Diagnostic Model Evaluation for Carbonaceous PM₂.₅ Using Organic Markers Measured in the Southeastern U.S.

PRAKASH V. BHAVE,*,†
GEORGE A. POULIOT,† AND MEI ZHENG‡

Atmospheric Sciences Modeling Division, Air Resources Laboratory, National Oceanic and Atmospheric Administration, Research Triangle Park, North Carolina 27711, and School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia 30332

Summertime concentrations of fine particulate carbon in the southeastern United States are consistently underestimated by air quality models. In an effort to understand the cause of this error, the Community Multiscale Air Quality model is instrumented to track primary organic and elemental carbon contributions from fifteen different source categories. The model results are speciated using published source profiles and compared with ambient measurements of 100 organic markers collected at eight sites in the Southeast during the 1999 summer. Results indicate that modeled contributions from vehicle exhaust and biomass combustion, the two largest sources of carbon in the emission inventory, are unbiased across the region. In Atlanta, good model performance for total carbon (TC) is attributed to compensating errors: overestimation of vehicle emissions with underestimations of other sources. In Birmingham, 35% of the TC underestimation can be explained by deficiencies in primary sources. Cigarette smoke and vegetative detritus are not in the inventory, but contribute less than 3% of the TC at each site. After the model results are adjusted for source-specific errors using the organic-marker measurements, an average of 1.6 μgC m^{-3} remain unexplained. This corresponds to 26–38% of ambient TC concentrations at urban sites and up to 56% at rural sites. The most likely sources of unexplained carbon are discussed.

Introduction

In the southeastern United States, carbon is the largest component of fine particulate matter (PM₂.₅) on an annual basis within urban areas (*1*) and the second largest contributor after sulfate to visibility degradation within Class I areas (*2*). During the summer months, average carbon concentrations in the Southeast are greater than those in any other U.S. region (*3*). However, these summertime carbon concentrations are systematically underestimated by current air quality models (*3–6*). Diagnosing the cause of this bias has proven difficult because particulate carbon originates from numerous primary sources and secondary formation pathways. Morris et al. (*7*) show that the underestimations could be explained by biogenic secondary formation pathways and polymerization processes that are not treated in most models. Zhang et al. (*8*) show that the model underestimations would be mitigated if primary organic aerosol emissions are increased. In both of the aforementioned studies, improvements in model performance are judged by comparison with PM₂.₅ measurements from routine monitoring networks that distinguish organic carbon (OC) and elemental carbon (EC) only. That distinction is insufficient to determine the relative contributions of different sources and formation pathways to the atmospheric loading of total carbon (TC), and hence, does not build confidence in the use of air quality models for regulatory applications in which an accurate understanding of source contributions is paramount. In contrast, organic speciation techniques provide tremendous insight into the origin of particulate carbon (*9–13*).

To date, air quality models have been evaluated directly against organic markers only in the Los Angeles metropolitan area (*14, 15*). Three obstacles have prohibited such evaluations over broader regions. First, ambient measurements of individual organic species are available at very few locations. Second, a detailed inventory of TC emissions from the most important sources is needed. Third, a regional-scale air quality model must be developed to simulate atmospheric concentrations of the organic markers. Recently, Zheng and co-workers compiled the first set of PM₂.₅ organic-marker measurements that span a multi-state geographic region (*13*). The U.S. Environmental Protection Agency assembled a detailed national emission inventory (NEI) for 1999 (*16*), the same year when the ambient organic-marker measurements were collected. The Community Multiscale Air Quality (CMAQ) model was instrumented to track primary OC and EC contributions from different source categories simultaneously, permitting the computation of speciated organic concentrations (*17*). In the present study, the CMAQ model results are evaluated against organic-marker measurements collected across the southeastern U.S. during summer 1999, and the source categories that contribute to model underestimations of TC are identified quantitatively.

Methodology

Ambient Measurements. A thorough description of the organic-marker measurements is provided by Zheng et al. (*13*). The subset of measurements used in the present study is described briefly here. PM₂.₅ samples were collected daily at eight sites in the Southeastern Aerosol Research and Characterization (SEARCH) network in July 1999 and during the Atlanta Supersite Experiment in August 1999. The SEARCH sites are located at one urban location and one suburban or rural location in four southeastern states: Mississippi, Alabama, Georgia, and Florida. The 24-h filter samples were analyzed for OC and EC using the thermal-optical reflectance (TOR) method, and monthly composites of the samples from each site were analyzed for 100 organic markers by gas chromatography/mass spectrometry (GC/MS). These markers were selected based on numerous earlier analyses of the major sources of ambient particulate carbon (*10, 12, 14, 18*). For example, hopanes and steranes are motor-vehicle-exhaust tracers, levoglucosan and resin acids are biomass-combustion tracers, nonanal and the selected alkenoic acids are food-cooking tracers, and the selected isoalkanes are cigarette smoke tracers. The average analytical uncertainty of these marker concentrations is approximately ±20% (*13*).

* Corresponding author phone: +1(919)541–2194; fax: +1(919)-541–1379; e-mail: prakash.bhave@noaa.gov.
† National Oceanic and Atmospheric Administration.
‡ Georgia Institute of Technology.

Model Description. In the present study, CMAQ v4.4 is used with the following options: Statewide Air Pollution Research Center (SAPRC99) gas-phase chemical mechanism, modal aerosol module v3 (AE3), Regional Acid Deposition Model (RADM) cloud module, eddy vertical diffusion, and piecewise parabolic method (PPM) advection. Emission inputs are described in detail in the Supporting Information. The aerosol module is described in detail by Binkowski and Roselle (*19*), with CMAQ v4.4 updates described by Bhave et al. (*20*) and references therein. The treatment of secondary organic carbon (OC_{sec}) in CMAQ v4.4 includes formation from aromatic and aliphatic hydrocarbons (i.e., anthropogenic OC_{sec}) and monoterpenes (i.e., biogenic OC_{sec}), as described by Yu et al. (*21*). The aerosol size distribution is modeled as a superposition of three lognormal modes that correspond nominally to diameter (D_p) < 0.1 μm (i.e., Aitken mode), 0.1 < D_p < 2.5 μm (i.e., accumulation mode), and D_p > 2.5 μm (i.e., coarse mode) size ranges. All model results discussed below are sums of the first two modes.

To compute the concentrations of organic markers, CMAQ v4.4 is instrumented to track primary OC and EC contributions from 15 different source categories: onroad and nonroad diesel and gasoline vehicles, anthropogenic biomass combustion, wild fires, coal, oil, and natural gas combustion, food cooking, paved road dust, other crustal material, aircraft exhaust, waste combustion, and miscellaneous industrial processes. In the base model, primary TC is tracked as four model species: Aitken-mode OC, Aitken-mode EC, accumulation-mode OC, and accumulation-mode EC. In the instrumented model, a copy of these four species is added for each of the 15 source categories. The new model species are internally mixed within their designated aerosol mode (i.e., Aitken or accumulation) and participate in advection, diffusion, deposition, condensational growth, and coagulation processes in a manner identical to the treatment of primary OC and EC species in the base model. A 10-day test simulation was conducted to verify that results from the instrumented model are practically identical to those from CMAQ v4.4. The carbon apportionment code is a publicly available model option in CMAQ v4.5−4.6.

The model described above was used to simulate gaseous and PM pollutant concentrations from June 12 to August 31, 1999. The first 3 days of model results are discarded to reduce the effects of clean initial conditions. The model domain covers the continental United States with a horizontal mesh of 32-km cells depicted in Figure 1b of reference *3*. Although it is possible to conduct the simulations with finer horizontal resolution, previous studies show that CMAQ model results for TC in the Southeast are not improved noticeably with 12-km or 8-km grid spacing (*5*, *6*). Meteorological inputs are identical to those described by Arnold and Dennis (*22*). Hourly CMAQ model results from the lowest vertical layer (height ≈ 38 m) in grid cells containing each monitoring location are aggregated to monthly averages for the present analyses. Model results are not spatially interpolated. The CMAQ outputs of TC from each source category are multiplied by organic speciation profiles (see the Supporting Information) to obtain modeled concentrations of each measured molecular marker.

Results

A comparison of model results with ambient TC concentrations reveals a model underestimation of 30% or more at all sites except Atlanta (see Table 1). This model bias is consistent with other recent model evaluations in the southeastern U.S. during summer (*3−6*) and is discussed in the Supporting Information. The following sections focus on comparisons of model results with organic markers.

Source Apportioned CMAQ Results. CMAQ model results for TC from each source category are summarized in the upper portion of Table 1. For the present discussion, sources with monthly averaged concentrations exceeding 0.1 μgC m^{-3} are regarded as major contributors of TC. Based on the CMAQ results, biogenic OC_{sec} and anthropogenic biomass combustion are major contributors at all sites. Onroad and nonroad diesel are major sources at all sites except in rural Alabama and Mississippi. Wild fires are major contributors at all sites except Centreville and both Mississippi sites. The very low wild-fire contribution at Centreville relative to nearby Birmingham is indicative of deficiencies in the spatial allocation of wild-fire emissions in the 1999 NEI (see the Supporting Information). Food cooking is a major source in Atlanta and Birmingham. Onroad gasoline, nonroad gasoline, paved road dust, and anthropogenic OC_{sec} are major sources only in Atlanta. Coal combustion is a major contributor only in Birmingham. Collectively, miscellaneous industrial sources make a major contribution in Atlanta, Birmingham, Oak Grove, and Yorkville. On the other hand, aircraft exhaust, waste combustion, oil combustion, natural-gas combustion, and crustal material make minor contributions to TC at all sites based on the monthly averaged CMAQ output. These results are generally consistent with patterns found in the regional-scale emission inventory, as discussed in the Supporting Information.

Other investigators have compared source-apportionment results from an air quality model with those from a chemical mass balance (CMB) receptor model (*23*). An analogous comparison could be made using the results in Table 1 and the CMB results in ref 13. Recent studies have shown that CMB results are very sensitive to the selection of source profiles and fitting species (*24*, *25*). A thorough analysis of the differences between CMAQ and CMB results would likely require several sensitivity tests using the CMB model, which is beyond the scope of this study. Rather than comparing Table 1 with CMB results, numerous insights are gained by comparing the model results directly with measured marker concentrations.

Model Evaluation: Organic Markers. Modeled concentrations of individual organic markers are computed by multiplying the source-specific TC concentrations in Table 1 by the published speciation profiles listed in the Supporting Information. The implicit assumption, examined below, is that all of the markers under consideration are neither degraded in the atmosphere nor formed by secondary processes (i.e., all markers are conserved tracers). In the ambient data set, concentrations of individual organic markers vary by 5 orders of magnitude (chrysene and/or triphenylene = 6 pg m^{-3} at Yorkville; dehydroabietic acid = 0.4 μg m^{-3} at Oak Grove). To evaluate model performance for all markers in a compact display, we focus on modeled/observed ratios. Assuming each species is a conserved tracer, instances where a ratio departs from unity imply errors in the emissions or transport of the given marker.

Zheng et al. (*13*) quantified 100 different organic markers in at least one of the nine summertime ambient samples. The R and S chiral forms of homohopanes, bishomohopanes, and trishomohopanes were quantified separately in the ambient samples, but each pair is summed together in the model evaluation because the chiral structures are not separately distinguished in many of the source profiles used to speciate the model results. This leaves 97 compounds and a total of 613 pairs of measured and modeled marker concentrations displayed in Figure 1. Analyzing the full suite of modeled/observed ratios yields insights about the TC concentrations at given locations, the validity of the conserved-tracer assumption for certain markers, and the accuracy of emission estimates for specific source categories. Major insights in these three areas are summarized below.

Site-Specific Comparisons. Anomalies in model performance which are unique to a monitoring site can be identified

TABLE 1. Monthly Averaged, Source-Apportioned Concentrations of TC [μgC m^{-3}] at all SEARCH Sites in July 1999 and Atlanta in August 1999[a]

	GFP	OAK	BHM	CTR	JST	JST (Aug)	YRK	PNS	OLF
CMAQ Model Results									
onroad diesel	0.11	0.04	0.25	0.04	0.58	0.64	0.16	0.10	0.10
nonroad diesel	0.18	0.07	0.30	0.06	0.80	0.92	0.20	0.16	0.16
onroad gasoline	0.03	0.01	0.08	0.01	0.24	0.26	0.05	0.04	0.04
nonroad gasoline	0.06	0.02	0.05	0.01	0.17	0.20	0.03	0.05	0.05
aircraft exhaust	0	0	0.01	0	0.01	0.01	0	0.01	0.01
anth. biomass comb.	0.22	0.21	0.47	0.16	0.68	0.70	0.45	0.36	0.36
wild fires	0.04	0.04	0.38	0.07	0.14	0.16	0.22	0.44	0.44
waste combustion	0	0	0.01	0	0.01	0.02	0.02	0	0
coal combustion	0.01	0.01	0.17	0.04	0.04	0.07	0.05	0.03	0.03
oil combustion	0.01	0	0.01	0	0.04	0.03	0.04	0.01	0.01
natural gas combustion	0.01	0.01	0.01	0	0.03	0.03	0.01	0	0
food cooking	0.05	0.02	0.13	0.01	0.46	0.52	0.07	0.06	0.06
paved road dust	0.04	0.01	0.06	0.01	0.11	0.12	0.03	0.04	0.04
crustal material	0.03	0.01	0.03	0.01	0.05	0.06	0.04	0.03	0.03
misc. industrial sources	0.04	0.15	0.48	0.04	0.61	0.65	0.24	0.05	0.05
other primary sources	0.01	0.01	0.01	0.01	0	0	0	0.05	0.05
anthropogenic OCsec	0.02	0.02	0.05	0.02	0.10	0.10	0.06	0.03	0.03
biogenic OCsec	0.43	0.59	1.48	0.81	1.69	2.04	1.30	0.43	0.43
total carbon (CMAQ)	1.30	1.23	3.97	1.33	5.77	6.56	2.96	1.88	1.88
Observational Estimates of Model Bias (Negative Values Indicate Model Overestimates)									
gasoline + diesel	b	c	0.57	c	−1.44	−2.01	−0.28[c,d]	b	−0.25
biomass combustion	0.49	1.36	b	0.18	b	b	b	−0.37	−0.55
food cooking	0.11	e	0.43	f	f	0.48	0.13	0.29	0.07
cigarette smoke	0.02	g	0.08	g	0.05	0.13	g	0.05	g
vegetative detritus	0.02	0.03	0.10	0.02	0.07	0.17	0.02	0.06	0.02
total carbon (CMAQ + biases)[h]	1.93	2.62	5.14	1.53	4.45	5.34	2.83	1.91	1.16
unexplained carbon	1.09	0.43	1.99	1.91	1.55	2.00	1.39	1.16	1.48
total carbon (observed)	3.03	3.05	7.13	3.44	5.99	7.33	4.22	3.07	2.64

[a] GFP = Gulfport, MS; OAK = Oak Grove, MS; BHM = N. Birmingham, AL; CTR = Centreville, AL; JST = Atlanta, GA; YRK = Yorkville, GA; PNS = Pensacola, FL; OLF = Outlying Field #8, FL. Upper section: modeled source contributions from 16 primary and 2 secondary categories. Lower section: model biases estimated from organic-marker evaluation. Bottom row: TC observations. Note that both Florida sites lie in the same 32-km grid cell. [b] Model results and observations agree within measurement uncertainty. [c] Hopanes and steranes not quantified in these samples. [d] Based on seven species for which modeled/min(observed) ratio exceeds unity. [e] Inconclusive, because the biomass deficit may contribute to nonanal underestimates. [f] Nonanal not quantified in these samples. [g] Cigarette-smoke tracers not quantified in these samples. [h] CMAQ results after adjusting for source-specific biases.

by searching Figure 1 for patterns where a particular symbol falls consistently above or below the remaining symbols for a set of markers. At the Oak Grove site, modeled/observed ratios are substantially lower than the ratios at all other sites for alkanoic acids with twenty or more carbon atoms (Figure 1b), retene (Figure 1d), resin acids (Figure 1e), cycloalkanes, nonanal, and levoglucosan (Figure 1f). This implies that the model is missing or underestimating a source near Oak Grove. Biomass-combustion smoke contains large quantities of alkanoic acids, retene, resin acids, and levoglucosan (26). Figure S2 indicates that episodic TC events on June 19, July 1, and July 3, which could result from local fires, are not captured in the model. The Oak Grove site is located 10 km south of Camp Shelby, a military installation encompassing over 500 km² of forested land. Prescribed burning is used sporadically in the Southeast for land management and is most intensive on military installations (24). In the 1999 NEI, prescribed-burning emissions are distributed with monthly resolution on a state-by-state basis. This allocation of emissions is not sufficiently detailed to capture the impact of prescribed fires near their source. Using an organic speciation profile obtained from prescribed-burning sites at two military installations in the Southeast (24) and the factor of 6.4 difference between modeled and measured levoglucosan concentrations (Figure 1f), we conclude that TC from biomass combustion is underestimated at Oak Grove by approximately 1.4 μgC m^{-3}.

At the Birmingham site, modeled/observed ratios are substantially lower than the ratios at all other sites for hopanes and steranes (Figure 1c), polycyclic aromatic hydrocarbons (PAH) (Figure 1d), and two aromatic ketones (Figure 1f).

Again, this implies that the model is missing emission sources near Birmingham and three explanations are provided. First, modeled concentrations for nine hopanes and steranes are lower than the observations by a factor of 1.86 on average (computed as the geometric mean of the modeled/observed ratios), implying that TC from vehicle exhaust in Birmingham is underestimated by up to 0.6 μgC m^{-3}. Diesel exhaust is also a significant source of PAH, which may explain a small portion of the PAH underestimation. Second, the underestimation of PAH may be due to incorrect speciation of coal-combustion TC, which is greater at Birmingham than at any other site (see Table 1). Due to the unavailability of coal combustion data, a Pittsburgh coke-production profile is used in the present study to speciate coal-combustion TC. Third, it is possible that PAH are emitted in large quantities from miscellaneous industrial sources near the Birmingham site (note the substantial contribution from this category in Table 1), which are predominantly iron and steel industries. These source contributions are not speciated in the model results because organic-marker profiles are unavailable.

Conserved-Tracer Assumption. An examination of Figure 1 reveals three distinct patterns that challenge the conserved-tracer assumption for certain markers. In the first pattern, modeled/observed ratios lie significantly above unity, indicative of species that are subject to volatilization or photochemical decay. The clearest example of this pattern is in the PAH results (Figure 1d), where several of the modeled/observed ratios approach 100. Atmospheric studies in Los Angeles in 1982 suggest that PAH with molar mass less than or equal to 252 g/mol (from fluoranthene to perylene in Figure 1d) are depleted by volatilization and/or chemical
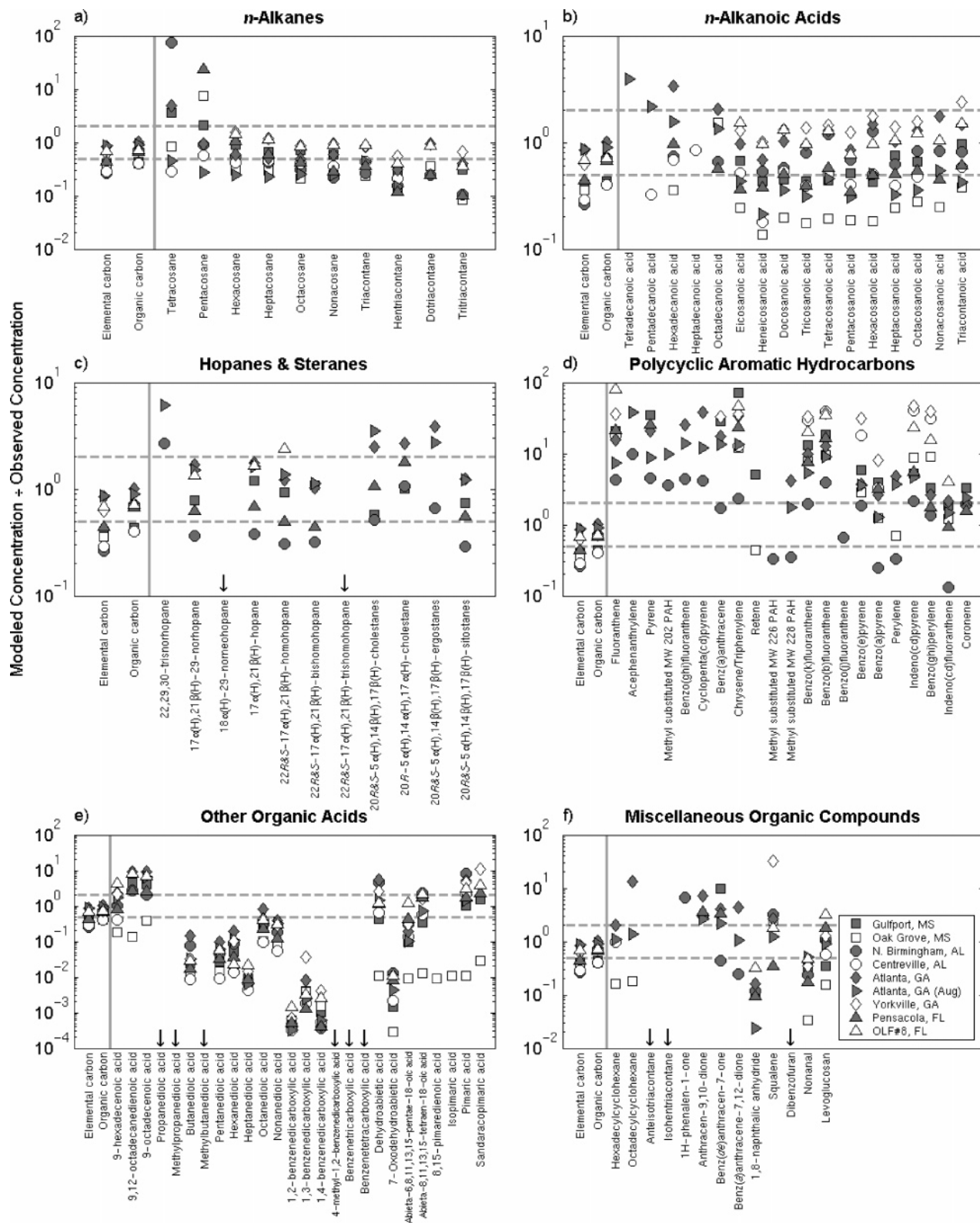
FIGURE 1. Evaluation of model results for individual organic markers (EC and OC displayed for reference). Each symbol represents a monthly averaged model result divided by an ambient measurement. Shaded symbols represent urban sites; unshaded symbols represent rural and suburban sites. Different shapes correspond to different states. Omitted symbols imply that the given compound was not quantified in the given ambient sample (e.g., benzo(a)pyrene in OLF#8). Downward-pointing arrows identify markers that were measured in ambient samples but not in the speciation profiles and thus have zero concentrations in the CMAQ results. Dashed horizontal lines bound all comparisons that agree within a factor of 2.

reaction, whereas the heaviest PAH are conserved in the particle phase (14). A 1993 field study during an intense photochemical episode in Los Angeles indicates that indeno-(cd)pyrene and indeno(cd)fluoranthene are also degraded

by atmospheric reactions (15). Laboratory experiments show that nine of the PAH plotted in Figure 1d, including indeno-(cd)pyrene and benzo(ghi)perylene, decay photochemically under typical atmospheric conditions (27). Present results
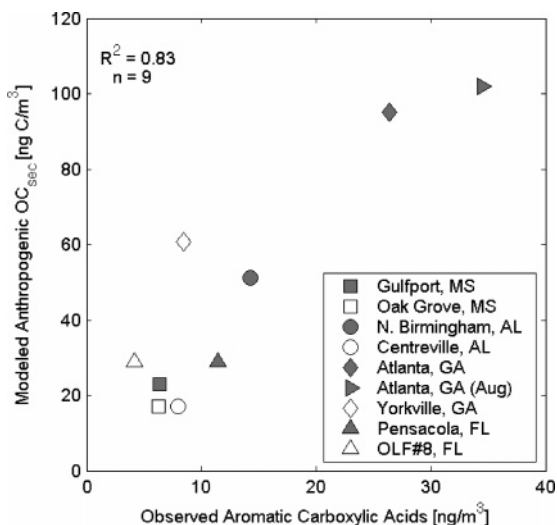
FIGURE 2. Model calculations of anthropogenic secondary organic carbon compared with the summed measurements of six aromatic carboxylic acids at all SEARCH sites in July 1999 and Atlanta in August 1999.

are in agreement with the earlier studies and indicate further that coronene, the heaviest PAH analyzed in the ambient samples (300 g/mol), may be subject to atmospheric degradation (geometric mean modeled/observed ratio equals 2.1). Another example of the pattern representative of volatility and atmospheric degradation is seen in the tetracosane and pentacosane results (Figure 1a). For both of these species, model results greatly exceed observations at three sites. Early source-reconciliation studies in Los Angeles indicate that these species are conserved in the particle phase (*10*), but Figure 1 of reference *28* demonstrates that these compounds are, in fact, semi-volatile, which is supported by the present results.

In the second pattern, modeled/observed ratios fall significantly below unity, suggesting that these species are formed in the atmosphere by secondary processes. Markers fitting this pattern include aliphatic dicarboxylic acids, 7-oxodehydroabietic acid, nonanal, aromatic carboxylic acids, and 1,8-napthalic anhydride, as shown in Figure 1e and f. Aliphatic dicarboxylic acids are known to be formed by secondary photochemical processes (see ref *10* and references therein). Laboratory studies show that 7-oxodehydroabietic acid is a photochemical degradation product of dehydroabietic acid (*29*), whereas nonanal is formed by the oxidation of 9-octadecenoic acid (*30*). Aromatic carboxylic acids are believed to be produced by secondary processes (*13*, *18*) following the oxidation of aromatic gaseous precursors. In the CMAQ model, aromatic OC_{sec} concentrations are simulated explicitly and, as shown in Figure 2, are strongly correlated with the summed measurements of six aromatic carboxylic acids. This correlation indicates that the spatial distribution of aromatic OC_{sec} across the SEARCH sites during summer is captured quite well in the CMAQ model. However, no conclusions can be drawn regarding the magnitude of the modeled OC_{sec} concentrations because the marker/OC_{sec} ratios are unknown.

The third pattern is illustrated by unshaded symbols lying above shaded symbols of the same shape. Markers exhibiting this pattern include heavy alkanoic acids at the Georgia and Florida sites (Figure 1b) and several heavy PAH (Figure 1d). On face, this pattern indicates that the CMAQ results overestimate ambient concentrations at rural sites to a greater degree than at urban sites. An alternative interpretation is that markers exhibiting this pattern are emitted in urban areas and are degraded by chemical reactions during

transport to rural areas. This was proposed recently by Robinson et al. to explain hopane concentrations in the Pittsburgh area (*31*).

Source-Specific Markers. Perhaps the most valuable information that can be obtained from Figure 1 is a quantitative assessment of the emission inventory for specific sources of TC. Such an assessment involves evaluating the CMAQ model results for organic species that are source-specific markers. The results of this assessment are summarized in the lower portion of Table 1. Modeled concentrations of hopanes and steranes, markers for gasoline- and diesel-vehicle emissions, are evaluated in Figure 1c. Model results fall within 8% of the measurements when averaged over all 42 modeled/observed pairs, implying that the regional emission inventory is unbiased for the sum of diesel and gasoline sources. Looking at each site individually, model results exhibit varying levels of agreement with the observations. Average modeled/observed ratios are 0.54 in Birmingham (discussed above), 1.76 in OLF#8, 1.82 in Atlanta during July, and 2.02 in Atlanta during August. Using the average modeled/observed ratios at each site, the amount of excess vehicle exhaust in the Atlanta and OLF#8 model results may be estimated (see Table 1).

Levoglucosan is commonly used as a marker for biomass combustion due to its source-specificity and high ambient concentrations (*32*). Excluding the Oak Grove sample for reasons discussed above, modeled and observed levoglucosan concentrations agree within 5% on average, indicating that the regional emission inventory is unbiased for biomass combustion. Looking at site-to-site variability, modeled/observed ratios are 0.34 in Gulfport, 0.57 in Centreville, 1.84 in Pensacola, and 3.22 in OLF#8. The modeled bias in TC concentrations from biomass combustion may be estimated from these ratios (see Table 1). Overestimations of biomass-combustion TC at the Florida sites may be due to poor spatial allocation of wild-fire emissions, as indicated in Figure S3.

At all sites except Oak Grove, model results exceed observations for two alkenoic acids—9,12-octadecanedioic acid and 9-octadecenoic acid—commonly used as food cooking markers, possibly due to photochemical oxidation that is not accounted for in the model (*31*). Cholesterol was not quantified in the 1999 SEARCH samples due to analytical limitations. Nonanal has been proposed as a food-cooking marker (*10*, *12*), but can be formed by secondary reactions as noted above and its precursor has been measured at prescribed-burning sites (*24*). Excluding the Oak Grove site, nonanal is underestimated by a factor of 3 on average. If we assume the nonanal underestimations are entirely due to missing food-cooking emissions, an upper estimate of the modeled deficits in food-cooking TC may be computed (see Table 1). The deficits estimated in this manner range from 3 to 10% of the observed TC at each site, so it may be a significant bias to explore in the future.

Although cigarette smoke is not included in the 1999 NEI, an upper estimate of its contributions can be obtained by dividing ambient measurements of anteisotriacontane and isohentriacontane by the respective marker/TC ratios from a published organic speciation profile (*33*). Both markers were quantified in the Birmingham and Pensacola samples, so cigarette-smoke contributions estimated from those measurements are averaged together. In the Gulfport and Atlanta samples, cigarette-smoke contributions are estimated from a single marker. Cigarette-smoke contributions estimated in this manner constitute less than 2% of the observed TC concentration at each site (see Table 1).

Among the normal alkanes shown in Figure 1a, hentriacontane and tritriacontane have the lowest modeled/observed ratios (geometric mean ratio = 0.2 for both species). These are often used as markers for vegetative detritus (*34*), which is another source of carbon that is missing from the

1999 NEI. Following the procedure of Hildemann et al. (*35*), contributions of vegetative detritus to ambient TC are estimated by assuming that the modeled deficit in both alkane species is entirely due to vegetative detritus. Results suggest that vegetative detritus contributes 2% or less of the observed TC at each site (see Table 1).

Aromatic ketones have been proposed as markers for natural gas combustion (*10*). Considering the 13 available measurements of aromatic ketones, modeled concentrations exceed the observations by a factor of 2.6 on average (see Figure 1f). This confirms that natural gas combustion makes a negligible contribution to ambient TC (see Table 1) and qualitatively verifies the recent correction to natural-gas-combustion $PM_{2.5}$ emission factors in the NEI (see the Supporting Information).

Discussion

When the CMAQ model results are supplemented with organic-marker based observational estimates that add missing vehicle exhaust, biomass smoke, food-cooking aerosol, cigarette smoke, and vegetative detritus, model results are brought closer to ambient TC concentrations at the Mississippi and Alabama sites (see Table 1). However, model results move further from the observed TC concentrations in all three Georgia samples and the OLF#8 sample when they are adjusted for source-specific biases. The departure of model results from observations in Georgia occurs because motor-vehicle contributions are overestimated at those sites in the CMAQ results. Thus, the good model performance at the Atlanta site is quite likely due to compensating errors: overestimations of vehicle emissions and underestimations of other sources.

After the CMAQ model results are adjusted for source-specific biases using the best available ambient measurements and speciation profiles, they fall below observed TC concentrations at all sites across the Southeast. Excluding the Oak Grove site where results are somewhat inconclusive, the average TC deficit is 1.6 μgC m^{-3}. The unexplained concentrations correspond to 56% of the observed TC at Centreville and OLF#8, and 26−38% at the other sites. Recent studies hypothesize that much of the unexplained carbon in the Southeast during summer may be biogenic OC_{sec} (*7*). This hypothesis cannot be tested with the present data set, but it will be explored in future studies using measurements of ^{14}C and organic markers for OC_{sec}. A number of primary sources which are not discussed above may also help explain the missing carbon. First, due to the lack of a source-specific profile, $PM_{2.5}$ emissions from open burning of household waste (SCC 2610030000) have been speciated in the 1999 NEI using a municipal-incinerator profile that fractionates only 4.1% of the total $PM_{2.5}$ emissions as carbon. Changing the speciation profile for this SCC to a foliage-burning profile (64−80% carbon by mass), which may be more representative than the current profile assignment based on the combustion characteristics of an open burn, would enhance the modeled TC concentrations in Atlanta (0.16 μgC m^{-3} in July, 0.20 μgC m^{-3} in August), Birmingham (0.09 μgC m^{-3}), and Yorkville (0.21 μgC m^{-3}). Second, it is shown in Table 1 that aircraft exhaust makes an infinitesimal contribution to modeled TC concentrations at all sites. In a recent study, aircraft exhaust was estimated to contribute 1 μgC m^{-3} around Atlanta (*36*). Our model results exceed 0.1 μgC m^{-3} in many grid cells over California (see Figure S4e). It is possible that the aircraft emission inventory submitted by California for the 1999 NEI is based on a more robust methodology than the default NEI methodology applied over the southeastern states, but this requires further exploration. Third, it has been proposed that pollen fragments and fungal spores, which are not in the NEI, may be an important source of TC in fine particles. Glycerophospholipid measurements taken in Toronto from

May−June 2000 indicate that these biological sources contribute between 12−22% of the OC in $PM_{2.5}$ (*37*). Those results are not directly applicable to the present study due to the differing location and season, but measurements of phospholipid species should be considered in future studies in the Southeast.

The results of this study point to several areas of future work that will help improve model performance for TC during summer months in the Southeast. Assuming biogenic OC_{sec} is a major component of the unexplained carbon, the CMAQ model should be updated to include OC_{sec} formation from isoprene, sesquiterpenes, and acid-catalyzed reactions (*7, 38*). Emission inventories for the Southeast should be improved with a specific focus on TC from aircraft exhaust, household waste burning, and possibly food cooking. Additional source characterization studies are needed to obtain organic speciation profiles for nonroad vehicles, coal-fired power plants, iron and steel industries, and household waste burning in the Southeast. Advancements in receptor modeling are needed to evaluate CMAQ model calculations for the split between gasoline and diesel exhaust and between onroad and nonroad sources. Ultimately, air quality model results should be reconciled against receptor-based models to build confidence in the source-apportionment results.

Supporting Information Available

Details about the emission inputs, organic speciation profiles, model evaluation against daily TC data, and spatial maps of the source-apportioned CMAQ results. This material is available free of charge via the Internet at http://pubs.acs.org.

Literature Cited

(1) Hansen, D. A.; Edgerton, E. S.; Hartsell, B. E.; Jansen, J. J.; Kandasamy, N.; Hidy, G. M.; Blanchard, C. L. The southeastern aerosol research and characterization study: Part 1-overview. *J. Air Waste Manage. Assoc.* **2003**, *53*, 1460−1471.

(2) Brewer, P. F.; Adlhoch, J. P. Trends in speciated fine particulate matter and visibility across monitoring networks in the southeastern United States. *J. Air Waste Manage. Assoc.* **2005**, *55*, 1663−1674.

(3) Yu, S.; Dennis, R. L.; Bhave, P. V.; Eder, B. K. Primary and secondary organic aerosols over the United States: estimates on the basis of observed organic carbon (OC) and elemental carbon (EC), and air quality modeled primary OC/EC ratios. *Atmos. Environ.* **2004**, *38*, 5257−5268.

(4) Morris, R. E.; McNally, D. E.; Tesche, T. W.; Tonnesen, G.; Boylan, J. W.; Brewer, P. Preliminary evaluation of the Community Multiscale Air Quality model for 2002 over the southeastern United States. *J. Air Waste Manage. Assoc.* **2005**, *55*, 1694−1708.

(5) Tesche, T. W.; Morris, R.; Tonnesen, G.; McNally, D.; Boylan, J.; Brewer, P. CMAQ/CAMx annual 2002 performance evaluation over the eastern US. *Atmos. Environ.* **2006**, *40*, 4906−4919.

(6) Zhang, Y.; Liu, P.; Queen, A.; Misenis, C.; Pun, B.; Seigneur, C.; Wu, S.-Y. A comprehensive performance evaluation of MM5-CMAQ for the summer 1999 Southern Oxidants Study episode—part II: Gas and aerosol predictions. *Atmos. Environ.* **2006**, *40*, 4839−4855.

(7) Morris, R. E.; Koo, B.; Guenther, A.; Yarwood, G.; McNally, D.; Tesche, T. W.; Tonnesen, G.; Boylan, J.; Brewer, P. Model sensitivity evaluation for organic carbon using two multi-pollutant air quality models that simulate regional haze in the southeastern United States. *Atmos. Environ.* **2006**, *40*, 4960−4972.

(8) Zhang, Y.; Liu, P.; Pun, B.; Seigneur, C. A comprehensive performance evaluation of MM5-CMAQ for the summer 1999 Southern Oxidants Study episode—part III: Diagnostic and mechanistic evaluations. *Atmos. Environ.* **2006**, *40*, 4856−4873.

(9) Rogge, W. F.; Mazurek, M. A.; Hildemann, L. M.; Cass, G. R. Quantification of urban organic aerosols at a molecular level: identification, abundance and seasonal variation. *Atmos. Environ.* **1993**, *27A*, 1309−1330.

(10) Schauer, J. J.; Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R. Source apportionment of airborne particulate matter using organic compounds as tracers. *Atmos. Environ.* **1996**, *30*, 3837−3855.

(11) Watson, J. G.; Fujita, E.; Chow, J. C.; Zielenska, B.; Richards, L. W.; Neff, W.; Dietrich, D. *Northern Front Range Air Quality Study final report and appendices*; 6580−685−8750.1F2; Desert Research Institute: Reno, NV, 1998.

(12) Schauer, J. J.; Cass, G. R. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. *Environ. Sci. Technol.* **2000**, *34*, 1821−1832.

(13) Zheng, M.; Cass, G. R.; Schauer, J. J.; Edgerton, E. S. Source apportionment of PM2.5 in the southeastern United States using solvent-extractable organic compounds as tracers. *Environ. Sci. Technol.* **2002**, *36*, 2361−2371.

(14) Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R.; Simoneit, B. R. T. Mathematical modeling of atmospheric fine particle-associated primary organic compound concentrations. *J. Geophys. Res.* **1996**, *101*, 19379−19394.

(15) Fraser, M. P.; Kleeman, M. J.; Schauer, J. J.; Cass, G. R. Modeling the atmospheric concentrations of individual gas-phase and particle-phase organic compounds. *Environ. Sci. Technol.* **2000**, *34*, 1302−1312.

(16) *Procedures Document for National Emission Inventory, Criteria Air Pollutants 1985−1999*; EPA-454/R-01-006; U.S. Environmental Protection Agency: Washington, DC, 2001.

(17) Bhave, P. V.; Pouliot, G. A.; Zheng, M. *Source Apportionment of Primary Carbonaceous Aerosol Using the Community Multiscale Air Quality Model.* In 27th NATO/CCMS International Technical Meeting on Air Pollution Modeling and Its Applications, Banff, Canada, October 24−29, 2004.

(18) Schauer, J. J.; Fraser, M. P.; Cass, G. R.; Simoneit, B. R. T. Source reconciliation of atmospheric gas-phase and particle-phase pollutants during a severe photochemical smog episode. *Environ. Sci. Technol.* **2002**, *36*, 3806−3814.

(19) Binkowski, F. S.; Roselle, S. J. Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component—1. Model description. *J. Geophys. Res.* **2003**, *108*, 4183.

(20) Bhave, P. V.; Roselle, S. J.; Binkowski, F. S.; Nolte, C. G.; Yu, S.; Gipson, G. L.; Schere, K. L. *CMAQ Aerosol Module Development: Recent Enhancements and Future Plans.* In CMAS Models-3 Users' Conference, Chapel Hill, NC, October 18−20, 2004.

(21) Yu, S.; Bhave, P. V.; Dennis, R. L.; Mathur, R. Seasonal and regional variations of primary and secondary organic aerosols over the continental United States: semi-empirical estimates and model evaluation. *Environ. Sci. Technol.* in review.

(22) Arnold, J. R.; Dennis, R. L. Testing CMAQ chemistry sensitivities in base case and emissions control runs at SEARCH and SOS99

(23) Held, T.; Ying, Q.; Kleeman, M. J.; Schauer, J. J.; Fraser, M. P. A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter. *Atmos. Environ.* **2005**, *39*, 2281−2297.

(24) Lee, S.; Baumann, K.; Schauer, J. J.; Sheesley, R. J.; Naeher, L. P.; Meinardi, S.; Blake, D. R.; Edgerton, E. S.; Russell, A. G.; Clements, M. Gaseous and particulate emissions from prescribed burning in Georgia. *Environ. Sci. Technol.* **2005**, *39*, 9049−9056.

(25) Robinson, A. L.; Subramanian, R.; Donahue, N. M.; Bernardo-Bricker, A.; Rogge, W. F. Source apportionment of molecular markers and organic aerosol. 3. Food cooking emissions. *Environ. Sci. Technol.* **2006**, *40*, 7820−7827.

(26) Simoneit, B. R. T. Biomass burning - A review of organic tracers for smoke from incomplete combustion. *Appl. Geochem.* **2002**, *17*, 129−162.

(27) Kamens, R. M.; Guo, Z.; Fulcher, J. N.; Bell, D. A. Influence of humidity, sunlight, and temperature on the daytime decay of polyaromatic hydrocarbons on atmospheric soot particles. *Environ. Sci. Technol.* **1988**, *22*, 103−108.

(28) Fraser, M. P.; Cass, G. R.; Simoneit, B. R. T.; Rasmussen, R. A. Air quality model evaluation data for organics .4. C-2-C-36 non-aromatic hydrocarbons. *Environ. Sci. Technol.* **1997**, *31*, 2356−2367.

(29) Corin, N. S.; Backlund, P. H.; Kulovaara, M. A. M. Photolysis of the resin acid dehydroabietic acid in water. *Environ. Sci. Technol.* **2000**, *34*, 2231−2236.

(30) Kieber, R. J.; Hydro, L. H.; Seaton, P. J. Photooxidation of triglycerides and fatty acids in seawater: Implication toward the formation of marine humic substances. *Limnol. Oceanogr.* **1997**, *42*, 1454−1462.

(31) Robinson, A. L.; Donahue, N. M.; Rogge, W. F. Photochemical oxidation and changes in molecular composition of organic aerosol in the regional context. *J. Geophys. Res.* **2006**, *111*, D0332.

(32) Simoneit, B. R. T.; Schauer, J. J.; Nolte, C. G.; Oros, D. R.; Elias, V. O.; Fraser, M. P.; Rogge, W. F.; Cass, G. R. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. *Atmos. Environ.* **1999**, *33*, 173−182.

(33) Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R. Sources of fine organic aerosol. 6. Cigarette smoke in the urban atmosphere. *Environ. Sci. Technol.* **1994**, *28*, 1375−1388.

(34) Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R.; Simoneit, B. R. T. Sources of fine organic aerosol. 4. Particulate abrasion products from leaf surfaces of urban plants. *Environ. Sci. Technol.* **1993**, *27*, 2700−2711.

(35) Hildemann, L. M.; Rogge, W. F.; Cass, G. R.; Mazurek, M. A.; Simoneit, B. R. T. Contribution of primary aerosol emissions from vegetation-derived sources to fine particle concentrations in Los Angeles. *J. Geophys. Res.* **1996**, *101*, 19541−19549.

(36) Unal, A.; Hu, Y. T.; Chang, M. E.; Odman, M. T.; Russell, A. G. Airport related emissions and impacts on air quality: Application to the Atlanta International Airport. *Atmos. Environ.* **2005**, *39*, 5787−5798.

(37) Womiloju, T. O.; Miller, J. D.; Mayer, P. M.; Brook, J. R. Methods to determine the biological composition of particulate matter collected from outdoor air. *Atmos. Environ.* **2003**, *37*, 4335−4344.

(38) Jang, M. S.; Czoschke, N. M.; Lee, S.; Kamens, R. M. Heterogeneous atmospheric aerosol production by acid-catalyzed particle-phase reactions. *Science* **2002**, *298*, 814−817.