SW) and longwave (LW) global cloud feedback. The intuition behind using spatial patterns of CCFs is motivated by the synoptic-scale atmospheric system within which the lifecycle of clouds – from formation to cessation – occurs, resulting in more robust predictions of global cloud feedback. Non-local features, such as large-scale patterns of sea-surface temperature anomalies and changes in the atmospheric circulation (e.g., convergence and divergence) are implicitly encoded using large spatial domains, which are not included in scalar CCF analysis despite their relevance for the context in which cloud development occurs (when considering monthly averaged data typically used for CCF analyses (Klein et al., 2017)). Altogether, considering larger-scale patterns resulted in better out-of-sample predictions, which consequentially tightened the cloud-induced uncertainty in general circulation model (GCM)-modelled ECS.

However, the framework introduced by CN21 highlighted an important limitation. As the same set of five CCFs were used for SW and LW analyses, their predictive skill was markedly stronger for global SW and net feedback components than for LW. Given that LW feedback is largely driven by high clouds, while SW feedback is instead predominantly driven by the oft-studied low clouds, we speculate the performance deficit may

be – at least to a degree – a symptom of CCF choice. Indeed, Zelinka et al. (2022) specifically recommend that the drivers of high cloud feedback must be targeted to reduce cloud-related uncertainty in ECS estimates.

To address these open questions, we use ridge regression to methodically assess candidate CCFs of high clouds within a range of spatial scales, aiming to inform CCF choice for future observational constraints on the ECS uncertainty. Here, we target LW cloud radiative anomalies (R_{LW}) as they are more directly associated with high clouds than SW (and consequently, net) radiative anomalies. We briefly assess implications of CCF choices on net anomalies, R_{NET} noting that, historically, LW and SW high-cloud radiative anomalies tend to offset each other, resulting in little net signal (R_{NET}) for thick clouds over monthly timescales. We therefore restrict our analysis to clouds with top pressures smaller than 680 hPa; future references to "R" are therefore specifically emanating from these non-low clouds (see Sect. 3.1 for the dataset used). Though radiative effects from midlevel clouds are also by definition included in our analysis, we collectively refer to radiative anomalies as "high" henceforth for simplicity (Zelinka et al., 2016).

We systematically assess static stability in the upper troposphere (S_{UT}), sub-cloud moist static energy (m), convective available potential energy (CAPE), convective inhibition (CIN) and upper tropospheric wind shear (ΔU for easterly shear) as CCFs ~~in on the basis of~~based on their physical relationships with high-cloud properties or convection, with an overview presented in Sect. 2. Aiming to inform choices for future *observational* constraint analyses, we only suggest CCFs that are readily available (or easily calculated from measurable quantities). Alternative variables, such as the radiatively-driven divergence, horizontal mass convergence, and gross moist stability, may also capture high-cloud properties but their derivation requires numerical modelling and hence we do not consider them here. Sections 3 and 4 discuss the data and methods we use, respectively, with combined results and discussion presented in Sect. 5. We first discern which CCF combinations ~~are able to~~can best predict out-of-sample grid-cell scale historical internal variability. We then investigate which combinations best predict out-of-sample globally-aggregated R_{LW}. Based on the results of our statistical testing, we physically interpret the coefficients for a single (optimal) ~~(deemed "optimal" by our analysis)~~ configuration of CCFs, and assess whether the spatial pattern, magnitude and variability of the cloud properties (i.e., cloud top pressure and cloud fraction) are accurately captured.

2 High-Cloud Controlling Factors

Ubiquitously present over the tropics, cirrus, cirrostratus and deep convective clouds are responsible for the largest annual-mean changes in global TOA LW flux (Chen et al., 2000). Tropical cirrus clouds develop through one of two mechanisms: outflow from deep convective cores, or in-situ ice formation that is not associated with convection (Gasparini et al., 2023; Kärcher, 2017). The former, referred to as "anvil cirrus" together with a mature cumulonimbus core, form tropical anvil clouds. "Thick" cirrus are both effective absorbers of upwelling LW radiation and also efficient reflectors of incident SW radiation. Over time, dynamical, radiative and microphysical processes can spread the thick anvil cirrus, extending anvil lifetime and resulting in larger cloud cover than the initial convective core (Luo and Rossow, 2004; Gasparini et al., 2023). Such processes can result in the formation of "thin" cirrus clouds, characterised by a relatively smaller SW cloud radiative forcing compared to LW (Jensen et al., 1994; McFarquhar et al., 2000). Though deep convective clouds presently have relatively small abundance (compared to other cloud types), their local radiative effects are large (Chen et al., 2000), and

therefore changes to their frequency of occurrence can have substantial impacts on cloud feedback. Despite this, most previous CCF analyses focused on low-cloud regimes so that the selection and design of CCFs were mainly motivated by meteorological situations driving cloud formation and cessation in those cloud regimes (Klein et al., 2017).

In CN21, a compromise was sought by considering classic CCFs such as T_{sfc} ~~sf~~, EIS ~~EIS~~, and RH_{700} ~~700~~ (relative humidity at 700 hPa), but by also using the vertical velocity at 500 hPa (ω_{500}) and upper-tropospheric relative humidity (UTRH, the vertically averaged relative humidity in the 200 hPa layer below the tropopause) as predictors in an attempt to additionally target high clouds. In the following, we will build on the CN21 CCF set-up, specifically targeting modifications and additions that are more likely to represent state variables important for the aforementioned high clouds. One-by-one we will motivate these CCF candidates physically and formally define, then test the prediction results of possible CCF combinations for high clouds in Sect. 5.

An overview of all CCFs considered and their scientific motivations is summarised in Table 1. We keep T_{sfc} ~~T$_{sfc}$~~, RH_{700} ~~RH$_{700}$~~, UTRH, and ω (at variable pressure levels) in all configurations, which we refer to as the "core" CCFs, as they jointly explain a large portion of historical variability in R_{LW}, and are each physically related to high-cloud formation. The large-scale distribution of tropical deep clouds is closely tied to the distribution of sea surface temperatures (SSTs) and upper-tropospheric relative humidity (Bony et al., 1997; Li et al., 2014), with research indicating that lower free-tropospheric relative humidity regulates the mean height of convective outflow (Sherwood et al., 2004). Vertical velocities (ω) indicate regions of subsidence or ascent, with enhanced ascending motion supporting thicker, higher cloud layers (Ge et al., 2021). Andersen *et al.* (2023) find that the magnitude of (local) sensitivity to ω is largest at 300 hPa, hence we test vertical velocity at both 300 hPa and 500 hPa (used in CN21) in this study.

Estimated boundary-layer inversion strength (EIS ~~EIS~~) is *not* typically regarded a controlling factor for high clouds specifically, despite its wide use in general and low-cloud CCF analyses. This results in relatively little literature interpreting high-cloud sensitivities to EIS ~~EIS~~. Despite this, CN21 used *only* the T_{sfc} ~~T$_{sfc}$~~ and EIS ~~EIS~~ sensitivities to observationally constrain global cloud feedback for both SW and LW components. These sensitivities are suitably decoupled from the clouds, and still achieve good (albeit poorer than SW and net) predictions for global LW feedback. We therefore suggest five candidate CCFs as replacements for EIS ~~EIS~~ that more directly represent convective processes or high-cloud formation, that are also sufficiently external to the clouds themselves and may be similarly used in constraints.

We list candidate CCFs (and EIS ~~EIS~~) and discuss them in turn below, with the exact definitions provided in Sect. 3.2:

- **Static stability** is the vertical gradient of potential temperature, measuring the stratification of the atmosphere (Grise et al., 2010). Upper-tropospheric static stability is robustly (negatively) correlated with upper-level cloud incidence over much of the global ocean (Li et al., 2014) and has been observationally linked with changes in tropical anvil cloud fraction through the "anvil iris" thermodynamic mechanism (Bony et al., 2016; Saint-Lu et al., 2020, 2022). We expect increases in local upper tropospheric static stability to result in local reductions in high cloud fraction, with suppressed vertical motion;
- **Moist static energy** characterises the energy of an air parcel in a moist environment, considering its internal energy (latent and sensible heat) and potential energy due to its

elevation. Sub-cloud moist static energy (m) may affect cloud formation, as higher levels of m signify increased potential for uplift and condensation. Additionally, when buoyant air from the boundary layer fills the free troposphere, it can inhibit initiation of convection in colder regions, setting a threshold that hinders further upward movement (Srinivasan and Smith, 1996; Zhang and Fueglistaler, 2020). We suggest that high m increases local high-cloudiness, while in contrast, we ~~speculate~~ hypothesize that non-local m can either decrease (due to convective thresholds) or increase cloudiness (depending on horizontal transport);

- **Convective available potential energy (CAPE)** is a measure of deep instability, describing the amount of energy available for an air parcel to rise freely through the atmosphere. CAPE offers insights into the onset, genesis and scale of atmospheric deep convection, and has been described as the fuel for a thunderstorm (Donner and Phillips, 2003; Jensen and Delgenio, 2006; Riemann-Campe et al., 2009). We speculate increased CAPE suggests an environment conducive to sustaining deep convection, and thus more high cloud;
- **Convective inhibition (CIN)**, a form of conditional instability and CAPE's opposing parameter, is a measure of the amount of energy required for a parcel to overcome a stable layer of air and initiate the development of deep convection. A large absolute value of CIN may indicate a stable atmosphere, and thus unfavourable conditions for the development of deep convective clouds (Louf et al., 2019). Note that high CIN is a required precursor for the buildup of CAPE. Once CIN has been overcome, conditions are favourable for deep convection;
- **Wind shear,** defined here as the vertical change in horizontal wind speed, is an important dynamical characteristic of the upper troposphere. Wind shear influences the organisation of convective storms and mesoscale convective systems in various ways, though understanding its relationship with cloud properties has proved historically challenging (Anber et al., 2014). However, studies suggest that wind shear can increase cloud-top turbulence, spread and stretch clouds horizontally through the advection of air at different levels and speeds, and hasten cirrus cloud dissipation (Lin and Mapes, 2004; Marsham and Dobbie, 2005; Jensen et al., 2011). We speculate wind shear mainly affects high-cloud fraction;
- **Estimated inversion strength (*EIS*~~EIS~~)** describes the strength of the boundary layer and is a dominant control for low-clouds (Wood and Bretherton, 2006; Andersen et al., 2022, 2023) and is widely used in general CCF analysis (CN21, (Klein et al., 2017). However, *EIS* ~~EIS~~ is not considered a driver of high-cloud incidence, but CN21 ~~speculated~~ suggested that *EIS* ~~EIS~~ may function as a proxy for factors relating to deep convection.

Note that several candidate CCFs are not independent. For example, high values of CIN are required for a buildup of CAPE, and a stable boundary layer may be represented by both high CIN and high *EIS*~~EIS~~.

Table 1. High-cloud controlling factors used in CN21 and proposed here, physical explanations connecting them to high clouds or convection, and the key studies supporting them. References to "clouds" in this table are for high clouds only. *EIS* ~~EIS~~ is not a core CCF, and therefore for conciseness we include *EIS* ~~EIS~~ under the "Candidate CCFs" subheading.

Cloud controlling factor	Physical explanation	Key studies
Core cloud controlling factors		
Surface temperature (T_{sfc} ~~T_{sfc}~~)	Warming surface temperature heats atmospheric column; large-scale distribution of clouds is tied to atmospheric profile of temperature; anvil clouds approximately rise with isotherms.	(Bony et al., 1997; Zelinka and Hartmann, 2011; Fueglistaler, 2019)
Free-tropospheric relative humidity (RH_{700})	Regulates mean height of convective outflow.	(Sherwood et al., 2004)
Upper-tropospheric relative humidity (UTRH)	Tropical clouds tied to spatial distribution of UTRH and lifetime of anvil clouds. A reciprocal relationship may exist; UTRH modulated by detrainment.	(Bony et al., 1997; Li et al., 2014)
Vertical pressure velocity (ω)	Indicates regions of ascent and subsidence. Enhanced ascending motion supports thicker clouds.	(Ge et al., 2021)
Candidate CCFs		
Estimated boundary layer inversion strength (EIS ~~EIS~~)	Limited literature; perhaps serves as a proxy for deep convective processes; strength of boundary layer inhibits convection.	CN21
Upper-tropospheric static stability (S_{UT})	Static stability associated with radiatively driven convergence; anvil altitude and amount collocate with peak convergence.	(Zelinka and Hartmann, 2010; Li et al., 2014; Bony et al., 2016; Saint-Lu et al., 2020, 2022)
Convective Available Potential Energy (CAPE)	Measure of deep instability; indicates energy available for convection.	(Donner and Phillips, 2003; Jensen and Delgenio, 2006; Chakraborty et al., 2016; Louf et al., 2019)
Convective Inhibition (CIN)	Shallow instability; indicates the energy required to leave stable boundary layer.	(Louf et al., 2019)
Sub-cloud moist static energy (m)	Moisture content of sub-cloud atmosphere fuels convection.	(Zhang and Fueglistaler, 2020)
Upper-tropospheric wind shear (ΔU_{300})	Influences organisation of convective storms; affects cloud-top turbulence and mesoscale anvil formation; affects cloud cover.	(Lin and Mapes, 2004; Marsham and Dobbie, 2005; Jensen et al., 2011).

Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left
Formatted: Left

3 Data

We use monthly-mean (unless explicitly mentioned otherwise) cloud property and CCF data, re-gridded to a common 5°x5° resolution. At these spatial and temporal scales, we expect the clouds to be approximately in equilibrium with their environment (Klein et al., 2017). To represent observed cloud-radiative data, we use

combined Moderate Resolution Imaging Spectroradiometer (MODIS) retrievals from both Aqua and Terra instruments, identified as MCD06COSP (Pincus et al., 2023). These retrievals are included as part of the CFMIP Observation Simulator Package (COSP, where CFMIP refers to the Cloud Feedback Model Intercomparison Project), which facilitates the evaluation of models against observations in a consistent manner (Bodas-Salcedo et al., 2011). For climate model data, we use eighteen GCMs that have run the International Satellite Cloud Climatology Project (ISCCP) simulator (Zelinka et al., 2012a) from the Coupled Model Intercomparison Project phases 5 and 6 (CMIP5/6). For a full list of CMIP models used in this research, see Supplementary Material Sect. S1. For the meteorological CCFs we use ERA5 reanalysis data at monthly resolution, ~~with the exception of~~except for CAPE and CIN which we first calculate using *daily* air temperature and relative humidity profiles, and then take the monthly mean. We use reanalysis data as a proxy for direct observations; henceforth, when "observed" results are discussed, we refer to predictions made for observed radiative anomalies using ERA5 meteorological CCFs.

We restrict the CMIP datasets to twenty years, aligned with the length of available observational record, though with slightly different time periods. For observations, data is available from July 2002 to June 2022. For the CMIP models, we use historical data from January 1981 to December 2000. We use this period because it is close to the present-day climate, under the constraint of availability of historical CMIP data (and noting that only a small set of models provide satellite simulator output for the RCP and SSP scenarios). For predictions of observed and modelled R_{LW}, we restrict our analysis from 60°S – 60°N. As is commonplace in CCF analysis, the seasonal cycles (climatological averages of each month) have been removed from the CCFs and radiative anomalies (Myers et al., 2021; Andersen et al., 2022). Prior to analysis, predictor variables are scaled to unit variance and zero mean to weight signals equally in the optimisation process (Scott *et al.*, 2020, CN21).

3.1 Cloud property histograms

Our analysis is based on histograms of cloud fraction as a joint function of cloud top pressure (CTP) and cloud optical depth (τ). Cloud-radiative kernels are used to convert binned cloud amount anomalies into top-of-atmosphere radiative flux anomalies, and to partition these into contributions from changes in cloud top pressure (CTP), cloud fraction (CF), and optical depth (τ), with a small residual contribution (Zelinka et al., 2012a, b, 2016). The cloud-radiative kernels we use here were first introduced in Zelinka et al. (2012a), with an improved decomposition method presented in Zelinka et al. (2016). Note that the same kernels (developed using ERA5 Interim temperature, humidity and ozone profiles) are used to decompose both the observed and modelled radiative anomalies. Cloud-radiative kernels are available from https://github.com/mzelinka/cloud-radiative-kernels.

3.2 Meteorological cloud controlling factors

Static stability is calculated using an interpolated monthly air temperature, ~~(T_i)~~ and pressure, p, profile~~s~~. The CMIP and ERA5 $T-p$ profiles are interpolated to 100 vertical levels using cubic spline interpolation to standard pressure levels. The static stability, S_p, at pressure level p is hence calculated using ~~S_p is~~

$$S_p = \frac{R_C}{C}\frac{T_p}{p} - \frac{dT}{dp} \quad (1)$$

Formatted: Font: (Default) Times New Roman, Font color: Auto, Pattern: Clear

Formatted: Font: (Default) Times New Roman, Font color: Auto, Pattern: Clear

where ~~where S_p is the static stability at pressure p,~~ C is the specific heat at constant pressure, and R_C the gas constant. We define upper tropospheric static stability, S_{UT}, ~~at~~ as an average over the interpolated ~~the standard~~ pressure level~~s closest~~ from ~~to the~~ the tropopause height in pressure units plus 50 hPa and plus 200 hPa~~in pressure units~~, where the monthly-mean tropopause is calculated using the standard WMO definition (Reichler et al., 2003). We interpolate the $T - p$ profile as standard pressure levels are too coarse to accurately calculate the second term in Eq. (1). ~~This~~ We vary the exact pressure levels that we average S_{UT} over ~~is~~ to ensure that our definition ~~of S_{UT}~~ accounts for the zonal distribution of tropopause height.

Moist static energy, CAPE and CIN are calculated using the Metpy V1.3.1 Python package (May and Bruick, 2019). Moist static energy is calculated at standard pressure levels using monthly air temperature and relative humidity datasets. To approximate sub-cloud moist static energy, m, we average moist static energy from the surface to (and including) 700 hPa. We use MetPy's "most unstable" CAPE and CIN function, which we calculate for all available CMIP models and ERA5. This involves calculating the most unstable air parcel from the temperature and humidity profiles, and hence calculating CAPE and CIN using this parcel. CAPE and CIN are first calculated using daily temperature, humidity and pressure values at standard CMIP pressure levels and then averaged for each month. Of the eighteen CMIP models, daily datasets for atmospheric temperature and humidity are only readily available for fourteen of the models (see Sect. S1 in Supplementary Material).

Free-tropospheric vertical wind shear is calculated as the difference in 925 hPa and 300 hPa easterly wind speeds, U, standardised by the change in geopotential height, z, where

$$\Delta U_{300} = \frac{U_{300} - U_{925}}{z_{300} - z_{925}} \quad (2)$$

with subscripts referring to the pressure levels for each variable (Chakraborty et al., 2016). Both easterly and northerly wind shear have been assessed, though we only discuss easterly shear here as overall performance metrics are relatively consistent between the directions of shear.

T_{sfc}, ω_{300}, ω_{500} and RH_{700} are directly observable or modelled quantities. We define EIS ~~EIS~~ and UTRH consistently with CN21. EIS ~~EIS~~ is a measure of lower-tropospheric stability, defined relative to the temperature-dependent moist adiabatic lapse rate (Wood and Bretherton, 2006) over global oceans. Over land, this is simply defined as the difference between the potential temperature at 700 hPa and the surface (Klein and Hartmann, 1993). UTRH is the vertically averaged relative humidity within the 200 hPa-layer below the tropopause (again defined using the WMO standard definition). Monthly-mean climatologies for all CCFs can be found in Fig. S1.

4 Method

4.1 Ridge regression

We use ridge regression to estimate sensitivities of cloud-radiative anomalies to changes in surrounding meteorological CCFs within two-dimensional spatial domains. While still being a linear least-squares regression approach, the inclusion of an L2-regularization penalty term means that the method can more effectively deal with high-dimensional regression problems than unregularized multiple linear regression (Hoerl and Kennard, 1970; CN21; Nowack et al., 2021). This, in turn, allows us to consider larger domains of CCFs as predictors in the first place, leading to improved generalized predictive skill. The spatial domain within which CCFs are used to predict

R at a central grid-cell, r, is referred to by the number of grid-cells in a longitude x latitude space (i.e., a 7x3 domain corresponds to 35° longitude x 15° latitude, see also Fig. 1). Five domain sizes are tested: 1x1, 7x3, 11x5, 15x9 and 21x11.

Statistical cross-validation is used to optimise the regression fit by minimising the cost function,

$$J_{ridge} = \sum_{t=1}^{n}\left(R(r)_t - \sum_{i=1}^{M} c_i X_{i,t}\right)^2 + \alpha \sum_{i=1}^{M} c_i^2 \quad (3)$$

which puts a penalty on overly large regression coefficients, c_i; where n is the number of datapoints; $X_{i,t}$ is the i-th CCF at time t; M is the number of dimensions in the model (i.e., for a 7x3 domain using five unique CCFs, $M = 7\ x\ 3\ x\ 5 = 105$); r the central-grid cell, and α the regularisation parameter.

The first term in Eq. (3) is the ordinary least squares regression error. We classically approximate $R(r)$ by a linear function of anomalies in the set of M cloud controlling factors,

$$dR(r) \approx \sum_{i=1}^{M} \frac{\partial R(r)}{\partial X_i}\ \partial d X_i. \quad (4)$$

We refer to

$$\Theta_i(r) = \frac{\partial R(r)}{\partial X_i} \quad (5)$$

as the sensitivities, $\Theta_i(r)$, of $R(r)$ to anomalies in the i-th CCF. See Fig. 1 for an example of the spatial pattern of for six CCFs using a 21x11 domain.

Using fivefold cross-validation, we determine the optimal value for the regularization parameter, α, where the second term on the right-hand side of~~in~~ Eq. (3) is the L2-regularization penalty. We split the timeseries into five ordered time slices and optimise α by fitting Eq. (3) to each of four slices at a time. Optimal α is hence found by evaluating predictions on the fifth time slice using the R^2 score independently for each location in the observed and modelled datasets.

For Sect. 5.1, 5.2 and 5.4 we use sensitivities to predict a ~~subsequently predict~~ two- year~~s of withheld data~~validation dataset. We repeat this process, rotating the ~~rotate the~~ withheld data~~set~~ every two years~~,~~ resulting~~resulting~~ in ten unique training-validation dataset combinations (see Supplementary Fig. S2 for a schematic of this process). Each of the ten two-year validation datasets ~~Predictions~~ are subsequently concatenated, resulting in a continuous twenty-year timeseries predicted "out-of-sample"~~,~~. The rotation of training-validation datasets ~~with no~~results in no datapoint ~~having~~ having been predicted using the same dataset that the model was trained on. Standard performance metrics (Pearson r correlation coefficient, R^2 score, and root mean squared error, RMSE) are calculated using the concatenated predictions and the original twenty-year dataset. For Sect. 5.3, we use the sensitivities estimated from ~~the~~ a full twenty-year dataset to visualise spatial distributions.

5 Results and Discussion

Here we present results for the CCF analyses for R_{LW}, including a systematic assessment and intercomparison of possible CCF configurations and domain sizes. "CCF configuration" refers to the combination of meteorological variables used to predict R_{LW}. Configurations are labelled based on which of the proposed CCFs (shown in Tab. 1) are used in addition to the following core retained factors T_{sfc}~~sfc~~, ω_{300}~~0~~, RH_{700}, and UTRH

(i.e., configuration S_{UT} refers to predictions made using T_{sfc}, ω_{300}, RH_{700}, ~~T_{sfc}, ω_{300}, RH_{700},~~ UTRH and S_{UT}). Where appropriate, we additionally point to the corresponding R_{NET} results in the Supplementary Material.

~~In the following, w~~We first compare CCF configurations using standard performance metrics for time series predictions. Since we learn separate CCF functions to predict R_{LW} at each 5° x 5° grid-point, we ~~first~~ briefly evaluate prediction performance of those functions individually, which we refer to as "local" predictions. We then average local performance metrics near-globally (i.e., for all available predictions, 60°S – 60°N inclusive), henceforth simply referred to as "globally" averaged, with grid-cells weighted by the cosine of their latitude. We also average metrics in the tropical ascent regions, which we define as grid-cells with observed climatological EIS ~~EIS~~ < 1 K, ω_{500} < 0 hPa s^{-1}, and latitude equatorward of 30° (Medeiros and Stevens, 2011).

~~We note that not all regions are equally cloudy, which leads to differently high levels of variance in R_{LW} and, thus, signals for the regression to learn from. As a result, performance metrics tend to be lower in subsidence regions where there are few high clouds. For the global performance averages, we have addressed this issue by weighting grid-cells based upon their climatological mean R_{LW} in addition to the cosine of their latitude, to avoid penalising average metrics from low scores in regions with relatively little signal.~~

Using the CCF framework, an observational constraint on global cloud feedback can be made using local R_{LW} predictions under a forcing (such as 4xCO_2) that are aggregated globally and normalised by the change in global mean surface temperature. Though we do not predict feedback here, we instead assess which CCF configuration best estimates the globally-aggregated R_{LW} by spatially averaging each local prediction and target value globally (and in tropical ascent regions) first, and then calculating the performance metrics. Henceforth, note a distinction between globally averaged *metrics* for local predictions (e.g., Fig. 2a-b) and metrics for globally-*aggregated* R_{LW} (e.g., Fig. 2c-d).

5.1 Predictive skill on observations

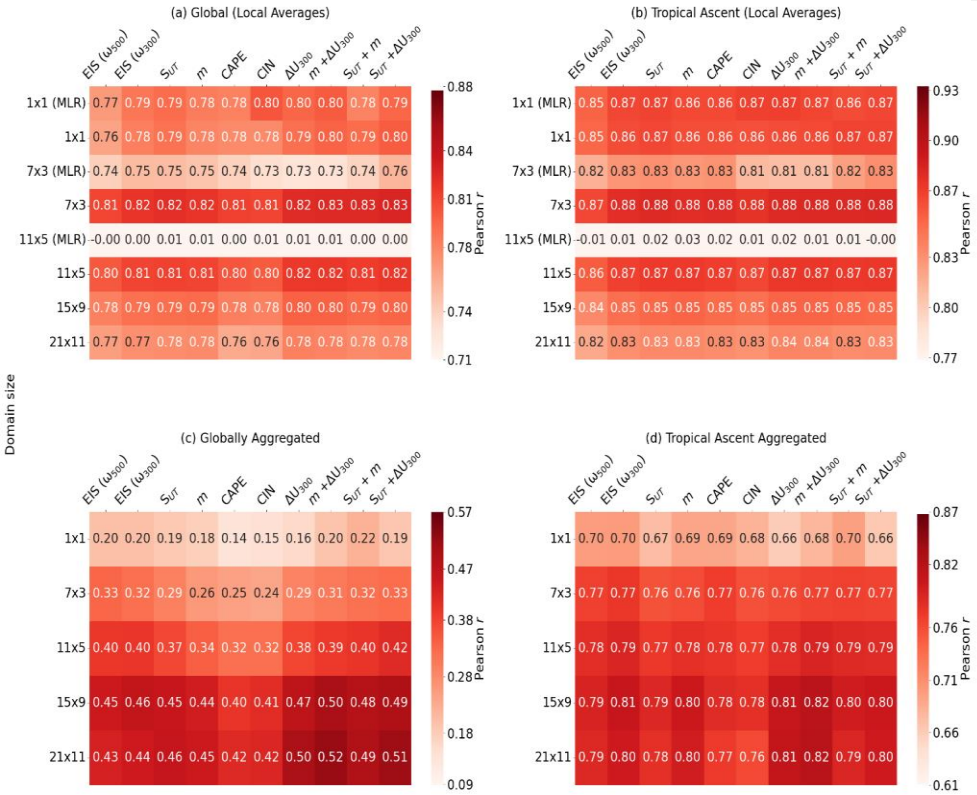

Figure 2. Matrices showing ~~skill metrics~~Pearson r for predictions made for the observed R_{LW} time series at each domain size using different "CCF configurations". A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}. Each configuration uses T_{sfc}, RH_{700}~~RH_{700}~~, UTRH and ω_{300}~~ω_{300}~~ (with the exception of the first column, where ω_{500}~~ω_{500}~~ is used instead) and a candidate CCF(s) (e.g., S_{UT}), which is used to label each column. Predictions are made locally, with the ~~RMSE~~ Pearson r averaged (a) globally and (b) in tropical ascent regions defined as grid-cells with observed climatological EIS ~~EIS~~ < 1 K, ω_{500}~~ω_{500}~~ < 0 hPa s^{-1}. Metrics are weighted by the cosine of latitude and monthly standard deviation of R_{LW} of each grid-cell (see Supplementary Material S2). Pearson r is also shown for aggregated predictions, (c) globally and (d) in the tropical ascent regions, and compared to similarly aggregated observations. All predictions are made using ridge regression, except for ~~row~~ rows~~1~~ 1x1 (MLR), ~~and~~ 7x3 (MLR) and 11x5 (MLR) in panels (a) and (b), which are made using multiple linear regression. Note different scales for each colorbar~~, with darker regions indicating higher skill (lower RMSE, higher Pearson r)~~.

We first assess~~assess~~ CCF configuration skill for local predictions, with results shown in Fig. 2a-b (with columns c-d showing globally-aggregated results). Using ridge regression, we confirm that all configurations predict out-of-sample local R_{LW} well at all domain sizes (with correlation matrices qualitatively consistent using R^2 and RMSE, not shown) ~~(shown in Fig. S2)~~. To demonstrate the strengths of ridge regression while using collinear predictors in high dimensions, we briefly compare our results to the traditional multiple linear regression (MLR) approach. Using a 1x1 domain, there is little difference in skill between predictions made with MLR and

ridge regression. Beyond 7x3, MLR coefficients become unstable, resulting in predictions that are not correlated with the observed (e.g., 11x5; results for larger domain sizes are not shown).

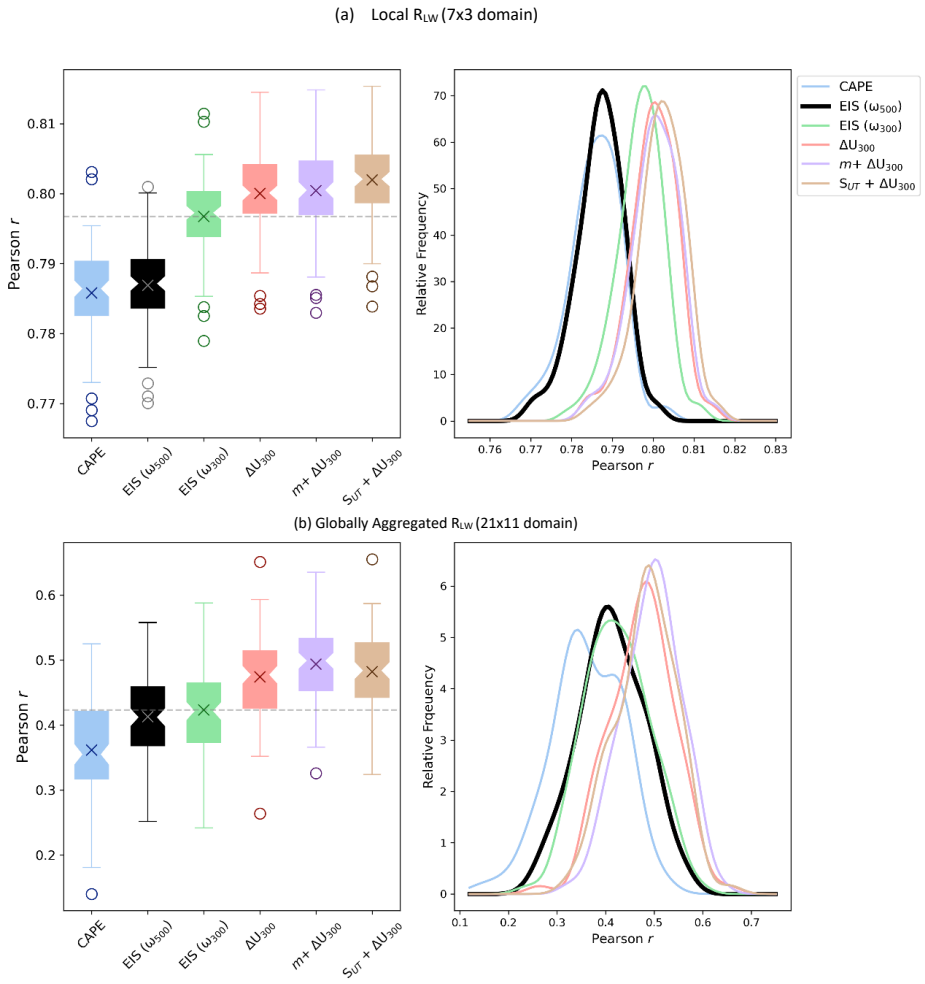

Figure 3. Box and whisker plots (left panels) showing the distribution of observed predictive skill based on 100 bootstrapped samples of R_{LW} for a selection of the CCF configurations. Crosses show the means, notches show the medians, and circles show the outliers. A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}, where each configuration uses T_{sfc}, RH_{700}, UTRH and ω_{300} (with the exception of the first box and whisker, where ω_{500} is used instead) and a candidate CCF(s) (e.g., S_{UT}), which is used to label each configuration. The right panels show the shapes of the distributions using a kernel density estimator. The top panels (a) show the distributions for local predictions at the 7x3 optimal domain size (analogous to Fig. 2a) and the bottom panels (b) show the distributions for the 21x11 globally-aggregated optimal domain size (analogous to Fig. 2c). EIS (ω_{300}) is highlighted in black to facilitate easier comparison between configurations.

We find local performance only slightly~~o~~ depends ~~both~~ on the CCF configuration, with EIS (ω_{500}) exhibiting the weakest performance (note that EIS (ω_{500}) is the configuration used in CN21).~~,~~ This is likely because a large proportion of the monthly variability is already explained using only T_{sfc}, ω_{300}, RH_{700} and UTRH

without the inclusion of additional CCFs (i.e., for 7x3, R^2 = 0.64 using core CCFs, compared with R^2 = 0.69 using EIS (w_{300}) in addition). Though changes in local skill (when globally averaged) between the CCF configurations are subtle, we find ~~in Sect. 5.2, showing~~ qualitatively consistent results for the CMIP models reaffirming that changes are robust. Predictive skill is instead more dependent on domain size, with metrics peaking at the 7x3 domain. We investigate this dependency on domain size in more detail in Section 5.1.1.

In line with Andersen et al. (2023) ~~(~~though note high-cloud radiative anomalies are not isolated in their study), we find the single largest improvement in R_{LW} predictive skill is achieved through changing ω from 500 hPa to 300 hPa, reflected by a large positive shift in the distributions shown in Fig. 3a. This suggests ω_{300} more effectively predicts deep convective and cirrus cloud radiative effects than ω_{500}, as we would expect (Ge et al., 2021). We do find that this results in a slight drop in performance for R_{NET} (Fig. S3) which likely is because 500 hPa instead better targets midlevel clouds which drive a shortwave contribution to R_{NET} that is not present for R_{LW}. However, comparing across configurations using the same vertical velocity reveals qualitatively similar heatmaps for R_{NET} and R_{LW} (note m performs slightly better for R_{NET} than R_{LW}). An additional prominent shift to the distribution arises through the inclusion of ΔU_{300}.

Given that raising the vertical pressure velocity results in a strong positive shift, we henceforth choose to replace ω_{500} with ω_{300} and compare further candidate CCF configurations with EIS (ω_{300}) as a new baseline for comparison (highlighted in black in ~~Fig. 3~~Fig. 3). ~~we confirm thatislyt Thus, at~~ At the optimal 7x3 domain, we find configuration $S_{UT} + \Delta U_{300}$ to reproduce observed local R_{LW} with the highest skill~~,~~ and we hence .

~~The~~show the spatial distributions ~~of~~ for predictive skill ~~arehence~~ in Fig. S4.

To quantify whether differences between configurations are statistically significant for the observed anomalies, we generate a distribution of Pearson r values using bootstrapping (Davison and Hinkley, 1997). We randomly sample the observed data (with replacement) 100 times, creating datasets equivalent in length to 18 years. Any remaining months are used as a validation dataset, where r is determined against predicted values. This process results in a distribution of 100 r values for each configuration, providing an estimate of predictive skill uncertainty, with a selection of the configurations shown in Fig. 3. The non-parametric Kruskal-Wallis test is hence used to identify statistical differences between all of the distributions. We find highly significant differences between all of the configurations ($p < 10^{-30}$). Accounting for its highest global median r, we pairwise test the predictive skill distribution for $S_{UT} + \Delta U_{300}$ with all other configurations (using an adjusted significance level of 0.5 % to account for multiple hypothesis testing). We find statistical similarity with only $m + \Delta U_{300}$ and ΔU_{300} ($p = 0.06$ and $p = 0.01$, respectively).

~~and, more prominently, on domain size (Fig. 2a-b) (note that EIS (ω_{500}) is the configuration used in CN21). Correlation matrices are qualitatively consistent between the performance metrics, though improvements are more pronounced for RMSE (at local scales) than for Pearson r and R^2 (likely owing to strict upper and lower limits for these metrics). We therefore show the RMSE in Fig. 2, with R^2 and Pearson r shown in Fig. S2. Though changes in local skill (when globally averaged) between the configurations are subtle, we reaffirm that they are indeed robust in Sect. 5.2, showing qualitatively consistent results for the CMIP models. In this section, we will discuss configurations for the optimal domain size for local predictions (7x3).~~

We now focus on predictive performance for the globally-*aggregated* R_{LW} time-series, with results shown in Fig. 2c-d and Fig. 3b. While local prediction performance peaks at 7x3 and is followed by a drop in skill, we find a discrepancy with the globally-aggregated performance, which instead *increases* with domain size.

Formatted: Not Highlight

Formatted: Default Paragraph Font, Not Superscript/ Subscript

Formatted: Indent: First line: 1.27 cm

For some configurations, r continues to increase beyond 21x11, though this begins to tail off (not shown). The relationship between domain size and predictive skill now aligns with the findings of CN21, where they show that the correlation between observed and predicted global cloud *feedback* increases with domain size. However, as domain size increases, so too do the model dimensions and thus the complexity. Owing to the trade-off between small improvements at even larger domain sizes and increased complexity, we restrict our analysis to 21x11 and below, and discuss globally-aggregated results using the 21x11 domain.

Here we find more marked improvements in predictive skill for most of the CCF configurations compared to EIS (ω_{500}), with performance again strongly dependent on domain size (Fig. 2c-d). However, we now find that changing the pressure level of ω no longer results in a substantial positive shift of the skill distributions, though inclusion of ΔU_{300} still results in improvements (Fig. 3b). We also note that performance metrics for globally-aggregated R_{LW} are comparatively worse than the globally-averaged *local* metrics. This is in line with accumulation of local errors and reduced variability in the globally-aggregated anomalies. In a comparison of all globally-aggregated distributions shown in Fig. 3b, ~~each~~there is evidence showing statistical differences ~~-~~at the 5 % significance level (with $p < 10^{-40}$). Here, $m +$ ΔU_{300} has the highest median r. In a pairwise comparison of $m + \Delta U_{300}$ with each other distribution, we find statistical differences with all configurations except $S_{UT} + \Delta U_{300}$ ($p = 0.02$) and ΔU_{300} ($p = 0.03$), again using an adjusted 0.5 % significance level owing to multiple statistical tests.

Neither CAPE nor CIN improve predictive skill at either scale compared to alternative candidate CCFs for most domain sizes. CAPE and CIN have been included as a CCF for their links to deep convection, which is not frequent outside of the warm tropics, resulting in being particularly poor predictors in the high-latitude extratropics (Fig. S4; for CMIP see S5). Additionally, literature hints at a potentially non-linear relationship between CAPE, CIN and high-cloudiness that would not be captured by the linear ridge regression. For example, in high-CAPE environments it is thought that there may generally be enough CAPE for convection to occur, indicating that the exact magnitude of CAPE is less important than passing a threshold for the onset of deep convection (Sherwood, 1999). The distribution of predictive skill also suggests there is a more complex relationship between CAPE (and CIN, not shown) and R_{LW}. Given that the distributions are calculated using randomly resampled datasets through bootstrapping with replacement, datapoints will be repeated. This reduces the diversity of the training data, which can result in poorer generalization of more complex or noisy relationships.

Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Not Highlight
Formatted: Keep with next

5.1.1 CCF importance at different spatial scales

We investigate ~~We suggest~~ the evolution of ~~the discrepancy~~ predictive skill with domain size for local and globally-aggregated predictions. ~~However, the EVF for ω_{300} substantially decreases when globally-integrating R_{LW} to only 6.3 %. This additionally explains why there is little difference between the performance of ω_{300} and ω_{500} in Fig. 2c-d, despi~~Owing to the linearity of ridge regression, we can partition the predicted local R_{LW} signal into contributions from each CCF, such that (for example)

$$R_{LW} = R_{LW(T_{sfc})} + R_{LW(RH_{700})} + \cdots + R_{LW(\Delta U_{300})}, \quad (6)$$

where $R_{LW(T_{sfc})}$ is the component of R_{LW} predicted using only T_{sfc} within the specified domain size ~~(here we use 21x11),~~ and so on for each CCF in the configuration. For each CCF, we calculate the explained variance fraction (EVF) for R_{LW} at each grid-cell. Equation (6) ~~can~~ is repeated for the global R_{LW} predictions, where local predictions are *first* globally-aggregated for each CCF and *then* summed. CCFs with higher EVFs are referred to

as more “important” for the predicted values (i.e., UTRH is typically the most important predictor for both local and globally-aggregated predictions (shown in Fig. 4)). Note that it is plausible that this may show bidirectional causality, where the presence of high cloud influences UTRH by modulating the moisture content in the upper troposphere (i.e., outflow from convective anvils), though our analysis cannot separate the direction.

The interaction between domain size, cloud controlling factors, and predictive skill is complex. We ~~We show an example of the EVF spatial distribution in the Supplementary Material (Fig Sxyz).~~summarise key points below:

- There is an emergent distinction between “local” and “non-local” predictors. For example, EVF for UTRH decreases with increasing domain size and, accordingly, we find that local UTRH sensitivities typically have strong magnitudes close to target grid-cell, with noisy, spatially incoherent coefficients further afield (see Fig. S6a-b for an example); thus, we describe UTRH as a “local” CCF (similarly for ω_{300} and RH_{700}).
- EVF for T_{sfc}, ΔU_{300}, and S_{UT} increases with domain size (i.e., “non-local” predictors), and each contribute a greater proportion of the globally-aggregated predictions compared to local predictions (Fig. S6c-d).
- Predictive skill is likely a trade-off between adding relevant information from “non-local” CCFs while adding superfluous information from “local” CCFs; i.e., too distant information does not provide additional predictive skill, at least to the degree that it would outweigh the corresponding increase in dimensionality of the regression problem.

- For globally-aggregated predictions, ω_{300} is the least important predictor (compared to the second most important for local predictions), thus explaining why the choice of pressure level of ω is less relevant at global scales (shown in Fig. 4) than locally.

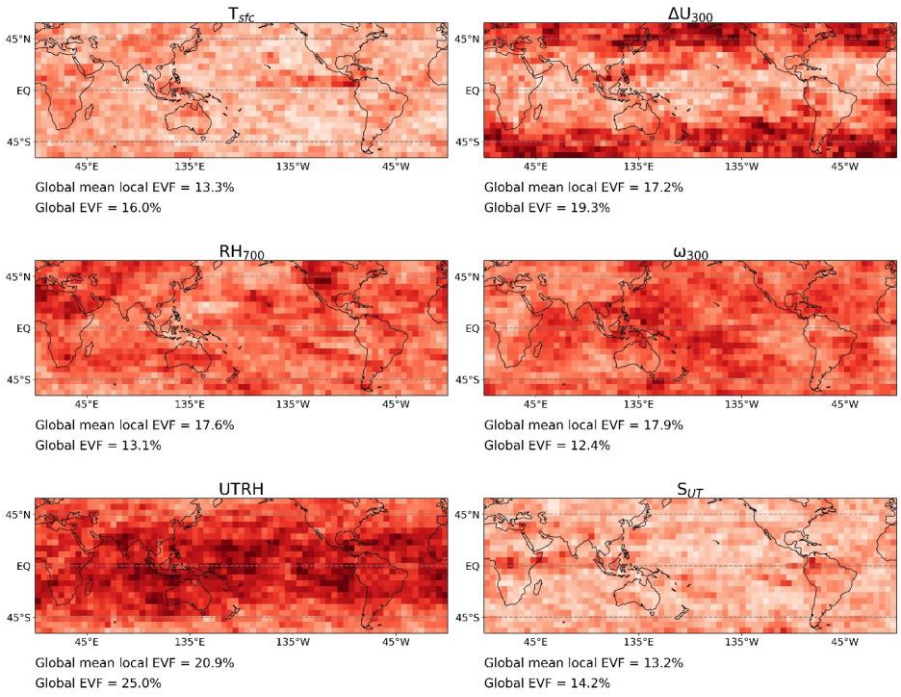

Figure 4. Maps showing the explained variance fraction (EVF) as a percentage for local predictions of R_{LW} using a 21x11 domain and using configuration $S + \Delta U_{300}$ (with T_{sfc}, RH_{700}, UTRH and ω_{300}). "Global mean local EVF" refers to the global mean EVF from local predictions, weighted by the cosine of each grid-cell's latitude. "Global EVF" refers to the EVF for each CCF's contribution to the globally-aggregated R_{LW}.

~~Figure 3. Maps showing the explained variance fraction (EVF) as a percentage for local predictions of R_{LW} using a 21x11 domain and using configuration $S + \Delta U_{300}$ (with T_{sfc}, RH_{700}, UTRH and ω_{300}). "Global mean local EVF" refers to the global mean EVF from local predictions, weighted by the cosine of each grid-cell's latitude. "Global EVF" refers to the EVF for each CCF's contribution to the globally-aggregated R_{LW}.~~

Our first three points involve the interaction between increasing model dimensions and the addition of potentially relevant context provided by the larger spatial domain. We discuss these points in more detail in Supplementary Material, Sect. S3. Addressing the last point, we note that

~~typically both (shown in Fig. Sxyz)(Fueglistaler, 2019)~~

~~Given that~~several studies point to thermodynamic changes dominating over dynamical effects for global cloud feedback, likely because dynamical effects cancel out at sufficiently large scales (Bony et al., 2004; Xu and Cheng, 2016; Byrne and Schneider, 2018). Conversely, thermodynamic and dynamical feedbacks have more comparable importance at more local scales. We find our results broadly analogous to this. The relatively large EVF for ω at local scales (17.9 %, the second highest in Fig. 4) explains why replacing ω_{500} with ω_{300} results in a positive shift to the skill distributions (Fig. 3a). In contrast, globally-aggregated EVF for ω_{300} is comparatively

Formatted: Caption

Formatted: Indent: First line: 1.27 cm

smaller (12.4 %, the lowest value in Fig. 4). This points to the cancellation of large-scale dynamically driven signals when globally-aggregated, thus explaining why there is little difference between the performance of ω_{300} and ω_{500} in Fig. 2c-d, despite resulting in a large improvement for a single CCF change at local scales. Finally, given that ω, an important "local" predictor, cancels out at globally-aggregated scales, the "non-local" predictors – such as T_{sfc} – contribute a larger proportion of the total predicted R_{LW}, thus explaining – at least in part – the discrepancy between globally-aggregated and local anomalies.

5.2 Predictive skill on CMIP models

Now, we briefly present results for the CCF configurations using the CMIP5/6 models. Key questions are whether the CCF approach performs similarly between models and observations, and if there are any obvious discrepancies that could point towards the analysis framework being less applicable than in observations. Performance metrics are first calculated locally for each GCM. Independently for each GCM, local metrics are meaned globally and in tropical ascent regions. The multi-model median result is then taken, with results analogous to ~~Fig. 2~~Fig. 2a-b shown in Fig. ~~Sxyz~~S7a-b. Finally, we ~~integrate~~ aggregate predictions (globally~~yy,~~ and in tropical ascent regions)~~,~~ independently for each GCM. The predicted global and tropical ascent-aggregated time series are compared against the similarly aggregated target values. Again, note a distinction between globally averaged, local performance *metrics*, and globally-*aggregated* R_{LW}~~.~~ throughout this discussion.

Figure 6~~5~~. Matrices showing the Pearson r score between the observed R_{LW} time series and predictions made at each domain size using different "CCF configurations". A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}. Each configuration uses T_{sfc}, RH_{700}, UTRH and ω_{300} (with the exception of the first column, where ω_{500} is used instead) and a candidate CCF(s) (e.g., S_{UT}), which is used to label each column. The median has been calculated from 14 of the CMIP models (excluding the additional 4 without CAPE or CIN). Predictions are made locally, with the RMSE averaged (a) globally and (b) in tropical ascent regions, defined as grid-cells with observed climatological EIS < 1 K, ω_{500} < 0 hPa s^{-1}. RMSE is weighted by the cosine of latitude and monthly standard deviation of R_{LW} of each grid-cell. Predictions are hence aggregated (c) globally and (d) in the tropical ascent regions and compared to similarly aggregated observations using Pearson r. Here, all predictions are made using ridge regression. Note different scales for each colorbar, with darker regions indicating higher skill (lower RMSE, higher Pearson r).

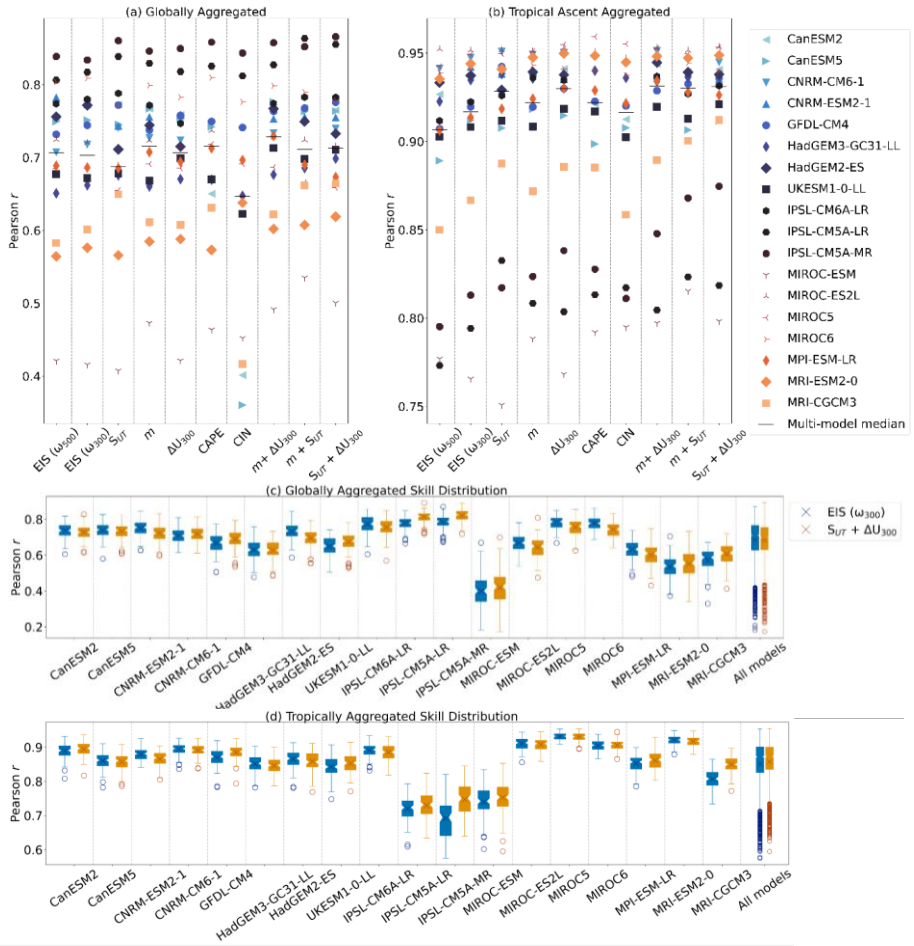

Figure 65. Pearson r scores for (a) globally- and (b) tropical ascent-aggregated predictions made at the 21x11 domain size using different "CCF configurations". A "CCF configuration" refers to the selection of cloud controlling factors used to predict R_{LW}. Each configuration uses T_{sfc}, T_{sfc}, RH_{700}, RH_{700}, UTRH and ω_{300} ω_{300} (with the exception of the first column, where ω_{500} ω_{500} is used instead) and a candidate CCF(s) (e.g., S_{UT}). The multi-model median Pearson r is shown from the 14 CMIP models where that have CAPE and CIN is availablecalculated. The bootstrapped (n = 100) predictive skill distributions for EIS (ω_{300} ω_{300}) and $S_{UT} + \Delta U_{300}$ are shown at the optimal 21x11 domain sizes for (c) globally-aggregated predictions and (d) tropical-ascent aggregated predictions. at the 21x11 domain. Global-mean local prediction skill distributions can be found in the Supplementary Material, Fig S xxx.

Formatted: Not Superscript/ Subscript

The CMIP Pearson r correlation matrices results for the performance metrics areare broadly analogous to the observationsobserved, where general patterns withinfound in Fig. 2 are also present in Fig. S7 are. We once again find a discrepancy in optimal domain size, with peak in local performance peaking at 7x3, and and predictive skill that increases with domain size for globally-aggregated R_{LW}, peaking at 21x11.. Differences include higher multi-model median skill metrics compared to the observations, which may be expected due to intrinsic knowledge of the meteorological conditions embedded within the CMIP models. Additionally, suppressed metrics for observed R_{LW} could be caused by slight mismatches between the observed radiative anomalies and the reanalysis meteorological variables. This therefore results in metrics that are more consistent

Field Code Changed

between CCF configurations than for the observed R_{LW}. In addition, smaller differences between configurations may in part be caused by higher metrics in the first place, leaving less room for improvement.

~~The spatial distributions of the performance metrics are qualitatively very similar to the observations, with the highest r and R^2 in the tropical ascent regions, and lowest over the Southern Ocean and subsidence regimes (see Fig. S5). While much of the Southern Ocean R^2 fell between 0.2 and 0.3 for the observations, the multi-model median R^2 generally exceeds 0.5 (though again, with outliers). This helps to support our attribution of lower predictive skill over the Southern Ocean in the observations, at least in part, to the known lower quality of reanalysis datasets in this region (Mallet et al., 2023).~~

~~For predictions of globally-aggregated R_{LW}, skill metrics are again higher for the CMIP models than for the observations, though performance similarly peaks at 21x11, and the general pattern of the correlation matrices shown in Fig. 5c-d is similar to Fig. 2c-d. Here, any configuration including S_{UT} predicts global R_{LW} with the highest correlation coefficient.~~We also find that CAPE performs comparatively better for the CMIP models than in the observations. This may be due to the way in which convection is parameterised in GCMs, thus resulting in stronger modelled relationships between cloud-radiative anomalies and CAPE than exist in the observations.

Highlighting uncertainties within the CMIP models themselves, there is a large spread in the skill metrics, shown for aggregated predictions in Figure 5a-b. We find that changes to the globally-aggregated performance do not imply similar changes to the tropical ascent-aggregated performance. For example, S_{UT} shows a slight decrease in the global multi-model median r compared to EIS (ω_{300}), despite showing a positive shift for predictions aggregated in the tropical ascent regions. Secondly, improvements to the multi-model median r do not imply that each GCM shows improvements independently. For example, the multi-model median r for tropical ascent-aggregated predictions made using configuration S_{UT} has improved compared to configuration EIS (ω_{300}); models such as MRI-CGCM3, GFDL-CM4 and IPSL-CM5A-MR have large leaps in Pearson r. Conversely, MIROC-ESM, CanESM5, and HadGEM2-ES show decreases. Opposing improvements and deteriorations of predictive skill is partially responsible for relatively small change in multi-model r between the configurations for the CMIP models.

In Section 5.1, we highlighted $S_{UT} + \Delta U_{300}$ as a possible optimal configuration. Here we identify whether differences between the CMIP-modelled predictive skill distributions for EIS (ω_{300}) and $S_{UT} + \Delta U_{300}$ are statistically significant. In a pairwise Kruskal-Wallis test on the combined Pearson r scores from all 18 models $(n = 1800)$, we find a significantly higher predictive skill distribution for $S_{UT} + \Delta U_{300}$ than EIS (ω_{300}) with $p < 10^{-11}$ (distributions not shown). This is unsurprising; 15 of the 18 individual CMIP models have a higher median r using $S_{UT} + \Delta U_{300}$ compared with EIS (ω_{300}).

Despite a slightly lower multi-model median, we find that the globally-aggregated distributions for all models combined are statistically similar at the 5 % significance level (shown in Fig. 5c, $p = 0.13$). Here, only half of the CMIP models have a higher median r using $S_{UT} + \Delta U_{300}$ compared with EIS (ω_{300}). However, visual inspection of the distributions for predictions aggregated in the tropical ascent regions (Fig. 5d) suggests that improvements found using $S_{UT} + \Delta U_{300}$ instead of EIS (ω_{300}) are more pronounced than any deteriorations. In summary, while the mean evolution of predictive skill within the CMIP models is broadly aligned with the observations, there are nuances which likely depend on the parameterization within the models themselves (Li et

al., 2012; Qu et al., 2014; Rio et al., 2019). This leads to a slightly different evolution of predictive skill with configuration between the CMIP models.

Formatted: Not Highlight

5.3 Physical interpretation of the cloud-radiative sensitivities

In addition to the statistical performance metrics shown above, we study the spatial distribution and magnitude of the sensitivities. Interpreting spatial sensitivities can be used in CCF analysis to justify predictor selection that is grounded in physical reasoning and can be done for any of the CCF configurations (e.g., Andersen et al., 2023). Though oOur analysis has identified two strong configurations,: $S_{UT} + \Delta U_{300}$ and $m + \Delta U_{300}$,. In this section, we only physically interpret the sensitivities of R_{LW} to the CCFs in configuration $S_{UT} + \Delta U_{300}$ in this section. using a 21x11 domain, derived from observations and CMIP models (shown in Fig. 6). We choose $S_{UT} + \Delta U_{300}$ over $m + \Delta U_{300}$ due based to on the wider literature on examining the relationship between high-cloud occurrence and static stability (e.g. Li et al., 2014) and due to the link between static stability and changes in tropical anvil cloud fraction through the "anvil iris" thermodynamic mechanism (Bony et al., 2016; Saint-Lu et al., 2020, 2022). , .Though we only show spatial sensitivities for one configuration here, We recommend similar physical interpretation of sensitivities be performed for should alternative configurations to be used in similar CCF applications, such as an constraining observational constraint on cloud feedback., we recommend similar physical interpretation of their sensitivities be performed.

For each CCF in the configuration, we sum each contribution Θ_i within the entire spatial domain (e.g., Eq. (5) for R_{LW}) and plot the total for each grid-cell.). This is the spatial sensitivity of the cloud-radiative anomaly to a given CCF, normalised for a one-standard deviation anomaly. Here, we derive the sensitivities using the full twenty-year datasets (with no dataset rotation or bootstrapping). There are several studies interpreting relationships between cloud-radiative anomalies and the core CCFs (e.g., CN21, Andersen et al., 2023), though not explicitly for high clouds. Therefore, we first briefly interpret our sensitivities to the core CCFs, shown in Fig.

~~6a~~Fig. 6a-d. We then assess the sensitivities for cloud properties (i.e., cloud top pressure and cloud fraction) before interpreting sensitivities for the additional CCFs, $S_{UT} + \Delta U_{300}$.

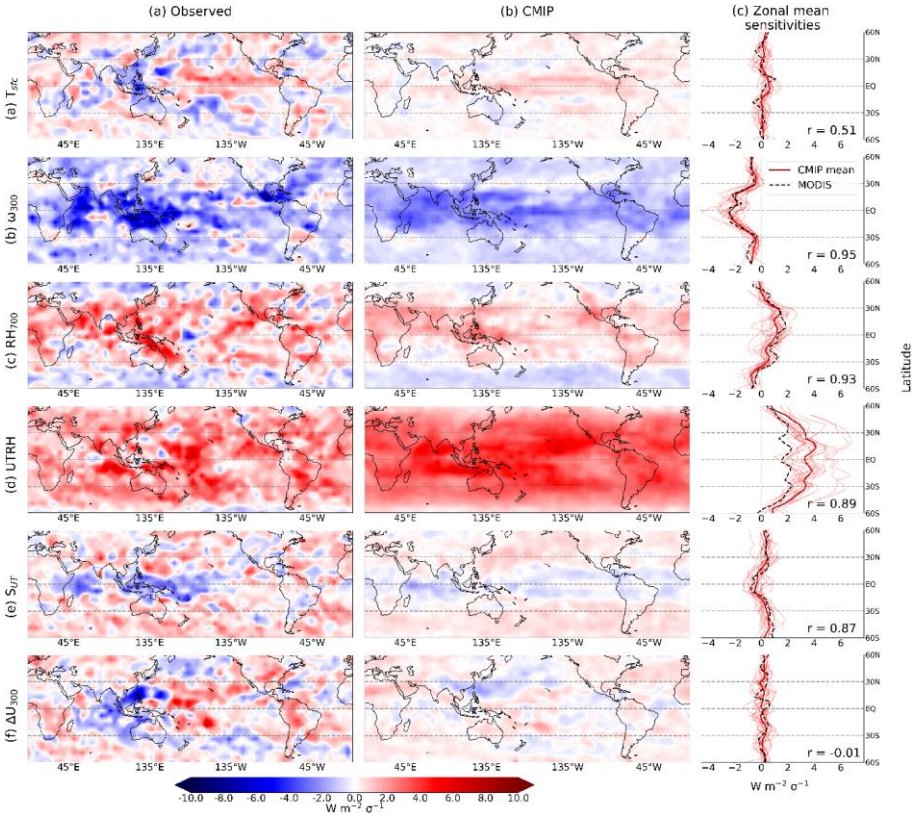

Figure ~~7.~~6. R_{LW} sensitivities ($\sum \theta_i$) to the cloud controlling factors in configuration $S_{UT} + \Delta U_{300}$ (also with T_{sfc} ~~T_{sfc}~~, RH_{700} ~~RH_{700}~~, UTRH and ω_{300} ~~ω_{300}~~), derived using a 21x11 domain and defined for a one-standard deviation anomaly in each CCF (scaled using ERA5 CCFs for visualisation purposes). To produce the maps, we sum all elements of the sensitivity vectors at each point r. The ~~left~~ column (a) shows observed sensitivities, and ~~the~~ column (b) ~~right~~ shows the multi-model mean. Column (c) shows zonal average sensitivity for the observations (dashed line), the multi-model mean (dark solid line) and individual CMIP model sensitivities. The Pearson r correlation coefficient for the zonal mean sensitivities is shown in the bottom corner of each panel.

The observed and multi-model mean spatial distributions for the core CCFs – T_{sfc}, ω_{300}, UTRH and RH_{700} – broadly align what we expect, and are qualitatively similar between the observations and multi-model means. We note that the observed global median regularization parameter, α, lies towards the upper-end of the inter-model spread~~, exceeded by only three models (CanESM5, CanESM2, and MIROC-ESM; see Fig. S8).~~ (not shown). We speculate that the CCFs in the CMIP models typically capture the variability in R_{LW} with greater skill than the observations, meaning less regularization is required on average. For all CCFs except UTRH, the magnitude of the modelled sensitivities are smaller than the observed (tropical ascent sensitivities~~not shown globally, though tropical ascent means are~~ shown in ~~Fig. 8~~Fig. 7). It is known that (CMIP5) GCMs underestimate the frequency of tropical anvil cloud and extratropical cirrus occurrence (Tsushima et al., 2013; Ceppi et al., 2017), and thus their radiative effects which can also ~~may be responsible for~~explain smaller sensitivities on average. ~~We~~

~~note that the observed global median regularization parameter, α, lies towards the upper end of the inter-model spread, exceeded by only three models (CanESM5, CanESM2, and MIROC-ESM; see Fig. S6). We speculate that the CCFs in the CMIP models typically capture the variability in R_{LW} with greater skill than the observations, thus requiring less regularization. It is also known that (CMIP5) GCMs underestimate the frequency of tropical anvil cloud and extratropical cirrus occurrence (Tsushima et al., 2013; Ceppi et al., 2017), and thus their radiative effects. This may also be responsible for smaller sensitivities on average. We also note that sensitivities do not explicitly imply causality, as certain relationships are known to be bidirectional (e.g., R_{LW} and UTRH).~~

The R_{LW} ~~–~~ T_{sfc} sensitivities (i.e., Eq. (5) summed for all $X = T_{sfc}$)~~,~~ shown in ~~Fig. 7a~~Fig. 6a, are generally small in magnitude, with regions of positive sensitivity in the Central and East Pacific (responsible for a slight positive peak in observed zonal mean sensitivity), and with negative (or, for the CMIP models, ~~a~~ reduced magnitude) sensitivity over the Maritime continent. ~~-~~The R_{LW} - UTRH sensitivities are ubiquitously positive~~,~~ and large in magnitude, consistent with increasing high-cloudiness with humidity, though ~~Fig. 7d~~ Fig. 6d ~~suggests~~ shows that CMIP-modelled sensitivities are consistently larger in magnitude than is observed. ~~We speculate that this~~This is possibly ~~may be~~ due to stronger coupling between upper-tropospheric humidity and cloud incidence in the CMIP models than in the observations, ~~(perhaps~~ owing to the parameterization of clouds in the models themselves (Li et al., 2012; Qu et al., 2014)~~)~~. The RH_{700} sensitivities are also widely positive (though negative at high latitudes), with smaller magnitude than UTRH (as we would expect for high clouds) ~~and~~ with the largest magnitudes in the deep tropics. ~~-~~Indicating increased high-cloudiness with increased ascent, the ω_{300} sensitivities are near-ubiquitously negative, with the strongest magnitudes broadly aligning with the tropical ascent regions in both observations and the CMIP models.

We ~~can also~~ use the decomposition of R_{LW} into its linear sum of contributions from changes in cloud top pressure (CTP), cloud fraction (CF), optical depth, and a small residual (with other components held fixed), to further interpret our sensitivities (Zelinka et al., 2012a, b, 2016). We do not show optical depth sensitivities here, owing to their small role in driving LW high-cloud radiative anomalies (see Fig. ~~S9~~S10~~1~~). LW radiative anomalies caused by changes in the cloud properties are henceforth referred to using an additional subscript, i.e., $R_{LW,CTP}$ is the contribution that changes in cloud top pressure (with no change in τ or CF) have on the total R_{LW}. Sensitivities for the decompositions can be found in the Fig. S~~89~~ S~~9107~~. We average the domain-summed sensitivities in the tropical ascent regions, shown in Fig. 7.~~Fig. 8.~~

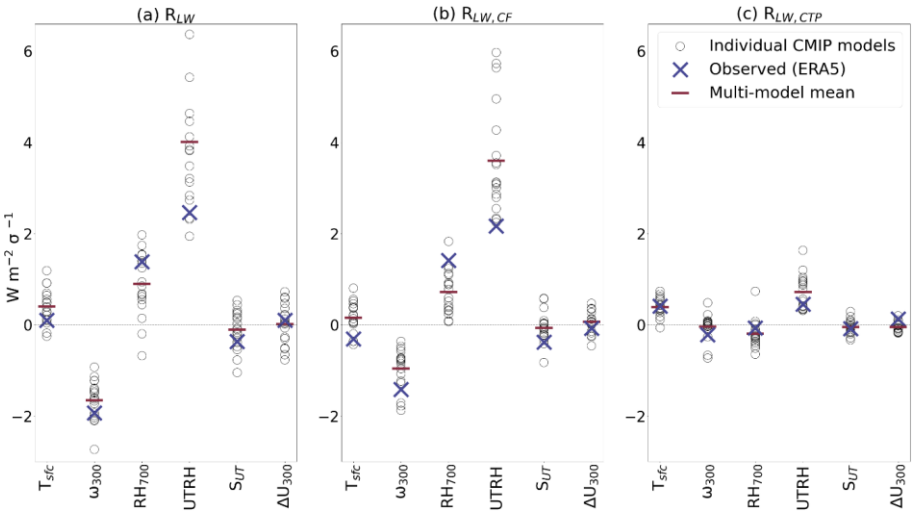

Figure 87. Observed and CMIP sensitivities to the cloud controlling factors in configuration $S_{UT} + \Delta U_{300}$ (with T_{sfc}, ~~T_{sfc}~~ RH_{700}~~RH$_{700}$~~, UTRH and ω_{300}~~(ω_{300})~~), derived using a 21x11 domain and defined for a one-standard deviation anomaly for each CCF, averaged over all tropical ascent grid-cells for (a) R_{LW}, (b) $R_{LW,CF}$, and (c) $R_{LW,CTP}$. The standard deviation used to scale each CCF has been calculated from the observed CCFs.

The R_{LW} - T_{sfc} ~~T_{sfc}~~ sensitivities average to approximately zero in the tropical ascent regions for ~~both~~ the observations ~~and CMIP models~~ (~~Fig. 8a~~Fig. 7a) with good agreement globally between the CMIP models and observations (zonal-mean $r = 0.52$). ~~However,~~ Fig~~-~~ure ~~7a~~ 6a shows a distinct positive sensitivity present over the Pacific Ocean, which we ascribe to an increase in high-cloud top pressure that is associated with warming sea surface temperature anomalies, thus radiating heat to space at cooler temperatures. We find that the spatial pattern of $R_{LW,CTP}$ - T_{sfc} sensitivities in the tropics are widespread positive (Fig. S~~9~~10), as we would expect (though more strongly positive in the models than the observations). ~~Indeed, we find that the spatial pattern of $R_{LW,CTP}$ - T_{sfc} sensitivities in this area are predominantly positive (Fig. 8c, Fig. S7), as we would expect.~~ Accordingly, the observed ~~mean~~ $R_{LW,CTP}$ - T_{sfc} sensitivit~~ies~~y in the tropical ascent regions ~~is~~ are positive, with larger magnitude than the similarly averaged and opposite-signed $R_{LW,CF}$ - T_{sfc} sensitivities (~~Fig. 8~~Fig. 7). This is despite a much smaller monthly signal for observed $R_{LW,CTP}$ than $R_{LW,CF}$. The modelled $R_{LW,CTP}$ - T_{sfc} sensitivities are stronger than the $R_{LW,CF}$ - T_{sfc} counterparts, resulting in the slightly more positive CMIP R_{LW} - T_{sfc} sensitivities.

The R_{LW} - ΔU_{300} sensitivity, shown in Fig. ~~7f~~6f, is more challenging to interpret than the core CCFs (Anber et al., 2014). This is partially due to the dynamic nature of wind shear; coefficients within the spatial domain capture dynamic variability signals, which may result in a range of positive and negative sensitivities, therefore cancelling in the summation over the 21x11 domain. ~~Nonetheless, we suggest reasons for both positive and negative sensitivity.~~ There is also less agreement between the observed and multi-model mean spatial distributions than all other CCFs, which ~~we speculate~~ may partially be caused by offset circulation cells in the CMIP models, resulting in different local sensitivities and dynamic signals (zonal-mean $r = -0.01$). ~~Nonetheless, w~~Here, we suggest reasons for both positive and negative sensitivity. Over the Maritime Continent and Indian Ocean, observed sensitivities are broadly negative. It is known that wind shear can hasten the dissipation of tropical tropopause cirrus (Jensen et al., 2011) which would result in decreased cloudiness, and thus

LW cooling. Conversely, there are many regions where the sensitivity is positive (such as the Central Pacific) which indicates LW warming with increased shear. ~~We speculate that this~~This ~~is~~ may be a result of shear spreading the high cloud, thus increasing cloud fraction (Lin and Mapes, 2004), and in turn reducing outgoing LW radiation. The role of wind shear may be sensitive to the pressure level relative to the tropopause (Chakraborty et al., 2016; Nelson et al., 2022). Given that we use the same shear height (i.e., the difference in 300 hPa and 925 hPa wind speeds) globally, ~~we speculate the~~it is likely that the zonal distribution of tropopause heights may cause the differing relationships. Despite differences between the spatial distributions, both observed and the multi-model mean sensitivities in the tropical ascent regions are consistent with each other, and average to approximately zero ~~(Fig. 8). , which w~~We ~~might~~ suspect this may due to shear being~~given that shear is~~ important for the organisation of convection, which is not represented in GCMs.

Finally, we address the S_{UT} sensitivities. Both the observed and multi-model mean R_{LW} - S_{UT} sensitivities, shown in ~~Fig. 7e~~Fig. 6e, are predominantly negative in the tropics, with largest magnitudes over the Central and West Pacific, and Maritime Continent (though more markedly so for the observations). Therefore, in the absence of changes in the other CCFs, anomalies in high cloud associated with an increase in S_{UT} would result in increased longwave emission to space over the tropics. This is what we expect, given the negative relationship between upper-tropospheric cloud incidence and static stability over tropical oceans (Li et al., 2014).

The observed $R_{LW,CF}$ - S_{UT} sensitivities are ~~also~~ negative across the tropics, most strongly in regions with high $R_{LW,CF}$ signals (see Fig. ~~S9 and S11, respectively~~S87). This ~~,~~ reveal~~ings~~ that LW cooling arises – at least in part – from a reduction in high-cloud fraction~~, , particularly in observations. This~~ qualitatively res~~e~~emblinge~~s~~ the anvil iris mechanism (Bony et al., 2016; Saint-Lu et al., 2020). As anvil clouds rise in response to global warming, their environment becomes more stable, owing to the dependency of static stability on atmospheric pressure (Saint-Lu et al., 2020, 2022). In a more stable atmosphere, the vertical pressure gradient associated with subsidence in clear-sky conditions is reduced. Owing to mass conservation, a reduction in the subsidence pressure gradient results in a reduction in anvil cloud fraction, caused by a decrease in horizontal convergence (Saint-Lu et al., 2020, 2022).

~~Though it is thought to be small in magnitude, the anvil cloud area feedback is subject to much uncertainty and considered to be underestimated by GCMs (Zelinka et al., 2022). Consistent with this, w~~We find that the ~~magnitude~~ mean~~of the~~ CMIP R_{LW} - S_{UT} ~~sensitivities~~ sensitivity in the tropical ascent regions ~~is substantially~~ smaller in magnitude than the observed, with considerable disagreement in sign (ranging from -1.0 to 0.44 Wm σ $^{-1}$; ~~(Fig. 8a~~Fig. 7a). ~~Indeed, the observed tropical ascent mean R_{LW} - S_{UT} sensitivity lies below the range of the individual CMIP models' sensitivities.~~ Most ~~A large fraction~~ of the total R_{LW} - S_{UT} sensitivity arises from the CF component (Fig. ~~8b~~7b), ~~which~~ where the ~~is consistently underestimated by the~~ CMIP model-mean approaches zero, though has a similarly large range.~~s.~~ Though it is thought to be small in magnitude, the anvil cloud area feedback is subject to much uncertainty and underestimated by GCMs (Zelinka et al., 2022), consistent with our results.~~This is unsurprising;~~ CMIP models ~~have been shown~~are known to ~~underestimate~~ predict a wide range of anvil cloud fraction ~~the negative anvil cloud fraction~~ feedbacks, including "unlikely" very positive feedback (Zelinka et al., 2022) which is perhaps reflected by strong positive tropical R_{LW} - S_{UT} sensitivities for two ~~models~~GCMs (Fig. S8~~e9~~). Given that static stability has been shown to robustly control high-cloud fraction (Saint-Lu et al., 2022), and based on our results, we ~~therefore speculate~~propose that the addition of S_{UT} into observational constraint frameworks may reduce some of the uncertainty arising from the anvil fraction feedback.

We also find that the spatial distributions for the $R_{LW,CTP}$ - S_{UT} observed and multi-model mean sensitivities are broadly similar with zonal-mean correlation $r = 0.60$~~similar to each other, mostly negative, and largest in magnitude in the tropics~~. For observations, the $R_{LW,CTP}$ - S_{UT} sensitivity is negative in the West Pacific and Maritime Continent, ~~This suggests~~indicating that an increase in S_{UT} results in LW cooling, arising from a change (i.e., a decrease) in cloud top pressure. Increased static stability results in suppressed vertical motion, which in turn prevents cloud tops from rising as high as they might in a more unstable environment (Zelinka and Hartmann, 2010, 2011; Saint-Lu et al., 2022). ~~Though~~ Negative sensitivities in the tropical ascent regions are less prevalent for the models, with negative sensitivities more widespread in the subtropics. ~~the spatial distributions are similar, the~~This results in a smaller magnitude of the CMIP $R_{LW,CTP}$ - S_{UT} sensitivities ~~are once again smaller than the observed~~ in the tropical ascent regions (~~Fig. 8c~~Fig. 7c).

As well as absorbing upwelling LW radiation, high clouds can reflect incident SW radiation depending on their optical depth. While ~~the~~ R_{LW} (and thus the sensitivities) is primarily driven by CF and CTP changes, R_{NET} is also driven by changes in optical depth, which predominantly affects SW radiative anomalies that we have not directly assessed. Thus, the net high-cloud radiative anomaly is comprised of complex interplay between competing LW and SW effects. ~~However, we note that the~~The magnitude of the observed R_{NET} ~~–~~ S_{UT} sensitivity is much smaller ~~in magnitude~~ than the $R_{LW} - S_{UT}$ component in the tropical ascent regions, though the spatial distribution is broadly similar (Fig. ~~S8~~S1~~2~~1), and negative in many high-cloud regions. This suggests that, assuming an increase in S_{UT} with warming (Bony et al., 2016),, high-clouds exert a negative (though weak) net feedback. ~~assuming an increase in S_{UT} with warming (Bony et al., 2016),~~ However, ~~the~~ the observed CMIP-mean tropical ascent $R_{NET,CF} - S_{UT}$ sensitivities average to approximately zero, indicating a very weak anvil cloud area feedback with increasing S_{UT}. While a weak anvil cloud feedback may be expected (McKim et al., 2024), it is also thought CMIP models tend to underestimate a negative anvil cloud fraction feedback (Zelinka et al., 2022). ~~This again reflects the known underestimation of the negative anvil cloud fraction feedback in CMIP models (Zelinka et al., 2022).~~ Additionally, Zelinka et al. (2022) show that eight CMIP models (including six of those used in this research) predict an “unlikely” positive feedback arising from changes in anvil cloud fraction. Therefore, the near-zero multi-model sensitivities may also arise due to cancellation of local sensitivities between the models~~.~~

.

5.4 Predicting radiative anomalies from cloud fraction and cloud top pressure changes

Based on the physical interpretation of the sensitivities, our results – combined with previous literature and theory – thus far support the use of $S_{UT} + \Delta U_{300}$ in high-cloud controlling factor frameworks. We have shown that $S_{UT} + \Delta U_{300}$ reproduces the ~~globally~~local and globally-aggregated ~~-aggregated~~ R_{LW} ~~time series with~~ with the ~~high~~est skill ~~in~~ for ~~both~~ observations, and performs well for the CMIP models~~,~~. Additionally, ~~t~~and ~~t~~he sensitivities shown in ~~Fig. 7~~ Fig. 6 suggest that the mechanisms driving high-cloud feedback – rising free-tropospheric clouds and reduction in anvil cloud fraction (Ceppi et al., 2017) – are captured by this selection of CCFs at the 21x11 domain. ~~Therefore, we f~~Finally, we question whether our approach captures the spatial pattern, temporal variability, and magnitude of these properties.

We predict twenty years of cloud-radiative anomalies induced by CF and CTP changes (with other components held fixed) for both observations and CMIP models, again using rotating eighteen-year datasets. We ~~The monthly radiative anomalies are~~globally aggregate~~d~~ the predicted anomalies ~~globally and in tropical ascent regions~~ (e.g., as in Fig~~s~~. 2c ~~and 4, c-d~~) and compare~~d~~ against similarly aggregated target values using the Pearson r correlation coefficient. ~~(to ensure trends are captured by our framework).~~ We do not assess optical depth-induced changes, owing to their small historical LW signal (see Fig. ~~S9~~S11~~0~~). Though optical depth is important for historical SW (and consequently, net) radiative anomalies, the high-cloud optical depth feedback is thought to be relatively small (Zelinka et al., 2022) and so we focus on CF and CTP. We place particular emphasis on the observations here, as the El Niño phase of the El Niño-Southern Oscillation (ENSO) from July 2015 to June 2016 saw anomalous warming in the East Pacific (see Fig. 8, ~~t~~op panel). ENSO is a dominant driver of natural ocean-atmosphere variability, resulting in regional tropical temperature and circulation anomalies that are accompanied by changes in cloud properties and the TOA radiation budget (Ceppi and Fueglistaler, 2021). Accordingly, July 2015 to June 2016 has one of the most anomalously warm annual mean surface temperatures in the 20-year record. We only highlight this El Niño event for the observed cloud properties, as it will be absent from the coupled historical simulations, and AMIP simulations do not reach 2016.

Out-of-sample globally-aggregated $R_{LW,CF}$ is well predicted ~~well by using configuration the~~ $S_{UT} + \Delta U_{300}$~~, configuration (Fig. 10a)~~ with a correlation coefficient of 0.63.~~.~~ The spatial distribution of the El Niño CF anomalies closely follows the R_{LW} distribution owing to its large signal, and ~~(shown in Fig. 9)~~ is ~~also well~~ reproduced accurately ~~,~~ with a correlation coefficient of r = 0.93 (Fig. 8). There is a positive $R_{LW,CF}$ anomaly in the East Pacific, overlapping the region of anomalous sea surface warming, indicating increased cloud fraction. Warmer SSTs enhance convection, resulting in increased upward motion, and thus increased high cloudiness. In the West Pacific, the SST anomaly is negative and smaller in magnitude, though there is a strong, negative $R_{LW,CF}$ anomaly, indicating a reduction in cloud fraction. Owing to the shift in circulation, suppressed convection can result in anomalous subsidence, hence reducing high cloudiness. Our configuration predicts $R_{LW,CF}$ with slight negative error in the East Pacific, indicating an underestimation of the increased cloud fraction. In the West Pacific, where there is an observed reduction in cloud fraction, our predictions have little error.

Field Code Changed

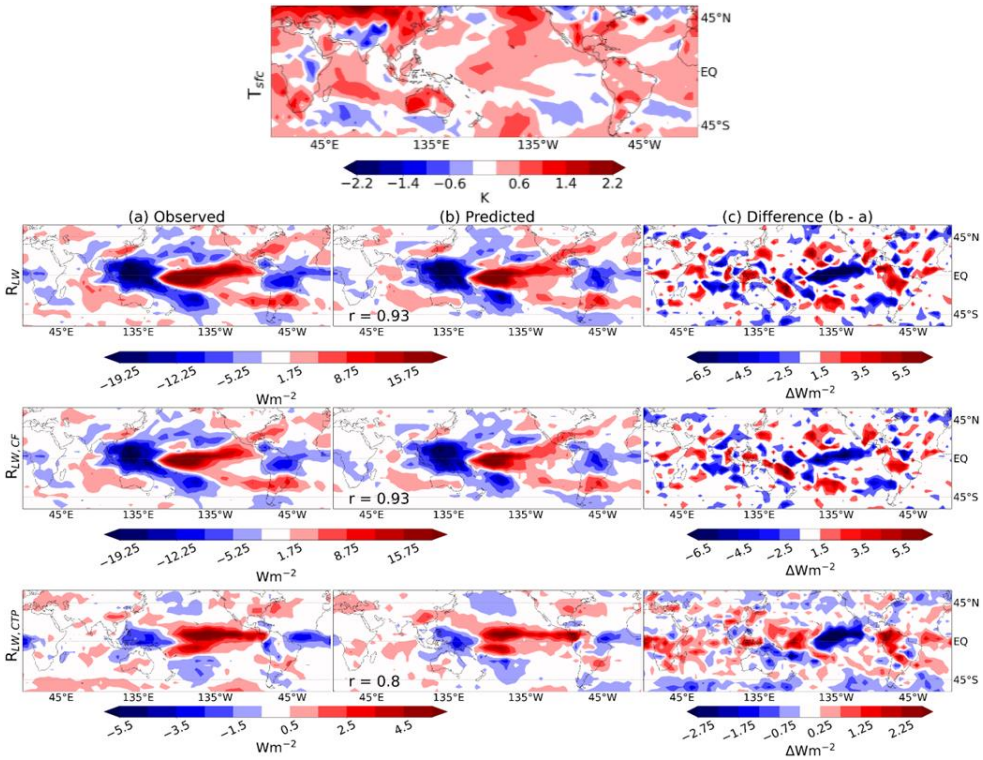

Figure ~~9~~8. Observed mean El Niño surface temperature anomaly (top) and radiative anomalies (panel (a)), averaged from July 2016 – June 2015 relative to the full twenty-year record. Predicted anomalies (panel (b)) made using a 21x11 domain and the configuration $S_{UT} + \Delta U_{300}$ (with T_{sfc}, RH_{700}, UTRH and ω_{300}) for the El Niño months. The difference (predicted – observed) is shown in the panel (c). The Pearson r spatial correlation between (a) and (b) is shown in the bottom left of panel (b). Note different colorbar ranges.

We also predict observed, globally-aggregated $R_{LW,CTP}$ well, though with slightly reduced correlation coefficients compared to $R_{LW,CF}$ (r = 0.68).~~.~~ Figure 9b ~~Figure 10b~~ shows that the magnitudes of strong positive and negative anomalies are slightly underestimated; this may be caused by a small signal for the regression model to learn from (see Fig. ~~S9~~S10~~1~~). ~~This may a~~Alternatively, this may hint towards a non-linear relationship between cloud top pressure and the CCFs, which would not be captured by ridge regression, and (in addition to CCF selection) could explain poorer skill for predicting LW

cloud feedback (CN21). Regardless, the spatial distribution of predicted El Niño $R_{LW,CTP}$ is again strongly correlated to the observed, here with r = 0.801 (shown in Figure 9a Fig. 9). Strong positive anomalies are present over the East Pacific, which we ascribe to a rise in cloud top pressure due to enhanced convection. As the atmosphere warms, a shift of the $R_{LW,CTP}$ distribution towards higher values, particularly in the tropics, may be expected owing to the rising of free-tropospheric clouds (Ceppi et al., 2017). We note that the globally-aggregated, annual mean $R_{LW,CTP}$ during this El Niño event is most extreme, positive anomaly in the observed twenty-year record, and is reproduced with small absolute error (-0.003 01 Wm σ^{-2}). Accordingly, we predict the most positive $R_{NET,CTP}$ annual anomaly with similarly small absolute error and correlation coefficient (absolute error -0.01 Wm^{-2}, $r = 0.57$; see Fig. S12). This is consistent with the extreme warmth during that period, and the associated rise of the tropopause. Despite potentially underestimating the amplitude of the monthly variability, our method does an excellent job capturing the most extreme (positive) positive annual anomaly out-of-sample (globally-aggregated shown in Fig. 9b Fig. 10; for tropical ascent aggregation, see Fig. S10). We also find that configurations S_{UT} and

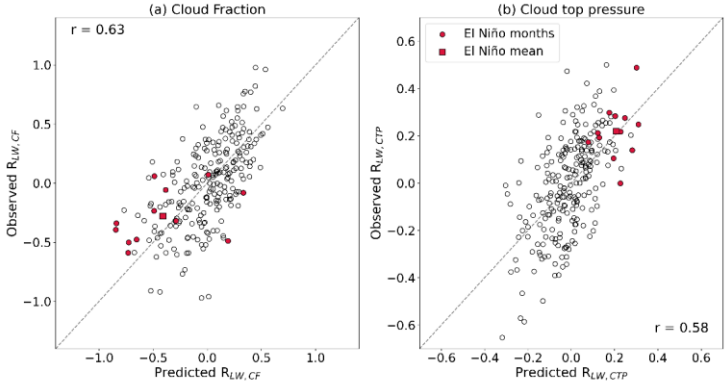

Figure 109. Scatter plot showing the correlation between observed and predicted monthly globally-aggregated (a) $R_{LW,CF}$ and (b) $R_{LW,CTP}$ time series using configuration $S_{UT} + \Delta U_{300}$ (in addition to T_{sfc}, T_{sfc}, RH_{700} RH$_{700}$, UTRH, and ω_{300}) and a 21x11 domain. El Niño months are shown using coloured circles, with the annual mean shown using a coloured square. Solid lines show y = x, and the dashed lines show the line-of-best fit through the points. For $R_{NET,CF}$ and $R_{NET,CTP}$, see Fig. S10S12.

$S_{UT} + \Delta U_{300}$ predicts the tropical mean El Niño $R_{LW,CTP}$ with the smallest absolute error (not shown).

We additionally confirm that our framework predicts out-of-sample globally-aggregated $R_{LW,CF}$ and $R_{LW,CTP}$ with good skill in the CMIP models, once again with slightly higher correlation coefficients than the observed (multi-model medians of r = 0.750.81 and 0.7877, respectively; see Fig. S11). To summarise, Though we only show results for the $S_{UT} + \Delta U_{300}$ configuration here, we have additionally assessed each configuration's ability to capture both $R_{LW,CF}$ and $R_{LW,CTP}$ for each month during the 2015 – 2016 El Niño (see coloured circles in Fig. 10 scatterplots). wWe have additionally shown that the spatial distribution of the observed El Niño anomalies are captured well, including the most extreme positive $R_{LW,CTP}$ annual anomaly, thus highlighting the strength of our proposed configuration $S_{UT} + \Delta U_{300}$. , in particular, as a high-cloud controlling factor. In a comparison of the El Niño months alone, we find that any configuration including S_{UT} reproduces tropically averaged $R_{LW,CTP}$ and $R_{LW,CF}$ anomalies with stronger positive correlations than alternative CCFs (not shown). We reiterate that the S_{UT} sensitivities (Fig. 7e Fig. 6e, Fig. S89e-S910e7) are physically congruent with previous literature, and appear to directly target the drivers of high-cloud feedback. We have additionally shown that the spatial distribution of the observed El Niño anomalies are captured well, including the extreme positive $R_{LW,CTP}$ anomaly, thus highlighting the strength of S_{UT}, in particular, as a high-cloud controlling factor.

6 Conclusion

Few studies directly assess cloud controlling factors for high clouds despite their substantial contributions to cloud feedback. Here, a~~A~~ selection of candidate cloud controlling factors (CCFs) ~~has~~ have been used to predict high-cloud radiative anomalies using ridge regression. We investigate five candidate CCFs: static stability in the upper troposphere, sub-cloud moist static energy, wind shear, convective available potential energy and convective inhibition, using the additional "core" meteorological drivers surface temperature, lower- and upper-tropospheric relative humidity, and upper-tropospheric vertical pressure velocity in each configuration. CCFs are used within a two-dimensional spatial domain to predict out-of-sample longwave cloud-radiative anomalies, R_{LW}. We assess configurations from local to globally-aggregated spatial scales, and physically interpret the spatial distribution of the sensitivities for the configuration $S_{UT} + \Delta U_{300}$. Finally, we assess the skill of $S_{UT} + \Delta U_{300}$ for predicting out-of-sample anomalies induced by changes in cloud top pressure and cloud fraction, including the El Niño event of 2015 – 2016.

We find that the optimal domain size and CCF combination is dependent on the temporal and spatial scales assessed, and we summarise the most relevant findings here:

1. All configurations predict out-of-sample historical variability for both R_{LW} (and R_{NET}) ~~anomalies~~ with good skill for observations and CMIP models at local scales. A domain of 7x3 optimises local predictions, where we show that ridge regression skill surpasses traditional multiple linear regression.
2. Converse to local predictions, predictive skill for globally-aggregated (i.e., a global spatial average) radiative anomalies *increases* with domain size, peaking at 21x11. We suggest a trade-off between local and non-local predictors is partially responsible for this domain size discrepancy between local and global predictions, though u~~U~~nravelling ~~this domain size discrepancy between local and global predictions remains~~this remains a key question for future research~~. Differences between the configurations are more pronounced at global scales, and based on these results we identify $S_{UT} + \Delta U_{300}$ as an optimal configuration that performs well at all scales.~~
2. ~~~~
3. The main mechanisms driving high-cloud feedback – rising of free-tropospheric clouds and reduction of anvil cloud fraction – appear to be captured by the core and candidate CCFs ~~sensitivities~~ in the $S_{UT} + \Delta U_{300}$ configuration. The spatial distributions of the R_{LW} sensitivities to the core CCFs and S_{UT} are physically consistent with our understanding and expectations, with observed and CMIP-modelled sensitivities qualitatively similar. There are larger differences between observed and the multi-model mean ΔU_{300} sensitivities, which are more complex to interpret than the core CCFs and S_{UT}.
4. Out-of-sample globally-aggregated anomalies induced by cloud top pressure and cloud fraction changes are predicted well using $S_{UT} + \Delta U_{300}$, in both observations and models. In particular, we obtain a quantitatively accurate out-of-sample prediction of the observed extreme anomalies in R_{LW}, $R_{LW,CF}$ and $R_{LW,CTP}$ during the 2015 – 2016 El Niño. The corresponding spatial distributions are also predicted with high correlation coefficients (r ~~>~~ 0.80).

Our systematic evaluation of high-cloud controlling factors highlights $S_{UT} + \Delta U_{300}$ as ~~an~~ a possible ~~important addition to C~~optimal configuration for CCF frameworks. Of course, our work is only the first attempt to assess candidates for high-cloud controlling factors so we welcome future work on additional candidate factors that might not have been considered here. We have also identified an important inconsistency regarding ideal domain size for CCF predictions on historical data locally, and globally aggregated. Given the strong out-of-sample predictive power of our framework, in future work we will use our optimal CCF configurations to constrain high-cloud feedback.

Data availability

ERA5 meteorological reanalysis data is freely available from the Copernicus Climate Change Service (C3S) Climate Data Store (Hersbach, H., *et al*., 2023a, DOI: 10.24381/cds.f17050d7, Hersbach, H., *et al*., 2023b, DOI: 10.24381/cds.6860a573 and Hersbach, H., *et al*., 2023c, DOI: 10.24381/cds.bd0915c6). Combined MODIS Aqua/Terra data are also freely available and downloaded monthly (Bodas-Salcedo *et al.*, 2011, DOI: 10.5067/MODIS/MCD06COSP_M3_MODIS.062). All CMIP5/6 data were obtained from the UK Center for Environmental Data Analysis portal (https://esgf-index1.ceda.ac.uk/search/cmip6-ceda/).

Acknowledgements

Firstly, we would like to thank the two anonymous reviewers for their thoughtful and helpful comments, which have improved this manuscript. S.W.K., P.N., P.C. and P.S. were supported through the UK Natural Research Environment Research Council (NERC) grant number NE/V012045/1. H.A., J.C. and P.S. have received funding from the European Union's Horizon 2020 research and innovation program under grant agreement no. 821205 (FORCeS) and H.A. and J.C. from the Deutsche Forschungsgemeinschaft (DFG) in the project Constraining Aerosol–Low cloud InteractionS with multi-target MAchine learning (CALISMA) under project number 440521482. This research was carried out on the High Performance Computing Cluster supported by the Research and Specialist Computing Support service at the University of East Anglia and additionally JASMIN, the UK's collaborative data analysis environment (https://jasmin.ac.uk).

Author contributions. S.W.K., P.N., and P.C., conceptualised this research. S.W.K. conducted the analysis and wrote the article with contributions from all authors on the text and interpretation of the results.

Competing interests. At least one of the (co-)authors is a member of the editorial board of Atmospheric Chemistry and Physics.

References

Anber, U., Wang, S., and Sobel, A.: Response of Atmospheric Convection to Vertical Wind Shear: Cloud-System-Resolving Simulations with Parameterized Large-Scale Circulation. Part I: Specified Radiative Cooling, Journal of the Atmospheric Sciences, 71, 2976–2993, https://doi.org/10.1175/JAS-D-13-0320.1, 2014.

Andersen, H., Cermak, J., Fuchs, J., Knutti, R., and Lohmann, U.: Understanding the drivers of marine liquid-water cloud occurrence and properties with global observations using neural networks, Atmospheric Chemistry and Physics, 17, 9535–9546, https://doi.org/10.5194/acp-17-9535-2017, 2017.

Andersen, H., Cermak, J., Fuchs, J., Knippertz, P., Gaetani, M., Quinting, J., Sippel, S., and Vogt, R.: Synoptic-scale controls of fog and low-cloud variability in the Namib Desert, Atmospheric Chemistry and Physics, 20, 3415–3438, https://doi.org/10.5194/acp-20-3415-2020, 2020.

Andersen, H., Cermak, J., Zipfel, L., and Myers, T. A.: Attribution of Observed Recent Decrease in Low Clouds Over the Northeastern Pacific to Cloud-Controlling Factors, Geophysical Research Letters, 49, e2021GL096498, https://doi.org/10.1029/2021GL096498, 2022.

Andersen, H., Cermak, J., Douglas, A., Myers, T. A., Nowack, P., Stier, P., Wall, C. J., and Wilson Kemsley, S.: Sensitivities of cloud radiative effects to large-scale meteorology and aerosols from global observations, Atmospheric Chemistry and Physics, 23, 10775–10794, https://doi.org/10.5194/acp-23-10775-2023, 2023.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J.-L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O.: COSP: Satellite simulation software for model assessment, Bulletin of the American Meteorological Society, 92, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

Bony, S., Lau, K.-M., and Sud, Y. C.: Sea Surface Temperature and Large-Scale Circulation Influences on Tropical Greenhouse Effect and Cloud Radiative Forcing, Journal of Climate, 10, 2055–2077, https://doi.org/10.1175/1520-0442(1997)010<2055:SSTALS>2.0.CO;2, 1997.

Bony, S., Dufresne, J.-L., Le Treut, H., Morcrette, J.-J., and Senior, C.: On dynamic and thermodynamic components of cloud changes, Climate Dynamics, 22, 71–86, https://doi.org/10.1007/s00382-003-0369-6, 2004.

Bony, S., Stevens, B., Coppin, D., Becker, T., Reed, K. A., Voigt, A., and Medeiros, B.: Thermodynamic control of anvil cloud amount, Proceedings of the National Academy of Sciences, 113, 8927–8932, https://doi.org/10.1073/pnas.1601472113, 2016.

Bretherton, C. S.: Insights into low-latitude cloud feedbacks from high-resolution models, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 373, 20140415, https://doi.org/10.1098/rsta.2014.0415, 2015.

Brient, F. and Schneider, T.: Constraints on Climate Sensitivity from Space-Based Measurements of Low-Cloud Reflection, Journal of Climate, 29, 5821–5835, https://doi.org/10.1175/JCLI-D-15-0897.1, 2016.

Byrne, M. P. and Schneider, T.: Atmospheric Dynamics Feedback: Concept, Simulations, and Climate Implications, Journal of Climate, 31, 3249–3264, https://doi.org/10.1175/JCLI-D-17-0470.1, 2018.

Ceppi, P. and Nowack, P.: Observational evidence that cloud feedback amplifies global warming, Proc. Natl. Acad. Sci. U.S.A., 118, e2026290118, https://doi.org/10.1073/pnas.2026290118, 2021.

Ceppi, P., Brient, F., Zelinka, M. D., and Hartmann, D. L.: Cloud feedback mechanisms and their representation in global climate models, WIREs Climate Change, 8, e465, https://doi.org/10.1002/wcc.465, 2017.

Chakraborty, S., Fu, R., Massie, S. T., and Stephens, G.: Relative influence of meteorological conditions and aerosols on the lifetime of mesoscale convective systems, Proceedings of the National Academy of Sciences, 113, 7426–7431, https://doi.org/10.1073/pnas.1601935113, 2016.

Chen, T., Rossow, W. B., and Zhang, Y.: Radiative Effects of Cloud-Type Variations, Journal of Climate, 13, 264–286, https://doi.org/10.1175/1520-0442(2000)013<0264:REOCTV>2.0.CO;2, 2000.

Davison, A. C. and Hinkley, D. V.: Bootstrap Methods and Their Application, Cambridge University Press, 606 pp., 1997.

Donner, L. J. and Phillips, V. T.: Boundary layer control on convective available potential energy: Implications for cumulus parameterization, Journal of Geophysical Research: Atmospheres, 108, https://doi.org/10.1029/2003JD003773, 2003.

Fuchs, J., Cermak, J., and Andersen, H.: Building a cloud in the southeast Atlantic: understanding low-cloud controls based on satellite observations with machine learning, Atmospheric Chemistry and Physics, 18, 16537–16552, https://doi.org/10.5194/acp-18-16537-2018, 2018.

Fueglistaler, S.: Observational Evidence for Two Modes of Coupling Between Sea Surface Temperatures, Tropospheric Temperature Profile, and Shortwave Cloud Radiative Effect in the Tropics, Geophysical Research Letters, 46, 9890–9898, https://doi.org/10.1029/2019GL083990, 2019.

Gasparini, B., Sullivan, S. C., Sokol, A. B., Kärcher, B., Jensen, E., and Hartmann, D. L.: Opinion: Tropical cirrus — From micro-scale processes to climate-scale impacts, Clouds and Precipitation/Atmospheric Modelling and Data Analysis/Troposphere/Physics (physical properties and processes), https://doi.org/10.5194/egusphere-2023-1214, 2023.

Ge, J., Wang, Z., Wang, C., Yang, X., Dong, Z., and Wang, M.: Diurnal variations of global clouds observed from the CATS spaceborne lidar and their links to large-scale meteorological factors, Clim Dyn, 57, 2637–2651, https://doi.org/10.1007/s00382-021-05829-2, 2021.

Grise, K. M., Thompson, D. W. J., and Birner, T.: A Global Survey of Static Stability in the Stratosphere and Upper Troposphere, Journal of Climate, 23, 2275–2292, https://doi.org/10.1175/2009JCLI3369.1, 2010.

Hoerl, A. E. and Kennard, R. W.: Ridge Regression: Biased Estimation for Nonorthogonal Problems, Technometrics, 12, 55–67, https://doi.org/10.1080/00401706.1970.10488634, 1970.

Jensen, E. J., Kinne, S., and Toon, O. B.: Tropical cirrus cloud radiative forcing: Sensitivity studies, Geophysical Research Letters, 21, 2023–2026, https://doi.org/10.1029/94GL01358, 1994.

Jensen, E. J., Pfister, L., and Toon, O. B.: Impact of radiative heating, wind shear, temperature variability, and microphysical processes on the structure and evolution of thin cirrus in the tropical tropopause layer, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2010JD015417, 2011.

Jensen, M. and Delgenio, A.: Factors Limiting Convective Cloud-Top Height at the ARM Nauru Island Climate Research Facility, J. Climate, 19, 2105–2117, https://doi.org/10.1175/JCLI3722.1, 2006.

Kärcher, B.: Cirrus Clouds and Their Response to Anthropogenic Activities, Curr Clim Change Rep, 3, 45–57, https://doi.org/10.1007/s40641-017-0060-3, 2017.

Klein, S. A. and Hartmann, D. L.: The Seasonal Cycle of Low Stratiform Clouds, Journal of Climate, 6, 1587–1606, https://doi.org/10.1175/1520-0442(1993)006<1587:TSCOLS>2.0.CO;2, 1993.

Klein, S. A., Hall, A., Norris, J. R., and Pincus, R.: Low-Cloud Feedbacks from Cloud-Controlling Factors: A Review, Surv Geophys, 38, 1307–1329, https://doi.org/10.1007/s10712-017-9433-3, 2017.

Li, J.-L. F., Waliser, D. E., Chen, W.-T., Guan, B., Kubar, T., Stephens, G., Ma, H.-Y., Deng, M., Donner, L., Seman, C., and Horowitz, L.: An observationally based evaluation of cloud ice water in CMIP3 and CMIP5 GCMs and contemporary reanalyses using contemporary satellite data, Journal of Geophysical Research: Atmospheres, 117, https://doi.org/10.1029/2012JD017640, 2012.

Li, Y., Thompson, D. W. J., Stephens, G. L., and Bony, S.: A global survey of the instantaneous linkages between cloud vertical structure and large-scale climate, Journal of Geophysical Research: Atmospheres, 119, 3770–3792, https://doi.org/10.1002/2013JD020669, 2014.

Lin, J.-L. and Mapes, B.: Wind shear effects on cloud-radiation feedback in the western Pacific warm pool, Geophysical Research Letters, 31, https://doi.org/10.1029/2004GL020199, 2004.

Louf, V., Jakob, C., Protat, A., Bergemann, M., and Narsey, S.: The Relationship of Cloud Number and Size With Their Large-Scale Environment in Deep Tropical Convection, Geophysical Research Letters, 46, 9203–9212, https://doi.org/10.1029/2019GL083964, 2019.

Luo, Z. and Rossow, W. B.: Characterizing Tropical Cirrus Life Cycle, Evolution, and Interaction with Upper-Tropospheric Water Vapor Using Lagrangian Trajectory Analysis of Satellite Observations, Journal of Climate, 17, 4541–4563, https://doi.org/10.1175/3222.1, 2004.

Marsham, J. H. and Dobbie, S.: The effects of wind shear on cirrus: A large-eddy model and radar case-study, Quarterly Journal of the Royal Meteorological Society, 131, 2937–2955, https://doi.org/10.1256/qj.04.122, 2005.

May, R. and Bruick, Z.: MetPy: An Community-Driven, Open-Source Python Toolkit for Meteorology, 2019, NS21A-16, 2019.

McFarquhar, G. M., Heymsfield, A. J., Spinhirne, J., and Hart, B.: Thin and Subvisual Tropopause Tropical Cirrus: Observations and Radiative Impacts, Journal of the Atmospheric Sciences, 57, 1841–1853, https://doi.org/10.1175/1520-0469(2000)057<1841:TASTTC>2.0.CO;2, 2000.

McKim, B., Bony, S., and Dufresne, J.-L.: Weak anvil cloud area feedback suggested by physical and observational constraints, Nat. Geosci., 1–6, https://doi.org/10.1038/s41561-024-01414-4, 2024.

Medeiros, B. and Stevens, B.: Revealing differences in GCM representations of low clouds, Clim Dyn, 36, 385–399, https://doi.org/10.1007/s00382-009-0694-5, 2011.

Myers, T. A. and Norris, J. R.: Reducing the uncertainty in subtropical cloud feedback, Geophysical Research Letters, 43, 2144–2148, https://doi.org/10.1002/2015GL067416, 2016.

Myers, T. A., Scott, R. C., Zelinka, M. D., Klein, S. A., Norris, J. R., and Caldwell, P. M.: Observational constraints on low cloud feedback reduce uncertainty of climate sensitivity, Nat. Clim. Chang., 11, 501–507, https://doi.org/10.1038/s41558-021-01039-0, 2021.

Nelson, T. C., Marquis, J., Peters, J. M., and Friedrich, K.: Environmental Controls on Simulated Deep Moist Convection Initiation Occurring during RELAMPAGO-CACTI, Journal of the Atmospheric Sciences, 79, 1941–1964, https://doi.org/10.1175/JAS-D-21-0226.1, 2022.

Nowack, P., Konstantinovskiy, L., Gardiner, H., and Cant, J.: Machine learning calibration of low-cost NO_2 and PM_{10} sensors: non-linear algorithms and their impact on site transferability, Atmospheric Measurement Techniques, 14, 5637–5655, https://doi.org/10.5194/amt-14-5637-2021, 2021.

Pincus, R., Hubanks, P. A., Platnick, S., Meyer, K., Holz, R. E., Botambekov, D., and Wall, C. J.: Updated observations of clouds by MODIS for global model assessment, Earth System Science Data, 15, 2483–2497, https://doi.org/10.5194/essd-15-2483-2023, 2023.

Qu, X., Hall, A., Klein, S. A., and Caldwell, P. M.: On the spread of changes in marine low cloud cover in climate model simulations of the 21st century, Clim Dyn, 42, 2603–2626, https://doi.org/10.1007/s00382-013-1945-z, 2014.

Qu, X., Hall, A., Klein, S. A., and DeAngelis, A. M.: Positive tropical marine low-cloud cover feedback inferred from cloud-controlling factors, Geophysical Research Letters, 42, 7767–7775, https://doi.org/10.1002/2015GL065627, 2015.

Reichler, T., Dameris, M., and Sausen, R.: Determining the tropopause height from gridded data, Geophysical Research Letters, 30, https://doi.org/10.1029/2003GL018240, 2003.

Rieck, M., Nuijens, L., and Stevens, B.: Marine Boundary Layer Cloud Feedbacks in a Constant Relative Humidity Atmosphere, Journal of the Atmospheric Sciences, 69, 2538–2550, https://doi.org/10.1175/JAS-D-11-0203.1, 2012.

Riemann-Campe, K., Fraedrich, K., and Lunkeit, F.: Global climatology of Convective Available Potential Energy (CAPE) and Convective Inhibition (CIN) in ERA-40 reanalysis, Atmospheric Research, 93, 534–545, https://doi.org/10.1016/j.atmosres.2008.09.037, 2009.

Rio, C., Del Genio, A. D., and Hourdin, F.: Ongoing Breakthroughs in Convective Parameterization, Curr Clim Change Rep, 5, 95–111, https://doi.org/10.1007/s40641-019-00127-w, 2019.

Saint-Lu, M., Bony, S., and Dufresne, J.-L.: Observational Evidence for a Stability Iris Effect in the Tropics, Geophysical Research Letters, 47, e2020GL089059, https://doi.org/10.1029/2020GL089059, 2020.

Saint-Lu, M., Bony, S., and Dufresne, J.-L.: Clear-sky control of anvils in response to increased CO2 or surface warming or volcanic eruptions, npj Clim Atmos Sci, 5, 1–8, https://doi.org/10.1038/s41612-022-00304-z, 2022.

Scott, R. C., Myers, T. A., Norris, J. R., Zelinka, M. D., Klein, S. A., Sun, M., and Doelling, D. R.: Observed Sensitivity of Low-Cloud Radiative Effects to Meteorological Perturbations over the Global Oceans, Journal of Climate, 33, 7717–7734, https://doi.org/10.1175/JCLI-D-19-1028.1, 2020.

Sherwood, S. C.: Convective Precursors and Predictability in the Tropical Western Pacific, Mon. Wea. Rev., 127, 2977–2991, https://doi.org/10.1175/1520-0493(1999)127<2977:CPAPIT>2.0.CO;2, 1999.

Sherwood, S. C., Minnis, P., and McGill, M.: Deep convective cloud-top heights and their thermodynamic control during CRYSTAL-FACE, Journal of Geophysical Research: Atmospheres, 109, https://doi.org/10.1029/2004JD004811, 2004.

Sherwood, S. C., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., Hegerl, G., Klein, S. A., Marvel, K. D., Rohling, E. J., Watanabe, M., Andrews, T., Braconnot, P., Bretherton, C. S., Foster, G. L., Hausfather, Z., von der Heydt, A. S., Knutti, R., Mauritsen, T., Norris, J. R., Proistosescu, C., Rugenstein, M., Schmidt, G. A., Tokarska, K. B., and Zelinka, M. D.: An Assessment of Earth's Climate Sensitivity Using Multiple Lines of Evidence, Reviews of Geophysics, 58, e2019RG000678, https://doi.org/10.1029/2019RG000678, 2020.

Siebesma, A. P., Bretherton, C. S., Brown, A., Chlond, A., Cuxart, J., Duynkerke, P. G., Jiang, H., Khairoutdinov, M., Lewellen, D., Moeng, C.-H., Sanchez, E., Stevens, B., and Stevens, D. E.: A Large Eddy Simulation Intercomparison Study of Shallow Cumulus Convection, Journal of the Atmospheric Sciences, 60, 1201–1219, https://doi.org/10.1175/1520-0469(2003)60<1201:ALESIS>2.0.CO;2, 2003.

Srinivasan, J. and Smith, G. L.: The Role of Heat Fluxes and Moist Static Energy in Tropical Convergence Zones, Monthly Weather Review, 124, 2089–2099, https://doi.org/10.1175/1520-0493(1996)124<2089:TROHFA>2.0.CO;2, 1996.

Tsushima, Y., Ringer, M. A., Webb, M. J., and Williams, K. D.: Quantitative evaluation of the seasonal variations in climate model cloud regimes, Clim Dyn, 41, 2679–2696, https://doi.org/10.1007/s00382-012-1609-4, 2013.

Wood, R. and Bretherton, C. S.: On the Relationship between Stratiform Low Cloud Cover and Lower-Tropospheric Stability, Journal of Climate, 19, 6425–6432, https://doi.org/10.1175/JCLI3988.1, 2006.

Xu, K.-M. and Cheng, A.: Understanding the tropical cloud feedback from an analysis of the circulation and stability regimes simulated from an upgraded multiscale modeling framework, Journal of Advances in Modeling Earth Systems, 8, 1825–1846, https://doi.org/10.1002/2016MS000767, 2016.

Zelinka, M. D. and Hartmann, D. L.: Why is longwave cloud feedback positive?, J. Geophys. Res., 115, D16117, https://doi.org/10.1029/2010JD013817, 2010.

Zelinka, M. D. and Hartmann, D. L.: The observed sensitivity of high clouds to mean surface temperature anomalies in the tropics, Journal of Geophysical Research: Atmospheres, 116, https://doi.org/10.1029/2011JD016459, 2011.

Zelinka, M. D., Klein, S. A., and Hartmann, D. L.: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part I: Cloud Radiative Kernels, Journal of Climate, 25, 3715–3735, https://doi.org/10.1175/JCLI-D-11-00248.1, 2012a.

Zelinka, M. D., Klein, S. A., and Hartmann, D. L.: Computing and Partitioning Cloud Feedbacks Using Cloud Property Histograms. Part II: Attribution to Changes in Cloud Amount, Altitude, and Optical Depth, Journal of Climate, 25, 3736–3754, https://doi.org/10.1175/JCLI-D-11-00249.1, 2012b.

Zelinka, M. D., Zhou, C., and Klein, S. A.: Insights from a refined decomposition of cloud feedbacks, Geophysical Research Letters, 43, 9259–9269, https://doi.org/10.1002/2016GL069917, 2016.

Zelinka, M. D., Klein, S. A., Qin, Y., and Myers, T. A.: Evaluating Climate Models' Cloud Feedbacks Against Expert Judgment, Journal of Geophysical Research: Atmospheres, 127, e2021JD035198, https://doi.org/10.1029/2021JD035198, 2022.

Zhang, Y. and Fueglistaler, S.: How Tropical Convection Couples High Moist Static Energy Over Land and Ocean, Geophysical Research Letters, 47, e2019GL086387, https://doi.org/10.1029/2019GL086387, 2020.