Faraday Discussions

Cite this: *Faraday Discuss.*, 2016, **189**, 635

Numerical modelling strategies for the urban atmosphere: general discussion

Rob MacKenzie, Alison Tomlin, Jörg Kleffmann, Thomas Karl,
C. N. Hewitt, Dwayne Heard, Karine Sartelet, Roberto Sommariva,
Urs Baltensperger, Roy Harrison, Sasha Madronich, Gordon McFiggans,
Spyros Pandis, John Wenger, Astrid Kiendler-Scharr, Neil M. Donahue,
Rachel Dunmore, Ruth Doherty, Sarah Moller, Iarla Kilbane-Dawe,
Brian McDonald, Andreas Wahner, Shupeng Zhu, Albert Presto,
Markus Kalberer, Matthew Hort, James Lee, Irina Nikolova,
Jose L. Jimenez, Lisa Whalley, Mohammed Salim Alam
and Andreas Skouloudis

DOI: 10.1039/C6FD90022D

Astrid Kiendler-Scharr opened a general discussion of the paper by Sasha Madronich: I understand that it's convenient to work with volatility distributions – why is it that we would assume that the chemistry actually leads to a population of each of the volatility bins? It seems that recent findings of auto-catalytic reactions in the atmosphere and oxidation reactions leading to highly oxidized molecules point to mechanisms where the consecutive formation of oxidation products with ever lower volatility during the course of a reaction sequence is not happening. Are these contradicting the use of the volatility basis set? Second, it seems that vapour pressures of compounds and compound classes are among the largest uncertainties of models describing partitioning - could you guide the experimentalists on what needs to be done to provide more precise vapor pressures?

Sasha Madronich responded: In explicit chemical models like GECKO-A, vapor pressures of all species (reagents and products) are calculated from molecular structures and can be mapped onto any discretized volatility basis set (VBS). However, decadal VBS bins may be inefficient if using several bins to represent species of very low volatility that are essentially all condensed, or may have insufficient resolution for species having saturation concentrations (vapor pressures converted to C*, $\mu g\ m^{-3}$) near ambient aerosol loading. If for some systems the vapor pressures drop more rapidly (*e.g.*, by autooxidation), these changes can still be represented within GECKO-A, but cannot be described by a VBS multi-generational stepping algorithm in which molecular identities have been anonymized. As for the determination of vapor pressures, it may be useful to note that the existing thermodynamic estimation methods agree quite well for pure hydrocarbons and mono-functional oxygenated species, but diverge from one another for increasingly functionalized molecules.[1]

[1] R. Valorso, B. Aumont, M. Camredon, T. Raventos-Duran, C. Mouchel-Vallon, N. L. Ng, J. H. Seinfeld, J. Lee-Taylor and S. Madronich, *Atmos. Chem. Phys.*, 2011, **11**, 6895–6910.

Gordon McFiggans asked: We have previously carried out investigations using different models (*e.g.* MCM and GECKO/A) and found significant differences in how the volatility bins are filled. Does the multigenerational approach, giving a continuous distribution, inevitably dictate that there will be very low sensitivity to the volatility distribution? How dependent is this continuity in the distribution on the SARs used in the model? Will the volatility distribution in the real atmosphere truly be as continuous as predicted by such a complex multigenerational model, or will the mixing of airmasses of different ages give more variability and dictate that a simple trajectory box modelling framework will always underestimate the sensitivity?

Sasha Madronich answered: We made a comparison of 3 vapor pressure SARs[1] and found very similar continuity in the distribution of the vapor pressure. This continuity is dictated by the multigenerational oxidation process producing a myriad of species with a large (and so *quasi*-continuous) spectrum of properties including vapor pressure. In our simulations of Mexico City, photochemical processing tends to generate a substantial amount of low-volatility compounds fairly quickly (*e.g.* see Fig. 3 of the paper), although of course this fraction grows over the course of several days. These early low-volatility compounds provide sufficient self-seeding to the mixture that the sensitivity exponent is only slightly above unity, *i.e.*, nearly linear. Any external sources of, *e.g.*, aged low-volatility particles would also contribute to the seeding and could result in even more linear sensitivity.

Neil Donahue said: To be provocative, I don't feel we really need to know vapor pressures to answer this. It is very important for explicit, predictive models to be able to predict partitioning, but within the VBS framework the partitioning behavior is empirical; the specific compounds are interesting but not essential to know. Furthermore, if we consider the whole system, other than very close to sources, such as roadways (which are very important of course in the urban context) I would not expect the system to be very sensitive to volatility (just as Sasha has shown).

Spyros Pandis asked: One of the limitations of the estimated organic aerosol volatility distributions based on thermodenuder measurements is that they cover only the material that is in the particulate phase. Therefore, they do not include organic material with higher volatilities that is in the gas phase during the period of the measurements. Could this effect explain the calculated increase of the nonlinear response of the organic aerosol system reported in this work, as one moves from the relatively clean conditions of the Paris campaign, to the more polluted conditions in Mexico City, and finally to the higher levels of organic aerosol in the smog chamber experiments?

Sasha Madronich answered: Estimation of volatility distributions from thermodenuder (TD) measurements is challenging for many reasons,[1] including ambiguity as to whether gas-particle equilibrium has been achieved during the experiment. For example, Cappa and Jimenez[2] used a TD residence time of 16 s for the Mexico City samples, and developed a complex model to simulate evaporation during this limited time, which presumably accounted for the initial absence of the gas phase from collected filters, but the experimental situation is obviously complex. Certainly, the effects would be largest in more polluted situations, consistent with the non-linearity rankings found in our study (smog chambers > Mexico City > Paris). However, the issues in interpreting TD observations are not relevant for the GECKO-A model results presented here (for the MILAGRO and BEACHON field campaigns), which are based on a theoretical estimate of vapor pressures.

1 M. Bilde, *Chem. Rev.*, 2015, **115**, 4115–4156.
2 C. D. Cappa and J. L. Jimenez, *Atmos. Chem. Phys.*, 2010, **10**, 5409–5424.

Spyros Pandis commented: If one is concerned not only about the local sensitivity but rather the response of the system to a change of the order of 50%, is the response similar for both an increase of the organic material and for a decrease? How does the addition of an extra volatility bin with a reasonable amount of material (not included in the thermodenuder measurements) affect this response?

Sasha Madronich replied: The relatively small non-linearities found here would indicate roughly proportional increases or decreases, even with respect to moderately large variations (*e.g.* a factor of 2) in available organic material. Local sensitivity could, however, be distinguished spatially from its regional values, or temporally between instantaneous or multi-day values, for example, noting that regional background seed aerosol on one day may often be the result of urban aerosol exported on previous days.

Urs Baltensperger opened a general discussion of the paper by Irina Nikolova: You conclude that there is probably a non-volatile core to the particles, otherwise there would be complete evaporation. In addition, the constant particle number concentration seems to indicate that every particle contains a non-volatile core. Can you speculate as to what that might be, specifically in the smallest fraction of 10 nm particles?

Irina Nikolova replied: Currently the composition of the non-volatile core is not fully understood. The relatively high sulphur content (less than 50 ppm) in the fuel may contribute to the constituents of this core. In addition, trace metals from the lubricating oil (such as Ca, Fe) may contribute to the composition of this core as well. Most of the hydrocarbon species evaporate from the unburned lubricating oil droplets heated at high temperatures. Such a metallic core could be coated with volatile species upon leaving the tailpipe. Metallic ash residues could also contribute. In our model, we do not specify the content of the non-volatile core; however, the health-related response to that core may be different depending on its chemical composition, and merits further investigation.

Alison Tomlin asked: It is maybe a bit surprising that the particle number concentrations don't change very much overall. Are coagulation and deposition processes so slow that they do not lead to overall particle losses or are new particles being formed?

Irina Nikolova responded: In our study we have examined the role of coagulation and deposition for a typical traffic induced particle number size distribution for three wind speeds (see Table 3 in our paper). All comparisons are against a base case steady-state solution without coagulation or deposition, but accounting for ultrafine particle emissions, condensation/evaporation, mixing with background air and air exchange between the street canyon and the rooftop. There is no new particle formation. For the ambient conditions and the street canyon geometry, we have estimated the time scales for dilution, coagulation and deposition as well. Coagulation and deposition are found to be slow processes, with a maximum reduction in the total number of particles of 8.2% for a wind speed of 1.5 m s^{-1} at rooftop and both processes acting on the particle size distribution. The reduction in particle number concentration decreases to 4.7% for a wind speed of 3 m s^{-1} at rooftop.

Sasha Madronich asked: Did you consider dry deposition of organic gases, particularly SVOCs, that are partitioning dynamically with the particle phase? Deposition velocities for some of these gases (especially those already partly oxidized) could be quite large, and could result in a significant loss of particle mass by evaporative readjustments.

Irina Nikolova replied: In our study we have considered the deposition of both gases and particulates. However, our model does not account for oxidation of SVOC gases due to the relatively large time scales involved in that process compared to the 20-minute model run-time we typically deploy. The emitted particles have only a very small component of partially oxidised compounds.

Neil Donahue said: I have a visceral reaction to the term non-volatile. While it is true that there may be a non-volatile core, I don’t believe it is necessary to explain what you observe. The vapor pressure of a 0.01 μg m^{-3} constituent is just 1 part per trillion. Furthermore, the system should equilibrate on a timescale roughly equal to the condensation sink, meaning pretty quickly in this case I should think. Have you considered an urban background concentration during your mixing calculations?

Irina Nikolova answered: Urban background concentration was considered during the simulations. In our calculations particle composition and vapour-phase concentration data were used based on measurements at traffic and background sites (please refer to Table 1 in our paper). Figure 1 in our paper provides the inlet background size distribution as well as the emission size distribution profile. Mixing between background particles and in-street canyon particles as well as vapour-phase mixing are parameterised using a dimensionless air exchange rate.[1]

[1] C.-H. Liu, D. Y. C. Leung and M. C. Barth, *Atmos. Environ.*, 2005, **39**, 1567–1574.

Neil Donahue said: At $d_p = 5$nm the Kelvin effect should take the saturation vapour pressure from 1 part per trillion to 10 parts per trillion. So I think things will go into steady state.

Albert Presto asked: I disagree with Neil as I think there are non-volatile cores present. We've done sampling in a traffic tunnel and when you heat to 250 °C you end up with particles remaining at 10 nm, which are definitely not organic. Other exhaust studies have used catalytic strippers and end up with material left around at 10 nm, so maybe they are metal because these cores don't burn in the stripper. They are at least functionally non-volatile and don't interact chemically.

Neil Donahue replied: I'm not saying that the particles do not have refractory cores, I'm saying that they don't need to. I agree that there is plenty of evidence that they do.

Gordon McFiggans addressed Irina Nikolova: The persistence of the finest particles, whilst being consistent with every particle containing a solid core, is also consistent with a very low concentration of very low volatility vapours being present in the background atmosphere. Why is it more likely that all particles have a solid involatile core? On the contrary, is it not more likely that the air that is mixing with the evolving and diluting plume contains a continuous distribution of volatility, including small mixing ratios of very low volatility components, sufficient to stop evaporation of the finest particles? This would appear to be consistent with the multigenerational modelling approach such as presented in the last talk. This calls into question the assumptions in the model boundary conditions in the present study. In order to draw any conclusions about the composition of the finest particles, I would suggest that it is necessary to conduct sensitivity simulations. These should include a range of initialisations with dilution air containing a continuous distribution of organic components down to low volatilities, to establish the required magnitude of the lowest volatility components that would be consistent with the persistence of the finest particles. This could be used to quantify the fraction of fine particles that need to have a solid involatile core for a given amount of low volatility components, hence establishing the likelihood of both explanations.

Irina Nikolova replied: Thank you for the valuable suggestions. Our current laboratory study with the GCxGC-ToF-MS (DOI: 10.1039/C5FD00185D) shows evidence of SVOCs originating from diesel fuel and lubricating oil contributing to the particle phase compounds. The largest concentration detected in the particle phase is found in the C23–C29 region, corresponding to the contribution of the lubricant oil. The diesel fuel contributes to the particulate composition with low molecular weight SVOCs (less than or equal to C20). A recent study by Gkatzelis *et al.*[1] shows that 40% of the traffic related particles evaporate in a thermodenuder at 400 °C, while the remaining 60% have a non-volatile core. In support of a non-volatile core are the studies referred to in our paper.[2–7] We currently use vapour phase SVOC concentrations measured in Birmingham in the model, and we have recently collected new data from London.

1 G. I. Gkatzelis, D. K. Papanastasiou, K. Florou, C. Kaltsonoudis, E. Louvaris and S. N. Pandis, *Atmos. Meas. Tech.*, 2016, **9**, 103–114.

2 S. Biswas, L. Ntziachristos, K. Moore and C. Sioutas, *Atmos. Environ.*, 2007, **41**, 3479–3493.

3 P. Tiitta *et al.*, *Atmos. Environ.*, 2010, **44**, 976–986.

4 T. Ronkko, A. Virtanen, J. Kannosto, J. Keskinen, M. Lappi and L. Pirjola, *Environ. Sci. Technol.*, 2007, **41**, 6384–6389.

5 T. Ronkko, T. Lahde, J. Heikkila, L. Pirjola, U. Bauschke, F. Arnold, H. Schlager, D. Rothe, J. Yli-Ojanpera and J. Keskinen, *Environ. Sci. Technol.*, 2013, **47**, 11882–11889.

6 U. Kirchner, V. Scheer, R. Vogt and R. Kägi, *J. Aerosol Sci.*, 2009, **40**, 55–64.

7 W. Birmili, K. Heinke, M. Pitz, J. Matschullat, A. Wiedensohler, J. Cyrys, H.-E. Wichmann and A. Peters, *Atmos. Chem. Phys.*, 2010, **10**, 4643–4660.

Sasha Madronich replied: The multigenerational GECKO-A model does indeed predict the presence of very low volatility components, initially with small mixing ratios but increasing as photochemical processing of gases from higher volatility bins continues. However, particle-phase processes such as oligomerization (not currently in GECKO-A) could contribute, and the model is not able to distinguish between its model construct of a *quasi*-ideal solution, and a solid involatile core.

Karine Sartelet asked: I think your study is very interesting and shows the impact of the Kelvin effect on organics. I was wondering what parameters (*e.g.* surface tension) you used to model the Kelvin effect?

Irina Nikolova responded: Surface tension was considered following the work of Queimada *et al.*[1]

1 A. J. Queimada, Isabel M. Marrucho, J. A. P. Coutinho, *Fluid Phase Equilib.*, 2001, **4639**, 1–10.

Karine Sartelet asked: Did you investigate the sensitivity of your results to the number of size sections used?

Irina Nikolova replied: Our goal is to evaluate in detail the physical behaviour and chemical composition of the ultrafine particles (with diameters of less than 100 nm). We have performed tests varying the number of the size bins (from 5 to 20). However, during the redistribution step after condensation/evaporation with a low number of size bins, the redistribution scheme becomes quite diffusive. This problem can be overcome by increasing the number of size bins and finding the balance between the computational load and the reduction in the numerical diffusivity. In our study, that balance occurred when using 15 discrete size bins.

Karine Sartelet asked: I wonder what the sensitivity to the number of size sections is, because processes such as deposition are known to be sensitive to the size distribution. Because 15 sections are used in this work, the sensitivity to the number of sections may be low.

Irina Nikolova replied: We have used the scheme of resistances to estimate the deposition velocities. Deposition velocities are higher in the nucleation mode particles (especially in the sub-10 nm size range) than in the Aitken mode particles due to the Brownian diffusion of the particles. Due to the evaporation of SVOCs, the particle size distribution shifts significantly to the lower end of the nucleation mode. A lower number of size bins (*e.g.* 5) on a logarithmic scale is too coarse in the nucleation mode and does not provide the detail we are interested in to investigate.

Rob MacKenzie opened a general discussion of the paper by Karine Sartelet and addressed Shupeng Zhu: Please could you comment on what, other than runtime constraints, determined your choice of size bins?

Shupeng Zhu responded: Yes, the number of size sections is limited by the computation cost. The choice of the size bounds was initially made based on a log size discretization (0.01, 0.0398, 0.1585, 0.6310, 2.5119 and 10.0) μm. However, as the measurement data provide information on PM_1 (diameter < 1 μm) and most of the BC mass observed in the measurements is less than 0.4 μm, we manually inserted two size bounds 0.4 μm and 1.0μm between 0.1585 μm and 2.5119 μm. We removed size bounds 0.0398 μm and 0.6310 μm in order to keep 5 size sections.

Spyros Pandis said: What was assumed regarding the mixing state and the size/composition distributions of the boundary conditions?

Shupeng Zhu replied: In the simulations of this paper, boundary conditions come from a larger-scale simulation, in which the internal mixing assumption is used. Therefore, we do not have information on the mixing state of the boundary conditions. Therefore, all particles coming from the boundary conditions are assumed to be aged and internally mixed. So for each size section, there is only one composition for particles for the boundary conditions.

Spyros Pandis commented: Can you please explain what was assumed regarding the size/composition distribution of the transportation emissions? Did all emitted particles have the same composition (*e.g,* the same black carbon, organic carbon, *etc.*)?

Shupeng Zhu replied: All emitted particles are considered as particles with pure substance. In this study only elemental carbon (*i.e.*, black carbon) and dust are directly emitted in particle phase. So each emitted particle is either made of pure elemental carbon or pure dust. All organic carbon (POA) is considered as semi-volatile and put into gas phase then allowed to condense into the particle phase within the SOA treatment.

Karine Sartelet added: For transportation, we assume that black carbon and dust are emitted from exhaust and resuspension respectively. We assume that there are two distinct populations. For SOA we assume that everything is in the gas phase, then let them partition.

Shupeng Zhu said: So all emissions in particle phase are considered as fully unmixed and only contain one chemical species. For particles originating from initial and boundary conditions, we assumed that they are internally mixed: within the same size section, they all have the same composition.

Roy Harrison remarked: I agree that the ATOFMS is an imperfect instrument with very different sensitivities to different elements. With careful calibrations against independent measurements, this could be compensated for to some degree, but you still cannot compensate for the fact that hit rates (the ratio of ionised particles to those entering the instrument) is different for particles of differing composition. There are also significant matrix effects for chemically-mixed particles, so there are major weaknesses in the technique and I am not entirely comfortable with using it as a basis for a simulation. However, having said that, I am not sure that there is any other technique available giving better information on single particles.

John Wenger commented: I agree with Prof. Harrison that the aerosol time of flight mass spectrometer (ATOFMS) has limitations. But if we're addressing the diversity of particles at an individual level, then the ATOFMS is the best technique for doing this. I also think that characterisation of the chemical diversity of particles is a useful approach which allows us to compare and contrast ambient particle mixing states at different locations around the world.

Gordon McFiggans addressed Shupeng Zhu and Karine Sartelet: A statement is made in the paper that a dynamic inorganic component treatment is used, with organics treated in equilibrium followed by redistribution across the particle sizes, due to the limitations of the H_2O model. That redistribution will put the organics into discrete size fractions of particles and hence dictate the mixing-state of all particles containing organics. Can you elaborate more on the redistribution? How confident are you that this redistribution puts them in the right place? Most importantly, how is the robustness of your conclusions dependent on the skill with which the organics are distributed? I would suggest that a sensitivity study is used to attribute error bars in the mixing state of organics to this redistribution.

Shupeng Zhu answered: Details about the redistribution method of organics species are given in the electronic supporting information (ESI) (see DOI: 10.1039/c5fd00175g). Basically, the mass redistributed after the equilibrium is related to both the number of particles and the condensation/evaporation kernel within each bin. This method tends to redistribute more organic aerosol (OA) mass into the smallest size bins due to their high number concentrations, and it leads to an overestimation of OA in those sizes. It is true that this redistribution method needs to be improved, and we also discussed one possible improvement by introducing the Kelvin effect into the weighting factor of the redistribution in the ESI. A sensitivity study is a good suggestion, but the best way would be to directly use a dynamic approach to compute the organics partitioning. Such approach already exists and we are planing to implant it into our model in future work.

Neil Donahue asked: I still don't understand how you have a large Kelvin influence for relatively large particles. How does this work out?

Shupeng Zhu responded: The particle diameter studied here is between 0.01 and 10 μm. For particles in the first size section, which have a mean diameter of around 0.029 μm, the Kelvin effect is not negligible. If the Kelvin influence is not taken into account, then more mass will be redistributed into the smallest size section during each time step. Even if this amount of mass is small for each single time step, the accumulated effect will be very significant as the over-estimated mass will feed back and produce a even larger weighting factor for the small size section in the next time step. Another factor is that we find the size distribution of inorganic species to be more reasonable as this is computed dynamically and includes the Kelvin effect. So, the ideal solution would be to use a dynamic computation for the organic species as well in our future work.

Astrid Kiendler-Scharr addressed Shupeng Zhu, Sasha Madronich and Irina Nikolova: I have a question to all three of you, coming from an experimental point of view. If you were to choose measurements to most efficiently improve your models and constrain the input to models, which measurements would these be? What is the largest gap in observational knowledge and which parameters should be measured?

Shupeng Zhu replied: I think a more precise measurement on a single particle level from some specifically designed gas chamber experiment would be very helpful for the improvement of our understanding of particle mixing, as well as the model design. The largest gap in observational knowledge is that we lack the means to monitor how the chemical composition evolved within individual particles during the mixing process. The parameter we are interested in is the mass fraction of each chemical component within each particle.

Sasha Madronich responded: One useful bit of information would be the distribution of specific functional groups per condensed molecule (*e.g.* various -OH, -CO, -COOH, -ONO_2, and -OOH moieties) and in particular the occurrence of multiple functional groups per molecule. FTIR quantification of some functional groups (*e.g.* Maria *et al.*[1]) has already shown potential for sorting particles of different origins (*e.g.* Russell *et al.*[2]). Among other things, such measurements could help assess to what extent the molecular identity of predicted gas phase photochemical products is retained in the condensed phase, and whether the particle bulk O/C, N/C, and H/C ratios can be reconstructed from the molecular scale information.

1 S. F. Maria, L. M. Russell, B. J. Turpin and R. J. Porcja , *Atmos. Environ.*, 2002, **36**, 5185–5196.
2 L. M. Russell, R. Bahadur and P. J. Ziemann, *Proc. Nat. Acad. Sci.*, 2012, **108**, 3516–3521.

Irina Nikolova responded: I agree with what has been said so far, that measurements of the mass fraction of each individual component and the distribution of specific functional groups per molecule are valuable inputs for modelling work. In order to understand the complex behaviour and interactions between particles in the atmosphere with multiple sources, there is a need for a clearer understanding on how individual particle composition changes per source. High spatial and temporal resolution measurements of particle composition in the ambient atmosphere are also highly needed.

Urs Baltensperger returned to a discussion of the paper by Sasha Madronich: I read from your modelling framework that you consider a mixture of organic species. Do you allow for inorganics and water as well, or do you just avoid mixing?

Sasha Madronich replied: Particle-phase inorganics and water were not considered in this study, but certainly should be in future work. Most current parameterizations of organic aerosols are still relatively simple, particularly those implemented in 3D chemistry transport models, and are largely based on smog chamber yields, where organic mass tends to overwhelm particle-phase inorganics (*e.g.*, those used as seeds), or on algorithms such as the multi-generational volatility basis set (VBS) that also assume a purely organic phase. Current thermodynamic models (*e.g.*, ISORROPIA, MADRID, MOSAIC and AIM) have limited representation of organics, and certainly not in the molecular detail provided by explicit gas-phase models such as MCM or GECKO-A. Neil Donahue's paper (DOI: 10.1039/C5FD00214A) showed that even different organic aerosols may or may not be miscible depending on their precise nature. As we gain better understanding of this complexity from laboratory and field measurements, improved parameterizations can hopefully be developed to represent mixing appropriately, rather than avoiding it.

Roy Harrison commented: I would like to question whether the Pankow model of absorption–adsorption into an organic medium is the correct paradigm. The atmosphere typically contains some liquid water associated with particles and our work on the mass closure of atmosphere aerosol clearly shows the presence of water. Since the more oxidised organic aerosol is likely to be hydrophilic, would a Henry's Law type model be more appropriate for the oxidised species?

Sasha Madronich responded: The Pankow paradigm (which for absorption into a liquid is essentially Raoult's law) has been used extensively to represent gas-particle partitioning of organics, especially in 3D models as it can be parameterized efficiently, *e.g.* in terms of chamber-measured yields, or with algorithms such as the multi-generational Volatility Basis Set (VBS). You are correct that this may not be always the best approach, given that water-soluble organic aerosols are at least as abundant as hydrophobic aerosol (especially after photochemical ageing), and that water is frequently a major constituent of the particles. The Henry's law limit is obviously relevant to organics diluted in cloud drops, (*e.g.*, Graedel and Goldberg[1] and Hermann *et al.*[2]) but it is also likely to be dominant for many aerosol systems, especially in high humidity environments such as the South Eastern U.S. and the Amazon.[3,4] Liquid–liquid phase separation within the aerosol is also possible (*e.g.*, Zuend and Seinfeld[5]). Unified models that encompass both Raoult's and Henry limits have been developed (*e.g.*, Pun *et al.*[6] and Bowman and Eskelson[7]) but have not been used widely in 3D models due to computational complexity and scarcity of experimental constraints. Better parameterizations of these processes could be a real opportunity to improve the predictive ability of the models.

1 T. E. Graedel, and K. I. Goldberg, *J. Geophys. Res.*, 1983, **88**, 10865–10882.
2 H. Herrmann, T. Schaefer, A. Tilgner, S. A. Styler, C. Weller, M. Teich and T. Otto, *Chem. Rev.*, 2015, **115**, 4259–4334.
3 B. Aumont, S. Madronich, I. Bey and G. S. Tyndall, *J. Atmos. Chem.*, 2000, **35**, 59–75.
4 B. Ervens, B. J. Turpin and R. J. Weber, *Atmos. Chem. Phys.*, 2011, **11**, 11069–11102.

5 A. Zuend and J. H. Seinfeld, *Atmos. Chem. Phys.*, 2012, **12**, 3857–3882.
6 B. K. Pun, R. J. Griffin, C. Seigneur and J. H. Seinfeld, *J. Geophys. Res.*, 2002, **107**, 4333.

Karine Sartelet remarked: In our model, for organics we distinguish hydrophobic from hydrophilic components. Each component is a surrogate for a class of organics. We do have water-soluble organics.

Gordon McFiggans addressed Karine Sartelet: There's a whole range of different treatments available for multicomponent thermodynamics and partitioning, with ion interaction models for aqueous electrolytes at one end, and absorptive partitioning for organic components at the other. Most widely used recent treatments include the AIOMFAC group contribution approaches to drive multicomponent inorganic and organic partitioning. It is readily possible to cover a wide range of different levels of complexity, dependent on the questions to be answered. I am not sure that it is the case that all large scale models use "Pankow" partitioning, and it's quite straightforward, if the activity coefficient reference state is correct, to include water soluble particle components. I think it is important to keep in mind these capabilities when choosing the appropriate tools to address the question being asked. Internally consistent frameworks do exist and may be more appropriate in investigations of the mixing of inorganic components with insoluble and water soluble organics.

Karine Sartelet replied: Although the activity coefficients were not modified to take into account the interactions between organics and inorganics in this work (only organic–organic interactions), the mixing state of organics is influenced by inorganics in our model H_2O, as soluble organics condense onto inorganics and insoluble organics do not. Separately from this work, we have also developed a model, SOAP,[1] to take into account the dynamics of condensation/evaporation of organics and organic–inorganic interactions using AIOMFAC. We did not use SOAP in this work, but we are planning to couple it to the mixing-state resolved model SCRAM to investigate the mixing state of inorganics with organics.

1 F. Couvidat, and K. Sartelet, *Geosci. Mod. Dev.*, 2015, **8**, 1111–1138.

Shupeng Zhu added: Our organic model differentiates between soluble and insoluble organics. In this work, activity coefficients for organic–organic interactions were taken into account using UNIFAC, but inorganic–organic interactions were not considered.

Rob MacKenzie opened a general discussion of the paper by Andreas Skouloudis: Please could you comment on whether you consider your results – showing no change in urban ozone over time – are consistent with the data interpretation and modelling of the paper presented by Christian Ehlers earlier (especially their Fig. 1, 2, and 26)?

Andreas Skouloudis answered: Although ozone precursors have been falling between the base year and the target year, the ratio of NO and NO_2 gases in the atmosphere has changed. We should also consider that Christian's manuscript examined only two sites (one at Mace Head in Ireland and one at

Hohenpeissenberg in Germany). The measurements we presented are in all monitoring sites at the ten modelling domains of the second AutoOil, and we show the actual situation all over Europe as was available at that time. Obviously, we did not examine the frequency of exceedances of the limit values because these limits depend on the climate conditions between the base and the target year, and because the number of hours of exceedance is subjective for different limit values, as you can easily see from Fig. 5 in our work. However, in constant conditions, there is a neat cycle in which nitrogen dioxide helps to form ozone and nitric oxide helps to break it apart. This means, basically, that we are taking away some of the ozone suppression and, as shown in Fig. 26 of Christian's work, this might result in increases of ozone concentrations. This indeed is in line with what we have observed in Fig. 10 of our work. This cycle appears to have been perturbed by control measures that were actually intended to remove the fine particles and soot in vehicle exhausts. It remains to be seen if these levels continue to stay high or if eventually we might see a reduction of ozone in future years. This is one of the reasons why models should become operational over larger scales, in order to account for what is happening in other regions of the world and identify properly the import of ozone and its precursors from outside. In the meantime, the ground-level of ozone might remain persistently high even in regions with restricted sunlight and not-so-hot temperatures.

Andreas Wahner commented: What was presented in the paper by C. Ehlers were measurements by us and statistics from the German federal agencies for German cities. In the paper, Fig. 1 shows measurements of background ozone at GAW Hohenpeisenberg, and there you hardly see any change in background ozone, except for maybe a small decline over the last 20 years. You were referring to many stations throughout Europe, and you can show that the long term trends at different stations are not the same. There seems to be no clear trend in ozone.

In the past years there have been advances in detailed model descriptions of aerosol concentration and formation, and in techniques for particle characterization. Do you see these advancements being used in the models you use in developing policy?

Andreas Skouloudis replied: Indeed, there have been many advancements focusing on aerosol science, their association with atmospheric chemistry and their size and number characterisation. There have been also been advancements in numerical techniques, as you can see from my other publication.[1] Yet, for policy development implementation there is still a lot of empiricism and the models used have incorporated parametrisations that hinder the generalisation of size and number distributions and create doubts on proposed solutions. Unfortunately, apart from inter-comparison of schemes, more research publications are needed for characterising nucleation, coagulation condensation deposition and other loss mechanisms of aero-disperse aerosols and the emissions from relevant anthropogenic sources as seen in M. Prank *et al.*[2], so far in coarse spatial geographical resolution.

1 A. N. Skouloudis and W. Eifler, Paper 7-46, Eurotrac Symposium Garmisch-Partenkirchen Germany, SPB Academic Publishing 11-15, 1994.

2 M. Prank *et al.*, *Atmos. Chem. Phys.*, 2016, **16**, 6041–6070.

Andreas Wahner remarked: In many cases of model and measurement comparison, simplified or parameterized models and limited datasets are used and models/ parameters are sometimes optimized to give better agreement. Should it be a task for the future, with growing accessibility to much observational data, to compare the full complexity of data with as detailed models as we know of, using inverse modelling techniques to find the weak point in our mechanism in an algorithmic optimized way? This may be possible in future if the advancements in computer power in the next 5 to 10 years are increasing as expected. Should our community not be prepared for these opportunities?

Matthew Hort said: I would agree with the need for high resolution modelling in a general sense but would urge caution in thinking that high resolution is the answer to everything. The resolution needs to be relevant for the question being "answered" and also needs to be considered in context of the other model inputs (*e.g.* emissions), parameters (*e.g.* chemical and physical) and available observation data for initiation, verification and validation. Also, in thinking about high resolution we have to think about the genuine predictability of the relevant phenomena, and if the only valid approach might be probabilistic. The exact scope will indeed depend on the application/task. I am coming from a large scale *i.e.* city and regional, suitable for air quality assessments and forecasts. At these scales the only viable approach is similar to that followed by Meteorological Services where measurement data are used to constrain the model initial state and verify, but this requires a sound model, daily detailed measurements and sound methods of comparison.

I think computers are decades away from being powerful enough to run models for real world applications, in the context I have stated above, at the kind of complexity Andreas has asked about (street resolving), for anything but highly idealised cases *i.e.* single street canyons or road junctions.

Alison Tomlin responded: I certainly agree with your first point. The topic of the paper is probabilistic modelling. For national scale forecasting, 1 km seems like a realistic objective in terms of scale. However, many pollution hot-spots might occur because of small scale issues such street topographies, particular traffic hot spots, localised emissions from combustion for heat applications *etc.* There is therefore a role for nested higher resolution, and building resolving models for such applications – assuming, as you say, that emissions data are available at such scales and that the models can be suitably validated. Even a CFD model such as that used in the present study can be run in less than 1 h on a standard PC so perhaps we are not decades away from using such models for particular case studies of importance. Models such as ADMS and OSPM, of course, can be run much faster than this and also have a role to play in nesting models for city applications. It will be important to carry out the types of validation and sensitivity studies such as the one described in our paper, to establish whether such models are fit for purpose for particular applications. If they can be shown to be relevant, then they could be used to answer some questions of local relevance that could not be answered by larger scale forecasting models.

Andreas Skouloudis replied: Although I agree with the necessity for better model inputs regarding emissions, sound physical and chemical models and initial conditions that are based on detailed hourly measurements, I do not think that computational capacity is the main bottleneck for the proliferation of advanced modelling applications, even nowadays. For this purpose, I would like to draw to your attention the article by Overpeck *et al.*[1] which shows that modelling potential will be steadily increasing until 2030 in parallel to initial measurements with satellite and sensor technologies. These are already, in 2015, resolving a lot of initialisation demands and eliminating the constrains of empirical assumptions. The main challenge ahead is not to increase the resolution of the modelling domain all the way down to street canyon or road junction level, but to demonstrate the success of modelling in resolving topics that otherwise could not be solved in various interlinked geographical scales (ideally two-way nested). The aim of our present work was to show the success of realistic regulatory forecasts on the basis of real measurements in the target year. This is perhaps a small effort in the direction of genuine predictability of the evolution of urban pollution. As for the uncertainty initial conditions, taking into consideration the revolution in meteorological services with the use of compact sensor technologies in association with satellite observations, and in association with sensors and RFIDs implemented in everyday human applications, I think that there is no need for caution. In reality we need, today, innovative modelling deployments that will substitute for the empiricism of parametrisations and the uncertainty of initial conditions, taking into account the advantages of actual real-time measurements. Certainly, we can do this even today in certain domains in Europe, otherwise all these increases of modelling potential will be wasted.

1 J. T. Overpeck, G. A. Meehl, S. Bony and D. R. Easterling, *Science* , 2011, **331**, 700–702.

Ruth Doherty asked: Going back to the previous questions regarding the idea to include all the chemical complexity we know about in air quality models – if we did this with the ADMS-Urban and other such models then the model wouldn't be able to run. This is because these urban models, which are typically local dispersion and chemistry models, were designed to be used by local authorities and planners and hence typically run on a single Windows or other PC. So as suggested previously by Dr Hort, we need to go back to the question/s asked and how we can best answer that. For example, if we are interested in the effect of long-term exposure to $PM_{2.5}$ concentrations, how would we deal with chemical complexity and what level of complexity would be needed to calculate chronic health impacts?

Alison Tomlin replied: There is a large body of work on methods for chemical model reduction (see Turanyi and Tomlin[1] for an extensive review). These methods have been very successfully used in combustion modelling but have been perhaps under-utilised in atmospheric chemistry models. Perhaps they would have something to offer in terms of providing a consistent way of reducing the size of reaction mechanisms or models whilst retaining the important chemical complexity.

1 T. Turanyi and A. S. Tomlin, *Analysis of Kinetic Reaction Mechanisms*, Springer-Verlag, Berlin, Heidelberg, 2014.

Nick Hewitt addressed Andreas Skouloudis: Concerning the geographical aggregation of data – Fig. 12 in your paper shows the log of concentration against the log of a population parameter. You define the functions in the text, but don't discuss the implication of the gradient of these lines for exposure and regulation across Europe. Could you elaborate?

Andreas Skouloudis responded: The details about how the local air quality concentrations were generalised all over the ten modelling domains, and can be found in our corresponding report[1]. These relationships were developed originally for the generalisation of the modelling results over urban domains in Europe and certainly link pollutant concentrations to human exposure. They have not yet been verified with individual exposure assessments since they refer to integrated values over large urban areas, but such verifications could certainly become feasible in the near future on the basis of real-time measurements from mobile monitoring devices. We are willing to discuss the implication of these relationships further, if suitable exposure data become available, and always on the basis of quantifiable parameters based on measurements.

1 EUR Report No EUR19725EN, European Commission, EU Bookshop ISSN: 1018-5593, 2000, pp 1–264.

Jörg Kleffmann opened a general discussion of the paper by Alison Tomlin: Did you include in your MISKAM model the turbulences which were induced by cars (the "piston effect")? How would this affect the high variability of NO_2 modelled inside the street canyon? We recently measured NO_x in a double lane traffic tunnel in Brussels,[1] showing homogenous mixing throughout the whole cross-section, which is also reasonable at least for the lower section (<5 m altitude) of a street canyon. Potential spatial variability (or not) is of high importance for defining locations for NO_2 measurements in urban environments.

1 M. Gallus *et al.*, *Build. Environ.*, 2015, **84**, 125–133.

Alison Tomlin responded: We did not include traffic induced turbulence in our simulations. Under low wind speeds and high traffic speeds this might be important. However, typical average vehicle velocities within the street varied between 10 and 20 km h^{-1} due to the levels of congestion experienced. These are unlikely to produce enough turbulence to make the canyon well mixed. The model was validated against observations for the low reactivity tracer carbon monoxide within an earlier study by Dixon *et al.*[1] That study showed a clear difference (up to a factor of 3) between concentrations in the leeward *vs.* windward canyon locations. The turbulent kinetic energy was also seen to vary between leeward and windward locations. For one particular wind direction where a low velocity convergence zone was shown to be formed by the model, there was a significant discrepancy between the model and observations. Traffic-produced turbulence could have been affecting the dispersion for these particular conditions.

1 N. S. Dixon, J. W. D Boddy, R. J. Smalley and A. S. Tomlin, *Atmos. Environ.*, 2006, **40**, 958–972.

Rob MacKenzie asked: It is interesting and very challenging to see the dependence of results on roof-top wind direction. What are the prospects of generating well-characterised roof-top wind directions for studies that do not measure it directly?

Alison Tomlin responded: This is a difficult challenge for many urban areas, since above roof meteorological measurements are not routinely obtained within most cities. In addition, in order to serve as a reliable reference measurement, the site needs to be far enough above the roof to reduce the influence of localised roof-top structures. In a study by Barlow *et al.*,[1] conducted during the DAPPLE campaign, a comparison was made between a near roof-top site and a site on the top of the BT Tower in terms of the ability of the reference data to be able to explain the in-street variability in channelling and recirculating flow components. In fact the higher measurement (190 m) explained more of the in-street flow than the site that was only 2 m above the roof-top; the latter showed interference from local roof-top structures. Long term measurements from the BT tower could form a suitable reference for London based studies and it would be very useful to perform a comparison between the use of data from the BT tower and that from nearby reference sites such as Heathrow. Some other cities may have equivalent sites; Leeds City Council, for example, has collected meteorological observations from an above-roof reference tower within the city centre for several years. These measurements are, however, piecemeal and there is no guarantee of their longevity due to funding challenges. A second option within a nested modelling scenario would be to use a coarser resolution model to obtain the boundary conditions for the street scale model. For this to be successful, however, the coarser model would need an accurate representation of the surface aerodynamic characteristics. This is a challenging area and there is a need for comparison between different types of reference measurements/models in terms of their ability to explain in-street flow characteristics for a wider range of case studies.

1 J. F. Barlow, A. Dobre, R. J. Smalley, S. J. Arnold, A. S. Tomlin and S. E. Belcher, *Atmos. Environ.*, 2009, **43**, 5536–5544.

Brian McDonald remarked: With regards to high resolution atmospheric modelling, we've been talking about chemical and meteorological complexity. What about emissions complexity? In order to run high-resolution models, high-resolution emission inventories are needed as input. To develop emission maps finer than 1 km is a big effort, and potentially fraught with large uncertainties. Let's not forget about the need to develop bottom-up emission inventories as well. We need better maps of sources.

Rob MacKenzie said to Alison Tomlin: Several comments from the floor – in response to a number of papers – have commented on the need to move to dynamic simulation of emissions. Please could you comment on this aspect of your own paper? Does this part of the model contribute much to variabilities in NO_2 concentration at street level?

Alison Tomlin responded: A dynamic traffic micro-simulation model was used in the work with the intention of assessing the impact of traffic demand on roadside increments in NO_2. What the paper showed was that for the lower demand scenarios the predicted concentrations were quite sensitive to the level of demand. Under these scenarios, transient queues built up at the signal controlled intersections, but they cleared between signals. For higher demands congestion could build up within the street, with permanent queues forming. Under these scenarios, predicted road-side increments in NO_2 were not very sensitive to changes in demand; larger reductions in demand were required in order to reduce permanent queue build up. This type of dynamic model therefore allows different traffic scenarios to be investigated. The challenge is to link the outputs of the traffic model to relevant real world emissions factors that accurately represent the effects of vehicle speed, acceleration, age, condition, mitigation technology *etc.* Such factors need to be representative of urban driving conditions that include a large degree of stop–start behaviour close to intersections.

Spyros Pandis asked: Our models are used both for improving our understanding of the present (often for scientific reasons), but also for predicting future changes due to changes in emissions (mostly for regulatory use). Are their sensitivities to different parameters, for example wind speed and direction, the same in these two cases? Clearly, the predicted concentration of a pollutant is quite sensitive to wind speed. Is, however, the change in this concentration for a change in emissions as sensitive to the wind? This issue is obviously relevant if one focuses on chemical mechanisms and specific reactions in there.

Alison Tomlin answered: This is an interesting question and there are perhaps different ways of looking at the issue. Firstly, one of the main conclusions of the paper was that despite the impacts of uncertainties in the model parameters on the width of predicted concentration distributions, it was still possible to assess the impact of potential traffic management measures, such as reductions in demand, using the HDMR method for variance decomposition. ANOVA type methods allow for the effects of individual parameters to be clearly seen within the scatter caused by other uncertainties in the model. Secondly, the treatment of model parameter uncertainty helps to avoid coming to the wrong conclusions in assessing the causes of concentration trends *i.e.* attributing changes in concentrations to the wrong processes. For example, it is common in the interpretation of pollution time series to factor out the effects of long term trends in wind speeds. If this is not done, then trends in concentration could be wrongly attributed to changes in emissions when they could be a result of differing climatology between different parts of the time series. If the site in question was within a street canyon, then changes in reference wind direction could have an even more dramatic effect on changes in measured concentrations than wind speeds. Therefore, if one time period had substantially differing background wind directions to another, these effects would need to be accounted for in the same way as wind speeds. For the assessment of the causes of short term episodes and intervention measures, this may be especially important. Finally, models are often used to test emissions inventories, which in turn are used in the assessment of pollution trends. Differences between model predictions and measurements are sometimes used to suggest errors within the inventory. However, they may also be due to a range of other model uncertainties, and these need to be taken into account before conclusions can be drawn about inventory problems. The treatment of uncertainties is therefore very important in inverse modelling.

Sasha Madronich remarked: You showed the sensitivity of model results to many parameters, including some arising from uncertainties in chemical reaction kinetics. The sensitivity to Arrhenius parameters (A-factor and activation energy, or E/R) is interesting in its own right, but should be distinguished from a sensitivity to the rate coefficient, $k(\mathrm{T}) = A\ \exp[-E/(RT)]$. When fitting plots of log $k(T)$ *vs.* $1/T$, covariance between intercept A and slope E/R makes their estimation less certain than the k data from which they are derived (*e.g.*, Cvetanovic *et al.*[1]). The IUPAC (your ref. 33) recommendation for the reaction NO + O_3 is $k(298\ \mathrm{K}) = 1.8 \times 10^{-14}\ \mathrm{cm^3\ s^{-1}}$ with an uncertainty of 8% ($\Delta\log = \pm 0.08$), and $k(T) = 1.4 \times 10^{-12}\ \mathrm{cm^3\ s^{-1}}$ (no uncertainty given) exp $[-(1310 \pm 200\ \mathrm{K})/T]$.

Propagating just the E/R range (1110–1510 K) through the Arrhenius equation gives a range of $k(298\ \mathrm{K}) = (0.9\ \mathrm{to}\ 3.3) \times 10^{-14}\ \mathrm{cm^3\ s^{-1}}$, indicating a factor of 2 uncertainty, much larger than the recommended 8% uncertainty of k near 298 K. It is the latter that controls the evolution of the chemical system – though the direct sensitivity to E/R certainly matters where temperature variability is an issue.

[1] R. J. Cvetanovic, D. L. Singleton and G. Paraskevopoulos, *J. Phys. Chem.* 1979, **83**, 50–60.

Alison Tomlin replied: Thanks for the comment. I think the uncertainty factor of 0.08 was for log(k) which leads to a 20% uncertainty in k at 298 K. That aside, I agree with you that treating correlations between Arrhenius parameters is an important topic. It was recently addressed in a very interesting paper by Nagy and Turanyi.[1] The problem is that these correlations are very rarely presented in the literature. They have to be derived from a wide variety of data that constrains the rate constant rather than individual studies. This requires the application of some kind of optimisation strategy and whilst such studies are beginning to emerge, correlation information is available for only a very limited number of systems. For combustion systems, older evaluations such as those of Atkinson and Baulch do often provide uncertainties in k at different temperatures, but these were not presented for the reaction under debate. Therefore we were left with using the uncertainty in E/R as presented. Treating correlations is a very interesting topic that should be considered in much more detail in the future.

[1] T. Nagy and T. Turányi, *Int. J. Chem. Kinet.*, 2011, **43**, 359–378.

Dwayne Heard commented: It is interesting to see that your sensitivity studies show that the model results are sensitive to the value of the rate coefficient for the reaction of NO with ozone, and also its temperature dependence. Coincidentally, I was recently asked about the kinetics of this reaction in connection with another modelling study, in particular the behaviour at lower temperatures. The kinetics of this reaction were studied in some detail in the 1970s and early 1980s, so more than 30 years ago. A more recent study of this reaction was published in 1998 by Moonen *et al.*[1], where it was stated that previous measurements had been made at low total pressures and/or with very large mixing ratios of NO and O_3 relative to those used in the atmosphere. Moonen *et al.* used 10 s of ppbs of the reagents, and studied the reaction at tropospheric temperatures and pressures. Although they found some differences in the activation energy and pre-exponential factors compared with the earlier studies, the rate coefficient at 298 K was similar to the recommended value, and the lack of a pressure dependence was confirmed. Bedjanian *et al.*[2] also published a rate constant in the same year, but it was only obtained at 298 K and at a low pressure of 1 Torr (in helium), agreeing with the recommended value. Although there is no reason to suspect that the kinetic data are incorrect, the kinetics of this reaction were established before the development of some of the modern techniques used for studying kinetics. Given the importance of this reaction as highlighted in this paper, and for atmospheric chemistry in general (it is difficult to think of a more important reaction for urban areas!), it may be prudent to embark on a fresh study of the kinetics of this reaction, particularly over a wide range of temperatures. I would like to thank Professor Neil Donahue for useful discussions regarding the kinetics of this reaction.

1 P. C. Moonen, J. N. Cape, R. L. Storeton-West and R. McColm, *J. Atmos. Chem.*, 1998, **29**, 299–314.
2 Y. Bedjanian, G. Le Bras and G. Poulet, *J. Phys. Chem.*, 1998, **102**, 10501–10511.

Alison Tomlin replied: I would agree with the comment that it may be useful to perform further detailed kinetic studies on this reaction. It would seem that there have been no recent studies, and none since the evaluation of Atkinson *et al.*[1] that determined the uncertainty in activation energy used in the sensitivity analysis of our work.

1 R. Atkinson, D. L. Baulch, R. A. Cox, J. N. Crowley, R. F. Hampson, R. G. Hynes, M. E. Jenkin, M. J. Rossi, and J. Troe, *Atmos. Chem. Phys.*, 2004, **4**, 1461–1738.

Thomas Karl communicated to Alison Tomlin: We have recently shown experimentally[1] that significant bias for treating chemical reactions can occur in RANS depending on Dahmköhler numbers and the underlying degree of heterogeneity of the reactant surface emissions. These complex and non-linear effects depend on spatial model resolution, chemical reactivity and scales of surface heterogeneity. Our experimental findings support recent LES modelling studies[2] performed over dense plant canopies, which are characterized by non-linear heterogeneity driven mixing effects similar to what we find in "urban canopies". Since the spatial scale of the presented modelling domain in your paper varies, the question arises as to how heterogeneity driven segregation of the reactants is treated as a function of Dahmköhler numbers in your modelling study.

1 L. Kaser *et al.*, *Geophys. Res. Lett.*, 2015, **42**, 10894–10903.
2 H. G. Ouwersloot, J. Vilà-Guerau de Arellano, C. C. van Heerwaarden, L. N. Ganzeveld, M. C. Krol and J. Lelieveld, *Atmos. Chem. Phys.*, 2011, **11**, 10681–10704.

Alison Tomlin replied: As explained in the paper, the model describes mixing by the Interaction by Exchange with the Mean (IEM) model where the mixing time-scale is determined by the mixing-constant and the ratio of turbulent kinetic energy to its dissipation. Therefore it is not possible to relate this directly to the Dahmköhler number. We accept that this is a simplification compared to considering the high temporal resolution intermittency that may be present in such situations. However, a compromise was sought since it was felt that LES would be too computationally expensive to study anything other than simple street canyon situations. And even with LES, a sub-grid scale turbulence model would be required below the 1 m grid resolution used. The model set-up was tested against experimental wind tunnel data (including a full global uncertainty analysis) for a NO_x plume dispersing in background ozone from an earlier paper by Ziehn *et al.*[1] What we found in this study was that the model gave good agreement with the measurements of mean NO_x and NO_2 concentrations but that the errors were higher for predictions of the concentration variances. We therefore feel that for the high model resolution used, the IEM model is a suitable compromise for predicting mean concentrations.

1 T. Ziehn, N. S. Dixon and A. S. Tomlin, *Atmos. Environ.*, 2009, **43**, 5978–5988.

Roberto Sommariva opened a general discussion of the paper by Lisa Whalley: Your models are run constrained to all measured species and parameters, except NO_2 and O_3. This approach is fine for short lived species, but not so much for long lived species, such as ozone. It takes a very long time for ozone to reach steady-state (7 days in the case of the MCM model!) and as a consequence, the concentrations of all the non-measured intermediates are very different at the end of the calculation of a single data point compared to at the beginning. This may lead to significant errors in the calculation of the ozone concentration and complicate the comparison with the O_3 observations, which is shown in your Fig. 3.

Lisa Whalley answered: In Fig. 3 we do not attempt to compare the steady state ozone concentration predicted by the box models to the observed ozone. We have avoided doing this, as it is unlikely in an urban centre such as London (with a multitude of NO_x and VOC emission points) that ozone will have reached steady state conditions. Instead we are comparing the steady state ozone concentrations predicted by the different mechanisms to each other, and assessing whether the differences observed are due to the model inputs used to constrain the mechanisms or due to the differences in the complexity of the chemistry mechanisms. We have chosen to run all the box models to steady state before comparison, so as to prevent any bias caused by variability in the length of time it may take for ozone concentrations to reach steady state when the complexity of the chemical mechanism differs. We have estimated *in situ* ozone production using the MCM box model after 1 h, which is the average time (since emission) it took an air-mass to reach the measurement site, and have compared this to the observed ozone and see that *in situ* production could account for ~ 60% of the ozone observed on average. During the transit from emission to the observation site oxidised VOC intermediates will have been generated within the air-mass. These oxidised products may not be detectable or identifiable by the GC systems used to measure the VOCs; nevertheless, the influence of these intermediate species should be considered when estimating the *in situ* ozone production.

Iarla Kilbane-Dawe remarked: I would like to come back to emissions modelling, rather than dynamics or concentrations. Given the residence time of the species in urban environments, the emissions are an important factor. It is really important to acknowledge how poor our understanding and characterisation of the emissions that are occurring in the urban environments are. Tim Murrels touched on some of these in a previous discussion session, but it is clear that many of the models and assumptions we use to implement spatially and temporally varying emissions, especially at high resolution in urban sites, are wrong – sometimes by as much as an entire order of magnitude for important species and partitioning of species within families. In many cases, important physical, chemical and socioeconomic effects are simply ignored.

The implication is that, since important emissions are wrong, then all of the adjustments being made to a model to make it agree with the verified concentration measurements are likely to be wrong. And by implication, much of the scientific effort we've seen today is being put in the wrong place.

Also, on the density point that Matthew Hort mentioned – I think it's timely to do high resolution modelling now, as I think we're about to see a great improvement in the quality of data available across cities.

Karine Sartelet commented: With regards to model uncertainty – we did an internal comparison of our model. We found that the internal uncertainties in the model were lower. Structural uncertainties were larger than parametric uncertainties. The uncertainty comes from the input data as well as the model itself.

Alison Tomlin replied: I agree that uncertainties originating from differences in model structures are also important. They are perhaps more difficult to quantify than parametric uncertainties since in principle an infinite number of model structures is possible (using, for example, different resolutions, turbulence models, sizes of chemical mechanisms, *etc.*). One approach taken in the climate change community is to perform regular model inter-comparison exercises. By including all available models run as ensembles, then an envelope of predictions can be given which at least traces uncertainties due to differences in model set-ups. The difficult next step is to trace the origins of these uncertainties back to specific parts of the models. For parametric uncertainty this is time consuming, but conceptually relatively straightforward using the ANOVA type approaches discussed in our paper. For structural uncertainties this is a bigger challenge.

Andreas Skouloudis said to James Lee: It would be interesting, also, to talk about the certainties of emissions inventories. For the London emissions inventory it would be very interesting for the younger generations, to carry out comparisons of inventories that were constructed with bottom-up synthesis, as well as with the disaggregation of top-down constructed inventories. But this work requires the synthesis of data from citizen monitoring, which has only been emerging in recent years together with the evolution of sensor technologies. This could inspire future research in this area and facilitate similar conclusions as we have seen in this important paper, and improve the accuracy of inventories from real-time data. I wonder if you would consider examining such types of comparisons in the future?

James Lee responded: I do not believe that, at present, data available from citizen monitoring (which will largely be based around small sensor technologies) is of sufficient quality to allow meaningful comparison with inventories. Even if the measurements prove that the concentrations measured by such systems are of good quality, this still does not give a direct comparison with emissions inventories. A model would be required to make the step from emission to concentration. I do think that personal monitors may have a role to play in the future in improving air pollutant data coverage across a city (especially one without such a large traditional network as London), something which could prove extremely useful for mitigation strategies.

Sarah Moller returned to the discussion of the papers by Andreas N. Skouloudis: You suggested that inventories at the local/city scale should be developed. That presents a major challenge to many areas in the UK. While the data available on sources and activity in London are reasonable, the same data are not collected for other cities and regions, for example the drive cycles developed for London and used in the London emission inventory. What are your thoughts on this? There is good data for some sources that could be used in regional/city scale inventories but there are also major pollution sources for which the geographical distribution in the UK national emission inventory is a poor representation of the actual situation. In these cases the activity data is often not available to give a better representation for inclusion in local inventories. Do you think it is feasible to expect cities to be able to produce reasonable inventories that are an improvement on what the national inventory can do? Are there good examples of this elsewhere in Europe that we could learn from?

Andreas Skouloudis answered: It is true that the emissions inventory data are not sufficiently good and there is always room for improvements. These improvements can be either directed towards improving the spatial resolution or even becoming more realistic on the temporal scale. It is true that in London human activities, transport, energy consumption *etc.* are detailed and very accurate. However, for the regulatory applications I have described in this work, it is necessary to have the same detail and temporary accuracy everywhere in Europe. Unfortunately, at the time when this work was carried out, the activity data were not all sufficiently reliable. In such cases, the comparison of emission inventories constructed from disaggregating top-down data with emission data from aggregating bottom-up data from different activities and sources can help in identifying the level of uncertainty. However, this is not going to resolve the actual situation of the modelling needs. We will need a process that calculates the emissions on the basis of citizens' actions in terms of real-time assessment of transport needs (*e.g.* from satellite and *in situ* images) and in terms of accurately monitoring the energy and telematic consumption everywhere. Since 2005 we have been working with big-data in automating this process and for this purpose it might be useful to consult our recent review publication[1.] This work shows some of the instruments that we could potentially use for improving the local activities and make it feasible to produce reasonable inventories. This may also be relevant to the comment made by Brian McDonald earlier.

1 P. Kumar, A. N. Skouloudis, M., M. Viana, M. Cristina Carotta, G. Biskos and L. Morawsk, *Sci. Total Environ.*, 2016, **560–561**, 150–159.

Brian McDonald remarked: On the question of bottom-up inventories and what we can do with them in the future: I think there's some hope here. I'm always amazed at how much data is available on human activity, from Google, our smartphones, *etc.* – "big data". It's not easy because some data may be proprietary, but there are opportunities here, and this kind of activity data could be useful in developing better maps of critical emission sources.

Gordon McFiggans commented: I disagree with Andreas Wahner that we should be attempting to include all known chemistry with full complexity in numerical models of the urban atmosphere. There are very many chemical, physical and dynamic processes competing for computational resources in numerical simulations. It will simply not be possible in a reasonable timeframe to explicitly treat all processes at the highest level of detail and complexity to answer all the questions of importance to our understanding of the urban atmosphere. Interpretation of kerbside and roadside exposure and street canyon variability of pollutant concentrations requires metre-scale resolution. A two-way coupled nested simulation covering the 5 domains shown in the first plot down to a 1 km resolution, including the intermediate chemical scheme in Lisa Whalley's paper (DOI: 10.1039/C5FD00218D) with detailed inorganic and organic size-resolved aerosol, runs at 5 times real time on the UK National Supercomputing Facility. The second plot, running at the 10 cm resolution required to capture steady-state CFD predictions of particle number concentration, including the semi-volatile redistribution of components in a street canyon, required 3 simulation months. These two very different approaches to answer very different questions of importance to urban air quality could not possibly include full complexity treatments of all processes. For the foreseeable future, it will be necessary to reduce the complexity of treatments in a manner appropriate to the questions being addressed. We need to choose the right tools for the job.

Andreas Skouloudis answered: I think that simulations with a CPU time of five times real-time can be still very useful. Especially if we identify periods of less than 4–5 days that are suitable for representing the annual mean of episodes with different percentile severity according to annual meteorology. We have been very successful in identifying such periods in several parts of the world and we are willing to help if asked.

Also, from your comment, it is not also clear if the experience of the modelling system is based on several processors and models that are designed, compiled and optimised to run on high performance parallel machines. Perhaps a scale between 1 km and 10 cm might be more rational, in particular if some of the meteorological/photochemical parameters become numerically unstable in spatial resolutions beyond 1 km. This is certainly an issue to be addressed elsewhere. Nevertheless, sometimes it might be better to wait between 20–25 days rather than taking regulatory decisions based on coarse models which in the end will lead to wrong abatement solutions or damaging regulatory interventions, as we have seen several times in the past two decades.

Iarla Kilbane-Dawe commented: Bottom-up inventories do make sense – there are quite rich data available from transport models, but there are difficulties in getting hold of them. We need to build relationships with local transport modellers. The scientific community needs to consider what the uncertainties are in the emissions factors tied to the activity data – and what the right models are to build these emissions factors. This is an effort where much greater effort is required.

Mohammed Salim Alam said: So what we have spoken about over the last few days is how there are too many unknowns in the characterisation of the emissions from diesel vehicles. The issue is that there is not much information in the literature on high molecular weights in particular, which is what we presented during the first session. A lot of what we currently "know" and have talked about is focused on low molecular weight species. There are far too many gaps in determining the composition of diesel exhaust emissions, particularly for high molecular weights, and I think that this is an area that should be focused on in order to determine its potential in SOA formation.

Matthew Hort remarked: Some of the challenge with mass sensors is that they can be of varying specification and also less than ideally located, resulting in varying quality of the return measurement. The meteorological community has been looking at tackling this issue, as high quality observations won't proliferate to the levels needed for high resolution modelling. One interesting development is in what is sometimes called opportunistic data (*i.e.* information from existing equipment that was built for another purpose). This is cheap, as it is an extra use of something already built, but it is often of high quality. This, I think, highlights an important additional dimension to cheap data, in that the sensors do not also need to be cheap. While this is a new area there may be lessons to be learnt from the Meteorological community and its extensive use of measurement networks.

The Met Office have a community data collection initiative and web site called WOW (Weather Observation Website):[1]Anyone can upload measurement data and information. The site now has thousands of additional measurement sites, over and above the official Met Office and other agency sites. This data, while from varying instruments and of varying quality and representivity, has proved very valuable in understanding the evolution of bad and intense weather events. The site is open to other sorts of measurements and in combining a range of environmental measurements it might offer a sensible and useful tool in expanding air quality observations.

1 Weather Observation Website, http://wow.metoffice.gov.uk.

James Lee commented: We are talking about two ways of using these cheap sensors. One way is to use a network of sensors to measure things like compliance, or to look at health effects of air pollution, and the other is using them for validation of models. The confidence of the scientific community in these sensors isn't high at the moment. I think people won't want to run computationally expensive models with measurements they don't have faith in.

Andreas Skouloudis answered: For the first aim, in order to show compliance, perhaps it will be sufficient to examine concentrations averaged over long time periods. For such purposes the passive sampler approach might not sufficient and certainly cheap. These can also be used for assessing chronic health effects.

However, for acute health effects there are no such cheap sensor products. Nanocrystalline sensors are expensive but sufficiently accurate. Their advantage is that they are mobile, and certainly more geographically representative, than expensive conventional monitoring.

Ruth Doherty opened a general discussion of the Concluding remarks by Jose-Luis Jimenez: Regarding your last slide on indoor air pollution, are the uncertainties large because we don't know enough about the indoor sources and their emissions?

Jose-Luis Jimenez replied: There has been very little study of indoor air chemistry, compared to outdoor air chemistry. If one searches, for example, Web of Science, the ratio of outdoor/indoor air chemistry publications is around 100. Many key variables (*e.g.* photolysis rates, OH, NO_3 radicals, HONO *etc.*) have never been measured indoors, or only once or twice. Many of the newer instruments developed for outdoor air research have never, or almost never, been used indoors either. Combined with the high diversity of indoor environments, materials, ways in which humans use these, *etc.*, and the importance of processes that are not relevant for outdoor air (emissions from materials, surface chemistry on materials and humans, impact of human breath) there are many questions about indoor chemistry that we know very little about. The Sloan Foundation in the US recognized these gaps and has started a funding program for indoor air chemistry,[1] which is how we have been able to do the initial measurements I discussed in my talk, but a much larger international effort is needed in this area in my opinion.

1 Alfred P. Sloan Foundation, STEM Research - Chemistry of Indoor Environments, http://www.sloan.org/major-program-areas/stem-research/chemistry-of-indoor-environments/.

Rachel Dunmore said: The graph comparing indoor and outdoor is normalised, can you comment on the scale of difference between the two when not normalised?

Jose-Luis Jimenez replied: The absolute total concentrations of the 115 measured acids were higher indoors, although the time periods were a few days apart, due to limitations on the location of the ToF-CIMS instrument for sampling indoor *vs.* outdoor air. This is consistent with prior literature results for formic and acetic acids, which are typically higher indoors. These results are being finalized for submission to ES&T, where the full details will be available.

Roberto Sommariva commented: You have talked about the problem of doing too many field campaigns and not spending enough time thinking about the data. This is in part due to human nature, and in part driven by the way the funding system works. Do you think there is something we – as individual scientists – can do to deal with this issue, or is it a structural problem of modern science that needs to be addressed at a higher level?

Jose-Luis Jimenez replied: This is a very good question, one that I have spent a considerable amount of time wondering about.

In terms of the causes: it seems to me that there are two main factors that drive the lack of sufficient funding for data analysis and modeling of field campaign data: one is that funding programs and the program managers who run them need "news" to justify funding of their particular program to their superiors, the broader government, and society (and to advance their own careers, in the case of the program managers). News of a new field campaign with an airplane, or ship, or containers and people in the field seem to sell a lot better than some additional papers from an "old" study done 4 years ago. A second factor is that for an individual scientist, leading a proposal for a new field campaign is a good way to advance one's own career, and also to increase one's chances of funding. The fact that things have been done this way for a long time has created a "culture" in the field, that almost all program managers and scientists seem to share, that, for example, "a field proposal gets you 3 years of funding". Of course one can write a proposal focused on data analysis, but the funding rate of those is typically much lower than of proposals to acquire new data. For example, for my own group, we have a 100% success rate on proposals to acquire new field measurements, but only a 25% success rate on (equally good, if not better) proposals to analyze data from past field studies. I have talked to multiple program managers about this problem, and they are always sympathetic – except when they call you and ask to cut your budget massively for the next field study (effectively greatly reducing analysis time), or risk being left out of the campaign.

This is a huge problem that leads to a lot of wasted effort and prevents some of the best science from being followed up on and published, as researchers are distracted into going into the field and producing new data, and end up abandoning all data. Even if one manages to get some extra funding for analysis of a given campaign, the fact that most other participants did not is a serious problem, as their responses to queries about their data and model drag, as they long ago ran out of funding for that campaign and are busy with other projects. In that situation, any analyses that would require significant time from other groups become unfeasible. As science becomes ever more complex and specialized, some of the most interesting findings happen on the 2nd or 3rd round of analysis, after the initial papers are out and people have had a chance to digest their own measurements and think about how they relate to findings from others in the same campaign, and the broader state of the field and other fields. The advent of very powerful but very complex instruments in our field (*e.g.* ToF-CIMS, FIGAERO-ToF-CIMS, TOF-PTRMS, SV-TAG *etc.*) will make this worse, as 3 years is woefully inadequate for datasets from such instruments during complex field (or lab) campaigns. Many such instruments produce data at the same rate (MB h^{-1}) as satellite instruments.

In terms of what can be done, I do believe that the funding agencies and the culture of our scientific field can be changed, even if incrementally. But many of us need to keep making the case many times, with specific examples of how the short timeline of funding is damaging to scientific output. This has happened in some meetings I have attended, but much more is needed. In the medium term, I think we will make some progress and lessen the problem, as long as enough people join in with these efforts. But the problem will probably remain serious by the time I retire.