Reliable Prediction of the Octanol–Air Partition Ratio

Sivani Baskaran, Ying Duan Lei, and Frank Wania*

Department of Physical and Environmental Sciences and Department of Chemistry, University of Toronto Scarborough, Toronto, Ontario, Canada

Abstract: The octanol–air equilibrium partition ratio ($K_{OA}$) is frequently used to describe the volatility of organic chemicals, whereby $n$-octanol serves as a substitute for a variety of organic phases ranging from organic matter in atmospheric particles and soils, to biological tissues such as plant foliage, fat, blood, and milk, and to polymeric sorbents. Because measured $K_{OA}$ values exist for just over 500 compounds, most of which are nonpolar halogenated aromatics, there is a need for tools that can reliably predict this parameter for a wide range of organic molecules, ideally at different temperatures. The ability of five techniques, specifically polyparameter linear free energy relationships (ppLFERs) with either experimental or predicted solute descriptors, EPISuite’s KOAWIN, COSMOtherm, and OPERA, to predict the $K_{OA}$ of organic substances, either at 25 °C or at any temperature, was assessed by comparison with all $K_{OA}$ values measured to date. In addition, three different ppLFER equations for $K_{OA}$ were evaluated, and a new modified equation is proposed. A technique’s performance was quantified with the mean absolute error (MAE), the root mean square error (RMSE), and the estimated uncertainty of future predicted values, that is, the prediction interval. We also considered each model’s applicability domain and accessibility. With an RMSE of 0.37 and a MAE of 0.23 for predictions of log $K_{OA}$ at 25 °C and RMSE of 0.32 and MAE of 0.21 for predictions made at any temperature, the ppLFER equation using experimental solute descriptors predicted the $K_{OA}$ the best. Even if solute descriptors must be predicted in the absence of experimental values, ppLFERs are the preferred method, also because they are easy to use and freely available. Environ Toxicol Chem 2021;40:3166–3180. © 2021 The Authors. Environmental Toxicology and Chemistry published by Wiley Periodicals LLC on behalf of SETAC.

Keywords: Quantitative structure–activity relationships; Organic contaminants; Environmental partitioning; Partitioning coefficient; Partitioning ratio

INTRODUCTION

The use and importance of equilibrium partition ratios have been well established in environmental chemistry and in chemical exposure and risk assessment (Mackay et al., 2015; Schwarzenbach et al., 2005), in agricultural chemistry (Lacoste et al., 2020), and in the pharmaceutical sciences (Lipinski et al., 1997). The octanol–air equilibrium partition ratio ($K_{OA}$) describes the distribution of a compound between $n$-octanol and the gas phase at equilibrium. The log $K_{OA}$ is frequently used to describe the volatility of organic chemicals or their tendency to be absorbed from the gas phase. The solvent $n$-octanol serves as a substitute for a variety of organic phases or sorbents. This includes organic matter in atmospheric particles (Finizio et al., 1997), soils (Finizio et al., 1998), and sludge (Cousins et al., 1998), as well as biological tissues including plant foliage (Paterson et al., 1990), lipids (Kelly & Gobas, 2003), blood, and milk (Batterman et al., 2002). It has also been applied to describe organic vapors partitioning into technical sorbents, such as polyurethane foam (Shoeb & Harner, 2002a) and silicone (Anderson et al., 2017), as well as phases such as organic films (Hamer et al., 2003), house dust (Bennett & Furtaw, 2004), and cotton and polyester (Saini et al., 2016). Accordingly, it has been increasingly used in chemical exposure and risk assessment including estimations of a chemical’s potential for bioaccumulation in air-breathing species (Gobas et al., 2003).

Laboratory measurements of the $K_{OA}$ by methods such as the generator column (Hamer & Mackay, 1995), headspace (Lei et al., 2019; Xu & Kropscott, 2012, 2013, 2014), and gas chromatography retention time techniques (Su et al., 2002; Wania et al., 2002; Zhang et al., 1999) are often time consuming, difficult, and/or expensive. Reliable prediction methods could be useful as an alternative to determining the $K_{OA}$ experimentally. Multiple quantitative structure–activity
relationships (QSARs) have been presented for predicting the $K_{OA}$ of organic chemicals. QSARs are commonly used in the pharmaceutical industry and in environmental science and chemistry to predict a property of a chemical based on its structure. Many of these QSARs are focused on predicting the $K_{OA}$ of a specific subset of closely related chemicals, for example, polychlorinated dibenzo-p-dioxins (PCDDs; Chen et al., 2002; Zeng et al., 2013), polybrominated diphenylethers (PBDEs; Chen et al., 2003a; Liu et al., 2013; Xu et al., 2007), polychlorinated naphthalenes (PCNs; Chen et al., 2003b), polychlorinated biphenyls (PCBs; Chen et al., 2003c; Chen et al., 2016; Li et al., 2020; Yuan et al., 2016), and polycyclic aromatic hydrocarbons (PAHs; Ferreira, 2001); thus they are too limited in their applicability domain for most purposes.

Not all prediction models are limited to a specific subset of compounds. Polyparameter linear free energy relationships (ppLFERs) use system constants that describe the characteristics of the bulk phases, whereas solute descriptors describe specific characteristics of a solute. If experimental solute descriptors are not available, they can be estimated from molecular structure using QSARs (e.g., ACD/Ab solv [Advanced Chemistry Development, 2021] and IFSQSAR [Brown et al., 2012]). Lack of structural diversity among the chemicals used in the calibration of the system constants can limit the applicability domain of a ppLFER equation. COSMOTHERM, OPERA, EPI SUITE™ (US Environmental Protection Agency [USEPA], 2012), and SPARC Performs Automated Reasoning in Chemistry (SPARC) can also estimate the $K_{OA}$ from molecular structure.

Different estimation software can yield divergent results and consequently may result in inconsistent regulatory decisions depending on the selected model (Zhang et al., 2010). It is important to compare the performance of various estimation techniques for $K_{OA}$ prior to implementation of predictions within regulatory procedures. For an estimation technique to be useful, it must be able to predict the $K_{OA}$ for a large variety of chemicals with quantifiable uncertainty, which is largely dependent on the diversity of its training and validation sets.

To our knowledge, no assessment has compared different prediction techniques for $K_{OA}$, but previous work has compared the performance of COSMOTHERM, A bsolv, and SPARC with respect to their ability to predict the hexadecane–air partition ratio (Bronner et al., 2010; Stenzel et al., 2012). Bronner et al. (2010) found that ppLFERs in combination with Absolv-predicted solute descriptors (root mean square error [RMSE]$=0.40$) were best able to predict the hexadecane–air partition ratio for bifunctional compounds, whereas for pesticides, drugs, and hormones the COSMOTHERM model had the smallest RMSE (0.97). Stenzel et al. (2012) also reported that COSMOTHERM (RMSE$=0.94$) and ppLFER/Absolv (RMSE$=0.99$) performed better than the SPARC model for predicting hexadecane–air partition ratios. Further research by Stenzel et al. (2014) in assessing the ability of these approaches to predict partitioning in four liquid/liquid systems again found that COSMOTHERM and ppLFER/Absolv had similar levels of performance, and performed much better than SPARC.

In the present study, using an exhaustive data set of experimentally determined $K_{OA}$ values recently compiled by Baskaran et al. (2021), we comparatively evaluated the ability of five methods, namely, EPI SUITE, COSMOTHERM, OPERA, and the ppLFERs with either experimental or estimated solute descriptors, to reliably predict the $K_{OA}$ of an organic compound both at $25 ^\circ C$ and at any temperature. The five different prediction techniques were selected based on their capacity to predict $K_{OA}$ for a wide range of chemicals.

**MATERIALS AND METHODS**

**Outline of the approach**

The basis of the evaluation was deviations of predicted from experimental $\log K_{OA}$. Specifically, we used RMSEs, residuals ($\log K_{observed} - \log K_{predicted}$), and mean absolute errors (MAEs) as numerical criteria of model performance. Positive residuals indicate the model is underpredicting the $\log K_{OA}$, and negative residuals indicate an overprediction. The MAE is the average of the absolute residuals:

$$MAE = \frac{\sum |\log K_{OA\ observed} - \log K_{OA\ predicted}|}{n}$$  

A 95% prediction interval for future estimates was calculated for each model using the mean ($x$) and standard deviation (SD) of the residuals. The prediction interval is a confidence interval that predicts the residual for a future estimate ($x_{new}$) made using the model, and can therefore give a range in which the true $\log K_{OA}$ value will likely fall.

$$x - 1.96 \times SD < x_{new} < x + 1.96 \times SD$$  

The performance of the model was compared over the whole range of experimental data, whereby the average of multiple measurements for the same chemical at the same temperature was taken ($n = 1453$). In addition, the model performance was assessed for the much smaller set of chemicals for which all five techniques succeeded in making $\log K_{OA}$ predictions at $25 ^\circ C$ ($n = 439$).

The methods EPI SUITE, OPERA, and ppLFERs with estimated solute descriptors also provide information on how well a chemical’s molecular structure fits into the model’s applicability domain and how reliable the prediction therefore is likely to be. We also evaluated separately the deviations from measurements of predictions belonging to different categories of reliability. Because the applicability domain for each model is defined differently, these categories of reliability cannot be compared between the different estimation techniques but are rather a measure of how reliable a model considers its prediction. For example, a prediction classified as “good” from a ppLFER using estimated solute predictions is not comparable to an OPERA estimated value marked “excellent.” Instead, these are independent assessments that tell us that the ppLFER system considers this prediction to be reliable and the OPERA model assesses this prediction to be very good. Although the rankings of different models cannot be compared, we
can compare the prediction performance for the different categories of reliability of one model.

Finally, we also compare the different techniques with respect to the size of their applicability domain, and the accessibility and ease-of-use of the modeling software.

**Experimental data**

The data used in the present study were taken from an extensive literature review that identified 2017 experimental $K_{OA}$ values in the literature (Baskaran et al., 2021). After $K_{OA}$ values for wet octanol ($K_{OA}$), chemical mixtures, ambiguous or nonorganic chemicals, and $K_{OA}$ values judged unreliable (Baskaran et al., 2021) were removed from the data set, 1950 experimental data points for 604 organic chemicals measured between −10 and 110 °C were included.

The temperature of each experimental log $K_{OA}$ was rounded to the nearest tenth. If there was more than one log $K_{OA}$ measurement for a chemical at a given temperature, the average experimental log $K_{OA}$ was used. This procedure yielded 475 and 1673 unique log $K_{OA}$ values at 25 °C and any temperature, respectively. The original calibration of ppLiFER and OPERA models was based on measured $K_{OA}$ data, and thus the measurements used to train and validate these models are included in our experimental data set. These values have not been removed from our analysis, and may bias the results of the assessment slightly in favor of models with more chemicals in the training and validation data sets. Note that the COSMOtherm and EPISuite models require no external calibration.

The simplified molecular-input line-entry system (SMILES) notations, CAS numbers, and names of the chemicals were included in the database by Baskaran et al. (2021). Structure data files (SDFs) having explicit hydrogen bonding and three-dimensional coordinates were generated using Open Babel (Ver 2.4.1; O’Boyle et al., 2011).

**Prediction techniques**

**ppLiFERS.** Two of the prediction techniques included in the comparison are ppLiFERS that use solute descriptors provided by the online UFZ-LSER Database (Ulrich et al., 2017). The ppLiFERS quantify a solute’s interactions with bulk phases using descriptors reflecting the properties of the solute and system constants characterizing the solvating phases (Endo & Goss, 2014). Solute descriptors include a chemical’s H-bond basicity (B), H-bond acidity (A), polarizability (S), excess molar refraction (E), McGowan molar volume (V), and log hexadecane–air partition ratio (L). The L and V terms describe the size of the chemical and general solute–solvent interactions, and the remaining terms quantify specific solute–solvent interactions (Abraham et al., 2001). Specifically, A and B terms describe hydrogen-bonding interactions (Abraham, 1993), S accounts for solute polarizability and dipolarity (Abraham et al., 1990), and E also describes polarizability and interactions between π and lone electron pairs (Abraham et al., 1990; Abraham, 1993).

Constants for the octanol–air system, determined using multiple linear regressions of measured log $K_{OA}$ values against solute descriptors, have been reported by Abraham and Acree (2008; Abraham et al., 2010; Ulrich et al., 2017; Equation 3) and Endo and Goss (2014; Equation 4). The standard error (SE) associated with each system constant is included in the parentheses.

$$
\log K_{OA} \text{ at } 25^\circ C = \left( -0.214 \pm 0.034 \right) \times E + \left( 0.561 \pm 0.042 \right) \\
\times S + \left( 3.507 \pm 0.072 \right) \times A + \left( 0.749 \pm 0.049 \right) \times B + \left( 0.943 \pm 0.007 \right) \times L \\
- \left( 0.147 \pm 0.022 \right)
$$

(3)

$$
\log K_{OA} \text{ at } 25^\circ C = \left( 0.470 \pm 0.027 \right) \times S + \left( 3.520 \pm 0.070 \right) \\
\times A + \left( 0.816 \pm 0.055 \right) \times B + \left( 0.071 \pm 0.052 \right) \times V + \left( 0.903 \pm 0.016 \right) \times L \\
- \left( 0.143 \pm 0.027 \right)
$$

(4)

The difference between Equations (3) and (4) is the use of either E or V as a solute descriptor. When used in conjunction, as in Equation (3), the E and L (or V) parameters together describe the cavity energy and van der Waals interactions (Goss, 2005). Because V describes the size of the cavity, it must inherently contain information on the cavity energy, which is the basis of ppLiFER equations of type 4 (Goss, 2005). The latter is often preferred because V is easily calculated.

For most organic compounds, V and L are highly correlated; exceptions are polyfluorinated compounds and organosilicon compounds such as the cyclic volatile methyl siloxanes (Endo & Goss, 2014; Supporting Information, Figure SI 4). The system constant for the V term in Equation (4) is very small, has a ρ value of 0.17, and has a high relative uncertainty, suggesting that this term may not be necessary. Further details are provided in the Supporting Information. Thus, using the same data set of 181 chemicals used by Endo and Goss (2014), we calibrated a ppLiFER equation that uses only the $S$, $A$, $B$, and $L$ parameters:

$$
\log K_{OA} \text{ at } 25^\circ C = \left( 0.422 \pm 0.035 \right) \times S + \left( 3.516 \pm 0.070 \right) \\
\times A + \left( 0.841 \pm 0.052 \right) \times B + \left( 0.923 \pm 0.006 \right) \times L - \left( 0.123 \pm 0.023 \right)
$$

(5)

Mintz et al. (2008) presented a ppLiFER for the enthalpy of solvation in octanol ($\Delta H_{OA}^\circ$ in units of kJ/mol):

$$
\Delta H_{OA}^\circ = \left( 6.044 \pm 1.104 \right) \times S - \left( 53.565 \pm 2.373 \right) \times A \\
- \left( 9.190 \pm 1.120 \right) \times B + \left( 1.566 \pm 1.397 \right) \times V \\
- \left( 9.663 \pm 0.336 \right) \times L - \left( 6.672 \pm 0.741 \right)
$$

(6)

Whereas $\Delta H_{OA}^\circ$ is temperature dependent, a $\Delta H_{OA}^\circ$ value obtained with Equation (6) is assumed to apply to the range from −10 to 45 °C (Mintz et al., 2008). We did not consider an alternative ppLiFER for $\Delta H_{OA}^\circ$ that uses the $V$ solute descriptor instead of $V$ (Mintz et al., 2007). The $\Delta H_{OA}^\circ$ can be converted to an internal energy of phase transfer between octanol and the gas phase ($\Delta U_{OA}^\circ$ in units of kJ/mol; Atkinson & Curthoys, 1978; Goss & Eisenreich, 1996) using $R$, the ideal gas constant ($8.314 \times 10^{-3}$ kJ K⁻¹ mol⁻¹), and the temperature $T$ at which $\Delta H_{OA}^\circ$ was measured (assumed to be 298.15 K):
The use of Equation (9) eliminates the need to use Equations (6–8) to derive a temperature-dependent estimate of the log $K_{OA}$.

Currently only the system constants for Equations (3) and (6) are integrated within the UFZ-LSER website (Ulrich et al., 2017). In the literature, one can also find ppLFERs for the partitioning between water-saturated (“wet”) octanol and the gas phase (Endo & Goss, 2014; Flanagan et al., 2005). Most measured $K_{OA}$ data are, however, for partitioning between “dry” octanol and the gas phase.

The UFZ-LSER Database (Ulrich et al., 2017) provides experimental and estimated solute descriptors. Chemicals were identified within the UFZ website using either the CAS number or, if not available, the universal SMILES format. For some chemicals the UFZ database contains multiple sets of solute descriptors from different peer-reviewed articles. In the present study we used three types of solute descriptors:

1. those experimentally derived and reported in the peer-reviewed literature, labeled as “UFZ-preselected published values,”
2. those that have been part of the Absolv database, which are for the most part experimentally derived values, but for which an explicit reference is not available, and
3. those that have been predicted with QSAR models built into the UFZ-LSER website and that rely on the Iterative Fragment Selection algorithm (Brown et al., 2012).

It is expected that the reliability of the solute descriptors decreases from 1 to 3. When we selected experimental solute descriptors for the chemicals with experimental $K_{OA}$ values, preference was therefore given to descriptors from peer-reviewed studies over the Absolv data set. In general, the “UFZ-preselected published values” were very similar to the Absolv solute descriptors. Experimental solute descriptors can be determined by measuring partitioning ratios, solubilities, and chromatographic data (Sprunger et al., 2007). It is assumed that if experimental solute descriptors exist for a given chemical, it is inherently within the applicability domain of the ppLFER model.

In addition, estimated solute descriptors were determined for all chemicals with experimental $K_{OA}$ values using the IFSQSARs (available from Brown, 2020; also see Brown, 2014; Brown et al., 2012). The IFSQSAR models have been integrated into the UFZ website (Ulrich et al., 2017). For PCDDs with a single substitution in the 2- or 7-position (i.e., PCDD 26, PCDD 29, and PCDD 50), the estimated $L$ value was corrected by +4.039 (T.N. Brown, personal communication, 2020). Ultimately, we evaluated the predictive performance of six ppLFERs: Equations (3), (4), and (5), each with either experimental or IFSQSAR-predicted solute descriptors. Of these six ppLFERs, we selected the best performing equations for both experimental and estimated solute descriptors for comparison with the other prediction techniques. The UFZ website provides information on how well a molecule fits within the applicability domain of the IFSQSAR models used to estimate solute descriptors. It is based on a chemical similarity score and leverage value. Using the IFSQSAR model directly provides an SE for each predicted solute descriptor (except $V$; Brown, 2020). Because SEs for the system constants in Equations (3)–(6) are also available, we applied a Monte Carlo analysis to calculate the overall error of a predicted log $K_{OA}$. The Supporting Information includes a sample calculation. Using the overall error of the prediction, we assigned a reliability score to each prediction (Table 1). Because the SEs for experimental solute descriptors are usually not available, only the error of the system constants could be considered in that case.

**EPISuite.** The KOAWIN model for predicting $K_{OA}$ is a part of the US EPA’s (2012) Estimation Programs Interface (EPI) Suite software. It predicts $K_{OA}$ from a thermodynamic triangle with the Henry’s law constant ($HLC$, Pa m$^3$ mol$^{-1}$) and the octanol–water equilibrium partition ratio ($K_{OW}$):

$$K_{OA} = K_{OW} \times RT/HLC$$

where $R$ is the ideal gas constant (8.314 Pa mol K$^{-1}$ m$^{-3}$) and $T$ is the temperature (K) of the HLC. If the dimensionless HLC air–water partition ratio ($K_{AW}$) is used, this simplifies to:

**TABLE 1:** Reliability score of polyparameter linear free energy relationship (ppLFER) predictions based on the overall error (OE) of the estimate

| Reliability score | Guideline |
|-------------------|-----------|
| Poor              | OE > 1    |
| Fair              | $OE \leq 1$ |
| Good              | $OE \leq 0.75$ |
| Excellent         | $OE \leq 0.5$ |
The $K_{OA}$ describes chemical partitioning between two mutually soluble solvents. By using measured or estimated $K_{OW}$ to derive $K_{OA}$ in a thermodynamic triangle, KOAWIN is calculating a partition ratio between water-saturated octanol and air, which we denote as $K_{OA}$.

The KOAWIN model can provide two $K_{OA}$ values for a compound, using either $K_{OW}$ and HLC values estimated from KOAWIN & HENRYWIN or by substituting any available experimental data for these estimated values. Although the latter is expected to provide more reliable results, for many chemicals EPISuite’s PhysProp database does not contain measured HLC values. Consequently, the value for $K_{OA}$ recommended by KOAWIN is often a value at least partially derived from estimated values.

The $K_{OW}$ and HLC at 25 °C are estimated with KOAWIN and HENRYWIN, respectively, which are both fragment-based QSARs. By default, KOAWIN uses HLCs predicted by the bond method within HENRYWIN, which can make predictions for a larger range of compounds than the group contribution method, although it is expected to be less accurate (Meylan & Howard, 1991). The KOAWIN model assumes that the $K_{OW}$ does not vary greatly with temperature, and the effect of temperature on $K_{OA}$ can be estimated from the temperature dependence of the HLC (Meylan & Howard, 2005). The HENRYWIN model uses Equation (11) to express the temperature dependence of the HLC (USEPA, 2012):

$$\ln HLC = A_h - B_h/T$$  \hspace{1cm} (11)

Whereas $A_h$ and $B_h$ for some compounds have been compiled from the literature, a slope analogy method is applied to estimate $B_h$ for most compounds, whereby classes of similar chemicals, such as different aldehydes or PCB congeners, share the same slope $B_h$ (USEPA, 2012). The $A_h$ is then obtained using the HLC at 25 °C, which is an experimental value, if available.

Version 1.11 of KOAWIN was run in batch mode using EPISuite Version 4.11 (USEPA, 2012) with the 2017 updated files provided on the US EPA website. The temperature dependence equations were obtained from HENRYWIN Version 3.21 in batch mode. Thus, two sets of estimated $K_{OA}$ values were obtained from EPISuite. The first set comprises predictions made at 25 °C only, using the HENRYWIN and KOAWIN predictions, hereafter referred to as EPISuite-25. The second set, referred to as EPISuite-T, uses the temperature-dependent equations provided by HENRYWIN to adjust the HLC in the calculation of $K_{OA}$. Because these equations use both experimental and estimated HLC values, the resulting $K_{OA}$ at 25 °C can differ from those made by EPISuite-25.

Because the $K_{OA}$ is determined using a thermodynamic triangle, KOAWIN does not have an applicability domain. However the applicability domains of the KOWWIN and HENRYWIN models should be considered. The applicability domains of HENRYWIN include both a range for the molecular mass (26.04–451.47 g/mol) and the log $K_{AW}$ (−11.64–2.92) of the chemicals in the training set (USEPA, 2012). The reported applicability domain of KOWWIN is based on the molecular mass range of chemicals in the training set (18.02–719.92 g/mol; USEPA, 2012). We assign a reliability score for every EPISuite estimation of $K_{OA}$ based on how a chemical fits within the applicability domains of HENRYWIN and KOWWIN. Equation (11) does not have an applicability domain limit, and in some cases the HLC value used is experimental, so the classifications of the predictions for EPISuite-25 and EPISuite-T are different (Table 2).

**TABLE 2: Reliability of the EPISuite-25 predictions determined using the applicability domain (AD) set by the KOAWIN and HENRYWIN models**

| EPISuite set | Reliability score | Guideline |
|-------------|------------------|-----------|
| EPISuite-25 | Poor             | Outside all 3 AD limits |
|             | Fair             | Outside 2 of the AD limits |
|             | Good             | Outside 1 of the AD limits |
|             | Excellent        | Inside all AD limits |
| EPISuite-T  | Poor             | Outside all 3 AD limits |
|             | Fair             | Outside 2 of the AD limits |
|             | Good             | Outside of the KOWWIN AD or uses slope analogy to obtain the HLC equation |
|             | Excellent        | Inside KOAWIN AD, experimental HLC equation |

HLC = Henry’s law constant.

**OPERA.** The command-line version of the OPEn structure–activity/property Relationship App (OPERA) model (Version 2.5) by Mansouri et al. (2018) was downloaded via GitHub (available from Mansouri, 2018; also see Mansouri & Williams, 2017). The OPERA model for the $K_{OA}$ at 25 °C is a QSAR model that uses molecular descriptors calculated using PaDEL and a weighted k-nearest neighbor approach (Mansouri & Williams, 2017). It was developed with the PhysProp database within EPISuite (Mansouri & Williams, 2017). The two PaDEL descriptors used for $K_{OA}$ prediction are the number of H-bond donors, expressing the chemical’s capacity for hydrogen bonding, and the log of the gas–hexadecane partitioning ratio (Mansouri & Williams, 2017). The number of H-bond donors plays a similar role to the solute descriptors $\beta$, $A$, and $S$ in a ppLFER, whereas the log hexadecane–air partition ratio is identical to the solute descriptor $L$. The OPERA model does not estimate the $K_{OA}$ at temperatures other than 25 °C.

The applicability domain of the model is assessed through a global and a local applicability domain index (Mansouri & Williams, 2017). The global index assesses whether a chemical fits within the space of the training set used to create OPERA (Mansouri et al., 2018), and the local index compares how similar the chemical is to the five nearest neighbors in the model space (Mansouri et al., 2018). The confidence level index assesses the reliability of the prediction based on the distances of the chemical to its nearest neighbors and the accuracy of the predictions for these nearest neighbors (Mansouri et al., 2018). For chemicals that fall within the global applicability domain, the confidence level index provides additional information regarding the reliability of the prediction (Mansouri et al., 2018). Each estimate from the OPERA model is
The BP formers were used by COSMOtherm (Ver 20.0.0) using the phase transfer (the different phases and subsequently the Gibbs free energy of which combined with statistical thermodynamics allows for the interaction energy of the chemical in octanol and the gas phase, different conformations of the compound to quantify the in-
COSMOtherm model uses the polar charge density of the compound in the virtual conductor (Klamt et al., 2009). The resulting electron density and geometry of the molecule based on its polar charge density and how those calculations, determines different possible conformations of a molecule.

In brief, COSMOconf with TURBOMOLE, using DFT density functional theory (DFT) and statistical thermodynamics COSMOconf with TURBOMOLE and COSMOtherm from Das-

| Reliability score | Global AD level | Local AD level | Confidence level index |
|-------------------|-----------------|----------------|-----------------------|
| Poor              | Outside         | <0.4           |                       |
| Fair              | Outside         | 0.4–0.6        |                       |
| Fair              | Inside          | <0.6           |                       |
| Good              | Outside         | ≥0.6           | <0.75                 |
| Good              | Inside          | ≥0.6           | ≥0.75                 |
| Excellent         | Inside          | ≥0.6           | ≥0.75                 |

categorized as excellent, good, fair, or poor, based on the importance of the global and local applicability domain level, and the confidence index (Table 3).

**COSMO-RS.** The use of the COnductor like Screening Model for Realistic Solvents (COSMO-RS) software suite requires both COSMOconf with TURBOMOLE and COSMOtherm from Dassault Systèmes. The approach uses both quantum chemical density functional theory (DFT) and statistical thermodynamics of the molecular interactions to predict \( K_{OA} \) (Klamt et al., 2009). In brief, COSMOconf with TURBOMOLE, using DFT/COSMO calculations, determines different possible conformations of a molecule based on its polar charge density and how those charges interact with a virtual conductor (Klamt et al., 2009). The resulting electron density and geometry of the molecule are used to identify the most energetically optimal state for the compound in the virtual conductor (Klamt et al., 2009). The COSMOtherm model uses the polar charge density of the different conformations of the compound to quantify the interaction energy of the chemical in octanol and the gas phase, which combined with statistical thermodynamics allows for the calculation of the chemical potential of the compound in the different phases and subsequently the Gibbs free energy of the phase transfer (\( \Delta G_{OA} \); Klamt et al., 2009).

The SDFs were entered into COSMOconf (Ver 20.0.0) with TURBOMOLE (Ver 4.5) to generate COSMO files. All conformers were used by COSMOtherm (Ver 20.0.0) using the BP-TZVPD-FINE+GAS parameterization to calculate the \( \Delta G_{OA} \) for each chemical at a given temperature, which is used to calculate the \( K_{OA} \):

\[
K_{OA} = e^{-\Delta G_{OA}/(R \times T)} \tag{12}
\]

Because this method, hereafter referred to as COSMOtherm, does not involve calibration using existing measurements or any experimental data, there is no applicability domain. Thus no reliability score can be assigned to the predictions.

**RESULTS AND DISCUSSION**

**Selecting a ppLFERs equation**

Predictions made with ppLFER Equations (3–5) and (9) using experimental descriptors showed similar performance when assessed against measured \( K_{OA} \) values (Table 4). This was expected, because Equations (3–5) were calibrated from similar training sets. Endo and Goss (2014) expanded the \( K_{OA} \) data set by Abraham and Acree (2008) with chemicals with reliable solute descriptors including some polyfluorinated and organosilicon compounds, and we subsequently used the Endo and Goss data set for calibrating Equation (5). The data set Jin et al. (2017) used to train their model is much larger than the others, because it includes \( K_{OA} \) values at temperatures other than 25 °C. However, the temperature-dependent \( K_{OA} \) values included by Jin et al. (2017) are similar to those used to parameterize Equation (6) by Mintz et al. (2008). Most of the chemicals included in the Jin et al. (2017) data set that are not included in the data sets used to develop Equations (3–6) are persistent organic pollutants such as PCBs and PCNs, as noted by Jin et al. (2017) and shown in the Supporting Information, Table SI 4.

Although predictions made at 25 °C by Equation (3) were slightly better than those made by the other three equations (Table 4), the difference was small. The residuals for each of the equations were also highly correlated, and Equations (4) and (5) performed better for fluorinated and organosilicon chemicals (see the Supporting Information). Because Equations (3) and (9) cannot predict a log \( K_{OA} \) for solutes without an E value, Equations (4) and (5) gave 10 more predictions at 25 °C and 32 more at all temperatures. Because all ppLFER system equations performed equally well and Equations (4) and (5) allow for predictions of more molecules, we only compared the results of Equation (5) with the other prediction techniques. Statistics for the residuals from the Abraham and Acree (2008), Endo and Goss (2014), and Jin et al. (2017) equations are included in the Supporting Information (Table SI 5).

**Comparing model performance in predicting \( K_{OA} \) at 25 °C**

We first compared the ability of different approaches to accurately predict \( K_{OA} \) at 25 °C. A violin plot shows the distribution of residuals for all prediction methods (Supporting Information, Figure SI 7). The EPISuite-25, OPERA, COSMOtherm, and

**TABLE 4: Performance of different polyparameter linear free energy relationships (ppLFERs) using experimental solute descriptors**

| Estimate                  | 25 °C | All temperatures |
|---------------------------|-------|-----------------|
|                           | No.   | MAE  | RMSE | PI width | No.   | MAE  | RMSE | PI width |
| Abraham & Acree, 2008     | 337   | 0.21 | 0.33 | 1.29     | 1363  | 0.22 | 0.32 | 1.23     |
| Endo & Goss, 2014         | 347   | 0.22 | 0.37 | 1.41     | 1395  | 0.20 | 0.32 | 1.25     |
| Modified (Equation 5)     | 347   | 0.23 | 0.37 | 1.43     | 1395  | 0.21 | 0.32 | 1.27     |
| Jin et al. 2017 (Equation 9)| 337   | 0.21 | 0.33 | 1.28     | 1363  | 0.22 | 0.33 | 1.28     |

MAE = mean absolute error; RMSE = root mean square error; PI = prediction interval.
TABLE 5: Statistics on the residuals, including the mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE), upper (P₈₅), and lower (P₀₅) prediction interval, when considering all log KOA estimates at 25 °C from all models

| Prediction tool | No. | Mean | MAE | Median | SD | RMSE | P₈₅ | P₀₅ |
|----------------|-----|------|-----|--------|----|------|------|------|
| ppLFER, experimental | 347 | -0.07 | 0.23 | -0.01 | 0.37 | 0.37 | 0.65 | -0.79 |
| ppLFER, estimated | 475 | -0.09 | 0.34 | -0.05 | 0.50 | 0.51 | 0.89 | -1.07 |
| EPISuite-25 | 475 | 0.00 | 0.58 | 0.06 | 0.78 | 0.78 | 1.54 | -1.54 |
| EPISuite-T | 474 | -0.01 | 0.54 | 0.01 | 0.74 | 0.74 | 1.43 | -1.45 |
| OPERA | 475 | 0.07 | 0.33 | 0.00 | 0.52 | 0.52 | 1.09 | -0.95 |
| COSMOtherm | 475 | 0.02 | 0.41 | -0.04 | 0.56 | 0.56 | 1.12 | -1.08 |

ppLFER = poyparameter linear free energy relationship.

PPFHER with estimated solute descriptors models can predict the KOA for all 475 chemicals with a measured value at 25 °C. For 128 chemicals (27%), the lack of empirical data prevented the KOA prediction with the ppLFER with experimental solute descriptors. The EPISuite-T model was not able to predict log KOA for one chemical, because the empirical HLC temperature equation reported for acetic acid (CAS # 64-19-7) to predict HLC at 25 °C appeared to be erroneous and did not match what was reported in the original work by Khan and Brimblecombe (1992) nor did the original equation produce a sensible result. Because the EPISuite-25 model already considers HLC values predicted with the bond method at 25 °C we did not substitute these HLC estimates in EPISuite-T.

Table 5 lists the numerical metrics of the comparison of predicted with measured KOA values. This comparison does not provide an entirely level playing field, because of the lower number of predictions for two of the techniques. There were 129 chemicals for which KOA could not be predicted by at least one of the six methods. We therefore also compared the predictive performance for the 346 measured KOA values at 25 °C, for which all six models were able to provide a prediction (Table 6).

By any of the metrics in Tables 5 and 6, the ppLFER with experimental solute descriptors performed the best in predicting the KOA at 25 °C, including if the comparison was restricted to the same set of chemicals (Table 6). The MAE and RMSE of the prediction were only approximately just over a fifth and a third of a log unit, respectively. When all possible predictions were considered (Table 5), the ppLFER equation with estimated solute descriptors slightly outperformed the COSMOtherm and OPERA models. If the 346 chemicals in Table 6 were considered, the OPERA model had the smallest RMSE of the three, although the differences were small. The ppLFER with experimental solute descriptors and COSMOtherm performed particularly well for volatile chemicals with a log KOA less than 6 (Figure 1).

The two EPISuite predictions had the largest deviations from the measured values, almost twice those of the best performing method. Invalidity of the assumption that wet and dry octanol have the same solvation properties (Abraham & Acree, 2008; Pimsawas et al., 1995) may contribute to the higher residuals. Wet octanol has a stronger hydrogen bonding acidity than dry octanol and is less capable of dissolving hydrophobic chemicals (Abraham & Acree, 2008). Thus KOA will typically be lower than KOA for nonpolar compounds and higher for more polar compounds (Abraham & Acree, 2008). However, when chemicals are grouped by their ability to undergo hydrogen bonding, as described by Baskaran et al. (2021), no correlation with the residual is observed (Supporting Information, Figure SI 9).

If we looked at the partition ratios used to calculate KOA in EPISuite (Supporting Information, Figure SI 10), we found that residuals were larger when the log KOW was greater than 5, and in these cases, the tendency for the EPISuite models to under- and overpredict KOA occurred more frequently when the log KOW was greater than 3 and less than –2, respectively. The KOAWIN model performed well for a small subset of chemicals with a log KOW between 2 and 5 and log KOW between –3 and 3. When the log KOW was less than 5, predictions for KOA were generally close to experimental values, except at very low KOW. This analysis suggests that when one is using estimated KOW and HLC values in a thermodynamic triangle, the log KOA is less reliable for more hydrophobic chemicals, as suggested by Abraham and Acree (2008). Only for 256 chemicals with a KOA measured at 25 °C did EPISuite contain experimental data for both log KOW and HLC. Although it was expected that the use of experimental values would improve

TABLE 6: Statistics on the residuals, including the mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE), upper (P₈₅), and lower (P₀₅) prediction interval, when considering only estimates for chemicals, for which all models could make predictions at 25 °C (n = 346)

| Prediction tool | Mean | MAE | Median | SD | RMSE | P₈₅ | P₀₅ |
|----------------|------|-----|--------|----|------|------|------|
| ppLFER, experimental | -0.07 | 0.23 | -0.01 | 0.37 | 0.37 | 0.65 | -0.79 |
| ppLFER, estimated | -0.05 | 0.31 | -0.01 | 0.45 | 0.45 | 0.83 | -0.94 |
| EPISuite-25 | 0.08 | 0.47 | 0.09 | 0.64 | 0.65 | 1.35 | -1.18 |
| EPISuite-T | 0.02 | 0.44 | 0.05 | 0.60 | 0.60 | 1.19 | -1.15 |
| OPERA | 0.00 | 0.26 | -0.01 | 0.44 | 0.44 | 0.86 | -0.86 |
| COSMOtherm | 0.00 | 0.34 | -0.04 | 0.46 | 0.46 | 0.90 | -0.90 |

ppLFER = polyparameter linear free energy relationship.
predictions of $K_{OA}$, it is difficult to see whether this was the case due to the sparsity of data (Supporting Information, Figure SI 11). The EPISuite-T model performed slightly better than EPISuite-25 because it incorporates some experimental values for HLC, whereas EPISuite-25 relies only on estimated $K_{OW}$ and HLC values. Using both estimated $K_{OW}$ and HLC, EPISuite was reported to have a standard deviation of 0.688 and MAE of 0.479, ($n = 310$; USEPA, 2012), which is very similar to our assessment (Table 5).

The good performance of the ppLFER models and OPERA may in part be attributable to an overlap between the data set that was used in their calibration and the data set of measured values we used for model evaluation. The training sets for the ppLFER and OPERA models consist of 181 (see the Supporting Information, Table SI 4) and 270 chemicals, respectively. Because our data set aimed to be comprehensive and include all reliable measured $K_{OA}$ values that have been reported in the literature (Baskaran et al., 2021), it is likely that almost all the log $K_{OA}$ data used to calibrate these models were also contained in the evaluation data set. Given the limited number of $K_{OA}$ data, particularly at 25°C, the overlap of the training chemicals with the chemicals used in the present assessment will reduce the error calculated with these prediction techniques.

Lampic and Parnis (2020) compared the prediction performance of ppLFERs with estimated solute descriptors, OPERA,
COSMOtherm, and EPISuite for various physical–chemical properties of per- and polyfluoroalkyl compounds at 25 °C. The OPERA model had the smallest reported RMSE and MAE for K_{OA} (Lampic & Parnis, 2020). The experimental data set included multiple measurements for the same compound, which can bias the statistical calculations. The data set also includes K_{OA} values for fluoroether alcohols, perfluorooctane sulfonamide ethanol (FOSE), and fluoroocane sulfonamide measured with the gas chromatography-retention time (GC-RT) technique (Lei et al., 2004). We have excluded those data from our data set because this technique is not suited for polar compounds (Baskaran et al., 2021).

**Comparing model performance in predicting K_{OA} at any temperature**

We next compared the prediction performance of the four models that can predict the K_{OA} at temperatures other than 25 °C. Because only 28% of measured K_{OA} values are for 25 °C, this data set is considerably larger; it comprises 1676 data points for 604 chemicals at temperatures ranging from −10 to 110 °C. The COSMOtherm and the ppLFER using estimated solute descriptors methods were able to predict K_{OA} corresponding to all 1676 literature values (Table 7). The ppLFERs equations using experimental solute descriptors were limited by the availability of the solute descriptors for 281 chemicals. The EPISuite-T model was able to predict K_{OA} values for all but one compound, acetic acid, as mentioned in the previous section, Comparing model performance in predicting K_{OA} at 25 °C. In total, 1394 measurements that could be compared against four models (Table 8).

All models predicted the log K_{OA} with a MAE less than 0.6 and SDs less than or equal to 0.8. The residuals consistently fell within 3 log units of the measured value (Figure 2 and Supporting Information, Figure SI 12), with some exceptions: the log K_{OA} predictions at 5 °C by COSMOtherm for N-ethyl FOSE and N-methyl FOSE, the prediction at 45 °C for 2,2′,3,4′,4′,5′-hexabromodiphenyl ether (BDE 183) and at 5 and 10 °C for endosulfan I by EPISuite-T, and the 25 °C prediction for 2′-methyl-2,4,4′-tribromodiphenyl ether (2′-MeO BDE 28) by the ppLFER equation using estimated solute descriptors. The ppLFERs using experimental solute descriptors had the lowest RMSE (0.32) and MAE (0.21) values. The use of ppLFERs with estimated solute descriptors performed marginally better than COSMOtherm.

The ΔH_{OA} is itself temperature dependent, and Equation (6) is meant to calculate ΔH_{OA} within the much narrower range of 10–45 °C (Mintz et al., 2008); so we explored whether the error of the prediction is dependent on temperature (Supporting Information, Figure SI 16). Although we saw no temperature dependence on the residuals, we evaluated the ppLFER equations within the applicability domain of the ΔH_{OA} equation (Supporting Information, Figure SI 19). There was little change in the model performance (Supporting Information, Table SI 7). In fact, the RMSE increased slightly to 0.33 and 0.44 for ppLFERs using experimental and estimated solute descriptors, respectively. In addition, the ΔH_{OA} and log K_{OA} at 25 °C predicted by the ppLFERs were highly correlated (R² = 0.98), which suggests that ΔH_{OA} could be estimated from log K_{OA} (Supporting Information, Figure SI 5). Further research is needed to understand the relationship between these properties using empirical data.

Because the equations in EPISuite are intended to predict HLC within the range from 0 to 50 °C, these considerations could equally apply to explain the relatively poor performance of EPISuite-T. However, the MAEs for EPISuite-T are not higher at extreme temperatures. The EPISuite information notes that a log K_{OA} at 10 °C estimated from an HLC at 10 °C and a K_{OW} at 25 °C can be expected to have an SD of approximately 0.575 and a MAE of 0.433, based on a sample size of 126 compounds (USEPA, 2012). In the present study we estimated higher SD and MAE for EPISuite-T predictions at any temperature of 0.80 and 0.59, respectively, based on a sample size of 1675. Limiting estimates to the temperature applicability domain of the model (0–50 °C) had little effect on the SD (0.81) and MAE (0.60; Supporting Information, Table SI 8).

---

**Table 7**: Statistics on the residuals of log K_{OA} predictions at temperatures between −10 and 110 °C, including the mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE), upper (Pl_u), and lower prediction interval (Pl_l).

| Prediction tool          | No.  | Mean | MAE | Median | SD | RMSE | Pl_u | Pl_l  |
|-------------------------|------|------|-----|--------|----|------|------|-------|
| ppLFER, experimental    | 1395 | 0.01 | 0.21 | 0.03   | 0.32| 0.32 | 0.64 | −0.63 |
| ppLFER, estimated       | 1676 | −0.02| 0.29 | 0.01   | 0.43| 0.43 | 0.82 | −0.86 |
| EPISuite-T              | 1675 | 0.13 | 0.59 | 0.10   | 0.80| 0.81 | 1.69 | −1.44 |
| COSMOtherm              | 1676 | 0.04 | 0.40 | 0.04   | 0.55| 0.56 | 1.12 | −1.05 |

ppLFER = polypar meter linear free energy relationship.

**Table 8**: Statistics on the residuals of log K_{OA} predictions at temperatures −10–110 °C that could be made with all models (n = 1394), including the mean absolute error (MAE), standard deviation (SD), root mean square error (RMSE), upper (Pl_u), and lower prediction interval (Pl_l).

| Prediction tool          | Mean | MAE | Median | SD  | RMSE | Pl_u | Pl_l  |
|-------------------------|------|-----|--------|-----|------|------|-------|
| ppLFER, experimental    | 0.01 | 0.21| 0.03   | 0.32| 0.32 | 0.64 | −0.63 |
| ppLFER, estimated       | 0.01 | 0.26| 0.02   | 0.38| 0.38 | 0.75 | −0.73 |
| EPISuite-T              | 0.11 | 0.51| 0.10   | 0.69| 0.69 | 1.46 | −1.23 |
| COSMOtherm              | 0.08 | 0.34| 0.05   | 0.45| 0.46 | 0.97 | −0.81 |

ppLFER = polypar meter linear free energy relationship.

© 2021 The Authors wileyonlinelibrary.com/ETC
Because COSMOtherm requires no internal calibration, it is particularly impressive that it predicted the log $K_{OA}$ at different temperatures so well. The COSMOtherm model was previously shown to systematically underpredict (with high residuals) the log $K_{OA}$ for substituted PAHs (Parnis et al., 2015). In addition, we found that COSMOtherm also systematically overpredicted the log $K_{OA}$ for the polar fluorinated compounds and PBDEs (Supporting Information, Figures SI 14 and SI 22). The ppLFER equations also tended to underpredict the log $K_{OA}$ for PBDEs.

It is also important to acknowledge that in some instances large residuals may be due to flawed measurements. Supporting Information, Figure SI 23, compiles measured data that cause absolute residuals greater than 0.75, when predicted with ppLFERs and COSMOtherm. The models all over- or underpredicted the $K_{OA}$ for the polar fluorinated compounds and PBDEs (Supporting Information, Figures SI 14 and SI 22). The ppLFER equations also tended to underpredict the log $K_{OA}$ for PBDEs.

The $K_{OA}$ values for N-nitrosodipropylamine and N-nitrosodibutylamine were measured using a static technique relating the volatility of analytes from fish tissue to octanol (Hiatt, 1997). Although many of the values reported in the present study did not stand out as erroneous, the SDs of some measurements were high (e.g., N-nitrosodipropylamine 20,000 ± 5100; N-nitrosodibutylamine 16,000 ± 8500).

The log $K_{OA}$ of 11.96 for BDE 183 measured using a generator column technique (Harner & Shoeib, 2002) may also be erroneous. This value is close to the limits of this technique, and it is possible that BDE 183 never reached equilibrium in the generator column. This would also explain why the reported $\Delta U_{OA}^n$ (referred to as $\Delta H_{OA}^n$) for BDE 183 is 10 kJ/mol lower than the $\Delta U_{OA}^n$ for BDE 153 (2,2′,4,4′,5,5′-hexabromodiphenyl ether) and BDE 156 (2,3,3′,4,4′,5-hexabromodiphenyl ether), even though these two PBDEs have measured log $K_{OA}$ values of 11.82 and 11.97 at 25 °C, respectively (Harner & Shoeib, 2002). A log $K_{OA}$ for BDE 183 closer to 13, as predicted by the ppLFER models and COSMOtherm, would also be more consistent with the other physical-chemical properties reported for this congener (Wania & Dugani, 2003).
Both β-HCH and δ-HCH had a measured log $K_{OA}$ of almost 9 at 25 °C using the generator column technique, whereas log $K_{OA}$ values for α-HCH and γ-HCH measured using the same technique were in the region of 7.5 and 8 (Shoeib & Harner, 2002b). As stereoisomers, these chemicals are unlikely to have log $K_{OA}$ values differing by more than 1 log unit, and the true log $K_{OA}$ for β-HCH and δ-HCH is likely closer to the value estimated by the ppLFER equations and COSMOtherm.

**Predictive performance and applicability domains**

Reliability scores for predictions were assigned as described in the Materials and Methods section. Most models consider chemicals with measured $K_{OA}$ values within their applicability domain as having a reliability score of excellent.

The reliability scores for EPISuite-25 predictions appeared to give a reasonable indication of the error associated with a prediction, namely, predictions falling outside of the 95% prediction interval (Figure 1) are scored as either good or fair. The 423 predictions made by EPISuite-25 that are scored excellent have an RMSE of 0.66 (Supporting Information, Table SI 9). The RMSE values for good and fair predictions were 2.24 ($n = 6$) and 1.29 ($n = 46$), respectively. The reliability scores of EPISuite-T predictions at 25 °C were generally lower than for the other models, with 46 judged good (RMSE = 1.22), and 268 fair (RMSE = 0.68). The lower reliability scores of the EPISuite-T predictions also occurred for predictions made at different temperatures (Supporting Information, Table SI 10). The RMSE of poor predictions (2.25 at 25 °C and 1.66 at all temperatures) was higher than the excellent predictions (0.49 at 25 °C, 0.55 at all temperatures). However, as with EPISuite-25, the RMSE of the good EPISuite-T predictions were in both cases higher than for the fair predictions. The good and fair categorizations used to describe the applicability domain of the EPISuite model appeared to be unreliable indicators of the uncertainty of the prediction, which may reflect the impact of a few outliers on the RMSE for a relatively small set of chemicals.

When ppLFER equations with experimental solute descriptors were used, the overall error of the prediction was always estimated to be less than 0.2, which meant all predictions were considered excellent. This means that considering only the uncertainty of the system constants and ignoring the error of the solute descriptors does not provide a metric suitable for judging the reliability of the prediction.

Most of the chemicals with a measured $K_{OA}$ at 25 °C fell within the applicability domain of the ppLFER model with estimated solute descriptors, with 362 chemicals falling into the excellent (RMSE = 0.40) and 96 into the good category (RMSE = 0.94). The reliability scores, determined using Monte Carlo analysis with the standard error of system constants and solute descriptors, gave a good indication of the error of the prediction, particularly for values at 25 °C. If we compare the reliability scores with the RMSE of the predictions at all temperatures (Supporting Information, Table SI 10), predictions with scores of excellent and poor had the smallest (0.33) and largest (0.81) RMSEs, respectively. The RMSEs for the good (0.65) and fair (0.41) categories suggest that these intermediate scores are less reliable indicators of prediction quality, possibly again because of the small sample size. The ppLFER equations generally had the fewest number of predictions that scored excellent and had an absolute residual greater than 1.
The OPERA model also judged most chemicals to have a good fit with its applicability domain, with predictions assigned categories of excellent and good for 417 and 58 chemicals, respectively. The RMSE values of the residuals for the good predictions were higher (0.64) than those for the excellent predictions (0.50), as would be expected. The OPERA model had 29 predictions that were within the global applicability domain, and had high local applicability domain and confidence levels, which had an absolute residual greater than 1. After EPISuite-25 (n = 53), this was the highest number of absolute residuals larger than 1 seen for excellent predictions at 25 °C.

There is no means of assessing whether a prediction made by COSMOtherm falls within the applicability domain of the model.

**Estimating the uncertainty of future predictions**

Our analysis makes it possible to estimate the possible error of future predictions with the investigated methods. If a model is used to make a new prediction, there is a 95% chance that the residual of a new prediction for a chemical within the applicability domain of the model is within the range of the prediction interval. A smaller prediction interval gives high confidence in future predictions. Figure 3 shows the prediction intervals for each technique at 25 °C and for the four models capable of predicting log $K_{OA}$ at other temperatures. Because the prediction interval is calculated using the SD of the residuals, the width of the prediction interval is directly correlated with the reported SDs. The prediction interval for each model is listed in Tables 5–8. For future predictions, we recommend using the prediction intervals as reported in Table 9.

**Accessibility and usability of models**

Another factor to consider when comparing the different prediction techniques is their usability and accessibility.

Although COSMOtherm performed very well, it is a licensed software that can be expensive to purchase. The COSMOconf calculations are very demanding in central processing unit (CPU) time. Using a supercomputer can reduce the time for calculation, but this requires both access to, and set-up, of the COSMOconf calculations on a supercomputer. On the other hand, once COSMO files for the different congeners of a chemical have been generated, any number of partition ratios at any temperature can be obtained with COSMOtherm with very limited additional CPU demand. On balance, the cost and time necessary to use COSMOtherm reduce its accessibility and usefulness as a routine prediction tool.

Experimental solute descriptors for ppLFER equations can be obtained from the UFZ website (Ulrich et al., 2017). Although Equations (3) and (6) are integrated into the website and it is possible to directly export the log $K_{OA}$ of a chemical at 25 °C and the $\Delta H^\text{OA}$, using Equations (4), (5), or (9) or the prediction of $K_{OA}$ values at temperatures other than 25°C require a simple spreadsheet for calculating the partition ratio and adjusting it for temperature. Equation (5) uses the fewest number of solute descriptors while still performing as well as Equations (3) or (9). This means that there is a higher likelihood that a complete set of experimental solute descriptors would be available for calculating log $K_{OA}$. If no experimental solute descriptors are available, the estimated solute descriptors can be calculated using the IFSQSARs implemented in the UFZ website (Ulrich et al., 2017), or by downloading the IFSQSAR prediction software by Brown (2020). Using the UFZ website for estimated solute descriptors will be sufficient for most people wanting to predict a log $K_{OA}$. The use of the stand-alone IFSQSAR model is only necessary to obtain the SE of the estimated solute descriptors.

The OPERA model is another freely available software that includes a graphic user interface that makes it easy to use (Mansouri, 2018). However, it is available only on Windows and Linux operating systems. The OPERA model has also been integrated into the CompTox Dashboard, which provides the prediction and details regarding the applicability domain and reliability of the prediction for a single chemical. When predicting the log $K_{OA}$ for multiple chemicals, obtaining these details from the CompTox Dashboard is not easy, and using the downloadable software is recommended.

The EPISuite models (EPISuite-25 and EPISuite-T) are also freely available from the US EPA website for computers running on a Windows operating system (USEPA, 2012). Of course, using the thermodynamic triangle approach does not require the use of the software, but the KOAWIN model can incorporate the use of experimental $K_{OW}$ and $HLC$ values when available. The models KOAWIN, HENRYWIN, and KOWWIN can all complete batch mode calculations, which is useful for large datasets. However, extracting the temperature equations for multiple chemicals from HENRYWIN for EPISuite-T can be time consuming, because the format of the equation can differ between chemicals.
CONCLUSIONS

If only the $K_{OA}$ at 25°C is required, the ppLFER equation using experimental solute descriptors has the best performance, although the ppLFER equation with estimated solute descriptors, OPERA, and COSMOtherm also give reasonably good estimates. Both EPISuite-2S and EPISuite-T consistently had the worst performance of all assessed prediction techniques. The ppLFER using estimated solute descriptors and COSMOtherm were able to reliably predict the $K_{OA}$ for the greatest number of chemicals and temperatures. Solute descriptors can be estimated for almost all neutral chemicals, and the COSMO-RS approach for estimating $K_{OA}$ can be applied to virtually any chemical. Both EPISuite-T and ppLFER equations using experimental solute descriptors were also able to predict the $K_{OA}$ for chemicals at various temperatures, but these are limited by the availability of HLC temperature correction equations and experimental solute descriptors, respectively.

For most models, the reliability of each of the prediction techniques as assessed based on the fit with the applicability domain only correlated somewhat with the residual of the actual prediction. In terms of usability, the ppLFER equations and the OPERA model are the easiest to use. The COSMOtherm model is more expensive and more complex, has higher costs and labor associated with its use.

In summary, a ppLFER using experimental solute descriptors is the best predictor of $K_{OA}$ regardless of temperature. If experimental descriptors are not available, a ppLFER with estimated solute descriptors or the COSMOtherm model are also well suited to predicting the $K_{OA}$. The OPERA model also works well for predicting $K_{OA}$ at 25°C.

Supporting Information—The Supporting Information is available on the Wiley Online Library at https://doi.org/10.1002/etc.5201.

Acknowledgment—We are grateful to A. Sangion for help with statistical analyses and to the European Chemical Industry Council for funding (project ECO-41 of the Long-range Research Initiative).

Data Availability Statement—Data, associated metadata, and calculation tools are available from the corresponding author (frank.wania@utoronto.ca).

REFERENCES

Abraham, M. H. (1993). Scales of solute hydrogen-bonding: Their construction and application to physicochemical and biochemical processes. Chemical Society Review, 22(2), 73–83. https://doi.org/10.1039/CS993220073

Abraham, M. H., & Acree, W. E. (2008). Comparison of solubility of gases and vapours in wet and dry alcohols, especially octan-1-ol. Journal of Physical and Organic Chemistry, 27(10), 823–832. https://doi.org/10.1002/poc.1374

Abraham, M. H., Gola, J. M. R., Cometto-Muñiz, J. E., & Cain, W. S. (2001). Solvation properties of refrigerants, and the estimation of their water-solvent and gas-solvent partitions. Fluid Phase Equilibria, 180(1), 41–58. https://doi.org/10.1016/S0378-3812(00)00511-2

Abraham, M. H., Smith, R. E., Luchtefeld, R., Boorern, A. J., Luo, R., & Acree, W. E. (2010). Prediction of solubility of drugs and other compounds in organic solvents. Journal of Pharmacologic Science, 99(3), 1500–1515. https://doi.org/10.1002/jphs.21922

Abraham, M. H., Whiting, G. S., Doherty, R. M., & Shuely, W. J. (1990). Hydrogen bonding. Part 13. A new method for the characterisation of G.C. stationary phases—The Laffort data set. Journal of the Chemical Society, Perkin Transactions, 2(8), 1451–1460. https://doi.org/10.1039/P29900001451

Advanced Chemistry Development. (2021). Calculate Abraham solvation parameters with ACD/Ab solvs. https://www.acdlabs.com/products/percepta/predictors/absol/index.php

Anderson, K. A., Points, G. L., Donald, C. E., Dixon, H. M., Scott, R. P., Wilson, G., Tidwell, L. G., Hoffman, P. D., Herbstman, J. B., & O’Connell, S. C. (2017). Preparation and performance features of wristband samplers and considerations for chemical exposure assessment. Journal of Exposure Science and Environmental Epidemiology, 27(6), 551–559. https://doi.org/10.1038/jes.2017.9

Atkinson, D., & Curthoys, G. (1978). The determination of heats of adsorption by gas-solid chromatography. Journal of Chemical Education, 55(9), 564. https://doi.org/10.1021/ed055p564

Baskaran, S., Lei, Y. D., & Wania, F. (2021). A database of experimentally derived and estimated octanol-air partition ratios ($K_{OA}$). Journal of Physical and Chemical Reference Data. https://doi.org/10.1063/5.0059652

Batterman, S., Zhang, L., Wang, S., & Franzblau, A. (2002). Partition coefficients for the trihalomethanes among blood, urine, water, milk and air. Science of the Total Environment, 284(1–3), 237–247. https://doi.org/10.1016/S0048-9697(01)00890-7

Bennett, D. H., & Furtaw, E. J. (2004). Fugacity-based indoor residential pesticide fate model. Environmental Science & Technology, 38(7), 2142–2152. https://doi.org/10.1021/es034287m

Bronner, G., Fenner, K., & Goss, K.-U. (2010). Hexadecane/air partition coefficients of multifunctional compounds: Experimental data and modeling. Fluid Phase Equilibria, 299(2), 207–215. https://doi.org/10.1016/j.fluid.2010.09.043

Brown, T. N. (2020). tnbrowncontam/ifsqsr. https://github.com/tnbrowncontam/ifsqsr

Brown, T. N. (2014). Predicting hexadecane-air equilibrium partition coefficients (L) using a group contribution approach constructed from high quality data. SAR QSAR Environmental Research, 25(1), 51–71. https://doi.org/10.1080/10698966.2013.841286

Brown, T. N., Amot, J. A., & Wania, F. (2012). Iterative fragment selection: A group contribution approach to predicting fish biotransformation half-lives. Environmental Science & Technology, 46(15), 8253–8260. https://doi.org/10.1021/es301182a

Chen, J. W., Harner, T., Schramm, K.-W., Quan, X., Xue, X. Y., Wu, W. Z., & Kettrup, A. (2002). Quantitative relationships between molecular structures, environmental temperatures and octanol–air partition coefficients of PCBDDFs. Science of the Total Environment, 300(1), 155–166. https://doi.org/10.1016/S0048-9697(01)11488-2

Chen, J. W., Harner, T., Schramm, K.-W., Quan, X., Xue, X. Y., & Kettrup, A. (2003a). Quantitative relationships between molecular structures, environmental temperatures and octanol–air partition coefficients of polychlorinated biphenyls. Computational Biology and Chemistry, 27(3), 405–421. https://doi.org/10.1016/S1476-9271(02)00089-0

Chen, J., Xue, X., Schramm, K.-W., Quan, X., Yang, F., & Kettrup, A. (2003b). Quantitative structure-property relationships for octanol–air partition coefficients of polychlorinated naphthalenes, chlorobenzenes and p,p′-DDT. Computational Biology and Chemistry, 27(3), 165–171. https://doi.org/10.1016/S0097-8485(02)00017-7

Chen, J. W., Harner, T., Yang, P., Quan, X., Chen, S., Schramm, K.-W., & Kettrup, A. (2003c). Quantitative predictive models for octanol-air partition coefficients of polychlorinated diphenyl ethers at different temperatures. Chemosphere, 57(7), 577–584. https://doi.org/10.1016/S0045-6535(03)00152-6

Chen, Y., Cai, X., Jiang, L., & Li, Y. (2016). Prediction of octanol-air partition coefficients for polychlorinated biphenyls (PCBs) using 3D-QSAR models. Ecotoxicology and Environmental Safety, 124, 202–212. https://doi.org/10.1016/j.ecoenv.2015.10.024

Cousins, I. T., Hartlieb, N., Teichmann, C., & Jones, K. C. (1997). Measured and predicted volatilisation fluxes of PCBs from contaminated sludge-amended soils. Environmental Pollution, 97(3), 229–238. https://doi.org/10.1016/S0048-9213(97)00096-1
Jin, X., Fu, Z., Li, X., & Chen, J. (2017). Development of polyparameter linear models for predicting the vapor pressures, octanol–air, and water–air partition coefficients for polyfluorinated sulfonamide, sulfonamidoethanols, and telomer alcohols. Journal of Chemical Engineering Data, 49(4), 1013–1022. https://doi.org/10.1016/j.jced.2016.09.003

Li, W., Ding, G., Gao, H., Zhuang, Y., Gu, X., & Peijnenburg, W. J. G. M. (2020). Prediction of octanol–air partition coefficients for PCBs at different ambient temperatures based on the solvation free energy and the dimer ratio. Chemosphere, 242, 125246. https://doi.org/10.1016/j.chemosphere.2019.125246

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development. Advances in Drug Delivery Review, 23(1–3), 3–25. https://doi.org/10.1016/S0169-409X(96)00423-1

Liu, H., Shi, J., Liu, H., & Wang, Z. (2013). Improved 3D-QSAR analysis of the predictive octanol–air partition coefficients of hydroxylated and methoxylated polybrominated diphenyl ethers. Atmospheric Environment, 77, 840–845. https://doi.org/10.1016/j.atmosenv.2013.05.068

Mackay, D., Celse, A. K. D., & Parmis, J. M. (2015). The evolution and future of environmental partitioning models. Environment Review, 24(1), 101–113. https://doi.org/10.1139/er-2015-0059

Mansouri, K. (2018). kmsansouri/OPERA. https://github.com/kmsansouri/OPERA

Mansouri, K., Grukle, C. M., Judson, R. S., & Williams, A. J. (2018). OPERA models for predicting physicochemical properties and environmental fate endpoints. Journal of Cheminformatics, 10(1). https://doi.org/10.1186/s13321-018-0263-1

Mansouri, K., & Williams, A. (2017). KOA model for the octanol/air partition coefficient prediction from OPERA models. QMRF Technical Report. https://doi.org/10.13140/RG.2.2.14409.54883/1

Meylan, W. M., & Howard, P. H. (1991). Bond contribution method for estimating Henry's law constants. Environmental Toxicology and Chemistry, 10(10), 1283–1293. https://doi.org/10.1002/etc.5620101007

Meylan, W. M., & Howard, P. H. (2005). Estimating octanol-air partition coefficients with octanol-water partition coefficient and Henry's law constants. Chemosphere, 61(5), 640–644. https://doi.org/10.1016/j.chemosphere.2005.03.029

Mintz, C., Burton, K., Ladlie, T., Clark, M., Acree, W. E., & Abraham, M. H. (2008). Enthalpy of solvation corrections for gaseous solutes dissolved in dibutyl ether and ethyl acetate. Thermochimica Acta, 470(1), 67–76. https://doi.org/10.1016/j.tca.2008.02.001

Mintz, C., Clark, M., Acree, W. E., & Abraham, M. H. (2007). Enthalpy of solvation correlations of gaseous solutes dissolved in water and in 1-octanol based on the abraham model. Journal of Chemical Informatics and Modeling, 47(1), 115–121. https://doi.org/10.1021/ci600402n

O’Boyle, N. M., Banck, M., James, C. A., Morley, C., Vandermeersch, T., & Hutchison, G. R. (2011). Open Babel: An open chemical toolbox. Journal of Cheminformatics, 3(1), 33. https://doi.org/10.1186/1758-2946-3-33

Odabasi, M., Cetin, E., & Sofuoglu, A. (2006). Determination of octanol–air partition coefficients and supercooled liquid vapor pressures of PAHs as a function of temperature: Application to gas–particle partitioning in an urban atmosphere. Atmospheric Environment, 40(34), 6615–6625. https://doi.org/10.1016/j.atmosenv.2006.05.051

Okeme, J. O., Rodgers, T. F. M., Parnis, J. M., Diamond, M. L., Bidlem, T. F., & Jantunen, L. M. (2020). Gas chromatographic estimation of vapor pressures and octanol–air partition coefficients of semivolatile organic compounds of emerging concern. Journal of Chemical Engineering Data, 65(5), 2467–2475. https://doi.org/10.1016/j.jced.2019.01.1126

Parnis, J. M., Mackay, D., & Hamer, T. (2015). Temperature dependence of Henry's law constants and KoA for simple and heteroatom-substituted PAHs by COSMO-RS. Atmospheric Environment, 110, 27–35. https://doi.org/10.1016/j.atmosenv.2015.03.032

Paterson, S., Mackay, D., Tam, D., & Shi, W. Y. (1990). Uptake of organic chemicals by plants: A review of processes, correlations and models. Chemosphere, 21(3), 297–331. https://doi.org/10.1016/0045-6535(90)90002-B

Pinsuwan, S., Li, A., & Yalkowsky, S. H. (1995). Correlation of octanol/water solubility ratios and partition coefficients. Journal of Chemical Engineering Data, 40(3), 623–626. https://doi.org/10.1021/je00019a019
Saini, A., Rauert, C., Simpson, M. J., Harrad, S., & Diamond, M. L. (2016). Characterizing the sorption of polybrominated diphenyl ethers (PBDEs) to cotton and polyester fabrics under controlled conditions. *Science of the Total Environment*, 563-564, 99–107. https://doi.org/10.1016/j.scitotenv.2016.04.099

Schwarzenbach, R. P., Gschwend, P. M., & Imboden, D. M. (2005). Partitioning: Molecular interactions and thermodynamics. In *Environmental Organic Chemistry* (pp. 57–96). John Wiley & Sons. https://doi.org/10.1002/0471649643.ch3

Shoeib, M., & Hamer, T. (2002a). Characterization and comparison of three passive air samplers for persistent organic pollutants. *Environmental Science & Technology*, 36(19), 4142–4151. https://doi.org/10.1021/es020635t

Shoeib, M., & Hamer, T. (2002b). Using measured octanol-air partition coefficients to explain environmental partitioning of organochlorine pesticides. *Environmental Toxicology and Chemistry*, 21, 984–990. https://doi.org/10.1002/etc.5620210513

Sprunger, L., Proctor, A., Acree, W. E., & Abraham, M. H. (2007). Characterization of the sorption of gaseous and organic solutes onto polydimethyl siloxane solid-phase microextraction surfaces using the Abraham model. *Journal of Chromatography A*, 1175(2), 162–173. https://doi.org/10.1016/j.chroma.2007.10.058

Stenzel, A., Endo, S., & Goss, K.-U. (2012). Measurements and predictions of hexadecane/air partition coefficients for 387 environmentally relevant compounds. *Journal of Chromatography A*, 1220, 132–142. https://doi.org/10.1016/j.chroma.2011.11.053

Stenzel, A., Goss, K.-U., & Endo, S. (2014). Prediction of partition coefficients for complex environmental contaminants: Validation of COSMOTHERM, ABSOLV, and SPARC. *Environmental Toxicology and Chemistry*, 33(7), 1537–1543. https://doi.org/10.1002/etc.2587

Su, Y., Lei, Y. D., Daly, G. L., & Wania, F. (2002). Determination of octanol–air partition coefficient (KOA) values for chlorobenzenes and polychlorinated naphthalenes from gas chromatographic retention times. *Journal of Chemical Engineering Data*, 47(3), 449–455. https://doi.org/10.1021/je015512n

Ulrich, N., Endo, S., Brown, T., Bronner, G., Abraham, M. H., & Goss, K.-U. (2017). UFZ-LSER database v 3.2. http://www.ufz.de/lserd

US Environmental Protection Agency. (2012). Estimation Programs Interface Suite™. https://www.epa.gov/tsca-screening-tools/download-epi-suitem-estimation-program-interface-v411

Wania, F., & Dugani, C. B. (2003). Assessing the long-range transport potential of polybrominated diphenyl ethers: A comparison of four multimedia models. *Environmental Toxicology and Chemistry*, 22(6), 1252–1261. https://doi.org/10.1002/etc.5620220610

Wania, F., Lei, Y. D., & Hamer, T. (2002). Estimating octanol–air partition coefficients of nonpolar semivolatile organic compounds from gas chromatographic retention times. *Analytical Chemistry*, 74(14), 3476–3483. https://doi.org/10.1021/ac0256033

Xu, H. Y., Zou, J. W., Yu, Q. S., Wang, Y. H., Zhang, J. Y., & Jin, H. X. (2007). QSAR/QSAR models for prediction of the physicochemical properties and biological activity of polybrominated diphenyl ethers. *Chemosphere*, 66(10), 1998–2010. https://doi.org/10.1016/j.chemosphere.2006.07.072

Xu, S., & Kropscott, B. (2012). Method for simultaneous determination of partition coefficients for cyclic volatile methylsiloxanes and dimethylsilanediol. *Analytical Chemistry*, 84(4), 1948–1955. https://doi.org/10.1021/ac202953t

Xu, S., & Kropscott, B. (2013). Octanol/air partition coefficients of volatile methylsiloxanes and their temperature dependence. *Journal of Chemical Engineering Data*, 58(1), 136–142. https://doi.org/10.1021/je301005b

Xu, S., & Kropscott, B. (2014). Evaluation of the three-phase equilibrium method for measuring temperature dependence of internally consistent partition coefficients (KOW, KOA, and KAW) for volatile methylsiloxanes and dimethylsilanol. *Environmental Toxicology and Chemistry*, 12, 2702–2710. https://doi.org/10.1002/etc.2754

Yuan, J., Yu, S., Zhang, T., Yuan, X., Cao, Y., Yu, X., Yang, X., & Yao, W. (2016). QSAR models for predicting generator-column-derived octanol-water and octanol/air partition coefficients of polychlorinated biphenyls. *Ecotoxicology and Environmental Safety*, 128, 171–180. https://doi.org/10.1016/j.ecoenv.2016.02.022

Zeng, X.-L., Zhang, X.-L., & Wang, Y. (2013). QSAR modeling of n-octanol/air partition coefficients and liquid vapor pressures of polychlorinated dibenzo-p-dioxins. *Chemosphere*, 91(2), 229–232. https://doi.org/10.1016/j.chemosphere.2012.12.060

Zhang, X., Brown, T. N., Wania, F., Heimstad, E. S., & Goss, K.-U. (2010). Assessment of chemical screening outcomes based on different partitioning property estimation methods. *Environment International*, 36(6), 514–520.https://doi.org/10.1016/j.envint.2010.03.010

Zhang, X., Schramm, K.-W., Henkelmann, B., Klimm, C., Kaune, A., Kettrup, A., & Lu, P. (1999). A method to estimate the octanol–air partition coefficient of semivolatile organic compounds. *Analytical Chemistry*, 71(17), 3834–3838. https://doi.org/10.1021/ac981103r