We thank both reviewers for thoughtful suggestions that have helped us improve the manuscript.

Reviewer 1

General Comments:

The study presents aerosol measurements over two years in Tucson, a major city surrounded by weakly populated desert. The measurements include particle size distributions, aerosol composition, CCN and total number concentrations. By means of several assumptions and simplified models, CCN closure is attempted. The data set is clearly unique for this location and also exceeds many other data sets that are often limited to a single season or few months at a time. A new clustering method has been used to sort size distributions based on their likely origin and history. Conclusions on the sources of aerosol particles and reasons for distribution shapes are drawn based on the observed evolution of size distributions in various seasons and the skills of the CCN closure models. Therefore, the current study exceeds previous ones in terms of the measurement period and tools that are used to interpret data. However, I think previous literature should be more carefully taken into account and discussed.

I have several more comments below that should be taken into account before this manuscript can be recommended for publication.

Major comments
1) Introduction
The introduction is very long and quite disorganized. It should review the current knowledge of data that are similar to those as used in the current study, e.g. data sets of size distributions and composition that are used to perform CCN closure. Details on specific organic aerosol properties such as surface tension etc. (p. 3867) distract only from the main focus of the current paper. In addition, it seems that the last paragraph on p. 3868 is redundant as it is repeated in the following.

>Response: The introduction has been shortened, removing less relevant sections relating to organic aerosol chemistry.

2) Wording
At many places, quite inaccurate or misleading terminology is used. E.g. 'aerosol chemistry' or 'particle chemistry' is often used and it does not become clear whether chemical processes or composition is meant.

>Response: Cases with ambiguous terms have been amended.

Other instances include p. 3865, l. 23: 'cloud droplet distribution’ – I don’t think that any of the studies cited here compared their data to cloud droplet distributions.

>Response: McFiggans et al. 2006 shows droplet distribution for an idealized salt activation case; however, we have changed “distribution” to “number” which better reflects the purpose of this statement.
p. 3866, l. 13; p. 3874, l. 18; p. 3880, l. 14; p. 3882, l. 28/9: which processes are referred to here?

>Response: These instances have either been removed or revised.

3) Data discussion
In the discussion part of the paper (Sections 3 and 4) often words like 'maybe', 'likely', 'probably' etc are used. While I understand that it might be difficult to give a clear and unambiguous interpretation of the data due to the somewhat limited number of measured parameters, a somewhat more detailed discussion should be given that weights the possible processes/effects in a more quantitative way.

>Response: It is simply not possible to quantify cause-and-effect relationships definitively using purely the observations that we have reported. We have tried to use language appropriate with the level of certainty associated with interpretations made on the reported data.

4) Previous literature
In the introduction, some previous CCN studies are cited together with their challenges and difficulties therein. However, in the discussion section not a single previous study is cited even though there are numerous studies that have been performed in regions where similar mixtures of fresh and aged aerosol exist. Also effects on number concentration, size distribution shapes etc due to daily, seasonal and source-dependent effects have been discussed there. One large aspect that has been highlighted in detail in many studies is the mixing state of fresh vs aged aerosol. The current study has to take into account findings from prior studies and put the current data set in their context. (see also next comment)

>Response: We thank the reviewer for pointing this out and while we are not entirely sure what specific references the reviewer had in mind, we have examined the archives of papers more strictly and try to make such references in the Discussion Section. A few representative papers are chosen that focused on urban environments and we added the following text:

“To put the TACO results in more context, fresh pollution aerosol in other urban areas such as Riverside and Houston could not be fully represented without knowledge of size-resolved composition (Cubison et al., 2008; Ervens et al., 2010). A number of other studies have shown that mixing state can help improve predictive capability of CCN behavior (Wex et al., 2010), including Atlanta (Padro et al., 2012) and during early morning rush hour near Mexico City (Lance et al., 2013); but studies also report that hydrophobic particles emitted in urban areas quickly (~ few hours) become internal mixtures via condensation of secondary hygroscopic species (e.g., Wang et al., 2010; Mei et al., 2013).”

5) Mixing state
In the current paper, mixing state is largely neglected and the inability of the simplified model approaches used here to predict CCN is explained by 'probably associated with the complexity of the aerosol mixing state' (p. 3880, l. 26). Given that the mixing state might play an important role for part of the current data set, the question arises how meaningful a single kappa is to capture the hygroscopicity of the total aerosol population. This approach should be either better justified or revised. Mixing state should be also discussed in the context of the relative role of various parameters that determine the activated fraction of an aerosol population, e.g., on p. 3881, l. 28/9 and p. 3882, l. 10.
Response: While the reviewer raises a good point regarding the physical meaning of parameters included in a simple model, we do not believe it is unjustified to use a knowingly oversimplified parameterization as a way of identifying the occasions when it is not satisfactory. We have added an extra discussion at the beginning of this section to identify that this is our intention: “One major simplification is the limitation of the treatment of hygroscopicity to a bulk measurement, which is permitted to vary temporally but does not isolate size dependent changes in hygroscopicity nor the hygroscopicity distribution, which may be an important component in relation to external mixing. These aspects are beyond the scope of these parameterizations and are likely to contribute to model shortfalls. Forthcoming work will separately study the degree of correspondence of hygroscopicity between the sub- and supersaturated regimes, size-dependent hygroscopicity and composition, and the closure of hygroscopicity from composition measurements.”

Minor comments
p. 3864, l. 15: ‘can be parameterized’ should be specified here.

Response: In handling comment 2, this should now be clearer.

p. 3865, l. 6: Cloud microphysical and optical properties are not only governed by aerosol number but also by the total amount of liquid water, which in turn is a function of cooling rates.

Response: We have replaced “governs” to “contributes to governing”

p. 3865, l. 25: This list should also include mixing state already. It is only discussed later even though it has been shown by detailed studies that it might be one of the determining factors in CCN closure studies.

Response: We have made this addition.

p. 3866, l. 24: In the two cited studies, CCN closure was quite satisfactory if mixing state was taken into account. This sentence should be reworded.

Response: Agreed, the current wording is ambiguous. The phrase “under assumptions of bulk hygroscopic properties has been added” to clarify the intended meaning.

p. 3866, l. 25-27: This sentence seems out of place here.

Response: We have revised this text.

l. 3867, l. 5: Did any of the cited studies indeed look at the effect of chemical processes and/or coagulation on size distributions and CCN properties?

Response: These references have been revised. Also the text was somewhat misleading in the manner it connected the statement with the references, so it has been adjusted to be more concise.

p. 3867, l. 12ff (cf. my comment 1)): This information is irrelevant for the current study. If you choose to keep it in the (already quite lengthy) introduction, a more balanced discussion should be given. E.g. sensitivity studies have shown that surface tension effects are rather negligible for CCN effects (Ervens et al., JGR, 2005).
Response: The introduction has been revised in the process of addressing Comment 1.

p. 3874, l. 25: Is there any explanation for the higher particle concentration during weekends?

Response: The CCN concentration is higher during the evening, of which we make mention, however, we do not have a full explanation for whether the enhancement is a result of different evening emissions patterns on weekends or some other pathway.

p. 3875, l. 13ff. (i) This paragraph should be a separate section.

Response: We respectfully disagree with this suggestions as this section is still part of the discussion on diurnal and weekly cycles.

(ii) Add to the numbers in parentheses ‘kappa = ’

Response: We have made this addition.

p. 3875, l. 1-12: Have the described effects such as a shift in size-distributions due to condensation of semivolatile compounds and the switch from the importance of semivolatiles to more biomass burning been observed in previous studies?

Response: Not specifically in Tucson. The presence of domestic wood burning emissions is unique to the winter season and is mentioned as a possible explanation for the difference seen in the diurnal cycle in winter which is not present in other seasons. There was no direct measurement of a tracer which could be used to separate quantitatively the contribution of biomass burning.

p. 3877, l. 14: Has it been observed previously that increasing partitioning of nitrate can indeed affect size distributions to an observable extent?

Response: Indirectly this is discussed in Andreae and Rosenfeld (2008, p29) referring to the tendency for submicron nitrate to partition onto existing particles (in their discussion, they cite Kleeman et al., (1999)), which increases activation ratio but does not increase CN.
We thank both reviewers for thoughtful suggestions that have helped us improve the manuscript.

Reviewer 2

General Comments:

This paper presents a detailed climatology of CCN concentrations, characteristics, and related aerosol properties based on two years of measurements in an “urbanized desert.” In addition, CCN closure analyses are conducted and the predictive skill of parameterized models of CCN concentrations are explored. This paper makes important contributions to the field by presenting long-term measurements in an under-studied environment; however, this paper would benefit from a more thorough discussion of some of the methodology and the broader implications of this work. I recommend publication after the consideration of the following comments.

Specific Comments

1. Two years of CCN measurements conducted at a supersaturation of 0.2% are presented. The authors should consider noting briefly why this value was chosen. Is this somewhat arbitrary or does it reflect a “typical” updraft condition/conditions specific to the climate of interest?

>Response: It is not an especially relevant supersaturation for the Tucson climate, however, it is a value that has been used in a number of other field measurements. It is also a worthy choice from the perspective of understanding CCN variability since it results in an activation diameter in the 100-200 nm range, for typical particle hygroscopicities. This range is close to the peak in the number distribution and hence CCN at 0.2% supersaturation is well suited to making comparisons with variability in the aerosol size distribution. The latter point is not relevant to the conclusions, and so we do not feel it is beneficial to attempt to try and offer further explanation for this choice in the manuscript.

2. Details regarding the cluster analysis are provided in the supplement; however, some of this information should be provided in the main text. In addition, further information is needed to describe this portion of the methodology, either in the supplement or the main text. Specifically:

- The reasoning for the selection of 4 clusters in the K-means clustering analysis should be included in the main text, as should a brief description of the definition of cluster associations/cluster assignment weights, especially given the presentation of this latter metric in Figure 5.

>Response: We have made this addition.

- Because the “fuzzying” of the cluster associations is outside of the more traditional application of K-means clustering, in which distributions are assigned to a single cluster, a justification of this choice should be provided. Would changing the number of clusters be another method by which the authors could achieve their goal of considering the transitions between/combination of physical process/regimes?

>Response: Part of this has been covered in the additional text added in response to the comment.
above. K-means is most successful in isolating patterns when there are distinct low-density regions separating high-density regions that are in a small neighborhood surrounding the centroid. In this case the data are far more continuous and thus the transition “region” between two clusters can have a significant effect on the cluster centroid and respective mean properties, since there are a large number of observations in this category. The implication of this is two-fold: 1) it is difficult to objectively determine the optimal number of clusters; and 2) the decision on how the cluster boundary is defined has an important effect on the cluster. Fuzzying helps to relieve the sensitivity to the latter, which could also be achieved by increasing the number of clusters, but this carries the penalty of more degrees of freedom i.e. 100% membership of C could be adequately approximated by 60% A + 40% B.

While the objective was to retain the smallest number of clusters, was there evidence that other potentially important physical processes/aerosol regimes contributed to variability in particle size distribution properties when more clusters were considered?

>Response: The answer to the question is “no”. Retaining the 4 clusters revealed the important result of the split seen in the summer between the “nucleation” cluster and “coagulation/condensation” cluster. Adding more clusters did not yield additional non-linear behavior (at least for moderate increases in the number of clusters) such that intermediate states could be well represented by combinations of the current clusters, which is permitted by the fuzzy boundaries.

3. Under some circumstances, cluster-derived parameters led to improvements in model performance. How do the authors envision the application of the presented clustering methodology in future field studies in other geographic regions and climates? I am particularly interested to know if there is evidence for the potential of a more generalized approach that could be applied to a large number of sites. In other words, some of the clusters presented here are likely to be specific to the region of interest or other very similar climates/sites. In their clustering analysis, did the authors see evidence for the potential for defining the clusters such that they more broadly represent the processes influencing size distributions at a wide number of sites?

>Response: We acknowledge that there could be many ways to implement this type of clustering to a more generalized problem. The method relies on using training data to first establish the cluster shapes – but following that, any data could be classified. If there is the potential for a regime not included in the training data to be important then some minimum threshold should be established to define an “undefined” cluster association such that it could be flagged that a particular site is not well reflected in the clusters. The critical aspect of establishing generalized clusters would be to capture variability that is sufficiently “extreme” to describe the salient modes in the size distribution, while sufficiently frequent to garner enough member associations for quality statistics. We do not claim to be able to optimally determine this, however, it is an interesting problem.

4. Along similar lines, with the aim of improving the representation of cloud properties and processes in large-scale models, in what ways does this work inform future similar campaigns in which long-term measurements of CCN and aerosol properties are measured? What measurements are crucial to this effort?

>Response: The community could benefit from studies of this type (arid and non-arid) regarding the feasibility of simplistic closure assumptions to identify how consistent this is. We would suggest that, at minimum, continuous CCN and aerosol size distribution data are needed to
perform sensitivity studies that assess the respective role of number, size and hygroscopicity on CCN variability. We feel no additional text is required in the manuscript to address this comment.

5. Particle composition measurements are for PM2.5, while size distribution measurements cover a range of 13 – 748 nm. Could differences in composition and/or mixing state for particles larger and smaller than 748 nm contribute to heterogeneity in the degree of variance in CCN concentrations explained by size and composition? Could the influence of this vary temporally, on hourly and seasonal scales?

>Response: Yes, without doubt, one could consider cases where size dependent composition effects significantly affect the inferences that can be made from PM2.5 speciation on CCN relevant composition. A good example of this in Tucson would be the prevalence of dust in PM2.5 as seen in the IMPROVE fine soil mass concentration. However, where we focus on emissions from urban mobile sources and secondary aerosol, we anticipate that the majority of the mass be attributable to sizes relevant to CCN. While we acknowledge that the mixing state and composition may be function of size within the CCN relevant range, we feel that it still offers some insight into CCN variability. Future efforts in the region should consider the use of measurements of size resolved composition. We feel that no additional text is required in the manuscript to address this comment.
On the competition among aerosol number, size and composition in predicting CCN variability: a multi-annual field study in an urbanized desert

Ewan Crosbie¹, Jong-Sang Youn³, Brian Balch³, Anna Wonaschütz³, Taylor Shingler³, Zhen Wang³, William C. Conant¹, Eric A. Betterton¹, Armin Sorooshian¹,²

1. Department of Atmospheric Sciences, University of Arizona, Tucson, AZ, USA.
2. Mel and Enid Zuckerman College of Public Health, University of Arizona, Tucson, AZ, USA.
3. Department of Chemical and Environmental Engineering, University of Arizona, Tucson, AZ, USA.
4. University of Vienna, Faculty of Physics, Vienna, Austria.

Correspondence to: A. Sorooshian (armin@email.arizona.edu)

Abstract

A two-year dataset of measured CCN concentrations at 0.2% supersaturation is combined with aerosol size distribution and aerosol composition data to probe the effects of aerosol number concentrations, size distribution and composition on CCN patterns. Data have been collected over a period of two years (2012-2014) in central Tucson, Arizona: a significant urban area surrounded by a sparsely populated desert. Average CCN concentrations are typically lowest in spring (233 cm⁻³), highest in winter (430 cm⁻³) and have a secondary peak during the North American Monsoon season (July to September; 372 cm⁻³). There is significant variability outside of seasonal patterns with extreme concentrations (1% and 99% levels) ranging from 56 cm⁻³ to 1945 cm⁻³ as measured during the winter, the season with highest variability.

Modeled CCN concentrations based on fixed chemical composition achieve better closure in winter, with size and number alone able to predict 82% of the variance in CCN concentration. Changes in aerosol chemical composition are typically aligned with changes in size and aerosol number, such that hygroscopicity can be parameterized even though it is still variable. In summer, models based on fixed chemical composition explain at best only 41% (pre-monsoon) and 36% (monsoon) of the variance. This is attributed to the effects of secondary organic aerosol (SOA) production, the competition between new particle formation and condensational growth, and the complex interaction of meteorology, regional and local emissions, and multi-phase chemistry during the North American Monsoon. Chemical composition is found to be an important factor for improving predictability in spring and on longer timescales in winter.

Parameterized models typically exhibit improved predictive skill when there are strong relationships between CCN concentrations and the prevailing meteorology.
and dominant aerosol *physicochemical processes*, suggesting that similar findings could be possible in other locations with comparable climates and geography.

1. Introduction

The influence of atmospheric aerosol particles on cloud properties and the consequential changes in radiative forcing carry the largest source of uncertainty in climate change prediction (IPCC, 2013). Cloud condensation nuclei (CCN) are the subset of aerosol particles that activate into droplets at a given supersaturation and their concentration therefore *contributes to governing* the microphysical and optical properties of clouds (Twomey, 1977; Albrecht, 1989). The global spatial and temporal variability of CCN concentrations consequently hold significant weight in predicting the droplet distribution in clouds and the ensuing microphysical and radiative properties (McFiggans et al., 2006; Andreae and Rosenfeld, 2008). Ultimately, CCN have been found to be a major factor in modulating cloud dynamics in both clean and polluted environments, with direct consequences on the hydrological cycle (Andreae et al., 2004; Altaratz et al., 2008; Stevens and Feingold, 2009).

While laboratory experiments involving the activation of single salt species (e.g., ammonium sulfate) or simple mixtures of organic compounds have offered satisfactory experimental validation (e.g., Brechtel and Kreidenweis, 2000) of the original underlying physical theory of droplet activation (Köhler, 1936), the extension to ambient atmospheric aerosol has proven more elusive (Covert et al., 1998; Chuang et al., 2000; Roberts et al., 2002; McFiggans et al., 2006; Ervens et al., 2010). Recent field studies (e.g., Broekhuizen et al., 2006; Dusek et al., 2006; Ervens et al., 2007; Hudson, 2007; Cubison et al., 2008; Quinn et al., 2008; Ervens et al., 2010; Burkart et al., 2011), spanning a range of aerosol scenarios, have not yet provided a comprehensive agreement on the relative importance of factors which affect CCN and the cloud droplet number, namely the following: the aerosol number, size distribution, composition, supersaturation and aerosol mixing state (Lance et al., 2004; Rissman et al., 2004; McFiggans et al., 2006; Andreae and Rosenfeld, 2008; Partridge et al., 2012).

During cloud formation, the supersaturation is driven by a combination of the aerosol related properties and dynamics (i.e., the updraft velocity) and therefore a complete description of the cloud system involves a two-way coupling of aerosol microphysics with circulation dynamics (Feingold, 2003). Modeling studies have shown that typically the supersaturation adjusts to large changes in aerosol properties (i.e., number, size and composition) to dampen the resulting variability observed in cloud droplet number concentration (Feingold, 2003); however, it has also been found that the distribution of CCN can have a significant impact on the cloud microphysics by affecting the droplet distribution (Feingold et al., 1999; McFiggans et al. 2006). The dynamics of initial droplet growth are affected by CCN properties (Feingold and Chuang, 2002; Raymond and Pandis, 2002, 2003; Chuang,
2003) and interstitial gas chemistry (Nenes et al., 2002; Lim et al., 2005) affecting
gas-particle partitioning through cloud processing.

Excluding the environmental factors that regulate supersaturation and droplet
growth kinetics, and focusing only on aerosol related properties that drive the initial
activation, yields important information relating to hygroscopicity. CCN closure
studies typically attempt to model the CCN concentration from measured aerosol
number, size and composition and then compare the modeled CCN to direct
measurements under a controlled set of supersaturated conditions (e.g., Dusek et al.,
2006; Ervens et al., 2007; Cubison et al., 2008; Bougiatioti et al., 2009; Lance et al.,
2009; Ervens et al., 2010; Jurányi et al., 2011, Martin et al., 2011; Levin et al., 2012;
Moore et al., 2012; Latham et al., 2013; Wu et al., 2013; Almeida et al., 2014). The
respective importance of composition and size distribution on CCN activation
remains an outstanding question. Closure studies have generally been successful for
background and remote sites (e.g., Jurányi et al., 2010), but less so in urban areas
(e.g., Burkart et al., 2012). The complexity of the aerosol composition and variability
in the aerosol mixing state are often the explanation for unsatisfactory closure.
under assumptions of bulk hygroscopic properties (Cubison et al., 2008; Ervens et
al., 2010). The single hygroscopicity parameter κ-Köhler Theory (Petters and
Kreidenweis, 2007, 2008) provides a theoretical framework to derive bulk
hygroscopicity for internal mixtures based on a volume weighted mixing rule. While
this simplicity is advantageous for closure models, this approach may not be
suitable for particles with complex morphology (e.g., Dusek et al., 2011; Hersey et
al., 2013).

Physical aging processes such as coagulation and condensational growth tend to
shift the aerosol population towards a more uniform mixing state when compared
to fresh emissions (Covert and Heintzenberg, 1993; Ervens et al., 2010). While
condensational growth processes increase CCN concentration by growing ultrafine
particles into the critical range for droplet activation, coagulation may result in
either increasing or decreasing CCN concentration since increased size comes at the
expense of aerosol number (Ripininen et al., 2011). Uncertainties in nucleation rates
and primary emissions have been shown to have significant impacts on global
estimates of CCN concentration (Pierce and Adams, 2009).

The study of CCN activation within an urban environment offers unique
opportunities to address the challenges associated with the inhomogeneity of
sources and aerosol aging, which gives rise to difficulties in predicting water uptake
behavior. Field studies purporting to quantify the influences of aerosol number, size
and compositional factors on CCN activity are often carried out over a limited, but
intense, period and hence offer a worthy characterization of the duration of the study
but perhaps lack climatological context, even related to sub-seasonal
variability. The current study addresses the two aforementioned issues by
reporting on long-term measurements of CCN, submicron size distributions and
composition taken jointly over multiple years in an urban area, specifically Tucson,
Arizona.

Ewan Crosbie 5/6/15 3:19 PM
Deleted: s

Ewan Crosbie 5/6/15 3:23 PM
Deleted: Weingartner et al., 2002;
Ewan Crosbie 5/6/15 3:19 PM
Deleted: 07; Andreae and Rosenfeld, 2008

Ewan Crosbie 4/18/15 3:35 PM
Deleted: -

Ewan Crosbie 4/18/15 3:35 PM
Deleted: The degree to which compositional and
compositional and mixing state variation
compressive properties of an
aerosol population in polluted air masses
seems to be strongly dependent on location and proximity to emissions (Covert
and... ... [1]

Ewan Crosbie 5/8/15 10:08 AM
Deleted: however, its simplicity is
advantageous for closure models

Ewan Crosbie 5/8/15 10:08 AM
Deleted: however, its simplicity is
advantageous for closure models

Ewan Crosbie 5/8/15 10:08 AM
Deleted: however, its simplicity is
advantageous for closure models

Ewan Crosbie 5/8/15 10:08 AM
Deleted: however, its simplicity is
advantageous for closure models
Tucson is located in the heart of the Sonoran Desert in the semi-arid southwestern United States. This location offers some unique opportunities for the study of CCN activation, primarily since there have been comparatively fewer documented measurements of CCN in arid regions. In addition, southern Arizona is situated in the region affected by the North American Monsoon (NAM) and as a result the highest monthly rainfall occurs during July and August and is accompanied by a strong influx of tropical moisture. The onset of the NAM in late June or early July leads to a rapid change from very hot and dry pre-monsoon conditions to the humid conditions associated with the monsoon and leads to changes in the aerosol properties (Sorooshian et al., 2011; Youn et al., 2013). Aside from the NAM, southern Arizona is situated in a relatively stable synoptic weather pattern, which gives rise to generally clear skies and light surface winds. The strong insolation produces a deep convective boundary layer in the afternoon and clear conditions lead to significant nocturnal cooling which together produce a significant but predictable diurnal cycle in temperature, humidity and convective boundary layer mixing.

The paper is subdivided as follows: (i) experimental methods and data collection are provided in Section 2; (ii) an overview of the "climatological" results is given in Section 3; (iii) the influence of size distribution and its relationship with composition is discussed in Section 4; (iv) CCN closure analysis is presented in Section 5; and (v) conclusions are presented in Section 6.

2. Data and Methods

2.1 Tucson Aerosol Characterization Observatory (TACO)

The study site is located at a rooftop location (approximately 30 m above ground) on the University of Arizona campus (32.229°N, 110.9538°W, 720 m ASL) in central Tucson (metro population ~1 million; U.S. Census Bureau, 2011). The sample inlet was located at rooftop level, approximately at the same height as nearby buildings, and 2 km northeast of downtown Tucson. The study period spanned more than two years (April 2012 – August 2014) and comprised long-term continuous measurements of CCN and related quantities, with a constant experimental setup.

2.2 Aerosol Instrumentation

Bulk CCN concentrations were measured using a CCN counter at fixed 0.2% supersaturation (CCN-100 Droplet Measurement Technologies; Roberts and Nenes, 2005). Particle size-resolved number concentrations were obtained using a scanning mobility particle sizer (SMPS 3080, TSI Inc.) coupled to a condensation particle counter (CPC 3772, TSI Inc.). The SMPS operated at 10:1 sheath-to-sample flow ratio and with a mobility diameter range from 13-748 nm. The integration of the size-resolved data over the entire range provided a measure of total...
condensation nuclei (CN). The CCN counter was calibrated twice during the study period using the method described in Rose et al. (2008) and exhibited a supersaturation of 0.192% ± 0.005% at the nominal 0.2% set-point value. A semi-continuous OC/EC analyzer (Sunset Laboratories Inc.) measured hourly organic carbon (OC) and elemental carbon (EC) concentrations in PM$_{2.5}$. Limits of detection were 0.2 µg/m$^3$ and 1.0 µg/m$^3$ for EC and OC, respectively. Water-soluble organic carbon (WSOC) was measured in PM$_{2.5}$ using a particle-into-liquid sampler (PILS, Brechtel Manufacturing Inc.) coupled to a total organic carbon analyzer (TOC; Sievers Model 800) (Sullivan et al., 2006; Duong et al., 2011; Wonaschütz et al., 2011). The overall measurement uncertainty associated with the reported WSOC concentrations is estimated to be approximately 10% with a limit of detection of 0.1 µg/m$^3$.

2.3 Local Meteorology

Collocated measurements of basic meteorological variables (including temperature, pressure, humidity, wind speed, wind direction and rainfall) were obtained at 5-second time resolution and archived as 1-minute and hourly averages. In addition, 1-minute direct normal irradiance (DNI) was obtained from the NREL Observed Atmospheric and Solar Information System (OASIS; http://www.nrel.gov/midc/ua_oasis/) site on an adjacent building on the university campus. SuomiNet GPS derived precipitable water vapor (PW) (Ware et al., 2000) data were obtained from the University of Arizona SA46 site (32.2298°N, 110.9539°W, 762 m ASL) resolved to 30-minute mean estimates. Finally, radiosonde data from the nearby National Weather Service were obtained from twice-daily balloon launches at 4 AM and 4 PM local time.

2.4 EPA IMPROVE

PM$_{2.5}$ aerosol composition measurements were obtained from two sites in the Inter-agency Monitoring of Protected Visual Environments (IMPROVE) network of filter samples (Malm et al., 1994). The Saguaro National Monument site (32.1742°N, 110.7372°W, 933 m ASL) is located within the foothills of the Rincon Mountains at the eastern extent of the Tucson metropolitan area and approximately 21 km east of TACO. The Saguaro West site (32.2486°N, 111.2178°W, 718 m ASL) is located on the western side of the topographically less prominent Tucson Mountains approximately 25 km west of TACO. 24-hour filter samples are collected at each site every three days. Data were obtained to coincide with as much of the study period as possible and were available up to December 2013 at the time of writing. Filter samples were analyzed for ions, metal and non-metal elements, and carbon (elemental and organic). Details on the extraction and analysis methodology are provided extensively elsewhere (http://vista.cira.colostate.edu/improve/Publications/IMPROVE_SOPs.htm). In addition to direct measurement, the IMPROVE network reports empirically derived
concentrations relevant to atmospheric aerosol including fine soil, sea salt, ammonium sulfate and ammonium nitrate (Malm et al., 1994).

2.5 Data Organization and Quality Control

All TACO data (CCN, SMPS, OC/EC and meteorology) are time synchronized and archived as averages at hourly increments. Sub-hourly variability in both the CCN concentration and the aerosol size distribution is highly influenced by localized intermittent sources, atmospheric turbulence and measurement related lags and noise. Since many of the metrics used in the interpretation of CCN variability involve ratios (or other non-linear functions) combining CCN and SMPS data, pre-filtering data to 1-hour reduces extraneous influences caused by sub-hourly covariance. All meteorological fields (except PW and radiosonde data) were additionally archived at 1-minute resolution. SMPS data from May and June 2013 are removed owing to sub-optimal data quality resulting from an instrument malfunction.

3. Climatological Results

3.1 Monthly and Seasonal Statistics

Monthly statistics of CN and CCN concentrations (henceforth referred to as CN and CCN) illustrate different trends as CN reveals a more stable annual cycle with minor reduction towards a minimum in June (Figure 1). CCN is more variable annually, and has two distinct peaks with a primary peak in December and a secondary peak in August. April has the lowest average CCN and also the lowest variability, as indicated by the interquartile range in Figure 1 for both CN and CCN. Conversely the interquartile range in CN for April is one of the highest, although in general CN exhibits significant sub-monthly variability when compared to the mean annual trends. OC and EC mass concentrations (Figure 1c) exhibit similar annual cycles, which suggests that aerosol related to urban combustion sources are ubiquitous; however, in summer the contribution is diluted by higher mixing heights (Figure 1f).

Seasonal temperature (T; Figure 1d), relative humidity (RH; Figure 1e) and direct normal irradiance (DNI; Figure 1f) illustrate the impact of the NAM on local meteorology, where strong increases in moisture are accompanied by slight temperature reductions and increased cloud cover.

Henceforth, data are grouped seasonally rather than monthly to analyze the annual cycle. Five seasons are defined to reflect the significant difference in meteorology between the pre-monsoon summer and the onset of the NAM. These are winter (W = DJF), spring (S = MA), pre-monsoon (PM = MJ), monsoon (M = JAS), and fall (F = ON). Table 1 provides a summary of seasonal CN and CCN statistics and includes only periods when both measurements are available. Winter and fall have the highest mean CN concentrations (~ 5200 cm⁻³), while pre-monsoon has the lowest with a mean just below 3900 cm⁻³. Extremes are quantified by 1% and 99% statistics and range between 749 cm⁻³ and 14406 cm⁻³ with winter showing the
highest variability. Average CN concentrations are typically lowest in spring (233 cm$^{-3}$), highest in winter (430 cm$^{-3}$) and have a secondary peak during the monsoon (372 cm$^{-3}$). Extremes in CN range between 56 cm$^{-3}$ and 1945 cm$^{-3}$ and winter variability far exceeds that of any other season.

Fine mode aerosol composition may help to explain the seasonal patterns in CCN and are illustrated using the IMPROVE data (Figure 2). Data are presented as an average of the two sites to the east and west of Tucson and can be interpreted as a suburban/semi-rural background reflecting regional scale aerosol composition onto which local urban sources are superimposed. Aerosol loading is highest during the pre-monsoon (PM) season, mainly due to the combined increase in the fine soil fraction, from windblown dust which occurs mainly in the spring and pre-monsoon seasons, and from the increase in sulfate during the pre-monsoon and monsoon (Sorooshian et al., 2013). Regional wildfire emissions are also most significant during pre-monsoon (Sorooshian et al., 2013). While dust particles may themselves act as CCN, they can also enhance the removal of CN and CCN by coalescence, while contributions from regional wildfire smoke may periodically enhance CN and CCN concentrations. Nitrate is more abundant in winter (~14%) compared to other seasons and may be a factor in the observed winter maximum in CCN concentrations. Sea salt contributes a modest fraction (~4.5%) of pre-monsoon aerosol when mid-tropospheric air originates mainly from the sub-tropical Pacific. The sum of the constituents presented in Figure 2 constitute between 93% and 101% of the seasonal average PM$_{2.5}$ as reported by gravimetric analysis.

The strong influence of urban sources on the fine mode carbonaceous aerosol in central Tucson is demonstrated by the elevated seasonal mean OC and EC mass concentrations at TACO versus the IMPROVE data (Table 2). This result is consistent with comparisons made by Sorooshian et al. (2011) for urban and rural sites in Arizona, which showed that carbonaceous mass concentrations varied strongly between urban and rural sites, whereas sulfate was more regionally homogenous.

3.2 Diurnal and Weekly Cycles

The diurnal cycle of CN illustrates a clear pattern involving a complex interaction of sources and sinks (Figure 3a). During weekdays, early mornings (7 AM to 9 AM) are characterized by traffic emissions, which increase the CN and EC concentrations (Figure 3d) indicative of fresh fossil combustion aerosol. Mean CN concentrations at 8 AM on weekdays (7925 cm$^{-3}$) are more than 160% of the equivalent weekend concentrations (4887 cm$^{-3}$). During the late morning, the convective boundary layer develops and dilutes the surface layer with relatively clean air from the free troposphere and/or residual layer leading to a marked drop in EC, OC (Figure 3d) and CN. Through the middle of the day, the convective boundary layer is still growing; however, a subtle reduction in the rate of decrease in CN (Noon to 2 PM) is suggestive of nucleation and growth of new particles which contribute as a source of CN. This is supported by the following: (i) concurrent enhancement in WSOC:OC.
By mid-afternoon (2 PM to 4 PM), the convective boundary layer reaches its peak depth and photochemical processes begin to slow down, leaving transport (vertical and horizontal) and coagulation as the dominant mechanisms, producing a net reduction in CN concentrations (Figure 3a) and increase in mean diameter (Figure 4b) while integrated aerosol volume concentration (used as a proxy for relative trends in PM$_{1}$) remains flat (Figure 4b). By late afternoon (4 PM to 6 PM) the convective boundary layer decouples from the surface and aerosol number and mass concentrations build again in the surface layer due to the evening peak in traffic emissions, with accompanying increases in EC and OC and reductions in mean diameter. During this time, secondary aerosol may still be influential once the boundary layer is decoupled, since residual ozone concentrations near the surface may still be sufficient to drive SOA production in the now thin surface layer.

The annualized diurnal cycle of CCN (Figure 3b) is less pronounced than that of CN mainly since CCN are typically unaffected by contributions from ultrafine particles with diameters less than 50 nm, which are highly variable. There is an increase in CCN during the evening, reaching a daily maximum at 10 PM and, interestingly, concentrations on weekends (429 cm$^{-3}$) are higher than on weekdays (380 cm$^{-3}$). There is a large range of CCN variability observed within each hour when compared to the hourly composite mean trend which is partially explained by the seasonal differences in the CCN diurnal cycle (Figure 3c). During winter, there is a significant diurnal cycle in CCN, while in other seasons the diurnal pattern is relatively flat. Due to reduced winter temperatures, semi-volatile organics are more likely to partition to the particle phase, which may incrementally shift the size distribution of freshly emitted particles associated with morning traffic towards larger sizes. In addition, nitrate also forms a larger component of the regional aerosol than in other seasons, which helps to increase the hygroscopicity and to reduce the diameter required for droplet activation. Both factors likely work in tandem with the diurnal emissions cycle, which results in a CCN pattern which more closely follows CN than other seasons. The other notable feature is that the peak CCN concentration occurs during the night in winter while it occurs during the afternoon in summer. In addition to partitioning of semi-volatiles, emissions from domestic wood burning are another potential contributor to CCN in the winter, while in summer it is likely SOA production, driven by photochemistry and moisture during the day (Youn et al., 2013).

A bulk hygroscopicity parameter ($\kappa$) is derived using the method of Petters and Kreidenweis (2007) and by assuming total activation above a critical activation diameter, such that the CCN concentration exactly matches the concentration of particles exceeding this critical diameter (Furutani et al., 2008; Burkart et al., 2011;
Wonaschütz et al., 2013). Hygroscopicity decreases concurrently with the morning traffic signature (Figure 4a) and then rebounds through the day to produce a peak between 2 PM and 4 PM matching expectations of organic aging and condensation growth by photochemically oxidized organics and sulfate. As expected, the morning minimum is less extreme on weekends ($\kappa = 0.15$) compared to weekdays ($\kappa = 0.10$) due to reduced traffic and this trend remains through the day with weekend maxima ($\kappa = 0.21$) exceeding weekday values ($\kappa = 0.19$). During the evening and night, the offset is far smaller ($\Delta\kappa \approx 0.005$). The $\kappa$ parameter tracks the diurnal pattern of activation ratio (Figure 4a), defined as the ratio of CCN to CN, which on first glance, together with the rather modest changes in mean aerosol diameter (Figure 4b), would indicate that chemical composition is driving the CCN variability at least on diurnal scales. However, two corollaries should be highlighted: a) the mean aerosol diameter is a rather simplistic representation of changes in the size distribution, and b) as mentioned earlier, the majority of the CCN variability is not described by composite mean hourly trends, at least in an annual sense, and thus, as will be examined in the forthcoming section, a more rigorous treatment of the size distribution is needed to better explain overall CCN variability.

4. Size distribution

Several studies (e.g., Conant et al., 2004; Dusek et al., 2006; Ervens et al., 2007) have suggested that the size distribution alone can explain CCN variability, however there are other examples (e.g., Hudson 2007; Burkart et al., 2011), which refute this particularly in cases where the aerosol is externally mixed. If the physical and chemical processes which govern size and composition changes are intrinsically tied to a single governing mechanism, a parameterization involving one component may suitably capture the variability in the other, at least when considering a fixed supersaturation. Furutani et al. (2008) reported the activation diameter to be well correlated with activation ratio during a ship-borne study in the eastern North Pacific, suggesting compositional changes as a result of aging (where size also increases) to be the major driver for CCN variability. In contrast, Burkart et al. (2011) examined the same relationship but found poor correlation between activation ratio and activation diameter in Vienna, Austria, suggesting a more complex relationship between size and composition.

The shape of the size distribution can be used to interpret physical processes (e.g., condensation, evaporation, nucleation, coagulation), while relative changes in CN concentration, combined with changes in shape, offer insight into atmospheric processes (e.g., advection and diffusion) and emissions. The well-established “K-means” clustering algorithm (Hartigan and Wong, 1979; Lloyd, 1982) was used here as a statistical tool to group size distributions by shape. The method was implemented with four clusters and the resulting four cluster centroids denoted archetypal size distribution shapes (Figure 5), to which the observations were assigned, according to their degree of association. The selection of four clusters struck a balance between capturing the salient patterns, while maintaining
simplicity; however, we do not claim that this choice was optimal for all applications. Cluster associations were “fuzzy”, and therefore an observation could be partially assigned to multiple clusters to reflect the continuity of transitions between clusters in the dataset. This provides the added advantage that smooth transitions in cluster properties can be represented without the additional complexity of defining intermediate clusters. A full description of the clustering method and the method by which associations are made is provided in Appendix A.

The mean diurnal cycle of cluster associations (Figure 5) and their mean properties (Table 3) provide a physical description of the clusters and are hereafter given the following identifiers, which are indicative of the physical process or ‘regime’ that is suggested by the cluster properties: nucleation (N), fresh fossil (FF), winter/nocturnal (WN), and coagulation/condensation (CC).

Winter (W) and summer (PM and M) exhibit substantially different patterns in cluster associations on diurnal scales, while the transition seasons (S and F) contain features of both winter and summer and are therefore more mixed in terms of the driving mechanisms. During winter (W), large swings in the size distribution shape are uncommon; however, with activation at 0.2% supersaturation occurring at diameters as low as 100 nm, the growth that accompanies a shift from FF to WN is sufficient to significantly increase the activation ratio. Unlike other seasons, it is likely that the main driver for size distribution changes occurring during winter is the equilibrium partitioning of semi-volatile species between gas and particle phase (e.g., nitrate). An additional contributor may result from the offset in emissions patterns between traffic (day) and domestic wood burning (night). Anomalously colder or more humid conditions tend to result in larger and more hygroscopic particle distributions and are typically also associated with more stable near-surface conditions leading to suppressed mixing and higher aerosol loading as seen in the WN CN, EC and OC concentrations (Table 3). In the extreme, the infrequent winter occurrence of the CC cluster is merely an extension of this trend occurring during the coldest winter nights where average hygroscopicity reaches κ=0.23 and average CCN concentrations are 811 cm⁻³. The fact that number, size and hygroscopicity tend to act in association is perhaps the reason why CCN variability is highest in winter on both synoptic and diurnal scales.

Conversely, in summer (PM and M) the shape of the size distribution is very variable and exhibits large swings between N and CC clusters (Figure 5). After primary emissions associated with the morning traffic peak (FF cluster) have been diluted through boundary layer mixing, competition between the N and CC cluster takes over. Unlike winter, there is no monotonic relationship between meteorology and size. Instead, hotter conditions with higher solar exposure tend to bifurcate the size distribution more between N and CC clusters with cooler and cloudy conditions favoring the retention of the intermediate FF or WN clusters. This suggests that the N and CC clusters are partially driven by photochemically produced secondary aerosol. Higher temperature and stronger direct normal irradiance (DNI) are likely coupled with higher hydroxyl concentrations, and ozone concentrations are typically 30-40% higher for N and CC clusters (Table 3), which accelerates the
production of reduced volatility oxidized organic vapors from precursor volatile organic compounds (VOCs). The partitioning of these vapors between condensation on existing particles and nucleation of new particles is likely a function of the aerosol surface area and the production rate of the low-volatility organics. Anomalously dry conditions are a feature of the N cluster, suggestive of reduced aerosol water reducing the available surface area. Another possible mechanism affecting the N cluster during the summer (PM and M) is the evaporation, or lack of condensation, of semi-volatile organic compounds associated with traffic emissions (Robinson et al., 2007) such that the FF cluster takes on some of the features of the N cluster. This mechanism would be supported by the anomalous contribution of EC to the N cluster during the PM and M seasons. Further analysis of the aerosol and gas phase composition is needed, before and during the monsoon, in order to fully understand the balance of regional and local processes in driving the preference of N and CC clusters.

Tucson often is under the influence of very light mean surface winds and so during the day, the predominant mechanism for ventilation of urban aerosol is through vertical mixing of the convective boundary layer, which is supported by measurements at a nearby mountain site (Shaw, 2007). Furthermore, the climatological mesoscale surface wind pattern, particularly in summer, is light southeasterly winds during the night and morning, followed by northwesterlies in the afternoon and evening, induced by regional topography (Philippin and Betterton, 1997). It is therefore possible for urban aerosol particles and precursor gases to be recycled over the site during the course of the day, through both these mechanisms. Processes which control the cluster associations may be also dependent on regional (e.g., nucleation of biogenic SOA) as well as local effects (e.g., recycling of urban emissions), which happened at an earlier time. The complex influences of this ‘memory effect’, together with the interaction of meteorology and emissions may be one of the contributing factors which cause evening and overnight CCN concentrations to be higher on weekends (Figure 3b).

5. CCN closure

Studies aimed at achieving a predictive model of CCN concentrations from measured number, size and composition (i.e., CCN closure) have shown mixed ability to predict CCN concentrations across a range of aerosol scenarios. To examine these dependencies, in the context of the present study, we consider the effect that simplifying assumptions have on the ability to predict CCN. Traditionally, closure studies aim to predict the hygroscopic properties from measured composition or sub-saturated growth factors, which are then combined with size distribution measurements to predict CCN (e.g., Ervens et al., 2010). With this method the inter-comparison of various scenarios, and the resulting degree to which CCN concentrations are predicted, is affected by both the model assumptions and the accuracy by which aerosol physicochemical properties are measured. Our focus here is to study the degree of CCN variability explained by incremental simplifications in a predictive model considered across a range of timescales. One
Major simplification is the limitation of the treatment of hygroscopicity to a bulk measurement, which is permitted to vary temporarily but does not isolate size dependent changes in hygroscopicity nor the hygroscopicity distribution, which may be an important component in relation to external mixing. These aspects are beyond the scope of these parameterizations and are likely to contribute to model shortfalls. Forthcoming work will separately study the degree of correspondence of hygroscopicity between the sub- and supersaturated regimes, size-dependent hygroscopicity and composition, and the closure of hygroscopicity from composition measurements.

Seven, highly simplified, predictive models are used to estimate CCN over the entire study period: (i) constant CCN (baseline); (ii) constant activation ratio (assesses the effect of number only); (iii) constant hygroscopicity (effect of number and size distribution); (iv) constant size distribution (effect of number and hygroscopicity); (v) measured number with size distribution shape and hygroscopicity derived from cluster associations; (vi) measured size and number with cluster derived hygroscopicity; and (vii) all parameters (a reconstruction, for reference only). The inclusion of models (v) and (vi) assesses whether the predictive skill can be improved by the use of a reduced order representation of the size distribution and hygroscopicity parameter (κ). Models (v) and (vi) can be considered an incremental refinement to models (ii) and (iii) where the assumption is that there is prior knowledge of expected cluster properties and associations.

Predicted CCN concentrations are compared to those measured and two performance metrics are evaluated: (i) "percentage variance explained" (VE) metric, which is the variance of the measured CCN explained by the model as determined by mean square residuals; and (ii) a "normalized mean error" (NME) metric, defined as the root-mean-square residual between modeled and measured CCN concentrations expressed as a percentage of the mean measured CCN concentration for the epoch. While both these metrics are connected, the VE is a better descriptor of the specific performance of the model, whereas the NME puts the model in the context of overall predictability. Models are first tested using (i) the cumulative dataset and (ii) for the five predefined seasons with model parameters set using seasonal best-fit values. The models (except (v) and (vi)) are then tested, using the same methodology, on data that have been filtered using a 24-hour running average and seven day average, with the underlying motivation to determine if environmental factors which control CCN predictability differ between diurnally and synoptically driven timescales.

The results (Table 4) show that when all seasons are considered, a constant hygroscopicity assumption explains more of the measured variance (~63% VE) than a constant size distribution (~44% VE) suggesting that overall the size distribution is generally a more important driver for CCN variability than composition. However, the goodness-of-fit (VE) is far lower than that presented by Dusek et al. (2006) and is probably associated with the complexity of the aerosol mixing state and spatiotemporal variability in composition, due to the proximity of
the TACO site to fresh emission sources as compared to the Dusek et al. (2006) study site. To put the TACO results in more context, fresh pollution aerosol in urban areas such as Riverside and Houston could not be fully represented without knowledge of size-resolved composition (Cubison et al., 2008; Ervens et al., 2010). A number of other studies have shown that mixing state can help improve predictive capability of CCN behavior (Wex et al., 2010), including Atlanta (Padro et al., 2012) and during early morning rush hour near Mexico City (Lance et al., 2013); but studies also report that hydrophobic particles emitted in urban areas quickly (~ few hours) become internal mixtures via condensation of secondary hygroscopic species (e.g., Wang et al., 2010; Mei et al., 2013).

In the daily and weekly filtered cases, the relative balance between size and composition is also similar. Using the submicron number concentration as a predictive model for CCN (i.e., a constant activation ratio assumption) performs poorly in all annual cases (and all seasonal cases except winter) since it is strongly affected by variability in nucleation and small Aitken mode particles from fresh emissions that do not contribute to CCN at the supersaturation levels considered here.

Compared to other seasons, the simplified predictive models perform the best in winter in terms of VE, however, this season also has far higher variability in CCN than any other season across the three timescales considered. Winter is also the only season where a constant activation ratio assumption offers any skill in CCN predictability suggesting that the modulation of CCN is more tied to bulk aerosol sources and sinks than compositional or size dependent changes or that these processes are strongly interlinked. Winter aerosol is mainly controlled by an interplay of urban emissions balanced by transport and mixing such that there is a strong correlation between the diurnal cycle of CN and EC, which serves as a combustion tracer. Strong nocturnal surface inversions, in conjunction with a lack of surface wind induced mixing, trap urban emissions close to the surface before the convective boundary layer develops, which happens later in the day than other seasons. Intermittent synoptic scale influences, such as frontal passages, affect aerosol sinks directly through wet scavenging, although this effect is presumably much weaker than less arid regions, and drive regional transport in the lower troposphere, which ventilates the urban plume. Synoptic systems affect column stability, which indirectly affects aerosol loading by regulating the extent of diurnally driven vertical mixing. Chemical aging processes and photochemically driven secondary aerosol formation are suppressed in winter compared to other seasons simplifying the diurnal changes in hygroscopicity and size distribution, although size and hygroscopicity appear to be tied to the diurnal cycle through temperature changes. Both size (constant κ, Model (iii)) and hygroscopicity (constant size distribution, Model (iv)) simplified models explain 82% and 73% of the CCN variance, respectively, reiterating that size and hygroscopicity changes are strongly coupled. The weekly filtered data indicate that hygroscopicity becomes marginally more influential than size changes over longer timescales and is perhaps
a consequence of regional sources associated with long-range transport competing with local emissions.

Regional scale transport is also an important feature of spring, which is a transition season where mid-latitude meteorology still affects the region, boundary layer mixing becomes more vigorous and surface winds are strongest on average. Dust loading is highest and temperature changes on diurnal and synoptic scales are also greatest which affects the partitioning of semi-volatile species (e.g., nitrate). The complex mixing state and highly variable aerosol composition makes CCN prediction difficult as reflected in the poor performance of the simplified models. The modeled predictability indicates that composition is far more important than size during spring and in fact the daily-filtered data suggests that using the size distribution (Model (iii)) to predict CCN is worse than assuming a constant seasonal average concentration, indicative of complex aerosol mixing states, morphology and scale-dependent mechanisms.

The pre-monsoon summer reveals a steady improvement in the model performance towards longer timescales (i.e., weekly) and the increasing relative importance of hygroscopicity. Intense solar radiation during this season increases the importance of VOC and SO$_2$ chemistry to form secondary aerosol species. Aerosol number may be strongly influenced by nucleation and therefore knowledge of the size distribution becomes essential on sub-diurnal scales. Over longer timescales all simplified approximations become reasonable suggesting a more stable meteorological pattern, which is typical of this season: as the jet migrates northward, synoptic steering becomes lighter and the circulation pattern becomes more driven by mesoscale circulations. The increased importance of hygroscopicity on timescales longer than a week is perhaps indicative of the influence of wildfire smoke and intermittent regional dust transport which periodically affect southern Arizona during this season.

The monsoon season exhibits the poorest performance of the simplified models out of all seasons, which is perhaps expected given the very complex meteorological pattern and the interplay between secondary aerosol production at the regional (e.g., biogenic SOA and sulfate) and local scale (e.g., urban SOA). Knowledge of the size distribution is essential since it is highly variable across all scales driven by both meteorological influences, in the form of monsoon thunderstorms, and secondary aerosol processes. Even considering size variability alone does not yield very satisfactory results implying that aerosol composition is very closely tied to changes in size distribution during the monsoon season. However, CCN variability is also lowest of all seasons, while the mean CCN concentration is relatively high implying partial cancellation in the effects caused by changes in size, number and composition. The consequence is that the NME metric is actually lowest in monsoon when a constant hygroscopicity model is used, which is the opposite of the situation during winter. Fall shows the opposite pattern to spring and pre-monsoon in that hygroscopicity has decreasing influence over longer timescales, and for the weekly
filtered case, the constant hygroscopicity model provides a very satisfactory model of CCN variability.

The inclusion of the cluster associations to estimate $\kappa$ (Model vi) provides an incremental improvement in the predictive skill (+3% to +15% additional %VE) when compared to a seasonally constant $\kappa$ (Model iii), with the exception of the pre-monsoon summer season, where a reduction in %VE was observed (~7%). Annually, the increase was approximately +5% on %VE. The comparison between the cluster-derived activation ratio (Model v) and a constant activation ratio (Model ii) was far more significant with an annual increase of +59% on %VE suggesting that a low-order representation of the size distribution shape, where other data is unavailable (e.g., from remote sensing methods), may offer a worthwhile improvement to the estimation of CCN concentration.

6. Conclusions

This study investigates the respective importance of aerosol number concentration, size distribution and composition in driving CCN variability in Tucson, Arizona. In doing so, a long-term characterization of the seasonal, weekly and diurnal patterns in aerosol number concentration, size distribution and selected particle speciation has been achieved. Seasonally, the average CN concentration exhibits a moderate trend towards a minimum during summer, while CCN concentrations exhibit significant winter and summer peaks. Weekday and weekend CN concentrations track the respective diurnal weekday and weekend EC and OC mass concentrations, indicating a strong influence of local combustion aerosol, predominantly from vehicle emissions but also, in winter, from domestic biomass burning. Activation ratio and hygroscopicity, as estimated by $\kappa$, track the morning peak in fossil fuel emissions, by concurrently showing a marked reduction, particularly on weekdays. This helps to support the notion that CCN concentrations are not significantly enhanced by fresh fossil emissions. The effects of local emissions are typically offset by those of boundary layer mixing; however, during the warmer and more photochemically active seasons, secondary aerosol processes become more influential.

During winter, the interplay between chemistry and dynamics is such that increasing size is accompanied by increasing hygroscopicity. This occurs most commonly at night and during anomalously cold periods, when boundary layer mixing is suppressed and aerosol loading is high, thus increasing CCN concentrations. Conversely, during the day and particularly during anomalously warm and dry periods, there is sufficient convective mixing to dilute the aerosol, evaporate hygroscopic semi-volatile species and generally promote the abundance of smaller particles, reducing CCN concentrations. The combined result of these effects is to increase the variability in CCN, since each of these contributing factors act together to enhance or suppress CCN concentrations. The added consequence is
that simplified models offer substantial predictive skill for CCN variability even though the observed changes in the size distribution are relatively subtle.

The summer is divided by the arrival of the North American Monsoon (July - September), which rapidly increases the abundance of moisture compared to the very hot and dry months that precede it (May – June). Secondary production of sulfate and organics becomes more influential during both summer seasons, and photochemically produced aerosol appears to be the mechanism responsible for an afternoon maximum in CCN concentration, compared to a nocturnal maximum in winter. The diurnal cycle of the boundary layer follows a similar pattern to other seasons, except that mixing heights are generally higher and nocturnal surface inversions are less pronounced, especially during the monsoon. While CN concentrations drop off during the day similar to other seasons, CCN concentrations remain relatively more stable indicating that condensed SOA and sulfate play a significant role in offsetting the loss in CCN caused by dilution.

Another important feature of the summer is the bifurcation in the size distribution shape, where the pattern swings back and forth from (i) an abundance of ultrafine particles that are potentially tied to a nucleation event to (ii) a deficiency of Aitken mode particles, and a growth in the number of particles larger than 100 nm that are more in line with a background aerosol population. While the meteorological conditions favoring both regimes are similar and likely explained by SOA and sulfate production, the mechanisms responsible for the bifurcation are still unclear. Possible mechanisms include aerosol water uptake, leading to increased aerosol surface area for condensation, which is supported by lower humidity on days when ultrafine particles are present, particularly before the monsoon. During the monsoon, regional biogenic SOA produced as a result of increased vegetation may explain the periodic import of small SOA particles into the urban plume. Finally, the role of the monsoon thunderstorms may also be responsible for erratic changes to the size distribution simply through the sporadic disruption of the local and regional circulation pattern.

The sensitivities of CCN concentration to changes in aerosol number, size and composition can be well represented in a theoretical framework as described by Köhler Theory and its various refinements. However, the extent to which these driving components vary, and the mechanisms through which they interact, is the primary limitation in consolidating parametric representations suitable for predictive models. Achieving satisfactory CCN closure using measurements of chemical composition and size has generally been most successful with background aerosol where substantial changes in composition are dampened by aging processes. However, the results of this study suggest that in certain regimes (e.g., during winter), where composition, size and number concentration have a more deterministic relationship, there are still opportunities for parametric simplifications to be successful even when chemical processes are relatively complex. Since the relationship can be explained by somewhat broad environmental mechanisms not entirely specific to Tucson, similar conclusions can
be drawn for other urban areas with comparable geographical and climatological settings.

Future work will focus on the predictability of κ using measurements of composition, patterns in the environmental conditions (e.g., emissions, meteorology and other auxiliary measures), and sub-saturated aerosol hygroscopicity with the primary goal being to determine if a single-parameter representation of CCN activation is suitable for this environment. In addition, we will focus on addressing the factors which control the summertime size distribution bifurcations and the extent to which they are influenced by biogenic and anthropogenic SOA production pathways.

Acknowledgements:

This research was supported in part by Grant 2 P42 ES04940-11 from the National Institute of Environmental Health Sciences (NIEHS) Superfund Research Program, NIH, the University of Arizona Foundation, the Institute of the Environment at the University of Arizona, and the Center for Environmentally Sustainable Mining through TRIF Water Sustainability Program funding. Funding is also acknowledged from NASA grants NNX14AK79H, NNX12AC10G, and NNX14AP75G. Glenn Shaw is acknowledged for helpful suggestions during the preparation of the manuscript. We acknowledge the sponsors of the IMPROVE network, the Pima County Department of Environmental Quality, NREL, UCAR and SuomiNet for data used in this study.

References


Petters, M. D., Prenni, A. J., Kreidenweis, S. M., DeMott, P. J., Matsunaga, A., Lim, Y. B., and Ziemann, P. J.: Chemical aging and the hydrophobic-to-hydrophilic conversion of


Table 1: Seasonal mean and extreme CN and CCN concentrations from hourly averaged data. Seasons are defined as follows: winter (W = DJF), spring (S=MA), pre-monsoon (PM = MJ), monsoon (M = JAS), fall (F = ON).

<table>
<thead>
<tr>
<th>Concentration (cm$^{-3}$)</th>
<th>W</th>
<th>S</th>
<th>PM</th>
<th>M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5189</td>
<td>4853</td>
<td>3872</td>
<td>4200</td>
<td>5200</td>
</tr>
<tr>
<td>Max (99%)</td>
<td>14406</td>
<td>13799</td>
<td>10869</td>
<td>11606</td>
<td>13682</td>
</tr>
<tr>
<td>Min (1%)</td>
<td>749</td>
<td>686</td>
<td>807</td>
<td>1070</td>
<td>853</td>
</tr>
<tr>
<td>CCN$_{SS=0.2%}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>430</td>
<td>233</td>
<td>301</td>
<td>372</td>
<td>303</td>
</tr>
<tr>
<td>Max (99%)</td>
<td>1945</td>
<td>809</td>
<td>667</td>
<td>741</td>
<td>951</td>
</tr>
<tr>
<td>Min (1%)</td>
<td>56</td>
<td>59</td>
<td>101</td>
<td>100</td>
<td>81</td>
</tr>
</tbody>
</table>
Table 2: Seasonal mean OC and EC concentrations, and associated standard deviations, at the TACO and IMPROVE sites.

<table>
<thead>
<tr>
<th>Site</th>
<th>Concentration (µg m⁻³)</th>
<th>W</th>
<th>S</th>
<th>PM</th>
<th>M</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>TACO</td>
<td>EC</td>
<td>0.69 ± 0.66</td>
<td>0.38 ± 0.38</td>
<td>0.27 ± 0.36</td>
<td>0.40 ± 0.34</td>
<td>0.54 ± 0.46</td>
</tr>
<tr>
<td></td>
<td>OC</td>
<td>6.96 ± 3.40</td>
<td>5.05 ± 2.25</td>
<td>4.87 ± 1.98</td>
<td>4.40 ± 1.60</td>
<td>5.31 ± 2.20</td>
</tr>
<tr>
<td>SAGUARO NM</td>
<td>EC</td>
<td>0.15 ± 0.07</td>
<td>0.11 ± 0.05</td>
<td>0.10 ± 0.05</td>
<td>0.12 ± 0.04</td>
<td>0.13 ± 0.07</td>
</tr>
<tr>
<td></td>
<td>OC</td>
<td>0.51 ± 0.18</td>
<td>0.50 ± 0.17</td>
<td>0.63 ± 0.33</td>
<td>0.63 ± 0.27</td>
<td>0.45 ± 0.20</td>
</tr>
<tr>
<td>SAGUARO WEST</td>
<td>EC</td>
<td>0.22 ± 0.13</td>
<td>0.12 ± 0.06</td>
<td>0.11 ± 0.05</td>
<td>0.13 ± 0.04</td>
<td>0.18 ± 0.08</td>
</tr>
<tr>
<td></td>
<td>OC</td>
<td>0.61 ± 0.30</td>
<td>0.49 ± 0.17</td>
<td>0.74 ± 0.32</td>
<td>0.69 ± 0.28</td>
<td>0.55 ± 0.20</td>
</tr>
</tbody>
</table>
Table 3: Seasonally derived mean cluster properties and associated environmental conditions (AR = activation ratio). Meteorological variables (T, RH and direct normal irradiance (DNI)) are presented as anomalies, based on departure from hourly means for each month. Entries in parentheses indicate that the cluster occurs less than 15% of the time in that season. An asterisk (*) next to EC denotes a case when the concentration is below LOD. O₃ data are obtained from a surface pollutant monitoring site (~ 9 km from TACO) operated by the Pima County Department of Environmental Quality (Children’s Park Station).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Season</th>
<th>CN</th>
<th>CCN</th>
<th>AR</th>
<th>k</th>
<th>EC</th>
<th>OC</th>
<th>WSOCC</th>
<th>O₃</th>
<th>ΔT</th>
<th>ARH</th>
<th>ΔDNI</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>4985</td>
<td>249</td>
<td>0.864</td>
<td>0.17</td>
<td>0.75</td>
<td>7.20</td>
<td>0.23</td>
<td>25</td>
<td>0.85</td>
<td>-1.7</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>5161</td>
<td>198</td>
<td>0.858</td>
<td>0.13</td>
<td>0.36</td>
<td>5.35</td>
<td>0.18</td>
<td>32</td>
<td>-0.30</td>
<td>-0.8</td>
<td>-3</td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td>4935</td>
<td>278</td>
<td>0.867</td>
<td>0.12</td>
<td>0.10</td>
<td>5.32</td>
<td>0.20</td>
<td>33</td>
<td>-0.77</td>
<td>-1.4</td>
<td>-49</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>5536</td>
<td>360</td>
<td>0.884</td>
<td>0.15</td>
<td>0.46</td>
<td>5.09</td>
<td>0.32</td>
<td>29</td>
<td>-0.64</td>
<td>1.7</td>
<td>-41</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>7256</td>
<td>282</td>
<td>0.858</td>
<td>0.14</td>
<td>0.56</td>
<td>5.55</td>
<td>0.26</td>
<td>19</td>
<td>-0.16</td>
<td>-1.2</td>
<td>-15</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>6337</td>
<td>490</td>
<td>0.993</td>
<td>0.19</td>
<td>1.79</td>
<td>11.0</td>
<td>0.18</td>
<td>16</td>
<td>-0.42</td>
<td>-0.1</td>
<td>-4</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>4980</td>
<td>278</td>
<td>0.871</td>
<td>0.16</td>
<td>0.36</td>
<td>5.63</td>
<td>0.18</td>
<td>29</td>
<td>-0.25</td>
<td>2.0</td>
<td>-43</td>
<td></td>
</tr>
<tr>
<td>WN</td>
<td>4042</td>
<td>334</td>
<td>0.998</td>
<td>0.15</td>
<td>0.07</td>
<td>5.09</td>
<td>0.20</td>
<td>35</td>
<td>-0.72</td>
<td>0.5</td>
<td>-40</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>4302</td>
<td>392</td>
<td>0.106</td>
<td>0.16</td>
<td>0.40</td>
<td>5.33</td>
<td>0.34</td>
<td>29</td>
<td>-0.96</td>
<td>5.8</td>
<td>-60</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>7743</td>
<td>363</td>
<td>0.880</td>
<td>0.16</td>
<td>0.62</td>
<td>5.94</td>
<td>0.33</td>
<td>16</td>
<td>-1.06</td>
<td>0.9</td>
<td>-26</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>6203</td>
<td>811</td>
<td>0.153</td>
<td>0.23</td>
<td>1.08</td>
<td>9.15</td>
<td>0.27</td>
<td>18</td>
<td>-0.39</td>
<td>5.6</td>
<td>-13</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>2659</td>
<td>267</td>
<td>0.124</td>
<td>0.18</td>
<td>0.17</td>
<td>4.63</td>
<td>0.19</td>
<td>44</td>
<td>1.30</td>
<td>-0.3</td>
<td>-21</td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>2412</td>
<td>349</td>
<td>0.166</td>
<td>0.15</td>
<td>0.99</td>
<td>5.03</td>
<td>0.28</td>
<td>46</td>
<td>0.41</td>
<td>1.3</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>M</td>
<td>2884</td>
<td>414</td>
<td>0.173</td>
<td>0.17</td>
<td>0.27</td>
<td>4.43</td>
<td>0.37</td>
<td>38</td>
<td>0.26</td>
<td>0.2</td>
<td>-24</td>
<td></td>
</tr>
<tr>
<td>F</td>
<td>3964</td>
<td>356</td>
<td>0.145</td>
<td>0.20</td>
<td>0.33</td>
<td>4.93</td>
<td>0.27</td>
<td>30</td>
<td>1.48</td>
<td>0.9</td>
<td>-25</td>
<td></td>
</tr>
</tbody>
</table>

Meteorological variables (T, RH and DNI) are presented as anomalies, based on departure from hourly means for each month.
Table 4: Closure model performance as quantified by variance explained (top) and normalized mean error (bottom). Models (i)-(iv) include holding constant either CCN, activation ratio (AR), κ, or size distribution (SD). Model (v) uses the cluster properties and associations (see Figure 5 and Table 3), Model (vi) uses the same assumptions as Model (iii) except that κ is determined from cluster associations, and Model (vii) is a reconstruction for reference only. A dash (·) indicates that the result is not available or performed so poorly it cannot be quantified by the metric.

<table>
<thead>
<tr>
<th></th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>(v)</th>
<th>(vi)</th>
<th>(vii)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Const. CCN</td>
<td>Const. AR</td>
<td>Const. κ</td>
<td>Const. SD</td>
<td>Clus. only</td>
<td>Clus. κ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>ALL</td>
<td>3.7</td>
<td>63.2</td>
<td>43.9</td>
<td>62.3</td>
<td>68.4</td>
<td>99.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>44.9</td>
<td>81.6</td>
<td>72.5</td>
<td>78.4</td>
<td>84.1</td>
<td>99.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>-</td>
<td>25.3</td>
<td>55.3</td>
<td>3.5</td>
<td>37.5</td>
<td>99.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>-</td>
<td>40.5</td>
<td>43.1</td>
<td>34.2</td>
<td>99.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>-</td>
<td>35.5</td>
<td>-</td>
<td>-</td>
<td>42.3</td>
<td>99.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-</td>
<td>40.3</td>
<td>31.1</td>
<td>3.4</td>
<td>54.6</td>
<td>99.4</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>(v)</th>
<th>(vi)</th>
<th>(vii)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Const. CCN</td>
<td>Const. AR</td>
<td>Const. κ</td>
<td>Const. SD</td>
<td>Clus. only</td>
<td>Clus. κ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Daily</td>
<td>ALL</td>
<td>6.1</td>
<td>70.0</td>
<td>47.0</td>
<td>-</td>
<td>-</td>
<td>99.4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>35.9</td>
<td>81.2</td>
<td>71.6</td>
<td>-</td>
<td>-</td>
<td>99.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>6.5</td>
<td>-</td>
<td>62.1</td>
<td>-</td>
<td>-</td>
<td>99.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>0.2</td>
<td>52.5</td>
<td>15.4</td>
<td>-</td>
<td>-</td>
<td>98.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>-</td>
<td>64.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>98.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-</td>
<td>59.9</td>
<td>17.8</td>
<td>-</td>
<td>-</td>
<td>98.1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>(i)</th>
<th>(ii)</th>
<th>(iii)</th>
<th>(iv)</th>
<th>(v)</th>
<th>(vi)</th>
<th>(vii)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Const. CCN</td>
<td>Const. AR</td>
<td>Const. κ</td>
<td>Const. SD</td>
<td>Clus. only</td>
<td>Clus. κ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weekly</td>
<td>ALL</td>
<td>7.1</td>
<td>67.7</td>
<td>43.3</td>
<td>-</td>
<td>-</td>
<td>99.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td>W</td>
<td>15.8</td>
<td>66.4</td>
<td>77.8</td>
<td>-</td>
<td>-</td>
<td>98.8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>6.0</td>
<td>33.7</td>
<td>74.1</td>
<td>-</td>
<td>-</td>
<td>98.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PM</td>
<td>45.4</td>
<td>72.9</td>
<td>75.8</td>
<td>-</td>
<td>-</td>
<td>96.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>M</td>
<td>-</td>
<td>43.9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>96.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>F</td>
<td>-</td>
<td>39.5</td>
<td>0.3</td>
<td>-</td>
<td>-</td>
<td>97.9</td>
<td></td>
</tr>
</tbody>
</table>

31
Figure 1: Monthly statistics of (a) CN, (b) CCN (0.2%), (c) OC and EC, (d) temperature, (e) RH, and (f) direct normal irradiance (DNI). Circles, diamonds, and the lines connecting them represent monthly averages. For (a) CN and (b) CCN, bars represent median and interquartile range of sub-monthly variability of the 1-hr averaged data. For (d) temperature and (e) relative humidity, bars represent monthly extremes, as measured by 5% and 95% levels of the 1-min average data. DNI is presented using 24-hour averages so that it includes the effect of the changing length of day with season, and peak mixing depth is calculated using the 4 PM radiosonde data.
Figure 2: Seasonal PM$_{2.5}$ speciation from the averaged Saguaro National Park and Saguaro West IMProVE sites. Six major groupings comprising the PM$_{2.5}$ mass are shown: FS = fine soil, OA = organic aerosol, EC = elemental carbon, AS = ammonium sulfate, AN = ammonium nitrate, SS = sea salt.
Figure 3: Hourly trends of (a) CN and (b) CCN (0.2%). Bars indicate median and interquartile range of the variability within each hour. Mean CN and CCN concentrations are shown for both weekdays (red) and weekends (blue). Hourly trends of CCN are shown in (c) for each season. Mean EC (solid) and OC (dashed) concentrations (d) are shown for weekdays (red) and weekends (blue).
Figure 4: Hourly trends of activation related properties, OC:EC ratio, and WSOC:OC ratio for weekdays (red) and weekends (blue). Note the applicability of the OC:EC ratio starts to become less well defined on weekends above 25 since EC concentrations are typically below LOD.
Figure 5: Size distribution cluster centroids, as derived by the K-means algorithm, and the hourly distribution of cluster associations separated by season. Clusters are assigned the following identifiers: Nucleation (N; blue), Fresh Fossil (FF; red), Winter/Nocturnal (WN; green) and Condensation/Coagulation (CC; black).
With regard to chemical factors, one of the largest uncertainties in predicting CCN concentration is the nature of the organic fraction of aerosol, which typically comprises many hundreds of compounds with a wide range of chemical structures, affecting solubility, volatility and water uptake (Saxena and Hildemann, 1996; Marcolli et al., 2004; Kanakidou et al., 2005). Furthermore, the existence of low-solubility organic species can, in parallel, cause surface tension depression (the surfactant effect) which enhances droplet activation more than their low solubility would suggest (Shulman et al., 1996; Facchini et al., 2000; Raymond and Pandis, 2002; Hartz et al., 2006; Moore et al., 2008). The photochemically driven oxidation of organic species in the particle phase (Rudich, 2003; Molina et al., 2004; Kanakidou et al., 2005), oligomerization of unsaturated hydrocarbons and carbonyls (Gao et al., 2004; Kalberer et al., 2004) and the partitioning of semi-volatile species from the gas phase (Pankow, 1994a,b; Seinfeld and Pankow, 2003) leads to changes in aerosol properties over timescales from hours to days and may enhance or suppress hygroscopicity (Petters et al., 2006; Michaud et al., 2009). Organics can also cause liquid phase separation with the formation of hydrophobic organic film coatings (Gill et al., 1983; Gilman et al., 2004) or metal-organic interactions which form hard shell insoluble coatings (Drozd et al., 2014), both of which would tend to oppose the surfactant effect, in terms of droplet activation.

The degree to which compositional and mixing state variation affects the hygroscopic properties of an aerosol population in polluted air masses seems to be strongly dependent on location and proximity to emissions (Covert and Heintzenberg, 1993; Cubison et al., 2008; Ervens et al., 2010).

under-studied type of environment, specifically a populated desert.

The present study takes place at a rooftop location in central Tucson. Tucson, Arizona is located in the heart of the Sonoran Desert in the semi-arid southwestern United States. This location offers some unique opportunities for the study of CCN activation primarily since there have been comparatively fewer documented measurements of CCN in arid regions. In addition, southern Arizona is situated in the region affected by the North American Monsoon (NAM) and as a result the highest monthly rainfall occurs during July and August and is accompanied by a strong influx of tropical moisture. The onset of the NAM in late June or early July leads to a rapid change from very hot and dry pre-monsoon conditions to the humid conditions associated with the monsoon and leads to changes in the aerosol properties (Sorooshian et al., 2011; Youn et al., 2013). Aside from the NAM, southern Arizona is situated in a relatively stable synoptic weather pattern, which gives rise to generally clear skies and light surface winds. The strong insolation produces a deep convective boundary layer in the afternoon and clear conditions
lead to significant nocturnal cooling which together produce a significant but predictable diurnal cycle in temperature, humidity and convective boundary layer mixing.

The paper is subdivided as follows: (i) experimental methods and data collection are provided in Section 2; (ii) an overview of the “climatological” results is given in Section 3; (iii) the influence of size distribution and its relationship with composition is discussed in Section 4; (iv) CCN closure analysis is presented in Section 5; and (v) conclusions are presented in Section 6.