Journal of Exposure Science and Environmental Epidemiology (2014), 1–10
© 2014 Nature America, Inc. All rights reserved 1559-0631/14
www.nature.com/jes

ORIGINAL ARTICLE

Statistical evaluation of the feasibility of satellite-retrieved cloud parameters as indicators of $PM_{2.5}$ levels

Chao Yu[1,2,3], Larry Di Girolamo[4], Liangfu Chen[1], Xueying Zhang[2] and Yang Liu[2]

The spatial and temporal characteristics of fine particulate matter ($PM_{2.5}$, particulate matter $<2.5\,\mu$m in aerodynamic diameter) are increasingly being studied from satellite aerosol remote sensing data. However, cloud cover severely limits the coverage of satellite-driven $PM_{2.5}$ models, and little research has been conducted on the association between cloud properties and $PM_{2.5}$ levels. In this study, we analyzed the relationships between ground $PM_{2.5}$ concentrations and two satellite-retrieved cloud parameters using data from the Southeastern Aerosol Research and Characterization (SEARCH) Network during 2000–2010. We found that both satellite-retrieved cloud fraction (CF) and cloud optical thickness (COT) are negatively associated with $PM_{2.5}$ levels. $PM_{2.5}$ speciation and meteorological analysis suggested that the main reason for these negative relationships might be the decreased secondary particle generation. Stratified analyses by season, land use type, and site location showed that seasonal impacts on this relationship are significant. These associations do not vary substantially between urban and rural sites or inland and coastal sites. The statistically significant negative associations of $PM_{2.5}$ mass concentrations with CF and COT suggest that satellite-retrieved cloud parameters have the potential to serve as predictors to fill the data gap left by satellite aerosol optical depth in satellite-driven $PM_{2.5}$ models.

Journal of Exposure Science and Environmental Epidemiology advance online publication, 23 July 2014; doi:10.1038/jes.2014.49

Keywords: $PM_{2.5}$; MODIS; SEARCH network; aerosol optical depth; cloud fraction; cloud optical thickness

INTRODUCTION

Numerous epidemiological studies have shown that fine particles ($PM_{2.5}$, particles with aerodynamic diameter $<2.5\,\mu$m) are associated with cardiovascular and respiratory morbidity and mortality.[1,2] Accurate $PM_{2.5}$ exposure estimates are crucial to air quality assessment and environmental health research. Measurements from ground-monitoring sites, which have high accuracy and reliable temporal coverage regardless of meteorological conditions, have been used in many epidemiological studies.[3] However, ground monitors are costly to operate, therefore have limited spatial coverage.

In 1999 and 2002, the National Aeronautics and Space Administration (NASA) launched its first two Earth Observing System satellites Terra and Aqua into polar orbits.[4,5] Since then, an increasing body of literature showed that ground-level $PM_{2.5}$ can be quantitatively estimated from satellite-retrieved column aerosol optical depth (AOD), after accounting for the impact of land use and meteorological parameters on the $PM_{2.5}$–AOD relationship.[6–10] However, a major issue of applying satellite data in $PM_{2.5}$ exposure modeling is missing data owing to cloud cover, as satellite-retrieved AOD values are only available in regions with little or no cloud cover. For example, MODIS-Terra AOD data were available only about 50% of the time owing to cloud cover and unfavorable surface conditions.[11] Missing data severely reduce the statistical power of $PM_{2.5}$ epidemiological models and can limit the choices of model structure (e.g., lag structure in $PM_{2.5}$ exposure estimates is difficult to implement), particularly as there is no reason to expect $PM_{2.5}$ to be the same for cloudy and clear days simply based on the meteorological controls on cloud formation and air pollution.[10] Spatial smoothing techniques with land use parameters, such as highway length and emission source locations, have been proposed to fill the data gaps.[12] However, these parameters are not temporally varying, and therefore cannot reflect the dynamic effect of cloud on daily $PM_{2.5}$ levels. Moreover, as spatial interpolation assumes a smooth and continuous transition of $PM_{2.5}$ levels from cloud-free regions to cloudy regions, it is likely that the resulted interpolated surfaces are overly smoothed and underestimate the true spatial variability in $PM_{2.5}$, especially when ground-level monitoring data are sparse.[13]

Previous studies showed that $PM_{2.5}$ level is probably related to cloud properties. Dawson et al.[14] found increases in cloud liquid water content, optical depth, and cloudy area led to decreases in simulated $PM_{2.5}$ concentrations over land in January and July in Pittsburgh and Atlanta, respectively, although the negative impacts were not significant. Tai et al.[15] showed a negative correlation between column cloud cover and total $PM_{2.5}$ concentrations in the Southeastern United States. Although Christopher et al.[11] reported that cloud cover is not a major problem for inferring monthly to annual $PM_{2.5}$ from space-borne sensors, their results indicated that mean $PM_{2.5}$ values under available satellite-derived AOD conditions are higher than from all ground measurements. Liu et al.[9] developed a two-stage generalized additive model to estimate daily $PM_{2.5}$ concentrations in cloud-free and

[1]State Key Laboratory of Remote Sensing Science, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China; [2]Department of Environmental Health, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [3]University of the Chinese Academy of Sciences, Beijing, China and [4]Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA. Correspondence: Dr. Yang Liu, Department of Environmental Health, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322, USA.
Tel.: +1 404 727 2131. Fax: +1 404 727 8744.
E-mail: yang.liu@emory.edu
Received 13 February 2014; accepted 27 May 2014

cloudy regions separately, and found significant differences in the spatial pattern of predicted $PM_{2.5}$ concentrations between cloud-free and cloudy regions.

Operational satellite aerosol remote sensing algorithms first identify pixels as "clear" or "cloudy" before attempting retrieval of aerosol microphysical or optical property. Retrieval of aerosol properties such as AOD is attempted on those pixels identified as "clear", and retrieval of cloud properties is attempted on those pixels identified as "cloudy". As such, the presence of satellite cloud data often signals missing AOD data, so they have the potential to fill the data gap left by satellite AOD in $PM_{2.5}$ statistical models. However, there has been little research to date on the association between cloud parameters and $PM_{2.5}$, and how to use this information in $PM_{2.5}$ exposure modeling. In this analysis, we explored the statistical associations of satellite-retrieved cloud properties with $PM_{2.5}$ mass and constituent concentrations using 11 years of data in the Southeastern United States. In this region, secondary ionic species and organic matter (OM) formed by photochemical reactions in the atmosphere contributed to $>50\%$ of the $PM_{2.5}$ mass.[16–18] The persistently high temperature (Temp) from May to October, accompanied by increased biogenic volatile organic compound emissions and sulfate precursor emissions result in highly active secondary production of $PM_{2.5}$.[17] Cloud cover might negatively affect $PM_{2.5}$ mass concentrations through the attenuation of photochemical reactions. Our objective is to evaluate the feasibility of satellite cloud parameters as statistical predictors of $PM_{2.5}$ concentrations in order to improve the spatial and temporal coverage of satellite-driven $PM_{2.5}$ statistical models.

MATERIALS AND METHODS

Cloud Data

The cloud data used in this study is Collection 5.1 Level-2 cloud products (MOD06), retrieved from the moderate resolution imaging spectro-radiometer (MODIS) sensor aboard NASA's Terra satellite. We used daytime Terra MODIS cloud data with best quality and processing flags (usefulness flag = 1 and confidence flag = 3) in this study.[19] Cloud parameters processed for the current analysis include cloud fraction (CF) and cloud optical thickness (COT) (Table 1). In our preliminary analysis, the correlations of $PM_{2.5}$ mass concentrations with CF and COT are more significant than other parameters such as cloud phase and cloud water path.

CF is calculated from each 5×5 group of 1-km resolution cloud mask pixels and has a 5-km spatial resolution at nadir, representing the fraction of a 25 km^2 area covered by clouds as observed from above by MODIS. For each $PM_{2.5}$ ground-monitoring site, mean CF is calculated from the 5×5 group of CF pixels, that is, a $25 \times 25\,km^2$ area centered at the site. COT is derived from MODIS 0.645 μm, 2.13 μm, and 3.75 μm bands over land, and represents the optical thickness of clouds at visible wavelengths.[19] COT data were resampled from its original 1-km resolution to 5-km resolution to match the spatial resolution of CF data, then mean COT was calculated from the 5×5 pixel group centered at each ground $PM_{2.5}$ monitoring site. This method is often used to reduce uncertainties in the instantaneous satellite observations when compared with more accurate and time-averaged, ground-based measurements.[11] The choice of a $25 \times 25\,km^2$ area also avoids overlapping of CF and COT calculation areas between neighboring $PM_{2.5}$ sites, except for Pensacola (PNS) and outlying landing field no.8 (OLF) (Figure 1), which are very close to each other.

We also obtained 11 years of Terra MODIS Collection 5.1 Level-2 aerosol products (MOD04) and examined a 5×5 group of the 10-km pixels centered at each site (Table 1). The larger AOD matching area ($50 \times 50\,km^2$) used to calculate mean AOD value for each ground site ensures clear separation of AOD coverage from cloud coverage. We divided the satellite-ground matched data records into three groups based on aerosol and cloud data availability at each site: Group 1 includes data records with only AOD retrievals and no cloud retrievals; Group 2 includes data records with both AOD and cloud retrievals; and Group 3 includes data records with only cloud retrievals and no AOD retrievals. The number of data records and mean values of $PM_{2.5}$ concentrations, AOD, and CF of each group are shown in Table 2, which shows that MODIS-Terra AOD data matched to the ground $PM_{2.5}$ sites were only available in 48.5% of the time. Because we are interested in how to use information provided by MODIS cloud products to supplement AOD data in $PM_{2.5}$ exposure models, Group 1 data records were excluded from further analysis. Table 2 also shows that nearly 25% of the $PM_{2.5}$ measurements are matched with both AOD and cloud retrievals. Having this subset may significantly increase the number of data records in small CF bins. However, the interaction between aerosols and clouds is complex. Satellite measurements of aerosols in the vicinity of clouds are influenced by several factors including aerosol hygroscopic growth in the humid air surrounding clouds, cloud-related particle changes, and 3-D radiative effects.[20,21] Because this is the first attempt to study the association between ground $PM_{2.5}$ levels and MODIS cloud parameters, we would like to focus on situations with clearly no AOD retrievals. Therefore, Group 2 data records were also excluded from data analysis to simplify the interpretation of our results.

Ground $PM_{2.5}$ Measurements and Meteorological Data

Daily mean $PM_{2.5}$ and speciation concentrations from 2000 to 2010 were obtained from the Southeastern Aerosol Research and Characterization (SEARCH) Network (Figure 1). SEARCH sites collect daily average total $PM_{2.5}$ mass concentrations measured with the federal reference method (FRM) every day, and speciation data including sulfate, nitrate, ammonium, OM, elemental carbon (EC), and major metal oxides (MMO) every day or every third day. SEARCH measurements are guided by a detailed quality control and assurance protocol, much of which is derived from EPA guidelines.[22] We selected daily FRM $PM_{2.5}$ measurement for

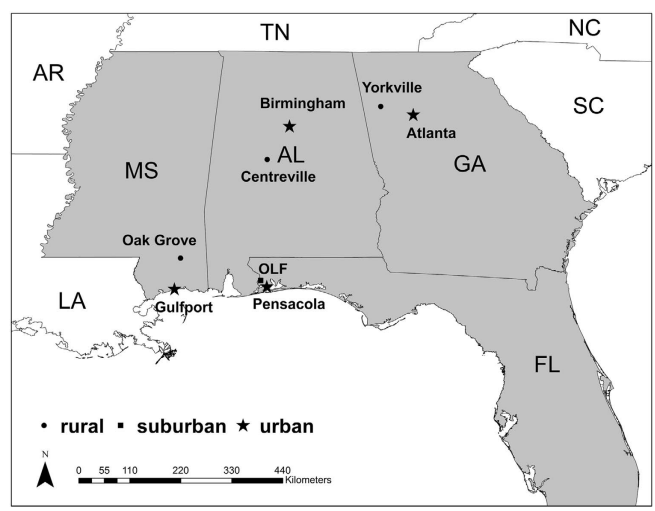

Figure 1. Geographic distribution of the eight monitoring sites in SEARCH network.

Table 1. The cloud and aerosol data products and parameters used in this study.

Parameter	*Satellite/product*	*SDS name*	*Resolution*	*QA*	*Reference*
CF	Terra/MOD06	"Cloud Fraction Day"	5 km	usefulness = 1 confidence = 3	19
COT	Terra/MOD06	"Cloud Optical Thickness"	1 km	usefulness = 1 confidence = 3	19
AOD	Terra/MOD04	"Optical Depth Land And Ocean"	10 km	None	27,28

© 2014 Nature America, Inc.

Table 2. Mean values of $PM_{2.5}$, AOD, and cloud fraction of three groups of MODIS aerosol/cloud observations for all data points and each site.

	Group 1			*Group 2*				*Group 3*		
	Num	*$PM_{2.5}$ (μg/m³)*	*AOD*	*Num*	*$PM_{2.5}$ (μg/m³)*	*AOD*	*CF (%)*	*Num*	*$PM_{2.5}$ (μg/m³)*	*CF (%)*
All	7347	12.65	0.078	5942	14.72	0.25	58.0	14106	11.20	89.6
ATL	894	14.48	0.099	719	18.17	0.29	59.4	1794	13.98	91.9
YRK	971	11.63	0.081	716	15.12	0.27	55.9	1790	11.90	90.6
BHM	934	16.51	0.061	846	19.94	0.27	60.0	1797	13.98	91.8
CTR	1103	11.67	0.077	737	13.54	0.25	56.2	1706	10.26	89.8
GFP	806	11.58	0.081	720	12.48	0.24	56.4	1857	9.65	86.8
OAK	984	11.33	0.064	658	12.54	0.23	58.9	1803	9.91	87.4
PNS	732	13.51	0.087	792	13.21	0.24	58.9	1499	10.47	89.1
OLF	923	10.91	0.075	754	12.02	0.24	58.2	1860	9.41	89.3

Group 1: Data points with only AOD retrievals. Group 2: Data points with both AOD and cloud retrievals. Group 3: Data points with only cloud retrievals.

Table 3. Description of the SEARCH sites.

Name	*Type*	*Setting*	*Latitude*	*Longitude*	*Days*	*$PM_{2.5}$ (μg/m³)*
Atlanta (ATL)	Urban	Industrial-residential	33.776	84.413	3556	15.00
Yorkville (YRK)	Rural	Forest-agricultural	33.931	85.046	3571	12.44
Birmingham (BHM)	Urban	Industrial-residential	33.553	86.815	3721	16.05
Centreville (CTR)	Rural	Forest	32.902	87.250	3653	11.37
Gulfport (GFP)	Urban	Residential	30.391	89.050	3464	10.72
Oak Grove (OAK)	Rural	Forest	30.985	88.932	3554	10.73
Pensacola (PNS)[a]	Urban	Residential	30.437	87.256	3165	11.96
Outlying Landing Field no.8 (OLF)	Suburban	Forest-grass	30.551	87.376	3683	10.39

Days is the number of days of data available between 1 January 2000 and 31 December 2010, $PM_{2.5}$ is the mean $PM_{2.5}$ mass concentrations ($\mu g/m^3$) for all days.
[a]$PM_{2.5}$ data from PNS site are from 2000 to 2009.

comparison with MODIS data because it is the national ambient air quality standard for $PM_{2.5}$ and there are twice as many FRM monitors nationwide as continuous monitors. Previous studies also indicated that satellite measurements were highly correlated with daily average ground measurements.[23,24] The eight SEARCH sites are arranged into four urban-rural (or suburban) pairs in four southeast states (Mississippi, Florida, Alabama, and Georgia). This configuration covers different land use categories (residential, forest, agriculture, and industrial) and allows the investigation of local and regional source influences, as well as coastal and inland conditions (Table 3).

The mean $PM_{2.5}$ mass concentrations of group 3 are lower than group 1 (except for Yorkville (YRK)) and group 2, which indicates that cloud cover might have a negative association with $PM_{2.5}$ mass concentrations (Table 2). In addition, the relationship between cloud cover and $PM_{2.5}$ concentrations might be associated with meteorological parameters such as wind speed (WS), Temp, relative humidity (RH), solar radiation (SR), and precipitation (PRECIP). Hourly meteorological measurements from the SEARCH sites were averaged to daily values and matched with $PM_{2.5}$ measurements.

Statistical Analysis

Relationships of $PM_{2.5}$ concentrations with CF and COT were examined respectively for all of the group 3 data points. The COT data were arranged into 25 value bins separately, with an equal number of data points in each bin. As $>30\%$ of the matched observations have a CF of 100%, we put these data points in an individual bin (bin 26) and arranged the rest of the CF data points into 25 bins. Detailed information for 26 value bins of CF and 25 value bins of COT is provided in Supplementary Materials (Supplementary Tables S1–S4). Mean values of $PM_{2.5}$ mass concentration, each major $PM_{2.5}$ constituent and meteorological variable within each CF or COT bin were calculated. As season, land use type, and locations were important factors in the satellite-driven statistical models of $PM_{2.5}$,[23] we also divided the data set into seasonal subsets (winter: DJF, spring: MAM, summer: JJA, fall: SON), land use subsets (urban sites: ATL, BHM, GFP, PNS; rural/suburban sites: YRK, CTR, OAK, OLF), or locational subsets (inland sites: ATL, YRK, BHM, CTR; coastal sites: GFP, OAK, PNS, OLF) according to the configuration of the SEARCH network. Data points within each subset were also sorted by the cloud parameter and arranged into 26 CF bins and 25 COT bins for regression analyses. We used linear piecewise regression to examine the associations between $PM_{2.5}$ mass concentrations and CF/COT for easier interpretation and comparison across regression results from seasonal, land use, and locational subsets. To further examine the impact of cloud cover on secondary particle generation, simple linear regression and piecewise linear regression were also used to examine the association between $PM_{2.5}$ constituent concentrations and CF and COT, respectively. Additional site-specific analysis and results on the comparison of various model formats are provided in Supplementary Materials.

RESULTS

A total of 14,106 data points with available cloud retrievals but missing AOD values were identified from 2000 to 2010 (Table 4). The average $PM_{2.5}$ concentrations vary by season, land use type, and location. Summer has the highest mean $PM_{2.5}$ levels, followed by spring, fall, and winter. Long-term average $PM_{2.5}$ levels at inland and urban sites are 20–25% higher than coastal and rural sites, respectively. Average CF and COT are $\sim 10\%$ higher in winter than the rest of the year. There are no significant differences in CF and COT values between urban and rural subsets, or between inland and coastal subsets.

$PM_{2.5}$ concentration is negatively correlated with CF ($R^2 = 0.74$, Slope $= -0.046$), and this association varies significantly by season (Figure 2). The negative slope between CF and $PM_{2.5}$ is stronger in summer ($R^2 = 0.65$, Slope $= -0.054$) and fall ($R^2 = 0.49$, Slope $= -0.054$) than in spring ($R^2 = 0.20$, Slope $= -0.022$), whereas the correlation in winter is statistically insignificant (P-value > 0.05). In contrast, the association between $PM_{2.5}$

© 2014 Nature America, Inc.

concentration and CF does not change substantially between urban (slope = −0.047) and rural sites (slope = −0.047) as shown in Figure 2c, except the constant discrepancy of $PM_{2.5}$ mass concentration levels between them. A *t*-test showed that the regression slope difference between urban and rural sites is insignificant (*P*-value >0.05). $PM_{2.5}$ mass concentrations decrease more significantly at inland sites (Slope: −0.057) than at coastal sites (Slope: −0.043) with increasing CF, indicating that CF is a more effective indicator of change in $PM_{2.5}$ levels at inland sites than at coastal sites. Site-specific plots of $PM_{2.5}$-CF relations also show similar results (Supplementary Materials, Supplementary Figure S1).

Figure 3 shows that $PM_{2.5}$ concentration is negatively associated with COT. However, the distribution of COT values is uneven, with ~75% of COT values below 20. We calculated the number, median, mean, and SD of COT, CF, $PM_{2.5}$ total mass, and constituent concentrations (nitrate, sulfate, ammonium, OM, EC, and MMO), and meteorological variables (WS, Temp, RH, SR, and PRECIP) for subsets of COT values larger and smaller than 20 (Supplementary Table S5 in the Supplementary Materials), and the statistical characteristics of COT values ≤20 differ substantially from that of COT values >20. For low COT values (≤20), the SD of COT is smaller than high COT values (>20), and the COT values are more concentrated. However, CF, $PM_{2.5}$ (especially for sulfate and nitrate), and SR all have substantially larger SDs than those with high COT values. Consequently, we defined an empirical cutoff of COT = 20 to fit a linear regression line for COT ≤20 and COT >20 separately. Compared with more complex model formats including power-law, parabola, and

Table 4. Descriptive statistics (mean ± SD) for $PM_{2.5}$, CF, and COT in each subset.

Subset	*Num*	*$PM_{2.5}$ (μg/m³)*	*CF (%)*	*COT*
All	14106	11.2 ± 6.2	89.6 ± 17.2	16.5 ± 17.5
Season				
Winter	3383	9.9 ± 5.5	95.8 ± 10.9	21.5 ± 20.3
Spring	3534	11.2 ± 5.4	88.8 ± 17.8	16.7 ± 18.1
Summer	4132	12.4 ± 6.9	85.2 ± 19.3	11.8 ± 13.3
Fall	3057	11.0 ± 6.5	89.5 ± 17.0	17.0 ± 16.7
Land use				
Urban	6947	12.1 ± 6.5	89.9 ± 16.3	16.3 ± 17.7
Rural	7159	10.3 ± 5.7	89.3 ± 18.0	16.7 ± 17.3
Location				
Inland	7087	12.6 ± 6.7	91.0 ± 16.4	18.0 ± 18.0
Coastal	7019	9.8 ± 5.3	88.1 ± 17.8	14.9 ± 16.8

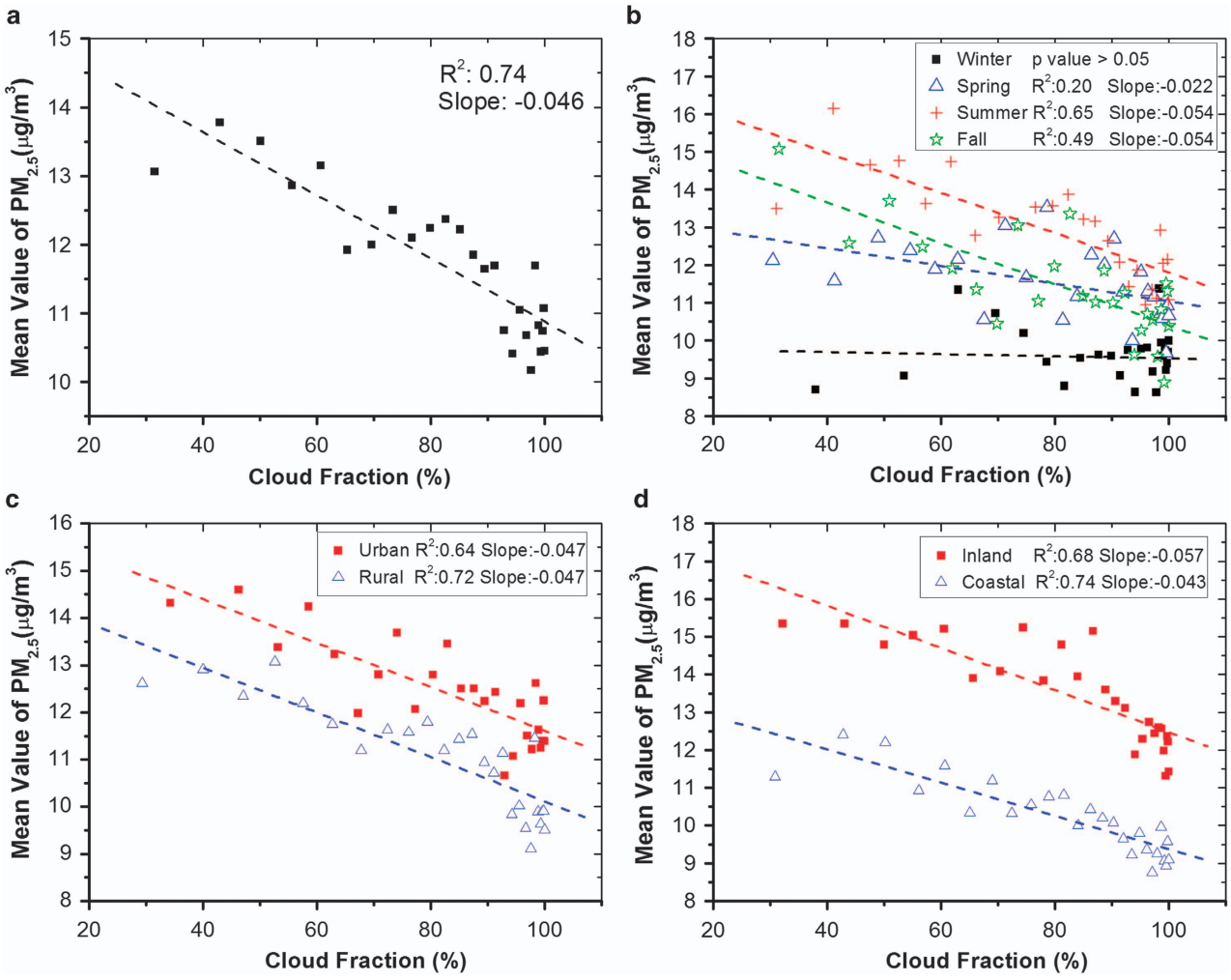

Figure 2. Relationships between $PM_{2.5}$ mass concentrations and cloud fraction for (a) all data, (b) seasonal subsets, (c) land use subsets and (d) location subsets. Dashed lines represent a linear fit of data points for individual data groups.

© 2014 Nature America, Inc.

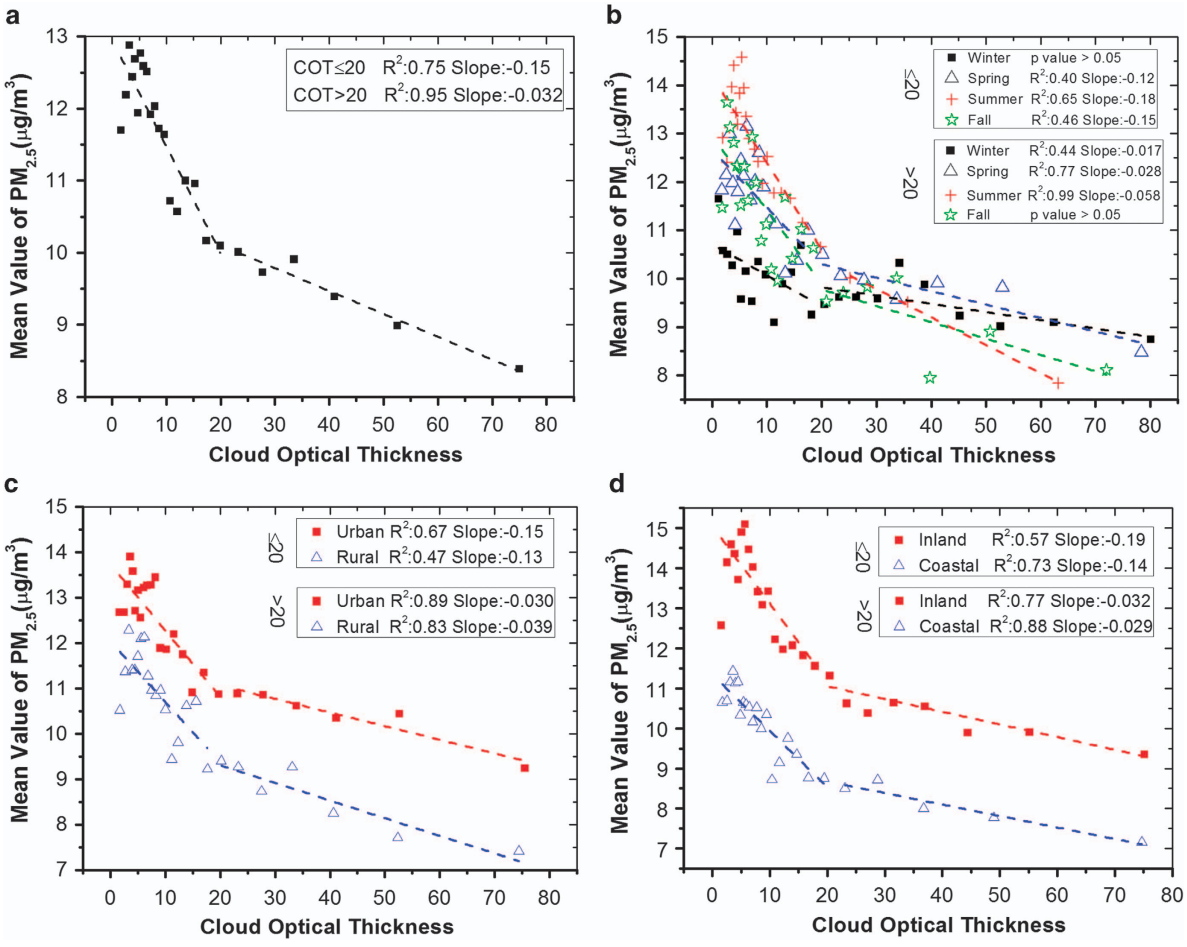

Figure 3. Relationships between $PM_{2.5}$ mass concentrations and cloud optical thickness for (a) all data, (b) seasonal subsets, (c) land use subsets and (d) location subsets. Dashed lines represent a linear fit of data points for individual data groups.

exponential equations, the piecewise linear function has the highest model R^2 and is easier to interpret (Supplementary Materials, Supplementary Figure S2 and Supplementary Table S6). Our results showed that the $PM_{2.5}$ concentration decreases more rapidly with COT when COT is ≤ 20 (Slope: -0.15) as compared with when COT is >20 (Slope: -0.032). Similar to CF, the association between $PM_{2.5}$ concentration and COT also varies by season (Figure 3b). The negative association of COT with $PM_{2.5}$ is significant in warmer seasons (summer: $R^2=0.65$, Slope $=-0.18$; fall: $R^2=0.46$, Slope $=-0.15$; spring: $R^2=0.40$, Slope $=-0.12$), whereas insignificant in winter (P-value >0.05). Figure 3c shows that the negative associations between COT and $PM_{2.5}$ levels are comparable at urban and rural sites, and are almost identical to that of the whole data set. The linear regression slopes in both land use types are neither significantly different from each other nor the slope derived using the entire data set at $\alpha=0.05$ level. As shown in Figure 3d, COT has a slightly greater impact on inland sites (slope: -0.19) than coastal sites (slope: -0.14) when COT is ≤ 20.

Figure 4 shows that secondary $PM_{2.5}$ constituents including sulfate ($R^2=0.73$, slope $=-0.021$) and OM ($R^2=0.37$, slope $=-0.011$) decrease significantly with increasing CF. Primary constituents such as EC and MMO are less affected as indicated by lower model R^2 values and close-to-zero slopes. Nitrate levels increase with CF ($R^2=0.31$, slope $=0.0019$). Figure 5 shows decreasing trends of sulfate, ammonium, OM with COT, but the negative slopes are greater at low COT levels (i.e., ≤ 20) (sulfate: $R^2=0.61$, slope $=-0.062$; ammonium: $R^2=0.34$, slope $=-0.013$; OM: $R^2=0.78$, slope $=-0.076$) than at high COT levels (>20) (sulfate: $R^2=0.91$, slope $=-0.020$; ammonium: $R^2=0.90$, slope $=-0.006$; OM: $R^2=0.64$, slope $=-0.014$). The levels of primary $PM_{2.5}$ constituents such as EC and MMOs decrease with increasing COT only at low COT levels.

Figure 6 shows that CF and COT are positively associated with WS, RH, and PRECIP. Stronger horizontal mixing with increasing WS as well as wet deposition are expected to lower $PM_{2.5}$ levels under clouds. In addition, Tai et al.[15] found a negative correlation of $PM_{2.5}$ with RH in the Southeast based on observed meteorological and $PM_{2.5}$ data. However, the mechanism for the relationship between RH and $PM_{2.5}$ is complex as individual $PM_{2.5}$ components show different correlations with RH.[15] The associations of meteorological parameters with COT are generally stronger at low COT levels as indicated by steeper negative slopes than high COT levels.

The seasonal variation of the relationship of $PM_{2.5}$ with CF and COT might be associated with temperature. Table 5 shows that sulfate and OM concentrations are the highest and nitrate concentrations the lowest in summer; these patterns are reversed in winter. As a result, the impact of cloud cover would be more pronounced in warm seasons than in cold seasons. The fractions of sulfate and OM in $PM_{2.5}$ are similar between urban and rural sites, which could result in comparable negative associations between cloud properties (CF, COT) and $PM_{2.5}$ levels at urban and rural sites. Owing to the slight differences in sulfate and OM mass

© 2014 Nature America, Inc.

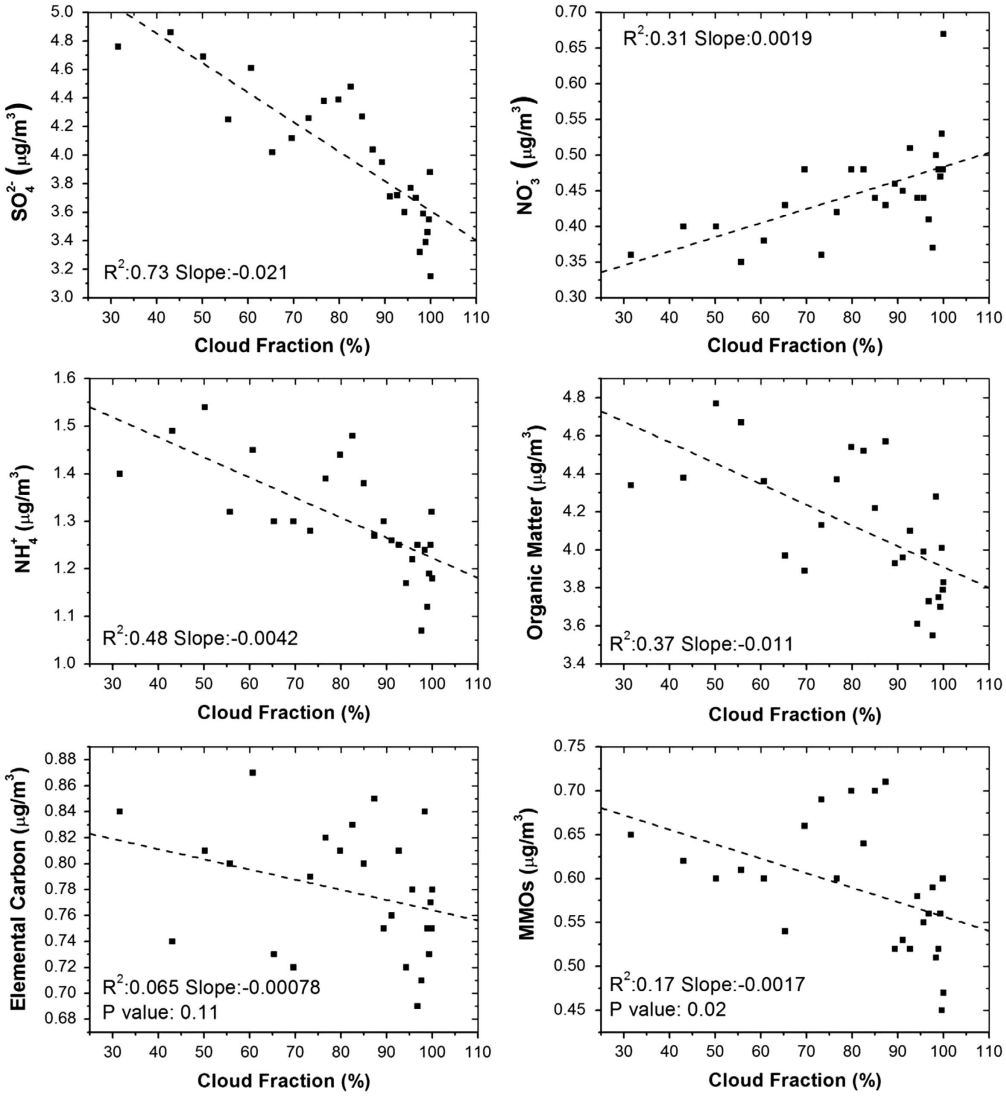

Figure 4. Relationships between $PM_{2.5}$ speciation and cloud fraction.

concentrations between inland sites and coastal sites, $PM_{2.5}$ mass concentrations decrease a little more significantly at inland sites than at coastal sites with increasing cloud cover.

DISCUSSION

A major limiting factor of current satellite-driven $PM_{2.5}$ exposure models in air pollution health effects research is missing data due to cloud cover. Using MODIS-retrieved cloud parameters and ground $PM_{2.5}$ mass concentrations, we showed that there are significant and systematic differences in $PM_{2.5}$ levels between cloudy and cloud-free regions, and $PM_{2.5}$ mass concentrations are negatively correlated with CF and COT. During our study period, secondary particles such as sulfate and OM account for up to 70% of the total $PM_{2.5}$ mass at each SEARCH site (Supplementary Figure S3, Supplementary Materials). The negative association between total $PM_{2.5}$ and cloud cover can be largely attributed to the negative response of sulfate and OM to cloud cover. The decrease in sulfate concentration is probably associated with the attenuation of photochemical reactions under cloud cover, although it may be partly offset by the positive influence from in cloud SO_2 oxidation at high RH.[15] For OM, this might be related to less frequent prescribed burns and wildfires under cloudy and high humidity conditions, which are major emissions sources of carbonaceous aerosols in this region.[25] The positive association of nitrate with cloud cover likely reflects the RH dependence of the ammonium nitrate formation equilibrium and decreased volatilization of ammonium nitrate with lower Temp under cloud cover.[14,26] As the counter-ion for sulfate and nitrate, the negative association of ammonium with cloud cover is likely due to the combined effects of sulfate and nitrate.[15] Our piecewise regression results, which showed significant different regression slopes between $PM_{2.5}$ and COT at COT ≤ 20 and COT > 20, might also be the reflection of the impact of cloud cover on secondary particle generation. Light transmittance of clouds is roughly equal to 1/COT for large COT values (personal communication with Dr. Steven Platnick). When COT is > 20, clouds might be thick enough to prevent direct sunlight from reaching the ground. As a result, the increase in COT no longer has a strong impact on $PM_{2.5}$ levels. The smaller regression slope at high COT values might also be related to the fact that lower Temp below clouds can lead to reduced photochemical reaction rates.

© 2014 Nature America, Inc.

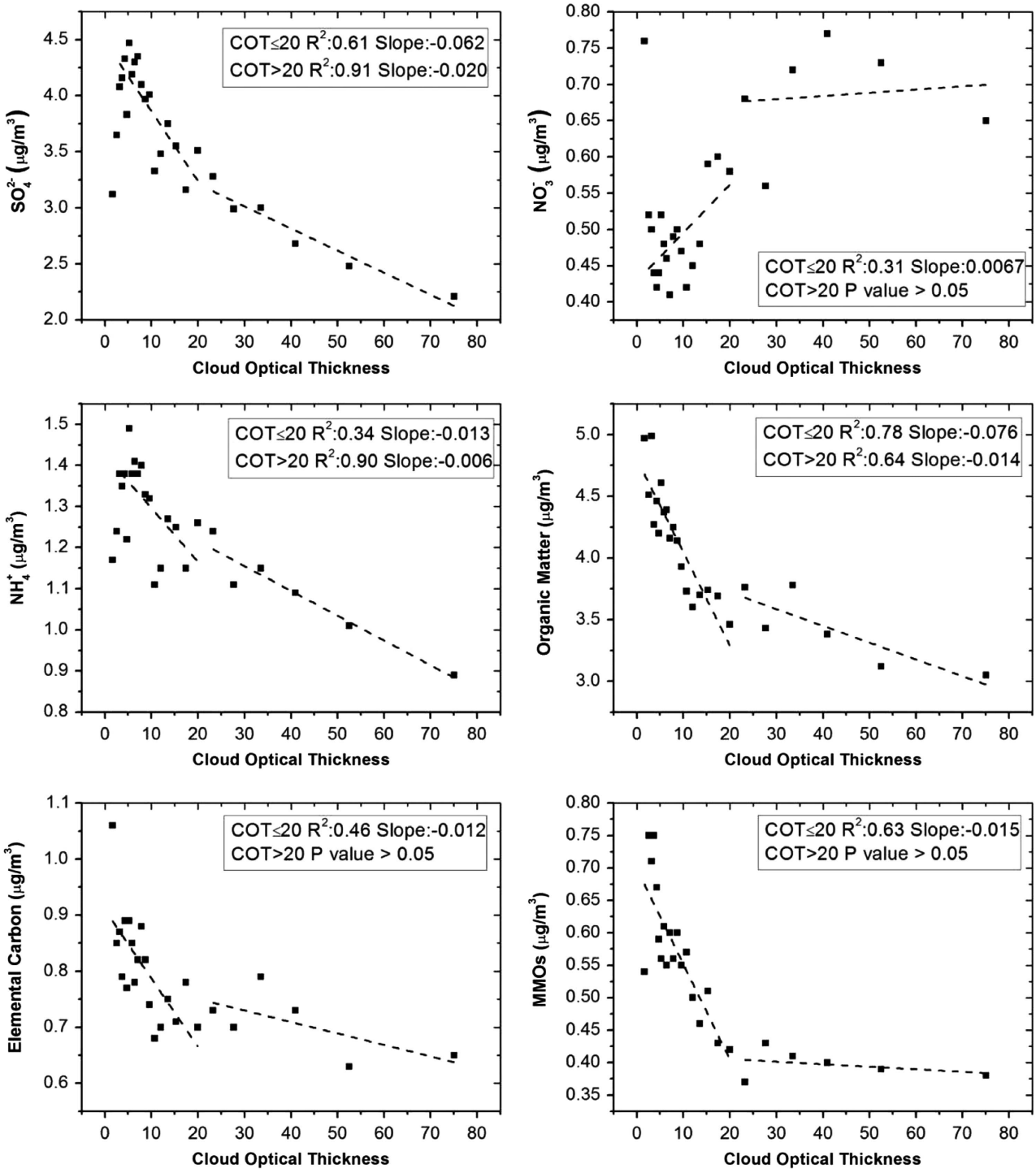

Figure 5. Relationships between $PM_{2.5}$ speciation and cloud optical thickness.

However, it should be noted that the production and removal of $PM_{2.5}$ is a complex interaction among various surface and meteorological variables, and these interactions are difficult to unravel with only the observation data sets used in this study. For example, the impact of PRECIP on $PM_{2.5}$ mass and constituent concentrations cannot be fully explored without additional information such as cloud type. Our empirical analysis cannot fully explain the different associations between $PM_{2.5}$ mass and constituent concentrations and low/high COT levels either. Further examination should be based on satellite-ground observations and atmospheric chemistry models to delineate the relationship between cloud conditions and ground $PM_{2.5}$.

Nonetheless, our empirical data analysis showed that MODIS cloud parameters are strongly associated with ground $PM_{2.5}$ mass concentrations. On one hand, satellite cloud parameters alone are insufficient to estimate $PM_{2.5}$ concentrations, as there is significant variability in $PM_{2.5}$ levels within each CF or COT value bin. On the other hand, almost all current satellite $PM_{2.5}$ models involve

© 2014 Nature America, Inc.

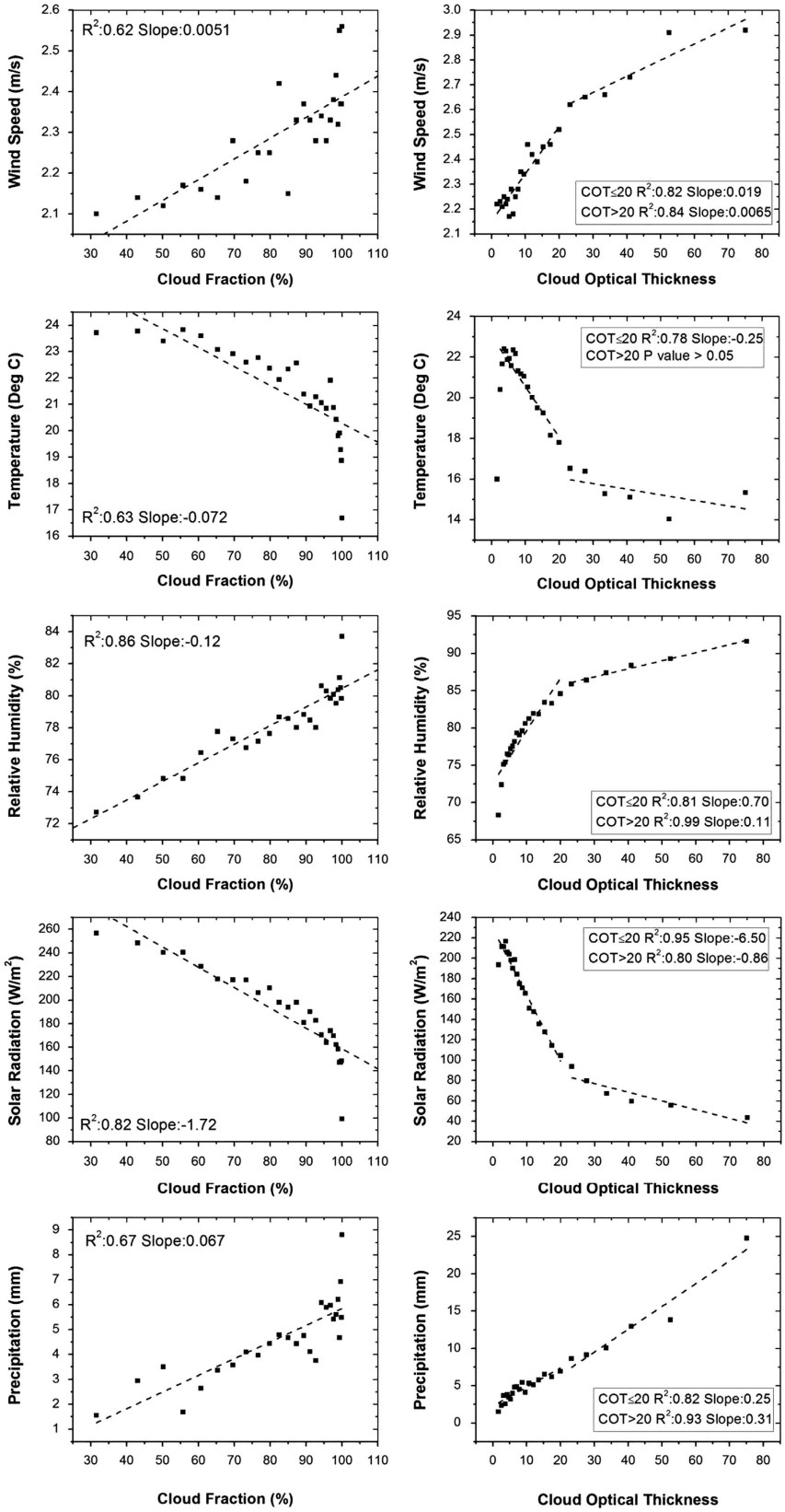

Figure 6. Relationships between meteorological variables and CF and COT.

multiple covariates such as meteorological and land use parameters to modify and strengthen the association between satellite AOD and $PM_{2.5}$.[9,12] Each of these covariates is correlated with $PM_{2.5}$, but is not robust enough to estimate $PM_{2.5}$ alone. However, together they form highly predictive models. Similarly, satellite cloud parameters will need to be used in conjunction with these

© 2014 Nature America, Inc.

Table 5. Mean values of $PM_{2.5}$ constituent concentrations ($\mu g/m^3$) for SEARCH subsets on cloudy days during 2000–2010.

	$PM_{2.5}$ ($\mu g/m^3$)	SO_4^{2-} ($\mu g/m^3$)	NO_3^- ($\mu g/m^3$)	NH_4^+ ($\mu g/m^3$)	EC ($\mu g/m^3$)	OM ($\mu g/m^3$)	MMOs ($\mu g/m^3$)
All	11.2	3.61	0.55	1.25	0.78	3.98	0.53
Season							
Winter	9.88	2.41	0.95	1.05	0.79	3.83	0.29
Spring	11.22	3.64	0.51	1.25	0.75	3.95	0.47
Summer	12.4	4.49	0.31	1.38	0.76	4.16	0.9
Fall	11.02	3.67	0.45	1.27	0.82	3.9	0.36
Type							
Urban	12.06	3.77	0.63	1.35	0.96	4.31	0.62
Rural	10.36	3.38	0.43	1.1	0.49	3.45	0.43
Location							
Inland	12.56	3.95	0.69	1.46	0.94	4.58	0.51
Coast	9.83	3.07	0.3	0.92	0.5	2.95	0.56

covariates to estimate $PM_{2.5}$ concentrations although a demonstration of such models is beyond the scope of this paper.

In conclusion, the statistically significant relationships between $PM_{2.5}$ mass concentrations and cloud properties (CF and COT) observed in our study serve an important purpose in $PM_{2.5}$ exposure modeling, especially for predicting $PM_{2.5}$ daily concentrations. Our paper is the first attempt to establish the feasibility of satellite-retrieved cloud parameters to be considered in similar ways to other meteorological and land use parameters in satellite $PM_{2.5}$ models. Given the global and almost daily coverage of satellite instruments such as MODIS, these cloud parameters may be developed into predictors of $PM_{2.5}$ levels. The regression slopes obtained in the current analysis can be used together with predicted $PM_{2.5}$ concentrations derived from satellite AOD in cloud-free conditions to estimate $PM_{2.5}$ concentrations under cloud cover, effectively doubling the coverage of satellite $PM_{2.5}$ models. The relationships between cloud properties and $PM_{2.5}$ mass concentrations described in this study could be also used as observational constraints for atmospheric chemistry models.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This work was partially supported by NASA Applied Sciences Program (grant no. NNX09AT52G, PI: Y Liu). In addition, this publication was made possible by USEPA grant R834799. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the USEPA. Further, USEPA does not endorse the purchase of any commercial products or services mentioned in the publication. The work of Chao Yu was supported by the China Scholarship Council (CSC) under the State Scholarship Fund. We thank Dr. Steven Platnick and Jerusha Barton for their technical support.

REFERENCES

1 Brook RD, Franklin B, Cascio W, Hong YL, Howard G, Lipsett M *et al.* Air pollution and cardiovascular disease - a statement for healthcare professionals from the expert panel on population and prevention science of the American heart association. *Circulation* 2004; **109**: 2655–2671.
2 Dominici F, Peng RD, Bell ML, Pham L, McDermott A, Zeger SL *et al.* Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA* 2006; **295**: 1127–1134.
3 Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K *et al.* Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA* 2002; **287**: 1132–1141.
4 Kaufman YJ, Herring DD, Ranson KJ, Collatz GJ. Earth Observing System AM1 mission to earth. *IEEE Trans Geosci Remote Sens* 1998; **36**: 1045–1055.
5 Parkinson CL. Aqua: an earth-observing satellite mission to examine water and other climate variables. *IEEE Trans Geosci Remote Sens* 2003; **41**: 173–183.
6 Hoff RM, Christopher SA. Remote sensing of particulate pollution from space: have we reached the promised land? *J Air Waste Manage Assoc* 2009; **59**: 645–675.
7 Kloog I, Koutrakis P, Coull BA, Lee HJ, Schwartz J. Assessing temporally and spatially resolved PM(2.5) exposures for epidemiological studies using satellite aerosol optical depth measurements. *Atmos Environ* 2011; **45**: 6267–6275.
8 Lee HJ, Liu Y, Coull BA, Schwartz J, Koutrakis P. A novel calibration approach of MODIS AOD data to predict PM(2.5) concentrations. *Atmos Chem Phys* 2011; **11**: 7991–8002.
9 Liu Y, Paciorek CJ, Koutrakis P. Estimating regional spatial and temporal variability of PM2.5 concentrations using satellite data, meteorology, and land use information. *Environ Health Perspect* 2009; **117**: 886–892.
10 Dey S, Di Girolamo L, van Donkelaar A, Tripathi SN, Gupta T, Mohan M. Variability of outdoor fine particulate (PM2.5) concentration in the Indian subcontinent: a remote sensing approach. *Remote Sens Environ* 2012; **127**: 153–161.
11 Gupta P, Christopher SA. Seven year particulate matter air quality assessment from surface and satellite measurements. *Atmos Chem Physics* 2008; **8**: 3311–3324.
12 Kloog I, Nordio F, Coull BA, Schwartz J. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal pm2.5 exposures in the Mid-Atlantic states. *Environ Sci Technol* 2012; **46**: 11913–11921.
13 Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T *et al.* A review and evaluation of intraurban air pollution exposure models. *J Expo Anal Environ Epidemiol* 2005; **15**: 185–204.
14 Dawson JP, Adams PJ, Pandis SN. Sensitivity of PM2.5 to climate in the Eastern US: a modeling case study. *Atmos Chem Physics* 2007; **7**: 4295–4309.
15 Tai APK, Mickley LJ, Jacob DJ. Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: implications for the sensitivity of PM2.5 to climate change. *Atmos Environ* 2010; **44**: 3976–3984.
16 Chen YJ, Zheng M, Edgerton ES, Ke L, Sheng GY, Fu JM. PM2.5 source apportionment in the southeastern U.S.: spatial and seasonal variations during 2001–2005. *J Geophys Res-Atmos* 2012; **117**: D08304.
17 Edgerton ES, Hartsell BE, Saylor RD, Jansen JJ, Hansen DA, Hidy GM. The southeastern aerosol research and characterization study: Part II. Filter-based measurements of fine and coarse particulate matter mass and composition. *J Air Waste Manage Assoc* 2005; **55**: 1527–1542.
18 Lee S, Russell AG, Baumann K. Source apportionment of fine particulate matter in the southeastern united states. *J Air Waste Manage Assoc* 2007; **57**: 1123–1135.
19 Platnick S, King MD, Ackerman SA, Menzel WP, Baum BA, Riedi JC *et al.* The MODIS cloud products: algorithms and examples from Terra. *IEEE Trans Geosci Remote Sens* 2003; **41**: 459–473.
20 Varnai T, Marshak A, Yang W. Multi-satellite aerosol observations in the vicinity of clouds. *Atmos Chem Physics* 2013; **13**: 3899–3908.
21 Chand D, Wood R, Ghan SJ, Wang MH, Ovchinnikov M, Rasch PJ *et al.* Aerosol optical depth increase in partly cloudy conditions. *J Geophys Res Atmos* 2012; **117**: D17207.
22 Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM *et al.* The southeastern aerosol research and characterization study: Part 1-overview. *J Air Waste Manage Assoc* 2003; **53**: 1460–1471.

© 2014 Nature America, Inc.

23 Liu Y, Franklin M, Kahn R, Koutrakis P. Using Aerosol Optical Thickness to Predict Ground-Level PM2.5 Concentrations in the St. Louis Area: a Comparison Between MISR and MODIS. *Remote Sens Environ* 2007; **107**: 33–44.
24 Wang J, Christopher SA. Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies. *Geophys Res Lett* 2003; **30**: 2095.
25 Park RJ, Jacob DJ, Logan JA. Fire and biofuel contributions to annual mean aerosol mass concentrations in the United States. *Atmos Environ* 2007; **41**: 7389–7400.
26 Stelson AW, Seinfeld JH. Relative-humidity and temperature-dependence of the ammonium-nitrate dissociation-constant. *Atmos Environ* 1982; **16**: 983–992.
27 Martins JV, Tanre D, Remer L, Kaufman Y, Mattoo S, Levy R. MODIS Cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Res Lett* 2002; **29**: MOD4-1–MOD04-4.
28 Remer LA, Kaufman YJ, Tanre D, Mattoo S, Chu DA, Martins JV *et al*. The MODIS aerosol algorithm, products, and validation. *J Atmos Sci* 2005; **62**: 947–973.

Supplementary Information accompanies the paper on the Journal of Exposure Science and Environmental Epidemiology website (http://www.nature.com/jes)

© 2014 Nature America, Inc.