Supplementary Information for : A Bayesian Approach to Calibrating High-Throughput Virtual Screening Results and Application to Organic Photovoltaic Materials

E.O. Pyzer-Knapp[a], G.N. Simm[b], and A. Aspuru Guzik[a *]

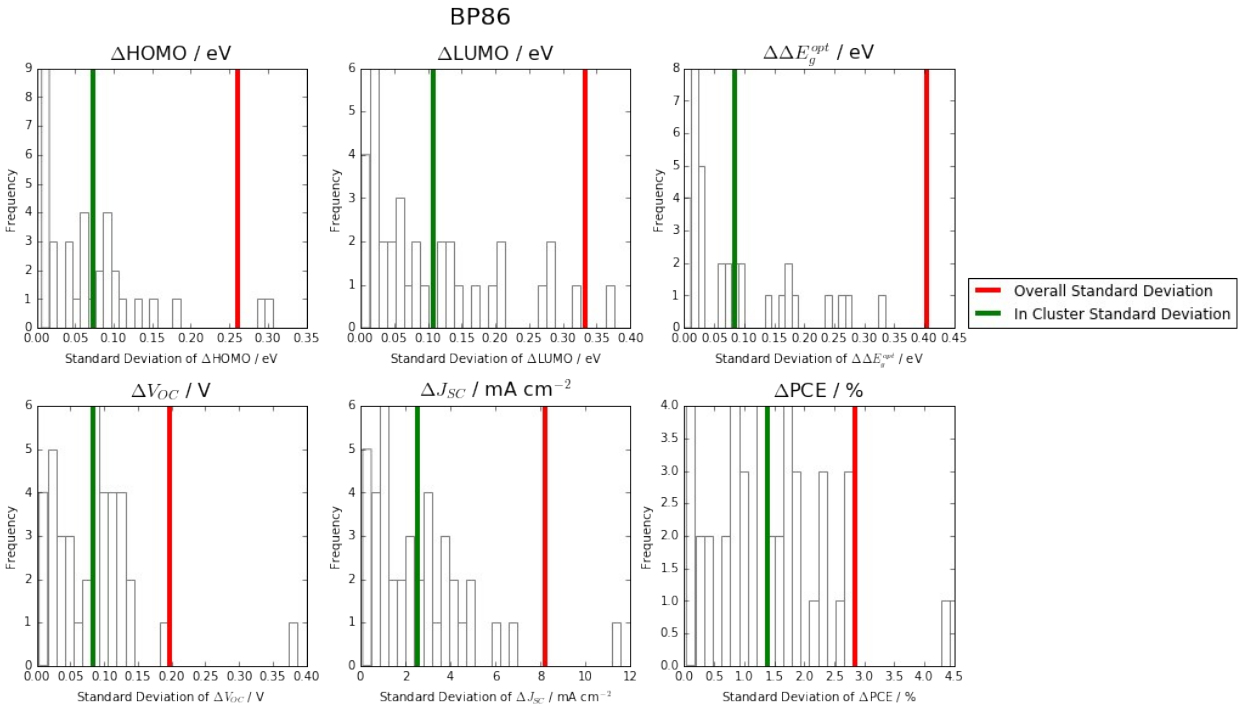

Figure S1: A histogram showing the distribution of values for the standard deviation of the 'delta' contained within each cluster for the BP86 /def2-SVP level of theory. If there were no structure-property relationship, we would expect the standard deviation of nodes contained within the average cluster (plotted as a green line) to have a similar value to the standard deviation of the entire set (plotted in red).

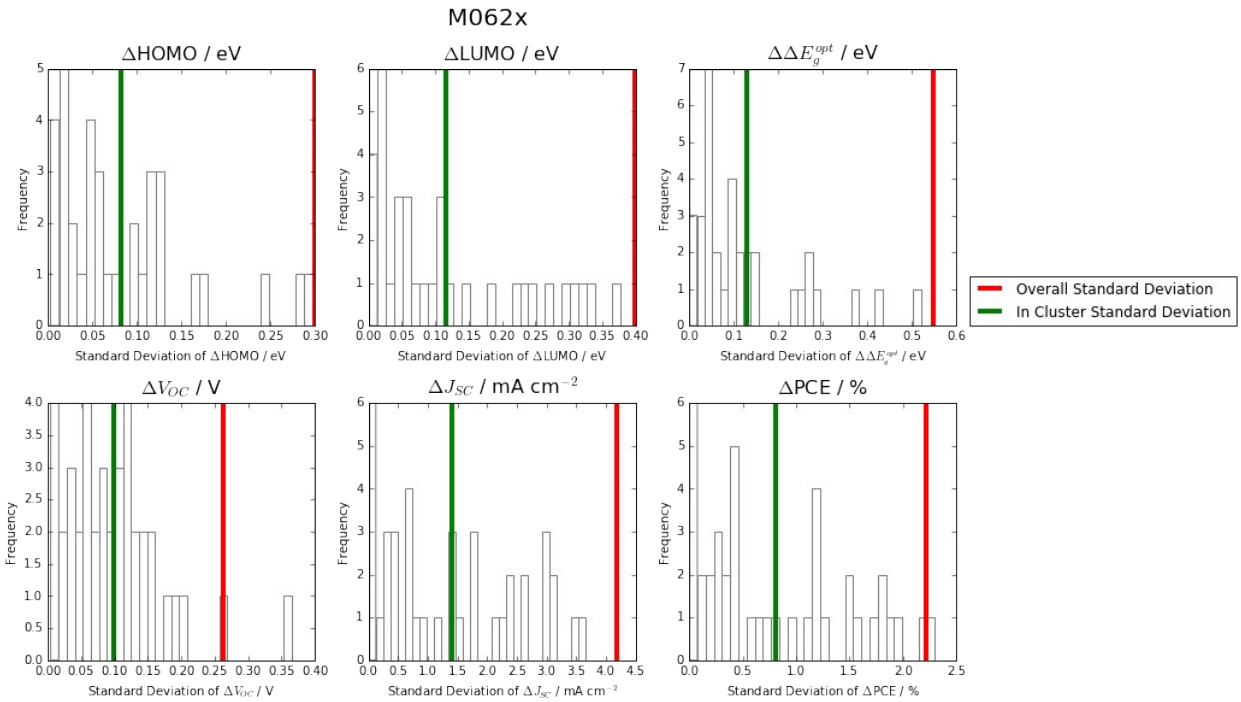

Figure S2: A histogram showing the distribution of values for the standard deviation of the 'delta' contained within each cluster for the M062x /def2-SVP level of theory. If there were no structure-property relationship, we would expect the standard deviation of nodes contained within the average cluster (plotted as a green line) to have a similar value to the standard deviation of the entire set (plotted in red).

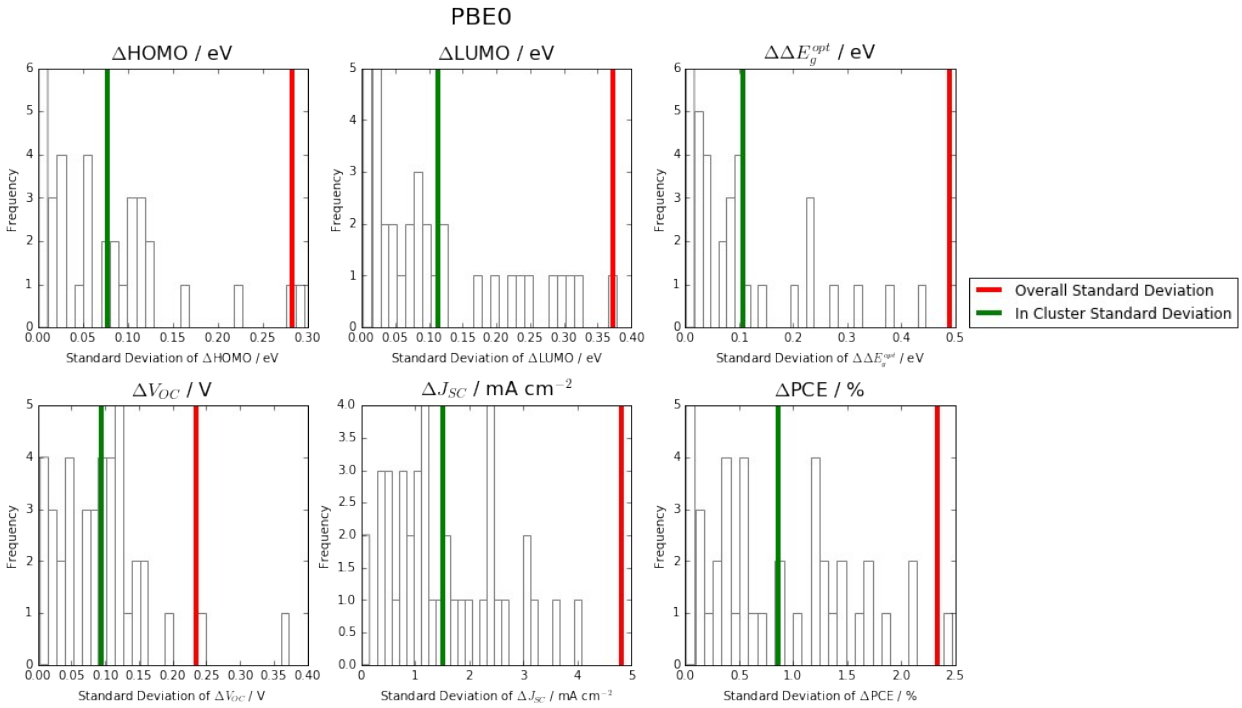

Figure S3: A histogram showing the distribution of values for the standard deviation of the 'delta' contained within each cluster for the PBE0 /def2-SVP level of theory. If there were no structure-property relationship, we would expect the standard deviation of nodes contained within the average cluster (plotted as a green line) to have a similar value to the standard deviation of the entire set (plotted in red).

M062x

Property	σ_f	σ_n	l
JSC	8.588417	1.987535	0.294803
VOC	4.282613	0.2	3.538542
Gap	0.720354	0.229187	0.916486
HOMO	0.066602	0.153097	0.475883
LUMO	0.465122	0.218979	0.971074

BP86

Property	σ_f	σ_n	l
JSC	68.74299	3.380612	0.428647
VOC	0.376849	0.18	1.837238
Gap	0.764586	0.2	1.674024
HOMO	0.051208	0.123001	0.429685
LUMO	0.136843	0.167237	0.506049

PBE0

Property	σ_f	σ_n	l
JSC	14.51363	2.220835	0.368901
VOC	2.032295	0.18	2.812411
Gap	0.385456	0.186988	0.682732
HOMO	0.076026	0.124657	0.463645
LUMO	0.181995	0.174804	0.532822

Table S1 Hyperparameters for the optimized Gaussian processes for calibrating the functionals M062x, BP86, and PBE0.

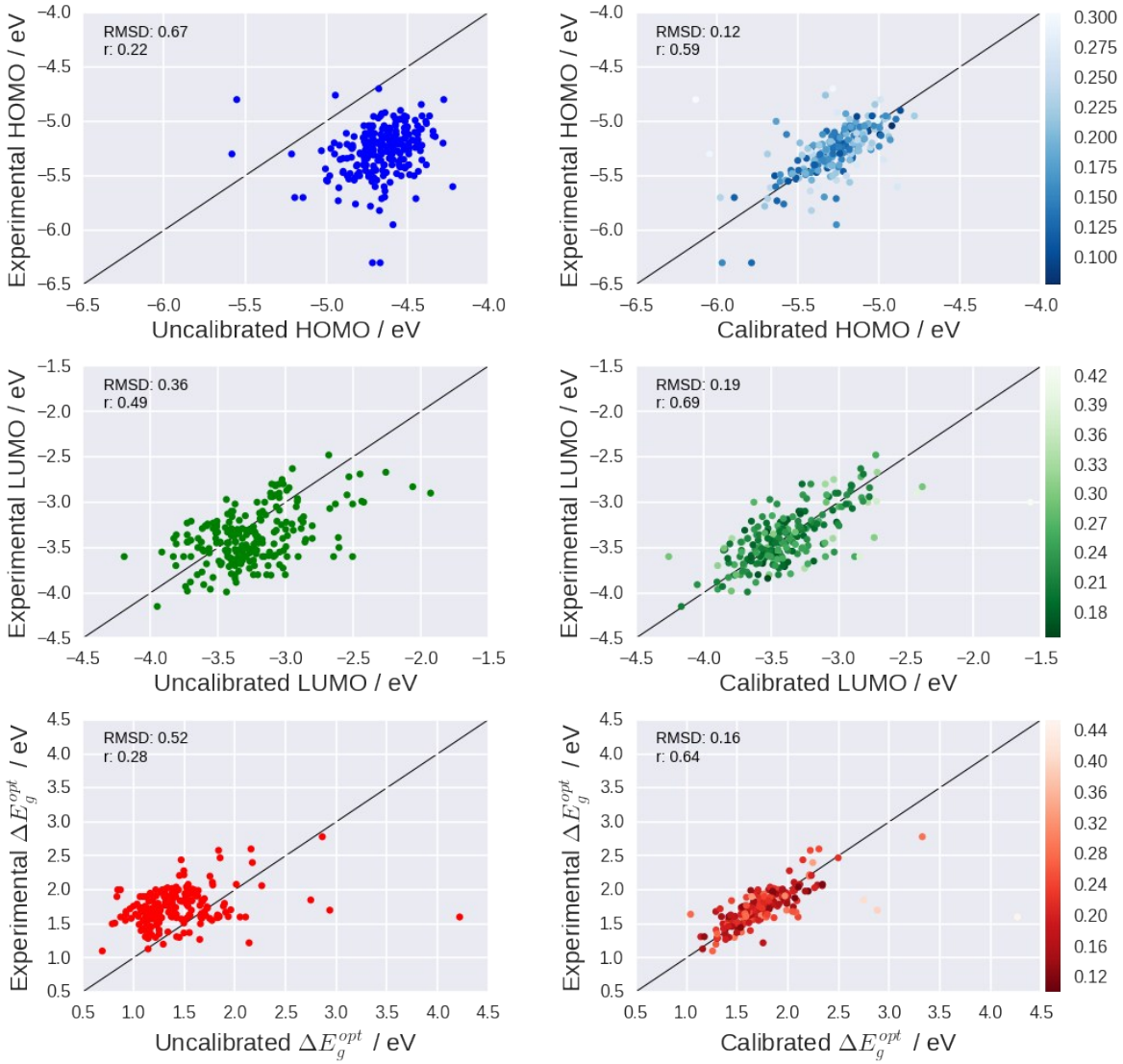

Figure S4: The results of calibrating BP86/def2-SVP quantum-chemical results for the Highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) and optical gap to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

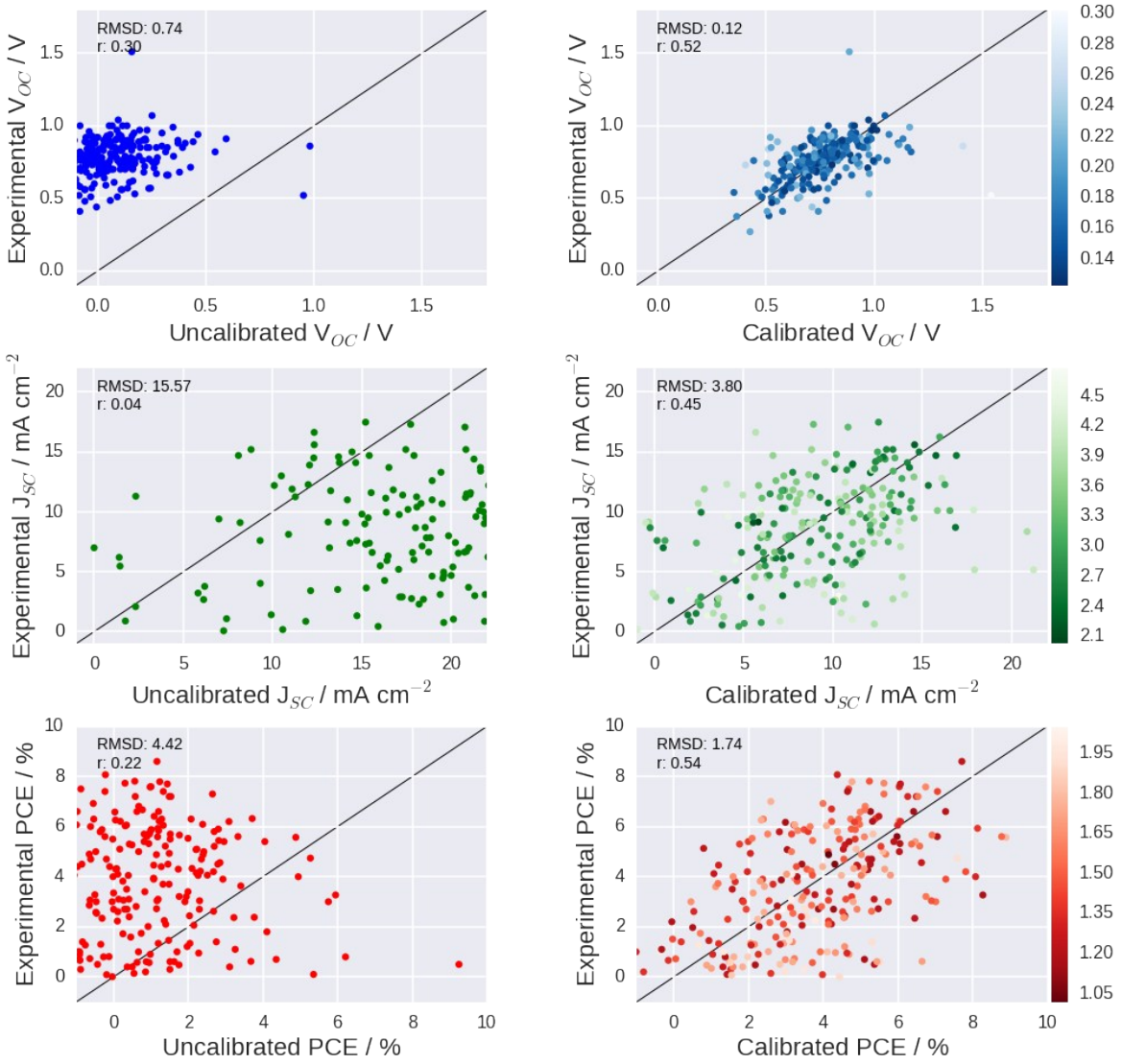

Figure S5: The results of calibrating BP86/def2-SVP quantum-chemical results for the open-circuit potential (VOC), short circuit current density (JSC), and power conversion efficiency (PCE) to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

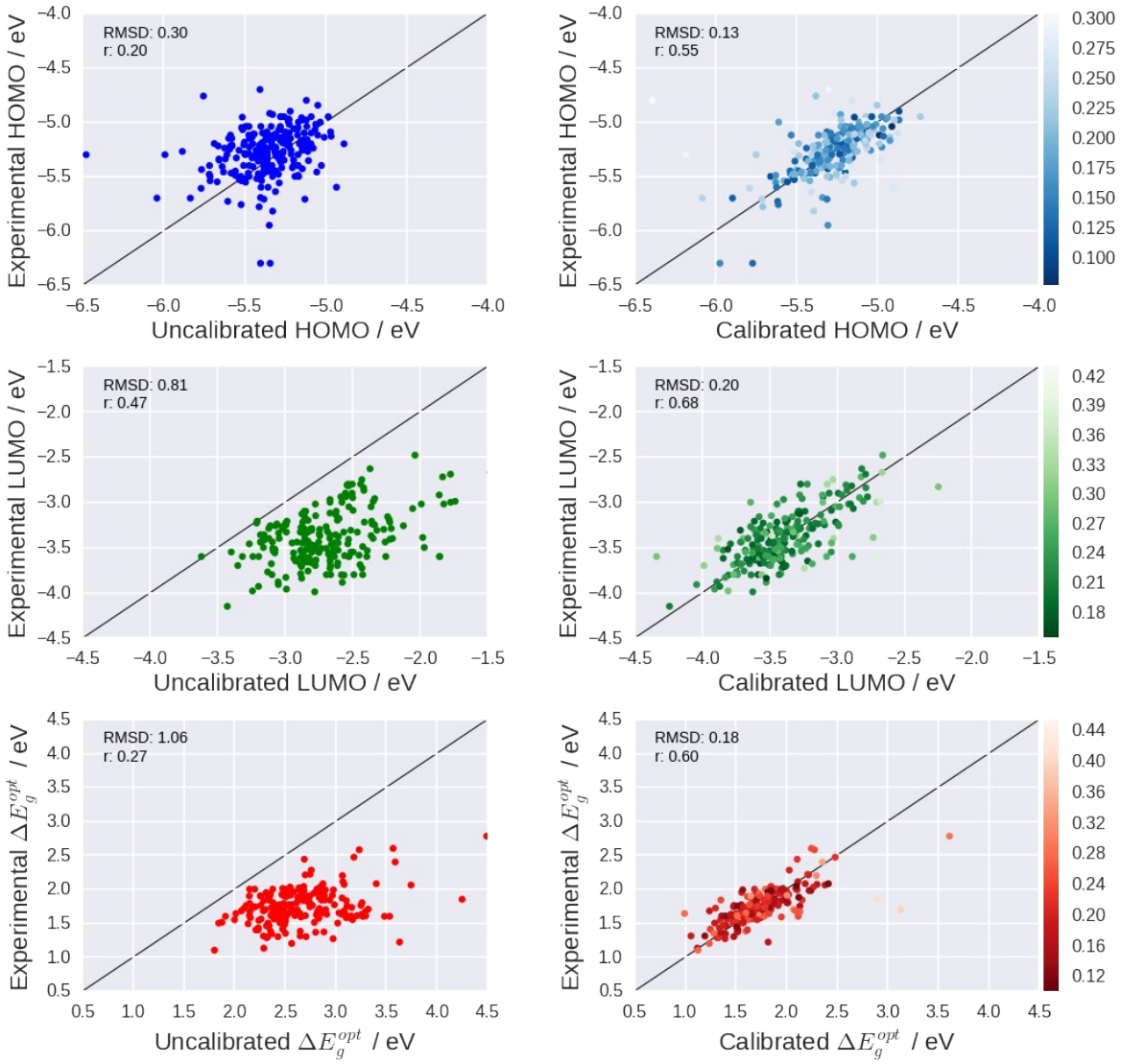

Figure S6: The results of calibrating PBE0/def2-SVP quantum-chemical results for the Highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) and optical gap to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

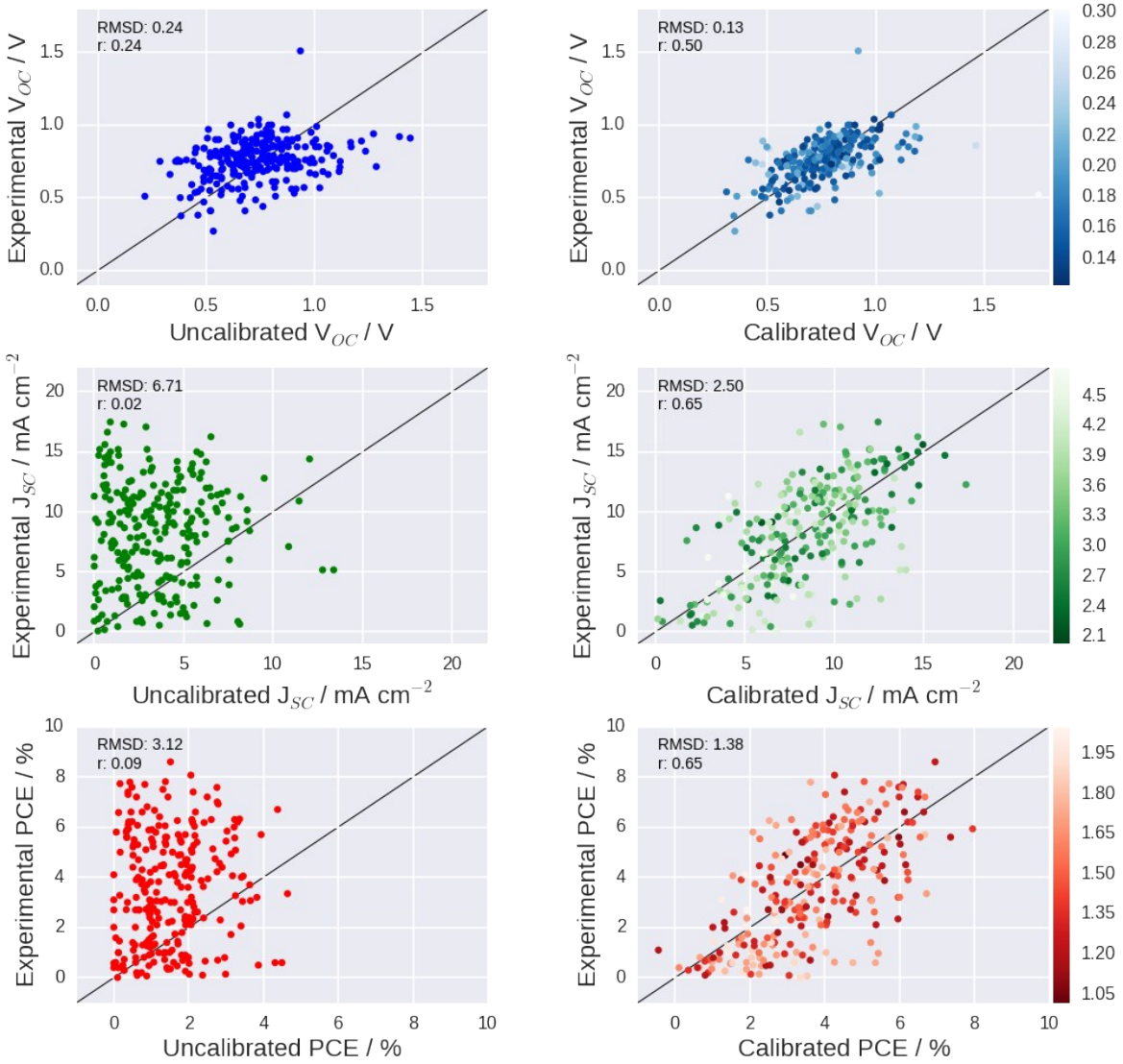

Figure S7: The results of calibrating PBE0/def2-SVP quantum-chemical results for the open-circuit potential (VOC), short circuit current density (JSC), and power conversion efficiency (PCE) to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

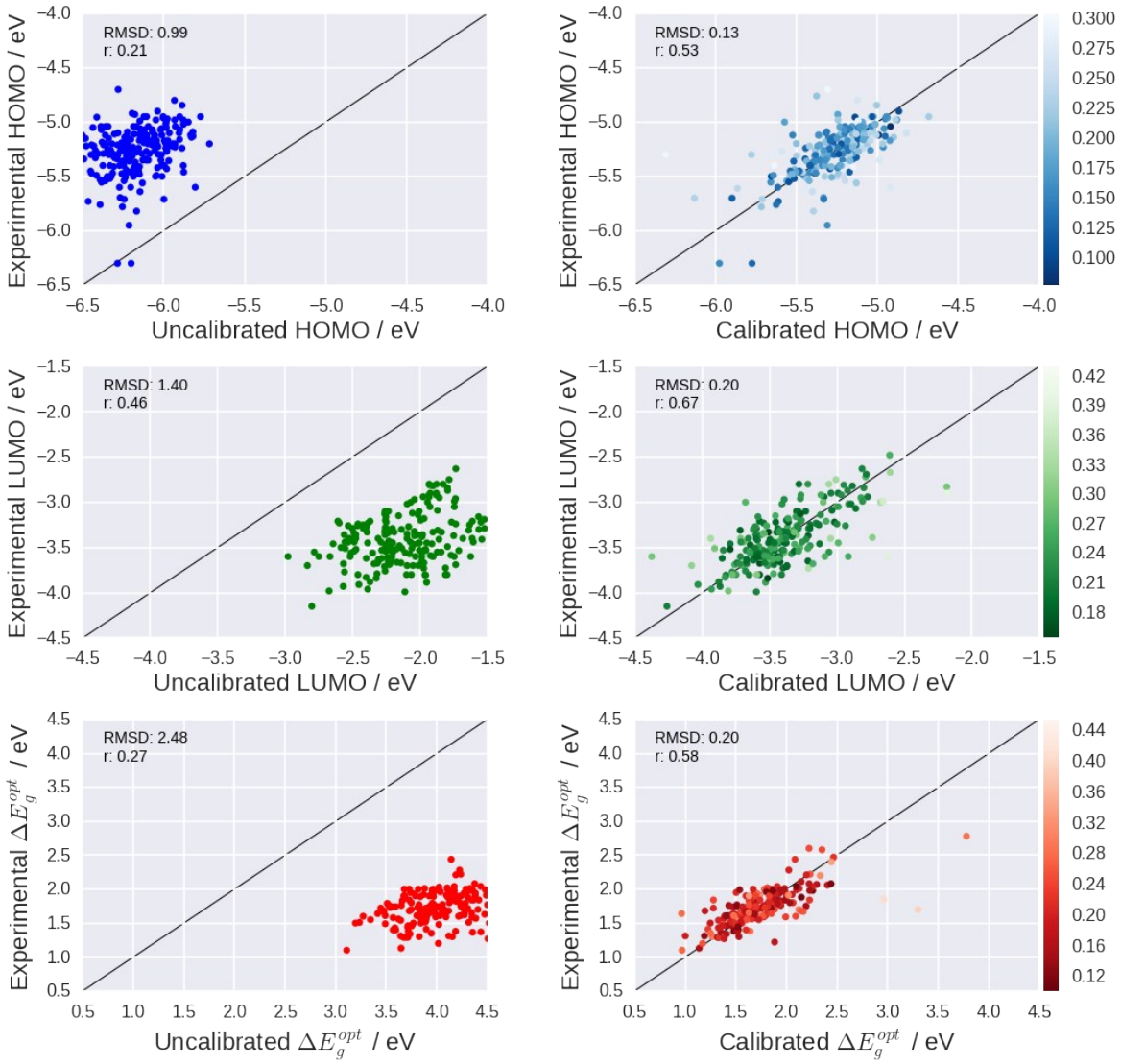

Figure S8: The results of calibrating M06-2X/def2-SVP quantum-chemical results for the Highest occupied molecular orbital (HOMO), lowest unoccupied molecular orbital (LUMO) and optical gap to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

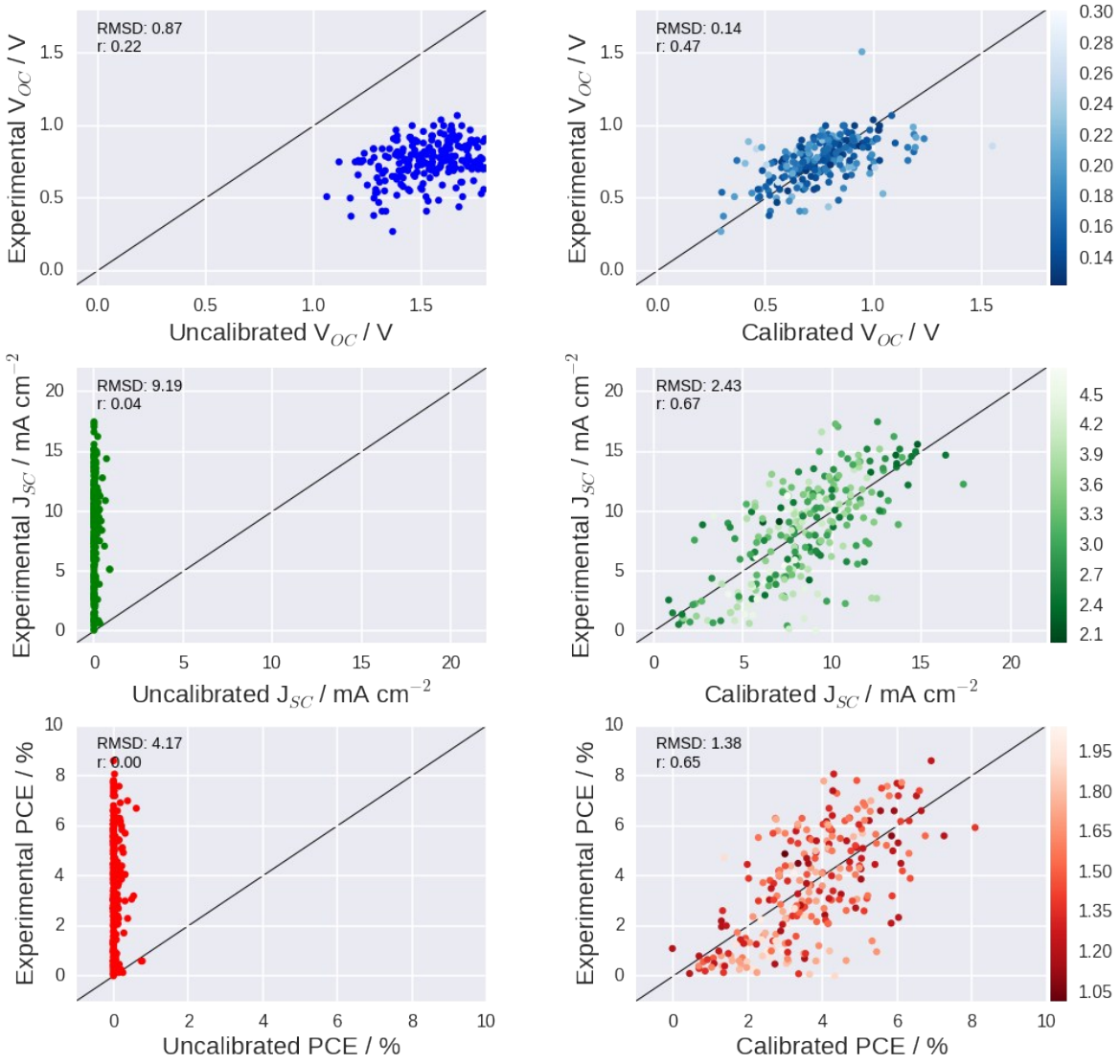

Figure S9: The results of calibrating M06-2X/def2-SVP quantum-chemical results for the open-circuit potential (VOC), short circuit current density (JSC), and power conversion efficiency (PCE) to the experimental HOPV15 data set. The uncertainty in the calibrated values is represented in the fill colour; the lighter the colour, the more uncertain the calibration.

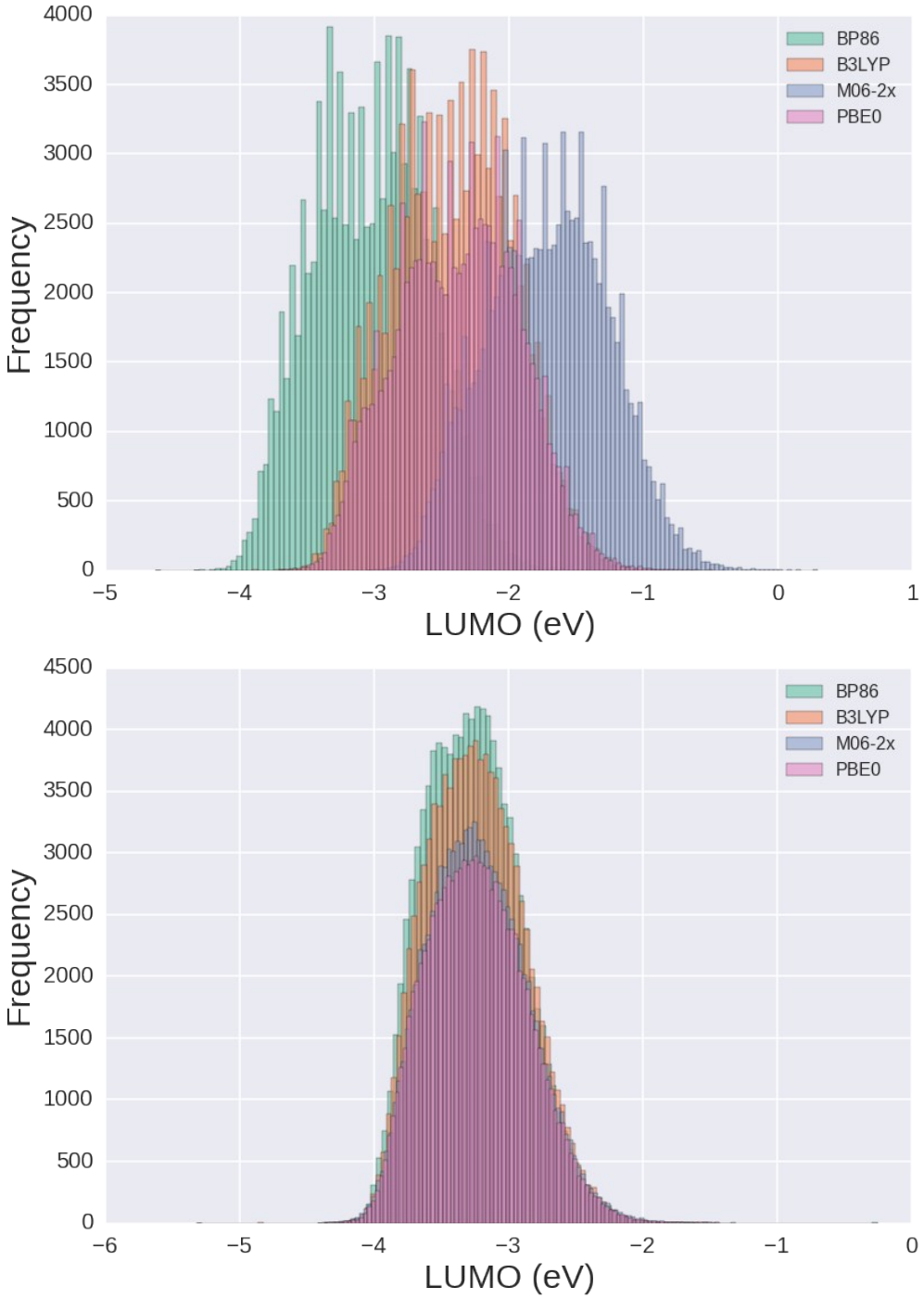

Figure S10: Lowest unoccupied molecular orbital (LUMO) energies (eV), Boltzmann averaged over conformers, calculated for 100,000 molecules from the Clean Energy Project Database (top) and the values for the same set of molecules after calibration (bottom).

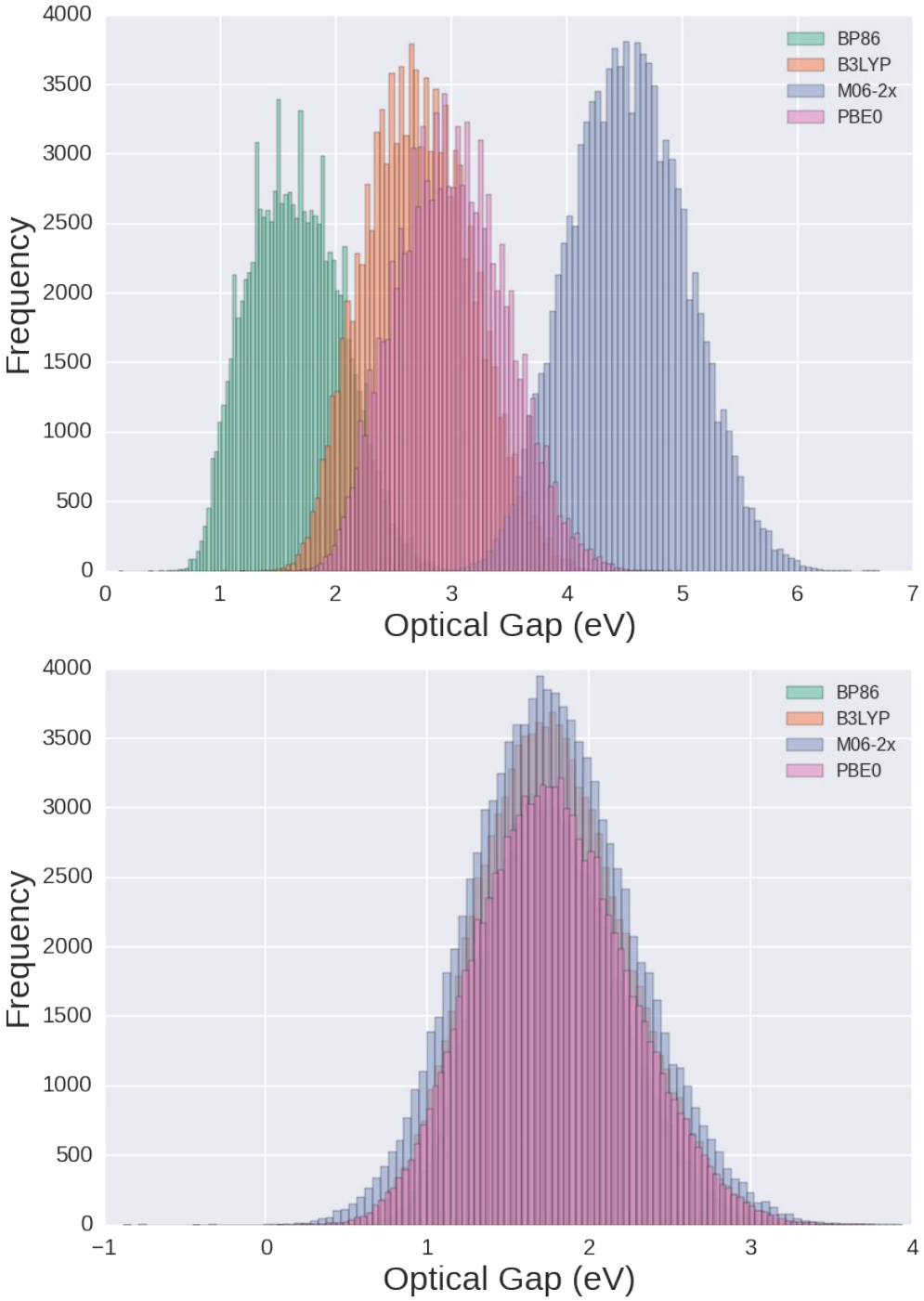

Figure S11: Optical gap energies (eV), Boltzmann averaged over conformers, calculated for 100,000 molecules from the Clean Energy Project Database (top) and the values for the same set of molecules after calibration (bottom).