

A Simple, Coherent Framework for Partitioning Uncertainty in Climate Predictions

STAN YIP ^{*} , CHRISTOPHER A. T. FERRO AND DAVID B. STEPHENSON

National Centre for Atmospheric Science, Exeter Climate Systems, University of Exeter, United Kingdom

ED HAWKINS

National Centre for Atmospheric Science, Department of Meteorology, University of Reading, United Kingdom.

^{*} *Corresponding author address:* Stan Yip, Harrison Building, University of Exeter, Exeter, Devon, EX4 4QF, UK.

E-mail: c.y.yip@exeter.ac.uk

ABSTRACT

A simple and coherent framework for partitioning uncertainty in multi-model climate ensembles is presented. The analysis of variance (ANOVA) is used to decompose a measure of total variation additively into scenario uncertainty, model uncertainty and internal variability. This approach requires fewer assumptions than existing methods and can be easily used to quantify uncertainty related to model-scenario interaction - the contribution to model uncertainty arising from the variation across scenarios of model deviations from the ensemble mean. Uncertainty in global mean surface air temperature is quantified as a function of lead time for a subset of the Coupled Model Intercomparison Project phase 3 ensemble and results largely agree with those published by other authors: scenario uncertainty dominates beyond 2050 and internal variability remains approximately constant over the 21st century. Both elements of model uncertainty, due to scenario-independent and scenario-dependent deviations from the ensemble mean, are found to increase with time. Estimates of model deviations that arise as by-products of the framework reveal significant differences between models that could lead to a deeper understanding of the sources of uncertainty in multi-model ensembles. For example, three models are shown diverging pattern over the 21st century, while another model exhibits an unusually large variation among its scenario-dependent deviations.

1. Introduction

Uncertainty in climate change prediction arises from three different sources. These are *model uncertainty*, *scenario uncertainty* and *internal variability*. *Model uncertainty* arises due to an incomplete understanding of the physical processes and the limitations of imple-

mentation of the understanding. *Scenario uncertainty* arises due to incomplete information about future emissions. *Internal variability* is the natural unforced fluctuation of the climate system. Internal variability is aleatoric and cannot be reduced by improvement of our scientific knowledge. However, Smith et al. (2007) demonstrated that a proper initialisation of climate predictions with observations can reduce the uncertainty for the next decade. An obvious and exploratory way to evaluate the total of these uncertainties is to calculate the spread of a multi-model ensemble. However, further statistical analysis is needed to quantify the contributions of particular sources of uncertainty, and to describe how a particular model reacts to a particular emissions scenario.

Various methods to decompose the total uncertainty into its sources have been suggested in climate science. Cox and Stephenson (2007) propose a conceptual framework for this purpose and illustrate its use with a single energy-balance model. Hawkins and Sutton (2009, hereafter HS09) and Hawkins and Sutton (2010) fit polynomial trend models over time and calculate various sources of uncertainty. These studies offer simple interpretations of uncertainty but the drawback is that the total uncertainty cannot be easily interpreted.

We use ANOVA to decompose sources of uncertainty (for a complete review, see Von Storch and Zwiers (2001), Chapter 9). ANOVA is a model-based approach that partitions the total variance into components due to different sources of variation, allowing a fuller interpretation. The seminal work of Madden (1976) suggests an ANOVA approach to test for the likelihood of potentially predictable long-range variability. Formal ANOVA models are used in Zwiers (1987) and Zwiers (1996) for analysing seasonal observations and an ensemble of climate simulations respectively, and discussed from a statistical point of view. Several previous papers use ANOVA extensively for evaluating model uncertainty from ensembles.

Räisänen (2001) uses ANOVA to divide the surface air temperature (SAT), precipitation and sea level pressure change into a common signal and variances associated with internal variability and model differences under the same forcing scenario. Hingray et al. (2007) use ANOVA to estimate the uncertainty in temperature and precipitation for a collection of atmosphere-ocean general circulation models (AOGCMs).

The remainder of this paper is organised as follows. Section 2 describes the available data and compares the methods of HS09 and ANOVA. We illustrate the results in Section 3 and present our discussion and conclusions in Section 4.

2. Data and methodologies

a. CMIP3 data

Following HS09, we illustrate the methodology by applying it to global, decadal mean SAT multi-model ensemble predictions for years 2001-2099 from the Coupled Model Inter-comparison Project (CMIP3) archive. The multi-model ensemble data is extracted from original monthly scale data to decadal scale using 10 years moving averaging (Table 1). The predictions are from $N_m = 7$ global climate models (GCMs) under $N_s = 3$ different future emissions scenarios (SRES A1B, A2 and B1) with $N_r = 2$ initial condition ensemble members for each model and scenario. This gives a total of $N_m \times N_s \times N_r = 42$ predictions for 2001-2099. These future scenarios are summarised in Solomon *et al.* (2007, chapter 10). The reason of using fewer GCMs than other studies such as Boer (2009) and Hawkins and Sutton (2009) is because only 7 GCMs have simulated all three scenarios with at least 2 ensemble

runs. Figure 1 shows time-series of all the simulation runs of all the models, scenarios and replicates. The total uncertainty increases with time since the simulation runs diverge with time.

b. Methodologies used in Previous Studies

The methodology in HS09 considers only one realisation per model per scenario ($N_r = 1$). Each prediction of SAT is fitted using a fourth-order polynomial model over years 1950 to 2099. The raw predictions (X) for each model (m), scenario (s) and year (t) are written as,

$$X(m, s, t) = z(m, s, t) + \mu_{\text{ref}}(m, s) + \epsilon(m, s, t), \quad (1)$$

where a reference temperature for each model-scenario combination is denoted by μ_{ref} , the polynomial fit of the projected change in global mean temperature is represented by z and the regression error (internal variability) is ϵ . The reference temperature used is the 1971-2000 mean for each model and scenario. The internal variability estimator is the multi-model mean of variance of the regression error $\epsilon(m, s, t)$,

$$V_{\text{HS}} = \frac{1}{N_m N_s T} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} \sum_{t=1}^T (\epsilon(m, s, t))^2. \quad (2)$$

The internal variability is considered to have constant variance in time. The model uncertainty estimator is the multi-scenario mean of inter-model variance of $z(m, s, t)$,

$$M_{\text{HS}}(t) = \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [z(m, s, t) - z(\cdot, s, t)]^2, \quad (3)$$

where $z(\cdot, s, t) = \sum_m z(m, s, t)/N_m$. The scenario uncertainty estimator is the variance of multi-model means of $z(m, s, t)$

$$S_{\text{HS}}(t) = \frac{1}{N_s} \sum_{s=1}^{N_s} [z(\cdot, s, t) - z(\cdot, \cdot, t)]^2, \quad (4)$$

where $z(\cdot, \cdot, t) = \sum_{m,s} z(m, s, t)/(N_m N_s)$. The sum of these sources of uncertainty is then defined to be the *total uncertainty*,

$$T_{\text{HS}}(t) = V_{\text{HS}} + M_{\text{HS}}(t) + S_{\text{HS}}(t). \quad (5)$$

Cox and Stephenson (2007) define the fractional uncertainty (noise-to-signal ratio) at time t to be

$$F(t) = \frac{1.65\sqrt{T_{\text{HS}}(t)}}{z(\cdot, \cdot, t)}. \quad (6)$$

HS09 also consider the fraction of variance defined as,

$$V_{\text{HS}}/T_{\text{HS}}(t) \quad (7)$$

$$M_{\text{HS}}(t)/T_{\text{HS}}(t) \quad (8)$$

$$S_{\text{HS}}(t)/T_{\text{HS}}(t), \quad (9)$$

where these refer to the fraction of variance of internal variability, model uncertainty and scenario uncertainty respectively.

c. Analysis of Variance Method

We adopt a model-based approach rather than descriptive or algorithmic approach since the use of a statistical model facilitates a coherent interpretation of uncertainty. We fit an ANOVA model on the projected temperature anomalies $x(m, s, r, t)$ for model m , scenario s and replicate r at time t from the 1971-2000 mean. First, we consider the following ANOVA model for global decadal SAT $x(m, s, r, t)$ for each time point t ,

$$x(m, s, r, t) = \mu(t) + \alpha(m, t) + \beta(s, t) + \gamma(m, s, t) + \epsilon(m, s, r, t), \quad (10)$$

where $\mu(t)$ is the overall effect representing the grand ensemble mean of all simulations at time $t = 1, 2, \dots, 99$; $\alpha(m, t)$ is scenario-independent deviation of model $m = 1, 2, \dots, 7$ from the overall ensemble mean $\mu(t)$; $\beta(s, t)$ is scenario deviation of emission scenario $s = 1, 2, 3$; the parameters $\alpha(m, t)$ and $\beta(s, t)$ are collectively called main effects; $\gamma(m, s, t)$ is the interaction term effect between model m and scenario s at time t which describes scenario-dependent deviation. The error term $\epsilon(m, s, r, t)$ is independent and identically distributed.

The notion of interaction is an important concept in ANOVA. Mathematically, interaction is said to occur if the separate effects do not combine additively (de González and Cox, 2007). For climate projections, it arises from how models react differently to emission scenarios which we call scenario-dependent model uncertainty, as opposed to scenario-independent model uncertainty. To demonstrate that there is a potential interaction term, we show the mean response at different lead times to different emission scenarios for each model in Figure 2. The lines are not parallel, indicating that there is an apparent interaction between some models and scenarios especially for long lead times. For example, in the year 2061-2099, the effect of changing emission scenario is different for the CCCMA model than for the MIUB ECHO model, as the lines joining the models cross.

The method of least squares is used for the parameter estimation. Applying constraints $\sum_{m=1}^{N_m} \hat{\alpha}(m, t) = 0$, $\sum_{s=1}^{N_s} \hat{\beta}(s, t) = 0$, $\sum_{m=1}^{N_m} \hat{\gamma}(m, s, t) = 0$, for $s = 1, \dots, N_s$ and $\sum_{s=1}^{N_s} \hat{\gamma}(m, s, t) =$

0, for $m = 1, \dots, N_m$, the parameter estimators $\hat{\mu}(t)$, $\hat{\alpha}(m, t)$, $\hat{\beta}(s, t)$ and $\hat{\gamma}(m, s, t)$ are

$$\hat{\mu}(t) = x(\cdot, \cdot, \cdot, t). \quad (11)$$

$$\hat{\alpha}(m, t) = x(m, \cdot, \cdot, t) - x(\cdot, \cdot, \cdot, t). \quad (12)$$

$$\hat{\beta}(s, t) = x(\cdot, s, \cdot, t) - x(\cdot, \cdot, \cdot, t). \quad (13)$$

$$\hat{\gamma}(m, s, t) = x(m, s, \cdot, t) + x(\cdot, \cdot, \cdot, t) - x(m, \cdot, \cdot, t) - x(\cdot, s, \cdot, t), \quad (14)$$

where $x(\cdot, \cdot, \cdot, t)$ is the overall mean at time t ; $x(m, s, \cdot, t)$ is the mean over all the members at time t for model m and scenario s ; $x(\cdot, s, \cdot, t)$ and $x(m, \cdot, \cdot, t)$ are means over the models, and replicates and the mean over the scenarios, and replicates at time t respectively.

We define all four sources of uncertainty in terms of the notion of variance. The ANOVA approach does not assume constant internal variability over time and is also not restricted to specify any type of trend for models. The internal variability $V(t)$ is the variance of each member around the model-scenario mean,

$$V(t) = \frac{1}{N_m N_s N_r} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} [x(m, s, r, t) - x(m, s, \cdot, t)]^2. \quad (15)$$

The scenario-independent model uncertainty $M(t)$ is the variance of model means around the ensemble mean,

$$M(t) = \frac{1}{N_m} \sum_{m=1}^{N_m} [x(m, \cdot, \cdot, t) - x(\cdot, \cdot, \cdot, t)]^2 = \text{Var}_m [\hat{\alpha}(m, t)]. \quad (16)$$

The scenario uncertainty $S(t)$ is the variance of scenario means around the ensemble mean,

$$S(t) = \frac{1}{N_s} \sum_{s=1}^{N_s} [x(\cdot, s, \cdot, t) - x(\cdot, \cdot, \cdot, t)]^2 = \text{Var}_s [\hat{\beta}(s, t)]. \quad (17)$$

The model-scenario interaction uncertainty $I(t)$ is the variance of model-scenario mean

around the sum of estimated main effects $\mu(t)$, $\alpha(m, t)$ and $\beta(s, t)$,

$$\begin{aligned}
I(t) &= \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} \{x(m, s, \cdot, t) - [x(\cdot, \cdot, \cdot, t) + (x(m, \cdot, \cdot, t) \\
&\quad - x(\cdot, \cdot, \cdot, t)) + (x(\cdot, s, \cdot, t) - x(\cdot, \cdot, \cdot, t))]\}^2 \\
&= \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [x(m, s, \cdot, t) + x(\cdot, \cdot, \cdot, t) - x(m, \cdot, \cdot, t) - x(\cdot, s, \cdot, t)]^2 \\
&= \text{Var}_{m,s}[\hat{\gamma}(m, s, t)].
\end{aligned} \tag{18}$$

The total uncertainty $T(t)$, is simply the variance of the ensembles, ,

$$\begin{aligned}
T(t) &= \frac{1}{N_s N_m N_r} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} [x(m, s, r, t) - x(\cdot, \cdot, \cdot, t)]^2 \\
&= \frac{1}{N_m N_s N_r} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} \sum_{r=1}^{N_r} [x(m, s, r, t) - x(m, s, \cdot, t)]^2 \\
&\quad + \frac{1}{N_m} \sum_{m=1}^{N_m} [x(m, \cdot, \cdot, t) - x(\cdot, \cdot, \cdot, t)]^2 \\
&\quad + \frac{1}{N_s} \sum_{s=1}^{N_s} [x(\cdot, s, \cdot, t) - x(\cdot, \cdot, \cdot, t)]^2 \\
&\quad + \frac{1}{N_s N_m} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [x(m, s, \cdot, t) + x(\cdot, \cdot, \cdot, t) - x(m, \cdot, \cdot, t) - x(\cdot, s, \cdot, t)]^2 \\
&= V(t) + M(t) + S(t) + I(t).
\end{aligned} \tag{19}$$

Since the multi-model ensembles used here are “ensembles of opportunity” (Tebaldi and Knutti, 2007), we do not think of the available models and scenarios as a sample from a wider population of possible models and scenarios. If a wider population could be envisaged from which the ensemble members form a sample then our ANOVA model could be adapted to include so-called random effects that would enable inferences about the population (Eisenhart, 1947). The simulation runs under the same model-scenario is thought as the samples drawn from a finite population. However, it is also appropriate to interpret internal variability as independent realisations from a infinite population. Then our estimation formula $V(t)$ will

be underestimated by 50% when $n = 2$. In that case using random effects model to capture such a setting (e.g. Gelman, 2005) is a possible solution but we retain simplicity by assuming finite population.

d. Connection to the estimates in HS09

The model-scenario interaction can be interpreted as a component of the model uncertainty defined in HS09. Consider the sum of the scenario-independent uncertainty $M(t)$ and the model-scenario interaction variance $I(t)$,

$$\begin{aligned}
& M(t) + I(t) \\
&= \frac{1}{N_m} \sum_{m=1}^{N_m} [x(m, \cdot, \cdot, t) - x(\cdot, \cdot, \cdot, t)]^2 \\
&\quad + \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [x(m, s, \cdot, t) + x(\cdot, \cdot, \cdot, t) - x(\cdot, s, \cdot, t) - x(m, \cdot, \cdot, t)]^2 \\
&= \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [x(m, \cdot, \cdot, t) - x(\cdot, \cdot, \cdot, t) + x(m, s, \cdot, t) + x(\cdot, \cdot, \cdot, t) - x(\cdot, s, \cdot, t) - x(m, \cdot, \cdot, t)]^2 \\
&= \frac{1}{N_m N_s} \sum_{m=1}^{N_m} \sum_{s=1}^{N_s} [x(m, s, \cdot, t) - x(\cdot, s, \cdot, t)]^2. \tag{20}
\end{aligned}$$

The sum of the terms $M(t)$ and $I(t)$ is analogous to how HS09 to define the model uncertainty $M_{HS}(t)$ in (3). The scenario uncertainty $S(t)$ and internal variability $V(t)$ are also similar to the definitions in (2) and (4).

3. Results

We now use the ANOVA approach described above to quantify the uncertainty in global mean, decadal mean SAT in a subset of the CMIP3 climate projections. Figure 3a shows how the different variance components vary with lead time. The scenario uncertainty ($S(t)$, thick dashed line) dominates the total uncertainty after year 2050, which agrees with HS09. Over the whole period, model uncertainty ($M(t)$, thick solid line) is also greater than the internal variability ($V(t)$, thin solid line), which itself is rather constant over time (as assumed by HS09). The model-scenario interaction variance ($I(t)$, thin dashed line) increases from less than 1×10^{-3} to $2.5 \times 10^{-2} K^2$, larger than the internal variability component, demonstrating that interaction is an important component of uncertainty, see Figure 3b.

Figure 4 presents a comparison of the fractions of uncertainty from the methodologies of HS09 and ANOVA. With our ANOVA approach, the scenario uncertainty dominates all the uncertainty after the year 2050. In the ANOVA method, the fraction of variance due to internal variability decreases rapidly in the first few decades, which may be due to a random fluctuation in the internal variability $V(t)$ in the first decade. Meanwhile, scenario uncertainty contributes slightly less in the ANOVA method for about the first 30 years. The model uncertainty dominates in the first few decades and has similar values in HS09 and the ANOVA approach. The model-scenario interaction variance increases from a small value and is saturated at about 5% after two decades.

Figure 5 shows how the various fitted ANOVA parameters evolve over time. In panel 5a, the separation between the model effects (α) is seen to increase with time and some of the models, such as ECHAM5 and PCM, give a larger contribution to the model uncertainty

than others. Systematic changes in mean deviations can be found in some models (like PCM) that contribute more to the model uncertainty than others. Such a pattern may be attributable to the fact that the PCM simulations in the historical period are not continuous with that on the 21st century SRES simulations. In panel 5b, the emission scenario SRES A2 overtakes SRES A1B in the year 2070 and they have a $0.5K$ separation in the year 2100. This separation can be understood as the response to the socio-economic difference in the emission scenario and also how model treat the forcings like aerosol differently. A plot of the interaction term (panel 5c) is helpful for understanding the contribution of model-scenario interaction which increases with lead time.

Figure 6 explores the interaction effect in more detail and shows that the interaction effect varies widely. For example, the variation to the different scenarios is the greatest for CCSM3, and the CCSM3 is relatively cool in SRES B1 and relatively warm in A2.

4. Discussion and conclusions

a. Methodology

We introduce a simple, coherent approach for the modelling of uncertainty in multi-model ensembles. The sources of uncertainty are estimated from the ANOVA model and add up to give total variance, which is a natural measure of global uncertainty. In contrast to the uncertainty decomposition constructed in previous studies, this approach does not need to specify a particular type of trend and noise distribution and does not assume constant internal variability over time. The ANOVA approach is a powerful way to quantify sources

of uncertainty and the results generated are often easy to interpret. It is easy to summarise the structure of all ensemble members under different scenarios with simple exploratory techniques. Another important feature is interaction. In this framework, model-scenario interaction is defined as a form of non-constancy of variance across scenarios in different models. The framework supports the decomposition of model uncertainty into a term that measures the uncertainty due to variation between scenario-independent model deviations and an interaction term that measures the uncertainty due to variation between scenario-dependent model deviations. Ignoring the significant interaction term in the analysis would lead to a dramatic impact on the interpretation of the data. The framework offers, along with some exploratory data analysis techniques, a more detailed interpretation on uncertainty and the contribution from a particular ensemble member.

b. Scientific Interpretation

Our results for uncertainty in the global decadal mean temperature change broadly agree with previous studies but some details are different especially for short lead times. There are some important findings:

- scenario uncertainty, conditional on the choice of scenarios, is of the greatest importance after year 2050.
- internal variability is constant over time but decreases rapidly as a fraction of the total.
- uncertainty from scenario-independent model deviations dominates uncertainty from scenario-dependent model deviations over the whole study period.

- the model-scenario interaction effect is an important contribution to uncertainty, especially at long lead times.

The first finding is fully in agreement with previous papers such as HS09 and Cox and Stephenson (2007). The second finding is an assumption in HS09 and is now validated using the ANOVA framework. The latter two findings are more closely tied to the presence of a significant interaction term. There are several possible reasons for this finding of certain models having large interaction terms. The most likely reason is that the same forcings are treated differently across the range of models (e.g. Kiehl, 2007). However, even if two models treat a forcing in exactly the same way, there could still be a contribution to the interaction if the models respond differently to the forcing, in other words if the models have different effective climate sensitivities.

c. Future Work

The ANOVA approach, because of its simplicity, is a good starting point to cope with some other more complex problems on attributing uncertainty from multi-model ensembles. Apart from global mean temperature, it is also interesting to investigate uncertainty for different space-time scales and other meteorological fields like precipitation and stratospheric ozone (Hawkins and Sutton, 2010 and Charlton-Perez *et al.*, 2010). It is possible to extend this approach into a more general class of models. These extensions are not currently common in the climate science community but have been used extensively in areas such as biology, epidemiology and financial modelling. For example, an obvious extension to climate science is a multivariate ANOVA (MANOVA, see details in Press, 1972, chapter 8) by incorporating

the relationship between different atmospheric fields. This is particularly useful because atmospheric fields are often correlated. A separate analysis of fields like temperature and precipitation may lead to a repeated use of data. For epidemiology applications, Zhang et al. (2009) develop the techniques of smoothed analysis of variance (SANOVA) to smooth spatial random effects by taking advantage of the spatial variation.

Acknowledgments.

S. Yip and E. Hawkins are funded by NERC National Centre for Atmospheric Science (NCAS). We thank Mat Collins and three anonymous reviewers for their thoughtful comments which helped improve the paper. We also thank the modelling groups, the Program for Climate Model Diagnosis and Intercomparison, and the WCRP's Working Group on Coupled Modelling for making CMIP3 multi-model dataset freely available (<https://esg.llnl.gov:8443/index.jsp>). Support of this dataset is provided by the Office of Science, U.S. Department of Energy.

REFERENCES

- Boer, G., 2009: Changes in Interannual Variability and Decadal Potential Predictability under Global Warming. *Journal of Climate*, **22** (11), 3098–3109.
- Charlton-Perez, A., et al., 2010: The potential to narrow uncertainty in projections of stratospheric ozone over the 21st century. *Atmos. Chem. Phys*, **10**, 9473–9486.
- Cox, P. and D. Stephenson, 2007: Climate change: A changing climate for prediction. *Science*, **317** (5835), 207.
- De González, A. and D. Cox, 2007: Interpretation of interaction: A review. *The Annals of Applied Statistics*, **1** (2), 371–385.
- Eisenhart, C., 1947: The assumptions underlying the analysis of variance. *Biometrics*, **3** (1), 1–21.
- Gelman, A., 2005: Analysis of variance: why it is more important than ever. *Annals of Statistics*, **33** (1), 1–31.
- Hawkins, E. and R. Sutton, 2009: The potential to narrow uncertainty in regional climate predictions. *Bulletin of the American Meteorological Society*, **90** (8), 1095–1107.
- Hawkins, E. and R. Sutton, 2010: The potential to narrow uncertainty in projections of regional precipitation change. *To appear in Climate Dynamics*.

- Hingray, B., A. Mezghani, and T. Buishand, 2007: Development of probability distributions for regional climate change from uncertain global mean warming and an uncertain scaling relationship. *Hydrology and Earth System Sciences*, **11 (3)**, 1097–1114.
- Kiehl, J., 2007: Twentieth century climate model response and climate sensitivity. *Geophysical Research Letters*, **34 (22)**, L22 710.
- Madden, R., 1976: Estimates of the natural variability of time-averaged sea-level pressure. *Monthly Weather Review*, **104 (7)**, 942–952.
- Press, S., 1972: *Applied multivariate analysis*. Holt, Rinehart and Winston.
- Räisänen, J., 2001: CO₂-induced climate change in CMIP2 experiments: Quantification of agreement and role of internal variability. *Journal of Climate*, **14 (9)**, 2088–2104.
- Smith, D., S. Cusack, A. Colman, C. Folland, G. Harris, and J. Murphy, 2007: Improved surface temperature prediction for the coming decade from a global climate model. *Science*, **317 (5839)**, 796.
- Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, 2007: *IPCC 2007. Climate Change 2007: the Physical Science Basis Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. New York: Cambridge University Press.
- Tebaldi, C. and R. Knutti, 2007: The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **365 (1857)**, 2053.

Von Storch, H. and F. Zwiers, 2001: *Statistical analysis in climate research*. Cambridge University Press.

Zhang, Y., J. Hodges, and S. Banerjee, 2009: Smoothed ANOVA with spatial effects as a competitor to MCAR in multivariate spatial smoothing. *Annals of Statistics*, **3** (4), 1805–1830.

Zwiers, F., 1987: A potential predictability study conducted with an atmospheric general circulation model. *Monthly Weather Review*, **115** (12), 2957–2974.

Zwiers, F., 1996: Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2. *Climate Dynamics*, **12** (12), 825–847.

List of Tables

- 1 Climate model data used in this study are obtained from the CMIP3 archive for years 2001-2099 for the A1B, B1 and A2 scenarios for GCMs with more than one ensemble member. 19

TABLE 1. Climate model data used in this study are obtained from the CMIP3 archive for years 2001-2099 for the A1B, B1 and A2 scenarios for GCMs with more than one ensemble member.

Acronym	Model
CCCMA	Canadian Centre for Climate Modelling and Analysis Coupled General Circulation Model, version 3.1
MIROC3.2	Model for Interdisciplinary Research on Climate 3.2 (medium resolution)
MIUB ECHO	ECHAM and the global Hamburg Ocean Primitive Equation Model
ECHAM5	ECHAM and the Max Planck Institute Ocean Model
CGCM	Meteorological Research Institute Coupled General Circulation Model, version 2.3.2
CCSM3	National Center for Atmospheric Research Community Climate System Model, version 3
PCM	National Center for Atmospheric Research Parallel Climate Model

List of Figures

- 1 Global, annual mean surface air temperature prediction from 7 different GCMs under three different emission scenarios from 2001 to 2099. Two ensemble members are shown here per model per emission scenario. 22
- 2 Interaction plots show the means of the SAT for the years 2001-2030; 2031-2060; 2061-2099; 2001-2099 against emission scenarios for all the models, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM). An interaction between model and scenario is present where the lines are not parallel. 23
- 3 Uncertainty in global decadal mean surface temperature projections is shown here by variances. Panel (a) shows all four components of uncertainty, i.e. scenario-dependent model uncertainty (thick solid line); scenario uncertainty (thick dashed line); internal variability (thin solid line); model-scenario interaction uncertainty (thin dashed line). Panel (b) shows the contribution of internal variability and the model-scenario interaction effect variance, i.e. internal variability (thin solid line); model-scenario interaction uncertainty (thin dashed line). The internal variability from an ANOVA model with absence of interaction (thick solid line) is also superposed in the same diagram. 24
- 4 Comparison of the fraction of variance for global, decadal mean surface air temperature using, (a) HS09 methodology and (b) the ANOVA-based approach. 25

- 5 Time-series plots of the fitted terms. Panel (a), (b) and (c) are the estimates $\hat{\alpha}(m, t)$, $\hat{\beta}(s, t)$ and $\hat{\gamma}(m, s, t)$ respectively, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM). 26
- 6 Time-series plots of the fitted model-scenario interaction effects $\hat{\gamma}(m, s, t)$. Panel (a), (b) and (c) are emission scenario B1, A1B and A2 respectively, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM). 27

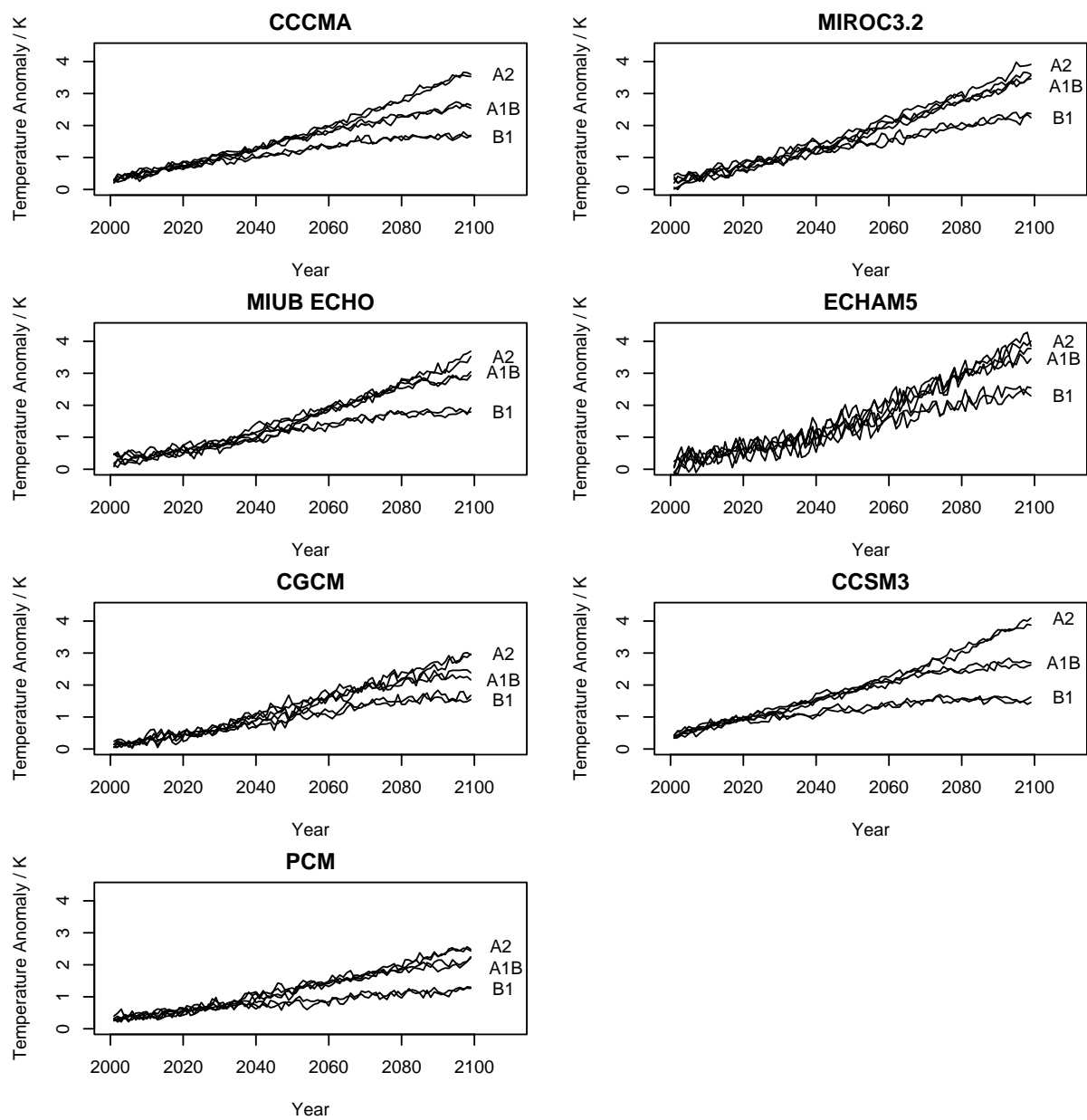


FIG. 1. Global, annual mean surface air temperature prediction from 7 different GCMs under three different emission scenarios from 2001 to 2099. Two ensemble members are shown here per model per emission scenario.

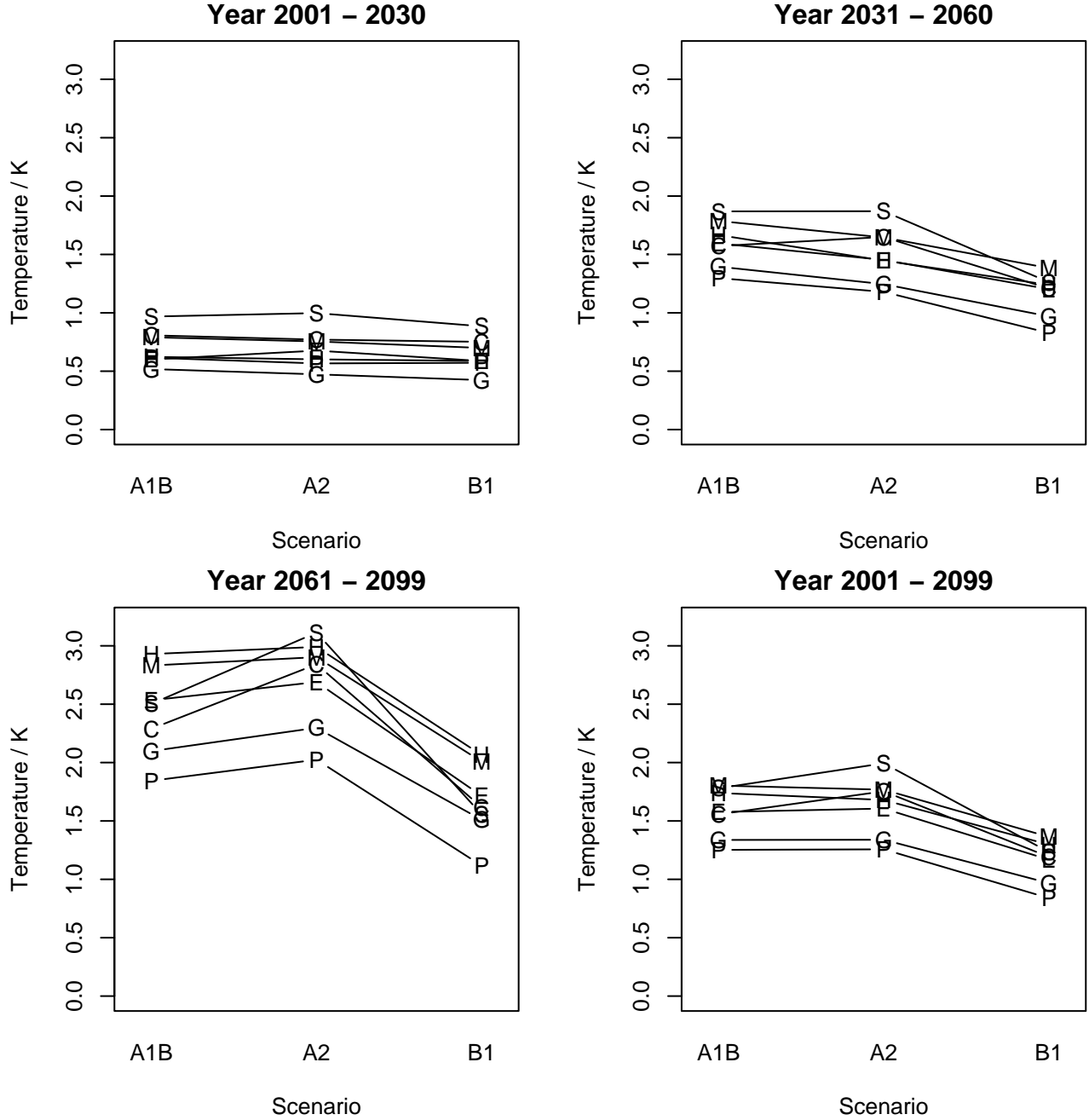


FIG. 2. Interaction plots show the means of the SAT for the years 2001-2030; 2031-2060; 2061-2099; 2001-2099 against emission scenarios for all the models, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM). An interaction between model and scenario is present where the lines are not parallel.

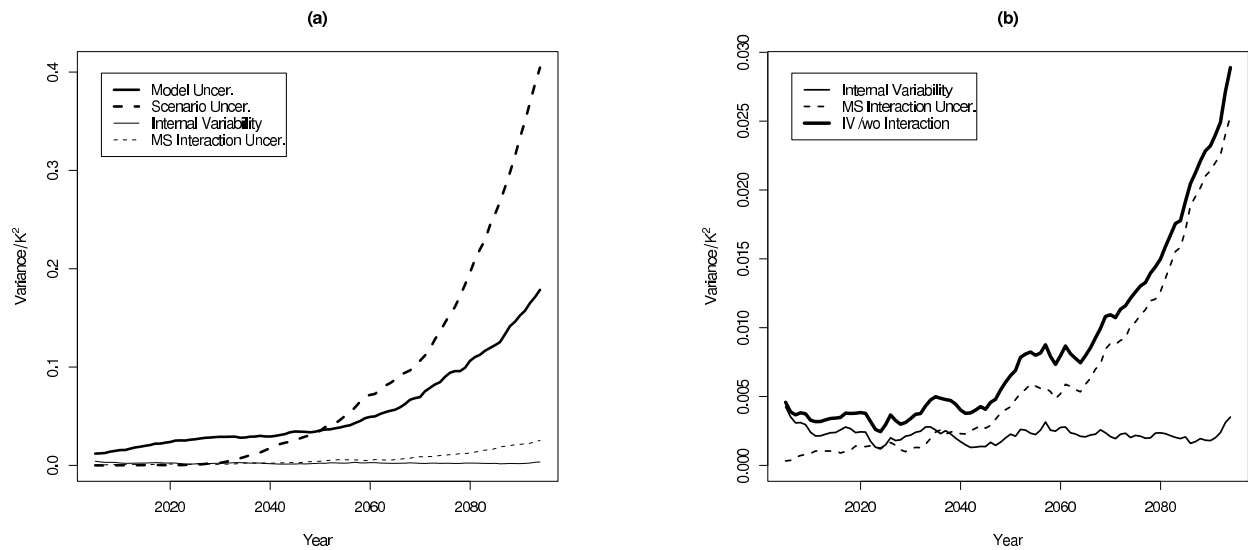


FIG. 3. Uncertainty in global decadal mean surface temperature projections is shown here by variances. Panel (a) shows all four components of uncertainty, i.e. scenario-dependent model uncertainty (thick solid line); scenario uncertainty (thick dashed line); internal variability (thin solid line); model-scenario interaction uncertainty (thin dashed line). Panel (b) shows the contribution of internal variability and the model-scenario interaction effect variance, i.e. internal variability (thin solid line); model-scenario interaction uncertainty (thin dashed line). The internal variability from an ANOVA model with absence of interaction (thick solid line) is also superposed in the same diagram.

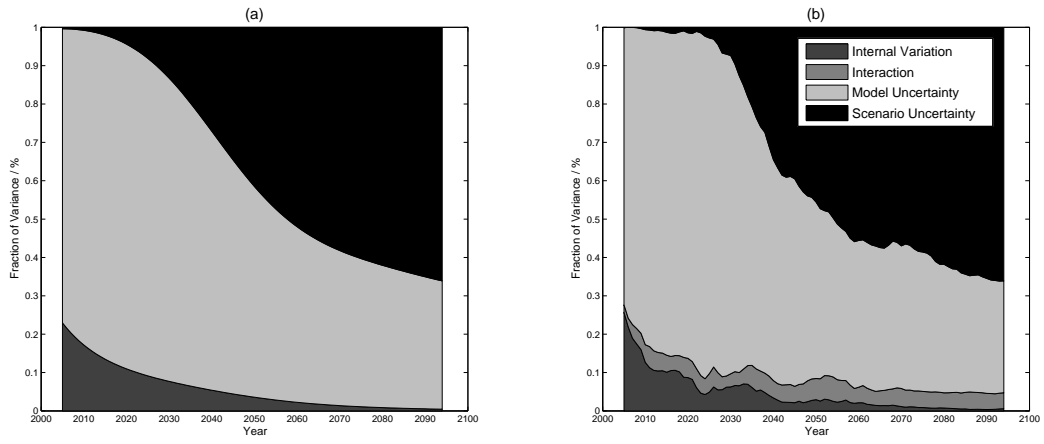


FIG. 4. Comparison of the fraction of variance for global, decadal mean surface air temperature using, (a) HS09 methodology and (b) the ANOVA-based approach.

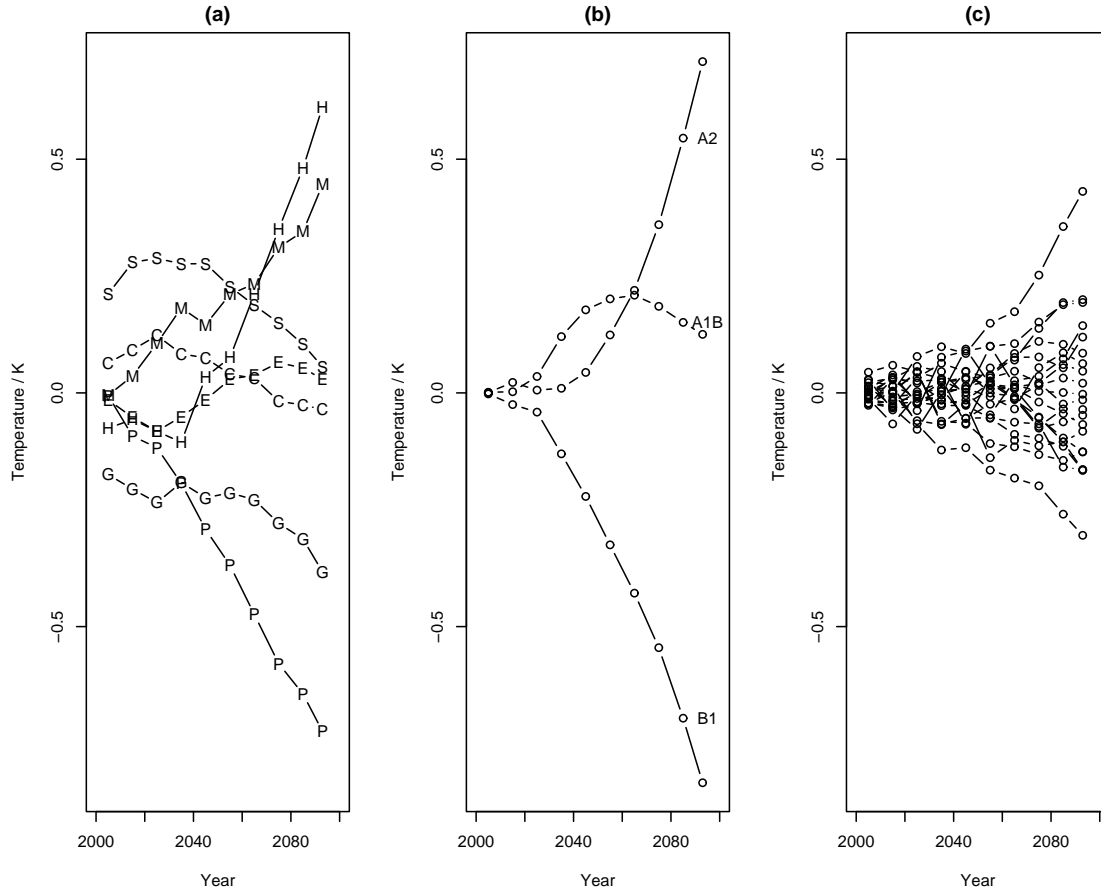


FIG. 5. Time-series plots of the fitted terms. Panel (a), (b) and (c) are the estimates $\hat{\alpha}(m, t)$, $\hat{\beta}(s, t)$ and $\hat{\gamma}(m, s, t)$ respectively, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM).

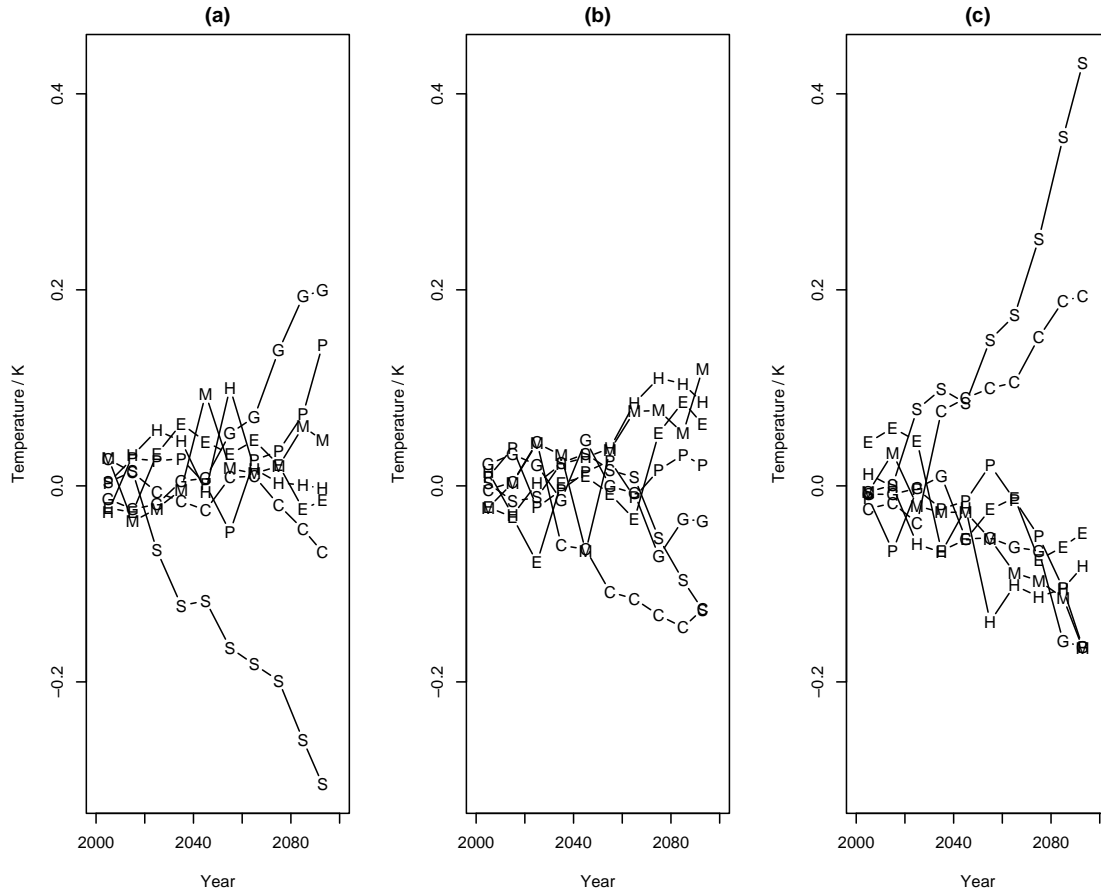


FIG. 6. Time-series plots of the fitted model-scenario interaction effects $\hat{\gamma}(m, s, t)$. Panel (a), (b) and (c) are emission scenario B1, A1B and A2 respectively, and models are represented by symbol C(CCCMA), M(MIROC3.2), E(MIUB ECHO), H(ECHAM5), G(CGCM), S(CCSM3) and P(PCM).