

Article

Estimation of Rate Constants for Reactions of Organic Compounds under Atmospheric Conditions

William P. L. Carter 

College of Engineering Center for Environmental Research and Technology (CE-CERT), University of California, Riverside, CA 92521, USA; carter@cert.ucr.edu

Abstract: Structure–activity (SAR) methods are presented for estimating rate constants at 298 K and approximate temperature dependences for the reactions of organic compounds with OH, NO₃, and Cl radicals and O₃, and O(³P) in the lower atmosphere. These are needed for detailed mechanisms for the atmospheric reactions of organic compounds. Base rate constants are assigned for the various types of H-abstraction and addition reactions, with correction factors for substituents around the reaction site and in some cases for rings and molecule structure or size. Rate constant estimates are made for hydrocarbons and a wide variety of oxygenates, organic nitrates, amines, and monosubstituted halogen compounds. Rate constants for most hydrocarbons and monofunctional compounds can be estimated to within ±30%, though predictions are not as good for multifunctional compounds, and predictions for ~15% of the rate constants are off by more than a factor of 2. Estimates are more uncertain in the case of NO₃ and O₃ reactions. The results serve to demonstrate the capabilities and limitations of empirical methods for predicting rate constants for the full variety of organic compounds that may be of interest. Areas where future work is needed are discussed.



Citation: Carter, W.P.L. Estimation of Rate Constants for Reactions of Organic Compounds under Atmospheric Conditions. *Atmosphere* **2021**, *12*, 1250. <https://doi.org/10.3390/atmos12101250>

Academic Editors: William R. Stockwell, Emily Saunders and Ajith Kaduwela

Received: 17 August 2021
Accepted: 23 September 2021
Published: 26 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: structure–activity relationships; volatile organic compounds; OH radicals; NO₃ radicals; Cl atoms; O(³P) atoms; atmospheric chemistry

1. Introduction

Many hundreds of types of volatile organic compounds (VOCs) are emitted into the atmosphere, from both anthropogenic and biogenic sources (e.g., [1–3]). Most of these will react in the atmosphere with OH radicals and (if present) Cl atoms, and many also react with other atmospheric oxidants such as ozone (O₃), NO₃ radicals, and O(³P) atoms [4,5]. The atmospheric reaction mechanisms for most of these compounds are complex, larger molecules can react in many ways and form large numbers of oxidized organic products, and most of these will also react in the atmosphere and form additional oxidized organic products [6–8]. Predictions of the lifetimes and fates of these compounds, and of their impacts on formation of secondary pollutants such as ozone, secondary organic aerosols, and toxic organic products, require a knowledge of the rate constants for the reactions of emitted organics or their products with atmospheric oxidants. Because of this, rate constants have been measured for many VOCs and atmospheric oxidants, and a number of compilations of these rate constants exist relevant to atmospheric modeling [5,9–14]. However, the available data do not include rate constants for many of the emitted compounds and for most of the oxidized products that are known or predicted to be formed. In addition, most compounds can react in more than one way, forming different intermediates and ultimately different products, and relatively few have data concerning rate constants or branching ratios for the individual processes. Therefore, it is necessary to estimate rate constants for all the possible reactions of a wide variety of compounds in order to assess atmospheric impacts of VOCs and develop comprehensive mechanisms to represent their reactions in models.

Various structure–reactivity (SAR) methods have been developed to estimate rate constants for the reactions of organics with atmospheric oxidants [15], though all are

limited in scope in terms of the types of compound or oxidant studied, some are not designed to predict rate constants at different sites of the molecule, and none incorporate all the available kinetic data that have been recently compiled [14]. Despite their limitations, various SARs have been used in the development of the current chemically detailed atmospheric reaction mechanisms, including the near-explicit Master Chemical Mechanism (MCM) [7,8,16–18] and the detailed SAPRC mechanisms [19–23], and also in the GECKO-A [6,24] and SAPRC [21,25,26] automated atmospheric chemical mechanism generation systems. Because of the complexity of the atmospheric chemistry of organics, mechanism generation systems provide the most practical way to develop and update detailed mechanisms [27], and serves as the basis for the continuing development of the SAPRC mechanisms [20–23] and for the ongoing update of the MCM and GECKO-A systems for the MAGNIFY project [24,28]. In the case of GECKO-A, the current version incorporates the recently published SARs of Jenkin et al. [29,30] for estimating rate constants for OH reactions, the SARs of Kerdouci et al. [31] for NO₃ reactions, and those of Jenkin et al. [32] for O₃ reactions.

In this paper, we describe the independently-developed SARs for estimating rate constants for initial reactions of organic compounds and their atmospheric oxidation products that are incorporated in the current SAPRC mechanism generation system [22,25,26], discuss their scope, predictive capability, and limitations, and compare them with the SARs recently developed for GECKO-A. The scope includes reactions of many organics with OH and NO₃ radicals, O₃, and Cl and (for a more limited number of compounds) O(³P) atoms. The objective is to utilize existing rate constant data to estimate rate constants for the full variety of organic compounds that might be emitted or formed in the lower atmosphere whose rate constants have not been measured. To be useful for mechanism development, branching ratios for all the non-negligible modes of reaction also need to be estimated. Therefore, the quantities estimated in this work are rate constants for reactions at various positions of the molecule, which are summed up to obtain the total rate constant for the compound that can be compared with experimental data. The approach is entirely empirical, utilizing methods and parameters to predict the available data as closely as possible to make predictions for as wide a variety of compounds as possible. Note that derivation and evaluation of these estimates against theoretically calculated rate constants or using product yield data are beyond the scope of the present study, though this would be useful for future work.

2. Scope and Rate Constant Database

The compounds covered by the estimation methods discussed by this work include the major classes of compounds that are emitted or formed in the atmosphere, and other compounds with similar functional groups. These include alkanes, alkenes, alkynes, alkylbenzenes, naphthalenes, oxygenates with –OH, –O–, and –CO– groups (and combinations thereof), oxidized nitrogen compounds with –NO₂, –ONO₂, and –ONO groups, amines, amides, and monohalogenated halocarbons (see Table 1). Compounds with atoms other than H, C, O, or N and polycyclic aromatics other than naphthalenes are beyond the scope of this work. The methods employed here did not perform satisfactorily for compounds with more than one halogen and for halogenated oxygenates, where long-range interactions, which are not considered here, appear to be important. Therefore, these compounds are not covered in this work. In addition, the limitations of our rate constant database for O(³P) reactions limit the scope of our estimates for these reactions. However, the methods employed should be applicable for most of the types of compounds that are emitted or formed in the atmosphere.

Table 1. List of groups contained in organic reactants that are used as a basis of the structure–reactivity estimates in this work.

Type	Groups				
Alkane	–CH ₃	–CH ₂ –	>CH–	>C<	–CH _x – [a]
Alkene	=CH ₂	=CH–	=C<	=CH _x [a]	=C=
Alkyne	≡CH	≡C–			
Aromatic [b]	–aCH–	–aC–	–pC–		
Oxygenate	–CHO	–C(O)–	–OH	–O–	
Nitrate, nitro, nitrite	–ONO ₂	–NO ₂	–ONO		
Amino	–NH ₂	–NH–	–N<		
Halogen	–F	–Cl	–Br	–I	

[a] This refers to any group where $x \geq 0$. [b] The prefixes indicate groups in one (“a”) or more than one (“p”) aromatic ring.

The estimation methods developed in this work focus on the estimation of rate constants at approximately 298 K, though approximate temperature dependences around this temperature are also estimated. Separate estimates are made for each reaction pathway that may be nonnegligible under atmospheric conditions, but estimates are not made for pathways expected to be too slow to be important at atmospheric temperatures and pressures. Therefore, these estimates are not applicable to combustion or extremely low temperature conditions. Because of limited data, extending the temperature range would require greatly reducing the types of compounds for which estimates could be made, thus reducing the utility of this work for detailed atmospheric chemical mechanism development. In addition, the focus of the estimates is on net rate constants for loss of the reactants at atmospheric pressures and temperatures, which are not always the same as rate constants for elementary reactions if these reactions can be reversed. All estimates are for atmospheric pressure only. However, except for well-studied reactions of smaller molecules where estimates are not really needed, most rate constants estimated in this work are believed to be at or near the high pressure limit.

The derivation of rate constant estimates requires as large and comprehensive a database as possible to derive and evaluate the estimation methods. Recently, McGillen et al. [14] compiled a comprehensive database of measured and evaluated rate constants for gas-phase reactions of a wide variety of organic compounds with OH, Cl, NO₃, and O₃ under atmospheric conditions, and that was utilized in this work. Rate constants for O(³P) reactions were not included in the McGillen et al. [14] database, so those used previously when developing the estimates for the SAPRC-07 mechanism [20,21] were used in this work with a few additions. The numbers of compounds in the database used for estimating the various types of reaction are summarized on Table 2. The database employed is discussed further in Section S1.1 of the Supplementary Information (SI), which also contains a graphic showing the numbers of different types of compounds (Figure S1). Most compounds were used to derive the estimation parameters, but some were used for evaluation only for reasons discussed below.

Table 2. Summary of numbers of rate constants and estimation parameters used in this work.

Reaction With	OH	Cl	NO ₃	O ₃	O(³ P)
Rate constants in database used	853	421	331	266	40
Rate constants used for parameter derivation	507	276	207	172	40
Rate constants used for evaluation only	346	145	124	94	0
Total parameters used for estimates (Table 3)	324	319	312	154	72
Parameters adjusted to fit data	95	72	58	56	16
Fraction of parameters estimated	71%	77%	81%	64%	78%
Fitting rate constants per adjusted parameter	4.3	3.8	3.6	3.1	2.5
Estimates from previous work (Section 4.3)	378		152	185	

Table 3. Summary of parameters used for estimation of rate constants in this work.

Parameter Name	Count [a]	Depends on	Description	Used in Eq.	Tables of Values
H abstraction from group G_1 by OH, Cl, or NO_3					
$k_{\text{H}}^{\text{base}}$	9–17	G_1	Base rate constant for H abstraction from G_1	1	S11
FH^{nbrs}	178–183	G_1 Substs	Correction for a each substituent on G_1	1–3	S12–S14
$\text{FH}^{\text{struct}}$	10–20	Ring	Ring strain correction if G_1 is in a ring	1	S15
Addition of OH, Cl, or NO_3 to $G_1=G_2$, $G_1\equiv G_2$, or $G_1=C=G_2$					
$k_{\text{A}}^{\text{base}}$	31–32	G_1, G_2	Base rate constant for addition to G_1	4, 5	S16
FA_1^{nbr}	36–43	G_1 Substs	Correction for a each substituent on G_1	4, 5	S17
FA_2^{nbr}	36–43	G_2 Substs	Correction for a each substituent on G_2	4, 5	S17
$\text{FA}^{\text{struct}}$	0–1	nC [b]	Correction based on molecule size	4, 6	see text
Addition of O_3 or $\text{O}(^3\text{P})$ to $G_1=G_2$, $G_1\equiv G_2$, $G_1=G_2-G_3=G_4$, or $G_1=G_2-X$ [c]					
$k_{\text{A}}^{\prime \text{base}}$	69–89	Gs, X	Addition to the double-bond system	7	S18
$\text{FA}^{\prime \text{nbr}}$	0–37	Substs [d]	Corrections for substituents around the bond	7	S19
$\text{FA}^{\prime \text{struct}}$	0–29	Ring [d]	Corrections if the Gs are in a ring	7	S19
$\text{FA}^{\prime \prime \text{struct}}$	0–1	Struct [d]	Steric correction	7, 8	see text
Additions of OH, Cl, or NO_3 to aromatic rings at aromatic group G					
$k_{\text{Aro}}^{\text{base}}$	3–6	G	Base rate constant for addition to G	9	S16, S18
FA^{loc}	36–80	Substs on ring	Correction factors for groups on the aromatic ring. Separate factors for <i>ortho</i> , <i>meta</i> , <i>para</i> , and <i>ipso</i>	9	S20
Additions to amino or iodo group G					
$k_{\text{A}}^{\prime \prime \text{base}}$	3–6	G	Base rate constant for addition to G	10	S16

[a] Minimum and maximum number of parameters for the applicable types of reactions. [b] This correction is only used for NO_3 reactions. nC is the number of carbons in the molecule. [c] X is a non-alkyl substituent on a double-bond system for which a separate base rate constant is derived. [d] These corrections used only for O_3 reactions. The steric correction is based on the number of groups in the β position relative to the double-bond system.

3. Methods

The general approach employed is to separately derive estimation methods for the rate constant at temperatures of approximately 298 K using the relatively large database of measured rate constants around this temperature, and to estimate temperature dependences separately based on the more limited database of temperature dependence measurements in the lower atmospheric temperature range. The 298 K rate constant estimates are made using group-additivity or empirically-based methods similar to those employed by Atkinson and co-workers [33,34], and applied more recently for developing SARs proposed for updating MCM [29–34]. These are based on estimating base rate constants for reactions at various positions in the molecule (determined in some cases by substituents present), multiplied by correction factors applied for certain substituents and also, in some cases, by correction factors based on the structure of the molecule. This approach has the advantage that it gives branching ratios for initial reactions at different positions of the molecule, which are needed for mechanism development, as well as the total rate constants. These base rate constants or correction factors are either adjusted to minimize differences between estimated and measured rate constants, or are estimated from other parameters or based on other considerations when sufficient data are not available, or when independent adjustment is not judged to be appropriate.

Note that the approach used in this work considers only the nature of the reaction site and the substituents or groups of substituents adjacent to it. Other than having different base rate constants for additions to 1,2-disubstituted double bonds with *cis* or *trans* configurations, it does not consider stereo-specificity or steric effects where rates of reactions are may be affected by the geometry of the molecule. While this limitation is not

expected to be significant for most compounds covered by the scope of this work, it can cause poor predictions in some cases, as indicated in Section 4.2.

The types of individual reactions for which estimates are made include H-abstraction reactions, additions to non-aromatic double or triple bonds, additions to aromatic rings, additions to the N-atom in amines and amides, and additions to iodide atoms. These are believed to be the main modes of reactions for these compounds at approximately 298 K, though the possibility for other types of reactions cannot be ruled out. In the case of additions to double bonds, the method used depended on the oxidant involved. In the case of additions by OH, NO₃, and Cl, separate base rate constants are assigned for additions to each side of the double or triple bond(s), which has the advantage of requiring the least number of parameters to cover all cases of interest. However, this was found to be unsatisfactory for predicting O₃ rate constants, where the initial reaction is believed to be cycloaddition rather than addition to either side of the bond. Cycloaddition is also expected to be the mechanism for the reactions of O(³P) with double bonds. Therefore, for O₃ and O(³P) predictions, the base rate constants are assigned for each possible conjugated multiple-bond structure. This is more appropriate for cycloaddition reactions, though it requires more parameters. Note that the limited database of O(³P) rate constants can be fit about as well by either method, but this would not be expected if the database were more comprehensive.

For the purpose of these structure–reactivity estimates, molecules are broken down into “groups” that are shown in Table 1. Most of the parameters used in estimates refer to these groups, either singly or in combination, being either base rate constants for reactions at these groups, or corrections to the base rate constants due to the presence of these groups next to the reaction site. Other parameters refer to corrections due to reactions at rings or in some cases corrections for the size or the extent of branching in the molecule. The various types of parameters used are summarized on Table 3 and the total numbers of parameters required for each type of reaction are included in Table 2. A large number of parameters are required because of the wide variety of structures of the molecules and the different ways they can react, as discussed for the various cases in the sections below.

Clearly, it is not possible to derive all these parameters by solely by optimizations (especially since there are more parameters than experimental rate constants in some cases), so many need to be estimated. In most cases, the estimates were based on assuming that sets of related parameters have the same values, but in some cases estimates were made using other methods, as discussed below and in the SI. A balance must be struck between having too few or too many fitting parameters. Using the absolute minimum number of parameters may be more satisfying, but may not give the best predictions that could be supported by the data. Using the maximum number of adjustable parameters that give the absolute best fits to the data might magnify the effects of measurement uncertainties and variabilities that could cause chemically unreasonable parameters and resulting in inappropriate extrapolations. There also may not be unique solutions to the optimization procedures, resulting in different sets of “best fit” parameter values giving significantly different predictions for other compounds. While using statistical or other formal means to derive the optimum number of parameters for this purpose is beyond the scope of this work, an attempt was made to strike this balance in this work based on expert or subjective opinions, as well as results of previous studies.

Although estimates or assumed relationships between parameters significantly reduce the numbers of parameters that have to be adjusted to fit the data, there are still large numbers of parameters to derive for each type of reaction. These were derived in a stepwise manner, with those involved in predicting rate constants using the smallest number of parameters being derived first (e.g., H abstractions from acyclic alkanes), followed by those for predictions involving the least number of additional parameters (e.g., additions to double bonds in alkenes, using the alkane-derived parameters for H abstractions from alkyl substituents, or H abstractions from acyclic compounds or additions to double bonds with single non-alkane substituents), and finally those involving additional considerations

(e.g., cyclic compounds). For this purpose, rate constants in the database are grouped by the types of compounds used in the sequential optimizations. Note that not all available rate constants were used to derive the parameters because some involve reactions of compounds with multiple functional groups, for which the applicability of the estimation methods used are much more uncertain. Therefore, rate constants for such compounds were not used to derive the parameters, but instead were used to test the applicability of these estimates to these types of compounds. These are indicated on Table 2 and elsewhere as “rate constants used for evaluation only”.

Although the preferred approach from a statistical and uncertainty analysis perspective would be to use a randomly-determined subset of the data for parameter derivation and the rest for evaluation and error analysis, this could not be performed in this work because of the large number of parameters relative to the size of our evaluation dataset. Therefore, only multifunctional and other compounds where the estimates are the most uncertain, and a small number of whose rate constants appeared to be inconsistent with related compounds, were excluded from the dataset. These cases are discussed in Section S1.2 of the SI.

Additional information about the optimization procedure is given in Section S1.2 of the SI. The estimation procedures discussed in this work were implemented and tested using the SAPRC mechanism generation system (MechGen) [25,26], with an Excel spreadsheet employed to organize and optimize the parameters and organize and analyze the results. Tables S1–S5 in the SI list the parameters that were optimized, giving the order in which they were derived and the groups of compounds used for each. The specific compounds used in the optimizations are included in the listings of experimental and estimated rate constants in the tables in Section S3 of the SI.

The following sections discuss the specific methods used to estimate the rate constants and temperature dependences for the various types of reactions. Examples of applications of these methods are given in Section S4 of the SI.

3.1. H-Abstraction Reactions

OH, NO₃, and Cl radicals can react with organic compounds by abstracting H atoms, forming the corresponding organic radical and H₂O, HNO₃, or HCl, respectively. H-atom abstraction reactions involving O₃ or O(³P) are assumed not to be important at atmospheric temperatures, so are not estimated in this work. These abstraction reactions can occur at any group that has H atoms, with the rate constant depending on type of group and also its neighboring groups and sometimes on the structure of the molecule around the reaction site. In most cases, the abstraction rate constants are estimated using

$$kH(\text{group, site, nbrs}) = kH^{\text{base}}(\text{group}) \times FH^{\text{struct}}(\text{site}) \times FH^{\text{nbrs}}(\text{nbrs, group}), \quad (1)$$

where $FH^{\text{nbrs}}(\text{nbrs, group})$ is a function of the set of $FH^{\text{nbr}}(\text{nbr, group})$ values as discussed below. Here, “H” is used to indicate that this is an H-abstraction process, and kH is the H-abstraction rate constant at 298 K; “group” is the group in the molecule from which H is abstracted, which in some cases can depend on the neighboring groups; “site” refers to characteristics of the molecule around the group where the reaction occurs, such as the group being in a ring; “nbrs” refers to the neighboring groups or α substituents (each designated “nbr”), which in some cases can depend on their neighbors (β substituents). The adjustable or estimated parameters are: $kH^{\text{base}}(\text{group})$, the rate constant for H abstraction from the group if all the neighbors were methyl groups; $FH^{\text{struct}}(\text{site})$, a correction factor used for reaction at certain sites in the molecule, if applicable; and $FH^{\text{nbrs}}(\text{nbr, group})$ is an overall substituent correction factor derived for the set of neighboring groups, which may also depend on the group where the reaction occurs.

Neighboring group correction factors are derived for the various possible substituent groups, which in some cases can also depend on the groups bonded to it as indicated on Tables S12–S14 in the SI. If there is no more than one substituent group containing atoms other than C or H (non-alkyl substituent), then the overall substituent correction

factor is given by the product of correction factors assigned for the individual substituent groups, i.e.,

$$FH^{n\text{brs}}(\text{nbrs}, \text{group}) = \prod_{\text{nbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}). \quad (2)$$

However, this may not be the best approach if there is more than one non-alkyl substituent, especially if the corrections are large. Although there are insufficient data to derive overall correction factors for all possible combinations of non-alkyl substituents, there are sufficient data to derive correction factors for some combinations. Table S7 and Figure S2 of the SI compares these with various functions of the individual single group correction factors derived for monofunctional compounds. Based on these results, as discussed in Section S1.3, for estimation purposes, we use the averages for the factors for the individual groups when their product is greater than 1, or the product otherwise, i.e.,

$$FH^{n\text{brs}}(\text{nbrs}, \text{group}) = \begin{cases} \prod_{\text{HCnbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}) \times \prod_{\text{nHCnbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}) \\ \text{if } \prod_{\text{nbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}) \leq 1, \text{ or} \\ \prod_{\text{HCnbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}) \times \text{Average}_{\text{nHCnbrs}} [FH^{n\text{br}}(\text{nbr}, \text{group})] \\ \text{if } \prod_{\text{nbrs}} FH^{n\text{br}}(\text{nbr}, \text{group}) > 1 \end{cases} \quad (3)$$

where $\text{HC}^{n\text{brs}}$ refers to the set of hydrocarbon substituents, if any, and $\text{nHC}^{n\text{brs}}$ refers to the set of other substituents. This is used in all cases with more than one α substituent on an alkyl group, even if a best fit pair factor could be derived.

Note that most previously developed SARs for H abstractions assume that $F^{n\text{br}}$ depends only on the substituent and not the group where the abstraction occurs, but in some cases much better fits are obtained if this is not assumed. For example, in general, it is necessary to assume different correction factors for abstractions from $\text{HC}(\text{O})-$ compared to abstractions from alkyl ($-\text{CH}_3$, $-\text{CH}_2-$ or $-\text{CH}<$) groups. However, in most (but not all) cases, the same factors are used for abstractions from the three types of alkyl groups, with the exceptions being cases where the data are sufficient to unambiguously derive separate values depending on the group, and also where the values so derived are significantly different.

In some cases, better predictions are obtained if base rate constants are defined for certain group and neighbor combinations, with the substituent effect incorporated in the base rate constant rather than using a correction factor. This is used in cases where very high values of $F^{n\text{br}}$ are required to fit the data, suggesting that the substituent may significantly affect how the reaction proceeds, and which would could give absurd estimates for some multifunctional compounds. This was found to be the case for $-\text{O}-$ substituents for abstractions from $\text{HC}(\text{O})-$ groups and $-\text{O}-$, $-\text{C}(\text{O})-$, and aromatic substituents on OH groups. In addition, large corrections for α and β $-\text{OH}$ and $-\text{O}-$ neighbors for abstractions from alkyl groups by NO_3 radicals are needed, and better results are obtained if these neighbors are incorporated in the base rate constant. This was not needed for abstractions by OH or Cl, where use of F^{sub} for α $-\text{OH}$ or $-\text{O}-$ substituents gave satisfactory results.

If the group where the abstraction is occurring is on a ring, then a ring strain correction factor is also applied as F^{struct} , which depends on the size of the ring. No other types of structure-dependent corrections were found to significantly improve estimates for abstractions. However, separate sets of ring strain correction factors were found to be necessary if the ring contains a $-\text{CO}-$, $\text{CO}-\text{O}-$, or $-\text{O}-$ groups for OH abstractions. In the case of Cl abstractions, there are sufficient data to derive only one set of ring strain correction factors, and in the case of NO_3 , there are insufficient data to derive any strain factors, and these are estimated based on those derived for OH. There are insufficient data to derive a comprehensive set of ring strain correction factors for multiple ring structures, so if the group where the reaction occurs is in more than one ring, the factor derived for the smallest ring is used.

Although almost 200 parameters were needed to estimate abstraction rate constants (Table 2), only approximately 25% (~10% in the case of NO_3) were treated as independent adjusted parameters that were optimized to fit the data. The various assumptions and estimates that were used are indicated in the tables in the SI listing the base rate constants

(Table S11) and correction factors (Tables S12–S14). Those that yielded the greatest reduction in parameters were as follows: (1) correction factors for reactions at $-\text{CH}_3$, $-\text{CH}_2-$ and $-\text{CH}<$ groups are the same; (2) no correction is used (factor = 1) when the data are insufficient and an estimate cannot be justified (resulting in a highly uncertain estimate); and (3) β substituents have no effect unless the data unambiguously indicate otherwise.

In addition, because of limitations of the available kinetic data relative to H atom abstractions by NO_3 and lower sensitivities of total NO_3 rate constants to abstraction estimates, it was necessary to estimate some for abstractions by NO_3 using correlations with those derived for abstractions by OH or Cl. These estimates are discussed in Section S1.4 of the SI.

3.2. Additions of Radicals to Non-Aromatic Double and Triple Bonds

Atmospheric radicals such as OH, NO_3 and Cl can react with some organic compounds by adding to double or triple bonds, forming an adduct on one side of the bond and a carbon-centered radical group on the other. For additions of these radicals to compounds with separated double or triple bonds, the estimates are made based on assigning group rate constants for each type of group about the bond where the addition occurs (which also depends on the type of group on the other side in most cases), plus corrections for substituents or structural characteristics. If G_1 and G_2 are groups on each side of the bond, the rate constant for addition to G_1 is estimated by

$$kA(G_1, G_2, \text{site}, \text{nbrs}) = kA^{\text{base}}(G_1, G_2) \times FA^{\text{struct}}(\text{site}) \times \prod_{\text{nbrs1}} FA_1^{\text{nbr}}(\text{nbr}) \times \prod_{\text{nbrs2}} FA_2^{\text{nbr}}(\text{nbr}), \quad (4)$$

where “A” indicates that this is an addition process, $kA^{\text{base}}(G_1, G_2)$ is the base rate constant for addition to G_1 when G_2 is on the other side of the bond, nbrs1 and nbrs2 refers to any non-alkyl group bonded to G_1 or G_2 , respectively, $FA^{\text{struct}}(\text{site})$ are structural corrections, where applicable, and $FA_1^{\text{nbr}}(\text{nbr})$ and $FA_2^{\text{nbr}}(\text{nbr})$ are correction factors for non-alkyl substituents on G_1 or G_2 . FA_1^{nbr} and FA_2^{nbr} can in general be different for a given nbr (though they are assumed to be the same if data are inadequate to determine them separately), but are assumed to be independent of G_1 and G_2 . The sum of $kA(G_1, G_2)$ and $kA(G_2, G_1)$ is then used to derive the total rate constant for addition to the bond, with their ratios determining the estimated branching ratios for additions to the different sides.

Equation (4) is also used to estimate rate constants for reactions at conjugated double bonds. The main difference is that groups that are bonded to a second group with a double bond is treated as a different type of group when defining G_1 and G_2 when determining base rate constants. For example, for additions to the structure $G_a=G_b-G_c=G_d$, the base rate constant for addition to G_a is derived using Equation (4) with G_1 based on G_a if it were in an isolated bond, and G_2 based on G_b in the middle of a conjugated bond system, whose base rate constants generally are different than that based on G_b in an isolated bond. Because of lack of data, estimates for conjugated triple bonds are made in the same way as for isolated bonds.

In the case of cumulated double bonds, defined as $G_1=C=G_2$, all the addition is represented as occurring at the central position, because that results in allylic stabilized radicals, while addition to the ends form much less favorable vinylic radicals. The addition rate constant is then estimated using

$$kA(G_1, G_2, \text{nbrs}) = kCA^{\text{base}}(G_1, G_2) \times \prod_{\text{nbrs1}} FA_1^{\text{nbr}}(\text{nbr}) \times \prod_{\text{nbrs2}} FA_2^{\text{nbr}}(\text{nbr}) \quad (5)$$

where “CA” refers to addition to a cumulated bond system, with the base rate constant, kCA^{base} , depending on the groups at ends of cumulated bonds. No structural corrections are used because of insufficient data to derive them (no compounds with cumulated double bonds in a ring are in the database, and they are unlikely to be important in any case). Because of insufficient data, the FA values are the same as derived for isolated double bonds. The only addition is that that F^{nbr} factors for $=\text{CH}$ or $=\text{C}<$ substituents are required

for estimates for compounds with adjacent conjugated and cumulated systems, though no corrections are made ($FA = 1$) for these cases because of insufficient data.

No structural corrections were used (i.e., $FA^{\text{struct}} = 1$) when estimating OH or Cl additions because there were found not to be clearly indicated by the available data. However, Kedouchi et al. [31] found it necessary to include various ring and molecule size corrections in their estimates for NO_3 additions. We found that the available data were insufficient to clearly indicate consistent or reasonable ring corrections, so these were not used in this work. On the other hand, significant improvement to predictions of NO_3 addition rate constants were obtained if the addition rate constant was assumed to increase with the size of the molecule, with

$$FA^{\text{struct}}(\text{nC}) = \{1 - \exp(-\min(\text{nC}, 12) \times FS^{\text{NO}_3}) / \{1 - \exp(-4 \times FS^{\text{NO}_3})\}\} \quad (6)$$

$$FS^{\text{NO}_3} = 0.101,$$

where nC is the number of carbons in the molecule and FS^{NO_3} is a parameter adjusted to fit the data, and yields $FA^{\text{struct}} = 0.79, 1, 1.19, 1.37, 1.57,$ and 1.67 for 3 through 8 carbons, respectively. The maximum number of carbons for this correction was set at 12 ($FA^{\text{struct}} = 2.11$), since there has to be an upper limit and using this for larger carbon numbers did not improve the predictions. Equation (6) was used regardless of the type of double-bond group, since the data were insufficient to clearly derive separate FS^{NO_3} values for different types of groups. Note that structural corrections were also found to be necessary for O_3 additions, as discussed in the next section.

Since the total rate constants for unsubstituted alkenes and alkynes depend only on the total of the two group rate constants but not the ratio, the ratios of group rate constants, or branching ratios for reactions at various types of multiple bonds, cannot be unambiguously derived from a database containing only total rate constants. These branching ratios were estimated as discussed in Section S1.5 of the SI. Briefly, the ratios for OH reactions were derived by assuming the group rate constants depended on the radical formed by the addition rather than the number of hydrogens on the group where the addition occurred, and on the yields of furan products from isoprene and 2,3-butadiene [9]. 100% addition to the most favorable position was assumed for NO_3 additions and equal addition to both positions was assumed for Cl additions, based on considerations of expected selectivity. However, these branching ratio estimates do not affect predictions of total rate constants that are discussed in this work.

The base rate constants and correction factors for these additions, and notes indicating how they were derived, are given in Tables S16 and S17 in the SI.

3.3. Additions of O_3 or $\text{O}(^3\text{P})$ to Non-Aromatic Double and Triple Bonds

Ozone and $\text{O}(^3\text{P})$ also add to double or triple bonds, with the rate-determining step assumed to be the formation of cyclic intermediates that subsequently react. Although Equation (4), above, could be used to estimate these addition reactions as well, the method developed by Jenkin et al. [32] is more appropriate for cycloadditions, and was found to perform significantly better for O_3 . An important difference between the two methods is that base rate constants are assigned separately for each different type of separated and conjugated double-bond system, with the presence of a non-alkyl α substituent defining additional double-bond systems where separate base rate constants are derived.

This method is not used for the radical addition estimates discussed in the previous section because it requires more parameters to make addition rate constant estimates, and has the additional disadvantage that it requires separate assignments of fractions of reactions at different positions in order to estimate mechanisms, which cannot be derived using rate constant data alone. In addition, it does not provide a means to estimate rate constants for more complex double-bond structures for which rate constant data are inadequate (e.g., more than two conjugated bonds). This is why we use Equation (4) for the radical addition reactions discussed above, since it performs almost as well in these cases. Both methods perform equally well in fitting the limited data for the $\text{O}(^3\text{P})$ rate constants,

but the same method is used as used for $O(^3P)$ as used for O_3 because their mechanisms are expected to be more similar to each other than to additions by radicals.

The total rate constant for addition of O_3 and $O(^3P)$ to any position in the bond, $kA'(\text{bond})$, is estimated using

$$kA'(\text{bond, site, nbrs}) = kA'^{\text{base}}(\text{bond}) \times FA'^{\text{struct}}(\text{site}) \times \prod_{\text{nbrs}} FA'^{\text{nbr}}(\text{nbr}), \quad (7)$$

where “bond” refers to each distinct set of groups around the multiple-bond structure, with separate base rate constants (kA'^{base}) for groups with non-alkyl α substituents if needed, and substituent corrections (FA'^{nbr}) being the same regardless of the position of the non-alkyl group about the bond. Note that substituent corrections are not used if the presence of the substituent is taken into account in defining kA'^{base} , so they are only needed if no base rate constant is assigned for this type of α substituent, or for compounds with more than one non-alkyl α substituent around the bond.

Although the only structural correction found to be necessary in estimating radical addition rate constants concerned a size correction for NO_3 additions, Jenkin et al. [32] also used ring corrections and corrections for branched structures, as well as size corrections in some cases for estimations of O_3 rate constants. We found that the available data did not support using as many structural corrections as used by Jenkin et al. [32], but some were still necessary. These included several ring corrections, with different factors derived for separated and conjugated bonds, as discussed in Section S1.6 of the SI, and included in Table S19. Note, however, the available data are not particularly well fit using correction factors that depend only on the ring sizes (see Figure S6), so estimates for rate constants for additions of O_3 to double bonds in cyclic compounds are quite uncertain.

Jenkin et al. [32] also found that use of steric corrections could improve predictions for compounds with branched structures. This is implemented in this work using a correction based on the numbers of β substituents around the double bond, i.e.,

$$FA'^{\text{struct}}(n\beta) = 1 - (F\beta^{O_3} \times \{(\min[4, \max(1, n\beta)] - 1)\}) \\ F\beta^{O_3} = 0.268 \quad (8)$$

where $n\beta$ is the number of β substituents other than H (either alkyl or otherwise), and $F\beta^{O_3}$ is a parameter adjusted to fit the data. This parameter value gave correction factors of 0.732 and 0.464 for two and 3 β substituents, respectively (3 being the maximum $n\beta$ in the current O_3 rate constant database). The available data were insufficient to derive structural corrections for $O(^3P)$ reactions, so no such corrections were used for $O(^3P)$.

The base rate constants for O_3 and $O(^3P)$ additions and notes documenting their derivations are given in Table S18 in the SI, and Table S19 gives the substituent and structure correction factors derived for O_3 additions. There were insufficient data to derive base rate constants for all the conjugated and α -substituted double-bond structures needed for comprehensive estimates, so some had to be estimated based on empirical relationships derived between base rate constants and numbers of substituents around the bonds, as discussed in Section S1.7 of the SI. Because of insufficient data, no correction factors are used for $O(^3P)$ additions, so $O(^3P)$ rate constant estimates for compounds with non-alkyl substituents, or where structural corrections might be appropriate, are more uncertain.

The equations and parameters discussed above provide a means for estimating total rate constants for reactions at a double-bond structure, but, unlike the methods used to estimate radical additions discussed in Section 3.2, do not provide a means of estimating branching ratios for the purposes of mechanism estimation if more than one reaction is possible. The assumptions made in this regard when implementing these estimates in the SAPRC mechanism generation system [25,26] are discussed in Section S1.8 of the SI. However, these considerations do not affect predictions of total rate constants that are discussed in this work.

3.4. Additions to Aromatic Rings

Atmospheric radicals can also add to aromatic rings, with this being a major process in the case of reactions of aromatics with OH radicals under atmospheric conditions, and a non-negligible process for the reactions of Cl with some aromatics. NO₃ radicals can also add to aromatic rings, though these reactions are relatively slow compared to other loss processes, except for phenols. Reasonably good fits to the rate constant data for aromatic hydrocarbons can be obtained if the rate constant for addition to each aromatic group is estimated using

$$k_{\text{Aro}}(\text{aGrp}, \text{aSubs}) = k_{\text{Aro}}^{\text{base}}(\text{aGrp}) \times \text{FA}^{\text{ipso}}(\text{aSub}_1) \times \text{FA}^{\text{ortho}}(\text{aSub}_2) \times \text{FA}^{\text{meta}}(\text{aSub}_3) \times \text{FA}^{\text{para}}(\text{aSub}_4) \times \text{FA}^{\text{meta}}(\text{aSub}_5) \times \text{FA}^{\text{ortho}}(\text{aSub}_6) \quad (9)$$

where $k_{\text{Aro}}^{\text{base}}$ is the base rate constant for addition to an aromatic $-\text{aCH}=\text{}$ or $-\text{sC}=\text{}$ group, aSub_i being the set of substituents on the ring, with aSub_i is the substituent at the i 'th position about the ring relative to where the addition occurs (1 being *ipso*), and FA^{ipso} , etc., being the aromatic ring substituent factor for a substituent at this position, where $\text{FA}^{\text{any}}(\text{H})$ (i.e., no substituent) set at 1.

Aromatic hydrocarbons can contain up to three types of aromatic groups, two occurring in single rings with ($-\text{aC}<$) and without ($-\text{aCH}-$) substituents, and the third being in two rings with polycyclic aromatic hydrocarbon (PAH) structures ($-\text{pC}<$, or PAH groups). Note that "substituent" in this case excludes the neighboring aromatic groups next to where the addition occurs. Since aromatic groups can have more than one *ortho*, *meta*, and (for PAHs) *para* substituents, multiple correction factors can be applied for substituent groups. The total rate constant for the addition to the aromatic ring is then the sum of the group rate constants for addition at each position, each multiplied by correction factors for substituents at various positions on the aromatic rings, with the yield of any particular OH-aromatic adduct being the ratio of the rate constant for OH addition at that position, divided by the total rate constant.

It is important to recognize that the rate constants derived using Equation (9) reflect the net effect of OH addition to form adducts that react further, so should not be used in models that include back decomposition of the adduct back to OH and the aromatic compound. That is because the base rate constants and correction factors were derived based on observed total OH + aromatic rate constants, not rate constants for elementary reversible reactions.

Use of Equation (9) does not give satisfactory results in the case of reactions of NO₃ with phenols, whose rate constants are many orders of magnitude greater than that for reactions of NO₃ with other aromatics. This can be attributed to the possibility that in most cases aromatic-NO₃ adducts rapidly back-decompose before they can react further, yielding a low rate constant for the overall reaction, while phenolic aromatic-NO₃ adducts in either the *ipso* or *ortho* position might undergo a rapid H shift from the -OH to the -ONO₂ group, yielding HNO₃ and a phenoxy radical [35]. Assuming this is the case, we use the approach of assigning separate (much higher) base rate constants for aromatic groups with -OH substituent in the *ipso* or *ortho* position, and set $\text{FA}^{\text{ipso}}(\text{OH}) = 1$. This results in fits to the data for the various phenolic compounds using Equation (9), though it is also necessary to use relatively high values of $\text{FA}^{\text{ortho}}(\text{OH}) = 23$ and $\text{FA}^{\text{para}}(\text{OH}) = 10$ (Table S20) to fit the data for catechols.

The base rate constants for the additions of OH, NO₃, and Cl radicals to groups in aromatic rings and notes documenting their derivations are included in Table S16 in the SI, and the correction factors for substituents on the aromatic ring, and their documentation, are given in Table S20. The available data are insufficient to derive separate aromatic correction factors for each of the four possible positions around the ring for all the possible substituents of interest, so estimates need to be made as indicated in the footnotes to Table S20. For example, if a substituent is found to deactivate addition (have FA values < 1) or have a relatively small effect, it was assumed that $\text{FA}^{\text{ortho}} = \text{FA}^{\text{meta}} = \text{FA}^{\text{para}} = \text{FA}^{\text{ipso}}$, so only one parameter had to be adjusted. If the substituent was found to have a positive

effect, as was the case for alkyl substituents, the factors for the positions found to be most important for alkyl substituents were adjusted, and the others were held at 1. There were sufficient data to adjust parameters for each position for alkyl substitution, but the optimized values were close to 1 except for *ortho* positions for OH and Cl or *ortho* or *para* positions for NO₃, so the other values used in those cases were set at 1 (see Table S20).

3.5. Other Addition Reactions

Certain addition reactions are assumed to consist of addition to an atom in the group, forming an intermediate centered on that group, at a rate constant that is independent of substituents or the structure of the molecule. Their rate constants are estimated using only assigned base rate constants, without corrections.

$$k_{\text{Add}}^{\text{group}}(G_1) = k_{\text{A}}^{\text{base}}(G_1), \quad (10)$$

where G_1 is a group where such an addition can occur. The cases where this is used are discussed below.

3.5.1. Reactions of Radicals with Amines

Although atmospheric radicals such as OH can react with amines with abstraction reactions, the predictions of the methods discussed in Section 3.1 are not consistent with the available data, as discussed in Section S1.9 of the SI. Instead, the available data are much better fit by treating these as addition reactions, which is kinetically equivalent to the initial reaction being the formation of a strong complex, as discussed in the review by Nielsen et al. [36]. This is implemented by assigning base rate constants for additions to $-\text{NH}_2$, $-\text{NH}-$, $-\text{N}<$, $\text{NH}_2(\text{CO})-$, $-\text{NH}(\text{CO})-$, and $-\text{N}(\text{CO})<$ groups. (Separate base rate constants were used for additions to amides because using a substituent correction factor for $-\text{CO}-$ did not give satisfactory estimates). No substituent corrections were used, so the estimated rate constants were equal to the base rate constants. There were sufficient data to estimate rate constants for Cl and OH additions but in the case of NO₃, data are available only for reactions N-alkyl carbamates and N,N-dimethyl amides, i.e., compounds with $-\text{NH}(\text{CO})\text{O}-$ groups, and $-\text{N}(\text{CO})<$ groups. These data, combined with the base rate constant data derived for the OH and Cl additions, were used to estimate base rate constants for NO₃ additions to other amine or amide groups as discussed in Section S1.4 in the SI. The base rate constants derived for the radical + amine addition reactions, and their associated documentation, are included in Table S16.

3.5.2. Reactions of O₃ and O(³P) with Amines

Rate constants have also been measured for reactions of O₃ and O(³P) with amines, and these are also estimated as addition reactions using base rate constants with no substituent corrections. These reactions are discussed further in Section S1.9 of the SI. These base rate constants and documentation notes are included in Table S18 in the SI. Note that there were insufficient data to include estimates for amides.

3.5.3. Reactions of NO₃ with Iodides

Our database includes rate constants for reactions of radicals with methyl and ethyl iodides, and in most cases these can be estimated using the H-abstraction estimation procedures of Section 3.1, with the correction factors indicating effects similar to other halogen substituents. However, their rate constants for reactions with NO₃ are orders of magnitude higher than for similar unsubstituted and are poorly predicted using Equation (1). The data are much more consistent with assuming that the reactions of NO₃ with iodides is initiated by adding or associating with the I group, then reacting further. Theoretical studies of reactions of iodides with NO₃ radicals are discussed by Bai et al. [37]. Therefore, for estimation purposes we treat reactions of NO₃ with iodides as additions, with the base rate constant that fits the data included in Table S16 in the SI. This treatment predicts that reactions of NO₃ with methyl and ethyl iodides would have the same rate constant,

though the rate constant for iodomethane is approximately 1.7-fold faster than iodoethane, perhaps due to steric effects or measurement uncertainties. However, there are insufficient data to refine these estimates. Possible mechanisms for these reactions are discussed by Bai et al. [37].

3.6. Estimating Temperature Dependences

The focus of this work is estimating rate constants at 298 K, though approximate temperature dependence estimates are also made to estimate rate constants for mechanisms applied to lower atmospheric conditions at different temperatures in the approximate range of 270–330 K. Estimates in this work are applicable for atmospheric pressure only, and no attempts are made to estimate rate constants for reactions with complex temperature or pressure dependences, or for large temperature ranges. All temperature dependence estimates in this work are based on the assumption that, at least for temperatures relevant to the lower atmosphere, the temperature dependence can be approximated by the two-parameter Arrhenius expression

$$k(T) = A \times \exp(-E_a/T) \quad (11)$$

where k is the rate constant at temperature T , A is the rate constant at the high temperature limit, and E_a is an overall activation energy in temperature units. Although temperature dependences for many reactions or for large temperature ranges are better described by more complex expressions with more parameters, deriving these additional parameters for estimation purposes is not straightforward and is beyond the scope of this work. Use of the two-parameter expression will limit the range of temperatures for which the estimates are valid in many cases, but this is not the largest source of uncertainty in our temperature-dependence estimates.

The data and methods used to derive the estimated A factors used in this work are discussed in Section S1.10 of the SI. As discussed there, although the rate constant database has 298 K rate constants for a wide variety of compounds, it has usable temperature-dependence information for only a subset of them, and for many types of compounds the A factors do not show consistent patterns usable for estimates. This means that temperature dependences derived for many types are highly uncertain, and should only be applied for temperatures near 298 K. Furthermore, although substituent and structure correction effects are likely to be temperature dependent, temperature dependences were only estimated for base rate constants. Despite the large uncertainties and many approximations made in these temperature dependence estimates, they are considered preferable to ignoring known temperature effects entirely when making rate constant estimates.

4. Results

4.1. Parameters and Rate Constants

The parameters for estimating rate constants and approximate temperature dependences for all types of reactions are given in various tables in Section S2 of the SI, where footnotes summarize how they were derived or estimated. Plots of estimated vs. experimental rate constants and error distribution plots for the estimates for five types of reactions are given in Figure 1, with separate plots for compounds used to derive the parameters and for those used for evaluations only. Distribution plots of factor errors for hydrocarbons, non-hydrocarbons used for parameter derivations, and compounds used for evaluation only are shown in Figure 2. Table 4 lists the compounds with the greatest estimation error, and also gives characteristics of their structures that may affect the reliability of the estimates. Figure 3 and Table 5 show the fractions and numbers of compounds where the estimates differed from the experimental data by various factors. Tables S21–S25 in Section S3 of the SI give the experimental and estimated rate constants and error factors for all compounds in the database, grouped by whether they were used for parameter derivations or not, and also by type of compound. Those tables also give the error factors

and the estimated and experimental A factors for those compounds that were used to derive the A factor estimates.

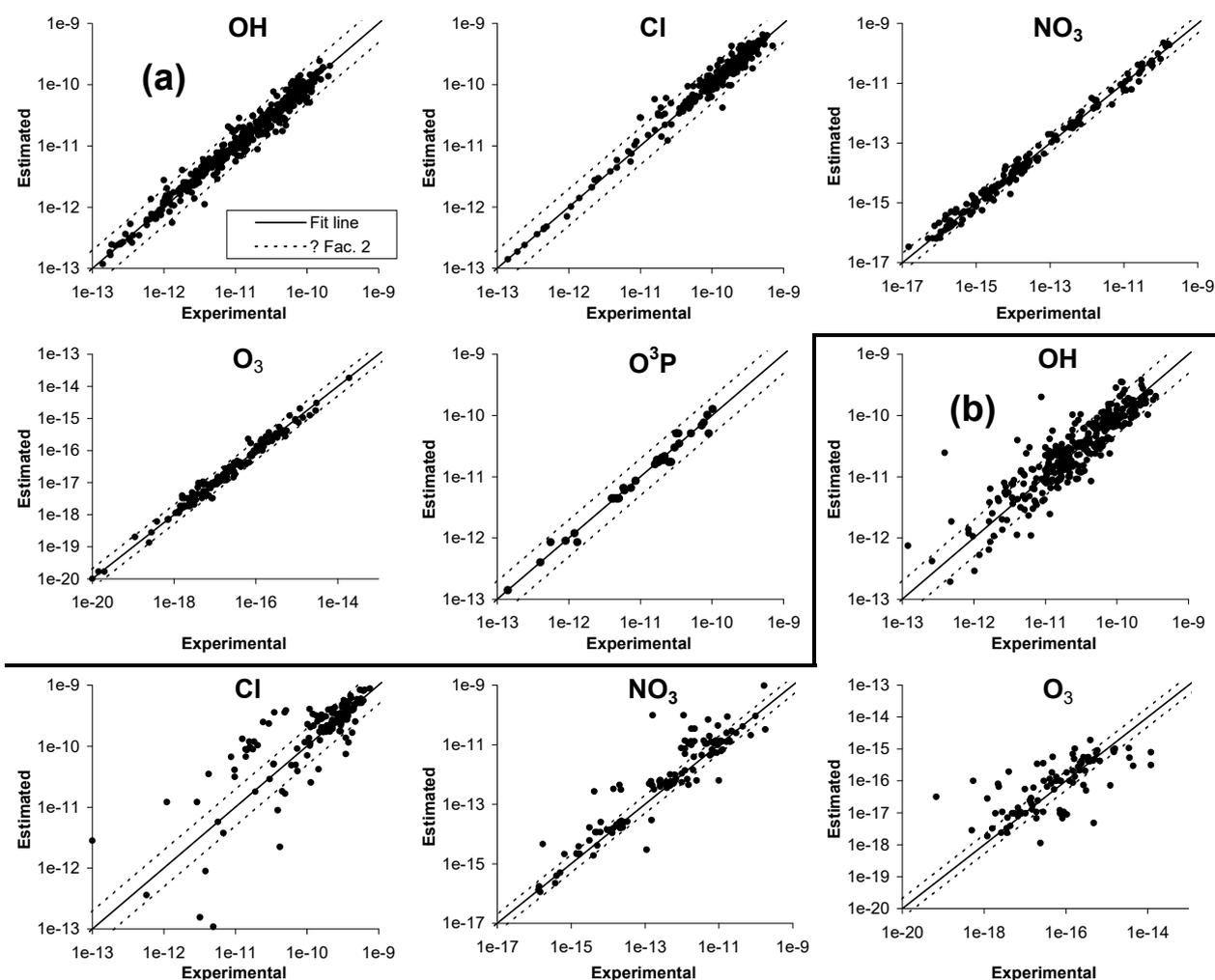


Figure 1. Plots of estimated vs. experimental 298 K rate constants for (a) compounds used for parameter derivations, and (b) compounds used for evaluation only.

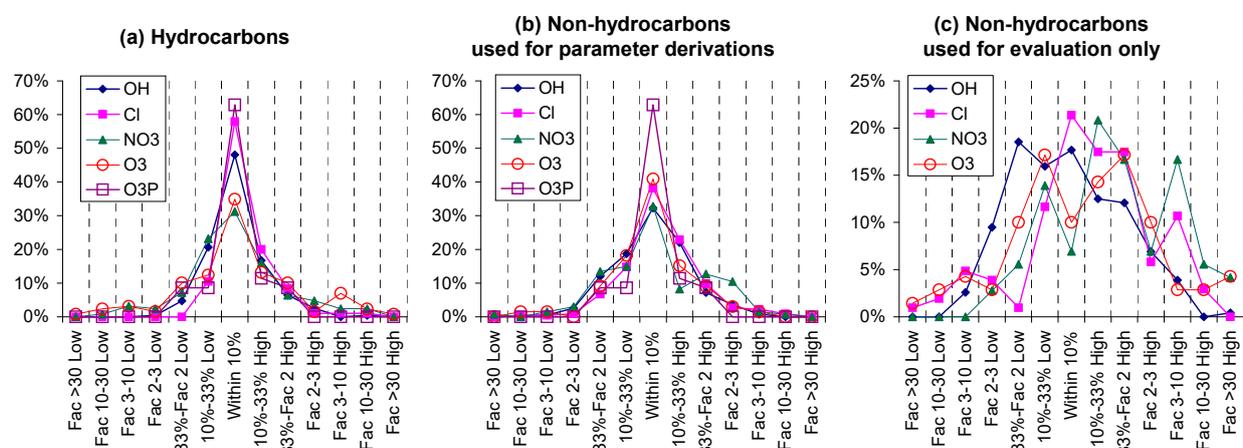


Figure 2. Distribution plots of factor errors in rate constant in estimates at 298 K. Rate constants for (a) hydrocarbons, (b) non-hydrocarbons used to derive parameters, and (c) non-hydrocarbons used for evaluation only.

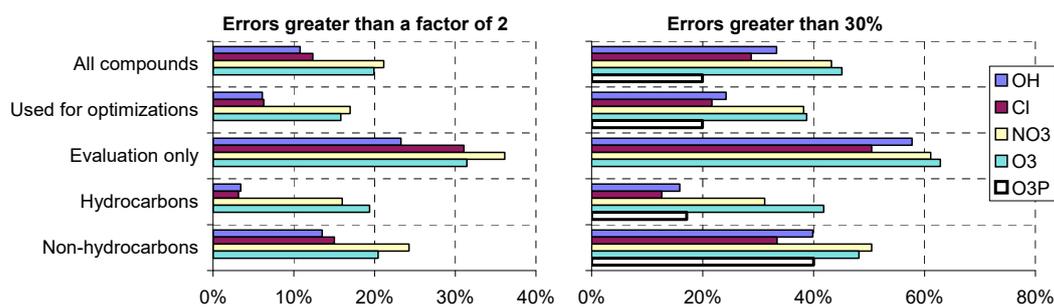


Figure 3. Fractions of compounds with errors greater than 30% or a factor of 2 for estimations of rate constants for reactions with OH, Cl, NO₃, and O₃.

Table 4. Compounds where estimates of the rate constants failed by the largest factors.

Name	k (cm ³ molec ⁻¹ s ⁻¹)			Features [b]				Notes [c]
	Expt	Est'd	Err [a]	Sub	DB	R	AR	
kOH								
maleic anhydride	3.9×10^{-13}	2.5×10^{-11}	63.0	2	1	1		1
fulvenallene	8.8×10^{-12}	2.0×10^{-10}	23.0		4	1		
1-naphthol	5.8×10^{-10}	5.4×10^{-11}	-10.6	1			2	2
gamma-crotonolactone	4.1×10^{-12}	4.0×10^{-11}	9.8	1	1	1		
pyruvic acid	1.2×10^{-13}	7.5×10^{-13}	6.2	2				3
1,4-dinitroxybutane	6.3×10^{-12}	1.1×10^{-12}	-5.7	2				
1-nitrocyclohexene	4.4×10^{-11}	8.5×10^{-12}	-5.2	1	1	1		
N-formyl pyrrolidinone	6.0×10^{-12}	3.0×10^{-11}	5.0	2		1		
3-methyl-2,4-pentanedione	1.2×10^{-11}	2.4×10^{-12}	-4.7	2a				4
4-methyl-2H-furan-5-one	1.3×10^{-11}	5.6×10^{-11}	4.5	1	1	1		
kCl								
oxirane	4.9×10^{-12}	1.1×10^{-13}	-44.9	1		1		3, 5
formyl acetate	1.0×10^{-13}	2.8×10^{-12}	28.9	2				6
methylene glycol diformate	5.7×10^{-13}	2.6×10^{-14}	-22.2	2a				
2-formyloxyethyl formate	3.2×10^{-12}	1.6×10^{-13}	-20.5	2				
2-nitrooxy-1-propanol	4.2×10^{-11}	2.2×10^{-12}	-18.8	2				
peroxypropionyl nitrate	1.1×10^{-12}	1.2×10^{-11}	11.2	1				6
1-nitrooxy-2-pentanol	1.3×10^{-11}	1.3×10^{-10}	10.6	2				
6-nitrooxy-1-hexanol	2.4×10^{-11}	2.5×10^{-10}	10.2	2				
furfural	3.5×10^{-11}	3.6×10^{-10}	10.2	2a	2	1		
naphthalene	4.2×10^{-12}	3.5×10^{-11}	8.3				2	7
kNO ₃								
eugenol	1.6×10^{-13}	9.8×10^{-11}	611	2	1		1	8
4-ethylguaiaicol	1.1×10^{-12}	9.8×10^{-11}	88.9	2			1	8
acetyl cedrene	4.2×10^{-15}	2.7×10^{-13}	64.3	1	1	4		
Coumaran	1.1×10^{-13}	3.0×10^{-15}	-36.8	1		1	1	
cineole	1.7×10^{-16}	4.6×10^{-15}	26.9	1		3		
methylenecyclopropane	1.4×10^{-14}	3.3×10^{-13}	23.6		1	1		9
2-methyl-2-propenyl acetate	2.0×10^{-14}	4.5×10^{-13}	22.2	1	1			
3(e)4-methylhex-3,5-dienal	1.8×10^{-12}	3.5×10^{-11}	19.3	1	2			
3(z)4-methylhex-3,5-dienal	2.2×10^{-12}	3.5×10^{-11}	15.8	1	2			
sabinene	1.0×10^{-11}	6.3×10^{-13}	-15.8		1	2		9
kO ₃								

Table 4. Cont.

Name	k (cm ³ molec ⁻¹ s ⁻¹)			Features [b]				Notes [c]
	Expt	Est'd	Err [a]	Sub	DB	R	AR	
gamma-crotonolactone	6.7 × 10 ⁻²⁰	3.2 × 10 ⁻¹⁷	468	1	1	1		10
4-methyl-2H-furan-5-one	5.4 × 10 ⁻¹⁹	1.0 × 10 ⁻¹⁶	185	1	1	1		10
1,1-dimethoxyethene	4.8 × 10 ⁻¹⁶	4.7 × 10 ⁻¹⁸	-101	2a	1			
3,4-diethyl-2-hexene	4.0 × 10 ⁻¹⁸	1.9 × 10 ⁻¹⁶	48.7		1			11
3,4-diethyl-2-hexene	4.0 × 10 ⁻¹⁸	1.9 × 10 ⁻¹⁶	48.7		1			11
beta-caryophyllene	1.2 × 10 ⁻¹⁴	3.1 × 10 ⁻¹⁶	-38.2		2	2		
acetylcedrene	2.2 × 10 ⁻¹⁸	8.1 × 10 ⁻¹⁷	36.7	1	1	4		
boisvelone	2.3 × 10 ⁻¹⁸	6.6 × 10 ⁻¹⁷	28.5	1	1	2		11
2-cyclohexen-1-one	1.2 × 10 ⁻¹⁸	2.8 × 10 ⁻¹⁷	23.4	1	1	1		10
n-butyl acrylate	2.4 × 10 ⁻¹⁷	1.1 × 10 ⁻¹⁸	-21.3	1	1			
(Z,Z)-1,3-cyclooctadiene	2.0 × 10 ⁻¹⁷	3.4 × 10 ⁻¹⁶	17.1		2	1		
kO(3P)								
1-methyl-1-cyclohexene	9.0 × 10 ⁻¹¹	5.1 × 10 ⁻¹¹	-1.8		1	1		12

[a] Err = Error factor = ratio of (estimated)/(experimental) if an overestimate, or -(experimental)/(estimated) if an underestimate. [b] "Sub" = number of non-alkyl substituent groups. "a" means the substituents are bonded to the same carbon. "DB" = number of double bonds. "R" = number of non-aromatic rings; "AR" = number of aromatic rings. [c] Notes regarding specific cases of poor prediction are as follows: 1. The ring may have aromatic characteristics that are not taken into account that may make additions less favorable. 2. The only other naphthol in the dataset is the 2-isomer, where the kOH prediction is much better. 3. Unique small molecule that may not be typical of larger molecules where estimates are needed. 4. This is expected to be in equilibrium between an unsaturated enol and a saturated keto form, with significantly different predicted rate constants. Rate constant overpredicted by a factor of ~9 for the enol form, so experimental rate constant may reflect a mixture of isomers. 5. Insufficient non-hydrocarbon ring compounds in database to derive ring correction factors. 6. Estimate for reaction at the methyl group alone causes overprediction even if the substituent completely deactivates reaction elsewhere. 7. Naphthalene itself was not used in deriving best fit parameters in order to improve predictions for substituted naphthalenes. 8. The experimental rate constants for eugenol and 4-ethylguaiacol are much lower than those for related compounds in our database, which are reasonably well fit by the estimates from this work, and are considered questionable (see discussion near the end of Section S1.2 of the SI). 9. The estimation methods do not make special corrections for exo double bonds. Note, however, that the estimation errors are in the opposite direction for the two such outlier compounds listed here. 10. Large corrections appear to be necessary for estimating rate constants with oxygenate groups in rings, but the correction factors that fit the limited data are highly inconsistent. Corrections much less than one are needed to fit compounds with -CO-O- in rings, while those with -O- groups in rings suggest factors much greater than one. This could be due to different hybridization and therefore different preferred bond angles for -CO-O- and -O- groups. Because of inadequate data to derive corrections for each type of oxygenated group, no optimizations were performed for compounds with oxygenated groups in rings. 11. The overestimate could be attributed to steric effects that are not taken into account. 12. No ring corrections are made because data for most cycloalkenes in the dataset do not indicate a consistent need for such corrections.

Table 5. Summary of fractions and numbers of compounds with error factors greater than specified amounts.

Error Factor	OH	Cl	NO ₃	O ₃	O(3P)
Fractions of compounds					
±30%	33%	29%	44%	45%	20%
2	11%	13%	21%	20%	None
5	1%	5%	7%	10%	
Numbers of compounds					
5	8	23	23	27	None
10	3	11	12	11	
25	1	2	4	2	

4.2. Performance of Estimates

The performance of our estimates is the best for O(3P) reactions, where all rate constants fit within a factor of 2 and where 80% are predicted within ±30%. This can be attributed to the relatively limited database used for O(3P) reactions and the fact that it does not include bifunctional or other types of compounds where estimates are more

uncertain. The results for the other types of reactions, where the databases are more comprehensive and include larger varieties of compounds, are more comparable and indicative of the overall performance of our estimates. In the case of both OH and Cl, approximately 70% of the rate constants are predicted within $\pm 30\%$ and almost $\sim 90\%$ within a factor of 2. The performance is not quite as good in this respect for NO_3 and O_3 , with approximately 55% being fit to within $\pm 30\%$ and approximately 80% to within a factor of 2. In terms of outliers where estimates are worse than a factor of 5, OH is the best with only 8 of the rate constants being in this group ($\sim 1\%$ of the total), and O_3 is the worst with 27 (almost 10%).

4.3. Comparison with Other Recent SARs

The SARs developed in this work are comparable to those previously developed by Jenkin et al. [29,30,32] for reactions of organics with OH and O_3 , and by Kerdouchi et al. [31] for reactions with NO_3 . These were also developed with the objective of predicting rate constants for a wide variety of compounds for use in developing detailed atmospheric mechanisms, and have been or are being in ongoing updates of the GECKO-A mechanism generation system and the next version of the MCM [24,28]. Their development was mostly independent of the work discussed here, though as discussed above some of the methods and results of the previous work were used to inform the work here on SARs for NO_3 and O_3 reactions.

Figure 4 shows a comparison of the estimated OH, NO_3 , and O_3 rate constants published by Jenkin et al. [29,30], Kerdouchi et al. [31], and Jenkin et al. [32] against those predicted in this work. Note that there are published estimates for only approximately 50% of the rate constants used in this work (see Table 2), since we used a larger database of experimental data. Figure 4 shows that this work gives estimates of very similar rate constants than those previously published, with the vast majority agreeing within a factor of 2, though there are several cases where greater differences are observed. The greatest prediction differences were for 3,4-diethyl-2-hexene, n-butyl acrylate, 3-carene, and ethyl methacrylate for O_3 , allyl acetate for NO_3 , and nitroethene for OH.

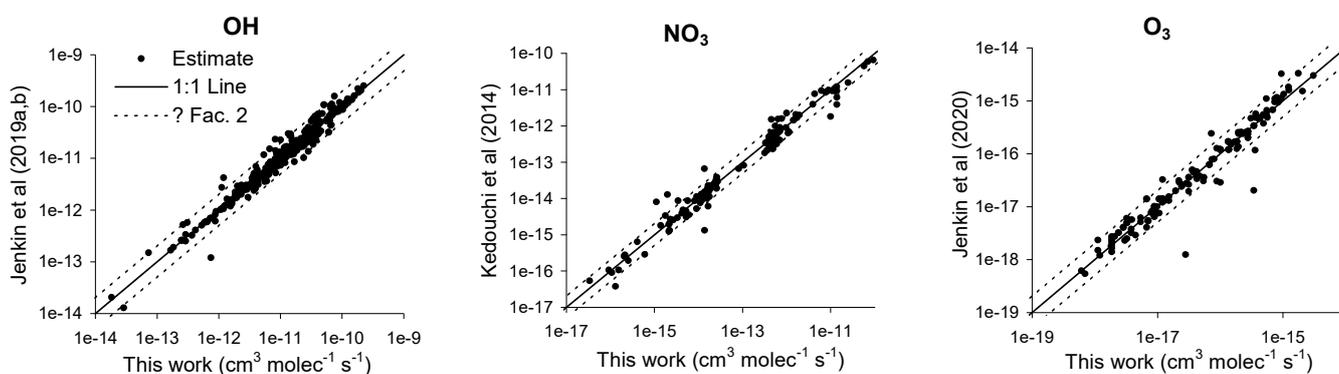


Figure 4. Comparison of estimates from Jenkin et al. [29,30,32] and Kerdouchi et al. [31] with those from this work.

Figure 5 shows distribution plots of factor errors comparing the performance of the previous estimates with this work in predicting the experimental data, and show very similar performance for all three types of reactions. These show distributions for all rate constants that were estimated in both this and the previous works, which are subsets of the rate constants estimated in this work. Additional plots are shown in Figure S11 in the SI, which have separate plots for compounds used for parameter derivation in this work and those that were used for evaluation only. The error distributions are almost identical for the compounds used to derive the parameters, reflecting the fact that estimates for these compounds tend to be somewhat less uncertain and therefore less sensitive to differences in the methods employed. The differences are somewhat greater for the compounds that were not used to derive the parameters in this work, with the published estimates being slightly better, but the differences are still relatively small considering the larger uncertainties.

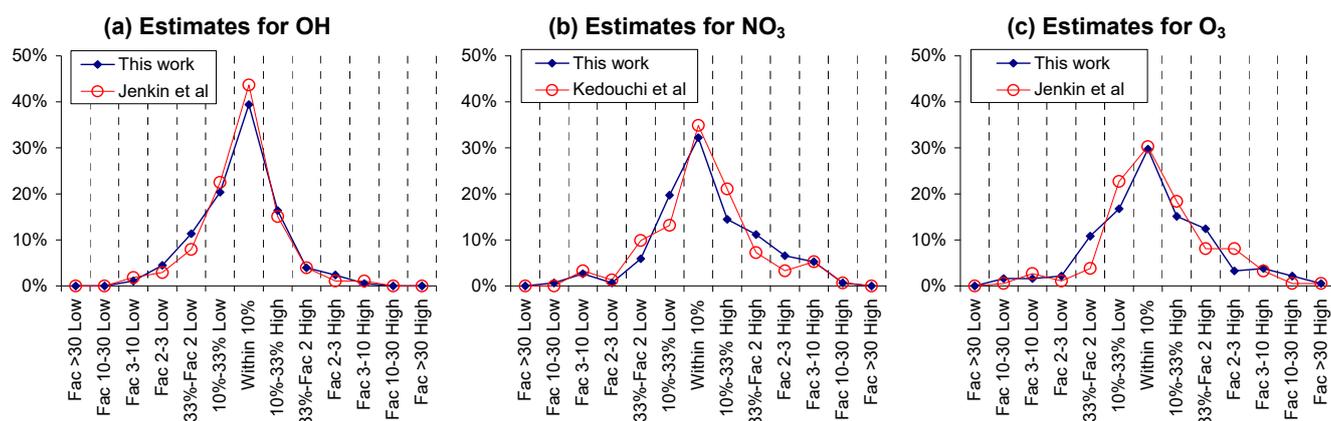


Figure 5. Comparison of distribution plots of factor errors for estimates from Jenkin et al. [29,30,32] and Kedouchi et al. [31] with those from this work.

Since the published estimates are only a subset of the estimates derived in this work, we also compared the distributions of our estimates for compounds also estimated in the previous works with those that are estimated in this work only, which are also shown in Figure S11 in the SI. Although the performance is somewhat better for the compounds where previous estimates are available, reflecting the fact that these estimates may be somewhat less uncertain, the performance for the other compounds is not significantly different. Therefore, it is likely that the previous estimation methods would perform similarly for the additional compounds in our database, at least for compounds that are within the scope of their methods.

5. Discussion

The estimation approach discussed in this work is entirely empirical, with the overall objective being to estimate rate constants and branching ratios for compounds for reactions or organic compounds with major atmospheric oxidants where no data are available. The methods employed are based largely on those SARs originally developed by Atkinson and co-workers and others for predicting rate constants for reactions with OH (e.g., [33,34]), and more recently by Jenkin et al. [29,30], Kerdouci et al. [31], and Jenkin et al. [32] for predicting rate constants for OH, NO₃ and O₃, and which are being used to update the GECKO-A mechanism generation system and the MCM [24]. The main contributions of the work discussed here include the ability to make estimates for more of the types of compounds that may be emitted or formed in the atmosphere, the extension of the methods to reactions of Cl and O(³P), and the use of the comprehensive rate constant database for atmospheric reactions recently compiled by McGillen et al. [14]. This also serves as partial documentation of the estimation methods used in the SAPRC mechanism generation system [25,26] for initial reactions of VOCs. In addition, the results of this work serve to demonstrate the limitations of the types of empirical methods used in these SARs for predicting rate constants for the full variety of organic compounds that may be of interest.

Estimates can be made for most types of C-, H-, O-, or N-containing organic compounds that are emitted or formed in the atmosphere, including almost all types of compounds whose mechanisms can be generated or predicted to be formed by the current near-explicit mechanisms such as the MCM [7,16–18] or by current mechanism generation systems [6,24–26]. This approach works extremely well, making predictions to within $\pm 30\%$ in most cases, for most hydrocarbons and reasonably well for most compounds with no more than one non-alkyl substituent, though there are some exceptions. As expected, this approach does not work as well for compounds with multiple substituents or where multiple corrections are needed, though most are still predicted to within a factor of two or better. This approach is not designed to handle cases where steric considerations are important, since the only structural characteristics taken into account are corrections when

the reaction site is in a ring (when sufficient data are available) or an empirical relationship between addition rate constants and molecule size in the case of NO₃ addition reactions, or a relationship between the number of β substituents and the rate constants in the case of O₃ additions. It is also not designed to handle long-range interactions, which are apparently important for many halogenated compounds.

Overall, approximately 2/3 of the rate constants in our large experimental database are predicted to within ±30% for reactions with OH and Cl, and somewhat over half in the case of reactions with NO₃ and O₃ (see Table 4). Predictions to within ±30% are probably as good as can be reasonably expected given uncertainties in rate constant measurements and the assumptions behind the predictions. Almost 90% of the rate constants in our experimental database are predicted within a factor of 2 for OH and Cl, and approximately 80% in the case of NO₃ and O₃. Except for O(³P), where all rate constants are predicted within a factor of 2 because of limitations in the experimental database, there are a number of outliers where estimates are off by more than a factor of 5, with the most being for O₃ and NO₃ and the least being for OH. These outliers can be attributed to a number of different factors, as discussed in Table 4. The poorer performance in the case of NO₃ and O₃ relative to OH and Cl might be attributed to increased importance of steric factors for the reactions of these larger oxidant species, and also the nature of the O₃ addition reactions, where more parameters are needed for estimates. The performance of our estimates for OH, NO₃, and O₃ is very close to the estimates developed for GECKO-A and MCM [28–32], which is expected since very similar methods were used.

One might question the relevance of estimating O(³P) rate constants for atmospheric mechanisms, since such reactions are generally minor sinks for consumptions of organic compounds in the atmosphere. However, reactions of O(³P) with alkenes can be a non-negligible process under higher NO_x conditions that might occur in some smog chamber experiments used to evaluate mechanisms (e.g., see [21]) or in plumes. Our estimates performed well against our database of O(³P) rate constants, but this is because few compounds in this more limited database required multiple correction factors where estimates are more uncertain. Therefore, estimates of O(³P) rate constants for the full variety of compounds in the atmosphere are either uncertain or beyond the scope of the estimates in this work. However, we would expect steric factors to be much less for O(³P) than is the case for O₃, so better performance for O(³P) than for O₃ might be expected if a more comprehensive database were available for use in its development, as is the case for OH and the other oxidants.

Although we believe that this work, along with recent works of Jenkin et al. [29,30,32] and Kedouchi et al. [31], represents an improvement in our ability to estimate the initial reactions of many types of compounds in the atmosphere, it is clear that more work is needed. Empirical methods such as these require the availability of a sufficiently comprehensive rate constant database to derive the full variety of parameters that are needed. Although we believe that the rate constant database used in this work for the OH, Cl, NO₃ and O₃ reactions is reasonably comprehensive in terms of the available literature and evaluations [14], it is not sufficient to derive all the parameters that are needed without making uncertain assumptions that are difficult to test. Rate constant measurements for targeted compounds could reduce the number of instances where estimated parameters that could not be experimentally verified had to be used, or could test the validity of the many assumptions and estimates had to be employed. (The comments in Tables S11–S20 discussing how certain parameters have to be estimated can indicate the types of compounds where more data are needed.) In addition, quantum calculation methods combined with theoretical kinetic analyses have reached the point where they may provide credible estimates for rate constants for individual reactions where no data are available, but such data were not used in this work. They could be particularly useful for multifunctional compounds that are difficult to obtain or study experimentally. This could allow for the development of better estimates for such compounds, and better methods to estimate base rate constants, substituent factors, or effects of multiple interacting factors, when no

experimental data are available. Use of high-quality theoretical data clearly represents one way this could be extended to improve and evaluate the estimates.

Halogenated compounds are important classes of organic emissions, so methods are needed to estimate their atmospheric rate constants. Unfortunately, although the current methods include hydrocarbons with single halogen substituents, they do not work for compounds with multiple halogens, nor for many halogenated oxygenated compounds. DeMore et al. [38] developed a method that works reasonably well for rate constants of abstractions of OH radicals from halocarbons, based on treating groups with multiple halogen substituents separately, but we found that these methods do not extend well to halogen-substituted ethers and other larger molecules with multiple halogen substituents. This is apparently due to longer range interactions between non-adjacent groups, and use of parameters representing only α or β substitution effects is not sufficient. Methods that take these longer range interactions need to be developed before rate constants for hydrocarbons with multiple halogens or oxygenated halogen compounds substituents can be estimated with sufficient reliability to be useful.

Aromatic oxygenates, nitrates, and nitro compounds are also emitted or formed in the atmosphere, and the methods developed here need to be extended to cover more than one such substituent about the aromatic ring. Although some classes of multisubstituted aromatics can be estimated, most are beyond the scope of the present work either because of limited data or because the current SAPRC mechanism generation system [25,26] cannot process reactions of such compounds. In principle, these methods could be extended if sufficient data were available, though some revisions in the methods employed may be needed.

Although this work includes approximate methods to estimate temperature dependences, the focus is on estimating rate constants near 298 K, which is relevant to atmospheric modeling. The temperature dependences estimated here are not adequate for reliable extrapolation to higher temperatures more relevant to combustion, or to the lower temperatures relevant to the upper atmosphere. More attention is needed to how to estimate temperature effects if a single method is to be reliably applied to such a wide temperature range. In the meantime, it might be more straightforward to separately develop SARs for specific temperature ranges, utilizing measured (or theoretically calculated or extrapolated) rate constants for the specific conditions of interest. This is the approach used in this work, and a similar approach could be used to develop SARs for combustion modeling.

The procedures developed in this work estimate rates of reactions at the different positions of the molecule, but have been developed and tested using only total rate constant measurements. However, data are also available concerning product yields for many of these reactions, which in many cases can be used to derive branching ratios for reactions at different positions. These could provide a useful test of branching ratio estimates and potentially allow some estimates to be refined. This would be especially useful for estimates for relatively minor routes that do not have large effects on the total rate constants, but may affect predictions of formation of toxic or non-volatile products that may be of interest. Initial work on this by Orlando and co-workers, discussed by Aumont et al. [24], compared branching ratios for reactions of ~60 organics with OH derived from product yield data with those predicted by OH SARs being developed for GECKO-A and MCM [24,28–32]. The results showed that the estimated branching ratios were reasonably reliable for alkanes (root mean squares error, RMSE of 2%), but the reliability decreased with increasing numbers of groups (RMSEs of 15% and 19% for monofunctional and bifunctional species, respectively). This appears to be consistent with the performance of the estimates for total rate constants discussed here, though a comprehensive statistical analysis has not been carried out. This work needs to be extended to other compounds and reaction where data may be available, and where possible the results used to improve estimates for minor pathways and a more comprehensive evaluation of the methods.

The empirical methods developed in this work require a relatively large number of parameters to take maximum advantage of the available data to make the estimates as

closely as possible, on the assumption that this will improve accuracy for estimates for other compounds. There are more parameters than can be reliably derived from the data, and use of too many adjustable parameters may result in multiple sets of parameters that fit the data, with each potentially giving significantly different predictions when applied to unstudied reactions, or in overfitting where normal variabilities and uncertainties in the measurement data affect parameter values and predictions in chemically unrealistic ways. Ideally, judgments on which parameters to adjust and which parameters to estimate or lump with other parameters should be aided by systematic statistical or mathematical analyses, with the criteria employed being clearly specified. However, in this work, these judgments were primarily based on expert opinion concerning whether there are sufficient data to unambiguously derive a parameter, and whether the parameters so derived are chemically reasonable. (If the optimization gave parameters judged to be unreasonable, they were estimated or lumped instead.) The development of the SARs for GECKO-A/MCM [24,28–32] employed similar expert judgments, and the fact that similar results are obtained tends to support the utility of this approach. However, use of more systematic procedures to decide what to optimize, lump, or estimate would give more scientifically justifiable results, and could aid uncertainty analysis.

A major omission in this work, which is also the case for the previously developed SARs, is a lack of methods developed to derive uncertainty ranges for the estimates. The wide variability in the quality of the predictions, the uncertainties in many of the estimated parameters, and the variable uncertainties in the measurement data used to derive the parameters indicate that uncertainty estimates are important when interpreting the results. This is an important area where further work is needed, and a logical next step in this effort.

Estimation methods such as those developed in this work should undergo thorough peer review to assure that it incorporates all the appropriate data and theoretical considerations and makes reasonable assumptions before its estimates are used in regulatory modeling and research. Because of this, the author is in a SAR evaluation working group to develop, review, and evaluate structure–reactivity estimation methods. The perspective of this group and the current status and outlook of SAR development are discussed elsewhere [15].

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/atmos12101250/s1>, additional discussion, tables of parameters, tables of rate constants, and calculation examples.

Funding: Partial support for this work was provided by the Coordinating Research Council through Project A-108, and additional support was provided by the California Air Resources Board through contract 11-761 and by the University of California Retirement System.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available in the Supplementary Information are available from the author in Excel spreadsheet format upon request.

Acknowledgments: The author gratefully acknowledges members of the SAR evaluation panel, particularly Max McGillen, John Orlando, Benedicte Picquet-Varrault, Wahid Mellouki, and Tim Wallington for helpful discussions, and particularly Max McGillen for leading the team to develop the rate constant database used in this work. The author also thanks Mike Jenkin for helpful discussions regarding the SARs he developed. The author also thanks one of the reviewers for very helpful comments and thorough editing of this manuscript and the supplement.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Guenther, A.B.; Jiang, X.; Heald, C.L.; Sakulyanontvittaya, T.; Duhl, T.; Emmons, L.K.; Wang, X. The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2.1): An extended and updated framework for modeling biogenic emissions. *Geosci. Model Dev.* **2012**, *5*, 1471–1492. [[CrossRef](#)]
2. Huang, G.; Brook, R.; Crippa, M.; Janssens-Maenhout, G.; Schieberle, C.; Dore, C.; Guizzardi, D.; Muntean, M.; Schaaf, E.; Friedrich, R. Speciation of anthropogenic emissions of non-methane volatile organic compounds: A global gridded data set for 1970–2012. *Atmos. Chem. Phys. Discuss.* **2017**, *17*, 7683–7701. [[CrossRef](#)]
3. Carter, W. Development of a database for chemical mechanism assignments for volatile organic emissions. *J. Air Waste Manag. Assoc.* **2015**, *65*, 1171–1184. [[CrossRef](#)]
4. Finlayson-Pitts, B.J.; Pitts, J.N., Jr. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*; Academic Press: San Diego, CA, USA, 1999.
5. Calvert, J.G.; Orlando, J.J.; Stockwell, W.R.; Wallington, T.J. *The Mechanisms of Reactions Influencing Atmospheric Ozone*; Oxford University Press: New York, NY, USA, 2015.
6. Aumont, B.; Szopa, S.; Madronich, S. Modelling the evolution of organic carbon during its gas-phase tropospheric oxidation: Development of an explicit model based on a self generating approach. *Atmos. Chem. Phys. Discuss.* **2005**, *5*, 2497–2517. [[CrossRef](#)]
7. Jenkin, M.E.; Saunders, S.M.; Wagner, V.; Pilling, M.J. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part B): Tropospheric degradation of aromatic volatile organic compounds. *Atmos. Chem. Phys. Discuss.* **2003**, *3*, 181–193. [[CrossRef](#)]
8. MCM. The Master Chemical Mechanism, Version 3.3.1. 2020. Available online: <http://mcm.york.ac.uk/> (accessed on 23 September 2021).
9. Calvert, J.G.; Atkinson, R.; Kerr, J.A.; Madronich, S.; Moortgat, G.K.; Wallington, T.J.; Yarwood, G. *The Mechanisms of Atmospheric Oxidation of Alkenes*; Oxford University Press: New York, NY, USA, 2000.
10. Calvert, J.G.; Atkinson, R.; Becker, K.H.; Kamens, R.M.; Seinfeld, J.H.; Wallington, T.J.; Yarwood, G. *The Mechanisms of Atmospheric Oxidation of Aromatic Hydrocarbons*; Oxford University Press: New York, NY, USA, 2002.
11. Calvert, J.; Mellouki, A.; Orlando, J.; Pilling, M.; Wallington, T. *The Mechanisms of the Atmospheric Oxidation of the Oxygenates*; Oxford University Press: New York, NY, USA, 2011.
12. NASA. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies*; Evaluation Number 18 JPL Publication 15-10; Jet Propulsion Laboratory: Pasadena, CA, USA, October 2015. Available online: <https://jpldataeval.jpl.nasa.gov/> (accessed on 23 September 2021).
13. IUPAC. Evaluated Kinetic and Photochemical Data; IUPAC Subcommittee on Gas Kinetic Data Evaluation for Atmospheric Chemistry. Web Version. 2019. Available online: <http://iupac.pole-ether.fr> (accessed on 23 December 2019).
14. McGillen, M.R.; Carter, W.P.L.; Mellouki, A.; Orlando, J.J.; Picquet-Varrault, B.; Wallington, T.J. Database for the kinetics of the gas-phase atmospheric reactions of organic compounds. *Earth Syst. Sci. Data* **2020**, *12*, 1203–1216. [[CrossRef](#)]
15. Vereecken, L.; Aumont, B.; Barnes, I.; Bozzelli, J.; Goldman, M.; Green, W.; Madronich, S.; McGillen, M.; Mellouki, A.; Orlando, J.; et al. Perspective on Mechanism Development and Structure-Activity Relationships for Gas-Phase Atmospheric Chemistry. *Int. J. Chem. Kinet.* **2018**, *50*, 435–469. [[CrossRef](#)]
16. Jenkin, M.E.; Saunders, S.M.; Pilling, M.J. The tropospheric degradation of volatile organic compounds: A protocol for mechanism development. *Atmos. Environ.* **1997**, *31*, 81–104. [[CrossRef](#)]
17. Saunders, S.M.; Jenkin, M.E.; Derwent, R.G.; Pilling, M.J. Protocol for the development of the Master Chemical Mechanism, MCM v3 (Part A): Tropospheric degradation of non-aromatic volatile organic compounds. *Atmos. Chem. Phys. Discuss.* **2003**, *3*, 161–180. [[CrossRef](#)]
18. Bloss, C.; Wagner, V.; Jenkin, M.E.; Volkamer, R.; Bloss, W.J.; Lee, J.D.; Heard, D.E.; Wirtz, K.; Martin-Reviejo, M.; Rea, G.; et al. Development of a detailed chemical mechanism (MCMv3.1) for the atmospheric oxidation of aromatic hydrocarbons. *Atmos. Chem. Phys. Discuss.* **2005**, *5*, 641–664. [[CrossRef](#)]
19. Carter, W.P. A detailed mechanism for the gas-phase atmospheric reactions of organic compounds. *Atmos. Environ. Part A Gen. Top.* **1990**, *24*, 481–518. [[CrossRef](#)]
20. Carter, W.P. Development of the SAPRC-07 chemical mechanism. *Atmos. Environ.* **2010**, *44*, 5324–5335. [[CrossRef](#)]
21. Carter, W.P.L. Development of the SAPRC-07 Chemical Mechanism and Updated Ozone Reactivity Scales. Final Report to the California Air Resources Board Contract No. 03-318. Available online: www.cert.ucr.edu/~wpcarter/SAPRC (accessed on 27 January 2010).
22. Carter, W.P.L. Documentation of the SAPRC-18 Mechanism, Draft Report to California Air Resources Board Contract No. 11-761. 29 May 2020. Available online: <http://www.cert.ucr.edu/~wpcarter/SAPRC/18/> (accessed on 23 September 2021).
23. Carter, W.P.; Heo, G. Development of revised SAPRC aromatics mechanisms. *Atmos. Environ.* **2013**, *77*, 404–414. [[CrossRef](#)]

24. Aumont, B.R.; Valorso, M.; Camredon, M.; Jenkin, M.; Rickard, A.; Brauer, B.; Newland, M.; Evans, M.; Mouchel-Vallon, C.; Madronich, S.; et al. Structure-Activity Relationships for the Development of MCM/GECKOA Mechanisms. Presented at the International Conference on Atmospheric Chemical Mechanisms, Davis, CA, USA, 5–7 December 2018; Available online: <https://acm.aqrc.ucdavis.edu/sites/g/files/dgvnsk3471/files/inline-files/Bernard%20Aumont.pdf> (accessed on 23 September 2021).
25. Carter, W.P.L. Documentation of the SAPRC-16 Mechanism Generation System. Draft Interim Report to California Air Resources Board Contract 11-761. 25 May 2019. Available online: <https://intra.engr.ucr.edu/~wpcarter/SAPRC/18/MechGen.pdf> (accessed on 25 September 2021).
26. Carter, W.P.L. Gateway to the SAPRC Mechanism Generation System. 2021. Available online: <http://mechgen.cert.ucr.edu/> (accessed on 12 August 2021).
27. Kaduwela, A.; Luecken, D.; Carter, W.; Derwent, R. New directions: Atmospheric chemical mechanisms for the future. *Atmos. Environ.* **2015**, *122*, 609–610. [[CrossRef](#)]
28. MAGNIFY. Mechanisms for Atmospheric Chemistry, Generation, Interpretation and Fidelity. 2021. Available online: <https://anr.fr/Project-ANR-14-CE01-0010> (accessed on 9 August 2021).
29. Jenkin, M.E.; Valorso, R.; Aumont, B.; Rickard, A.R.; Wallington, T.J. Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aliphatic organic compounds for use in automated mechanism construction. *Atmos. Chem. Phys. Discuss.* **2018**, *18*, 9297–9328. [[CrossRef](#)]
30. Jenkin, M.E.; Valorso, R.; Aumont, B.; Rickard, A.R.; Wallington, T.J. Estimation of rate coefficients and branching ratios for gas-phase reactions of OH with aromatic organic compounds for use in automated mechanism construction. *Atmos. Chem. Phys. Discuss.* **2018**, *18*, 9329–9349. [[CrossRef](#)]
31. Kerdouci, J.; Picquet-Varrault, B.; Doussin, J.-F. Structure–activity relationship for the gas-phase reactions of NO₃ radical with organic compounds: Update and extension to aldehydes. *Atmos. Environ.* **2014**, *84*, 363–372. [[CrossRef](#)]
32. Jenkin, M.E.; Valorso, R.; Aumont, B.; Newland, M.J.; Rickard, A.R. Estimation of rate coefficients for the reactions of O₃ with unsaturated organic compounds for use in automated mechanism construction. *Atmos. Chem. Phys. Discuss.* **2020**, *20*, 12921–12937. [[CrossRef](#)]
33. Atkinson, R. A structure-activity relationship for the estimation of rate constants for the gas-phase reactions of OH radicals with organic compounds. *Int. J. Chem. Kinet.* **1987**, *19*, 799–828. [[CrossRef](#)]
34. Kwok, E.; Atkinson, R. Estimation of hydroxyl radical reaction rate constants for gas-phase organic compounds using a structure-reactivity relationship: An update. *Atmos. Environ.* **1995**, *29*, 1685–1695. [[CrossRef](#)]
35. Bolzacchini, E.; Bruschi, M.; Hjorth, J.; Meinardi, S.; Orlandi, M.; Rindone, B.; Rosenbohm, E. Gas-phase reaction of phenol with NO₃. *Environ. Sci. Technol.* **2001**, *35*, 1791–1797. [[CrossRef](#)] [[PubMed](#)]
36. Nielsen, C.J.; Herrmann, H.; Weller, C. Atmospheric chemistry and environmental impact of the use of amines in carbon capture and storage (CCS). *Chem. Soc. Rev.* **2012**, *41*, 6684–6704. [[CrossRef](#)]
37. Bai, F.-Y.; Sun, Y.-Q.; Wang, X.; Pan, X.-M. Atmospheric chemistry of alkyl iodides: Theoretical studies on the mechanisms and kinetics of CH₃I/C₂H₅I + NO₃ reactions. *RSC Adv.* **2015**, *5*, 88087–88095. [[CrossRef](#)]
38. DeMore, W.B. Experimental and Estimated Rate Constants for the Reactions of Hydroxyl Radicals with Several Halocarbons. *J. Phys. Chem.* **1996**, *100*, 5813–5820. [[CrossRef](#)]