Electronic Supplementary Material (ESI) for Environmental Science: Water Research & Technology.
This journal is © The Royal Society of Chemistry 2021

Supplementary Material for

A comparative study on removal of the dye wastewater by nascent state manganese dioxide and ferric hydroxide under acid condition

Yujia Tan [a], Zhiwen Cheng [a], Yawei Liu [a], Xiaoping Gao [a], Shiqiang Liu [a], Yuanyang Ren [a], Zhemin Shen [a,b,c*]

a. School of Environmental Science and Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, China.

b. Shanghai Institute of Pollution Control and Ecological Security, Shanghai, 200092, China.

c. Department of Developmental and Behavioral Pediatrics, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, 200127, P. R China

***Corresponding author contact.**

Zhemin Shen

Dongchuan Road, Minhang District, Shanghai, 200240, China.

Tel: +86-21-54745262, Fax: +86-21-54742863

Email: zmshen@sjtu.edu.cn

Totally 25 pages including 3 Text, 10 Tables.

*Corresponding author. Tel.: +86 21 54741065
E-mail address: zmshen@sjtu.edu.cn.

Text S1 Calculation details of parameters

A total of 20 quantum chemical parameters were calculated by using Gaussian 09 and Material Studio 7.0 based on the density functional theory (DFT). In Gaussian 09, the calculation method and basis set were B3LYP and 6-311G (d,p), respectively. The parameters included the total energy of a molecule (E_{B3LYP}), dipole moment (μ), energy of the highest occupied molecular orbital (E_{HOMO}), energy of the lowest unoccupied molecular orbital (E_{LUMO}), the gap energy (E_{GAP}), the most positive partial charge on a hydrogen atom (qH^+), negative partial charge on a carbon atom (qC^-), and positive partial charge on a hydrogen atom linked with a carbon atom ($q(CH)^+$). In Material Studio 7.0, the B3LPY method was selected to match the results of Gaussian 09. The maximum force, displacement and iterations were set as 0.004 Ha/Å, 500 and 0.3 Å, respectively. The SCF tolerance was set to 10-6 a.u. The bond order values (BO_n, BO_x) and Fukui indexes ($f(+)_x/_n$, $f(-)_x/_n$, $f(0)_x/_n$) were obtained from the out mol files. In addition, structural information to distinguish compound properties was also recorded, such as the relative molecular mass (MM) of organic compounds, the number of carbon atoms (N_C), the number of hydrogen atoms (N_H), the number of oxygen atoms (N_O), the number of nitrogen atoms (N_O), and the ratio of oxygen and carbon atoms ($N_{O:C}$), the ratio of carbon and hydrogen atoms ($N_{C:H}$), the ratio of nitrogen and carbon atoms ($N_{N:C}$) in a molecule.

Text S2 Y test procedure

First, the dependent variables (R_{exp}) were rearranged in the training set twenty times randomly. Then the original independent variable matrix was kept unchanged to get 20 new models. Finally, the R^2 and Q_{INT}^2 values of the 20 new models were obtained according to the method mentioned above. If R^2 and Q_{INT}^2 of the new models are smaller than that of the original model, it means that the original model has good validity and robustness.

Text S3 Applicability domain

According to the OECD (Organisation for Economic Co-operation and Development) principles, a QSAR model should have a defined domain of applicability [1]. The domain of application for a QSAR model describes whether the

model will predict an endpoint for a specific chemical with a given reliability. It helps the users of the model to judge whether the prediction for a new chemical is reliable or not, the predictions for only those compounds that fall into this domain could be considered reliable [2]. In this study, the applicability domain (APD) was visualized by using the Williams plot [3]. It helps estimate whether the compound is in the applicability domain according to the standardized residuals (σ) and leverage (h^*) values, and then judge the accuracy of the prediction [4, 5]. The standardized residual (σ) is defined as (S-E1).

$$\sigma = \frac{(y_i - \tilde{y}_i)}{\sqrt{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2 / (n-1)}} \quad \text{(S-E1)}$$

Where y_i is the R_{exp} value of the ith compound and $\tilde{y}_i$ is the R_{pre} value; n is the number of compounds in the training set. Leverage can be calculated as (S-E2).

$$h_i = x_i(X^T X)^{-1} x_i^T (i = 1,2,3.....n) \quad \text{(S-E2)}$$

Where x_i is the vector of descriptor for the ith compound while *X* is the descriptor matrix from the training set. The warning leverage (h^*) is defined as (S-E3).

$$h^* = 3(m+1)/n \quad \text{(S-E3)}$$

Where m is the number of predictor variables, and n is the number of training compounds.

Table S1 The dyes used in the experiment.

No.	Dyes	Abbreviation	Molecular formula	CAS No.	Molecular mass	Purity	Source
1[b]	Tris 4-aminophenyl Methanol	T4-AM	$C_{19}H_{19}N_3O$	467-62-9	305.37	AR	Sinopharm Chemical Reagent Co., Ltd.
2	Methylene Blue Trihydrate	MBT	$C_{16}H_{24}ClN_3O_3S$	7220-79-3	373.90	AR	Sinopharm Chemical Reagent Co., Ltd.
3	Azure B	AB	$C_{15}H_{16}ClN_3S$	531-55-5	305.83	AR	Sinopharm Chemical Reagent Co., Ltd.
4	Curcumin	Curcumin	$C_{21}H_{20}O_6$	458-37-7	368.38	AR	Sinopharm Chemical Reagent Co., Ltd.
5[a]	Basic Blue 26	BB26	$C_{33}H_{32}N_3Cl$	2580-56-5	506.09	AR	Sinopharm Chemical Reagent Co., Ltd.
6[ab]	Mordant Yellow 1	MY1	$C_{13}H_8N_3NaO_5$	584-42-9	309.21	AR	Sinopharm Chemical Reagent Co., Ltd.
7[a]	Crystal Violet	CV	$C_{25}H_{30}ClN_3$	548-62-9	407.99	AR	Sinopharm Chemical Reagent Co., Ltd.
8	Acid Orange 7	AO7	$C_{16}H_{11}N_2NaO_4S$	633-96-5	350.32	AR	Sinopharm Chemical Reagent Co., Ltd.
9	Acid Orange 10	AO10	$C_{16}H_{10}N_2Na_2O_7S_2$	1936-15-8	452.37	AR	Sinopharm Chemical Reagent Co., Ltd.
10	Methyl Red	MR	$C_{15}H_{15}N_3O_2$	493-52-7	269.30	AR	Sinopharm Chemical Reagent Co., Ltd.
11[b]	Fuchsin Basic	FB	$C_{20}H_{20}ClN_3$	3248-93-9	301.38	AR	Sinopharm Chemical Reagent Co., Ltd.
12[a]	Mordant Black 17	MB17	$C_{20}H_{13}N_2O_5SNa$	2538-85-4	416.39	AR	Sinopharm Chemical Reagent Co., Ltd.
13	Isatin	Isatin	$C_8H_5NO_2$	91-56-5	147.13	AR	Sinopharm Chemical Reagent Co., Ltd.
14	Acid Sky Blue A	ASBA	$C_{37}H_{35}N_2NaO_6S_2$	3486-30-4	690.80	AR	Sinopharm Chemical Reagent Co., Ltd.
15	Acid Chrome Blue K	ACBK	$C_{16}H_9N_2Na_3O_{12}S_3$	3270-25-5	586.40	AR	Sinopharm Chemical Reagent Co., Ltd.
16[b]	Eriochrome Black T	EBT	$C_{20}H_{12}N_3NaO_7S$	1787-61-7	461.38	AR	Sinopharm Chemical Reagent Co., Ltd.
17	Rhodamine B	RB	$C_{28}H_{31}ClN_2O_3$	81-88-9	479.01	AR	Sinopharm Chemical Reagent Co., Ltd.
18[a]	Carmine	Carmine	$C_{22}H_{20}O_{13}$	1390-65-4	492.39	AR	Sinopharm Chemical Reagent Co., Ltd.
19	Metanil Yellow	MY	$C_{18}H_{14}N_3NaO_3S$	587-98-4	375.38	AR	Sinopharm Chemical Reagent Co., Ltd.
20	Cresol Red	CR	$C_{21}H_{17}NaO_5S$	1733-12-6	382.43	AR	Sinopharm Chemical Reagent Co., Ltd.
21[b]	Methyl Orange	MO	$C_{14}H_{14}N_3NaO_3S$	547-58-0	327.33	AR	Sinopharm Chemical Reagent Co., Ltd.
22	Reactive Black 5	RB5	$C_{22}H_{16}N_2O_{11}S_3Na_2$	12225-25-1	626.55	AR	Sinopharm Chemical Reagent Co., Ltd.
23	Bromophenol Blue	BB	$C_{19}H_{10}Br_4O_5S$	115-39-9	669.96	AR	Sinopharm Chemical Reagent Co., Ltd.
24[a]	Fuchsin Acid	FA	$C_{20}H_{17}N_3Na_2O_9S_3$	3244-88-0	585.54	AR	Sinopharm Chemical Reagent Co., Ltd.

25	Brilliant Red X 3B	BRX-3B	$C_{19}H_{10}Cl_2N_6Na_2O_7S_2$	17804-49-8	615.33	AR	Sinopharm Chemical Reagent Co., Ltd.
26[b]	Acid Black 1	AB1	$C_{22}H_{14}N_6Na_2O_9S_2$	1064-48-8	616.49	AR	Sinopharm Chemical Reagent Co., Ltd.
27	Ponceau S	PS	$C_{22}H_{12}N_4Na_4O_{13}S_4$	6226-79-5	760.57	AR	Sinopharm Chemical Reagent Co., Ltd.
28	Acid Blue 93	AB93	$C_{37}H_{27}N_3Na_2O_9S_3$	28983-56-4	799.80	AR	Sinopharm Chemical Reagent Co., Ltd.
29	Cresol Purple	CP	$C_{21}H_{18}O_5S$	2303-01-7	382.43	AR	Sinopharm Chemical Reagent Co., Ltd.
30[a]	Indigo	Indigo	$C_{16}H_{10}N_2O_2$	482-89-3	262.26	AR	Sinopharm Chemical Reagent Co., Ltd.
31[b]	Basic Brown 1	BB1	$C_{18}H_{20}Cl_2N_8$	10114-58-6	419.31	AR	Sinopharm Chemical Reagent Co., Ltd.
32	Acid Green 1	AG1	$C_{30}H_{15}FeN_3Na_3O_{15}S_3$	19381-50-1	878.45	AR	Sinopharm Chemical Reagent Co., Ltd.
33	Acid Orange 74	AO74	$C_{16}H_{12}N_5NaO_7S$	10127-27-2	441.35	AR	Sinopharm Chemical Reagent Co., Ltd.
34	Acid Scarlet	AS	$C_{18}H_{14}N_2Na_2O_7S_2$	3761-53-3	480.42	AR	Sinopharm Chemical Reagent Co., Ltd.
35	Direct Red 28	DR28	$C_{32}H_{22}N_6Na_2O_6S_2$	573-58-0	696.66	AR	Sinopharm Chemical Reagent Co., Ltd.
36[b]	Disperse Red 16	DR16	$C_{24}H_{26}IN_5O_7$	61968-52-3	519.94	AR	Sinopharm Chemical Reagent Co., Ltd.
37[a]	Bromocresol Green	BG	$C_{21}H_{14}Br_4O_5S$	76-60-8	698.01	AR	Sinopharm Chemical Reagent Co., Ltd.
38	Reactive Yellow 3	RY3	$C_{27}H_{18}ClN_8Na_3O_{10}S_3$	6539-67-9	815.09	AR	Sinopharm Chemical Reagent Co., Ltd.

[a]: Samples in the external test set at pH=4, NSMD.

[b]: Samples in the external test set at pH=4, NSFH.

Table S2 The inorganic compounds used in the experiment.

No.	Compounds	Molecular formula	CAS No.	Molecular mass	Purity	Source
1	Anhydrous Sodium Sulfate	Na_2SO_3	7757-83-7	126.04	AR	Sinopharm Chemical Reagent Co., Ltd.
2	Potassium Permanganate	$KMnO_4$	7722-64-7	158.03	AR	Sinopharm Chemical Reagent Co., Ltd.
3	Ferric Chloride	$FeCl_3$	7705-08-0	162.20	AR	Sinopharm Chemical Reagent Co., Ltd.
4	Sodium Hydroxide	NaOH	1310-73-2	40.00	AR	Sinopharm Chemical Reagent Co., Ltd.
5	Sulfuric Acid	H_2SO_4	7664-93-9	98.08	AR	Sinopharm Chemical Reagent Co., Ltd.

Table S3 8 structural parameters and 20 quantum chemical parameters used in this study.

No.	Descriptor	Note	Unit
1	MM	Relative molecular mass	g/mol
2	N_C	The number of carbon atom	-
3	N_H	The number of hydrogen atom	-
4	N_O	The number of oxygen atom	-
5	N_N	The number of nitrogen atom	-
6	$N_{O:C}$	The ratio of oxygen atom to carbon atom	-
7	$N_{C:H}$	The ratio of carbon atom to hydrogen atom	-
8	$N_{N:C}$	The ratio of nitrogen atom to carbon atom	-
9	$E_{(B3LYP)}$	The total energy of a molecule (E_{B3LYP})	kcal/mol
10	μ	Dipole moment	Debye
11	qH^+	The most positive partial charge on a hydrogen atom	e
12	$q(CH^+)_n$	Minimum value of positive partial charge on a hydrogen atom linked with a carbon atom	e
13	$q(CH^+)_x$	Maximum value of positive partial charge on a hydrogen atom linked with a carbon atom	e
14	$q(C^-)_n$	Minimum value of negative partial charge on a carbon atom	e
15	$q(C^-)_x$	Maximum value of negative partial charge on a carbon atom	e
16	E_{HOMO}	Energy of the highest occupied molecular orbital	eV
17	E_{LUMO}	Energy of the lowest unoccupied molecular orbital	eV
18	E_{SUM}	Sum of E_{LUMO} and E_{HOMO}, $E_{SUM}=E_{LUMO}+E_{HOMO}$	eV
19	E_{GAP}	Gap of E_{LUMO} and E_{HOMO}, $E_{GAP}=E_{LUMO}-E_{HOMO}$	eV
20	E_{GAP}^2	The square of E_{GAP}	-
21	BO_n	Minimum value of bond order	-
22	BO_x	Maximum value of bond order	-
23	$f(+)_n$	Minimum value of nucleophilic Fukui index	e
24	$f(+)_x$	Maximum value of nucleophilic Fukui index	e

25	$f(-)_n$	Minimum value of electrophilic Fukui index	e
26	$f(-)_x$	Maximum value of electrophilic Fukui index	e
27	$f(0)_n$	Minimum value of free radical Fukui index	e
28	$f(0)_x$	Maximum value of free radical Fukui index	e

Table S4 The structural parameters of 38 dyes.

No.	Dyes	Mm	N_C	N_H	N_O	N_N	$N_{O:C}$	$N_{C:H}$	$N_{N:C}$
1	T4-AM	305.37	19	19	1	3	0.0526	1.0000	0.1579
2	MBT	373.90	16	24	3	3	0.1875	0.6667	0.1875
3	AB	305.83	15	16	0	3	0.0000	0.9375	0.2000
4	Curcumin	368.38	21	20	6	0	0.2857	1.0500	0.0000
5	BB26	506.09	33	32	0	3	0.0000	1.0313	0.0909
6	MY1	309.21	13	8	5	3	0.3846	1.6250	0.2308
7	CV	407.99	25	30	0	3	0.0000	0.8333	0.1200
8	AO7	350.32	16	11	4	2	0.2500	1.4545	0.1250
9	AO10	452.37	16	10	7	2	0.4375	1.6000	0.1250
10	MR	269.30	15	15	2	3	0.1333	1.0000	0.2000
11	FB	301.38	20	20	0	3	0.0000	1.0000	0.1500
12	MB17	416.39	20	13	5	2	0.2500	1.5385	0.1000
13	Isatin	147.13	8	5	2	1	0.2500	1.6000	0.1250
14	ASBA	690.80	37	35	6	2	0.1622	1.0571	0.0541
15	ACBK	586.40	16	10	7	2	0.4375	1.6000	0.1250
16	EBT	461.38	20	12	7	3	0.3500	1.6667	0.1500
17	RB	479.01	28	31	3	2	0.1071	0.9032	0.0714
18	Carmine	492.39	22	20	13	0	0.5909	1.1000	0.0000
19	MY	375.38	18	14	3	3	0.1667	1.2857	0.1667
20	CR	382.43	21	17	5	0	0.2381	1.2353	0.0000
21	MO	327.33	14	14	3	3	0.2143	1.0000	0.2143
22	RB5	626.55	22	16	11	2	0.5000	1.3750	0.0909
23	BB	669.96	19	10	5	0	0.2632	1.9000	0.0000
24	FA	585.54	20	17	9	3	0.4500	1.1765	0.1500

25	BRX-3B	615.33	19	10	7	6	0.3684	1.9000	0.3158
26	AB1	616.49	22	14	9	6	0.4091	1.5714	0.2727
27	PS	760.57	22	12	13	4	0.5909	1.8333	0.1818
28	AB93	799.80	37	27	9	3	0.2432	1.3704	0.0811
29	CP	382.43	21	18	5	0	0.2381	1.1667	0.0000
30	Indigo	262.26	16	10	2	2	0.1250	1.6000	0.1250
31	BB1	419.31	18	20	0	8	0.0000	0.9000	0.4444
32	AG1	878.45	30	15	15	3	0.5000	2.0000	0.1000
33	AO74	441.35	16	12	7	5	0.4375	1.3333	0.3125
34	AS	480.42	18	14	7	2	0.3889	1.2857	0.1111
35	DR28	696.66	32	22	6	6	0.1875	1.4545	0.1875
36	DR16	519.94	24	26	7	5	0.2917	0.9231	0.2083
37	BG	698.01	21	14	5	0	0.2381	1.5000	0.0000
38	RY3	815.09	27	18	10	8	0.3704	1.5000	0.2963

Table S5-1 The Quantum chemical descriptors of 38 dyes.

No.	Dyes	$E_{(B3LYP)}$	μ	qH^+	$q(CH^+)_n$	$q(CH^+)_X$	$q(C^-)_n$	$q(C^-)_X$	E_{HOMO}	E_{LUMO}	E_{SUM}
		(a.u)	(Debye)	(e)	(e)	(e)	(e)	(e)	(eV)	(eV)	(eV)
1	T4-AM	-975.1907	3.1488	0.457	0.194	0.218	-0.257	0.285	-2.3021	-2.6068	-4.9089
2	MBT	-1643.4005	12.0854	0.290	0.191	0.290	-0.366	0.256	-3.7279	-1.2844	-5.0123
3	AB	-1604.0848	11.8733	0.384	0.187	0.236	-0.367	0.306	-5.2272	-3.4667	-8.6939
4	Curcumin	-1263.8883	3.3578	0.480	0.169	0.219	-0.353	0.484	-6.1225	0.4054	-5.7170
5	BB26	-1902.8361	9.3990	0.419	0.175	0.252	-0.344	0.199	-6.1442	0.1497	-5.9946
6	MY1	-1203.0840	7.7001	0.457	0.198	0.250	-0.297	0.791	-6.1334	-3.5538	-9.6871
7	CV	-1595.4592	4.0784	0.219	0.174	0.219	-0.350	0.193	-6.2912	1.1456	-5.1456
8	AO7	-1587.4971	7.9904	0.489	0.200	0.230	-0.251	0.387	-6.6123	-1.2762	-7.8885
9	AO10	-2373.1568	14.6745	0.491	0.211	0.266	-0.289	0.406	-6.8545	-2.3646	-9.2191
10	MR	-895.5266	6.1273	0.474	0.185	0.226	-0.351	0.818	-6.5715	-0.2694	-6.8408
11	FB	-1398.9075	22.1782	0.420	0.198	0.265	-0.596	0.271	-3.4585	-3.7769	-7.2354
12	MB17	-1816.4024	6.8500	0.490	0.198	0.240	-0.254	0.381	-6.4408	-0.9660	-7.4068
13	Isatin	-513.1963	5.9116	0.409	0.206	0.219	-0.254	0.622	-7.3687	0.0952	-7.2735
14	ASBA	-2790.3819	22.4467	0.483	0.179	0.252	-0.600	0.182	-4.0789	-0.3020	-4.3810
15	ACBK	-3309.2926	5.2536	0.503	0.202	0.239	-0.282	0.406	-3.2109	-2.8463	-6.0572
16	EBT	-2022.1368	11.5672	0.476	0.188	0.244	-0.252	0.447	-6.1987	-0.5170	-6.7157
17	RB	-1881.4300	8.7880	0.482	0.190	0.280	-0.581	0.438	-6.2340	-0.3270	-6.5610
18	Carmine	-1828.8050	8.2625	0.510	0.148	0.238	-0.596	0.814	-6.6504	-3.7306	-10.3810
19	MY	-1645.0402	8.6591	0.383	0.199	0.238	-0.2064	0.201	-6.7130	-0.4027	-7.1157
20	CR	-1585.5559	4.9301	0.467	0.201	0.224	-0.60-oiu5	0.346	-6.7647	0.2993	-6.4653
21	MO	-1492.4559	8.0654	0.230	0.186	0.230	-0.351	0.211	-6.6667	-0.4354	-7.1021
22	RB5	-3286.1672	8.8369	0.441	0.228	0.277	-0.591	0.477	-4.3565	-5.4449	-9.8014
23	BB	-11801.1500	7.6560	0.479	0.212	0.245	-0.257	0.290	-6.8600	-1.8510	-8.7110

24	FA	-3133.2784	2.5151	0.499	0.203	0.239	-0.599	0.225	-3.1891	-3.3116	-6.5007
25	BRX-3B	-3626.9843	12.4315	0.502	0.206	0.243	-0.231	0.656	-3.5701	-3.1755	-6.7456
26	AB1	-2973.6926	10.0412	0.428	0.210	0.246	-0.289	0.491	-7.0449	-3.8395	-10.8844
27	PS	-4285.0131	12.4768	0.437	0.215	0.243	-0.305	0.440	-2.8844	-4.2313	-7.1157
28	AB93	-3787.2394	13.6871	0.499	0.205	0.243	-0.327	0.244	-7.1538	-0.8435	-7.9973
29	CP	-1585.6311	4.7888	0.463	0.196	0.236	-0.592	0.356	-6.8000	0.0408	-6.7592
30	Indigo	-875.9119	0.0002	0.434	0.203	0.215	-0.251	0.507	-7.0096	0.2068	-6.8028
31	BB1	-1135.0034	4.9925	0.423	0.192	0.224	-0.337	0.242	-5.1374	-1.9075	-7.0449
32	AG1	-5388.7400	16.5020	0.286	0.259	0.286	-0.222	0.429	-6.3700	-0.3810	-6.7510
33	AO74	-1977.9964	7.5116	0.492	0.208	0.266	-0.607	0.552	-6.9959	-3.4667	-10.4626
34	AS	-2451.7962	9.1499	0.499	0.191	0.227	-0.584	0.354	-6.4300	0.1742	-6.2558
35	DR28	-3134.0328	12.9468	0.396	0.192	0.262	-0.262	0.262	-6.4925	-0.5633	-7.0558
36	DR16	-2152.9718	10.9835	0.391	0.158	0.258	-0.692	0.852	-6.5823	-3.6871	-10.2694
37	BG	-11879.7800	6.1390	0.477	0.212	0.249	-0.598	0.294	-6.6970	-1.7970	-8.4940
38	RY3	-3032.0056	9.0962	0.491	0.210	0.401	-0.279	0.683	-3.2490	-3.2463	-6.4953

Table S5-2 The Quantum chemical descriptors of 38 dyes.

No.	Dyes	E_{GAP}	E_{GAP}^2	BO_n	BO_x	$f(+)_n$	$f(+)_X$	$f(-)_n$	$f(-)_X$	$f(0)_n$	$f(0)_X$
		(eV)	-	-	-	(e)	(e)	(e)	(e)	(e)	(e)
1	T4-AM	-0.3048	0.0929	0.9381	1.4235	0.001	0.049	0.004	0.080	0.003	0.051
2	MBT	2.4435	5.9709	0.9081	1.5532	0.003	0.110	0.011	0.084	0.009	0.078
3	AB	1.7606	3.0995	0.9080	1.5582	0.002	0.113	0.010	0.112	0.009	0.080
4	Curcumin	6.5279	42.6137	0.8652	1.8016	0.001	0.075	0.004	0.058	0.004	0.056
5	BB26	6.2939	39.6132	0.9107	1.5019	0.001	0.084	0.006	0.045	0.005	0.045
6	MY1	2.5796	6.6544	0.1034	1.6288	0.002	0.153	0.005	0.088	0.009	0.082
7	CV	7.4368	55.3055	0.9152	1.4864	0.003	0.089	0.007	0.051	0.007	0.048
8	AO7	5.3361	28.4737	0.1506	1.8294	0.005	0.113	0.004	0.077	0.004	0.077
9	AO10	4.4898	20.1584	0.1031	1.8270	0.002	0.096	0.004	0.074	0.003	0.073
10	MR	6.3021	39.7161	0.1311	1.8636	0.009	0.113	0.002	0.086	0.006	0.071
11	FB	-0.3184	0.1014	0.9467	1.4845	0.003	0.096	0.006	0.102	0.005	0.071
12	MB17	5.4749	29.9740	0.1448	1.8283	0.005	0.105	0.005	0.052	0.005	0.065
13	Isatin	7.4640	55.7110	0.8848	1.9188	0.026	0.164	0.020	0.119	0.035	0.124
14	ASBA	3.7769	14.2649	0.8928	1.8400	-0.002	0.089	-0.003	0.079	-0.002	0.053
15	ACBK	0.3646	0.1330	0.8272	1.8365	0.004	0.102	0.003	0.067	0.003	0.073
16	EBT	5.6817	32.2812	0.2292	1.8300	0.003	0.148	0.004	0.062	0.005	0.080
17	RB	5.9070	34.8926	0.9620	1.5300	-0.005	0.063	-0.007	0.038	-0.006	0.062
18	Carmine	2.9197	8.5249	0.1254	1.8978	-0.002	0.085	-0.003	0.062	0.002	0.063
19	MY	6.3102	39.8190	0.9382	1.8329	0.004	0.127	0.002	0.100	0.003	0.085
20	CR	7.0640	49.8998	0.8053	1.9669	-0.013	0.115	-0.010	0.069	0.002	0.064
21	MO	6.2313	38.8293	0.9328	1.8232	0.006	0.125	0.003	0.099	0.005	0.088
22	RB5	-1.0884	1.1847	0.7973	1.8947	0.000	0.087	0.000	0.082	0.000	0.063
23	BB	5.0090	25.0901	0.9500	1.4150	-0.008	0.095	-0.010	0.039	0.002	0.053

24	FA	-0.1224	0.0150	0.1005	1.9097	0.002	0.103	0.002	0.090	0.003	0.085
25	BRX-3B	0.3946	0.1557	0.2293	1.8445	0.007	0.084	0.001	0.160	0.005	0.088
26	AB1	3.2055	10.2749	0.1051	1.8392	0.003	0.078	0.004	0.062	0.003	0.047
27	PS	-1.3469	1.8143	0.1417	1.7902	0.001	0.087	0.002	0.136	0.002	0.073
28	AB93	6.3102	39.8190	0.8733	1.9361	0.000	0.084	0.002	0.063	0.001	0.064
29	CP	6.8408	46.7972	0.7973	1.9465	-0.009	0.101	-0.011	0.069	0.003	0.057
30	Indigo	7.2164	52.0758	0.9396	1.7672	0.012	0.096	0.008	0.071	0.015	0.080
31	BB1	3.2299	10.4325	0.8921	1.6153	0.003	0.078	0.007	0.048	0.006	0.053
32	AG1	5.9890	35.8681	0.6380	1.6190	-0.050	0.056	-0.030	0.057	-0.060	0.057
33	AO74	3.5293	12.4557	0.1305	1.8315	0.001	0.143	0.003	0.091	0.004	0.082
34	AS	6.6041	43.6143	0.8274	1.8532	0.003	0.114	0.004	0.068	0.004	0.090
35	DR28	5.9293	35.1563	0.9240	1.8159	0.002	0.077	0.004	0.053	0.003	0.052
36	DR16	2.8953	8.3825	0.2245	1.8366	0.000	0.129	0.001	0.099	0.001	0.074
37	BG	4.9000	24.0100	0.9540	1.4190	-0.012	0.071	-0.009	0.035	0.002	0.056
38	RY3	0.0027	0.0000	0.8418	1.9493	0.000	0.083	0.000	0.154	0.000	0.096

Table S6 The removal efficiency by NSMD and NSFH at pH=4.

No.	Dyes	λ_{max}	NSFH		NSMD	
			R_{exp}	error	R_{exp}	error
1	T4-AM	538	0.0054	±0.0014	0.8993	±0.0538
2	MBT	661	0.1065	±0.0209	0.9984	±0.0014
3	AB	605	0.1184	±0.0141	0.9315	±0.0238
4	Curcumin	264	0.1865	±0.0398	0.7891	±0.0481
5	BB26	615	0.2096	±0.0321	0.9623	±0.0218
6	MY1	352	0.2528	±0.0612	0.6956	±0.1241
7	CV	589	0.2657	±0.0475	0.9996	±0.0002
8	AO7	484	0.2743	±0.0398	0.8287	±0.0034
9	AO10	477	0.2876	±0.0022	0.8728	±0.0532
10	MR	265	0.2894	±0.0413	0.7328	±0.0734
11	FB	541	0.3154	±0.0053	0.9998	±0.0002
12	MB17	438	0.3170	±0.0374	0.8129	±0.0193
13	Isatin	301	0.3178	±0.0572	0.6521	±0.0323
14	ASBA	600	0.3623	±0.0133	0.9856	±0.0108
15	ACBK	542	0.3654	±0.0322	0.7312	±0.0165
16	EBT	614	0.3678	±0.0793	0.9461	±0.0485
17	RB	553	0.3765	±0.0281	0.9732	±0.0211
18	Carmine	506	0.3949	±0.0133	0.7621	±0.0681
19	MY	444	0.3975	±0.0740	0.8864	±0.1120
20	CR	435	0.4316	±0.0290	0.7087	±0.0243
21	MO	463	0.4373	±0.0633	0.7844	±0.0293
22	RB5	597	0.4537	±0.0276	0.6521	±0.0198
23	BB	437	0.4740	±0.0557	0.8756	±0.0281

24	FA	544	0.4765	±0.0032	0.5431	±0.0219
25	BRX-3B	537	0.4985	±0.0465	0.8533	±0.0593
26	AB1	618	0.5090	±0.0355	0.7983	±0.0298
27	PS	506	0.5342	±0.0576	0.7966	±0.0568
28	AB93	626	0.5437	±0.0215	0.8624	±0.0281
29	CP	433	0.5673	±0.0475	0.6733	±0.0438
30	Indigo	609	0.5673	±0.0086	0.8943	±0.0548
31	BB1	462	0.5846	±0.0681	0.8477	±0.0981
32	AG1	285	0.6358	±0.0772	0.9942	±0.0027
33	AO74	479	0.6754	±0.0391	0.6425	±0.0472
34	AS	510	0.6855	±0.0281	0.9136	±0.0812
35	DR28	599	0.7266	±0.0387	0.8854	±0.0483
36	DR16	462	0.7469	±0.0714	0.7545	±0.0958
37	BG	616	0.7544	±0.0775	0.8645	±0.0412
38	RY3	424	0.8674	±0.0576	0.9576	±0.0281
Average value	-	-	0.4311	±0.0417	0.8358	±0.0596

Table S7 The correlation coefficients between the value of R_{exp} and the 28 selected parameters.

No.	Parameters	Pearson Correlation coefficient	
		NSMD	NSFH
1	BO_x	-0.610**	0.331
2	$N_{O:C}$	-0.482**	0.364*
3	$q(C^-)_x$	-0.478**	0.184
4	$f(0)_x$	-0.411**	0.084
5	$f(-)_x$	-0.398*	0.074
6	qH^+	-0.357*	0.130
7	N_O	-0.342*	0.481**
8	$f(+)_x$	-0.300	-0.076
9	$q(CH^+)_x$	-0.246	0.381*
10	$f(0)_n$	-0.231	-0.262
11	$f(+)_n$	-0.201	-0.270
12	$N_{N:C}$	-0.168	0.179
13	$N_{C:H}$	-0.167	0.356*
14	N_N	-0.116	0.356*
15	$q(CH^+)_n$	-0.112	0.289
16	$E_{(B3LPY)}$	-0.074	-0.399*
17	E_{HOMO}	-0.070	-0.181
18	$f(-)_n$	-0.050	-0.383*
19	MM	-0.033	0.583**
20	μ	0.432**	0.043
21	N_H	0.378*	-0.085
22	E_{LUMO}	0.372*	-0.116
23	BO_n	0.334	-0.035
24	E_{SUM}	0.322	-0.286
25	N_C	0.276	0.273
26	E_{GAP}	0.268	0.027
27	E_{GAP}^2	0.204	0.018
28	$q(C^-)_n$	0.163	-0.245

Note: '**' and '*' indicate that when the confidence level are 0.01 and 0.05, the correlation is statistically significant.

Table S8 Checking statistical data of models at pH=4.

Model No.	Variables	Regression coefficients	| t |	Sig.	VIF
M-1	BO_x	-0.422	5.408	0.000	1.024
	μ	0.016	5.794	0.000	1.126
	$q(CH^+)_x$	-1.589	4.233	0.000	1.106
M-2	BO_x	-0.460	7.031	0.000	1.051
	μ	0.017	7.210	0.000	1.130
	$q(CH^+)_x$	-1.248	3.842	0.001	1.209
	E_{LUMO}	0.024	3.601	0.001	1.147
M-3	BO_x	-0.426	6.183	0.000	1.207
	μ	0.017	7.440	0.000	1.141
	$q(CH^+)_x$	-1.185	3.680	0.001	1.233
	E_{LUMO}	0.020	2.752	0.011	1.390
	$f(-)_x$	-0.553	1.385	0.179	1.440
M-4	BO_x	-0.368	4.860	0.000	1.556
	μ	0.016	7.364	0.000	1.166
	$q(CH^+)_x$	-1.283	4.041	0.001	1.279
	E_{LUMO}	0.037	2.914	0.008	4.639
	$f(-)_x$	-0.682	1.729	0.097	1.501
	E_{GAP}^2	-0.002	1.621	0.119	5.778
F-1	MM	0.001	7.362	0.000	1.832
	μ	-0.016	3.013	0.006	1.363
	$f(0)_x$	5.255	3.679	0.001	1.266
	E_{HOMO}	-0.058	3.355	0.003	1.338
F-2	MM	0.001	7.950	0.000	1.836
	μ	-0.015	2.996	0.006	1.376

	$f(0)_x$	5.865	4.285	0.000	1.325
	E_{HOMO}	-0.060	3.739	0.001	1.347
	$q(C^-)_n$	-0.291	2.117	0.045	1.057
F-3	MM	0.001	7.875	0.000	1.990
	μ	-0.017	3.759	0.001	1.404
	$f(0)_x$	4.674	3.649	0.001	1.490
	E_{HOMO}	-0.072	4.848	0.000	1.460
	$q(C^-)_n$	-0.413	3.205	0.004	1.194
	N_N	0.034	2.797	0.010	1.781
Criterion	-	-	>2.045	<0.05	<5

Table S9 The experimental and predicted values of removal rate by NSMD.

No.	Dyes	R_{exp}	Model M-1		Model M-2		Model M-3		Model M-4	
			R_{pre}	Diff.	R_{pre}	Diff.	R_{pre}	Diff.	R_{pre}	Diff.
1	T4-AM	0.8993	0.4967	0.0391	0.5086	0.0510	0.4880	0.0304	0.4916	0.0340
2	MBT	0.9984	0.6656	0.1225	0.6155	0.0724	0.6190	0.0759	0.6499	0.1068
3	AB	0.9315	0.7356	0.0931	0.6990	0.0565	0.7016	0.0591	0.6927	0.0502
4	Curcumin	0.7891	0.7479	0.0958	0.7758	0.1237	0.7487	0.0966	0.7214	0.0693
5	BB26	0.9623	0.7127	0.0606	0.6313	-0.0208	0.6496	-0.0025	0.6320	-0.0201
6	MY1	0.6956	0.6912	0.0179	0.7215	0.0482	0.7242	0.0509	0.7214	0.0481
7	CV	0.9996	0.8496	0.1540	0.8134	0.1178	0.8100	0.1144	0.8012	0.1056
8	AO7	0.8287	0.7039	-0.0048	0.7357	0.0270	0.7373	0.0286	0.7349	0.0262
9	AO10	0.8728	0.7403	0.0091	0.7069	-0.0243	0.7188	-0.0124	0.7533	0.0221
10	MR	0.7328	0.7635	0.0307	0.7874	0.0546	0.7785	0.0457	0.7772	0.0444
11	FB	0.9998	0.8017	0.0472	0.7604	0.0059	0.7591	0.0046	0.7511	-0.0034
12	MB17	0.8129	0.7641	0.0020	0.7099	-0.0522	0.7301	-0.0320	0.7341	-0.0280
13	Isatin	0.6521	0.8052	0.0208	0.8300	0.0456	0.8134	0.0290	0.8047	0.0203
14	ASBA	0.9856	0.7565	-0.0326	0.7938	0.0047	0.7951	0.0060	0.8030	0.0139
15	ACBK	0.7312	0.8690	0.0724	0.8128	0.0162	0.7907	-0.0059	0.7791	-0.0175
16	EBT	0.9461	0.8046	0.0063	0.7545	-0.0438	0.7736	-0.0247	0.7663	-0.0320
17	RB	0.9732	0.7677	-0.0452	0.7817	-0.0312	0.7941	-0.0188	0.8007	-0.0122
18	Carmine	0.7621	0.8014	-0.0273	0.8056	-0.0231	0.8049	-0.0238	0.8059	-0.0228
19	MY	0.8864	0.8533	0.0056	0.8455	-0.0022	0.8556	0.0079	0.8769	0.0292
20	CR	0.7087	0.8454	-0.0079	0.8124	-0.0409	0.7746	-0.0787	0.7844	-0.0689
21	MO	0.7844	0.8268	-0.0356	0.8476	-0.0148	0.8572	-0.0052	0.8439	-0.0185
22	RB5	0.6521	0.9147	0.0502	0.9267	0.0622	0.9385	0.0740	0.9212	0.0567
23	BB	0.8756	0.8521	-0.0207	0.8493	-0.0235	0.8567	-0.0161	0.8459	-0.0269

24	FA	0.5431	0.9471	0.0715	0.9581	0.0825	0.9675	0.0919	0.9452	0.0696
25	BRX-3B	0.8533	0.8355	-0.0499	0.8733	-0.0121	0.8845	-0.0009	0.8785	-0.0069
26	AB1	0.7983	0.7979	-0.0885	0.8264	-0.0600	0.8100	-0.0764	0.7989	-0.0875
27	PS	0.7966	0.7236	-0.1707	0.7527	-0.1416	0.7463	-0.1480	0.7319	-0.1624
28	AB93	0.8624	0.9143	0.0150	0.8931	-0.0062	0.8814	-0.0179	0.8986	-0.0007
29	CP	0.6733	0.8146	-0.0990	0.8530	-0.0606	0.8520	-0.0616	0.8490	-0.0646
30	Indigo	0.8943	0.9684	0.0369	0.9363	0.0048	0.9161	-0.0154	0.9059	-0.0256
31	BB1	0.8477	0.8361	-0.1100	0.8669	-0.0792	0.8723	-0.0738	0.8756	-0.0705
32	AG1	0.9942	0.9272	-0.0351	0.9870	0.0247	0.9885	0.0262	0.9730	0.0107
33	AO74	0.6425	0.8610	-0.1122	0.9173	-0.0559	0.9273	-0.0459	0.9135	-0.0597
34	AS	0.9136	0.9932	0.0076	1.0424	0.0568	1.0384	0.0528	1.0681	0.0825
35	DR28	0.8854	0.9374	-0.0568	0.9987	0.0045	1.0018	0.0076	0.9796	-0.0146
36	DR16	0.7545	0.8881	-0.1103	0.9272	-0.0712	0.9170	-0.0814	0.9360	-0.0624
37	BG	0.8645	0.9010	-0.0986	0.9688	-0.0308	0.9603	-0.0393	0.9373	-0.0623
38	RY3	0.4576	1.1183	0.1185	1.1018	0.1020	1.0877	0.0879	1.0620	0.0622

Table S10 The experimental and predicted values of removal rate by NSFH.

No.	Dyes	R_{exp}	Model F-1		Model F-2		Model F-3	
			R_{pre}	Diff.	R_{pre}	Diff.	R_{pre}	Diff.
1	T4-AM	0.0054	-0.0555	-0.0609	-0.1318	-0.1372	-0.0849	-0.0903
2	MBT	0.1065	0.0946	-0.0119	0.0783	-0.0282	0.1056	-0.0009
3	AB	0.1184	0.1274	0.0090	0.1154	-0.0030	0.1588	0.0404
4	Curcumin	0.1865	0.2520	0.0655	0.2145	0.0280	0.2106	0.0241
5	BB26	0.2096	0.2365	0.0269	0.1958	-0.0138	0.2941	0.0845
6	MY1	0.2528	0.2607	0.0079	0.2271	-0.0257	0.2788	0.0260
7	CV	0.2657	0.2479	-0.0178	0.2057	-0.0600	0.3135	0.0478
8	AO7	0.2743	0.2986	0.0243	0.2498	-0.0245	0.2731	-0.0012
9	AO10	0.2876	0.2868	-0.0008	0.2538	-0.0338	0.2760	-0.0116
10	MR	0.2894	0.2135	-0.0759	0.1882	-0.1012	0.2681	-0.0213
11	FB	0.3154	-0.1918	-0.5072	-0.1359	-0.4513	-0.0956	-0.4110
12	MB17	0.3170	0.3099	-0.0071	0.2532	-0.0638	0.2914	-0.0256
13	Isatin	0.3178	0.4196	0.1018	0.3998	0.0820	0.3467	0.0289
14	ASBA	0.3623	0.1347	-0.2276	0.1823	-0.1800	0.2174	-0.1449
15	ACBK	0.3654	0.3602	-0.0052	0.3085	-0.0569	0.3049	-0.0605
16	EBT	0.3678	0.3442	-0.0236	0.3003	-0.0675	0.3420	-0.0258
17	RB	0.3765	0.3138	-0.0627	0.3519	-0.0246	0.4272	0.0507
18	Carmine	0.3949	0.3650	-0.0299	0.4084	0.0135	0.4224	0.0275
19	MY	0.3975	0.3609	-0.0366	0.3216	-0.0759	0.3708	-0.0267
20	CR	0.4316	0.3202	-0.1114	0.3638	-0.0678	0.3857	-0.0459
21	MO	0.4373	0.3354	-0.1019	0.3226	-0.1147	0.3795	-0.0578
22	RB5	0.4537	0.3569	-0.0968	0.3949	-0.0588	0.4475	-0.0062
23	BB	0.4740	0.5119	0.0379	0.4504	-0.0236	0.4386	-0.0354

24	FA	0.4765	0.4649	-0.0116	0.5100	0.0335	0.5701	0.0936
25	BRX-3B	0.4985	0.3739	-0.1246	0.3244	-0.1741	0.4228	-0.0757
26	AB1	0.5090	0.3994	-0.1096	0.3463	-0.1627	0.5471	0.0381
27	PS	0.5342	0.3998	-0.1344	0.3614	-0.1728	0.4103	-0.1239
28	AB93	0.5437	0.6200	0.0763	0.5922	0.0485	0.6694	0.1257
29	CP	0.5673	0.2877	-0.2796	0.3232	-0.2441	0.3525	-0.2148
30	Indigo	0.5673	0.3772	-0.1901	0.3231	-0.2442	0.3635	-0.2038
31	BB1	0.5846	0.2039	-0.3807	0.1596	-0.4250	0.4142	-0.1704
32	AG1	0.6358	0.5714	-0.0644	0.5100	-0.1258	0.5677	-0.0681
33	AO74	0.6754	0.4458	-0.2296	0.5040	-0.1714	0.6723	-0.0031
34	AS	0.6855	0.4679	-0.2176	0.5248	-0.1607	0.5687	-0.1168
35	DR28	0.7266	0.4273	-0.2993	0.3712	-0.3554	0.5503	-0.1763
36	DR16	0.7469	0.4028	-0.3441	0.4835	-0.2634	0.6598	-0.0871
37	BG	0.7544	0.5705	-0.1839	0.6082	-0.1462	0.6355	-0.1189
38	RY3	0.8674	0.6505	-0.2169	0.6158	-0.2516	0.7813	-0.0861

References

[1] OECD, Guidance Document On The Validation of (Quantitative) Structure-activity Relationships [(Q)SAR] Models., (2007).

[2] G. Melagraki, A. Afantitis, Enalos InSilicoNano platform: an online decision support tool for the design and virtual screening of nanoparticles, RSC Advances, 4 (2014) 50713-50725.

[3] P. Gramatica, Principles of QSAR models validation: internal and external, QSAR & Combinatorial Science, 26 (2007) 694-701.

[4] S. Gupta, N. Basant, Modeling the pH and temperature dependence of aqueousphase hydroxyl radical reaction rate constants of organic micropollutants using QSPR approach, Environmental Science and Pollution Research, 24 (2017) 24936-24946.

[5] E.V. Ortiz, D.O. Bennardi, D.E. Bacelo, S.E. Fioressi, P.R. Duchowicz, The conformation-independent QSPR approach for predicting the oxidation rate constant of water micropollutants, Environmental Science and Pollution Research, 24 (2017) 27366-27375.