Global climate models and climate data: A user guide for economists

Article · March 2011

4 authors, including:

GLOBAL CLIMATE MODELS AND CLIMATE DATA:

A USER GUIDE FOR ECONOMISTS[1]

Maximilian Auffhammer
University of California Berkeley & NBER

Solomon M. Hsiang
Columbia University & NBER

Wolfram Schlenker
Columbia University & NBER

Adam Sobel
Columbia University

February 2011

Abstract

This paper provides an introduction to climate data sets, climate models, and the output they provide. We identify five major issues related to the use of gridded historical weather products and climate model output: Spatially specific prediction error due to the coarse resolution of climate model output; spatially varying correlations between different weather indicators; differences in cross sectional versus time series variation, strong spatial correlation in weather variables; and finally the endogeneity of weather coverage. We discuss the conceptual issues arising from these five issues for economic impact studies of climate change.

Keywords: Global Climate Models, Downscaling, Climate Change Impacts

JEL Code: Q54

---

1. INTRODUCTION

The Intergovernmental Panel on Climate Change (IPCC) defines climate "[…] in a narrow sense […] as the "average weather", or more rigorously, as the statistical description in terms of the mean and variability of relevant quantities over a period of time ranging from months to thousands or millions of years. The classical period is 30 years, as defined by the World Meteorological Organization (WMO). These quantities are most often surface variables such as temperature, precipitation, and wind. Climate in a wider sense is the state, including a statistical description, of the climate system." While average temperature, rainfall and windspeed may describe many gross features of climate, they are certainly not sufficient statistics. Yet, while there are plausibly a very large number of statistics and additional variables needed to fully describe a climate (e.g. the probability distribution of many variables -- such as solar radiation in the ultraviolet -- for all moments in time through a year), some approximations must be made for practical reasons. Climatologists study the statistics and dynamics exhibited in this vast state space. Economists, on the other hand, would like to distill the knowledge of climatologists to those aspects of climate that seem most immediately relevant to society. Unfortunately, we still have very little grasp of the relative social importance of very basic climatological parameters; thus it is still unknown what approximations are acceptable. For instance, we are only beginning to understand simple impacts of average temperatures; or the impact of different mean-preserving temperature distributions on society (e.g., Schlenker and Roberts, 2009; Hsiang, 2010). Until we gain a better understanding of what aspects of the climate can be ignored from a social scientist's perspective, we should recognize climate as the conditional distribution of a large random vector of environmental parameters, only some of which can be included in any analysis for practical reasons.

Climate science and economics share some basic features. Both disciplines predominantly have to rely on observational data that are used to test hypotheses derived from mathematical models. At the same time, both the tools used in research and the obstacles to explanatory success can differ strikingly in the two disciplines, e.g., the interpretation of uncertainty or the salience of complex simulations. Describing some of these differences and why they are important for

economic analyses is the goal of the next section.   In the third section, we discuss climate data more generally and point out five statistical concerns users should be aware of before using observational or model data in economic analysis.

2. WHAT IS CLIMATE AND WHAT ARE CLIMATE MODELS?

"Weather" refers to the instantaneous state of the atmosphere, or to the atmosphere's evolution over short periods of time (i.e., days).   Because climate is the expected distribution of weather, a weather model can be used to predict the climate if run for sufficiently long periods to allow adequate sampling of that distribution.   Climate models most commonly discussed and used in modern climate science are, in their essence, just this: weather models run for a long time.   It is nonetheless also possible to produce a climate model that does not predict daily weather outcomes but instead models some moments of the distributions of key variables (typically means).   For example, the most primitive climate models predict global mean temperature by assuming an atmosphere in which all variables are spatially uniform and temporally constant across the globe.   In equilibrium, the rate at which solar radiation is absorbed by the earth has to equal the rate at which the earth emits infrared radiation back into space.

Modern climate or weather models examine the time evolution of the atmosphere and ocean by differential equations founded in the laws of classical physics.   In general it is impossible to derive closed-form solutions of these equations.   Models therefore rely on numerical methods to obtain approximate solutions.   General Circulation Models (GCMs) are the most sophisticated climate models and are the type of model currently used by the Intergovernmental Panel on Climate Change (IPCC).   Heuristically, GCMs break up the three-dimensional atmosphere and ocean into a grid of rectangular "boxes," each of which has a state that is summarized by only a handful of numbers.   This is an approximation because properties of a real fluid vary in space on a wide range of scales and thus have variations at scales smaller than the grid box. Yet, variations on a smaller scale than the "boxes" cannot be represented explicitly in the GCM. Just as with the resolution of a computer monitor or television, the resolution of a GCM (as defined by the size of the grid boxes) constrains the accuracy with which it can represent different aspects of reality, a point we will further discuss below.

To produce a climate simulation with a GCM, one needs not only the model itself, but also *initial conditions* and *forcings*.[2] Initial conditions describe the state of the atmosphere at the starting time of the simulation. Forcings are factors external to the climate system (and thus not predicted by the model) that influence it, e.g., incoming solar radiation and the concentrations of greenhouse gases.   The definition of a given quantity as either an externally specified forcing or a predicted variable is not immutable, but represents a modeling choice. Standards about these choices change over time - generally in the direction of modeling more variables explicitly and treating fewer as forcings - as science and technology evolve.   For example, the transport of heat by the ocean was a specified forcing in the so-called "slab mixed layer" models commonly used for early climate change projections, while it is predicted explicitly in "fully coupled" models that are more common today.

At the beginning of GCM model run all "grid boxes" of the ocean and atmosphere are set to initial conditions. It then uses an approximation of the differential equations governing the fluid between those boxes to model how they will evolve with time. For example, if two boxes next to one another have different temperatures at time t, there is an equation that will describe how both their temperatures, together with other state variables, determine their temperature at time $t+1$.[3] The state of the atmosphere and ocean in each subsequent period is determined by the governing equations, external forcings, and its previous state.

Weather and climate models have consistently been running on some of the world's most powerful super computers ever since the mathematician John Von Neumann started the first large scale weather modeling initiative in 1946 at Princeton's Institute for Advanced Study (Aspray, 1990).   Computational power has been the binding constraint on models given the large number of pixels needed to model the atmosphere and ocean and the large number of

---

[2] The term (radiative) forcing is defined by the IPCC as: "a measure of the influence a factor has in altering the balance of incoming and outgoing energy in the Earth-atmosphere system and is an index of the importance of the factor as a potential climate change mechanism. In this [AR4] report radiative forcing values are for changes relative to preindustrial conditions defined at 1750 and are expressed in watts per square meter (W/m2)."

[3] For example, the strength of the wind, which blows air from one grid box to the next.

computations needed for each time step. It would in principle be possible to construct models of almost arbitrarily high resolution, but computations would be so long that that time *t+1* would arrive in the real world before it does in the simulated world of the GCM, limiting the usefulness of a "forecast" that arrived after the actual realization it was predicting.   Models have always been designed to complete simulations in reasonable times, given the computing power available, a fact that limits their resolution and realism.

A typical GCM in the fourth and latest IPCC assessment report models the atmosphere using grid-boxes that are about 2 degrees latitude by 2 degrees longitude by 1 km high.   At the equator, the base of this box, which represents a single "pixel" in the global climate system, will cover forty-nine-thousand square kilometers (slightly larger than Massachusetts, Connecticut and Rhode Island combined).   Modeling all variation in rainfall, clouds, temperature and winds as homogenous over such a large area is obviously a simplification, but variations on smaller scales cannot be explicitly captured by the discrete representation used in such a GCM.   For many applications, higher resolution estimates of climate changes are needed.   When suh high-resolution estimates are needed for a limited area, "downscaling" is a computationally feasible approach.   One approach to downscaling involves creating a smaller sub-model of high resolution that is embedded in an otherwise coarser resolution of the GCM.   This allows smaller scale phenomena in the region of interest to be captured as they respond to global scale changes. Alternatively, statistical approaches use observed data at a finer scale and impose the observed within grid variability around the grid mean obtained from GCM.

Many of the climate phenomena that are of interest to economists are too small to be captured by the resolution of GCMs.   For example, GCMs are unable to realistically simulate hurricanes, which themselves may be large enough to dominate a few pixels, but which have an underlying structure that is too small to be simulated realistically.   Another example is the urban heat-island effect, which is governed by the thermal properties of built structures.   Both of these phenomena are subjects of ongoing work that try to understand their dynamics in a changing climate by using downscaling techniques (Bender et al., 2010, Rosenzweig et al., 2005). Unfortunately, there is no general and reliable way to extract these kinds of small-scale phenomena from coarse resolution global models.   Economists interested in working on these

or other small-scale phenomena should consult the relevant literature since the proper techniques may be phenomenon and location specific.

Reanalysis and missing data

Weather outcomes, i.e., realizations of climate (the underlying distribution of environmental states) are directly observed. The use of weather observations in climate models and economic research is the focus of this section. For applied economists looking to estimate climate response functions, the observational record of historical weather is sometimes frustratingly incomplete. The absence of data is most pronounced over poor regions with governments that do not prioritize weather data collection and regions with few inhabitants, such as deserts or over oceans. One approach that climate scientists have developed for filling in the holes of observationally sparse regions is "data assimilation". Data assimilation is used to produce data sets known as "reanalyses" in the context of climate studies (as opposed to weather prediction). These data sets contain particular advantages for applied economists studying the developing world (e.g. Guiteras, 2010, Schlenker and Lobell, 2010, Hsiang et al., 2011), but have not been widely adopted.

Data assimilation is the process by which observational data are combined with a GCM-like model to form an estimate of the state of the atmosphere and ocean that combines the strengths of both. The model propagates information from locations where observations exist to more data-sparse regions, providing an estimate of the state in data-sparse regions that is based on the model's physical laws as well as observations elsewhere. Whenever high quality observations exist, they are used to "correct" the model as further described below. The result is an estimate of the state of the system, which is discretized uniformly on the model grid, is more uniform in quality and realism than the observations are, and yet closer to the state which actually occurred than any pure model prediction would be.

Heuristically, the process of data assimilation is not unlike an economist's use of statistical regressions to interpolate missing observations. However, there are many important differences. First, data assimilation tries to minimize some loss function subject to an

enormous set of difference equations.[4]   On the upside, this means observations are reconstructed using some information from fundamental physical principles.   On the downside, it is unlikely that the large number of real observations can be fitted to a single, unified, dynamically-consistent state of the global environment at sequential moments in time (or even at a single moment, for that matter). This points to the second major difference from statistical interpolation: data assimilation tries to match a panel of observations to a complex dynamic model whose state at time $t+1$ is conditioned on its previous state at time t.   In economics, we are familiar with the notion that dynamic panel estimations require extra care since the dynamic aspect of the data generating process produces new issues.   Errors in estimating the state of the global environment at time t may propagate through the dynamic components of the model and affect the estimated state at time $t+1$.   Fortunately, observations at time $t+1$ can be used to correct for some of these errors if the model is adjusted to account for them.

Data assimilation is also used to produce initial conditions for weather forecasts.   Both the underlying models and the data assimilation methods evolved over time.   Sequences of such initial conditions produced over many years have artificial temporal inhomogeneities, which cannot be cleanly distinguished from true variability in the climate, causing problems for long-term climate studies.   The method known as "reanalysis" addresses this by assimilating all observations from a long period using a single GCM and data assimilation system (Kalnay et al. 1996).   Reanalysis tries to optimally estimate the state of the global environment over a large sequence of periods by optimally fitting a dynamic model to all those periods simultaneously. This process is difficult and costly, and therefore only a few research centers complete regularly updated data sets. The National Center for Environmental Prediction (NCEP) in the United States and the European Center for Medium-range Weather Forecasting (ECMWF) produce the two most commonly used reanalysis products.

Some final notes on the implications of using reanalysis data: First, reanalysis output is "pixilated" in the same sense as GCM output.   Both use similar finite-difference and finite-element techniques. Second, reanalysis output cannot be forced to perfectly match

---

[4] These difference equations approximate differential equations.

observational data.   The former has both limited resolution and is influenced by the GCM even when observations are present.   Third, there are two opposing ways to interpret differences between observations and reanalysis output: one is that reanalysis output is an approximation of "correct" observations, the other is that observations contain measurement errors that reanalysis "corrects"; the former is probably more widely accepted than the latter, but both contain some truth.   Forth, reanalysis is computed with models that are imperfect and contain systematic biases.   Constraining these models with data "fed" into them does not always satisfactorily correct for the model's built-in biased behavior.   In regions where observations are sparse or of poor quality, reanalysis provides estimates that may be better than what would otherwise be available.   Nonetheless, it is important to keep in mind that reanalysis output in such regions is basically a model prediction, which may be less accurate than in more observation-rich regions. Fifth, although the underlying climate model and data assimilation system are both held constant in the reanalysis process, the locations and types of observations being used still changes over time due to historical changes in the global observation system (e.g., the opening or closing of weather stations, launching or failure of satellites, etc.).   This can cause spurious temporal changes in reanalysis output that changes with the observational system it is fitted against.   One should hence be cautious in using such data to compute long-term trends.   The introduction of satellite observations around the late 1970s appears to be particularly problematic.

What are the differences between GCMs?
The most recent Fourth Assessment Report (AR4) of the IPCC evaluated the climate change projections of twenty GCMs that are built and run by research groups around the world.   The term "projection" is used to denote a simulation in which a particular set of assumptions is made about how the forcings – particularly greenhouse gas concentrations, both anthropogenic and natural – will evolve over time.   Projections made by different models for the same set of forcing assumptions often disagree even when two models are produced by the same group. Understanding the basis of these disagreements is sometimes important for economists using them.   The number of available models and the output they provide varies along a number of dimensions.   From a user's perspective, the models differ by the time step, geographic

resolution, and the indictors they provide.[5] Chapter 8 of Working Group 1 report for the IPCC's 4th assessment report provides discussion of the models and the variability in their predictions. An excellent description of the available output is given by Meehl et al (2007). A public access depository of GCM output, the WCRP CMIP3 multi-model database, is housed at Lawrence Livermore National Laboratories (https://esg.llnl.gov:8443/index.jsp). Figure 1 summarizes for models used in AR4 what data are available. Some models have not been run for all possible forcing scenarios, which may limit the choice of model. All models have output for at least one of the IPCC SRES scenarios and some even provide this output at a 3 hour time step out to 2100. Below we focus on the conceptual differences between climate models and what this implies for the economist user of the model output.

The IPCC has adopted an array of scenarios that project different types of future economic development and specify anthropogenic climate forcings accordingly (IPCC SRES, 2000). Each SRES scenario specifies the carbon-dioxide emissions of the human population for all years. The three primary driving forces behind these scenarios are population, world GDP, and the ratio of per capita income between Annex 1 countries (developed and transition economies) and non-Annex I countries (developing countries). Population in the year 2100, for example, varies from 7.0 to 15.1 billion for the "A" scenario family and 7.0 to 10.4 for the "B" scenario family. World GDP ranges from 243 trillion dollars to 550 trillion dollars for the "A" family and 235 to 328 trillion dollars for the "B" family. The range of income ratios is 1.5 to 4.2 for the A class scenario and 1.8 to 3.5 for the "B" scenarios. For comparison, the ratio was 16.1 in 1990, i.e., the models assume significant convergence. Secondary driving forces all relate to the energy mix, i.e., the final energy intensity, primary energy used, share of coal in primary energy, and share of zero carbon in primary energy.[6] Scenarios also specify the emissions of several other greenhouse gases other than carbon-dioxide, such as methane (SPM table 2a/b). These other gases have differing warming potentials and atmospheric lifetimes.

---

[5] 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 climate model properties.

[6] The IPCC SRES Summary for Policymakers Tables SPM 1a/b provide detailed descriptions of these scenarios: http://www.grida.no/publications/other/ipcc_sr/?src=/climate/ipcc/emission/008.htm

The SRES report, however, does not specify scenarios for other pollutants produced by humans and how they should be incorporated in the GCM.   For some pollutants, such as aerosols, which impact climate on a local and regional scale, emission trajectories are left to the discretion of the modeling group.   The circulation of these emissions in the atmosphere is also left up to the modeling group.   The degree to which these chemical reactions are incorporated varies greatly. Some models prescribe the behavior of chemicals exogenously while others include systems of equations for the evolution of chemical processes and solve for chemical dynamics "endogenously."   Different modeling assumption can result in significantly different predictions (e.g. Son et al 2008).   Modeling groups have to choose what dimensions, e.g., chemical processes, to model in more detail.   The limits of computational power are frequently binding constraints that don't allow for a complete modeling of all dimensions.

Some of the other qualitative ways in which models differ include their ocean, vegetation, and hydrological dynamics and how they model land use changes, extreme events, as well as ice and snow.   Because of the different choices made by different groups, some models are known to simulate certain types of phenomena better than others.   For example, the model produced by NASA's Geophysical Fluid Dynamics Laboratory (GFDL) is known for giving a good reconstruction of the 1960-80's drought in the Sahel of Africa (Held et al. 2005, Biasutti and Giannini, 2006).   Economists interested in specific phenomena should consult the literature to understand which models are considered best for the topics in which they are interested.

So far we have outlined what modelers choose to include or exclude from their models.   There are also important mechanical differences in how they implement their simulations of the system. Model resolution is one mechanical difference that was discussed earlier.   In the AR4, atmospheric resolutions ranged from 1.1x1.1-degree to 4x5-degree grid cells; ocean components are generally higher resolution.[7]   Higher resolution models are generally thought to perform

---

[7] In some models, space is not actually discretized using grid boxes.   Instead, a "spectral" decomposition using a set of smooth basis functions is used.   In such models, "resolution" refers to the order of terms retained in spectral decomposition and not to the size of discrete pixels, though a rough correspondence can be made; e.g. "T42" corresponds roughly to a 2.8 degree grid box size.

better, but high resolution alone cannot guarantee a model's performance. Most of the other mechanical details of models are of lesser concern to economists. Economists should consult model evaluations when selecting models (Meehl et al. 2007, Reichler and Kim, 2008).

Uncertainty in GCMs

The way in which climate modelers think about uncertainty differs substantially from the way economists generally think about it, regardless of whether one is a "frequentist" or a "bayesian". Climate modelers are most concerned about three types of uncertainty: *forcing uncertainty, model uncertainty* and *uncertainty over initial conditions*.

Forcing uncertainty arises because the SRES scenarios that describe future emissions and human activity are simply guesses based on fundamentally different assumptions about future human behavior. This is probably the largest uncertainty in future climate projections on time scales that exceed a few decades. In general, climate modelers have not tried to explicitly address this uncertainty beyond providing simulations for each scenario, perhaps because it is viewed as within the purview of the social sciences. Although there is some recent work by economists in this domain (Auffhammer and Steinhauser, forthcoming), additional work on this problem would be valuable.

Model uncertainty is somewhat analogous to specification uncertainty in applied economics: regardless of the quantity or quality of data one has, if a modeler chooses the wrong equation into which she feeds excellent data, she is likely to find an incorrect result. The array of modeling choices described above demonstrated that it is difficult to know when we have the "right" climate model. Not only do models sometimes contain different dynamics, but even models that contain similar dynamics may describe them in a different mathematical setup. Selecting the equations and parameters used to model a certain relationship involves discretionary choices. For example, in the current generation of models, some of the largest model uncertainties involve the sub-gridscale parameterizations of clouds, aerosols, land surfaces, ice and snow. One way that the modeling community has dealt with these kinds of uncertainty is to use a "multi-model mean" as the estimate for projections. The mean output from several different modeling groups generally matches observations better than any of the

individual models do (see Figure 1 in Reichler and Kim, 2008). Why this should be the case is not fully understood, but it suggests that model errors have a significant random component. Whether it extends to future climate projections is not testable. Nonetheless, the practice seems to work for the historical record and is one of the underlying motivations for having a large number of groups model future climate. For economists with the necessary resources, averaging output of climate models before employing them may be a useful technique (e.g. Hsiang, 2011). Alternatively, one can separate the distribution of predicted impacts into (i) a distribution based on the uncertainty of the economic parameters that fixes predicted climate change at the average forecast; and (ii) a distribution of impacts due to the uncertainty on how the climate will evolve that samples from the distribution of predicted changes and fixes the economic parameters at the mean estimate (e.g. Schlenker and Lobell, 2010).

The third type of uncertainty that concerns climate scientists is uncertainty in initial conditions. This is especially relevant as the atmosphere and oceans exhibit *deterministic chaos*. The realization that perfectly deterministic systems may exhibit unpredictable behavior (Lorenz, 1963) is considered one of the greatest scientific discoveries of the twentieth century, rivaling the Theory of Relativity and quantum mechanics. For a wide range of nonlinear dynamical systems, of which the atmosphere and ocean are members, any uncertainty in the initial state - however small - grows sufficiently rapidly in time that it is impossible to predict the instantaneous state of the system beyond some finite time horizon. Thus, even if there is no model uncertainty, our ability to predict the future will still be bounded by a theoretical limit, believed to be about two weeks for the atmosphere. Weather prediction - the prediction of the *instantaneous* atmospheric state – beyond this horizon is believed to be impossible, regardless of any technical improvements. On longer time scales, it can still be possible to predict slower variations in the climate – that is, changes in *statistics* of the atmospheric state. These variations can be predictable to the extent that they are forced by components of the system that evolve more slowly. External forcings are an example of a slow component, and anthropogenic climate change can be described as a shift in the distribution of states that is in principle predictable on time scales much longer than those of weather prediction. Another slow component is the ocean. The ocean's predictability horizon is longer than that of the atmosphere. This allows for *seasonal* climate variability on a time scale of 3-12 months. It is

currently speculated that memory of initial conditions of the may perhaps result in some useful on the scale of decades.   Nonetheless, on long enough time scales, the ocean is also chaotic and unpredictable.   For predictions centuries in the future, chaos in the ocean is a non-negligible source of uncertainty.

Climate modelers have generated "ensemble" simulations to manage chaos.   This technique is similar to Monte-Carlo methods used by economists when model parameters are uncertain. Small differences in initial conditions used in climate simulations rapidly amplify to produce weather states as different as what one would have obtained if an arbitrary realization of the climate had been chosen.   A separate model run is conducted for a number of initial conditions that are drawn from a distribution, which is an attempt to characterize the distribution of states in the future.   Each initial condition produces a projection of the global environment over time, called an "ensemble member".   The set of all ensemble members run for a given model defines the ensemble.   In the AR4, a typical model reported results from ensembles with 3-5 members. If the ensemble is large enough, natural variability due to chaos, being effectively random, will average out.   The response of climate variables of interest to imposed forcing can then be approximated by changes in the ensemble mean.

Finally, model uncertainty and uncertainty over initial conditions can be combined.   Recall that chaos-induced uncertainty is managed by averaging ensembles generated from different initial conditions with the same model, while model uncertainty is managed by averaging output from multiple models, each of which may contain multiple ensemble members.   The result is referred to a "multi-model ensemble," where outcomes from ensembles of different models are averaged. While these averaging techniques are considered to estimate the *climate* most accurately, it is important to recall that climate is a simplified description of the underlying distribution of weather.   For example, if an economist were to estimate crop losses due to temperature changes over time and the loss function is nonlinear and concave (e.g. Schlenker and Roberts, 2009), the losses calculated using temperatures from a multi-model ensemble mean would likely be too small because it omits the variance in temperatures around the mean.

3. What to keep in mind when using Climate Data in Economic Applications

Aggregation Issues

General Circulation Models divide the earth's surface into a discrete grid. As far as the GCM is concerned, there is variation in climate across these discrete grid cells, yet within each such cell climate statistics are homogenous. For example, if one uses a climate model providing output at a monthly time scale, temperatures within the month and among all locations in the grid cell are considered constant. Both temporal and spatial aggregation might be inappropriate. A 2x2-degree cell is "small" from the perspective of the global climate, but not from the perspective of human systems. A 2x2-degree grid spacing at the equator is equivalent to a grid width of 222 kilometers (138 miles). It is not hard to imagine a stretch of this length with vastly varying climates (e.g. driving east from San Diego's coastal climate to El Centro's dry and hot desert climate). This becomes especially relevant if the underlying topography is not flat but mountainous or is located near the ocean. Figure 2 shows the severity of this bias for the lower 48 states. Here we contrast average temperatures predicted by the Hadley III GCM to a fine-scaled (2.5x2.5 mile grid) weather data set for the contiguous United States (PRISM, 2009). The figure plots the difference in the average daily maximum temperature in the month of July in the years 1960-1999 between the GCM, which has the coarser resolution, and the fine scale weather grid. A positive number indicates that the GCM grid average exceeds the PRISM average, which is based on interpolated station data. The map shows quite clearly that this bias is most significant in mountainous areas, which usually are also less populated areas. At the extremes, we see that this bias can reach +25°C at some mountain tops. This is not surprising, since surface temperatures tends to fall about 7°C per 1000m in elevation, so mountains are much colder than areas at lower elevations in the same grid cell. This bias does not only exist in remote mountainous regions but also in heavily populated areas, which are often located near oceans. The map shows that for the month of July, the entire western seaboard has biases, which are significantly greater than any predicted warming. The average absolute difference across the entire United States is 3.0°C, while the root mean squared error is 4.0°C, both of

which are significantly larger than the average predicted changes under the SRES climate change scenarios by the end of the century.   While the magnitude of the bias varies by location, it also varies by the climate indicator one is using.   For example, if we use the annual mean temperature instead of the average daily maximum July temperature, the absolute error reduces to 1.8°C, and the root mean squared error to 2.4°C.

This is especially relevant for studies of economic impacts of climate change.   In these studies one generally parameterizes a response function between *e.g.* electricity demand and temperature, using observations from a weather station based dataset and observed electricity demand.   In order to obtain counterfactual electricity demand under a scenario with climate change, one requires a baseline climate and a counterfactual climate.   What Figure 2 indicates is that if one were to use an average of observed micro-level weather as the baseline climate and level predictions of climate taken from a GCM as counterfactual climate at a future date, the resulting estimated impacts would be due to both simulated warming and the bias displayed in Figure 2.   If the response function is nonlinear in weather/climate, as has been shown in agriculture (e.g. Schlenker and Roberts, 2009) and electricity demand (e.g. Aroonruengsawat and Auffhammer, 2011) this bias may be amplified or offset depending on the nature of the non-linearity.   In either case, the resulting impact estimates are contaminated.

The literature has suggested several mechanisms to correct such biases. In addition to using downscaled climate models as suggested in section 2, the most commonly used approach is based on regression methods.   One establishes a correlation between the historical grid values from the GCM and local station based data and uses the fitted regression relationship with future values of GCM output to arrive at downscaled GCM predictions.   Innovations of this approach have involved non-linear estimation, neural networks and Bayesian methods. Wilby and Wigley (1997) provide a review of the main approaches used in practice and compare their performance at selected locations. As Fowler, Blenkinsop and Tebaldi (2007) note, there is a large literature examining the performance of different downscaling approaches for different regions and climate variables.   There is no single best approach for all variables (e.g. maximum temperature, rainfall and wind speed) and locations.   Further, 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 matter for many economic applications such as agriculture and electricity demand, one requires a downscaled version of a climate model delivering daily output, which are available for some regions such as California (CEC, 2006), or at coarser time steps nationally (e.g. Maurer et al, 2007) and globally (e.g. Maurer 2009).

In the absence of an appropriate downscaled dataset for the region and time step of interest, the most common practice is to derive predicted *changes* for each (coarse) GCM grid and add these to an average of the historic baseline data used in the parameterization of the response function, thereby preserving within-GCM grid variation. By doing this, one subtracts out the location specific bias only if this bias is stationary in time. This approach, however, shifts the historic time-series at a location by the predicted change, leaving its variance unchanged. In case researchers are worried about predicted changes in the mean *and* the variance, 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 compared to the baseline. It should, however, be noted that there is much less consensus on the predicted changes in the variance among models.

In summary, one should not simply use GCM output as a direct forecast of future climate in impact estimation relative to a weather station-based baseline climate, even if the forcing scenario on which the simulation is based were accurate.

Panel versus Cross-section

Studies on the economic impacts of climate change on agriculture have used two distinct sources of identification in order to estimate response functions: Early studies relied on cross sectional variation in weather or climate in various locations to the outcome variable of interest (e.g., Mendelsohn, Nordhaus and Shaw, 2004, Kelly, Kolstad and Mitchell, 2005). A potential concern with a cross-sectional approach is that there may be unobservable variables which vary across these spatial units and which are correlated with the climate/weather indicator used. Recent studies therefore focus on a panel data analysis, which controls for space and time fixed effects (e.g., Greenstone and Deschenes 2007, Schlenker and Roberts 2009). Fixed effects estimators rely on variation across time within a spatial unit (e.g. county) instead of variation across these spatial units. The underlying identification therefore relates time series deviations

from the location specific mean in the climate indicators to deviations in the outcome variable of interest. While the economic implications of either approach (e.g. long-run versus short-run adaptation) have been discussed elsewhere, there is a practical issue relating to the choice of weather dataset used, which has received no attention. It is established that the 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. However, the datasets do not agree as well in the timing or magnitude of deviations from this mean. This problem is larger in areas with a small number of weather stations. We illustrate this point using two of the most commonly used global gridded weather data sets:

1) The CRU 2.1 data set from the University of East Anglia uses a statistical interpolation procedure and gives monthly minimum and maximum temperature on a 0.5x0.5 degree grid (Mitchell and Jones, 2005).

2) The reanalysis data from the National Center for Environmental Prediction (NCEP/NCAR) gives daily average temperature and total precipitation on a non-uniform grid (1.875 degrees longitude, and roughly 1.90 degrees latitude, although the latter is not evenly spaced).

CRU and NCEP use very different techniques in order to arrive at a temperature grid with global coverage. While the CRU data base is created using a statistical interpolation procedure, the reanalysis data from NCEP also incorporates other dynamic constraints to make the data consistent with underlying physical principles of the climate system as discussed in Section 2. In the following we focus on two variables: annual average temperature as well as total precipitation by country. We calculate country averages by taking a weighted average across grid cells. We calculate the weight given to each grid as what share of the land area of a country it covers.

We derive the average outcome by country over the period 1960-1999 as recent studies have used country-level aggregates and averages (Dell et al. 2008, Schlenker and Lobell, 2010, Hsiang 2010). The correlation between average temperatures across countries is 0.99 between the CRU and NCEP data bases, and 0.88 for total precipitation. The two data sources hence give similar estimates to which areas of the world are hot and which are cold, which is reassuring

for studies relying on cross sectional variation across countries. It is, however, more difficult to predict how weather variables change year-to-year. In a second step we construct annual deviations from the country-specific mean over the 40-year period in each data set, therefore resulting in a panel data sample 40 times larger. The correlation coefficients reduce to 0.74 for average temperature and 0.30 for precipitation. The correspondence in deviations across datasets for precipitation is very low, which is likely due to the fact that precipitation is less smooth in space and hence the exact smoothing routine employed becomes more important.

These average correlations mask considerable heterogeneity by country. If we construct weather shocks by weighting each grid cell by the amount of maize that is grown within it (Monfreda et al., 2005), which is commonly done if one is interested in agricultural impacts of climate change, the correlation coefficient among weather deviations in the United States increases to 0.89 for average temperature and 0.70 for precipitation. Presumably because of the good observational network in the United States, the two data sources again roughly agree. On the other hand, weather shocks constructed over the maize growing area in Nigeria have correlation coefficients of 0.69 and 0.26, respectively. In regions with very limited monitoring networks, which is generally the developing world, the weather shock used to identify response coefficients in econometric estimation hence varies significantly by what data source is used.

In summary, panel studies that rely on deviations from averages should be careful in which data source they use, as measurement error - and related statistical concerns - are amplified by demeaning in the climate change context.

Correlation of weather variables

Many economic studies, including, but not limited to, the estimation of climate change impacts, have focused on the impact of one weather variable in isolation, e.g., regressing income on precipitation shocks only (Miguel et al., 2004). While precipitation shocks are exogenous and hence a plausible instrument for income, it should be noted that the coefficient on precipitation will measure the combined effect of precipitation and temperature if the two variables are correlated. In the climate change context, this is important if the estimated coefficient is used to extrapolate what would happen under a different climate. In order to parse out unbiased effects

for changes in precipitation and temperatures, which historically covary, both have to be included in the regression equation. This is especially important if the correlation is predicted to change in the future.

Figure 3 plots the Pearson correlation coefficient between annual average temperature and total precipitation for each of the CRU grid cells over the years 1960-1999. Notices that the correlations vary greatly and there are areas with both significant positive as well as negative correlation between precipitation and temperature. This implies that if one only controls for one of the two weather variables in a regression, the sign of the omitted variables bias depends on the location under study. Hot areas generally show negative correlation, as more precipitation and the associated evaporation results in cooling and lower average temperatures. The negative correlation can be as large as -0.7. On the other hand, a positive correlation is observed in cooler areas where increased precipitation is associated with 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 (US, Russia, France, Spain).

It is crucial to note that climatic variables other than temperature and precipitation, e.g., relative humidity, solar radiation, wind speed and direction, may contaminate empirical estimates through a classical omitted variable problem. The presence of these other phenomena and their correlation with temperature or precipitation may be location specific. For example, hurricane activity is correlated with temperatures in the Caribbean (Hsiang, 2010).

In summary, if temperature, precipitation and other atmospheric variables are correlated, a study that seeks to extrapolate based on an estimated response function what would happen under a different climate has to include all of these variables to get an unbiased estimate of the effect of each variable.

Spatial correlation

Climate variables are inherently spatial in nature. While variation in weather is often considered random across time, variation across space displays significantly less "randomness".

Because weather and climate variations have inherent spatial scales which can be quite large, some of the weather or climate variables which we use in econometric estimation are highly spatially correlated. Figure 4 calculates the average correlation of annual mean temperature at each CRU grid cell in 1960-1999 with the eight surrounding grid cells. Note that we have chosen a highly nonlinear scale (correlation to the power of 100) as all correlations are extremely close to one. As described above, errors might prorogate from one box to the next in both interpolated station data as well as data assimilation methods. It is imperative to take into account this spatial correlation in econometric estimation, which generally results 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. If a test for spatial dependence in the residuals rejects the null of independence, there are three common ways to account for spatial correlation: (i) use a spatial weighting matrix, which is most efficient in the case there the weighting matrix is known, but will result in biased estimates if the weighting matrix is misspecified; (ii) use the nonparametric approach provided by Conley (1991), which does not require one to specify a weighting matrix; or (iii) use a grouped bootstrap where years are resampled with replacement. The latter requires that year-to-year fluctuations are random, while errors within a year can be correlated. In many areas of the world, the independence of year-to-year variation is questionable given planetary scale oscillations, such as the El Niño-Southern Oscillation, which may be autoregressive (Hsiang et al., 2011).

Endogenous weather coverage

Another set of economic questions examines how the relationship between weather variables and economic variables of interest might change with large policy changes, for example the independence of a country or a extreme exogenous shock like a natural disaster (Kahn, 2005). The natural approach is a difference-in-difference analysis. One concern here is that if the degree of measurement error varies between the pre and post intervention date, the treatment effect estimate is very likely biased due to traditional attenuation concerns. If weather variables are consistently measured, the difference in difference regression design will be free of this bias. However, the researcher has to ensure that weather station coverage does not change with the policy change, which might introduce measurement error and result in downward bias in the estimated coefficients in the post-intervention period. This concern is strongly related to the

discussion on spatially varying correlation due to different interpolation methods over a sparse data matrix.

To illustrate this point we downloaded daily data from the Global Summary of the Day database maintained by the National Climatic Data Center (NCDC) of the National Oceanographic and Atmospheric Administration (NOAA). We count the number of days a weather station within a country has non-missing observations and sum it across all stations, giving us the total count of day-station observations by country per year. While most countries show an upward trend over time, the results for some transition countries are striking. For example, Romania, like most other countries, 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 rapidly decreased until it stabilized around 11,000 station-days in 2003-2007, decreasing coverage by a factor of six. Results from a difference-in-difference analysis of how, for example, farmers respond to weather shocks pre- and post the fall of the iron curtain would have to be interpreted with caution. If one uses any of the gridded data products available, it is crucial to determine whether the underlying station data has changed if one attaches economic significance to the estimated weather variables.

4. Conclusions

In closing, we would like to point out that when one uses gridded datasets of historical or future climate, it is crucial that one consider the fact that both types of datasets are very different from observed weather. Historical gridded data products are very convenient since they provide often highly disaggregated weather for large geographic regions over long time periods. This increased coverage comes at a cost: the birth and death of weather stations and spatial correlation introduced by the extrapolation algorithms create potential biases in the estimated coefficients and standard errors if one uses these weather products as covariates. When one uses Global Climate Model output as a counterfactual of future climate, the choice of model will have significant ramifications as to the sign and magnitude of the estimated impacts. Further, the failure to account for location specific biases of each model may further bias impact estimates.

References

Aroonruengsawat, Anin and Maximilian Auffhammer. 2011. Impacts of Climate Change on Residential Electricity Consumption: Evidence From Billing Data. in *The Economics Of Climate Change: Adaptations Past and Present*, Gary Libecap and Richard H. Steckel, editors. University of Chicago Press.

Aspray, William. 1990. *John Von Neumann and the Origins of Modern Computing*. Cambridge, MA: MIT Press.

Auffhammer, Maximilian and Ralf Steinhauser. forthcoming. Forecasting the Path of US CO2 Emissions Using State-Level Information. *Review of Economics and Statistics*.

Biasutti, Michela and Alessandra Giannini. 2006. "Robust Sahel drying in response to late 20th century forcings," *Geophysical Research Letters*, 33:1-4.

Bender, Morris A., Thomas R. Knutson, Robert E. Tuleya, Joseph J. Sirutis, Gabriel A. Vecchi, Stephen T. Garner, Isaac M. Held. 2010. "Modeled Impact of Anthropogenic Warming on the Frequency of Intense Atlantic Hurricanes." *Science.* 327(5964): 454 – 458.

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. 2008. "Climate Change and Economic Growth: Evidence from the Last Half Century," *NBER Working Paper 14132*.

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): 354-385.

Fowler, H.J., Blenkinsop, S. and Tebaldi, C. 2007. Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, 27(12), 1547-1578.

Franco, Guido and Alan H. Sanstad. 2008. Climate Change and Electricity Demand in California. *Climatic Change*. 87: 139–151.

Guiteras, Raymond. 2010. "The Impact of Climate Change on Indian Agriculture." *Working paper*.

Held, I. M., T. L. Delworth, J. Lu, K. L. Findell, and T. R. Knutson. 2005. "Simulation of Sahel drought in the 20th and 21st centuries," *Proceedings of the National Academy of Sciences,* 102(50): 17891-17896.

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):15367-15372.

Hsiang, Solomon M., Kyle C. Meng and Mark A. Cane. 2011. "Global Climate Influences Civil Conflict: Evidence from El Niño." *Working paper*.

Hsiang, Solomon. 2011. "Global Economic Exposure to Future Temperature Changes," *Working Paper.*

Intergovernmental Panel on Climate Change. 2000. *Special Report Emissions Scenarios (SRES).*

Kahn, Matthew. 2005. "The death toll from natural disasters: the role of income, geography, and institutions." *Review of Economics and Statistics*. 87(2): 271-284.

Kelly, David L., Charles D. Kolstad, and Glenn T. Mitchell. 2005. "Adjustment costs from environmental change." *Journal of Environmental Economics and Management*, 50(3): 468-495.

Lorenz, E. N., 1963. "Deterministic nonperiodic flow". *Journal of the Atmospheric Sciences*, **20**, 130-141.

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

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

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: 1383-1394.

Mendelsohn, Robert, William D. Nordhaus and Daigee Shaw. 1994. "The Impact of Global Warming on Agriculture: A Ricardian Analysis," *American Economic Review*, 84(4): 753-771.

Miguel, Edward, Shanker Satyanath and Ernest Sergenti. 2004. "Economic Shocks and Civil Conflict: An Instrumental Variables Approach." *Journal of Political Economy*, 112(4): 725-753.

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. 2008. "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.

PRISM. 2009. "Parameter-elevation Regressions on Independent Slopes Model." Oregon State University (http://www.prism.oregonstate.edu/). Accessed: 2009.

Reichler, Thomas and Junsu Kim. 2008. "How Well Do Coupled Models Simulate Today's Climate?" *Bulletin of the American Meteorological Society*, 89: 303-311.

Rosenzweig, Cynthia, William D. Solecki, Lilly Parshall, Mark Chopping, Gregory Pope, and Richard Goldberg. 2005. "Characterizing the urban heat island in current and future climates in New Jersey." *Global Environmental Change Part B: Environmental Hazards*, 6(1): 51-62.

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): 15594-15598.

Son, S. W., L. M. Polvani, D. W. Waugh, H. Akiyoshi, R. Garcia, D. Kinnison, S. Pawson, E. Rozanov, T. G. Shepherd, and K. Shibata. 2008. "The Impact of Stratospheric Ozone Recovery on the Southern Hemisphere Westerly Jet." *Science*, 320(5882): 1486-1489

Wilby, R. L. and T. M. L. Wigley. 1997. "Downscaling general circulation model output: a review of methods and limitations." *Progress in Physical Geography*, 21(4): 530–548.

Figure 1: Data Availability by Forcing and Climate Model in the WCRP CMIP3 Multi-model Database
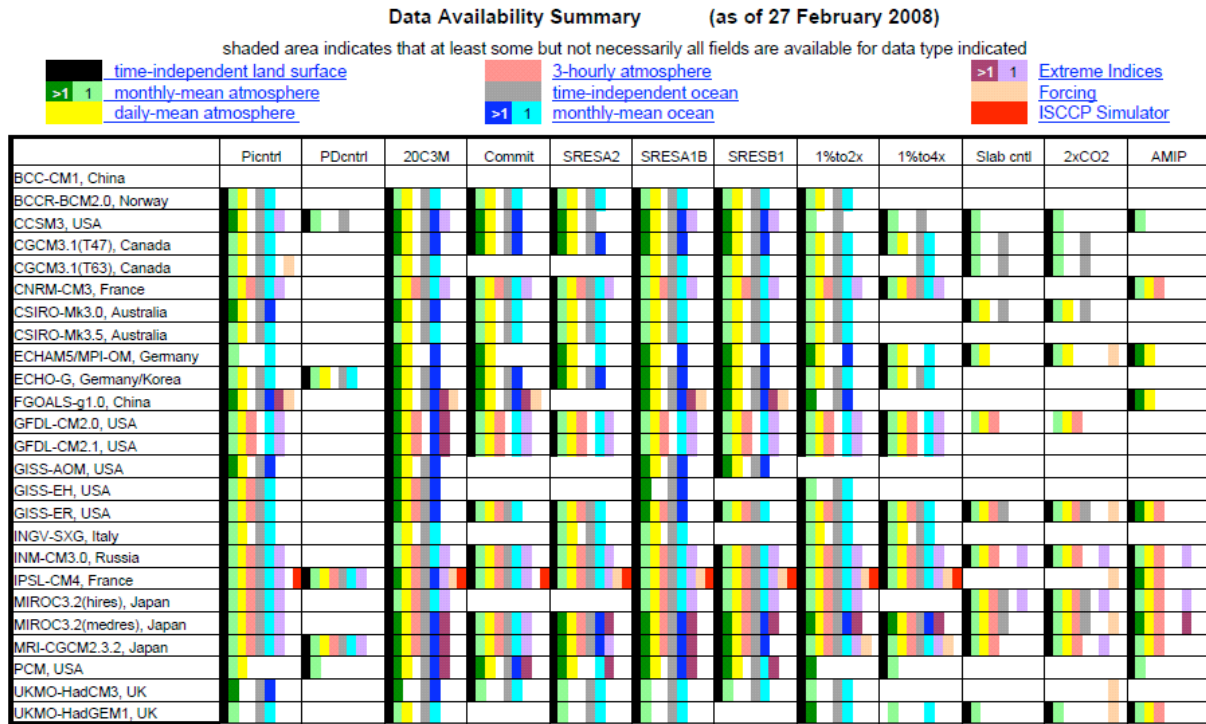
Figure 2: Aggregation Bias: Hadley Grid Averages (1961-1999) versus PRISM Grid Averages
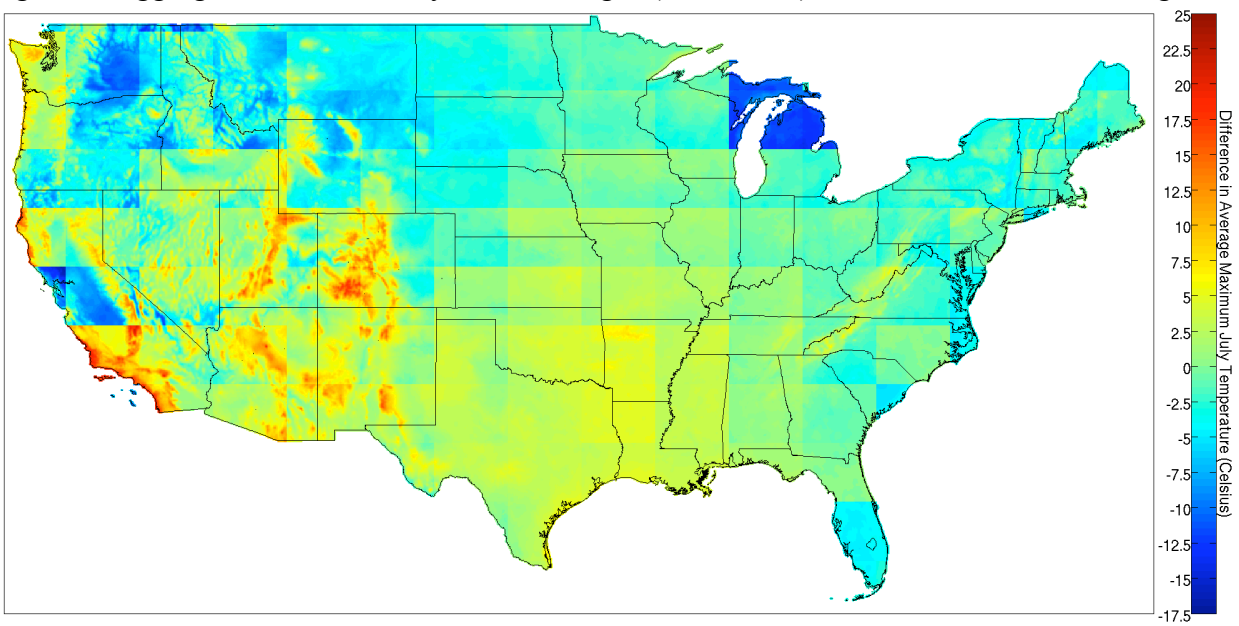
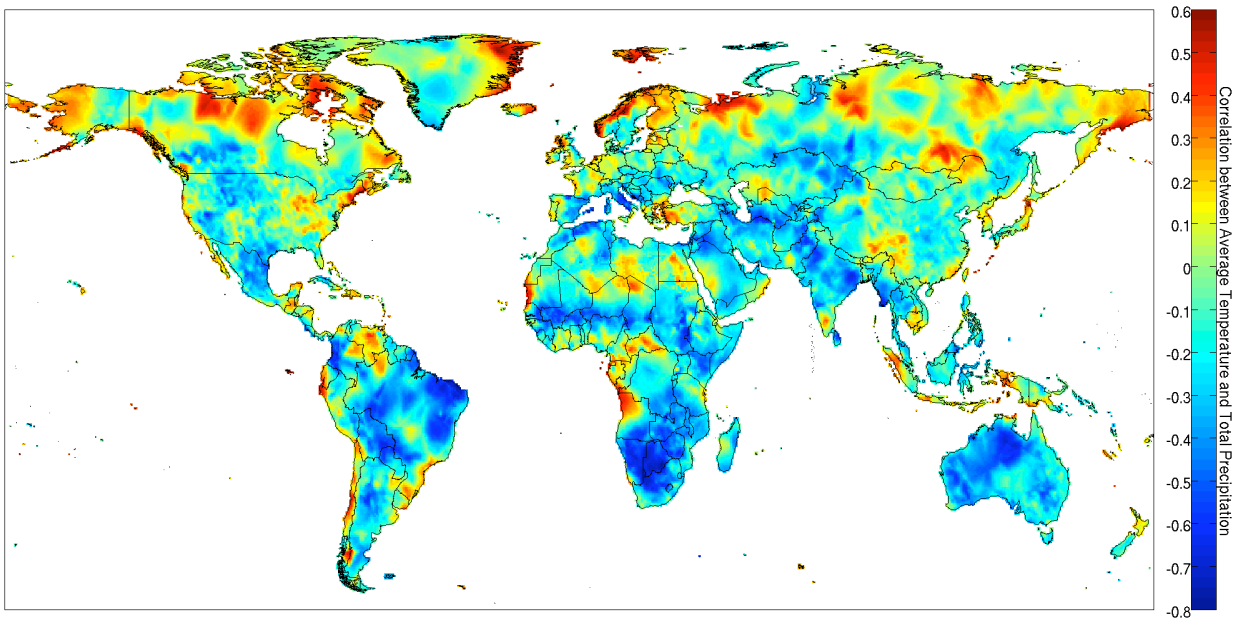Figure 3: Correlation between Annual Average Temperature and Total Precipitation (CRU 2.1 data 1960-1999)

Figure 4: Correlation of Average Annual Temperature at CRU grid with surrounding Eight Grid Cells (1960-1999).