Clim Dyn (2017) 48:3035–3045
DOI 10.1007/s00382-016-3249-6

The influence of model resolution on temperature variability

Jeremy M. Klavans[1,2] · Andrew Poppick[3] · Shanshan Sun[4] · Elisabeth J. Moyer[4]

Received: 2 December 2015 / Accepted: 3 June 2016 / Published online: 5 August 2016

Abstract Understanding future changes in climate variability, which can impact human activities, is a current research priority. It is often assumed that a key part of this effort involves improving the spatial resolution of climate models; however, few previous studies comprehensively evaluate the effects of model resolution on variability. In this study, we systematically examine the sensitivity of temperature variability to horizontal atmospheric resolution in a single model (CCSM3, the Community Climate System Model 3) at three different resolutions (T85, T42, and T31), using spectral analysis to describe the frequency dependence of differences. We find that in these runs, increased model resolution is associated with reduced temperature variability at all but the highest frequencies (2–5 day periods), though with strong regional differences. (In the tropics, where temperature fluctuations are smallest, increased resolution is associated with increased variability.) At all resolutions, temperature fluctuations in CCSM3 are highly spatially correlated, implying that the changes in variability with model resolution are driven by alterations in large-scale phenomena. Because CCSM3 generally overestimates temperature variability relative to reanalysis output, the reductions in variability associated with increased resolution tend to improve model fidelity. However, the resolution-related variability differences are relatively uniform with frequency, whereas the sign of model bias changes at interannual frequencies. This discrepancy raises questions about the mechanisms underlying the improvement at subannual frequencies. The consistent response across frequencies also implies that the atmosphere plays a significant role in interannual variability.

Keywords Climate variability · Variability · Model resolution · CCSM3 · Spectral analysis

1 Introduction

Human activities are influenced not only by mean climate, but also by variability around climatic means, which can affect food production, human health, water supply, and even mortality (Thornton et al. 2014; IPCC 2014, and references therein). Understanding potential variability changes as the planet warms in response to elevated greenhouse gas concentrations is therefore a current research priority. Because the observational record under changing climatic conditions is short, general circulation models (GCMs) are a principal tool for this purpose (IPCC 2012, 2013). However, it is widely recognized that existing GCMs do not perfectly reproduce observed variability, and therefore may not perfectly simulate future variability (e.g. IPCC 2013; Leeds et al. 2015). Improving the representation of both current and future variability is therefore an impetus for continued model development (e.g. Reichler and Kim 2008; Delworth et al. 2012).

Electronic supplementary material The online version of this article (doi:10.1007/s00382-016-3249-6) contains supplementary material, which is available to authorized users.

✉ Elisabeth J. Moyer
moyer@uchicago.edu

[1] Center for Robust Decision-making on Climate and Energy Policy, University of Chicago, Chicago, IL, USA

[2] Present Address: Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, FL, USA

[3] Department of Statistics, University of Chicago, Chicago, IL, USA

[4] Department of the Geophysical Sciences, University of Chicago, Chicago, IL, USA

Increasing model resolution is often assumed to be a key factor in improving the simulation of climate variability (e.g. Bacmeister et al. 2013; Delworth et al. 2012; Hansen et al. 2012; Yeager et al. 2006). If the model physics is reasonably realistic, altering model resolution could be expected to affect variability primarily through either of two mechanisms, which we term "spatial averaging effects" and "non-local effects". In the first mechanism, increasing resolution allows the model to resolve topographic features or other spatially heterogeneous processes, capturing spatially uncorrelated variations that coarser models would artificially suppress. Spatial averaging effects under this definition can only increase variability as resolution increases. In the second mechanism, model resolution may alter the properties of variations that are broad-scale and spatially correlated. For example, improving the representation of topography may also alter large-scale meteorological patterns and thereby affect variability across many grid cells. Variability may be affected through either or both mechanisms whenever increasing model resolution allows explicit treatment of processes that were formerly parameterized (e.g. clouds, ocean eddies, convection, sea ice formation) (Delworth et al. 2012; DeWeaver and Bitz 2006).

Few studies have systematically examined the relationship between model resolution and variability. Most studies of the impact of increasing model resolution have focused on climate means. Both global and regional studies have suggested that increased resolution improves seasonal mean temperature and precipitation in a variety of models, e.g. GFDL (Delworth et al. 2012), IPSL (Marti et al. 2010), ECHAM (Roeckner et al. 2006) and CCSM3 (Gent et al. 2010), but Hack et al. (2006) find little effect on global mean temperature (GMT) in CCSM3 and Kirtman et al. (2012) find an increase in bias in global mean sea surface temperature in CCSM3.5 (changing ocean resolution only). The sign of the effect of resolution differs across models, with increases in global mean temperature with resolution in CCSM3.5 and GFDL (Kirtman et al. 2012; Delworth et al. 2012) but a statistically insignificant decrease in CCSM3 (Hack et al. 2006). If increased resolution appears to improve the representation of individual regions, those improvements may still not be uniform across all regions or seasons. For example, Iorio et al. (2004) find that in the atmosphere-only CCM3 model, higher resolution improves seasonal mean precipitation over the United States in local autumn and winter, but degrades it in spring and summer, when precipitation is predominantly convective.

In the subset of research that evaluates how model resolution affects climate variability, most studies focus on the large-scale effects associated with an individual meteorological phenomenon: for example, monsoons, jet stream location, cyclones, the Madden-Julian Oscillation (MJO), and the El Nino-Southern Oscillation (ENSO). Results are somewhat inconsistent, but sources of variability at a variety of frequencies appear sensitive to horizontal resolution. Multiple studies find that increasing resolution yields more realistic simulations of both the North American and Asian monsoon systems (specifically in regard to the magnitude and distribution of precipitation) (Collier and Zhang 2007; Kobayashi and Sugi 2004; Gent et al. 2010; Delworth et al. 2012). Several studies find that the location of the jet stream (Guemas and Codron 2011) and the number of simulated cyclones (Kobayashi and Sugi 2004) change monotonically with increasing resolution, but neither change is classified as an improvement. The simulation of MJO is found either to not improve (Hack et al. 2006) or to actually degrade with resolution (Gualdi et al. 1997). Climate models have long-standing difficulties capturing ENSO periodicity (Deser et al. 2006; Navarra et al. 2008); increasing resolution appears to improve both the periodicity and amplitude of ENSO (e.g. Neale et al. 2008; Guilyardi et al. 2004), as well as the "extratropical response" to the oscillation (Deser et al. 2006).

The statistics of climate variability is an active area of research at present because of concerns that future conditions of higher CO_2 may involve shifts in variability. Observational analyses have sought to detect shifts in temperature variability by estimating either changes in marginal distributions (e.g. Hansen et al. 2012; Huntingford et al. 2013) or changes in the frequency of exceeding thresholds of extreme conditions (e.g. Alexander et al. 2006; Morak et al. 2013) (a metric that can convolve changes in variability and means). Analyses of model projections also commonly focus on threshold exceedance (e.g. Meehl et al. 2000; Wehner et al. 2010), though some recent articles have evaluated changes in the variance of temperature time series (e.g. Holmes et al. 2015; Schneider et al. 2014). Both marginal distribution and threshold exceedance approaches often convolve multiple timescales of variability or study only narrow frequency ranges.

In this work we seek to systematically examine how resolution can influence the broad statistics of temperature variability: that is, how global variability can change with resolution at all timescales. We use model output from a single model, CCSM3, at different resolutions, and evaluate their differences in global temperature variability at all frequencies, from days to decades, and in all regions, from single pixels to global averages. Some frequency-resolved variability studies do exist in the literature, most commonly in the context of interannual variability; often, studies use power spectra to compare model reconstructions to paleoclimate proxies or long-term observations (e.g. IPCC 2013; Jones et al. 2013; Laepple and Huybers 2014a, b). Recently, spectral methods have been used to compare variability in model simulations of current and future climates (Leeds et al. 2015). In this work, we apply a similar

approach to compare model output at different resolutions. Our goals are to (1) identify how model resolution may affect variability through either/both improved representation of small-scale heterogeneity and/or altered large-scale physics, and (2) determine which frequencies and regions may be most sensitive to resolution.

2 Data and methods

2.1 Model output and data

2.1.1 CCSM3

This study takes advantage of publicly available output from the Community Climate System Model 3 (CCSM3), a fully-coupled Atmosphere-Ocean General Circulation Model (AOGCM) developed by researchers at the National Center for Atmospheric Research (NCAR) (Collins et al. 2006, and references therein). CCSM3 links four separate components: the Community Atmosphere Model version 3 (CAM3), the Parallel Ocean Program (POP), the Community Land Surface Model version 3 (CLM3) and the Community Sea Ice Model version 5 (CSIM5). We use CCSM3 rather than a more recent model version because of the availability of daily temperature output from runs with identical forcing at multiple horizontal resolutions of the same model. (Note however that some parameters must be adjusted in a model version when grid size is altered.) These runs were specifically designed for use in comparative studies of resolution sensitivity (Hack et al. 2006; Yeager et al. 2006).

We use one 100-year segment from each of three ~800 year-long pre-industrial (PI) control simulations at T31, T42 and T85 atmospheric resolution: 3.75°, 2.8°, and 1.4°, respectively. The model grids are aligned: each T42 grid cell covers 2 × 2 T85 cells, and 3 × 3 T31 cells are 4 × 4 T42 or 8 × 8 T85 cells. The T85 and T42 resolution atmospheric models are coupled to a one-degree ocean model; the T31 atmosphere is coupled to a three-degree ocean. (For this reason, we primarily compare T85 and T42.) The CLM3 resolution follows that of the atmospheric component.

We analyze segments near the end of each run to ensure approximate stationarity, i.e. fully equilibrated conditions. (The fitted trend in GMT is $< \pm 0.03\,^{\circ}$K/century in all three run segments used; the trends in mean ocean temperature are 0.01, −0.04, and $-0.05\,^{\circ}$K/century for the T31, T42, and T85 atmospheric resolutions, respectively (Yeager et al. 2006).) As mentioned above, GMT does not change significantly across resolutions (Hack et al. 2006;

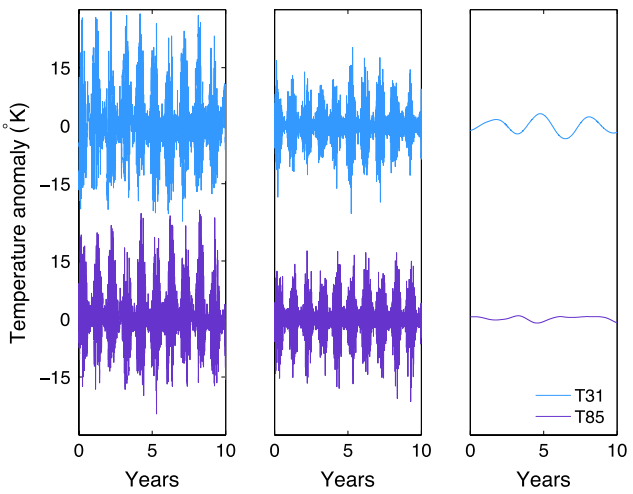

Fig. 1 Example of decomposition of variability into different frequency bands, using a 10-year time series of deseasonalized CCSM3 temperature output from a single grid cell (Northwest Territories, Canada) for T85 (*purple*) and T31 (*blue*) resolutions. *Left* Raw time series of daily mean reference temperature. Although the mean seasonal cycle has been removed, seasonal variation in variability remains evident, with highest values in winter. *Center* the high-frequency (period < than five days) component of the same time series. In this example, variability in this frequency band is larger at T85 resolution than at T31 (by ~8 %). *Right* the low-frequency (inter-annual) component of the same time series. Variability in this frequency band is smaller at T85 resolution than at T31. (Over 100 years of data, the difference is 10 %)

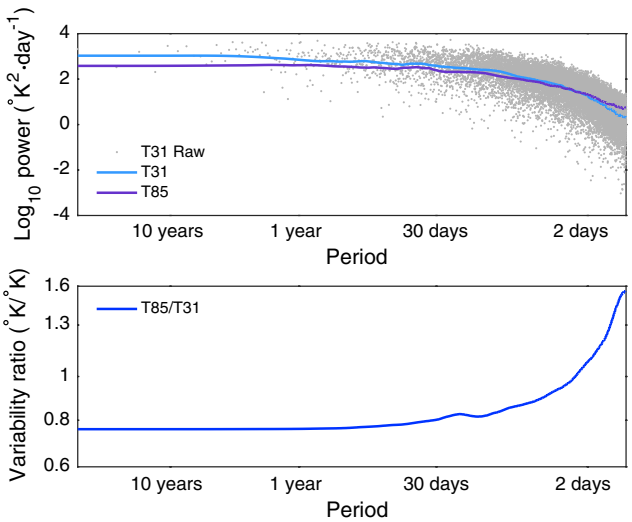

Fig. 2 Example of representation of variability in frequency space, using the daily mean temperature time series shown in Fig. 1 (only now using 100 years of model output). *Upper* raw (*gray points*) and smoothed (*light blue*) power spectra from the T31 time series, and for comparison, the smoothed power spectrum from the T85 time series (*purple*). Note the greater number of estimates at the high frequencies. Because the raw estimate is noisy, smoothing is useful for comparing estimates of spectral density. *Lower* difference in variability between resolutions shown as the ratio of spectral densities (the smoothed ratio of the T85/T31 power spectra). Variability increases with resolution only at the highest frequencies

Otto-Bliesner et al. 2006). All temperature output reported is at reference (2-m) atmospheric height unless otherwise specified. For more information about the model runs, see Hack et al. (2006), supplemental online material or description online at http://www.cesm.ucar.edu/experiments/ccsm3.0/. Because we consider the seasonal cycle to be a component of mean climate, we pre-process all runs by removing 12 seasonal harmonics.

2.1.2 NCEP/DOE reanalysis

We compare model output to a gridded observational product, the NCEP/DOE AMIP II Reanalysis (Kanamitsu et al. 2002), which covers the time period 1979–2014 (limited by the availability of satellite data). In this dataset, model physics and algorithms were frozen at the time of publication (2002) and data up to the present day assimilated into the existing system (Saha et al. 2010). The product is gridded at ~1.8°.

2.2 Methods and definitions

2.2.1 Defining variability

To analyze climate variability we decompose the time series of daily temperature at each model grid point into its frequency components (see Fig. 1). We apply a discrete Fourier transform to each deseasonalized time series to obtain an estimate of the raw power spectrum with units of $°K^2\ day^{-1}$ (variance per frequency interval). In this paper, we present variability either as a power spectrum or as "integrated variability" over a specific range of frequencies. We generally report integrated variability in units of °K, i.e. as the standard deviation of temperature associated with the given range of frequencies. We use units of $°K^2$ when discussing the contribution of different components to an overall observed effect, because the variances in units of $°K^2$ sum to the total variance.

In computing an estimate of a power spectrum, we smooth the log of a raw power spectrum with a variable-bandwidth kernel of the form described in Poppick et al. (2016), supplemental material. (See Fig. 2, top, for comparison of smoothed and unsmoothed spectra). We reduce the bandwidth of the kernel estimator at longer periods (where the Fourier frequencies, at integral multiples of $1/n\ day^{-1}$, are more sparse) in order to reduce bias across interannual frequencies and to avoid contamination by the higher-frequency parts of the spectrum. This reduced smoothing is justified because these spectra do typically show more variation at the lower frequencies. We do not apply a smoothing function when computing integrated variability, which involves integrating the power spectrum over a broad range of frequencies.

2.2.2 Comparing variability

We compare variability between resolutions using ratios, either of corresponding power spectra or of corresponding estimates of integrated variability (Fig. 2, bottom). In the former case, we smooth the log ratio of spectra rather than the individual spectra, using the same variable-bandwidth kernel described above. In the latter case, we simply calculate integrated variability for each frequency bin and take a ratio. Although most figures in this work present variability ratios in units of standard deviations (°K/°K), all are originally calculated in units of variances ($°K^2/°K^2$).

It is important to note that our estimated differences in variability between model resolutions will be dominated by wintertime effects, because outside of the tropics, temperature variability is generally stronger in winter than in summer. (See Fig. 1.) Overall differences in variability between resolutions are therefore implicitly weighted towards wintertime differences. It would be possible, with minor extensions of the procedure, to remove this implicit weighting by not only removing the mean seasonal cycle but to also demodulating the seasonal changes in variability. In this case the final estimate of variability difference would reflect the mean fractional difference normalized for seasonal variability changes (see Leeds et al. 2015).

2.2.3 Comparing resolutions

Comparing model output across multiple resolutions requires that we choose appropriate pixel comparison strategies, given that, by definition, pixels do not match. This paper utilizes two different pixel comparison strategies, each of which provides unique information about differences in variability but incurs some tradeoffs.

When "pixel matching", each model pixel is compared to a single nearby pixel at a different resolution, regardless of the difference in pixel size. This approach may be most appropriate for impacts assessments, as it utilizes the raw model output and will include variability attributable both to spatial averaging and to non-local effects. However, if each high-resolution pixel is compared to one at a lower resolution, then the resulting analysis requires repeating information. (For example, four different T85 pixels would all be compared to the same T42 pixel.) Conversely, if each low-resolution pixel is to be compared to higher-resolution output, one could either choose a single high-resolution pixel for comparison and discard all unmatched pixels, or compare to the mean variability of all potential matching pixels. Unless otherwise noted, figures presented in this work are obtained via pixel matching to a single low resolution pixel, so as to preserve all of the possible effects of variability on resolution. (That is, we do not average over any pixel-length spatial heterogeneity.)

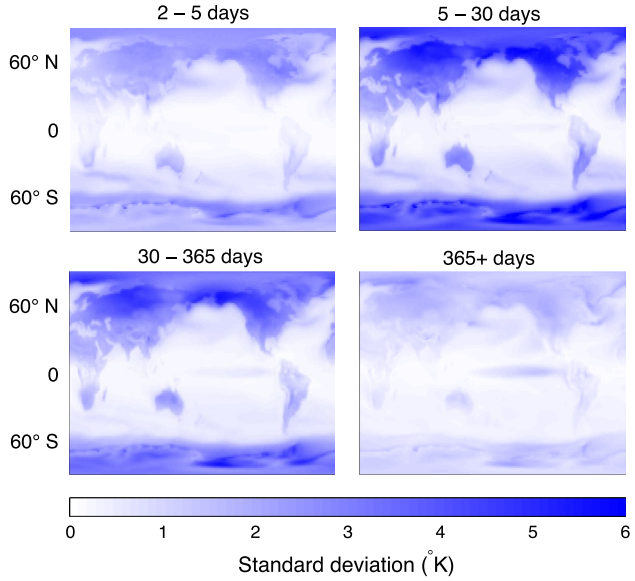

Fig. 3 Maps of integrated variability (°K) in four frequency bands for CCSM3 at T85 resolution. In general, variability is higher over land than ocean and over high than low latitudes (as seen in both observations and models). Frequency bins account for different proportions of global variance (°K^2): 2–5 day: 14 %, 5–30 day: 49 %, 30–365 day: 30 % and 365+ days: 6 %. The value reported at each pixel is the square root of the integrated power spectrum for the given range of frequencies

The alternative approach of "pixel averaging" isolates non-local effects by eliminating, or averaging over, spatial heterogeneities across the length scale of model pixel. In this strategy, each low-resolution pixel is compared to the analogous output the high-resolution model would produce over the same area, i.e. to the mean temperature timeseries over the equivalent area. We indirectly use this approach when considering the spatial characteristics of contributions to variability.

Our conclusions are generally robust to the choice of pixel comparison strategies. See Figs. S5–S6, which repeat Fig. 4 for all comparison strategies, and Table S2, which quantifies variability for each comparison strategy.

3 Results

3.1 Absolute variability

Integrated variability follows expected patterns in all three resolutions of CCSM3 and broadly resembles observations (Fig. 3). Variability is generally larger over land than ocean and larger at high latitudes (>66°) than at low latitudes (<23°). These patterns are seen in both observations and models (Leeds et al. 2015; Holmes et al. 2015). Figure 3 shows variability in T85 output in the four frequency bands: 2–5, 5–30, 30–365, and > 365 days. Patterns are largely similar in all frequency bands, though at interannual frequencies land/ocean contrast is reduced and ENSO variability appears. We selected these particular frequency bands to highlight differences between resolutions. The bands do not divide variability equally, but comprise 14, 50, 30, and 6 % of total variance, respectively (°K^2). We display results in these bands throughout this paper.

Table 1 Global average integrated variability (°K) in the four frequency bins from CCSM3 at three resolutions (T31, T42 and T85) and NCEP–DOE reanalysis ("Obs.")

Res.	Decomposed variability (°K)				
	2–5 days	5–30 days	30–365 days	365+ days	Total
T31	1.03	2.24	1.78	0.78	3.14
T42	1.11	2.24	1.75	0.81	3.15
T85	1.12	2.07	1.63	0.75	2.96
Obs.	1.07	1.97	1.57	0.85	2.86

As in the example of Figs. 1 and 2, variability increases with model resolution only at the highest frequencies (2–5 days). Note that the sum of the decomposed integrated variabilities (°K) does not equal the rightmost column; their squares sum to the total global variance (°K^2). (For this information disaggregated by region, see Table S1)

3.2 Relative variability

Despite their broad similarities, model runs at different resolutions do show differences in variability, and those differences are frequency-dependent (Table 1). While spatial averaging effects would cause variability to increase with resolution, overall variability in CCSM3 decreases with resolution. In the global average, variability increases with resolution only at the highest frequencies (2–5 days), with all other frequencies showing decreases. Changes in global average variability are generally monotonic but not linear with resolution. At high frequencies, changes in variance (°K^2) are stronger from T31 to T42 (14 %) than from T42 to T85 (2 %); in other frequency bands, changes in variance are stronger from T42 to T85. Those differences may be informative of the mechanisms driving changes.

Differences in variability with resolution also exhibit distinct regional patterns, generally with strong latitudinal dependence (Fig. 4). At all frequencies, variability generally increases with resolution in the tropics and decreases with resolution at mid-latitudes. (There are some local deviations: we reproduce the decrease in variability in the Niño 3.4 region, a proxy for ENSO amplitude, found by Deser et al. (2006) in this same dataset, but note that because the timeseries is only 100 years long, signal-to-noise at interannual frequencies is relatively low.) Frequency dependence is most apparent at high latitudes, where variability increases with resolution at the highest frequencies (2–5

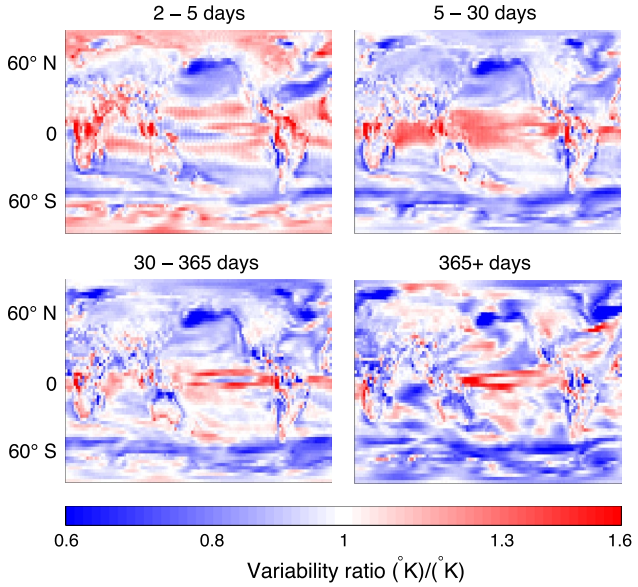

Fig. 4 Differences in integrated temperature variability at different CCSM3 resolutions, shown as the ratio (°K/°K) of T85:T42, for four frequency bins. Output at the two resolutions are compared by pixel matching to the lower resolution, yielding a T42 figure. (See Supplementary Figs. S1–S2 for equivalent comparisons made with other resolution combinations; all yield qualitatively similar results.) High latitude land, where variability is strongest, dominates the global average of Table 1: variability here generally increases with resolution (*red*) at the highest frequencies and decreases (*blue*) at lower frequencies. In the Arctic, isolated patches of strong variability differences across all frequencies are due to discrepancies in sea ice extent between T85 and T42 resolutions

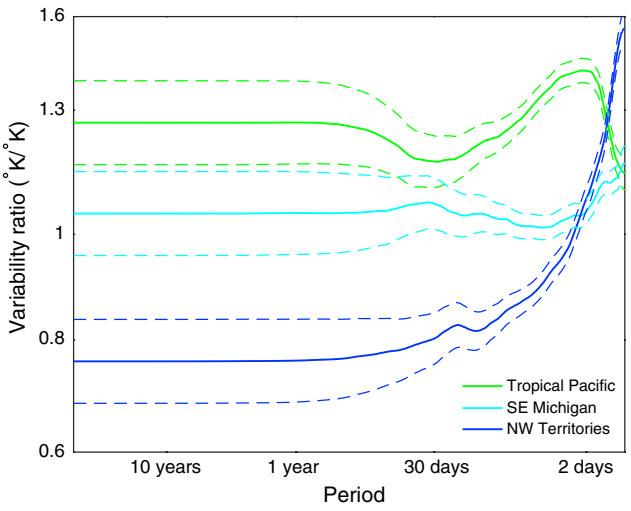

Fig. 5 The smoothed ratio (T85:T31) of power spectra for three example CCSM3 grid points (*green* Tropical Pacific, *light blue*: U.S. Midwest and *blue* Northwest Territories, Canada). *Dashed lines* of the same color represent two standard errors around the estimate of the ratio. At low freqeuncies, decreases in variability with resolution in the Northwest Territories and increases in the tropics are significant to two standard errors. At high frequencies, all three regions show significant increases. (These examples were chosen for the large effect they exhibit; see the Supplemental Material for significance testing across all regions and Figs. S3–S4 for iterations of this plot at other resolutions.)

day period), especially over land, but decreases at all lower frequencies. Because the high latitudes have the largest absolute variability, behavior in this region drives the global results of Table 1. Although relative changes in tropical variability are large, absolute variability is small there, limiting the tropical contribution to the global average.

We highlight the latitudinal differences in variability between model resolutions in Fig. 5, which shows the ratio of power spectra from three representative land pixels at low, medium, and high latitudes. (For maximum clarity, we compare T85:T31 in this figure.) To approximate the uncertainty in these values, we include bands for ± two standard errors, constructed based on the width of the smoothing kernel. The high- and low-latitude pixels show differences between T31 and T85 resolutions significant to two standard errors at all frequencies, with high latitudes pixels exhibiting the strongest frequency dependence: strong increases with resolution at periods of days, transitioning to decreases at periods of weeks or longer.

Applying this significance test to the T85:T42 comparison, the latitudinal banding of Fig. 4 is significant to two standard errors; across all pixels, at periods of two weeks, 83 % of high-latitude grid points and 68 % of mid-latitude grid points show significant variability decreases from T42 to T85, and 58 % of low-latitude grid points show significant variability increases. (See Tables S3–S5 for global and regional significance testing at additional frequencies.)

3.3 Spatial contributions to variability

Given the significant differences in variability between model resolutions, it is useful to consider what processes drive these changes. As mentioned above, resolving small-scale spatial heterogeneity can produce only increases in variability with resolution, but overall global mean variability in T85 is lower, not higher, than at coarser resolutions (∼3 % lower standard deviation or ∼5 % lower variance than in T42). (For this and the following section, we report values as variances with units of $°K^2$, the more natural units when considering components of a total.) This significant decrease cannot be explained solely through spatial averaging effects. Furthermore, although some regions show strong local increases in variability with resolution (Fig. 4), those changes could result from spatial averaging effects or from some other mechanism.

To help characterize the drivers of variability changes, we estimate the potential contribution of spatial averaging effects based on the spatial correlation of temperatures

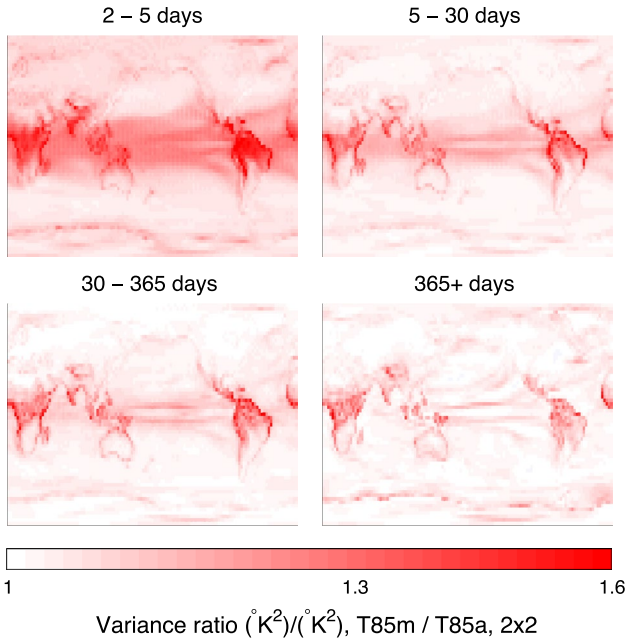

Fig. 6 *Spatial averaging test, variance ratio:* average temperature variance in four adjacent T85 grid cells divided by the variance of the average temperature in that 2×2 block. If all temperature fluctuations were spatially uncorrelated at this scale, the resulting ratio would be n pixels or 4; if fluctuations were strongly spatially correlated, the ratio would be close to 1. Actual variance ratios are ∼1.05 for non-tropical regions and ∼1.14 for the tropics (±15 degrees latitude); note that the tropics contribute negligibly to the global average ratio (∼1.07). These results suggest that temperature fluctuations in CCSM3 T85 exhibit strong spatial correlation at T42 spatial scale (∼300−200 km in tropics/midlatitudes)

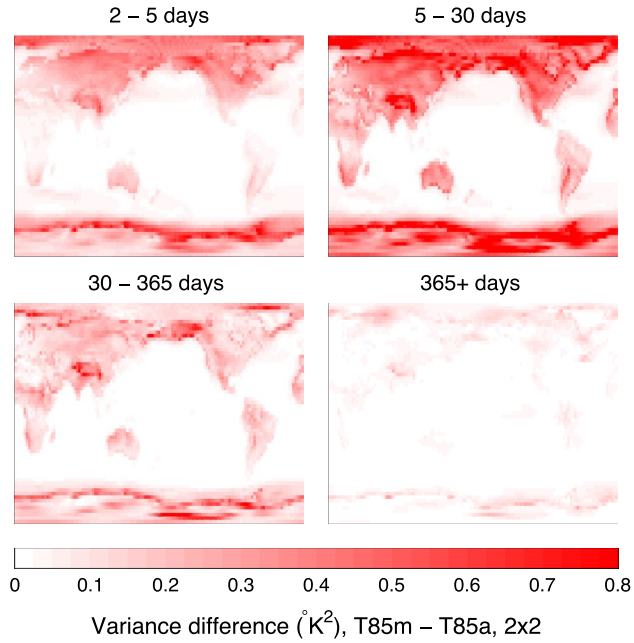

Fig. 7 *Spatial averaging test, variance difference:* Average variance in four adjacent T85 grid cells minus the variance of the average temperature in that 2×2 block. The resulting pattern is similar to that of absolute variability, with strong land/ocean contrast (compare to Fig. 3), suggesting that the contribution of spatially uncorrelated variability largely scales with a spatially correlated component. The larger tropical contribution of spatially uncorrelated variability seen in Fig. 6 could be explained by a small, global, additional additive "noise" contribution of $\sim 0.1\,°K^2$ variance

in a single model run. That is, we quantify the amount of variability that would be expected to be averaged away in a lower resolution model due purely to spatial averaging effects. For this test we degrade a T85 resolution run (pixel size ∼100−150 km in tropics/midlatitudes) to T42 resolution (∼200−300 km) by averaging temperatures in 2×2 adjacent pixels, and then compare the artificial low-resolution output to the original output. To avoid artifacts from topographic features, here we compare not to the variance of a single high-resolution pixel but to the average variance of the four original T85 pixels.

We find that averaging produces only small reductions in variance, implying that temperature fluctuations are strongly correlated in adjacent T85 grid cells (Fig. 6; see also Fig S7 to directly compare to Fig. 4). In the global average, variance is reduced by only about 7 %. Strong spatial correlation appears to hold even at larger length scales: T42 model output also shows only weak effects of spatial averaging, with a reduction in variance of 12 % (Fig. 8, left column 3rd panel). If each temperature timeseries could be decomposed into two components, one perfectly correlated across 2×2 grid cells and another perfectly uncorrelated, these results would imply that ∼ 91 % of total variance in T85 derives from fluctuations spatially correlated at length scales of >200−300 km, and 85 % of variance in T42 derives from fluctuations spatially correlated over scales >400−600 km.

The spatial properties of temperature fluctuations are not however globally uniform. The fractional effect of spatial averaging is small and relatively constant for over much of the globe, but becomes larger in the tropics, where total variance is lowest (Fig. 6). Spatial averaging of T85 output reduces variance by ∼5 % in the extratropics (>15° latitude) but 12 % in the tropics. The larger spatial averaging effect in the tropics would be consistent, for example, with an additional globally uniform component of spatially uncorrelated noise of $\sim 0.1\,°K^2$ variance that would be significant in the tropics (with mean variance $0.85\,°K^2$) but negligible in the extratropics (with mean variance $11.5\,°K^2$) (See Fig. 7).

The strong spatial correlation of temperature fluctuations in all versions of CCSM3 means that increased resolution of spatial heterogeneity could not explain the differences in variability between CCSM3 model resolutions. In those regions where the model variability increases with resolution (i.e. is higher at T85 than at T42), the small implied

contribution of spatial averaging could not explain the large changes in variability with resolution. And in regions where model variability is actually lower at T85 than at T42 resolution, some other phenomenon must be outweighing the small spatial averaging effects. These results suggest that the effects of model resolution on variability occur predominantly though altered large-scale phenomena (correlated over many grid cells), not directly through improved resolution of small-scale phenomena.

3.4 Spatial effects as function of altitude

We would expect that the spatial correlation of temperature fluctuations should increase with altitude: the less the interaction with the surface, the more spatially homogeneous the atmosphere should become. In an examination of T42 output we find that this does in fact hold. (We use T42 output to investigate altitude dependence since the T85 output used here has temperature data only up to reference height.) As in the previous example, we estimate the spatially uncorrelated component of variability by averaging 2×2 pixels to artificially degrade resolution and then comparing to the original-resolution model (Fig. 8, right column, which shows variance across all frequencies). As expected, averaging effects decline monotonically with altitude: averaging reduces variance by 17 % at the surface, 12 % at reference height (as mentioned earlier), and 5 % at 300 mb. This gradual decline in the relative contribution of spatially uncorrelated fluctuations contrasts with the trend in variability itself. (Temperature variability is roughly constant from surface through 850 mb and then drops abruptly above the boundary layer; see Tables S7–S9).

At all altitudes, spatial averaging effects do not appear to explain the differences in variance between model resolutions. Figure 8, left panel, compares variance at T42 and T31 resolutions from surface to 300 mb. At near-surface altitudes, T42 and T31 models show similar global average variance (to <1 %), but with strong regional differences. These patterns diminish with altitude; by 300 mb, T42 variance is uniformly 13 % greater than that in T31. This difference is larger than can be simply explained as a spatial averaging effect, given the small apparent contribution of spatially uncorrelated fluctuations. The averaging test of Fig. 8, right panel, implies that resolving spatial heterogeneities when increasing resolution from T31 to T42 would produce only a 4 % increase in variance. (To convert values in right column to the corresponding expected change between T42 and T31, multiply by 2/3). We show in Fig. 8 only total integrated variance, but these discrepancies hold at all frequencies when disaggregated. At all altitudes, the changes in temperature variability associated with model resolution appear to involve changes in large-scale phenomena.

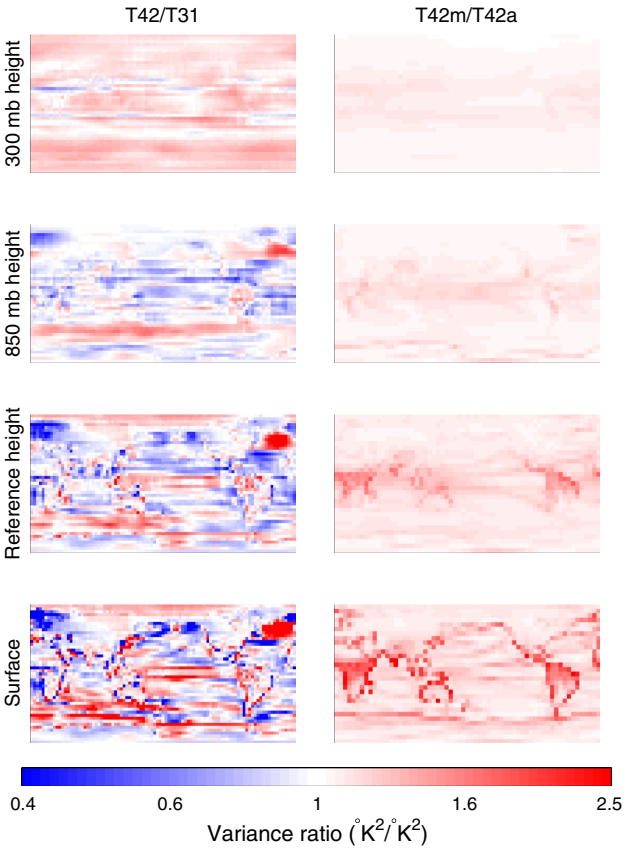

Fig. 8 *Left:* Comparison of overall variance with resolution at four altitudes, shown as the ratio ($^\circ K^2/^\circ K^2$) of T42:T31. Output at the two resolutions are compared by pixel matching to the lower resolution, yielding a T31 figure. Regional patterns of variability increases and decreases are strongest at the surface and diminish with altitude. By 300 mb, the higher-resolution model shows relatively uniform 13 % increase in variance. *Right* Ratio of average variance in T42 to that in the same dataset degraded to $4\times$ coarser resolution (below T31). As expected, the effects of spatial averaging decrease with altitude. The ratio of variance in raw to averaged model output is 1.21, 1.15, 1.08, and 1.05 from surface to 300 mb. (Note that the degraded-from-T42 resolution is lower than that of T31; multiply by 2/3 for comparison with left panel.) In no case can spatial averaging effects explain variance differences with resolution

3.5 Model fidelity

One of the rationales for running climate models at finer spatial resolution, despite the increased computational demand, is the expectation that improved resolution will produce improved model fidelity. We examine that assumption for CCSM3 by comparing temperature variability in model output with that in the NCEP–DOE reanalysis product, which we take as ground truth. CCSM3 generally overestimates total global temperature variability relative to the NCEP–DOE reanalysis (Table 1). Because increasing resolution in this model reduces global variability, it therefore also improves average model fidelity. Furthermore, at a regional level, increasing resolution tends to reduce biases

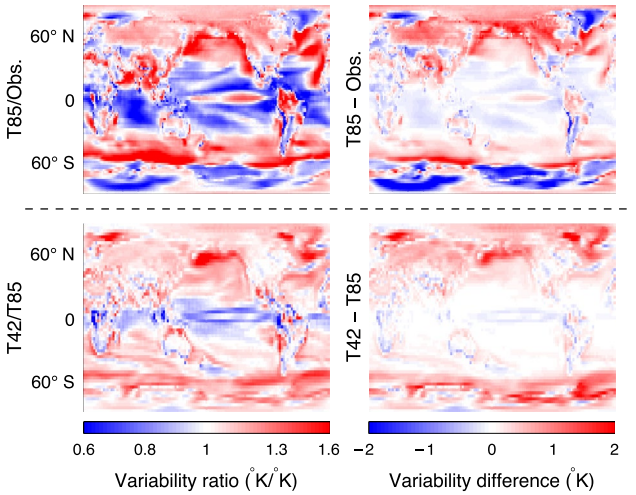

Fig. 9 *Top* Total variability (°K) in the T85 model compared to reanalysis. *Bottom* Same comparison for T42:T85 resolution, for reference. Unlike previous figures we show the ratio of low- to high-resolution versions, so that for each column, the places where colors match indicate that increasing resolution improves model representation of variability. All plots are made by "pixel matching" to the lower resolution. Note that, in this figure we return to plotting variability, rather than variance, to allow a more intuitive understanding of model/observation discrepancies

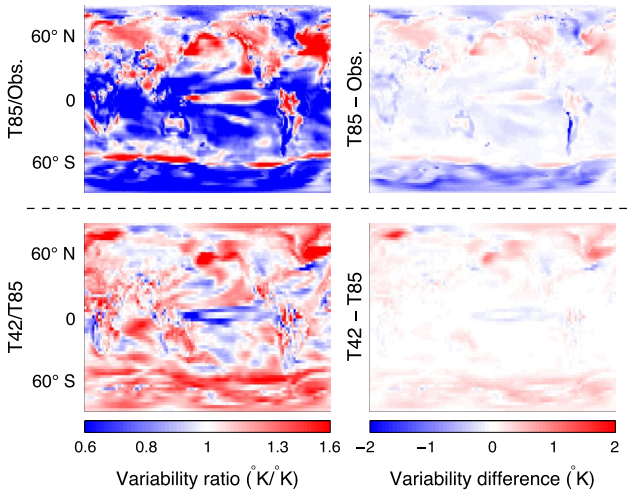

Fig. 10 *Top* Interannual variability (periods: 1–10 years; °K) in the T85 model compared to reanalysis. *Bottom* Same comparison for T42:T85 resolution, for reference (Compare also to Fig. 9). CCSM3 consistently underestimates interannual variability, and increasing resolution exacerbates this bias

in total variability regardless of the sign of those biases. That is, the pattern of changes produced by increasing resolution—lower variability at high latitudes and higher variability in the tropics—roughly matches the pattern of observation-model discrepancies (Fig. 9, bottom row; see also Fig S8). The result is that at most locations, increasing CCSM3 spatial resolution appears to improve fidelity in total integrated temperature variability.

However, the improvements in fidelity with model resolution do not hold for all frequencies. As shown in Fig. 4, increasing model resolution produces similar effects over most of the frequency range sampled (from 5 days to interannual). But observation-model discrepancies are very different for subannual and interannual frequencies (compare Figs. 9, 10 and S9). While CCSM3 overestimates subannual variability, it underestimates interannual variability. (Underestimation of long-period variability is common in climate models. Laepple and Huybers (2014a) compare historical CMIP5 model runs to gridded instrumental observations, and find a large spread in model behavior, but with the multi-model mean biased low for interannual variability.) In CCSM3, increasing model resolution exacerbates this systematic bias. These results are significant to two standard errors despite the relatively short datasets (100 years for model output and 35 years for observations). (See Table S6.) Given that different physical processes may govern variability occurring at short and long timescales, it remains unclear why altering model resolution produces such consistent effects across most frequencies.

4 Discussion

In this study of a single climate model, we find that altering the spatial resolution of the atmospheric component affects temperature variability in all regions and at all frequencies, even while means are relatively unaffected. Globally averaged near-surface temperature variability decreases with increasing resolution, at all but the highest frequencies (periods of 2–5 days). The global average is dominated by near-universal decreases in variability at mid and high latitudes; the tropics show a consistent band of variability increase with resolution. This complex spatial pattern of variability differences associated with resolution is evidently related to surface interactions, since its strength decreases with altitude. However, while surface topography representation seems to be important in the boundary layer, it is likely not the only meaningful driver, since resolution-driven changes in variability are evident even at 300 mb. At this altitude, increasing resolution from T31 to T42 produces a fairly uniform 13 % increase in variability (larger than is consistent with simple averaging effects).

The changes in variability induced by changes in model resolution (at all regions, altitudes, and frequencies) appear to be driven by alterations in large-scale circulations rather than by simple spatial averaging effects. Spatial averaging effects are negligible since the overwhelming bulk of temperature variability is spatially correlated over large distances, at all model resolutions and at all periods from days to decades. Changes in variability between model resolutions must therefore result from changes in large-scale phenomena.

The finding that differences in variability with resolution are nonlinear (Table 1) is also consistent with a limited impact of spatial averaging effects. If variability (standard deviation) were perfectly uncorrelated in space, then it would scale inversely with grid cell length. If variability is dominated by large-scale phenomena, the effects of altering model spatial resolution will be difficult to predict in advance (see also Jung et al. 2012). Nonlinear responses to changes in resolution have been seen in prior studies, though these involve smaller grid cell sizes than tested here. These studies also suggest that there may be some resolution threshold at which model performance no longer alters with resolution (Jung et al. 2012; Kinter et al. 2013, and references therein). Results here showing nonlinear differences at coarser resolutions would also be consistent with a range of resolution effects impacting different phenomena.

Improvements in model fidelity are often cited as a motivation for increasing resolution. In CCSM3, the changes in temperature variability produced by increasing resolution do improve model fidelity in most regions and at most frequencies. At subannual frequencies, increasing resolution reduces both low biases in the tropics and high biases at high latitudes. However, at interannual frequencies, increasing resolution degrades fidelity, exacerbating a low bias in CCSM3. The discrepancy raises some questions about the mechanisms driving variability changes between model versions.

It is generally assumed that the physical drivers of variability differ at sub-annual and interannual frequencies. (Differences in model biases are also consistent with this assumption.) Sub-annual temperature variability is likely driven in large part by eddy transport. Interannual temperature variability will include some contribution from high-frequency processes (which can produce statistical effects at long timescales) but should be dominated by intrinsically low-frequency variability in ocean temperatures. It is therefore surprising that increasing spatial resolution in the atmosphere/land component of a climate model alone produces effects in interannual variability similar to those at higher frequencies, both in spatial patterns and in magnitude. Some recent studies do suggest that alterations in atmospheric model resolution should influence variability over a wide range of frequencies (see Clement et al. 2015; Zhang et al. 2014). At least in the model we study here, atmospheric (and possibly also land and subsurface) processes may play an even stronger role in modulating interannual variability than would be typically expected.

The results of this paper are largely descriptive, but offer suggestions for future research directions. First, systematic study across model resolution can offer insight into the mechanisms governing temperature variability. Our results suggest that those mechanisms are synoptic- and larger-scale atmospheric phenomena, that may respond to smaller-scale changes in model resolution. Multi-model comparisons would be needed to determine both whether this conclusion is robust across models and whether resolution robustly improves model fidelity with observations. Comparisons across different parametrizatons could help diagnose whether changes are due directly to resolution or indirectly via associated parameter changes. The nonlinear effects of resolution at different frequencies in CCSM3 (Table 1) also suggest that study over a far greater range of resolutions would be helpful to determine if trends are maintained as models move to still higher spatial resolutions and to clarify whether the large computational demands of increased resolution are warranted. Finally, the results here suggest that identifying differences in marginal distributions are not sufficient for identifying mechanisms. Frequency-resolved studies appear to be necessary tools for identifying and characterizing model responses to changes in spatial resolution.

Acknowledgments We thank Matthew Huber, Robert Jacob, Ben Kirtman, Leonard Smith, and Michael Stein for helpful comments on this paper. This research was performed as part of the Center for Robust Decision-making on Climate and Energy Policy (RDCEP) at the University of Chicago, funded by a grant from the National Science Foundation (NSF) Decision Making Under Uncertainty program (SES-0951576). Model runs were completed by NCAR and are available publicly on www.earthsystemgrid.com. This work was completed in part with resources provided by the University of Chicago Research Computing Center.

References

Alexander LV, Zhang X, Peterson TC et al (2006) Global observed changes in daily climate extremes of temperature and precipitation. J Geophys Res Atmos 111(D5):D05,109

Bacmeister JT, Wehner MF, Neale RB et al (2013) Exploratory high-resolution climate simulations using the community atmosphere model (CAM). J Clim 27(9):3073–3099

Clement A, Bellomo K, Murphy LN et al (2015) The Atlantic multidecadal oscillation without a role for ocean circulation. Science 350(6258):320–324

Collier JC, Zhang GJ (2007) Effects of increased horizontal resolution on simulation of the North American monsoon in the NCAR CAM3: an evaluation based on surface, satellite, and reanalysis data. J Clim 20(9):1843–1861

Collins WD, Bitz CM, Blackmon ML et al (2006) The community climate system model version 3 (CCSM3). J Clim 19(11):2122–2143

Delworth TL, Rosati A, Anderson W et al (2012) Simulated climate and climate change in the GFDL CM2.5 high-resolution coupled climate model. J Clim 25(8):2755–2781

Deser C, Capotondi A, Saravanan R et al (2006) Tropical Pacific and Atlantic climate variability in CCSM3. J Clim 19(11):2451–2481

DeWeaver E, Bitz CM (2006) Atmospheric circulation and its effect on Arctic Sea Ice in CCSM3 simulations at medium and high resolution. J Clim 19(11):2415–2436

Gent PR, Yeager SG, Neale RB et al (2010) Improvements in a half degree atmosphere/land version of the CCSM. Clim Dyn 34(6):819–833

Gualdi S, Navarra A, von Storch H (1997) Tropical intraseasonal oscillation appearing in operational analyses and in a family of general circulation models. J Atmos Sci 54(9):1185–1202

Guemas V, Codron F (2011) Differing impacts of resolution changes in latitude and longitude on the midlatitudes in the LMDZ atmospheric GCM. J Clim 24(22):5831–5849

Guilyardi E, Gualdi S, Slingo J et al (2004) Representing El Niño in coupled ocean–atmosphere GCMs: the dominant role of the atmospheric component. J Clim 17(24):4623–4629

Hack JJ, Caron JM, Danabasoglu G et al (2006) CCSM–CAM3 climate simulation sensitivity to changes in horizontal resolution. J Clim 19(11):2267–2289

Hansen J, Sato M, Ruedy R (2012) Perception of climate change. Proc Natl Acad Sci 109(37):E2415–E2423

Holmes CR, Woollings T, Hawkins E et al (2015) Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. J Clim 2015:2221–2236

Huntingford C, Jones PD, Livina VN et al (2013) No increase in global temperature variability despite changing regional patterns. Nature 500(7462):327–330

Iorio JP, Duffy PB, Govindasamy B et al (2004) Effects of model resolution and subgrid-scale physics on the simulation of precipitation in the continental United States. Clim Dyn 23(3–4):243–258

IPCC, 2012: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. In: Field CB, Barros V, Stocker TF, Qin D, Dokken DJ, Ebi KL, Mastrandrea MD, Mach KJ, Plattner GK, Allen SK, Tignor M, Midgley PM (eds). Cambridge University Press, Cambridge, UK and New York, USA, p 582

IPCC (2013) Evaluation of climate models. In: Flato G, Marotzke J, Abiodun B et al (eds) Climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change. Cambridge University Press, Cambridge, UK and New York, USA, pp 741–866

IPCC (2014) Human health: impacts, adaptation, and co-benefits. In: Smith KR, Woodward A, et al. (eds) Climate change 2014: impacts, adaptation, and vulnerability. Cambridge University Press, Cambridge, UK and New York, USA, pp 709–754

Jones GS, Stott PA, Christidis N (2013) Attribution of observed historical near-surface temperature variations to anthropogenic and natural causes using CMIP5 simulations. J Geophys Res Atmos 118(10):4001–4024

Jung T, Miller MJ, Palmer TN et al (2012) High-resolution global climate simulations with the ECMWF model in project Athena: experimental design, model climate, and seasonal forecast skill. J Clim 25(9):3155–3172

Kanamitsu M, Ebisuzaki W, Woollen J et al (2002) NCEP–DOE AMIP-II reanalysis (R-2). Bull Am Meteorol Soc 83(11):1631–1643

Kinter JL, Cash B, Achuthavarier D et al (2013) Revolutionizing climate modeling with project Athena A: multi-institutional, international collaboration. Bull Am Meteorol Soc 94(2):231–245

Kirtman BP, Bitz C, Bryan F et al (2012) Impact of ocean model resolution on CCSM climate simulations. Clim Dyn 39(6):1303–1328

Kobayashi C, Sugi M (2004) Impact of horizontal resolution on the simulation of the Asian summer monsoon and tropical cyclones in the JMA global model. Clim Dyn 23(2):165–176

Laepple T, Huybers P (2014a) Global and regional variability in marine surface temperatures. Geophys Res Lett 41(7):2528–2534

Laepple T, Huybers P (2014b) Ocean surface temperature variability: large model-data differences at decadal and longer periods. Proc Natl Acad Sci USA 111(47):16,682–16,687

Leeds WB, Moyer EJ, Stein ML (2015) Simulation of future climate under changing temporal covariance structures. Adv Stat Climatol Meteorol Oceanogr 1(1):1–14

Marti O, Braconnot P, Dufresne JL et al (2010) Key features of the IPSL ocean atmosphere model and its sensitivity to atmospheric resolution. Clim Dyn 34(1):1–26

Meehl GA, Zwiers F, Evans J et al (2000) Trends in extreme weather and climate events: issues related to modeling extremes in projections of future climate change. Bull Am Meteorol Soc 81(3):427–436

Morak S, Hegerl GC, Christidis N (2013) Detectable changes in the frequency of temperature extremes. J Clim 26(5):1561–1574

Navarra A, Gualdi S, Masina S et al (2008) Atmospheric horizontal resolution affects tropical climate variability in coupled models. J Clim 21(4):730–750

Neale RB, Richter JH, Jochum M (2008) The impact of convection on ENSO: from a delayed oscillator to a series of events. J Clim 21(22):5904–5924

Otto-Bliesner BL, Tomas R, Brady EC et al (2006) Climate sensitivity of moderate- and low-resolution versions of CCSM3 to preindustrial forcings. J Clim 19(11):2567–2583

Poppick A, McInerney DJ, Moyer EJ et al (2016) Temperatures in transient climates: improved methods for simulations with evolving temporal covariances. Ann Appl Stat 10(1):477–505

Reichler T, Kim J (2008) How well do coupled models simulate today's climate? Bull Am Meteorol Soc 89(3):303–311

Roeckner E, Brokopf R, Esch M et al (2006) Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. J Clim 19(16):3771–3791

Saha S, Moorthi S, Pan HL et al (2010) The NCEP climate forecast system reanalysis. Bull Am Meteorol Soc 91(8):1015–1057

Schneider T, Bischoff T, Plotka H (2014) Physics of changes in synoptic midlatitude temperature variability. J Clim 28(6):2312–2331

Thornton PK, Ericksen PJ, Herrero M et al (2014) Climate variability and vulnerability to climate change: a review. Glob Chang Biol 20(11):3313–3328

Wehner MF, Smith RL, Bala G et al (2010) The effect of horizontal resolution on simulation of very extreme US precipitation events in a global atmosphere model. Clim Dyn 34(2–3):241–247

Yeager SG, Shields CA, Large WG et al (2006) The low-resolution CCSM3. J Clim 19(11):2545–2566

Zhang H, Clement A, Di Nezio P (2014) The South Pacific meridional mode: a mechanism for ENSO-like variability. J Clim 27(2):769–783