Internal variability of all-sky and clear-sky surface solar radiation on decadal timescales

Boriana Chtirkova[1], Doris Folini[2], Lucas Ferreira Correa[1], and Martin Wild[3]

[1]ETH Zurich
[2]ETH Zurich, Institute for Atmospheric and Climate Science
[3]ETH Zürich

November 22, 2022

Abstract

Internal variability comprises all processes that occur within the climate system without any natural or anthropogenic forcing. Climate driving variables like the surface solar radiation (SSR) are shown to exhibit unforced trends (i.e. trends due to internal variability) of magnitudes comparable to the magnitude of the forced signal even on decadal timescales. We use annual mean data from 50 models participating in the pre-industrial control experiment (piControl) of the Coupled Model Intercomparison Project – Phase 6 (CMIP6) to give quantitative grid-box specific estimates of the magnitudes of unforced trends. To characterise a trend distribution, symmetrical around 0, we use the 75th percentile of all possible values, which corresponds to a positive trend with 25% chance of occurrence. For 30-year periods and depending on geographical location, this trend has a magnitude between 0.15 and 2.1Wm^{-2}/decade for all-sky and between 0.04 and 0.38Wm^{-2}/decade for clear-sky SSR. The corresponding area-weighted medians are 0.69Wm^{-2}/decade for all-sky trends and 0.17Wm^{-2}/decade for clear-sky trends. The influence of internal variability is on average 6 times smaller in clear-sky, compared to all-sky SSR. The relative uncertainties of these estimates, derived from the CMIP6 inter-model spread, are ±32% for all-sky and ±43% for clear-sky SSR trends. Reasons for differences between models like horizontal resolution, aerosol handling and the representation of atmospheric and oceanic phenomena are investigated. The results can be used in the analysis of observational time series by attributing a probability for a trend to comprise a component due to internal variability, given its magnitude, length and location.

Internal variability of all-sky and clear-sky surface solar radiation on decadal timescales

Boriana Chtirkova[1], Doris Folini[1], Lucas Fereira Correa[1], Martin Wild[1]

[1]Institute for Atmospheric and Climate Sciences, ETH Zürich, Zürich, Switzerland

Key Points:

- Internal variability of all-sky and clear-sky surface solar radiation and associated decadal-scale unforced trends are quantified probabilistically.
- Cloud variability is responsible for a large fraction of the unforced SSR trends: internal variability is ~ 6 times smaller in clear-sky trends compared to all-sky.
- Unforced SSR trends are particularly strong in some regions, the regional pattern being different for clear-sky and all-sky trends.

Corresponding author: Boriana Chtirkova, `boriana.chtirkova@env.ethz.ch`

Abstract

Internal variability comprises all processes that occur within the climate system without any natural or anthropogenic forcing. Climate driving variables like the surface solar radiation (SSR) are shown to exhibit unforced trends (i.e. trends due to internal variability) of magnitudes comparable to the magnitude of the forced signal even on decadal timescales. We use annual mean data from 50 models participating in the pre-industrial control experiment (piControl) of the Coupled Model Intercomparison Project – Phase 6 (CMIP6) to give quantitative grid-box specific estimates of the magnitudes of unforced trends. To characterise a trend distribution, symmetrical around 0, we use the 75th percentile of all possible values, which corresponds to a positive trend with 25% chance of occurrence. For 30-year periods and depending on geographical location, this trend has a magnitude between 0.15 and 2.1 Wm^{-2}/decade for all-sky and between 0.04 and 0.38 Wm^{-2}/decade for clear-sky SSR. The corresponding area-weighted medians are 0.69 Wm^{-2}/decade for all-sky trends and 0.17 Wm^{-2}/decade for clear-sky trends. The influence of internal variability is on average 6 times smaller in clear-sky, compared to all-sky SSR. The relative uncertainties of these estimates, derived from the CMIP6 inter-model spread, are ±32% for all-sky and ±43% for clear-sky SSR trends. Reasons for differences between models like horizontal resolution, aerosol handling and the representation of atmospheric and oceanic phenomena are investigated. The results can be used in the analysis of observational time series by attributing a probability for a trend to comprise a component due to internal variability, given its magnitude, length and location.

1 Introduction

Internal variability has been identified together with model response and emission scenario uncertainty as one of the sources of uncertainty in climate projections (e.g. Hawkins and Sutton (2009); Deser et al. (2010)). In fact, on time scales of typically a few decades into the future, internal variability is a major contributor to the overall uncertainty in climate projections. The advances in numerical modelling in recent decades (IPCC, 2013, 2021; Eyring et al., 2016) work towards a better estimation of models' responses to radiative forcing, but internal variability has been ruled out as an irreducible uncertainty source. Due to its dominant relevance on shorter time scales, internal variability has the potential to diminish or even reverse the long-term effect of anthropogenic forcing over a limited period (Hawkins & Sutton, 2009). Being stochastic in nature, internal variability restricts us from making a direct comparison between historical model simulations and observations, and concerning future projections, it interferes with the time of emergence of the forced signal. Since the effect of these natural fluctuations in the climate system cannot be neglected for decadal time scales, statistical approaches of quantifying and understanding them need to be considered.

The relevance of internal variability for a climate variable may be gathered from the ratio between the variable's response to a forced signal and the fluctuations which it exhibits due to the chaotic nature of the climate system, also known as the signal-to-noise ratio. A climate variable which is suspected to have a low signal-to-noise ratio, but still maintain a pronounced response to anthropogenic forcing on the decadal timescale, is the downward surface solar radiation (SSR) (Wild et al., 2005; Wild, 2009; Folini et al., 2017). This makes anthropogenically forced trends challenging to identify in both observations (Sanchez-Lorenzo et al., 2008; Wild, 2016) and models (Storelvmo et al., 2018; Moseid et al., 2020). The multi-decadal changes of SSR, termed dimming and brightening (Wild et al., 2005), are known to affect the climate system thorough different mechanisms, concerning the hydrological cycle (e.g. Wild et al. (2008); Ramanathan (2001)), global warming (Wild et al., 2007), the cryosphere (Ohmura et al., 2007; Wild et al., 2008; Wild, 2009), the terrestrial biosphere and carbon cycle (e.g. Jones and Cox (2001); Mercado et al. (2009); Farquhar (2003)). Global dimming and brightening trends are observed for example in China from 1958-1999, the sustained trend magnitudes for the 41-year period being $-7.4\mathrm{Wm}^{-2}$/decade for all-sky and $-9.7\mathrm{Wm}^{-2}$/decade for clear-sky SSR (S. Yang et al., 2019). In order to disentangle the effects of forced trends versus internal variability, one needs to make a distinction between the processes which control them. The most widely accepted factors responsible for the anthropogenic effect have been narrowed down to the emissions of anthropogenic aerosols (Streets et al., 2006; Gidden et al., 2019; Wang et al., 2021) and their effect on clouds (e.g. Quaas et al. (2008)), both of which bear an enormous uncertainty; in addition, human-caused warming contributes to changes in the amount and lifetime of radiatively active gases relevant for short wave radiation like water vapour (Santer et al., 2007; Hodnebrog et al., 2019). Internal variability, on the other hand, arises from the non-linear dynamics of the atmosphere and ocean and possesses the characteristics of a random stochastic process (Deser et al., 2010), which can be described and quantified through its corresponding distribution function.

We take the internal variability analysis as a starting point for understanding the behavior of a climate variable and apply it more specifically to SSR, a key component of the global energy balance (Wild et al., 2014). The unforced control runs (piControl) of the latest generation of climate models, participating in the sixth phase of the Coupled model intercomparison project – CMIP6 (Eyring et al., 2016), comprise a collection of processes that occur solely due to internal variability of the climate system. Models also give us the unique opportunity to separate the effect of clouds (and their associated uncertainty) from other factors affecting SSR through the distinction between all-sky SSR (including cloudy conditions) and clear-sky SSR (removing clouds from the radiative transfer calculations). We follow a methodology developed by Folini et al. (2017) and based on CMIP – Phase 5 models (Taylor et al., 2012a), which links the probability of occurrence of an unforced all-sky SSR trend with a certain length to the standard deviation of the underlying SSR time series. In the current work, this methodology is applied to the latest generation of climate model output data (CMIP6) and extended also to clear-sky SSR. Clear-sky is of interest as there are no cloud effects, thus a major source of internal variability (noise) is absent, suggesting that any anthropogenic contribution should be more easily identifiable than under all-sky conditions.

The structure of this paper is as follows: we briefly present the methods and data in section 2; we test where in the world and for what temporal scales trends of varying length can be statistically linked, present results and estimate uncertainty from the multi-model ensemble in section 3; and discuss possible applications in the analysis of observational time series in section 4.

2 Methods and Data

To examine trends caused by internal variability only, we use an analytical model, described in Thompson et al. (2015) and applicable to data that possess an approximately Gaussian distribution and are stationary in time, and applied to SSR by Folini et al. (2017). The model links the standard deviation of the distribution of all possible N-year trends σ_N to the standard deviation of the underlying annual time series σ_{ts} through $\sigma_N \approx \sqrt{12}N^{-3/2}\sigma_{ts}$ (Weatherhead et al., 1998; Tiao et al., 1990; Nishizawa & Yoden, 2005; Hinkelman et al., 2009). A trend is defined as the linear regression slope for a specified N-year period in the variable time series. The mathematical model relies on the following two key assumptions: (1) the distribution of all possible trends with length N derived from the time series is Gaussian; (2) the variable does not exhibit significant autocorrelation in time. Our approach towards the problem is to first test whether, where and for what time scales assumptions (1) and (2) are valid and then analyse the trend magnitudes within these spatial and temporal constrains. We ensure a Gaussian distribution through the Kolmogorov-Smirnov (K-S) and Anderson-Darling (A-D) normality tests with a significance level $\alpha =$ 0.05, as done in Folini et al. (2017). The tests are performed for trends with lengths of 10, 30, 50 and 100 years. Autocorrelation in time is checked using the Pearson correlation coefficient calculated for the whole time series against the same time series shifted by one time increment.

In order to build a statistical distribution of unforced SSR trends over decadal time scales, one needs hundreds of years of data without anthropogenic forcing – a task that can only be achieved with climate models. A suitable data source for our analysis are the pre-industrial control runs (also known as piControl) of each model participating in CMIP6. The piControl experiment is designed to evaluate the climate models' unforced variability through keeping constant greenhouse gases and anthropogenic aerosol concentrations representative for the period prior to the 1850s (Eyring et al., 2016). The simulation lengths vary from a few hundred up to 2000 years between different models, usually only one ensemble is provided per model. To ensure statistical robustness, we take only models, which have submitted more than 450 years of a piControl simulation, which leaves us with 54 simulations from 50 models. Due to autocorrelation considerations of atmospheric and oceanic phenomena relevant to SSR, we mainly work with yearly average data. To obtain a suitable multi-model comparison, we interpolate the SSR variables ("rsds" and "rsdscs" in the CMIP6 archive) to a common grid using second order conservative remapping. The common grid is chosen as the finest grid among CMIP6 models, namely the CNRM-CM6-1-HR model's grid with a spatial increment $\sim 0.5°$ or 360 latitude and 720 longitude points. The statistical analysis is applied to each individual grid box, both for all-sky and clear-sky SSR. An example of the statistical distribution per model in one grid cell (the grid cell containing Lindenberg, Germany) is given through the model data-derived probability density functions (PDFs) in Figure 1. The trend distributions are calculated for a 30-year period and are centered around 0. The 30-year trends all-sky distribution with a multi-model median for that grid cell of $\sigma_{30,as} = 1.16$ Wm^{-2}/decade is notably wider than that of clear-sky $\sigma_{30,cs} = 0.16$ Wm^{-2}/decade (the values correspond to the CMIP6 multi-model median, taken over the 30-year trend PDFs per model). The multi-model median PDF (solid black line), calculated as the median value per bin of the trends distribution, closely follows the Gaussian, analytically calculated from the multi-model median of σ_{ts} and a mean value of 0 (dashed black line). The two lines diverge for larger absolute trends in clear-sky SSR due to the large inter-model spread in that part of the distribution, thus the multi-model median PDF cannot follow the analytical Gaussian. This spread is mainly caused by the flatter distribution of EC-Earth models. In the upcoming sections, we aggregate this analysis per grid box and focus on the differences between grid boxes.

To justify our modelling study, and to infer whether internal variability of SSR in numerical models is sound, we use the following two sources of observational evidence: point measurements from the Baseline Surface Radiation Network (BSRN; (Ohmura et al., 1998)) and gridded surface fluxes from the Clouds and the Earth's Radiant Energy System (CERES) Energy Balanced and Filled (EBAF) surface data product (Kato et al., 2018), edition 4.1. The CERES-EBAF surface fluxes are satellite-derived and validated against surface measurements, they cover the period after 2000. For our analysis, the data is interpolated to the same 0.5° grid as the models.

3 Results

We apply the statistical method described above per grid box for each model - first we turn to the assumptions that are behind the approximate analytical link between the statistics of the SSR time series and the statistics of SSR trends of different lengths. For each gird box of each model we examine how well these assumptions are fulfilled (section 3.1); we then turn to the regional differences at grid box level (section 3.2); in section 3.3 we analyse the contribution of clear-sky variability to all-sky variability; lastly we look at differences among models in CMIP6, while considering observations-derived data and also extending back to CMIP5; we consider potential reasons for these differences and estimate the general uncertainty of our analysis in Section 3.4.

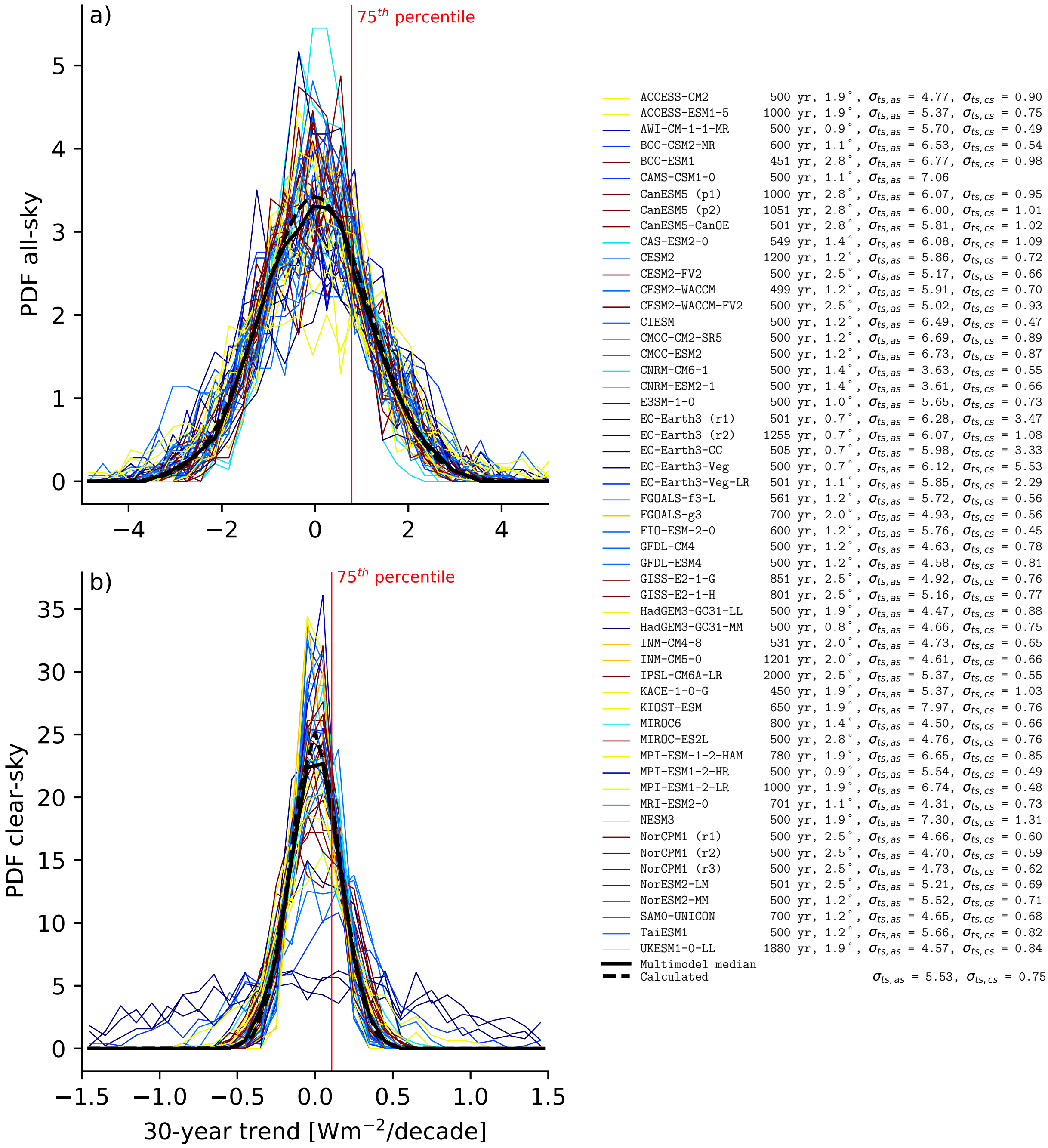

Figure 1. PDFs of trend distributions derived per model for one grid box, corresponding to Lindenberg, Germany (52.21°N, 14.122°E). The legend comprises the names of the models used in the study (first column), the length of the control runs in years (second column), the approximate horizontal increment in degrees (third column), the standard deviation of the underlying annual time series for all-sky $\sigma_{ts,as}$ and for clear-sky $\sigma_{ts,cs}$ in Wm^{-2} (forth and fifth columns). The continuous black line represents the multi-model PDF, obtained via taking the median value per bin. The dashed black line is the Gaussian calculated using the median of σ_{ts} taken over all models as standard deviation to obtain σ_{30} (see text). Vertical red lines indicate the 75th percentile of the distribution when taking both positive and negative values.

3.1 Temporal and spatial scale of applicability

The CMIP6 piControl data is tested per grid box for autocorrelation in time and whether the trend distributions for different trend lengths are normal with a significance level $\alpha = 0.05$. The trend distribution is gathered through linear regressions of yearly averaged data. The control runs of the models are first checked for linear trends, as climate models are known to drift in at least some variables (Gupta et al., 2013; Irving et al., 2020). Repeating the analysis using linearly de-trended data, where the linear trend is calculated and subtracted from the entire control simulation, is found to make no significant difference in the distribution of SSR trends (not shown); all subsequent results are obtained using the original data. Grid cells, where the statistical tests for Gaussian distribution fail, are marked as unsuitable for applying analytical relations that link the statistics of the SSR time series (σ_{ts}) with the statistics of associated N-year long SSR trends (σ_N).

We use the K-S test to compare the grid-box trend distribution with a sample of a Gaussian distribution with a mean value of zero and standard deviation as the one of the model data in the specific grid cell. For all-sky SSR the K-S test is failed in 2% of the grid boxes for the 10-year trends distributions, failures occur in 5% of the grid boxes for 30-year trends, 10% of the grid boxes for 50-year trends, and 23% of the grid boxes for 100-year trends. This percentage value represents the median among models, while its relative standard deviation does not go beyond 3%. For clear-sky SSR the K-S test is failed slightly more often: in 2% of the grid boxes for 10-year trends, 6% of the grid boxes for 30-year trends, 11% the of grid boxes for 50-year trends, and 25% the of grid boxes for 100-year trends. This proves that in the majority of grid boxes for both all-sky and clear-sky SSR the distribution is Gaussian with a mean value of zero. The A-D statistical test, which is more sensitive to the tails of the distribution (Stephens, 1974), shows a larger percentage of rejections for both all-sky and clear-sky SSR.

We adopt a similar approach of testing the underlying time series for autocorrelation in time and check whether it is larger than a certain value. The lag-1 autocorrelation for yearly average data of all-sky SSR is more than 0.1 in 25% of grid boxes, above 0.2 in 7% of grid boxes, and above 0.3 in 3% of grid boxes. The corresponding percentage values for clear-sky are slightly larger: 45%, 17% and 7% for 0.1, 0.2 and 0.3 respectively. Performing the same tests for the monthly anomalies results in significantly larger autocorrelation values and deems our statistical approach inapplicable for most regions, concerning the clear-sky variable.

One can obtain an idea of the spatial regions where the theory is not applicable from the hatched areas in Figure 2 – left column for all-sky, right column for clear-sky SSR. They represent the regions, which fail the K-S test for the 30-year distribution and have an autocorrelation above 0.2. We note that these regions still maintain unforced trends, though we cannot link trends of different lengths N and with a given percentile p analytically. These regions shrink if we test for shorter trend periods and compare the autocorrelation coefficient against a larger value. Consequentially, the regions extend to a greater area when we are interested in longer periods and/or smaller autocorrelation values (small autocorrelation of the time series results in a smaller uncertainty of the results (Thompson et al., 2015)). As the given percentages suggest, the majority of the inapplicability results from a high autocorrelation in time. The corresponding regions are the Tropical Pacific, where the conditions are dominated by the El NiñoSouthern Oscillation (ENSO), and areas covered by sea ice. Rejections of the K-S test occur almost uniformly on the planet with a slightly higher concentration for some models in the Northern polar regions. The clear-sky SSR generally shows higher autocorrelation values.

In the following, when we refer to the applicable region, we mean the grid boxes in which we can analytically link σ_{ts} to σ_N for a trend distribution of any length N, i.e.

the grid boxes that possess a Gaussian distribution and approximately zero autocorrelation in time, obtained from the CMIP6 multi-model median.

3.2 Regional differences

To infer the regional differences of trend statistics, we obtain the distribution of all possible trends per grid box. The distribution of all possible trends of the same length, which is centered around 0, can be intuitively represented through its 75th percentile. This value represents a positive trend with a probability of occurrence of 25%, when considering both positive and negative trends (i.e. there is a 50% chance that the absolute magnitude of the trend exceeds this value, see vertical red lines on Figure 1). Since the distribution is symmetrical around zero, the corresponding negative trend with a 25% probability of occurrence has the same magnitude with an opposite sign. The percentiles p of trends with different lengths N with a mean value of 0 can be expressed through $t(p, N) = \sigma_N Z(p)$, where $Z(p)$ is the appropriate entry from the Z-table, which links a percentile value to the standard deviation in a standard normal distribution, for example $Z(75) = 0.674$. Throughout the paper, we use the 75th percentile of the distribution of all possible 30-year trends $t(p{=}75, N{=}30)$, which is consistent with the preceding study of Folini et al. (2017) and comprises the time scales of global dimming and brightening (Wild et al., 2005). The CMIP6 multi-model median of this value is presented for all-sky and clear-sky on Figure 2 a) and b), respectively. In the applicable regions, discussed in Section 3.1, the all-sky trends with a global mean value of 0.69 $\mathrm{Wm^{-2}}$/decade are generally larger than those of clear-sky with 0.165 $\mathrm{Wm^{-2}}$/decade. On average, all-sky SSR shows larger magnitude trends over oceans (0.75 $\mathrm{Wm^{-2}}$/decade), compared to land (0.58 $\mathrm{Wm^{-2}}$/decade). On the contrary, clear-sky SSR trends are more significant over land (0.179 $\mathrm{Wm^{-2}}$/decade) compared to ocean (0.157 $\mathrm{Wm^{-2}}$/decade).

The spatial distribution of trends is also illustrated in a probabilistic manner on Figure 2 c) and d). The transformation to probability per grid point for fixed values of t and N can be derived using either σ_{ts} or σ_N for any trend period where the distribution of all possible trends remains Gaussian. In this example, we use the distribution derived from all possible 30-year trends per model per grid box and take the CMIP6 multi-model median of $\sigma_{N=30}$ for both all-sky and clear-sky SSR trends per grid box. We then extract the percentile (probability) that corresponds to a certain t through $Z(p) = \frac{t(N)}{\sigma_N}$. From this transformation, one can tell the probability of occurrence of a trend with magnitude t over a period $N = 30$ years at a certain location. For all-sky SSR, we choose $t = 1.5$ $\mathrm{Wm^{-2}}$/decade (the choice is purely arbitrary, but is of comparable magnitude to observed SSR trends (e.g. Wild (2009))) and the analysis tells us there is up to a 32% chance to observe an unforced trend of this magnitude or larger in the regions around the Intertropical Convergence Zone (ITCZ). The probabilities are notably larger over the open ocean areas, Central America, the east coasts of South America and Africa and Australia. The probability of observing it in Europe can be up to 15%, and in China – 19% (given are the maximum values over the CMIP6 multi-model median for the respective regions). Central Europe and Spain have higher probability values than the north and east parts of the continent. The probability of observing such a trend tends towards zero in the subtropical regions with no clouds.

We choose $t = 0.3$ $\mathrm{Wm^{-2}}$/decade for clear-sky SSR, which has the highest probability of occurrence in the subtropical desert regions (up to 22% in the Sahara desert). Observing a trend of such magnitude in Europe has a probability of up to 12%, and in China – 13%. In contrast to the all-sky distribution, the northernmost and easternmost parts of Europe show higher probabilities for clear-sky trends than the central part. The probability tends towards 0 in the Amazon, Antarctica, Greenland and large areas of the open ocean.

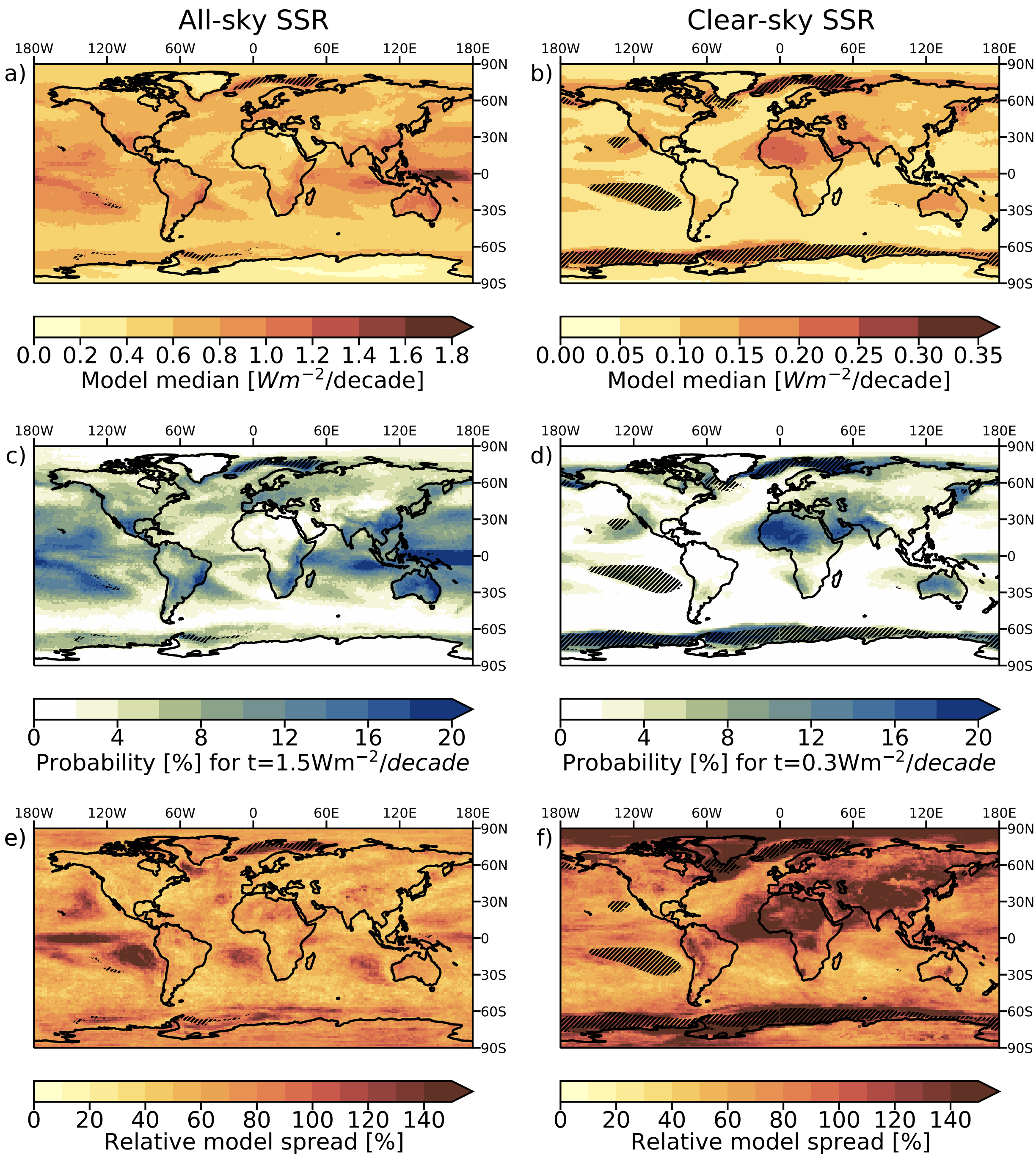

Figure 2. CMIP6 multi-model median maps of the 75th percentile of all 30-year trends for all-sky (a) and clear-sky (b) SSR. Maps showing the probability of occurrence for an all-sky SSR trend with magnitude 1.5 Wm^{-2}/decade (c) and a clear-sky SSR trend with magnitude 0.3 Wm^{-2}/decade (d) over a period of 30 years. Relative model spread (the difference between the 90th and 10th percentile, divided by the multi-model median) of the 75th percentile per grid box for all-sky (e) and clear-sky SSR trends (f). The hatched areas represent the grid boxes, which fail the K-S test for 30-year trends for more than 40% of considered models or have a median absolute value of the autocorrelation above 0.2.

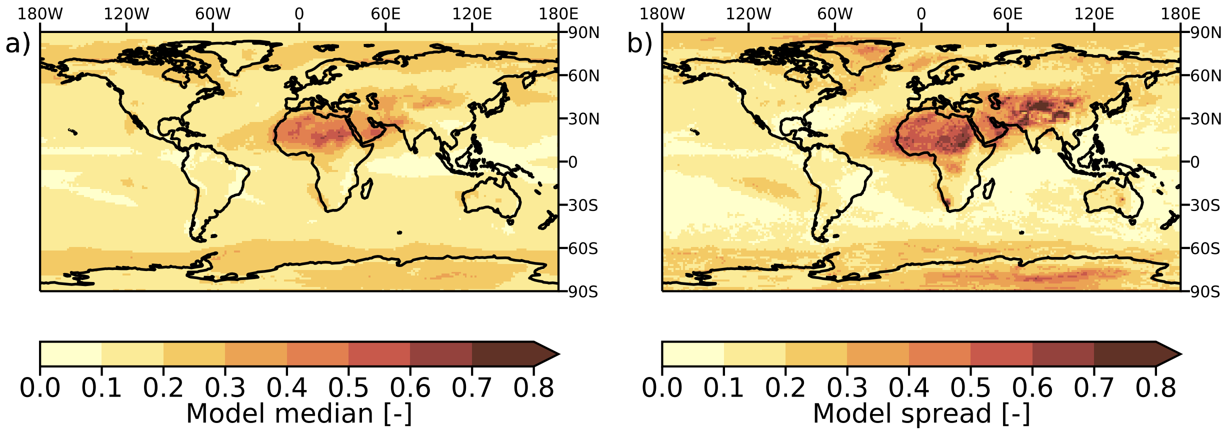

Figure 3. Ratio between clear-sky and all-sky SSR trends (75th percentile of all trends): CMIP6 multi-model median (a) and CMIP6 model spread (c).

In contrast to the open ocean tendencies of having stronger trends in all-sky SSR compared to land, semi-enclosed seas with less inter annual SST variability (in CMIP6 models), like the Mediterranean sea and the Black sea, show the opposite: lower magnitude trends in all-sky SSR compared to the neighbouring lands.

We note that the maps of t(75, 30) (Figure 2 a-b) and the probability maps (Figure 2 c-d) are two ways of presenting the patterns of internal variability of SSR trends: regions where the probabilities are higher tend to experience trends of larger magnitudes.

An initial measure of the uncertainty of the given results can be obtained from the range of estimates CMIP6 models produce for $t(75, 30)$, which is an indicator of the level of disagreement among models. The relative spread, shown on Figure 2 e) and f), is defined as the difference between the 90th and 10th percentile of $t(75, 30)$ per grid point, divided by the multi-model median (i.e. the 50th percentile) of $t(75, 30)$ at that point. Generally, the patterns of disagreement among models resemble the patterns of trends: the spread is larger where the median magnitudes of $t(75, 30)$ are larger. The median value (across all grid boxes, weighted by their cell area) of the relative spread of all-sky trends is 63% (0.38 Wm^{-2}/decade in absolute units), yielding an uncertainty based on model differences of ±32%. For clear-sky, the spread is notably larger, with a median of 86% (0.09 Wm^{-2}/decade in absolute units), yielding an uncertainty of ±43%. All-sky SSR trends have a more prominent spread above the oceans while clear-sky SSR - above land, which is mostly influenced by desert areas (mineral dust). Regions covered by sea ice, which are also prone to autocorrelation in time for clear-sky SSR, have a larger inter-model spread. For Europe, the patterns of the model spread closely resemble the trend magnitudes, whereas for China we observe opposite patterns in the western and eastern part: Northwest China has a lower potential for trends, but exhibits a larger spread among models in both all-sky and clear-sky SSR. Reasons for differences among models are further discussed in Section 3.4.

3.3 Relationship between all-sky and clear-sky SSR trends

We further analyze the ratio between clear-sky and all-sky SSR trends, which generally takes values between 0 and 1, where lower values mean the trends in surface radiation are almost entirely due to clouds and larger values indicate that trends similar in magnitude are observed both in all-sky and clear-sky. Spatial patterns of this ratio, obtained by dividing the 75th percentile of clear-sky by the 75th percentile of all-sky SSR trends per model and then taking the multi-model median per grid box, are shown on Figure 3 a). From it, we see that the influence of clouds dominates above oceans, the ITCZ and Southwest China. On the other hand, the ratio is close to 1 in large desert areas in Africa and Asia. It is worth noting that in desert areas with more prominent clear-sky trends, the negative correlation between clear-sky SSR and integrated water vapour content (not shown) is weaker than on other places on the globe. On average, the ratio between clear-sky and all-sky SSR is 0.166, which translates to an increase of SSR variability by a factor of 6 due to cloud variability. The inter-model spread, shown on Figure 3 b), is more prominent over deserts and areas, where models differ in their clear-sky SSR variability. On average the inter-model spread is 0.172 in absolute units or 100% in relative units.

3.4 Model differences within CMIP6, comparison with CMIP5 and CERES

Turning back to the considerable model spread, we opt to investigate differences among models as one potentially relevant factor. Such differences range from horizontal resolution, physical parametrizations, differences in the emission inventories and treatment of aerosols to the models' overall ability to represent processes internal to the climate system, which are of relevance to SSR like the ENSO (e.g. (Li et al., 2015; Y. Yang et al., 2016; Pinker et al., 2017)), Pacific Deacadal Oscillation (PDO), North Atlantic Oscillation (NAO, e.g. (Chiacchio & Wild, 2010)), etc.

Comparing the PDFs for models with different horizontal resolution does not show a clear distinction among models with a smaller or larger spatial increment. Each PDF on Figure 1 is derived from one grid cell, the size of which depends on the model's spatial increment (no interpolation to a finer grid is performed). This dependency is also checked for other grid boxes, as well as spatially averaged data (not shown). The Pearson correlation coefficient between the model's spacial increment and spatially averaged σ_{ts} per model is close to 0 (-0.03 for all-sky and -0.15 for clear-sky SSR). The fact that there is no clear distinction between them suggests that either (1) models have the ability to account for the variability at smaller spatial scales with microphysical parametrizations, (2) factors other than grid resolution are the main determinants for internal variability in climate models or (3) internal variability, relevant for trends longer than 10 years, occurs at spatial scales larger than 2.8° (the coarsest CMIP6 model resolution). A further investigation of the spatial scales of internal variability is beyond the scope of the current paper, but a potential topic for a future study.

Another possible reason for differences in SSR among models is the aerosol treatment, which is known to be as important as aerosol emission inventories (Persad et al., 2014). We distinguish CMIP6 models, which use prescribed aerosols and those coupled to an aerosol model, thus allowing for interactive aerosol treatment. The results are presented on Figure 4, where the top panel (a-b) shows multi-model median of $t(75, 30)$, derived only from models with prescribed aerosols, while the bottom panel (c-d) is derived from ESM models that also include an aerosol model. The overall impression is that all-sky patterns of trend magnitudes are similar in both ESM model groups, but the models with interactive aerosol yield larger trends above desert areas and sediment mountain ranges like the Himalayas. Numerically, the average clear-sky $t(75, 30)$ value is 0.172 Wm^{-2}/decade for models with interactive aerosols versus 0.156 Wm^{-2}/decade for models using prescribed aerosols, i.e. the aerosol models in CMIP6 attribute to 10% stronger clear-sky trends on average. This relative difference is more pronounced above desert areas: 80% for the Sahara, 60% for the Sahel, 67% for the Syrian desert, 40% for the Arabian peninsula, 11% for the Kalahari desert, 30% for the Gobi desert. Models with an interactive representation of aerosols have a larger all-sky yearly autocorrelation in time above the

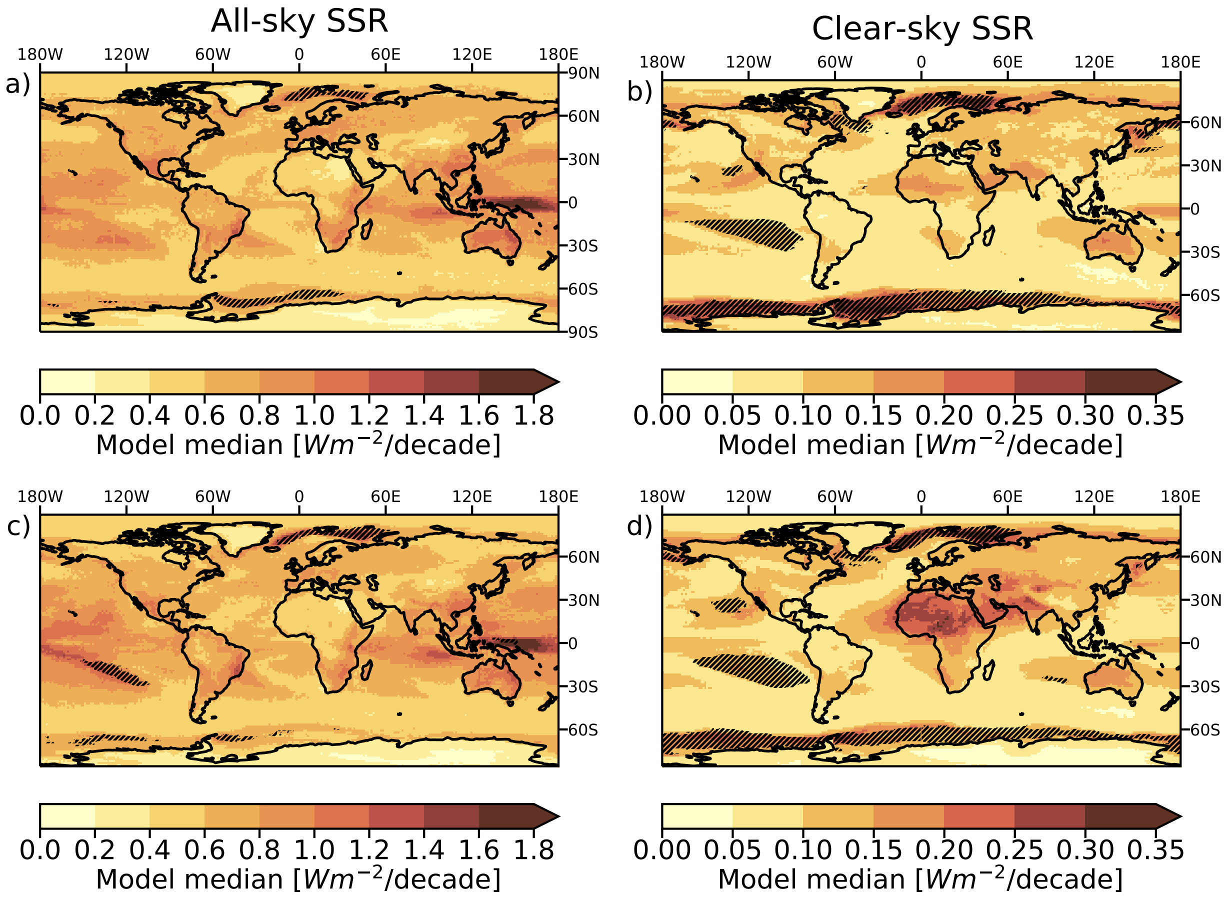

Figure 4. Multi-model median maps of the 75th percentile of all 30-year trends for CMIP6 models using prescribed aerosols (a, b) and models coupled to an aerosol model (c, d). The hatched areas represent the grid boxes, which fail the K-S test for 30-year trends for more than 40% of considered models or have a median absolute value of the autocorrelation above 0.2.

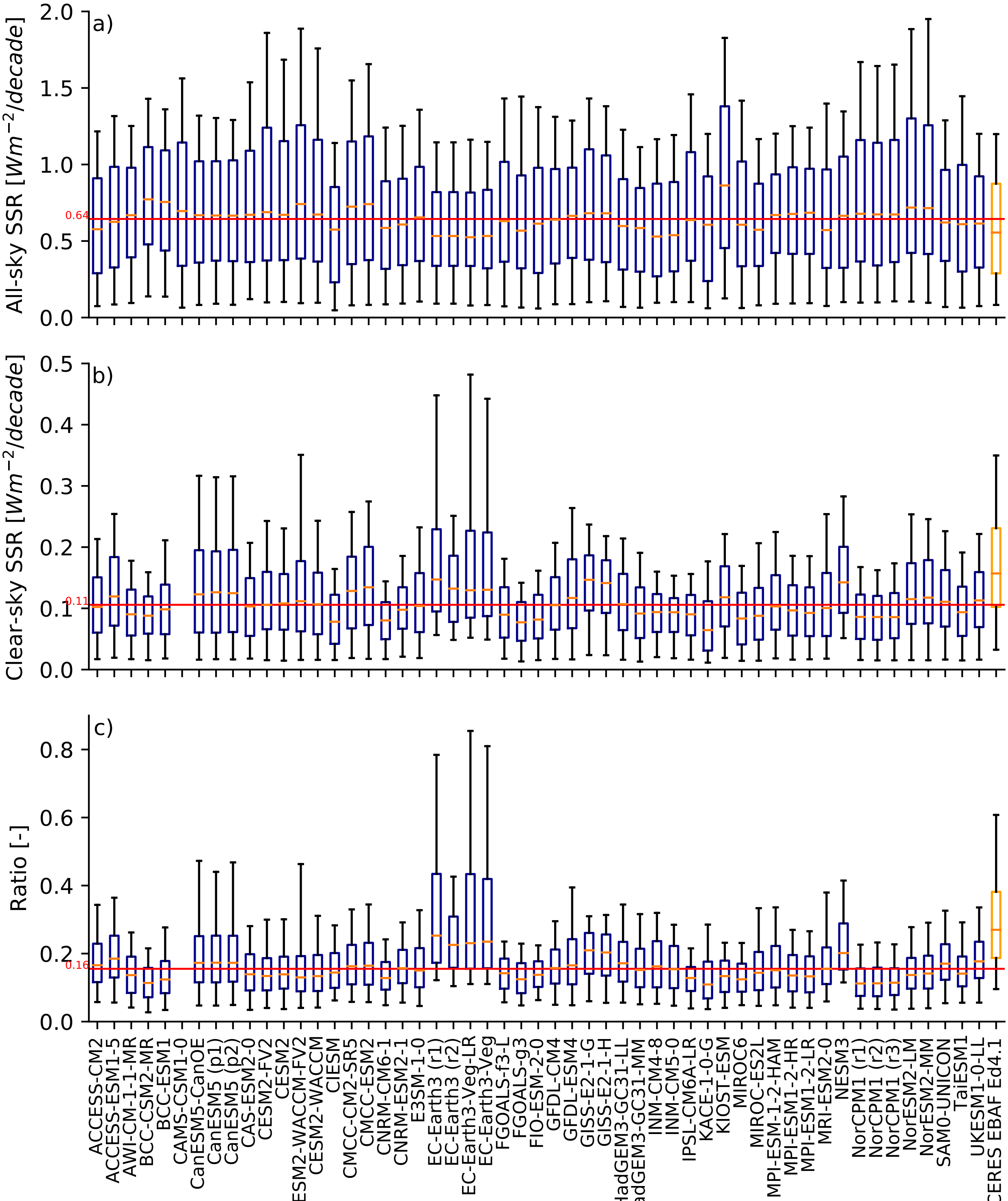

Figure 5. Box plots, representing $t(75, 30)$ per model for all-sky (a), clear-sky (b) SSR trends and the ratio between the two (c), when covering all grid points of a model. Rightmost orange box represents $t(75, 30)$, calculated from σ_{ts} of CERES EBAF Ed4.1. Data is weighted to the grid box area. Short orange lines indicate the median value among all grid cells, blue boxes mark the 25th and 75th percentiles and the whiskers – the 10th and 90th percentiles. Horizontal red line indicates the mean value of all CMIP6 model medians, numerical value is given on the left.

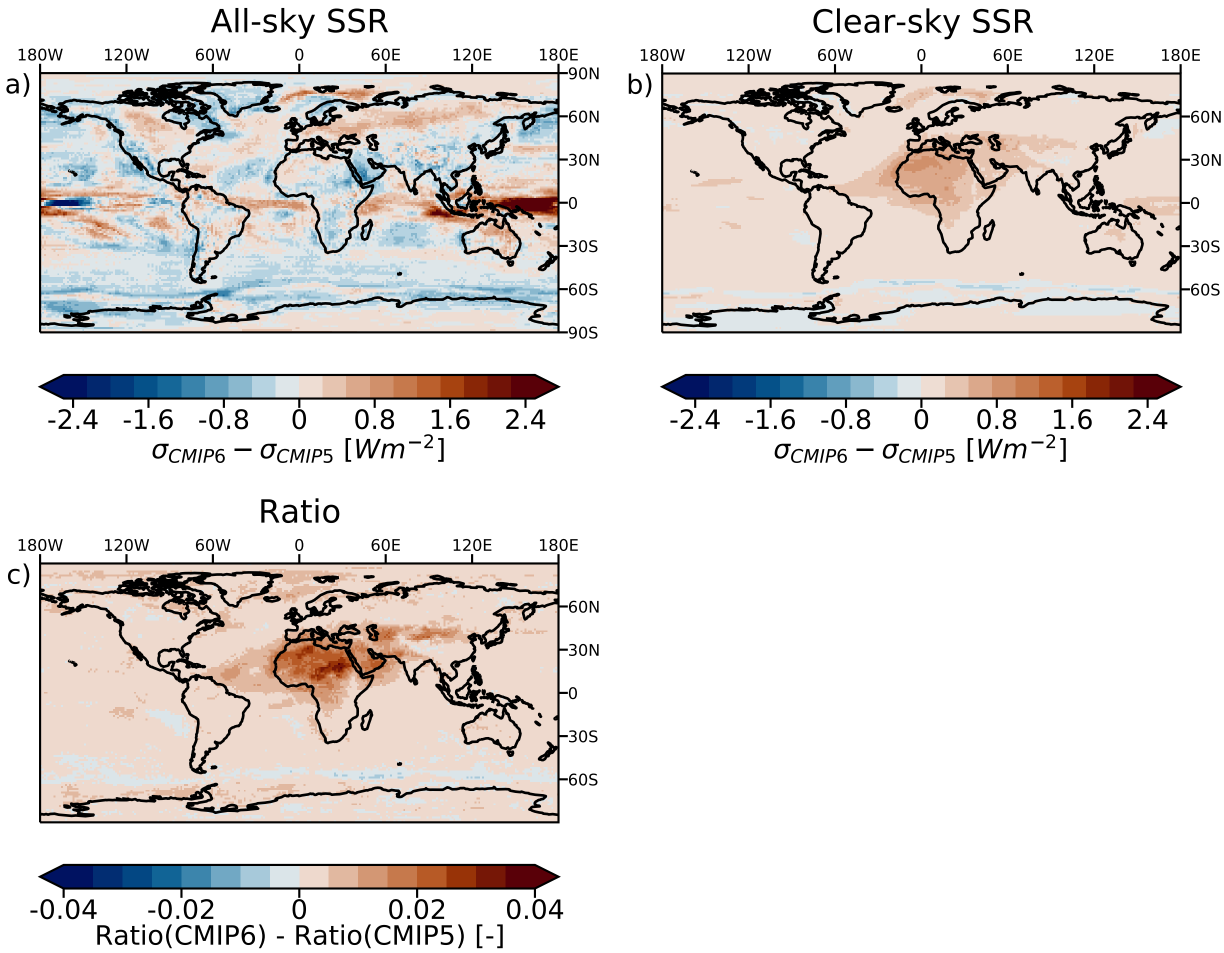

Figure 6. Difference between CMIP6 and CMIP5 multi-model median of the standard deviation for annual mean all-sky SSR (a) and for annual mean clear-sky SSR (b); Difference between CMIP6 and CMIP5 median values of the ratio between 30-years' clear-sky to all-sky trends (c).

Table 1. A comparison of the standard deviation of the SSR time series for all-sky – $\sigma_{ts,as}$, clear-sky - $\sigma_{ts,cs}$ and their ratio globally (top) and for Lindenberg, Germany (bottom). Global means are calculated as the mean of the standard deviations of annual mean data per grid cell, weighted by the grid cell area. For Lindenberg, the value from the corresponding box from the 0.5° model grid is used (nearest neighbour remapping). CMIP6 values are derived from the multi-model median.

	$\sigma_{ts,as}$[Wm^{-2}]	$\sigma_{ts,cs}$[Wm^{-2}]	$\frac{\sigma_{ts,cs}}{\sigma_{ts,as}}$[-]
	Global		
CERES EBAF Ed4.1	4.40	1.29	0.13
CMIP6 piControl	4.85	0.74	0.15
	Lindenberg		
BSRN station no. 12	5.15	1.61	0.31
CERES EBAF Ed4.1	5.64	1.37	0.24
CMIP6 piControl	5.53	0.75	0.14

Tropical Pacific Ocean and in sea ice areas, compared to models with prescribed aerosols. On the contrary, for clear-sky SSR, models with interactive aerosols show smaller in size areas with autocorrelation of the yearly averaged values. In general, there is no standardized way of prescribing aerosol loads (Meehl et al., 2020) and models with prescribed aerosol can resemble the behaviour of those using an interactive treatment, thus the similarities in the patterns of clear-sky SSR trends between Figures 4 b) and d).

The representation of atmospheric and oceanic modes of variability in models is an essential part of the description of the overall variability of the climate system. However, such modes are often subject to autocorrelation in time, which adds an additional uncertainty to the analytical link between the magnitudes of trends of different lengths (Thompson et al., 2015). Regions with a high autocorrelation in time generally show trends of larger magnitudes and possess a larger inter-model spread in both all-sky and clear-sky SSR trends, as it can be seen from Figure 2. The hatched areas in the Tropical Pacific for clear-sky SSR is due to a positive autocorrelation in time above 0.3 and suggests a relationship between ENSO events and clear-sky SSR. The Pearson correlation coefficient, calculated between the Nino 3.4 index (Trenberth & National Center for Atmospheric Research Staff (Eds), 2020) and clear-sky SSR (annual means), shows a strong negative correlation between the two variables in the Tropical Pacific regions with values below -0.7. Further analyzing the shapes of the distributions in the Tropical Pacific area, we find that there are more grid cells which pass the statistical tests (i.e. have a Gaussian distribution) in clear-sky SSR than in all-sky SSR. A thorough analysis of the relationship between SSR variability and climate modes of variability may provide interesting further causes of decadal scale SSR variability but is no longer pursued in the present paper.

A next step in the analysis of model differences is to look into individual models and place them into an observational context. While the models' unforced control runs include only processes internal to the climate system, the observational data includes in addition anthropogenic forcing (e.g. aerosol emissions from industrial units) and natural forcing (e.g. volcanic eruptions), therefore one would expect trends of greater magnitude (i.e. larger σ_{ts}) from the latter. Another difference is the length of the time series over which the standard deviation is calculated – the piControl simulations range from 500 to 2000 years, while the observational period captures only a few decades, which would result in a larger uncertainty of σ_{ts}. A third difference is that clear-sky SSR in models is calculated from a dedicated call of the radiation scheme, while in observations it is derived, based on the SSR in cloud-free days (Wild et al., 2018). As observational reference, we use the satellite-derived gridded data from CERES EBAF Ed4.1. for Earth's surface: the value $t(75, 30)$ is calculated, based on σ_{ts} of the yearly average SSR per grid box for the period 2001-2020. To compare individual models and observations-derived gridded data, we introduce box plots, comprised of $t(75, 30)$ within all grid boxes. Results for all-sky, clear-sky SSR and their ratio are shown in Figure 5.

For all-sky SSR, model median values (short orange lines) are close to their multi-model mean (horizontal red line) with a relative standard deviation of 11%, yielding $t(75, 30) = 0.64 \pm 0.07$ Wm^{-2}/decade. Contra intuitively, but in line with the findings of Folini et al. (2017) for CMIP5, the distribution of all-sky SSR trends in CERES-EBAF data (with a median value of $t(75, 30) = 0.56$ Wm^{-2}/decade) lies lower than the multi-model median of CMIP6 unforced simulations. On average, $\sigma_{ts,as}$ in CERES-EBAF is 9% less than the one for CMIP6 multi-model median (numerical values are shown in Table 1). This discrepancy can be attributed to the shortness of the observational period, which results in a large uncertainty in the calculation of σ_{ts}. Focusing solely on individual models, they tend to have different upper and lower limits of trend magnitudes: the relative standard deviation from the 90th percentile mean is 0.16 (0.22 Wm^{-2}/decade in absolute units), while from the 10th it is 0.21 (0.02 Wm^{-2}/decade). Models with higher upper limits like CESM2, CMCC, CNRM-ESM2-1, NorCPM2 and NorESM2, treat aerosols interactively. A further step in analyzing these differences is to look at the spatial patterns of trends for each model individually – shown on Figure A1. The largest unforced all-sky trends for all models occur in the Tropical Pacific. It is interesting to note that models differ also in the extent of areas, where the trends distribution is not Gaussian. These are the CMCC and EC-Earth models in the North Atlantic region. Almost all models, with the exception of CESM2, show autocorrelation in time of yearly averaged all-sky SSR around Antarctica.

In the clear-sky SSR trends, the absolute differences are small, but the relative differences (computed via dividing by the multi-model median per grid box) among models are larger. The relative standard deviation from the multi-model mean is 0.18 (0.02 Wm^{-2}/decade in absolute units), from the mean of all 90th percentiles – 0.32 (0.07 Wm^{-2}/decade), and from the mean of all 10th percentiles – 0.52 (0.01 Wm^{-2}/decade). The large spread in the lower limit is mostly due to the EC-Earth model family, which stands out with stronger clear-sky trends. Further examining the spatial patterns per model on Figure A2, one can see that unlike the all-sky trends, the places on the globe with the strongest clear-sky trends differ substantially from one model family to the next. Earth system models, coupled to an aerosol model, tend to have stronger trends above desert regions, even though BCC-ESM1 and the GFDL models use prescribed aerosols and show a similar pattern. The CanESM5 models additionally exhibit stronger trends in Southeast Asia. The EC-Earth models stand out with very large trends in the entire Northern Hemisphere; they also show a lot of fine structure, possibly due to their relatively high spatial resolution, compared to other models. The hatched areas on Figure A2, which show the regions of applicability for individual models, are mainly due to autocorrelation in time of clear-sky SSR above oceans, but the CMCC and EC-Earth models again get rejected by the statistical tests for Gaussian distribution in the North Atlantic. The KACE-1-0-G model shows high autocorrelation in time (up to 0.7) in most of the globe. In summary, the spatial patterns of clear-sky trends differ substantially among models in comparison to the all-sky trends, which explains the large spread on Figure 2 f). For the clear-sky case, CERES-EBAF data indicate trends of larger magnitudes compared to unforced trends in models as the CMIP6 multi-model median is close to the lower 25th percentile of CERES-EBAF trends (0.10 Wm^{-2}/decade). On average, $\sigma_{ts,cs}$ in in CERES-EBAF is 74% more than the one for CMIP6 multi-model median (see Table 1) – this difference may be attributed to the presence of anthropogenic forcing in CERES-EBAF data.

Turning to the ratio between clear-sky and all-sky trends, shown on Figure 5 c), one can see that the absolute and relative differences among models are of the same magnitude as those of clear-sky (Figure 5 b). As in the clear-sky case, the ratio between clear-sky and all-sky trends in CERES-EBAF is larger than that of the CMIP6 models. The relative standard deviation from the multi-model mean is 0.21 (0.03 Wm^{-2}/decade), from the 90th percentile mean – 0.40 (0.14 Wm^{-2}/decade), and from the 10th – 0.39 (0.02 Wm^{-2}/decade). The lower limit (with the exception of EC-Earth models) tends towards 0, indicating that the contribution of cloud variability to decadal all-sky trends is dominating in a larger fraction of the planet. If we look at absolute upper limits, instead of 90th percentile values, all models contain areas, where the clear-sky trends are as large as the all-sky trends, but no more than 10% of the globe represent such areas (i.e. where the ratio is close to 1). Within the EC-Earth models, the clear-sky trends are more dominant, attributing to a fraction of 0.8 to the all-sky trends in 10% of the globe. Figure A3 shows the spatial patterns of the ratio. It is evident that almost all CMIP6 models (with the exception of BCC-CSM2. CNRM-CM6, FGOALS and GISS) show a prevalence of clear-sky trends in the Sahara region. The ratio reaches different maximum values within models, but the patterns over land are similar. The EC-Earth model family stands out with a large clear-sky to all-sky ratio above land areas in the Northern Hemisphere, the Arctic and Antarctica.

Continuing further the analysis of model differences, we turn to the previous generation of climate models – CMIP5 (Taylor et al., 2012b). Differences between the 5th and 6th generations of climate models involve the parametrization of supercooled liquid water in clouds in some of the models. Thus, CMIP6 models are subject to larger cloud feedbacks (especially in the extratropical areas), which result in an increased equilibrium climate sensitivity (Zelinka et al., 2020; Dong et al., 2020). We check whether associated model changes result in different all-sky SSR trends in unforced CMIP6 simulations. For CMIP5, Folini et al. (2017) estimate $t(75, 30)$ between 0.15 and 1.07 Wm^{-2}/decade. Using the same methodology, we obtain ranges of the multi-model median, taken over all grid boxes, between 0.15 and 2.1 Wm^{-2}/decade, suggesting that CMIP6 models have regions with larger all-sky SSR variability than CMIP5 in their piControl runs. Similarly, we obtain the effective range of $t(75, 30)$ for clear-sky SSR to be $[0.04, 0.29]$ Wm^{-2}/decade for CMIP5 and $[0.04, 0.38]$ Wm^{-2}/decade for CMIP6, suggesting again regions with larger variability in CMIP6. The spread among models has increased from CMIP5 to CMIP6 for both all-sky and clear-sky SSR.

We next explore the regional differences between the two model generations. The differences between the multi-model median of the standard deviations of annual mean SSR in CMIP6 and CMIP5 are shown in Figure 6 for all-sky (a) and clear-sky (b). It is evident that all-sky SSR experiences a stronger variability in CMIP6 in the equatorial ocean areas with an area of decreased variability in the Central Pacific, which are possibly linked to adjustments of the representation of the ITCZ in CMIP6 (Tian & Dong, 2020). Notably, the desert areas show less all-sky SSR variability in the latest generation of models. The global average value of the difference between σ_{ts} of the two model generations (subtraction is performed per grid box) is close to 0, but with a standard deviation of 0.60 Wm^{-2}. The difference in clear-sky SSR generally has the same sign with CMIP6 models showing slightly larger clear-sky SSR variability than CMIP5. The mean value of the standard deviation $\overline{\sigma_{ts,cs}} = 0.14$ Wm^{-2} can be translated to $t(75, 30) =$ 0.0199 Wm^{-2}/decade. The largest differences between CMIP5 and CMIP6 clear-sky variability are observed in the desert areas with differences in $\sigma_{ts,cs}$ up to 0.95 Wm^{-2} in the Sahara region. This is possibly due to an increased number of models with an explicit aerosol representation in CMIP6. The relative contribution of clear-sky variability to all-sky variability has increased from CMIP5 to CMIP6. This effect is evident on Figure 6 c), which shows the difference of the ratio of clear-sky to all-sky SSR trends (as in Section 3.3) between the two model generations. The increased prevalence of clear-sky trends has increased the most in subtropical deserts, where the contribution of aerosols is significant and clouds are rare. A reduction in the clear-sky SSR variability around the Antarctic low pressure belt is also notable. On average, the standard deviation of all-sky SSR (thus the magnitude of unforced trends) remains almost unchanged (a decrease by 0.4% from CMIP5 to CMIP6), while the standard deviation of clear-sky trends is 24% larger in CMIP6, compared to CMIP5.

4 Discussion

The results of the present work give a quantitative estimate of the likelihood that a trend of a given length and magnitude at a specific location is entirely due to internal variability, based on the information contained in the unforced piControl runs from a large number of state-of-the-art climate models. We further want to explore the possibility of a given trend to be amplified or dampened by internal variability and demonstrate cases, in which the results from the present work can be applied. For this, we turn back to the statistical distribution of unforced trends for one grid box, corresponding to Lindenberg, Germany, on Figure 1. Knowing the statistics of the underlying distribution for that specific grid box, we can calculate the probability of occurrence of an unforced trend of a given magnitude and over a given time period. This is shown for all-sky and clear-sky SSR on Figure 7 a) and b) respectively. The contour lines are calculated using the CMIP6 multi-model median σ_{ts} for the grid box, corresponding to Lindenberg. It is evident that the probability decreases on longer timescales, but is still of importance for trends of smaller magnitudes. The magnitudes of clear-sky trends caused by internal variability are considerably smaller than all-sky trends for this location.

In order to put our theoretical analysis of unforced model simulations into a real-world context, we compare the standard deviations of the underlying time series of Lindenberg ground-based observations, CERES-EBAF and the CMIP6 piControl multi-model median at the same grid point (bottom panel in Table 1). The ground-based data from the BSRN station Lindenberg for both all-sky and clear-sky SSR data is available for a period of 18 years from 1995 until 2012. (clear-sky time series are derived using Long and Ackerman (2000)). We again note that the first two comprise forcing factors, while the CMIP6 piControl does not. For all-sky SSR, both the BSRN observational site and CERES-EBAF data show less variability than CMIP6 piControl, but the three values differ by no more than 9%. This is again counter-intuitive, but can be explained by one of the following: (1) the 17 years of ground observations at Lindenberg are not enough to capture the internal variability that occurs on longer timescales, (2) we have a compensating effect of the forced trends. For clear-sky SSR, the observational site shows the largest value of σ_{ts} – almost twice as much as the one estimated from CMIP6 piControl. Overall, at this location the BSRN and CERES-EBAF data agree well and the comparison to CMIP6 piControl is in line with what was previously discussed for the global average values in Section 3.4.

Bearing in mind that there are discrepancies in our expectations for model data and observations, but in absolute terms they are of comparable magnitude, we compare our 2-dimensional probability function with yearly-averaged observations from Lindenberg (shown as gray points on Figure 7 c-d). Since the observational time series fall entirely within the brightening period in Europe (Sanchez-Lorenzo et al., 2015), we calculate the linear regression for the whole period. The resulting trend is +0.28 $\mathrm{Wm^{-2}}$/year for all-sky and +0.052 $\mathrm{Wm^{-2}}$/year for clear-sky SSR. Both trends are represented by the solid red lines on Figure 7. To compare the resulting trend with internal variability, we transform the 2-dimensional probabilistic functions from Figure 7 a) and b) to absolute value coordinates (instead of trends) and center them at the beginning of the trend lines on Figure 7 c) and d). The resulting plot shows how steep the trend line should be to escape the influence of internal variability in both exacerbating and suppressing manner. The probability for the all-sky trend to be entirely due to internal variability, inferred using statistics from CMIP6 unforced control runs, is 13%; the probability of

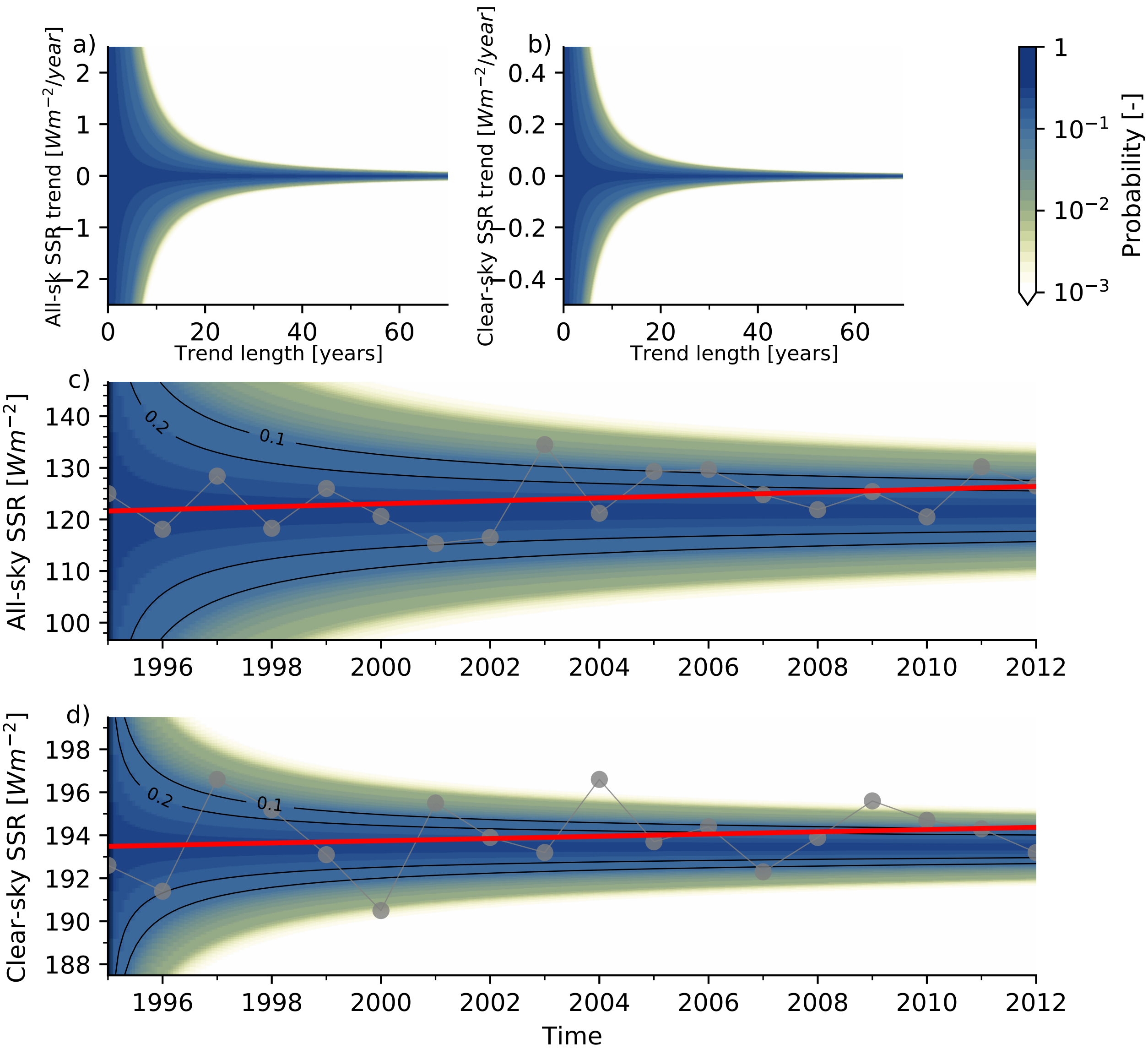

Figure 7. Probability of occurrence of a trend entirely due to internal variability as a function of trend length (x-axis) and trend magnitude (y-axis) for all-sky (a) and clear-sky (b) SSR for the grid box corresponding to Lindenberg, Germany (52.21°N, 14.122°E). Middle and bottom panels show the same probability coloring superimposed over observational all-sky (c) and clear-sky (d) SSR time series from the corresponding BSRN station – gray points indicate yearly average values, the solid red line is obtained from their linear regression; the probability color scheme is centered at the intersect between the linear fit and the beginning of the observational period.

the trend in clear-sky to be entirely due to internal variability is 6% (a script to calculate this for any trend magnitude and length, depending on the geographical location is linked in Acknowledgements). This difference in probabilities is because clear-sky shows 7.4 times less variability than all-sky SSR at this location according to the CMIP6 multi-model median.

We should also note that internal variability can act not only as a sole reason for SSR trends, but also exacerbate or suppress forced trends. If we assume linear additivity of trends, for Lindenberg there is 29% probability that the total 17-year all-sky trend is composed of +0.14 $\mathrm{Wm^{-2}}$/year forced trend plus an enhancing internal variability trend of +0.14 $\mathrm{Wm^{-2}}$/year. For the clear-sky case, the probability of 1:1 contribution of an anthropogenic trend (+0.026 $\mathrm{Wm^{-2}}$/year) and unforced trend (+0.026 $\mathrm{Wm^{-2}}$/year) is 22%. Likewise, it is probable that the observed trend at the BSRN station Lindenberg was enhanced by internal variability. However, internal variability can also have a suppressing effect and with the same probabilities are the observed trends could be 1.5 times larger (i.e. +0.42 $\mathrm{Wm^{-2}}$/year for all-sky and +0.078 $\mathrm{Wm^{-2}}$/year for clear-sky) if it were not for internal variability.

Until now, we explored two ways of reducing the effect of internal variability: looking at longer time series and looking into clear-sky SSR. We briefly address another way, which is analyzing composite time series from several location, as spatial averaging is supposed to also limit the effect of internal variability. Taking all-sky observations from 56 locations in Europe, Sanchez-Lorenzo et al. (2015) estimate European dimming and brightening trends to be: +0.96 $\mathrm{Wm^{-2}}$/year for 10 years (early brightening); −0.25 $\mathrm{Wm^{-2}}$/year for 35 years (dimming); and +0.32 $\mathrm{Wm^{-2}}$/year for 26 years (brightening) until 2012. We test the probability for each of these trends to be entirely due to internal variability, taking spatially averaged all-sky SSR data from unforced CMIP6 simulations.

The annual mean all-sky SSR piControl time series per model are interpolated to the 56 individual locations, used by Sanchez-Lorenzo et al. (2015). Afterwards, the time series are combined into one composite time series per model and for each model σ_{ts} is calculated. The CMIP6 multi-model median of $\sigma_{ts,as}$ is 3.05 $\mathrm{Wm^{-2}}$ with a spread of ±22% between models. For comparison, averaging out $\sigma_{ts,as}$, obtained from each station location in CMIP6 individually, yields 4.56 $\mathrm{Wm^{-2}}$ (ranging from 3.31 $\mathrm{Wm^{-2}}$ for Postdam, Germany to 6.06 $\mathrm{Wm^{-2}}$ for Lerwick, Great Britain), thus taking the composite time series of these 56 locations reduces the model-calculated internal variability by $\frac{1}{3}$ from the non-aggregated value. Using $\sigma_{ts,as}$ of the composite time series (3.05 $\mathrm{Wm^{-2}}$), we calculate that the probability the early brightening is entirely due to internal variability is around 0.2%, for the dimming period - 0.00005%, and for the recent brightening until 2012 it is around 0.002%. For comparison, we calculate the same probabilities using the mean $\sigma_{ts,as}$ of all individual locations (4.56 $\mathrm{Wm^{-2}}$), which yields 2.7%, 0.05%, 0.36% for the three periods respectively. The resulting probabilities are low due to their strong dependence on trend length ($N^{-3/2}$, see Figure 7-a) and the relatively long periods discussed in Sanchez-Lorenzo et al. (2015). Therefore, based on climate models' representation of internal variability, it is very unlikely that the entire periods of dimming and brightening in Europe are exclusively due to internal variability. However, it cannot be ruled out that trends at specific locations have been partially influenced by internal variability.

5 Summary and Conclusion

We quantify how much internal variability of all-sky and clear-sky SSR can contribute to decadal SSR trends at individual locations. Even though internal variability is regarded as a source of uncertainty in climate simulations, we analyze it as a stochastic process based on physical interactions within the climate system. For the statistical analysis, we use unforced multi-century CMIP6 simulations (piControl), which do not include any natural (e.g. volcanoes) or anthropogenic (e.g. greenhouse gases, aerosols) forcing. We construct the distribution of all possible SSR trends from 54 CMIP6 control simulations, each extending beyond 500 years of length. Trends are calculated as linear regressions from yearly averaged data. The resulting distribution is Gaussian and centered around 0 in 89% of the grid cells for both all-sky and clear-sky trends when looking at trends periods of 50 years or less. Additionally, no significant autocorrelation in time of the underlying annual SSR time series is found apart from the Tropical Pacific and areas covered by sea ice. Having a Gaussian distribution of trends and no autocorrelation in time of the original time series implies that trends of arbitrary length (and percentile) and σ_{ts} are analytically related.

For the analysis of regional differences in trends, we present a 30-year positive trend with 25% chance of occurrence (the 75th percentile of the distribution of all possible 30-year trends, $t(75,30)$) per grid cell. This variable is linked to trends of different lengths N and percentiles through: $t(p,N) = \sigma_N Z(p) \approx \sqrt{12}N^{-3/2}\sigma_{ts}Z(p)$. It is found that the magnitude of unforced trends is strongly dependent on the geographical region, taking values for $t(75,30)$ between 0.15 and 2.1 Wm^{-2}/decade for all-sky SSR and between 0.04 and 0.38 Wm^{-2}/decade for clear-sky SSR. The respective medians are 0.69 Wm^{-2}/decade for all-sky trends and 0.17 Wm^{-2}/decade for clear-sky trends. The variability of all-sky SSR is slightly more pronounced over the oceans in comparison to land areas, while clear-sky SSR shows larger variability above land, especially large dry desert areas with high natural aerosol content. Additionally, global climate models with an explicit aerosol representation show substantially larger decadal trends above deserts (up to 80%) compared to models with prescribed aerosols. Analyzing the ratio between clear-sky and all-sky trends provides an estimate of how relevant cloud variability is with respect to unforced all-sky trends. Regions, where cloud variability dominates (i.e. the ratio is close to 0), include large ocean areas, the ITCZ and Southwest China. On the other hand, clear-sky trends account for a larger fraction of the total trends (i.e. where the ratio is close to 0) above the large deserts in Africa and Asia due to low cloud amounts and high natural aerosol forcing. Averaged over all grid cells, the ratio between clear-sky and all-sky trends is 0.166, or unforced clear-sky trends are 6 times smaller than all-sky trends; this property is preserved at different trend lengths.

The CMIP6 inter-model spread can be used as an indication of the uncertainty of the present analysis. For all-sky SSR trends, the relative spread is ±32% and for clear-sky trends, it is ±43%. inter-model differences additionally provide information on which processes relevant for SSR trends the models disagree on and where in the world they diverge the most. Our analysis suggests that the largest disagreement among models is found in the regions with the largest magnitudes of trends. The absolute differences among models are the largest when comparing all-sky SSR trends, while the relative differences are more substantial in clear-sky trends and the ratio of the two. A larger spread is observed in models with interactive aerosols, suggesting that SSR trends in climate models are affected by the insufficient knowledge concerning aerosol processes.

Finally, we discuss applications of the current work to the analysis of observational time series. The quantitative estimates allow us to assign probabilities to the anthropogenic and the unforced fraction that must exist in a given observed trend. The benefit of using time series comprised of multiple locations to suppress the effect of internal variability is demonstrated by the $\frac{1}{3}$ decrease of the standard deviation of the SSR trends distribution when spatially averaging the SSR time series over 56 locations across Europe. Using σ_{ts} from CMIP6 unforced control runs, we are able to show that internal variability is extremely unlikely to have been the sole cause of dimming and brightening in Europe, but its influence at individual locations for shorter periods should not be neglected. A further analysis of the spatial scales of internal variability and its contribution in composite time series is a possibility for a future study. The results from the current work can be used to attribute a probability to a trend of certain length, magnitude, and location to be due to internal variability.

Appendix A Model specific maps

Acknowledgments
This work was funded by Swiss National Science Foundation grant no. 200020_188601. We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies who support CMIP6 and ESGF. The CMIP6 data may be accessed at `https://esgf-node.llnl.gov/projects/cmip6/`. The CERES data may be obtained at `https://ceres.larc.nasa.gov/data/`. Observational data from the BSRN network is available at `http://ftp.bsrn.awi.de/`. The data analysis and visualization have been performed using exclusively free and open source software (including Linux, Python and CDO): we would like to thank the community for building powerful tools from which everyone can benefit. The script used to calculate the probability of an unforced trend of a certain length and magnitude at a specific location is available at `https://gitlab.com/mjade/internal-variability-of-all-sky-and-clear-sky-ssr-on-decadal-timescales`

References

Chiacchio, M., & Wild, M. (2010). Influence of nao and clouds on long-term seasonal variations of surface solar radiation in europe. *Journal of Geophysical Research: Atmospheres*, *115*(D10). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2009JD012182` doi: https://doi.org/10.1029/2009JD012182

Deser, C., Phillips, A., Bourdette, V., & Teng, H. (2010, dec). Uncertainty in climate change projections: the role of internal variability. *Clim Dyn*, *38*(3-4), 527–546. Retrieved from `https://doi.org/10.1007%2Fs00382-010-0977-x` doi: 10.1007/s00382-010-0977-x

Dong, Y., Armour, K. C., Zelinka, M. D., Proistosescu, C., Battisti, D. S., Zhou, C., & Andrews, T. (2020). Intermodel spread in the pattern effect and its contribution to climate sensitivity in cmip5 and cmip6 models. *Journal of Climate*, *33*(18), 7755 - 7775. Retrieved from `https://journals.ametsoc.org/view/journals/clim/33/18/jcliD191011.xml` doi: 10.1175/JCLI-D-19-1011.1

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016, may). Overview of the coupled model intercomparison project phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.*, *9*(5), 1937–1958. Retrieved from `https://doi.org/10.5194%2Fgmd-9-1937-2016` doi: 10.5194/gmd-9-1937-2016

Farquhar, G. D. (2003, mar). ATMOSPHERIC SCIENCE: Pinatubo, diffuse light, and the carbon cycle. *Science*, *299*(5615), 1997–1998. Retrieved from `https://doi.org/10.1126%2Fscience.1080681` doi: 10.1126/science.1080681

Folini, D., Dallafior, T. N., Hakuba, M. Z., & Wild, M. (2017, jan). Trends of surface solar radiation in unforced CMIP5 simulations. *J. Geophys. Res. Atmos.*, *122*(1), 469–484. Retrieved from `https://doi.org/10.1002%2F2016jd025869` doi: 10.1002/2016jd025869

Gidden, M. J., Riahi, K., Smith, S. J., Fujimori, S., Luderer, G., Kriegler, E., ... Takahashi, K. (2019, April). Global emissions pathways under different socioeconomic scenarios for use in CMIP6: a dataset of harmonized emissions trajectories through the end of the century. *Geoscientific Model Development*, *12*(4),

Figure A1. Maps of the 75th percentile of 30-year all-sky SSR trends for each model considered. The hatched areas represent the grid boxes, which fail the K-S test for the 30-year trends or have an autocorrelation in time of the original time series above 0.2. Asterisk next to model name indicates the use of prescribed aerosols. The individual plots are produced in the native model grid (without interpolation).