COMMENTARY:

# Overestimated global warming over the past 20 years

John C. Fyfe, Nathan P. Gillett and Francis W. Zwiers

**Recent observed global warming is significantly less than that simulated by climate models. This difference might be explained by some combination of errors in external forcing, model response and internal climate variability.**

Global mean surface temperature over the past 20 years (1993–2012) rose at a rate of 0.14 ± 0.06 °C per decade (95% confidence interval)[1]. This rate of warming is significantly slower than that simulated by the climate models participating in Phase 5 of the Coupled Model Intercomparison Project (CMIP5). To illustrate this, we considered trends in global mean surface temperature computed from 117 simulations of the climate by 37 CMIP5 models (see Supplementary Information). These models generally simulate natural variability — including that associated with the El Niño–Southern Oscillation and explosive volcanic eruptions — as well as estimate the combined response of climate to changes in greenhouse gas concentrations, aerosol abundance (of sulphate, black carbon and organic carbon, for example), ozone concentrations (tropospheric and stratospheric), land use (for example, deforestation) and solar variability. By averaging simulated temperatures only at locations where corresponding observations exist, we find an average simulated rise in global mean surface temperature of 0.30 ± 0.02 °C per decade (using 95% confidence intervals on the model average). The observed rate of warming given above is less than half of this simulated rate, and only a few simulations provide warming trends within the range of observational uncertainty (Fig. 1a).

The inconsistency between observed and simulated global warming is even more striking for temperature trends computed over the past fifteen years (1998–2012). For this period, the observed trend of 0.05 ± 0.08 °C per decade is more than four times smaller than the average simulated trend of 0.21 ± 0.03 °C per decade (Fig. 1b). It is worth noting that the observed trend over this period — not significantly
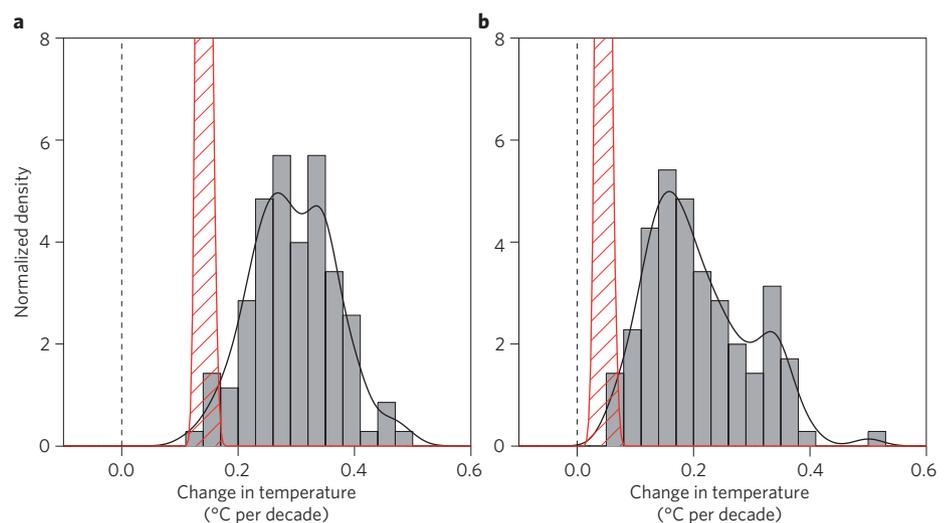


**Figure 1 |** Trends in global mean surface temperature. **a**, 1993–2012. **b**, 1998–2012. Histograms of observed trends (red hatching) are from 100 reconstructions of the HadCRUT4 dataset[1]. Histograms of model trends (grey bars) are based on 117 simulations of the models, and black curves are smoothed versions of the model trends. The ranges of observed trends reflect observational uncertainty, whereas the ranges of model trends reflect forcing uncertainty, as well as differences in individual model responses to external forcings and uncertainty arising from internal climate variability.

different from zero — suggests a temporary 'hiatus' in global warming[2–4]. The divergence between observed and CMIP5-simulated global warming begins in the early 1990s, as can be seen when comparing observed and simulated running trends from 1970–2012 (Fig. 2a and 2b for 20-year and 15-year running trends, respectively).

The evidence, therefore, indicates that the current generation of climate models (when run as a group, with the CMIP5 prescribed forcings) do not reproduce the observed global warming over the past 20 years, or the slowdown in global warming over the past fifteen years. This interpretation is supported by statistical tests of the null hypothesis that the observed and model mean trends are equal,

assuming that either: (1) the models are exchangeable with each other (that is, the 'truth plus error' view); or (2) the models are exchangeable with each other and with the observations (see Supplementary Information). Differences between observed and simulated 20-year trends have *p* values (Supplementary Information) that drop to close to zero by 1993–2012 under assumption (1) and to 0.04 under assumption (2) (Fig. 2c). Here we note that the smaller the *p* value is, the stronger the evidence against the null hypothesis. On this basis, the rarity of the 1993–2012 trend difference under assumption (1) is obvious. Under assumption (2), this implies that such an inconsistency is only expected to occur by chance once in 500 years, if
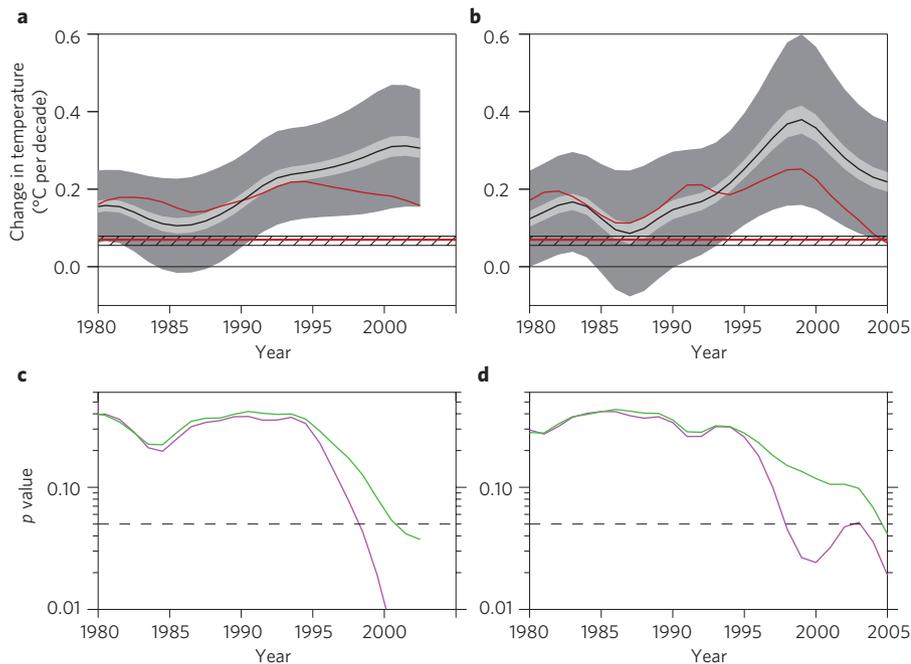
**Figure 2 |** Global mean surface temperature trends and *p* values. **a,b**, 20-year (**a**) and 15-year (**b**) running trends. Black curves are ensemble-averaged trends over the 37 sets of model simulations. Dark-grey shading indicates the 2.5–97.5% ranges of the simulated estimates. Light-grey shading shows the 95% uncertainty ranges of the ensemble means, derived by dividing the 2.5–97.5% ranges by the square root of the number of models. Red curves are the observed trends averaged over 100 realizations and the horizontal red lines show the observed 1900–2012 trends averaged over 100 realizations. Black cross-hatchings are the 95% uncertainty ranges for simulated 1900–2012 ensemble mean trends. Note that the observed and simulated long-term trends are very similar to one another, and are smaller than the short-term trends. **c,d**, 20-year (**c**) and 15-year (**d**) *p* values on trend differences between the simulations and observations for assumption (1) (purple curves), or assumption (2) (green curves). The horizontal dashed lines indicate the threshold below which we reject the null hypothesis.
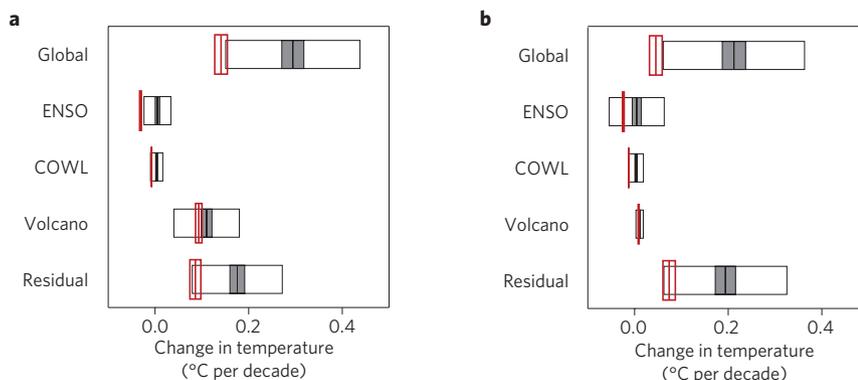


**Figure 3 |** Trends in global mean surface temperature and in associated natural and residual time series. **a**, 1993–2012. **b**, 1998–2012. The 2.5–97.5% ranges for observed estimates are shown by the red boxes. The 2.5–97.5% ranges for simulated estimates are represented by the open black boxes, with the 95% ranges on ensemble mean trends indicated by grey shading. The estimated natural signals shown are associated with the El Niño-Southern Oscillation (ENSO), dynamically induced atmospheric variability (cold ocean–warm Earth; COWL) and major explosive volcanic eruptions (Volcano). Trends in global mean surface temperature without these estimated natural signals are shown at the bottom (Residual).

variation might combine differently in observations than in models. For example, the forced trends in models are modulated up and down by simulated sequences of ENSO events, which are not expected to coincide with the observed sequence of such events. For this reason the moderating influence on global warming that arises from the decay of the 1998 El Niño event does not occur in the models at that time. Thus we employ here an established technique to estimate the impact of ENSO on global mean temperature, and to incorporate the effects of dynamically induced atmospheric variability and major explosive volcanic eruptions[5,6]. Although these three natural variations account for some differences between simulated and observed global warming, these differences do not substantively change our conclusion that observed and simulated global warming are not in agreement over the past two decades (Fig. 3). Another source of internal climate variability that may contribute to the inconsistency is the Atlantic multidecadal oscillation[7] (AMO). However, this is difficult to assess as the observed and simulated variations in global temperature that are associated with the AMO seem to be dominated by a large and concurrent signal of presumed anthropogenic origin (Supplementary Fig. S1). It is worth noting that in any case the AMO has not driven cooling over the past 20 years.

Another possible driver of the difference between observed and simulated global warming is increasing stratospheric aerosol concentrations. Results from several independent datasets show that stratospheric aerosol abundance has increased since the late 1990s, owing to a series of comparatively small tropical volcanic eruptions[8]. Although none of the CMIP5 simulations take this into account, two independent sets of model simulations estimate that increasing stratospheric aerosols have had a surface cooling impact of about 0.07 °C per decade since 1998[8,9]. If the CMIP5 models had accounted for increasing stratospheric aerosol, and had responded with the same surface cooling impact, the simulations and observations would be in closer agreement. Other factors that contribute to the discrepancy could include a missing decrease in stratospheric water vapour[10] (whose processes are not well represented in current climate models), errors in aerosol forcing in the CMIP5 models, a bias in the prescribed solar irradiance trend, the possibility that the transient climate sensitivity of the CMIP5 models could be on average too high[11,12] or a possible unusual episode of

20-year periods are considered statistically independent. Similar results apply to trends for 1998–2012 (Fig. 2d). In conclusion, we reject the null hypothesis that the observed

and model mean trends are equal at the 10% level.

One possible explanation for the discrepancy is that forced and internal

internal climate variability not considered above[13,14]. Ultimately the causes of this inconsistency will only be understood after careful comparison of simulated internal climate variability and climate model forcings with observations from the past two decades, and by waiting to see how global temperature responds over the coming decades. ❑

*John C. Fyfe[1]\*, Nathan P. Gillett[1] and Francis W. Zwiers[2] are at [1]Canadian Centre for Climate Modelling and Analysis, Environment Canada, Victoria, British Columbia V8W 2Y2, Canada, [2]Pacific Climate Impacts Consortium, University of Victoria, Victoria, British Columbia V8W 3R4, Canada.*
*\*e-mail: john.fyfe@ec.gc.ca*

### References
1. Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. *J. Geophys. Res.* **117**, D08101 (2012).
2. Easterling, D. R. & Wehner, M. F. *Geophys. Res. Lett.* **36**, L08706 (2009).
3. Knight, J. *et al. Bull. Am. Meteorol. Soc.* **90**, S75–S79 (2009).
4. Fyfe, J. C. *et al. Geophys. Res. Lett.* **38**, L22801 (2011).
5. Thompson, W. J., Wallace, J. M., Jones, P. D. & Kennedy, J. J. *J. Clim.* **22**, 6120–6141 (2009).
6. Fyfe, J. C., Gillett, N. P. & Thompson, D. W. J. *Geophys. Res. Lett.* **37**, L16802 (2010).
7. Schlesinger, M. E. & Ramankutty, N. *Nature* **367**, 723–726 (1994).
8. Solomon, S. *et al. Science* **333**, 866–870 (2011).
9. Fyfe, J. C., von Salzen, K., Cole, J. N. S., Gillett, N. P. & Vernier, J-P. *Geophys. Res. Lett.* **40**, 584–588 (2013).
10. Solomon, S. *et al. Science* **327**, 1219–1223 (2010).
11. Stott, P., Good, P., Jones, G., Gillett, N. & Hawkins, E. *Environ. Res. Lett.* **8**, 014024 (2013).
12. Otto, A. *et al. Nature Geosci.* **6**, 415–416 (2013).
13. Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A. & Trenberth, K. E. *Nature Clim. Change* **1**, 360–364 (2011).
14. Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I. & Asif, M. *Nature Clim. Change* **3**, 649–653 (2013).

# COMMENTARY:

# Uncertainty analysis in climate change assessments

## Richard W. Katz, Peter F. Craigmile, Peter Guttorp, Murali Haran, Bruno Sansó and Michael L. Stein

**Use of state-of-the-art statistical methods could substantially improve the quantification of uncertainty in assessments of climate change.**

Because the climate system is so complex, involving nonlinear coupling of the atmosphere and ocean, there will always be uncertainties in assessments and projections of climate change. This makes it hard to predict how the intensity of tropical cyclones will change as the climate warms, the rate of sea-level rise over the next century or the prevalence and severity of future droughts and floods, to give just a few well-known examples. Indeed, much of the disagreement about the policy implications of climate change revolves around a lack of certainty. The forthcoming Intergovernmental Panel on Climate Change (IPCC) Fifth Assessment Report (AR5) and the US National Climate Assessment Report will not adequately address this issue. Worse still, prevailing techniques for quantifying the uncertainties that are inherent in observed climate trends and projections of climate change are out of date by well over a decade. Modern statistical methods and models could improve this situation dramatically.

Uncertainty quantification is a critical component in the description and attribution of climate change. In some circumstances, uncertainty can increase when previously neglected sources of uncertainty are recognized and accounted for (Fig. 1 shows how uncertainty can increase for projections of sea-level rise). In other circumstances, more rigorous quantification may result in a decrease in the apparent level of uncertainty, in part because of more efficient use of the available information. For example, despite much effort over recent decades, the uncertainty in the estimated climate sensitivity (that is, the long-term response of global mean temperature to a doubling of the $CO_2$ concentration in the atmosphere) has not noticeably decreased[1]. Nevertheless, policymakers need more accurate uncertainty estimates to make better decisions[2].

Detailed guidance provided to authors of the IPCC AR5 and the US National Climate Assessment Report emphasizes the use of consistent terminology for describing uncertainty for risk communication. This includes a formal definition of terms such as 'likely' or 'unlikely' but, oddly, little advice is given about what statistical techniques should be adopted for uncertainty analysis[3,4]. At the least, more effort could be made to encourage authors to make use of modern techniques.

Historically, several compelling examples exist in which the development and application of innovative statistical methods resulted in breakthroughs in the understanding of the climate system (for example, Sir Gilbert Walker's research in the early twentieth century related to the El Niño–Southern Oscillation phenomenon[5]). We anticipate that similar success stories can be achieved for quantification of uncertainty in climate change.

Although climate observations and climate model output have different sources of error, both exhibit substantial spatial and temporal dependence. Hierarchical statistical models can capture these features in a more realistic manner[6]. These models adopt a 'divide and conquer' approach, breaking the problem into several layers of conceptually and computationally simpler conditional statistical models. The combination of these components produces an unconditional statistical model, whose structure can be quite complex and realistic.