

Time of emergence of climate signals

E. Hawkins¹ and R. Sutton¹

The time at which the signal of climate change emerges from the noise of natural climate variability (Time of Emergence, ToE) is a key variable for climate predictions and risk assessments. Here we present a methodology for estimating ToE for individual climate models, and use it to make maps of ToE for surface air temperature (SAT) based on the CMIP3 global climate models. Consistent with previous studies we show that the median ToE occurs several decades sooner in low latitudes, particularly in boreal summer, than in mid-latitudes. We also show that the median ToE in the Arctic occurs sooner in boreal winter than in boreal summer. A key new aspect of our study is that we quantify the uncertainty in ToE that arises not only from inter-model differences in the magnitude of the climate change signal, but also from large differences in the simulation of natural climate variability. The uncertainty in ToE is at least 30 years in the regions examined, and as much as 60 years in some regions. Alternative emissions scenarios lead to changes in the both the median ToE (by a decade or more) and its uncertainty. The SRES B1 scenario is associated with a very large uncertainty in ToE in some regions. Our findings have important implications for climate modelling and climate policy which we discuss.

NOTE: This is the author's version of the paper accepted for publication in GRL. Copyright 2011 American Geophysical Union. Further reproduction or electronic distribution is not permitted.

1. Introduction

The signal of anthropogenic climate change is emerging against the background of natural climate variability. Only when the signal of change is of sufficient magnitude relative to this background variability can we be confident that a significant change has been detected. Such detection is a necessary step in the process of attributing a particular change to a specific cause, such as the observed rise in greenhouse gas concentrations [Hegerl *et al.*, 2007]. A headline conclusion from the IPCC AR4 was that “most of the observed increase in global average temperatures since the mid-20th century is very likely due to the observed increase in anthropogenic greenhouse gas concentrations”. Of greater importance for adaptation to climate change are changes on the regional and local scales that affect people, economies and ecosystems. However, on these smaller scales natural variability is larger, making detection more difficult. Some detection and attribution studies that have addressed these scales have obtained positive results [Hegerl *et al.*, 2007; Stott *et al.*, 2010], but for many regions and variables the signal of anthropogenic change has yet to clearly emerge from the ‘noise’ of natural climate variability.

So when will the signal emerge? And where and how? These are key questions for adaptation policy and planning in particular. Much attention has focused on the absolute magnitude of future climate change, and uncertainties in this magnitude [e.g. Randall *et al.*, 2007]. But in many situations it is not the absolute magnitude that matters so much as the magnitude of change relative to the background levels of variability. The reason is that many natural and human systems are inherently adapted to the local background level of variability. It is when changes move outside this range that major impacts are most likely to arise. For example, Lobell and Burke [2008] demonstrate that uncertainty in future crop yields is often dominated by uncertainty in projections of temperature, rather than precipitation, because the changes in temperature are far further outside the range of natural variability to which the crop is adapted.

The question of when a significant climate change signal will emerge in different regions was discussed and presented in tabular form in the IPCC AR4 [Table 11.1 of Christensen *et al.*, 2007]. The general pattern of results is that the Time of Emergence (ToE) is soonest (~ 10 years) for low latitude regions, longest (20-40 years) for mid-latitude regions and of intermediate duration (15-20 years) for polar regions. These findings were recently confirmed by Mahlstein *et al.* [2011], who additionally provided results at country scale, and demonstrated that the earliest emergence of significant warming occurs in the summer season in low latitude countries.

The results presented in AR4 and Mahlstein *et al.* [2011], whilst valuable, leave many open questions about ToE. Basic but important questions concern the choices involved in defining both signal and noise. In some studies the definition of noise has been broadened to include the uncertainty in the climate response to anthropogenic forcing and the uncertainty in future anthropogenic emissions [Giorgi and Bi, 2009; Hawkins and Sutton, 2009, 2011]. In this study we focus on the natural internal variability of climate as the key source of noise relevant for ToE. A major motivation for our study is to address the question, *what is the uncertainty in estimates of ToE?* This is a key question for adaptation policy, as it is fundamental to the risk assessments on which such policy must rely. To address it we develop a methodology for estimating ToE from individual climate models, rather than relying on a multi-model mean. Our approach recognises that current climate models show substantial differences not only in their simulation of the climate change signal (e.g. in their climate sensitivity), but also in their simulation of the natural internal variability of climate. We demonstrate that both these dimensions of uncertainty directly influence the uncertainty in ToE. We also consider the question *how is ToE sensitive to alternative scenarios for anthropogenic emissions?* This question has obvious relevance to mitigation policy.

In Section 2 we describe the methods and data used. Section 3 explores the natural variability across the models and observations, and in Section 4 we show our ToE estimates and discuss the sensitivity to various choices in the analysis. We summarise in Section 5.

2. Methods and Data

Estimating the Time of Emergence (ToE) of any climate signal requires estimates for the climate change signal (S)

¹NCAS-Climate, University of Reading, UK.

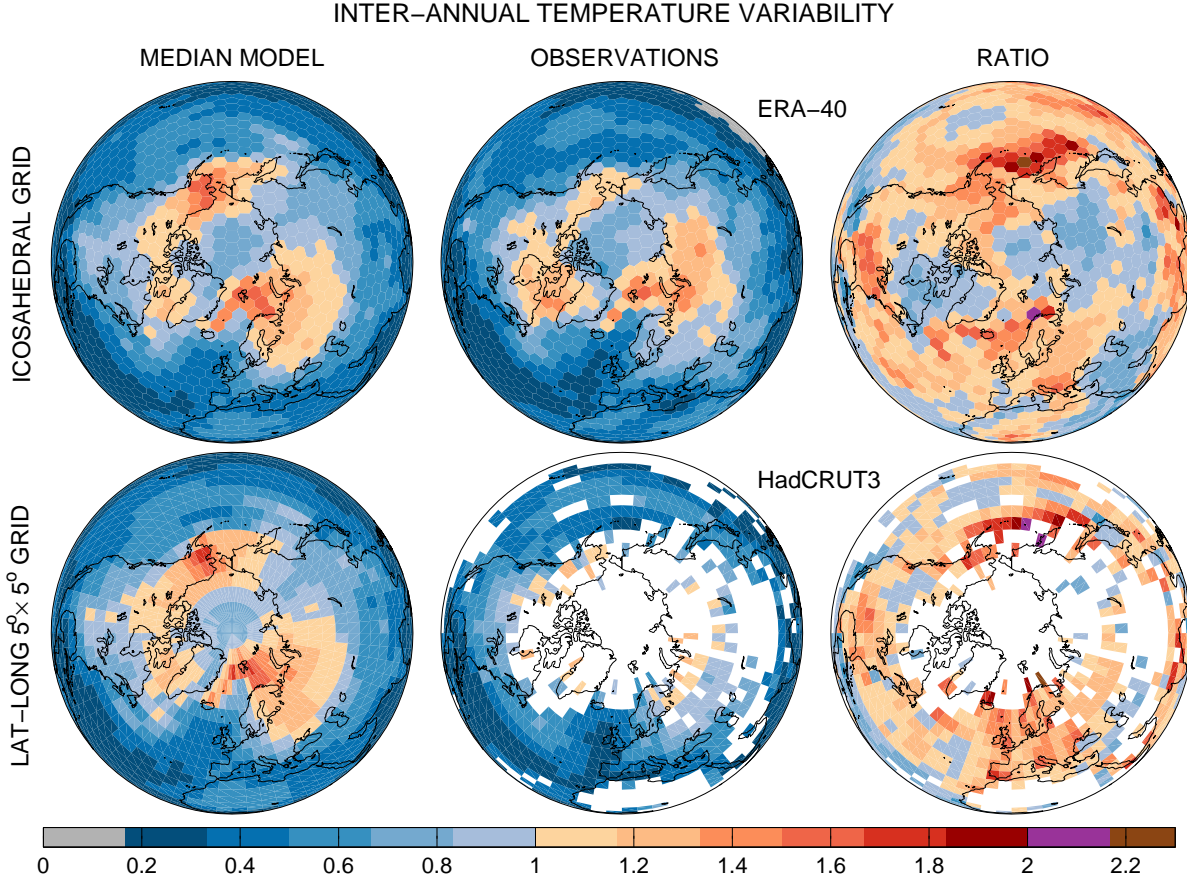


Figure 1. The median model interannual variability using annual means on different spatial grids: an icosahedral grid with roughly equal areas (top) and a regular $5^\circ \times 5^\circ$ grid (bottom). The observations are shown in the middle column for ERA-40 (top) and HadCRUT3 (bottom, with missing values in white). These four panels have units of $[K]$. The ratio of the median model to the observations shows the GCMs tend to overestimate the variability in the northern extra-tropical regions (right), with units of $[K/K]$. The two observational estimates also agree well where data exists for HadCRUT3.

and for the variability (or ‘noise’ - N). We define the ToE as the first year in which the signal-to-noise ratio, S/N , crosses particular threshold values (such as 1 and 2), but for certain impacts other values may be more appropriate.

To estimate the signal for surface air temperature (SAT) we use data from 1950-2099 from single ensemble members from a set of 15 GCMs which have run all three SRES scenarios (A2, A1B, B1). We define S by first regressing SAT from a single ensemble member for each model in each grid box (T_{local}) onto a smoothed version of the global mean SAT projection ($\tilde{T}_{\text{global}}$), making the assumption that the local changes scale with global temperature. This assumption works well in the past [e.g. *van Oldenborgh et al., 2009*], but may break down in the future if the aerosol and greenhouse gas signals become decoupled. We estimate the regression coefficients (α, β) between T_{local} and $\tilde{T}_{\text{global}}$, and define

$$S(t) = \alpha \tilde{T}_{\text{global}}(t) + \beta \quad (1)$$

for each grid point. The smoothing for $\tilde{T}_{\text{global}}$ is a fourth order polynomial, fitted over the period 1950-2099 [*Hawkins and Sutton, 2009*]. All estimates are made relative to the reference period 1986-2005 - we choose the recent past as being of greatest relevance for adaptation. We test this method using an ensemble of 10 simulations with a single model, producing an ensemble standard deviation in ToE for each of the three regions considered later in Section 4 of less than 3 years (not shown).

To estimate the noise in SAT we utilise each GCM’s pre-industrial control simulations. We define N as the interan-

nual standard deviation of seasonal (or annual) means. Note that our choice contrasts with previous studies, which have used the variability of 20-year averages [*Christensen et al., 2007; Giorgi and Bi, 2009*]. Because variance decreases with averaging our choice results in later estimates for ToE, but we argue that it has greater relevance to adaptation policy. As noted in the Introduction our definition of ToE differs further from that of *Giorgi and Bi [2009]* in that their definition of N combines an estimate of the intermodel variance in the signal strength with an estimate of the internal variability. Our view is that these two sources of variance are very different in origin - in particular the latter is a fundamental property of the climate system, whilst the former is not - and therefore it is helpful to keep them separate. Furthermore, we suggest that ToE - as we define it - is also a fundamental property of the climate system. The true values of ToE are unknown, but we can use climate models to estimate it.

For observational comparisons of the variability we use the HadCRUT3 dataset [*Brohan et al., 2006*] from 1900-2010 and ERA-40 [*Uppala et al., 2005*] from 1958-2001. For HadCRUT3, the interannual variability is derived by requiring all 12 months of data to be present in a year at each grid point to create an annual average, and at least 80 years present since 1900 to estimate the interannual variability. The variability is the standard deviation of the detrended annual means, where HadCRUT3 is detrended as above by regressing out the signal due to increases in the global mean temperature fitted with a 4th order polynomial. We use

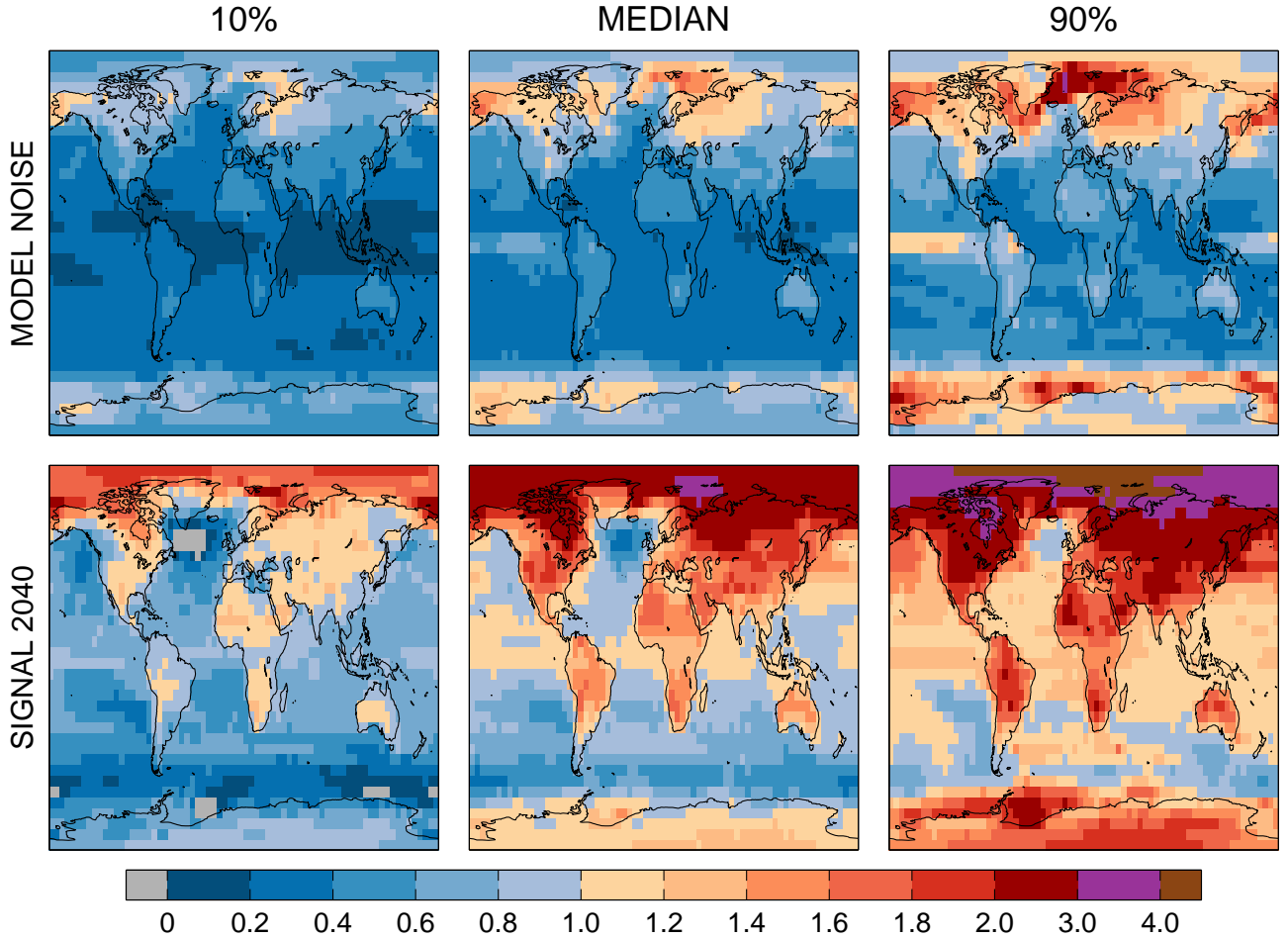


Figure 2. Inter-annual variability (top) and signal in 2040 (bottom) using annual means from the GCMs. The estimates include the 10th, 50th and 90th percentiles for each gridpoint (columns). Note the non-linear colour scale at high values.

a linear detrending at each grid point for ERA-40, which is spatially and temporally complete. Note that these estimates are likely to be slightly too large as there will be residual variability due to non-linearities in the forced trend. Cleanly separating the trend from the variability is non-trivial, and we do not consider it further here.

3. Surface air temperature variability

Different GCMs show a very wide range of interannual variability [e.g. *Randall et al.*, 2007]. These differences in variability characteristics across GCMs raises a key issue: what estimate for the variability should be used to estimate ToE? One option is to use the historical observations, but this is complicated by the trends in temperature, and short time-series for many regions. We later test the sensitivity to other options which include using each GCM's own estimate of internal variability, or using a multi-model median estimate.

3.1. Sensitivity to spatial grid used

Before considering the differences in variability, another question arises concerning the spatial grid for analysis. On a regular latitude-longitude grid, the grid cells in the polar regions have a far smaller spatial area than in the tropics. This could distort the analysis as smaller areas are likely to have larger variability solely due to their size and would ar-

tificially inflate ToE at the poles, relative to tropical regions for instance.

To examine this sensitivity we compare the interannual variability for the high northern latitudes using the native GCM data remapped on two different grids before analysis (Fig. 1). The median standard deviation of annual mean temperature on an icosahedral grid, with roughly equal areas in each grid cell (top left), and a regular latitude-longitude $5^\circ \times 5^\circ$ grid (bottom left) show very similar patterns, but there are two key points to note. First, there is a local minimum in SAT variability in the vicinity of the North pole, which is perhaps not widely appreciated [e.g. *Christensen et al.*, 2007; *Mahlstein et al.*, 2011]. Secondly, there are small differences between the results for the different grids, with the icosahedral grid showing generally smaller variability in some high latitude regions as would be expected if the grid size is having an effect. However, the differences are small and for the ToE analysis we focus on the regular latitude-longitude $5^\circ \times 5^\circ$ grid. Fig. S2 shows maps of ToE on the icosahedral grid for comparison.

3.2. Comparing models and observations

Our observational estimates of the interannual variability of annual mean SAT are shown in Fig. 1. When comparing the observations to the median model estimate, it is seen that the GCMs tend to have too much variability in the northern extra-tropical region, by up to 20% [also see *Mahlstein et al.*, 2011]. Considering that the observations are likely to be an over-estimate of the true internal

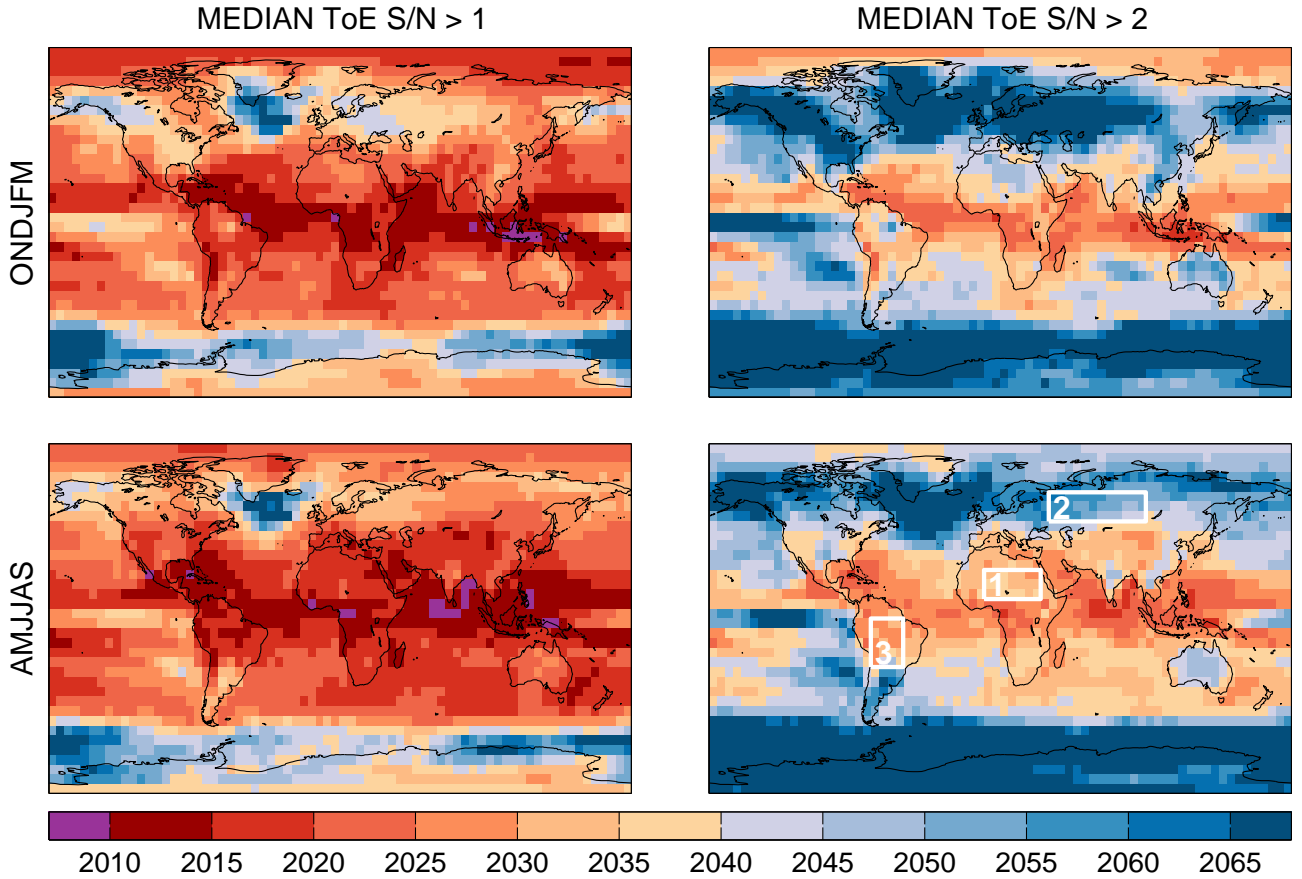


Figure 3. Median Time of Emergence for surface air temperatures for October-March (top) and April-September (bottom). First year when temperature has expected $S/N > 1$ (left) and $S/N > 2$ (right). The regions indicated by the white boxes are used in Figs. 4, 5.

variability (see Section 2), this finding would appear to be robust. This bias will affect estimates of the ToE by making it later than it would be if an observational estimate was used. In other regions it appears the GCMs have too little variability, although this may be an artefact of imperfectly detrending the observations. In all examples, the variability in SAT is generally greater at higher latitudes (excepting the local minimum over the pole), and over land (Fig. 2). Exceptions are the ENSO region, and the region of the Barents Sea, both of which show unusually high variability.

4. The Time of Emergence for regional temperatures

The signal of SAT change in 2040 shows a wide range across the GCMs (Fig. 2), with a range (10-90% quantiles) of up to 4°C at high latitudes and typically 1°C at lower latitudes. In addition, the range (10-90% quantiles) of the GCM estimates of the interannual variability can be up to a factor of 3 in standard deviation. Much previous work has focussed on the large range of climate sensitivity, but the ranges in variability are arguably as important across the GCMs considered, contributing to uncertainty in S/N .

As noted in Section 2, we define the Time of Emergence (ToE) as the first year when the S/N is larger than a particular threshold. ToE is estimated for each model separately, and maps of the median ToE for SAT under SRES A1B, for two half-years, are shown in Fig. 3, using two different S/N thresholds. Fig. S1 shows maps of ToE for individual GCMs.

For many tropical regions, the median ToE for $S/N > 1$ is within the next decade, and before 2030 for $S/N > 2$. The early emergence in these regions, and the tendency for earlier emergence in boreal spring/summer (April-September) than in boreal autumn/winter (October-March), is consistent with the results of *Mahlstein et al.* [2011]. For extra-tropical regions the ToE times are delayed by several decades, with ToE for $S/N > 2$ later than 2060 for many locations. In the northern extratropics, ToE is again generally later for autumn/winter than spring/summer. However, an interesting contrast to this seasonal variation is seen over the central Arctic, where the median ToE occurs 5-10 years *earlier* in autumn/winter than in spring/summer (Fig. S2), primarily because the signal is stronger.

Histograms of ToE for $S/N > 2$ for three equal area land regions (as indicated in Fig. 3) are shown in Fig. 4, for half-years, using the SRES A1B scenario. Note that because the noise variance decreases with averaging, the ToE for an area mean is not the same, and will generally be earlier, than the mean ToE for that area. The blue bars show the ToE using each GCM's own estimate of the variability, and the green bars show the ToE using the median estimate of the variability as N . The red bars show the projected ToE using the median signal and median noise. Fig. S3 shows the histograms using the median signal and each GCM's variability.

The blue bars demonstrate the large uncertainty in estimates of ToE from different climate models. This uncertainty is at least 30 years in all three regions, and as much 60 years in some regions. Also evident in Fig. 4 is the seasonal variation, with all three regions showing a shift to earlier

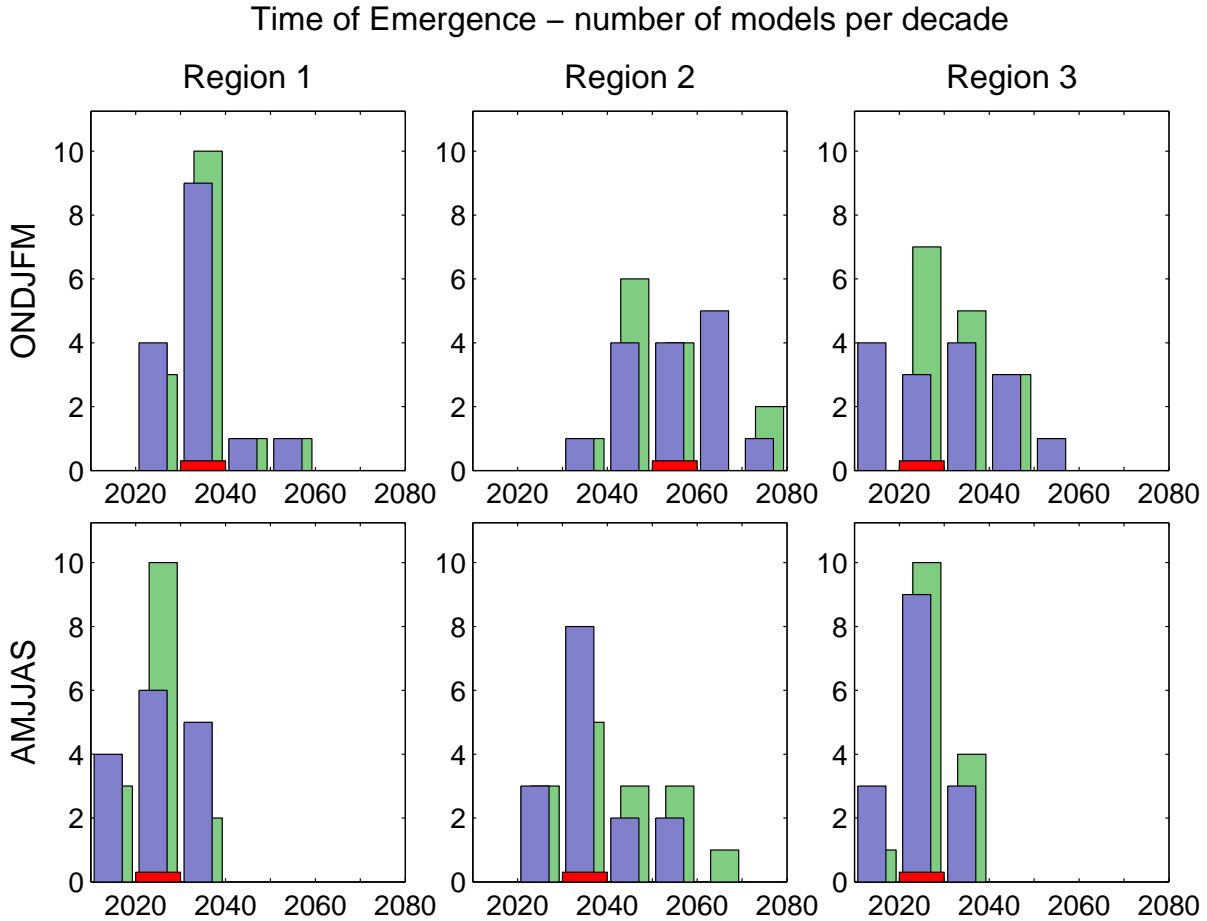


Figure 4. Time of emergence for surface air temperatures for three equal area regions as indicated in Fig. 3. October–March (top) and April–September (bottom). Histograms representing the number of models which cross the $S/N > 2$ temperature threshold in each decade, using each GCM’s own variability (blue) and median variability (green). The red bars indicate the decade of ToE when using both the median signal and median noise.

ToE in boreal spring/summer (even though region 3 is in the Southern hemisphere). The distributions of ToE show a clear central peak for some regions and seasons (e.g. region 1 for boreal autumn/winter), but in other cases are rather flat or even bimodal (e.g. region 3 in boreal autumn/winter and region 1 in boreal spring/summer).

The differences between the blue and green bars indicate the change in uncertainty in ToE when using a single (median) estimate for the noise, N . Whilst the distributions change there is no consistent reduction, or increase, in spread. The same is true when using the median signal together with estimates of the noise from individual models (see Fig. S3). These results demonstrate that the inter-model spread in both signal and noise contribute significantly to the spread in ToE.

The final sensitivity on ToE we consider is the dependence on emissions scenario (Fig. 5). Overall, SRES A1B shows a slightly earlier ToE for the various regions, because it is actually the warmest scenario in the near-term [e.g. *Meehl et al.*, 2007], whereas A2 is warmer at the end of the 21st century. The B1 scenario shows a shift towards later ToE; more striking, however, is a large increase in the uncertainty in ToE under this scenario. This is because the lower levels of emissions result in a more gradual rise in temperatures, producing a wider range of threshold crossing times [also see *Joshi et al.*, 2011].

Lastly we consider the ‘Global Temperature for Regional Emergence’ (GTRE), where the emergence is expressed as

the change in global mean temperature (from 1986–2005) required to produce a regional threshold crossing in S/N [in a similar way to *Mahlstein et al.*, 2011]. Fig. S4 demonstrates that this metric, which may be more useful for mitigation decisions, effectively reduces the emissions uncertainty, but still produces wide ranges in GTRE for different models and regions.

5. Discussion and Conclusions

We have developed a new methodology for estimating Time of Emergence (ToE) for individual climate models, which has enabled us to estimate the uncertainty in ToE that arises from intermodel differences in their simulations of natural variability (noise) as well as differences in their simulations of the signal of climate change. We applied our methods to simulations of surface air temperature in the CMIP3 ensemble of climate models. Our major findings are:

1. Consistent with many other studies there is a large spread in the magnitude of the climate change signal simulated by the models when forced by the same forcing scenario. In some higher latitude regions the magnitude of the signal in 2040 differs by more than 3°C .
2. Less widely recognised, there is also a very large spread in the amplitude of natural variability simulated by the models. In some regions the amplitude of variability differs by a factor of 3 or 1.5°C .

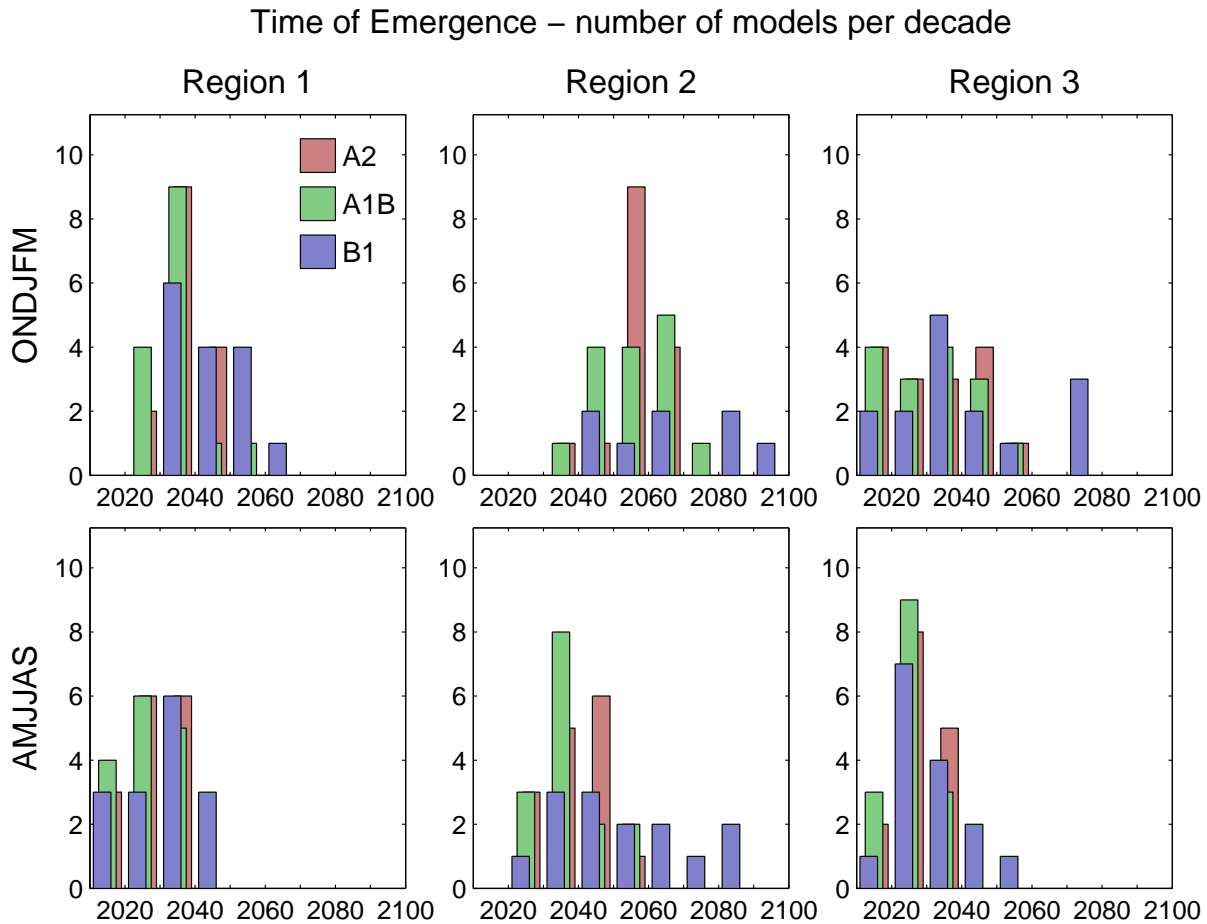


Figure 5. Time of emergence for surface air temperatures for three equal area regions as indicated in Fig. 3. October-March (top) and April-September (bottom). Histograms representing the number of models which cross the $S/N > 2$ temperature threshold in each decade, using each GCM's own variability and different scenarios (blue - B1, green - A1B, red - A2).

3. The median ToE occurs several decades sooner in low latitudes, particularly in boreal spring/summer, than in mid-latitudes. The median ToE in the Arctic occurs 5-10 years sooner in boreal autumn/winter than in boreal spring/summer.

4. There is a very large uncertainty in ToE arising from the inter-model differences in simulating both signal and noise. This uncertainty is at least 30 years in all regions and as much 60 years in some regions.

5. Alternative emissions scenarios lead to changes in the both the median ToE (by a decade or more) and its uncertainty. The SRES B1 scenario is associated with a very large uncertainty in ToE in some regions.

Our findings, especially the large uncertainty in ToE, have important implications for climate policy. We have argued that in many cases major impacts (e.g. widespread crop failure) are likely to be associated with crossing thresholds in signal-to-noise. To develop robust adaptation strategies, policy makers and planners must take into account the large uncertainty in *when* these events may take place [also see Joshi *et al.*, 2011].

A limitation of our study is that it is based on the analysis of a particular climate model ensemble of opportunity: CMIP3. This ensemble is unlikely to span the full range of uncertainty in ToE, particularly in light of recent studies which demonstrate that different climate models are not independent [e.g. Knutti *et al.*, 2010; Pennell and Reichler, 2010]. These points highlight the importance of improving model simulations of natural climate variability, and of exploring the potential for identifying observational constraints on ToE. There is also need to estimate ToE for a

wider range of climate and climate-related variables. In view of the importance of ToE for climate adaptation and mitigation policy, we suggest that this research, and the related development work to improve climate models, should be accorded high priority.

Acknowledgments. EH and RS are supported by NCAS-Climate. We thank Hilary Weller for technical help with the icosahedral grid and Jonathan Gregory for useful suggestions. We also acknowledge Irina Mahlstein and an anonymous reviewer for their constructive comments which helped improve the paper.

References

- Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850, *J. Geophys. Res.*, *111*, D12106, doi:10.1029/2005JD006548.
- Christensen, J. H., et al. (2007), *Regional Climate Projections. In: Climate Change 2007: The Physical Science Basis.*, Cambridge University Press, Cambridge, UK.
- Giorgi, F., and X. Bi (2009), Time of emergence (TOE) of GHG-forced precipitation change hot-spots, *Geophys. Res. Lett.*, *36*, L06709, doi:10.1029/2009GL037593.
- Hawkins, E., and R. Sutton (2009), The potential to narrow uncertainty in regional climate predictions, *Bull. Amer. Met. Soc.*, *90*, 1095–1107, doi:10.1175/2009BAMS2607.1.
- Hawkins, E., and R. Sutton (2011), The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dyn.*, *37*, 407–418, doi:10.1007/s00382-010-0810-6.

- Hegerl, G., F. Zwiers, P. Braconnot, N. Gillett, Y. Luo, J. Marengo Orsini, N. Nicholls, J. Penner, and P. Stott (2007), *Understanding and attributing climate change. In: Climate Change 2007: The Physical Science Basis.*, Cambridge University Press, Cambridge, UK.
- Joshi, M., E. Hawkins, R. Sutton, J. Lowe, and D. Frame (2011), Projections of when temperature change will exceed 2°C above pre-industrial levels, *Nature Climate Change*, *1*, 407–412, doi:10.1038/nclimate1261.
- Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. Meehl (2010), Challenges in combining projections from multiple climate models, *J. Climate*, *23*, 2739–2758, doi:10.1175/2009JCLI3361.1.
- Lobell, D. B., and M. B. Burke (2008), Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation, *Environ. Res. Lett.*, *3*, 034,007, doi:10.1088/1748-9326/3/3/034007.
- Mahlstein, I., R. Knutti, S. Solomon, and R. Portmann (2011), Early onset of significant local warming in low latitude countries, *Env. Res. Lett.*, *6*, 034,009, doi:10.1088/1748-9326/6/3/034009.
- Meehl, G. A., et al. (2007), *Global climate projections. In: Climate Change 2007: The Physical Science Basis.*, Cambridge University Press, Cambridge, UK.
- Pennell, C., and T. Reichler (2010), On the effective number of climate models, *J. Climate*, *24*, 2358–2367, doi:10.1175/2010JCLI3814.1.
- Randall, D. A., et al. (Eds.) (2007), *Climate Models and Their Evaluation. In: Climate Change 2007: The Physical Science Basis.*, Cambridge University Press, Cambridge, UK.
- Stott, P., N. Gillett, G. Hegerl, D. Karoly, D. Stone, X. Zhang, and F. Zwiers (2010), Detection and attribution of climate change: a regional perspective, *WIREs: Climate Change*, *1*, 192–211, doi:10.1002/wcc.34.
- Uppala, S. M., et al. (2005), The ERA-40 re-analysis, *QJRM*, *131*, 2961–3012, doi:10.1256/qj.04.176.
- van Oldenborgh, G. J., S. Drijfhout, A. van Ulden, R. Haarsma, A. Sterl, C. Severijns, W. Hazeleger, and H. Dijkstra (2009), Western europe is warming much faster than expected, *Climate of the Past*, *5*(1), 1–12, doi:10.5194/cp-5-1-2009.

E. Hawkins and R. Sutton, Department of Meteorology, University of Reading, Reading.
RG6 6BB, UK. (e.hawkins@reading.ac.uk)

Time of emergence of climate signals

Ed Hawkins^{1,*} and Rowan Sutton¹

¹NCAS-Climate, Department of Meteorology, University of Reading, UK.

Supplementary Information

Time of Emergence for each GCM

Fig. S1 shows the Time of Emergence (ToE) for $S/N > 2$ for each of the GCMs considered, vividly demonstrating the broad range in this quantity.

Sensitivity to analysis grid

Fig. S2 shows the ToE for the northern hemisphere on different spatial grids for two half-year means. The differences between the grids is small, but the short emergence times for the pole are clear in this projection, especially in boreal winter.

The spatial regridding was performed from the native model grids before analysis using the *remapcon* routine from the *cdo* library.

Sensitivity to choice of signal and noise

The distributions of ToE for the three regions discussed in the main paper are shown in Fig. S3. The blue bars are using each GCM's own estimate of both signal and variability, whereas the green bars use the median estimate of the signal. This is the companion figure to Fig. 4 of the main paper, which used the median estimate of the noise instead of the signal. These results demonstrate that the intermodel spread in both signal and noise contribute significantly to the spread in ToE.

Temperature of Emergence

In Fig. S4 we show histograms of the global temperature change at which the $S/N > 2$ threshold is expected to be crossed for the three regions, as a direct comparison with Fig. 5. The emissions uncertainty is reduced in this metric, which may be more relevant for mitigation purposes.

*To whom correspondence should be addressed. Email: e.hawkins@reading.ac.uk

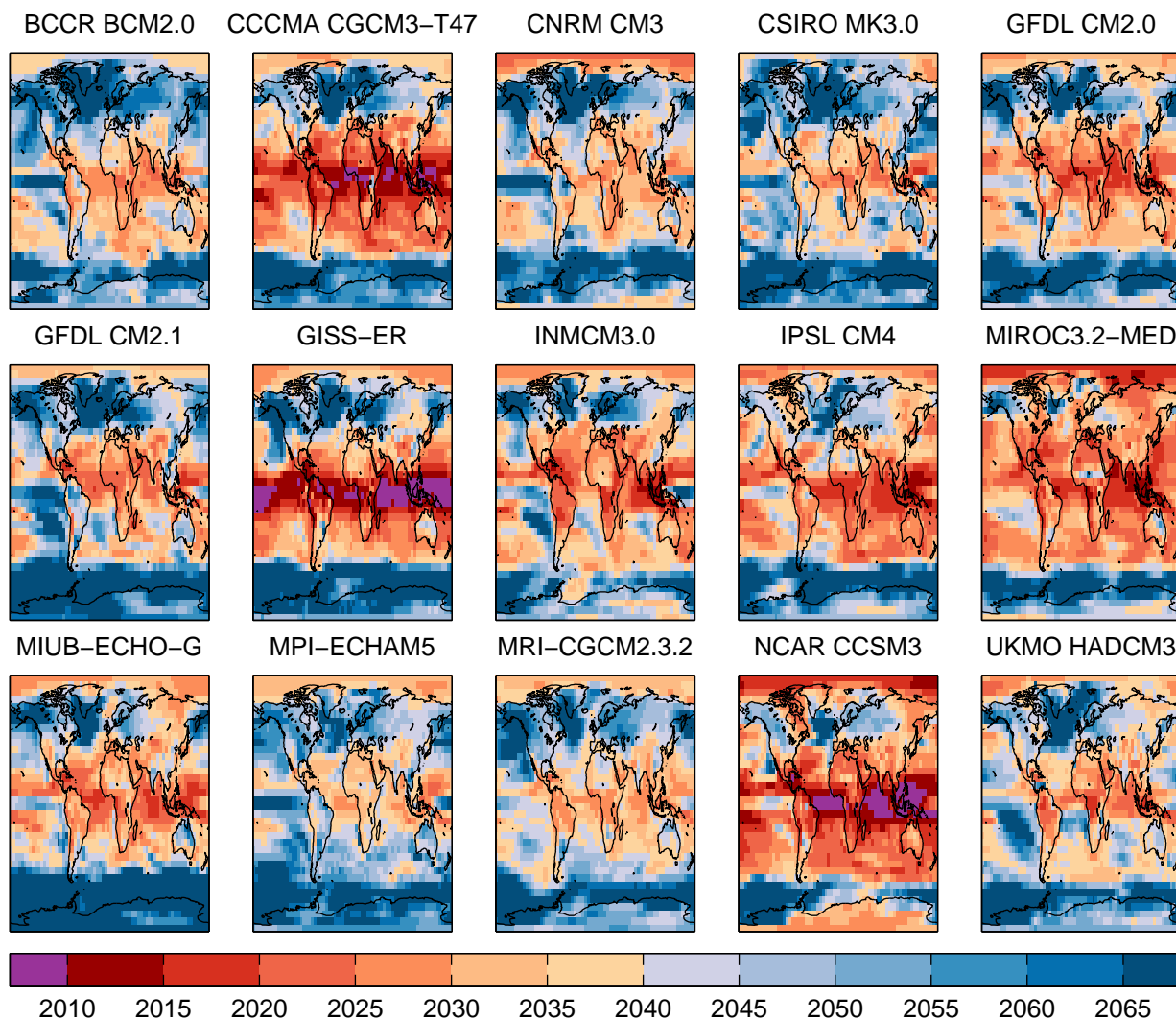


Figure S1: Time of emergence for surface air temperatures. First year when temperature has $S/N > 2$ for an annual mean for each GCM.

ToE GRID COMPARISON

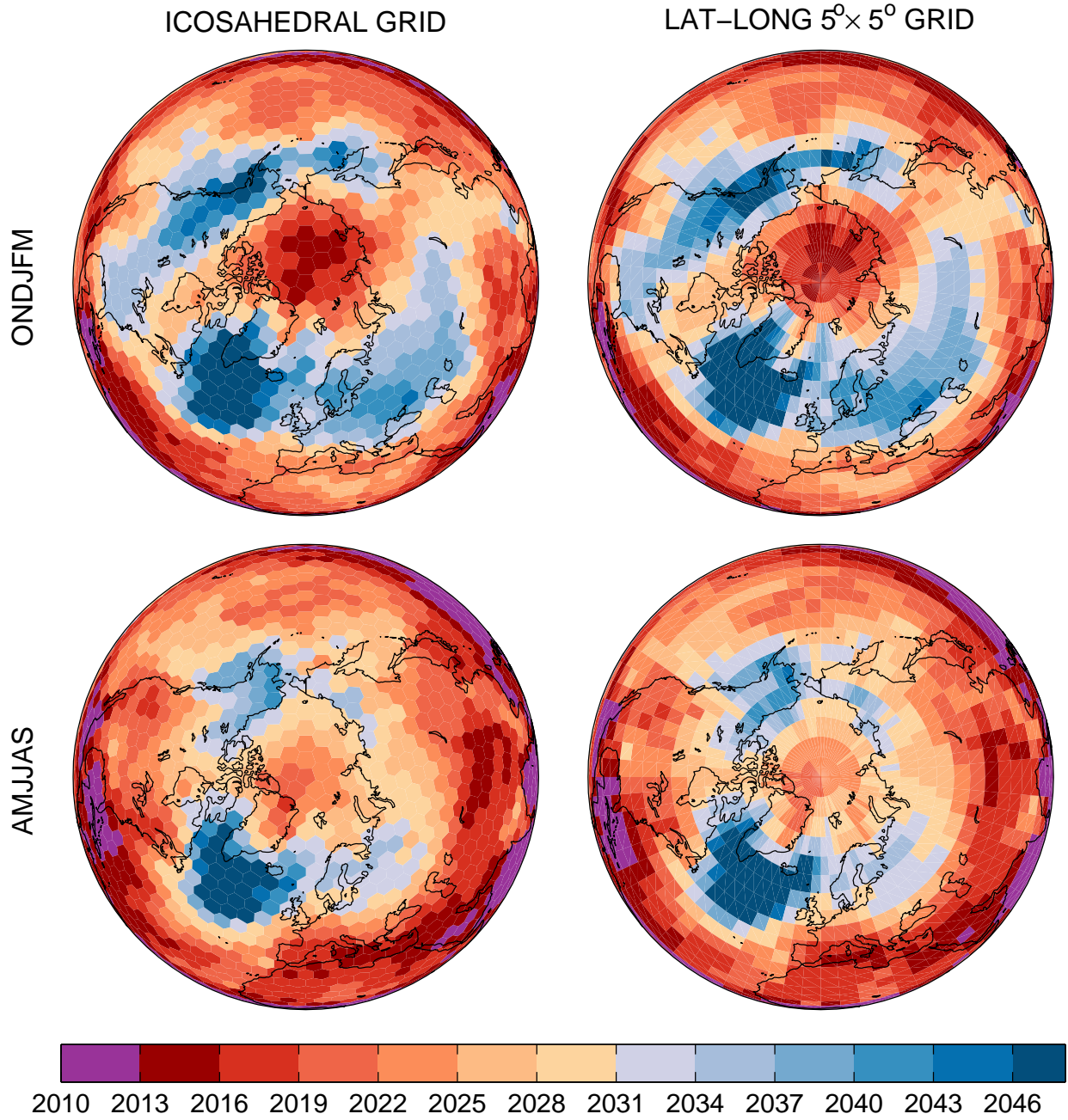


Figure S2: The sensitivity of the median ToE for surface air temperatures to the style of grid used for analysis, focussing on the Arctic, using $S/N > 1$. Note the different colour scale from the main paper to highlight the nearer term in more detail.

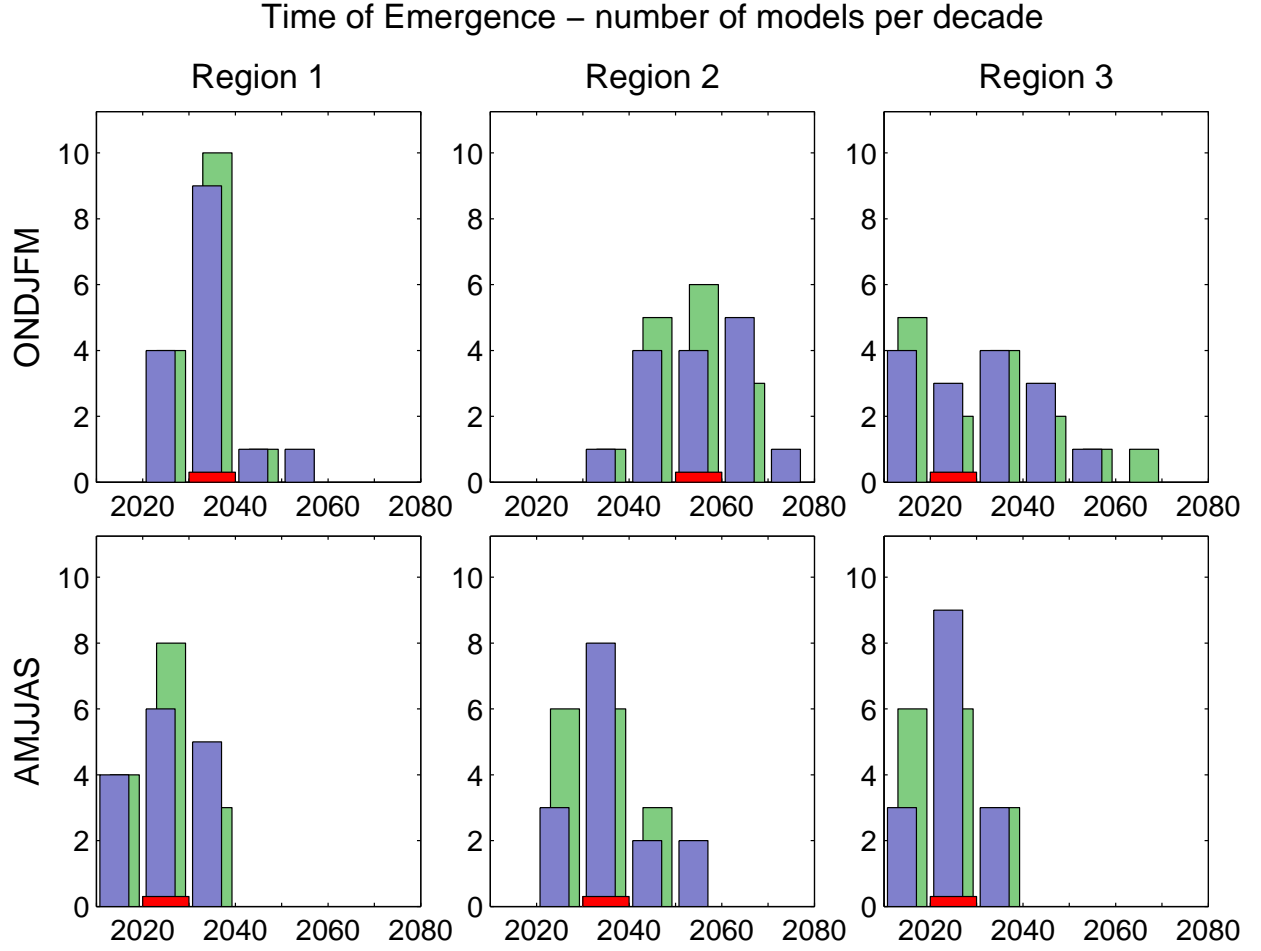


Figure S3: Time of emergence for surface air temperatures for three equal area regions as indicated in Fig. 4. October-March (top) and April-September (bottom). Histograms representing the number of models of when the first year temperature has expected $S/N > 2$, using each GCM's own variability (blue) and median signal (green).

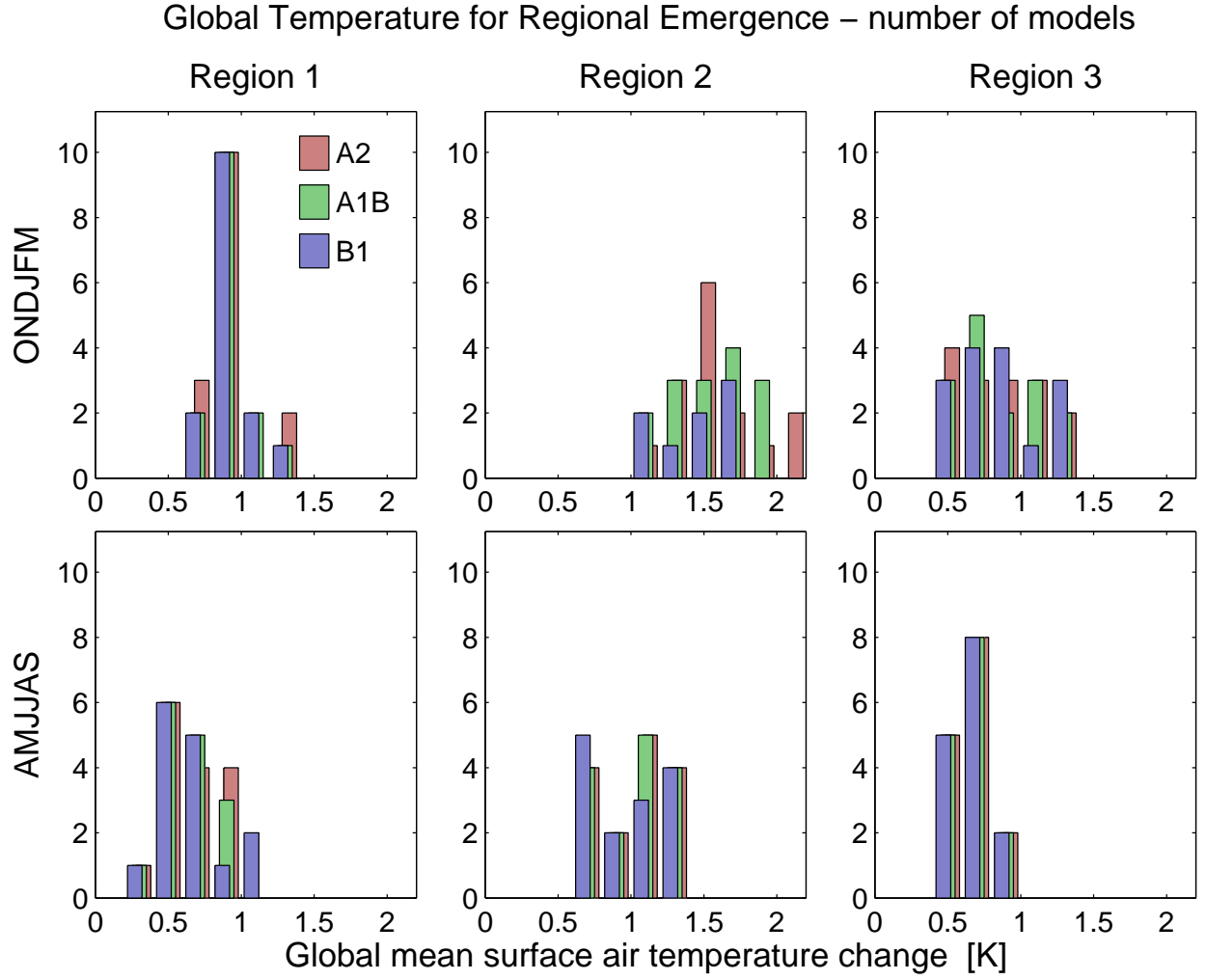


Figure S4: As Fig. 5, but for the global temperature for regional emergence (GTRE). Histograms represent the number of models for each global temperature which have an expected $S/N > 2$. Note that models which cross the threshold after 2100 are not included.