Using Weather Data and Climate Model Output in Economic Analyses of Climate Change

Maximilian Auffhammer*, Solomon M. Hsiang†,
Wolfram Schlenker§ and Adam Sobel‡

Introduction

There is a long history of using weather measures as explanatory variables in statistical models. For example, Fisher (1925) examined the effects of rainfall on wheat yields, and Wright (1928) used weather as an instrumental variable to identify a demand function for oils. Because weather is exogenous and random in most economic applications, it acts like a "natural experiment" and thus in some settings allows researchers to identify statistically the causal effect of one variable on an economic outcome of interest (Angrist and Krueger 2001). The relatively recent literature on the economic impacts of climate change has turned the spotlight onto quantifying the effect of climate on a number of economic outcomes of interest (e.g., agricultural yields, mortality rates, electricity and water demand). This literature has often found a nonlinear relationship between climate and these outcomes, with extremely warm temperatures being especially important (e.g., Schlenker and Roberts 2009). Climate is a long average of weather at a given location. To identify the causal effect of climate on these outcomes, the literature has generally relied on either climate normals (i.e., long averages of observed weather in a cross-sectional setting) or day-to-day (or year-to-year) fluctuations in observed weather as explanatory variables across time and space. The econometrician's choice of a weather versus a climate measure as an explanatory variable critically affects the interpretation of the estimated coefficients in the econometric model: that is, whether the outcome is a true climate response or a short-run weather elasticity.

*University of California, Berkeley, and National Bureau of Economic Research (NBER); e-mail: auffhammer@berkeley.edu.

†Princeton University, Princeton, NJ, and NBER; e-mail: shsiang@princeton.edu.

‡Columbia University, New York, NY; e-mail: ahs129@columbia.edu.

§Columbia University, New York, NY; University of California, Berkeley, and NBER; e-mail: schlenker@berkeley.edu.

We thank Charlie Kolstad, four referees, and Suzanne Leonard for very helpful comments. We acknowledge support from DOE grant DE-FG02-08ER64640 (Auffhammer and Schlenker), EPA Grant FP-916932 (Hsiang), and NSF Grant NSF SES-1048946 (Sobel). All remaining errors are ours.

Review of Environmental Economics and Policy, volume 7, issue 2, 2013, pp. 181–198
doi:10.1093/reep/ret016
Advance Access published on June 28, 2013

Anyone who has ever struggled with station-level weather data is well aware of the fact that since the beginning of systematic weather monitoring in the 1800s, stations are born and die, and almost all have a large number of missing observations. Further, there is not necessarily a weather station in each location of interest to the economist. In order to overcome these issues, a number of gridded weather data sets have been developed; these provide complete coverage over land by extrapolating existing weather information from monitors over a grid. Although many of these data sets are free and easily imported into formats used by economists, there are five pitfalls that empirical researchers should be aware of before using any of these data sets in econometric settings. First, while many of the gridded weather products that are available reproduce very similar *average* temperatures for the majority of grid cells, *the derived deviations around the mean can be significantly different.* Second, if one is interested in creating a weather series for a geographic region, simply averaging nonmissing weather station data for stations in the region introduces measurement error, which has well understood econometric consequences. Third, the correlation between weather variables (e.g., rainfall and temperature) across space varies significantly in sign and magnitude. This can lead to the classic problem of indeterminate omitted variables bias in applications that fail to control for the full suite of weather indicators. Fourth, weather indicators often display significant spatial correlation due to the underlying data-generating process as well as the extrapolation methods employed. This may lead to significant multicollinearity, which in turn may lead to inflated standard errors on included weather variables. Finally, because the weather stations used to construct the gridded products come in and out of existence, there may be artificial variation and breakpoints in the temperature series, which the econometrician needs to examine, especially when working on a small geographic region.

The majority of recent economic studies use the statistically estimated causal effect of weather on the economic outcome of interest to simulate the future impacts of climate change, based on the output of global climate models (GCMs), on that outcome. GCMs[1] are physics-based models that provide long-run predictions of climate. These models are sometimes also called atmosphere-ocean GCMs (AOGCMs), or simply, and most commonly, climate models. However, our experience has been that most economists have limited understanding of GCMs and often make critical mistakes in using their output. Thus our goal here is to provide economic researchers considering the use of weather and climate model output with a guide to what products are available and, most importantly, with a discussion of the most common mistakes and how to avoid them. We begin in the next section with an introduction to weather data, including a summary of the types of data sets available. Next we provide a more detailed discussion of the five common pitfalls mentioned earlier. This is followed by a brief overview of GCMs—how they work and what output they provide—as well as a number of suggestions for further reading. Perhaps most importantly, we identify two common and significant errors that often occur when GCM output is used to simulate the future impacts of climate change on an outcome of interest, which are related to GCM model selection and spatial and temporal aggregation of GCM output. We present a summary and conclusions in the final section.

[1]Early on these were known as general circulation models (see e.g., Phillips 1956).

An Introduction to Weather Data

As noted in the introduction, the difference between weather and climate is basically a matter of time. Weather is the condition of the atmosphere over a short period of time, whereas climate is the behavior of the atmosphere over a relatively long period of time. Since roughly 1850, weather outcomes have been measured and recorded through a global network of weather stations and, more recently, satellites. Daily weather data at stations throughout the world are freely available from the US National Oceanic and Atmospheric Administration (NOAA 2011a). Additional raw station data (with varying degrees of spatial and temporal coverage and temporal resolution) can be found at NOAA (2011b). However, these sources do not provide a complete record because many countries regard their weather data as proprietary and often charge high fees for such data (e.g., India), thus effectively limiting their availability. Moreover, the spatial and temporal coverage of weather stations varies greatly across the globe, with higher spatial density and longer time series at stations in countries with historically higher incomes (e.g., the United States and the European Union 15).

Gridded Weather Data Products

Gridded weather data sets use interpolation across space and time to combine available weather station data into a balanced panel of observations on a fixed spatial scale or grid. This approach deals with the problem of missing observations at a given station or missing data because a station does/did not exist at a particular location. One such product, the parameter-elevation regressions on independent slopes model (PRISM 2009), produces monthly estimates of weather on a 2.5×2.5-mile scale for the contiguous United States. Each "grid" approximates a weather measure for the spatial unit by interpolating the daily station data while accounting for elevation, wind direction, rain shadows, and many other factors. This elaborate procedure is possible in the United States because there are several thousand weather stations that produce daily records for many different weather indicators.

The Climatic Research Unit at the University of East Anglia produces a global gridded weather data set (Climatic Research United [CRU] 2013) that provides monthly estimates on a 0.5×0.5-degree scale. This scale corresponds roughly to grids that are 35 miles across at the equator. Willmott, Matsuura, and Legates (2010) provide another gridded data product (often referred to as the "Delaware," or "UDEL," data set because it was produced by the University of Delaware), which has the same spatial and temporal resolution as the CRU (version TS2.1) product, but uses a somewhat different data set and extrapolation algorithm. Most notably, the CRU product provides data on both monthly minimum and maximum temperatures (i.e., the average of all daily minimums and maximums), while the Delaware data set provides only the monthly average temperature.[2]

Many data products include the number of stations and the dates of coverage for each grid. The most pronounced absence of data is for poor regions whose governments do not prioritize weather data collection and for regions with few inhabitants, such as deserts or over oceans. In fact, there are some grids covering land areas that do not have a single weather station.

[2]The CRU data set (version TS2.1) ended in 2002. The updated data set version TS 3.2 extends coverage through 2011. In early 2013, the Delaware data set coverage ended in 2008, but it is currently being extended.

Data Assimilation

An alternative approach to the spatial extrapolation algorithms just discussed that climate scientists have developed for filling in the "holes" for observationally sparse regions is "data assimilation," which produces data sets that the climate community generally calls "reanalyses." Data assimilation is the process by which observational data are combined with a physics-based model (similar to a climate model, which is discussed later). The model increases the extent of information from locations where observations exist to more data-sparse regions, thus providing estimates of weather/climate for data-sparse regions that are based on physical laws described by the model as well as observations elsewhere. These types of data sets have been used by applied economists studying the developing world (e.g., Guiteras 2010; Hsiang, Meng, and Cane 2011; Schlenker and Lobell 2010), but they have not been widely adopted.

The process of data assimilation is not unlike an economist's use of a structural model to interpolate missing observations. Data assimilation seeks to minimize a loss function subject to a large set of difference equations, which are derived from fundamental physical principles (e.g., the conservation of energy). More recently, such reanalysis efforts have tried to estimate the state of the global environment over a long sequence of periods by optimally fitting a single dynamic model to all those periods simultaneously. This process is difficult and costly, and thus only a few research centers offer regularly updated data sets. The National Center for Environmental Prediction (NCEP) in the United States (Kistler et al. 2001) and the European Center for Medium-Range Weather Forecasting (ECMWF 2010) produce the two most commonly used reanalysis products.

It is important to note that reanalysis output cannot be forced to perfectly match observational data. This is because reanalysis output has both limited resolution and is influenced by the GCM even when observations are present. Moreover, reanalysis is conducted with models that, like economic models, are imperfect and contain systematic biases. Constraining these models with the data that are "fed" into them does not always correct satisfactorily for the model's built-in biased behavior. Although reanalysis provides estimates that may be better than what would otherwise be available for regions where observations are sparse or of poor quality, the reanalysis output for such regions is still basically a model prediction, which is likely to be less accurate than for more observation-rich regions.

Five Potential Pitfalls

We turn next to a discussion of the five main pitfalls of using these weather data products in econometric settings and how to avoid them. In order to examine these issues it is important to understand that studies on the economic impacts of climate change on economic sectors (e.g., agriculture) have used two distinct approaches to estimate response functions. First, early studies relied on cross-sectional variation in weather or climate in different locations to explain variation in the outcome variable of interest (e.g., Kelly, Kolstad, and Mitchell 2005; Mendelsohn, Nordhaus, and Shaw 1994). However, one limitation of the cross-sectional approach is that there may be unobservable variables that vary across these spatial units, which are likely correlated with the climate/weather indicator used. Therefore, recent studies have adopted a second approach that focuses on a panel data analysis, which controls for space and time-fixed effects (e.g., Auffhammer, Ramanathan, and Vincent 2006; Deschenes and

Greenstone 2007; Schlenker and Roberts 2009). Fixed-effects estimators rely on variation across time *within* a spatial unit (e.g., a county) as the source of identifying variation rather than variation *across* these spatial units. This means that the underlying identification relates time series deviations from the location-specific mean in the climate indicators to deviations in the outcome variable of interest.

Pitfall 1: The Choice of Weather Data Set

Although the economic implications of either approach (i.e., long-run versus short-run adaptation in panel versus cross-sectional studies) have been discussed elsewhere (e.g., Lobell and Burke 2010), the practical issue of which weather data set to use has received no attention in the literature. As we will show here, most gridded weather data sets agree on the average value of weather variables across space (i.e., places that are on average hot or cold), but they are not in full agreement about the timing or magnitude of deviations from this mean, which is the source of identifying variation in panel data studies. This is a more serious problem for areas with a small number of weather stations because the data must be interpolated from stations that are further removed and hence might experience idiosyncratic shocks. We show this lack of correlation in the deviations using the three global gridded weather data sets discussed earlier:

(1) The CRU data set (version TS2.1), which uses a statistical interpolation procedure without reanalysis and gives monthly minimum and maximum temperature as well as precipitation on a 0.5×0.5-degree grid (Mitchell and Jones, 2005).
(2) The Willmott et al. (2010) UDEL data set, which uses a statistical interpolation procedure without reanalysis and gives monthly average temperature as well as precipitation on a 0.5×0.5-degree grid.
(3) The reanalysis data from NCEP/National Center for Atmospheric Research (NCAR) (Kistler et al. 2001), which gives daily minimum and maximum temperature and total precipitation on a nonuniform grid (1.875 degrees longitude and roughly 1.90 degrees latitude, although the latter is not evenly spaced).[3]

CRU and UDEL are statistically interpolated, whereas NCEP uses data assimilation with a physical model as discussed earlier. We focus here on two variables that are available in all three data sets: average temperature and total precipitation. We calculate country averages by taking a weighted average across grid cells that overlap a country's boundary for the months of the primary maize growing season (Sacks et al. 2010). We define the growing season as extending from the first of the month in which it begins to the end of the month when it ends because two of the three weather data sets provide only monthly values. In this way we are able to average observations over the same time period for all three data sets. Next, we calculate the weight given to each grid in a county as the share of the country's land area that the grid covers. This allows us to derive the average temperature (the average between the minimum and maximum temperatures for those data sets that report the minimum and maximum) as well as total precipitation by country over the period 1960–1999. Several recent studies have used

[3]This data set is sometimes called the NCEP/NCAR/DOE (i.e., US Department of Energy) reanalysis.

similar country-level aggregates and averages (Dell, Jones, and Olken 2012; Hsiang 2010; Schlenker and Lobell 2010).

Correlations of country-level climate normals across data products

First, we compare average outcomes across locations by deriving average temperature and precipitation over 1960–1999 to get one observation per country. We find that the correlation between the data based on the statistical interpolation procedures (CRU and UDEL) for average temperatures is 0.998, while it is 0.990 between NCEP and either CRU or UDEL. For total precipitation, the correlation between CRU and UDEL for average season-total precipitation is 0.985 and 0.882 between NCEP and CRU (and 0.883 between NCEP and UDEL). This indicates that the three data sources provide similar estimates concerning which areas of the world are hot and which are cold on average. This is a reassuring finding for studies that rely on cross-sectional variation across countries. For both average temperature and average precipitation, the correlation is slightly lower between the reanalysis data (NCEP) and the two statistical interpolation techniques (CRU and UDEL).[4]

Correlations of country-level annual fluctuations across data products

It is difficult to predict how weather variables change year to year when weather is not observed in a specific location or time period. To illustrate this point, we construct *annual* deviations from the country-specific mean in each data set over the 1960–1999 period. This provides us with a 40-year *panel* rather than a single cross section of normals. We find that for average temperature, the correlation coefficients between models are significantly lower compared to those discussed earlier. The pairwise correlation coefficients are CRU and UDEL: 0.917; NCEP and CRU 0.742; NCEP and UDEL: 0.724. For precipitation, the correlation coefficients across data sets are even lower. This is likely due to the fact that precipitation is less smooth than temperature in space and time, which makes the extrapolation algorithm employed more important.[5] The pairwise correlation coefficients are CRU and UDEL: 0.698; NCEP and CRU: 0.299; and NCEP and UDEL: 0.269. While the correlations are especially low when we compare deviations in the reanalysis data (NCEP) to the statistical interpolation methods (CRU and UDEL), the drop to a correlation coefficient below 0.7 for CRU-UDEL is significant as well because both methods are statistical interpolation routines using raw station data. So whether an outcome is above or below normal—and by how much—depends crucially on which weather data set is being used.

Across-country heterogeneity in correlations of annual fluctuations

The average correlations for both cross-section and panel data mask considerable heterogeneity by country. To illustrate this point, we construct weather shocks by weighting each grid cell

[4]We also examined a fourth data set, NCC (Ngo-Duc, Polcher, and Laval 2005), which is a hybrid of NCEP and CRU. That is, it scales the NCEP reanalysis output using a constant monthly factor so that the 1948–2000 average equals the CRU average. Not surprisingly, for our 1960–99 sample, the correlation between NCC and CRU for average season-total precipitation exceeds 0.99.

[5]For some reanalysis products, precipitation is generated by the model even if precipitation observations exist.

by the amount of maize that is grown within it (Monfreda, Ramankutty, and Foley 2005), a common approach for examining the agricultural impacts of climate change. We find that the pairwise correlation coefficients among weather deviations for season-total precipitation in the United States are CRU and UDEL: 0.963; NCEP and CRU: 0.758; and NCEP and UDEL: 0.714. The deviations are much more highly correlated than when all countries are included, presumably because of the good observational network in the United States. In contrast, precipitation shocks constructed over the maize-growing area in Mexico have pairwise correlation coefficients of CRU and UDEL: 0.726; NCEP and CRU: 0.069; and NCEP and UDEL: 0.307. This illustrates that in regions with limited monitoring networks, which is generally the case in the developing world, the weather shock used to identify response coefficients in econometric estimation varies significantly depending on which data source is used.

In summary, when economists are conducting panel studies that rely on deviations from averages, they should be careful about which data source they use because measurement errors—and related statistical concerns such as attenuation bias—are amplified by demeaning explicitly or via fixed effects (Fisher et al. 2012). Conducting sensitivity checks by using more than one data source can be helpful in determining whether the results are robust.

Pitfall 2: Averaging Daily Station-Level Data across Space

Another pitfall of using weather data products in econometric estimation concerns averaging station-level data across space. Several economic studies that link economic outcomes to weather (or control for weather) use inverse distance-weighted averages for the closest available weather stations (see, e.g., Deschenes and Greenstone 2007; Mendelsohn et al. 1994). As with the panel versus cross-section data issue, such an approach works well for a cross-sectional analysis but becomes problematic when fixed effects are included in a panel data setting, especially when both location and time-fixed effects are included. This is because weather station data are sometimes missing (i.e., not only do weather stations come in and out of existence, they are also often turned off or values are simply not recorded). A time series of inverse distance-weighted averages of weather station data is likely to include variation from the birth and death of stations and observations that are missing for a given period. When location fixed effects remove average weather outcomes at the interpolated location, and temporal fixed effects are included, the remaining weather variation is greatly diminished and the variation that is due to stations coming in and out of the sample can potentially account for a significant share of the overall variance. For example, Fisher et al. (2012) provide an example where the noise-to-signal ratio after removing location and temporal fixed effects is 7:1; that is, the measurement error greatly exceeds the variation used in the identification, which is likely to result in significant attenuation bias in estimation.

A possible alternative to averaging weather station data that report weather indicators on a given day is to first fill in missing weather station data by regressing it on the closest surrounding stations and then to predict missing observations at a station (Auffhammer and Kellogg 2011; Schlenker and Roberts 2009). Then the full weather record is derived by interpolating a balanced panel of "patched" weather station data. This approach keeps the set of stations that are used in the interpolation constant and ensures that the resulting variation is not caused by variation in station coverage.

Pitfall 3: Correlation of Weather Variables

The third pitfall relates to the classic omitted variables problem. Many economic studies, including (but not limited to) those estimating climate change impacts, have focused on the impact of one weather variable in isolation, for example, regressing income only on precipitation shocks (Miguel, Satyanath, and Sergenti 2004). While precipitation shocks are exogenous and hence a plausible instrument for income, it is important to note that to the extent that precipitation and temperature are correlated, the coefficient on precipitation will measure the combined effect of the two variables. This is particularly important in the climate change context if the estimated coefficient is used to estimate climate change impacts under a climate influenced by human activity. In order to obtain unbiased estimates of the effects of changes in precipitation and temperatures, which are historically correlated, both variables must be included in the regression equation, especially if the correlation is predicted to change in the future.

To underline the importance of this observed correlation between different climate indicators, figure 1 shows the correlation coefficients between annual average temperature and total precipitation for each of the CRU (version TS2.1) time series (TS) grid cells for the years 1960–1999. The map indicates clearly that the correlations vary greatly and that there are regions with both significant positive and significant negative correlation between precipitation and temperature. This implies that if one controls for only one of the two weather variables in a regression, the sign of the omitted variable bias will depend on the location under study. Hot areas generally show negative correlation (as high as −0.7) because more precipitation and the associated evaporation results in cooling and lower average temperatures. In contrast, a positive correlation is generally observed in cooler areas because increased precipitation is associated with the import of warm and humid tropical air, and cloud cover keeps the underlying surface warmer. It is noteworthy that some large and not-so-large countries have areas of both negative and positive correlation (e.g., United States, Russia, France, and Spain).

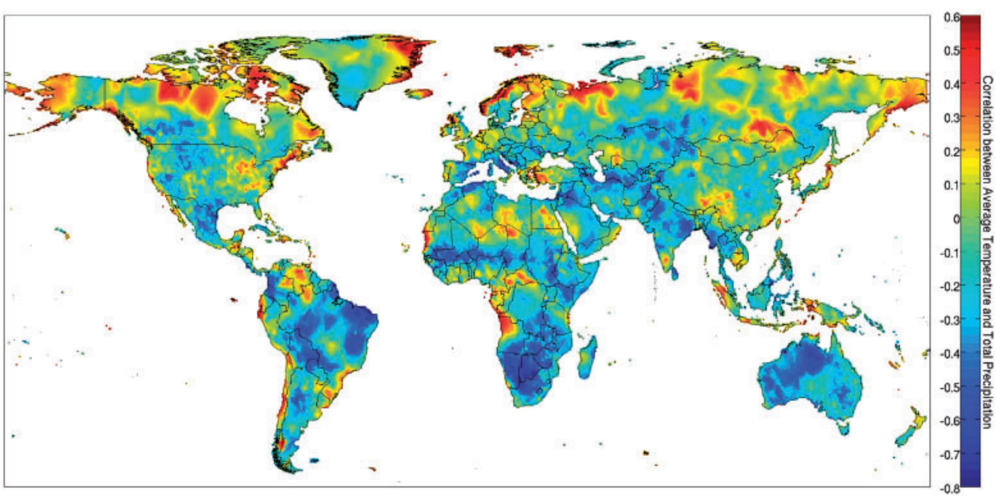

Figure 1 Correlation between annual average temperature and total precipitation in each grid cell (CRU TS 2.1 Data 1960–1999)
Source: Authors' calculations.

It is also important to note that climatic variables other than temperature and precipitation (e.g., relative humidity, solar radiation, wind speed and direction) may bias empirical estimates through the classic omitted variables problem. The existence of these other variables and their correlation with temperature or precipitation may be location specific. For example, in a panel regression with country and year fixed-effects and country-specific trends, Hsiang (2010) finds that exposure to hurricane winds in Caribbean Basin countries is correlated over time with a country's local surface temperature, with each 1 degree Celsius increase in a country's summer surface temperature being correlated with a 2.6 (± 1.2) meter per second increase in area-averaged wind exposure in that country. This increase in wind exposure is substantial, since it raises expected hurricane damages by 20 percent (Hsiang and Narita 2012), suggesting potentially biased estimates of temperature impacts if wind exposure is excluded from the analysis.

In summary, if temperature, precipitation, and other atmospheric variables are correlated, a study that seeks to extrapolate (based on an estimated response function) potential climate impacts must include all of these variables in order to obtain an unbiased estimate of the effect of each variable.

Pitfall 4: Spatial Correlation

Climate variables are inherently correlated across space and time. While variation in weather is often considered random across time, variation across space displays significantly less "randomness," especially at smaller spatial scales. This means that some of the weather or climate variables that we use in econometric estimation are highly spatially correlated and that estimates of standard errors will be biased unless steps are taken to correct for spatial correlation.

To provide a sense of the degree of spatial correlation in these data sets, figure 2 shows the average correlation of annual mean temperature at each CRU (version TS 2.1) grid cell

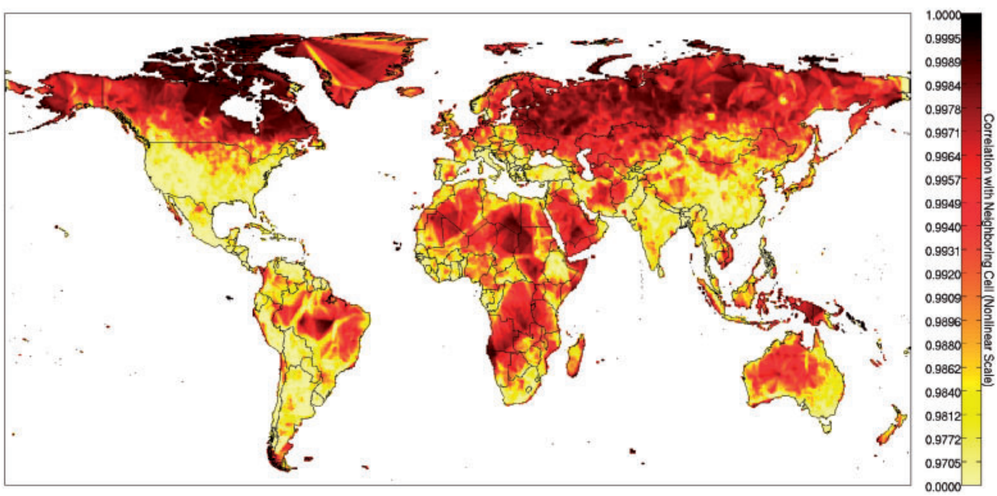

Figure 2 Correlation of average annual temperature at CRU grid with surrounding eight grid cells (CRU TS 2.1 Data 1960–1999)

Note: We have chosen a highly nonlinear scale (correlation to the power of 100) because all correlations are extremely close to 1.

Source: Authors' calculations.

with the eight surrounding grid cells for the 1960–1999 period. As discussed earlier, errors might propagate from one grid cell to the next for both interpolated station data and data assimilation methods. If the model correctly accounts for all weather variables, the spatial dependence of the regressors will not be a problem. Most economic studies to date control only for temperature and precipitation. However, other weather variables such as wind direction, humidity, and vapor pressure might also have an impact, and these omitted variables are presumably spatially correlated as well. If they have a causal effect on the outcome of interest, as for example, the effect of vapor pressure deficit (which is closely related to relative humidity) on crop yields (Roberts, Schlenker, and Eyer 2013), then they become part of the error term, which will then also be spatially correlated. Thus it is imperative to take this spatial correlation into account in econometric estimation.[6]

There are three main approaches to account for spatial correlation:

(1) Use a spatial weighting matrix. This is most efficient when the weighting matrix is known, but it will result in biased estimates if the weighting matrix is misspecified;
(2) Use the nonparametric approach provided by Conley (1999), which does not require one to specify a weighting matrix; or
(3) Use a grouped bootstrap where years are resampled and replaced. This approach requires that year-to-year fluctuations be random, but errors within a year can be correlated.[7]

Finally, it is important to note that many of the gridded weather data sets we have discussed simply interpolate station data. In data-sparse regions, several grids might be linked to the same set of weather stations. This will lead to highly multicollinear weather variables that do not allow for proper identification (especially in a panel setting where grid averages are removed) because the remaining variation is simply due to the fact that slightly different weights have been used for different weather stations.

In summary, one has to adjust for spatial correlation to obtain unbiased standard errors and valid confidence bands.

Pitfall 5: Endogenous Weather Coverage

The final pitfall concerns why weather stations are observed in some areas and time periods and not in others. One strand of the economics literature examines how the relationship between weather variables and economic variables of interest might change due to large policy changes, such as a country becoming independent, or an extreme exogenous shock, such as a natural disaster (Kahn 2005). The most obvious method for accounting for such changes is the now standard difference-in-difference analysis. One concern with this approach is that if the degree of measurement error varies between the pre- and postintervention (or event) date, the treatment effect estimate will very likely be biased because of classic attenuation bias concerns. However, if weather variables are measured consistently, the difference-in-difference

[6]This will generally result in significantly larger standard errors. For example, Schlenker and Roberts (2009) find that accounting for spatial correlation increases standard errors by a factor of 6.

[7]However, in many areas of the world, the independence of year-to-year variation is questionable because of planetary-scale climate oscillations, such as the El Niño-Southern Oscillation, which may be autoregressive (Hsiang et al. 2011).

regression design will be free of this bias. Thus it is important that weather station coverage not change with the policy change (or exogenous shock) because it could introduce measurement error and result in a downward bias in the estimated coefficients in the postintervention period.[8]

To examine this issue in more detail, we downloaded daily data from the Global Summary of the Day database maintained by NOAA's National Climatic Data Center (NOAA 2011a), counted the number of days a weather station within a country had nonmissing observations, and summed it across all stations. This provides the total count of daily station-level observations in a country. While most countries show an upward trend in this measure over time, the results for some transition countries are striking. For example, Romania had an upward trend until it peaked at 67,727 station-days in 1988. Following the fall of the Iron Curtain in 1989, the number decreased rapidly until it stabilized around 11,000 station-days in 2003–2007, decreasing coverage by a factor of 6. This suggests that the results from a difference-in-difference analysis of how, for example, farmers responded to weather shocks before and after the fall of the Iron Curtain would have to be interpreted with caution.

In summary, when using any of the gridded data products available, it is crucial to determine whether the underlying station data have changed over time (i.e., before and after a major shock or event).

Climate Models and Their Output

If the econometrician has successfully estimated the causal "dose–response" relationship between socioeconomic outcomes and historical weather or climate data, often the logical next step is to use that estimated relationship to predict future impacts due to anthropogenic climate change. This step requires making forecasts of future climate under the assumption of heightened atmospheric concentrations of greenhouse gases, which is usually accomplished by employing output from a spatially explicit physics-based model of the global climate, which, as discussed in the introduction, is known as a GCM. This section provides an overview of GCMs and discusses some of the major potential pitfalls of using these models in the simulation of future economic impacts of climate change.

Components and Properties of GCMs

Although GCMs have several components that are parameterized using statistical procedures, the core of every GCM is a set of deterministic mathematical equations that describe the laws of motion for a fluid. These laws were derived in fluid mechanics laboratories over centuries, and GCMs use numerical approximations of these laws. To solve these equations, GCMs approximate the atmosphere and ocean, which are continuous fluids, with some form of numerical discretization. The simplest way to visualize this procedure (although it is less sophisticated than what is typically used in current practice) is a three-dimensional grid

[8]This issue is closely related to the discussion about spatial correlation that is due to different interpolation methods over a sparse data matrix.

of "boxes," each of which possesses several state variables, for example temperature or air pressure, which vary across space from one box to the next and evolve over time but are uniform within each box.[9] Given a three-dimensional structure of these state-variables at time t, a GCM solves for the variables' structure at time $t+1$ using the model's numerical representation of fluid-mechanical laws. Following an initialization that specifies the structure of these variables in the very first time period, GCMs iteratively repeat this calculation for time-steps of about thirty minutes, gradually constructing a projection for the future state of the world.[10]

GCMs typically take forecasts of human activity as exogenous. To make climate projections across different GCMs comparable, modelers simulate future climate outcomes under a set of standardized "scenarios" that exogenously prescribe a time series of future greenhouse gas emissions, aerosols, and other short-lived pollutants based on demographic, economic, and regulatory assumptions.[11]

When the emissions scenario is held fixed, GCMs differ primarily in their numerical representations of the climate's state and its various processes. Having discretized the atmosphere and ocean with grid cells (the previously mentioned "boxes") of various resolutions,[12] GCMs selectively represent processes that occur on spatial scales smaller than these grid cells—known as "sub-grid scales"—using "parameterizations," which are formulations that are not based as directly on the known laws of physics as are the resolved fluid dynamics, but incorporate a greater degree of empiricism or theoretical construction. For example, chemical reactions, vegetation responses, cloud formation, and rainfall are all sub-grid scale processes whose numerical representations may vary across GCMs (Sect. 8.2 in IPCC 2007). Unlike the core fluid-mechanical equations that have a standard representation in a discretized global model, there is no standard representation of these sub-grid scale processes, and thus the improvement of their representation in GCMs continues to be an active area of research. There have been various community efforts to try to accelerate advances in this area by comparing the performance of models, most notably the Coupled Model Intercomparison Project, or CMIP (see Meehl et al. 2007), and conducting studies that attempt to score the forecast ability, known as "skill," of different models along various dimensions (see Reichler and Kim 2008, and Sect. 8.3 and 8.4 of IPCC 2007). Different models have different "skill," and thus we advise economics researchers who are studying specific regions or processes and are interested in selecting a GCM projection to first consult the appropriate literature as well as specialists in the field.[13]

[9]The Intergovernmental Panel on Climate Change (IPCC 2011a) provides a brief description and graphic to illustrate this structure.

[10]For introductory materials on the structure of these models, see Tebaldi and Knutti (2010) and Section 8.1.3 of IPCC (2007). For more advanced descriptions, consult Warner (2011) or IPCC (2007). Donner, Schubert, and Somerville (2011), and Weart (2008, 2011) provide detailed histories of GCM development.

[11]These assumptions and their resulting scenarios were established in the IPCC's Special Report on Emissions Scenarios (SRES) (IPCC 2000) and are summarized in IPCC (2011b).

[12]See the supplemental tables to Reichler and Kim (2008) or IPCC Scientific Basis Table 8.1 (http://www.ipcc.ch/publications_and_data/ar4/wg1/en/ch8s8-2.html) for a concise summary of these climate model properties.

[13]For researchers seeking intuition for the numerical setup of GCMs, they can download a one-dimensional climate model tutorial built by the Goddard Institute of Space Studies for teaching purposes (http://icp.giss.nasa.gov/education/geebitt/). More ambitious researchers can download and run a full open source GCM, the Community Earth System Model, produced by the National Center for Atmospheric Research (http://www.cesm.ucar.edu/models/cesm1.0/).

Differences in Model Predictions

There are over twenty well-known climate models, all with readily available output.[14] This section discusses how the choice of GCM affects estimates of climate impacts. In their survey of economic assessments of climate change impacts, Burke et al. (2011) found that 50 percent of the studies used the model developed by the Hadley Centre to calculate economic climate change impacts across a variety of sectors, of which 17 percent used only the Hadley model. Among health impact studies, 38 percent relied on the Hadley model alone. However, there is no evidence that the Hadley model or, in fact, any other model should be the preferred climate model to use. This is supported by the fact that for some climate indicators, such as precipitation, the predictions for certain regions vary dramatically across models. In the extreme, some models predict wetter summers for West Africa and others predict drier summers—all using the same SRES scenarios.

One way to address the challenge of having to choose one GCM is to use model or ensemble averages (e.g., Tebaldi and Knutti 2007). This decreases the reliance on a single model. However, we believe it is important to either report the impacts for a number of climate models separately or to average them and indicate the variability in impacts across models. This is not difficult to do, and, given the low costs of data storage on personal computers and the access to free bandwidth for most academics, there is no reason not to. Alternatively, if predicted changes within a study area vary more across than within climate models, then presenting a set of uniform scenarios might be informative and also highlight the sensitivity of the results.

We next turn to a set of issues that arise when one tries to match the time and spatial scale of the GCM to that of the econometric model for simulation purposes.

Aggregation Bias

As described earlier, GCMs effectively divide the earth's surface into a discrete grid, where there is variation in climate across discrete grid cells, but climate statistics are homogenous within each cell. For example, if one uses a climate model that provides output on a monthly basis, it is assumed that temperatures within the month and among all locations within the grid cell are constant.[15] Such temporal and spatial aggregation might be inappropriate and produce biased impact estimates. While many models are being run at a resolution that is higher than 2×2 degree (for the next IPCC [AR5] report), most of the economic impact studies in the existing literature use model output at a 2×2-degree or coarser resolution. While a 2×2-degree cell may be "small" from the perspective of the global climate, it is not small from the perspective of *human* systems. For example, a 2×2-degree grid spacing at the equator is equivalent to a grid width of 222 kilometers (138 miles). It is not hard to imagine that a stretch of this length will have vastly varying climates (e.g., driving east from San Diego's coastal

[14]Climate projections from GCMs running IPCC's Special Report on Emissions Scenarios are available free of charge, and model output can be downloaded from the IPCC's data distribution website (http://www.ipcc-data.org/) or the CMIP data distribution website (https://esg.llnl.gov:8443/index.jsp). For summaries of climate projections from GCMs running SRES scenarios, see IPCC (2007), chapter 10, for global summaries and chapter 11 for regional summaries. The IPCC also provides an interactive data visualization application online (http://www.ipcc-data.org/ddc_visualisation.html).
[15]Some models have within-grid deterministic variation, but this is a relatively recent effort.

climate to El Centro's dry and hot desert climate). This aggregation issue becomes especially problematic if the underlying topography is mountainous or located near the ocean.

Quantifying aggregation bias

To examine the severity of this aggregation bias, we compare average temperatures predicted by the Hadley III GCM to a fine-scaled (2.5 × 2.5-mile grid) weather data set (PRISM 2009) for the 48 contiguous United States (see figure 3). Figure 3 shows quite clearly that this bias is most significant in mountainous areas, which are also usually less populated areas. At the extremes, we see that the bias can reach +25°C at some mountaintops. This is not surprising because surface temperatures tend to fall about 7°C per 1,000 m in elevation, which means that mountains are much colder than areas at lower elevations in the same grid cell. The aggregation bias exists not only for remote mountainous regions but also for heavily populated areas, which are often located near oceans. In fact, figure 3 indicates that the entire Western Seaboard has biases, and that those biases are significantly greater than any predicted warming. The average absolute difference in temperature across the entire United States is 3.0°C and the root mean squared error is 4.0°C, which are both substantially larger than the average predicted changes by the end of the century under the SRES forced climate change scenarios. This means that if one simply interprets GCM output at a grid cell as an unbiased forecast of climate at any location in that grid cell, the bias may be a much larger driver of projected impacts than actual warming.

Moreover, while the severity of the aggregation bias varies by location, it also varies by the climate indicator one is using. For example, if we use the *annual* mean temperature rather than the average daily maximum July temperature, the absolute error reduces to

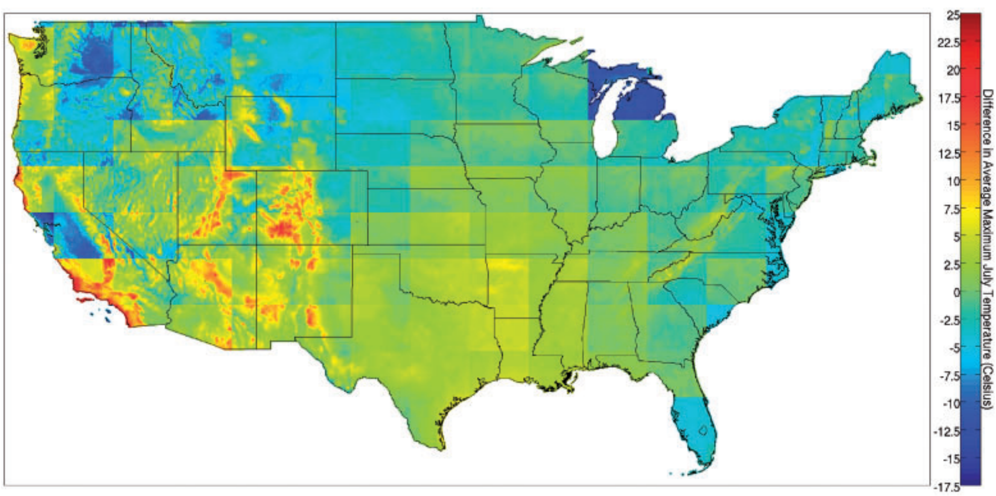

Figure 3 Aggregation bias: Hadley grid averages versus PRISM grid averages in each PRISM grid (1961–1999)

Notes: The figure plots the difference in the average daily maximum temperature in the month of July in the years 1960–1999 between the GCM (Hadley III), which has the coarser resolution, and the fine-scale weather grid (PRISM 2009). A positive number indicates that the GCM grid average exceeds the PRISM average, which is based on interpolated station data.

Source: Authors' calculations.

1.8°C and the root mean squared error to 2.4°C. Thus the magnitude of the bias varies by location and indicator used.

This bias is especially relevant for studies of the economic impacts of climate change. These studies generally parameterize a response function between, for example, electricity demand and temperature, using observations from a weather station–based data set and observed electricity demand. In order to calculate the counterfactual electricity demand under a scenario with climate change, one must have a baseline climate and a counterfactual climate. However, if one uses an average of observed gridded weather products as the baseline climate and predictions of climate from a GCM as the counterfactual climate at a future date, the resulting estimated impacts will be due to both the simulated warming *and* the bias displayed in figure 3. If the response function is nonlinear in weather/climate, as has been shown to be the case in agriculture (e.g., Schlenker and Roberts 2009) and electricity demand (e.g., Auffhammer and Aroonruengsawat 2011, 2012), then this bias may be amplified or offset depending on the nature of the nonlinearity. However, in either case, the resulting impact estimates will be biased. We next turn to a simple approach, which overcomes this issue.

Correcting aggregation bias

The literature has suggested several ways to correct such biases. In addition to using climate models with finer resolutions, the most commonly used approach is based on regression methods, whereby the researcher establishes a correlation between the historical grid values from the GCM and local station-based data and then uses this fitted regression relationship with future values of GCM output to arrive at “downscaled” GCM predictions.[16] Fowler, Blenkinsop, and Tebaldi (2007) provide a review of the main approaches used in practice and compare their performance at selected locations. They note that there is a large literature examining the performance of different downscaling approaches for different regions and climate variables. They conclude that there is no single best approach for all variables (e.g., maximum temperature, rainfall, wind speed) and locations. Moreover, they find that downscaled versions of all GCMs at a desired temporal resolution covering all regions of interest are simply not available. If one is interested in daily values, which are important for many economic applications, including agriculture and electricity demand, then a downscaled version of a climate model delivering daily output is needed. Such data sets are available for some regions, such as California (Cayan 2009), or at coarser time resolution nationally (e.g., Maurer et al. 2007) and globally (e.g., Maurer 2009).

In the absence of an appropriate downscaled data set for the region and time resolution of interest, the most common practice is to derive predicted *changes* for each (coarse) GCM grid and then add these to an average of the historic baseline data used in the parameterization of the response function, thereby preserving within-GCM grid variation. This approach subtracts out the location-specific bias only if this bias is stationary in time. However, this approach shifts the historic time-series at a location by the predicted change, leaving its variance unchanged. If researchers are concerned about predicted changes in the mean *and* the variance, then the fine-scaled historic deviations from location-specific averages can be rescaled by the ratio of the predicted variance at the GCM grid in the future relative to the baseline. It should be noted,

[16]Innovations on this basic approach have involved nonlinear estimation, neural networks, and Bayesian methods.

however, that there is much less consensus among models concerning the predicted changes in the variance than in the mean.

In summary, it is crucial that economists not simply use GCM output as a direct forecast of future climate when estimating impacts relative to a weather station–based baseline climate. One simple solution is to simply add the predicted change in weather to the baseline climate when calculating impacts.

Conclusions

This article has reviewed the most common gridded weather products and outlined five pitfalls when using them as regressors in econometric models. More specifically, we have emphasized that weather anomalies (deviations from normal) vary greatly between data sources and are highly correlated between weather measures and across space. Researchers need to address these issues when constructing and using weather shocks. We have also discussed the basic features of GCMs and examined issues related to spatial scale when using these models in the estimation of the economic impacts of climate change.

In closing, we want to emphasize that when using gridded data sets of historical or future climate, it is important to recognize that both types of data sets are very different from observed weather. Moreover, although historical gridded data products are very convenient because they often provide highly disaggregated weather for large geographic regions over long time periods, this increased coverage comes at a cost. That is, the birth and death of weather stations, the frequent occurrence of missing values, and the spatial correlation introduced by extrapolation algorithms all create potential biases in the estimated coefficients and standard errors if one uses these weather products as independent variables in econometric analyses. In addition, when using GCM output as a counterfactual of future climate, the choice of model has significant implications for the sign and magnitude of the estimated impacts. This means it is important to account for the location-specific biases of each model in order to avoid causing further biases in estimates of the economic impacts of climate change.

References

Angrist, Joshua D., and Alan B. Krueger. 2001. Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives* 15: 69–85.

Auffhammer, Maximilian, and Anin Aroonruengsawat. 2011. Simulating the impacts of climate change, prices and population on California's residential electricity consumption. *Climatic Change* 109 (S1): 191–210.

———. 2012. Erratum: Simulating the impacts of climate change, prices and population on California's residential electricity consumption. *Climatic Change* 113 (3–4): 1101–4.

Auffhammer, Maximilian, and Ryan Kellogg. 2011. Clearing the air? The effects of gasoline content regulation on air quality. *American Economic Review* 101: 2687–2722.

Auffhammer, Maximilian, V. Ramanathan, and Jeffrey R. Vincent. 2006. Integrated model shows that atmospheric brown clouds and greenhouse gases have reduced rice harvests in India. *Proceedings of the National Academy of Sciences* 103: 19668–72.

Burke, Marshall, John Dykema, David Lobell, Edward Miguel, and Shanker Satyanath. 2011. Incorporating climate uncertainty into estimates of climate change impacts, with applications to U.S. and African agriculture. NBER Working Paper 17092.

Cayan, Dan, Mary Tyree, Tapash Das, Mike Dettinger, Hugo Hidalgo, Ed Maurer, Peter Bromirski, Nicholas Graham, and Reihard Flick. 2009. Climate change scenarios and sea level rise estimates for the California 2009 climate change scenarios assessment. California Energy Commission PIER Program Report CEC-500-2009-014-F.

Climatic Research Unit (CRU). University of East Anglia. 2013. CRU Time Series (TS) high resolution gridded datasets 3.2. NCAS British Atmospheric Data Centre. http://badc.nerc.ac.uk/view/badc.nerc.ac.uk__ATOM__dataent_1256223773328276.

Conley, Timothy G. 1999. GMM estimation with cross sectional dependence. *Journal of Econometrics* 92 (1): 1–45.

Dell, Melissa, Benjamin F. Jones, and Benjamin A. Olken. 2012. Climate change and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics* 4 (3): 66–95.

Deschenes, Olivier, and Michael Greenstone. 2007. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather. *American Economic Review* 97 (1): 354–85.

Donner, L., W. Schubert, and R. Somerville. 2011. *The development of atmospheric general circulation models.* New York: Cambridge University Press.

European Center for Medium-Range Weather Forecasting (ECMWF). 2010. IFS documentation—Cy36r1. Part II: Data assimilation. http://www.ecmwf.int/research/ifs-docs/CY36r1/ASSIMILATION/IFSPart2.pdf.

Fisher, Anthony C., W. Michael Hanemann, Michael Roberts, and Wolfram Schlenker. 2012. The economic impacts of climate change: Evidence from agricultural output and random fluctuations in weather: Comment. *American Economic Review* 102 (7): 3749–60.

Fisher, Ronald A. 1925. The influence of rainfall on the yield of wheat at Rothamsted. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character* 213: 89–142.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi. 2007. Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology* 27 (12): 1547–78.

Guiteras, Raymond. 2010. The impact of climate change on Indian agriculture. University of Maryland Working Paper.

Hsiang, Solomon M. 2010. Temperatures and tropical cyclones strongly affect economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences* 107 (35): 15367–72.

Hsiang, Solomon M., Kyle C. Meng, and Mark A. Cane. 2011. Civil conflicts are associated with the global climate. *Nature* 476: 438–41.

Hsiang, Solomon M., and Daiju Narita. 2012. Adaptation to cyclone risk: Evidence from the global cross-section. *Climate Change Economics* 3 (2): 1–28.

Intergovernmental Panel on Climate Change (IPCC). 2000. *Special report emissions scenarios* (SRES).

———. 2007. Climate change 2007: Working group I: The physical science basis. http://www.ipcc.ch/publications_and_data/ar4/wg1/en/.

———. 2011a. What is a GCM? http://www.ipcc-data.org/ddc_gcm_guide.html.

———. 2011b. *Environmental data and scenarios.* http://www.ipcc-data.org/ddc_envdata.html.

Kahn, Matthew. 2005. The death toll from natural disasters: The role of income, geography, and institutions. *Review of Economics and Statistics* 87 (2): 271–84.

Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell. 2005. Adjustment costs from environmental change. *Journal of Environmental Economics and Management* 50 (3): 468–95.

Kistler, Robert, Eugenia Kalnay, William Collins, Suranjana Saha, Glenn White, John Woollen, Muthuvel Chelliah, Wesley Ebisuzaki, Masao Kanamitsu, Vernon Kousky, Huug van den Dool, Roy Jenne, and Michael Fiorino. 2001. The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society* 82: 247–67.

Lobell, David B., and Marshall Burke. 2010. *Climate change and food security*. Chapters 5 and 6. New York: Springer.

Maurer, Edwin P. 2009. http://www.engr.scu.edu/~emaurer/global_data/.

Maurer, Edwin P., Levi Brekke, Tom Pruitt, and Phillip B. Duffy. 2007. Fine-resolution climate projections enhance regional climate change impact studies. *EOS Transaction—American Geophysical Union* 88 (47): 504.

Meehl, Gerald A., Curt Covey, Thomas Delworth, Mojib Latif, Bryant McAvaney, John F. B. Mitchell, Ronald J. Stouffer, and Karl E. Taylor. 2007. The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society* 88: 1383–94.

Mendelsohn, Robert, William D. Nordhaus, and Daigee Shaw. 1994. The impact of global warming on agriculture: A Ricardian analysis. *American Economic Review* 84 (4): 753–71.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti. 2004. Economic shocks and civil conflict: An instrumental variables approach. *Journal of Political Economy* 112 (4): 725–53.

Mitchell, Timothy D., and Philip D. Jones. 2005. An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology* 25 (6): 693–712.

Monfreda, Chad, Navin Ramankutty, and Jonathan A. Foley. 2005. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Global Biogeochemical Cycles* 22: 1–19.

Ngo-Duc, T., J. Polcher, and K. Laval. 2005. A 53-year forcing data set for land surface models. *Journal of Geophysical Research* 110: 1–13.

National Oceanic and Atmospheric Administration (NOAA). 2011a. Daily global historical climatology network. ftp://ftp.ncdc.noaa.gov/pub/data/ghcn/daily.

———. 2011b. Daily surface weather data. http://www.ncdc.noaa.gov/oa/climate/climatedata.html#surface.

Phillips, Norman A. 1956. The general circulation of the atmosphere: A numerical experiment. *Quarterly Journal of the Royal Meteorological Society* 82 (352): 123–64.

PRISM. 2009. Parameter-elevation regressions on independent slopes model. Oregon State University (http://www.prism.oregonstate.edu/).

Reichler, Thomas, and Junsu Kim. 2008. How well do coupled models simulate today's climate? *Bulletin of the American Meteorological Society* 89: 303–11.

Roberts, Michal J., Wolfram Schlenker, and Jonathan Eyer. 2013. Agronomic weather measures in econometric models of crop yield with implications for climate change. *American Journal of Agricultural Economics* 95 (2): 236–43.

Sacks, William J., Delphine Deryng, Jonathan A. Foley, and Navin Ramankutty. 2010. Crop planting dates: An analysis of global patterns. *Global Ecology and Biogeography* 19: 607–20.

Schlenker, Wolfram, and David B. Lobell. 2010. Robust negative impacts of climate change on African agriculture. *Environmental Research Letters* 5 (1): 1–8.

Schlenker, Wolfram, and Michael J. Roberts. 2009. Nonlinear temperature effects indicate severe damages to U.S. crop yields under climate change. *Proceedings of the National Academy of Sciences* 106 (37): 15594–98.

Tebaldi, C., and R. Knutti. 2007. The use of the multi-model ensemble in probabilistic climate projections. *Philosophical Transactions of the Royal Society A* 365 (1857): 2053–75.

———. 2010. Climate models and their projections of future changes. In *Climate change and food security*, ed. D. Lobell, and M. Burke. New York: Springer.

Warner, T. T. 2011. *Numerical weather prediction.* New York: Cambridge University Press.

Weart, S. 2008. *The discovery of global warming.* Cambridge, MA: Harvard University Press.

———.2011. *General circulation models of climate.* http://www.aip.org/history/climate/GCM.htm.

Willmott, C. J., K. Matsuura, and D. R. Legates. 2010. Terrestrial air temperature: 1900–2008 gridded monthly time series. http://climate.geog.udel.edu/~climate/html_pages/download.html.

Wright, Philip G. 1928. *The tariff on animal and vegetable oils.* New York: Macmillan.