Prediction of Reaction Rates of VOCs with NO$_3$ Radicals

Science For A Better Life

Michael Schindler, Sebastian Schmidt

Bayer AG Crop Science Division, Environmental Safety, Alfred-Nobel-Str. 50, D-40789 Monheim, Germany

E-mail contact: sebastian.schmidt1@bayer.com

Abstract

A QSAR for the prediction of rate constants for the degradation of volatile organic compounds by nitrate radicals is developed using the Partial Least Squares technique. The QSAR is based on experimental data published in the literature for 260 compounds. They are modeled by a set of calculated descriptors from standard descriptor generation tools and from quantum chemistry.

A diversity-based data set of 99 compounds turned out to be the optimum choice with regard to simplicity and performance. The final QSAR model is characterized by r^2=0.831 (fit) and q^2=0.823 (prediction), and by an r^2_{pred}=0.862 for the n=155 external validation set. The QSAR needs 3 latent variables. The most important descriptors for the QSAR are the ionization potential, obtained from density functional theory, and the energy of the highest occupied molecular orbital, which are modulated by fingerprints indicating the presence of specific molecular fragments like functional groups or ring systems.

Introduction

The lifetime of Volatile Organic Compounds (VOCs) in the troposphere depends on their degradation either by direct photo-transformation or by indirect processes, such as their reaction with radicals. It is part of the process of approval of plant protection products to estimate the lifetime in air in order to assess the potential for long range transport (Regulation 1107/2009 EC). Therefore methodology is required to support the determination of the half-life of chemicals in the troposphere. The most important degradation route is via reaction with hydroxyl radicals, predominantly occurring during daytime. At night, the reaction with the (photochemically unstable) nitrate radicals occurs.

The reaction rates of hydrocarbons with NO$_3$ radicals are approximately 30 times lower than with OH radicals, however, this may well be compensated by the increased concentration of NO$_3$ radicals at night-time. It has been found that NO$_3$ contributed 28% to the overall VOC initiated oxidation in the continental boundary layer over Europe. Especially for phenols the reaction with NO$_3$ radicals appears to be an important degradation route.
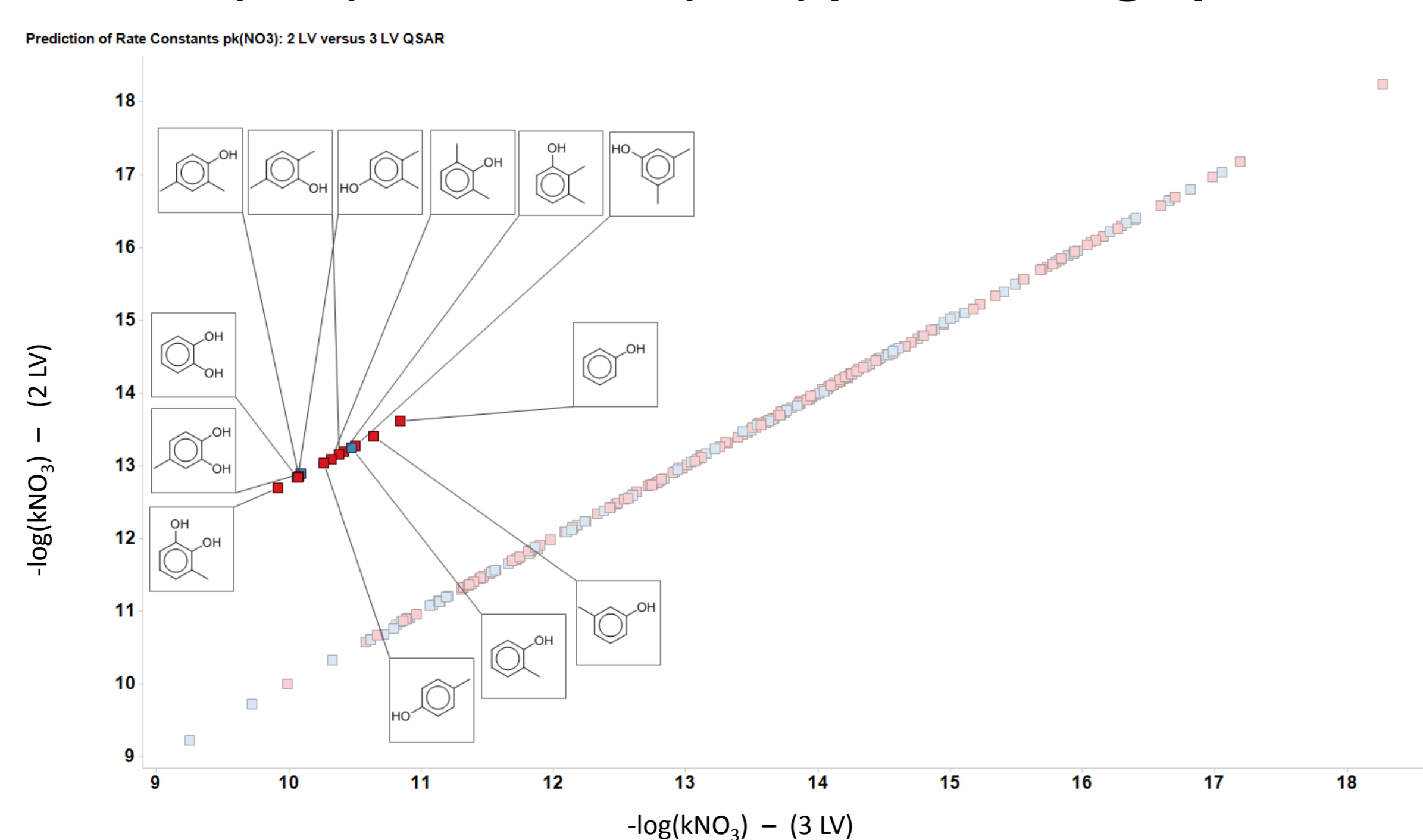
Results

QSAR models for the prediction of rate constants (kNO$_3$) for the reaction of nitrate radicals with VOCs have been developed [1] by using the PLS approach [2]. They are based on compound selections guided by diversity considerations. The optimum QSAR uses 99 out of 260 compounds as the work set, needs three latent variables (LVs), is characterized by r^2=0.831 (fit), q^2=0.823 (prediction), by an r^2_{pred}=0.862 for the validation set of n=155 compounds and has a larger applicability domain (AD) than previous models [3].

Graph 1. Experimental versus predicted rate constants for the full data set from the final 3-LV-PLS model

PLS QSAR (n=99) of -log k(NO3), Prediction vs. Experiment, n=260

y = 2.13 + (0.84 * x)
r2 = 0.848
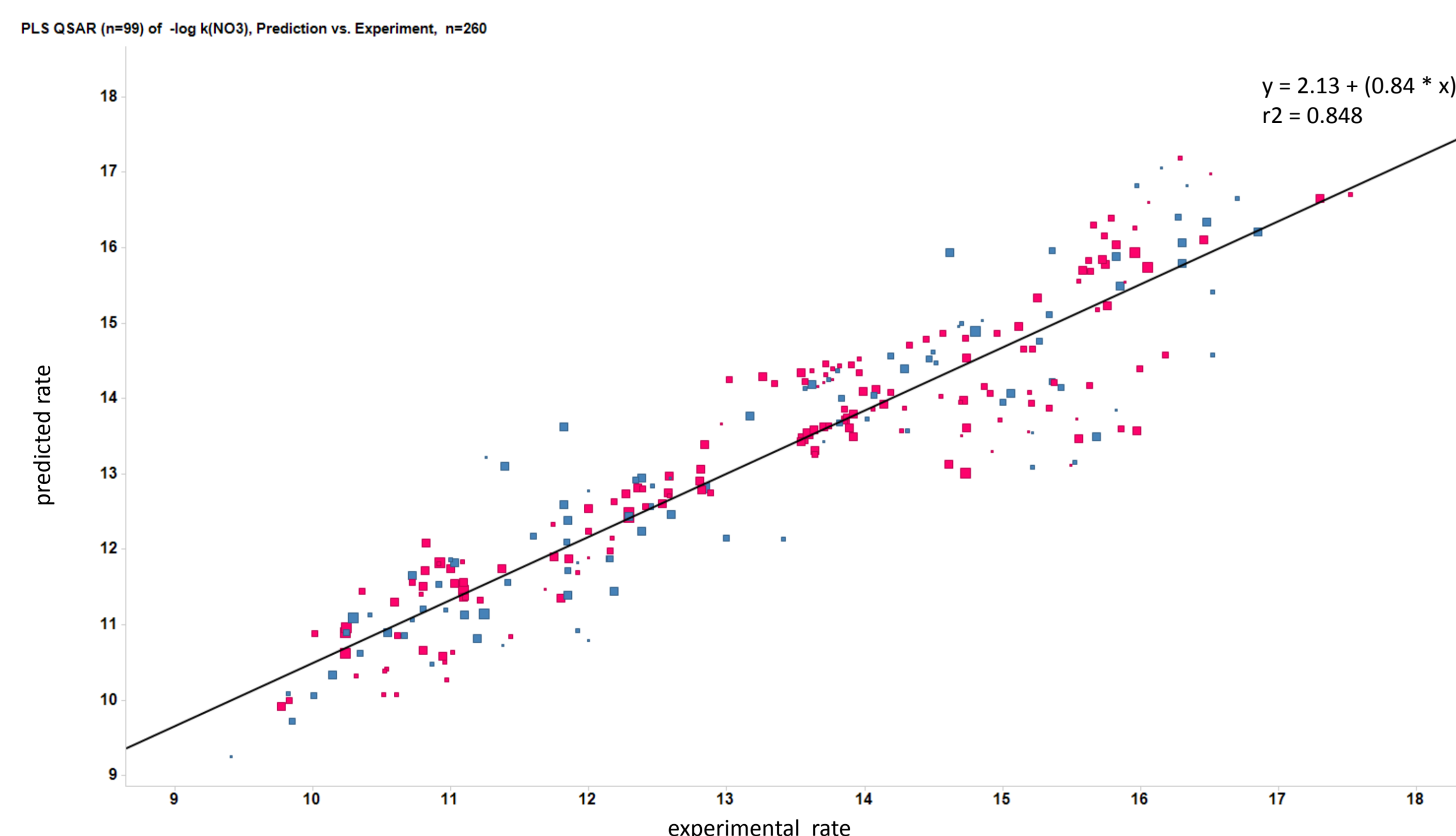
The most important descriptors for the QSAR are the ionization potential (IP) and the energy of the highest occupied molecular orbital (HOMO) which are modulated by fingerprints indicating the presence of specific molecular fragments. The effect of insufficiently differentiating between phenolic- and non-phenolic aromatic compounds is shown in Graph 2.

Graph 2: Comparison of predicted rate constants by two QSAR models with (3 LV) and without (2 LV) phenolic fingerprints

Prediction of Rate Constants pk(NO3): 2 LV versus 3 LV QSAR

$$-\log(kNO_3) = -7.19 - 2.77 \times \mathbf{f320} + 2.15 \times \mathbf{f405} - 1.00 \times \mathbf{HOMO} + 1.49 \times \mathbf{IP} \quad \text{(3 LV)}$$

$$-\log(kNO_3) = -7.21 \qquad\qquad + 2.14 \times \mathbf{f405} - 1.05 \times \mathbf{HOMO} + 1.43 \times \mathbf{IP} \quad \text{(2 LV)}$$

The validity of a QSAR is affected by high-leverage data points. Care has to be taken to avoid so-called 'bad high-leverage points' (Graph 3). However, 'good high-leverage points' can stabilize a model.

Graph 3: Williams plot of the 2-LV- and the final 3-LV-QSAR model
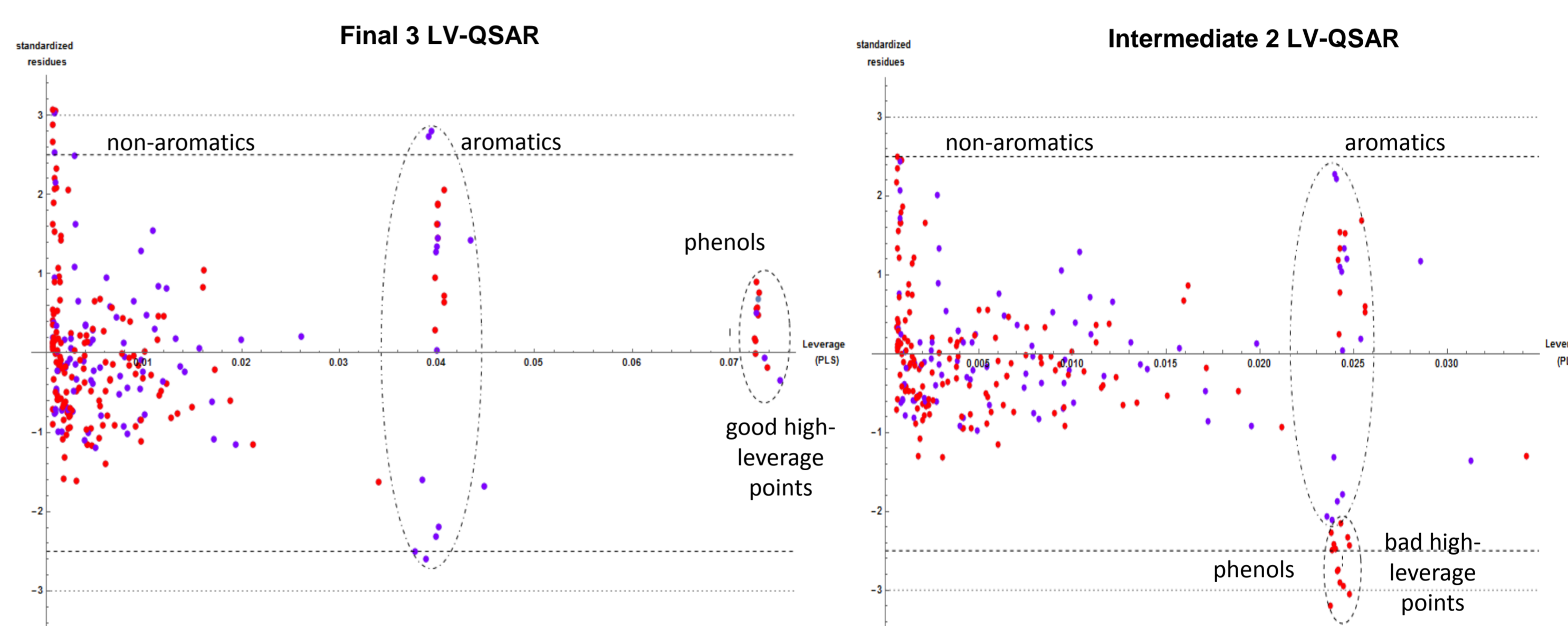
Dotted and dashed lines denote ±3σ and ±2.5σ. Aromatic and phenolic compound clusters are designated by dot-dashed and dashed ellipses; work- and test-set compounds are shown in blue and red

Conclusions

A robust QSAR model for the prediction of reaction rates of VOCs with nitrate radicals has been developed, extending the range of applicability to hydroxybenzenes. Before using the QSAR for predictions, it is important to check whether the new compounds belong to the AD of the QSAR by using the statistical membership probability criteria and the descriptor value ranges.

The best validation of a QSAR is the comparison of predictions with new experimental data. This is shown below for methoxyphenols [4], demonstrating the reliability of the present QSAR.

Table 1. External validation by comparison with new experiments

Methoxphenol	kNO$_3$	QSAR Parameters				−log(kNO$_3$)	
(mp)	exptl. data	f405	f320	HOMO [eV]	IP [ev]	experiment	prediction
4-mp	1.375E-10	1	1	-8.64	5.81	9.86	9.52
3-mp	1.15E-11	1	1	-8.94	6.08	10.94	10.23
2,6-dmp	1.584E-10	1	1	-8.61	5.87	9.80	9.59
2-mp	2.69E-11	1	1	-8.69	5.97	10.57	9.82
2-m-4-mep	8.41E-11	1	1	-8.61	5.84	10.08	9.55

References

[1] M.Schindler, A QSAR for the prediction of rate constants for the reaction of VOCs with nitrate radicals, Chemosphere 154 (2016) 23-33

[2] Umetrics, 2014. SIMCA-P+ 14. Umetrics AB, Umea, Sweden

[3] E.Papa, E.Gramatica, Externally validated QSPR modelling of VOC tropospheric oxidation by NO$_3$ radicals. SAR QSAR Environ.Res. 19, 655-668 (2008)

[4] A.Lauraguais et al., Kinetic Study of the Gas-Phase Reactions of Methoxyphenol Compounds: Experimental and Theoretic Studies, J. Phys. Chem. A , 2016, 120 (17), 2691–2699