PCCP

ROYAL SOCIETY OF CHEMISTRY

TUTORIAL REVIEW

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2023, **25**, 2717

Received 25th August 2022,
Accepted 14th December 2022

DOI: 10.1039/d2cp03937k

rsc.li/pccp

The computational road to reactivity scales

Maike Vahl and Jonny Proppe *

Reactivity scales are useful research tools for chemists, both experimental and computational. However, to determine the reactivity of a single molecule, multiple measurements need to be carried out, which is a time-consuming and resource-intensive task. In this Tutorial Review, we present alternative approaches for the efficient generation of quantitative structure–reactivity relationships that are based on quantum chemistry, supervised learning, and uncertainty quantification. First published in 2002, we observe a tendency for these relationships to become not only more predictive but also more interpretable over time.

1 Reactivity scales in a nutshell

Reactivity is a kinetic concept and defined as a molecule's propensity to undergo a chemical reaction: ''A species is said to be more reactive or to have a higher reactivity in some given context than some other (reference) species if it has a larger rate constant for a specified elementary reaction.''[1] Reactivity *scales*, on the other hand, interrelate the reactivity of different molecules in a quantitative way.[2] They are useful research tools in several aspects, three of which we would like to highlight:

Institute of Physical and Theoretical Chemistry, Technische Universität Braunschweig, Gaußstraße 17, 38106 Braunschweig, Germany. E-mail: j.proppe@tu-braunschweig.de

Insight. By integrating reactivity scales into the analysis of time-resolved experimental data, relative heights of reaction barriers can be estimated, which facilitates the elucidation of reaction mechanisms.

Efficiency. The search for optimal starting materials is usually a time-consuming and resource-intensive process. By combining reactivity scales with chemical intuition, one can speed up this process through targeted selection.

Reliability. Reactivity scales can also serve as experimental benchmarks when assessing the performance of computational models. Better computational models yield better predictions.

This Tutorial Review does not cover all kinds of reactivity scales. Instead, we focus on two of the most prevalent kinds of

Maike Vahl

Maike Vahl completed her BSc in Chemistry in 2019 and her MSc in Chemistry in 2021 at Georg-August University in Göttingen (Germany). For her master's thesis, she worked on the development of neural network potentials including dispersion interactions for n-alkanes under the supervision of Prof. Jörg Behler in Göttingen. Since 2022, she is a PhD student in the Computational Materials Design group of Prof. Jonny Proppe at TU Braunschweig. Her work focuses on the data-driven design of reactive systems for the energy transition.

Jonny Proppe

Jonny Proppe studied Chemistry at the University of Hamburg and obtained his doctoral degree in 2018 from ETH Zürich (Prof. M. Reiher), for which he was awarded the IBM Research prize in 2020. After postdoctoral stays with Prof. A. Aspuru-Guzik (Harvard University and University of Toronto) and Prof. R. A. Mata (Georg-August University, Göttingen), he was appointed as assistant professor of Computational Materials Design at TU Braunschweig in October 2021. He and his team develop and apply various methods (quantum chemistry, kinetic modelling, machine learning, and uncertainty quantification) to uncover general principles of chemical reactivity.

reactivity complementing each other: *electrophilicity* and *nucleophilicity*.

The genesis of reactivity scales dates back to Lewis's electron theory of valency and the general acid–base theory of Lowry and Brønsted (both mid-1920s) on the basis of which Ingold introduced the concepts of *electrophiles* and *nucleophiles* in the 1930s.[3]

In 1953, Swain and Scott, who studied S_N2 reactions at the time, presented the first systematic approach to constructing a nucleophilicity scale.[4]

$$\log\frac{k}{k_0} = sn \tag{1}$$

Here, the dependent variable k refers to the bimolecular rate constant of the reaction under study at a given temperature. The constant k_0, by contrast, refers to the bimolecular rate constant of a reference reaction at the same temperature. The parameter n is defined as the reactivity (nucleophilicity) of the nucleophile under study and s is an electrophile-specific scaling factor indicating the sensitivity of the rate constant k to variations in the nucleophile. As the rate constant has a natural upper bound (the diffusion limit) and a practical lower bound, the equation by Swain and Scott is limited to a rather narrow range of nucleophilicity values.

In 1972, Ritchie proposed a nucleophilicity scale for reactions with carbocations and diazonium ions.[5]

$$\log\frac{k}{k_0} = N_+ \tag{2}$$

Contrary to the work by Swain and Scott, (i) the right-hand side of the equation is independent of the nature of the electrophile, and (ii) the reference rate constant k_0 is electrophile-specific with water serving as the reference nucleophile. The latter feature allowed Ritchie to widen the range of accessible nucleophilicity values. The relationship shown in eqn (2) is also referred to as "constant-selectivity relationship", because the difference $N_+^{(1)} - N_+^{(2)} = \log(k_1/k_2)$, reflecting the reactions of two nucleophiles with the same electrophile, is independent of the reactivity of that electrophile. (We recommend ref. 3 for a more detailed account of the early history of reactivity scales.)

These pioneering achievements ultimately led to the Mayr–Patz equation (MPE),[3]

$$\log k = s_N(E + N) \tag{3}$$

which constitutes the workhorse of reactivity studies in polar organic chemistry since 1994.[2] The MPE approximates the decadic logarithm of the bimolecular rate constant of an electrophile–nucleophile reaction at 20 °C by three parameters: electrophilicity E, nucleophilicity N, and a nucleophile-specific sensitivity factor s_N. In view of eqn (1) and (2), the term $s_N E$ is equivalent to $\log k_0$. Indeed, eqn (1)–(3) are special cases of the general equation of electrophile–nucleophile combinations,[2]

$$\log k = s_E s_N(E + N) \tag{4}$$

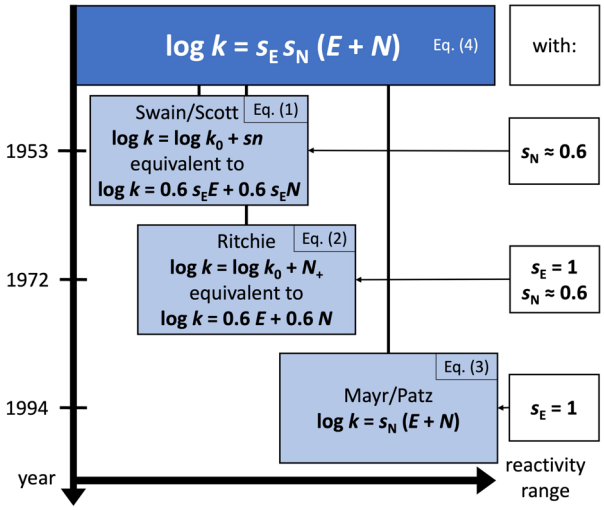


Fig. 1 The general equation of electrophile–nucleophile combinations (top), special cases of which are resolved with respect to time and the range of accessible electrophilicity/nucleophilicity values.

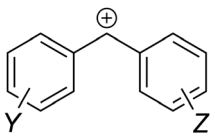


Fig. 2 Lewis structure of a generic benzhydrylium ion. Substituents *Y* and *Z* are to be placed at *meta* and/or *para* positions to avoid mixing of electronic and steric effects.

where s_E, by analogy with s_N, is an electrophile-specific sensitivity factor. An overview classifying the different equations with respect to time and applicability is provided in Fig. 1.†

The MPE [eqn (3)] was developed on the basis of a rich body of data on reactions involving benzhydrylium ions (Fig. 2) and π-nucleophiles (including alkenes, arenes, allylsilanes, *etc.*). Contrary to the work by Ritchie, Mayr and Patz relaxed the assumption of constant selectivity by introducing the sensitivity factor s_N. Therefore, the order of selectivity of two nucleophiles reacting with a weak electrophile may be reversed when reacting with a strong electrophile (see Table 1 and Fig. 3).

Moreover, the scales of electrophilicity and nucleophilicity developed by Mayr and coworkers were no longer bound; by combining strong (weak) electrophiles with weak (strong) nucleophiles, it could be ensured that all reactions are observable ($k > 10^{-5}$ M^{-1} s^{-1}) without reaching the diffusion limit ($k < 10^{8}$ M^{-1} s^{-1}). It is this leap in adaptability that has warranted the persistence of Mayr's reactivity scale approach for the past three decades.

Meanwhile, the applicability of the MPE has also been confirmed for other types of carbocations and for Michael

† The left-hand sides of eqn (3) and (4) refer to the logarithm of a quantity that is not dimensionless ($\log k$). As long as it is ensured that the reactivity parameters strictly correspond to a specific set of units (here, $[k] = M^{-1}$ s^{-1}), this expression is unambiguous.

Table 1 An example of reversed selectivity ordering for species with reactivity parameters documented in previous work.[6] While the weaker electrophile **E15** is more selective toward the stronger but less sensitive nucleophile **N7**, the stronger electrophile **E20** is more selective toward the weaker but more sensitive nucleophile **N40**

	N7 (s_N = 1.00, N = 1.70)	**N40** (s_N = 1.46, N = 0.49)
E15 (E = 0.00)	log k = 1.70	log k = 0.72
E20 (E = 3.59)	log k = 5.29	log k = 5.96

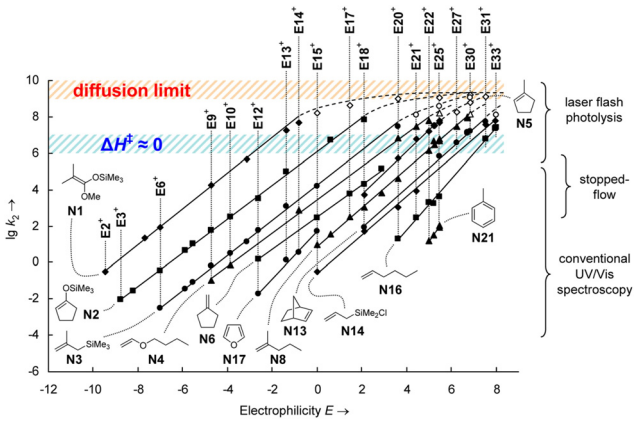


Fig. 3 log k (here, lg k_2) as a function of the electrophilicity parameter E as defined in eqn (3). Each line refers to a series of reactions of a single nucleophile with various electrophiles. The nucleophilicity parameter N can be determined from the intersection of a line with the abscissa (log k = 0 ⇒ E = − N). Nonparallel lines indicate nucleophiles with different values of the sensitivity factor s_N. Reprinted with permission from J. Ammer, C. Nolte and H. Mayr, *J. Am. Chem. Soc.*, 2012, **134**, 13902–13911. Copyright 2012 American Chemical Society.

acceptors, as well as for a diverse range of nucleophiles including amines, enamines, alcohols, enol ethers, ylides, phosphines, hydride donors, and organometallic compounds.[7] Mayr's Database of Reactivity Parameters (https://www.cup.lmu.de/oc/mayr/reaktionsdatenbank2/) currently holds values of reactivity parameters for 352 electrophiles (E) and 1251 nucleophiles (s_N, N).[8,9]

It is this broad application space that eventually caught the attention of computational chemists in 2002.[10,11] Two directions have been followed since then: quantum chemical calculations of log k (Section 2) and regression of reactivity parameters on chemical descriptors (Section 3). The overall workflow embedding both experimental and computational approaches to constructing reactivity scales is shown in Fig. 4. A recent development involves uncertainty quantification of both Mayr–Patz-type reactivity parameters (E, N, s_N) and log k to increase the reliability of the reactivity scale approach (Section 4). In Section 5, we discuss limitations of the computational reactivity scale approach and conclude with a look ahead.

2 Computational prediction of log k

Box 1 summarizes established approaches to determining log k in wet laboratories. This section, on the other hand, discusses approaches to determine log k by means of quantum chemical calculations. The motivation is twofold. Calculations have the potential to (i) accelerate research as they are more resource-efficient and simpler to automate than experiments, and (ii) deepen understanding as they can resolve details that are hidden in experimental data.

The Eyring equation [eqn (5)] is the common link between the molar activation free energy in solution – obtained

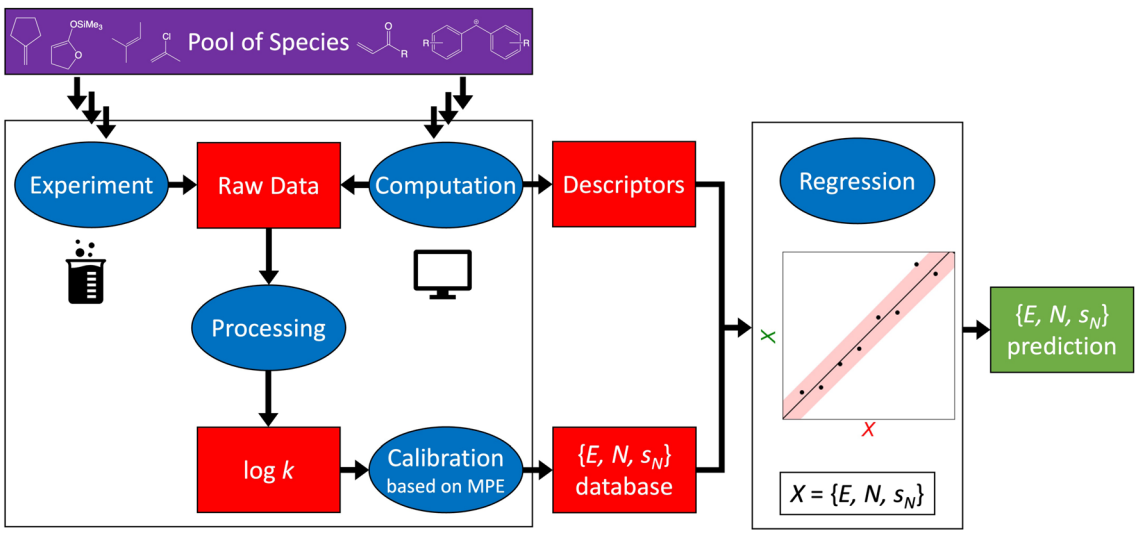


Fig. 4 Schematic workflow illustrating the possible approaches to constructing reactivity scales. A pool of species (electrophiles and nucleophiles) defines the starting point. Raw kinetic data on electrophile–nucleophile reactions can be obtained by means of both experiment (Box 1) and computation (Section 2), which is processed to yield a rate constant – here, in the form of log k – of the reaction under study. From the resulting set of log k values, Mayr–Patz-type reactivity parameters, which are subsequently stored in Mayr's database,[8,9] can be determined through calibration (Box 2). Once a sufficient number of reactivity parameters exists, more efficient access to new reactivity parameters is possible *via* regression of the more expensive database parameters on computationally generated chemical descriptors (Section 3).

Box 1 Experimental determination of log *k*.

Conductometry

Photometry

Conventional UV/vis

Stopped-flow UV/vis

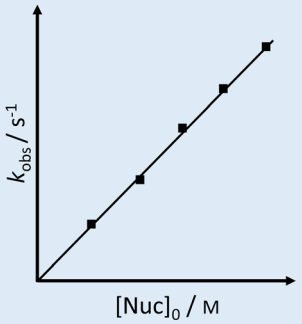
Laser flash photolysis

The method of measurement must be chosen depending on the rate of the reaction in progress. In general, the decrease in concentration of the electrophile after adding a nucleophile is observed by photometry or conductometry.[42]

$A_t = A_0 \exp(-k_{obs}t) + C$ → $k_{obs} = k_{exp}[Nuc]_0$

Pseudo first-order kinetics can be assumed if the initial concentration of the nucleophile is much higher than that of the electrophile, $[Nuc]_0 \gg [Elec]_0$.

A least-squares fit of the exponential function yields k_{obs}, from which the second-order rate constant, k_{exp}, can be determined by linear regression. Each of the two steps requires multiple measurements.[42]



Box 2 Reactivity parameters from log *k*.

reference ↔ reference	non-ref. ↔ reference
reference ↔ non-ref.	non-ref. ↔ non-ref.

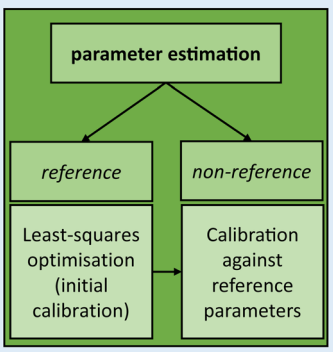
Mayr's Database of Reactivity Parameters[8,9] contains two different types of species, *reference* and *non-reference* species, which span a matrix as symbolically shown.

Parameters (s_N, N, E) for **reference** species are obtained via *least-squares optimisation* on a set of measured log *k* values* (Box 1). The starting points of the reactivity scales must be defined by **two anchor species** for two parameters to allow the previously underdetermined three-parameter equation $\log k = s_N(E + N)$ to be solvable and thus applicable. Since 2012, the fixed values $\boldsymbol{E = 0.00}$ for $(4\text{-MeOC}_6\text{H}_4)_2\text{CH}^+$ and $\boldsymbol{s_N = 1.00}$ for allyltrimethylsilane are employed for this reason.[43]

Parameters for **non-reference** species are obtained by *calibration* of a series of measured log *k* values** against the corresponding reference parameters.

* Combination of reference electrophile with reference nucleophile.

** Combination of reference elec. (nuc.) with non-reference nuc. (elec.)



by quantum chemical calculations –, $\Delta G^{\ddagger}_{sol}$, and the rate constant, k.[12]

$$k = \kappa \frac{k_B T}{h c^{\circ}_{sol}} \exp\left(-\frac{\Delta G^{\ddagger}_{sol}}{RT}\right) \quad (5)$$

Here, k_B, h, R, T, c°_{sol}, and κ denote Boltzmann's constant, Planck's constant, the gas constant, absolute temperature, the standard state in solution (1 M), and a transmission coefficient, respectively. The transmission coefficient κ takes values between zero and one and is sometimes referred to as fudge factor because of its *ad hoc* nature. It is usually set to one, which implies the no-recrossing assumption. That is, once the transition state has been passed along the reaction coordinate, the product inevitably forms without passing through the transition state again.

The molar activation free energy in solution, in turn, is a multilevel composite quantity.

$$\Delta G^{\ddagger}_{sol} = G^{TS}_{sol} - \left(G^{electrophile}_{sol} + G^{nucleophile}_{sol}\right) \quad (6)$$

On the highest level, it is composed of the molar solution free energies of the reactants and the transition state (TS) connecting the isolated reactants with the adduct to be formed. Each of these components can be decomposed into three major terms.

$$G_{sol} = G_{gas} + \Delta G_{solv} + RT \ln\left(\frac{c^{\circ}_{sol}}{c^{\circ}_{gas}}\right) \quad (7)$$

The first term, the molar free energy in the gas phase, G_{gas}, is obtained from a combination of electronic-structure calculations and equilibrium statistical mechanics. The second term, ΔG_{solv}, takes the change of free energy into account when embedding the molecule in gas phase into a solvent environment. It is usually obtained by a parametrised implicit solvation model.[13] The third term, $RT \ln\left(c^{\circ}_{sol}/c^{\circ}_{gas}\right)$, is a crucial constant taking into account the change of the standard state when transitioning from gas phase to solution phase. The standard state of the solution is set to $c^{\circ}_{sol} = 1$ M, whereas an

ideal gas is assumed for the standard state of the gas phase at 293.15 K (20 °C) and 1 atm: $c^{\circ}_{gas} = 4.157 \times 10^{-2}$ M. Neglecting this correction term leads to an offset of almost 2 kcal mol^{-1} in barriers of bimolecular reactions.

The internal energy U and the entropy S are the variables determining the free energy.

$$G_{gas} = U_{gas} + RT - TS_{gas} \quad (8)$$

The dominant contribution to the internal energy is the electronic energy of the ground state, $E^{(0)}_{el}$, which is obtained from electronic-structure calculations.

$$U_{gas} = [E^{(0)}_{el} + E^{(0)}_{vib} + E^{therm}_{vib} + E^{therm}_{rot} + E^{therm}_{trans}]_{gas} \quad (9)$$

Other factors determining the internal energy are the zero-point vibrational energy, $E^{(0)}_{vib}$, obtained from frequency calculations (requiring the Hessian of the electronic energy), and thermal contributions due to vibration, rotation, and translation obtained from equilibrium statistical mechanics.[14] It is assumed here that the first electronic excitation energy is much greater than k_BT, *i.e.*, all excited states are assumed to be inaccessible at any temperature.

The entropy is also composed of electronic, vibrational, rotational, and translational contributions derived from equilibrium statistical mechanics.[14]

$$S_{gas} = [S_{el} + S_{vib} + S_{rot} + S_{trans}]_{gas} \quad (10)$$

Because the harmonic approximation to S_{vib} leads to numerical instabilities when low-frequency vibrations are present, it has become established to employ a quasi-harmonic approximation[15] instead where the low-lying modes are treated as rotational contributions.

The "jigsaw puzzle of quantum chemistry"

It is a challenging task to predict log k from quantum chemical calculations.[16,17] Many of the variables involved in the aforementioned equations cannot, in practical terms, be determined with arbitrary accuracy. Assumptions and approximations are necessary to render their calculation feasible. For instance, electronic and zero-point-vibrational energies require the choice of an electronic-structure method and a basis set; the free energy of solvation, ΔG_{solv}, is usually obtained from parametrised implicit solvation models; and vibrational and rotational motion are most often approximated by harmonic oscillators and rigid rotors, respectively.

On the other hand, the equations relating these variables to each other are approximations, too. For instance, the Eyring equation follows from canonical transition state theory (TST), which is not an exact theory.[18] A deviation of one order of magnitude in the rate constant can be expected from this simplification alone.[19] Furthermore, it is assumed that translational, rotational, vibrational, and electronic motion are separable; basis sets, convergence thresholds, integration grids, truncation and rounding all introduce numerical noise; and the existence of conformational degrees of freedom is fully neglected in these equations.

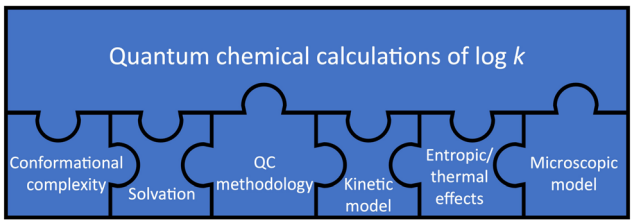


Fig. 5 The "jigsaw puzzle of quantum chemistry". The large puzzle piece representing the general approach to calculating log k by means of quantum chemistry must be connected to each of the six smaller puzzle pieces representing different aspects.

These limitations suggest that quantum chemical calculations of log k require careful consideration of several aspects, which we represent as puzzle pieces in Fig. 5, in the style of work on homogeneous catalysis by Harvey *et al.*[16] These aspects are also highlighted in a recent review by Jorner *et al.*, who discuss a related topic (organic reactivity in the context of machine learning).[20] The jigsaw puzzle suggests that accurate quantum chemical calculations of log k are not possible if one piece of the puzzle is not considered correctly or at all.

In the following, the different puzzle pieces are explained in the context of the computational studies highlighted in this section, an overview of which is provided in Table 2.[21–27] For a more detailed account of general molecular reactions in solution, we refer to the best-practice guide for thermochemistry/kinetics calculations by Harvey and colleagues.[16]

Quantum chemical methodology. The various computational studies are based on Kohn–Sham density functional theory (KS-DFT), better known as simply DFT.[28] It is an approximate theory as the exact exchange–correlation (XC) functional is unknown. KS-DFT is the workhorse of quantum chemistry due to its supposedly favourable trade-off between accuracy and efficiency. Quantum chemists target so-called chemical accuracy (1 kcal mol^{-1}) required for quantitative predictions in thermochemistry and chemical kinetics. To put this number into context: an error of 1 kcal mol^{-1} in $\Delta G^{\ddagger}_{sol}$ leads to a shift of 0.75 units in log k. While no approximate XC functional offers universal chemical accuracy, contemporary approximations are capable to reach or even pass this limit for well-defined classes of chemical systems.[29,30]

Among the selected XC functionals in the reviewed field are hybrid functionals[21,22,24–27] and double-hybrid functionals.[23] Dispersion corrections were considered occasionally.[27] Stationary points of the potential energy surface (reactants, adducts, and transition states) were verified by frequency calculations in all studies. In the majority of cases, transition states were verified to be connected with reactants and adducts through IRC (intrinsic reaction coordinate) following.[22–27]

Despite the known difficulties XC functionals introduce, such as self-interaction and static-correlation errors,[31] KS-DFT remains the method family of choice. While post-Hartree–Fock wave function methods, with CCSD(T) often being referred to as gold standard, offer several advantages over KS-DFT (*e.g.*, systematic improvability, generally higher accuracy), their

Table 2 Overview of the computational studies highlighted in Section 2. For each publication, electrophiles and/or nucleophiles are listed. Additional information is provided in the right-most column

Publication	Data set	Remarks
Wang *et al.*[21]	π-Nucleophiles	The first to have used calculated $\log k$ values to determine *N*. Demonstration of the importance of steric interactions.
Zhuo *et al.*[22]	Benzhydrylium ions, Michael acceptors, diverse nucleophiles	Derivation of MPE *via* FMO theory and quantification of *E* and *N via* LUMO and HOMO energies.
Allgäuer *et al.*[23]	Michael acceptors	Additional experimental prediction of $\log k$.
Jangra *et al.*[24]	Cycloadditions of aryldiazomethanes with Michael acceptors	Additional experimental prediction of $\log k$.
Mayer *et al.*[25]	Reaction of phenolate ions with benzhydrylium ions and quinone methides	Regioselectivity study of phenolate reactions, C *vs.* O attack, including the prediction of kinetic or thermodynamic reaction control. Additional experimental prediction of $\log k$.
Li *et al.*[26]	Heteroallenes	Additional experimental prediction of $\log k$.
Zhang *et al.*[27]	Reactions of iminium ions with diphenyldiazomethane and aryldiazomethanes	Additional experimental prediction of $\log k$.

computational scaling is often much steeper, also because larger basis sets are required to obtain converged properties.[32] To ensure high-quality results from KS-DFT calculations, best-practice guides have been published.[33,34]

Entropic and thermal effects. The common approximations (implemented as default in all established quantum chemistry codes) were employed. No information is provided on the use of hindered-rotor treatments[35] or the free-rotor-based quasi-harmonic approximation for the calculation of the vibrational entropy.[15] In most cases, thermal effects were calculated for a temperature of 298.15 K (25 °C), 5 K above the reference temperature.[22–27] Only Wang *et al.*[21] adhered to the temperature stipulated by Mayr's approach (20 °C). The effect of this deviation on $\log k$ is much weaker than the choice of the XC functional, though.

Solvation. Most of the studies make use of implicit solvation models.[21,23–27] While these models are convenient, they are known to yield more pronounced errors for charged than for neutral species.[36] Due to the extensive presence of cations in Mayr's database, the contribution of this choice to the overall error is to be examined in future work. A microsolvation study that embeds a few explicit solvent molecules in a polarizable continuum (an implicit solvent environment) may serve as an appropriate starting point.[37] It should be noted that the computational cost of explicit and explicit/implicit hybrid models increases not only because of the increased system size but also because of the need for sampling various supramolecular conformations. This renders implicit models the solvation models of choice in most practical cases.

Kinetic model. The majority of studies took the route outlined in the Eyring equation [eqn (5)] to arrive at $\log k$.[21,23–27] Since rate constants are exponential functions of reaction barriers, even small errors introduced by quantum chemical calculations can lead to rate constants that differ by orders of magnitude. As long as thermal relaxation is fast with respect to reaction time scales, and if quantum effects such as proton tunneling can be neglected, canonical TST and, hence, the Eyring equation are reasonable approximations. More sophisticated approaches based on microcanonical TST,[18,38] the master equation,[38] or quantum rate theories[18] should be considered otherwise. The application of these approaches is often limited to the description of reactions of small molecules in the gas phase because of their extreme computational cost. There exist corrections to TST taking account of quantum effects such as tunneling, which in many, but not all, cases lead to results that agree reasonably well with quantum mechanical rate theories.[16] These approaches, however, are not practical for the study of reactivity scales that are based on reactions of bulky electrophiles and nucleophiles in solution.

Zhuo *et al.* took an alternative route to calculating $\log k$.[22] They computed Mayr–Patz-type reactivity parameters from frontier orbital energies to arrive at $\log k$. For this approach to work, the parameters need to be calculated in a recursive fashion. That is, each parameter is a function of all previously determined parameters. As a consequence, the error introduced to the first parameter by the quantum chemical treatment propagates to all of the following parameters, rendering the method increasingly inaccurate. The authors claim that this method is equivalent to the approach followed by Mayr and co-workers, which is not true (*cf.* Box 2). In a recent study, frontier molecular orbitals have been revisited for the calculation of the nucleophilicity parameter of enamines.[39]

Microscopic model. A microscopic model includes information about the positions and identities of the atomic nuclei of the system under study. It can comprise the species of interest in isolation, but also counter ions as well as solvent molecules (which may be considered implicitly as mentioned above). For a careful reflection on the microscopic structure model, all relevant species need to be considered under the target experimental conditions.[16] In the studies discussed here, the microscopic structure model was not studied in detail. Different solvents were tested by Wang *et al.*[21] They found the C–C intermolecular distance in the transition state of benzhydrylium–π-nucleophile reactions to correlate with the reactivity parameters of the MPE [eqn (3)]. The combined experimental–computational studies usually considered the same solvent in which they performed the experiment in their calculations.[23–27] Counter ions were taken into account by Mayer *et al.* in a regioselectivity study of phenolate reactions in aprotic solvents.[25] Their computational results suggest that the counter ion has a negligible effect on the regioselectivity in that specific case.

Conformational complexity. Both reactants and transition states may exhibit conformational degrees of freedom. Each

conformation is associated with an individual free energy, thereby affecting the reaction barrier and, hence, the rate constant of the reaction. Zhuo *et al.*[22] performed a conformational analysis using the MMFF94x force field[40] for three nucleophiles. The ten lowest conformations were further investigated and the energies and structures of the lowest are reported. Jangra *et al.*[24] screened the conformational space for each point along the energy profile of different reactions. Mayer *et al.*,[25] Li *et al.*,[26] and Zhang *et al.*[27] performed conformational analyses with the OPLS3 force field.[41] In each of the studies mentioned in this paragraph, the lowest-energy species was selected for further calculations, which is a popular choice as it corresponds to the highest populated state at equilibrium. A more refined approach including averaging over several conformations should be considered if these are energetically close to the minimum.

3 Computational prediction of reactivity scales

Box 2 describes how the reactivity parameters E, N, and s_N are determined on the basis of experimental (or, alternatively but yet hypothetically, computational) $\log k$ data.[6,42,43] This section, on the other hand, discusses approaches to determine these parameters by means of regression analysis. The motivation to do so stems from the fact that determining reactivity parameters from $\log k$ is expensive for two reasons. First, accurate measurements – both physical (Box 1) and virtual (Section 2) – are expensive *per se*, and second, multiple measurements are required to determine a single parameter. Using relatively inexpensive, computable descriptors that correlate well with these parameters, the aforementioned drawbacks can be mitigated. We define descriptors as representations (descriptions) of molecules, electrophiles or nucleophiles in the present work, the specific forms of which will be discussed later in this section.

The regression analyses reviewed in this section (see also Fig. 6) have in common that the descriptors function as independent variables (x), whereas the reactivity parameters of Mayr's database function as dependent variables (y). Regression of y on x yields a model that can be used to make efficient predictions ($f(x) = \hat{y}$) of otherwise expensive-to-obtain reactivity parameters for new molecules based on descriptors representing them.

The regression model can comprise a single variable (x) or multiple variables ($\mathbf{x} = (x_1, x_2, \ldots, x_M)$), depending on the

Authors (Year) of Publication	Data Set	Descriptor Class	Computational Protocol		Regression Type
			Method	Environment (code)	
Schindele et al. (2002) [11]	Elec: **12** (Benzhydrylium ions)	MAA, hydroxide/hydride anion affinities	B3LYP/6-311++G(3df,2pd)//B3LYP/6-31G(d,p)	Gas phase (Gaussian)	Linear
Pérez et al. (2002) [10]	Elec: **20** (Benzhydrylium ions)	Electronic	B3LYP/6-31G(d)//B3LYP/6-31G(d)	Gas phase (Gaussian)	Linear
Haas et al. (2022) [50]	Elec: **16** (Carboxylic acids), Nuc: **22** (Primary alkyl amines)	Electronic, topologic, measurement shifts	M062X/6-311++G(d,p)//B3LYP/6-311++G(d,p)	SMD(THF) (Gaussian)	Multivariate linear
Allgäuer et al. (2017) [23]	Elec: **44** (Michael acceptors), Nuc: **8** (Reference)	Electronic, MAA, MAA*	B3LYP/6-31G(d,p)//B3LYP/6-31G(d,p) and B3LYP/6-311++G(3df,2pd)//B3LYP/6-31G(d,p)	Gas phase/SMD(DMSO) (Gaussian)	Linear
Pereira et al. (2011) [46]	Elec: **64** (Uncharged)	Electronic, empiric	PM6// B3LYP/6-31G*	Gas phase (MOPAC209, GAMESS; ADRIANA.CODE 2.2.2, CDK Descriptor Calculator 1.1.1, Dragon Professional 5.5)	Multivariate linear, random forest, SVM, NN
Kiyooka et al. (2013) [47]	Elec+Nuc: **65**	Electronic	MP2/6-31G(d,p)//MP2/6-31+G(d,p) and MP2/6-31+G(d,p)//MP2/6-31+G(d,p)	Gas phase (GAMESS, Gaussian)	Linear, multiple linear, quadratic
Wang et al. (2010) [21]	Elec: **6** (Benzhydrylium ions); Nuc: **60** (π-nucleophiles)	Electronic, topologic	BH&HLYP/6-311++G(3df,2p)//B3LYP/6-311+G(d,p)//B3LYP/6-31G(d) (best)	PCM(DCM)/UAHF (Gaussian)	Linear
Mood et al. (2020) [57]	Elec: **75** (28 of 32 functional groups)	MAA*	PBE0-D3BJ/def2-TZVP//PBE0-D3BJ/def2-TZVP	COSMO(∞), SMD(DMSO) (Turbomole)	Linear
Zhuo et al. (2012) [22]	Elec: **33** (Benzhydrylium ions, Michael acceptors), Nuc: **85**	Electronic	B3LYP/6-31G(d)//B3LYP/6-31G(d)	Gas phase (Gaussian)	Linear
Kadish et al. (2021) [58]	Nuc: **98**	MCA*	PBE0-D3BJ/def2-TZVP//PBE0-D3BJ/def2-TZVP	COSMO(∞) (Turbomole)	Linear
Hoffmann et al. (2020) [48]	Elec: **111** (Carbocations, Michael acceptors, arenes)	Electronic, topologic, QTAIM	M06-2X/6–311++G(2d,2p)//M06-2X/6–311++G(2d,2p)	IEFPCM (DCM) (Gaussian), (QTAIM: AIMAll software)	Linear, LaR, PCA, GPR, RBF, GBDT
Orlandi et al. (2021) [49]	Data points: **341** (Elec+Nuc combinations)	Electronic, topologic, proton affinity	M06-2X/def2-TZVP//M06-2X/def2-TZVP	SMD(diff. solvents) (Gaussian)	Multivariate linear
Lee et al. (2020) [54]	Elec: **216,** Nuc: **726**	Electronic, topologic	B3LYP/def2-TZVP//B3LYP/def2-TZVP	Gas phase (Gaussian)	Linear, NN
Boobier et al. (2021) [55]	Nuc: **904**	Electronic, topologic, solvent PCA descriptors	M06-2X-D3/def2-SVP//M06-2X-D3/def2-SVP; PM6	Gas phase/PCM (DCM, MeCN, DMSO, H_2O) (Gaussian, CIRpy, Multiwfn, ACS GCIPR Solvent Tool)	Multivariate linear, PLS, NN, SVM, RF, ET, Bag, GPR

Fig. 6 Overview of studies examining correlations between chemical descriptors and reactivity parameters. Selected aspects are: the data set used for training and testing, the selected descriptors, the computational protocol applied for obtaining those descriptors, and the examined types of regression. Colour gradients indicate the complexity of the described aspect, with darker colours referring to higher complexity. List of abbreviations: MAA*, methyl anion affinity in solution; MCA*, methyl cation affinity in solution; LaR, lasso regression; PCA, principal component analysis; GPR, Gaussian process regression; RBF, radial basis function; GBDT, gradient boosting decision tree; PLS, partial least squares; SVM, support vector machine; RF, random forest; ET, extra trees; Bag, bagging.

dimensionality of the applied descriptor, and its flexibility can range from low (linear) to high (arbitrarily nonlinear).[44] While nonlinear and multivariate models require more training data (pairs of x and y), they have the potential to yield more accurate predictions. The downside of nonlinear models is that they are more difficult to interpret than linear models. In addition, multivariate models are more prone to overfitting training data compared to univariate models, which diminishes both predictive power and understanding.

A general/multivariate linear model is *linear* with respect to its *coefficients* $c_0,c_1,c_2,\ldots,c_M$.

$$y \approx f(\mathbf{x}) = c_0 + \sum_{m=1}^{M} c_m x_m \tag{11}$$

The variables can be arbitrarily nonlinear without affecting the linearity of the model. In a polynomial model, for instance, where $x_m = x^m$, all variables are constructed from a single variable x. The form of eqn (11) still qualifies the model as being linear. The coefficient c_0 is known as *bias*. In the case of $M = 1$, a univariate linear model is obtained, where c_0 and c_1 are referred to as *intercept* and *slope*, respectively.

Nonlinear models of contemporary prominence are neural networks (NNs),[44] support vector machines, (SVMs)[44] and Gaussian processes (GPs).[45] These model types are associated with the term *machine learning*.[44] They are trained – like any regression model – in a *supervised* fashion. That is, in addition to the input values (x), the target values (y) are presented to the model. Regression is a subdomain of supervised learning. Among other nonlinear models, NNs, SVMs, and GPs have been examined in the descriptor-based regression analysis of reactivity parameters.[46–50] A particular model class we would like to highlight is the class of *kernel* models to which GPs belong. The kernel model shown in eqn (12) resembles the general linear model shown in eqn (11).

$$y \approx f(\mathbf{x}_*) = \sum_{n=1}^{N} \alpha_n k(\mathbf{x}_*, \mathbf{x}_n) \tag{12}$$

Although the kernel model appears to be linear, we note important differences. First, the sum runs over the index n, with N representing the number of training points, instead of the index m, with M representing the number of coefficients. Indeed, a data point of interest ($\mathbf{x}_*$) is compared to all training points ($\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N$) *via* the kernel function $k(\mathbf{x}_*,\mathbf{x}_n)$, which, for this very reason, is also referred to as similarity measure. The kernel itself carries its own coefficients, termed *hyperparameters*, which require optimisation on their own. Therefore, kernel models are effectively nonlinear. If the kernel is exponential (*e.g.* Gaussian), the regression function $f(\cdot)$ can assume any shape,[45] which renders this type of model highly flexible and the bias term ("α_0") superfluous. GPs, in addition, provide the standard deviation of the predictions they make. This feature not only is interesting for assessing the reliability of particular predictions, it can also be exploited to systematise the selection of new training points for the sake of model improvement in a data-economic fashion (*active learning*[51]), as we did in previous work on dispersion corrections in collaboration with the Reiher group at ETH Zürich.[52]

Over the last 20 years, the correlation of many different descriptors to one of the three reactivity parameters has been studied extensively. The descriptors discussed here can be assigned to different classes. The first and most prominent class contains **electronic** descriptors mostly based on conceptual density functional theory (CDFT),[53] which include (i) global molecular descriptors and (ii) atomic descriptors.[10,21–23,46–50,54,55] They are defined by mathematical combinations of LUMO and HOMO energies within in the frontier molecular orbital (FMO) approximation. The atomic descriptors additionally account for local effects by using local functions such as the Fukui function and dual descriptor. Examples for (i) are the electrophilicity index ω and the electron affinity μ^{+}_{FMO}, as well as the local electrophilicity $\Delta\rho^{+}_{\mathrm{FMO_A}}$ and the minimum composite Fukui function μ for (ii). Atomic descriptors based on the quantum theory of atoms in molecules (QTAIM)[56] are also presented.[48] The quantum chemical calculation of electronic descriptors is relatively expensive. However, these descriptors provide fundamental insight that can be applied to many other problems and thus offer great benefits.

The second class of descriptors includes different **affinities**. Most prominent examples are methyl anion affinities (MAA) as well as methyl cation affinities (MCA) and proton, hydride and hydroxide affinities.[11,23,57,58] These descriptors are obtained from quantum chemical calculations of reaction energies. They are more expensive than the electronic descriptors of the first class as reaction energies require the characterisation of both a reactant state (isolated species) and a product state (adduct). While affinities do not provide conceptual insight, they involve complete information of the electron density, explaining their strong correlation with reactivity parameters.

Topological descriptors define the third class and include, amongst others, spatial parameters such as buried volume, bond lengths, and more.[48–50,54,55] The fourth class of descriptors contains **empirical** ones from cheminformatics such as the fractional negative surface area (FNSA) descriptor, as well as measurement shifts from NMR or IR spectroscopy and specific solvent descriptors.[46,50,55]

In addition to the choice of the descriptor class and the descriptor itself, the selection of training data is crucial for the creation of a representative model. The number of electrophiles and nucleophiles with known reactivity parameters, as well as their diversity, has a major impact on the generalisation and application capabilities of the model created. The more diverse the data set for training the model, the greater the applicability after successful training. In recent years, the data set size has been expanded to a number of reference structures greater than 750 to learn correlations between different descriptors and reactivity parameters, which enable highly accurate predictions of reactivity. Saini *et al.*[59] report their best result using a NN for 752 structures, while Tavakoli *et al.*[60] use methyl anion/cation affinities in solution (MAA* and MCA*, see below) as inputs for over 2421 structures to train a graph attention network. A graph

neural network was trained by Nie *et al.*[61] on nearly 900 nucleophiles from Mayr's database, in which electronic and, additionally, solvent descriptors were employed.

To provide an overview of the regression studies reviewed in this section, the following aspects are listed in the table of Fig. 6: the data set used for training and testing, the selected descriptors, the computational protocol applied for obtaining those descriptors, and the examined types of regression.

The publications are ordered by increasing complexity and diversity of the data set used, as indicated by the yellow colour gradient in the data set column, leading to more complex and diverse data sets in dark yellow. In the same way, the columns describing the descriptor class (orange) and regression type (green) are coloured. The column representing the computational protocol (blue) distinguishes between the method, environment, and code used. Light blue represents gas phase calculations, and dark blue represents calculations involving implicit solvation.

A discussion of the correlation results as well as their generalisability is beyond the scope of this work and can be found in the respective publications. Instead, we developed a colour scheme, which we propose as a rule of thumb for assessing the quality of a regression analysis, both in terms of correlation result and generalisability: the higher the number of darker colours in the different columns for a study, the higher the generalisation capabilities for the combination of descriptor and regression model.

In our opinion, the following advances of the past few years are particularly noteworthy. Mood *et al.* found that the correlation between the methyl anion affinity (MAA*) and the electrophilicity parameter E can be greatly improved or made possible in the first place by taking into account solvent effects for structurally diverse data sets.[57] Kadish *et al.* found the same qualitative results for the methyl cation affinity (MCA*) and the nucleophilicity parameters N.[58] Previously, correlation studies on individual structure classes concluded that solvent effects can be neglected as these were found to be approximately constant within a given class.[10,11,21] However, as soon as structurally diverse data sets are applied, good correlation results can no longer be expected without taking solvent effects into account. Another highlight is the work by Hoffmann *et al.*, who created a way to systematise the choice of descriptor and regression model through comprehensive study and comparison of many different examples in both categories.[48] Last but not least, Orlandi *et al.* were able to use multivariate linear regression not only to significantly improve the prediction of the nucleophilicity parameter, but also to create an interpretable model. The main influence on the nucleophilicity N comes from the proton affinity of the nucleophile as well as the solvation energy of both the nucleophile and the adduct. Steric effects also have an effect on N, but to a minor extent.[49] Boobier *et al.* came to a very similar conclusion in their study including 904 nucleophiles.[55] A possible explanation for the weak steric influence could be a bias in the database, as the majority of its nucleophiles does not suffer from significant steric hindrance.

4 Uncertainty quantification

So far, we reviewed the various known computational approaches to determining reactivity parameters (E, N, s_N) of polar species and the bimolecular rate constants of their reactions (in the form of $\log k$). This section addresses the question what the *uncertainty*, or accuracy, of those predictions is. While the aforementioned studies provided measures of the average error/performance of their prediction models, these measures neglect that prediction errors are generally nonuniformly distributed.

This nonuniformity of errors is taken into account, in a qualitative way, in Mayr's database[8,9] by a star-rating system. Each species listed in the database is assigned up to five stars, depending on the number and status (either reference or non-reference, see Box 2) of reaction partners that were employed to determine the reactivity parameter(s) of that species. The more stars, the higher the expected accuracy of the corresponding parameter(s). Five stars are reserved for reference species. Therefore, the most accurate predictions of rate constants derived from eqn (3) can be expected for reactions of reference electrophiles with reference nucleophiles. As a consequence, computational models trained on reference species tend to yield better predictions than those trained on a more diverse set of species, which reflects Mayr's uncertainty principle of reactivity: prediction accuracy (of $\log k$) and chemical diversity cannot be maximised at the same time.

Inspired by the qualitative star-rating approach, we developed an uncertainty quantification (UQ) framework[6] to estimate the uncertainty associated with the prediction of a *specific* parameter/rate constant. Case-specific uncertainties carry more information about the underlying problem than average measures of error/performance.

For instance, a desired nucleophile (in terms of reactivity) may be expensive or difficult to synthesise in contrast to a less suitable nucleophile that is readily available. In such a scenario, it would be preferable to know whether the predicted difference in reactivity is statistically significant in order to decide whether the extra work or money is worth the effort. By transforming the associated reactivity parameters into probability distributions by our UQ framework, the aforementioned significance can be quantified by calculating the normalised overlap integral (NOI) of their distributions. The possible values of the NOI range from zero (no overlap at all) to unity (identical distributions).

Another field of application is optimal design,[62] which is concerned with the challenge of achieving a certain goal with a minimum number of experiments. Given the amount of resources that are required to determine a single $\log k$ value, optimal design is certainly a task of interest. The probability distributions generated by our UQ framework can function as decision-making tool in this context. The larger the uncertainty of $\log k$ for a specific electrophile–nucleophile combination is, the more information is gained by measuring $\log k$ for this combination. By visualising an uncertainty matrix spanned by the reference electrophiles and reference nucleophiles of

Mayr's database, we could derive a rule for selecting new combinations of benzhydrylium ions and π-nucleophiles in a systematic fashion.[6] Verification by experimental colleagues is pending.

Besides synthesis planning and optimal design, case-specific uncertainties open up new possibilities for benchmark studies of computational models. Benchmarking is a crucial assessment tool for computational chemists.[63] Its success depends on the availability of highly accurate reference data, whose generation is typically resource-intensive and time-consuming and, therefore, poses a critical bottleneck for exploring broader domains of chemical space. As our framework enables the quantification of uncertainty in $\log k$ for electrophile–nucleophile combinations for which experimental $\log k$ values do not yet exist, the resulting predictions can be employed as pseudo-benchmarks.

While these novel possibilities may sound appealing, UQ only generates added value to synthesis planning, optimal design and benchmarking if high accuracy of the uncertainty estimates can be ensured. Therefore, it is important to estimate and assess uncertainty in concert. By reporting our uncertainty estimates as 95% confidence intervals, it is straightforward to determine their overall validity. We find that 99% (instead of 95%) of the actual prediction errors are located within their corresponding 95% confidence interval, suggesting a tendency to overestimate the prediction uncertainty. A more elaborate analysis by Pernot confirms this trend.[64] It also reveals that the overestimation increases with decreasing magnitude of the uncertainty estimates. The finding by Pernot is an interesting starting point for further investigation as the relative magnitude of individual contributions to the overall uncertainty correlates with its absolute magnitude. That is, the smaller the magnitude of the overall uncertainty, the larger the relative magnitude of the model error (the other contribution being parameter uncertainty).

Taking into account the latest findings, we are currently working on a second version of our UQ framework, the first version of which is available as a Python package (MAYRUQ: **https://gitlab.com/jproppe/mayruq**) under an open-access license.[65]

5 Where will the road take us?

Scales of electrophilicity and nucleophilicity have been strongly rooted in physical organic chemistry for decades. They link the reactivity of molecules – in the form of empirical parameters – to the bimolecular rate constant of polar reactions. Here, we reviewed computational approaches to constructing reactivity scales, including (i) the quantum chemical calculation of such rate constants, (ii) the estimation of reactivity parameters by means of regression analysis, and (iii) the quantification of uncertainty for these quantities. In this last section, we discuss general aspects of the computational approach, including both limitations and ideas for future directions.

As reactivity parameters are the key to predicting rate constants of polar reactions, their efficient estimation is of particular interest in cases where high-frequency data streams are processed in order to investigate chemical reactions. This scenario is not hypothetical. Reaction network exploration,[66–68] which seeks to identify reaction mechanisms with microscopic detail, is an active field of research processing extensive amounts of data. Such an undertaking is only feasible in a recursive manner where each elementary reaction of a given search generation is assessed in terms of its kinetic and thermodynamic relevance.[69] If labelled irrelevant, it will no longer be considered in future generations in order to save resources. The hundreds, thousands, or even more transition state optimisations required to evaluating kinetic importance (in the form of $\log k$) present a critical bottleneck of this procedure.

We envision a workflow in which those quantum chemical characterisations of $\log k$ are primarily replaced by Mayr–Patz estimates [eqn (3)] in combination with efficient regression-based predictions of reactivity parameters. The efficiency of these predictions is bound by the ease with which the chemical descriptors to be correlated with the reactivity parameters can be generated. At present, most of the descriptors applied in this context result from expensive quantum chemical calculations. However, in the past decade, a considerable number of inexpensive but information-rich chemical descriptors have been developed, which can be constructed from the positions and charges of a system's atomic nuclei alone.[70]

Besides efficiency, reliability is a key aspect when considering the replacement of quantum chemical calculations with regression-based predictions. As these predictions cannot be more accurate than the data they have been trained on (''garbage in, garbage out'', a popular saying among computer scientists), it is crucial to ensure high accuracy of the reactivity parameters listed in Mayr's database.[8,9] It is known that the accuracy of these parameters is not uniformly distributed across the database. For instance, reactivity parameters of reference species yield more reliable predictions of $\log k$ than those of non-reference species (*cf.* Box 2). We identify two general strategies toward enhancing the overall accuracy of database/training parameters, which are partially interdependent. The first strategy is to extending the pool of reference species. This may appear like a simple task requiring nothing more than a repetition of the least-squares optimisation task presented in Box 2. However, a successful reparametrisation also depends on the number and distribution of available $\log k$ values. Therefore, the second strategy addresses the acquisition of new $\log k$ data. To avoid redundant experimental work, optimal-design approaches coupled with uncertainty quantification can guide the planning of new experiments.[6]

In time-resolved (non-equilibrium) contexts, such as molecular dynamics simulations, different challenges arise since $\log k$ is an intrinsically time-averaged quantity. In pioneering work on this topic, Hoffmann *et al.* examined the impact of structural fluctuations on the electrophilicity parameter of Michael acceptors.[71] They provided evidence that the prediction of electrophilicity can be improved (with respect to the values listed in Mayr's database) and even made temperature-

dependent, allowing for better comparability with actual experimental data.

While the prediction of experimental observables is an undoubtedly desirable task, it does not deepen our understanding of organic reactivity *per se*. Recent studies by Hoffmann *et al.*,[48] Van Vranken and co-workers,[57,58] and Orlandi *et al.*,[49] which present attempts to systematise the search for factors governing reactivity, mark a starting point in this direction. We conclude with a quote by Herbert Mayr, which we consider a signpost guiding theoretical chemists along the computational road to reactivity scales: "We still consider the long linearity of the log *k versus E* correlations, which we find again and again, as a challenge to theory. It should be noted that this amazing linearity is completely independent of the validity or nonvalidity of [the Mayr–Patz equation]. [...] I repeat my appeal to theoreticians. There is a lack in our understanding of organic reactivity."[7]

Author contributions

Maike Vahl: conceptualisation, investigation, visualisation, writing – original draft, writing – review and editing. Jonny Proppe: conceptualisation, investigation, project administration, supervision, writing – original draft, writing – review and editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors acknowledge funding by Germany's joint federal and state program supporting early-career researchers (WISNA) established by the Federal Ministry of Education and Research (BMBF).

Notes and references

1 P. Muller, *Pure Appl. Chem.*, 1994, **66**, 1077–1184.
2 H. Mayr, *Tetrahedron*, 2015, **71**, 5095–5111.
3 H. Mayr and M. Patz, *Angew. Chem., Int. Ed. Engl.*, 1994, **33**, 938–957.
4 C. G. Swain and C. B. Scott, *J. Am. Chem. Soc.*, 1953, **75**, 141–147.
5 C. D. Ritchie, *Acc. Chem. Res.*, 1972, **5**, 348–354.
6 J. Proppe and J. Kircher, *ChemPhysChem*, 2022, **23**, e202200061.
7 H. Mayr, *Angew. Chem., Int. Ed.*, 2011, **50**, 3612–3618.
8 H. Mayr and A. R. Ofial, *SAR QSAR Environ. Res.*, 2015, **26**, 619–646.
9 H. Mayr and A. R. Ofial, Mayr's Database of Reactivity Parameters, **https://www.cup.lmu.de/oc/mayr/reaktionsdatenbank2/**, last accessed on 24 August 2022.
10 P. Pérez, A. Toro-Labbé, A. Aizman and R. Contreras, *J. Org. Chem.*, 2002, **67**, 4747–4752.
11 C. Schindele, K. N. Houk and H. Mayr, *J. Am. Chem. Soc.*, 2002, **124**, 11208–11214.
12 H. Eyring, *J. Chem. Phys.*, 1935, **3**, 107–115.
13 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3094.
14 J. W. Ochterski, *Thermochemistry in Gaussian*, Gaussian, Inc, 2000.
15 S. Grimme, *Chem. – Eur. J.*, 2012, **18**, 9955–9964.
16 J. N. Harvey, F. Himo, F. Maseras and L. Perrin, *ACS Catal.*, 2019, **9**, 6803–6813.
17 J. Proppe, T. Husch, G. N. Simm and M. Reiher, *Faraday Discuss.*, 2016, **195**, 497–520.
18 P. Hänggi, P. Talkner and M. Borkovec, *Rev. Mod. Phys.*, 1990, **62**, 251–341.
19 A. S. Petit and J. N. Harvey, *Phys. Chem. Chem. Phys.*, 2011, **14**, 184–191.
20 K. Jorner, A. Tomberg, C. Bauer, C. Sköld and P.-O. Norrby, *Nat. Rev. Chem.*, 2021, **5**, 240–255.
21 C. Wang, Y. Fu, Q.-X. Guo and L. Liu, *Chem. – Eur. J.*, 2010, **16**, 2586–2598.
22 L.-G. Zhuo, W. Liao and Z.-X. Yu, *Asian J. Org. Chem.*, 2012, **1**, 336–345.
23 D. S. Allgäuer, H. Jangra, H. Asahara, Z. Li, Q. Chen, H. Zipse, A. R. Ofial and H. Mayr, *J. Am. Chem. Soc.*, 2017, **139**, 13318–13329.
24 H. Jangra, Q. Chen, E. Fuks, I. Zenz, P. Mayer, A. R. Ofial, H. Zipse and H. Mayr, *J. Am. Chem. Soc.*, 2018, **140**, 16758–16772.
25 R. J. Mayer, M. Breugst, N. Hampel, A. R. Ofial and H. Mayr, *J. Org. Chem.*, 2019, **84**, 8837–8858.
26 Z. Li, R. J. Mayer, A. R. Ofial and H. Mayr, *J. Am. Chem. Soc.*, 2020, **142**, 8383–8402.
27 J. Zhang, Q. Chen, R. J. Mayer, J.-D. Yang, A. R. Ofial, J.-P. Cheng and H. Mayr, *Angew. Chem., Int. Ed.*, 2020, **59**, 12527–12533.
28 W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133–A1138.
29 N. Mardirossian and M. Head-Gordon, *Mol. Phys.*, 2017, **115**, 2315–2372.
30 L. Goerigk, A. Hansen, C. Bauer, S. Ehrlich, A. Najibi and S. Grimme, *Phys. Chem. Chem. Phys.*, 2017, **19**, 32184–32215.
31 A. J. Cohen, P. Mori-Sánchez and W. Yang, *Science*, 2008, **321**, 792–794.
32 R. A. Friesner, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6648–6653.
33 P. Morgante and R. Peverati, *Int. J. Quantum Chem.*, 2020, **120**, e26332.
34 M. Bursch, J.-M. Mewes, A. Hansen and S. Grimme, *Angew. Chem., Int. Ed.*, 2022, **61**, e202205735; *Angew. Chem.*, 2022, **134**, e202205735.
35 E. Dzib and G. Merino, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2022, **12**, e1583.
36 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2013, **9**, 609–620.
37 G. N. Simm, P. L. Türtscher and M. Reiher, *J. Comput. Chem.*, 2020, **41**, 1144–1155.
38 D. R. Glowacki, C.-H. Liang, C. Morley, M. J. Pilling and S. H. Robertson, *J. Phys. Chem. A*, 2012, **116**, 9545–9560.

39 Y. Li, L. Zhang and S. Luo, *ACS Omega*, 2022, **7**, 6354–6374.
40 T. A. Halgren, *J. Comput. Chem.*, 1996, **17**, 490–519.
41 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.
42 H. Mayr, T. Bug, M. F. Gotta, N. Hering, B. Irrgang, B. Janker, B. Kempf, R. Loos, A. R. Ofial, G. Remennikov and H. Schimmel, *J. Am. Chem. Soc.*, 2001, **123**, 9500–9512.
43 J. Ammer, C. Nolte and H. Mayr, *J. Am. Chem. Soc.*, 2012, **134**, 13902–13911.
44 C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, New York, 2006.
45 C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge (MA), United States, 2006.
46 F. Pereira, D. A. R. S. Latino and J. Aires-de-Sousa, *J. Org. Chem.*, 2011, **76**, 9312–9319.
47 S.-i Kiyooka, D. Kaneno and R. Fujiyama, *Tetrahedron*, 2013, **69**, 4247–4258.
48 G. Hoffmann, M. Balcilar, V. Tognetti, P. Héroux, B. Gaüzère, S. Adam and L. Joubert, *J. Comput. Chem.*, 2020, **41**, 2124–2136.
49 M. Orlandi, M. Escudero-Casao and G. Licini, *J. Org. Chem.*, 2021, **86**, 3555–3564.
50 B. C. Haas, A. E. Goetz, A. Bahamonde, J. C. McWilliams and M. S. Sigman, *Proc. Natl. Acad. Sci. U. S. A.*, 2022, **119**, e2118451119.
51 B. Settles, *Active Learning*, Morgan & Claypool, San Rafael (CA), United States, 2012, vol. 18.
52 J. Proppe, S. Gugler and M. Reiher, *J. Chem. Theory Comput.*, 2019, **15**, 6046–6060.
53 P. Geerlings, F. De Proft and W. Langenaeker, *Chem. Rev.*, 2003, **103**, 1793–1874.
54 B. Lee, J. Yoo and K. Kang, *Chem. Sci.*, 2020, **11**, 7813–7822.
55 S. Boobier, Y. Liu, K. Sharma, D. R. J. Hose, A. J. Blacker, N. Kapur and B. N. Nguyen, *J. Chem. Inf. Model.*, 2021, **61**, 4890–4899.
56 R. F. W. Bader, *Atoms in Molecules: A Quantum Theory*, Oxford University Press, Oxford, UK, 1990.
57 A. Mood, M. Tavakoli, E. Gutman, D. Kadish, P. Baldi and D. L. Van Vranken, *J. Org. Chem.*, 2020, **85**, 4096–4102.
58 D. Kadish, A. D. Mood, M. Tavakoli, E. S. Gutman, P. Baldi and D. L. Van Vranken, *J. Org. Chem.*, 2021, **86**, 3721–3729.
59 V. Saini, A. Sharma and D. Nivatia, *Phys. Chem. Chem. Phys.*, 2022, **24**, 1821–1829.
60 M. Tavakoli, A. Mood, D. Van Vranken and P. Baldi, *J. Chem. Inf. Model.*, 2022, **62**, 2121–2132.
61 W. Nie, D. Liu, S. Li, H. Yu and Y. Fu, *J. Chem. Inf. Model.*, 2022, **62**, 4319–4328.
62 F. Pukelsheim, *Optimal Design of Experiments*, Society for Industrial & Applied Mathematics, Philadelphia, PA, 2006.
63 R. A. Mata and M. A. Suhm, *Angew. Chem., Int. Ed.*, 2017, **56**, 11011–11018.
64 P. Pernot, *J. Chem. Phys.*, 2022, **156**, 114109.
65 J. Proppe, Uncertainty Quantification of Reactivity Scales, **https://gitlab.com/jproppe/mayruq**, last accessed on 24 August 2022.
66 A. L. Dewyer, A. J. Argüelles and P. M. Zimmerman, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**, e1354.
67 J. P. Unsleber and M. Reiher, *Annu. Rev. Phys. Chem.*, 2020, **71**, 121–142.
68 M. Steiner and M. Reiher, *Top. Catal.*, 2022, **65**, 6–39.
69 J. Proppe and M. Reiher, *J. Chem. Theory Comput.*, 2019, **15**, 357–370.
70 F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi and M. Ceriotti, *Chem. Rev.*, 2021, **121**, 9759–9815.
71 G. Hoffmann, V. Tognetti and L. Joubert, *Chem. Phys. Lett.*, 2019, **724**, 24–28.

