Novel informatics software for automated individual aerosol component property predictions and complex ensemble predictions – an online community facility

D. Topping[1,3], M. Bane[2], M. Barley[3], D. Lowe[3], R. Pinning[2] and G. McFiggans[3]
[1]National Centre for Atmospheric Science, University of Manchester, Manchester, UK
[2]IT Services, University of Manchester, Manchester, UK
[3]Centre for Atmospheric Science, University of Manchester, Manchester, UK
Keywords: Aerosol modeling, Organic aerosols, Informatics.
Presenting author email: david.topping@manchester.ac.uk

Introduction:
Atmospheric aerosols are known to be an important yet highly uncertain component of the earth's climate and air quality (Solomon et al 2007). Owing to the complexity and diversity of atmospheric aerosol components, quantification of the properties that determine their highly uncertain climatic and human health impacts requires the development of novel technological applications. Inorganic material is restricted to a few well-understood compounds. However, organic material can comprise many thousands, as yet largely unidentified, compounds with a vast range of properties.

We must be able to predict however many thousands of components can exist in particulate matter. Predicting the evolution of aerosol requires calculating the distribution of all components between the gas and aerosol phase. This requires knowledge of all component vapour pressures and other thermodynamic properties. The many thousands of individual aerosol components ensure that explicit manual calculation of these properties is laborious and time- consuming. The emergence of explicit automatic mechanism generation techniques (Aumont, 2005; Jenkin, 2003) including up to many millions of individual gas phase products as aerosol precursors renders the process impossible and automation is necessary.

It can require significant resource to develop methods for automating these processes, but also to develop efficient numerical frameworks that drive the complex calculations. In this presentation we showcase development and application of a new online facility that combines benchmark aerosol models with open source informatics software to provide a new community service.

The model system:
The facility is written in python, a high-level language is widely used as a scripting language for web applications. It supports both fixed-point and floating-point numeric calculations, string manipulation and facilities for structures and multi-dimensional arrays. It is distributed under a permissive free software license and is available for all major operating systems, which will enable our software to be highly portable. Also, a PYTHON interface has been developed for the chemical informatics software of our choice (Open Babel) as described by O'Boyle et al (2008).
The SMILES chemical informatics syntax is used to automate extraction of molecular substructure information from individual compound identifiers (figure 1). Commonly employed in commercial and public software, it can be imported by most molecular editors and has a wide base of software support and extensive theoretical backing (www.daylight.com).

We will demonstrate the need for developing libraries that capture all functionality of compounds expected in the atmosphere, a problem that occurs in other available online services used by the community.

Example applications:
Important uncertainties are in those parameters which dictate the aerosol water content and gas / aerosol partitioning. To this end we will demonstrate the use of this facility to calculate fundamental properties, such as pure component vapour pressures, for an unlimited set of compounds. Absorptive partitioning simulations carried out on the facility will demonstrate the ability to predict secondary organic aerosol loading and composition using output from complex gas phase degradation mechanisms. Specifically we will use output from the Master Chemical Mechanism (MCM Jenkin (2003)) to simulate the detailed composition of aerosol in multiple scenarios.

Future outlook:
The facility is also designed to act as a springboard for collaborative purposes, the inclusion of specific visualization techniques enabling the modeling and measurement community to directly compare outputs from models and instruments. In addition we will present progress on using GPGPUs (General purpose graphical processing units) to obtain large increases in computational performance on the complex models hosted on the website. This is an exciting area of research and is already being explored for chemical models that are often heavily reduced to reduce numerical complexity.

This work is supported by the National Centre for Atmospheric Science (NCAS), NERC grant NE/H002588/1 and NERC grant NE/J013471/1.

Aumont, B., Szopa, S., and Madronich, S. Atmos. Chem. Phys., 5, 2497-2517, (2005).
Jenkin, M. E., Saunders, S. M., Wagner, V., and Pilling, M. J.: Atmospheric Chemistry and Physics, 3, 181-193, (2003).
O'Boyle,N.M. Morley, C., and Hutchison, G. R.: Chem. Cent. J., 2, 1–7. (2008).
Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L., IPCC: Climate change 2007