Summary of changes

We thank both reviewers for their constructive comments and detailed reading. In response to the suggestions by two reviewers we have

- updated the plots, enhancing Fig. 6 and SFig. 1
- revised text throughout the manuscript to bring out key findings better, differentiate local and global, continental vs. oceanic variability changes and improve clarity
- enhanced the discussion on ENSO changes, temperature-precipitation relationships and potential hydroclimate constraints
- added three new supplementary figures to support the additional discussion
- corrected spelling.

A detailed response to the helpful remarks of the referees is given below.

1 Reply to the first reviewer

(Original report cited in italics) The analysis seeks to quantify changes in variability in both temperature and precipitation as a function of the baseline climate state by using simulations from a broad range of climate experiments. The manuscript is well written and logically organized. The core scientific objectives are well articulated. The approach and simulations used are appropriate for the science questions posed. The figures are well-designed generally and nicely illustrate some new and useful results. We thank the reviewer for this positive assessment.
My main concerns relate to the summaries provided for the results and the manner in which the authors gloss over some of what I find to be the key findings. I also suggest a reconsideration of the figure layouts to more strongly align with the structure of the text (Figs 2,3).

In the revised document we reworked the summary paragraphs and discussion in order to bring out the key findings better. We introduce the comparison across states, that underlies the layout of our figures earlier (in particular Fig. 2 and 3), which should improve the alignment between visuals and text.

Action: Changed introduction to results section to support the figure layout.

I recommend splitting the abstract into two paragraphs to make the summary clearer - one for what is done and one for what is found.

We agree with the reviewer that this improves clarity and have split the abstract accordingly.

Action: Done.

- p1-10: based on Fig 2, I’m not sure that I find this summary of a decrease in variance for increasing temperature to be true, particularly over land. For example if I am interpreting them correctly, Figs 2g/h show strong positive increases, particularly over land from 40N-S where they are associated with significant impacts. There are various statements in the text that seem at odds with the passage as well (e.g. p12-24-30). Would it make sense to parse this statement a bit more to relate clearly to specific results and distinguish between regions of coherent change that contrast (i.e. land/ocean).

We thank the reviewer for her/his careful inspection of the results. Indeed, temperature variability does not show a uniform decrease globally, and our statements will be made more specific in the revised document. We observe differing trends over continents and oceans at mid-to-low latitudes, with warming associated with increasing temperature variance over the continents of Africa, South America, the maritime continent, Australia and South-East Asia, as well as over South-West Europe and the Southern United States. Decreasing temperature variance is found over the high-latitude continents and the world’s oceans except for the Northern Atlantic and over the Indian Ocean. Globally averaged, the local changes point at a decrease in temperature variance. Local changes are generally consistent across timescales, which we demonstrate in the revised document with land/ocean spectra.

Action: We added a clarifying sentence in the abstract, including the land/ocean discrepancy. To support the results and discussion section, three new supplementary figures were introduced.

p1-14: is ‘dominating rainfall variability’ appropriate - do they explain the vast majority of variance? across what timescales?
Indeed, this is an ambiguous phrase and difficult to pinpoint in models or data. We now clarify that our analysis is at the annual timescale, and to demonstrate that the large-scale modes influencing variability at annual timescales remain stable. We have revised the text in the abstract to better capture the intended meaning: ‘By compositing extreme precipitation years across the ensemble, we demonstrate that the same large-scale modes influencing rainfall variability in Mediterranean climates persist throughout palaeoclimate and future simulations.’ These modes correspond to extreme precipitation years in the ensemble, identified with composites over the fifty-year time series. Thus, the patterns robustly capture the atmospheric state during annual precipitation extremes (in this case, one standard deviation above/below the mean in annually-averaged precipitation). These are the patterns that emerge, but since this is not an analysis based on Empirical Orthogonal Functions (EOFs) it does not provide the amount of variance it explains. We clarify this in the revised document.

Action: Revised sentence in p1-14. Revised corresponding results section.

- p2-8: the sentence seems to suggest that internal variability is distinct from natural variability?

That was not our intention and we have revised this sentence. We had meant to use ‘internal’ when discussing models and ‘natural’ when discussing the real system.

Action: Changed.

- p2-21: scale linearly? - wording seems to suggest so

We wrote ‘In any region, damages do, however, scale with increased variability (Katz and Brown, 1992; Alexander and Perkins, 2013)’. This should imply a direction, but not a qualitative statement on linearity. Depending how damages are estimated (e.g., Katz and Brown (1992) base their statements on a threshold model for crop yields) the increase may follow a different (e.g., exponential) form. Therefore we rephrased this to ‘are expected to increase with greater variability’.

Action: Changed.

- p2-30: isn’t there also evidence for increases in variability on some timescales? such as ENSO teleconnections?

We agree with the reviewer that, in the literature, increases in variability have been discussed for specific climate variables, time scales and regions. This includes several studies that suggest ENSO variability may increase (e.g., Cai et al., 2018). We will include a brief discussion of potential ENSO variability changes in the revised document, as the ENSO timescale does show unclear changes (c.f. the power spectra in Fig.6). In particular, it is noticeable that in this power-spectral range some models show a shift in the ENSO frequency, resulting in a peak-and-trough-pattern in the temperature spectral ratio. A change in the ENSO pattern, on the other hand, would not necessarily show up in the spectrum, if the overall variance at the timescale does not change. Amplitude changes are similarly
difficult to pinpoint and account for between models. Therefore, for a full discussion of the ENSO changes in CMIP6/PMIP4 we will refer the reader to the ENSO-centered paper currently in open discussion: Brown et al. (2020).

**Action:** Expanded discussion section.

- **p3-8:** Precipitation changes are also strongly linked inversely to temperature. Wouldn’t this therefore be a source of increased temperature variability? There is associated literature on the topic that should be discussed and cited.

Indeed, the inverse link between temperature and precipitation has been discussed in the literature (e.g. in the landmark studies of Allen and Ingram, 2002; Adler et al., 2008; Trenberth and Shea, 2005). It is clear that, in particular at daily to interannual timescales soil moisture plays a relevant role in the precipitation feedback on temperature variability (Vidale et al., 2007; Fischer and Knutti, 2013). It is, however, also clear that models have difficulties representing these feedbacks at the land surface, in particular on longer timescales (Rehfeld and Laepple, 2016). The detail of representation of sub-grid-scale convective processes could also determine whether a local feedback is modeled positively or negatively (Hohenegger et al., 2009). We appreciate the suggestion and add a section on the precipitation-temperature linkage to the discussion.

**Action:** Amended introduction and the discussion section.

- **p7:** For many of these experiments, multiple ensemble members are available. I don’t see mention of how many members are used? If only 1, that should be made explicit on page 3. If more than 1, that too should be made explicit and the approach for avoiding overweighting individual models should be described. Despite the additional work, there does seem to be merit in consideration of all available members to address various questions on the role of noise in the results - some listed below.

We are sorry if we had failed to specify this. We have only used a single ensemble member for each model (generally r1i1p1f1) – this information will be included in the revised document. This approach has been adopted for two reasons. Firstly, it is cleaner, as the reviewer notes it does not overweight individual models in the computation of the ensemble means. Secondly, there is a very low number of the palaeoclimate simulations which have multiple ensemble members.

**Action:** Statement on ensemble members added.

- **p8-10:** It seems sufficient to merely cite the CVDP rather than each script invoked by it since the CVDP documentation covers this.

OK, the script names were removed from the text.

**Action:** Done.

- **p12-14:** again referencing the work here that has been done on temperature precipitation relationships seems appropriate.
Thank you for the suggestion. We add a reference to the established literature at this point.

Action: Done.

- p12-30: I don’t think the global pattern correlation tells the key components of the story. From 40N-40S it seems clear that the PCs are positive.

Our aim was to provide the global pattern correlations to provide additional support for the ensemble-mean figures that we show. In the revised manuscript we complement the global statements by a refined local view, in particular for the tropical to subtropical land areas. To this effect we have added three new figures to the supplementary material which underline these statements.

Action: As indicated above, the discussion was amended.

- p13: The caption for Fig 2 should be explicit regarding whether it is the global mean temperature change that is used to compute the ratio or the regional change.

We rephrased the caption to clarify that we are using the gridbox-scale change for the ratio.

Action: Done.

- a number of the CVDP variable names are used - which are long and likely not familiar for many readers. I’d recommend creating acronyms for these so that they can be shortened and are more intuitive.

We agree with the reviewer that the CVDP variable names are long. We have removed the names of the CVDP scripts from the text, so the variables names now only occur on Fig. 3. We feel that the caption of Fig. 3 is sufficiently detailed to make definition of the acronyms superfluous.

Action: No change.

- Fig 4: how do you estimate your degrees of freedom in computing P-values? There is mention of 500-DOF in discussion of LGM but that is clearly excessive given the strong mutual dependencies across models, no? Perhaps a more stringent estimate is warranted?

Indeed, we assume 500 degrees of freedom for the spatial field pattern correlations (out of a total of $180 \times 360 = 64800$ grid boxes) across the fields shown in Fig. 2 and Fig. 3. However, for the regression lines in Fig. 4, which shows the changes in (mode) variability against global mean temperature change in the simulations, we assumed all model simulations and models to be independent. This results in 60 degrees of freedom (with 61 simulations contributing to the regression). We clarify this in the caption in the revision.

Action: Done.
- I have a general suggestion regarding the structure of the figures. Since the text is structured to discuss T/P of each experiment why organize the figures to show only T for all experiments and then P for all experiments. Particularly given the mutual relationships that exist, I find merit in having one figure of 4 panels for each experiment - for a total of 4 figures of 4 panels rather than 2 figures of 8.

The reviewer is correct, we currently do not follow the subpanel figure order in the discussion. However, our focus is on general relationships across different modeled states. Therefore we add a paragraph in response to this in the beginning of the results section that discusses the idea of comparing relationships across the experiments (from cold to warm, from mean to variance change, from temperature to precipitation). With 4 figures of 4 panels it is more difficult to make out the similarities/differences across the experiments and variables.

Action: Revised introduction to results.

- Fig 4: what is the contribution of internal variability versus model structural contrasts to the scatter in each panel? can multiple ensemble members, where available, be used to estimate a contribution range? I think this would provide key context for interpreting the figure.

We appreciate the suggestion. Multiple ensemble members are, unfortunately, not available for the palaeoclimate simulations. Nevertheless, we agree that the contribution of internal variability is an important factor to consider. Therefore we utilize the long preindustrial control experiments to estimate the contribution of internal variability. This is then added to Fig. 4 as confidence intervals around the unity line.

Action: Estimated magnitude of internal variability and added it to Fig. 4.

- Figure 5: I imagine the “W” in the titles corresponds to West? If so I’d spell it out to avoid confusion with “Wet”. Also what justifies the selection of the regions? They are much smaller than the climate zones they are intended to represent. Their small size suggests they may be particularly subject to internal variability rather than structural differences across models or experiments.

We have revised the figure titles as suggested. The regions are based on the Köppen climate classification of Mediterranean climates, and in particular the western boundaries of continents wherein the extratropical climate appears to cause precipitation anomalies of different signs than the global mean change from the pre-industrial (see Fig. 3). The selected boxes are actually larger than these regions strictly defined (and in all cases encompass multiple grid boxes), and our analysis in fact shows that the same modes of variability are important across different climate states. In the revised document we add supplementary figures that show that the circulation patterns are robust across climates (except for Western South Africa, where there is no signal).
- Figure 6: I suspect that the global mean again masks some important regional effects. How might the results change for land between 40N-S? Indeed, as suspected by the reviewer and as the total variance changes in Fig. 2 in the manuscript make clear, the global mean variance change differs from that over the low-to-mid-latitude land areas. Averaging the spectra over land areas between 40S and 40N we have less clear changes, and most importantly find indications for higher temperature variance in the warm experiments than in the preindustrial control. There is, however, also slightly more temperature variance over these areas in the LGM (cold) experiment than in the preindustrial control across the spectrum. Definitive statements are complicated by the fact that there is less intermodel agreement. We expanded the results section, and the discussion to take this into account. To support this discussion we added three figures to the supplementary material which show the power spectra over land areas, globally and from 40S to 40N, as well as the ocean average to support the discussion of Fig. 6.

Action: Revised.

- p12-13: what is meant by “meridional atmospheric gradient modes of variability”? Is this referring back to results in Section 3.4? Might make this reference more explicit.

Indeed, this is a reference to SAM/NAM and NAO results in Sect. 3.4. We make this more clear in the revised document.

Action: Done.

- p17-11: Is the lack of consistency the result of the choice of such small regions?

The lack of consistency only occurs for the South African case. This suggests that the Mediterranean regions are generally appropriately sized. Please refer to the response above regarding Fig. 5.

Action: as above.

- p17: There doesn’t seem to be any rationale for the organization of paragraphs. Perhaps make one for each region? - p17: After reading Section 3.4 I don’t seem to have much of an understanding of the robustness from past to future climate - the stated goal of the section.

We agree, the previous section title was misleading. We have changed this and revised the text to better explain the results. In addition, we have included additional supplementary figures that better illustrate the robustness of the relevant patterns from past to future climates (with the exception of South Africa). The organization of this section is based on that of the relevant figure, and proceeds between the regions from southwestern South America to the western Mediterranean.

Action: Revised paragraph.

- p19-5: Why combine land/ocean regions? I think a distinction should be made, par-
We agree with the reviewer that a distinction enhances the discussion and therefore added three new supplementary figures as indicated in the response above.

Action: See above.

- p22-14: What is meant by a reduced ENSO? reduced variance in Niño3.4 SSTs?
  Yes. We shall specify that in the revised manuscript.
  Action: Corrected.

- p22-22: “mean strength”? Note that much of the analysis examined changes in generic variance and then changes in the indices themselves. What is left out is the change in teleconnection strength. Shouldn’t this be considered? and isn’t it perhaps more important than the changes in the indices themselves?
  We understand the reviewer’s comment and sympathise with its sentiment. However, analysing the changes in the teleconnection patterns is not trivial and requires specifying certain decisions that may not be appropriate for all modes. We feel that to do so rigorously could require an individual manuscript for each mode, and therefore is not appropriate for this paper. We further explain this issue using ENSO. The teleconnection changes in these simulations have been explored in Brown et al. (2020) ENSO teleconnections are often computed using composites - there were serious issues dealing with the changes in the mean state conflating with the teleconnections changes (after Cai et al. (2014)). Additionally ENSO teleconnections may not be reciprocal for the different phases.
  Action: Amended discussion.

- p22/23: Perhaps cite the figures and panels that support each statement as such references are at times unclear. Some figures seem to clearly contradict the statements made.
  We worked through the manuscript again to ensure that the statements cover both the global and the regional scale to avoid misunderstandings.
  Action: Added reference to figures

2 Reply to the second reviewer

(Original report cited in italics)

Thank you for asking me to review paper: “Variability of surface climate in simulations of past and future” by Rehfeld et al. Do please accept my apologies for the delay in returning this review. Any high-quality paper on climate variability is useful, and to say the obvious, it is changes to inter-annual variability that could as much of an effect on
society as background climatic changes. This paper provides quite critical information on how the climate system might evolve by a careful scanning across available climate model simulations. The Abstract is clear and captures what the analysis does. The paper builds on what is an under-utilised resource of paleoclimate simulations. The Reference list is comprehensive, and that in itself makes the paper useful to the climate modelling community.

We thank the reviewer for this positive assessment.

A few comments: The research has been undertaken well, and so I can only really offer a few points which the authors might like to consider. (1) The decrease in local variability as global temperatures increase is always a fascinating feature of the climate system. This reduction also goes against much-perceived wisdom that a warmer world will be a more climatically-volatile world. The authors might like just to note that, possibly in the discussion?

Indeed, the findings at the global scale do contradict the intuitive expectation (rooted in the molecular physics of gases, perhaps, in the Maxwell-Boltzmann-distribution?). In the revised manuscript we will differentiate more strongly between the global and regional scales in the discussion, and will take up this suggestion.

Action: Discussion expanded.

(2) The approach taken is predominantly statistical, which is correct and proper. However, ultimately it would be nice to understand better the background physical processes behind all of the discovered correlations and features. This understanding is not easy when using outputs from climate models developed at research centres across the global, because it can be difficult to “get inside” the models for extra diagnostics. However, a few sentences saying that this analysis could trigger future investigations of the driving processes might help (and possibly with references). For instance, one suggestion is that lower sea-ice coverage in a warmer world will suppress yearly variations in temperature – fitting with the findings of this paper. Other authors have investigated “teleconnections” between the key oceanic forcings and related adjustments to meteorology over land areas. Some of these authors will have offered how atmospheric advection has a role to enforce such connections.

We absolutely agree with the reviewer in that a better understanding of the physical mechanisms of changing climate variability is crucial to understand our results. Some research on this exists, but a conclusive view across regions, seasons and timescales is difficult. On interannual timescales, sea-ice extent has been shown to correlate with global temperature variability (Huntingford et al., 2013). However, it remains unclear whether this would remain to be the case if a summer-ice-free Arctic has been reached, and how it influences low-latitude climate variability. A key role from the seasonal (Holmes et al., 2016) to the millennial (Rehfeld et al., 2018) timescale is certainly played by the meridional temperature
gradients that modulate atmospheric flows. However, due to the turbulent nature of the atmosphere, changes to the contributions of latent and sensible heat transport to mid-to-high latitude temperature variability are difficult to assess (Schneider et al., 2015). Therefore, as the reviewer notes, better understanding of the background physical processes behind the correlations and features is required. Our analysis and results therefore clearly calls for extending future research on the driving processes of variability changes. We add this to the discussion and conclusion of the manuscript.

Action: Added in discussion.

3) As so much of this paper describes common features between Earth System Models, then maybe at least some sort of mention should be made of the Emergent Constraint (EC) technique? ECs could potentially use the discovered inter-model agreements, in tandem with any additional contemporary measurement, to constrain future projections? Just a sentence or two hinting at this might be useful.

This is a good suggestion that we will adopt in the revised manuscript. There have so far been few examples of variability-based observational constraints (e.g., Cox et al., 2018). We will add this idea as motivation in the introduction, and then return to the theme in the discussion.

Action: Added.

4) There are substantial sets of paleo measurements that are rarely used by the climate modelling community. Again, maybe for Discussion, but this paper, with its thoughtful aligning of both paleo and future climate simulations, illustrates their huge potential to constrain climate projection. In other words, if the past can tell us more about the future (e.g. Figure 1, hydrological sensitivity is a valid statistic both for the past and the future), then any past records of simultaneous precipitation and temperature estimates provide valuable extra information.

Again this is very useful comment. It is something that we have started thinking seriously about. Highlighting the potential in our revised discussion will not only make the manuscript stronger, but help motivate our own future research. There are some methodological issues that need to be resolved before it can be deployed in earnest though. Crucially, obtaining joint (or closeby) and robust estimates of temperature and precipitation from proxy data is a fundamental challenge (Rehfeld and Laepple, 2016; Rehfeld et al., 2016).

Action: Considered in discussion.

5) One thing I especially like about the manuscript is the emphasis on oceanic modes of variability (ENSO, IPO, IOD etc). And this is obviously important given the paper is about variability. The authors will know (i.e. in numerical code) where the boundaries are. Would it be appropriate to give a map somewhere, with each of the oceanic modes of
oscillation marked? Most will know where ENSO is, but some of the others are less well known.

We agree with the reviewer that both atmospheric and oceanic modes of variability are important to consider. We have provided Supplementary Fig. 1 with the boundaries of the modes, and will highlight it in the manuscript revision.

Action: Highlight added.

(6) Do please work through the paper checking clarity. In general, the manuscript reads well, but in some places, it takes time to fully appreciate the analysis, along with a risk of ambiguity. In addition, the captions should be self-contained. As an example, the Caption for Figure 5, it takes some time to realise that the key point is for each location (as in the subplot headers) corresponds to high rainfall amounts The vague “selected regions” should be expanded more. Or even mark the epic-centre of each region with an annotated arrow for instance.

We have modified the caption. We also note that here, as in the original submission, the regions are marked by green boxes. Furthermore, we worked through the manuscript again to ensure each figure/caption is more self-explanatory.

Action: Caption modified.

(7) Some sentences are difficult to read. For instance, in the Conclusions “Global mean precipitation increases with temperature from cold to as-warm-as-preindustrial to warm scenarios.”. Maybe better something like: “Modelled global mean precipitation is found to increase as global temperatures also increase. This finding is valid for simulations from pre-industrial periods into a future warmer world, as adjusted by the burning of fossil fuels. However, our paleo-simulations also show this finding to be true, in the transition from colder periods to the warmer period at the beginning of the industrial revolution”.

We thank the reviewer for the detailed reading and this suggestion. We will re-phrase this sentence in the revision and are checking through the entire manuscript again to ensure more clarity.

Action: Rephrased.

(8) The diagrams are good and informative, but a little attention to formatting and detail could turn them into something exceptional. Just check the basics, such that in each, all annotation are clear and in sufficiently large font size. Figure 6, make it standard format - so remove the dotted lines maybe?

Thank you for this suggestion. The revised version of Fig. 6 follows a more standard aspect ratio, includes boxes around the panels and consistent label sizes.

Action: Formatting of Fig. 6 adjusted.
References


Variability of surface climate in simulations of past and future

Kira Rehfeld¹, Raphaël Hébert², Juan M. Lora³, Marcus Lofverstrom⁴, and Chris M. Brierley⁵

¹Institute of Environmental Physics, Ruprecht-Karls-Universität Heidelberg, INF 229, 69120 Heidelberg, Germany
²Alfred-Wegener Institute Helmholtz-Center for Polar- and Marine Research, Telegrafenberg A45, 14473 Potsdam, Germany
³Department of Geology and Geophysics, Yale University, 210 Whitney Ave, New Haven, CT 06511, US
⁴Department of Geosciences, University of Arizona, 1040 E. 4th Street, Tucson, AZ 85721, US
⁵Department of Geography, University College London, London, WC1E 6BT, UK

Correspondence: Kira Rehfeld (krehfeld@iup.uni-heidelberg.de)

Abstract. It is virtually certain that the mean surface temperature of the Earth will continue to increase under realistic emission scenarios. Yet, comparatively little is known about future changes in climate variability. We explore This study explores changes in climate variability over the large range of climates simulated by the Coupled Model Intercomparison Project Phases 5 and 6 (CMIP5/6) and the Paleoclimate Modeling Intercomparison Project Phase 3 (PMIP3). This consists of, including time slices of the Last Glacial Maximum, the Mid Holocene and idealized warming experiments (1% CO₂ and abrupt4 × CO₂), and encompasses climates with a range of 12°C of global mean temperature change. We examine climate variability from different perspectives: the perspectives of local interannual change, coherent climate modes, and through compositing extremes.

The change in the interannual variability of precipitation is strongly dependent upon the local change in the total amount of precipitation. Meanwhile only over tropical land is the change in the interannual temperature variability positively correlated to temperature change, and then weakly. In general At the global scale, temperature variability is inversely related to mean temperature change—with analysis of power spectra demonstrating that this holds from on intra-seasonal to multi-decadal timescales. This decrease is stronger over the oceans, while there is an increased temperature variability over subtropical land areas (40° S - 40° N) in warmer simulations. We systematically investigate changes in the standard deviation of modes of climate variability, such as including the North Atlantic Oscillation, the El Niño-Southern Oscillation and the Southern Annular Mode, with global mean temperature change. While several climate modes do show consistent relationships (most notably the Atlantic Zonal Mode), no generalisable pattern emerges. By compositing extreme precipitation events years across the ensemble, we demonstrate that the atmospheric drivers dominating same large-scale modes influencing rainfall variability in Mediterranean climates persist throughout palaeoclimate and future simulations. The robust nature of the response of climate variability, between both cold and warm climates and across multiple timescales, suggests that observations and proxy reconstructions could provide a meaningful constraint on climate variability in future projections.

1 Introduction
Slow and sustainable Knowledge of slow and sustained changes in mean climate conditions are important to understand for understanding climatic risks and uncertainties (IPCC-AR5, 2013). However, understanding changes in the variability around this the mean is at least as pressing as the of an issue as understanding of changes in mean climate for society and agriculture (Katz and Brown, 1992). This is because societal (Alexander and Perkins, 2013; Katz and Brown, 1992; Hsiang et al., 2013) and ecosystem (Seddon et al., 2016; Stenseth, 2002) impacts scale with climate variability, and increasing variability leads to increasing extreme events (IPCC-AR5, 2013; Schär et al., 2004).

Climate variability can be defined as variations in the mean state and other statistics (e.g. standard deviations, the frequency of occurrence of extremes) of temperature, precipitation and atmospheric circulation on spatial and temporal scales beyond individual weather events (Qin et al., 2014; Xie et al., 2015). Internal variability arises due to complex (often nonlinear) internal processes within the atmosphere-ocean-biosphere-cryosphere system (Deser et al., 2012a; Olonscheck and Notz, 2017) (Deser et al., 2012a; Olonscheck and Notz, 2017; Lofverstrom, 2020), or as forced variability in response to changes in natural or anthropogenic forcing (Foster and Rahmstorf, 2011). However, the actual evolution of climate combines anthropogenic forcing and natural climate variability (Horton et al., 2016), with internal variability dominating the local-to-regional synoptic evolution (e.g., Deser et al., 2012a; Wallace et al., 2015). In a simple stochastic model (Hasselmann, 1976), internal variability is proportional to climate sensitivity, and has been used to derive emergent constraints from temperature variability over the historical era (Cox et al., 2018). A core focus of research has been the investigation of major climate phenomena, modes of climate variability (Qin et al., 2014), such as the El-Niño/Southern Oscillation (Walker and Bliss, 1932; Bjerknes, 1966), and their contemporary change and representation by climate models (Deser et al., 2010, 2012a; Phillips et al., 2014). Their projected changes, and relevance for future regional climate evolution remain uncertain (Xie et al., 2015; Christensen et al., 2013). At the same time, atmospheric circulation changes contribute strongly to internal climate variability and, inherently, uncertainty of future projections (Thompson et al., 2015).

Trends established based on the instrumental record are uncertain, and both increasing (Hansen et al., 2012) or decreasing (Rhines and Huybers, 2013; Lenton et al., 2017) trends in temperature variability have been established reported. These trends differ amongst world regions (Rhines and Huybers, 2013; Huntingford et al., 2013): More economically underdeveloped areas were found to be more affected by are generally more vulnerable to increases in temperature variability than the more high-latitude developed regions (Bathiany et al., 2018). In any region, damages do, however, scale with increased climate impacts are expected to increase with greater variability (Katz and Brown, 1992; Alexander and Perkins, 2013). Therefore, there is a need to better understand changes to climate variability under warming. A warming similar to that projected for the next centuries (IPCC-AR5, 2013) occurred from between the Last Glacial Maximum (LGM, 27-19 thousand years before present, 27-19 kyrs BP) before until apparently stable Holocene climate conditions were reached (since 11.7 kyrs BP). Along with this warming, a reduction in centennial to millennial-scale temperature variability to a quarter of the glacial level was estimated based on palaeoclimate proxy data, and linked to the reduction of the local meridional temperature gradients (Rehfeld et al., 2018). Based on this mechanistic link, a continued decrease in temperature variability at the global scale could
be expected at long timescales (Rehfeld et al., 2018). It is, however, unclear how these long timescales link to affect the synoptic to decadal variability, which is not generally observable with palaeoclimate proxies. There is corroborating evidence based on model simulations for decreases in variability at interannual (Holmes et al., 2016) and longer (Brown et al., 2017) timescales. In particular, the contemporary decline observed reduction in Arctic sea-ice extent has been linked to declines in temperature variability at a global scale (Huntingford et al., 2013; Olonscheck and Notz, 2017; Bathiany et al., 2018). At the seasonal scale, higher temperature variability over Northern Hemisphere (NH) land in summer (Holmes et al., 2016) has been observed (Holmes et al., 2016), consistent with increases in summer extremes (Coumou and Rahmstorf, 2012; Pfleiderer et al., 2019).

Clearly, changes in hot Changes in warm temperature extremes are linked to the local mean temperature change (Rhines and Huybers, 2013), but increasing synoptic variability could contribute to more frequent heat waves (Horton et al., 2016) and circulation changes to larger winter temperature variability (Screen and Simmonds, 2014) and persistence of weather patterns (Francis and Vavrus, 2012). Increasing precipitation, For example, observed increases of both mean precipitation and precipitation variability, have been linked to warming (Pendergrass et al., 2017; Collins et al., 2013; Allen and Ingram, 2002; Held and Soden, 2006). In most climate models, precipitation variability was found to increase over land for future warming scenarios, with variability increasing at the same or a a similar or higher rate than the mean (Pendergrass et al., 2017). At synoptic to interannual timescales, local temperature and precipitation variability are negatively correlated over continental areas (Trenberth and Shea, 2005; Rehfeld and Laepple, 2016). Conversely, at longer timescales and at the global scale, a positive relationship has been found (Rehfeld and Laepple, 2016; Adler et al., 2008; Allen and Ingram, 2002). Precipitation changes are, however, strongly linked to changes in circulation and internal variability, which are not fully that remain poorly understood (Hawkins, 2011; Christensen et al., 2013; Deser et al., 2012a).

Here we investigate the linkage between mean-state and variability changes of temperature and precipitation across a wide range of global mean temperatures. In particular, we examine changes in climate variability and on interannual to multidecadal timescales in simulations conducted in the framework of the Palaeoclimate Modeling Intercomparison Project phase 3 (Braconnot et al., 2012, PMIP3), as well as the Coupled Model Intercomparison Project phase 5 (CMIP5, Taylor et al., 2012) and phase 6 (CMIP6, Eyring et al., 2016a) to which it is affiliated. We contrast changes in interannual variability across simulations for the LGM (denoted lgm in the following), the mid-Holocene (midHolocene; 6kyrs BP) and for the 6,000 years before present and idealized warming scenarios with 1%CO2 increase per year (1pctCO2) as well as an abrupt quadrupling of CO2 (abrupt4xCO2). Section 2 gives details on these experiments as well as the data preprocessing and the comparison metrics. Section 3 examines changes in local interannual variability, modes of variability, the drivers of extreme precipitation changes and in the spectrum of variability. In Sect. 4 we discuss how this compares to previous findings, and identify key uncertainties. We conclude, in Sect. 5 with a discussion on prospects for validation of modeled climate variability.
Figure 1. Hydrological sensitivity across the past and future model ensemble. The change in global mean temperature from the PI is plotted against the percentage change in global mean precipitation rate. Symbols indicate the different climate models, following Table 1. Colours show the different experiments. The line indicates 2% change in precipitation per Kelvin temperature change.

2 Data and Methods

2.1 Model simulations

The core aim of this study is to compare past and future climate simulations, and to assess the similarities – or differences – in climate variability across different Earth system states. We consider a range of state-of-the-art climate models (listed in Table 1). Therefore, this analysis is based on climate model experiments coordinated by the Coupled Model Intercomparison Project (CMIP) phase 5 (CMIP5; Taylor et al., 2012) and phase 6 (CMIP6; Eyring et al., 2016a) as well as the corresponding Palaeoclimate Model Intercomparison Project phase 3 (PMIP3; Braconnot et al., 2012). There are 25 climate models considered in this study (Table 1 and Fig. 1). We use a single ensemble member for each model (generally r1i1p1f1) and climate state. The preindustrial control (piControl) simulations represent constant preindustrial (PI) conditions and are the baseline for comparison in all our analyses. We analyze the air surface temperature (‘tas’), precipitation (‘pr’), sea surface temperature (SST) and sea-level pressure (SLP) variables.
2.2 The Last Glacial Maximum experiment (lgm)

The last glacial maximum (lgm) experiment represents conditions of 21,000 years ago. Globally averaged surface temperature was about 3-5 degrees colder than today (Annan and Hargreaves, 2013; Shakun and Carlson, 2010) in response to a global mean radiative forcing of about \(-4\text{W/m}^2\) (Broccoli, 2000) by reduced greenhouse gas concentrations (GHG), large continental ice-sheets, and a lower sea-level (Clark and Mix, 2002; Broccoli, 2000; Annan and Hargreaves, 2015). A standard set of forcings (orbit, GHG) and surface boundary conditions (ice sheets) was set out in PMIP3 (Braconnot et al., 2012; PMIP3, 2010) and PMIP4 (Taylor et al., 2012). In particular, the ice sheet extent and height is modified with respect to the piControl configurations, to reflect the extensive LGM Northern hemisphere ice sheet cover. The CO\(_2\) concentrations are fixed at 185ppm, CH\(_4\) at 350ppb and N\(_2\)O at 200ppb (PMIP3, 2010), whereas the solar constant, vegetation and aerosols follow the preindustrial control setup (Taylor et al., 2012). Overall, insolation was higher-than-preindustrial in winter in both hemispheres, and lower-than-preindustrial summer in both hemispheres (up to \(-12\text{W/m}^2\) in NH high latitudes) (Otto-Bliesner et al., 2006). This corresponds to a reduced seasonal contrast in the top-of-atmosphere radiation. The multi-model mean, shown in Fig. 2a, shows global cooling, but strongest cooling in the polar regions and above ice sheets (Fig. 2a).

2.3 The mid Holocene experiment (midHolocene)

The midHolocene experiments represent conditions of at 6,000 years before present, during the peak warmth of the current interglacial (Taylor et al., 2012; Braconnot et al., 2012). The different orbital configuration (with higher-than-present-day obliquity and eccentricity) led to an enhanced seasonal contrast in insolation, with stronger insolation in June to September from the high northern latitudes down to 30\(^\circ\)S (up to \(32\text{W/m}^2\) in NH summer), stronger insolation in September to November in SH spring (up to \(+48\text{W/m}^2\) in 30\(^\circ\)S to 90\(^\circ\)S), and negative insolation anomalies of similar magnitude in the other months of the year (Otto-Bliesner et al., 2006). This lead to a weak global mean insolation anomaly. Greenhouse gas concentrations in the PMIP3 ensemble were prescribed as for the piControl simulation (\(~280\text{ppm CO}_2, 650\text{ppb CH}_4, 270 \text{ppb N}_2\text{O}\)), as were the configurations of vegetation, aerosols, ice sheets, topography and coastlines (PMIP3, 2010). In previous model intercomparison exercises, global mean temperatures were found to be similar to today (Otto-Bliesner et al., 2006), but proxy data from the Northern Hemisphere support warmer temperatures (Wanner et al., 2015; Marcott et al., 2013).

2.4 The warming experiments 1pctCO2 and abrupt4xCO2

To complement the palaeoclimate simulations, we analyze two baseline experiments each model in CMIP5 and CMIP6 has performed: the idealized warming experiments, 1pctCO2 and abrupt4xCO2 (Taylor et al., 2012; Eyring et al., 2016a). In the abrupt4xCO2 experiment, atmospheric CO\(_2\) concentrations are abruptly quadrupled from preindustrial conditions to analyze fast feedbacks and climate sensitivity (Eyring et al., 2016a). The simulations are continued for at least 150 years. We analyze the years 100-150 for all simulations. [Note that we follow the naming scheme of CMIP5 (abrupt4xCO2; Taylor et al., 2012), while in CMIP6 the experiment name is abrupt-4xCO2 (Eyring et al., 2016a). The experimental protocols are equivalent...]

...
between the CMIP generations (Taylor et al., 2012). The CO₂ concentrations in the 1pctCO2 simulations are prescribed to increase by 1% per year in a compound fashion starting from preindustrial conditions (Eyring et al., 2016a). The change in global mean temperature at the time of CO₂ doubling in this experiment is called the transient climate response (TCR; Andrews et al., 2012). This compound increase achieves a quadrupling of carbon dioxide after 140 years, but the climate system is still highly transient. The 1pctCO2 simulations are continued between 140-160 years, of which we analyze the final 50 years. The realized warming in the 1pctCO2 scenarios is less than in the abrupt4xCO2 runs (Table 1), as the system is still farther from equilibration.

2.5 Preprocessing of model simulations

The model output is treated in a consistent fashion across all the analyses. We always analyze the final 50 years of each simulation (the final 50 in all but, except for in the abrupt4xCO2 experiment, where the years 100 to 149 are analyzed. For the variability analyses, output is converted to anomalies with respect to the monthly climatology over the 50 years using the ncl function rmMonAnnCycTLL. These anomalies are then linearly detrended at each gridpoint using the ncl function dtrend_msg_n. This process removes the changing mean state in the transient simulations and is based on the conventions of the Climate Variability Diagnostics Package (CVDP, Phillips et al., 2014; Eyring et al., 2016b). For the analyses performed here, done at annual resolution, primarily focus on interannual timescales, therefore we do not apply the PaleoCalAdjust software to account for the output-averaging calendar effects (Bartlein and Shafer, 2019).

2.6 Comparisons across the ensemble

All model output used in the study is available for download on the Earth System Grid Federation (Eyring et al., 2016b). Each model is weighted equally during ensemble averaging. These experiments provide a large range of global mean temperature (GMT) changes (Fig. 1), ranging from -6 to +6.5 K with respect to the preindustrial state. Over this range of 12K in GMT, the area-weighted global mean precipitation (GMP) varies between -12% for the lgm experiments, and +12% for the abrupt4xCO2 experiments. The slope of the relationship between temperature change and precipitation change is known as the hydrological sensitivity (HS, O’Gorman et al., 2011). For CMIP5 and CMIP3 models, values between 2 and 3 % K⁻¹ have been established (Li et al., 2013; O’Gorman et al., 2011; Allen and Ingram, 2002). Based on the mean temperature and precipitation values for each model (Fig. 1), we calculate HS individually for each GCMs (Table 1) and explore ensemble wide relationships (sections 3.1 & 3.2).

2.7 Diagnosing variability changes

This research spans across several different definitions of variability described in the literature. We term the kind of variability analysed by e.g. Huntingford et al. (2013) and Pendergrass et al. (2017) as “local variability”, in that it considers the year-to-year variations at an individual location. There has been a concerted effort to investigated the preferred spatial patterns and
<table>
<thead>
<tr>
<th>Model</th>
<th>CMIP6</th>
<th>ECS (K)</th>
<th>midHolocene</th>
<th>Igm</th>
<th>1pctCO2</th>
<th>abrupt4xCO2</th>
<th>Hydro. Sens. ((\eta_a), %/K)</th>
</tr>
</thead>
<tbody>
<tr>
<td>bcc-csm1-1</td>
<td>False</td>
<td>3.1</td>
<td>-0.1</td>
<td>-</td>
<td>3.7</td>
<td>4.9</td>
<td>1.8</td>
</tr>
<tr>
<td>CCSM4</td>
<td>False</td>
<td>2.9</td>
<td>-0.2</td>
<td>-4.9</td>
<td>4.3</td>
<td>4.9</td>
<td>1.8</td>
</tr>
<tr>
<td>CESM2</td>
<td>True</td>
<td>5.2</td>
<td>-</td>
<td>-</td>
<td>4.9</td>
<td>7.2</td>
<td>1.6</td>
</tr>
<tr>
<td>CNRM-CM5</td>
<td>False</td>
<td>3.3</td>
<td>0.1</td>
<td>-2.7</td>
<td>4.0</td>
<td>5.1</td>
<td>1.7</td>
</tr>
<tr>
<td>COSMOS-ASO</td>
<td>False</td>
<td>N/A</td>
<td>-</td>
<td>-5.7</td>
<td>-</td>
<td>-</td>
<td>2.2</td>
</tr>
<tr>
<td>CSIRO-Mk3-6-0</td>
<td>False</td>
<td>4.1</td>
<td>-0.0</td>
<td>-</td>
<td>3.7</td>
<td>5.7</td>
<td>1.9</td>
</tr>
<tr>
<td>CSIRO-Mk3L-1-2</td>
<td>False</td>
<td>3.1</td>
<td>0.0</td>
<td>-3.2</td>
<td>-</td>
<td>-</td>
<td>1.5</td>
</tr>
<tr>
<td>EC-EARTH-2-2</td>
<td>False</td>
<td>4.2</td>
<td>-0.1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FGOALS-f3-L</td>
<td>True</td>
<td>3</td>
<td>-</td>
<td>-5.0</td>
<td>5.2</td>
<td>5.2</td>
<td>1.8</td>
</tr>
<tr>
<td>FGOALS-g2</td>
<td>False</td>
<td>3.7</td>
<td>-0.8</td>
<td>-4.5</td>
<td>3.3</td>
<td>5.7</td>
<td>2.0</td>
</tr>
<tr>
<td>FGOALS-s2</td>
<td>False</td>
<td>4.5</td>
<td>-0.1</td>
<td>-</td>
<td>4.4</td>
<td>6.0</td>
<td>2.0</td>
</tr>
<tr>
<td>GISS-E2-1-G</td>
<td>True</td>
<td>2.7</td>
<td>-</td>
<td>-</td>
<td>2.0</td>
<td>4.2</td>
<td>1.2</td>
</tr>
<tr>
<td>GISS-E2-R</td>
<td>False</td>
<td>2.1</td>
<td>-0.1</td>
<td>-4.8</td>
<td>3.0</td>
<td>3.4</td>
<td>1.5</td>
</tr>
<tr>
<td>HadGEM2-CC</td>
<td>False</td>
<td>4.5</td>
<td>0.2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>HadGEM2-ES</td>
<td>False</td>
<td>4.6</td>
<td>0.2</td>
<td>-</td>
<td>4.6</td>
<td>6.2</td>
<td>1.1</td>
</tr>
<tr>
<td>INM-CM4-8</td>
<td>True</td>
<td>2.1</td>
<td>-</td>
<td>-3.0</td>
<td>3.0</td>
<td>3.6</td>
<td>1.7</td>
</tr>
<tr>
<td>IPSL-CM5A-LR</td>
<td>False</td>
<td>4.1</td>
<td>-0.2</td>
<td>-4.7</td>
<td>4.2</td>
<td>6.1</td>
<td>2.4</td>
</tr>
<tr>
<td>IPSL-CM6A-LR</td>
<td>False</td>
<td>4.5</td>
<td>-</td>
<td>-</td>
<td>5.7</td>
<td>7.3</td>
<td>1.4</td>
</tr>
<tr>
<td>MIROC-ES2L</td>
<td>True</td>
<td>2.66</td>
<td>-</td>
<td>-</td>
<td>3.6</td>
<td>4.1</td>
<td>1.3</td>
</tr>
<tr>
<td>MIROC-ESM</td>
<td>False</td>
<td>4.7</td>
<td>-0.6</td>
<td>-5.3</td>
<td>4.1</td>
<td>6.0</td>
<td>1.9</td>
</tr>
<tr>
<td>MPI-ESM-P</td>
<td>False</td>
<td>3.5</td>
<td>-0.2</td>
<td>-4.4</td>
<td>4.2</td>
<td>5.8</td>
<td>1.7</td>
</tr>
<tr>
<td>MRI-CGCM3</td>
<td>False</td>
<td>2.6</td>
<td>-0.1</td>
<td>-4.7</td>
<td>3.4</td>
<td>4.4</td>
<td>2.5</td>
</tr>
<tr>
<td>MRI-ESM2-0</td>
<td>True</td>
<td>3.1</td>
<td>-</td>
<td>-</td>
<td>3.6</td>
<td>4.8</td>
<td>1.6</td>
</tr>
<tr>
<td>NESM3</td>
<td>True</td>
<td>3.7</td>
<td>-</td>
<td>-</td>
<td>5.9</td>
<td>6.9</td>
<td>1.7</td>
</tr>
</tbody>
</table>

Table 1. Details of the models and experiments involved in the analysis. Each experiment provides the global mean change in surface temperature from the preindustrial control simulation (\(\Delta T\)). The (actual) hydrological sensitivity \(HS\) is the global mean percentage change in precipitation divided by the temperature change. It is was calculated via linear regression through all the simulations if available. Where fewer experiments existed, it was calculated as the directed average of the values, excluding the midHolocene simulation.
temporal variations that account for large-scale features in variance in the climate system. We term these as "modes of climate variability", and they are considered as the product of a specific spatial pattern and an associated index time series (Qin et al., 2014). They are diagnostic measures for teleconnections or surface climate patterns, and defined on pressure, temperature or precipitation fields. Here, we investigate the changes of ten modes of variability from the detrended time series following the workflow of the Climate Variability Diagnostics Package (CVDP, Phillips et al., 2014). We investigate seven atmospheric-oceanic coupled modes defined as predominant SST patterns, and three extratropical atmospheric modes with large-scale teleconnection patterns (Deser et al., 2010; Nigam, 2003).

2.7.1 Local variability

Local variability is computed as the standard deviation of the annual mean temperature or precipitation. In all simulations, a 50 year subset was selected (often typically) the final 50 years, Sec. 2.5), anomalies with respect to the simulations climatology computed and then a linear trend removed. Variance ratios are detrended. Standard deviation ratios were computed on the individual model grids and interpolated bilinearly onto a common 1×1° latitude/longitude grid prior to ensemble averaging.

2.7.2 The El Niño/Southern Oscillation (ENSO)

The El Niño/Southern Oscillation (Bjerknes, 1966) is an atmosphere-ocean coupled mode of variability with large-scale changes in SST, SLP, precipitation and winds as well as the ocean thermocline depth in the equatorial pacific varying semi-periodically with a timescale of 2-10 years (Philander, 1983). ENSO is one of the main drivers of global mean temperature variability, with global teleconnections (Bjerknes, 1969), and a pronounced impact on the global energy balance and global mean temperature (Trenberth and Fasullo, 2012; Foster and Rahmstorf, 2011). The SLP oscillation in the South Pacific (‘Southern Oscillation’) was first described by Walker and Bliss (1932), and the link between atmospheric oscillating patterns and local ocean circulation first described by Bjerknes (1966). Here we use the Niño3.4 and Niño4 indices, which are the equatorial (5°S–5°N) area-averaged SST anomalies over the regions 170°W–120°W and 160°E–150°W, respectively (Trenberth, 1997; Deser et al., 2010, 2012b). Computations are based on the ncl-script sst.indices.ncl (Phillips et al., 2014).

2.7.3 The Interdecadal Pacific Oscillation (IPO)

The Interdecadal Pacific Oscillation (IPO) shows a pattern of SST change similar to ENSO (equatorial warming), but with different impacts (Power et al., 1999; Meehl and Hu, 2006). Here we construct a monthly index timeseries based on the first principal component of 13-yr low pass filtered Pacific (40°S–60°N, 110°E–70°W) area-weighted SST anomalies, where the global mean SST anomaly has been removed at each timestep.
2.7.4 The Indian Ocean Dipole (IOD)

The Indian Ocean Dipole (IOD) is an irregular pattern of SST variability in the Indian Ocean independent of ENSO in the Pacific (Webster et al., 1999). In an IOD— a negative IOD event, the western region warms and eastern region cools. The opposing pattern, with a decrease in the zonal temperature gradient, is the positive IOD mode (IOD+), a positive IOD event. The associated changes in surface pressure and rainfall lead to rainfall modulation and extreme precipitation events at the western/eastern boundaries (Webster et al., 1999). Its subdecadal variability is modulated on decadal to multi-decadal timescales (Ashok et al., 2004). Here, the index time series is calculated using script sst.indices.ncl based on the CVDP (Phillips et al., 2014) as the difference of the area-averaged SST anomaly between the regions $50^\circ$E–$70^\circ$E, $10^\circ$S – $10^\circ$N and $90^\circ$E–$110^\circ$E, $10^\circ$S–equator (Saji et al., 1999).

2.7.5 The Atlantic Meridional Mode (AMM)

The Atlantic Meridional Mode (AMM), sometimes called the Atlantic dipole mode or gradient mode, is a leading mode of SST variability in the equatorial Atlantic (Servain et al., 1999). The SST pattern, with opposing anomalies on either side of the equator, modulates the meridional gradient of the sea surface temperature anomaly in the tropical Atlantic, and hence the movement of the Intertropical Convergence Zone (ITCZ) and associated precipitation (Xie and Carton, 2004). The SST gradient is complemented by cross-equatorial atmospheric flow, strengthened by wind-evaporation-surface temperature feedbacks (Xie and Carton, 2004). The AMM has been linked with hurricane activity in the area (Vimont and Kossin, 2007) and impacts rainfall over tropical Atlantic/NE Brazil/Sahel (Kushnir et al., 2006). Following Doi et al. (2010), the AMM state is defined here as the basin-wide, area average, detrended SST anomaly difference between the two hemispheres regions $15^\circ$N–$5^\circ$N, $50^\circ$W–$20^\circ$W minus the average of $15^\circ$S–$5^\circ$S, $20^\circ$W–$10^\circ$E (Phillips et al., 2014).

2.7.6 The Atlantic Zonal Mode (ATL3)

Atlantic Zonal Mode (ATL3) is an equatorial coupled mode, similar to ENSO (Zebiak, 1993), therefore sometimes referred to as ‘Atlantic Niño’ (Xie and Carton, 2004). Calculation of the mode in the CVDP follows Zebiak (1993), and is based on the area average of the detrended SST anomaly over the region $3^\circ$N - $3^\circ$S, $20$ - $0^\circ$W. The ATL3 displays interannual variations with roughly a four-year period. Its variations are linked to rainfall variability in the Sahel region (Giannini et al., 2003).

2.7.7 The Pacific Decadal Oscillation (PDO)

The Pacific Decadal Oscillation (PDO), also termed Interdecadal Pacific Oscillation (Power et al., 1999), is the leading mode of variability of monthly SST anomalies over the North Pacific after global mean anomaly is removed. It emerges as a mode partially driven by ENSO and independent, stochastically emerging variations (Deser et al., 2010; Mantua et al., 1997; Schneider and Cornuelle, 2005). However, no clear spectral peak has been identified (Deser et al., 2010), as it arises from a superposition of SST fluctuations with different dynamical origins (Schneider and Cornuelle, 2005; Deser et al., 2010). The PDO was first described in
1997 as recurring climate pattern of ocean-atmosphere variability over North Pacific and linked to impacts on Salmon production and coastal surface temperatures on the west coast of the North American continent and the adjacent sea surface (Mantua et al., 1997). The index is associated with temperature/precipitation changes over western and eastern edges of North Pacific and displays positive correlation with winter precipitation in California (Mantua et al., 1997). The pattern is generally similar to ENSO variations but with a weaker Southern Pacific imprint (Deser et al., 2010). We calculate a monthly index time series from the leading principal component of the area-weighted SST anomalies in the box 20–70°N to 110°E–100°W, where the global mean SST anomaly for each time step has been removed (Deser et al., 2010) based on the script pdo.ncl from the CVDP (Phillips et al., 2014).

2.7.8 North Atlantic Oscillation (NAO) and the Northern Annular Mode (NAM)

The North Atlantic Oscillation (NAO) is a quasi-periodic spatial pattern of sea-level pressure changes between the Arctic and the Arctic and subtropical North Atlantic (Stephenson et al., 2003; Walker and Bliss, 1932). NAO variations impact the atmospheric circulation over North Atlantic and the strength of the Westerly inflow into Europe, influencing storm tracks, temperature and precipitation, in particular in boreal winter (Hurrell, 1995; Hurrell and Deser, 2010). It varies on a seasonal, interannual to decadal timescale (Hurrell, 1995). In positive NAO phases, a large difference in SLP between the high and mid-latitudes implies a strong SLP gradient and strong westerly inflow into central Europe. In negative NAO phases, the smaller difference in pressure is associated with a southerly shift in the North Atlantic storm tracks and enhanced precipitation in the Mediterranean and North Africa. Here we calculate the NAO index using the script pslnam_nao.ncl (Phillips et al., 2014), based on the first principal component of the boreal winter (DJF) area-weighted annual SLP average over the box 20–80°N, -90 – 40°E (Hurrell and Deser, 2010). Given that this calculation results in a normalised time series, to investigate changes in NAO variability, we consider the spatial standard deviation of the Empirical Orthogonal Function (EOF) over the box instead (Power et al., 2013).

The Northern Annular Mode (NAM) describes the zonal SLP deviations in the zonal pressure gradient between the polar regions and the subtropics. This gradient governs synoptic (5-day-mean) variability of sea-level pressure in the northern hemisphere NH (Lorenz, 1951). By definition, it is related to the NAO. Here, it is calculated as the leading EOF of the area-weighted monthly-mean SLP anomalies over the latitudes 20–90°N (Hurrell and Deser, 2010), with its variability measured by the spatial standard deviation of this EOF (Power et al., 2013).

2.7.9 Southern Annular Mode (SAM)

The Southern Annular Mode (SAM) index gives the strength of the sea-level pressure gradient in the Southern Hemisphere mid-latitudes (Karoly, 1990). It is a distinctive pattern of climate variability in the Southern Hemisphere, in particular in winter (Karoly, 1990; Marshall, 2003). The variations in the SLP gradient impact regional temperatures, precipitation (Marshall, 2003; Gillett et al., 2006) as well as the circulation of the Southern Ocean. Negative values of SAM have been linked with weakenings of the polar vortex, and an increasing occurrence of hot and dry extremes in Australia
(Lim et al., 2019). SAM impacts latitudinal rainfall distribution from the subtropics to Antarctica, with recent trends towards a more positive mode than over the last 1000 years, and links to an Antarctic interior cooling/peninsula warming (Abram et al., 2014). Here, we calculate the PDO variability using the script psl.sam_psa.ncl from the CVDP (Phillips et al., 2014). Seasonal/annual PSL—Monthly PSL anomalies averages are formed over the latitudes 20–90°S, and a square root of the cosine of latitude weighting is applied. The leading EOF is considered to give the pattern for the SAM (Thompson and Wallace, 2000), and spatial standard deviation of this pattern (Power et al., 2013) is used as our measure of its variability.

### 2.8 Changes in precipitation extremes

We investigate the major large-scale patterns of variability associated with precipitation variability across climates. Based on Fig. 3 we find that, in many regions, past and future precipitation variability shows opposing signs. We select five regions with Mediterranean-type climates (Seager et al., 2019): (1) the southwestern tip of South America, (2) southwestern South Africa, (3) southwestern Australia, (4) coastal western North America, and, (5) the western Mediterranean. These regions, in the present, lie between the poleward edge of the winter Hadley cell and equatorward edge of the mid-latitude storm tracks, and have climates. The climate is therefore characterized by wintertime precipitation and summertime dryness associated with subtropical subsidence, and display substantial interannual variability (Seager et al., 2019).

For each region, model and experiment we (a) first calculate the climatological average, annual mean precipitation and, as an individual threshold, the interannual standard deviation of local precipitation. We (b) then identify where, in the 50-year timeslice, precipitation falls above or below 1 standard deviation and (c), and composite sea-level pressure, surface air temperature and precipitation for these extreme precipitation years across all experiments and model simulations.

### 2.9 Timescale-dependence of the variability changes

The power spectrum, \( P(\tau) \), of a climate variable describes how its variability is distributed over the timescales \( \tau \), with the integral over the entire spectrum yielding the total variance of the signal (Chatfield, 2004). Here we use multitaper power spectrum (Thomson, 1990) with linear detrending, and investigate the area-weighted mean spectra of the local (grid-box) time series. The scaling exponent, \( \beta \), is used to summarize the scaling relationship of variance with timescale, or equivalently frequency which relates to timescale as \( f = 1/\tau \), assuming that the spectrum approximately follows \( P(f) \sim \tau^\beta \). The scaling exponent \( \beta \) is estimated as the linear slope between the logarithm of the power spectral density and the logarithm of timescales; the fit is performed between 4 months to 20 years. Uncorrelated white noise has no autocorrelation, and the scaling exponent is zero (\( \beta = 0 \)). For \( \beta > 0 \) (\( \beta < 0 \)), the underlying stochastic process displays positive (negative) autocorrelation. Positive autocorrelation for temperature can be expected (Fredriksen and Rypdal, 2016), while precipitation and pressure have lower, or negative values (Fraedrich et al., 2009).
3 Results

3.1 Hydrological sensitivity across the ensemble

Fig. 1 shows the range of global mean temperature change and precipitation change from the piControl simulations. The lgm ensemble has a mean temperature anomaly of 4.2 (range of -2.5 to -6) K, and precipitation anomalies range from -6 to -12%/K. The midHolocene ensemble shows no large, consistent global mean changes. However, those models that show models with wetter conditions show positive global mean temperature anomalies. The 1pctCO2 simulations display temperature anomalies from +3 to +7K, and precipitation increases between 3 and 12%. The abrupt4xCO2 warming simulations are slightly warmer (+4 to +7K) and wetter (+5 to +12%/K). For the entire ensemble, we estimate an overall mean HS of 1.73±0.005 (one standard error of the slope) taking into account all models weighted equally. The equilibrium experiments (lgm and midHolocene) fall consistently on the 2%/K-line (Allen and Ingram, 2002), whereas the transient warming experiments fall below. We find no discernible difference between the precipitation scaling between the CMIP5 and CMIP6 models. We find no systematic relationship between ECS and HS. We note: Additional investigations (not shown) demonstrate that our findings hold with and without calendar adjustment.

3.2 Changes in local interannual variability

We aim at a comparison of mean state and variability changes across multiple climate states. Changes in temperature, and temperature variability (Fig. 2) do show some consistent progression from the palaeoclimate experiments to the idealized warming. As expected, we find globally cooler conditions for the LGM. These are highly consistent across the ensemble, as the stippling—there is scarcely any stippling in Fig. 2a, indicating that at least 2/3rds of the considered models show do agree on the same sign as the mean, spreads across the entire field (Fig. 2a). Comparing this to Fig. 2e, which shows the change in simulated temperature variability in the lgm experiment vs. the piControl as the ratio of standard deviations of the annual means shows that the interannual temperature variance is high in areas which experienced much colder conditions (at the sea-ice edges), and where the lower sea level led to more exposed shelves (e.g., Indonesia) as well as at the edges of the large continental ice sheets (Laurentide, European, and Eurasian). The simulated lgm temperature variability is higher in the mid-to-high latitudes of both hemispheres, but large areas of the tropics show decreases in interannual temperature variance against the piControl experiment, in particular the ENSO region, South America, Southern Africa and the West Pacific Warm Pool show decreases in interannual temperature variance against the piControl experiment. Overall, the mean-change pattern of the lgm experiment is weakly anticorrelated with the variance-change pattern of standard deviation changes (r=−0.12, p<0.05 based on area-weighted Pearson correlation and a one-sided t-test conservatively assuming 500 degrees of freedom, accounting for the high degree of spatial autocorrelation in the fields).

The local changes in mean precipitation for the lgm simulations (Fig. 3a) are overall negative, consistent with the globally decreased precipitation (Fig. 1). We find consistent shifts towards higher precipitation in the continental areas of both hemispheres affected by subtropical cyclonic subsidence precipitation, over northern Africa, southern Africa, across the subtropical southern Atlantic, as well as southwestern North America. Interannual precipitation variance in the lgm simulations is lower.
Figure 2. The change in mean annual temperature (a-d) and its variability (e-f) across multiple climate experiments. Each panel shows the ensemble average difference. The changes in the mean temperature are calculated as the experiment minus preindustrial control gridbox annual means. The changes in variability are given based on the ratio of the standard deviation of annual mean temperature in the experiment, over that of the piControl experiment. Ratios above 1 indicate higher variability in the experiment than in the piControl. The contours in each panel show the ensemble-mean pattern in the preindustrial control. Contour variations are due to the different number of models available for individual experiments, as the preindustrial ensemble-mean is only computed from models in each experiment. Stippling indicates where the sign of the change agrees disagrees for more than 2/3rds of the ensemble.

than in the control simulations with the exception of the areas which have higher mean precipitation, where variability also
increases (Fig. 3e). Across the multimodel field, mean and variance change are positively correlated (r=0.63, p<0.01).

The midHolocene simulations show weak but consistent (sub)tropical cooling, and moderately warmer conditions in the annual mean temperatures (Fig. 2b), consistent with as expected given the positive high-latitude insolation forcing (Sect. 2.3). Overall, the interannual temperature variance shows patterns of higher and lower-than-piControl variance with modest degrees of inter-model consistency. Similar to the lgm variance ratio field, there are reductions in the tropical Atlantic temperature variance, consistent collocated with a local increase in precipitation (Fig. 3b), and precipitation variance (Fig. 3f). Precipitation variance appears lower in the Pacific, and higher over the Atlantic and Indian Ocean sector, with a strong positive precipitation anomaly over Northern Africa. Mean and variance change standard deviation changes are strongly correlated for precipitation (r=0.55, p<0.01), but only weakly correlated for temperature (r=0.09, p<0.05).

Mean temperature change for the 1pctCO2-scenario is consistently positive with stronger warming over the continents and amplified warming in the high Northern latitudes (Fig. 2c). Interannual temperature variance The interannual temperature standard deviation ratio (Fig. 3g) shows consistent increases in temperature variability over South-Western southwestern North America, South America, Africa, Australia, the Indian Peninsula and China as well as over the North Atlantic, and decreases in temperature variance variability against piControl over Northern northern North America, Scandinavia, the Tibetan Plateau, Northeast China as well as across the Arctic. Surrounding Antarctica, decreasing temperature variance variability is observable south of the polar circle, but moderate increases in temperature variance standard deviations are observable over East Antarctica. Overall, the mean change and variance standard deviation change patterns are anticorrelated (r=-0.23, p<0.01), meaning that where we find stronger warming we also observe lower simulated temperature variability.

Mean precipitation change across the 1pctCO2-ensemble is positive (Fig. 1-Inspecting-). However, inspecting Fig. 3 indicates, however, that this increase affects primarily primarily affects the high latitudes and the Equatorial area region. In South America, no clear change in precipitation is discernible, whereas the Sah and Arabian Sea are wetter. Mean and variance standard deviation change fields are positively correlated (r=0.67, p<0.01). Patterns of temperature and precipitation changes in the abrupt4xCO2-scenario (Fig. 2d) and 3d) are highly consistent with those for the 1pctCO2-scenario (r=0.94,p<0.01 for precipitation, r=0.98, p<0.01 for temperature). In mean and variance variability, a stronger amplification of the warming patterns (Fig. 2h), over the continents, the North Atlantic, the Indopacific and the areas-Indo-pacific and the locations- of the subtropical high are discernible. The polar and continental amplification of the temperature change patterns of the lgm-scenario are mirrored in the areas of warming in the 1pctCO2 and abrupt4xCO2-scenarios (r=-0.65 resp. and r=-0.64 respectively, p<0.01). In particular in the west-coast mid-latitudes where higher precipitation is simulated at the LGM, it appears lower in the warming scenarios of the Northern HemisphereNH.
Figure 3. The change in mean annual precipitation (a-d) and its variability (e-f) across multiple climate experiments. Each panel shows the ensemble average difference (as percentage changes with respect to each model’s respective piControl). The changes in variability are based on the ratio of the standard deviation of annual mean precipitation in the experiment, over that of the piControl experiment. Ratios above 1 indicate higher variability in the experiment than in the piControl experiment. The contours in each panel show the ensemble-mean pattern in the preindustrial control (in mm/day). Contour variations are due to the different number of models available for individual experiments, as the preindustrial ensemble-mean is only computed from models in each experiment. Stippling indicates where the sign of the change agrees disagrees for more than 2/3rds of the ensemble.
3.3 Changes in modes of variability

3.3.1 Changes in the global mean

Global mean precipitation increases with global mean temperature across the ensemble (Fig. 1). However, across the multi-model ensemble we find a tendency across the models for the variance of global mean temperature to decrease decreases with the global mean state, resulting in lower variance than in the piControl for the majority of models considered in the idealized warming scenarios and higher-than-preindustrial variance for the lgm experiment (Fig. 4a). At the same time, the standard deviation of global mean precipitation increases with approximately 3%/K (Fig. 4b), hence at a higher rate than the global mean precipitation (Fig.1). Comparing these temporal changes against the spatial expression in Figs. 2 and 3 we find that the global reduction of temperature variability with warming is dominated by the ocean and high-latitude signal, whereas the mid-latitude continental areas show consistent increases in temperature variability with warming. At the same time, the precipitation increase precipitation is more inhomogeneous in spatial location and magnitude (Fig. 3d,h).

3.3.2 Changes in SST-based modes

Changes in the SST-based modes of variability across the ensemble are given in Fig. 4c–h. The majority of models (6/9) show a lower-than-preindustrial NINO3.4 and NINO4 standard deviation for the lgm and for the midHolocene (9/14), and a higher-than-preindustrial ENSO-index variance for the idealized warming scenarios (10/8 and 7/11, Fig. 4c,d). Nonetheless, there is no statistically significant association link between global mean temperature and ENSO variability increase (e.g. Christensen et al., 2013). Preliminary findings from the new PMIP4 simulations appear to confirm these conclusions about the palaeoclimate time periods (Brown et al., 2020). This fits with palaeoENSO reconstructions of suppressed activity during the mid-Holocene, yet with potential changes in ENSO variability during the LGM (Lu et al., 2018). There are no systematic changes in standard deviation across the ensemble for the PDO (Fig.4e) or the IPO (Fig 4f), although both are not well resolved by the short records analysed here. For the IOD (Fig 4g) there are no tendencies in the lgm-ensemble, with about as many models showing an increased in standard deviation as showing a decrease. However, a majority of models show suppressed IOD activity under the warming scenarios corresponding with the reduced temperature variability over the Arabian Sea upwelling (Fig. 2), which may be a response to the increased ocean stratification seen in the transient simulations (Oyarzún and Brierley, 2019). In the tropical Atlantic, weak but negative trends for the AMM (Fig.4h) and the ATL3 (Fig. 4i) variability for warmer conditions are found. This fits with the findings of Brierley and Wainer (2018), and is not inconsistent with the increased future rainfall variability over both the Amazon and West Africa (Fig.3g,h) - it it indicates a diminished influence of Atlantic climate variability in the regions.

3.3.3 Changes in atmospheric modes of variability

Let us now consider the atmospheric modes of variability (Fig.4j-l). In the lgm experiments, the simulated temperature gradient in the Northern hemisphere is stronger NH is higher than in the preindustrial piControl - all but one model (Fig. 4) show
reduced variability for the NAM and the NAO. Conversely, in the idealized warming scenarios, with their reduced temperature gradient, which have a reduced meridional temperature gradient, more models show increasing standard deviations. Whether a reduced standard deviation indicates a more stable storm track or a more spatially-constrained one requires further investigation and possibly moving away from EOF-based mode definitions.

The Southern Annular Mode shows a tendency towards reduced standard deviations for the idealized warming scenarios (Fig. 4i), but also for . This also occurs in the lgm experiments. This counter-intuitive response may arise from the competing influences of variability in of the Antarctic sea ice edge (Fig. 2) and the hydrologically-related variability within the storm tracks (Fig. 3).

3.4 Testing the stationarity of modes Circulation patterns underlying extratropical precipitation extremes

Precipitation changes in Mediterranean-type climates oppose on the western edges of continents in the extratropics display opposite signs in their precipitation anomalies, with respect to pre-industrial, and relative to the global mean change across the ensemble . To (see Fig. 5). Given this difference, we assess whether the drivers of precipitation in these regions, shown by boxes in the lefthand panels of Fig. 5, atmospheric drivers of such regional precipitation are consistent from past to future climates, we to better inform the relevance of variability in paleoclimates to future climate change in these susceptible semi-arid regions.

We investigate sea-level pressure and surface air temperature anomalies associated with high precipitation anomalies (Fig and low annual precipitation by compositing over years with regional precipitation above or below one standard deviation around the mean (following sect. 5). Exploratory investigations uncovered no statistically significant changes between individual climate states (not shown). The following proceeds by the regions in Fig. 5.

High precipitation years in Patagonia southwestern South America (Patagonia) are associated with an increased SLP gradient between the region and the Antarctic continent (Fig. 5a), indicative of positive SAM conditions, a moderate cooling in the South Eastern Pacific sector southeastern Pacific sector, and warmer conditions in the South Atlantic and Southern Indian ocean. The reverse situation can be found for years with low precipitation anomalies in Patagonia (SFig the same region (supplementary Fig 2a-SF2a,b). This is true for the entire ensemble (Fig. SF3, as well as for individual climate states (not shown). The global precipitation composites for regional high and low-precipitation years in SFigFig. 3-SF3 show that years with high-precipitation anomalies in the region are also associated with lower-than-average precipitation in the ENSO regions (SFigFig 2b–SF3b).

There is no inter-model and inter-experiment consistency in the interannual atmospheric conditions for high precipitation years in western South Africa with regards to SLP and temperature (Fig. 5c,d), indicating that the drivers of variability are more complex in this region and cannot be explained by a single climate mode. Precipitation variability may therefore be controlled more by differences between individual storms than by persistent large-scale modes of the atmosphere. Indeed, there are no coherent large-scale structures even within individual climate states, though a regional signature in temperature exists (Fig SF2, SF3).

The composite plots for Western southwestern Australia (Fig. 5e,f; Fig. SF2, SF3) show, similarly to the southwestern South American composites, that increased precipitation is found for years with a strong SAM and an increased SLP gradient between
Figure 4. Relationship of the standard deviation of climate indices and modes to the change in global mean temperature from preindustrial conditions. Colours indicate the different experiments: CMIP5 and CMIP6 models are not differentiated. (a) Change in the standard deviation of the global, annual mean surface temperature. (b) Change in the standard deviation of the global, annual mean precipitation rate. Changes in the standard deviation (i.e. amplitude of the mode) of (c) ENSO based on the NINO3.4 index and (d) based on the NINO4 region, (e) the PDO, (f) the IPO, (g) the IOD, the meridional (h, AMM) and zonal (i, ATL3) modes of equatorial Atlantic SST variability, and (j) the Northern Annular Mode, (k) the boreal winter NAO and (l) the Southern Annular Mode. All modes Dashed horizontal lines are calculated by the Climate Variability Diagnostics Package (Phillips et al., 2014) given from \( \pm 2\sigma_{\text{fit mode}} \). See Sect. 2.7 for details on the individual modes, and how any changes in mean climate state between the experiments are removed prior to calculation. Linear unweighted fits to the mode changes and the corresponding p-values are given in each panel without censoring for significance. P-values assume 60 degrees of freedom.

Australia and Antarctica. The higher pressure and temperatures in the North Pacific sector for both Western southwestern South American and Western Australian composites could indicate stable teleconnection patterns across the experiments. Cooler conditions prevail throughout the tropics in high-precipitation years, suggesting a decreased southern hemisphere Southern Hemisphere meridional temperature gradient. Precipitation composites (SFig Fig. 3eSF3 e,f) show a dipole-like structure reminiscent of ENSO, with more precipitation in Western Australia associated with increased precipitation in South-East southeast Asia, and less-than-average precipitation in the Equatorial Pacific.
High precipitation in **Western southwestern** North America is associated with enhanced low-pressure and higher-than-average SLP over the **eastern North Pacific, the** North Atlantic and Greenland (Fig. 5g; Fig. SF2) as well as locally warmer conditions (Fig. 5h; Fig. SF3), and drier conditions to the North and South (Alaska/Mexico, SFig Fig. 3h SF3h). These patterns suggest a consistent influence of the PDO and the NAM on interannual precipitation variability in the region.

This is-response is structurally highly similar to the patterns observable for the **Western western** Mediterranean, where high precipitation anomalies are associated with an increased pressure gradient between the mid- and high latitudes (Fig. 5i; Fig. SF3), cooler conditions on the Iberian Peninsula and Eurasia and warmer conditions over the Arctic regions of North America and the Labrador Sea (Fig. 5j; Fig. SF3). For both Western North America and the Western Mediterranean, high annual precipitation years are associated with positive precipitation anomalies in the Equatorial Pacific (SFig Fig. 3h SF3h, j).

Therefore **In summary**, in both **southwestern** South and North America, anomalous precipitation is associated with sea-level pressure variations over the eastern Pacific in the respective hemisphere (low pressure during wet years, **and** high pressure during dry years) illustrative of circulation patterns that are more or less conducive to water delivery to the continent. In the **South Southern Hemisphere**, this is also associated with a standing wave structure in surface air temperatures at mid-latitudes, as well as an equatorial Pacific signature reminiscent of ENSO. Precipitation variability over **western southwestern** Australia is also linked to equatorial Pacific temperatures, as well as pressure variations in the Indian and south Pacific oceans, while precipitation variability over the western Mediterranean is **clearly more strongly** linked to variability over the North Atlantic (likely the NAO), as well as the North Pacific and eastern equatorial Pacific (the latter is suggestive of ENSO).

### 3.5 Changes in the spectrum of variability

We investigate the globally averaged, area-weighted power spectra of local monthly temperature (Fig. 6 a,b) and precipitation (Fig. 6 c,d) anomalies. We find that, in the **global mean**, the spectrum of temperature shows overall higher local temperature variability in the **lgm** experiments, and lower temperature variability for the warm experiments (Fig. 6a), consistent with the findings for total variance (Figs. 2, 3 and 4a,b). **Around** On the ENSO timescale (around 3.5–3.7 years), the **decrease of reduced** variance is less important for the warm experiments, but more important for the **midHolocene** experiment, thus leading to small changes in the **scaling before and after that timescale variance on longer and shorter timescales**. Overall, the scaling of **intraannual intra-annual** to decadal temperature variability is rather consistent for all experiments (ranging from $\beta=0.26$ to $\beta=0.35$) and The scaling changes little with respect to the piControl experiment, as can also be seen by the **rather relatively** flat spectral ratio curves (Fig. 6b). The **lgm** curve however shows a small decrease in scaling since the variance increases more on the side of smaller timescales. We also find a remnant—There is also an annual peak in the **idealized warming scenario for the LpCO2 temperature temperature spectra**, which could be due to an incomplete detrending of a changing seasonal cycle.

**The increase in ENSO-band variance** The global picture of more variable surface temperatures in the **lgm** experiments, and decreased variance at all timescales for the warmer experiments is weakened over land (SF6), but holds over the oceans (SF7). However, temperature variance increases over low-to-mid-latitude continental areas in the warm experiments (SF8, spectrum across $40^\circ$S to $40^\circ$N). Generally there is no strong inter-model consistency over the low-to-mid-latitude continental areas, as also indicated by Fig. 2.
The variance around the 3-5-year timescale in the warm experiments is more apparent for the local precipitation anomalies than for temperature (Fig. 6c) than for temperature, and in addition, This is consistent with the findings of Cai et al. (2014), who found an increasingly frequent ENSO occurrence in warming experiments. Conversely, we find that it also decreases for the midHolocene and lgm experiments. Overall, the precipitation variance increases rather fairly consistently over all timescales for the warm experiments with respect to the piControl runs, and likewise decreases for the lgm and midHolocene experiments. The precipitation spectral ratios with respect to the piControl simulations (Fig. 6d) outline these patterns clearly. These coherent changes in the global mean spectra are also corroborated by a high degree of consistency in the scaling patterns of surface temperature, precipitation and surface pressure (SFig SF 4), which show ‘white’, or flat, spectra over the continents and ‘red’ spectra with variance concentrated at longer timescales over the oceans, particularly along the equator. There is a reddening of the variability over areas where sea-ice is lost in the warm experiments. This could be attributed to the open seas dampening the high-frequency variability more with warming. There is a similar blueing in the lgm over the Fram Strait and the Barents Sea where sea-ice cover is extended (SFig SF 4). However, there is a reddening over the Arctic for sea-level pressure in the lgm.

4 Discussion

4.1 Changes in climate variability with global mean temperature

Using a wide range of model simulations has allowed us to examine the relationship between changes in global mean temperature and climate variability from the perspective of the mean and variance, standard deviation fields, changes in modes of variability, and the timescale-dependency of temperature and precipitation changes. We find that globally averaged temperature variability decreased from the cold to the warm experiments. This is true for global mean temperature, and the global mean of regional temperatures. We find that also consistent changes in temperature variance are more localized than changes in the mean fields. From the cold to the warm simulations, temperature variability increases over land, and tends to decrease over the oceans. This pattern also holds across timescales. Temperature variability reduction is particularly strong in the at high latitudes, where seasonality and interannual temperature variability are particularly high (Huybers and Curry, 2006). This suggests that changes in temperature variability, in both directions, affect areas which also undergo a large mean-state change.

We find clear indications for shifts in relationships between global mean temperature and precipitation variability as well as in tropical Atlantic modes of variability in the tropical Atlantic, with the zonal and meridional modes both strongly varying in the lgm experiments, and shifting towards weaker variability in the warmer scenarios. This is consistent with the recent findings of Brierley and Wainer (2018), who investigated tropical Atlantic sea surface temperature variability using a similar model ensemble, but also including the historical era. The zonal gradient mode in the Indian Ocean, IOD, shows a tendency for lower variability in the midHolocene (and thus, for tropical weak cooling) and future warming scenarios, and is therefore not systematically changing with global mean temperature.
The reduced ENSO variability for the *midHolocene* experiments shows that the ENSO mode strength firstly strength of ENSO primarily links to the tropical temperature changes, and only secondly-secondarily to global mean temperature change. This finding is corroborated by the clear decrease of ENSO-related variance in the global mean spectra for the *mid-Holocene* experiments. Beyond the ENSO-related timescales, however, changes in temperature and precipitation variability scale across the experiments without strong regard for timescales. In the ENSO power spectral range of 3-7 years we notice a peak-and-trough pattern of some models, which might represent a change in the ENSO frequency impacting global-scale climate variability. Changes in ENSO spatial patterns, or the event amplitude, however, would not be visible in the spectrum if the overall variance at the timescale did not change. Previous studies have suggested that ENSO variability might increase with global warming (e.g., Cai et al., 2018; Timmermann et al., 1999), but we do not find clear evidence supporting this finding (in agreement with Brown et al., 2020).

The meridional atmospheric gradient modes of variability in both hemispheres (SAM, NAM and NAO) show a weak tendency towards more positive (poleward) displacements of the subtropical high with global mean temperature increases in our experiments. This is consistent with the findings of precipitation reductions in Mediterranean climates at the Western continental edges in both hemispheres. It is, however, unclear to what extent the annular mode (and the westerly jet position) shift due to changes in global mean temperature and the general circulation, or due to ice-sheet height and sea-ice changes that might, to some extent, be independent of the change in the mean (Chavaillaz et al., 2013).

### 4.2 Temperature vs. precipitation scaling

We find that, globally averaged, precipitation variability increases Global averaged precipitation variability increases with the global mean temperature of the experiments in our analyses (Figs. 2,3,4b). There is a larger degree of correlation between mean and variability change for precipitation: The drier climates dry regions in the *lgm* experiment are spatially extensive, and highly correlated with areas of lower precipitation variability (Fig. 3). Conversely, wetter regions in the idealized warming scenarios are also those which also show higher precipitation variability. Yet, we find no relationship between the sensitivity propensity of a model for precipitation increase and its sensitivity to warming under GHG increase, to its propensity for precipitation increase increased GHG forcing (Tab. 1). The overall scaling of 1.7%/K that we find across the model ensemble is somewhat lower than the 2%/K Li et al. (2013) found for a similar (although somewhat smaller) set of CMIP5 models and experiments and. This is also smaller than what has been established for earlier models (Allen and Ingram, 2002; Held and Soden, 2006). We note, however, that mean precipitation in the *lgm- to midHolocene-* experiments scales with the respective temperature anomalies by 2%/K, and it is the idealized transient warming scenarios that fall below these lines. This could indicate that the temperature in these experiments temperature changes are changes faster than precipitation responses and, would the experiments be continued, they would get closer to the expected line (Samset et al., 2018; Myhre et al., 2018; Andrews et al., 2010). Indeed, Samset et al. (2018) found that the precipitation increase over the global oceans is markedly slower than that over land, which perhaps explains why we find a scaling that our scaling is closer to the terrestrial response in equilibrium experiments (1.8%/K, Li et al. (2013))(1.8%/K, Li et al., 2013). Andrews et al. (2010) established that the atmospheric response correlates strongly with the atmospheric component of the radiative forcing, whereas the slow response is, independent of
the mechanism leading to the global temperature change, of 2-3%/K. It is unclear, also unclear, how precipitation variability relates to precipitation extremes, as they typically operate on much shorter timescales. O’Gorman et al. (2011) found, based on CMIP3 model simulations, that extratropical precipitation extremes increase with 6%/K, and hence at a rate closer to the thermodynamic rate of 7%/K (Allen and Ingram, 2002; Held and Soden, 2006). Global mean precipitation rates are, therefore, increasing with warming. At daily to interannual timescales, soil moisture plays a relevant role in the precipitation feedback on temperature variability (Vidale et al., 2007; Fischer and Knutti, 2013). It is, however, also clear that models have difficulties representing these feedbacks at the land surface, in particular on longer timescales (Rehfeld and Laepple, 2016). The detail of representation of sub-grid-scale convective processes could also determine whether a local feedback is modeled positively or negatively (Hohenegger et al., 2009). The observed negative coupling between local temperature and precipitation variability at short timescales (e.g., Trenberth and Shea, 2005, found local correlations up to -0.7) therefore should feed back onto higher temperature variability. To what degree, however, we cannot assess from this analysis, as as synoptic-scale processes are not resolved in the monthly data available.

4.3 Comparison to climate reconstructions and observations

Analysis of instrumental records has shown that the number of record-breaking rainfall events has been increasing over the instrumental era (Lehmann et al., 2015). This is consistent with an ongoing increase in the global mean precipitation rate. Evidence for continental-scale colder/drier conditions at the LGM comes from a variety of terrestrial proxies (Kohfeld and Harrison, 2000; Bartlein et al., 2011), as well as oceanic proxies (MARGO project members, 2009). The sampling rate and resolution of proxies for palaeohydrology is are, however, often not sufficient to investigate changes in precipitation variability. A high-resolution speleothem record allowed (Luetscher et al., 2015) to relate shifts in the LGM westerly storm tracks in Europe, which are consistent with our finding of enhanced precipitation in the lgm experiments. Koutavas and Joanides (2012) suggested that ENSO variability was higher at the LGM than in the Holocene. It is, however, unclear how this relates to our finding of more La-Niña-like conditions in most model simulations, but a reduced ENSO variance in ENSO 3.4 region SSTs has been corroborated by isotope proxies and isotope-enabled modeling (Zhu et al., 2017). Other studies found ENSO variability to become more persistent with GHG-induced warming (Cai et al., 2014). Especially for precipitation, our results suggest that mean state and variability changes are coupled for both temperature and precipitation. Notwithstanding methodological challenges (Rehfeld et al., 2016; Rehfeld and Laepple, 2016), if robust joint/co-located estimates of past temperature and precipitation can be obtained from proxy data, these could potentially serve as constraints on future projections (Schmidt et al., 2014).

4.4 Limitations

We have shown that patterns of temperature and precipitation variability in palaeoclimate and future simulations mirror each other one another, bringing together equilibrium and transient experiments. Nevertheless, there are important limitations that preclude a direct interpretation for future projections (Christensen et al., 2013). Firstly, the snapshots we have been able to analyze are short analyzed here are short (50 years), and therefore many longer slower modes of variability, operating on
decadal or longer timescales. Furthermore, we are not able to investigate the variability in the index time series, but only their mean strength. Analysis of multidecadal modes, and driving mechanisms of variability changes (e.g., from the ocean circulation, or sea-ice mechanisms) in future studies could provide critical insights, and would strongly benefit from the availability of stored model diagnostics (e.g., AMOC strength) and longer simulation output. We have limited the analysis to linear properties of the surface climate fields, and therefore cannot distinguish whether local changes are remotely forced, e.g., due to changing teleconnections from ENSO. We find that, while temperature variability decreases in the model simulations from the $l_{gm}$ to the future $1pctCO_2$ and $abrupt4xCO_2$ scenarios, the magnitude of change is far lower than that observed in proxy data on longer timescales (Rehfelld et al., 2018). This could be due to models underestimating regional variability beyond the multidecadal timescale (Laepplle and Huybers, 2014b; Rehfelld et al., 2018).

At the global scale, climate models do, however, capture correct levels of intraannual to multi-decadal temperature variability (Laepplle and Huybers, 2014a; Pages2k-Consortium, 2019).

5 Conclusions

We have investigated the simulated changes in surface climate variability across a wide range of climates based on the PMIP3/CMIP5/CMIP6 model ensembles. Across the ensemble, we find global patterns of changes which are roughly opposite between cold ($l_{gm}$) and warm ($1pctCO_2/abrupt4xCO_2$) experiments. Global mean precipitation increases with temperature from cold to as warm as preindustrial to warm scenarios. While the simulated global mean precipitation is found to increase with global temperatures. This is true for the change from pre-industrial conditions into idealized CO$_2$-induced warming scenarios. We also find a similar increase from the cold period of the LGM to the pre-industrial reference period. Simulated temperature variability is, at the global scale, higher in the $l_{gm}$ scenarios, and decreases globally with temperature, precipitation variability with temperature. Precipitation variability, on the other hand, is lower in the cold state, and higher for the warmer scenarios. In There are regions which display opposite patterns: in both hemispheres, precipitation changes at the mid-latitude western coasts of the continents (California, Patagonia, South Africa, Southern Australia, and the Mediterranean) are the inverse of the global mean change in precipitation. They display more precipitation variability in the $l_{gm}$ scenario, and consistently lower precipitation, and precipitation variability, in the $1pctCO_2$ and $abrupt4xCO_2$ scenarios. The circulation modes that affect these regions remain consistent across the model ensemble. We investigated, but did not find, an universal relationship between the variability of climate modes and global mean temperature change. No model shows a reduction in temperature variance as large as that for centennial-to-millennial timescales observed in palaeoclimate data for the Last Glacial to Holocene transition, but this could be due to the much shorter timescales we have investigated here. Yet, on intraannual seasonal to multidecadal timescales, we find evidence of scaling, and that changes in variability appear to occur proportionally across these timescales. Interannual precipitation variability across these simulations appears to robustly, and linearly, relate the relative change in regional variance and the relative change in the mean precipitation. This relationship, and the consistency across timescales, could imply that hydroclimate proxy reconstructions at decadal to centennial timescales provide an additional constraint on simulated past and future precipitation variability changes.
Code and data availability. Model data is freely available on the ESGF. The Climate Variability Diagnostic Package is available at http://www.cesm.ucar.edu/working_groups/CVC/cvdp/. Processed data is available on the PMIPVarData website at www.past2future.org. Scripts and code are available on request.

Author contributions. KR and CB devised the project. All coauthors jointly planned the analyses. CB performed the data standardization and computed the climate modes. KR performed the correlation and regression analyses. ML and JL performed the extratropical composites. RH performed spectral analyses. All authors prepared figures. KR wrote the paper with contributions from all coauthors. All authors commented on the final draft.

Competing interests. The authors declare no competing interests.

Acknowledgements. We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP, the US Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led the development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. K.R. acknowledges funding by the German Research Foundation (DFG, code RE3994-2/1) and the Heidelberg Center for the Environment for providing a venue for discussion. We are thankful for the support of colleagues in two working groups: PAGES Climate Variability Across Scales (CVAS) and PMIP Past2Future: insights from a constantly varying past. Special thanks are due to Julia Hargreaves, Darrell Kaufman, and Jarmo Kikstra. We thank the two reviewers for their thoughtful comments.
References


Lorenz, E. N.: Seasonal and Irregular Variations of the Northern Hemisphere Sea-Level Pressure Profile, https://doi.org/10.1175/1520-0469(1951)008<0052:saivot>2.0.co;2, 1951.


Figure 5. Sea-level pressure (in hPa) and surface air temperature (in °C) anomaly composites for high precipitation years in five regions with Mediterranean climates (indicated by green boxes on the left-hand panels). Sea-level pressure anomaly composites (panels a,c,e,g,i on the left) and surface air temperature anomaly composites (panels b,d,f,h,j on the right) show the large-scale patterns across models and experiments for composites over years of anomalously high precipitation. In the selected regions, years with (defined as one standard deviation above the average) were composited for each simulation. Green boxes show the regions of interest on the left-hand panels. Stippling shows regions wherein fewer than two-thirds or more of the simulations agree on the sign of the pattern. SFigs 2 and 3 show the corresponding composites for anomalously low precipitation, and composites for the precipitation change in these years.
Figure 6. Changes of the global mean of the power spectra between the experiments. For temperature (a), variance across all timescales and for most models is highest in the \textit{lgm} experiment, and decreases for the warmer experiments. This is the opposite for precipitation (c), which sees moderate increases in precipitation variability with warming. For each model, we took the ratio of the global mean spectra of each experiment over the \textit{piControl} for both temperature (b) and precipitation (d), thus showing the timescale-dependency of the local variance change. Shaded confidence intervals are based on the entire range of the model ratios.