Combinations of Natural and Anthropogenic Forcings in Twentieth-Century Climate

GERALD A. MEEHL, WARREN M. WASHINGTON, CASPAR M. AMMANN, JULIE M. ARBLASTER, T. M. L. WIGLEY, AND CLAUDIA TEBALDI
National Center for Atmospheric Research,* Boulder, Colorado

(Manuscript received 18 July 2003, in final form 16 March 2004)

ABSTRACT

Ensemble simulations are run with a global coupled climate model employing five forcing agents that influence the time evolution of globally averaged surface air temperature during the twentieth century. Two are natural (volcanoes and solar) and the others are anthropogenic [e.g., greenhouse gases (GHGs), ozone (stratospheric and tropospheric), and direct effect of sulfate aerosols]. In addition to the five individual forcing experiments, an additional eight sets are performed with the forcings in various combinations. The late-twentieth-century warming can only be reproduced in the model with anthropogenic forcing (mainly GHGs), while the early twentieth-century warming is mainly caused by natural forcing in the model (mainly solar). However, the signature of globally averaged temperature at any time in the twentieth century is a direct consequence of the sum of the forcings. The similarity of the response to the forcings on decadal and interannual time scales is tested by performing a principal component analysis of the 13 ensemble mean globally averaged temperature time series. A significant portion of the variance of the reconstructed time series can be retained in residual calculations compared to the original single and combined forcing runs. This demonstrates that the statistics of the variances for decadal and interannual time-scale variability in the forced simulations are similar to the response from a residual calculation. That is, the variance statistics of the response of globally averaged temperatures in the forced runs are additive since they can be reproduced in the responses calculated as a residual from other combined forcing runs.

1. Introduction

Previous studies with global coupled models have shown that a specification of the major known forcings that acted on the climate system in the twentieth century [e.g., such as greenhouse gases (GHGs), solar, volcanoes, etc.] can reproduce, to first order, many aspects of the observed time series of globally averaged temperature for that time period (Stott et al. 2000; Meehl et al. 2003; Ammann et al. 2003; Broccoli et al. 2003). It has been assumed in a number of these studies that the response of the climate system to forcing agents is additive such that the response to a combination of forcings is equivalent to the sum of those forcings (e.g., Cubasch et al. 2001). One way to verify that this is correct (or at least a good approximation) is to run a global coupled climate model with single forcings as well as with equivalent combinations of forcings. For example, if AB is the result for A and B together, then one could compare the difference (A + B) − B with A [or (A + B) − A with B]. The better the agreement or similarity, the better is the additivity assumption.

An alternative proposed by Wigley et al. (2004b, manuscript submitted to J. Climate, hereafter WIGb) is to compare A + B − (A + B) with the null forcing (control) case. They address the additivity issue in terms of hemispheric mean temperature trends over the twentieth century in single and combined forcing runs with the Parallel Climate Model (PCM), a global coupled climate model. They conclude that if apparently spurious century–time-scale trends in the control run are accounted for, corresponding trends in the response to forcings are, in most cases, additive. Here we address another aspect of the additivity issue with regards to the decadal and interannual time-scale response of globally averaged surface air temperature.

2. The model and experiments

We use the fully coupled Department of Energy (DOE) PCM described by Washington et al. (2000) and used in the studies of Ammann et al. (2003), Meehl et al. (2003), WIGb, and Santer et al. (2003a,b). The resolution of the atmosphere is T42, or roughly 2.8° × 2.8°, with 18 levels in the vertical. Resolution in the ocean is roughly 1/8° down to a 1/2° in the equatorial Tropics, with 32 levels. No flux corrections are used in
Table 1. Forcings used in PCM twentieth-century climate simulations (each represents a four-member ensemble). Abbreviations for forcings used in Table 2 are in parentheses.

Forcing experiments, 1890–2000, for four-member ensembles:

1) Solar (Hoyt and Schatten 1993) (So)
2) Sulfate aerosol direct effect (S)
3) GHG (CO$_2$, water vapor, O$_3$, CH$_4$, N$_2$O, CFC12, CFC11) (G)
4) Ozone (tropospheric and stratospheric) (O)
5) Volcano (Ammann et al. 2003) (V)
6) Volcano + solar (VS)
7) GHG + sulfate + ozone (GSO)
8) GHG + sulfate (GS)
9) Solar + GHG + sulfate + ozone (SoGSO)
10) Volcano + solar + ozone (VSO)
11) Volcano + solar + GHG + sulfate + ozone (VSOG)
12) Solar + ozone (SO)
13) GH + ozone (GO)

Figure 1. (a) The four-member ensemble mean (red line) and ensemble member range (pink shading) for globally averaged surface air temperature anomalies ($^\circ$C; anomalies are formed by subtracting the 1890–1919 mean for each run from its time series of annual values) for volcanic forcing; the solid blue line is the ensemble mean and the light blue shading is the ensemble range for globally averaged temperature response to volcanic forcing calculated as a residual [(volcano + solar) – solar]; the black line is the observations after Folland et al. (2001); (b) same as (a) except for solar forcing, and a solar residual [(solar + GHG + sulfate + ozone) – (GHG + sulfate + ozone)]; (c) same as (a) except for sulfate forcing, and sulfate residual [(GHG + sulfate + ozone) – (GHG + ozone)]; (d) same as (a) except for ozone forcing, and ozone residual [(GHG + sulfate + ozone) – (GHG + sulfate)]; (e) same as (a) except for GHG forcing and GHG residual [(GHG + sulfate) – sulfate].
response (ensemble mean and range) for single forcings compared to equivalent residual responses calculated as differences from other forcing runs. If there is good agreement between the two time series, the global-mean response to the forcing can be assumed to be additive. Though all show seemingly relatively good agreement by eye both for the ensemble means and with the overlap of the ensemble ranges for each simulation, we will quantify similarity in section 4. The observed time series of globally averaged temperature is also included in each panel for comparison.

For volcanic forcing (aerosols from volcanic eruptions act to cool the surface temperatures for a few years after a given eruption), active volcanic eruptions prior to about 1915 and after 1960 (Ammann et al. 2003) keep the globally averaged temperature anomalies for those two time periods close to or below $0^\circ$C (Fig. 1a). The global cooling signature from individual volcanic eruptions in the observations can be seen in the model simulations as documented by Ammann et al. (2003) and seen in Wigley et al. (2004a, manuscript submitted to J. Climate). From about 1910 to 1960, the lack of volcanic activity results in globally averaged temperature anomalies of around $+0.05^\circ$ to $+0.15^\circ$C, though the volcanic temperature response remains mostly below the observed values after about 1930. The temperature response to the other natural forcing (solar) is shown in Fig. 1b. As noted by Meehl et al. (2003), the low-frequency increase of solar forcing over the first four decades of the twentieth century produces warming of around $+0.2^\circ$C in the 1940s compared to the beginning of the century, followed by a relative cooling of about $0.1^\circ$C in the 1950s–70s, and then increases to greater than $+0.2^\circ$C after about 1980. The two natural forcings combine to contribute to midcentury warming, though, as documented by Meehl et al. (2003), most of this signal is from the solar forcing. At the end of the century, however, negative volcanic forcing counteracts the effects of positive solar forcing (Ammann et al. 2003).

For the anthropogenic forcings, the temperature response to direct sulfate aerosol forcing in Fig. 1c is negative, with the greatest cooling from increasing sulfate aerosol loadings occurring after about 1950 of about $-0.2^\circ$ to $-0.3^\circ$C compared to the beginning of the century. Changes in tropospheric and stratospheric ozone in the model produce an apparent positive temperature response of about $0.1^\circ$C after 1970, though this signal is small compared to the noise as seen in Fig. 1d. By far the largest temperature response is to the GHGs in Fig. 1e, with slow warming occurring in the first half of the twentieth century up to about $+0.1^\circ$C in the 1940s, but then accelerating after about 1970 to values of around $+0.8^\circ$C at the end of the century. It is clear that the only forcing that approximates the large observed late-century warming is the response to anthropogenic GHGs in Fig. 1e.

Figure 2 shows the responses to specific forcing combinations, together with the equivalent residuals, com-
pared to observations. The simulation with the natural forcings combined (volcanic and solar) in Fig. 2a appears to have characteristics of the sum of the responses of the forcings by themselves in Figs. 1a,b. There is warming greater than about +0.2°C in the mid-1940s, and somewhat less warming in the late century. Additionally, there are abrupt temperature reductions of several tenths of a degree in association with volcanic eruptions (notably Agung in 1963, El Chichón in 1982, and Mt. Pinatubo in 1991). The observed time series falls within the ensemble ranges for the first half of the twentieth century for all but a few years in the mid-1940s. However, natural forcings alone cannot reproduce the observed late-century warming. In fact, in the natural forcings run, the last decades are about 0.1°C cooler than the peak temperatures in the 1940s.

For the anthropogenic forcings combined in Fig. 2b (GHGs, sulfates, ozone), the dominance of the GHGs in Fig. 1e in comparison to the other two contributes to warming of about +0.6°C at the end of the century. The cooling effect of the sulfates noted in Fig. 1c reduces the warming from ozone (Fig. 1d) and GHGs (Fig. 1e). Anthropogenic forcings alone do a poor job of simulating the observed midcentury warming, with most years from about 1935 to 1970 lying outside the ensemble range, but do a very good job in capturing the late-century warming after 1970. If the responses from the individual forcings are added and their sum is compared to the simulation where they are all included together, there is a close correspondence as shown in Fig. 2c. This suggests that, for globally averaged temperature, additivity holds and responses from individual forcings can be combined linearly. The good correspondence between the model simulations and the observations should not be overinterpreted. If the assumed forcings were correct, then this agreement would indicate that the model’s climate sensitivity was realistic. Forcing uncertainties, however, admit a quite wide range of sensitivity possibilities. When the temperature response from the simulation with natural forcings is compared to the simulation where natural and anthropogenic forcings are combined (see Figs. 2a,c,d), it can be seen that natural forcings contribute most to the early twentieth-century warming. The simulation with all forcings closely tracks the large observed warming in the late twentieth century. It can be seen from Fig. 2b that, in the model, this is mainly a consequence of the climate system response to anthropogenic forcings. This confirms previous studies that have shown similar results (e.g., Stott et al. 2000; Meehl et al. 2003; Ammann et al. 2003).

4. Quantifying linearity

To address the issue of linearity of the model’s globally averaged temperature response to different forcings, we can rigorously quantify the agreement on the decadal time scale of the globally averaged temperature responses to single and combined forcing simulations with the corresponding temperature responses calculated as residuals.

We do this by performing a principal component–EOF analysis on the 13 forcing experiments in Table 1 (110-yr time series, each representing the ensemble mean of a four-member ensemble experiment, 10-yr running averages). After quantifying the relative importance of the 13 components thus derived, we use a few dominant ones as basis functions to reconstruct the original and residual temperature responses and quantify their agreement.

We define X as the 110 × 13 matrix whose columns contain the (centered, i.e., having mean zero) time series, and define S as its variance–covariance matrix, i.e., $S = XX^T$. We derive the matrix $U = XB$ of principal components $U1, U2, \ldots, U13$, through a singular value decomposition of the matrix X. By construction $UU^T = L$, a diagonal matrix containing the eigenvalues of the matrix S. Thus, the relative size of the values in L can be used as an indication of which of the components should be retained, and which can be discarded. By this analysis, we find that as few as three principal components (PCs) are associated with 99% of the variance observed in the matrix X of the original 13 experiments. In particular, U1 explains 90%, U2 explains 8%, and U3 1% of the total variance.

The reconstructions of the time series $X1, X2, \ldots, X13$, by linear combinations of these three PCs, have $R^2$ value above 90% in all cases but for the ozone only forcing experiment. For the latter, the $R^2$ value is 82%, but it should be pointed out that the total variation of the ozone-only temperature response is very small, and the mean standard error of the regression is only 0.018°C, well in line with the other 12 mean standard errors, ranging between 0.007°C and 0.020°C.

Thus we have found that the 13 experiments can be accurately summarized, both in terms of individual and combined variance, and in terms of $R^2$ and root-mean-square error, by the linear combination of just three 110-yr-long time series. We show the three PCs, U1, U2, and U3, in Fig. 3a (and we list in Table 3 the three coefficients for each of the 13 reconstructions, $X = U^{-1}*$, where the asterisk is a reminder that these are reduced matrices $110 \times 3$ and $3 \times 13$, respectively, compared to the original B and U).

We consider this result of interest in itself, because it definitely suggests the intrinsic linearity of the temperature response to different forcing combinations, linearity that we quantify next. The same result also identifies three fundamental modes of variation in the temperature responses, whose relative importance can be interpreted in light of the original combined forcing time series in the context of the original forcing response in Fig. 1. For example, PC1 in Fig. 3a clearly relates to the dominant role of GHGs shown in Fig. 1e. This can be seen in both the response to GHGs in Fig. 1e and PC1 in Fig. 3a as a gradual rise in the first two-thirds
of the time series, followed by a steep increase in the last third. For GHGs, this reflects the large increase in fossil fuel consumption in the latter part of the twentieth century and associated GHG production. PC2 in Fig. 3a represents most of the elements of the globally averaged temperature time series from volcano and solar forcing in Figs. 1a and 1b, respectively, with some of the late-century cooling from sulfate aerosols (Fig. 1c) represented as well. As noted by Meehl et al. (2003), solar was likely the dominant forcing that produced the temperature increases in the first part of the twentieth century. Thus, in PC1 and PC2, explaining 90% and 8%, respectively, of the variance of the original time series, the elements of the two major features of the globally averaged surface air temperature over the twentieth century are represented: the early century warming associated mainly with solar forcing, and the late-century warming produced mainly by increased GHGs in the model. PC3, representing only 1% of the variance, shows low-amplitude variations that have elements of both the early century and late-century warming and cannot be associated directly with any single forcing time series in Fig. 1. Had the PC analysis shown a gradual decay of variance explained among the 13 PCs with no obvious cutoff, we would not advocate this approach over a simple mean standard error computation between original forcings and residuals. But the obvious dominance of just three components seems to us to justify this simple additional step, and adds insight to the analysis of linearity.

Having quantified the high degree of accuracy by which linear combinations of the three dominant PCs approximate the original time series, we now perform the comparison between original forcing time series and the residual time series simply by comparing the corresponding sets of three coefficients. By exploiting the orthogonality of the three PCs, we simply compute differences between corresponding coefficients and multiply them by the variances associated with the three components, in order to quantify how much of the original variance is lost by approximating the temperature response from the original forcing experiment by the residual time series.

As shown in Fig. 4a, the variance loss values for all 20 pairs of original forcings and residuals listed in Table 2 are less than 1%. Therefore, we can conclude that about 99% or more of the variance on the decadal time scale is reproduced using the residual response to the forcings compared to the original forced response.

We include two additional calculations in Fig. 4a to illustrate how much variance is lost from forcing experiments that are clearly not additive in Fig. 2. For the simulations with all forcings combined compared to the simulation with only natural forcings (Fig. 2b) the variance lost in the reconstruction is 11.57%, mainly due to the simulations diverging in the late twentieth century (the loss of variance being associated mainly to the discrepancy in the coefficients associated with PC1). For another combination, the anthropogenic forcing in Fig. 2b and the natural forcing in Fig. 2a, the results in Fig. 4a show the significant loss of 11.63% of the variance for anthropogenic compared to natural. Again, the loss is accounted for by the different modulation of PC1, which quantifies the qualitative behavior shown in Fig. 2d.

A similar PC calculation is performed on the unsmoothed annual mean data. In this context a PC analysis addresses the additional need of making the results robust with respect to the inherent noise in the time series, identifiable as interannual variability or single year responses from individual volcanoes. As an example of this noise, ENSO events occur in an unrelated fashion in individual ensemble members, and the use of the ensemble average may smooth some of the noise.
but not entirely. Conversely, the model response to volcanic eruptions is forced the same way in every ensemble member, and should be represented in a PC representation of the time series.

In this case the PC loadings are 80% for PC1, 10% for PC2, and 2% each for PC3 and PC4, the remaining fraction being distributed uniformly among PC5–10. Here we reconstruct the original time series by retaining the four principal components, shown in Fig. 3b. Similar to what we found earlier, the individual reconstructions have a high degree of accuracy. The mean standard error between the reconstruction and the original time series is between 0.03 and 0.04 in all cases.

As in Fig. 3a, the PCs in Fig. 3b can be interpreted in terms of the temperature response time series in Figs. 1 and 2. PC1 corresponds most closely to the GHG simulation in Fig. 1e, and PC2 to the natural forcings (solar plus volcanoes) in Fig. 2d. Note that the individual volcanic eruptions show up in PC2 as sharp downward excursions of the time series (e.g., for El Chichón and Mt. Pinatubo in 1982 and 1991, respectively). PC3 and PC4 are similar and both pick up small-amplitude low-frequency variance not accounted for by the first two dominant PCs.

Figure 4b shows very similar results to Fig. 4a, with the variance lost in the reconstructions by the residual time series being less than 1% in all 20 cases but one, where the loss is only 1.1%. As could be expected from the decadal calculation in Fig. 4a, for the two cases where there is clear divergence of the signals (cases 21 and 22) there is a greater than 10% loss of variance (10.9% and 12.7%, respectively). Again, the loss is accounted for in almost its entirety by the difference in the coefficients of PC1.

<table>
<thead>
<tr>
<th>PC2</th>
<th>PC3</th>
<th>PC4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90</td>
<td>0.20</td>
<td>0.09</td>
</tr>
<tr>
<td>-0.18</td>
<td>0.00</td>
<td>-0.45</td>
</tr>
<tr>
<td>0.43</td>
<td>-0.23</td>
<td>0.28</td>
</tr>
<tr>
<td>0.09</td>
<td>-0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>-0.06</td>
<td>0.37</td>
<td>-0.11</td>
</tr>
<tr>
<td>0.03</td>
<td>0.57</td>
<td>0.10</td>
</tr>
<tr>
<td>0.32</td>
<td>-0.14</td>
<td>-0.44</td>
</tr>
<tr>
<td>0.25</td>
<td>-0.07</td>
<td>-0.41</td>
</tr>
<tr>
<td>0.40</td>
<td>0.22</td>
<td>0.10</td>
</tr>
<tr>
<td>0.06</td>
<td>0.49</td>
<td>0.26</td>
</tr>
<tr>
<td>0.37</td>
<td>0.28</td>
<td>-0.35</td>
</tr>
<tr>
<td>0.13</td>
<td>0.13</td>
<td>-0.02</td>
</tr>
<tr>
<td>0.53</td>
<td>-0.17</td>
<td>0.33</td>
</tr>
</tbody>
</table>
5. Conclusions

Thirteen forcing simulations for twentieth-century climate are run in four-member ensembles, and the ensemble means are compared to test for additivity of the forcings compared to residual calculations of the globally averaged surface air temperature response. We do this by comparing the similarity of the statistics of the variance at the decadal and interannual time scales. The late-twentieth-century warming can only be reproduced in the model if anthropogenic forcing (dominated by GHGs) is included, while the early twentieth-century warming requires the inclusion of natural forcings in the model (mostly solar). The signature of simulated globally averaged temperature at any time in the twentieth century is a direct consequence of the sum of both natural and anthropogenic forcings. The similarity of the response to the forcings on the decadal and interannual time scale is tested by performing a PC analysis with the 13 ensemble mean globally averaged temperature time series. A significant portion of the variance (mostly greater than 99%) of the reconstructed time series can be retained in residual calculations compared to the original single and combined forcing runs, thus demonstrating similarity in terms of the statistics of the variability being additive for the decadal and interannual time scale in the forced simulations.

These results pertain specifically to the globally averaged surface air temperature response to different forcings. However, Meehl et al. (2003) showed that the regional precipitation response in the Tropics to solar forcing in the early twentieth century was different than that for GHG forcing in late twentieth century. This was mainly due to coupled feedbacks that affected the strength of climatological precipitation maxima in the convergence zones over the oceans and in the monsoon regions. The results in the present paper indicate that, for global scales, additivity is a good approximation of the response to various forcings. But at regional scales, Meehl et al. (2003) show that coupled feedbacks may affect the response in different ways. The relationship between global and regional responses is the now the subject of a further investigation.

Acknowledgments. This work was supported in part by the Weather and Climate Impact Assessment Initiative at the National Center for Atmospheric Research. Additionally, portions of this study were supported by the Office of Biological and Environmental Research, U.S. Department of Energy, as part of its Climate Change Prediction Program, and the National Science Foundation.

REFERENCES