## Journal of the American Statistical Association

# Comment

T. Storelvmo and T. Leirvik

T. Storelvmo is Assistant Professor, Department of Geology and Geophysics, Yale
University, New Haven, CT 06511. T. Leirvik is Ph.D. Student, Department of Economics,
University of Lugano, Lugano 6900, Switzerland.

PLEASE SCROLL DOWN FOR ARTICLE

# Comment

## T. STORELVMO and T. LEIRVIK

### 1. INTRODUCTION

In their pioneering study applying statistical and econometric analysis to a climate dataset, Magnus, Melenberg, and Muris (2011), hereafter MMM, present a decomposition of the temperature trend over the last four decades of the 20th century into a solar radiation component and a greenhouse gas component. While it is well known that atmospheric greenhouse gas (GHG) concentrations are increasing steadily in response to anthropogenic fossil fuel burning and thereby heating the planet, the general public is less aware of a corresponding downward trend in solar radiation reaching the surface. The latter is very likely due to higher concentrations of particles (so-called aerosols) in the atmosphere, also as a result of anthropogenic fossil fuel and biomass burning. Aerosols can both (1) directly increase the amount of solar radiation scattered back to space and (2) brighten clouds such that they become more reflective to solar radiation. Hence the observed reduction in solar radiation reaching the surface during the second half of the 20th century, which is sometimes referred to as "global dimming" but will be termed "the aerosol effect" for the remainder of this discussion.

### 2. THE DRIVERS OF CLIMATE CHANGE: GREENHOUSE GASES VERSUS AEROSOLS

The aerosol effect acts to cool the climate and partly compensate for the warming due to increasing GHG concentrations, and is currently one of the most uncertain aspects of projections of future climate. This is largely because of the complicated and poorly understood processes involved in the emissions and atmospheric lifetimes of aerosols, and their interactions with clouds and radiation while airborne. In Storelvmo et al. (2009), four different methods for calculating the aerosol brightening of clouds in global climate models (GCM) were compared, and yielded coolings that ranged from negligible to comparable in magnitude to the warming due to increasing greenhouse gases. This wide uncertainty range for the aerosol effect is problematic for the following reason: The GCMs that were employed to simulate future climate in the last report from the Intergovernmental Panel for Climate Change (IPCC) (23 in total; Randall et al. 2007) report vastly different so-called Equilibrium Climate Sensitivities (ECSs). The ECS is defined as the change in mean surface air temperature in response to a doubling of atmospheric $CO_2$. The ECS ranged from 2.1°C to 4.4°C for the 23 models that participated in this Climate Model Intercomparison Project (*http://www-pcmdi.llnl.gov/ipcc/about_ipcc.php*). Surprisingly, despite the wide range of ECSs among the models, they were all able to reasonably reproduce the observed temperature record for the 20th century. However, in a recent article by Kiehl (2007) an explanation for this puzzle was offered; they found that there was a strong negative correlation between the magnitude of the aerosol effect and the ECS for a subset of the 23 models discussed above. This can be understood as follows: A GCM with anomalously high climate sensitivity may yield an exaggerated warming trend in response to increasing GHGs, but can currently compensate for this by choosing an aerosol effect from the upper end of the uncertainty range (corresponding to a relatively strong cooling that masks much of the warming due to GHGs). Similarly, a GCM with anomalously low climate sensitivity can currently choose a negligible aerosol effect. The ECS is therefore largely unconstrained by the observed temperature record of the past, with the result that future climate cannot be predicted with any confidence (global mean surface air temperatures are projected to rise by anything from 1°C to 6°C by year 2100, according to the 4th assessment report from IPCC, hereafter IPCC AR4).

### 3. COMPARING THE STATISTICS OF OBSERVATIONAL VERSUS MODELING DATASETS

In this context, the interdisciplinary approach taken by MMM is timely and an important contribution. By estimating the parameters of a relatively simple climate model, they are able to make quantitative statements about what fraction of the warming due to increasing GHGs was likely masked by an aerosol effect. This is referred to as a radiation effect in their study, however the only plausible explanation for the observed trend in solar radiation reaching the surface is the increase of atmospheric aerosol concentrations over the examined time period (Wild, Ohmura, and Makowski 2007). This study opens up for the exciting possibility of constraining the aerosol effect in GCMs used to project future climate. However, there are several potential cavities in such an approach, some of which are: (1) Observations of the climate variables relevant for the simplified climate model in MMM are taken from unevenly distributed weather stations, which generally tend to undersample sparsely populated and remote regions with harsh climates; (2) Radiation data is missing for certain time periods and weather stations, complicating the comparison to models further; and (3) The typical resolution of a GCM is currently approximately $2.5 \times 2.5$ degrees. In MMM radiation data from individual GEBA weather stations were assigned temperatures from the CRU data set, given on a $0.5 \times 0.5$ degrees grid. In other words, for each GCM data point, there will 25 observational data points. The latter two issues are illustrated in Figure 1.

All of the above may result in differences between GCMs and the data analyzed by MMM that are related to data quality rather than actual climate. As an illustration, we have randomly selected one of the GCMs that were used to simulate

T. Storelvmo is Assistant Professor, Department of Geology and Geophysics, Yale University, New Haven, CT 06511 (E-mail: *trude.storelvmo@yale.edu*). T. Leirvik is Ph.D. Student, Department of Economics, University of Lugano, Lugano 6900, Switzerland.
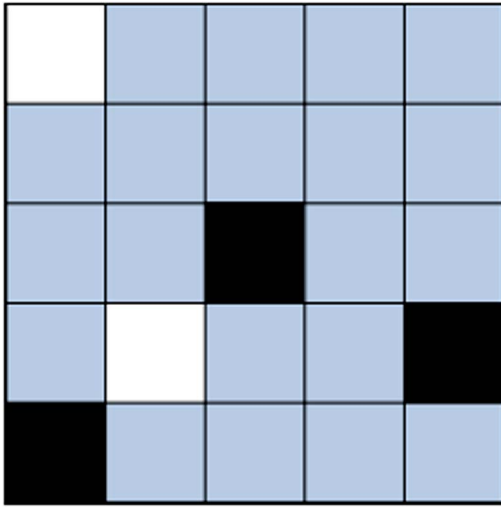
Figure 1. Illustration of one GCM grid box covering 2.5 degrees in the latitude and longitude directions, which corresponds to a maximum of 25 data values from the observational grid. However, as indicated by the black boxes, data is missing in some of the observational grid boxes not containing any weather stations. Furthermore, some observational grid boxes have temperature (TEMP) data but are missing radiation (RAD) data for parts of the time series, as illustrated by the white boxes. The online version of this figure is in color.

future climate for IPCC AR4, and analyzed GCM data from the same time period as considered in MMM (i.e., 44 years, from January 1959 to December 2002). The spatial resolution of this specific GCM is $2.8 \times 2.8$ degrees. We have selected only GCM grid boxes that have a minimum of 20% land, and excluded Antarctica, as in MMM. The threshold of 20% is arbitrary, but sensitivity tests revealed little sensitivity to choice of land threshold. The statistics obtained for the variables RAD (downward solar radiation at the surface in $Wm^{-2}$) and TEMP (surface air temperature in °C), are displayed in Tables 1 and 2. For comparison, we have also included the observational data from Tables 1 and 2 in MMM (for the case of a complete panel for TEMP).

As evident from Table 1, the mean temperature from the GCM is significantly lower (by 4.57°C) than the observational mean. This is not surprising, considering that the observations are undersampling colder regions (see Figure 2 in MMM). The sensitivity of the temperature statistics to the actual weather stations sampled is also illustrated by the difference between the complete and unbalanced panels in MMM. For the same reason, the minimum temperature from the GCM is much lower (by 11.36°C) than the corresponding observational value, while the maximum temperature is almost identical. Because the GCM samples a wider range of temperatures, the location variance is larger in the GCM dataset. However, the year-to-year variability (given by the annual standard deviation) is much smaller in the GCM.

RAD is somewhat larger in the GCM than in the observational dataset. In this case it is less obvious that unevenly distributed observations are causing the difference, but observational data points are clustered at Northern Hemisphere midlatitudes, less dense in the Tropics and very sparse at high latitudes. Another plausible explanation for the discrepancy between the GCM and observations could be poorly represented clouds and aerosols in the GCM. The variance in the GCM dataset is higher, as expected from the wider range of locations sampled.

Table 2 shows the statistics for the time differences in temperature from the GCM and the observations (taken from Table 2 in MMM). The overall mean temperature trend in the GCM is about 1/3 of that observed. Again, because of the unevenly distributed observations differences are expected. However, Arctic regions, that have been observed and projected to warm at a faster rate than the rest of the Globe, are not well represented in the observations. Hence, the discrepancy between the simulated and observed temperature trend is not easily explained by uneven observational sampling. Another alarming difference is that while MMM finds a strong negative trend in RAD over the time period considered [Figure 4(a) in MMM], and concludes that this effect has likely masked 58% of the warming due to greenhouse gases, the GCM simulates no trend in RAD at all (slope coefficient: $-0.012$, $R^2 = 0.049$) over the same time period. In other words, this particular GCM simulates no masking effect, in strong contrast to the finding in MMM. However,

Table 1. Annual mean sample statistics for TEMP and RAD. The label *overall* refers to a statistic of the entire dataset spanning both time and space, while the between and within labels refer to cross-sectional (averaged over time) and time series (averaged over locations) statistics of the dataset, respectively. The OBS label refers to the observational dataset used in MMM, while the GCM label refers to data generated by a Global Climate Model simulating climate over the time period 1959–2002

| Variable | | Mean | Std. | Min | Max |
|---|---|---|---|---|---|
| TEMP OBS | overall | 13.40 | 8.90 | −22.04 | 31.23 |
| | between | 13.40 | 8.89 | −19.96 | 29.75 |
| | within | 13.40 | 0.34 | 12.91 | 14.14 |
| TEMP GCM | overall | 8.83 | 13.79 | −33.4 | 31.22 |
| | between | 8.83 | 13.77 | −25.0 | 30.53 |
| | within | 8.83 | 0.21 | 8.43 | 9.29 |
| RAD OBS | overall | 160.9 | 42.46 | 52.00 | 324.00 |
| | between | 160.9 | 44.68 | 55.46 | 316.00 |
| | within | 160.9 | 9.09 | 148.77 | 183.21 |
| RAD GCM | overall | 176.3 | 58.8 | 53.7 | 303.0 |
| | between | 176.3 | 58.5 | 65.6 | 300.2 |
| | within | 176.3 | 0.61 | 174.5 | 177.6 |

Table 2. Sample statistics for time differences in temperature, that is, the annual mean temperature change from one year to the next. See Table 1 for explanations of labels

| Variable | | Mean | Std. | Min | Max |
|---|---|---|---|---|---|
| ΔTEMP OBS | overall | 0.0142 | 0.731 | −4.925 | 5.158 |
| | between | 0.0142 | 0.016 | −0.058 | 0.094 |
| | within | 0.0142 | 0.258 | −0.514 | 0.573 |
| ΔTEMP GCM | overall | 0.0046 | 1.040 | −7.829 | 7.377 |
| | between | 0.0046 | 0.037 | −0.165 | 0.173 |
| | within | 0.0046 | 0.144 | −0.298 | 0.237 |

it should be kept in mind that these results are from a single GCM, which may well be an outlier.

## 4. OUTLOOK

To determine whether GCMs in general are systematically simulating a much weaker aerosol masking effect than the one reported in MMM, further study involving new results from numerous GCMs is required. A fifth report from IPCC is currently underway (due in 2014), and model results from a new CMIP are becoming available. An interesting extension to MMM and the present discussion would be to do equivalent analyses of all models contributing to the CMIP, to determine whether simulated relationships between the aerosol and greenhouse effect are comparable to the one found in MMM. Ultimately, this may aid the climate modeling community in their efforts to determine the true climate sensitivity. Certainly, such a study would have to be interdisciplinary in nature, bringing experts from the statistics and econometrics community and the field of climate science together in an effort to solve a problem that has been identified as one of the major scientific puzzles of our time (Kerr 2005).

## ADDITIONAL REFERENCES

Kerr, R. A. (2005), "How Hot Will the Greenhouse Be?" *Science*, 309, 100. [467]

Kiehl, J. T. (2007), "Twentieth Century Climate Model Response and Climate Sensitivity," *Geophysical Research Letters*, 34, *doi:10.1029/2007GL031383*. [465]

Randall, D. A., Wood, R. A., Bony, S., Colman, R., Fichefet, T., Fyfe, J., Kattsov, V., Pitman, A., Shukla, J., Srinivasan, J., Stouffer, R. J., Sumi, A., and Taylor, K. E. (2007), "Climate Models and Their Evaluation," in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller, Cambridge, U.K./New York: Cambridge University Press. [465]

Storelvmo, T., Lohmann, U., and Bennartz, R. (2009), "What Governs the Spread in Shortwave Forcing," *Geophysical Research Letters*, 36, *doi:10.1029/2008GL036069*. [465]

Wild, M., Ohmura, A., and Makowski, K. (2007), "Impact of Global Dimming and Brightening on Global Warming," *Geophysical Research Letters*, 34, *doi:10.1029/2006GL028031*. [465]

# Rejoinder

Jan R. MAGNUS, Bertrand MELENBERG, and Chris MURIS

We thank the editors for inviting this discussion, and Trude Storelvmo and Thomas Leirvik (henceforth SL) for their kind remarks and thoughtful comments. Our empirical analysis, like almost all statistical analyses, can be criticized from four angles: choice of model, estimation technique, suitability of the data, and applicability. We shall briefly touch on each of these, although SL focus on the applicability aspect. Their principal concern is whether our approach, in particular the decomposition of the temperature change into a greenhouse effect and a solar radiation effect (called the "aerosol effect" by SL), can be used in calibrating global climate models (GCMs). SL raise and illustrate a number of issues that might complicate such an application. This is an important concern and it was not discussed in our article. So we welcome the opportunity to discuss it now. We first consider the tension between model complexity and statistical feasibility, and clarify the difference between a GCM and our model, then we turn to the applicability issue, and finally we mention some other potential limitations.

A GCM is a detailed model of the climate system. It allows for interaction and feedback effects, and typically depends on many (unknown) parameters. A GCM is complex—too complex to allow a "proper" statistical analysis. There are two reasons why this is so. First, given the available data, such as observations on surface temperature, the parameters of a GCM are typically not identifiable: simulating the climate with different combinations of parameter values might result in the same temperature series, so that the observed temperature data does not

Jan Magnus is Professor of Econometrics (E-mail: *magnus@uvt.nl*) and Bertrand Melenberg is Professor of Econometrics and Quantitative Finance (E-mail: *b.melenberg@uvt.nl*), Department of Econometrics and OR, Tilburg University, P.O. Box 90153, 5000LE Tilburg, The Netherlands. Chris Muris is Postdoctoral Fellow, Courant Research Center "Poverty, equity and growth in developing and transition countries," University of Göttingen, Wilhelm-Weber-Str. 2, D-37073, Germany (E-mail: *chrismuris@gmail.com*).