

Assessing and predicting water quality index with key water parameters by machine learning models in coastal cities, China

Jing Xu^{1*}, Yuming Mo², Senlin Zhu¹, Jinran Wu³, Guangqiu Jin⁴,

You-Gan Wang⁵, Qingfeng Ji¹, Ling Li⁶

1. College of Hydraulic Science and Engineering, Yangzhou University,
Yangzhou, China

Emails: xujing7503@yzu.edu.cn

slzhu@yzu.edu.cn

qfji@yzu.edu.cn

2. School of Naval Architecture and Ocean Engineering, Jiangsu University
of Science and Technology, Zhenjiang, China

Email: yumingmo@just.edu.cn

3. The Institute for Learning Sciences and Teacher Education, Australian
Catholic University, Queensland, Australia

Email: ryan.wu@acu.edu.au

4. College of Water Conservancy and Hydropower Engineering, Hohai
University, Nanjing, China

Emails: jingq@hhu.edu.cn

5. School of Mathematics and Physics, The University of Queensland,
Queensland, Australia

Emails: ygwanguq2012@gmail.com

6. Key Laboratory of Coastal Environment and Resources of Zhejiang
Province (KLaCER), School of Engineering, Westlake University, Hangzhou,
China

Email: liling@westlake.edu.cn

Submitted to *Heliyon* on 21 January 2024

Resubmitted to *Heliyon* on 22 April 2024

* Corresponding author

Table S1. Weights and normalization factors of the parameters used in the calculation of the water quality index.

Water quality parameters	Weight	Normalization factor										
		100	90	80	70	60	50	40	30	20	10	0
WT	1	[16,21)	(15,16)/ [21,23)	(14,15)/ [23,25)	(12,14)/ [25,27)	(10,12]/ [27,29)	(5,10]/ [29,31)	(0,5]/ [31,35)	(-2,0]/ [35,36)	(-4,-2]/ [36,40)	(-6,-4]/ [40,45)	≤ -6 / ≥ 45
pH	1	7	(7,8]	(8,8.5]	(8.5,9]	[6.5,7)	[6,6.5)/ (9,9.5]	[5,6)/ (9.5,10]	[4,5)/ (10,11]	[3,4)/ (11,12]	[2,3)/ (12,13]	[1,2)/ (13,14]
DO	4	≥ 7.5	(7,7.5]	(6.5,7]	(6,6.5]	(5,6]	(4,5]	(3.5,4]	(3,3.5]	(2,3]	[1,2]	< 1
EC	1	< 750	< 1000	< 1250	< 1500	< 2000	< 2500	< 3000	< 5000	< 8000	$< 12,000$	$\geq 12,000$
Tur	2	< 5	< 10	< 15	< 20	< 25	< 30	< 40	< 60	< 80	≤ 100	> 100
COM	3	< 1	< 2	< 3	< 4	< 6	< 8	< 10	< 12	< 14	≤ 15	> 15
AN	3	< 0.01	< 0.05	< 0.1	< 0.2	< 0.3	< 0.4	< 0.5	< 0.75	< 1	≤ 1.25	> 1.25
TP	1	< 0.01	< 0.02	< 0.05	< 0.1	< 0.15	< 0.2	< 0.25	< 0.3	< 0.35	≤ 0.4	> 0.4
TN	2	< 0.1	< 0.2	< 0.35	< 0.5	< 0.75	< 1	< 1.25	< 1.5	< 1.75	≤ 2	> 2

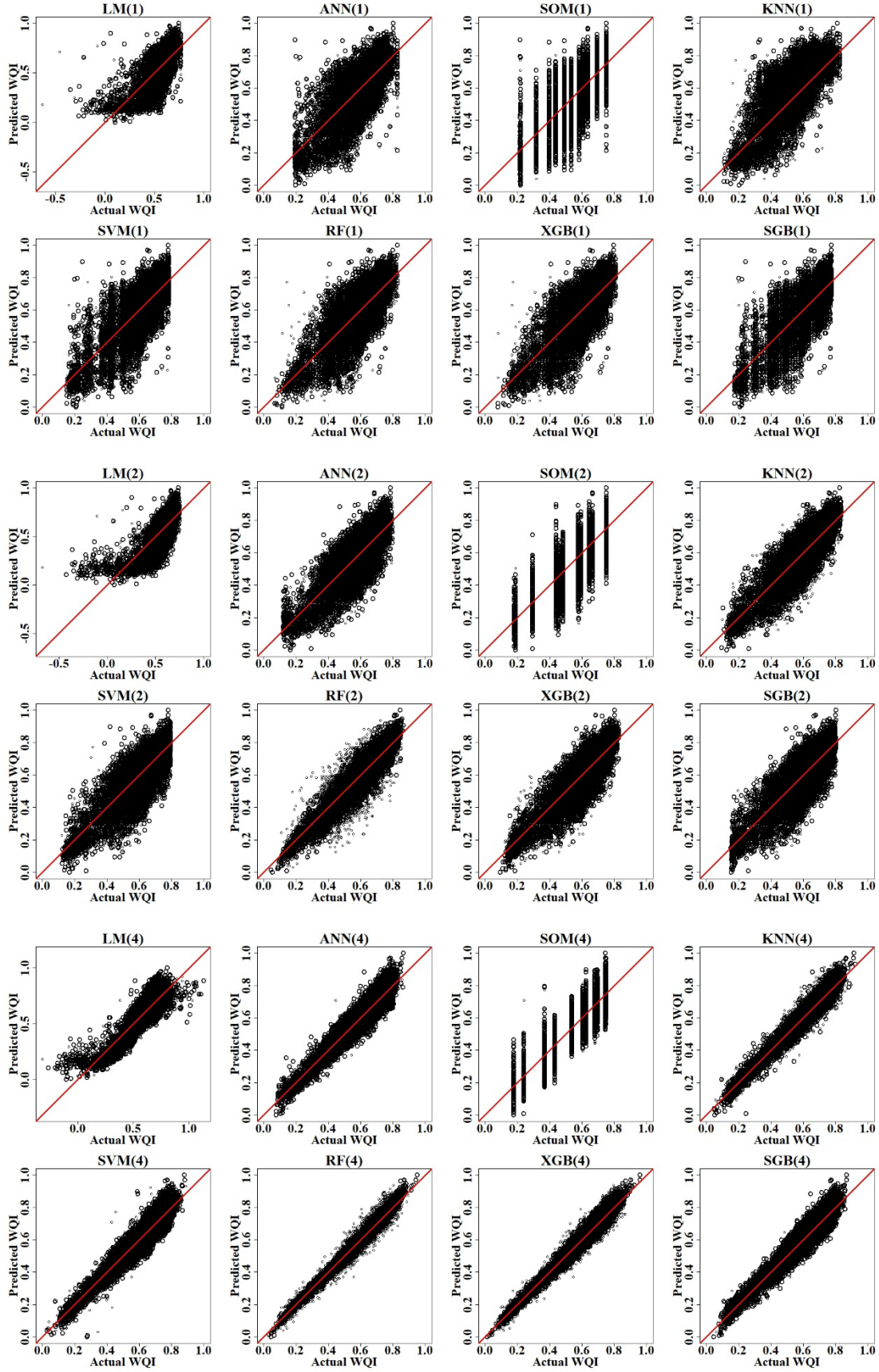
Table S2. The WQI values and corresponding grades of water quality.

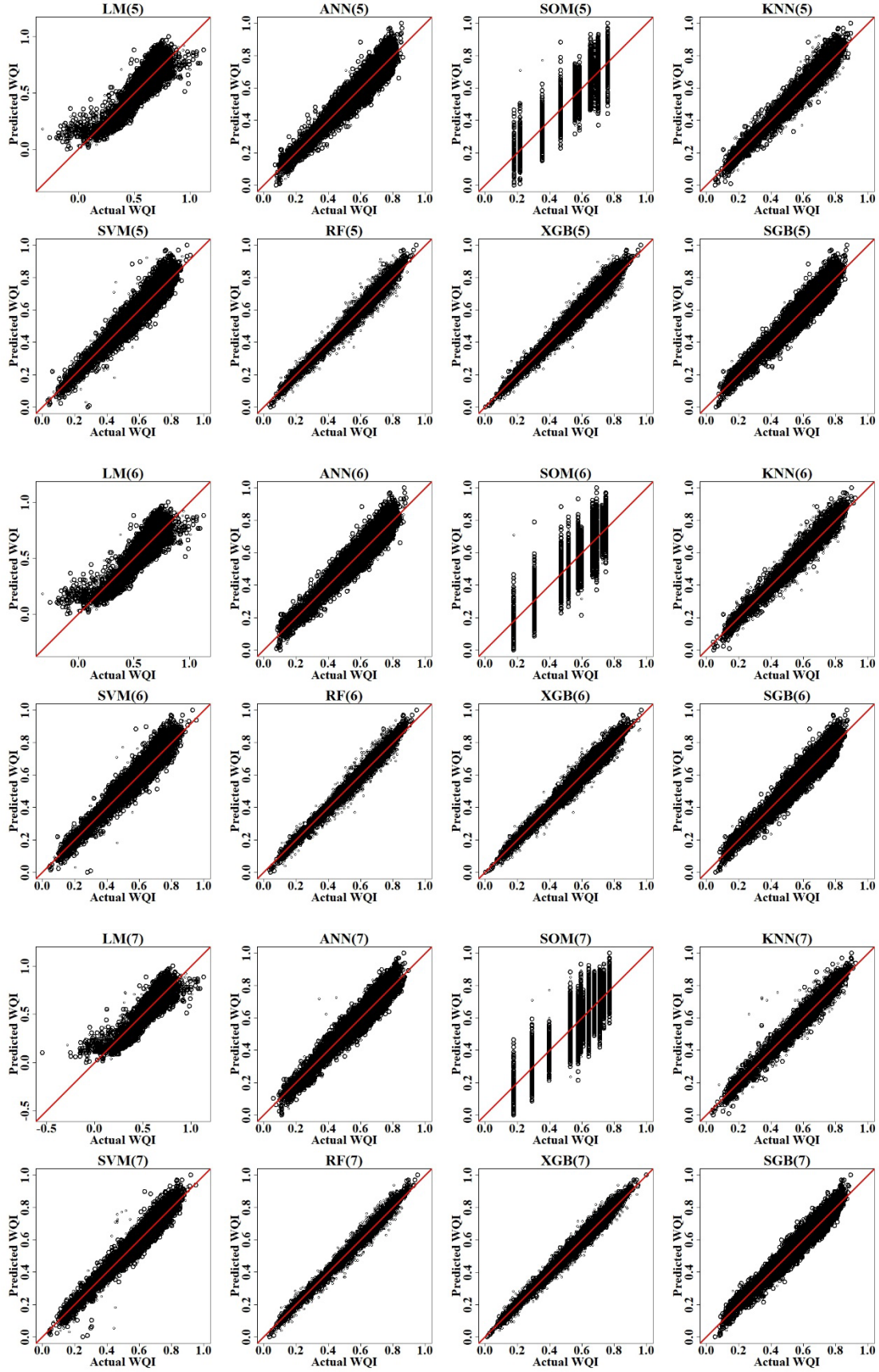
WQI value	WQI levels
91-100	Excellent water quality
71-90	Good water quality
51-70	Medium water quality
26-50	Low water quality
0-25	Bad water quality

Table S3. Evaluation of nine models with different groups of input variables to predict the WQI.

Model		MAE		NSE		RMSE		R ²	
		Train	Test	Train	Test	Train	Test	Train	Test
LM	1	0.081	0.081	0.513	0.508	0.106	0.107	0.513	0.508
	2	0.078	0.078	0.557	0.561	0.101	0.101	0.557	0.562
	3	0.063	0.063	0.711	0.716	0.082	0.081	0.711	0.717
	4	0.055	0.055	0.783	0.793	0.071	0.069	0.783	0.793
	5	0.055	0.054	0.786	0.796	0.070	0.069	0.786	0.797
	6	0.055	0.054	0.786	0.797	0.070	0.069	0.786	0.797
	7	0.054	0.053	0.792	0.799	0.069	0.068	0.792	0.799
	8	0.049	0.048	0.821	0.833	0.064	0.062	0.821	0.833
	9	0.048	0.048	0.825	0.837	0.064	0.062	0.825	0.838
ANN	1	0.078	0.077	0.568	0.575	0.100	0.099	0.568	0.575
	2	0.071	0.069	0.662	0.665	0.088	0.088	0.6625	0.665
	3	0.040	0.041	0.885	0.882	0.052	0.052	0.885	0.882
	4	0.035	0.036	0.912	0.912	0.045	0.045	0.912	0.912
	5	0.035	0.036	0.913	0.912	0.045	0.045	0.913	0.912
	6	0.035	0.035	0.916	0.916	0.044	0.044	0.916	0.916
	7	0.033	0.033	0.928	0.926	0.041	0.042	0.928	0.926
	8	0.024	0.024	0.959	0.958	0.031	0.031	0.959	0.958
	9	0.019	0.019	0.972	0.972	0.025	0.025	0.972	0.972
SOM	1	0.078	0.077	0.565	0.571	0.100	0.099	0.565	0.571
	2	0.073	0.072	0.633	0.635	0.092	0.092	0.633	0.635
	3	0.054	0.054	0.797	0.798	0.069	0.069	0.797	0.799
	4	0.052	0.052	0.816	0.813	0.065	0.066	0.816	0.813
	5	0.054	0.055	0.793	0.792	0.069	0.069	0.793	0.792
	6	0.059	0.059	0.751	0.757	0.076	0.075	0.751	0.757
	7	0.060	0.060	0.744	0.746	0.077	0.077	0.744	0.746
	8	0.060	0.061	0.743	0.743	0.077	0.077	0.743	0.743
	9	0.058	0.057	0.767	0.772	0.073	0.073	0.767	0.772
KNN	1	0.094	0.096	0.412	0.381	0.117	0.120	0.542	0.514
	2	0.057	0.069	0.773	0.660	0.072	0.089	0.774	0.661
	3	0.029	0.039	0.936	0.892	0.038	0.050	0.937	0.892
	4	0.024	0.032	0.956	0.924	0.032	0.042	0.957	0.924

SVM	5	0.023	0.031	0.959	0.925	0.030	0.042	0.961	0.926
	6	0.019	0.029	0.972	0.935	0.025	0.039	0.972	0.935
	7	0.017	0.026	0.977	0.943	0.023	0.037	0.977	0.943
	8	0.015	0.021	0.981	0.961	0.021	0.030	0.982	0.961
	9	0.015	0.021	0.981	0.959	0.021	0.031	0.981	0.960
	1	0.077	0.076	0.573	0.578	0.099	0.099	0.574	0.579
	2	0.066	0.068	0.678	0.667	0.086	0.088	0.680	0.669
	3	0.037	0.039	0.898	0.890	0.048	0.051	0.898	0.890
	4	0.032	0.034	0.923	0.915	0.042	0.045	0.923	0.915
RF	5	0.031	0.033	0.928	0.919	0.041	0.043	0.928	0.919
	6	0.029	0.031	0.935	0.928	0.039	0.041	0.935	0.928
	7	0.026	0.028	0.950	0.939	0.034	0.038	0.950	0.939
	8	0.015	0.016	0.984	0.976	0.019	0.024	0.984	0.976
	9	0.013	0.015	0.987	0.979	0.018	0.022	0.987	0.979
	1	0.075	0.078	0.605	0.562	0.096	0.101	0.605	0.563
	2	0.037	0.067	0.899	0.666	0.048	0.088	0.902	0.668
	3	0.016	0.036	0.979	0.905	0.022	0.047	0.980	0.905
	4	0.013	0.029	0.987	0.937	0.017	0.038	0.987	0.937
XGB	5	0.012	0.027	0.988	0.942	0.017	0.037	0.988	0.942
	6	0.012	0.026	0.989	0.949	0.016	0.034	0.990	0.950
	7	0.009	0.022	0.993	0.961	0.013	0.030	0.993	0.961
	8	0.005	0.013	0.998	0.986	0.007	0.018	0.998	0.986
	9	0.005	0.012	0.998	0.987	0.007	0.017	0.998	0.988
	1	0.075	0.077	0.601	0.573	0.096	0.099	0.601	0.574
	2	0.059	0.067	0.753	0.682	0.076	0.086	0.754	0.682
	3	0.031	0.038	0.931	0.900	0.040	0.048	0.931	0.900
	4	0.019	0.031	0.970	0.932	0.026	0.039	0.970	0.932
SGB	5	0.018	0.029	0.974	0.937	0.024	0.038	0.974	0.937
	6	0.016	0.027	0.979	0.946	0.022	0.036	0.979	0.946
	7	0.014	0.023	0.986	0.959	0.018	0.030	0.986	0.959
	8	0.006	0.010	0.997	0.992	0.008	0.013	0.997	0.992
	9	0.005	0.008	0.998	0.995	0.006	0.011	0.998	0.995
	1	0.077	0.076	0.577	0.581	0.099	0.099	0.577	0.581
	2	0.067	0.068	0.687	0.681	0.085	0.086	0.687	0.681
	3	0.038	0.039	0.900	0.893	0.048	0.050	0.900	0.893
	4	0.033	0.033	0.925	0.922	0.042	0.043	0.925	0.922
	5	0.032	0.033	0.927	0.924	0.041	0.042	0.927	0.924
	6	0.031	0.032	0.932	0.929	0.039	0.040	0.932	0.930
	7	0.028	0.029	0.945	0.943	0.036	0.036	0.945	0.943
	8	0.013	0.014	0.987	0.986	0.017	0.018	0.987	0.986
	9	0.011	0.011	0.991	0.990	0.014	0.015	0.991	0.990





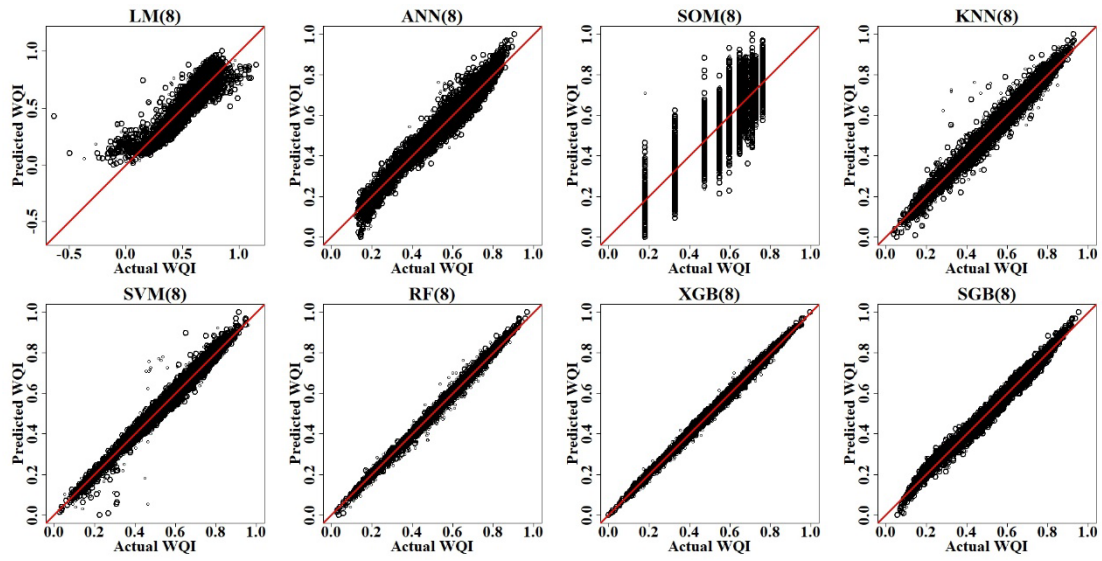


Figure S1. The actual vs. predicted WQI which were predicted by different input groups by statistical and machine learning models.