Linear solvation energy relationship development for adsorption of synthetic organic compounds by carbon nanomaterials: an overview of the last decade

Journal:	*Environmental Science: Water Research & Technology*
Manuscript ID	EW-FRO-07-2020-000644.R1
Article Type:	Frontier

Water Significance Statement

for

Linear solvation energy relationship development for adsorption of synthetic organic compounds by carbon nanomaterials: an overview of the last decade

Onur G. Apul, Francois Perreault, Gamze Ersan, Tanju Karanfil

Owing to the discovery and upsurge of engineered carbon nanomaterials (CNMs) in 2000s and 2010s, researchers witnessed the flooding of scientific literature by articles demonstrating the adsorption of synthetic organic compounds (SOCs) by CNMs. Because of this rapid expansion, predictive modelling for adsorption of SOCs by CNMs was made possible in early 2010s. Due to the application potential of engineered CNMs in water treatment and necessity to understand their interactions with SOCs and potential health and environmental implications after their release to the natural aquatic systems, these models are essential. The broad predictive capabilities of linear solvation energy relations models also offer a powerful framework to better understand the interaction of organic molecules with surfaces.

Linear solvation energy relationship development for adsorption of synthetic organic compounds by carbon nanomaterials: An overview of the last decade

Onur G. Apul[1,4*], Francois Perreault[2], Gamze Ersan[3], Tanju Karanfil[3]

[1]Civil and Environmental Engineering, University of Maine, Orono ME 04469.
[2]School of Sustainable Engineering and the Built Environment, Arizona State University, Tempe AZ 85281.
[3]Environmental Engineering and Earth Sciences, Clemson University, Clemson SC 29625.
[4]Civil and Environmental Engineering, University of Massachusetts Lowell, Lowell MA 01854.

Prepared for Submission
to
ES: WRT (2020)

*Corresponding author

Abstract

The production of carbon nanomaterials has been increasing due to their outstanding properties and innovation potential in various industrial applications. Hence, their discharge into the aquatic environment from point and non-point sources is inevitable. Therefore, the natural and built aquatic environments that contain thousands of synthetic organic contaminants, would facilitate adsorptive interaction between synthetic organic molecules and carbon nanomaterial surfaces. Today's literature rapidly developed to contain a great number of adsorption data for synthetic organic compounds onto carbon nanomaterials. The availability of adsorption data enables generation of structure-activity relationships by utilizing adsorbate and adsorbent characteristics in simple/controlled aquatic solutions. These models not only create practical predictive tools but also provide mechanistic insights about intermolecular interactions between organic molecules and carbon nanomaterials. This review analyzes 86 linear solvation energy relationship models published in the last decade. Our work summarizes and compares key independent variables associated with adsorbate properties (e.g., molecular size, H-bonding ability, polarizability) and identifies critical factors for selecting dependent variables (e.g., adsorbent characteristics, saturation concentration). The study also discusses intermolecular interactions unraveled by the modeling efforts at different adsorbate concentrations. This work aims providing an analysis of literature to identify key gaps in today's model training efforts and list guiding principles for ongoing and future linear solvation energy relationship studies.

Keywords: Adsorption, Carbon Nanomaterials, Linear Solvation Energy Relationship, LSER, Predictive Model, QSAR

1. General aspects of LSER model development for adsorption of synthetic organic compounds by engineered carbon nanomaterials

Engineered carbon nanomaterials (CNM) are graphitic nano-structures (e.g., graphene nanosheets, carbon nanotubes) with outstanding thermal, dielectric, mechanical, and chemical properties [1, 2]. Due to their aptitude in a wide array of industrial applications, they are being produced with an increasing rate [3]. Their upsurge in production makes their eventual release into the natural aquatic environmental almost inevitable. Once released into the environment, CNMs would interact with over 80,000 synthetic organic compounds (SOC) in either natural or built aquatic systems. In addition, CNMs share the same graphitic backbones with traditional carbonaceous adsorbents and they form bundles (aggregates) in water that resemble the porous activated carbon (AC) particles. Therefore, CNMs are often compared to traditional adsorbents and are tailored to remove SOCs selectively and efficiently from water. Hence, CNM have been accepted as promising adsorbents with versatile physicochemical properties. It is imperative to investigate the adsorption of SOCs by CNMs and eventually generate structure-activity relationships to cut the time, cost and labor requirements for adsorption experiments. These models would also provide mechanistic insights into the SOC-CNM interactions and enable application of CNMs in water treatment by decoding the complex adsorptive interactions in engineered systems.

Since Kamlet and co-workers' published their seminal work in 1985; numerous quantitative structure-activity relationship (QSAR) models have been developed for adsorption of SOCs by AC in aquatic backgrounds [4-16]. These statistical models are trained to predict partitioning of SOCs onto ACs and they assist decoding molecular-level adsorption interactions. Owing to the discovery and upsurge of "Nobel Prize decorated" engineered CNMs in 2000s and 2010s; researchers witnessed the flooding of scientific literature by articles demonstrating the adsorption of SOCs by CNMs (e.g., carbon nanotubes [17-32] and graphene nanosheets [33, 34]). As a consequence of the rapid expansion of the literature, QSAR modeling for adsorption of SOCs by CNMs was made possible in early 2010s.

Multiple mechanisms of varying relative importance have been proposed to control the adsorption of SOCs by carbonaceous adsorbents. The quantification of these individual contributions is a challenging task, and it can partially be addressed by QSAR modeling. QSAR modeling that utilizes solvation theory is an accepted approach. These specific approaches are referred as linear free energy relationship (LFER) or linear solvation energy relationships (LSER) in the literature [35-37]. To train an LSER model, solvatochromic descriptors are utilized and each descriptor can assist explaining adsorptive interactions between the adsorbate, adsorbent surface, and the solvent. The first LSER model for adsorption of organic compounds by multi-walled carbon nanotubes (MWCNT) was published in 2010 by Nature Nanotechnology [30]. Since this first article, dozens of LSER equations were developed for several CNM types for a multitude of SOCs at different saturation concentrations. Figure 1 shows the milestones and a timeline of the LSER modeling literature.

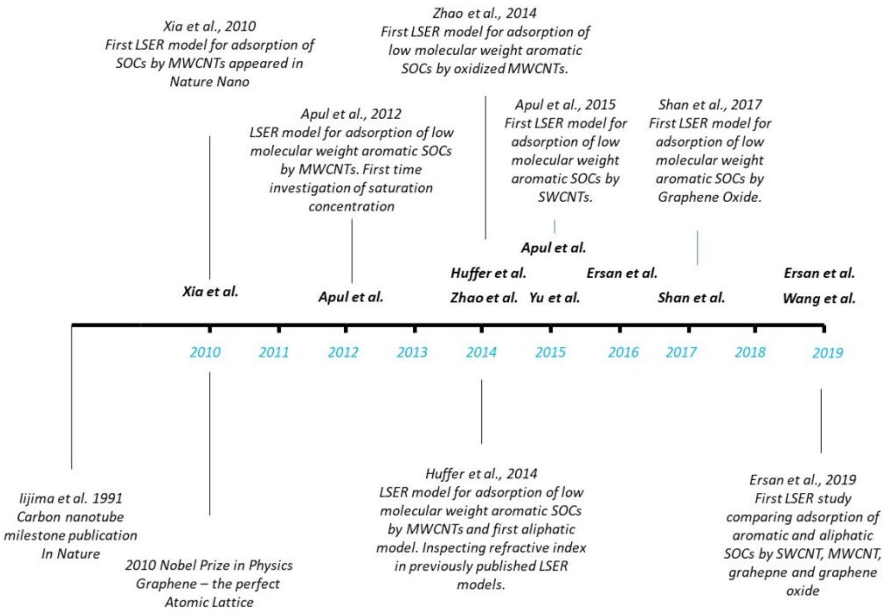


Figure 1. Timeline of LSER modeling studies in the last decade [23-34, 38].

A typical LSER model (eq. 1) uses predetermined set of adsorbate properties i.e., Abraham solvatochromic parameters as independent variables [39-44] and adsorption coefficient as the independent variable. The independent parameters are established based on the solvation theory because of the mechanistic association of solvation and adsorption [35-37].

$$\mathrm{Log}\ K_i = aA + bB + vV + pP + rR + c \qquad \text{[eq. 1]}$$

Here, ‘K’ is the adsorption coefficient at saturation concentration of ‘i’ obtained by taking the ratio of solid phase concentration to aqueous phase concentration (q_e/C_e) at equilibrium. Parameter ‘A’ and ‘B’ indicate hydrogen bond donating and accepting ability, respectively, ‘V’ represents the molecular volume or McGowan’s volume, ‘P’ is the polarizability/dipolarity, ‘R’ stands for the excess molar refraction, ‘c’ is the regression constant and ‘a’, ‘b’, ‘v’, ‘p’, and ‘r’ are regression coefficients [45]. Regression coefficients indicate relative influence of each parameters on adsorption. Positive and negative values of coefficient indicate direct and inverse proportionality, respectively.

LSER model development workflow includes the compilation of available adsorption data (i.e., either experimental data collection or data mining from published isotherms), collection of solvatochromic parameters of the adsorbate molecules, categorization of the datasets with respect to adsorbent properties (e.g., graphenes vs. carbon nanotube dataset), adsorbate properties (e.g., aliphatic vs. aromatic datasets), and aquatic matrix (e.g., distilled water vs. natural organic matter solution). Once the datasets are formed, modeling hypotheses can be tested by generating models via multiple linear regression. This workflow and details of dataset generation has been previously published elsewhere [46]. LSER model development is a vigilant balance between data fitting and prediction strength. An efficient predictive model aims for the (i) minimization of least squared errors (good data fitting ability) and, (ii) maximization of operational width and predictive strength. According to the Topliss and Costello rule [47], to minimize the risk of chance correlations, the ratio of training set SOCs to descriptors should be at least 5:1. Furthermore, this requirement is sometimes violated, when the similar SOCs (a very narrow range of $\log C_{sw}$, $\log K_{ow}$ or molar volume) are selected [26, 27]. In addition to the statistical validity, a strong predictive model is expected to provide mechanistic insights to the adsorption mechanisms by attempting to identify/quantify the intermolecular interactions between adsorbate molecules and adsorbent surfaces. This article (i) summarizes the LSER development efforts for adsorption of SOCs by carbon nanomaterials in the past decade, (ii) critically reviews the ongoing debates in the field and (iii) provide future research recommendations. This review is the first article that surveys the recent literature to provide a framework and workflow for the elements of LSER model development in the field to the best of our knowledge. This study also presents a perspective to merge findings of the last decade and guide future work using the existing knowledge presented on a unified platform.

2. Selecting independent variables

There are 86 LSER models published in 14 articles in the last decade capturing adsorption of organic pollutants by CNMs predominantly aromatic SOCs [21-34]. Figure 2a shows a summary of parameters (i.e., highest statistical significance within the model) for adsorption of SOCs by CNMs. Their coefficients are shown in Figure 2b as a box and whiskers plot. Majority of the models reported *V* parameter as the most prevalent independent variable in the LSER equations with an average coefficient of *2.70±1.78*. Despite the variability of compound and adsorbent properties, the effect of molecular size (*V*) has been repeatedly reported to be the controlling molecular property for adsorption that is positively correlated. This was attributed to the energy requirement of adsorbates for cavity formation among water molecules with increasing molecular size causing stronger hydrophobic repulsion from water onto carbon surfaces. In addition, as molecular sizes increase, non-specific attraction (e.g., van der Waals Forces) can contribute to the adsorption of molecules by carbon nanomaterials. To date, there are no models distinguishing the two intermolecular interactions from each other. Additional experimental work suppressing hydrophobic repulsion via utilization of different solvent may be conducted to gain more insight to the molecular size effect on adsorption.

Hydrogen bond accepting property (*B*) is reported as the most predominant parameter in LSER equations in 30% of the studies with an average coefficient of *-2.90±3.21*, which is inversely proportional to adsorption as indicated by its negative coefficient. This indicates molecules with higher basicity showing stronger tendencies to form hydrogen bonding with water owing to the acidic property of water defined by Bronsted concept. Thus, hydrogen bond accepting molecules have less participation onto the carbon surface. However, 84% of the models that report *B* as the most predominant parameter are trained either by solely aliphatic molecules or datasets that contain aliphatic molecules. Therefore, the selection of adsorbate molecules for LSER training can skew the model parameters, which will be discussed in the following section. Aliphatic molecules tested in literature are usually short chain hydrocarbons

(2-4 C) and their molecular sizes do not cover a wide range to provide a fair comparison with aromatic molecules. Literature lacks comprehensive LSER models for adsorption of aliphatic molecules by CNMs.

Hydrogen bond donating property (*A*), polarizability (*P*), and refractive index (*R*) were relatively less significant with average coefficients *-0.70±1.29*, *0.54±1.37* and, *0.64±1.50*. Although, solvation theory prescribes a set of independent variables, there are multiple interpretations of LSER modeling where less significant parameters are excluded. Some of the studies only consider a subset of statistically significant LSER parameters per statistical analysis. Regardless of the analytical approach, picking the independent variables either by excluding insignificant parameters (elimination), or by adding parameters (stepwise or forward addition) play a critical role on model development. It is debatable whether to remove or keep these insignificant terms because, insignificant variables also capture the subtle intermolecular interactions on adsorption. In addition, level of statistical significance (usually α=0.95 or 0.99) is an important consideration and can create subjectivity. Therefore, models with and without the accepted statistically insignificant parameters must be compared side-by-side in order to objectively show the influence of these parameters. In addition, exclusion of insignificant parameters can result in inflated significance of the model coefficients potentially skewing the results.

The refractive index (*R*) parameter was removed from some models that contain only aromatic molecules due to triviality of excess molar refraction caused by shielded sigma electrons inside π electron clouds [17, 18, 24]. However, later studies reported that the contribution of *R* parameter depends on solute saturation concentration as well as the adsorbent type [26]. Influence of *R* parameter comprises induction of dipoles and polarizable aromatic molecules may show a statistically significant impact on adsorption [29]. In addition, some studies have reported that the type of adsorbent (multi-walled vs. single-walled carbon nanotubes) indicated different importance of *R* parameter [28,31]. Despite the variability of significant parameters, *R* parameter has not been reported to be the most predominant term in LSER models so far [28].

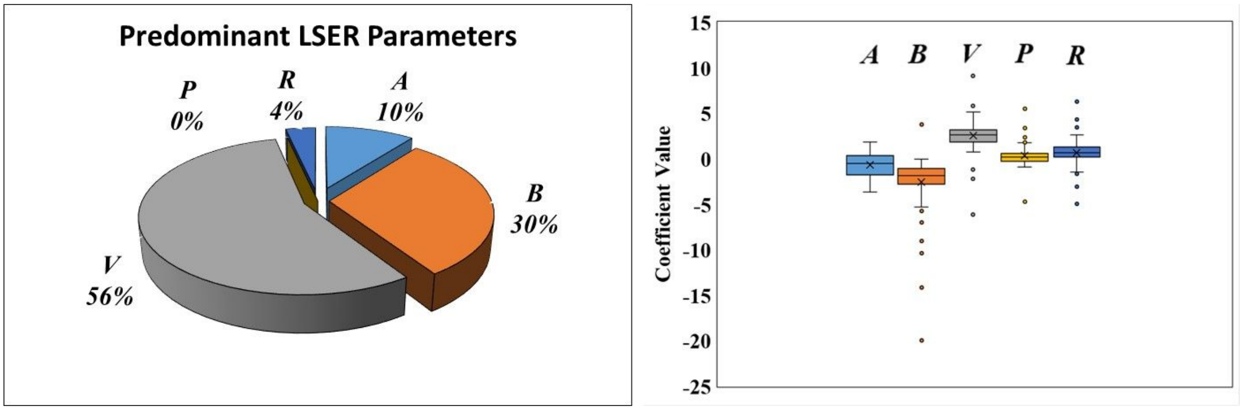


Figure 2. Summary of all predominant LSER parameters based on all reported 86 LSER equations (on left), their coefficient ranges as box, and whiskers plot (on right).

3. Dataset generation and pre-processing prior to model development

To date dozens of articles reported adsorption of 140 organic molecules' by carbon nanomaterials [33]. The virtual limit of LSER model development is availability of published adsorption data unless researchers undertake experimental data generation. Further, collection, indexing and selection of the data to be utilized in LSER equation is a tedious task. There are two components of the dataset: (i) independent variables i.e., solvatochromic descriptors and (ii) dependent variables i.e., adsorption descriptors. The compilation of independent variables are rather simple owing to the advances in computation chemistry software (e.g., ACD ADME Suite) [21-34]. However, the collection and pre-processing of adsorption descriptors is a time consuming task. Data is usually reported as adsorption isotherms and equilibrium (or pseudo-equilibrium) saturation concentration, which depends on the initial concentration. Therefore, extracting the data from the published data (e.g., isotherms) at the desired saturation concentration needs to be performed within the actual experimental range to minimize extrapolation-based errors.

The intermolecular interactions that are governing the adsorption of SOCs by CNMs profoundly depend on compound properties. Combining different classes of compounds (e.g., aromatic and aliphatic compounds, ionizable vs. non-ionizable) in one dataset may increase the operating range of the model; however, the model linearity and parameter interpretation vigor would be compromised. Therefore, modeling efforts must embrace subsets of the database that are categorized with a meaningful approach. Although there are several approaches to classify SOCs (e.g., size, polarity, formal charge, polarizability, and aromaticity); adsorbate molecule's size and complexity are critical categorization metrics for three reasons. First, size of a molecule dictates non-specific attractive force between the compound and the surface. Larger molecules generally have more time-varying uneven electron distribution within molecule's structure that enables interactions with adjacent molecules, which increases the intensity of intermolecular forces. Similarly, larger molecules require more energy to partition in water (unless hydrogen bonding functional groups exist in its structure) causing larger cavity formation energies in water. Second, large and more complex molecules contain more functional groups (e.g., carboxylic acids, ketones, phenolic groups, amines) than small and simple molecular structures. The complexity of the molecular structure complicates the intermolecular interactions from a modeling perspective i.e., hydrogen bonding, molecule-molecule interactions. Therefore, size dictated cut-offs may enable LSER models for adsorption of SOCs by CNMs. Third, SOC adsorption strictly depends on the physical accessibility of molecules into the final sorption sites in the pore structures, thus similar size molecules will have similar pore access. Although, molecular complexity and size is an essential metric for categorization, researchers need to consider other physicochemical SOC classifications based on their datasets prior to presenting their models. Because, some molecules have the ability to configure their structure and pack in narrow pores, on the contrary, rigid and stable molecules lack this ability. In addition, if an ionizable molecule is included in the LSER model, the degree of protonation at the corresponding solution pH must then be taken into account to properly assess molecular properties.

The aqueous concentration of the adsorbate is another critical factor that influences LSER modeling. Molecules in a dilute solution are assumed not to interact with each other regardless of molecular level thermodynamic considerations simply due to extremely low likelihood of molecules to coincide in bulk solution (Figure 3a). Therefore, the adsorption interactions are limited to (i) attractive forces between the molecules and the adsorbent surface e.g., non-specific attraction, H-bonding, electrostatic attraction, π-π bonding (ii) repulsive forces between the solvent and the molecules e.g., hydrophobic repulsion, salting out. On the other hand, saturated solutions exert additional molecule-molecule interactions in solution or on the surface (Figure 3b). Therefore, complex intermolecular interactions are likely to be captured by solvatochromic parameters changing the model parameters significantly. The most adequate approach to take the potential bias of saturation concentration into account would be comparing LSER equations at same aqueous concentration; however, the compounds used in LSER equations display a very wide range of solubility (from $1\text{x}10^{-1}$ to $5\text{x}10^{5}$ mg/L). Considering the orders of magnitude difference in solubility of SOCs, it is unlikely to find an overlapping (pseudo-)equilibrium concentration. Therefore, it is typical to normalize the equilibrium concentration with respect to solubility of the compound and report adsorption affinity (K_D) at a fixed percent solubility concentration (e.g., adsorption affinity at 1% of compound solubility: q_e/C_e at 1% solubility = $K_{D,\ 0.01}$).

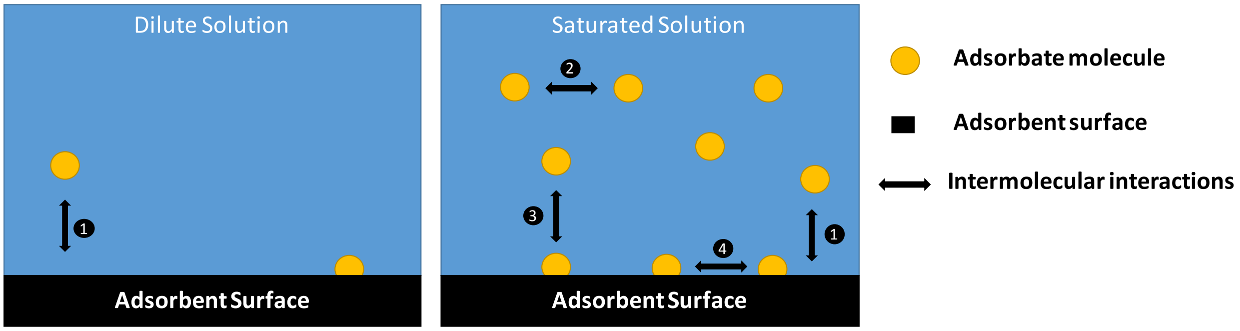


Figure 3. Intermolecular interactions in dilute and saturated solutions: (1) adsorbate molecule - surface interactions (2) adsorbate-adsorbate interactions in bulk solution (3) adsorbed molecule – dissolved molecule interactions (4) interaction of adsorbed molecules

4. Effect of CNM properties on LSER modeling

Despite the lack of adsorbent properties in quintessential LSER modeling, it is clear that CNM properties play a defining role for adsorption. There are several approaches to account for the impact of adsorbent in LSER modeling. The first and foremost is realizing the effects of predominant nanomaterial properties on adsorption such as specific surface area, pore volume, pore size distribution, outer diameter, length, number of graphene layer. Specific surface area measurement is generally conducted by N_2 adsorption onto the adsorbent in bulk phase under vacuum at 77 K and subsequently the data is modeled according to Brunauer–Emmett–Teller gas adsorption theory [48]. Although aqueous phase adsorption conditions are much different from those of N_2 gas adsorption, the standardized surface area measurement is considered as the benchmark tool by most researchers. Preliminary efforts indicated that normalizing the MWCNT adsorption descriptors by specific surface area had no impact on LSER modeling for low SOC concentration models (e.g., 0.02% of C_s) [23]. This can be attributed to LSER adsorption descriptors capturing adsorption affinity at dilute conditions rather than adsorption capacity, for which surface area is used as a surrogate parameter [23]. In other words, availability of total surface area is not critical for adsorption of a small number of SOC molecules. Instead, other physical features of CNMs may be influential on adsorption. For example, at low concentrations (1 μg/L), larger diameter MWCNTs were shown to adsorb more phenanthrene molecules than smaller diameter MWCNTs per unit surface area [49]; however, this relationship disappears at higher SOC concentration (>100 μg/L). This observation was attributed to the lower surface curvature (or large diameter) of MWCNTs thermodynamically enabling stronger π- π interactions with flat and polyaromatic phenanthrene molecules via better alignment and stacking. Overall, these observations indicate that surface area normalization to account for the differences of CNMs in LSER modeling may not be necessary for particularly low SOC concentrations. Similar to specific surface area, total pore volume or a defined sub-fraction of pore volume (e.g., pores smaller than 2 nm or micropores) can also be used as an indicator for adsorption. However, the same argument persists for the inability of pore volume to assess LSER modeling at low SOC concentrations. Previous work indicated that SWCNTs tend to form tighter aggregates than MWCNTs due to stronger non-specific interparticle attraction, creating microporous SWCNT bundles in water [50]. Difference between SWCNT and MWCNT are structural are not mechanistically captured in a single LSER; therefore, the approach taken in the scientific community has been historically to develop models for general classes of CNMs such as SWCNTs, MWCNTs, graphenes, and graphene oxides. This resulted in an increase the model linearity and strength of models within the prediction domain.

In addition to physical characteristics, there are also chemical properties of CNMs that influence the LSER modeling. The surface chemistry is described by a variety of parameters such as total surface oxygen content (elemental analysis), oxygen-containing functional groups (Boehm titration), pH point of zero charge (pH_{pzc}), and water vapor adsorption, which aim to indicate the surface polarity. The surface polarity increases the intermolecular H-bonding interactions between CNM surfaces and water molecules creating clusters of water on the surface minimizing the interactions of hydrophobic SOCs with the surface. Other efforts indicated that the CNM functional groups cause increased π-π electron donor-acceptor interactions due to localization of π electrons on the graphitic surfaces especially for aromatic SOCs [51]. In addition, increased H-bonding interaction between polar functional groups of SOCs and CNM surfaces were reported [52, 53]. More recently, incident impact of air bubbles on favorable polar interactions for adsorption of amphiphilic molecules by CNTs were demonstrated [54]. Therefore, adsorptive interactions strongly rely on the surface chemistry of CNMs and are important to account for in LSER models. The relative contribution of H-bond donating and accepting abilities of molecules are captured in LSER models but no mechanistic approach is widely accepted for capturing the same ability of CNM adsorbents. Since the distinguishing factor between 'pristine' and 'oxidized' CNMs is the overall oxygen content, researchers define the practical cutoff values for adsorbents' oxygen content (e.g., <5wt.% oxygen) to classify them. The "pristine" CNMs contain minute amounts of oxygen on the imperfections of the basal plane, wrinkles, folds, edges and corners; therefore, during LSER modeling, a common approach is creating subsets of CNMs with different oxygen content improve the model fit and prediction ability [23]. Although the model captures the polar interactions, CNM physical structure is intertwined with the surface chemistry. In contrast to rigid and traditional graphitic adsorbents (e.g., activated carbons), higher surface polarity may result in looser aggregation of CNTs or larger sheet-sheet distances for graphenes [50, 55, 56]. Further, new approaches that couples surface chemistry and physical properties of CNMs can be developed to unravel the adsorptive interactions of SOCs with functionalized CNMs [57].

5. Summary and recommendations for future work

Various engineering applications, technology development opportunities and/or scientific curiosity motivated our field to create the critical mass for developing statistical predictive models for adsorption of SOCs by CNMs. This literature survey summarizes the recent findings while researchers pursue their debates about "creating the perfect model" for adsorption of synthetic organic contaminants by carbon nanomaterials. Whether the perfect model exists or not, it is evident that the studies significantly contributed to our current understanding of this research area.

Today's literature largely focuses on LSER models for carbon nanotubes and graphene derivatives. Although, the chemical interactions of the SOCs and all carbon adsorbents will have similarities in water, physical properties of carbon adsorbent (e.g., chirality, end groups, morphology, and defects) and their aggregation must be systematically investigated. Although, the final form of nanomaterial micro-assemblies in water is difficult-to-predict, it will be utterly different from their bulk phase bundle structures. Researchers must recognize that the surface area measurement of CNMs are generally determined under nitrogen gas ambient at 77 K via

BET and DFT models. Detected surface area and pore structure may not be accurate in water because of the aggregation behavior of nanomaterials. Therefore, researchers need to reveal the aggregation mechanism of CNMs in water and its relationship to adsorptive interactions.

Moving forward, a better understanding of how LSER models are changing in function of the surface chemistry and morphology of the carbon surface will be important in predicting the adsorption capacity of a broader range of carbon nanostructure. Previous studies have shown that properties such as tube curvature, oxygen content, and type of oxygen functionalization can all affect the affinity of organic contaminants to carbon nanomaterials [49, 58-60]. How these different properties affect the predictive capacity of LSER models is still not completely understood. A recent study on LSER development for graphenes and carbon nanotubes showed that while the main solvation descriptors of importance remained unchanged, the relative importance of each parameter in the LSER model differed in function of the carbon nanostructure [33]. Right now, LSER models are ill-equipped to describe how the adsorbent properties can change the LSER model prediction. To address this knowledge gap, a systematic approach studying well-characterized nanostructures of progressively varying properties could identify the boundaries where the properties of the carbon nanomaterials start to significantly affect the predictive capacity of the LSER models.

Another challenge in hand is the missing bridge between predictive models that are developed in distilled and deionized waters or similar simplified aquatic matrices and real-world applications. The community must focus on developing or affirming predictive models in complex ambient (controlled or uncontrolled) to address the real-world application of these models. As our understanding of LSER models for CNMs becomes more and more complete, there will be opportunities for designing other ultrahigh surface area nano-adsorbents. These novel and promising adsorbent materials would employ LSER models and guide material development for their efficient use for contaminant removal. While mechanistic model development studies focus on compiling and collecting experimental data, strategies beyond multiple linear regression such as support vector machine and artificial neural networks [24, 61] will be capable of improving the models. These traditional tools will further be improved within the context of artificial intelligence and machine learning algorithms if integrated into the developing “big data” vision. Therefore, researchers are encouraged to provide their experimental data and detailed information in the supporting information section of their publications and support the growing efforts of connecting user-fed databases globally. The improvement of quantifiable metrics for model success must always be complemented with mechanistic models to gain insights into intermolecular interactions between SOCs and CNMs.

Providing a better predictive capability for the adsorption of organic contaminants on surfaces can have important implications for the fate and transport of organic contaminants in the environment and improve engineered water treatment systems if we can address the existing research gaps. The broad predictive capabilities of LSER models offer a powerful framework to better understand the interaction of organic molecules with surfaces and develop appropriate remediation or control strategies.

Conflicts of Interest

There are no conflicts to declare

Acknowledgements

This work was partially supported by research grants from the National Science Foundation (CBET 1511826 and EEC-1449500) and by faculty start-up grants from University of Massachusetts Lowell. However, the manuscript has not been subjected to the peer and policy review of the agencies and therefore does not necessarily reflect their views.

References

1. Mauter, Meagan S., and Menachem Elimelech. "Environmental applications of carbon-based nanomaterials." *Environmental Science & Technology* 42.16 (2008): 5843-5859.

2. Menéndez, J. A., et al. "Microwave heating processes involving carbon materials." *Fuel Processing Technology* 91.1 (2010): 1-8.

3. Janković, N. Z., and D. L. Plata. "Engineered nanomaterials in the context of global element cycles." *Environmental Science: Nano* 6.9 (2019): 2697-2711.

4. Luehrs, D.C., et al., *Linear Solvation Energy Relationship of the Limiting Partition Coefficient of Organic Solutes between Water and Activated Carbon.* Environmental Science & Technology, 1996. **30**(1): p. 143-152.

5. Kamlet, M.J., et al., *Linear solvation energy relationships. 33. An analysis of the factors that influence adsorption of organic compounds on activated carbon.* Carbon, 1985. **23**(5): p. 549-554.

6. Blum, D.J.W., I.H. Suffet, and J.P. Duguet, *Quantitative structure-activity relationship using molecular connectivity for the activated carbon adsorption of organic chemicals in water.* Water Research, 1994. **28**(3): p. 687-699.

7. Brasquet, C. and P. Le Cloirec, *QSAR for organics adsorption onto activated carbon in water: what about the use of neural networks?* Water Research, 1999. **33**(17): p. 3603-3608.

8. Shih, Y. and P.M. Gschwend, *Evaluating Activated Carbon−Water Sorption Coefficients of Organic Compounds Using a Linear Solvation Energy Relationship Approach and Sorbate Chemical Activities.* Environmental Science & Technology, 2009. **43**(3): p. 851-857.

9. Dickenson, E.R.V. and J.E. Drewes, *Quantitative structure property relationships for the adsorption of pharmaceuticals onto activated carbon.* Water Science and Technology, 2010. **62**(10): p. 2270-2276.

10. Burg, P., et al., *The characterization of an active carbon in terms of selectivity towards volatile organic compounds using an LSER approach.* Fuel, 2000. **79**: p. 1041–1045.

11. Burg, P., et al., *Comparison of three active carbons using LSER modeling: prediction of their selectivity towards pairs of volatile organic compounds (VOCs).* Carbon, 2002. **40**: p. 73–80.

12. Burg, P. and D. Cagniant. *Study of the influence of the soluble organic fraction of an exhaust diesel soot by a linear solvation energy relationship approach.* Carbon, 2003. **41** (5): p. 1031−1035.

13. Zhang, S., X. Liu, and T. Karanfil, *Applicability of the linear solvation energy relationships in the prediction for adsorption of aromatic compounds on activated carbons from aqueous solutions.* Separation and Purification Technology, 2013. **117**: p. 111–117.

14. De Ridder, D. J., et al., *Modeling equilibrium adsorption of organic micropollutants onto activated carbon,* Water Research, 2010. **44**: p. 3077–3086.

15. Abe, I., et al., *Prediction of adsorbability of organic compounds from aqueous solution on activated carbon by means of the linear free-energy relationship*, Bulletin of the Chemical Society of Japan, 1983. **56**: p. 1002–1005.

16. Chen, W., L. Duan, and D. Zhu, *Adsorption of polar and nonpolar organicchemicals to carbon nanotubes.* Environmental Science and Technology, 2007. **41(24):** p. 8295-8300.

17. Roy, J., et al., *Predictive quantitative structure–property relationship (QSPR) modeling for adsorption of organic pollutants by carbon nanotubes (CNTs).* Environmental Science: Nano, 2019. **6**: p. 224–247.

18. Rahimi-Nasrabadi, M., et al., *Predicting adsorption of aromatic compounds by carbon nanotubes based on quantitative structure property relationship principles*. Journal of Molecular Structure, 2015. **1099**: p. 510–515.

19. Chayawan, V., *Quantum-mechanical parameters for the risk assessment of multiwalled carbon-nanotubes: a study using adsorption of probe compounds and its application to biomolecules.* Environmental Pollution, 2016. **218**: p. 615–624.

20. Hassanzadeh, Z., et al., *Combining radial basis function neural network with genetic algorithm to QSPR modeling of adsorption on multi-walled carbon nanotubes surface.* Journal of Molecular Structure, 2015. **1098**: p. 191–198.

21. Li, M. S., et al., *Linear free energy relationships for the adsorption of volatile organic compounds onto multiwalled carbon nanotubes at different relative humidities: comparison with organoclays and activated carbon.* Environmental Science: Processes & Impacts, 2017. **19(3)**: p. 276-287.

22. Wang, Y., et al., *Modeling adsorption of organic pollutants onto single-walled carbon nanotubes with theoretical molecular descriptors using MLR and SVM algorithms*. Chemosphere, 2019. 214: p. 79-84.

23. Apul, O.G., et al., *Predictive Model Development for Adsorption of Aromatic Contaminants by Multi-Walled Carbon Nanotubes.* Environmental Science & Technology, 2013. **47**(5): p. 2295-2303.

24. Wang, Q.L., et al., *Development of a 3D QSPR model for adsorption of aromatic compounds by carbon nanotubes: comparison of multiple linear regression, artificial*

neural network and support vector machine. RSC Advances, 2013. **3**(46): p. 23924-23934.

25. Wu, W., et al., *Correlation and prediction of adsorption capacity and affinity of aromatic compounds on carbon nanotubes.* Water Research, 2016. **88**: p. 492-501.

26. Zhao, Q., et al., *Concentration-dependent polyparameter linear free energy relationships to predict organic compound sorption on carbon nanotubes.* Scientific Reports, 2014. **4**: p. 3888.

27. Apul, O.G., Y. Zhou, and T. Karanfil, *Mechanisms and modeling of halogenated aliphatic contaminant adsorption by carbon nanotubes.* Journal of Hazardous Materials, 2013. **295:** p. 138–144.

28. Ersan, G., O. G. Apul, and T. Karanfil, *Linear solvation energy development for adsorption of organic contaminants by carbon nanotubes.* Water Research, 2016. **98**: p. 28–38.

29. Hüffer, T., et al., *Prediction of sorption of aromatic and aliphatic organic compounds by carbon nanotubes using polyparameter linear free–energy relationships*. Water Research, 2014. **59**: p. 295–303.

30. Xia, X.R., A. Monteiro-Riviere, and J. E. Riviere, *An index for characterization of nanomaterials in biological systems*. Nature Nanotechnology, 2010. **5 (9)**: p. 671–675.

31. Yu, X., W. Sun, and J. Ni, *LSER model for organic compounds adsorption by single walled carbon nanotubes: comparison with multi-walled carbon nanotubes and activated carbon.* Environmental Pollution, 2015. **206**: p. 652–660

32. Zhao, Y., et al., *Prediction of adsorption properties for ionic and neutral pharmaceuticals and pharmaceutical intermediates on activated charcoal from aqueous solution via LFER model*. Chemical Engineering Journal, 2019. **362**: p. 199-206.

33. Ersan, G., O.G. Apul, and T. Karanfil, *Predictive models for adsorption of organic compounds by Graphene nanosheets: comparison with carbon nanotubes.* Science of The Total Environment, 2019. **654**: p. 28-34.

34. Shan, S., et al., *Linear solvation energy relationship to predict the adsorption of aromatic contaminants on graphene oxide.* Chemosphere, 2017. **185**: p. 826-832.

35. Endo, S. and Goss, K.-U, *Applications of Polyparameter Linear Free Energy Relationships in Environmental Chemistry.* Environmental Science & Technology, 2014. **48 (21):** p. 12477-12491.

36. Nguyen, T. H., K.-U, Goss, and W.P. Ball, *Polyparameter linear free energy relationships for estimating the equilibrium partition of organic compounds between water and the natural organic matter in soils and sediments.* Environmental Science & Technology, 2005. **39 (4)**: p. 913−924.

37. Goss, K.-U. and R.P. Schwarzenbach, *Linear free energy relationships used to evaluate equilibrium partitioning of organic compounds.* Environmental Science & Technology, 2001. **35 (1)**: p. 1−9.

38. Lijima, S, *Helical microtubules of graphitic carbon,* Nature, 1991. **354**: p. 56–58.

39. Abraham, M. H, and J.C. McGowan, *The use of characteristic volumes to measure cavity terms in reversed phase liquid chromatography.* Chromatographia 1987, **23 (4)**: p. 243−246.

40. Abraham, M. H., P.L. Grellier, and R.A. McGill, *Determination of olive oil-gas and hexadecane-gas partition coefficients, and calculation of the corresponding olive oil-water and hexadecane-water partition coefficients.* Journal of the Chemical Society, Perkin Transactions 2, 1987. **6**: p. 797−803.

41. Abraham, M. H., et al., *Hydrogen bonding. Part 13. A new method for the characterization of GLC stationary phases: The Laffort data set.* Journal of the Chemical Society, Perkin Transactions 2, 1990. **8**: p. 1451−1460.

42. Abraham, M. H., et al., *Hydrogen bonding. XVI. A new solute solvation parameter, π2 H, from gas chromatographic data.* Journal of Chromatographia, 1991. **587 (2)**: p. 213−228.

43. Abraham, M. H., *Hydrogen-bonding. 31. Construction of a scale of solute effective or summation hydrogen-bond basicity.* Journal of Physical Organic Chemistry, 1993. **6 (12)**: p. 660−684.

44. Abraham, M. H., S*cales of solute hydrogen-bonding: Their construction and application to hysicochemical and biochemical processes.* Chemical Society Reviews, 1993. **22** (**2**): p. 73−83.

45. Brusseau, M.L., et al., *Comprehensive retention model for PFAS transport in subsurface systems.* Water Research, 2019. **148**: p. 41-50.

46. Apul, O.G. *Predictive model development for adsorption of organic contaminants by carbon nanotubes*. 2014, Doctoral Dissertation, Clemson University, Clemson, SC.

47. Topliss, J. G. and J.D. Costello, *Chance correlation in structure-activity studies using multiple regression analysis.* Journal of Medicinal Chemistry, 1972. **15**: p.1066–1069.

48. Brunauer, S, Emmett, P., and Teller, E. *Adsorption of gases in multimolecular layers*. Journal of the American chemical society, 1938. **60.2**: 309-319.

49. Apul, O.G., et al., *Impact of carbon nanotube morphology on phenanthrene adsorption*. Environmental toxicology and chemistry, 2011. **31 (1):** p. 73-78.

50. Zhang, S., Shao, T., Bekaroglu, S. S. K., & Karanfil, T. *The impacts of aggregation and surface chemistry of carbon nanotubes on the adsorption of synthetic organic compounds*. Environmental science & technology, 2009 **43(15)**, 5719-5725.

51. Lin, D, and Baoshan X. *Adsorption of phenolic compounds by carbon nanotubes: role of aromaticity and substitution of hydroxyl groups.* Environmental science & technology 42.19 (2008): 7254-7259.

52. Chen, W, Duan, L. and Zhu. D. *Adsorption of polar and nonpolar organic chemicals to carbon nanotubes*. Environmental science & technology 41.24 2007: 8295-8300.

53. Chen, W., Duan, L., Wang, L. and Zhu, D. Adsorption of hydroxyl- and amino-substituted aromatics to carbon manotubes. Environmental Science & Technology, 2008. 42(18): p. 6862-6868.

54. Meng, P., et al. *Role of air bubbles overlooked in the adsorption of perfluorooctanesulfonate on hydrophobic carbonaceous adsorbents*. Environmental science & technology 48.23. 2014: 13785-13792.

55. Zhao, J., Wang, Z., White, J. C., & Xing, B. *Graphene in the aquatic environment: adsorption, dispersion, toxicity and transformation*. 2014, Environmental science & technology, 48(17), 9995-10009.

56. Saleh, N. B., L.D. Pfefferle, and M. Elimelech, *Aggregation kinetics of multiwalled carbon nanotubes in aquatic systems: Measurements and environmental implications.* Environmental. Science & Technology, 2008, **42**: p. 7963–7969.

57. Dominguez-Medina, S., Kisley, L., Tauzin, L. J., Hoggard, A., Shuang, B., DS Indrasekara, A. S., ... & Zubarev, E. R. *Adsorption and unfolding of a single protein triggers nanoparticle aggregation*. 2016. ACS nano, 10(2), 2103-2112.

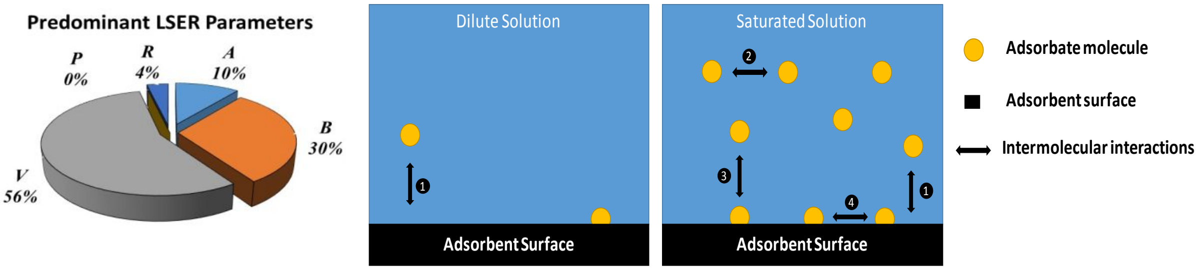
58. Li, Y., et al., *Size effect of single-walled carbon nanotube on adsorption of perfluorooctanesulfonate.* Chemosphere, 2013. **91 (6):** p. 784-790.

59. Gotovac, S., et al., *Effect of Nanoscale Curvature of Single-Walled Carbon Nanotubes on Adsorption of Polycyclic Aromatic Hydrocarbons*. Nano Letter, 2007**. 7 (3):** p. 583-587.

60. Wei, J., et al., *Comparing the effects of different oxygen-containing functional groups on sulfonamides adsorption by carbon nanotubes: Experiments and theoretical calculation*. Chemical Engineering Journal, 2017. **312:** p. 167-179.

61. Zhang, K., Zhong, S., & Zhang, H. J. *Predicting aqueous adsorption of organic compounds onto biochars, carbon nanotubes, granular activated carbons, and resins with machine learning*. Environmental Science & Technology, 2020.

$$\text{Log } K = aA \underline{+\ bB + vV +} pP + rR + c$$



The LSER models for adsorption of SOCs by CNMs profoundly depend on compound properties, adsorbent characteristics and the background solution.