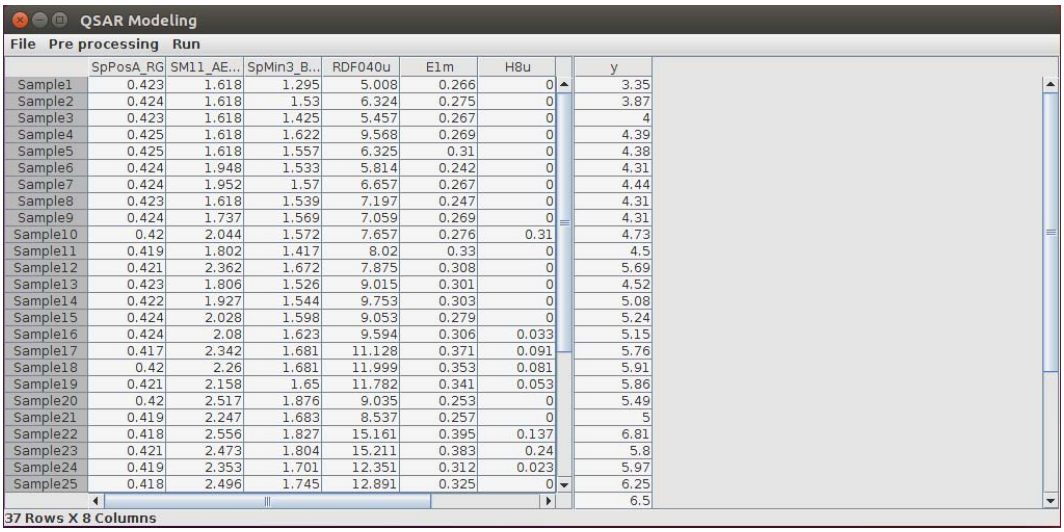
Supplementary Information

QSAR modeling: a new open source computational package to generate and validate QSAR models

João Paulo A. Martins e Márcia M. C. Ferreira* marcia@iqm.unicamp.br
Instituto de Química, Universidade Estadual de Campinas, CP 6154, 13084-971 Campinas – SP, Brasil



	SpPosA_RG	SM11_AE...	SpMin3_B...	RDF040u	E1m	H8u	y
Sample1	0.423	1.618	1.295	5.008	0.266	0	3.35
Sample2	0.424	1.618	1.53	6.324	0.275	0	3.87
Sample3	0.423	1.618	1.425	5.457	0.267	0	4
Sample4	0.425	1.618	1.622	9.568	0.269	0	4.39
Sample5	0.425	1.618	1.557	6.325	0.31	0	4.38
Sample6	0.424	1.948	1.533	5.814	0.242	0	4.31
Sample7	0.424	1.952	1.57	6.657	0.267	0	4.44
Sample8	0.423	1.618	1.539	7.197	0.247	0	4.31
Sample9	0.424	1.737	1.569	7.059	0.269	0	4.31
Sample10	0.42	2.044	1.572	7.657	0.276	0.31	4.73
Sample11	0.419	1.802	1.417	8.02	0.33	0	4.5
Sample12	0.421	2.362	1.672	7.875	0.308	0	5.69
Sample13	0.423	1.806	1.526	9.015	0.301	0	4.52
Sample14	0.422	1.927	1.544	9.753	0.303	0	5.08
Sample15	0.424	2.028	1.598	9.053	0.279	0	5.24
Sample16	0.424	2.08	1.623	9.594	0.306	0.033	5.15
Sample17	0.417	2.342	1.681	11.128	0.371	0.091	5.76
Sample18	0.42	2.26	1.681	11.999	0.353	0.081	5.91
Sample19	0.421	2.158	1.65	11.782	0.341	0.053	5.86
Sample20	0.42	2.517	1.876	9.035	0.253	0	5.49
Sample21	0.419	2.247	1.683	8.537	0.257	0	5
Sample22	0.418	2.556	1.827	15.161	0.395	0.137	6.81
Sample23	0.421	2.473	1.804	15.211	0.383	0.24	5.8
Sample24	0.419	2.353	1.701	12.351	0.312	0.023	5.97
Sample25	0.418	2.496	1.745	12.891	0.325	0	6.25
							6.5

Figure 1S: The main screen of *QSAR modeling* program

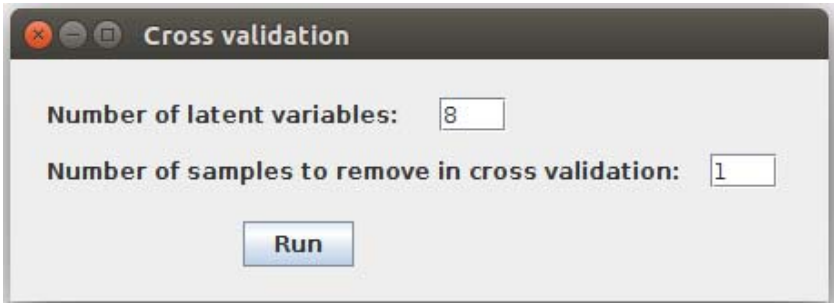


Figure 2S. Window from *QSAR Modeling* where the user chooses the maximum number of latent variables and the number of samples to be removed during cross validation.

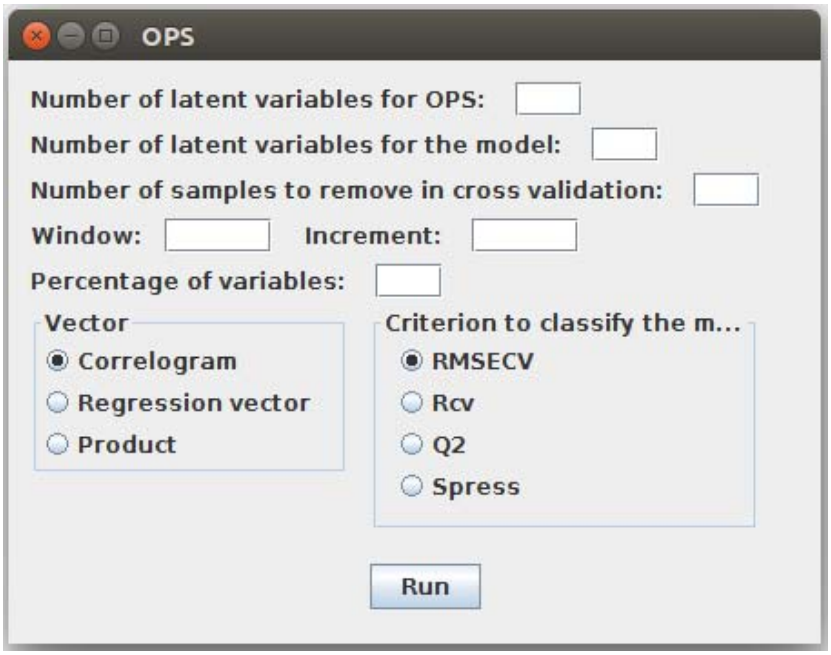


Figure 3S: *QSAR modeling* window where the user can choose the options to run the OPS algorithm

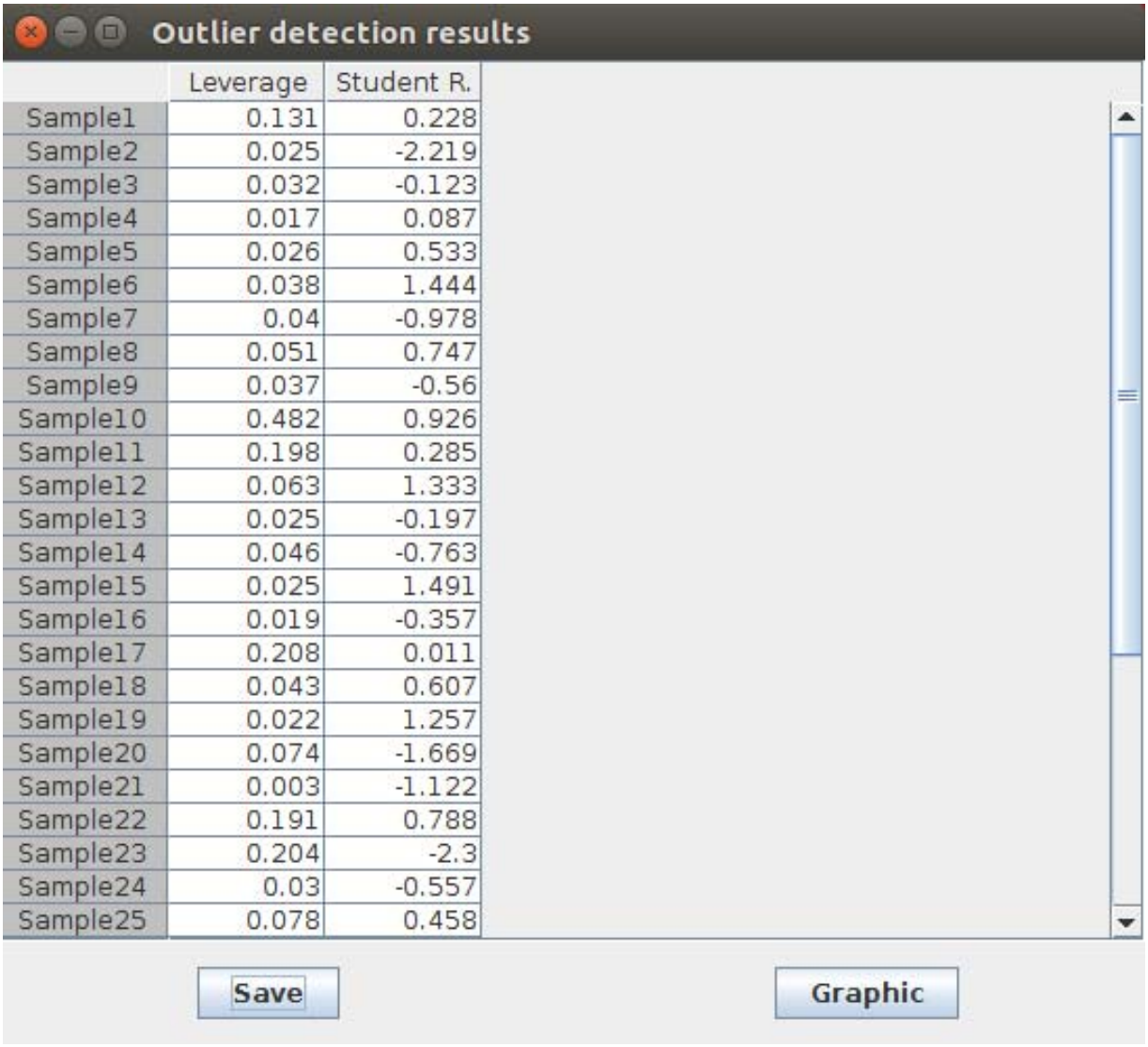


Figure 4S. Results from outlier detection including Leverage and Studentized residuals values for compounds from the training set.

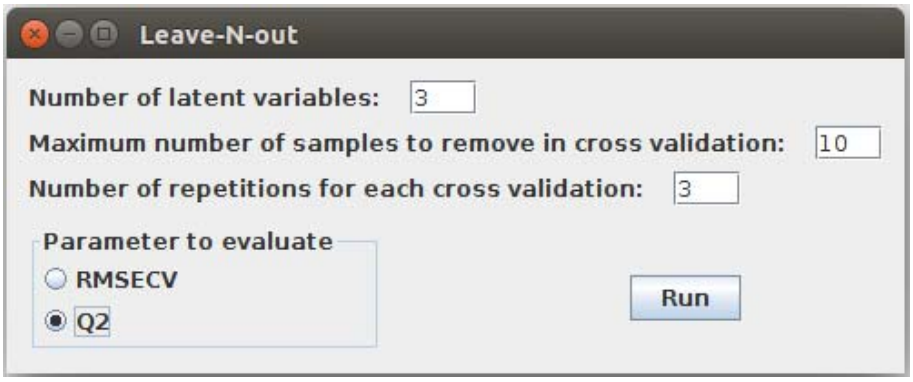


Figure 5S: Leave-*N*-out validation procedure to assess the robustness of a model using *QSAR modeling*.

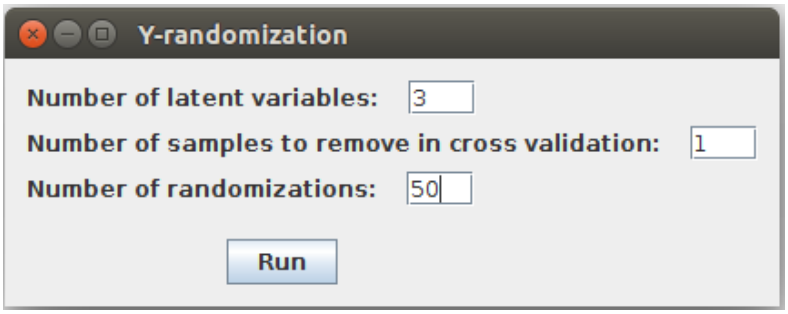


Figure 6S. **y**-randomization validation procedure to verify chance correlation of a model using *QSAR modeling*.