Comments on “A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions”

PAUL J. NORTHROP *

University College London, UK

*Corresponding author address: Paul Northrop, Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT.
E-mail: p.northrop@ucl.ac.uk
ABSTRACT

In a recent article Yip et al. present a simple, yet powerful, framework, for partitioning uncertainty in multimodel climate ensembles. Simplicity is achieved by extracting from climate prediction ensembles a single subset with a convenient structure. We extend Yip et al.’s analysis of predictions of 21st century global temperature by analysing all possible subsets. Findings can depend strongly on the particular dataset chosen, so relying on a single dataset may be misleading.
Yip et al. (2011) (hereafter YFSH11) use analysis of variance (ANOVA) to partition uncertainty in multimodel climate ensembles, illustrating their framework using predictions of twenty-first century global temperature from the Coupled Model Intercomparison Project phase 3 (CMIP3) archive. They extract from the data a subset with a convenient structure: 7 models, each with 2 ensemble runs under each of the emissions scenarios A1B, A2 and B1. Having equal numbers of runs per model-scenario combination means that the ANOVA decomposition of variability is straightforward.

The subset of data used by YFSH11 is one of many possibilities. In the CMIP3 archive these 7 models have simulated 5, 3, 3, 3, 5, 4, and 4 ensemble runs under all three scenarios. Thus, there are 97,200 datasets with the desired structure. (YFSH11’s use of 2 ensemble runs suggests that they had fewer runs available.) We examine how greatly the choice of runs affects findings by repeating YFSH11’s analysis for 2046–2055 (approximately when they found that scenario uncertainty overtakes model uncertainty) for each of the 97,200 datasets. Figure 1 compares the estimated model uncertainty $M$, scenario uncertainty $S$, model-scenario interaction $I$ and internal variability $V$, expressed as percentages of total uncertainty (c.f. figure 4 of YFSH11).

The plot of $I$ against $V$ has 6 groups of points, corresponding to the inclusion/exclusion of individual predictions from the PCM model. Under scenario B1 the PCM runs (temperature anomalies of 1.52, 0.91, 0.76 and 1.24 K) exhibit unusually large variability: there are two large values and two small values. In the ‘2 large PCM B1 runs’ group the estimated percentages attributable to $I$ are much larger, and to $V$ are slightly larger, than in the ‘2 small PCM B1 runs’ group. Correspondingly, the percentage contributions from $M$ and $S$ tend to be smaller for datasets in which both large PCM B1 runs appear. The findings are influenced strongly by which PCM runs are included. Moreover, if the PCM runs are fixed, the plot of $M$ vs. $S$ shows that the choice of the other runs has an appreciable effect on the findings. In the ‘2 small PCM B1 runs’ group (which contains YFSH11’s dataset) the relative magnitudes of $M$ and $S$ vary quite considerably.
YFSH11’s framework can quantify the effect on findings of particular runs in the dataset. If the initial choice of runs to include in the dataset is arbitrary, it is also important to check that findings are not sensitive to this choice.
REFERENCES

List of Figures

1 Estimates of $I$ vs. $V$ (top) and $M$ vs. $S$ (bottom), for 2046–2055. Each point represents one of the possible 97,200 datasets. A different shade of grey is used for each of the 6 ways of choosing 2 runs from the 4 PCM runs.
Fig. 1. Estimates of $I$ vs. $V$ (top) and $M$ vs. $S$ (bottom), for 2046–2055. Each point represents one of the possible 97,200 datasets. A different shade of grey is used for each of the 6 ways of choosing 2 runs from the 4 PCM runs.