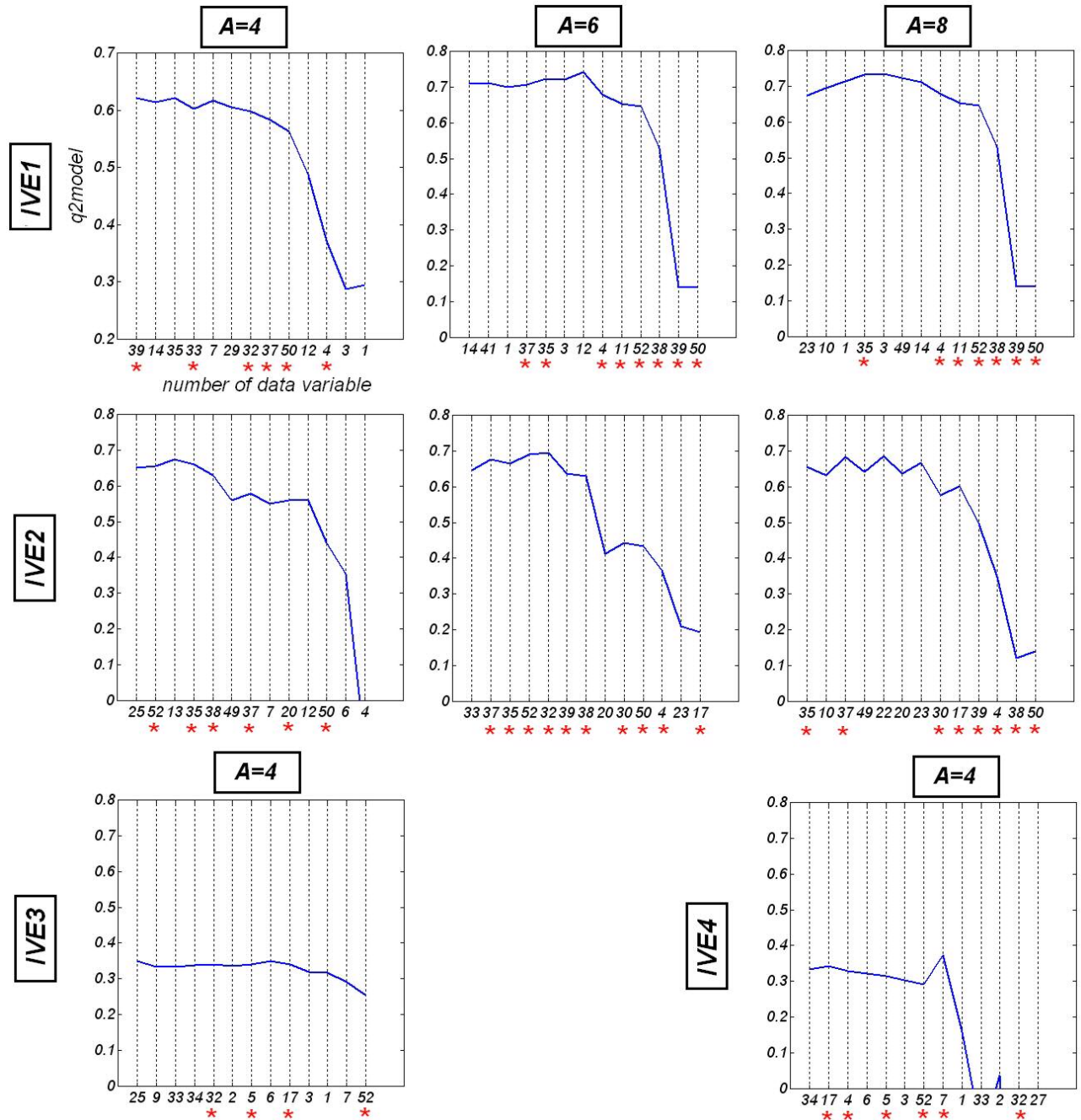


Supporting Information Available

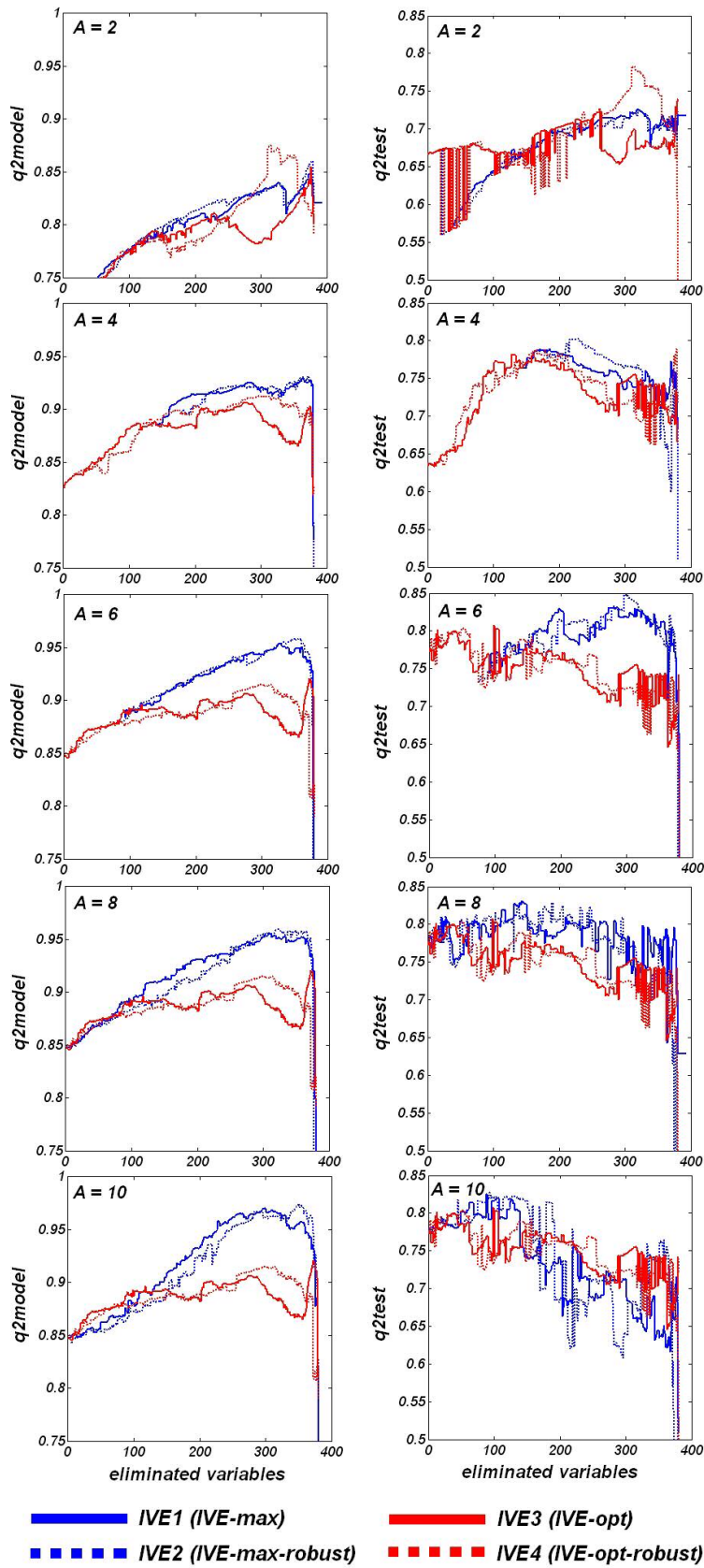
The performance of IVE1-4 variable elimination for the Selwood dataset.

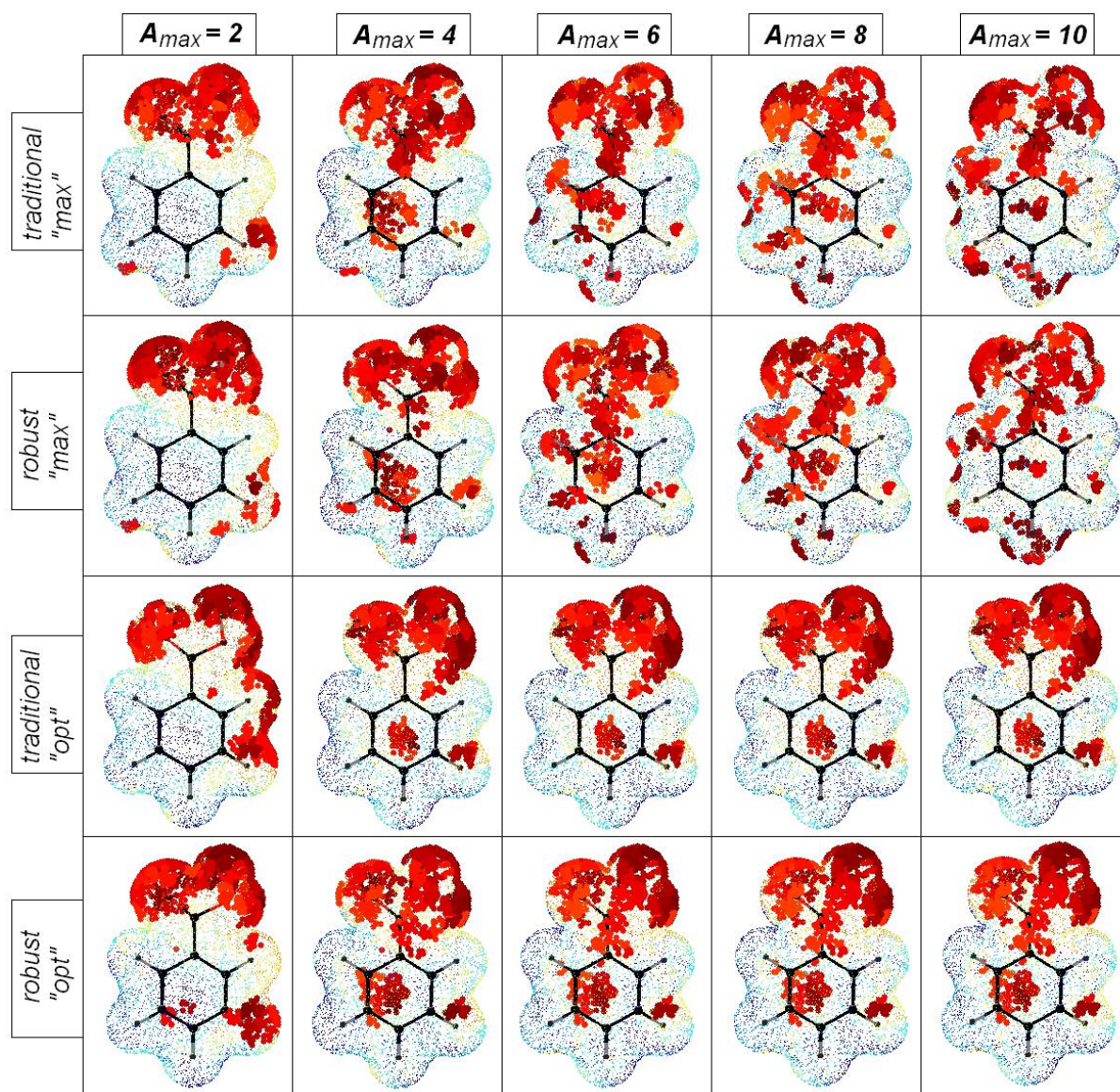


Variable elimination performance for the Selwood dataset. The numbers on the x axis show the respective variables surviving the respective IVE elimination at the end phase of such processing. The y axis shows the decreasing q^2 estimator. The asterisks indicate the variable

numbers reported to survive variable elimination reported in the literature.^{19, 20} The increase of A does not change the elimination curve profiles for *IVE3* and *IVE4*.

The performance of the PLS CoMSA (molecules without halogens: AM1 charges; molecules containing any halogens in structure: PM3 charges) modeling coupled with the *IVE1* to *IVE4* tested by q^2_{test} and q^2_{model} performance measures.





Interaction contour plots revealed by *IVE1* to *IVE4* for the different model complexity, using the PM3 charges. Colors indicate the order of eliminated variables, where blue means the early eliminated variables and maroon these that tend to survive the process.