Statistics, Politics, and Policy

Manuscript 1004

Atmospheric Circulations do not Explain the Temperature-Industrialization Correlation

Ross McKitrick, *University of Guelph*

Atmospheric Circulations do not Explain the Temperature-Industrialization Correlation

Ross McKitrick

Abstract

Gridded land surface temperature data products are used in climatology on the assumption that contaminating effects from urbanization, land-use change and related socioeconomic processes have been identified and filtered out, leaving behind a "pure" record of climatic change. But several studies have shown a correlation between the spatial pattern of warming trends in climatic data products and the spatial pattern of industrialization, indicating that local non-climatic effects may still be present. This, in turn, could bias measurements of the amount of global warming and its attribution to greenhouse gases. The 2007 report of the Intergovernmental Panel on Climate Change (IPCC) set aside those concerns with the claim that the temperature-industrialization correlation becomes statistically insignificant if certain atmospheric circulation patterns, also called oscillations, are taken into account. But this claim has never been tested and the IPCC provided no evidence for its assertion. I estimate two spatial models that simultaneously control for the major atmospheric oscillations and the distribution of socioeconomic activity. The correlations between warming patterns and patterns of socioeconomic development remain large and significant in the presence of controls for atmospheric oscillations, contradicting the IPCC claim. Tests for outlier influence, spatial autocorrelation, endogeneity bias, residual nonlinearity and other problems are discussed.

KEYWORDS: global warming, data quality, industrialization, spatial autocorrelation

Author Notes: I thank Werner Antweiler and Glen Waddell for assistance in the econometric programming. Financial support from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

1 INTRODUCTION

1.1 Overview

This paper examines a claim in the most recent report of the Intergovernmental Panel on Climate Change (IPCC) concerning temperature data quality. At issue is whether trends in climate data sets over land are solely due to global climatic change, or are to some extent measuring local non-climatic effects such as urbanization, land surface modification, instrument changes, etc. The latter factors are supposed to have been filtered out of climate data products. Inadequate filtering would imply a form of data contamination. In its most recent report, the IPCC claimed that residual non-climatic effects are negligible in the data upon which they base their main conclusions. But published evidence has contradicted this on the grounds that, in the climate data relied upon by the IPCC, the spatial pattern of warming over land correlates strongly with the spatial pattern of industrialization and economic growth, a pattern not predicted as a feature of general climatic warming. The IPCC dismissed this as a spurious effect attributable to large-scale atmospheric circulation systems, which, they claim, renders the temperature-industrialization correlations statistically insignificant. But they provided no evidence for this position. The purpose of this paper is to evaluate their assertion in a statistical framework.

1.2 Background

Numerous studies have shown that land-use change, such as urbanization, removal of forest cover and introduction of irrigated agriculture, introduce warming biases into local surface temperature data records (e.g. Jones et al. 2009, Christy et al. 2006, Pielke Sr. et al. 2002, Mahmood et al. 2010, etc., see review in McKitrick and Michaels 2007). Since these local changes are not related to global atmospheric climate changes, they need to be filtered out of climate data sets. But de Laat and Maurellis (2004, 2006, collectively denoted DM) and McKitrick and Michaels and (2004a, 2007, collectively denoted MM) showed that the spatial pattern of temperature trends in gridded surface climate data products is strongly correlated with indicators of industrial and socioeconomic development, which are broadly called anthropogenic surface processes. MM additionally test for and establish significant effects from measures of data inhomogeneity, or variations in measurement quality, on data sets that ostensibly have been adjusted to remove such effects. The aggregate effects of these influences are not small: DM and MM07 each estimate non climate-related effects in post-1980 surface temperature data amounting to between one-third and one-half of the observed warming trend over the global land surface.

The presence of non-climate-related trends in climatic data sets is a form of contamination that may be overstating atmospheric temperature trends (Klotzbach et al. 2009) and leading to misattribution of temperature changes to greenhouse gas effects. Pielke Sr. et al. (2002) found land surface changes produce regional climatic modifications that are not accounted for in the standard radiative forcing metric, further supporting the possibility of misattribution of spatiotemporal variability in gridded surface data.

Benestad (2004) and Schmidt (2009) both argued that the evidence of data contamination can be dismissed as artifacts of spatial autocorrelation, but neither provided a formal test. The spatial autocorrelation issue is examined in McKitrick and Nierenberg (2009) and will be discussed below. Benestad (2004) and McKitrick and Michaels (2004c) debated the extent to which global-scale patterns should be replicated in subsamples. McKitrick and Michaels (2007) performed random subsampling experiments and found consistently strong replication (see Section 3.5 below). Schmidt (2009) showed that the results for individual coefficients are weaker when using the satellite reanalysis product of Mears et al. (2001) rather than the Spencer and Christy (1990) data. McKitrick and Nierenberg (2009) showed that the joint significance tests on which the main conclusions are based remain significant regardless of which satellite data product is used.

Empirical papers in climatology rely strongly on the assumption that climate data products are free of effects from surface processes and measurement inhomogeneity. To take one example, in a comparison of warming trends in climate models and climatic data, Jun et al. (2008, p. 935), state:

> Inhomogeneities in the data arise mainly due to changes in instruments, exposure, station location (elevation, position), ship height, observation time, urbanization effects, and the method used to calculate averages. However, these effects are all well understood and taken into account in the construction of the data set.

Later, after observing discrepancies between model-generated and observed trends (which they denote D_i), when explaining why they do not attribute them to data contamination but rather assume they are all attributable to climate model biases, they state: "…climate scientists have fairly strong confidence in the quality of their observational data compared with the climate model biases. Therefore, we assume that the effect of observational errors to D_i is negligible." (Jun et al. 937)

The same assumption is made in the most recent report of the Intergovernmental Panel on Climate Change (IPCC 2007), and indeed is fundamental to their interpretation of the surface temperature data. Confining the data contamination question only to urban heat island (UHI) effects, though the underlying issue is in fact broader, the IPCC states (p. 244)

> In summary, although some individual sites may be affected, including some small rural locations, the UHI effect is not pervasive, as all global-scale studies indicate it is a very small component of large-scale averages.

The quoted statement is misleading since studies of UHI effects are inherently local, whereas the global-scale studies of DM and MM looked at more general issue of surface processes and data inhomogeneities, and did find large effects.

This paper focuses on the treatment of the contamination problem by the IPCC in its 2007 Fourth Assessment Report, where the issue was raised but dismissed as follows.

> McKitrick and Michaels (2004) and De Laat and Maurellis (2006) attempted to demonstrate that geographical patterns of warming trends over land are strongly correlated with geographical patterns of industrial and socioeconomic development, implying that urbanisation and related land surface changes have caused much of the observed warming. However, the locations of greatest socioeconomic development are also those that have been most warmed by atmospheric circulation changes (Sections 3.2.2.7 and 3.6.4), which exhibit large-scale coherence. **Hence, the correlation of warming with industrial and socioeconomic development ceases to be statistically significant.** In addition, observed warming has been, and transient greenhouse-induced warming is expected to be, greater over land than over the oceans (Chapter 10), owing to the smaller thermal capacity of the land.

(IPCC 2007 Chapter 3 page 244, emphasis added). The emphasized sentence makes a specific statistical claim: temperature-industrialization correlations cease to be statistically significant once account is taken of atmospheric circulation effects. Numerically, a result ceases to be significant if its *P* value rises above 0.05, and loses marginal significant when *P* goes above 0.1. The IPCC did not cite any published *P* values, or any published evidence of any kind, in support of their claim, and indeed none exists; nor were any new statistical calculations presented in the IPCC report itself. It is also noteworthy that the claim was not subject to the IPCC's peer review process since the paragraph in question did not appear in either of the two drafts that were circulated for expert review. It appeared for the first time in the final published version.[1]

Since the quality of the land surface temperature data is integral to so many reports and studies on climate change it is important to assess the IPCC's claim of statistical insignificance by means of a proper testing. In this paper I examine whether the results in MM04 and MM07 become insignificant once the effects of atmospheric oscillations on the surface temperature field are entered into the models in a reasonable way. I show that after augmenting the models in McKitrick and Michaels (2004, 2007) with four major atmospheric circulation indexes, the correlations in question remain highly significant and continue to indicate that urbanisation, related land surface changes and data inhomogeneities can account for much of the post-1980 warming over land. In order to establish the robustness of these findings I test the augmented McKitrick and Michaels (2007, hereinafter "MM07") model for spatial autocorrelation, endogeneity bias, error misspecification, outlier effects, and overfitting. No evidence for any of these problems emerges, supporting an overall conclusion that the IPCC conjecture was not only presented without support, but is also untrue, and that evidence pointing to significant contamination of climate data over land should therefore not have been dismissed.

2 EMPIRICAL TESTING OF CIRCULATION PATTERN EFFECTS

2.1 MM 2004 Model

Most of the temperature data used in this paper is grouped into 5 degree-by-5 degree grid cells on the Earth's surface. The unit of observation is a linear trend (degrees C per decade) through monthly "anomalies," or deviations from local averages. Hence the regressions are cross-sectional, and seek to explain the spatial pattern of warming and cooling trends over land.

By including the spatial pattern of socioeconomic variables such as population growth and GDP growth, as well as the spatial pattern of geographical and climatological factors, McKitrick and Michaels (2004a) sought to test the hypothesis that the spatial pattern of warming is independent of the socioeconomic influences that climatologists claim have been identified and removed from the data set. Their first regression equation was:

[1] IPCC drafts and review comments are available at http://www.ipcc-wg3.de/publications/assessment-reports/ar4/forth-assessment-review-comments

$$
\begin{aligned}
STREND_i = \alpha &+ \gamma_1 PRESS_i + \gamma_2 WATER_i + \gamma_3 COSABLAT_i + \beta_1 POP_i \\
&+ \beta_2 SCALE79_i + \beta_3 COAL80_i + \beta_4 COALGROW_i + \beta_5 INC79_i \\
&+ \beta_6 GDPGROW_i + \theta_1 SOVIET_i + \theta_2 SURFMISS_i + \theta_3 LIT79_i + \varepsilon_i \qquad [1]
\end{aligned}
$$

where $STREND_i$ is the 1979-2000 trend in weather station data from the Goddard Institute of Space Studies (GISS, http://data.giss.nasa.gov/gistemp/station_data), $PRESS_i$ is mean air pressure, $COSABLAT_i$ is cosine of absolute latitude, $WATER_i$ is a dummy (0,1) variable representing proximity to an ocean coast or a large body of water, POP_i is the GISS estimate of the local population, $SCALE79_i$ is the product of 1979 local per capita income and local population (i.e. a measure of the total scale of measured economic activity in the station's vicinity income of the station location), $COAL80_i$ is the 1980 total national coal consumption in million short tons, $COALGROW_i$ is the average (compound) annual increase in total coal consumption from 1980 to 1998, $INC79_i$ is 1979 real per capita income, $GDPGROW_i$ is average growth in annual real national Gross Domestic Product (GDP) from 1979 to 2000, $SOVIET_i$ is a dummy for membership in the former Soviet Union, $SURFMISS_i$ is the number of months between 1979:1 and 2000:12 in which the observation is missing and $LIT79_i$ is the 1979 average literacy rate for the country.

The use of weather station data was intended to provide a benchmark for the use of climatic data. Local weather station data are not adjusted for the contaminating influences of nearby socioeconomic activity, so it was expected that the various socioeconomic coefficients would be significant, as indeed they were. Then [1] was re-estimated using a different variable, $GTREND_i$, as the dependent variable. This is the linear trend over 1979-2000 in the Climatic Research Unit (CRU) grid cell temperature anomaly data (Brohan et al. 2006) for the same locations as each station in equation [1], where $i = 1,\ldots, 218$ is the location index. The CRU data are used by the IPCC and others on the assumption that the socioeconomic effects have been removed through various ad hoc adjustments described in Brohan et al (2006). Further details on all data sources are in McKitrick and Michaels (2004a). Equation [1] was estimated using Generalized Least Squares applying a correction for heteroskedastic errors. Application of an additional correction for clustering of error terms is discussed below. The working assumption of climate data users implies that the socioeconomic coefficients should vanish when $GTREND_i$ is the dependent variable. But MM found that the while the coefficients had somewhat smaller magnitude they were still highly significant. On this basis they concluded that the corrections for non-climatic biases in the data were inadequate. MM extended their data and testing framework in a 2007 paper, as described the next section.

Variable	MM04s	MM04sc	MM04g	MM04gc
press	0.016	0.025	0.011	0.016
	1.49	*2.22*	*2.88*	*3.45*
water	0.098	0.100	0.003	0.009
	1.80	*1.85*	*0.10*	*0.29*
cosablat	-0.381	**-0.555**	**-0.602**	**-0.660**
	-1.61	*-2.32*	*-5.95*	*-6.23*
pop	0.002	0.004	-0.002	-0.001
	0.56	*1.14*	*-1.07*	*-0.75*
scale79	-0.001	-0.003	0.001	0.001
	-0.31	*-0.87*	*0.47*	*0.25*
coal80	**-0.454**	**-0.507**	**-0.309**	**-0.278**
	-2.73	*-2.69*	*-3.98*	*-3.33*
coalgrow	-0.005	-0.006	0.001	0.001
	-1.28	*-1.32*	*0.23*	*0.34*
inc79	**0.040**	**0.037**	**0.018**	**0.014**
	3.56	*3.05*	*3.50*	*2.76*
gdpgrow	**0.085**	**0.080**	**0.026**	0.019
	4.12	*3.83*	*2.50*	*1.80*
soviet	**0.489**	**0.431**	0.129	0.080
	3.90	*3.13*	*1.88*	*1.22*
surfmiss	0.000	0.000	-0.003	-0.002
	0.19	*0.29*	*-0.48*	*-0.23*
lit79	**-0.005**	**-0.006**	**-0.003**	**-0.002**
	-3.18	*-3.22*	*-3.09*	*-2.77*
ao		**1.023**		0.343
		2.94		*1.61*
nao		**-0.946**		-0.218
		-2.96		*-1.12*
pdo		0.148		0.150
		1.04		*1.59*
so		-0.125		0.021
		-1.06		*0.24*
Constant	-15.392	**-24.357**	**-10.536**	**-14.995**
	-1.43	*-2.15*	*-2.69*	*-3.28*
P(X = 0)	**0.0021**	**0.0050**	**0.0004**	**0.0219**
P(Circ = 0)		**0.0046**		0.1061
N	218	218	205	205
r^2	0.26	0.30	0.38	0.41

TABLE 1. Results from McKitrick and Michaels (2004b,c) re-done introducing atmospheric circulation measures AO, NAO PDO and SO. Number in italics is *t*-statistic for coefficient immediately above. **Bold** denotes coefficients that are significant at 5%. First column (MM04s): reproduces original results, dependent variable is station trends, estimator is GLS with heteroskedasticity correction. Second column (MM04sc) introduces circulation indexes as correlation coefficients (variables ao—so). Third column (MM04g) reproduces MM04 results using grid cell trend as dependent variable. Fourth column (MM04gc) introduces circulation indexes as correlation coefficients (variables ao—so). P(X=0) is test that non-climatic effects (*Pop* through *Lit79*) are jointly zero. P(Circ=0) is test that atmospheric circulation effects (AO, NAO PDO, SO) are jointly zero. *N* is sample size, r^2 is coefficient of determination.

In order to assess the IPCC claim that the MM (2004a) result is spurious due to atmospheric circulation changes, I first obtained a set of correlation fields using the National Oceanic and Atmospheric Administration web site (NOAA, (http://www.cdc.noaa.gov/Correlation) to generate relevant terms for addition to the regression model. Various atmospheric circulation patterns are discussed in the IPCC Report, chiefly the Arctic Oscillation (AO), North Atlantic Oscillation (NAO), Pacific Decadal Oscillation (PDO) and the Southern Oscillation Index (SO), also known as the El Niño cycle. Each one is an oscillating pattern in air pressure or ocean-atmosphere interactions operating over long (multidecadal) timescales which are known to have strong influences on prevailing weather patterns over land. Indexes measuring the state of each of these oscillations are constructed, typically using pressure gradients across fixed measurement points, or principal components of pressure data over the region of interest.

The index values themselves cannot be used since there is only one monthly value for the whole world. What we are instead interested in is the influence of each oscillation pattern on ground temperatures around the world. The relationship at the gridcell level between surface air temperatures over the 1979-2001 interval and, respectively, the indexes of AO, NAO, PDO and the SO were obtained. These are the most appropriate measures to use for testing the IPCC claim, since they represent the component of temperature changes within a grid cell that are most directly associated with changes in the standard index of the state of the oscillation pattern.

Associations can be measured using a slope from a regression of each gridcell temperature series on the global index, or a Pearson correlation coefficient between the same measures. Both types of coefficient were tried and the one most favorable to the IPCC hypothesis was selected for each model. All results, data and code are available in Supplementary Information. Figure 1 illustrates the correlation field values for the AO.

Equation [1] was re-estimated on both station and gridded data after augmenting with AO_i, NAO_i, PDO_i and SO_i, each of which denotes the correlation in grid cell *i* between the indicated oscillation pattern and grid cell temperatures. The results are in Table 1. In the station data sample (columns 1—2) the socioeconomic coefficients retain their approximate size and significance levels and an *F* test shows they remain jointly significant (P = 0.005) after the circulation indexes are introduced (note that all joint parameter tests herein use *F* tests of linear restrictions). The AO and NAO terms are individually significant and the four circulation indexes are jointly significant (joint P = 0.005).

With gridded trends as the dependent variable (columns 3—4), there is no support for introducing circulation indexes into the model as they fail to achieve individual or joint significance (P = 0.106). If they are included anyway, three of the four significant socioeconomic coefficients remain significant and the fourth (GDP growth) falls to marginal significance, while the group remain jointly significant ($P = 0.022$).

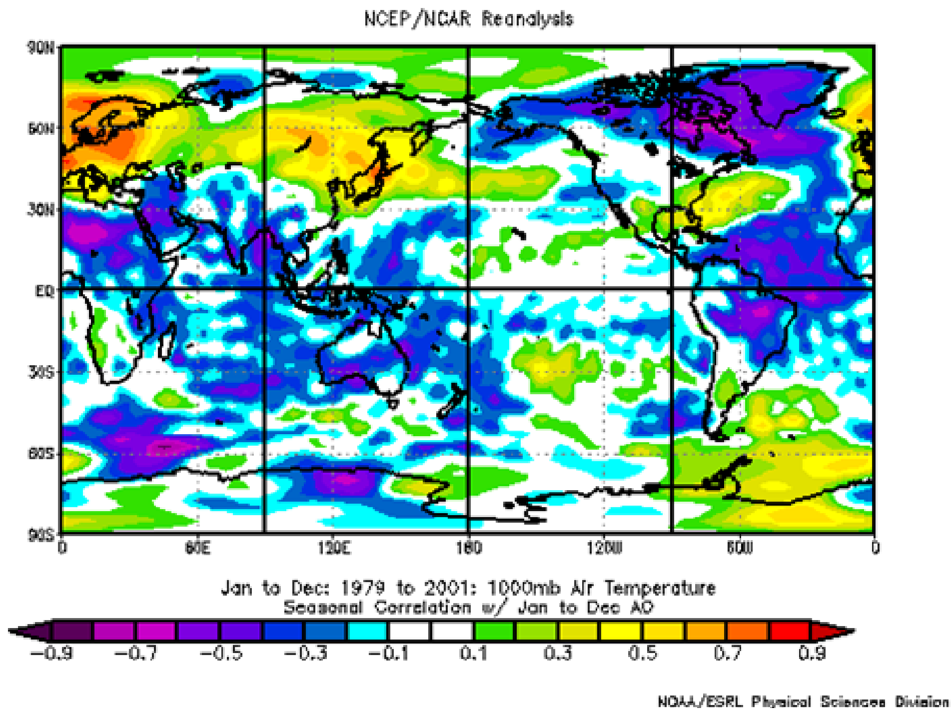

Figure 1. Partial correlation between surface air temperature and the AO index 1979-2001. Source: http://www.cdc.noaa.gov/Correlation.

Applying a correction for clustered standard errors (Moulton 1990) increases the error variances slightly. Atmospheric circulation indexes remain individually and jointly insignificant (joint $P = 0.124$) and thus their inclusion in the model is not supported. Other inferences remain the same in the station data. In the gridded data, GDP growth falls to insignificance but the economic variables (*POP* through *GDPgrow*) retain joint significance ($P = 0.035$) and all the socioeconomic variables together remain jointly marginally significant (P=0.077).

Note that McKitrick and Michaels (2004) included a suite of model specification tests, including sensitivity to removal of influential outliers, out-of-sample predictive ability, robustness to re-specification of the dependent variable as surface-troposphere trend differences, and insensitivity to stepwise inclusion of independent variables. Since the tests reported herein show that atmospheric circulation indexes are not supported in the model these test results remain valid.

MM07 undertook a wider set of model specification tests, and the effects of including atmospheric circulation indexes will be discussed in Section 3.

2.2 MM 2007 Model

MM07 developed a new and larger data set to re-test their 2004 results. They obtained temperature data for all 469 land-based grid cells with temperature data over 1979:1 to 2002:12 in the CRU 'crutem2v' edition data set, and corresponding socioeconomic data, and estimated the regression

$$\begin{aligned}\theta_i = \beta_0 &+ \beta_1 TROP_i + \beta_2 PRESS_i + \beta_3 DRY_i + \beta_4 DSLP_i + \beta_5 WATER_i \\ &+ \beta_6 ABSLAT_i + \beta_7 p_i + \beta_8 m_i + \beta_9 y_i + \beta_{10} c_i + \beta_{11} e_i + \beta_{12} g_i \\ &+ \beta_{13} x_i + u_i \qquad [2]\end{aligned}$$

where θ_t is the linear (Ordinary Least Squares) trend through monthly temperature anomalies in 5x5 degree grid cells (data obtained from http://ipcc-ddc.cru.uea.ac.uk), $TROP_i$ is the time trend of Microwave Sounding Unit (MSU)-derived temperatures in the lower troposphere in the same grid cell as θ_t over the same time interval (Spencer and Christy 1990), obtained from http://vortex.nsstc.uah.edu/data/msu/t2lt, $PRESS_i$ is as above, DRY_i is a dummy variable denoting when a grid cell is characterized by predominantly dry conditions (which is indicated by the mean dewpoint being below 0 ^{o}C), $DSLP_i$ is DRY_i x $PRESS_i$, $WATER_i$ is a dummy variable indicating the grid cell contains a major coastline, $ABSLAT_i$ denotes the absolute latitude of the grid cell, p_i is 1979-1999 population growth, m_i is the 1979 to 1999 percent change in real GDP per capita, y_i is the corresponding percent change in national GDP, c_i is the corresponding growth of national coal consumption, g_i is the GDP density (GDP per square km) as of 1979 and e_i is the average level of educational attainment as late in the interval as possible; each in the country where gridcell i is located; and x_i is the number of missing months in the observed temperature series for gridcell i over the interval 1979—2002. Complete details on data sources and model derivation are in MM07. Educational attainment is included as a measure of the difficulty of recruiting qualified technical staff to operate the meteorological monitoring network.

Equation [2] was estimated on the subsample (n = 440) excluding Antarctica and gridcells with too many missing values. They used Generalized Least Squares, controlling for heteroskedasticity and error clustering. Table 2 reports the MM07 results, and also shows the results from augmenting the model with indicators of atmospheric circulation patterns as above. Only the PDO is individually significant, and the four together are jointly significant (P = 0.0002).

Variable	MM07	MM07circ-r
trop	**0.8631**	**0.8279**
	8.62	*9.16*
slp	0.0044	0.0060
	1.02	*1.34*
dry	0.5704	-0.1276
	0.10	*-0.02*
dslp	-0.0005	0.0002
	-0.09	*0.04*
water	-0.0289	-0.0301
	-1.37	*-1.48*
abslat	0.0006	0.0008
	0.51	*0.60*
g	**0.0432**	**0.0520**
	3.36	*4.17*
e	**-0.0027**	**-0.0025**
	-5.14	*-5.57*
x	0.0041	0.0041
	1.66	*1.63*
p	**0.3839**	**0.4376**
	2.72	*3.20*
m	**0.4093**	**0.4463**
	2.39	*2.80*
y	**-0.3047**	**-0.3231**
	-2.22	*-2.54*
c	**0.0062**	**0.0057**
	3.45	*3.51*
ao		0.1319
		1.03
nao		-0.1101
		-1.32
pdo		**0.1845**
		2.31
so		0.1510
		1.25
constant	-4.2081	-5.8078
	-0.96	*-1.29*
P(X=0)	**7.1E-14**	**1.2E-15**
P(Circ=0)		**0.0002**
N	440	440
r^2	0.53	0.54

Table 2. Results from McKitrick and Michaels (2007) re-done introducing atmospheric circulation measures AO, NAO PDO and SOI. Number in italics is t-statistic for coefficient immediately above. Bold denotes coefficients that are significant at 5%. First column (MM07): reproduces main original results, dependent variable is grid cell trends, estimator is GLS with heteroskedasticity correction and error clustering. Second column (MM07circ-r) uses same dependent variable and estimator, and introduces atmospheric circulation variables measured using regression coefficients (ao—so). P(X=0) is test that non-climatic effects (g through c) are jointly zero. P(Circ=0) is test that atmospheric circulation effects (AO, NAO, PDO, SO) are jointly zero. N is sample size, r2 is coefficient of determination.

The remaining coefficients are quite robust to their inclusion. Five of the six significant socioeconomic indicators gain size and/or significance, and the joint significance *P* value falls to the 10^{-15} scale. The estimated average surface warming trend after removing contaminating effects falls from 0.30 C decade^{-1} to 0.17 in MM07 and to 0.16 C decade^{-1} in this case (see MM07 for filtering methodology).

3 FURTHER SPECIFICATION TESTS

The calculations in Section 2 falsify the assertion in IPCC (2007) that the strong correlations between indicators of industrialization and surface temperature trends become statistically insignificant upon controlling for the influence of atmospheric circulation patterns on temperatures. In this section I apply a battery of tests to the results from the augmented MM07 regressions, to check for the overall robustness of the findings and to rule out various other attempts to dismiss the findings as spurious.

3.1 Spatial Autocorrelation

Schmidt (2009) and Benestad (2004) argued that spatial autocorrelation (SAC) of the climate trend field might lead to exaggerated significance in the above regressions. The fact that a dependent variable exhibits SAC is not a problem for hypothesis testing if the model on the right hand side explains it and leaves an uncorrelated residual. Rewrite the regression model [2] in matrix notation as

$$\mathbf{y} = \mathbf{X}\beta + u \qquad [3]$$

where **y** is the linear trend in the temperature series for each of 440 surface grid cells, **X** is the matrix of climatic and socioeconomic covariates, β is the vector of least-squares slope coefficients and u is the residual vector. SAC in the residual vector can be treated using

$$u = \lambda \mathbf{W} u + e \qquad [4]$$

where λ is the autocorrelation coefficient, **W** is a symmetric nxn matrix of weights that measure the influence of each location on the other, and e is a vector of homoskedastic Gaussian disturbances (Pisati 2001). A test of H_0:λ=0 measures whether the error term in [2] is spatially independent. Anselin et al. (1996) shows that, if the alternative model allows for possible spatial dependence of the **y** variables, i.e.

$$\mathbf{y} = \phi\mathbf{Z}\mathbf{y} + \mathbf{X}\beta + e \qquad [5]$$

where **Z** is a matrix of spatial weights for **y** and may not be identical to **W**, standard adaptations of Wald and Lagrange Multiplier (LM) formulae yield tests that are severely biased towards over-rejection of the null. Anselin et al. (1996) propose a $\chi^2(1)$ Lagrange Multiplier (LM) test of λ=0 robust to possibly nonzero ϕ in [5], which has substantially superior performance in Monte Carlo evaluations compared to the non-robust LM test.

Hypothesis tests, and any subsequent parameter estimations, are conditional on the assumed form of the spatial weights matrix **W** in [4]. I consider three possibilities. Denote the great circle distance between the grid cell centers from which observation i and observation j are drawn as g_{ij}. Weighting matrix 1 (W1) is computed such that each element is $1/g_{ij}$ and the rows are standardized to sum to one. Weighting matrix 2 (W2) is computed such that each element is $1/\sqrt{g_{ij}}$ and the rows are standardized to sum to one. Weighting matrix 3 (W3) is computed such that each element is $1/g_{ij}^2$ and the rows are standardized to sum to one. Matrix W1 assumes the influence of adjacent cells diminishes at a hyperbolic rate. Matrix W2 assumes the inter-cell influence declines more slowly with distance while W3 assumes it declines more rapidly with distance.

Table 3 shows the LM test values (robust and non-robust) for weighting matrices W1—W3 applied to equation [2], with and without the atmospheric circulation terms. For the robust LM statistic (Panel *a*) in none of the three cases is there evidence of significant SAC in the residuals of [2]. The W2 weighting rule, which allows for the slowest decline in the influence of adjacent grid cells as the distance increases, shows the largest test score, though it is still insignificant. These results reverse for the non-robust test (Panel *b*) where W2 yields the smallest test scores. The other two weighting schemes reject the null in the MM07 case, though with the addition of the atmospheric circulation terms the score in W1 becomes insignificant.

While the evidence thus supports not treating for SAC, as a precaution we can nonetheless re-estimate [2] augmented with atmospheric circulation terms and applying [4] as the error term model, to make sure that important conclusions do not hinge on the decision about SAC. The parameter estimates change very little and the filtering method still causes the estimated average trend over land to fall from ~0.30 C/decade to ~0.18 C/decade (~0.27 C/decade to ~0.14 C/decade if gridcells are cosine-weighted). The joint socioeconomic effects remain highly significant (see Panel *c*).

Interestingly, if the results from the non-robust LM test under W3 are invoked to recommend adding the controls for SAC, the results in Panel *c* show that while the surface process and inhomogeneity effects remain highly significant, the atmospheric circulation effects become jointly insignificant under *a*

Weighting Matrix	Robust LM Score (χ^2 (1), *P* value)	
	MM07	MM07+circulations
W1 (Inverse-linear)	0.032 (0.858)	0.041 (0.840)
W2 (Inverse-square root)	2.564 (0.109)	2.796 (0.095)
W3 (Inverse-squared)	0.094 (0.759)	0.669 (0.413)

b

Weighting Matrix	Non-Robust LM Score (χ^2 (1), *P* value)	
	MM07	MM07+circulations
W1 (Inverse-linear)	**4.240 (0.039)**	2.69 (0.101)
W2 (Inverse-square root)	0.030 (0.862)	0.003 (0.954)
W3 (Inverse-squared)	**16.032 (0.000)**	**10.824 (0.001)**

c

Weighting Matrix	Joint Tests in Augmented MM07 Model Applying Controls for Spatial Autocorrelation	
	Socioecon (χ^2 (7), *P* value)	Circulation (χ^2 (4), *P* value)
W1 (Inverse-linear)	**48.05 (0.000)**	**10.13 (0.038)**
W2 (Inverse-square root)	**68.88 (0.000)**	**11.44 (0.022)**
W3 (Inverse-squared)	**23.64 (0.001)**	6.49 (0.166)

TABLE 3. Hypothesis tests for spatial autocorrelation in model [2] of surface temperature trends and inhomogeneity-anthropogenic surface process biases. **Bold** denotes coefficients that are significant at 5%. Panel *a*: Robust LM test on null hypothesis of no spatial dependence in the model residuals for MM07 model and MM07 augmented with atmospheric circulation variables. Panel *b*: Non-robust LM test on null hypothesis of no spatial dependence in the model residuals for MM07 model and MM07 augmented with atmospheric circulation variables. Panel *c*: Linear restrictions test on null hypothesis of no joint significance of, respectively, socioeconomic and circulation index variables in MM07 augmented with atmospheric circulation variables, controlling for spatial autocorrelation.

that specification. Likewise under W1, the non-robust LM hints at SAC, but not if circulation controls are added (Panel *b*). Hence there is no configuration of tests and models in Table 3 that indicates both spatial autocorrelation and significant atmospheric circulation effects; moreover the conclusions of MM07 concerning the significance of the non-climatic effects are unaffected by the decision to control for spatial autocorrelation or not.

3.2 Influence of Outliers

As in MM07, the results of the model with circulation indexes added were checked to make sure that outliers are not driving the conclusions. Observations were removed if the corresponding diagonal element of the OLS hat matrix exceeded twice the mean of the diagonal elements (Kmenta 1986, pp. 424-426). This resulted in removal of 21 observations, leaving a sample size of 419. The coefficients of the model without outliers were quite similar to those in MM07

Table 2, and the tests of contaminating influences remained highly significant. The Hausman chi-squared statistic was used to test for systematic change in the model parameters, yielding a $\chi^2(18)$ score of 16.39, which is insignificant (P = 0.565). Consequently there is no evidence that the conclusions are dependent on outlier observations.

3.3 Overfitting and Collinearity

Overfitting refers to the fact that if there are n observations and k independent variables, as k approaches n the model converges to a perfect fit even if none of the independent variables actually have any explanatory power. The indication that overfitting may be a problem is a significant joint F statistic for all independent variables even though none of them are individually significant. This can also arise if two or more independent variables are highly correlated, or collinear. The results reported herein clearly do not exhibit this problem, since many independent variables are individually significant. MM07 noted that correlations among the model variables were low, and variance inflation factors (VIF) indicated that the explanatory variables were nearly independent of one another. Augmenting the model with atmospheric circulation indexes leaves these results unchanged: the socioeconomic variables exhibit VIFs below 10, indeed most were below 3.

3.4 Regression Error Specification (RESET) and Endogeneity Tests

The RESET score was insignificant in MM07 (Sct. 4.3); likewise with the circulation indexes added it remains so (P = 0.564) indicating no evidence of untreated nonlinearity in the model structure. The Hausman endogeneity score also remained insignificant (P = 0.97) providing evidence that the model findings are not spurious effects of reverse causality (see discussion in MM07 Sct. 4.4).

3.5 Out-of-Sample Prediction Test

A good check against spurious results is the ability of a model to predict the values of a portion of the sample withheld during estimation. MM07 (Sct. 4.5) applied a test as follows. Thirty percent of the data were randomly selected and removed, then the model was re-estimated on the remaining 70%. The model was then used to predict the values of the withheld sample. If the model is perfectly accurate then a regression of the observations on the predicted values would yield a 45 degree line (constant = 0, slope = 1). An F-test of these coefficient restrictions is thus an exact test of whether the model systematically fails to predict out-of-sample.

In MM07 this procedure was repeated 500 times. The same number of repetitions were applied after the model was augmented with the atmospheric circulation terms. The mean value of the constant was 0.016, the mean slope coefficient was 0.946, the mean R^2 was 0.498 and the mean P value on the test H0:(constant = 0, slope = 1) was 0.373. These scores were almost identical to those in MM07, indicating that the additional terms neither added nor detracted from the model's stability for out-of-sample prediction.

3.6 Tropospheric Pattern

One of the tests in MM07 (Section 4.6) involved an alternative estimation in which the UAH-derived tropospheric trends were removed from the right hand side and used as the dependent variable instead of the surface trends. Had the socioeconomic coefficients retained their size and significance it would suggest that the surface results were spurious, since we would not expect the surface processes to have much effect at the height measured by the satellites (approximately 5 to 15 km aloft). The results in MM07 showed that the surface process variables did indeed lose size and significance in the troposphere, in line with expectations. However, that specification leaves a strong SAC pattern in the residuals. Augmenting the regression with the oscillation variables and applying an SAC correction (using the likelihood-maximizing inverse square weights) yields the expected reduction in magnitude of all the socioeconomic coefficients. The surface process variables become individually and jointly insignificant. Two of the variables that measure surface data quality become much smaller in size but remain (or become) significant, indicating that they are acting as a proxy for some regional effect, since the underlying variables cannot affect the satellite trends. The missing observation count becomes significant, but has a positive value only in 5% of the sample (mainly in the tropics) and is insignificant in the full model anyway. The educational attainment coefficient falls by three-quarters and changes sign, yet remains significant. In this case the educational attainment variations apparently overlap some aspect of the spatial trend pattern in the tropospheric record. But since the tropospheric trends are included in the full model anyway, this portion of the variance in the education variable is conditioned out in the full model, as indicated by the sign change. Finally, the SAC coefficient in this regression is extremely large (>0.994) indicating that a single spatial lag for this test regression is likely inadequate, and that the variances are likely overstated a bit. Hence the apparent significance levels may simply be due to underestimated variances.

4 Conclusions

Regional patterns of industrialization, land-use change and variations in the quality of temperature monitoring have been shown by several groups of authors to leave significant imprints on climate data, adding up to a widespread net warming bias that may account for as much as half the post-1980 warming over land. The Fourth Assessment Report of the IPCC dismissed this evidence with the claim that "the correlation of warming with industrial and socioeconomic development ceases to be statistically significant" upon controlling for atmospheric circulation patterns. This claim was presented without any supporting statistical evidence. The models in this paper implement a reasonable way of augmenting the original regressions with the relevant oscillation data, and the results contradict the IPCC claim. The temperature-industrialization correlations in question are quite robust to the inclusion of standard measures of the effects of atmospheric circulation patterns on temperatures, confirming the presence of significant extraneous signals in surface climate data on a scale that may account for about half the observed upward trend over land since 1980.

As discussed in the underlying papers by deLaat and Maurellis and McKitrick and Michaels, socioeconomic activity can lead to purely local atmospheric modifications (such as changes in water vapour and fine particle levels), which, along with other land-surface modifications and data inhomogeneities, can cause apparent trends in temperature data that are not attributable to general climatic changes. As was noted half a century ago by J. Murray Mitchell Jr., referring to the use of temperature observations for measuring climatic trends, "The problem remains one of determining what part of a given temperature trend is climatically real and what part the result of observational difficulties and of artificial modification of the local environment." (Mitchell Jr., 1953). The results herein show that this concern is still valid, and the conjecture invoked by the IPCC to dismiss it is not supported by the data. A substantial fraction of the post-1980 trends in gridded climate data over land are likely not "climatically real" but arise from measurement quality problems and local environmental modifications.

References

Anselin, L, A.K. Bera, R. Florax and M.J. Yoon (1996) Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26: 77—104.

Benestad R.E. (2004) Are temperature trends affected by economic activity? Comment on McKitrick & Michaels. (2004). *Climate Research* 27:171–173

Brohan, P., J.J. Kennedy, I. Harris, S.F.B. Tett and P.D. Jones (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *Journal of Geophysical Research* **111**, D12106, doi:10.1029/ 2005JD006548

Christy, J.R., W.B. Norris, K. Redmond, and K.P. Gallo (2006) Methodology and Results of Calculating Central California Surface Temperature Trends: Evidence of Human-Induced Climate Change? *Journal of Climate* Volume 19, Issue 4 (February 2006) DOI: 10.1175/JCLI3627.1

De Laat, A.T.J., and A.N. Maurellis (2004) Industrial CO_2 emissions as a proxy for anthropogenic influence on lower tropospheric temperature trends. *Geophysical Research Letters* Vol. 31, L05204, doi:10.1029/2003GL019024.

De Laat, A.T.J., and A.N. Maurellis (2006), Evidence for influence of anthropogenic surface processes on lower tropospheric and surface temperature trends. *International Journal of Climatology.* 26:897—913.

IPCC (2007) *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.).. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jun, Mikyoung, Reto Knutti and Douglas W. Nychka (2008) Spatial analysis to quantify numerical model bias and dependence: How many climate models are there? *Journal of the American Statistical Association* 108 No. 483 934—947 DOI 10.1198/016214507000001265.

Jones, P. D., D. H. Lister, and Q. Li (2008) Urbanization effects in large-scale temperature records, with an emphasis on China. *Journal of Geophysical Research 113*, D16122, doi:10.1029/2008JD009916.

Mahmood, R., R.A. Pielke Sr., K.G. Hubbard, D. Niyogi, G. Bonan, P. Lawrence, B. Baker, R. McNider, C. McAlpine, A. Etter, S. Gameda, B. Qian, A. Carleton, A. Beltran-Przekurat, T. Chase, A.I. Quintanar, J.O. Adegoke, S. Vezhapparambu, G. Conner, S. Asefi, E. Sertel, D.R. Legates, Y. Wu, R. Hale, O.W. Frauenfeld, A. Watts, M. Shepherd, C. Mitra, V.G. Anantharaj, S. Fall,R. Lund, A. Nordfelt, P. Blanken, J. Du, H.-I. Chang, R. Leeper, U.S. Nair, S. Dobler, R. Deo, and J. Syktus, (2010) Impacts of land use land cover change on climate and future research priorities. *Bulletin of the American Meteorological Society* in press

McKitrick, R.R. and P.J. Michaels (2004a) A test of corrections for extraneous signals in gridded surface temperature data. *Climate Research* 26(2) pp. 159-173

McKitrick, R.R. and P. J. Michaels (2004b) Erratum, *Climate Reearch.* 27(3) 265—268.

McKitrick, R.R. and P. J. Michaels (2004c) Are Temperature Trends Affected by Economic Activity? Reply to Benestad (2004) *Climate Research* 27:175-176.

McKitrick, R.R. and P.J. Michaels (2007) Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data. *Journal of Geophysical Research*, 112, D24S09,doi: 10.1029/2007JD008465.

McKitrick, R.R. and N. Nierenberg (2009) Correlations between Surface Temperature Trends and Socioeconomic Activity: Toward a Causal Interpretation**.** submitted to *International Journal of Climatology*.

Mitchell Jr., J.M. (1953) On the causes of instrumentally-observed secular temperature trends. *Journal of Meteorology* 10, pp. 244-261.

Moulton, B.R. (1990) An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units. *The Review of Economics and Statistics* 72(2) 334-338.

Pielke R.A. Sr., G. Marland, R.A. Betts, T.N. Chase, J.L. Eastman, J.O. Niles, D.D.S. Niyogi and S.W. Running (2002) The influence of land-use change and landscape dynamics on the climate system: Relevance to climate-change policy beyond the radiative effect of greenhouse gases. *Philosophical Transactions of the Royal Society of London Series A*. A360:1705-1719

Pisati, M. (2001) Tools for spatial data analysis. *Stata Technical Bulletin* STB-60, March 2001, 21—37.

Schmidt, M. (2009) Spurious correlations between recent warming and indices of local economic activity. *International Journal of Climatology* DOI:10.1002/joc.1831.

Spencer, R.W. and J.C. Christy (1990) Precise monitoring of global temperature trends from satellites. *Science* 247:1558—1562.