**Dynamic and chemical controls on new particle formation occurrence and characteristics
from in situ and satellite-based measurements**

Sullivan, R. C.* and Pryor, S. C.^a

*^a Department of Earth and Atmospheric Sciences, Cornell University

306 Tower Rd, Ithaca, NY 14853, United States of America

* Corresponding author: rcs365@cornell.edu

^a email: sp2279@cornell.edu

9   **Abstract**

10   We analyze the association between satellite-based measurements of chemical conditions
11 (sulfur dioxide (SO₂), nitrogen dioxide (NO₂), and formaldehyde (HCHO) concentrations),
12 insolation (UV), and aerosol particle properties (aerosol optical depth (AOD) and Ångström
13 exponent (AE)); and the occurrence of new particle formation (NPF), formation rates (J₆),
14 growth rates (GR), and survival probabilities (SP) using particle size distribution measurements
15 taken during two extended field campaigns at a forested location in southern Indiana. When
16 conditionally sampled by event occurrence and non-occurrence the satellite-derived parameters
17 exhibit significant differences and also show some degree of skill in predicting NPF though
18 logistic regression analysis. During leaf-on measurement periods, NPF occurrence exhibits
19 strong seasonality (NPF is more frequent in spring vs. summer) and is associated with a low
20 condensational sink, while leaf-off NPF occurrence is associated with high near-surface UV
21 receipt. Multiple linear regression equations of J₆, GR, and SP using the chemical conditions as
22 predictors exhibit some significant r^2 values ($p < 0.1$), but are relatively unstable and many of the
23 regression coefficients do not differ significantly from zero.

24   Key words: Nucleation; NPF intensity; growth rates; survival probabilities; remote sensing;
25   predictability

26   **1   Introduction**

27   New particle formation (NPF) events enhance aerosol particle (hereafter particles)
28 concentrations (Kulmala et al., 2004) and occur on regional scales (Crippa and Pryor, 2013).
29 They thus have the potential to impact regional and possibly global climate (Merikanto et al.,
30 2009). For newly formed particles (~1-3 nm) to become climate relevant, they must survive and
31 grow to sufficient diameters to act as cloud condensation nuclei (~100 nm). NPF events have

been estimated to contribute up to 50% of particles > 100 nm on a regional scale over northeastern North America (Pierce et al., 2014), but this contribution is likely to vary regionally (Kulmala et al., 2004; Westervelt et al., 2013) and such estimates are sensitive to the data analysis and model assumptions. The climate impact of NPF events is determined by factors such as the in situ particle population, frequency of NPF events, nucleation rate (i.e. how many new particles form, described here using the formation rate of 6-30 nm particles, J_6), growth rates (GRs), and survival probabilities (SPs, i.e. percentage of new particles that grow to climate relevant size). Herein we use statistical analyses of in situ and remote sensing data to diagnose the dependence of NPF occurrence, J_6 , GR, and SP on parameters known, or thought, to play a key role in dictating each of those variables. These conditions are not necessarily mutually inclusive, i.e. conditions conducive to NPF event occurrence may be unfavorable for high formation rates and/or GR. For example, low condensational sink (CS) was found to be conducive to NPF occurrence and high J_6 , while high CS was associated with high GR in the Yangtze River Delta in China (Qi et al., 2015). Clean air masses are also conducive to NPF in northeastern North America, but polluted air masses are associated with higher J_{10} (Pierce et al., 2014).

Specific objectives of the analyses herein are to characterize:

1. Chemical and physical regimes associated with NPF occurrence.
  2. The association between different chemical regimes and J_6 , GR, and SP of NPF events.
- We develop a framework that can be uniformly applied to all in situ observational data sets of ultrafine particle size distributions (PSDs). As in situ trace gas measurements are not always available, we draw predictor variables from satellite-borne radiometers. It is acknowledged that these observations are associated with larger uncertainty than in situ measurements, and are

typically not continuous through time. However, they provide consistent information about atmospheric composition across a range of spatial scales, and are uniformly available at least at low- and mid-latitudes. Further, they quantify columnar conditions and thus may be more strongly linked to regional-scale NPF than point measurements made at/or near the surface.

2 Methods

2.1 Data sets

PSD measurements used to describe the occurrence and characteristics of NPF events were taken during two extended field campaigns at a forested location in southern Indiana (in the Morgan Monroe State Forest, MMSF; 39.317°N, 86.417°W; Figure 1) (Pryor et al., 2014, 2010):

- Data collected during Jan. 2007 – Apr. 2009 using a Scanning Mobility Particle Sizer (SMPS comprised of an Electrostatic Classifier, nano-DMA, and Condensation Particle Counter) (Pryor et al., 2010). The measurements were taken sequentially every ~2.5 min for 10 minutes out of every half hour at 2 m, 34 m, and 46 m for a particle diameter range of 3.22 – 105.5 nm (due to low tubing penetration efficiencies, only the 81 bins > 6 nm are used here).
- Data collected during Mar. 2012 – Dec. 2013 using a TSI Fast Mobility Particle Sizer (FMPS 3091; 32 logarithmically spaced bins 6 – 520 nm) (Pryor et al., 2014). Measurements were taken sequentially for 10 minutes out of every half hour at each of three sample heights: 12 m, 20 m, and 28 m. The FMPS reports at 1 Hz, but the data have been averaged to 2 min.

Analyses presented here employ predictands derived using PSD measurements from 34 m for 2007–2009 and 28 m for 2012–2013; both heights are above the canopy, but represent near surface concentrations.

The selection of predictor variables was based on the following reasoning:

- While uncertainty remains regarding the relative importance of different nucleation

mechanisms (Boy et al., 2007) (e.g. cluster activation (Kulmala et al., 2006), kinetic (McMurry, 1980), ion mediated (Yu and Turco, 2000), binary (Jaecker-Voirol and Mirabel, 1989), or ternary nucleation (Kulmala et al., 2000)), most studies indicate a key role for sulfuric acid (H_2SO_4) with either ammonia (NH_3) or organic molecules contributing to the critical clusters (Almeida et al., 2013; O'Dowd et al., 2002; Sipilä et al., 2010). Although not uniformly observed, our further expectation is that NPF likelihood and intensity will tend to be reduced under conditions of a high CS, and that near-surface detection of NPF will be enhanced under conditions of high turbulence intensity (since there is some evidence that nucleation is concentrated aloft from the surface (Crippa et al., 2012; Pryor et al., 2011)).

- The relative contribution of different species to particle growth may vary over the course of an event. Prior research at MMSF has indicated initial particle growth is dominated by ammonium and sulfate although later growth exhibited evidence of some contribution from organics (Pryor et al., 2011). In Pittsburgh H_2SO_4 also dominated formation and initial growth, followed by growth from ammonium, and organics contributing to particle growth later in the event (Zhang et al., 2004).
- A newly formed particle can grow to climate relevant size, given there are sufficient condensable vapors present and that it does not coagulate with other particles. Thus SP may be increased in conditions favorable for high GR and reduced under high CS or in conditions favorable for high J_6 and thus greater coagulation loss.

Based on the above, statistical models of NPF occurrence and properties (predictands are NPF occurrence, J_6 , GR, and SP) are built using the following predictor variables:

- $\text{LAI} \times T_{\text{max}}$ (where T_{max} is the daily maximum hourly air temperature and LAI is the leaf area index as measured every sixth day at MMSF). LAI and T_{max} used here are from in situ

measurements, but proxies of both (e.g. skin temperature and LAI) are available from the Moderate Resolution Imaging Spectroradiometer (MODIS)) (Huete et al., 1999; Wan, 1999). LAI $\times$ T_{max} is used as a proxy for biogenic volatile organic compound (BVOC) emissions (Guenther et al., 1993), the oxidation products of which may be involved in nucleation and/or growth of freshly nucleated particles (Laaksonen et al., 2008). This variable is thus anticipated to have a positive association with NPF occurrence, J₆, GR, and SP.

- Formaldehyde (HCHO). Although satellite retrievals of HCHO are highly uncertain (Barkley et al., 2013), HCHO is one of very few products of VOC oxidation that is detected by satellite-borne radiometers, and while it is not directly relevant to NPF or ultrafine particle growth, satellite-derived HCHO concentrations are used as an index of the abundance of relatively low-volatility oxidized organics of both anthropogenic and biogenic origin (Chance et al., 2000; Henze and Seinfeld, 2006). Given contradictory evidence that high isoprene concentrations may either suppress (Kiendler-Scharr et al., 2009) or participate in NPF (Surratt et al., 2006) and that yields of HCHO from isoprene oxidation are non-linearly dependent on NO_x (Wolfe et al., 2015), it is difficult to assert an a priori expectation of the relationship between HCHO and NPF occurrence or intensity. Since HCHO is a product of oxidation of a wide array of VOCs and generation of less volatile products it is anticipated to have a positive association with GR and SP.
- Sulfur dioxide (SO₂). SO₂ is used here as a proxy for H₂SO₄, a key NPF precursor (Sipilä et al., 2010) and a key component of condensational growth (Pryor et al., 2011). It is thus anticipated to have a positive association with NPF occurrence, J₆, GR, and SP.
- Nitrogen dioxide (NO₂). NO₂ is used as a proxy for anthropogenic gas and particle emissions (Russell et al., 2012). Given NO_x emissions are associated with both increased CS (due to co-

emitted particles) and possible enhancement of condensable vapors (e.g. collocation of NO₂ and SO₂ from major industrial complexes) it is anticipated to have a negative association with NPF occurrence and positive association with J₆, GR, and SP.

- Ultraviolet radiation (UV). UV is used as an index of oxidant production, and is anticipated to have a positive association with NPF occurrence, J₆, GR, and SP.
- UV×SO₂ is used as a proxy for the production of H₂SO₄ from SO₂.
- Aerosol optical depth (AOD) × Ångström exponent (AE). AE 1/∞ particle size, and with AOD (AOD×AE) is used here to represent particle surface area and as an index of CS, and thus the competition for condensable vapors. It is therefore anticipated to have a negative association with NPF occurrence, J₆, GR, and SP.

SO₂ (column density assuming center of mass at 2.5 km), NO₂ (tropospheric column density), HCHO (total column amount), and local solar noon spectral irradiance at 310 nm (UV) are retrieved from the Ozone Monitoring Instrument (OMI) on the AURA satellite (overpass ~1345 local standard time (LST)) (Chance, 2002). All OMI data flagged as impacted by row anomalies are excluded from analyses (OMI Team, 2012) and a cloud screen of 0.3 is used to remove cloud contaminated retrievals. AOD and AE are retrieved using the ‘dark-target’ algorithm from MODIS on the Terra (~1030 LST) satellite (Levy et al., 2010). We use the Level 2 products at 13 km × 24 km and 10 km × 10 km at nadir for OMI (version 005) and MODIS (collection 051), respectively (see time series of the predictors in Figure 2).

Given NPF is frequently observed to occur on regional scales and aerosol particle optical properties are coherent on the mesoscale (Anderson et al., 2003) or regional scale (Sullivan et al., 2015), to increase the data availability all valid satellite retrievals within a 100 km radius of MMSF are averaged for each day to generate a value for each predictor. On average, ~60 and

~30 (~50) pixels within this radius are averaged for MODIS and OMI trace gases (UV), respectively. This averaging masks only moderate variability of AOD and AE as indicated by the mean and average standard deviation of both: $\langle \text{AOD} \rangle = 0.19$ ($\sigma = 0.06$) and $\langle \text{AE} \rangle = 1.27$ ($\sigma = 0.28$). UV measurements also did not vary greatly within the averaging radius (mean and average σ of $56.2 \text{ mW m}^{-2} \text{ nm}^{-1}$ and $5.9 \text{ mW m}^{-2} \text{ nm}^{-1}$). The trace gas measurements exhibit higher heterogeneity within the averaging radius with mean (average σ) SO_2 , NO_2 , and HCHO of -0.03 DU (0.51 DU), $3.72 \times 10^{15} \text{ mlc cm}^{-2}$ ($1.42 \times 10^{15} \text{ mlc cm}^{-2}$), and $1.21 \times 10^{16} \text{ mlc cm}^{-2}$ ($6.43 \times 10^{15} \text{ mlc cm}^{-2}$). Negative SO_2 values are considered valid within retrieval uncertainty and highlight the challenge in observing SO_2 from satellites (e.g. due to high ozone absorption of UV) (Veefkind et al., 2011).

Although some controls of NPF vary at the synoptic scale (e.g. high near-surface UV receipt), the meteorological context for nucleation detection in the near-surface layer is likely manifest at smaller scales (Crippa et al., 2012). Thus, while dynamical conditions are also compared for event and non-events days (Table 1), they are not included in the regression or principal component analyses due to lack of physical interpretability of components with high weighting on variables with a large scale discrepancy (i.e. in situ turbulence measurements versus satellite derived concentrations averaged over a 100 km footprint). The physical parameters included in the analysis derive from measurements at 46 m and are:

- Friction velocity (u_*) is used as a metric of turbulent mixing, which can both initiate nucleation (Wehner et al., 2010), and is associated with eroding of the nocturnal boundary layer and entrainment of freshly nucleated particles and/or precursors (e.g. $\text{SO}_2/\text{H}_2\text{SO}_4$) into the near surface layer (Crippa et al., 2012; Pryor et al., 2011). It is thus anticipated to be higher on NPF event days.

- Sensible heat flux (H) is used as an index of buoyancy-driven turbulent mixing and BVOC emissions being flushed upward from the canopy. It is thus anticipated to be higher on NPF event days.

Mean u_* and H are calculated for the duration of each NPF event (defined in Sec 2.2; mean event start – end for non-event days).

2.2 Event classification, growth rates, formation rates, and survival probabilities

We use the following automated methodology to identify the occurrence of clear NPF events followed by sustained growth, and to compute the event variables (J₆, GR, and SP):

- Only measurement days with > 95% (~23 hours) of data available are included.
- The geometric mean diameter for particles < 100 nm (D_g) and for nucleation mode particles (< 30 nm; D_{gNuc}) is used to define the start of NPF when $D_g - D_{gNuc}$ is first ≤ 10 nm within 2 hours of the maximum nucleation mode concentration, and ends when $D_g - D_{gNuc} > 10$ nm (Crippa and Pryor, 2013).
- J_6 is calculated as the rate change in nucleation mode particle concentration (N_{nuc} ; 6 – 30 nm), corrected for particle loss to coagulation (F_{coag}) and growth (F_{growth}) (Westervelt et al., 2013):

$$J_6 = \frac{dN_{nuc}}{dt} + F_{coag} + F_{growth} \dots (1)$$

$$\text{where } F_{coag} = \sum_{i=6 \text{ nm}}^{i=30 \text{ nm}} \left[\frac{1}{2} K(T)_{ii} N_i + \sum_{j=i+\Delta D_p}^{j=\max(D_p)} (K(T)_{ij} N_j) \right] \dots (2),$$

$$F_{growth} = \sum_{i=6 \text{ nm}}^{i=30 \text{ nm}} \left[\frac{N_i}{30 \text{ nm} - D_{pi}} * GR \right] \dots (3),$$

$K(T)$ is the Fuch's form Brownian coagulation coefficient calculated using the ambient T, N_i and D_{pi} are the number concentration and median diameter of bin i , and ΔD_p is the bin width.

- GR is calculated using linear regression fits to the time series of D_{gNuc} for 0 – 3 hours after

the event start (Pryor et al., 2014, 2010).

- All events with negative GR or J_6 were removed, as were those with duration of ≤ 1 hour.
- SP estimates are derived from the timescale for particle growth (τ^{growth}) from one bin into the next relative to the timescale for coagulation loss (τ^{coag}) to larger particles (Westervelt et al., 2013):

$$SP = \prod_{i=6}^{i=100} \frac{nm - \Delta D_p}{nm} \exp \left(- \frac{\tau_{D_{pi}, D_{pi+\Delta D_p}}^{\text{growth}}}{\tau_{dp_i}^{\text{coag}}} \right) \dots (4)$$

$$\tau_{D_{pi}, D_{pi+\Delta D_p}}^{\text{growth}} = \frac{(D_{pi+\Delta D_p} - D_{pi})}{GR} \dots (5)$$

$$\tau_{dp_i}^{\text{coag}} = \frac{1}{\frac{1}{2} K(T)_{ii} N_i + \sum_{j=i+\Delta D_p}^{\max(D_p)} K(T)_{ij} N_j} \dots (6)$$

NPF events are not distributed evenly across seasons (Figure 2), and the mechanisms of and controls on NPF may differ across seasons (Yu et al., 2015), thus analyses are conducted independently for leaf-on ($\text{LAI} \times T_{\max} \geq 50^{\text{th}}$ percentile; $\sim 460 \text{ m}^2 \text{ m}^{-2} \text{ K}$) and leaf-off periods ($\text{LAI} \times T_{\max} < 50^{\text{th}}$ percentile).

Although there is no significant difference in GR, estimated J_6 is significantly higher in 2007–2009 than in 2012–2013 (Figure 3). This difference may derive from differing instrumentation, sampling protocols (e.g. tubing length and loss correction; Pryor et al., 2014, 2010), or physical causes (e.g. from the drought of 2012 (Pryor et al., 2014)). For example, although inter-comparison of the FMPS and SMPS data indicate Pearson correlation of hourly average total particle number concentration > 0.9 , N_{nuc} from the SMPS was systematically higher (Pryor et al., 2010), potentially leading to higher estimated J_6 . Lab inter-comparison of these instruments also indicate that, to a degree that is a function of composition, the SMPS exhibits higher modal peak concentrations and lower modal diameter than the FMPS (Hornsby and Pryor, 2014). Thus, the data sets from the two experimental campaigns are treated separately.

2.3 Statistical Methods

None of the predictor variables pass an Anderson-Darling test for normality ($\alpha = 0.05$), so the concentrations are standardized prior to regression analyses and a Wilcoxon rank sum test (Wilks, 2011) is used to test the null hypothesis that event and non-event conditions are from the same populations and thus address Objective 1.

Event occurrence is a categorical variable, so logistic regression is employed to predict the logit (natural logarithm of the odds) of NPF event occurrence (Peng et al., 2002):

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 x_1 + \dots + \beta_i x_i \dots (7)$$

$$\text{and thus } p = \frac{e^{\alpha + \beta_1 x_1 + \dots + \beta_i x_i}}{1 + e^{\alpha + \beta_1 x_1 + \dots + \beta_i x_i}} \dots (8)$$

where p is the probability of event occurrence, α is the regression constant, and β_i is the regression coefficient for predictor variable x_i . As residual variance cannot be quantified for a binary variable, a generalized coefficient of determination (pseudo- r^2 , r_p^2) is used to assess model fit (Nagelkerke, 1991):

$$r_p^2 = \frac{1 - \left(\frac{L_I}{L_M}\right)^{\frac{2}{n}}}{1 - L_I^{\frac{2}{n}}} \dots (9)$$

where L_I is the likelihood of the null model (intercept only), L_M is the likelihood of the logistic regression model, and n is the number of observations. Goodness of fit is quantified using a likelihood ratio χ^2 test, with degrees of freedom = number of predictor variables:

$$\chi^2 = -2 \ln \left[\frac{L_I}{L_M} \right] \dots (10)$$

Multiple linear regression (MLR) analysis is used to determine the association between different chemical regimes and J_6 , GR, and SP of NPF events.

Prior to conducting the regression analyses, principal component analysis (PCA) is

applied to the predictor variable correlation matrices to evaluate the co-variability of the predictor variables. Since application of PCA requires a complete data set, missing observations are filled with the mean of all available measurement days (~30-35%, 15%, and 35% of days (~20-25%, 15%, and 25% of event days) are missing data from OMI (SO₂, NO₂, and HCHO), OMI (UV), and MODIS (AOD×AE), respectively).

All predictor variables are mean centered and standardized prior to application of the PCA and regression analyses. Although not a strict requirement for use of PCA, the results are generally more robust under the condition of multivariate normality (Kim and Kim, 2012), thus HCHO and NO₂ concentrations were also log-transformed. Principal components with eigenvalues greater than one are retained (Kaiser, 1960) (i.e. the first three components) and rotated using a Varimax (orthogonal) rotation (Richman, 1986).

3 Results

3.1 Principal component analysis

Some degree of multicollinearity is expected amongst the predictor variables, and was diagnosed using PCA to provide context for interpretation of the regression analyses. The first three principal components explain 84% and 77% (69% and 64%), and 76% and 78% (85% and 63%) of the variability in the predictor variables in 2007–2009 events and non-events, and 2012–2013 events and non-events during leaf-on (leaf-off).

The three orthogonal modes (Table 1) for all data sub-sets exhibit high loadings on:

- SO₂ and UV×SO₂. This mode represents variability in the availability of SO₂ (and by association H₂SO₄), and (at least for orthogonal components) indicates this variability is decoupled from the other predictors.
- LAI×T_{max}, UV, and NO₂(-). Although it was postulated that LAI×T_{max}, UV, and NO₂ could

be proxies for BVOC emissions, photochemical production of oxidants, and primary emissions, respectively, the co-variability of these predictors likely indicate their temporal variability is dominated by seasonality (with higher NO₂ in the cooler months, hence the negative weighting).

- HCHO and AOD×AE. This mode indicates co-variability of the proxy of VOC oxidation and CS, and may reflect the impact of stagnation in causing a build up of both HCHO and CS or the contribution of BVOC oxidation products to secondary organic aerosol formation (and thus increased AOD).

The three primary modes of variability (as manifest in the PCs) are very consistent between the two experimental periods, during leaf-on and leaf-off, and during NPF event and non-event days (Table 1). This implies that it may be possible to build more parsimonious models (using only a few of the predictors) but also that the chemical environments on event vs. non-event days are not readily differentiated.

3.2 Event versus non-event conditions

The automated event detection scheme indicates sustained NPF with clear evidence of growth (equivalent to A-class event days as defined by (Dal Maso et al., 2005)) on 16% and 20% of days with valid PSD measurements in 2007–2009 and 2012–2013, respectively. These frequencies are slightly lower than those derived using a subjective classification scheme (18% and 27%, respectively) (Pryor et al., 2014, 2010), likely due to stringent requirements for data completeness applied here to ensure the validity of calculated statistical metrics (J₆, GR, and SP).

Differences in dynamical and chemical conditions during event and non-event days are largely consistent with a priori expectations, indicate some consistency across the two experimental periods, and exhibit a clear dependence on season (leaf-on vs. leaf-off) (Figure 4;

281 Table 1). During leaf-on periods, u_* , H, and NO_2 are significantly ($\alpha = 0.05$) (NO_2 not significant
 282 in 2012–2013) higher on NPF event days. Higher u_* and H on event days may reflect turbulent
 283 initiation of NPF, or entrainment of nucleated particles or $\text{SO}_2/\text{H}_2\text{SO}_4$ into the near-surface layer
 284 during destabilization of the nocturnal boundary layer (Crippa et al., 2012; Pryor et al., 2011;
 285 Wehner et al., 2010). Higher NO_2 on NPF event days may indicate that anthropogenic precursor
 286 emissions are more important to NPF than the associated increase in CS due to primary particle
 287 emissions, or may be indicative of co-variability of NO_2 and the hydroxyl radical (OH). Also
 288 during the leaf-on period, $\text{LAI} \times T_{\text{max}}$, HCHO, and $\text{AOD} \times \text{AE}$ (HCHO not significant and $\alpha = 0.1$
 289 for $\text{AOD} \times \text{AE}$ in 2007 – 2009) are significantly lower on event days. Lower $\text{LAI} \times T_{\text{max}}$ on event
 290 days is likely reflective of the higher NPF event frequency in spring than summer, while lower
 291 HCHO may reflect the role of high isoprene concentrations in reducing oxidant availability (and
 292 thus suppressing NPF) (Kiendler-Scharr et al., 2009) or the co-variability of HCHO and
 293 $\text{AOD} \times \text{AE}$ (see discussion in section 3.1). Lower $\text{AOD} \times \text{AE}$ is also observed on event days
 294 consistent with high CS inhibiting NPF. Contrary to a priori expectations, neither UV nor SO_2
 295 differed significantly during event and non-event days in the leaf-on seasons. We postulate that
 296 the lack of difference in SO_2 concentrations may reflect the relatively high uncertainty in
 297 remotely-sensed SO_2 concentrations (or higher sensitivity to high-elevation SO_2) (OMI Team,
 298 2012; Veefkind et al., 2011) and/or that although H_2SO_4 plays a key role in NPF (Crippa et al.,
 299 2013; Pryor et al., 2011) due to continuing very high regional emissions of SO_2 (Pryor et al.,
 300 2010) it may not be a limiting factor in southern Indiana. The lack of apparent dependence on
 301 UV receipt during the leaf-on season suggests sufficient UV may be present during this period
 302 and other predictors become more significant, or may simply reflect the confounding influence
 303 of season (with highest NPF frequency in spring, when UV receipt is comparatively low).

During leaf-off periods, significantly higher $\text{LAI} \times T_{\max}$, H, and UV occurred on event days (H not significant in 2012–2013), which may be linked to photochemical production of H_2SO_4 driving cold-season nucleation and/or the coming of spring. There were no significant differences in u_* , SO_2 , NO_2 , HCHO, or $\text{AOD} \times \text{AE}$ on event versus non-event days during leaf-off in both measurement periods.

Thus apart from SO_2 , all of the proposed predictor variables of NPF occurrence, J_6 , GR, and SP exhibited significant differences in magnitude on event and non-event days. Hence all were included in the regression analyses.

Logistic regression equations with NPF occurrence as the predictand have a r_p^2 of 0.19 (0.19) and 0.13 (0.28) for 2007–2009 and 2012–2013 leaf-on (leaf-off), and all 4 models had χ^2 p-values < 0.01 indicating skill in predicting event occurrence over climatology (Table 2). Consistent with the above discussion, during leaf-on, $\text{AOD} \times \text{AE}$ has a significant (p-value < 0.1) negative regression coefficient for event occurrence, indicating increased CS suppresses NPF. $\text{LAI} \times T_{\max}$ also has a negative regression coefficient, consistent with the higher event occurrence in spring than summer. The negative coefficient on SO_2 , but positive coefficient on $\text{UV} \times \text{SO}_2$ in the logistic regression models of NPF occurrence may reflect the key importance of H_2SO_4 production during transport into the region, although these variables only have significant coefficients in MLR of data from 2007–2009. During leaf-off, there is a significant positive regression coefficient on UV, supporting the postulate that production of H_2SO_4 drives cold-season nucleation occurrence.

3.3 Influence of ambient conditions on growth rates, formation rates, and survival probabilities

MLR equations for GR exhibited the lowest r^2 values (0.13 to 0.65) and also the smallest

number of significant regression coefficients – with only UV exhibiting a non-zero coefficient for leaf-on during 2007-2009 (Table 2). MLR equations for J_6 are associated with higher r^2 values during leaf-on, but still relatively few predictors are associated with significant regression coefficients (Table 2). For example, the regression equation for J_6 in the leaf-on period of 2007-2009 has a significant negative weighting on HCHO and the intercept only (positive), but is associated with an r^2 of 0.52, while there are no significant regression coefficients for J_6 during leaf-on in 2012-2013. For the same experimental period, the coefficient in the equation for SP is significant and positive on HCHO which may reflect the covariation of HCHO and AOD $\times$ AE manifest in the PCA, and thus that HCHO is reflecting increased CS and suppression of J_6 due to competition for the low volatility vapors, and that the particles that do form have higher SP due to lower Aitken mode number concentrations (e.g. less self coagulation and/or competition for condensable vapors with other newly formed particles). SO₂ and UV $\times$ SO₂ have significant regression coefficients for equations for SP during both leaf-on experiments and leaf-off in 2012-2013, but the signs are not consistent (positive (negative) for SO₂ (UV $\times$ SO₂) for 2007-2009 leaf-on and 2012-2013 leaf-off, but negative (positive) for 2012-2013 leaf-on) making interpretation challenging.

4 Discussion and concluding remarks

PSD measurements made in a deciduous forest in the Midwestern USA in 2007–2009 and 2012–2013 are used to detect NPF and characterize those events (J_6 , GR, and SP). The resulting data sets are then used to determine the degree to which those variables can be ‘predicted’ using in situ measurements of atmospheric dynamics and satellite-based measurements of atmospheric composition. Previous work has indicated some potential for using of satellite-driven proxy algorithms for daily nucleation mode particle concentrations (Kulmala

et al., 2011) and/or ultrafine particle concentrations (Crippa et al., 2013). Results presented herein indicate that many of the proposed predictor variables: $LAI \times T_{\max}$, NO_2 , $HCHO$, UV , $AOD \times AE$, and $UV \times SO_2$ exhibit different median values when conditionally sampled by event versus non-event days (Table 1), and exhibit some skill (as measured by r^2 values of regression equations, Table 2) in predicting event occurrence (by logistic regression) and the magnitudes of J_6 , GR , and SP . However, the regression coefficients are relatively unstable (i.e. some differ in sign between the two experimental periods) and generally do not differ significantly from zero, indicating only modest explanation of the variability of NPF characteristics. The generalizability of these findings merits further investigation, as do possible causes of these results which include: (i) decoupling of columnar gas concentrations from those at the surface where the PSD measurements are made, (ii) the relatively large uncertainty in the satellite measurements, (iii) the spatial averaging used (approx. 100 km, necessitated by the relatively low numbers of satellite retrievals in the pixel containing the measurement site), (iv) selection of inappropriate or inadequate predictor variables (e.g. NH_3 is key to NPF, but only measured ~weekly from the Tropospheric Emission Spectrometer (TES), and key gases (e.g. H_2SO_4) are present at low concentrations and/or characterized by absorption bands that are not unique), (v) relatively small sample size of event days, and/or (vi) although NPF occurs on regional scales, event metrics (J_6 , GR , and SP) may exhibit heterogeneity on finer scales.

5 Acknowledgements

This work was supported under the NASA Earth and Space Science Fellowship Program - Grant “14-EARTH14F-0207” and a grant to SCP from NSF (1102309, 1517365). The satellite data used in this study were acquired as part of the NASA Earth-Sun System Division and Science Mission Directorate, and archived and distributed by the MODIS Level 1 and

373 Atmosphere Archive and Distribution System and by the Goddard Earth Sciences Data and
374 Information Services Center. The authors acknowledge K. Novick for maintaining the MMSF
375 AmeriFlux site.

376 **6 References**

- 377 Almeida, J., Schobesberger, S., Kürten, A., Ortega, I.K., Kupiainen-Määttä, O., Praplan, A.P.,
378 Adamov, A., Amorim, A., Bianchi, F., Breitenlechner, M., 2013. Molecular understanding
379 of sulphuric acid-amine particle nucleation in the atmosphere. *Nature* 502, 359–363.
- 380 Anderson, T.L., Charlson, R.J., Winker, D.M., Ogren, J.A., Holmén, K., 2003. Mesoscale
381 Variations of Tropospheric Aerosols*. *J. Atmos. Sci.* 60, 119–136.
- 382 Barkley, M.P., Smedt, I. De, Van Roozendaal, M., Kurosu, T.P., Chance, K., Arneth, A.,
383 Hagberg, D., Guenther, A., Paulot, F., Marais, E., 2013. Top-down isoprene emissions over
384 tropical South America inferred from SCIAMACHY and OMI formaldehyde columns. *J.*
385 *Geophys. Res. Atmos.* 118, 6849–6868.
- 386 Boy, M., Bonn, B., Kazil, J., Lovejoy, N., Turnipseed, A., Greenberg, J., Karl, T., Mauldin, L.,
387 Kusciuch, E., Smith, J., 2007. Relevance of several nucleation theories in different
388 environments, in: *Nucleation and Atmospheric Aerosols*. Springer, pp. 87–91.
- 389 Chance, K., 2002. OMI algorithm theoretical basis document, volume IV: OMI trace gas
390 algorithms. Available at [http://eosps0.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-](http://eosps0.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-04.pdf)
391 [04.pdf](http://eosps0.gsfc.nasa.gov/sites/default/files/atbd/ATBD-OMI-04.pdf).
- 392 Chance, K., Palmer, P.I., Spurr, R.J.D., Martin, R. V, Kurosu, T.P., Jacob, D.J., 2000. Satellite
393 observations of formaldehyde over North America from GOME. *Geophys. Res. Lett.* 27,
394 3461–3464.
- 395 Crippa, P., Petäjä, T., Korhonen, H., Afandi, Gse., Pryor, S.C., 2012. Evidence of an elevated
396 source of nucleation based on model simulations and data from the NIFTy experiment.
397 *Atmos. Chem. Phys.* 12, 8021–8036. doi:10.5194/acp-12-8021-2012
- 398 Crippa, P., Pryor, S.C., 2013. Spatial and temporal scales of new particle formation events in
399 eastern North America. *Atmos. Environ.* 75, 257–264. doi:10.1016/j.atmosenv.2013.04.051
- 400 Crippa, P., Spracklen, D., Pryor, S.C., 2013. Satellite-derived estimates of ultrafine particle
401 concentrations over eastern North America. *J. Geophys. Res. Atmos.* 118, 9968–9981.
402 doi:10.1002/jgrd.50707

403 Dal Maso, M., Kulmala, M., Riipinen, I., Wagner, R., Hussein, T., Aalto, P.P., Lehtinen, K.E.J.,
 404 2005. Formation and growth of fresh atmospheric aerosols: eight years of aerosol size
 405 distribution data from SMEAR II, Hyytiälä, Finland. *Boreal Environ. Res.* 10, 323–336.

406 Guenther, A.B., Zimmerman, P.R., Harley, P.C., Monson, R.K., Fall, R., 1993. Isoprene and
 407 monoterpene emission rate variability: model evaluations and sensitivity analyses. *J.*
 408 *Geophys. Res. Atmos.* 98, 12609–12617.

409 Henze, D.K., Seinfeld, J.H., 2006. Global secondary organic aerosol from isoprene oxidation.
 410 *Geophys. Res. Lett.* 33. doi:10.1029/2006GL025976

411 Hornsby, K.E., Pryor, S.C., 2014. A laboratory comparison of real-time measurement methods
 412 for 10–100-nm particle size distributions. *Aerosol Sci. Technol.* 48, 571–582.

413 Huete, A., Justice, C., Van Leeuwen, W., 1999. MODIS vegetation index (MOD13). Algorithm
 414 Theor. basis Doc. Available http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf.

415 Jaecker-Voirol, A., Mirabel, P., 1989. Heteromolecular nucleation in the sulfuric acid-water
 416 system. *Atmos. Environ.* 23, 2053–2057.

417 Kaiser, H.F., 1960. The application of electronic computers to factor analysis. *Educ. Psychol.*
 418 *Meas.* 141–151.

419 Kiendler-Scharr, A., Wildt, J., Dal Maso, M., Hohaus, T., Kleist, E., Mentel, T.F., Tillmann, R.,
 420 Uerlings, R., Schurr, U., Wahner, A., 2009. New particle formation in forests inhibited by
 421 isoprene emissions. *Nature* 461, 381–384.

422 Kim, D., Kim, S.-K., 2012. Comparing patterns of component loadings: Principal Component
 423 Analysis (PCA) versus Independent Component Analysis (ICA) in analyzing multivariate
 424 non-normal data. *Behav. Res. Methods* 44, 1239–1243.

425 Kulmala, M., Arola, A., Nieminen, T., Riuttanen, L., Sogacheva, L., Leeuw, G. de, Kerminen,
 426 V.-M., Lehtinen, K.E.J., 2011. The first estimates of global nucleation mode aerosol
 427 concentrations based on satellite measurements. *Atmos. Chem. Phys.* 11, 10791–10801.

428 Kulmala, M., Lehtinen, K.E.J., Laaksonen, A., 2006. Cluster activation theory as an explanation
 429 of the linear dependence between formation rate of 3nm particles and sulphuric acid
 430 concentration. *Atmos. Chem. Phys.* 6, 787–793.

431 Kulmala, M., Pirjola, L., Mäkelä, J.M., 2000. Stable sulphate clusters as a source of new
 432 atmospheric particles. *Nature* 404, 66–69.

433 Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V.-M., Birmili, W.,
 434 McMurry, P.H., 2004. Formation and growth rates of ultrafine atmospheric particles: a
 435 review of observations. *J. Aerosol Sci.* 35, 143–176.

436 Laaksonen, A., Kulmala, M., O'Dowd, C.D., Joutsensaari, J., Vaattovaara, P., Mikkonen, S.,
 437 Lehtinen, K.E.J., Sogacheva, L., Maso, M.D., Aalto, P., 2008. The role of VOC oxidation
 438 products in continental new particle formation. *Atmos. Chem. Phys.* 8, 2657–2665.

439 Levy, R.C., Remer, L.A., Kleidman, R.G., Mattoo, S., Ichoku, C., Kahn, R., Eck, T.F., 2010.
 440 Global evaluation of the Collection 5 MODIS dark-target aerosol products over land.
 441 *Atmos. Chem. Phys.* 10, 10399–10420.

442 McMurtry, P.H., 1980. Photochemical aerosol formation from SO₂: A theoretical analysis of
 443 smog chamber data. *J. Colloid Interface Sci.* 78, 513–527.

444 Merikanto, J., Spracklen, D. V, Mann, G.W., Pickering, S.J., Carslaw, K.S., 2009. Impact of
 445 nucleation on global CCN. *Atmos. Chem. Phys.* 9, 8601–8616.

446 Metzger, A., Verheggen, B., Dommen, J., Duplissy, J., Prevot, A.S.H., Weingartner, E.,
 447 Riipinen, I., Kulmala, M., Spracklen, D. V, Carslaw, K.S., 2010. Evidence for the role of
 448 organics in aerosol particle formation under atmospheric conditions. *Proc. Natl. Acad. Sci.*
 449 107, 6646–6651.

450 Nagelkerke, N.J.D., 1991. A note on a general definition of the coefficient of determination.
 451 *Biometrika* 78, 691–692.

452 O'Dowd, C.D., Aalto, P., Hmeri, K., Kulmala, M., Hoffmann, T., 2002. Aerosol formation:
 453 Atmospheric particles from organic vapours. *Nature* 416, 497–498.

454 OMI Team, 2012. Ozone monitoring instrument (OMI) data user's guide. OMI-DUG-5.0.
 455 Available at: [http://disc.sci.gsfc.nasa.gov/Aura/data-](http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/additional/documentation/README.OMI_DUG.pdf)
 456 [holdings/additional/documentation/README.OMI_DUG.pdf](http://disc.sci.gsfc.nasa.gov/Aura/data-holdings/additional/documentation/README.OMI_DUG.pdf).

457 Peng, C.-Y.J., Lee, K.L., Ingersoll, G.M., 2002. An introduction to logistic regression analysis
 458 and reporting. *J. Educ. Res.* 96, 3–14.

459 Pierce, J.R., Westervelt, D.M., Atwood, S.A., Barnes, E.A., Leaitch, W.R., 2014. New-particle
 460 formation, growth and climate-relevant particle production in Egbert, Canada: analysis from
 461 1 year of size-distribution observations. *Atmos. Chem. Phys.* 14, 8647–8663.

462 Pryor, S.C., Barthelmie, R.J., Sørensen, L.L., McGrath, J.G., Hopke, P., Petäjä, T., 2011. Spatial
 463 and vertical extent of nucleation events in the Midwestern USA: insights from the
 464 Nucleation In ForesTs (NIFTy) experiment. *Atmos. Chem. Phys.* 11, 1641–1657.
 465 doi:10.5194/acp-11-1641-2011

466 Pryor, S.C., Hornsby, K.E., Novick, K.A., 2014. Forest canopy interactions with nucleation
 467 mode particles. *Atmos. Chem. Phys.* 14, 11985–11996. doi:10.5194/acp-14-11985-2014

468 Pryor, S.C., Spaulding, A.M., Barthelmie, R.J., 2010. New particle formation in the Midwestern
 469 USA: Event characteristics, meteorological context and vertical profiles. *Atmos. Environ.*
 470 44, 4413–4425. doi:10.1016/j.atmosenv.2010.07.045

471 Qi, X.M., Ding, A.J., Nie, W., Petäjä, T., Kerminen, V.-M., Herrmann, E., Xie, Y.N., Zheng,
 472 L.F., Manninen, H., Aalto, P., 2015. Aerosol size distribution and new particle formation in
 473 the western Yangtze River Delta of China: 2 years of measurements at the SORPES station.
 474 *Atmos. Chem. Phys.* 15, 12445–12464.

475 Richman, M., 1986. Rotation of principal components. *J. Climatol.* 6, 293–335.

476 Russell, A.R., Valin, L.C., Cohen, R.C., 2012. Trends in OMI NO₂ observations over the United
 477 States: effects of emission control technology and the economic recession. *Atmos. Chem.*
 478 *Phys.* 12, 12197–12209.

479 Sipilä, M., Berndt, T., Petäjä, T., Brus, D., Vanhanen, J., Stratmann, F., Patokoski, J., Mauldin,
 480 R.L., Hyvärinen, A.-P., Lihavainen, H., 2010. The role of sulfuric acid in atmospheric
 481 nucleation. *Science* 327, 1243–1246.

482 Sullivan, R.C., Levy, R.C., Pryor, S.C., 2015. Spatiotemporal coherence of mean and extreme
 483 aerosol particle events over eastern North America as observed from satellite. *Atmos.*
 484 *Environ.* 112, 126–135. doi:10.1016/j.atmosenv.2015.04.026

485 Surratt, J.D., Murphy, S.M., Kroll, J.H., Ng, N.L., Hildebrandt, L., Sorooshian, A., Szmigielski,
 486 R., Vermeylen, R., Maenhaut, W., Claeys, M., 2006. Chemical composition of secondary
 487 organic aerosol formed from the photooxidation of isoprene. *J. Phys. Chem. A* 110, 9665–
 488 9690.

489 US Geological Survey, 2011. Gap Analysis Program (GAP) National Land Cover, Version 2.

490 Veefkind, J.P., Boersma, K.F., Wang, J., Kurosu, T.P., Krotkov, N., Chance, K., Levelt, P.F.,
 491 2011. Global satellite analysis of the relation between aerosols and short-lived trace gases.
 492 *Atmos. Chem. Phys.* 11, 1255–1267.

493 Wan, Z., 1999. MODIS land-surface temperature algorithm theoretical basis document. MODIS
 494 Land-surface Temp. Algorithm Theor. Basis Doc. (LST ATBD). Available
 495 http://modis.gsfc.nasa.gov/data/atbd/atbd_mod11.pdf.

496 Wehner, B., Siebert, H., Ansmann, A., Ditas, F., Seifert, P., Stratmann, F., Wiedensohler, A.,
 497 Apituley, A., Shaw, R.A., Manninen, H.E., 2010. Observations of turbulence-induced new
 498 particle formation in the residual layer. *Atmos. Chem. Phys.* 10, 4319–4330.

499 Westervelt, D.M., Pierce, J.R., Riipinen, I., Trivitayanurak, W., Hamed, A., Kulmala, M.,
 500 Laaksonen, A., Decesari, S., Adams, P.J., 2013. Formation and growth of nucleated
 501 particles into cloud condensation nuclei: model–measurement comparison. *Atmos. Chem.*
 502 *Phys.* 13, 7645–7663.

- 503 Wilks, D.S., 2011. Statistical methods in the atmospheric sciences. Academic press.
- 504 Wolfe, G.M., Kaiser, J., Hanisco, T.F., Keutsch, F.N., de Gouw, J.A., Gilman, J.B., Graus, M.,
505 Hatch, C.D., Holloway, J., Horowitz, L.W., 2015. Formaldehyde production from isoprene
506 oxidation across NO_x regimes. *Atmos. Chem. Phys. Discuss.* 15, 31587–31620.
- 507 Yu, F., Luo, G., Pryor, S.C., Pillai, P.R., Lee, S.H., Ortega, J., Schwab, J.J., Hallar, A.G.,
508 Leaitch, W.R., Aneja, V.P., Smith, J.N., Walker, J.T., Högrefe, O., Demerjian, K.L., 2015.
509 Spring and summer contrast in new particle formation over nine forest areas in North
510 America. *Atmos. Chem. Phys. Discuss.* 15, 21271–21298. doi:10.5194/acpd-15-21271-
511 2015
- 512 Yu, F., Turco, R.P., 2000. Ultrafine aerosol formation via ion-mediated nucleation. *Geophys.*
513 *Res. Lett.* 27, 883–886.
- 514 Zhang, Q.I., Stanier, C.O., Canagaratna, M.R., Jayne, J.T., Worsnop, D.R., Pandis, S.N.,
515 Jimenez, J.L., 2004. Insights into the chemistry of new particle formation and growth events
516 in Pittsburgh based on aerosol mass spectrometry. *Environ. Sci. Technol.* 38, 4797–4809.

517 Tables

518 Table 1. Median dynamic and chemical conditions measured during NPF event and non-event days (2007–2009 and 2012–2013) at
 519 MMSF. $\text{LAI} \times T_{\text{max}}$, u_* , and H are from in situ measurements and UV, SO_2 , NO_2 , HCHO, and $\text{AOD} \times \text{AE}$ are satellite measurements
 520 from OMI and MODIS averaged over a 100 km radius around the research site (Figure 1). As events are not distributed evenly across
 521 seasons, and the mechanisms of and controls on NPF may differ across seasons (Yu et al., 2015) (Figure 2), leaf-on and leaf-off
 522 periods are treated independently. **Bold** (*Italic*) values indicate rejection of the null hypothesis that event and non-event days are from
 523 the same population (Wilcoxon rank sum test $\alpha = 0.05$ (0.10)). Also shown are variables with high ($|\text{loading}| \geq 0.5$) loadings on the first
 524 (single underline), second (double underline), and third (dashed underline) principal components of the PCA of the seven data-sets
 525 (negative loadings are indicated by the shading).

			n	$\text{LAI} \times T_{\text{max}}$ ($\text{m}^2 \text{m}^{-2} \text{K}$)	u_* (m s^{-1})	H (W m^{-2})	UV ($\text{mW m}^{-2} \text{nm}^{-1}$)	SO_2 (DU)	NO_2 $\times 10^{15}$ (mlc cm^{-2})	HCHO $\times 10^{16}$ (mlc cm^{-2})	AOD $\times \text{AE}$	UV $\times \text{SO}_2$ ($\text{DU mW m}^{-2} \text{nm}^{-1}$)
Leaf-on	2007	Non-event	129	<u>1300</u>	0.33	45	<u>88</u>	<u>0.047</u>	<u>2.77</u>	<u>1.3</u>	<u>0.37</u>	<u>4.8</u>
	–											
	2009	Event	27	<u>1100</u>	0.48	120	<u>87</u>	<u>0.020</u>	<u>3.64</u>	<u>1.2</u>	<u>0.33</u>	<u>1.3</u>
Leaf-off	2012	Non-event	205	<u>1200</u>	0.44	76	<u>83</u>	<u>-0.008</u>	2.34	<u>1.4</u>	<u>0.19</u>	<u>-0.2</u>
	–											
	2013	Event	53	<u>1100</u>	0.59	130	<u>78</u>	<u>0.022</u>	2.89	<u>1.2</u>	<u>0.06</u>	<u>1.8</u>
Leaf-off	2007	Non-event	257	<u>230</u>	0.62	66	<u>22</u>	<u>-0.115</u>	<u>5.10</u>	<u>8.9</u>	0.13	<u>-2.8</u>
	–											
	2009	Event	47	<u>250</u>	0.58	100	<u>41</u>	<u>0.008</u>	4.55	<u>8.5</u>	<u>0.16</u>	<u>0.5</u>
Leaf-off	2012	Non-event	97	<u>310</u>	0.53	60	22	<u>-0.049</u>	<u>3.75</u>	<u>1.0</u>	<u>0.11</u>	<u>-1.2</u>
	–											
	2013	Event	20	<u>370</u>	0.58	72	<u>35</u>	<u>-0.245</u>	5.52	<u>1.1</u>	<u>0.14</u>	<u>-6.3</u>

526
527

Table 2. Regression coefficients (c) and p-values for predictor variables used in the logistic (event occurrence) and multiple linear (J₆, GR, and SP) regression models. Also shown is the generalized coefficients of determination (Nagelkerke, 1991) and likelihood ratio χ^2 test p-values for the logistic regressions, and variance explained (r^2) and F-test p-values for the multiple linear regression models. **Bold** indicates significant (p-value < 0.1) regression coefficients. Sample size is shown in the second column (event, non-event).

Leaf-on			Event Occurrence		GR (nm h ⁻¹)		J ₆ (cm ⁻³ s ⁻¹)		SP	
			c	p-value	c	p-value	c	p-value	c	p-value
2007-2009 n=27, 130	Intercept		-1.88	0.00	2.13	0.00	4.10	0.00	0.06	0.00
	LAI×T _{max} (m ² m ⁻² K)		-0.51	0.05	-0.02	0.92	0.61	0.34	0.01	0.50
	SO ₂ (DU)		-2.76	0.02	-0.38	0.63	2.53	0.31	-0.19	0.01
	NO ₂ (mlc cm ⁻²)		0.34	0.25	0.21	0.35	-0.68	0.32	0.01	0.74
	HCHO (mlc cm ⁻²)		0.36	0.25	0.40	0.15	-2.07	0.02	0.04	0.09
	UV (mW m ⁻² nm ⁻¹)		0.39	0.19	0.41	0.08	-0.07	0.92	0.02	0.35
	AOD×AE		-0.73	0.05	-0.21	0.41	-0.30	0.71	-0.01	0.52
	UV×SO ₂ (DU mW m ⁻² nm ⁻¹)		2.64	0.02	0.09	0.91	-2.93	0.24	0.14	0.04
	r _p ² , r ² (p-value)		0.19 (0.008)		0.25 (0.51)		0.52 (0.03)		0.49 (0.05)	
			c	p-value	c	p-value	c	p-value	c	p-value
2012-2013 n=55, 206	Intercept		-1.54	0.00	2.01	0.00	1.47	0.00	0.06	0.00
	LAI×T _{max} (m ² m ⁻² K)		-0.35	0.08	0.08	0.65	-0.06	0.72	-0.01	0.70
	SO ₂ (DU)		-0.13	0.69	0.63	0.13	-0.34	0.39	0.07	0.03
	NO ₂ (mlc cm ⁻²)		0.16	0.40	0.06	0.73	-0.24	0.20	0.00	0.94
	HCHO (mlc cm ⁻²)		0.11	0.59	0.26	0.16	0.18	0.31	0.00	0.79
	UV (mW m ⁻² nm ⁻¹)		0.28	0.18	0.04	0.80	-0.03	0.86	0.01	0.62
	AOD×AE		-0.83	0.00	-0.13	0.42	0.23	0.14	-0.03	0.06
	UV×SO ₂ (DU mW m ⁻² nm ⁻¹)		0.29	0.38	-0.46	0.25	0.45	0.24	-0.07	0.04
	r _p ² , r ² (p-value)		0.13 (0.002)		0.13 (0.41)		0.23 (0.07)		0.15 (0.31)	
			c	p-value	c	p-value	c	p-value	c	p-value

		Event Occurrence		GR (nm h ⁻¹)		J ₆ (cm ⁻³ s-1)		SP		
		c	p-value	c	p-value	c	p-value	c	p-value	
Leaf-off	2007-2009 n=47, 256	Intercept	-1.96	0.00	1.59	0.00	3.90	0.00	0.12	0.00
		LAI×T _{max} (m ² m ⁻² K)	-0.06	0.71	0.24	0.16	-0.88	0.22	0.08	0.00
		SO ₂ (DU)	0.14	0.71	0.11	0.80	-0.12	0.95	0.06	0.29
		NO ₂ (mlc cm ⁻²)	0.09	0.59	-0.13	0.43	-0.76	0.29	-0.02	0.32
		HCHO (mlc cm ⁻²)	-0.08	0.62	0.00	1.00	0.20	0.80	0.03	0.15
		UV (mW m ⁻² nm ⁻¹)	0.93	0.00	-0.26	0.15	0.93	0.21	-0.07	0.00
		AOD×AE	-0.12	0.46	0.26	0.14	-0.52	0.47	-0.01	0.66
		UV×SO ₂ (DU mW m ⁻² nm ⁻¹)	-0.16	0.62	-0.03	0.94	0.61	0.74	-0.05	0.37
		r _p ² , r ² (p-value)	0.19 (<0.001)		0.13 (0.55)		0.13 (0.59)		0.35 (0.01)	
		2012-2013 n=18, 96		c	p-value	c	p-value	c	p-value	c
	Intercept		-2.08	0.00	1.47	0.00	1.03	0.00	0.05	0.00
	LAI×T _{max} (m ² m ⁻² K)		0.54	0.06	-0.03	0.87	-0.46	0.17	0.09	0.00
	SO ₂ (DU)		-0.41	0.48	0.34	0.39	1.13	0.10	-0.09	0.06
	NO ₂ (mlc cm ⁻²)		0.55	0.14	-0.11	0.61	-0.10	0.78	-0.01	0.53
	HCHO (mlc cm ⁻²)		0.37	0.29	0.13	0.50	-0.26	0.40	0.03	0.18
	UV (mW m ⁻² nm ⁻¹)		0.84	0.01	0.13	0.61	0.83	0.07	-0.07	0.02
	AOD×AE		-0.29	0.35	0.35	0.13	-0.15	0.67	0.06	0.02
	UV×SO ₂ (DU mW m ⁻² nm ⁻¹)		0.01	0.98	-0.33	0.35	-1.55	0.02	0.10	0.02
	r _p ² , r ² (p-value)		0.28 (0.005)		0.65 (0.08)		0.67 (0.07)		0.75 (0.02)	

536

537

538

539

Figure captions

Figure 1. USGS land use surrounding the Morgan Monroe State Forest research site in Southern, IN (39.317°N, 86.417°W; red star) (US Geological Survey, 2011).

Figure 2. Daily biological, dynamic, radiative, chemical, and aerosol properties on days with PSD measurements (2007–2009 and 2012–2013). $LAI \times T_{\max}$, u_* , and H are from in situ measurements and UV, SO_2 , NO_2 , $HCHO$, and $AOD \times AE$ are satellite measurements from OMI and MODIS averaged over a 100 km radius around the Morgan Monroe State Forest research site (Figure 1). Event days are shown in red and non-event days are in blue.

Figure 3. Cumulative probability plot of GR (blue/cyan) and J_6 (red/magenta) as derived from measurements with the SMPS in 2007–2009 and FMPS in 2012–2013. Systematically higher nucleation mode number concentration measured by the SMPS relative to the FMPS (Pryor et al., 2010) may explain the ~2 times higher J_6 measured in 2007–2009 relative to 2012–2013.

Figure 4. Cumulative probability plots of the predictor variables during events (starred, thick lines) and non-events (thin lines) (2007–2009 and 2012–2013). Statistical differences in event versus non-event distributions are quantified using a Wilcoxon rank sum test and presented in Table 1.

Figure 1

Figure 2

Figure 3

Figure 4