ENVIRONMENTAL RESEARCH
LETTERS

Changes to population-based emergence of climate change from CMIP5 to CMIP6

Changes to population-based emergence of climate change from CMIP5 to CMIP6

Hunter C. Douglas[1], Luke J. Harrington[2], Manoj Joshi[3,4], Ed Hawkins[5], Laura E. Revell[6], and David J. Frame[6,7]

[1] New Zealand Climate Change Research Institute, Victoria University of Wellington Te Herenga Waka, Wellington, New Zealand
[2] Te Aka Mātuatua School of Science, University of Waikato, Hamilton, New Zealand
[3] Climatic Research Unit, University of East Anglia, Norwich, U.K.
[4] School of Environmental Sciences, University of East Anglia, Norwich, U.K.
[5] Department of Meteorology, University of Reading, Reading, U.K.
[6] School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand
[7] School of Earth and Environment, University of Canterbury, Christchurch, New Zealand

November 20, 2022

Key Points

- CMIP6 models show that emergence of anomalous average annual temperatures will occur earliest and strongest in tropical low latitudes.
- Changes to forcings and model responses cause signal-to-noise ratios to increase compared to CMIP5, with notable exceptions in some highly populated regions.
- The tropics are disproportionately home to lower-income nations, which are also projected to experience higher population growth under the Shared Socio-economic Pathways (SSP).

Abstract

The Coupled Model Intercomparison Project Phase 6 (CMIP6) model ensemble projects climate change emerging soonest and most strongly at low latitudes, regardless of the emissions pathway taken. In terms of signal-to-noise (S/N) ratios of average annual temperatures, these models project earlier and stronger emergence under the Shared Socio-economic Pathways (SSPs) than the previous generation did under corresponding Representative Concentration Pathways (RCPs). Spatial patterns of emergence also change between generations of models; under a high emissions scenario, mid-century S/N is lower than previous studies indicated in Central Africa, South Asia, and parts of South America, West Africa, East Asia, and Western Europe, but higher in most other populated areas. We show that these global and regional changes are caused by a combination of higher effective climate sensitivity (ECS) in the CMIP6 ensemble, as well as changes to emissions pathways, component-wise effective radiative forcing (ERF), and region-scale climate responses between model generations. We also present

the first population-weighted calculation of climate change emergence for the CMIP6 ensemble, quantifying the number of people exposed to increasing degrees of abnormal temperatures now and into the future. Our results confirm the expected inequity of climate change-related impacts in the decades between now and the 2050 target for net-zero emissions held by many countries. These findings underscore the importance of concurrent investments in both mitigation and adaptation.

1 Introduction

Achieving net-zero emissions by the mid-2050s is required to limit global warming to less than 1.5 K (with limited overshoot) (IPCC, 2022), and several countries have set net-zero targets for the decade 2041-2050 (Hale et al., 2022). Whether or not substantive action is taken to reduce emissions, the climate will continue to change until the point net-zero emissions are reached (Allen et al., 2009, Zickfeld et al., 2012, MacDougall et al., 2020). Understanding when and how the climate change signal emerges from the noise of natural variation during this period is important for assessing the likely impacts of climate change and how to mitigate, prepare for, and adapt to them.

Signal-to-noise ratio is an established metric for emergence (Hawkins and Sutton, 2012, Frame et al., 2017, Hawkins et al., 2020). S/N has commonly been applied to seasonal or longer-term average temperatures to assess when and how the impacts of climate change will be experienced, indicated by the time at which areas exceed S/N thresholds or the magnitude of the S/N ratio at a given time (e.g. Mahlstein et al., 2011, Hawkins and Sutton, 2012). This approach has also been applied to projections of precipitation, drought, and ocean parameters (Giorgi and Bi, 2009, King et al., 2015, Chen et al., 2021 Section 1.4.2.2), and to observational data (Mahlstein et al., 2012, Hawkins et al., 2020). Annual average temperature signal-to-noise has been found to be strongest in the tropics due to the lower internal variability in these regions (Figure S1, Mahlstein et al., 2011, Harrington et al., 2017). Because local ecosystems are adapted to the lower variability in these regions, the same increase in annual temperatures can lead to greater impacts (Walther et al., 2002, Williams et al., 2007, Beaumont et al., 2011, Mora et al., 2013). We note that emergence occurs much sooner in average annual temperatures than, say, monthly or daily, due to smaller internal variability as timescales lengthen (Harrington, 2021).

Frame et al., 2017 analysed data from 25 models in the Coupled Model Intercomparison Project, phase 5 (CMIP5) alongside population data to assess when the world would be exposed to annual average temperatures that crossed S/N ratio thresholds (i.e. numbers of standard deviations from the mean) of 1, 2, and 3, relative to a recent baseline of 1986-2005. The authors designated these thresholds “unusual”, “unfamiliar”, and “unknown” climates, respectively. They found that the world’s population would be exposed to different climates faster than the surface area on average and that the changes would be experienced earlier and more severely in lower-income regions. Hawkins et al., 2020 added the designation “inconceivable” for S/N values above 5.

Many updates have been made to the ensemble of global climate models between CMIP5 and CMIP6. These include increased resolution, more participating models, and updated parameterisations of sub-gridcell-scale physical processes that more closely align with the latest understanding of climate drivers such as radiative transfer, cloud microphysics, aerosol chemistry, sea ice dynamics, land cover, and stochasticity (Chen et al., 2021 Section 1.5.3.1, Eyring et al., 2021 Section 3.8.2). These have led to better agreement with observational datasets and reanalyses (Bock et al., 2020). Past warming over the instrumental period is often well simulated by these models, with the multi-model average being close to the best estimate from observations and reanalyses (Arias et al., 2021), although many higher sensitivity models struggle to simulate aspects of the satellite period and deep-time paleoclimate periods (Bock et al., 2020, Kageyama et al., 2021, Otto-Bliesner et al., 2021). CMIP6 models exhibit a wider range of effective climate sensitivity (ECS), primarily due to updates in the representation of extratropical cloud feedbacks and aerosol interactions (Meehl et al., 2020, Zelinka et al., 2020). Such a range of model responses represents the main source of uncertainty for projections of future temperatures under high-emissions scenarios, whereas uncertainty in the effects of short-lived forcings like aerosols dominate for low-emissions scenarios (Arias et al., 2021). The emissions pathways specified for the SSPs for CMIP6 were not intended to reproduce those in the RCPs for CMIP5 (O'Neill et al., 2016), though the net radiative forcing is very similar over time in corresponding scenarios (Gidden et al., 2019). However, forcing due to individual components can be considerably different due to the different emissions pathways of each (see Meinshausen et al., 2020 and Figure S2). Considering the changes in greenhouse gas emissions pathways between CMIP5 and CMIP6, the CMIP6 scenarios exhibit higher projected CO_2 emissions relative to their CMIP5 counterparts for most of the century. CH_4 emissions are slightly higher for SSP1-2.6 and SSP2-4.5, and considerably lower for SSP5-8.5. N_2O emissions are generally lower for all scenarios, particularly so in SSP5-8.5. The net effect of these changes isn't immediately apparent, and will differ from model to model and across timescales.

Considering aerosol emissions, the CMIP6 ensemble exhibits a greater spread in projected emissions across scenarios (Gidden et al., 2019). SO_2 emissions are generally lower in SSP1-2.6 and generally higher in SSP2-4.5 and SSP5-8.5, while black carbon (BC) emissions are generally lower in SSP1-2.6, higher in SSP5-8.5, and vary in SSP2-4.5. Aerosols are not well-mixed in the atmosphere, and so have regional impacts on temperature. Recent studies have assessed the forcing due to aerosols prescribed for CMIP6, taking into account transport (e.g. Lund et al., 2019), though directly comparable studies between model generations that account for differing model responses are not yet available.

The CMIP5 model ensemble exhibited systematic biases in their response to climate forcings, including a warm bias in the Southern Ocean attributed to deficiencies in cloud processes (Hyder et al., 2018). Modelling groups implemented different improvements to address biases, such as new planetary boundary layer and convection schemes in the NASA GISS model (Stanfield et al., 2015), updated aerosol optical properties and natural emission rates in CanESM5 (Swart, Cole, et al., 2019), and including aerosol indirect effects in BCC-CSM (Wu et al., 2019). These changes have resulted in improved agreement with observations in aerosol- and cloud-related metrics (Cherian and Quaas, 2020), but quantifying overall improvement between model generations remains challenging (Szopa et al., 2021 Section 6.4). Models' responses to individual forcing agents can, however, be quantified in terms of effective radiative forcing (ERF), a simulation-derived measure of the effect of an agent on the earth's radiative budget.

Here, we present an analysis of population-based exposure to unusual climates, updating the approach used in Frame et al., 2017 with results from CMIP6. The SSPs provide projections for country-level population estimates that vary over time and scenario. This level of detail was not available for the RCPs. We show how the climatic and population changes projected in the SSPs interact, how analysis using these updated data compares to the findings of earlier studies, and what factors cause the observed changes.

2 Methods

We obtained monthly average temperature climate model output data from the World Climate Research Programme's CMIP (Phase 6) (Eyring et al., 2016). We selected five scenarios from ScenarioMIP (O'Neill et al., 2016) that span the range of future outcomes: SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. Three of these have corresponding scenarios from the previous generation of RCPs: RCP2.6, RCP4.5, and RCP8.5. Other scenarios represented by fewer than 15 models were excluded. The first two SSP scenarios (SSP1-1.9 and SSP1-2.6) result in global warming of approximately 1.5 and 2.0 K at 2100, respectively, in line with Paris Climate Agreement targets. Results for all scenarios plus the historical and pre-industrial control (piControl) simulations were available for 37 climate models, with the exception of SSP1-1.9, for which only 15 models' results were available. The CMIP6 models used are listed in Table 1.

For comparison, we analysed results for 29 climate models from the CMIP5 generation that used the RCPs (Taylor et al., 2012), listed in Table 2. To assess the statistical significance of changes between model ensembles, we applied a two-sided student's T-test with a 90% threshold at each gridpoint and adjusted the threshold to account for spatial autocorrelation using a False Discovery Rate (FDR) control procedure, following Wilks, 2016. To help diagnose the causes of changes between CMIP generations, we also repeated the analysis using the 25 CMIP6-era models with published ECS within the same range as CMIP5-era models (2.08-4.67 K). We selected these 25 models based on ECS values published by Meehl et al., 2020; Nijsse et al., 2020; Schlund et al., 2020; and Zelinka et al., 2020. ECS values were not published for three of the 37 models, which we excluded. Fyfe et al., 2021 used two generations of the CanESM model to disentangle the effects of changes in the model parameterisation and the forcings applied from CMIP5 to CMIP6, finding that the different forcings have significant impacts. We similarly disentangled causes for the observed differences by calculating signal and noise on three sets of results: CanESM2 run on CMIP5 forcings, CanESM5 run on CMIP5 forcings, and CanESM5 run on CMIP6 forcings.

We processed monthly mean, near-surface (2 m) air temperature data to create continuous timeseries from January 1850 to December 2100. We defined noise and signal following Frame et al., 2017: noise at each gridpoint is the standard deviation in annual temperatures from the last 200 years of each model's piControl simulation, and signal is degrees Kelvin change from a 1986-2005 baseline. We additionally de-trended the piControl data before calculating noise, as we found that multiple models exhibited unexpected, statistically significant trends in annual temperatures, possibly due to insufficient spin-up time in the control simulation (Figure S3). The choice of baseline will affect results, with higher S/N ratios for earlier baselines. Our choice of a relatively recent baseline aligns with prior work and expresses change relative to living memory for a large proportion of the world's population. We tested sensitivity to this choice by alternatively using an earlier baseline of 1961-1990. Signal, noise, and signal-to-noise are calculated for each realisation of a given model and emissions scenario before averaging across realisations. The global mean surface temperature (GMST) signal (change in annual average GMST since this same baseline) against which local data are regressed is smoothed

Model	SSP Experiments					ECS	Reference
	1-1.9	1-2.6	2-4.5	3-7.0	5-8.5		
ACCESS-CM2	0	3	3	3	3	4.72	Dix et al., 2019
ACCESS-ESM1-5	0	10	19	10	10	**3.87**	Ziehn et al., 2019
AWI-CM-1-1-MR	0	1	1	5	1	**3.16**	Semmler et al., 2019
BCC-CSM2-MR	0	1	1	1	1	**3.04**	Xin et al., 2019
CAMS-CSM1-0	2	2	2	2	2	**2.29**	Rong, 2019
CanESM5	50	50	50	50	50	5.62	Swart, Cole, et al., 2019
CAS-ESM2-0	0	2	2	2	2	**3.51**	Chai, 2020
CESM2	0	3	3	3	3	5.16	Danabasoglu, 2019a
CESM2-WACCM	0	1	5	3	5	4.75	Danabasoglu, 2019b
CMCC-CM2-SR5	0	1	1	1	1	**3.52**	Lovato and Peano, 2020
CMCC-ESM2	0	1	1	1	1	-	Lovato et al., 2021
CNRM-CM6-1	0	6	10	6	6	4.83	Voldoire, 2019a
CNRM-CM6-1-HR	0	1	1	1	1	**4.28**	Voldoire, 2019b
CNRM-ESM2-1	5	5	10	5	5	4.76	Seferian, 2019
EC-Earth3	10	11	11	11	18	**4.30**	Consortium (EC-Earth), 2019a
EC-Earth3-Veg	3	7	8	6	8	**4.31**	Consortium (EC-Earth), 2019b
EC-Earth3-Veg-LR	3	3	3	3	3	-	Consortium (EC-Earth), 2020
FGOALS-f3-L	0	1	1	1	1	**3.00**	Yu, 2019
FGOALS-g3	1	4	4	5	4	**2.88**	Li, 2019
GFDL-ESM4	1	1	3	1	1	**2.60**	John et al., 2018
GISS-E2-1-G	6	11	20	18	11	**2.72**	NASA/GISS, 2020a
GISS-E2-1-H	2	10	10	6	10	**3.11**	NASA/GISS, 2020b
IITM-ESM	0	1	1	1	1	-	Panickal and Narayanasetti, 2020
INM-CM4-8	0	1	1	1	1	1.83	Volodin et al., 2019a
INM-CM5-0	0	1	1	5	1	1.92	Volodin et al., 2019b
IPSL-CM6A-LR	6	6	11	11	6	**4.56**	Boucher et al., 2019
KACE-1-0-G	0	3	3	3	3	**4.48**	Byun et al., 2019
MCM-UA-1-0	0	1	1	1	1	**3.65**	Stouffer, 2019
MIROC-ES2L	4	10	30	10	3	**2.68**	Tachiiri et al., 2019
MIROC6	1	10	50	3	50	**2.61**	Shiogama et al., 2019
MPI-ESM1-2-HR	0	2	2	10	2	**2.98**	Schupfner et al., 2019
MPI-ESM1-2-LR	0	8	10	10	7	**3.00**	Wieners et al., 2019
MRI-ESM2-0	1	1	10	5	1	**3.15**	Yukimoto et al., 2019
NorESM2-LM	0	1	3	3	1	**2.54**	Seland et al., 2019
NorESM2-MM	0	1	2	1	1	**2.50**	Bentsen et al., 2019
TaiESM1	0	1	1	1	1	**4.31**	Lee and Liang, 2020
UKESM1-0-LL	5	16	17	16	5	5.34	Good et al., 2019

Table 1: Number of realisations analysed for CMIP6 models. ECS values in bold are within the CMIP5 range of 2.08-4.67 K. Shaded cells indicate models with aerosol optical depth data used in the analysis.

using a fourth-order polynomial fit. We tested sensitivity to this smoothing approach by alternatively using two other techniques: a 20-year rolling average and a 41-year lowess filter, as per Hawkins et al., 2020. We compared results against a global one-eighth degree gridded population dataset with projections for each of the SSPs (Jones and O'Neill, 2016, v1.01). For the RCPs, we applied the population pathway of the corresponding SSP. Data processing is further described in Supplementary Data.

Following the categorisations in Frame et al., 2017, we assessed exposure to signal-to-noise thresholds for different socioeconomic and geographic groupings of countries. These groupings are outlined in Table 3. There is some overlap between groupings (e.g. Indonesia is in both Association of Southeast Asian Nations and Global Emerging Markets).

Model	RCP Experiments			ECS	References
	2.6	4.5	8.5		
BNU-ESM	1	1	1	4.04	Ji et al., 2014
CCSM4	6	6	6	2.94	Meehl et al., 2012
CESM1-CAM5	3	3	3	-	Gent et al., 2011
CESM1-WACCM	3	3	3	-	Calvo et al., 2012
CNRM-CM5	1	1	5	3.25	Voldoire et al., 2013
CSIRO-Mk3-6-0	10	10	10	4.09	Rotstayn et al., 2012
CanESM2	5	5	5	3.70	Arora et al., 2011
EC-EARTH	2	11	8	-	Hazeleger et al., 2012
FGOALS-g2	1	1	1	3.38	Li et al., 2013
FIO-ESM	3	3	3	-	Qiao et al., 2013
GFDL-CM3	1	3	1	3.97	Donner et al., 2011
GFDL-ESM2G	1	1	1	2.43	Dunne et al., 2012
GFDL-ESM2M	1	1	1	2.44	
GISS-E2-H	3	16	5	2.31	Schmidt et al., 2006
GISS-E2-R	3	17	5	2.12	
HadGEM2-AO	1	1	1	-	Martin et al., 2011
HadGEM2-ES	4	4	4	4.61	Collins et al., 2011
IPSL-CM5A-LR	4	4	4	4.13	Dufresne et al., 2013
IPSL-CM5A-MR	1	1	1	4.12	
MIROC-ESM	1	1	1	4.67	S. Watanabe et al., 2011
MIROC-ESM-CHEM	1	9	1	-	
MIROC5	5	5	5	2.72	M. Watanabe et al., 2010
MPI-ESM-LR	3	3	3	3.63	Giorgetta et al., 2013
MPI-ESM-MR	1	3	1	3.46	
MRI-CGCM3	1	1	1	2.61	Yukimoto et al., 2012
NorESM1-M	1	1	1	2.80	Iversen et al., 2013
NorESM1-ME	1	1	1	-	
bcc-csm1-1	1	1	1	2.83	Wu, 2012
bcc-csm1-1-m	1	1	1	2.89	

Table 2: Number of realisations analysed for CMIP5 models. Shaded cells indicate models with aerosol optical depth data used in the analysis.

Group	Full name	States	Approx. 2010 population	Description
ASEAN	Association of Southeast Asian Nations	10	650,000,000	
AOSIS	Alliance of Small Island States	39	61,000,000	
GEM	Global Emerging Markets	23	3,700,000,000	Those countries in the G20 that are not in OECD90
LDC	Least Developed Countries	58	1,500,000,000	Countries with 2020 Human Development Indices lower than India's (Conceição, 2020)
OECD90	Organisation for Economic Co-operation and Development (1990)	24	1,000,000,000	Member states of the OECD as of 1990

Table 3: Country groupings

3 Results and Discussion

3.1 Mid-century signal-to-noise

Figure 1 depicts the geospatial emergence of temperature signal-to-noise in the mid-twenty-first century, 2040-2060 (M21C). The findings are qualitatively similar to previous studies that

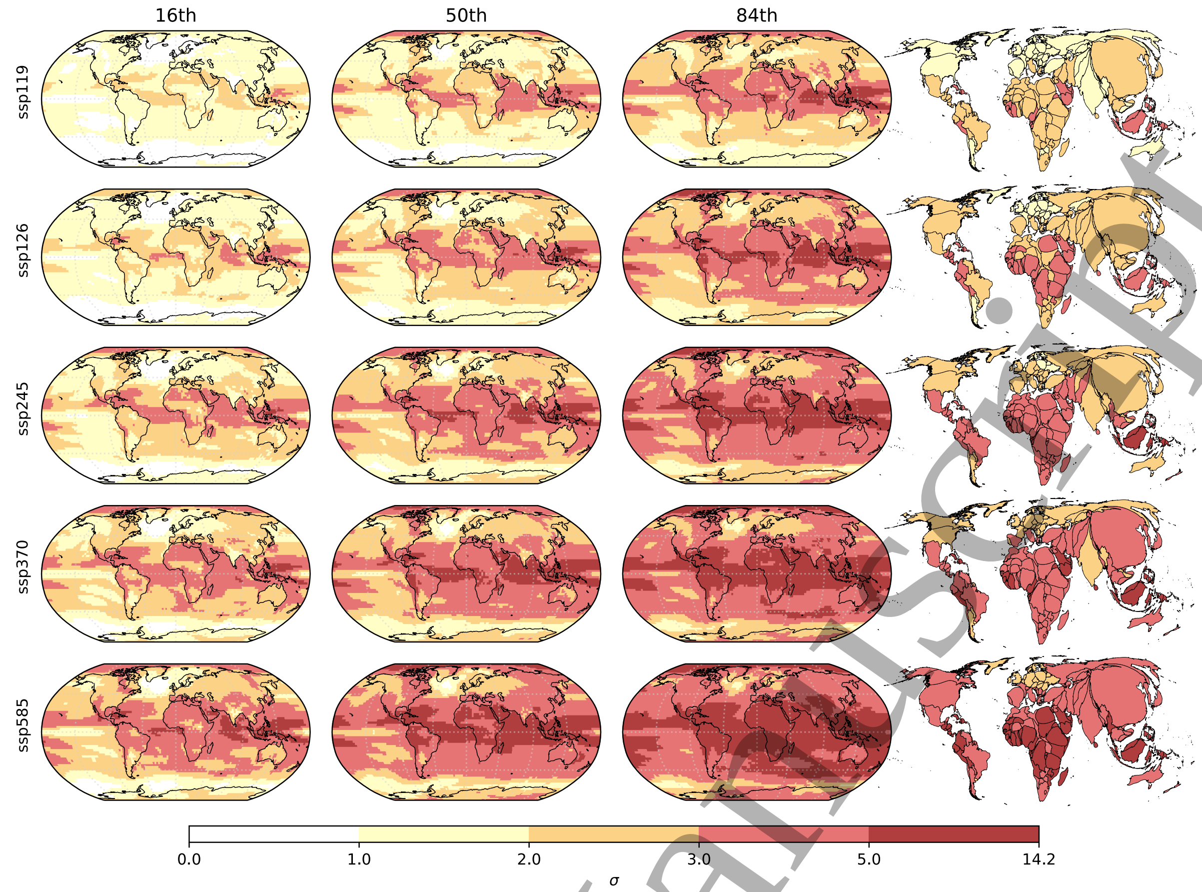

Figure 1: Multi-model signal-to-noise ratios in 2040-2060 average annual temperatures. From left to right, the columns show the 16$^{\text{th}}$, 50$^{\text{th}}$ (median), and 84$^{\text{th}}$ percentile results across models, and the fourth column shows country-averaged S/N (median) in population-weighted cartograms. The rows, from top to bottom, show results for SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0, and SSP5-8.5. Since both S and N have dimensions of K, S/N is dimensionless, but can be expressed as multiples of the noise (i.e. numbers of standard deviations, σ). Colours correspond to the named S/N ratio thresholds of 1, 2, 3, and 5, i.e. "unusual", "unfamiliar", "unknown", and "inconceivable" climates.

used earlier model generations (e.g. Mahlstein et al., 2011, Hawkins and Sutton, 2012, Frame et al., 2017) in that S/N is most pronounced in low latitudes due in large part to these areas' low inter-annual variation (noise) in annual mean temperatures. Despite the greater absolute warming near the poles, these regions also exhibit higher noise, resulting in comparatively low S/N ratios (e.g. Hawkins et al., 2020). See Figure S1 for noise and signal calculated individually. In low and mid-latitudes, both signal and noise are greater over land than the adjacent ocean, resulting in less land/sea contrast for S/N than for signal or noise individually. Scenarios with higher radiative forcing exhibit predictably higher M21C S/N across all regions.

These results hold qualitatively when signal-to-noise is computed for the warmest monthly average temperatures each year instead of annual average (Figure S4), though the magnitude is depressed due to higher variation in monthly temperatures, especially over land. Figure S5 shows equivalent results for CMIP5 models and RCPs, and Figure S6 shows equivalent results for the late-twenty-first century period of 2071-2100 used in previous studies. Using the earlier baseline of 1961-1990 uniformly increases signal across the globe in all scenarios, due to the lower GMST at that time (not shown). The results are slightly sensitive to the GMST smoothing technique. Using the alternative 41-year lowess filter approach resulted in faster apparent emergence. Global-average M21C S/N is 12% higher for SSP1-1.9 and 4% higher for SSP5-8.5 using this approach. However, similar changes apply to the RCPs, and the spatial patterns of emergence remain unchanged. We report results using the 4th-order polynomial approach throughout.

The range of results across the ensemble of models is represented in Figure 1 by the columns with the 16th, 50th, and 84th percentile results. The 16/50/84th-percentile S/N values are calculated and displayed at each gridpoint, as opposed to showing all gridpoints for the model with 16/50/84th-percentile global-average S/N values. Comparing results across columns in Figure 1 thus provides a conservative estimate of model uncertainty. See Figure S7 for the percent of gridpoints represented by each model in the 16/50/84th-percentile plots. More sensitive models are more represented in the 84th-percentile plot and less sensitive models in the 16th-percentile plot, but most models contribute data to all plots. There is generally a bigger difference in S/N between results in adjacent columns than between those in adjacent rows, which are representative of scenario uncertainty. While other percentiles and scenarios could validly be chosen, this indicates that model uncertainty is comparable to or greater than scenario uncertainty as of mid-century. Scenario and model uncertainties have been found to be equal at around 50 years from outset when calculating global, long-term, mean near-surface air temperature for CMIP3 (Hawkins and Sutton, 2009) and CMIP6 (Lehner et al., 2020).

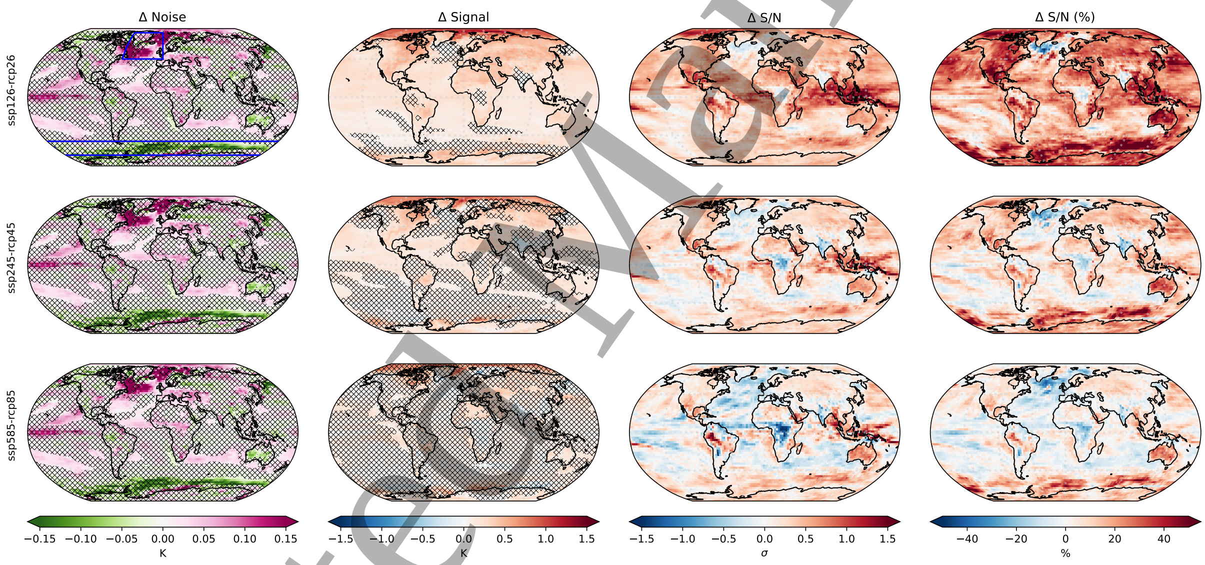

Figure 2: Changes in baseline-period noise (column 1), 2040-2060 average signal (column 2), and S/N ratio (column 3) from CMIP5 RCP scenarios to corresponding CMIP6 SSPs (multi-ensemble, multi-model medians, as per Figure 1). Areas are hatched where the change is not statistically significant under a two-sided student's T-test at the 90% level, adjusted for spatial autocorrelation. No such testing is applied to the S/N ratio, an inherent measure of significance. Results are multi-model medians for SSP1-2.6 - RCP2.6 (top row), SSP2-4.5 - RCP4.5 (middle row), and SSP5-8.5 - RCP8.5 (bottom row). The fourth column shows changes in S/N as a percentage of CMIP5-era S/N. The regions outlined in blue in the top left panel are the North Atlantic and Southern Ocean regions discussed throughout.

Figure 2 shows the differences in multi-model median noise and M21C signal and S/N between the CMIP6 and CMIP5 ensembles across the globe. Global, area-weighted average changes are summarised in Table 4: noise increases by 2.9% in CMIP6 and signal increases by 6.9-27% (depending on the scenario), resulting in S/N increases of 2.7-22%. Overall, M21C S/N is higher in the newer generation of models. Table 4 and Figure S8 include the same values calculated for the sub-population of 25 CMIP6 models with ECS in the range of CMIP5 models.

CMIP6 Ensemble	Scenario	Δ Noise - K (%)	Δ Signal - K (%)	Δ S/N - σ (%)
All CMIP6 models (37)	SSP126-RCP26	0.0099 (2.9%)	0.23 (27%)	0.44 (22%)
	SSP245-RCP45		0.16 (13%)	0.23 (8.6%)
	SSP585-RCP85		0.12 (6.9%)	0.075 (2.7%)
ECS in CMIP5 range (25)	SSP126-RCP26	0.012 (4.1%)	0.20 (22%)	0.28 (15%)
	SSP245-RCP45		0.13 (10%)	0.058 (3.1%)
	SSP585-RCP85		0.069 (3.9%)	-0.17 (-3.2%)

Table 4: Global (area-weighted) average changes from CMIP5 to CMIP6 in baseline-period noise and 2040-2060 average signal and S/N ratio. Absolute and percent changes are shown. Each CMIP5 RCP scenario is paired with its corresponding CMIP6 SSP scenario based on nominal radiative forcing. Global-average signal and noise values for CMIP5 and CMIP6 are also shown in Figure S1.

The first column of Figure 2 shows that between the CMIP5 RCP scenarios and the corresponding CMIP6 SSPs, variability increases slightly across much of the globe (most strongly in the Northern Atlantic), and decreases in the Southern Ocean. These changes are statistically significant in a few regions, including the North Atlantic. Over land, this includes a decrease in Western Australia and increases in Southern Europe and the northern Middle East, Africa just south of the Sahel, and an area near the China-Russia-Mongolia border. On a global average basis, variability changes very little, as reported in Table 4.

The temperature change signal (column 2) increases across almost all the globe for all scenarios, excepting decreases in South Asia. These changes are more statistically significant in the lower emissions scenarios, with very little of the change in signal passing the significance test for SSP5-8.5. The M21C S/N ratio (column 3) generally increases from CMIP5 to CMIP6, with geographic differences arising from the changes in noise and signal. The changes are more heterogeneous under higher emissions scenarios; SSP5-8.5 (and to a lesser extent, SSP2-4.5) shows higher S/N in most land areas, but lower S/N in Central and West Africa, Western Europe, South and East Asia, and some areas of South America. Many ocean areas also exhibit decreases in S/N. These patterns of reduced S/N arise in areas with increased noise and small (if any) increases in signal.

3.2 Drivers of differences between model generations

Figure 2 captures the result of changes to both external forcing and model response. The depressed signal in the North Atlantic may be due in part to the greater weakening of the Atlantic meridional overturning circulation (AMOC) exhibited across the CMIP6 ensemble (Weijer et al., 2020). Less transport of warm water from low to high latitudes results in a smaller temperature increase in the high-latitude North Atlantic. In the Southern Ocean, the CMIP6 ensemble projects lower noise, increased signal, and a resulting higher S/N ratio compared to the CMIP5 ensemble, for all scenarios. However the changes in signal are not statistically significant for most of the region. The Southern Ocean changes may be due to developments in cloud process representation, an identified source of the warm bias in this region (Hyder et al., 2018), and/or improvements in ocean circulation and surface winds compared to CMIP5, including a weaker Antarctic Circumpolar Current (ACC) (Beadling et al., 2020). CMIP6 models, on average, exhibit a more positive shortwave feedback for extratropical clouds (Arias et al., 2021, Zelinka et al., 2020). It is notable that from CMIP5 to CMIP6, in most parts of the globe, an increase in noise accompanies an increase in signal, resulting in a smaller increase to the S/N ratio.

The highest magnitude of change in noise is observed in high-latitude oceans, suggesting a possible sea-ice influence. We compared available sea ice area data in the piControl period for 21 CMIP5 and 30 CMIP6 models. There were changes in both the average amount and inter-annual variation of sea ice area, but these were not statistically significant and were not consistently correlated with the observed changes in the noise field. We also compared ACC strength in both model generations, but did not find any correlation with noise in the Southern Ocean region. One of the main changes between the piControl simulations for CMIP5 and CMIP6 is that the latter includes a protocol for volcanic aerosols (Fyfe et al., 2021). We compared average aerosol optical depths for these two ensembles, as per Figure 3. There were changes between model generations, including a slight increase in aerosol variability in the North Atlantic and a slight decrease in the Southern Ocean, but these were not found to be statistically significant. The differences in noise, then, are likely driven more by the differing responses to aerosols and other forcing agents between the model generations.

Table 4 and Figure S8 show model inter-generational changes considering just the sub-population of 25 CMIP6 models with ECS in the same range as CMIP5 models. Comparing these results to the earlier ones shows the changes due primarily to differences in the emissions pathways (if we assume that model parameterisation differences manifest as changes in ECS). In the sub-population, noise changes similarly to the full ensemble, and the temperature change signal still increases for all scenarios. On a global-average basis, noise increases by 4.1% in CMIP6 and signal increases by 3.9-22% (depending on the scenario), resulting in S/N changes of -3.2 to +15%. Figure S1 shows separate plots of noise and signal. These results show that the increases in temperature for the CMIP6 ensemble are due to both increased ECS and changes to the emissions pathways. This agrees with single-model and reduced complexity model studies that have isolated the difference due to emissions pathways (Wyser et al., 2020, Nicholls et al., 2020, Fyfe et al., 2021). For changes in S/N ratios, forcing differences are more significant for the lower-emissions scenarios. For the high emissions scenario, global-average S/N even decreases slightly (see Table 4).

We repeated signal and noise calculations on results from two generations of the CanESM model run on both generations' emissions to disentangle forcing and model response influences. For signal, forcing changes account for 44%, 54%, and 38% of the model inter-generational differences on a global average basis in SSP1-2.6, 2-4.5, and 5-8.5, respectively. The differences in noise are generally less significant (see Figure 2), but on a global-average basis, forcing differences account for 88% of change. In the regions of greatest change in noise, however, (the North Atlantic and Southern Ocean regions), changes in model response account for 105 and 103% of the difference in noise, respectively. While these results are just from one modelling centre, they illustrate that changes to both applied forcings and model parameterisation are significant between CMIP5 and CMIP6.

In order to control for global model response and so assess regional differences, we compared noise, signal, and signal-to-noise ratios across model generations at global warming levels (GWL) of 1.5, 2.0, 2.5, and 3.0 K (see Figures S9, S10). We compared average temperatures for the 20 years preceding the year at which the backwards-looking 20-year rolling average crosses each GWL, averaging across all scenarios for which this occurs. The spatial patterns of S/N across GWLs closely match those across emissions scenarios, with emergence strongest in low latitudes (Figure S9). Comparing S/N between model generations at the same GWLs largely controls for signal, so the spatial patterns in changes to the signal-to-noise ratio are dominated by the changes in the noise field (Figure S10).

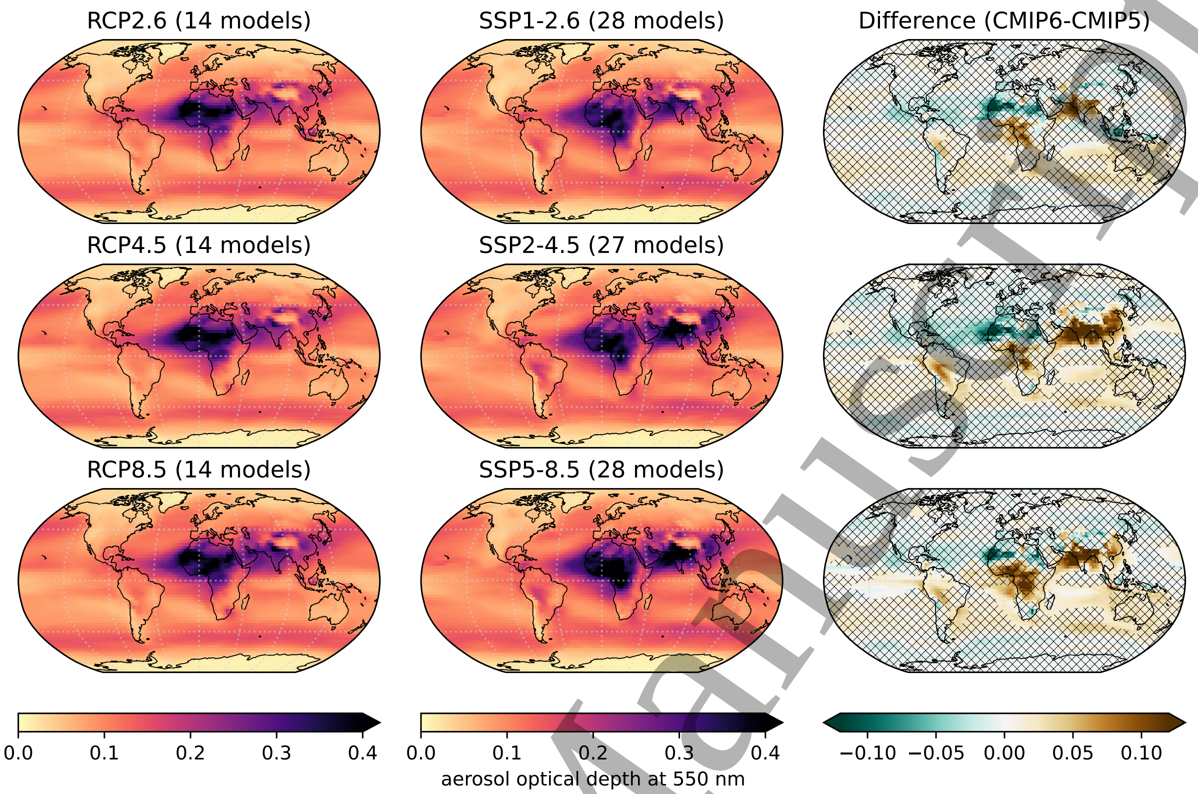

Figure 3: 2040-2060 average aerosol optical depths at 550 nm for CMIP5 models, CMIP6 models, and the difference between these. Areas are hatched where the change is not statistically significant under a two-sided student's T-test at the 90% level, adjusted for spatial autocorrelation. Results are multi-model medians for SSP1-2.6 - RCP2.6 (top row), SSP2-4.5 - RCP4.5 (middle row), and SSP5-8.5 - RCP8.5 (bottom row).

Aerosol emissions have regional impacts on surface temperature due to their short atmospheric residence time. There remains considerable uncertainty in the magnitude of aerosol forcing, which has a significant impact on modelled global temperature (Dittus et al., 2020). While different aerosol species have different direct and indirect effects on net radiation balance, aerosols in aggregate lower insolation, so one would expect greater aerosol concentrations to cause lower temperatures, all else being equal (Zelinka et al., 2014, Smith et al., 2020, Szopa et al., 2021 Section 6.4). To assess the impacts of changes in aerosol forcing between model generations, we calculated geospatial differences in M21C-average ambient aerosol optical depth at 550 nm as a proxy for aerosol concentration. Figure 3 shows the multi-model median results for the two model generations. Note that only a subset of models used in the temperature analysis had optical depth data available (see Tables 1 and 2). Noting this limitation, there are notable consistencies between the aerosol and temperature fields. Compared to CMIP5, CMIP6 models exhibit statistically significant increases in aerosol optical depth above regions in South and East Asia, South America, and south-western Africa, particularly in the moderate and high emissions scenarios. These changes are due to changes in both the prescribed aerosol emissions and the models' handling of aerosols (e.g. circulation, deposition, chemistry, etc.). The slight decreases in signal and more pronounced decreases in signal-to-noise in South Asia shown in Figure 2 correlate well with the aerosol pattern. The significant aerosol increases in South America and East Asia correspond less well with changes in signal-to-noise. In North and Central Africa, while the changes in aerosol optical depth are not statistically significant, they do correspond well with the observed changes in signal-to-noise. It is reasonable to conclude that changes in aerosol loading are responsible for a significant part of the regional differences in signal-to-noise ratios between model generations.

The changes in the greenhouse gas (GHG) and aerosols emissions pathways between model generations have competing effects and differ between scenarios (e.g. higher CO_2 and lower CH_4 under SSP5-8.5 compared to RCP8.5, versus comparable emissions of both under SSP1-2.6 and RCP2.6). Methods to aggregate these effects rely on singular measures of ERF for each forcing agent (e.g. Meinshausen et al., 2020), though it has been shown that these differ by model (Zelinka et al., 2014, Smith et al., 2020, Zelinka et al., 2020).

Considering aerosol forcing, Zelinka et al., 2014 estimated ERF due to year-2000 aerosol emissions compared to pre-industrial in the CMIP5 ensemble, while Smith et al., 2020 performed the equivalent analysis for the CMIP6 ensemble (albeit with more models and 2014-level emissions). Both studies calculated ERF in terms of shortwave and longwave aerosol-radiation interactions and aerosol-cloud interactions. Comparing the two, net aerosol ERF is less negative in the CMIP6 ensemble: -1.01 ± 0.23 Wm^{-2} versus -1.17 ± 0.30 Wm^{-2} ($\pm$ 1σ). This is due primarily to less negative shortwave aerosol-cloud interactions, in line with Zelinka et al., 2020. While a less negative aerosol ERF could contribute to the greater warming we identified in the CMIP6 ensemble, these studies are not directly comparable. 2014-prescribed black carbon and sulfur emissions were 25% higher and 2.5% lower, respectively, than the 2000-prescribed emissions (Moss et al., 2010, Riahi et al., 2017, Gidden et al., 2019). New experiments that directly compare ERF for CMIP5 and CMIP6 models would assist in diagnosing the drivers of model inter-generational differences.

The increase in annual average temperature signal from CMIP5 to CMIP6 shown in Figure 2 and Table 4 is also consistent with findings from Zelinka et al., 2020. There, authors applied a radiative kernel technique to calculate ECS and ERF due to a doubling of CO_2 in the CMIP5 and CMIP6 ensembles and decomposed the feedbacks to diagnose the factors influencing the changes between generations. They found that ECS increased in both mean and variance, while ERF increased slightly in mean and decreased slightly in variance. The increase in ECS was due primarily to stronger positive feedbacks in extra-tropical low clouds. Based on this alone, we should expect a warmer globe in the CMIP6 ensemble at the same concentration of CO_2, which aligns with our results.

3.3 Population exposure

Figure 4 integrates S/N across the globe but adds the dimension of time to show when different proportions of the globe's land area (left column) and population (middle column) cross different S/N thresholds. Also shown is global population over time under each of the five scenarios. Under a moderate emissions scenario (SSP2-4.5), most models agree, nearly half of the population (48%) will be experiencing "unknown" (S/N>3) annual mean temperatures by 2050, with more than 90% of people over the "unfamiliar" threshold of S/N>2. The fraction of the population exposed to an unknown climate at 2050 varies from 10% under SSP1-1.9 to 87% under SSP5-8.5, again emphasising the significant influence of emissions pathway on the

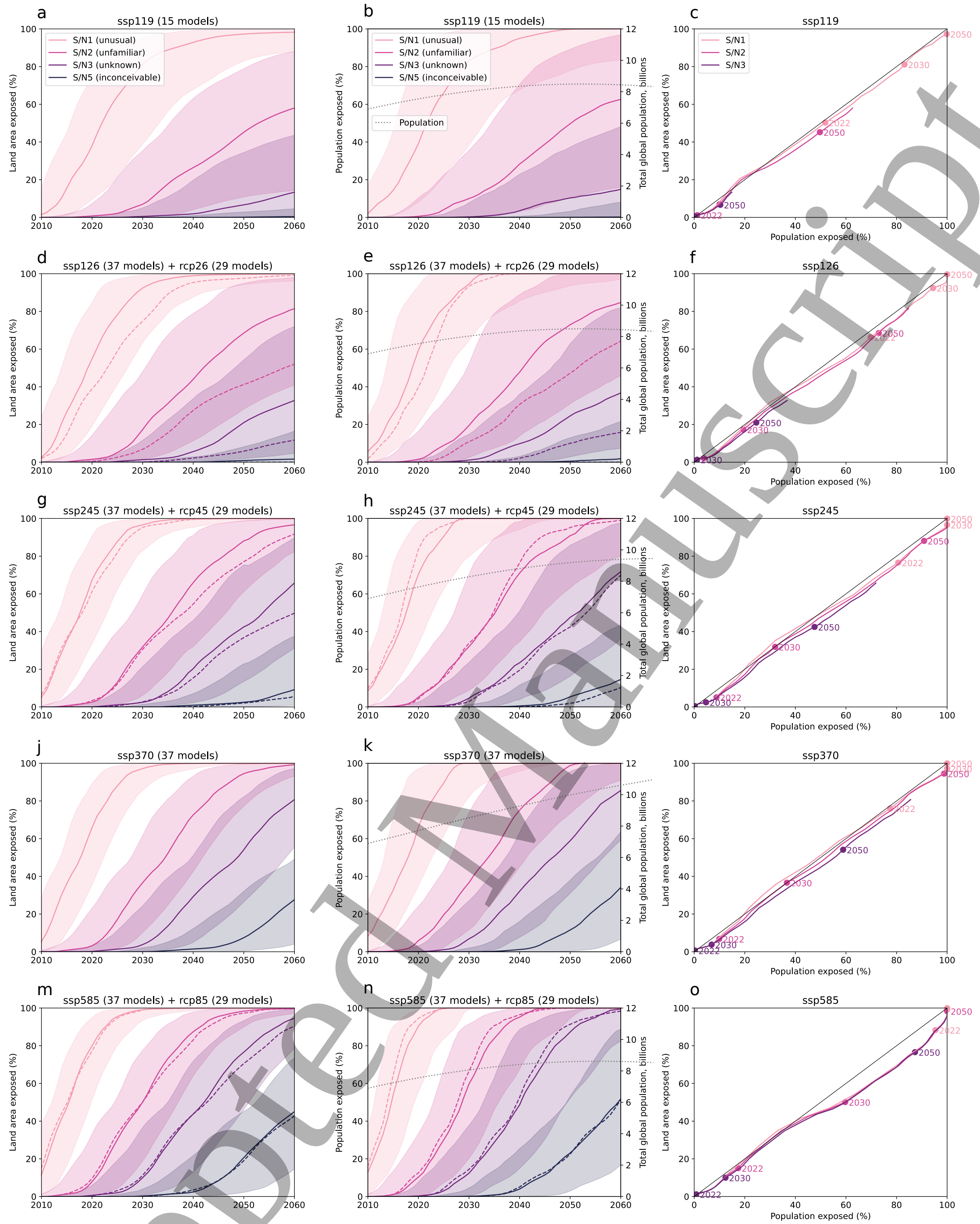

Figure 4: Crossing of average annual temperature S/N thresholds of 1, 2, 3, and 5 over time, shown as a proportion of Earth's land area (left column) and population (middle column). Median results are shown as solid lines, and the ranges between the 16$^{\text{th}}$ and 84$^{\text{th}}$ percentile results are shown as the shaded regions. Median results for corresponding CMIP5 RCPs are shown as dashed lines, and the global population over time is shown as a dotted grey line. The right column compares S/N emergence by area and population by plotting the median results against each other. Results for 2022, 2030, and 2050 are shown as dots, and the 1:1 line is shown in black.

projected mid-century climate.

As the historical data under CMIP6 extend only to the end of 2014, the values shown for 2022 are projections. The multi-model median projections are that 52-95% of the global population is currently experiencing an unusual climate (S/N>1) as of 2022, depending on the scenario. Some (1.2-17%) are already experiencing an unfamiliar climate (S/N>2) by this scale. This is compared to a recent baseline of 1986-2005, which emphasises the rapidity of change that we're experiencing. (Using an earlier baseline of 1961-1990 gives higher S/N values, with 87-99% of the population already experiencing S/N>1 and 11-22% with S/N>2, not shown.) A version of Figure 4 with the higher emissions scenarios extended to 2100 is included as Supplementary Figure S11. The results shown in Figure 4 are not very sensitive to the GMST smoothing technique until mid-century. However, when emissions turn net-negative in the lower-emissions scenarios, the choice of technique does affect exposure calculations.

Comparing scenarios, it appears that the high emissions scenarios have lower associated uncertainty. This is in part an artefact of the threshold selection; under a higher emissions scenario, these thresholds are passed earlier in time, and all models agree that the lower thresholds are passed before 2050. The lower emissions scenarios, in contrast, pass the thresholds later and have S/N peaks that are close to the threshold values. A similar absolute spread between scenarios at 2050 thus appears as a greater uncertainty for the low emissions scenarios. This illustrates an important point: we can have more confidence that a high emissions future will lead to "inconceivable" climates than that low emissions future will prevent "unknown" ones. This highlights the importance of investing concurrently in both mitigation and adaptation.

The right column of Figure 4 compares S/N emergence by area and population. Most scenarios show that temperatures will rise for the global population faster than for overall land area, shown by S/N threshold exceedances falling mostly below the 1:1 line. That is, average annual temperatures will change faster in areas where people live than where they don't, in agreement with Frame et al., 2017. Increased temperature change where people live compared to the global average is more often explored in terms of the land-sea contrast (e.g. Joshi et al., 2013), so it is noteworthy that this holds when only comparing to overall land area. This trend is most pronounced for SSP5-8.5 and least pronounced for SSP3-7.0, with the difference mainly due to population projections; under SSP3-7.0, the Global Emerging Markets (GEM) grouping of countries exhibits continued growth through the century, unlike in the other scenarios. The Least Developed Countries (LDC) grouping is the other major driver of global population; its population grows under all scenarios.

Comparing the results to the corresponding RCPs from CMIP5, Figure 4 shows that SSP1-2.6 and SSP2-4.5 project more rapid emergence of unusual to unknown climatic conditions by land area than do the previous generation of RCPs, while SSP5-8.5 is broadly comparable. Both RCP4.5 and RCP8.5 show more rapid emergence by population than their corresponding SSPs for lower thresholds. This is due to the different spatial pattern of emergence; these two RCPs project stronger and more rapid emergence of unknown annual temperatures in the heavily populated regions of South Asia, West/Central Africa, and parts of Western Europe, as shown in Figure 2.

Figure 5 shows S/N values and the proportion of global population exceeding them as of M21C, broken into the five groupings from Table 3. Using S/N in annual average temperatures as a proxy for climate change impacts, this figure illustrates the disparity in impacts between dif-

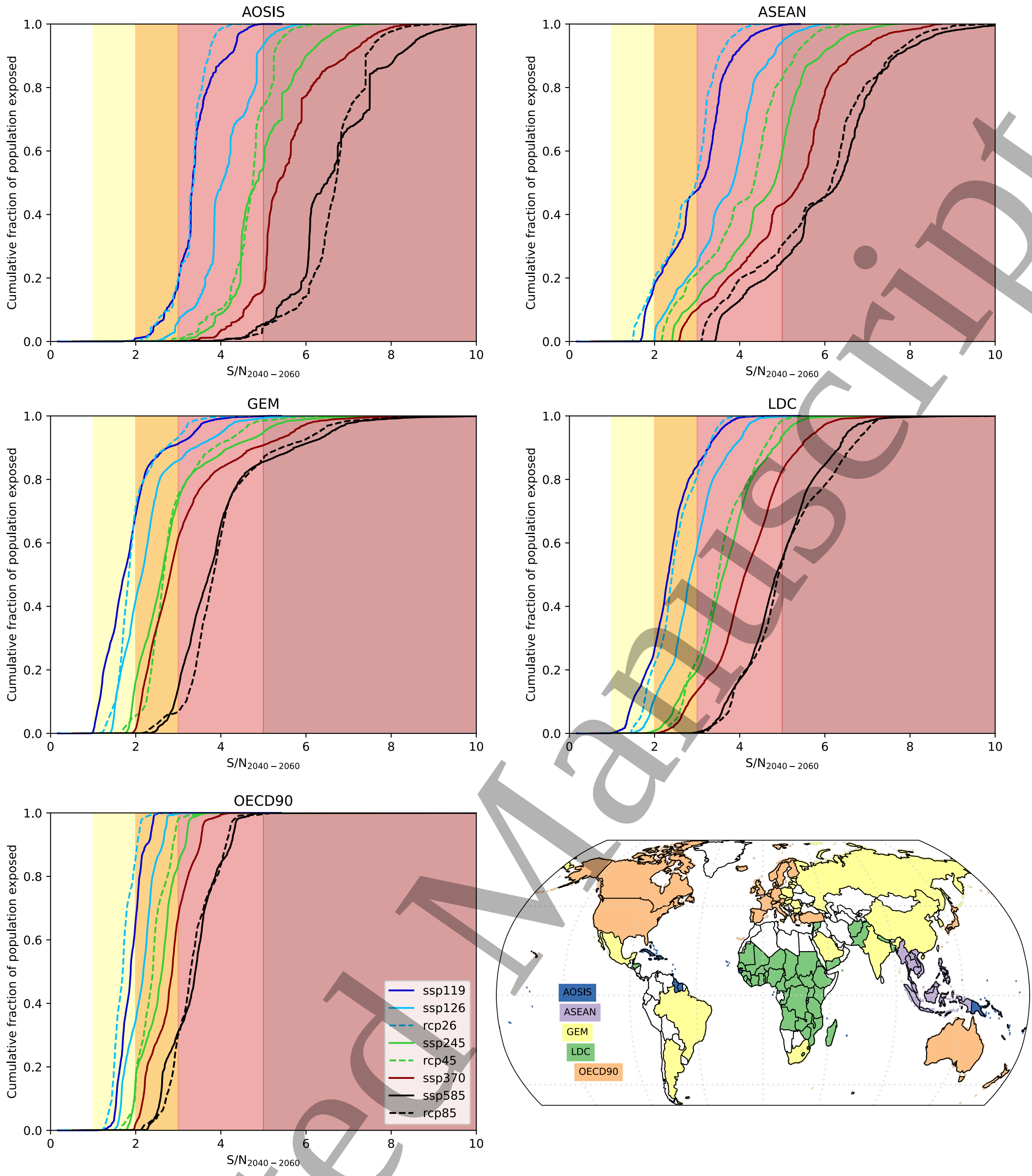

Figure 5: The proportion of population exposed to various levels of 2040-2060 average S/N for annual average temperature. Results are shown across five groupings of countries (defined in Table 3: ASEAN, AOSIS, GEM, LDC, and OECD90. Results for each of five SSPs are distinguished by colour, and results for corresponding CMIP5 RCPs are shown as dashed lines. Unusual, unfamiliar, unknown, and inconceivable annual temperatures are indicated by the background shading. The geographic distribution of the groupings is shown bottom right, with countries in more than one grouping coloured by the first alphabetically.

ferent socioeconomic and geographic groups. The position of each curve along the x-axis indicates the degree of impacts experienced, and the slope of the curve is a measure of the uniformity of impacts across the group. From this we can see that the Organisation for Economic Co-operation and Development (1990) (OECD90) grouping has both the lowest impacts and the most equally distributed impacts across its population.

Considering just the two low-emissions scenarios, SSP1-1.9 and SSP1-2.6 (which are consistent with temperature rises of 1.5 K and 2.0 K, respectively) gives an indication of the difference in impacts between two aspirational warming levels. The impacts for the OECD90 grouping at 2.0 K are comparable to those for the Least Developed Countries (LDC) grouping at 1.5 K, and lower than those for the Association of Southeast Asian Nations (ASEAN) and Alliance of Small Island States (AOSIS) groupings at 1.5 K (see Supplementary Figure S12).

These findings agree with earlier studies in finding that climate change impacts are expected to be unequally distributed, with less developed countries and those with higher projected population growth rates experiencing greater changes than developed ones, on average (Frame et al., 2017, Harrington et al., 2017, Harrington and Otto, 2018, King and Harrington, 2018, Frame et al., 2019). The grouping with the most unusual M21C climate is AOSIS, closely followed by ASEAN. Both groupings are characterised by relatively small land masses, proximity to the equator, and having many states located in the maritime continent, with climates dominated by the surrounded ocean. The ocean's thermal inertia contributes to this region having generally lower noise (and so higher S/N) than, say equatorial Africa or South America (see Figure S1). The small island nations are also particularly vulnerable to rising sea levels (Hooijer and Vernimmen, 2021). The risk of compounding impacts is thus particularly acute for these states.

4 Conclusions

We analysed the emergence of unknown annual average temperatures due to climate change projected by the SSPs of CMIP6. The results showed expected patterns of stronger and earlier emergence under higher emissions scenarios, with the emergence pattern strongest in the tropics. All scenarios project that a significant proportion of the world's population was already experiencing "unusual" (S/N>1) annual temperatures as of 2022, and most models agree that around half of the globe will be experiencing "unknown" annual temperatures (S/N>3) by 2050 under the moderate emissions scenario of SSP2-4.5. Inter-model uncertainty suggests we can have more confidence that a high emissions future will lead to an "unknown" climate by mid-century than that a low emissions future will prevent this.

In general, CMIP6 shows earlier and stronger emergence of anomalous annual mean temperatures (higher S/N ratios) than the corresponding scenarios from CMIP5, though there are notable decreases at the regional level. CMIP6 models exhibit lower S/N in Central Africa and South Asia under all scenarios, and the higher emissions scenarios also show lower S/N over parts of South America, West Africa, Western Europe, and East Asia. These regional decreases in densely populated areas mean than population-based emergence is actually slightly weaker and later under SSP5-8.5 than it was under RCP8.5. Noise increases in most areas, accompanying increases in the signal. Differences in S/N between generations arise from changes in both model responses and applied forcing, with the newer models using emissions pathways of the SSPs and the older using RCPs. To separate the effect of the higher mean and range of ECS in CMIP6 models, we repeated the analysis for a subset of models with ECS in the same range as that of CMIP5. We found that the increase in temperature is not due solely to increased model sensitivity. Other factors that explain some of the observed differences include changes to aerosol optical depths (particularly for Central Africa and South Asia), different

GHG emissions, changes in the ERF of models to radiative forcing agents, and large-scale climate responses such as Southern Ocean cloud behaviour and weakening of the AMOC. None of these causes alone accounts for all the observed differences, and quantifying their relative influence is the task for targetted experiments, coordinated across modelling groups.

We also incorporated nation-scale, dynamic population datasets aligned with emissions pathways to assess exposure to climate change. We found that unusual annual temperatures emerge earlier in areas where people live than where they don't, and that the nations least equipped to adapt to climate change will be disproportionately affected, regardless of the emissions pathway taken. That this conclusion holds despite more granular projections demonstrates that earlier findings were likely not a result of oversimplification or overly broad assumptions about future population distributions.

Acknowledgements

We acknowledge funding from the New Zealand Ministry for Business, Innovation & Employment's Endeavour Fund Whakahura programme (Grant ID: RTVU1906). EH is supported by the UK National Centre for Atmospheric Science and by the NERC REAL PROJECTIONS and EMERGENCE projects. We also acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which coordinated and promoted CMIP6. We further thank the climate modelling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the funding agencies that support CMIP6 and ESGF.

Data Availability

Code for the analysis presented in this paper is available at https://github.com/hdouglas/CMIP6emergence. CMIP6 simulation results: https://esgf-node.llnl.gov/projects/cmip6/. CMIP5 simulation results: https://esgf-node.llnl.gov/projects/cmip5/. Emissions data: https://esgf-node.llnl.gov/search/input4mips/, IIASA RCP Database, Version 2.05: https://tntcat.iiasa.ac.at/RcpDb/dsd, IIASA SSP Database, Version 2.0: https://secure.iiasa.ac.at/web-apps/ene/SspDb/. Global one-eighth degree gridded population dataset, v1.01: https://sedac.ciesin.columbia.edu/data/set/popdynamics-1-8th-pop-base-year-projection-ssp-2000-2100-rev01. Geopolitical boundary shapefiles: https://www.naturalearthdata.com/downloads/. Human Development Index data: http://hdr.undp.org/en/content/download-data. Region masking algorithm: https://github.com/regionmask/regionmask. Regridding algorithm: https://github.com/JiaweiZhuang/xESMF (Zhuang, 2020).