Predicting arene rate coefficients with respect to hydroxyl and other free radicals in the gas-phase: a simple and effective method using a single topological descriptor


1School of Earth, Atmospheric and Environmental Sciences, The University of Manchester, The Sackville Building, Sackville Street, P.O. Box 88, Manchester M60 1QD, UK
2Department of Air Quality and Climate Change, Energy Research Centre for the Netherlands, P.O. Box 1, 1755 ZG Petten, The Netherlands
3Biogeochemistry Research Centre, School of Chemistry, The University of Bristol, Cantock’s Close BS8 1TS, UK

Received: 13 November 2006 – Published in Atmos. Chem. Phys. Discuss.: 26 February 2007
Revised: 23 May 2007 – Accepted: 2 July 2007 – Published: 6 July 2007

Abstract. The reactivity of aromatic compounds is of great relevance to pure and applied chemical disciplines, yet existing methods for estimating gas-phase rate coefficients for their reactions with free radicals lack accuracy and universality. Here a novel approach is taken, whereby strong relationships between rate coefficients of aromatic hydrocarbons and a Randić-type topological index are investigated, optimized and developed into a method which requires no specialist software or computing power.

Measured gas-phase rate coefficients for the reaction of aromatic hydrocarbons with OH radicals were correlated with a calculated Randić-type index, and optimized by including a term for side chain length. Although this method is exclusively for use with hydrocarbons, it is more diverse than any single existing methodology since it incorporates alkenylbenzenes into correlations, and can be extended towards other radical species such as O(3P) (and tentatively NO3, H and Cl). A comparison (with species common to both techniques) is made between the topological approach advocated here and a popular approach based on electrophilic substituent constants, where it compares favourably.

A modelling study was carried out to assess the impact of using estimated rate coefficients as opposed to measured data in an atmospheric model. The difference in model output was negligible for a range of NOx concentrations, which implies that this method has utility in complex chemical models.

Strong relationships (e.g. for OH, $R^2=0.96$) between seemingly diverse compounds including benzene, multisubstituted benzenes with saturated, unsaturated, aliphatic and cyclic substitutions and the nonbenzenoid aromatic, azulene suggests that the Randić-type index presented here represents a new and effective way of describing aromatic reactivity, based on a quantitative structure-activity relationship (QSAR).

1 Introduction

Since its conception, the term “aromatic” used in description of a specific class of organic compound has been subjected to many refinements, and where ambiguity remains (e.g. where Hückel’s rule is not obeyed or where the compound in question is of nonplanar geometry) consensus is met chiefly on the empirical grounds of the characteristic reactivity of aromatic compounds in general. The present study uses a quantitative structure-activity relationship to describe the reactivity of a subset of these compounds, the arenes, whose gas-phase rate coefficients have up to now appeared to have no discernable correlation from knowledge of their constitution alone.

The reactivity of aromatic hydrocarbons with free radicals is of direct importance to a variety of chemical subdisciplines including atmospheric chemistry, chemical epidemiology and combustion processes. Motor vehicle emissions and fuel spillage are a major source of arenes in the urban environment (Calvert et al., 2002), resulting from the high proportion of arenes present in gasoline (especially in European and diesel formulations) and the close relationship between fuel composition and exhaust composition with respect to aromatics (Leppard et al., 1993, 1995). Arenes are classified as carcinogenic by the International Agency for Research on Cancer (IARC) and exposure has been linked to dermal
absorption and subsequent biotransformation into harmful metabolites (Pellack-Walker and Blumer, 1986); leukaemia (Carletti and Romano, 2002), cytogenetic damage (Celik et al., 2003) and persisting neurobehavioural disturbances in humans (Mikkelsen, 1997) and laboratory rats (Gralewicz and Wiaderna, 2001).

Aromatic hydrocarbons are common trace components of the urban atmosphere, and given their associated health risk, their primacy as tropospheric ozone precursors (Derwent et al., 1996) and the contributions of their oxidation products towards secondary organic aerosol (SOA) formation (Odum et al., 1997), it is highly pertinent to investigate factors affecting the atmospheric lifetime (e.g. their reactivity towards free radicals) and the volatility of oxidation products (e.g. the branching ratio between radical addition and radical abstraction reactions) of aromatic hydrocarbons.

The present study describes an accurate and accessible method by which unknown rate coefficients can be estimated for the reaction of arenes with OH and O(3P) radicals and provisional estimation methods for certain arene species with H and Cl (where data are sparse).

2 Methodology

Methods for calculating the Randić index for saturated and unsaturated hydrocarbons are described elsewhere in the literature (Randić, 1975; McGillen et al., 2006a, b). The present method is an extension of that used to calculate the Randić index for unsaturated hydrocarbons (McGillen et al., 2006b), whereby a Kekulé structure is assumed for a phenyl group. Although not a true representation of the benzene ring, the alternating single and double bond arrangement of the Kekulé structure assigns the same hydrogen-suppressed valency to each carbon atom of the aromatic ring and the valency value that is assigned is therefore unimportant, since any error associated with carbon valence is systematic and is of no concern to a correlation study such as this.

Table 1. Examples of bicyclic and monocyclic arenes belonging to \( \sum x°_{0-2} \).

<table>
<thead>
<tr>
<th>Arene</th>
<th>( \sum x°_{0-2} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>indene</td>
<td>0</td>
</tr>
<tr>
<td>toluene</td>
<td>0</td>
</tr>
<tr>
<td>indane</td>
<td>1</td>
</tr>
<tr>
<td>ethylbenzene</td>
<td>1</td>
</tr>
<tr>
<td>tetralin</td>
<td>2</td>
</tr>
<tr>
<td>propylbenzene</td>
<td>2</td>
</tr>
</tbody>
</table>

The Randić index \( R \) is calculated using Eq. (1)

\[
R = \sum(mn)^{-0.5}
\]

where \( m \) and \( n \) are valencies of adjacent vertices joined by an edge. The following is a worked example for the structural graph of indene, (see Fig. 1).

The valencies are 4 for vertices 1 and 5, 2 for vertex 4 and 3 for vertices 2, 3, 6, 7, 8 and 9. Using Eq. (1), \( R \) is calculated:

\[
R = (4 \times 3)^{-0.5} + (3 \times 3)^{-0.5} + (3 \times 3)^{-0.5} \\
+ (3 \times 2)^{-0.5} + (2 \times 4)^{-0.5} + (4 \times 4)^{-0.5} \\
+ (4 \times 4)^{-0.5} + (4 \times 3)^{-0.5} + (3 \times 3)^{-0.5} \\
+ (3 \times 3)^{-0.5} + (3 \times 3)^{-0.5} + (3 \times 3)^{-0.5} \\
+ (3 \times 3)^{-0.5} + (4 \times 4)^{-0.5}
\]

Where \( m \) and \( n \) correspond to vertices 1 and 2, 2 and 3, 2 and 3, 3 and 4, 4 and 5, 5 and 1, 5 and 1, 5 and 6, 6 and 7, 6 and 7, 7 and 8, 8 and 9, 9 and 8, 9 and 9 and 9 and 1 respectively, giving the result 4.46.

One drawback of using the Kekulé structure in this approach is that certain arenes (notably those containing an ortho-substitution) will generate a slightly different index (typically \( \pm 0.01 \)) depending on which of the two resonance structures is used in calculation of the index. Although such a minor difference is considered inconsequential for predictive purposes, for thoroughness these two values have been averaged in this study.

Similar to a phenomenon observed previously for alkenes (McGillen et al., 2006b), species plot on separate trendlines

![Fig. 1. Structural graph of indene.](image-url)
Table 2. Measured arene + radical room temperature rate coefficients used in this study, alongside Randić-type index and $\Sigma x^o$ subgroup for each respective arene.

<table>
<thead>
<tr>
<th>arene</th>
<th>$k_{OH}$</th>
<th>$k_{O(3P)}$</th>
<th>$k_{NO_3}$</th>
<th>$k_H$</th>
<th>$k_{Cl}$</th>
<th>$R$</th>
<th>$\Sigma x^o$</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene</td>
<td>$1.25 \times 10^{-12}$ (a)</td>
<td>$3.16 \times 10^{-14}$ (l)</td>
<td>$5.65 \times 10^{-14}$ (l)</td>
<td>$1.30 \times 10^{-15}$ (v.)</td>
<td>$3.000$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>toluene</td>
<td>$5.70 \times 10^{-12}$ (b)</td>
<td>$7.63 \times 10^{-14}$ (l)</td>
<td>$7.00 \times 10^{-17}$ (q.)</td>
<td>$1.35 \times 10^{-15}$ (l.)</td>
<td>$3.366$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>p-xylene</td>
<td>$1.30 \times 10^{-11}$ (c)</td>
<td>$2.15 \times 10^{-13}$ (l)</td>
<td>$4.97 \times 10^{-16}$ (r.)</td>
<td>$5.80 \times 10^{-13}$ (u.)</td>
<td>$3.732$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>o-xylene</td>
<td>$1.14 \times 10^{-11}$ (a)</td>
<td>$1.83 \times 10^{-13}$ (m.)</td>
<td>$4.13 \times 10^{-16}$ (r.)</td>
<td>$1.50 \times 10^{-10}$ (v.)</td>
<td>$3.741$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>m-xylene</td>
<td>$2.09 \times 10^{-11}$ (c)</td>
<td>$3.98 \times 10^{-13}$ (m.)</td>
<td>$2.60 \times 10^{-16}$ (q.)</td>
<td>$1.50 \times 10^{-10}$ (v.)</td>
<td>$3.732$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>mesitylene</td>
<td>$5.75 \times 10^{-11}$ (c)</td>
<td>$2.60 \times 10^{-12}$ (n.)</td>
<td>$8.76 \times 10^{-16}$ (r.)</td>
<td>$1.40 \times 10^{-10}$ (v.)</td>
<td>$4.098$</td>
<td>$0$</td>
<td></td>
</tr>
<tr>
<td>hemimellitene</td>
<td>$3.27 \times 10^{-11}$ (c)</td>
<td>$1.13 \times 10^{-12}$ (n.)</td>
<td>$1.86 \times 10^{-15}$ (r.)</td>
<td>$4.116$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi$-cumene</td>
<td>$3.25 \times 10^{-11}$ (c)</td>
<td>$1.02 \times 10^{-12}$ (n.)</td>
<td>$1.81 \times 10^{-15}$ (r.)</td>
<td>$4.107$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>styrene</td>
<td>$5.86 \times 10^{-11}$ (d)</td>
<td>$4.50 \times 10^{-12}$ (a.)</td>
<td>$1.51 \times 10^{-13}$ (d.)</td>
<td>$3.49 \times 10^{-12}$ (l.)</td>
<td>$3.60 \times 10^{-10}$ (v.)</td>
<td>$4.309$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$-methylstyrene</td>
<td>$5.30 \times 10^{-11}$ (e)</td>
<td>$4.32 \times 10^{-12}$ (l.)</td>
<td>$4.32 \times 10^{-12}$ (l.)</td>
<td>$4.444$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$-dimethylstyrene</td>
<td>$3.30 \times 10^{-11}$ (f)</td>
<td>$4.45 \times 10^{-12}$ (l.)</td>
<td>$4.58 \times 10^{-12}$ (g.)</td>
<td>$4.951$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trans-polyphenylbenzene</td>
<td>$5.99 \times 10^{-11}$ (e)</td>
<td>$4.32 \times 10^{-12}$ (l.)</td>
<td>$4.45 \times 10^{-12}$ (l.)</td>
<td>$4.444$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2-propenylbenzene</td>
<td>$7.80 \times 10^{-11}$ (g.)</td>
<td>$4.10 \times 10^{-12}$ (g.)</td>
<td>$4.10 \times 10^{-12}$ (g.)</td>
<td>$4.455$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>indene</td>
<td>$2.72 \times 10^{-10}$ (h.)</td>
<td>$3.90 \times 10^{-10}$ (h.)</td>
<td>$3.90 \times 10^{-10}$ (h.)</td>
<td>$4.982$</td>
<td>$0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ethylbenzene</td>
<td>$6.07 \times 10^{-12}$ (a)</td>
<td>$5.84 \times 10^{-14}$ (l.)</td>
<td>$2.49 \times 10^{-13}$ (l.)</td>
<td>$3.927$</td>
<td>$1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>toluene, p-ethyl</td>
<td>$1.36 \times 10^{-11}$ (i.)</td>
<td>$8.58 \times 10^{-16}$ (s.)</td>
<td>$8.58 \times 10^{-16}$ (s.)</td>
<td>$4.293$</td>
<td>$1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>toluene, o-ethyl</td>
<td>$1.32 \times 10^{-11}$ (i.)</td>
<td>$8.58 \times 10^{-16}$ (s.)</td>
<td>$8.58 \times 10^{-16}$ (s.)</td>
<td>$4.299$</td>
<td>$1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>toluene, m-ethyl</td>
<td>$2.24 \times 10^{-11}$ (i.)</td>
<td>$7.35 \times 10^{-15}$ (g.)</td>
<td>$7.35 \times 10^{-15}$ (g.)</td>
<td>$4.445$</td>
<td>$1$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>indane</td>
<td>$1.91 \times 10^{-11}$ (g.)</td>
<td>$6.61 \times 10^{-12}$ (i.)</td>
<td>$6.61 \times 10^{-12}$ (i.)</td>
<td>$4.427$</td>
<td>$2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n-propylbenzene</td>
<td>$5.71 \times 10^{-12}$ (j.)</td>
<td>$6.77 \times 10^{-14}$ (p.)</td>
<td>$6.77 \times 10^{-14}$ (p.)</td>
<td>$4.309$</td>
<td>$2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cumene</td>
<td>$1.51 \times 10^{-11}$ (k.)</td>
<td>$1.00 \times 10^{-15}$ (k.)</td>
<td>$1.00 \times 10^{-15}$ (k.)</td>
<td>$4.675$</td>
<td>$2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>tetralin</td>
<td>$3.44 \times 10^{-11}$ (d.)</td>
<td>$8.50 \times 10^{-15}$ (q.)</td>
<td>$8.50 \times 10^{-15}$ (q.)</td>
<td>$4.951$</td>
<td>$2$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>benzene, tert-butyl</td>
<td>$4.85 \times 10^{-12}$ (l.)</td>
<td>$6.64 \times 10^{-14}$ (n.)</td>
<td>$6.64 \times 10^{-14}$ (n.)</td>
<td>$4.616$</td>
<td>$3$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


according to the total number of primary (1°), secondary (2°) and tertiary (3°) carbon atoms ($\Sigma x^o$) present in the side chain(s), examples of these subgroups based on $\Sigma x^o$ are given in Table 1.

Correlation was improved using a nonlinear regression method to model these data, and incorporating a term for $\Sigma x^o$ using Eq. (2)

$$\ln k = a_0 \ln R + a_1 + b_0 \ln (D + 1)$$

(2)

where $R$ is the Randić index, $D$ is $\Sigma x^o$ and $a_0$, $a_1$ and $b_0$ are the model parameters, which can be determined by minimizing the $l^2$-norm of the difference between the rate coefficient measurements and regression model output.

3 Results

Table 2 lists all known measured room temperature rate coefficients for the reaction of C6-C10 arenes with OH, O(3P), NO3, H and Cl radicals together with the respective $\Sigma x^o$ subgroup and calculated Randić index for each species.

3.1 OH dataset

The OH dataset is the most comprehensive of those covered, and its range of reactivity extends over two orders of magnitude. Rate coefficients for a large variety of species are measured, and the dataset therefore offers the most insight into the distribution of $\Sigma x^o$ subgroups.

Figure 2 is a plot of Randić index against the logarithm of the room temperature rate coefficient ($\log k_{298}$). When
which results from contributions to the rate coefficients from rate coefficients of $m$ contribute most towards the scatter, shown by the enhanced is observed. Species possessing respectively. An excellent correlation coefficient of 0.96 ing Eq. (2) where $a$ and $b_0$ and $b_1$ are 13.15, −46.04 and −1.62 respectively. Again, a good correlation is observed and it is expected that with supplementary data for $\Sigma x_{0-3}^\circ$, a better correlation could be achieved by improving constraints on model parameters and spacing between $\Sigma x_{0-3}^\circ$.

### 3.2 O($^3P$) dataset

The second best correlation is observed for the O($^3P$) dataset with $R^2=0.92$, furthermore the dataset possesses a similar range of reactivity to the OH dataset, although does not cover as wide a range in terms of structural diversity.

Figure 4 is a plot of log $k_{298}$ against log $k_{298}$ modelled using Eq. (2) where $a_0$, $b_0$ and $b_1$ are 13.15, −46.04 and −1.62 respectively. Again, a good correlation is observed and it is expected that with supplementary data for $\Sigma x_{0-3}^\circ$, a better correlation could be achieved by improving constraints on model parameters and spacing between $\Sigma x_{0-3}^\circ$.

### 3.3 NO$_3$ dataset

The NO$_3$ dataset covers the largest range of reactivity, extending over seven orders of magnitude and is the second most extensive dataset in this study.

Figure 5 is a plot of Randić index against log $k_{298}$ with respect to NO$_3$. Unlike the other datasets, an inflection is observed in the data between alkylbenzenes and alkenylbenzenes. Treated as a single trendline, this phenomenon clearly contributes to scatter within the correlation of $\Sigma x_{0-3}^\circ$. Although data appear to segregate into $\Sigma x_{0-3}^\circ$ subgroups as

---

**Fig. 2.** A plot of Randić index against measured arene + OH log $k_{298}$. Data is seen to subdivide into approximately evenly spaced, parallel subgroups, dependent on $\Sigma x_{0-3}^\circ$.

**Fig. 3.** Log-log plot of arene + OH $k_{298}$ modelled using Eq. (2) against measured $k_{298}$.

**Fig. 4.** Log-log plot of arene + O($^3P$) $k_{298}$ modelled using Eq. (2) against measured $k_{298}$.

plotted with an unmodified Randić index, data are observed to subdivide into parallel and equally offset lines, similar to correlations observed in alkene data (McGillen et al., 2006b). Scatter within respective $\Sigma x_{0-3}^\circ$ is minimal and the gradients of trendlines are in good agreement.

Figure 3 is a plot of log $k_{298}$ against log $k_{298}$ modelled using Eq. (2) where $a_0$, $b_0$ and $b_1$ are 10.21, −38.51 and −1.11 respectively. An excellent correlation coefficient of 0.96 is observed. Species possessing $m$-substituents appear to contribute most towards the scatter, shown by the enhanced rate coefficients of $m$-xylene, mesitylene and $m$-ethyltoluene, which results from contributions to the rate coefficients from the stability of the resonance structures of $m$-substituted radical transition states. The alkenylbenzene, $\beta$-dimethylstyrene is not included in these correlations because of its highly anomalous rate coefficient. The unusually slow rate coefficient of this compound was noted by the original experimental investigators (Chiorboli et al., 1983), who attributed the non-planar geometry of this compound’s two methyl groups as a contributor to the steric hindrance of the molecule with respect to its otherwise reactive substituent unsaturated bond.
4 Discussion

The hydroxyl radical dataset is most comprehensive both in terms of number of rate coefficients and range of arene substitutions studied. As such, it is the most insightful with respect to the distribution of $\Sigma x^\circ$ subgroups and is used as a template for other radicals whose behaviour is assumed to be analogous (e.g. $O(^3P)$, H and Cl).

The hydroxyl dataset is likewise most useful for testing the effectiveness of the topological approach, since the established arene rate coefficient estimation method of Zetzsch (1982) and optimized further by Atkinson (1991), based on electrophilic substituent constants of Brown and Okamoto (1958) appears to have been developed solely for OH-arene reactions. A direct comparison between the topological approach and Zetzsch’s method is impossible, since the two techniques predict for a different range of aromatic compounds: the topological approach can describe alkenyl substitution whereas Zetzsch’s method cannot, and Zetzsch’s method considers oxygenated and nitrated substitution whereas the topological approach, at the time of writing has not been developed to do so. However, where overlap exists between the two methods (i.e. the alkenylbenzenes), a comparison can be made, as is shown in Table 3. It is considered most intuitive to assess the predictive capabilities of the two methods by expressing a ratio relationship, $r$, between the estimated rate coefficient and the measured rate coefficient using Eq. (3).

$$ r = \frac{k_{\text{estimated}}}{k_{\text{measured}}} \tag{3} $$

To ensure a direct comparison, where the measured rate coefficient exceeds the estimated rate coefficient, $r$ is inverted. As is evident from histograms of these data shown in Figs. 6a–b, the predictive capability of the topological approach is significantly better than Zetzsch’s method, which exhibits a larger range of error and a more diffuse distribution across this range.

The range of arenes incorporated into the hydroxyl radical correlation surpasses Zetzsch’s method in the sense that it includes alkenyl substituted species. This is a surprising result since the literature suggests that alkenyl substituted arenes react almost exclusively through radical addition to the substituted unsaturated bond (Calvert et al., 2002; personal communication, R. Atkinson, 20061). A corollary of this hypothesis is that reactions involving alkenyl substituted arenes ought to react via a distinctly different mechanism to the remaining arenes, where radical addition to the aromatic ring, hydrogen abstraction from the substituent group(s) and hydrogen abstraction from the aromatic ring are possible mechanisms. The correlations observed in Figs. 2 and 3 appear to contradict this hypothesis, since alkenyl substituted arenes

---

1Atkinson, R.: University of California, Riverside, California, CA, personal communication, 2006.
and Cl datasets, it appears that these radicals behave in an analogous way to the OH radical, as observed previously by Wayne et al. (1990). This is especially apparent in the $O(3P)$ correlation, where a similar range of reactivity is observed, and members of $\Sigma x_{\text{r,3}}$ plot in their respective fields on a scatter graph of $\log k_{298}$ vs. Rapić index. However, each $\Sigma x_{\text{r,0}}$ contains only one data point per subgroup, assuming there is scatter among these subgroups, the position of the respective trendlines of $\Sigma x_{\text{r,0}}$ is therefore rather conjectural at the time of writing, but could be constrained more accurately with further kinetic measurements of $\Sigma x_{\text{r,0}}$ species. With this in mind, the method used for modelling the hydroxyl radical rate coefficients cannot be adapted easily to the other classes of compound with the present dataset because of uncertainties associated with deriving the model parameters in

---

Table 3. Comparison between the predictive capabilities of the topological approach and the Zetzsch method using the ratio relationship, $r$ between rate coefficients of the respective estimation method and measured rate coefficients.

<table>
<thead>
<tr>
<th>arene</th>
<th>$k_{\text{OHmeas.}}$</th>
<th>$k_{\text{OHmod.}}$</th>
<th>$r$</th>
<th>$k_{\text{OHZetzsch}}$</th>
<th>$r$</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene</td>
<td>$1.19 \times 10^{-12}$</td>
<td>$1.71 \times 10^{-12}$</td>
<td>1.44</td>
<td>$2.30 \times 10^{-12}$</td>
<td>1.93</td>
</tr>
<tr>
<td>toluene</td>
<td>$5.70 \times 10^{-12}$</td>
<td>$4.61 \times 10^{-12}$</td>
<td>1.24</td>
<td>$5.90 \times 10^{-12}$</td>
<td>1.04</td>
</tr>
<tr>
<td>$p$-xylene</td>
<td>$1.30 \times 10^{-11}$</td>
<td>$1.24 \times 10^{-11}$</td>
<td>1.05</td>
<td>$7.10 \times 10^{-12}$</td>
<td>1.83</td>
</tr>
<tr>
<td>$o$-xylene</td>
<td>$1.14 \times 10^{-11}$</td>
<td>$1.27 \times 10^{-11}$</td>
<td>1.12</td>
<td>$7.10 \times 10^{-12}$</td>
<td>1.61</td>
</tr>
<tr>
<td>$m$-xylene</td>
<td>$2.09 \times 10^{-11}$</td>
<td>$1.24 \times 10^{-11}$</td>
<td>1.68</td>
<td>$1.50 \times 10^{-11}$</td>
<td>1.39</td>
</tr>
<tr>
<td>mesitylene</td>
<td>$5.75 \times 10^{-11}$</td>
<td>$3.34 \times 10^{-11}$</td>
<td>1.72</td>
<td>$3.82 \times 10^{-11}$</td>
<td>1.51</td>
</tr>
<tr>
<td>hemimellitene</td>
<td>$3.27 \times 10^{-11}$</td>
<td>$3.51 \times 10^{-11}$</td>
<td>1.07</td>
<td>$1.83 \times 10^{-11}$</td>
<td>1.79</td>
</tr>
<tr>
<td>$\psi$-cumene</td>
<td>$3.25 \times 10^{-11}$</td>
<td>$3.42 \times 10^{-11}$</td>
<td>1.05</td>
<td>$1.83 \times 10^{-11}$</td>
<td>1.78</td>
</tr>
<tr>
<td>ethylbenzene</td>
<td>$6.07 \times 10^{-12}$</td>
<td>$5.26 \times 10^{-12}$</td>
<td>1.15</td>
<td>$5.60 \times 10^{-12}$</td>
<td>1.08</td>
</tr>
<tr>
<td>toluene $p$-ethyl</td>
<td>$1.36 \times 10^{-11}$</td>
<td>$1.54 \times 10^{-11}$</td>
<td>1.13</td>
<td>$7.10 \times 10^{-12}$</td>
<td>1.92</td>
</tr>
<tr>
<td>toluene $o$-ethyl</td>
<td>$1.32 \times 10^{-11}$</td>
<td>$1.58 \times 10^{-11}$</td>
<td>1.20</td>
<td>$7.10 \times 10^{-12}$</td>
<td>1.86</td>
</tr>
<tr>
<td>toluene $m$-ethyl</td>
<td>$2.24 \times 10^{-11}$</td>
<td>$1.54 \times 10^{-11}$</td>
<td>1.46</td>
<td>$1.43 \times 10^{-11}$</td>
<td>1.57</td>
</tr>
<tr>
<td>$n$-propylbenzene</td>
<td>$5.71 \times 10^{-12}$</td>
<td>$7.24 \times 10^{-12}$</td>
<td>1.27</td>
<td>$7.50 \times 10^{-12}$</td>
<td>1.31</td>
</tr>
<tr>
<td>cumene</td>
<td>$6.61 \times 10^{-12}$</td>
<td>$5.09 \times 10^{-12}$</td>
<td>1.30</td>
<td>$4.63 \times 10^{-12}$</td>
<td>1.25</td>
</tr>
<tr>
<td>$p$-cymene</td>
<td>$1.51 \times 10^{-11}$</td>
<td>$1.53 \times 10^{-11}$</td>
<td>1.01</td>
<td>$1.21 \times 10^{-11}$</td>
<td>1.25</td>
</tr>
<tr>
<td>tetralin</td>
<td>$3.44 \times 10^{-11}$</td>
<td>$3.50 \times 10^{-11}$</td>
<td>1.01</td>
<td>$1.14 \times 10^{-11}$</td>
<td>3.02</td>
</tr>
<tr>
<td>benzene, tert-butyl</td>
<td>$4.85 \times 10^{-12}$</td>
<td>$6.78 \times 10^{-12}$</td>
<td>1.40</td>
<td>$4.30 \times 10^{-12}$</td>
<td>1.13</td>
</tr>
</tbody>
</table>


---

Fig. 6. (a) Histogram of the ratio relationship, $r$ between rate coefficients estimated using the topological approach and measured rate coefficients. (b) Histogram of the ratio relationship, $r$ between rate coefficients estimated using the Zetzsch method and measured rate coefficients.

fit seamlessly into the $\Sigma x_{\text{r}}$ trendline, a phenomenon that is observed in all datasets studied except for that of the nitrate radical. Short of comprehensive product analyses, this study can only offer a preliminary interpretation of this result, viz. that there are common mechanisms between alkenyl substituted arenes and other arenes. This hypothesis is supported by electron density calculations, which were executed using the MOPAC 5.0 (Stewart, 1983) package, implementing PM3 parameterizations. Such calculations indicate that aromatic rings possess comparable electron density to their alkenyl substituents, and ought therefore to possess similar reactivity.

From the strong correlations observed in the $O(3P)$, H and Cl datasets, it appears that these radicals behave in an analogous way to the OH radical, as observed previously by Wayne et al. (1990). This is especially apparent in the $O(3P)$ correlation, where a similar range of reactivity is observed, and members of $\Sigma x_{\text{r,3}}$ plot in their respective fields on a scatter graph of $\log k_{298}$ vs. Rapić index. However, each $\Sigma x_{\text{r,0}}$ contains only one data point per subgroup, assuming there is scatter among these subgroups, the position of the respective trendlines of $\Sigma x_{\text{r,0}}$ is therefore rather conjectural at the time of writing, but could be constrained more accurately with further kinetic measurements of $\Sigma x_{\text{r,0}}$ species. With this in mind, the method used for modelling the hydroxyl radical rate coefficients cannot be adapted easily to the other classes of compound with the present dataset because of uncertainties associated with deriving the model parameters in
Eq. (2). However, the position of the $\Sigma x^0$ trendline is unlikely to change significantly, and it is recommended that, as a provisional method for $\Sigma x^0$ species, a simple linear regression fit between experimental rate coefficient data and the Randić index will produce estimated rate coefficients of satisfactory accuracy, given by the equation of the straight line, $y = mx + c$ of the linear regression. Where $y$ is $\log k_{298}$, $x$ is the Randić index of the compound in question, the gradient $m$ is 1.648, 1.359 and 1.052 for $O(P)$, H and Cl respectively and the $y$-intercept $c$ is $-18.628$, $-17.359$ and $-13.836$ for $O(P)$, H and Cl respectively. For smaller datasets, H and Cl, this method should be used with caution, since the small sample size precludes a confident analogy with the hydroxyl radical.

The nitrate radical dataset exhibits different and more complex behaviour than the OH, O($P)$, H and Cl radical datasets. This is to be expected, since the branching ratio between the reaction pathways of the nitrate radical with arenes are known to differ considerably from those of the hydroxyl radical and likely other radicals of similar behaviour - the major difference being the negligibility of the reaction of the NO$_3$ radical with arene ring systems (Wayne et al., 1990). Evidence for this lack of interaction with the ring system is presented by the reversal of the trend of reactivity in ortho-, meta- and para-substituted alkylbenzenes observed in OH-arene reactions (Wayne et al., 1990), whereby meta-substituted species are slowest, despite increased activation of reactive sites in the ring system. In contrast, it appears that interactions with the aromatic ring system are of primary importance in the reaction of arenes with the hydroxyl radical, since meta-substitution enhances rates of reaction significantly. Since NO$_3$ attack must therefore occur predominantly on arene substituent sites, the comparison between alkyl and alkényl substituted arenes becomes analogous to a comparison between NO$_3$-alkane and NO$_3$-alkene reactions, which are kinetically and mechanistically dissimilar. It is therefore reassuring to note that the $\Sigma x^0$ trendline of the NO$_3$ dataset (see Fig. 5) shows significant scatter, which results from the deviation of alkényl substituted arenes, styrene and indene. It is also notable that the difference in gradient in the correlation of $\log k_{298}$ and Randić index between alkylbenzene and alkénylbenzene trendlines, denoted by $\theta$ in Fig. 5, is almost identical to that of alkanes and alkenes (not presented here) calculated using the same method (0.42 and 0.43 respectively).

The emergence of several trends in the NO$_3$ dataset precludes accurate estimation of NO$_3$-arene rate coefficients at this time, and it is expected that further experimental data especially from alkényl substituted arenes and $\Sigma x^0$ will provide greater insight into the interaction of nitrate radicals with arenes.

Despite the complex behaviour observed in the NO$_3$ dataset, in general, correlations with radical-arene rate coefficients are excellent. Since the Randić-type index used only encodes the valency of carbon atoms with respect to other carbon atoms, their adjacency and a term for $\Sigma x^0$, other molecular properties are not incorporated in the index, and given the high level of correlation, appear to be of minor importance in comparison.

Despite the simplicity of the method used, this topological approach generates very strong relationships with the overall rate coefficients of arenes. Although an exact physical interpretation of what the Randić index measures remains subject to discussion (Estrada, 2002), the factors governing arene reactivity are simple enough that this index can describe them in adequate detail. The Randić index calculation requires few input parameters, and the result is a description of the connectivity of carbon atoms within the molecule. Other factors that might be expected to be of importance such as the geometry and 3-dimensional spatial configuration of a molecule appear to be minor in comparison and by considering the details that the Randić index omits, some qualitative insight may be gained regarding the relative importance of factors contributing towards arene reactivity.

5 Atmospheric implications

The impact of using rate coefficients estimated through our method versus the recommended values of Calvert et al. (2002) upon the output of an atmospheric model was assessed. The model has been described in detail before (Utembe et al., 2005) and uses the CRI mechanism described in detail by Jenkin et al. (2002). Initial comparisons were made using the full MCM mechanism and the CRI mechanism, where it emerged that the two mechanisms yielded the same results. The reduced mechanism was therefore adopted for the complete set of integrations, since it is more efficient to run. Briefly, the model was integrated for four days using a Gear type solver. Emissions of NO$_x$, CO, SO$_2$, CH$_4$ and some 124 VOCs (detailed in previous papers) were incorporated into the model using the National Atmospheric Emissions Inventory (NAEI) (http://www.naei.co.uk). The box model was assumed to be static in space but was subjected to diurnally varying emissions, deposition (where this had a diurnally varying component), temperature and sunlight. A range of NO$_x$ conditions appropriate to the urban environment were required for this study, the NO$_x$ emissions were scaled by factors of 0.1, 0.5, 1.0, 2.0, and 3.0 and covered NO$_x$ concentrations from around 500 ppb to 20 ppb. Rate coefficients of arenes with respect to OH were obtained from measurement data as recommended by Calvert et al. (2002). The rate of change of each arene is dependent on its initial concentration, the level of emission, and the concentrations of atmospheric oxidants (mainly NO$_3$ and OH). The emissions of NO$_x$, CO, SO$_2$, and VOCs into the model are based on the UK totals for 2001 reported by the National Atmospheric Emissions Inventory. A range of NO$_x$ conditions appropriate to the urban environment were required for
this study, the NO\textsubscript{x} emissions were scaled by factors of 0.1, 0.5, 1.0, 2.0, and 3.0.

The reduced mechanism was restricted to C\textsubscript{\leq 8} arenes, therefore the list of arenes for which rate coefficients were altered are: benzene, toluene, \textit{p}-xylene, \textit{o}-xylene, \textit{m}-xylene, mesitylene, hemimellitene, \textit{\psi}\text{-}cumene, styrene, ethylbenzene, \textit{p}-ethyltoluene, \textit{o}-ethyltoluene, \textit{m}-ethyltoluene, cumene, and \textit{n}\text{-}propylbenzene. Figure 7 shows that over the whole range of NO\textsubscript{x} conditions, our rate coefficients estimates have a minor effect upon the modelled mixing ratios of ozone in comparison to using measurement data. Altering the rate coefficients between those predicted in this work and those recommended by Calvert et al. (2002) have a small impact on the relative rate of ozone production arising from the oxidation of these arenes (around 15–20%) and have a negligible impact on the absolute rate of production (<5%). The reason for the small impact on the absolute rate of production is that for those key arenes that contribute most to ozone formation (either because they are the highest emission or the fastest rate coefficient or both), the agreement between predicted and measured is very good (typically within 20%). Furthermore, only minor effects were observed upon the concentrations of other major species (such as NO\textsubscript{x} and HO\textsubscript{x}), especially under conditions of low NO\textsubscript{x}. Differences in the concentrations of other important species e.g. formaldehyde (HCHO) and peroxyacetyl nitrate (PAN) mixing ratios only began to emerge at NO\textsubscript{x} emission factors greater than 1.0.

Further modelling studies were conducted to compare the estimates using the topological approach with those predicted using the Zetzsch method. As Zetzsch’s method does not include alkenyl substituted arenes, the rate coefficient of styrene was not altered. The results suggest that rate coefficients predicted by the topological approach compare to measured values as well as, if not better than the values predicted by Zetzsch. The differences between the Zetzsch method versus the topological approach become most apparent for [PAN] and [HCHO] under high NO\textsubscript{x} conditions (see Fig. 8).

As a result of these modelling studies, it is clear that rate coefficients perform well in a simplified version of the MCM, and it is suggested that the topological approach is most suitable for inclusion in the full MCM mechanism, where advantage could be made of the range of arene rate coefficients that can be predicted. It is further suggested that there are many species of arene for which kinetics with respect to atmospheric oxidants are not available that are present in vehicle exhaust emissions (e.g. AQIRP, 1995) and which therefore are of potential importance in urban environments, and it
is probable that failure to account for these minor species will result in a greater error associated with model output than the error associated with the use of modelled versus measured rate coefficient data.

6 Conclusions

Strong relationships are observed in correlations between the Randić-type index of this study and log $k_{298}$ for the reaction of arenes with OH, $O(\Delta P)$, H and Cl radicals. Correlation is particularly strong with the OH radical, whose dataset is superior in both number and range of compounds for which measurements are available. In the correlations of the remaining radical species considered (excluding NO$_3$), behaviour is assumed to be analogous with the OH radical, and rate coefficients should therefore be predicted using the same method. Although, for these species, further rate coefficient measurements (preferably possessing as large a range in reactivity as possible) from $\Sigma x_{\beta0}$ are necessary in order to constrain the model parameters of Equation (II) accurately, facilitating accurate rate coefficient prediction for arenes belonging to $\Sigma x_{\beta0}$.

The present correlations with the NO$_3$ radical appear to be complicated by mechanistic dissimilarity between alkyl and alkenyl substituted arenes, resulting in several relationships which are not sufficiently strong to enable accurate prediction of rate coefficients. Further measurements of both alkyl and alkenyl substituted arenes are required before this method can be used to predict NO$_3$ rate coefficients accurately. However, the trends observed in the NO$_3$ data suggest that the Randić-type index used in this study is able to identify mechanistically distinct behaviour.

With the exception of the nitrate reaction, correlations of alkyl-substituted arenes accord with a general trend incorporating all arene species. This suggests that the reactions of alkyl-substituted arenes and alkyl substituted arenes possess very similar mechanisms, contrary to the prevailing opinion in the literature, which suggests that reaction occurs almost exclusively on the alkyl substituent (Calvert et al., 2002), an argument which appears to be founded upon the magnitude of alkyl substituted species relative to their alkyl homologues and the two existing product studies of the OH + styrene reaction (Bignozzi et al., 1983; Tuazon et al., 1993). Here, the failure of NO$_3$ to exhibit the same tendency as the other radicals provides a reference point by which mechanistically distinct behaviour can be recognized in these correlations, the relationships of the other radicals afford no such discrimination.

Modelling studies identified the topological approach as suitable for inclusion in complex atmospheric models such as the MCM, which at present uses measured rate coefficients for model input, and which may benefit from the range of species that can be predicted accurately through this approach.

In conclusion, this study finds the topological approach to be eminently suitable for the estimation of arene-radical gas-phase rate coefficients, because it surpasses the methodology of Zetzsch (1982) and Atkinson (1991) both in terms of its diversity and its accuracy, since this method is able to predict every rate coefficient within a factor of two of the measured value (with the exception of the anomalously slow rate coefficient of $\beta$-dimethylstyrene). Furthermore, despite its simplicity, the topological approach represents an accurate description of arene reactivity, and as such an approximation of the important factors contributing towards this reactivity.

Edited by: A. Hofzumahaus

References


Bignozzi, C. A., Maldotti, A., Chiorboli, C., Bartocci, C., and Carassiti, V.: Kinetics and mechanism of reactions between aro-


tribution of organic compounds during the August 2003 ozone episode in the southern UK, Faraday Discuss., 130, 311–326, 2005.