The effect of vegetation on surface temperature: A statistical analysis of NDVI and climate data

R. K. Kaufmann,[1] L. Zhou,[2] R. B. Myneni,[1] C. J. Tucker,[3] D. Slayback,[4] N. V. Shabanov,[1] and Jorge Pinzon[4]

Received 25 July 2003; revised 15 September 2003; accepted 23 October 2003; published 20 November 2003.

[1] We use statistical techniques to quantify the effect of interannual variations in vegetation within land covers on surface temperature in North America and Eurasia from satellite measures of surface greenness and ground based meteorological observations. During the winter, reductions in the extent of snow cover cause (in a statistical sense) temperature to rise. During the summer, increases in terrestrial vegetation within land covers cause (in a statistical sense) temperature to fall. Temperature-induced increases in vegetation have slowed increases in surface temperature, but this feedback may be limited by the range over which temperature has a positive effect on vegetation.
INDEX TERMS: 1610 Global Change: Atmosphere (0315, 0325); 1615 Global Change: Biogeochemical processes (4805); 1620 Global Change: Climate dynamics (3309); 1640 Global Change: Remote sensing. **Citation:** Kaufmann, R. K., L. Zhou, R. B. Myneni, C. J. Tucker, D. Slayback, N. V. Shabanov, and J. Pinzon, The effect of vegetation on surface temperature: A statistical analysis of NDVI and climate data, *Geophys. Res. Lett.*, *30*(22), 2147, doi:10.1029/2003GL018251, 2003.

1. Introduction

[2] Changes in terrestrial vegetation can modify local, regional, and global climate at diurnal, seasonal, and long-term scales [e.g., *Bounoua et al.*, 2000; *Bonan*, 1997; *Dickinson and Henderson-Seller*, 1988]. These vegetation-induced changes imply that warming enhanced vegetation growth and lengthened growing season [*Zhou et al.*, 2001] may offset a portion of anthropogenic warming. Here we use statistical techniques to quantify the feedback effect of vegetation on surface temperature in North America and Eurasia from satellite measures of surface greenness and ground based meteorological observations. Results indicate that reductions in the extent of snow cover increase temperature while increases in vegetation within land covers reduce temperature. Temperature-induced increases in vegetation have slowed increases in temperature, but this feedback may be limited by the range over which temperature has a positive effect on vegetation.

2. Methodology

[3] We analyze the relationship among surface greenness (which we interpret as a proxy for photosynthetically active vegetation), as measured by the GIMMS NDVI data (including AOD and solar zenith angle) [*Zhou et al.*, 2001], temperature [*Hansen et al.*, 1999], and precipitation [*Xie and Arkin*, 1997]. This data set has satisfied several quality criteria and has been used in several analyses [*Zhou et al.*, 2003; *Zhou et al.*, 2001]. The data are compiled by season; winter (Jan.–March), spring (April–May), summer (June–Aug.), and autumn (Sept.–Oct.) and georeferenced to a $2^\circ \times 2^\circ$ box. We use a land cover classification map with 8 km resolution to assign vegetated pixels to one of thirteen land covers [*DeFries et al.*, 1998]. Pixels that belong to the same land cover are averaged to generate values for each box. These data constitute a panel with eighteen observations (1982–1999) for 445 boxes in North America and 980 boxes in Eurasia.

[4] To determine whether terrestrial vegetation affects surface temperature, we use the notion of Granger causality [*Granger*, 1969]. Granger causality is based on the notion of predictability. Granger causality tests whether past values of a variable X contain statistically meaningful information about the current values of variable Y that is not contained in past values of variable Y and other relevant information. Should past values of variable X contain information about current values of variable Y beyond the information contained in the Y sequence and the other variables in the information set, variable X is said to "Granger cause" variable Y. The detection of Granger causality does not necessarily imply a physical causal mechanism between the two variables. Furthermore, the detection of Granger causality depends on the conditioning variables.

[5] To test for a causal relation from NDVI to surface temperature, we estimate the following equation:

$$\begin{aligned} T_{Sit} = {} & \alpha + \beta_1 Year + \beta_2 Lat_i + \beta_3 NDVI_{Sit-1} + \beta_4 T_{Sit-1} + \beta_5 P_{Sit-1} \\ & + \sum_{j=0}^{1} \delta_j SZA_{Sit-j} + \sum_{j=0}^{1} \gamma_j AOD_{Sit-j} \\ & + \sum_{j=0}^{1} \theta_j \left[\frac{1}{N-1} \sum_{k=1,k\neq i}^{N} T_{Si,t-j} \right] \\ & + \sum_{j=0}^{1} \phi_j \left[\frac{1}{N-1} \sum_{k=1,k\neq i}^{N} P_{Sit-j} \right] \end{aligned} \qquad (1)$$

[1] Department of Geography, Boston University, Boston, Massachusetts, USA.
[2] School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia, USA.
[3] Biospheric Sciences Branch, Code 923, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.
[4] Science Systems and Applications Inc., Biospheric Sciences Branch, Code 923, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

0094-8276/03/2003GL018251$05.00

Table 1. Regression Results for Equations 1, 2, and 3

	Eurasia					North America				
	β_3	S_{2a}	S_{3a}	Obs	ADF	β_3	S_{2a}	S_{3a}	Obs	ADF
Winter										
Cover 1	5.34**	−3.61**	−4.69**	583	−41.9**	4.00**	−4.66**	−5.06**	319	−25.8**
Cover 3	3.89**	3.78**	3.34**	219	−30.6**					
Cover 4	3.30**	−2.09*	−3.11**	283	−29.9**	1.82$^+$	−3.57**	−4.08**	149	−19.1**
Cover 5	2.81**	−1.79$^+$	−2.87*	431	−34.8**	3.85**	−3.68**	−0.57	197	−21.2**
Cover 6	6.12**	1.67	0.25	269	−34.4**	2.53*	−1.29	−1.53	242	−21.5**
Spring										
Cover 1	−2.08**	−3.39**	−3.24**	572	−29.6**	3.26**	−1.98*	−2.56*	273	−8.9**
Cover 3	−0.70	−0.84	−1.00	178	−14.4**					
Cover 4	−1.53**	−2.78**	−3.82**	307	−22.9**	1.44*	−1.91$^+$	−1.48	130	−5.94**
Cover 5	−2.66**	−3.41**	−4.07**	426	−27.0**	2.59**	−1.89$^+$	−1.75$^+$	178	−8.24**
Cover 6	−2.11*	−2.08*	−2.29*	182	−13.7**	0.82	0.72	1.74$^+$	102	−3.31*
Summer										
Cover 1	−1.53**	−1.50	−2.51*	668	−30.4**	−1.20*	−2.72*	−1.78$^+$	361	−18.6**
Cover 3	−4.71**	−1.06	−1.14	332	−19.3**					
Cover 4	−0.09	−0.86	−0.28	414	−26.1**	0.27	0.14	0.30	190	−18.9**
Cover 5	−0.74*	−0.85	−1.41	526	−27.5**	−0.04	2.84*	3.97**	226	−17.8**
Cover 6	−2.77**	−3.91**	−3.19**	510	−26.4**	−2.32**	−1.33	−1.76$^+$	292	−11.0**
Autumn										
Cover 1	0.40	−0.47	−0.45	666	−32.8**	−0.65	−0.76	−1.62	355	−19.2**
Cover 3	−0.46	0.35	2.00	307	−32.9**					
Cover 4	−1.32**	−0.50	−0.18	408	−23.9**	0.30	−0.80	1.12	187	−7.41**
Cover 5	−0.77*	−0.41	−0.26	521	−30.7**	−0.13	0.58	0.24	224	−10.8**
Cover 6	−0.22	−0.33	0.69	469	−29.9**	−1.24*	−0.39	−0.29	275	−22.0**

Cover refers to land cover. Obs refers to the number of grid boxes for which data are available (N in equation 1). Coefficients are statistically significantly different from zero at the: **1%, *5%, +10% level. The null hypothesis of the ADF test is no cointegration [*Pedroni*, 1999]. F tests indicate all equations can be estimated using OLS.

in which T is observed temperature during season S for grid box i at time t, Lat is latitude, P is precipitation, SZA is solar zenith angle, AOD is aerosol optical depth, and N is the number of boxes that contain a particular land cover in North America or Eurasia. The variables in equation (1) clarify the causal relationship between NDVI and temperature. Concerns that the NDVI data are contaminated by changes in solar zenith angle and aerosol optical depth are addressed by including SZA and AOD. Precipitation is included because it can affect temperature directly and indirectly via soil moisture and transpiration [*Bounoua et al.*, 2000].

[6] Granger causality from NDVI to temperature is indicated by the statistical significance of β_3. Rejecting the null hypothesis $\beta_3 = 0$ indicates that the lagged value of NDVI has information about the current value of temperature beyond that contained in the lagged value of temperature and the other variables in equation (1). This would indicate that NDVI "Granger causes" temperature. The nature of this effect is indicated by the sign on β_3. A negative value indicates that increases in NDVI cool surface temperature.

[7] Equation (1) is estimated with each season's data using all boxes for each of five land-covers; evergreen needleleaf forests (1), deciduous needleleaf forests (3), deciduous broadleaf forests (4), mixed forests (5), and woodlands (6), in North America or Eurasia. This separation implies that we measure the temperature effect of vegetation changes for specific land covers, not changes between land covers. The equations can be estimated using a variety of techniques. If the relationship among variables is spatially homogeneous (i.e., slopes and intercepts are the same across boxes), ordinary least squares can be used. A fixed or random effects estimator can be used if only the intercepts vary among boxes. Finally, if the slopes and intercepts vary among boxes, equation 1 can be estimated using a random coefficient model. We choose among these estimators using standard statistical procedures [*Hsiao*, 1986]. The number of lags (1) is the maximum value that allows us to perform these tests on the eighteen observations per box.

[8] We extend the analysis of Granger causality by testing whether equation 1 (unrestricted model) generates a more accurate out-of-sample forecast than a restricted version of equation 1 (restricted model), in which the lagged value of NDVI is eliminated by imposing $\beta_3 = 0$ [*Granger and Huang*, 1997]. To compute the out-of-sample forecast, we eliminate one box from the sample and estimate the unrestricted and restricted versions of equation (1) using observations from the remaining boxes. The regression results for the unrestricted and restricted models are used to generate an out-of-sample forecast for the eighteen temperature observations for the box excluded from the sample. This process is repeated for each box so that we have an out-of-sample temperature forecast for each box (N*18).

[9] We compare the accuracy of the two out-of-sample forecasts with tests for predictive accuracy using the following loss function:

$$d_t = \left[T_{Sit} - \hat{T}_{SitU}\right]^2 - \left[T_{Sit} - \hat{T}_{SitR}\right]^2 \qquad (2)$$

in which $\hat{T}_{SitU}$ is the out-of-sample temperature forecast generated by the unrestricted version of equation 1 and $\hat{T}_{SitR}$ is the out-of-sample temperature forecast generated by the restricted version of equation (1). The values of d_t are

Table 2. The Correlation Between Snow Cover Extent and NDVI

	Eurasia					North America				
	Winter			Spring		Winter			Spring	
	Jan	Feb	March	April	May	Jan	Feb	March	April	May
Land Cover 1	−0.055	−0.0040+	0.0004	−0.00647	−0.00032	−0.00023	−0.0038*	−0.0037*	−0.0146*	0.0072
Land Cover 3	−0.0032	−0.0037*	0.0040*	−0.0059*	0.0015					
Land Cover 4	−0.0037	−0.0032	−.0009	−0.0056	−0.0046	−0.00067	−0.0062*	−0.0051*	−0.0190*	0.0095
Land Cover 5	−0.004	−0.0045+	0.0003	−0.0071	−0.0025	−0.0018	−0.0056*	−0.0041*	−0.0182*	0.0085
Land Cover 6	−0.0072*	−0.0048*	0.004*	−0.0059	−0.0002	0.00041	−0.0029	−0.00042	−0.00591	0.0011

Coefficients are statistically significantly different from zero at the: **1%, *5%, +10% level. The relation between NDVI and SCE is estimated using OLS as follows: $NDVI_t = \alpha + \sum_{i=1}^{n}\beta_i SCE_{it} + \mu_t$ in which i is month and t is year. Data on snow cover extent from ftp://ftp.ncep.noaa.gov/pub/cpc/wd52dg/snow/snw_cvr_area/.

weighted and summed to generate the S_{2a} and S_{3a} statistic [*Diebold and Mariano*, 1995] as follows:

$$S_{2a} = \frac{\sum_{t=1}^{N} I_+(d_t) - 0.5N}{\sqrt{0.25N}}$$

$$S_{3a} = \frac{\sum_{t=1}^{N} I_+(d_t) rank(|d_t|) - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

$$I_+(d_t) = 1 \quad if\ d_t > 0 \quad = 0\ otherwise \tag{3}$$

The S_{2a} and S_{3a} statistics test the null hypothesis that the accuracy of the out-of-sample forecasts are equal. The more accurate model is identified by the sign on the test statistic which can be evaluated against a student's t with degrees of freedom equal to (N*18-1). If the forecast errors simulated by the unrestricted model are smaller (absolute terms), the test statistic will be negative and exceed the five percent critical value. This result would indicate that eliminating NDVI from equation 1 reduces the accuracy of the out-of-sample forecast and that NDVI "Granger causes" temperature.

3. Results

[10] Regression results indicate that the presence/absence of a causal relation from NDVI to surface temperature varies by season and region. For summer, β_3 generally is negative and the S_{2a} and S_{3a} statistics generally are negative and exceed the five percent critical threshold (Table 1). For winter, β_3 generally is positive and the S_{2a} and S_{3a} statistics generally are negative and exceed the five percent critical threshold. For spring, β_3 generally is positive in North America and negative in Eurasia. For both areas, the S_{2a} and S_{3a} statistics generally are negative and exceed the critical threshold. For autumn, β_3 and the S_2a and S_3a statistics generally are not statistically different from zero. The lack of any relation for autumn is consistent with results described by *Bonan* [1997].

4. Discussion and Conclusions

[11] Seasonal changes in the sign for β_3 are consistent with two mechanisms by which surface features affect temperature. Negative values for β_3 indicate that higher values for NDVI reduce surface temperature. During summer, NDVI is positively correlated with vegetation. The negative value of β_3 represents a cooling effect of terrestrial vegetation. Conversion of forest to short vegetation may warm temperatures due to increased sensible heat flux in relative to latent heat flux [*Eltahir*, 1996] or may cool the land surface in high latitudes due to oceanic influences associated with changes in land surface properties [*Bonan*, 1997; *Bonan et al.*, 1992]. Our finding of a summer-time cooling is consistent with the former and also with those generated by *Bounoua et al.* [2000], who focus on interannual changes in NDVI within land-cover types, but contradict the latter. *Bonan et al.* [1992] show that deforestation cools summer temperature because (1) the increased extent of sea ice due to the colder winter climate reinforces the cooling due to higher ocean albedo, and (2) the thermal lag effect of oceans inhibits warming in summer. *Bonan* [1997]

Table 3. Temperature Feedback Effect

	Eurasia					North America			
	1	3	4	5	6	1	4	5	6
Summer									
β_1	0.0078*	0.0106*	0.005*	0.0067*	0.0106*	0.0055*	0.0055*	0.0055*	0.0055*
β_2	−0.0009*	0.0005	−0.0009*	−0.0008*	−0.001*	0.0012*	0.0014*	0.0009*	0.0004*
β_3	−1.53*	−4.712*	−0.087	−0.744*	−2.771*	−1.202*	0.274	−0.035	−2.318*
Turning point	4.33		2.78	4.19	5.3				
NDVI effect	0.0069	0.011	0.0041	0.0059	0.0096	0.0067	0.0069	0.0064	0.0059
Temperature Feedback	−0.011	−0.052	0.0005	−0.0044	−0.0266	−0.0081	0.00189	−0.0002	−0.0137
% Feedback	−1.06	−5.23	0.054	−0.44	−2.66	−0.805	0.18906	−0.0224	−1.367

β_1 and β_2 refer to the coefficients associated with the linear and quadratic terms for summer temperature as estimated by *Zhou et al.* [2003]. As such, the effect of temperature on NDVI is calculated as: $\beta_1 T + \beta_2 T^2$ Turning point refers to the temperature increase at which the relation between summer temperature and summer NDVI changes from positive to negative. The turning point is given by $-(\beta_1/(2\beta_2)$. β_3 refers to the coefficient estimated from equation 1.

*Coefficients are statistically different from zero at the 5% level.

shows that replacing trees with crops alters surface properties such as roughness length, leaf, stomatal physiology, and surface albedo. The absolute and relative changes in these variables probably are large relative to the corresponding changes associated with interannual variations in NDVI within land-cover types. As such, our results may not be comparable to those that alter land-covers.

[12] Positive values for β_3 during winter imply that increases in NDVI increase surface temperature. NDVI is weakly correlated with vegetation during winter. Instead, winter values for NDVI are negatively correlated with the extent (not depth) of snow cover (Table 2). Consistent with these correlations, calculations based on the equations for radiative transfer indicate that NDVI is negatively related to snow cover. A reduction in snow cover extent (SCE) reduces surface albedo. This reduction increases the absorption of solar radiation, which increases near surface temperature. This creates a positive relation between NDVI and surface temperature. This result is consistent with those described by *Bounoua et al.* [2000] and *Bonan et al.* [1992].

[13] Differences in the magnitudes of these two mechanisms may be responsible for regional differences in the sign associated with β_3 for spring. The negative values of β_3 estimated for Eurasia imply that the cooling effect of vegetation dominates the warming effect of reduced SCE. The relative size of these two effects may be reversed in North America, where spring values for β_3 are positive. Consistent with this hypothesis, spring values for NDVI and SCE are correlated ($p < .05$) in North America, but there is no correlation between NDVI and SCE in Eurasia (Table 2).

[14] We evaluate the extent to which vegetation can damp warming by comparing the effect of temperature on NDVI with the effect of the resultant change in NDVI on temperature. We calculate the effect of a 1°C increase in summer temperature (relative to the 1982–1999 mean) on summer NDVI using the coefficients estimated statistically from this data set [*Zhou et al.*, 2003]. A 1°C rise increases summer NDVI by .0009–.0011 (Table 3). This increase in NDVI reduces the initial rise in temperature by 0.2% to 5.2% (Table 3).

[15] These results indicate that intact land-covers slowed the increase in surface temperature at mid and high latitudes in North America and Eurasia over the last twenty years. But this feedback may be effective over a narrow range of temperature increases. If temperature increases too rapidly, or if temperature increases beyond some critical value, vegetation may decline. Consistent with this potential, statistical estimates indicate that there is an inverted U shaped relation between summer temperature and summer NDVI for most land covers in Eurasia [*Zhou et al.*, 2003]. The turning points of these quadratic relations are 3°C–5°C greater than the 1982–1999 average. Beyond these turning points, further temperature increases would reduce NDVI. Such reductions would reinforce further temperature increases.

[16] **Acknowledgments.** This work was funded by the NASA Earth Science Enterprise.

References

Bonan, G. B., Effects of land use on the climate of the United States, *Clim. Change*, *37*, 449–486, 1997.
Bonan, G. B., D. Pollard, and S. L. Thompson, Effects of boreal forest vegetation on global climate, *Nature*, *359*, 716–718, 1992.
Bounoua, L., et al., Sensitivity of climate to changes in NDVI, *J. Clim.*, *13*, 2277–2292, 2000.
DeFries, R. S., et al., Global land cover classification at 8 km spatial resolution: the use of training data derived from Landsat imagery in decision tree classifiers, *J. Rem. Sens.*, *19*, 3141–3168, 1998.
Dickinson, R. E., and A. Henderson–Seller, Modeling tropical deforestation: A study of GCM land surface parameterizations, *Q. J. Meteorol. Soc.*, *114*, 439–462, 1988.
Diebold, F. X., and R. S. Mariano, Comparing predictive accuracy, *J. Bus. and Econ. Stat.*, *13*, 253–263, 1995.
Eltahir, E. A. B., Role of vegetation in sustaining large-scale atmospheric circulations in the tropics, *J. Geophys. Res.*, *101*(D2), 4255–4268, 1996.
Granger, C. W. J., Investigating causal relations by econometric models and cross spectral models, *Econometrica*, *37*, 424–438, 1969.
Granger, C. W. J., and L. Huang, Evaluation of Panel Data Models: Some suggestions from time series, Department of Economics, University of California, San Diego, 1997.
Hansen, J., et al., GISS analysis of surface temperature change, *J. Geophys. Res.*, *104*, 30,997–31,022, 1999.
Hsiao, C., Analysis of Panel Data, Cambridge: Cambridge University Press, 1986.
Pedroni, P., Critical values for cointegration tests in heterogeneous panels with multiple regressors, *Oxford Bulletin of Economics and Statistics, Special issue*, 653–670, 1999.
Xie, P., and P. A. Arkin, Global precipitation: A 17–year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs, *Bull. Am. Meteorol. Soc.*, *78*, 2539–2558, 1997.
Zhou, L., et al., Variations in northern vegetation activity inferred from satellite data of vegetation index during 1981 to 1999, *J. Geophys. Res.*, *106*, 20,069–20,083, 2001.
Zhou, L., et al., Relation between interannual variations in satellite measures of vegetation greenness and climate between 1982 and 1999, *J. Geophys. Res.*, *108*(D1), 4004, doi:10.1029/2002JD002510, 2003.

R. K. Kaufmann, R. B. Myneni, and N. V. Shabanov, Department of Geography, 675 Commonwealth Avenue, Boston University, Boston, MA 02215, USA. (kaufmann@bu.edu)
L. Zhou, School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA.
C. J. Tucker, Biospheric Sciences Branch, Code 923, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.
D. Slayback and J. Pinzon, Science Systems and Applications Inc., Biospheric Sciences Branch, Code 923, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA.