

## Supplementary Information

|                                                                                                   |    |
|---------------------------------------------------------------------------------------------------|----|
| Supplementary Figure 1: Friedman random forest r-statistic .....                                  | 6  |
| Supplementary Figure 2: Friedman random forest RvE .....                                          | 7  |
| Supplementary Figure 3: Friedman random forest r-statistic Q-Q plot .....                         | 7  |
| Supplementary Figure 4: Friedman linear ridge regression r-statistic .....                        | 8  |
| Supplementary Figure 5: Friedman linear ridge regression RvE .....                                | 8  |
| Supplementary Figure 6: Friedman linear ridge regression r-statistic Q-Q plot .....               | 9  |
| Supplementary Figure 7: Friedman GPR r-statistic .....                                            | 9  |
| Supplementary Figure 8: Friedman GPR RvE.....                                                     | 10 |
| Supplementary Figure 9: Friedman GPR r-statistic Q-Q plot.....                                    | 10 |
| Supplementary Figure 10: Friedman Bayesian GPR r-statistic .....                                  | 11 |
| Supplementary Figure 11: Friedman Bayesian GPR RvE .....                                          | 11 |
| Supplementary Figure 12: Friedman Bayesian GPR r-statistic Q-Q plot.....                          | 12 |
| Supplementary Figure 13: Friedman GPR single predictor r-statistic (bootstrap UQ) .....           | 12 |
| Supplementary Figure 14: Friedman GPR single predictor RvE (bootstrap UQ) .....                   | 13 |
| Supplementary Figure 15: Friedman GPR single predictor r-statistic Q-Q plot (bootstrap UQ) .....  | 14 |
| Supplementary Figure 16: Diffusion random forest r-statistic.....                                 | 14 |
| Supplementary Figure 17: Diffusion random forest RvE .....                                        | 15 |
| Supplementary Figure 18: Diffusion random forest r-statistic Q-Q plot .....                       | 15 |
| Supplementary Figure 19: Diffusion linear ridge regression r-statistic .....                      | 16 |
| Supplementary Figure 20: Diffusion linear ridge regression RvE .....                              | 16 |
| Supplementary Figure 21: Diffusion linear ridge regression r-statistic Q-Q plot .....             | 17 |
| Supplementary Figure 22: Diffusion GPR r-statistic .....                                          | 17 |
| Supplementary Figure 23: Diffusion GPR RvE .....                                                  | 18 |
| Supplementary Figure 24: Diffusion GPR r-statistic Q-Q plot.....                                  | 18 |
| Supplementary Figure 25: Diffusion Bayesian GPR r-statistic .....                                 | 19 |
| Supplementary Figure 26: Diffusion Bayesian GPR RvE .....                                         | 19 |
| Supplementary Figure 27: Diffusion Bayesian GPR r-statistic Q-Q plot .....                        | 20 |
| Supplementary Figure 28: Diffusion GPR single predictor r-statistic (bootstrap UQ) .....          | 20 |
| Supplementary Figure 29: Diffusion GPR single predictor RvE (bootstrap UQ) .....                  | 21 |
| Supplementary Figure 30: Diffusion GPR single predictor r-statistic Q-Q plot (bootstrap UQ) ..... | 21 |
| Supplementary Figure 31: Perovskite random forest r-statistic .....                               | 22 |
| Supplementary Figure 32: Perovskite random forest RvE.....                                        | 22 |
| Supplementary Figure 33: Perovskite random forest r-statistic Q-Q plot.....                       | 23 |
| Supplementary Figure 34: Perovskite linear ridge regression r-statistic.....                      | 23 |
| Supplementary Figure 35: Perovskite linear ridge regression RvE .....                             | 24 |
| Supplementary Figure 36: Perovskite linear ridge regression r-statistic Q-Q plot.....             | 24 |
| Supplementary Figure 37: Perovskite Bayesian GPR r-statistic.....                                 | 25 |
| Supplementary Figure 38: Perovskite Bayesian GPR RvE .....                                        | 25 |
| Supplementary Figure 39: Perovskite Bayesian GPR r-statistic Q-Q plot .....                       | 26 |
| Supplementary Figure 40: Friedman with 0.1 sigma gaussian noise r-statistic.....                  | 26 |
| Supplementary Figure 41: Friedman with 0.1 sigma gaussian noise RvE.....                          | 27 |
| Supplementary Figure 42: Friedman with 0.1 sigma gaussian noise r-statistic Q-Q plot.....         | 27 |
| Supplementary Figure 43: Friedman with 0.2 sigma gaussian noise r-statistic.....                  | 28 |
| Supplementary Figure 44: Friedman with 0.2 sigma gaussian noise RvE.....                          | 28 |
| Supplementary Figure 45: Friedman with 0.2 sigma gaussian noise Q-Q plot.....                     | 29 |
| Supplementary Figure 46: Friedman with 0.3 sigma gaussian noise r-statistic.....                  | 29 |
| Supplementary Figure 47: Friedman with 0.3 sigma gaussian noise RvE.....                          | 30 |
| Supplementary Figure 48: Friedman with 0.3 sigma gaussian noise r-statistic Q-Q plot .....        | 30 |

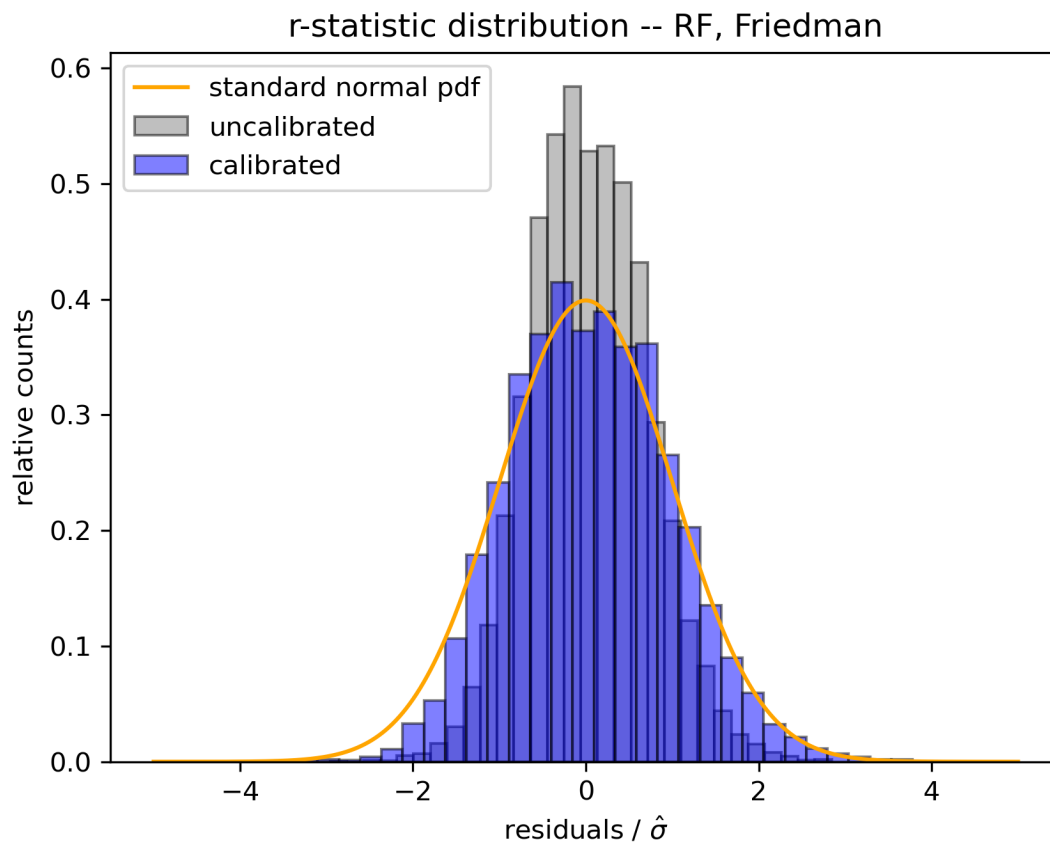
|                                                                                                                                                                                                                                                                                                                                                                                          |    |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| Supplementary Figure 49: Friedman with 0.4 sigma gaussian noise r-statistic.....                                                                                                                                                                                                                                                                                                         | 31 |
| Supplementary Figure 50: Friedman with 0.4 sigma gaussian noise RvE.....                                                                                                                                                                                                                                                                                                                 | 31 |
| Supplementary Figure 51: Friedman with 0.4 sigma gaussian noise r-statistic Q-Q plot.....                                                                                                                                                                                                                                                                                                | 32 |
| Supplementary Figure 52: Friedman with 0.5 sigma gaussian noise r-statistic.....                                                                                                                                                                                                                                                                                                         | 32 |
| Supplementary Figure 53: Friedman with 0.5 sigma gaussian noise RvE.....                                                                                                                                                                                                                                                                                                                 | 33 |
| Supplementary Figure 54: Friedman with 0.5 sigma gaussian noise r-statistic Q-Q plot.....                                                                                                                                                                                                                                                                                                | 33 |
| Supplementary Figure 55: Friedman with 1.0 sigma gaussian noise r-statistic.....                                                                                                                                                                                                                                                                                                         | 34 |
| Supplementary Figure 56: Friedman with 1.0 sigma gaussian noise RvE.....                                                                                                                                                                                                                                                                                                                 | 34 |
| Supplementary Figure 57: Friedman with 1.0 sigma gaussian noise r-statistic Q-Q plot.....                                                                                                                                                                                                                                                                                                | 35 |
| Supplementary Figure 58: Friedman with 2.0 sigma gaussian noise r-statistic.....                                                                                                                                                                                                                                                                                                         | 35 |
| Supplementary Figure 59: Friedman with 2.0 sigma gaussian noise RvE.....                                                                                                                                                                                                                                                                                                                 | 36 |
| Supplementary Figure 60: Friedman with 2.0 sigma gaussian noise r-statistic Q-Q plot.....                                                                                                                                                                                                                                                                                                | 36 |
| Supplementary Figure 61: Convergence of a for Friedman with random forest .....                                                                                                                                                                                                                                                                                                          | 37 |
| Supplementary Figure 62: Convergence of b for Friedman with random forest .....                                                                                                                                                                                                                                                                                                          | 37 |
| Supplementary Figure 63: Convergence of a for Friedman with linear ridge regression.....                                                                                                                                                                                                                                                                                                 | 38 |
| Supplementary Figure 64: Convergence of b for Friedman with linear ridge regression.....                                                                                                                                                                                                                                                                                                 | 38 |
| Supplementary Figure 65: Convergence of a for Friedman with GPR.....                                                                                                                                                                                                                                                                                                                     | 39 |
| Supplementary Figure 66: Convergence of b for Friedman with GPR.....                                                                                                                                                                                                                                                                                                                     | 39 |
| Supplementary Figure 67: Convergence of a for Diffusion with random forest .....                                                                                                                                                                                                                                                                                                         | 40 |
| Supplementary Figure 68: Convergence of b for Diffusion with random forest .....                                                                                                                                                                                                                                                                                                         | 40 |
| Supplementary Figure 69: Convergence of a for Diffusion with linear ridge regression.....                                                                                                                                                                                                                                                                                                | 41 |
| Supplementary Figure 70: Convergence of b for Diffusion with linear ridge regression .....                                                                                                                                                                                                                                                                                               | 41 |
| Supplementary Figure 71: Convergence of a for Diffusion with GPR.....                                                                                                                                                                                                                                                                                                                    | 42 |
| Supplementary Figure 72: Convergence of b for Diffusion with GPR.....                                                                                                                                                                                                                                                                                                                    | 42 |
| Supplementary Figure 73: Convergence of a for Perovskite with random forest .....                                                                                                                                                                                                                                                                                                        | 43 |
| Supplementary Figure 74: Convergence of b for Perovskite with random forest.....                                                                                                                                                                                                                                                                                                         | 43 |
| Supplementary Figure 75: Convergence of a for Perovskite with linear ridge regression .....                                                                                                                                                                                                                                                                                              | 44 |
| Supplementary Figure 76: Convergence of b for Perovskite with linear ridge regression .....                                                                                                                                                                                                                                                                                              | 44 |
| Supplementary Figure 77: Diffusion Neural Network (ensemble of 25) r-statistic .....                                                                                                                                                                                                                                                                                                     | 45 |
| Supplementary Figure 78: Diffusion Neural Network (ensemble of 25) RvE .....                                                                                                                                                                                                                                                                                                             | 45 |
| Supplementary Figure 79: Bulk modulus random forest r-statistic.....                                                                                                                                                                                                                                                                                                                     | 46 |
| Supplementary Figure 80: Bulk modulus random forest RvE.....                                                                                                                                                                                                                                                                                                                             | 46 |
| Supplementary Figure 81: Double perovskite gap random forest r-statistic .....                                                                                                                                                                                                                                                                                                           | 47 |
| Supplementary Figure 82: Double perovskite gap random forest RvE .....                                                                                                                                                                                                                                                                                                                   | 47 |
| Supplementary Figure 83: Heusler random forest r-statistic .....                                                                                                                                                                                                                                                                                                                         | 48 |
| Supplementary Figure 84: Heusler random forest RvE .....                                                                                                                                                                                                                                                                                                                                 | 48 |
| Supplementary Figure 85: Piezo random forest r-statistic.....                                                                                                                                                                                                                                                                                                                            | 49 |
| Supplementary Figure 86: Piezo random forest RvE.....                                                                                                                                                                                                                                                                                                                                    | 49 |
| Supplementary Figure 87: Steel strength random forest r-statistic.....                                                                                                                                                                                                                                                                                                                   | 50 |
| Supplementary Figure 88: Steel strength random forest RvE .....                                                                                                                                                                                                                                                                                                                          | 50 |
| Supplementary Figure 89: Superconductor random forest r-statistic .....                                                                                                                                                                                                                                                                                                                  | 51 |
| Supplementary Figure 90: Superconductor random forest RvE.....                                                                                                                                                                                                                                                                                                                           | 51 |
| Supplementary Figure 91: Thermal conductivity random forest r-statistic .....                                                                                                                                                                                                                                                                                                            | 52 |
| Supplementary Figure 92: Thermal conductivity random forest RvE.....                                                                                                                                                                                                                                                                                                                     | 52 |
| Supplementary Figure 93: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out non-transition metal host” domain test where the Al, Ca and Mg hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data..... | 53 |
| Supplementary Figure 94: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out 3d transition metal host” domain test where the Fe, Cu and Ni hosts were left out and the                                                                                                                                                                  |    |

|                                                                                                                                                                                                                                                                                                                                                                                                     |    |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data.....                                                                                                                                                                                                                                     | 53 |
| Supplementary Figure 95: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out 4d/5d transition metal host” domain test where the Ag, Pd, Ir, Pt and Au hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data. .... | 54 |
| Supplementary Figure 96: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out refractory metal host” domain test where the Mo, W and Zr hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data.....                 | 54 |

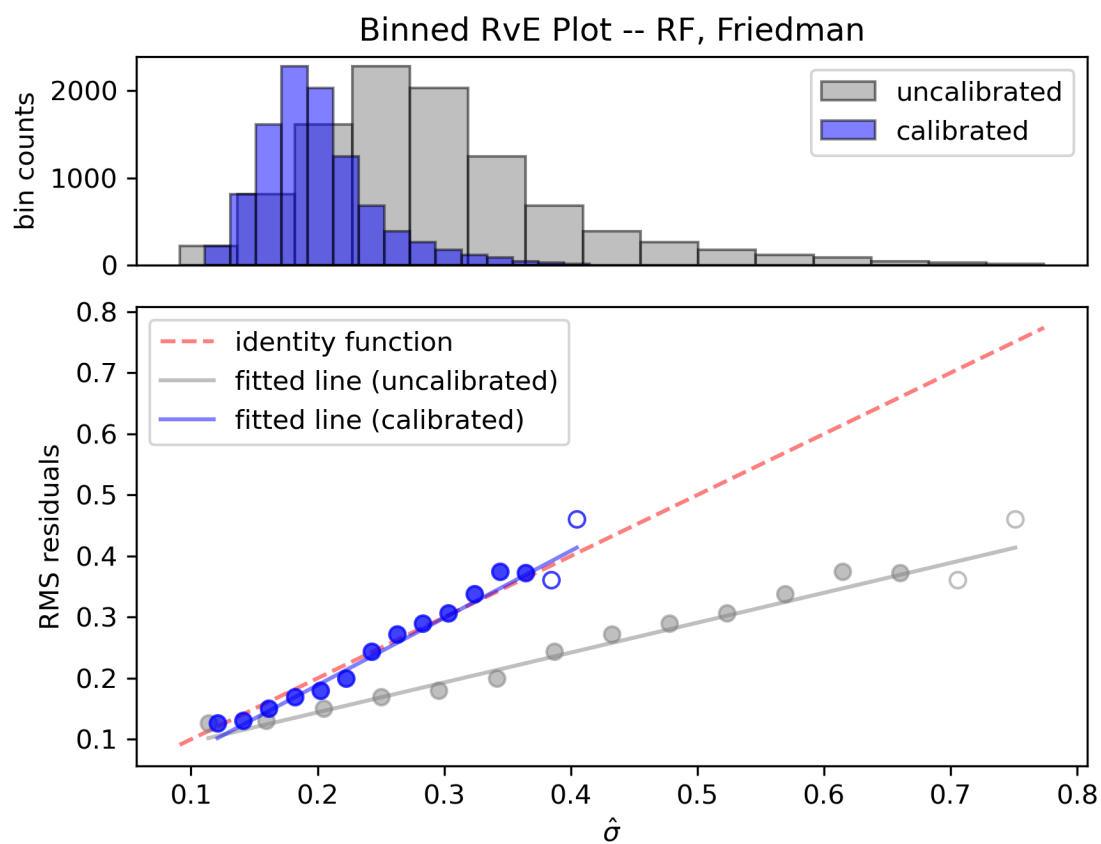
**Supplementary Table 1:** Summary of values from all *r*-statistic and RMS residual vs. uncertainty estimates in the SI. Given are the mean and standard deviation of the *r*-statistic distribution for uncalibrated and calibrated error estimates, the slope, intercept, and *R*<sup>2</sup> values for the linear fits on the RMS residual vs. uncertainty estimate plots, and the calibration factors *a* and *b*, which scale the uncalibrated uncertainty estimates *U*(*x*) as *aU*(*x*)+*b*.

|                                          |              | r-stat<br>mean | r-stat<br>stdev | RvE<br>slope | RvE<br>intercept | RvE <i>R</i> <sup>2</sup> | <i>a</i>            | <i>b</i>             | Overall<br>RMSE |
|------------------------------------------|--------------|----------------|-----------------|--------------|------------------|---------------------------|---------------------|----------------------|-----------------|
| RF,<br>Friedman                          | uncalibrated | 0.043          | 0.675           | 0.489        | 0.046            | 0.964                     | 0.445               | 0.070                | 0.196           |
|                                          | calibrated   | 0.060          | 0.938           | 1.099        | -0.031           | 0.964                     |                     |                      |                 |
| RF,<br>Diffusion                         | uncalibrated | -0.022         | 0.707           | 0.660        | 0.017            | 0.971                     | 0.647<br>±<br>0.047 | 0.033<br>±<br>0.025  | 0.368           |
|                                          | calibrated   | -0.027         | 0.979           | 1.010        | -0.013           | 0.970                     |                     |                      |                 |
| RF,<br>Perovskite                        | uncalibrated | -0.014         | 0.766           | 0.875        | -0.023           | 0.972                     | 0.807<br>±<br>0.038 | 0.002<br>±<br>0.003  | 0.377           |
|                                          | calibrated   | -0.019         | 0.941           | 1.080        | -0.024           | 0.953                     |                     |                      |                 |
| LR,<br>Friedman                          | uncalibrated | 0.645          | 7.878           | 12.182       | -0.274           | 0.582                     | 8.213               | -0.061               | 0.556           |
|                                          | calibrated   | 0.101          | 1.091           | 1.483        | -0.184           | 0.582                     |                     |                      |                 |
| LR,<br>Diffusion                         | uncalibrated | 0.001          | 1.574           | 0.575        | 0.279            | 0.561                     | 0.927<br>±<br>0.115 | 0.154<br>±<br>0.041  | 0.481           |
|                                          | calibrated   | 0.002          | 1.058           | 0.597        | 0.193            | 0.521                     |                     |                      |                 |
| LR,<br>Perovskite                        | uncalibrated | -0.080         | 4.566           | 2.567        | 0.240            | 0.698                     | 4.447<br>±<br>0.154 | -0.034<br>±<br>0.017 | 0.637           |
|                                          | calibrated   | -0.021         | 1.107           | 0.580        | 0.258            | 0.685                     |                     |                      |                 |
| GPR,<br>Friedman                         | uncalibrated | 0.093          | 1.582           | 2.027        | -0.008           | 0.929                     | 1.875               | -0.006               | 0.079           |
|                                          | calibrated   | 0.061          | 0.984           | 1.081        | -0.001           | 0.929                     |                     |                      |                 |
| GPR,<br>Diffusion                        | uncalibrated | -0.094         | 2.182           | 1.779        | 0.047            | 0.963                     | 1.867<br>±<br>0.207 | 0.048<br>±<br>0.029  | 0.310           |
|                                          | calibrated   | -0.039         | 0.963           | 0.905        | 0.015            | 0.929                     |                     |                      |                 |
| GPR<br>Bayesian,<br>Friedman             | uncalibrated | 0.041          | 0.542           | 1.505        | -0.069           | 0.891                     | 0.839               | -0.019               | 0.059           |
|                                          | calibrated   | 0.064          | 0.907           | 1.794        | -0.035           | 0.891                     |                     |                      |                 |
| GPR<br>Bayesian,<br>Diffusion            | uncalibrated | 0.001          | 0.873           | 0.151        | 0.212            | 0.154                     | 0.214<br>±<br>0.122 | 0.227<br>±<br>0.043  | 0.269           |
|                                          | calibrated   | -0.001         | 0.901           | 0.343        | 0.164            | 0.149                     |                     |                      |                 |
| GPR<br>Bayesian,<br>Perovskite           | uncalibrated | -0.006         | 1.097           | 0.850        | 0.049            | 0.814                     | 1.114<br>±<br>0.036 | 0.000<br>±<br>0.000  | 0.310           |
|                                          | calibrated   | -0.005         | 0.992           | 0.756        | 0.052            | 0.829                     |                     |                      |                 |
| GPR<br>Single<br>Predictor,<br>Friedman  | uncalibrated | 0.111          | 1.184           | 1.462        | -0.004           | 0.868                     | 1.440               | -0.006               | 0.059           |
|                                          | calibrated   | 0.106          | 1.008           | 1.015        | 0.002            | 0.868                     |                     |                      |                 |
| GPR<br>Single<br>Predictor,<br>Diffusion | uncalibrated | -0.078         | 2.004           | 1.101        | 0.106            | 0.869                     | 1.545<br>±<br>0.208 | 0.071<br>±<br>0.031  | 0.269           |
|                                          | calibrated   | -0.032         | 0.932           | 0.650        | 0.073            | 0.846                     |                     |                      |                 |
| RF,<br>Friedman,<br>0.1 noise            | uncalibrated | 0.037          | 0.775           | 0.394        | 0.101            | 0.926                     | 0.394               | 0.105                | 0.225           |
|                                          | calibrated   | 0.047          | 0.979           | 1.001        | -0.005           | 0.926                     |                     |                      |                 |
| RF,<br>Friedman,<br>0.2 noise            | uncalibrated | 0.020          | 0.914           | 0.292        | 0.187            | 0.914                     | 0.290               | 0.194                | 0.288           |
|                                          | calibrated   | 0.008          | 0.977           | 1.008        | -0.009           | 0.914                     |                     |                      |                 |
| RF,<br>Friedman,<br>0.3 noise            | uncalibrated | 0.038          | 1.017           | 0.279        | 0.256            | 0.804                     | 0.253               | 0.264                | 0.366           |
|                                          | calibrated   | 0.036          | 1.004           | 1.103        | -0.035           | 0.804                     |                     |                      |                 |
| RF,<br>Friedman,<br>0.4 noise            | uncalibrated | -0.072         | 0.978           | 0.188        | 0.347            | 0.750                     | 0.202               | 0.369                | 0.435           |
|                                          | calibrated   | -0.068         | 0.936           | 0.928        | 0.005            | 0.750                     |                     |                      |                 |

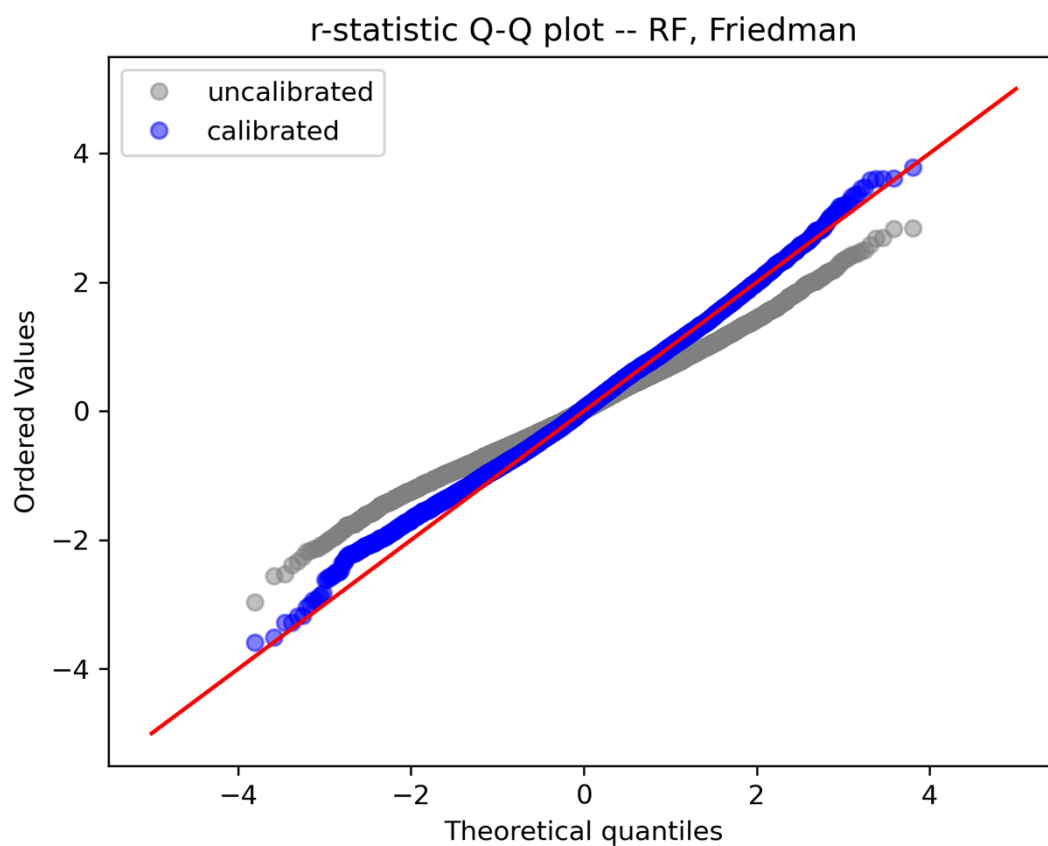
|                               |              |        |       |        |        |       |                     |                      |       |
|-------------------------------|--------------|--------|-------|--------|--------|-------|---------------------|----------------------|-------|
| RF,<br>Friedman,<br>0.5 noise | uncalibrated | 0.039  | 1.081 | 0.086  | 0.451  | 0.166 | 0.173               | 0.405                | 0.493 |
|                               | calibrated   | 0.026  | 1.009 | 0.495  | 0.250  | 0.166 |                     |                      |       |
| RF,<br>Friedman,<br>1.0 noise | uncalibrated | 0.046  | 1.137 | 0.056  | 0.702  | 0.115 | 0.017               | 0.744                | 0.741 |
|                               | calibrated   | 0.036  | 0.980 | 3.316  | -1.764 | 0.115 |                     |                      |       |
| RF,<br>Friedman,<br>2.0 noise | uncalibrated | -0.030 | 1.107 | 0.193  | 0.789  | 0.577 | -0.021              | 0.978                | 0.964 |
|                               | calibrated   | -0.011 | 1.006 | -9.036 | 9.622  | 0.577 |                     |                      |       |
| NN,<br>Diffusion              | uncalibrated | -0.101 | 2.309 | 1.53   | 0.10   | 0.88  | 1.874<br>±<br>0.216 | 0.042<br>±<br>0.030  | 0.325 |
|                               | calibrated   | -0.046 | 1.040 | 0.79   | 0.07   | 0.91  |                     |                      |       |
| RF, Bulk<br>modulus           | uncalibrated | 0.026  | 1.102 | 0.78   | 0.11   | 0.88  | 0.700<br>±<br>0.084 | 0.139<br>±<br>0.044  | 0.491 |
|                               | calibrated   | 0.023  | 1.034 | 1.06   | -0.01  | 0.90  |                     |                      |       |
| RF,<br>Double<br>perov. gap   | uncalibrated | 0.007  | 0.774 | 0.72   | 0.01   | 0.99  | 0.747<br>±<br>0.039 | 0.002<br>±<br>0.009  | 0.236 |
|                               | calibrated   | 0.009  | 1.026 | 0.96   | 0.01   | 0.98  |                     |                      |       |
| RF,<br>Heusler                | uncalibrated | 0.008  | 1.040 | 1.36   | -0.14  | 0.93  | 1.073<br>±<br>0.059 | -0.029<br>±<br>0.024 | 0.644 |
|                               | calibrated   | 0.008  | 1.042 | 1.25   | -0.09  | 0.93  |                     |                      |       |
| RF, Piezo                     | uncalibrated | 0.018  | 1.356 | 0.57   | 0.38   | 0.62  | 0.364<br>±<br>0.060 | 0.526<br>±<br>0.051  | 0.804 |
|                               | calibrated   | -0.003 | 1.016 | 1.24   | -0.20  | 0.42  |                     |                      |       |
| RF, Steel<br>strength         | uncalibrated | 0.005  | 1.118 | 0.93   | 0.06   | 0.93  | 0.905<br>±<br>0.068 | 0.043<br>±<br>0.021  | 0.493 |
|                               | calibrated   | 0.003  | 1.069 | 1.03   | 0.01   | 0.94  |                     |                      |       |
| RF,<br>Super-<br>conductor    | uncalibrated | 0.010  | 1.051 | 0.85   | 0.04   | 1.00  | 0.852<br>±<br>0.019 | 0.035<br>±<br>0.006  | 0.372 |
|                               | calibrated   | 0.007  | 1.005 | 1.00   | 0.01   | 1.00  |                     |                      |       |
| RF,<br>Thermal<br>conduct.    | uncalibrated | -0.082 | 2.338 | 0.86   | 0.21   | 0.94  | 1.042<br>±<br>0.125 | 0.151<br>±<br>0.031  | 0.511 |
|                               | calibrated   | -0.032 | 1.058 | 0.81   | 0.10   | 0.93  |                     |                      |       |



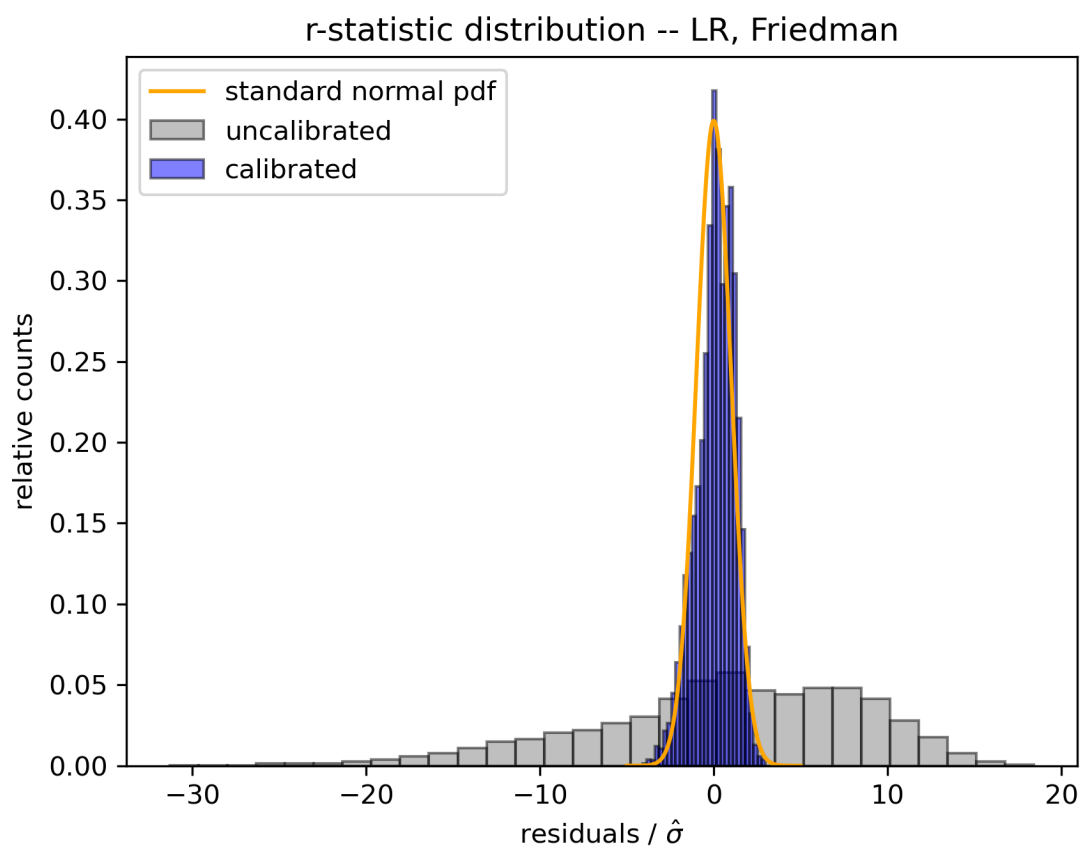
Supplementary Figure 1: Friedman random forest r-statistic



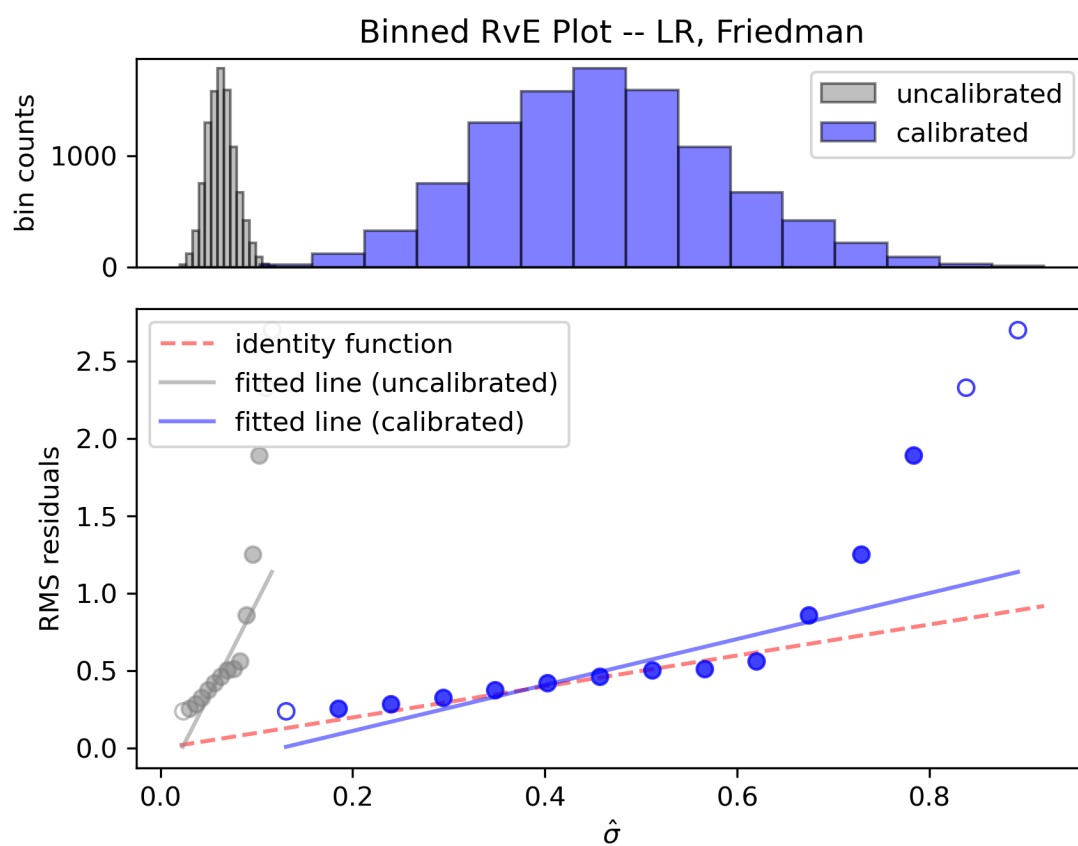
Supplementary Figure 2: Friedman random forest RvE



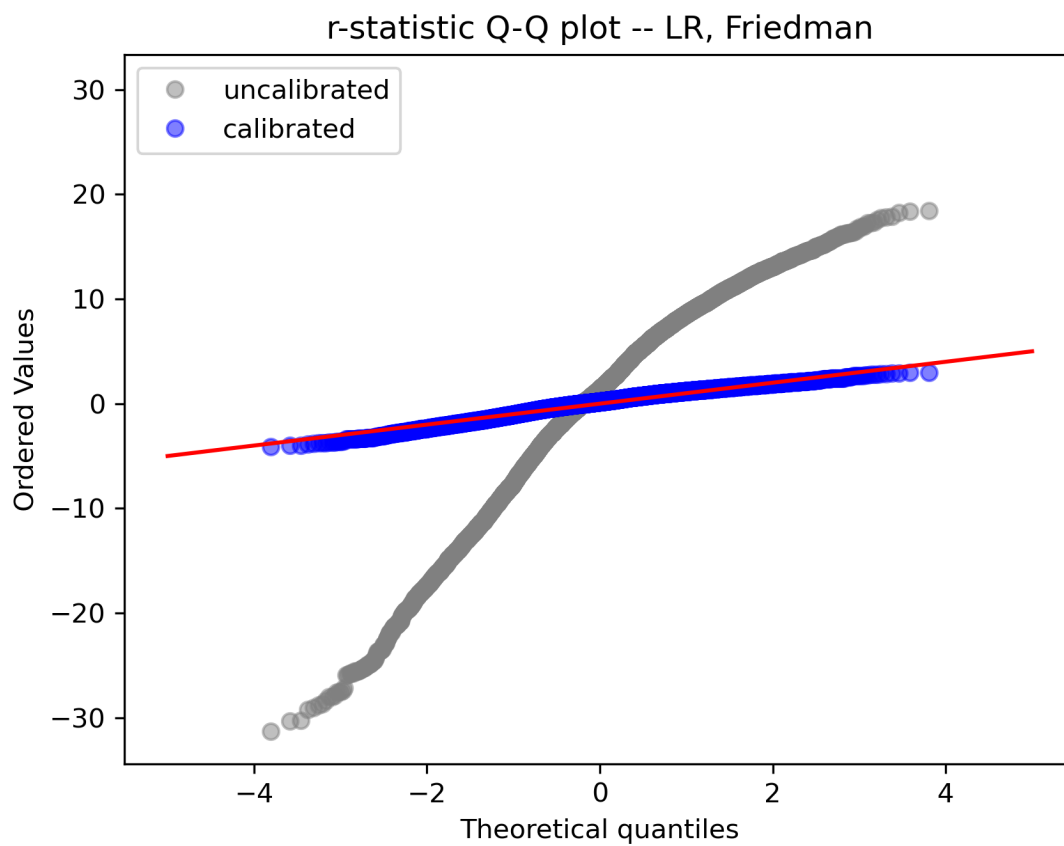
Supplementary Figure 3: Friedman random forest r-statistic Q-Q plot



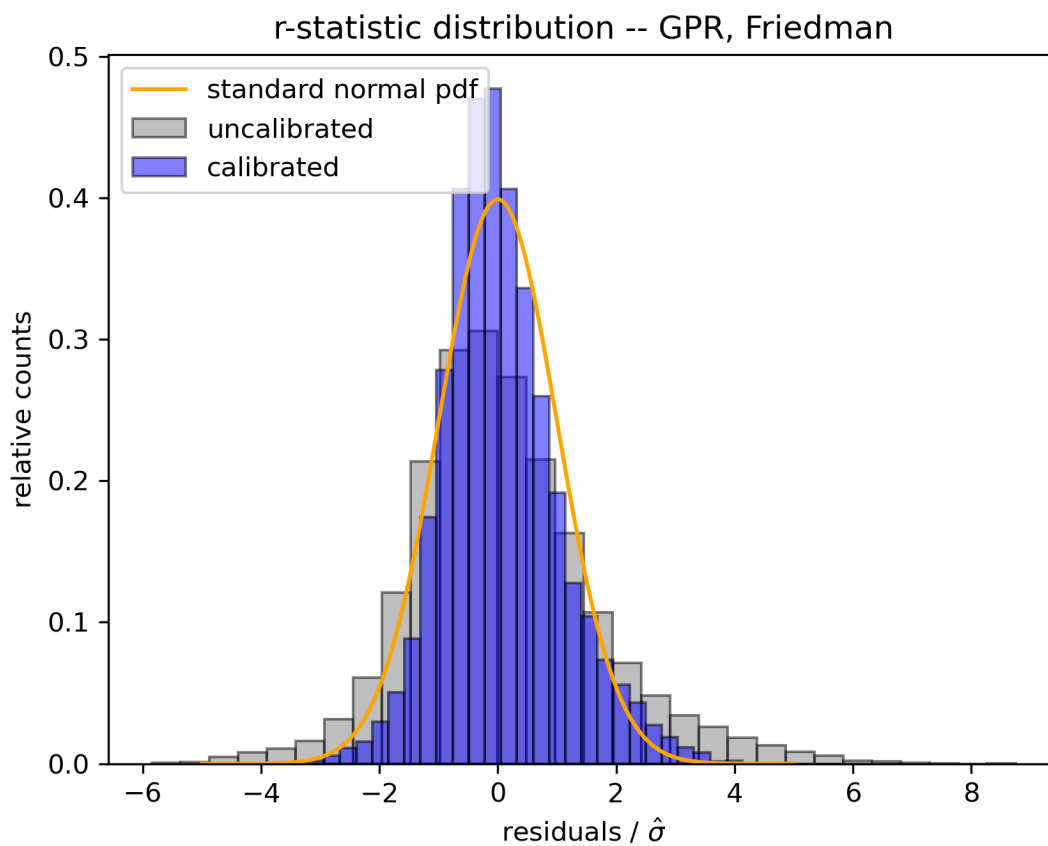
Supplementary Figure 4: Friedman linear ridge regression r-statistic



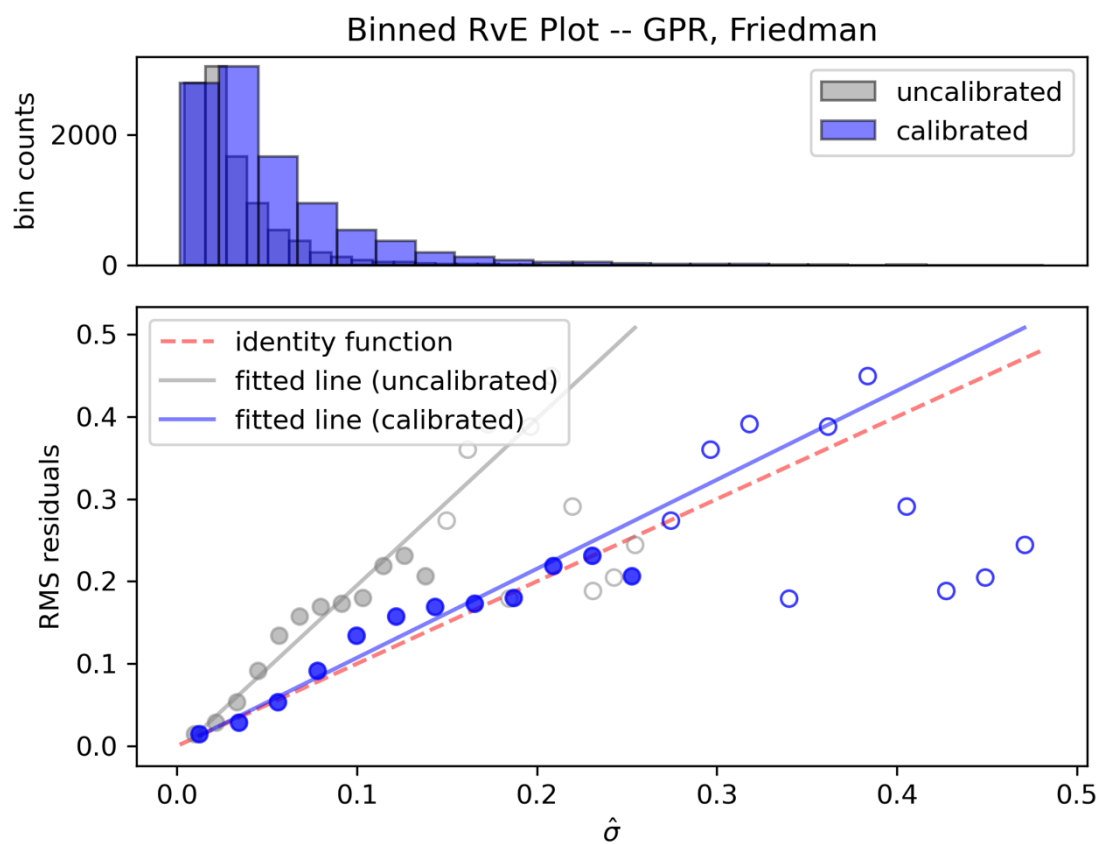
Supplementary Figure 5: Friedman linear ridge regression RvE



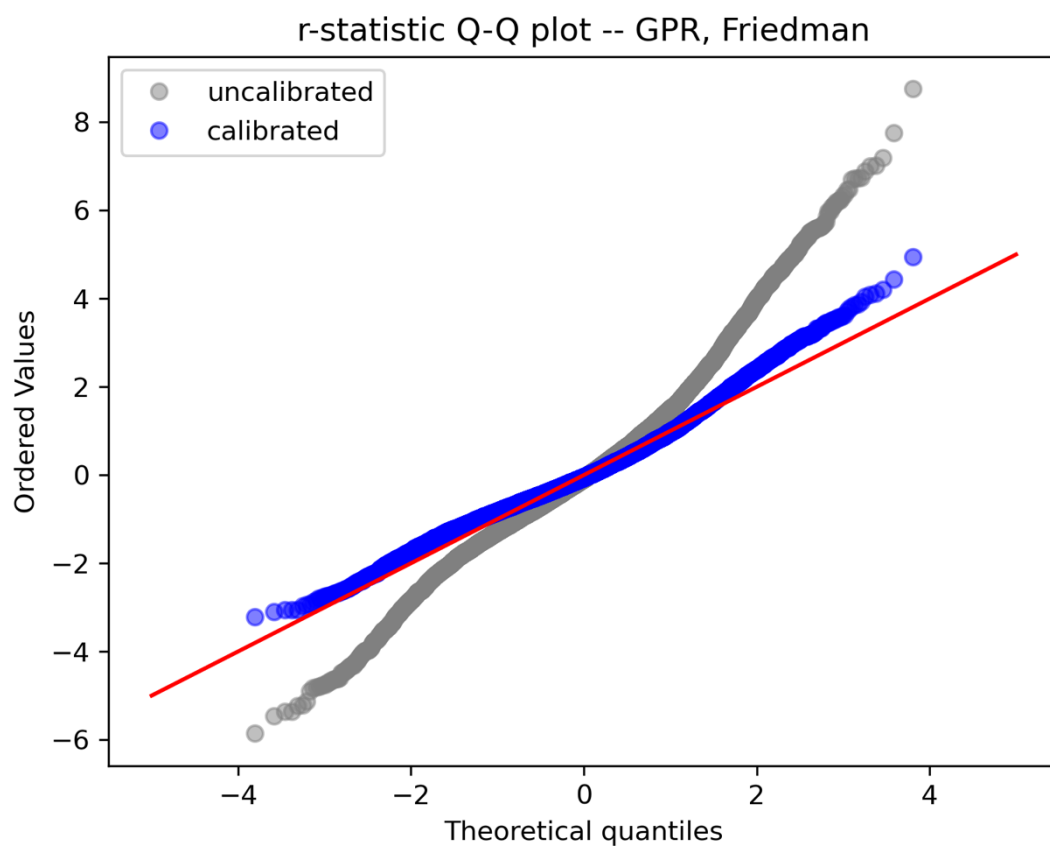
Supplementary Figure 6: Friedman linear ridge regression r-statistic Q-Q plot



Supplementary Figure 7: Friedman GPR r-statistic

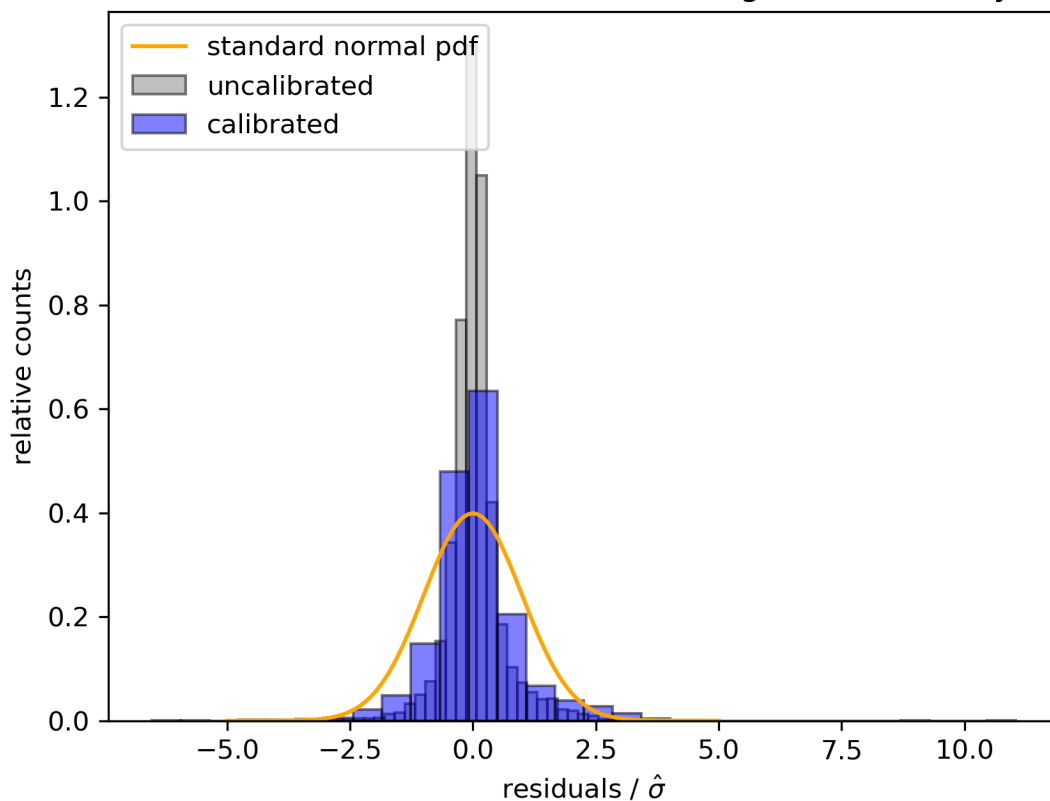


Supplementary Figure 8: Friedman GPR RvE



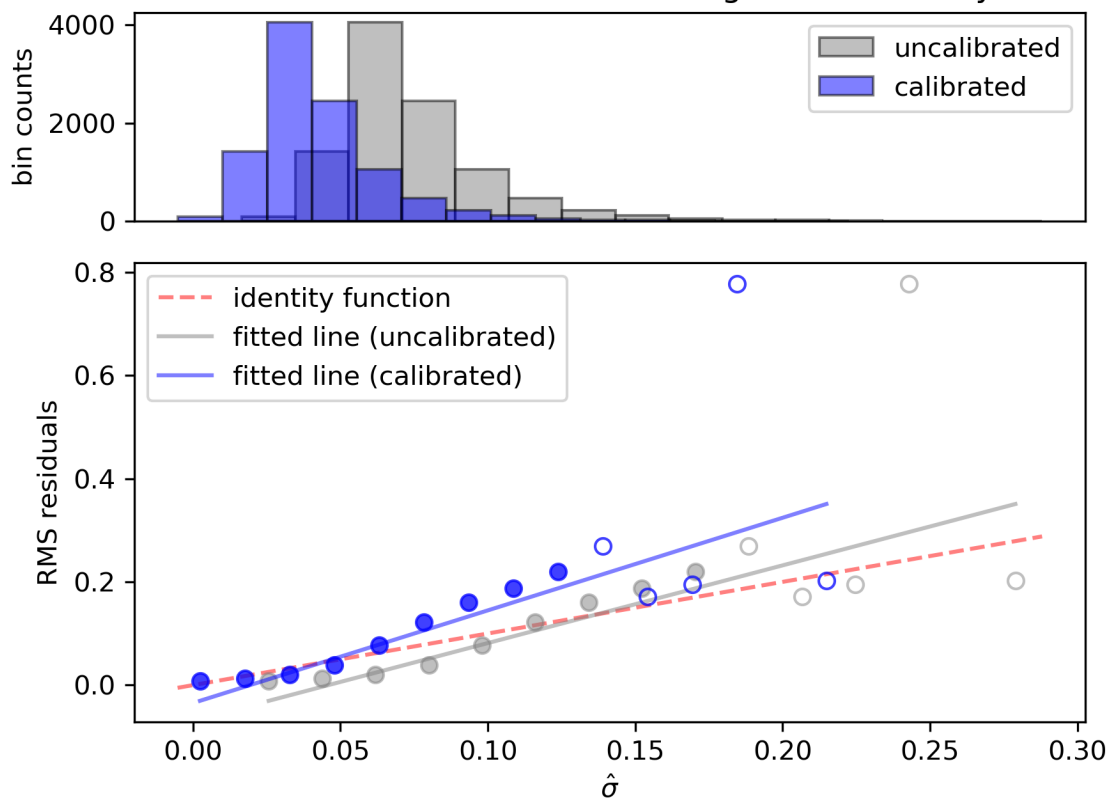
Supplementary Figure 9: Friedman GPR r-statistic Q-Q plot

### r-statistic distribution -- GPR Friedman Single Predictor Bayesian



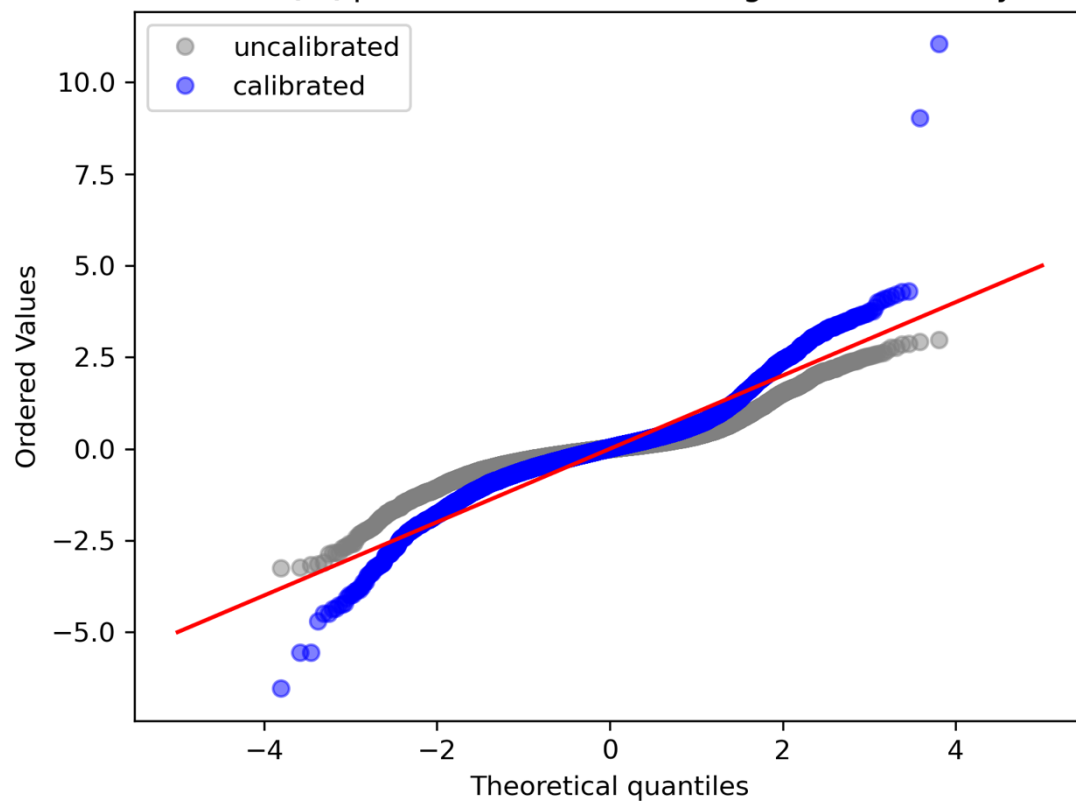
Supplementary Figure 10: Friedman Bayesian GPR r-statistic

### Binned RvE Plot -- GPR Friedman Single Predictor Bayesian



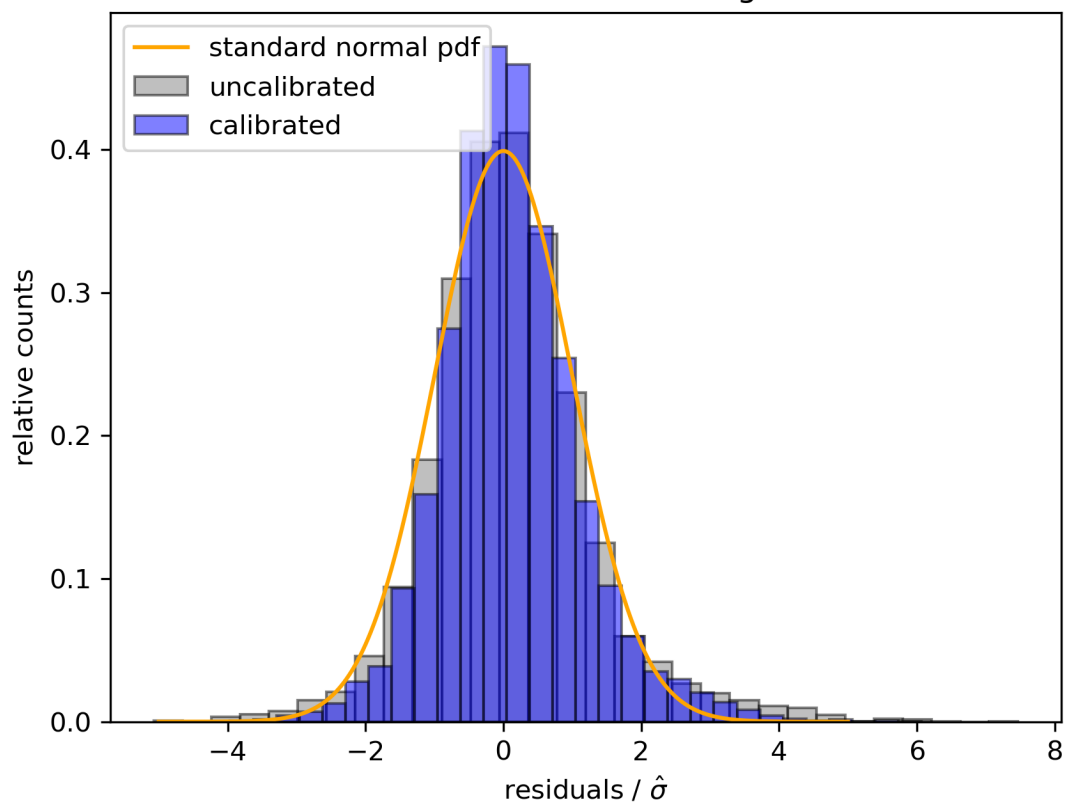
Supplementary Figure 11: Friedman Bayesian GPR RvE

r-statistic Q-Q plot -- GPR Friedman Single Predictor Bayesian

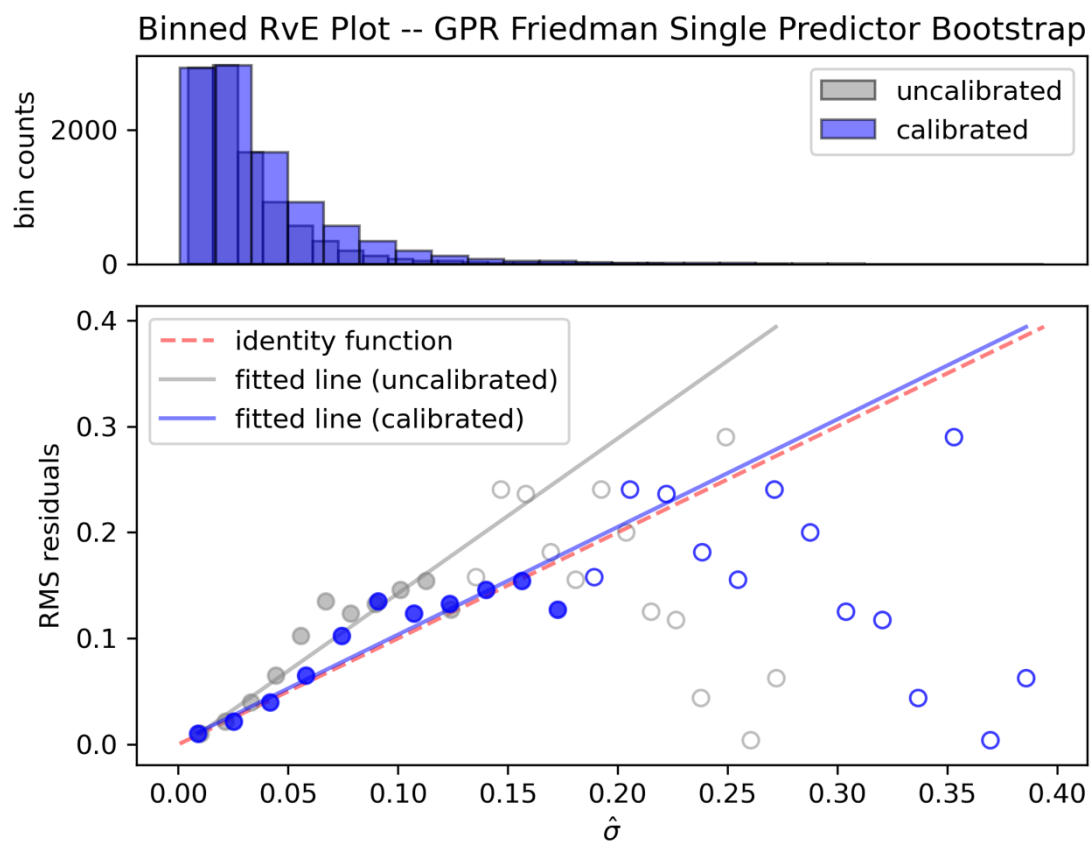


Supplementary Figure 12: Friedman Bayesian GPR r-statistic Q-Q plot

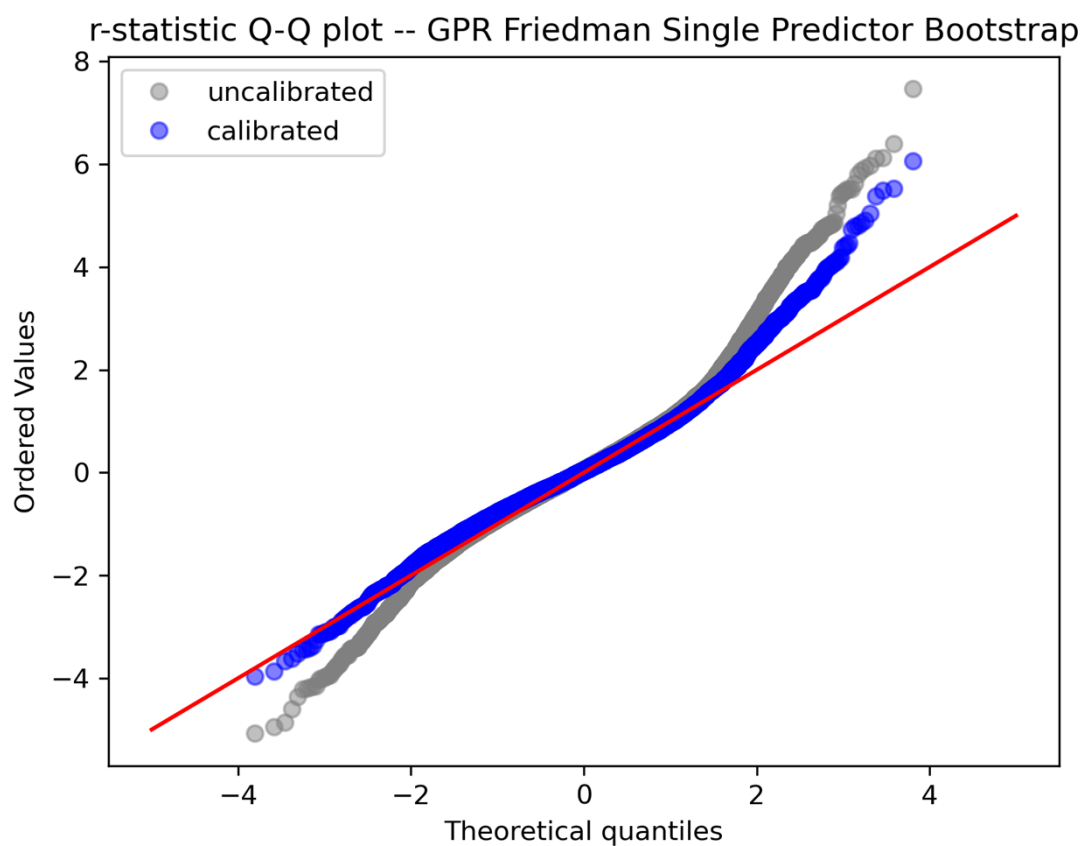
r-statistic distribution -- GPR Friedman Single Predictor Bootstrap



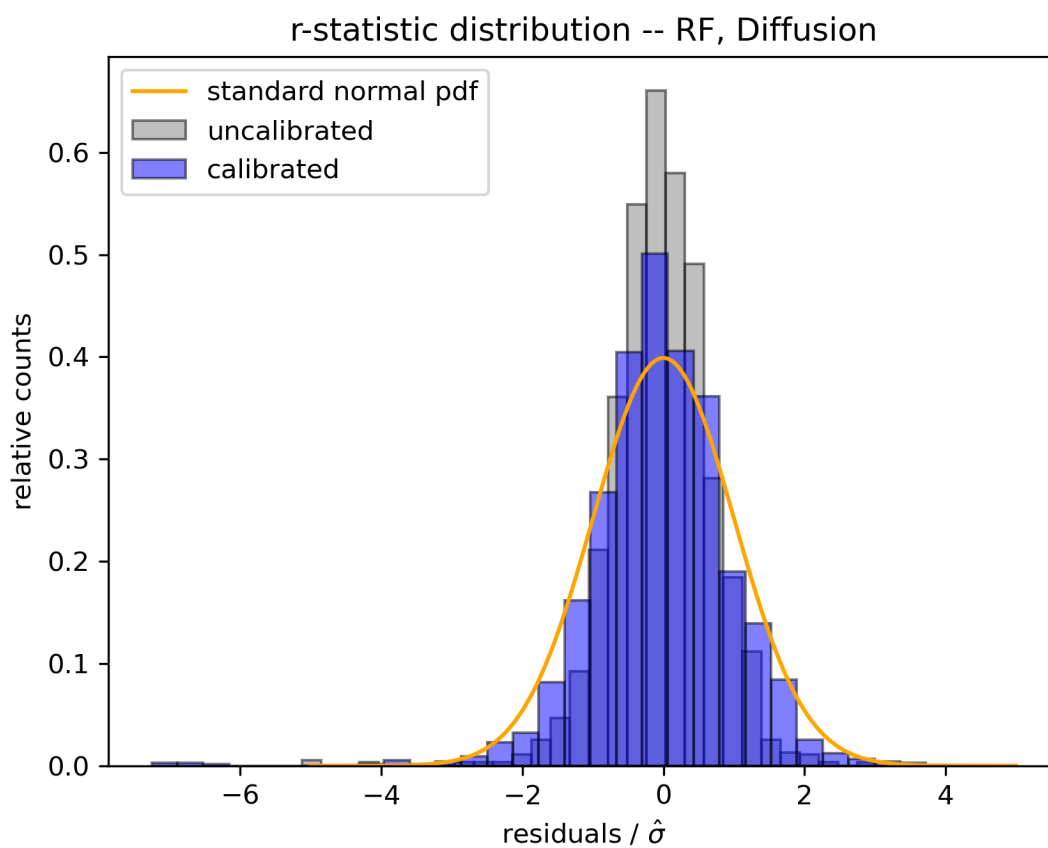
Supplementary Figure 13: Friedman GPR single predictor r-statistic (bootstrap UQ)



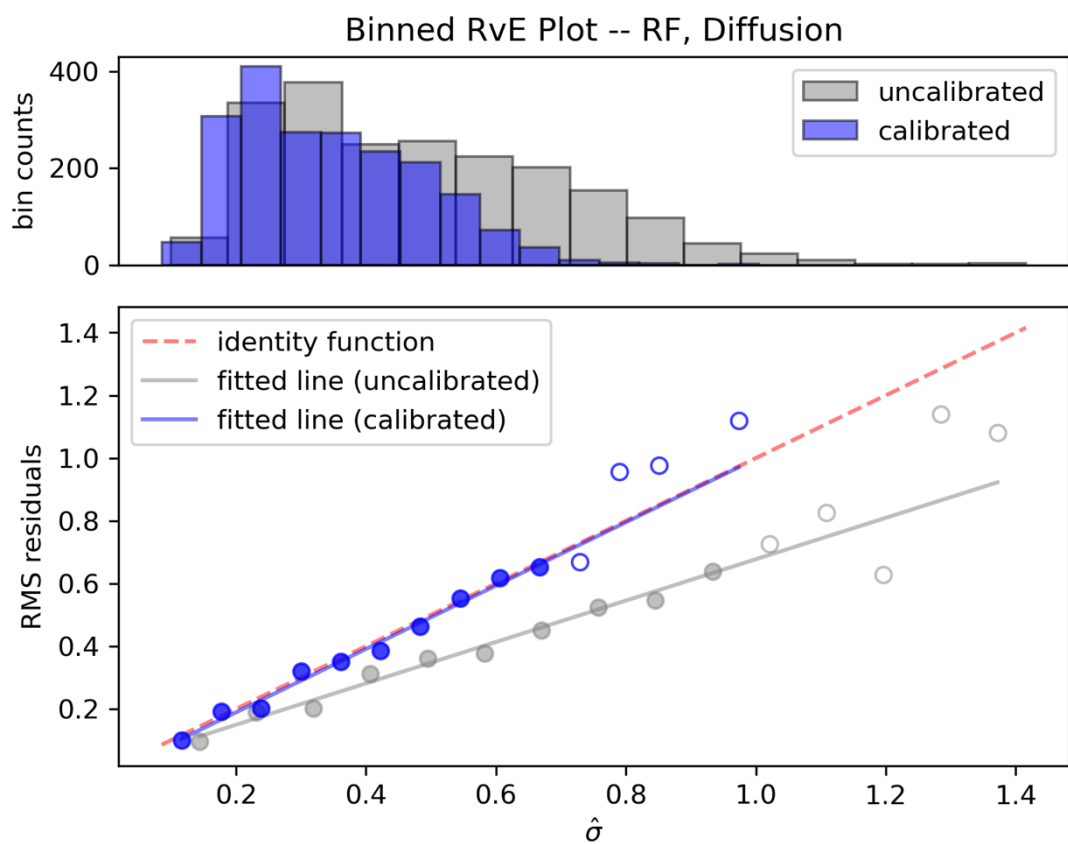
Supplementary Figure 14: Friedman GPR single predictor RvE (bootstrap UQ)



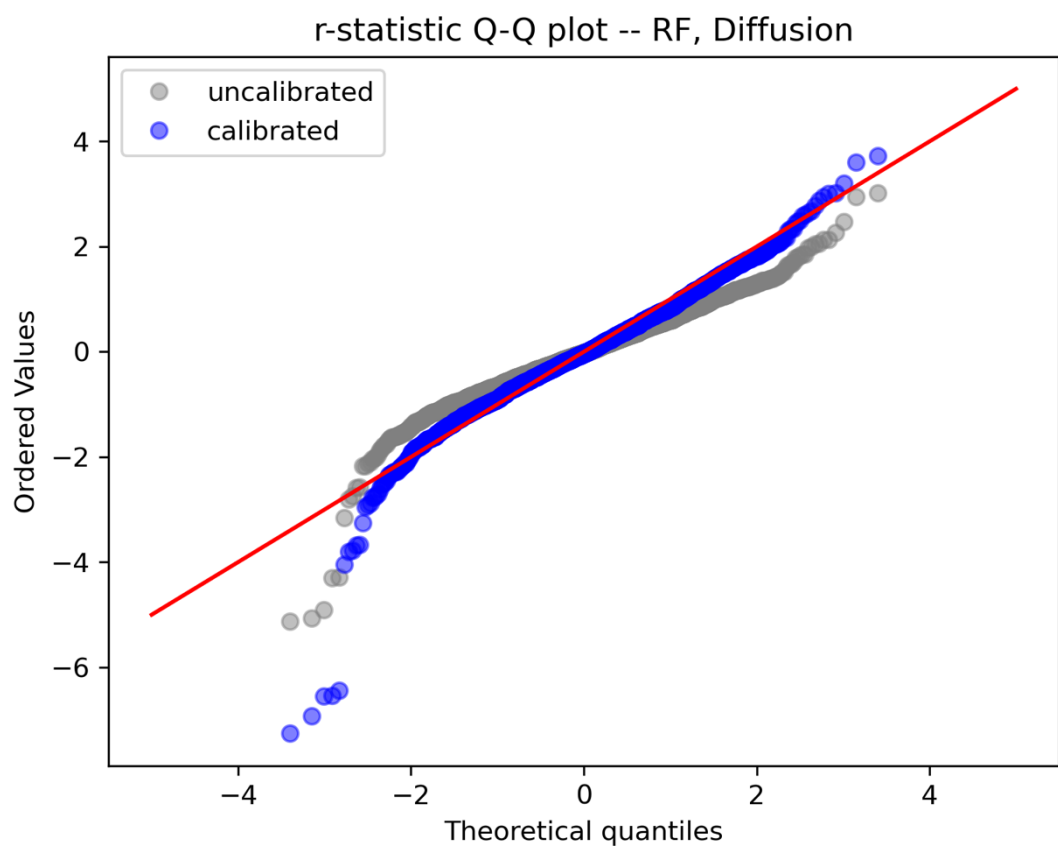
Supplementary Figure 15: Friedman GPR single predictor  $r$ -statistic Q-Q plot (bootstrap UQ)



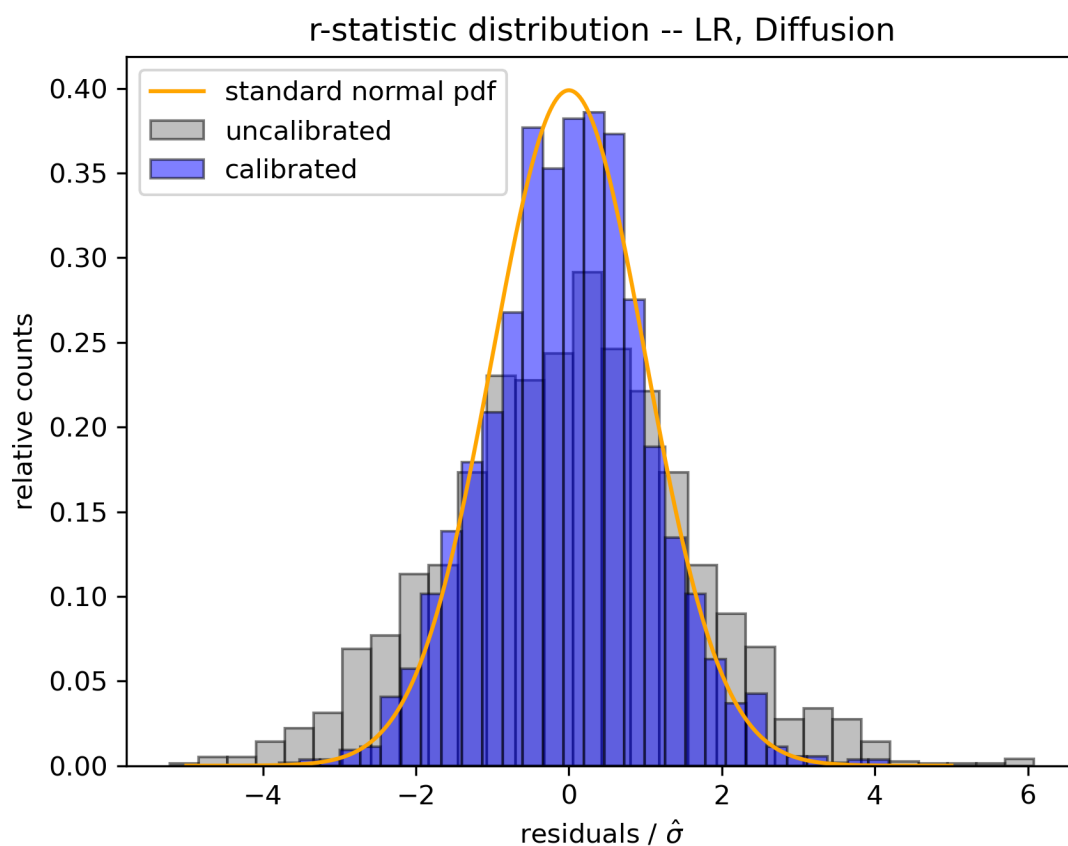
Supplementary Figure 16: Diffusion random forest  $r$ -statistic



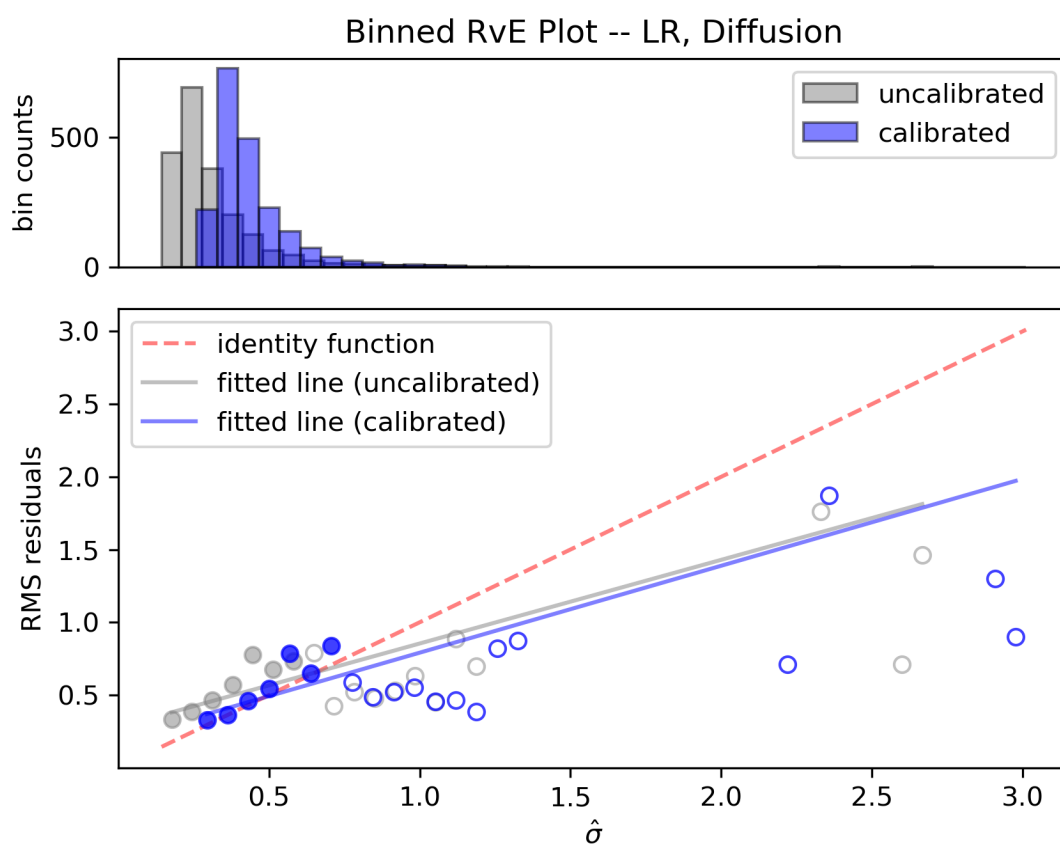
Supplementary Figure 17: Diffusion random forest RvE



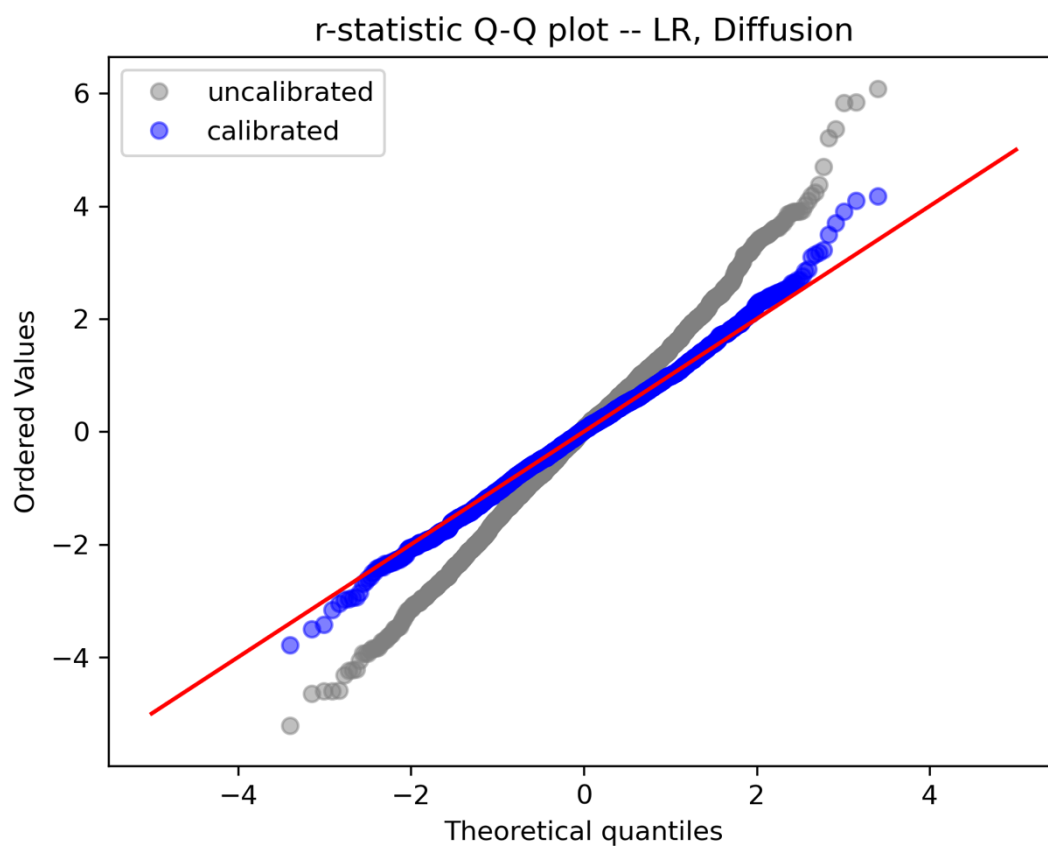
Supplementary Figure 18: Diffusion random forest r-statistic Q-Q plot



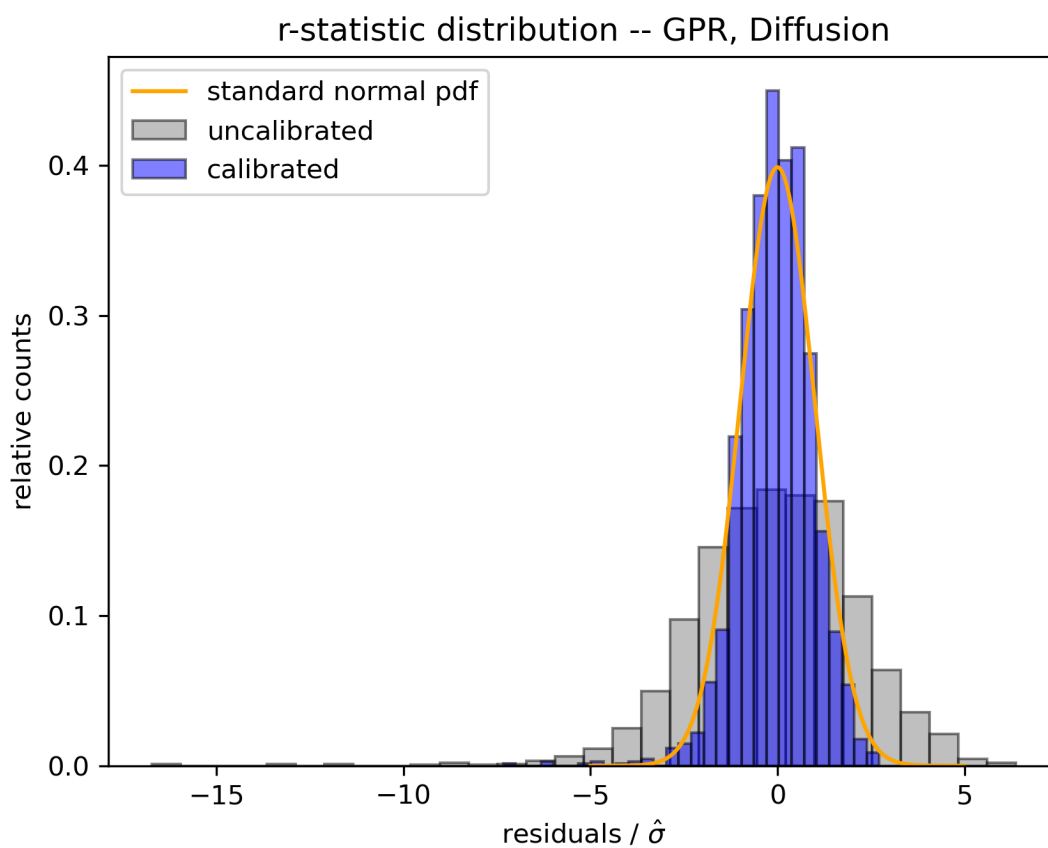
Supplementary Figure 19: Diffusion linear ridge regression  $r$ -statistic



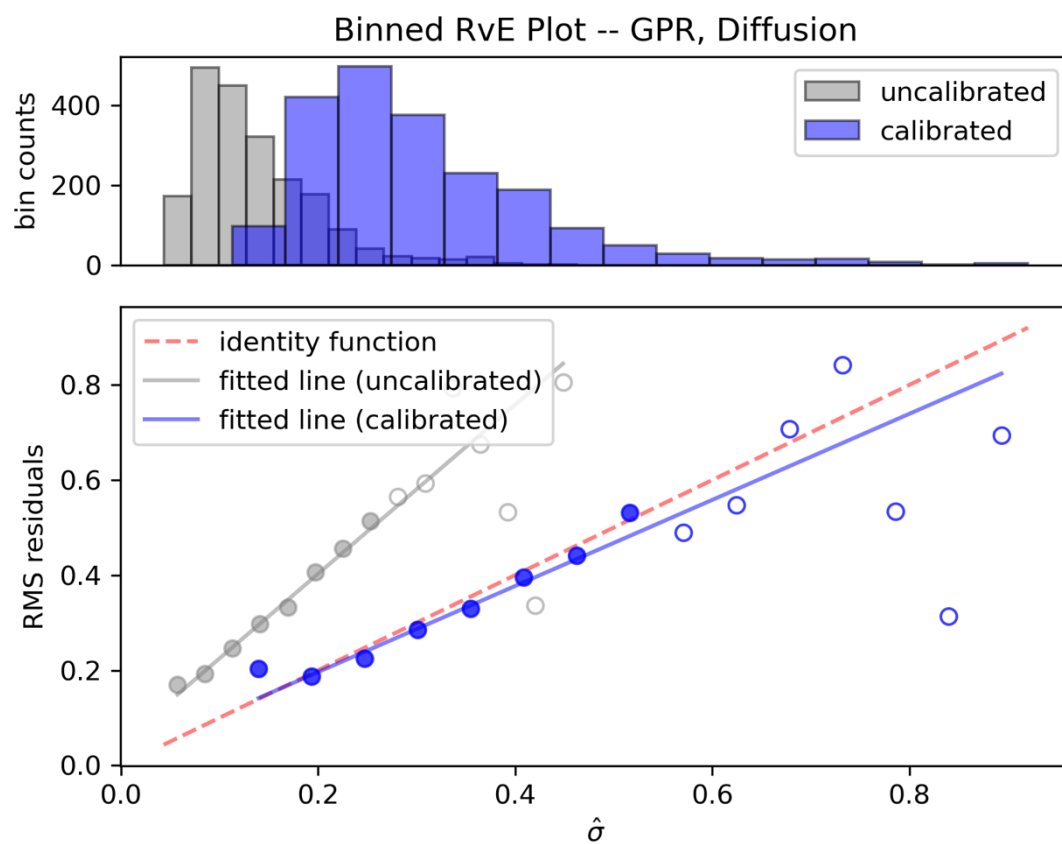
Supplementary Figure 20: Diffusion linear ridge regression RvE



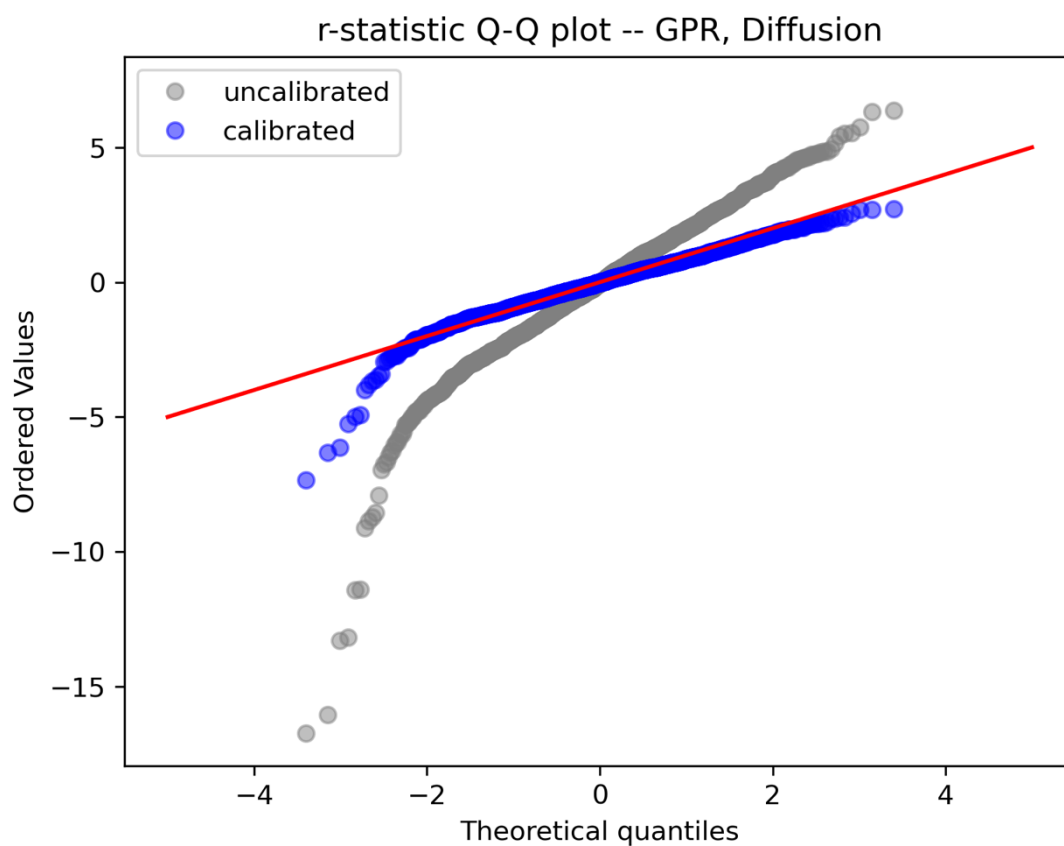
Supplementary Figure 21: Diffusion linear ridge regression *r*-statistic Q-Q plot



Supplementary Figure 22: Diffusion GPR *r*-statistic

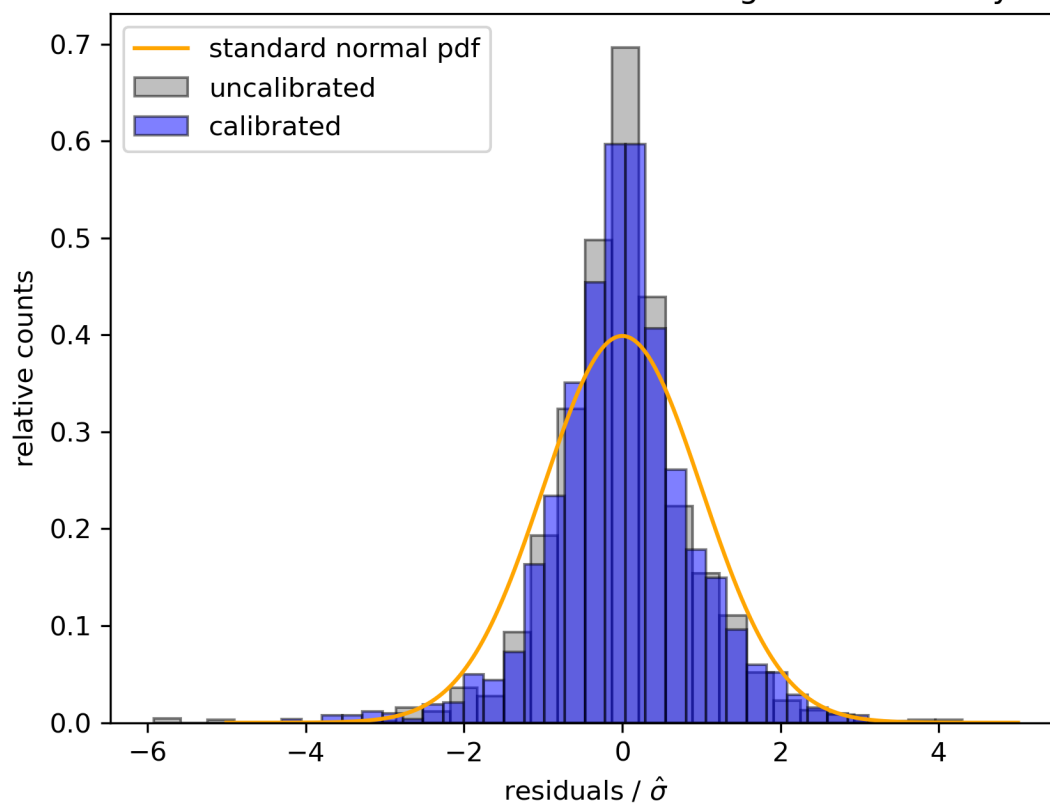


Supplementary Figure 23: Diffusion GPR RvE



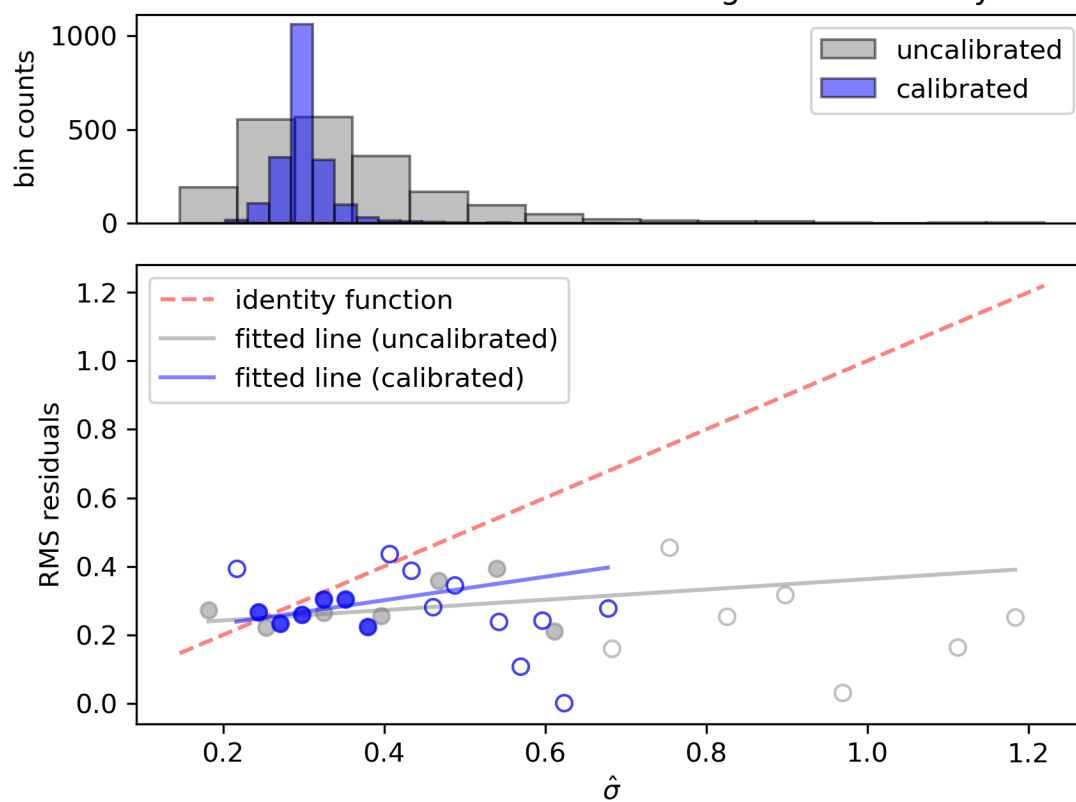
Supplementary Figure 24: Diffusion GPR r-statistic Q-Q plot

# r-statistic distribution -- GPR Diffusion Single Predictor Bayesian



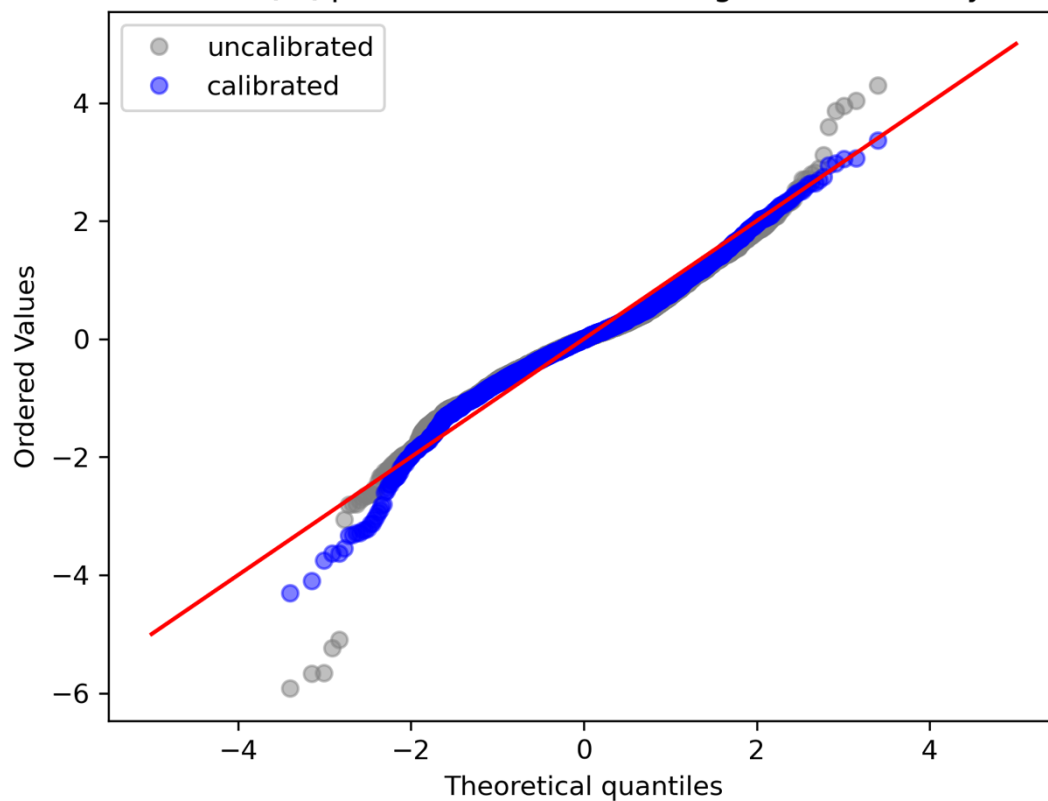
Supplementary Figure 25: Diffusion Bayesian GPR r-statistic

# Binned RvE Plot -- GPR Diffusion Single Predictor Bayesian



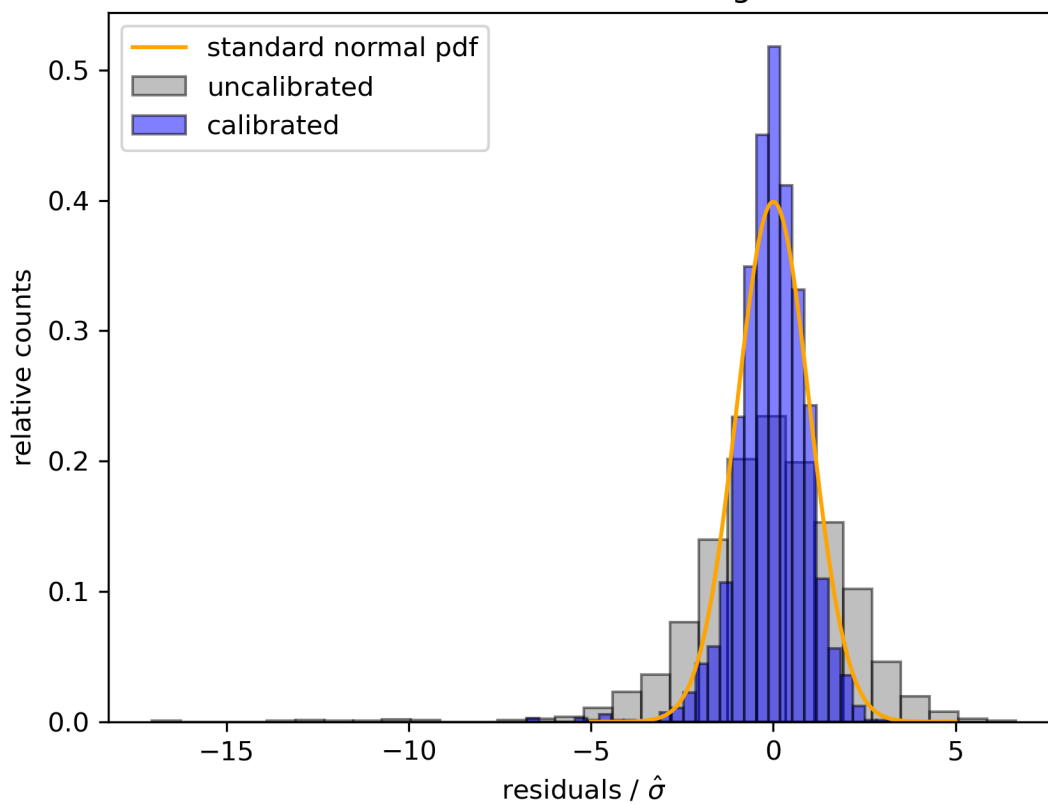
Supplementary Figure 26: Diffusion Bayesian GPR RvE

r-statistic Q-Q plot -- GPR Diffusion Single Predictor Bayesian

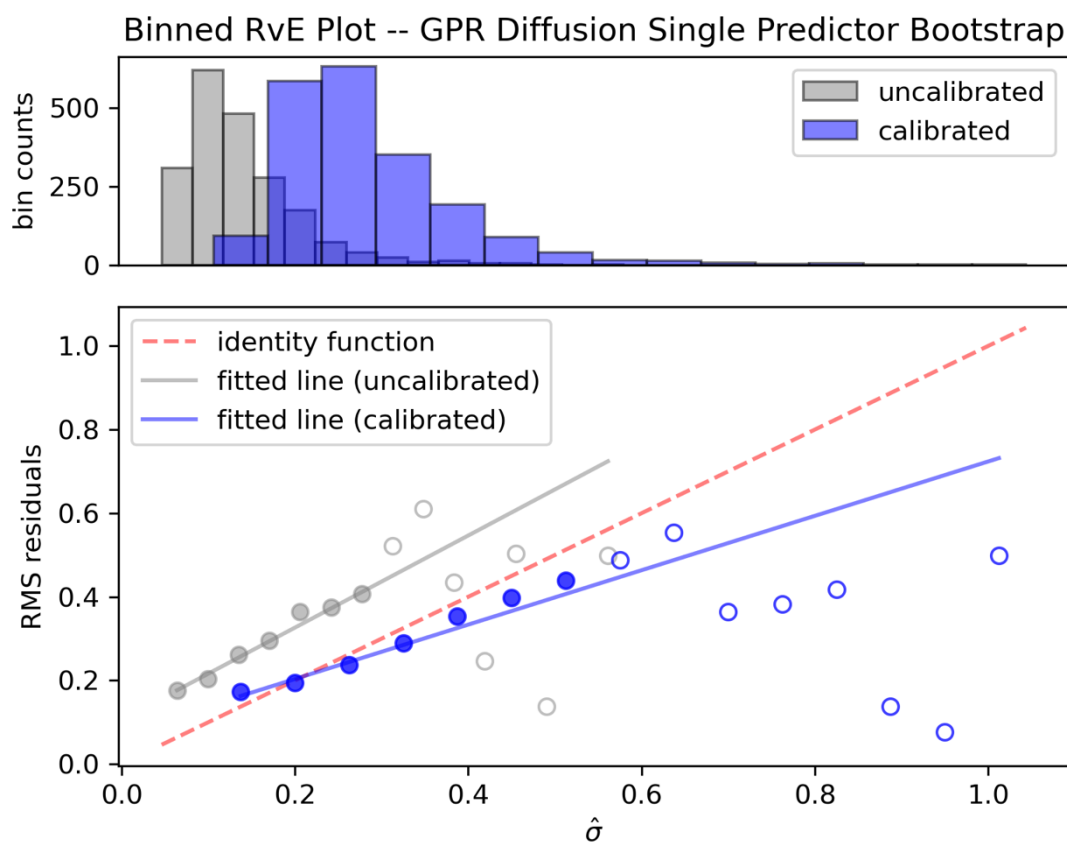


Supplementary Figure 27: Diffusion Bayesian GPR r-statistic Q-Q plot

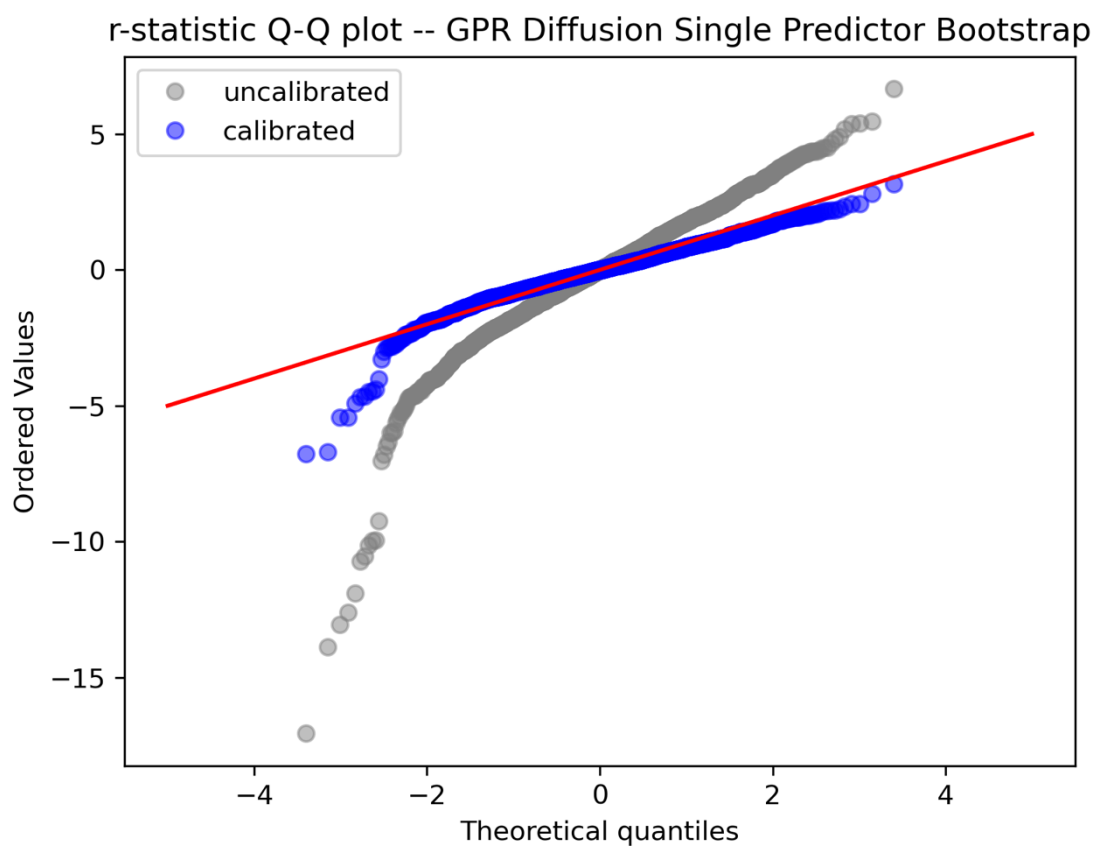
r-statistic distribution -- GPR Diffusion Single Predictor Bootstrap



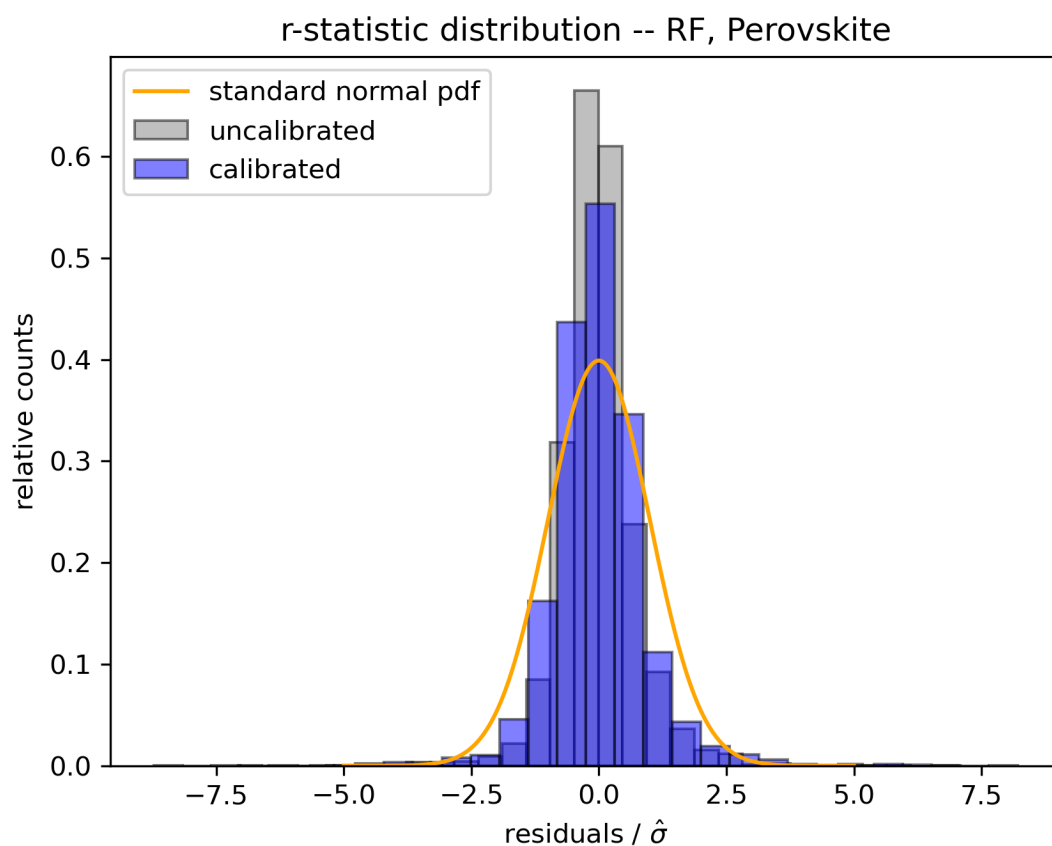
Supplementary Figure 28: Diffusion GPR single predictor r-statistic (bootstrap UQ)



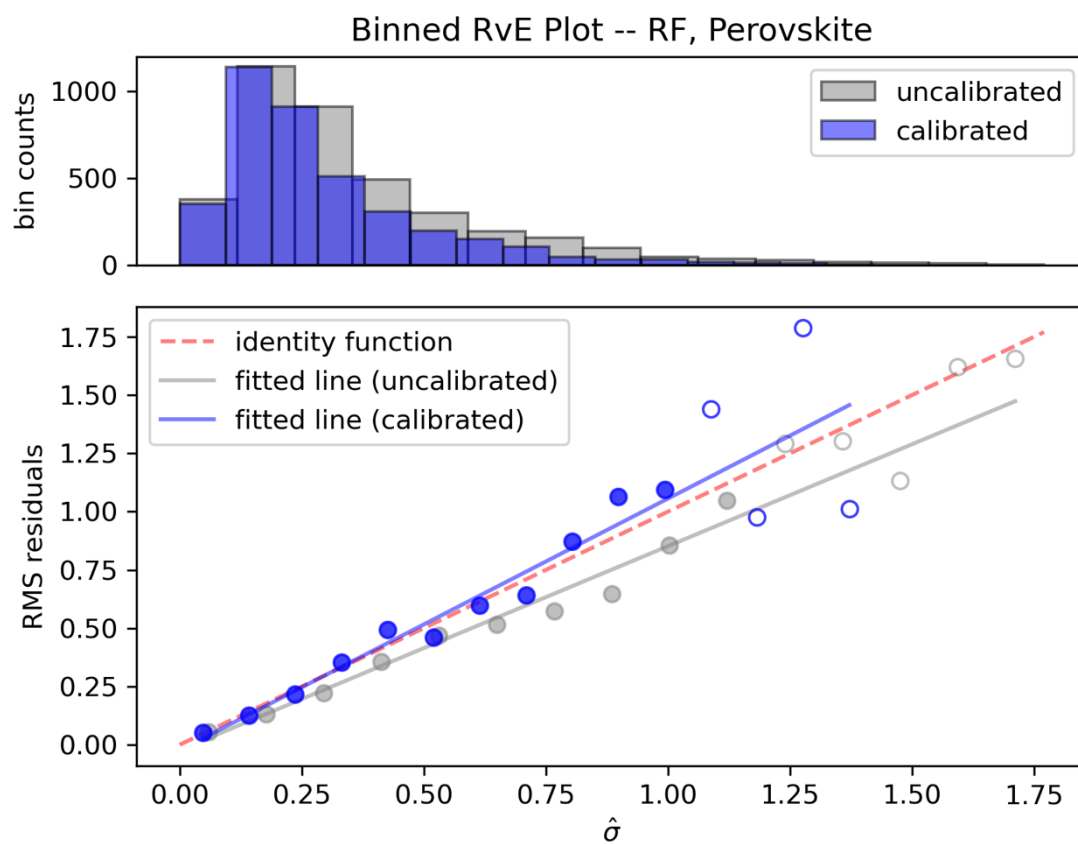
Supplementary Figure 29: Diffusion GPR single predictor RvE (bootstrap UQ)



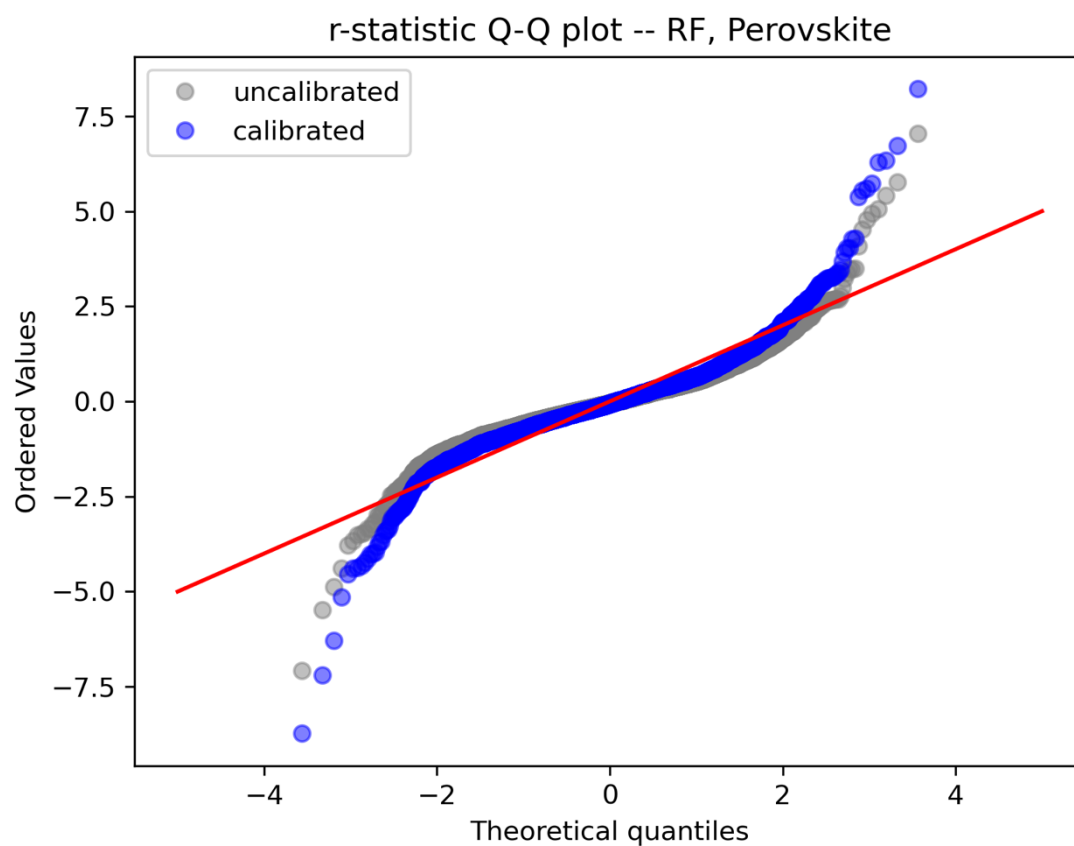
Supplementary Figure 30: Diffusion GPR single predictor r-statistic Q-Q plot (bootstrap UQ)



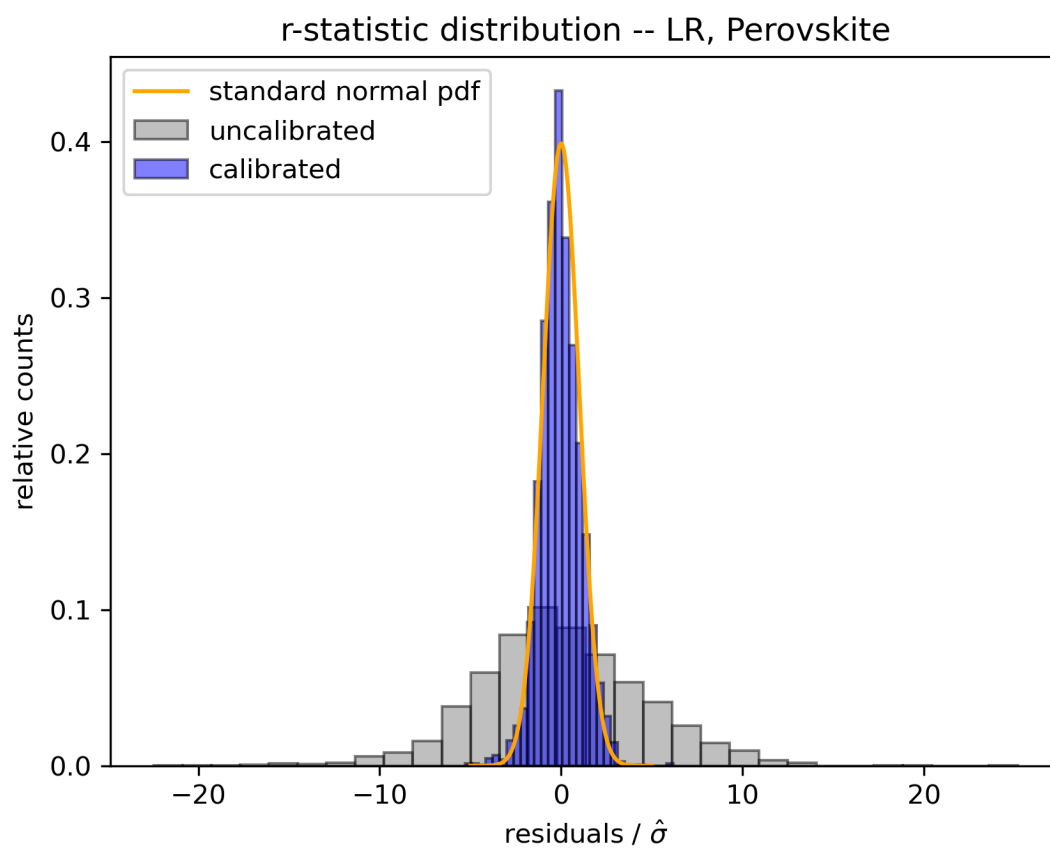
Supplementary Figure 31: Perovskite random forest r-statistic



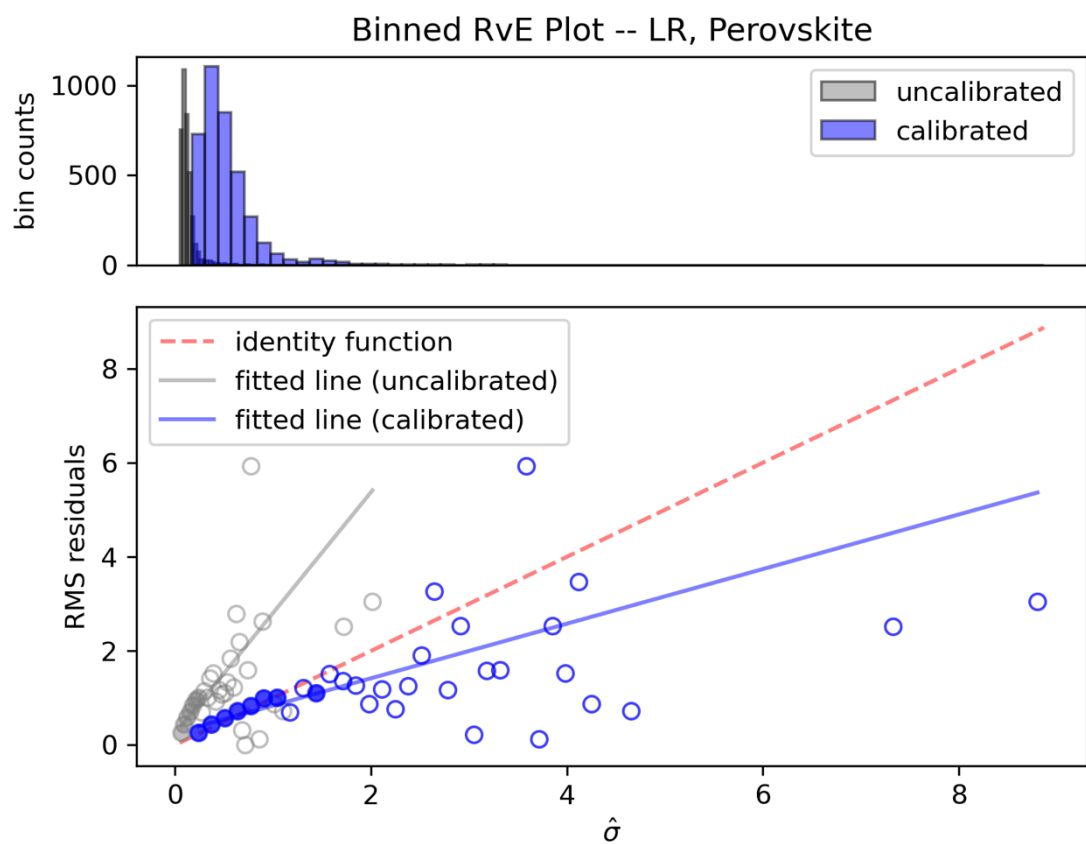
Supplementary Figure 32: Perovskite random forest RvE



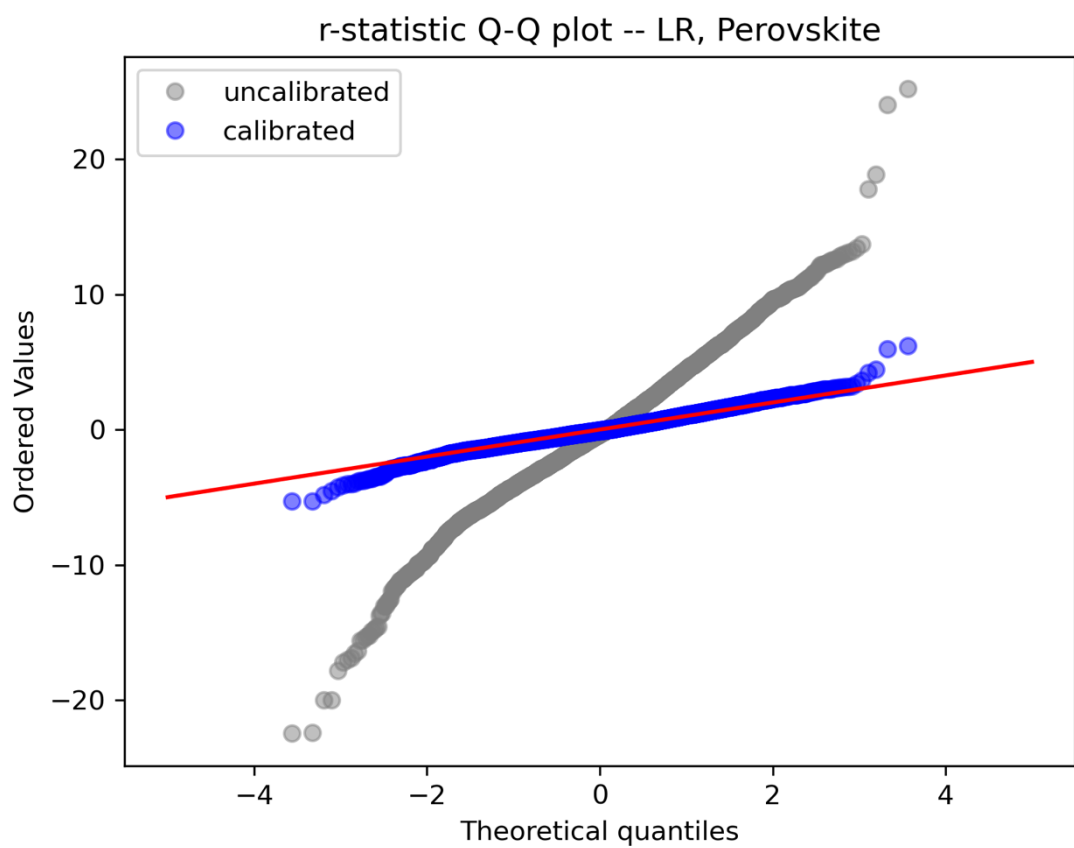
Supplementary Figure 33: Perovskite random forest r-statistic Q-Q plot



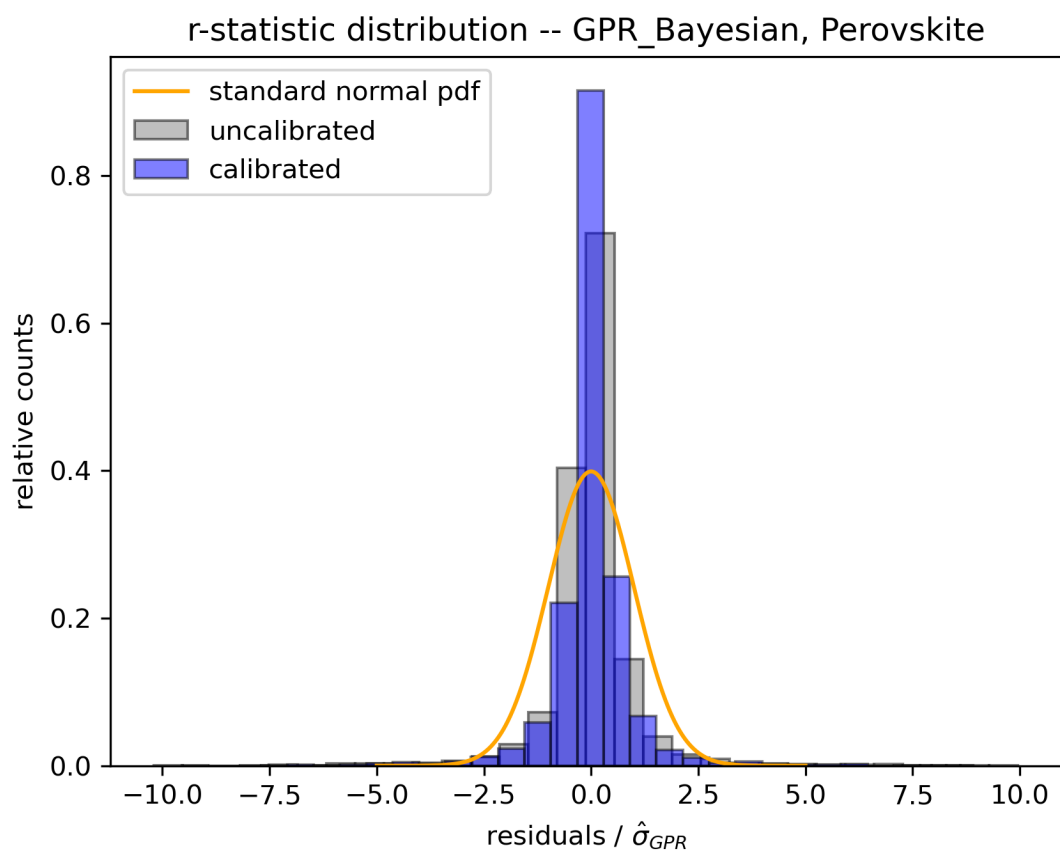
Supplementary Figure 34: Perovskite linear ridge regression r-statistic



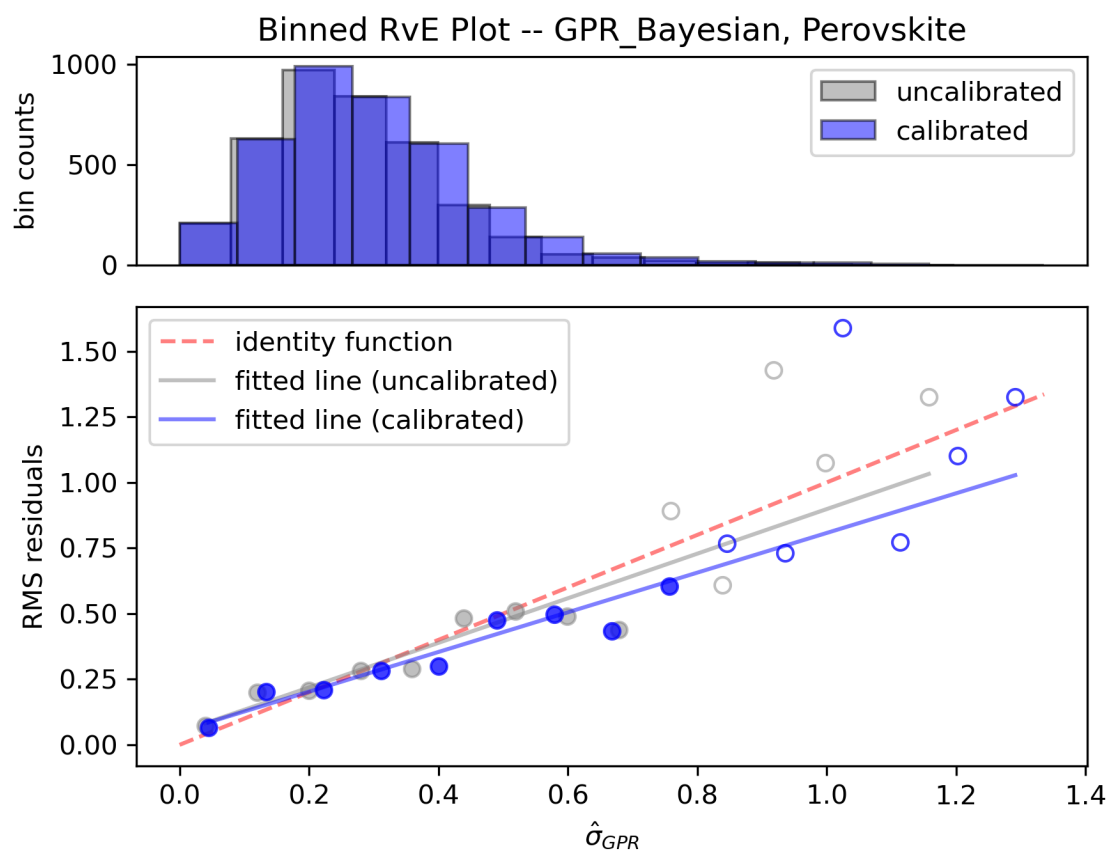
Supplementary Figure 35: Perovskite linear ridge regression RvE



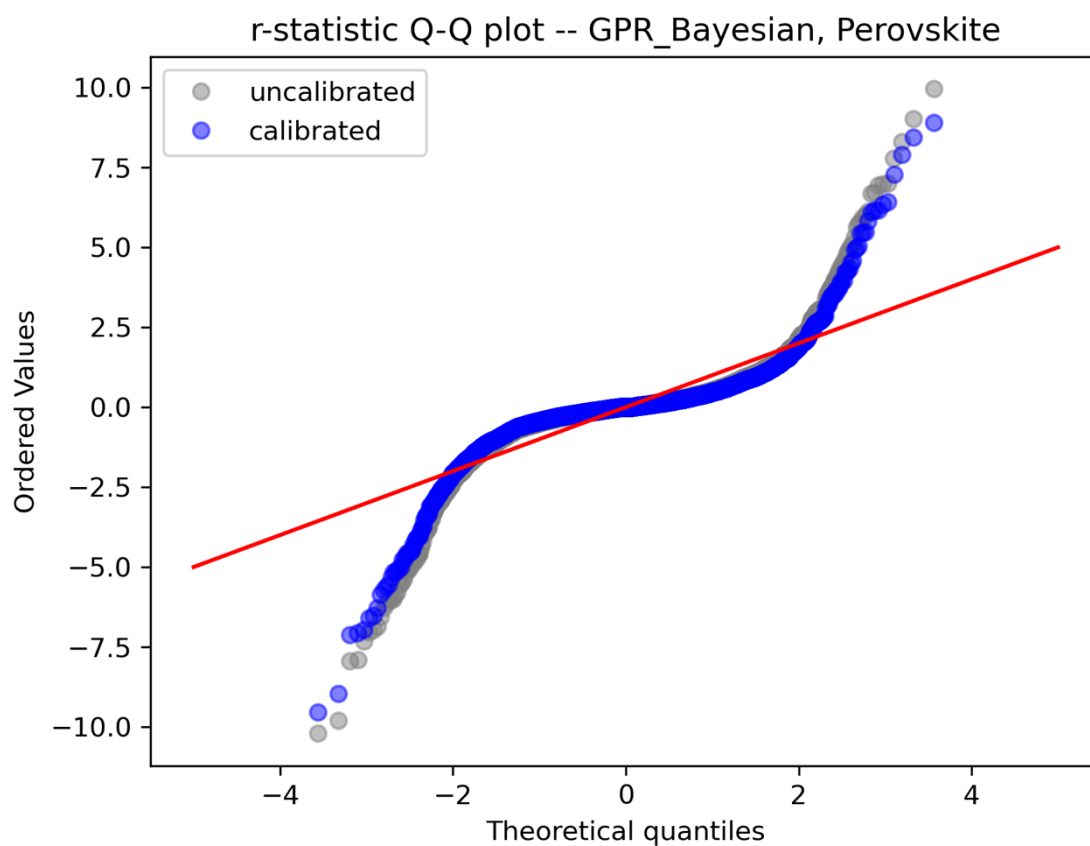
Supplementary Figure 36: Perovskite linear ridge regression r-statistic Q-Q plot



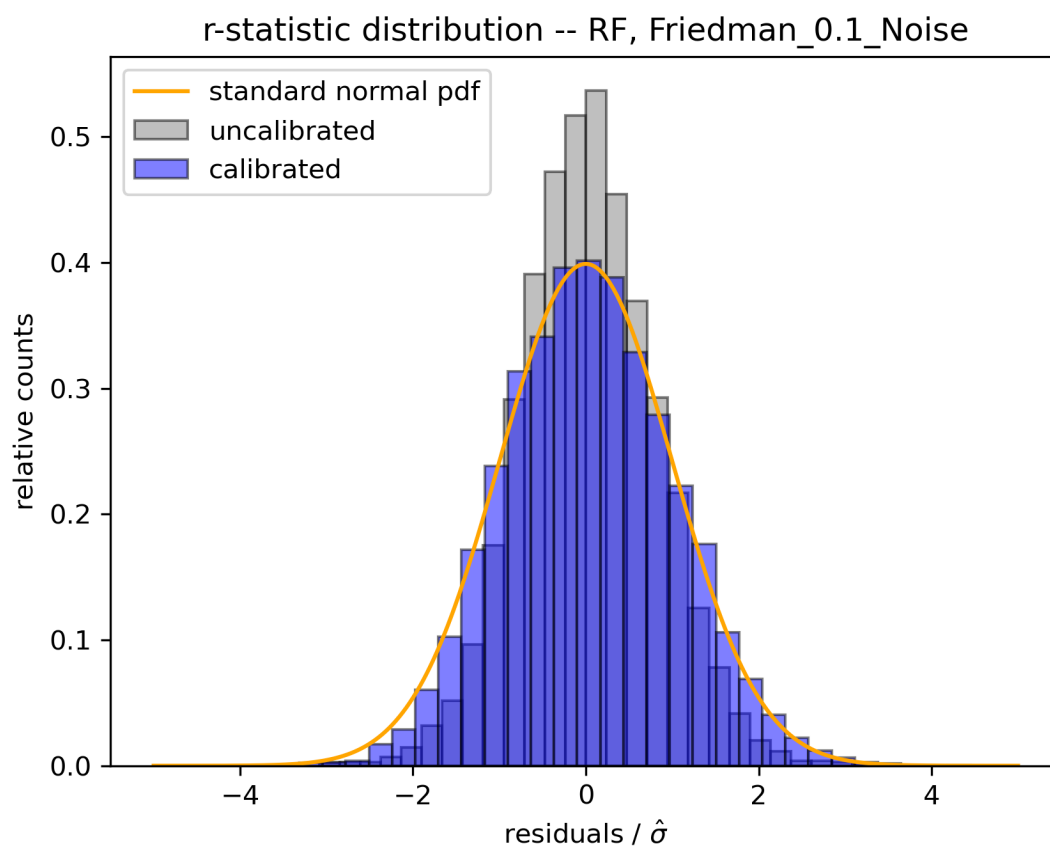
Supplementary Figure 37: Perovskite Bayesian GPR r-statistic



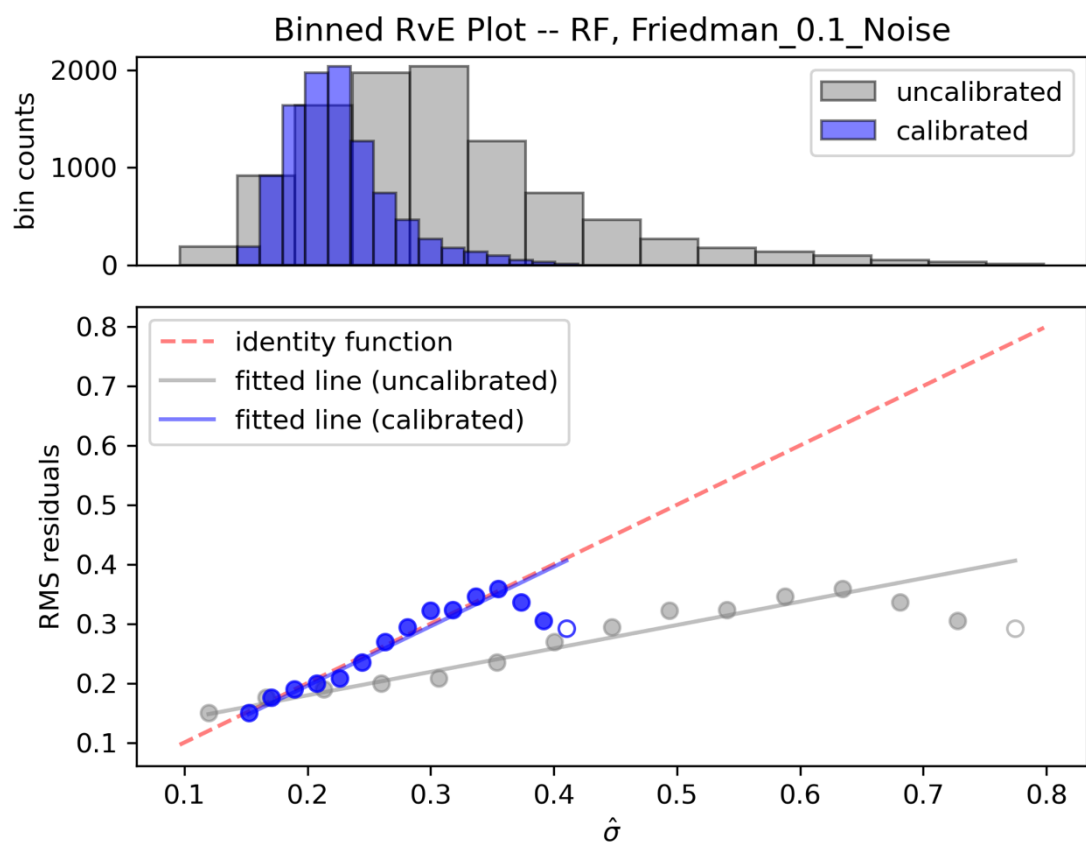
Supplementary Figure 38: Perovskite Bayesian GPR RvE



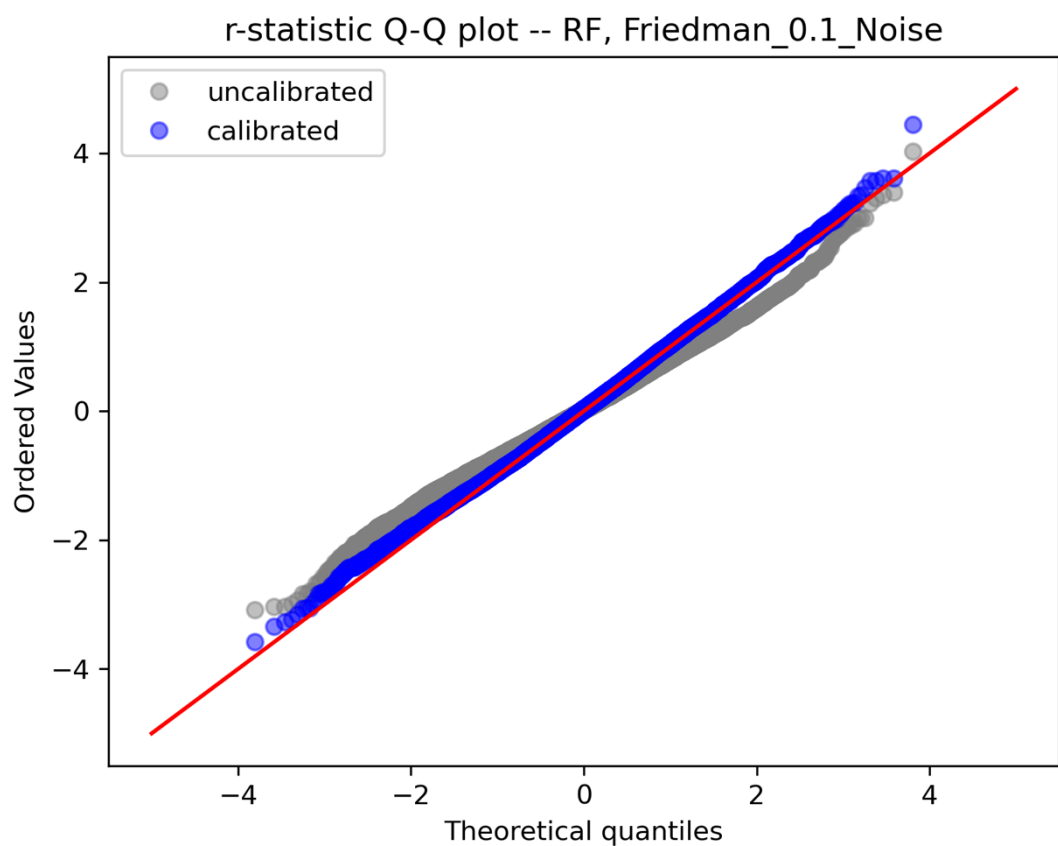
Supplementary Figure 39: Perovskite Bayesian GPR r-statistic Q-Q plot



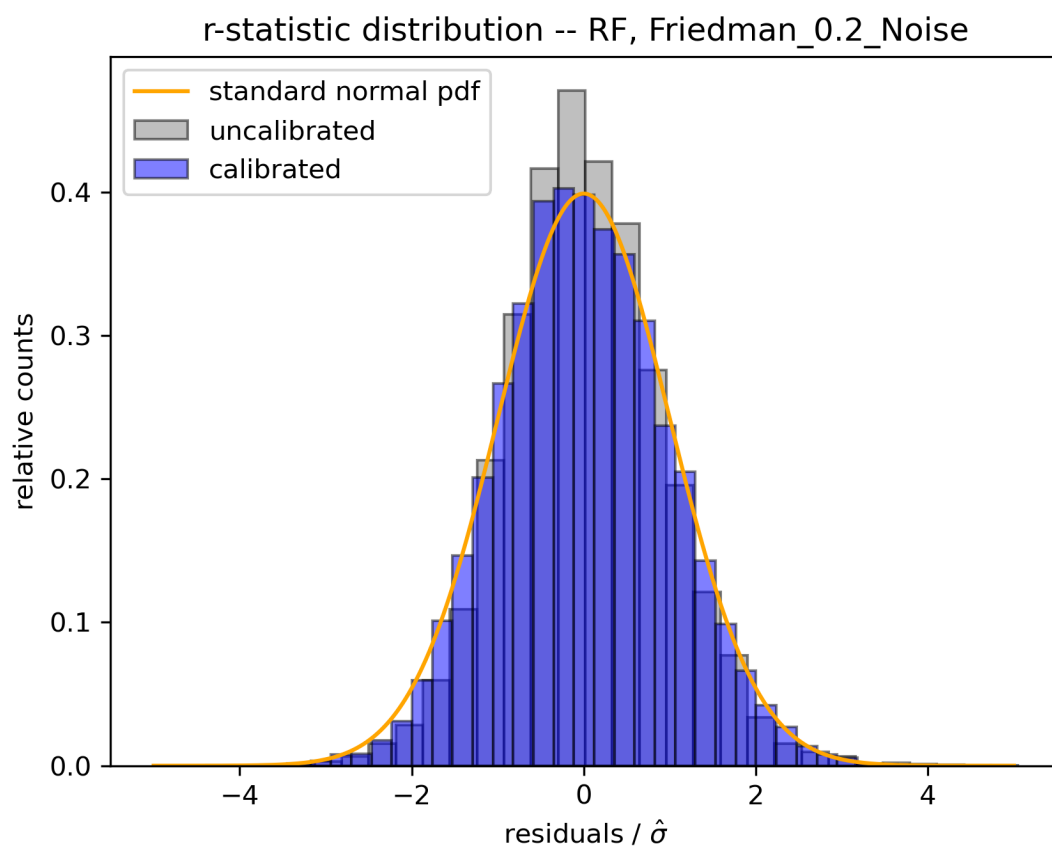
Supplementary Figure 40: Friedman with 0.1 sigma gaussian noise r-statistic



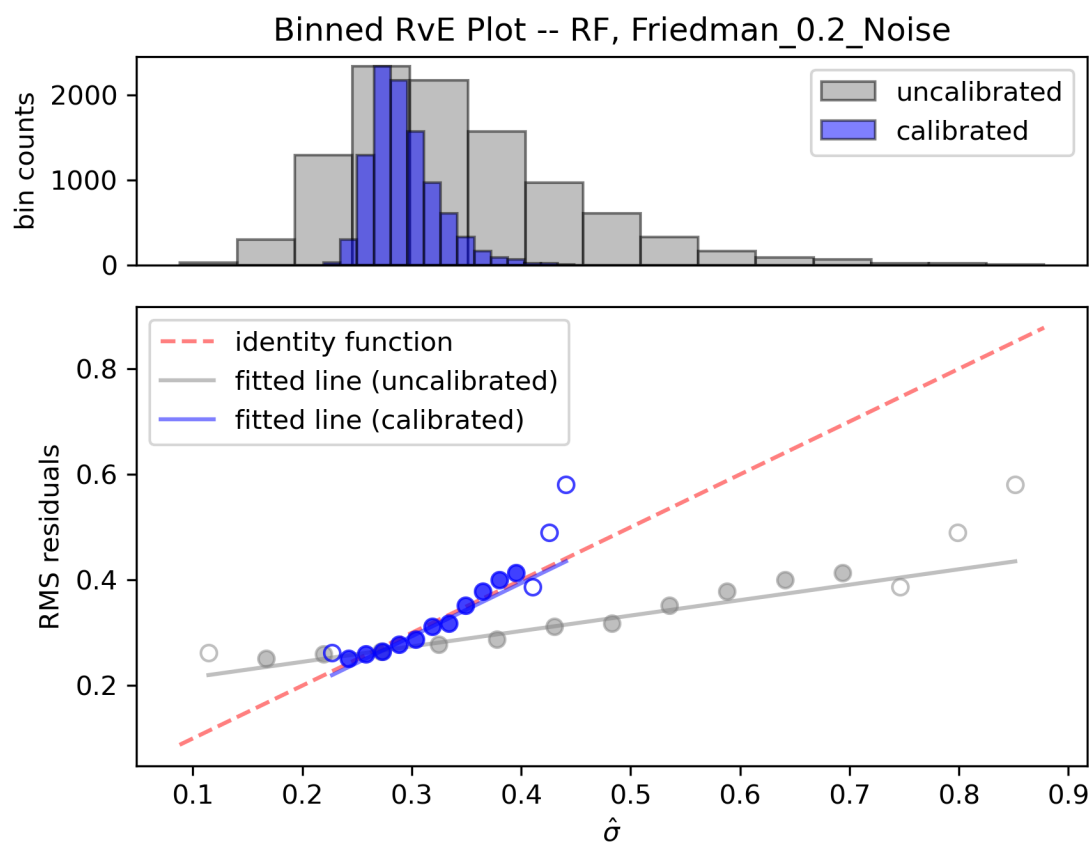
Supplementary Figure 41: Friedman with 0.1 sigma gaussian noise RvE



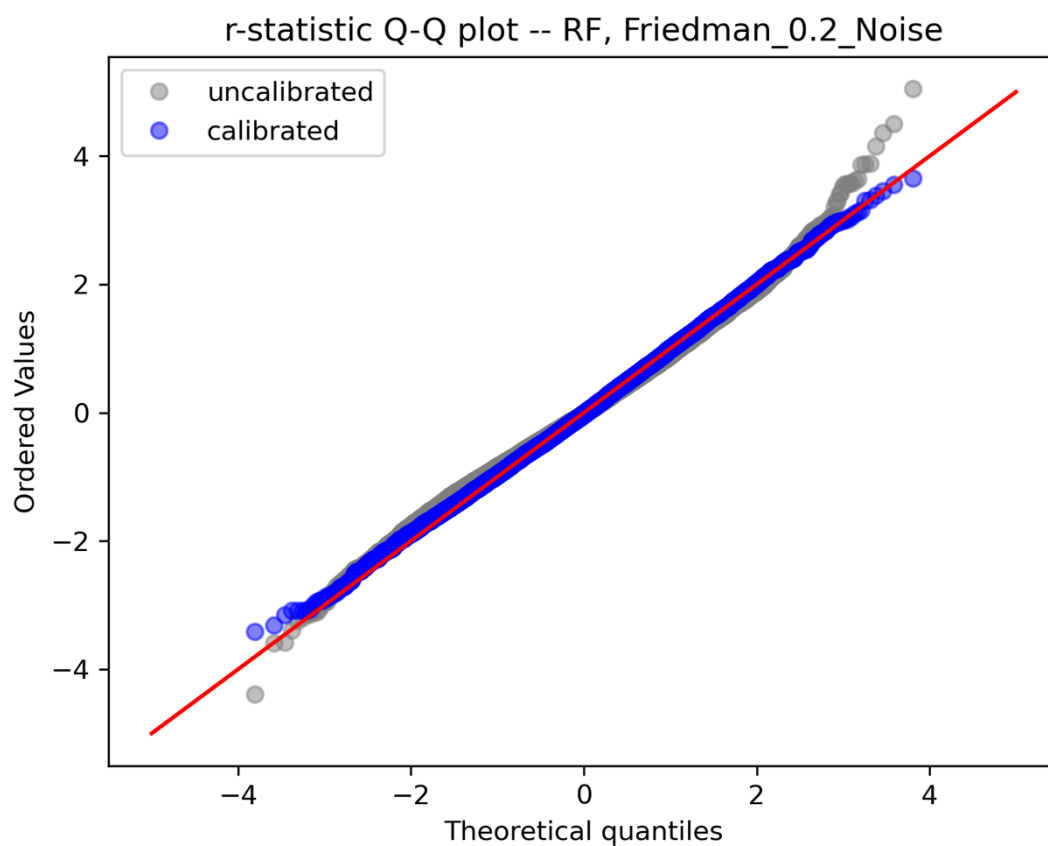
Supplementary Figure 42: Friedman with 0.1 sigma gaussian noise r-statistic Q-Q plot



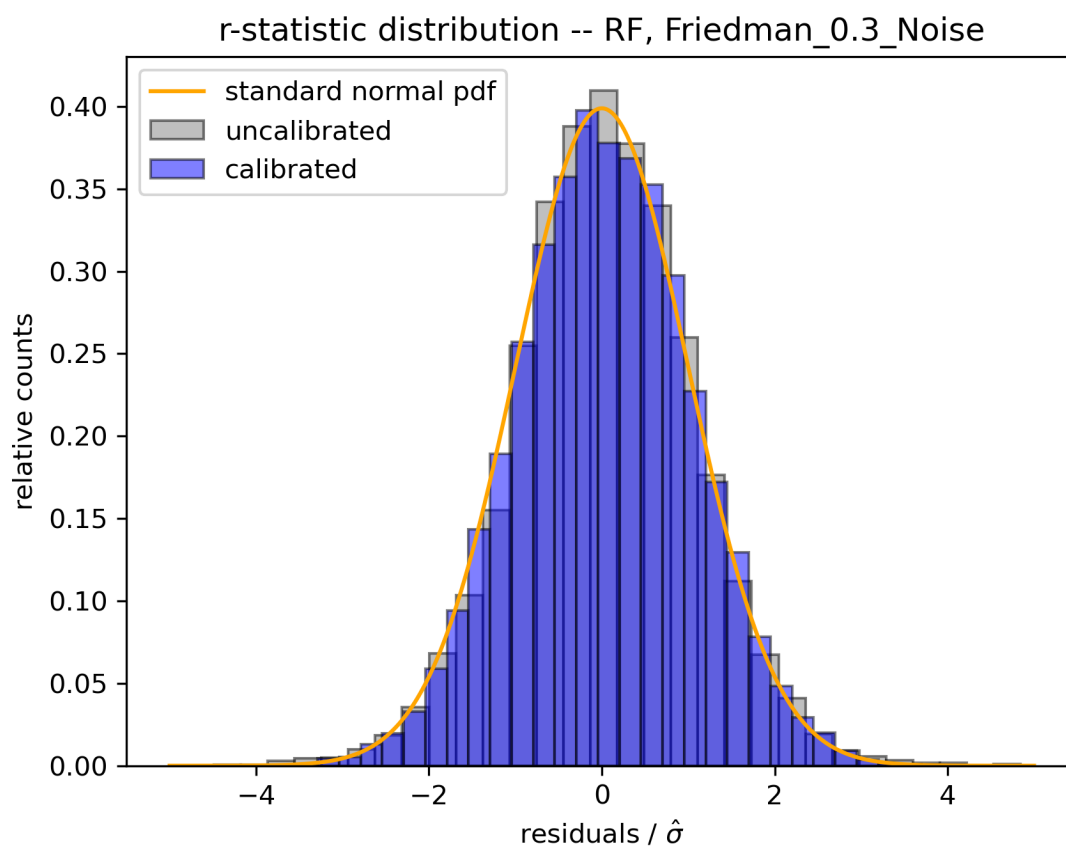
Supplementary Figure 43: Friedman with 0.2 sigma gaussian noise r-statistic



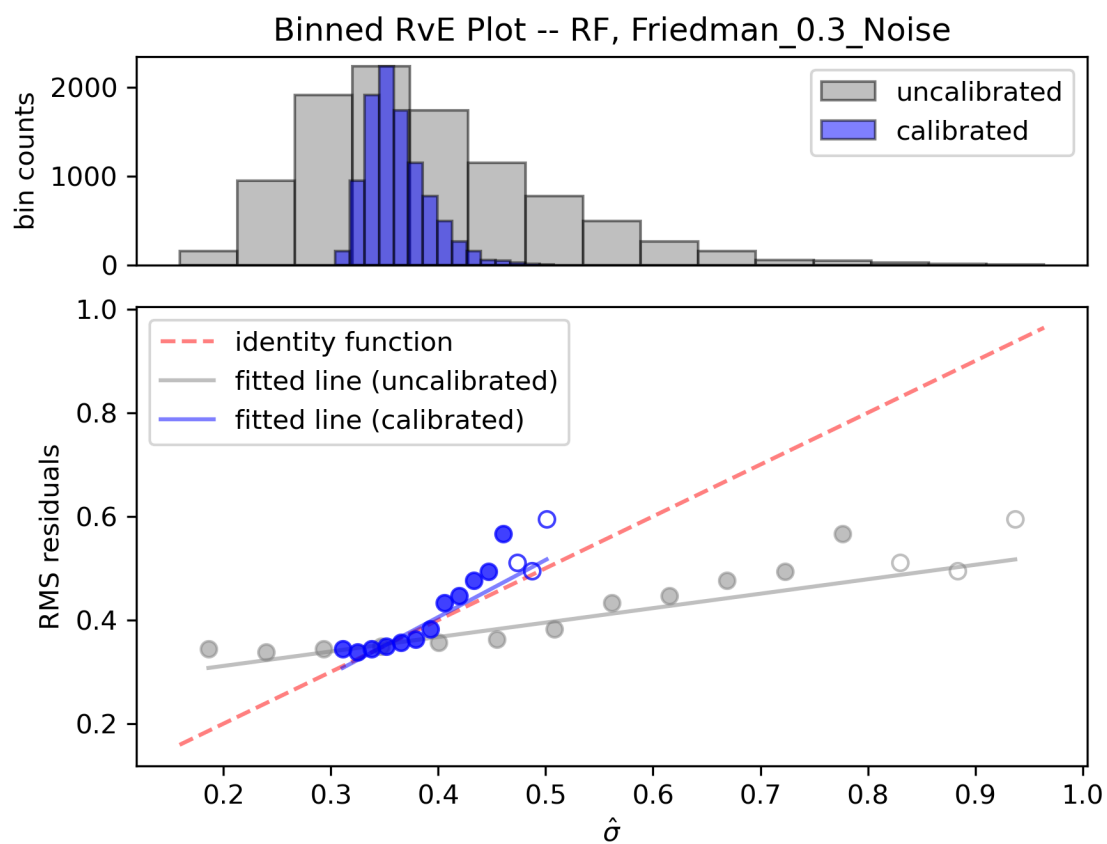
Supplementary Figure 44: Friedman with 0.2 sigma gaussian noise RvE



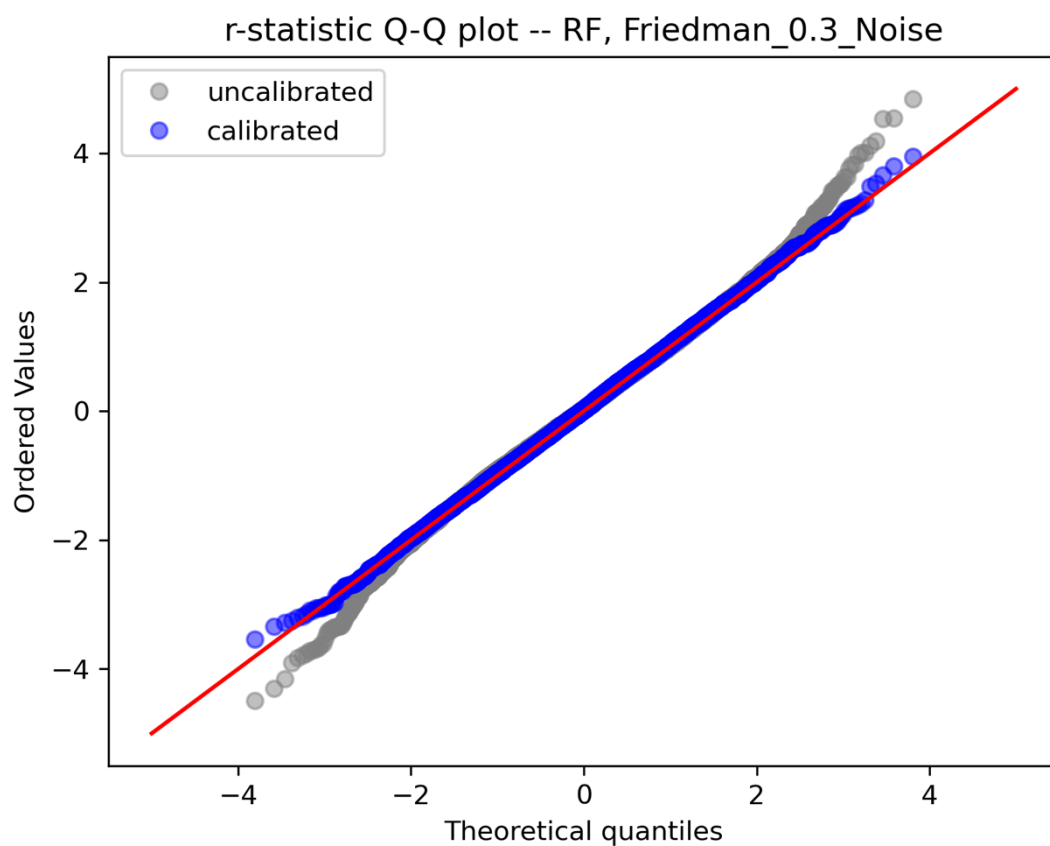
Supplementary Figure 45: Friedman with 0.2 sigma gaussian noise Q-Q plot



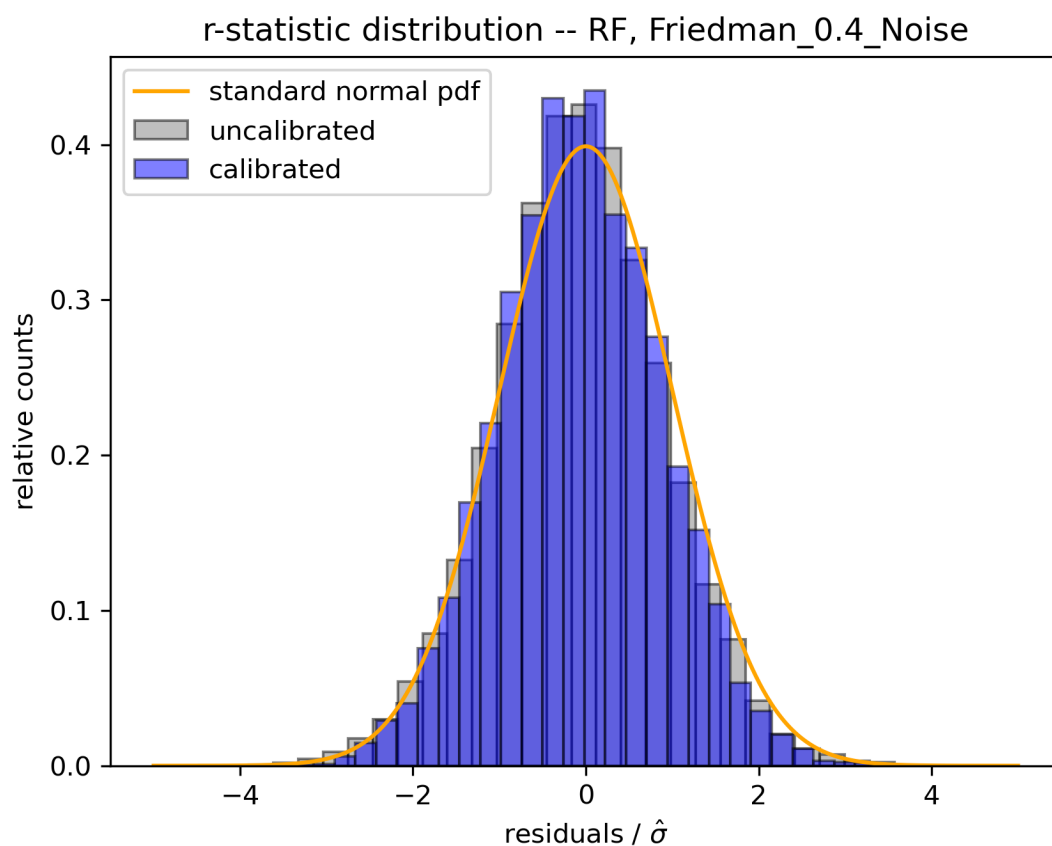
Supplementary Figure 46: Friedman with 0.3 sigma gaussian noise r-statistic



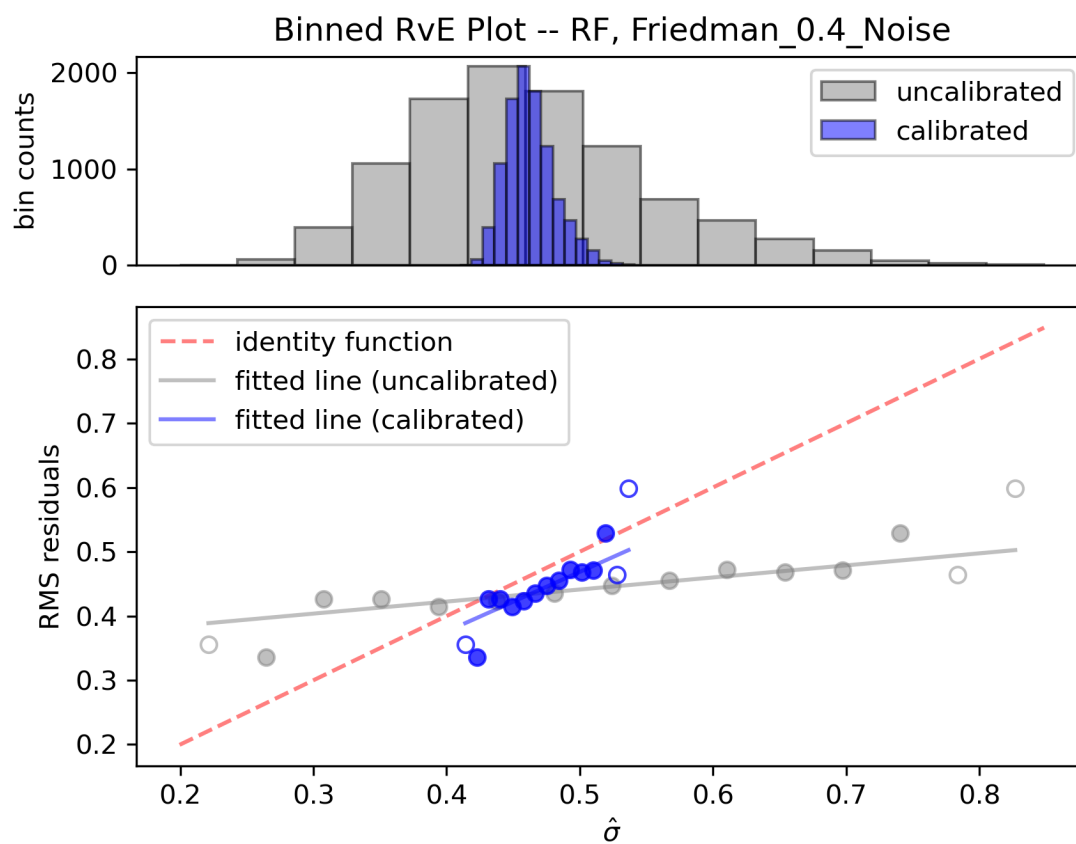
Supplementary Figure 47: Friedman with 0.3 sigma gaussian noise RvE



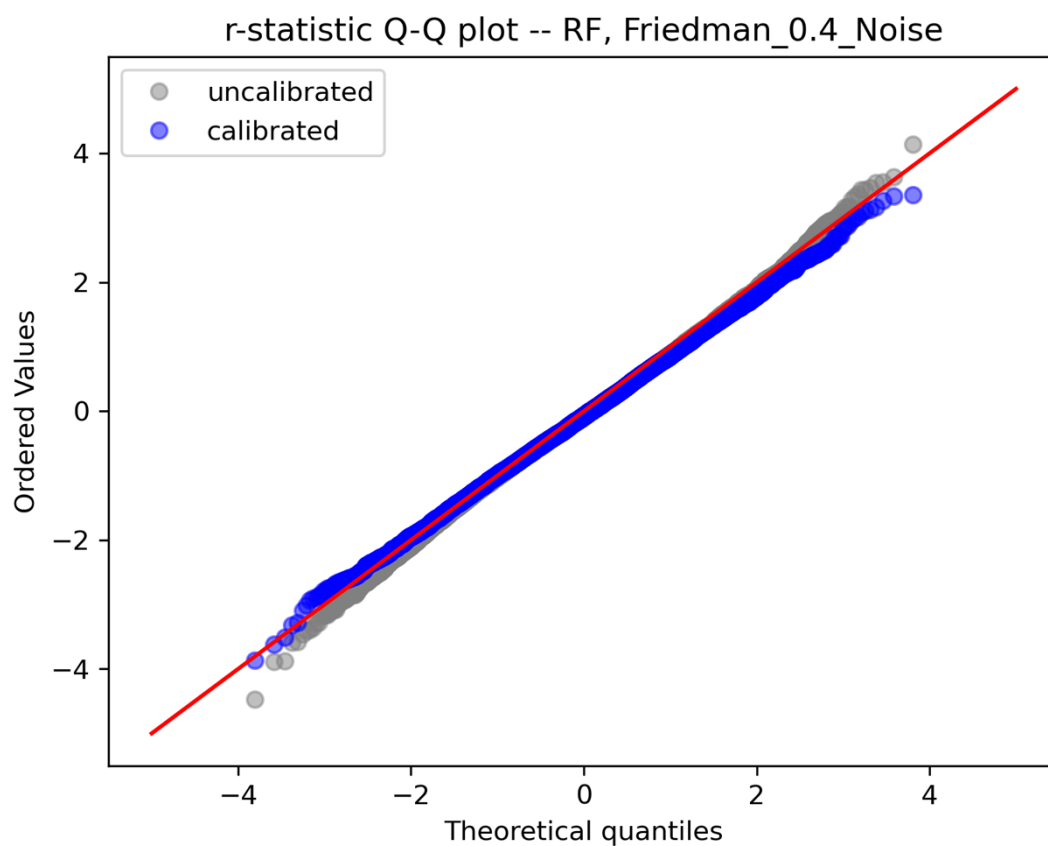
Supplementary Figure 48: Friedman with 0.3 sigma gaussian noise r-statistic Q-Q plot



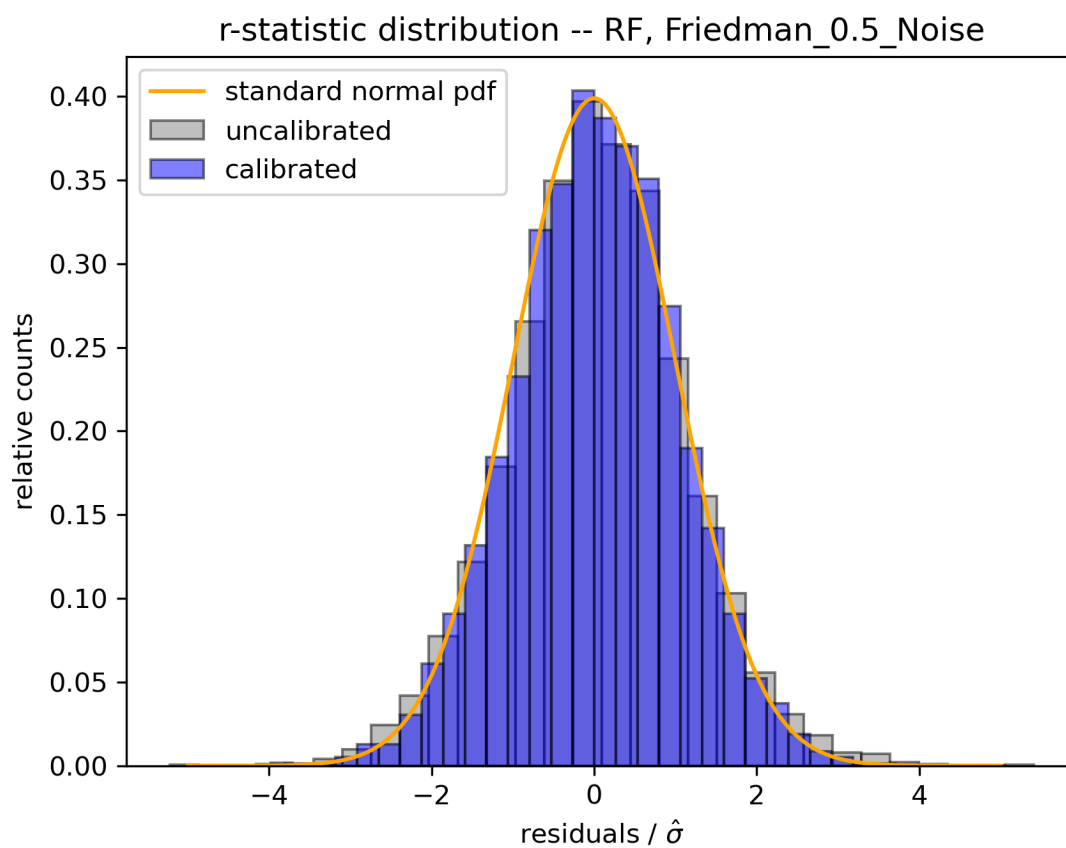
Supplementary Figure 49: Friedman with 0.4 sigma gaussian noise r-statistic



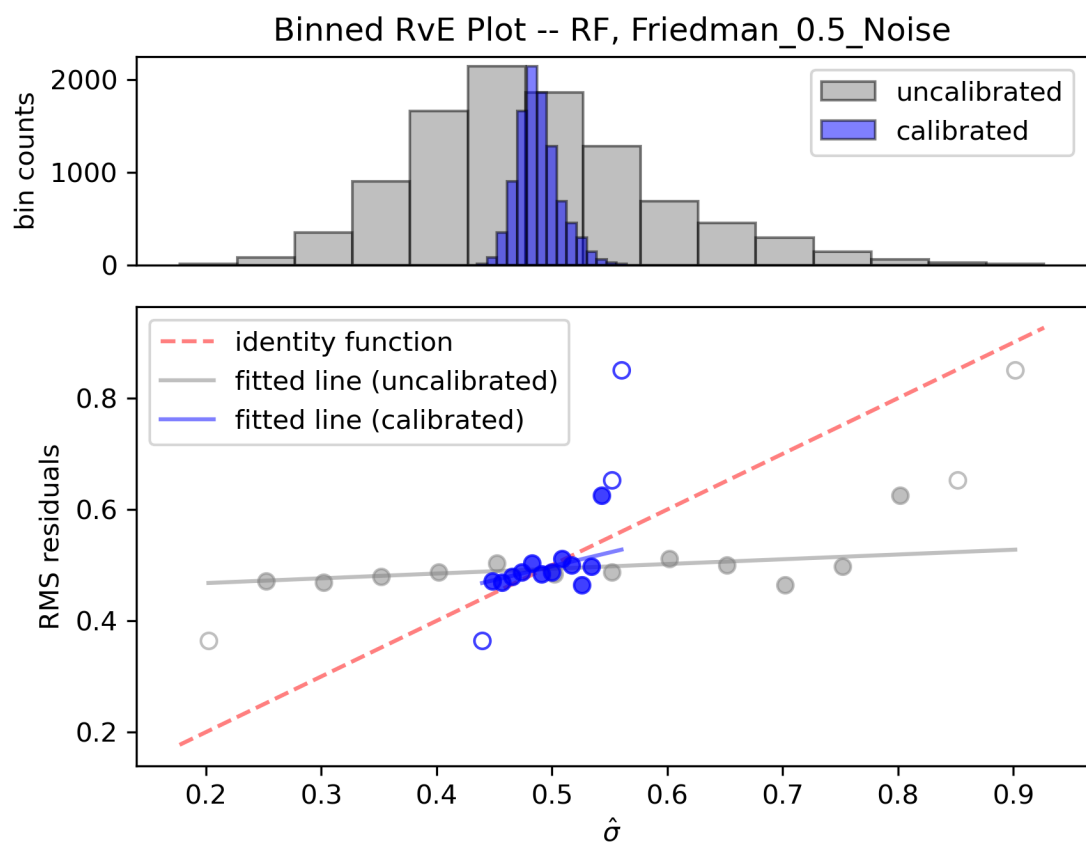
Supplementary Figure 50: Friedman with 0.4 sigma gaussian noise RvE



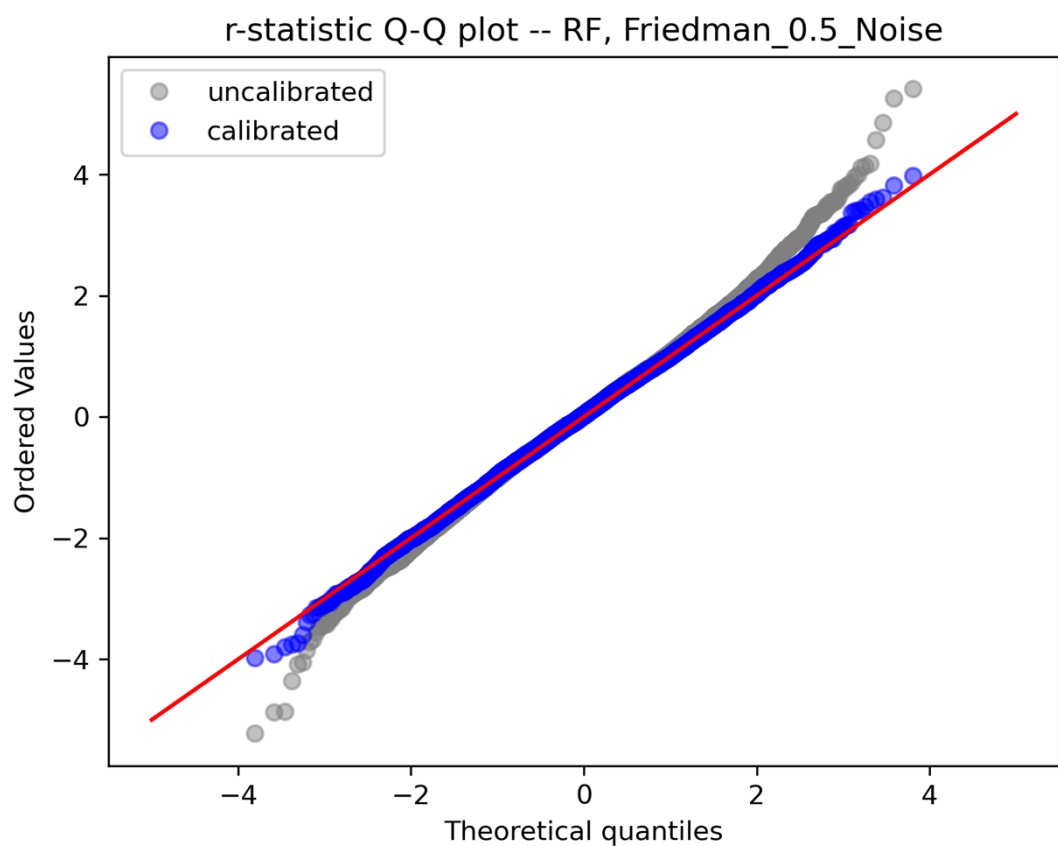
Supplementary Figure 51: Friedman with 0.4 sigma gaussian noise r-statistic Q-Q plot



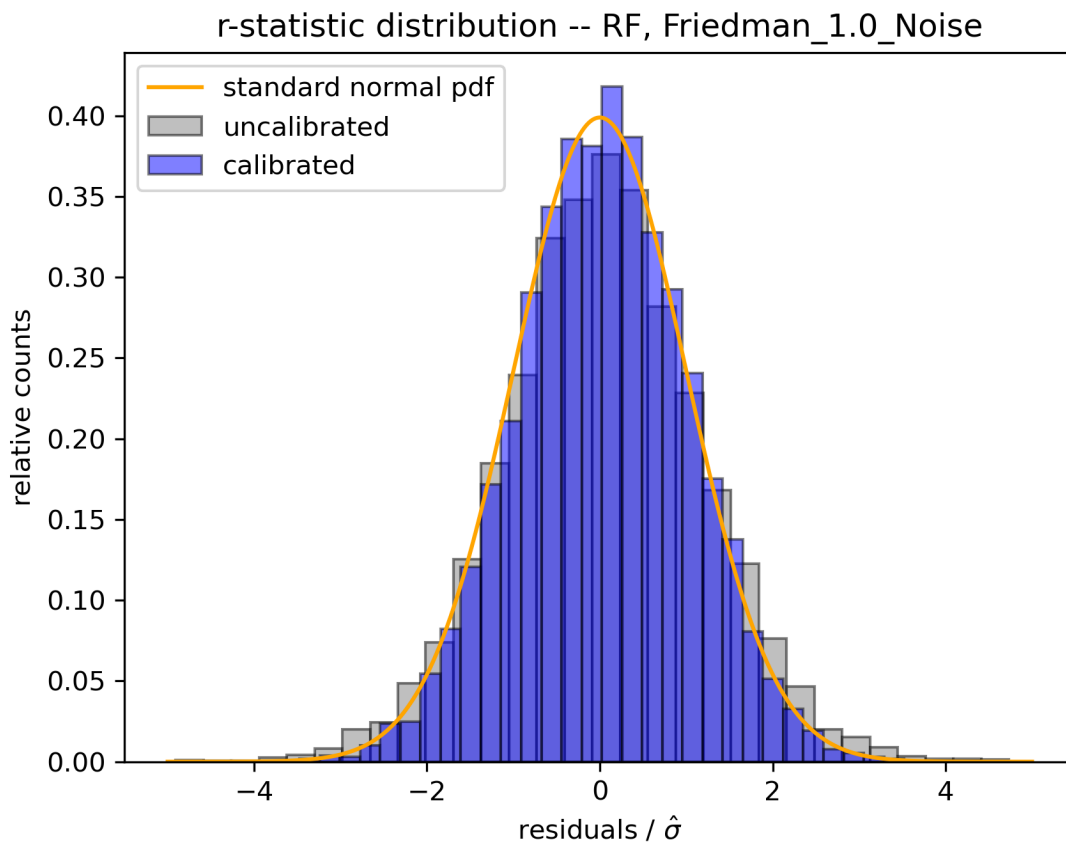
Supplementary Figure 52: Friedman with 0.5 sigma gaussian noise r-statistic



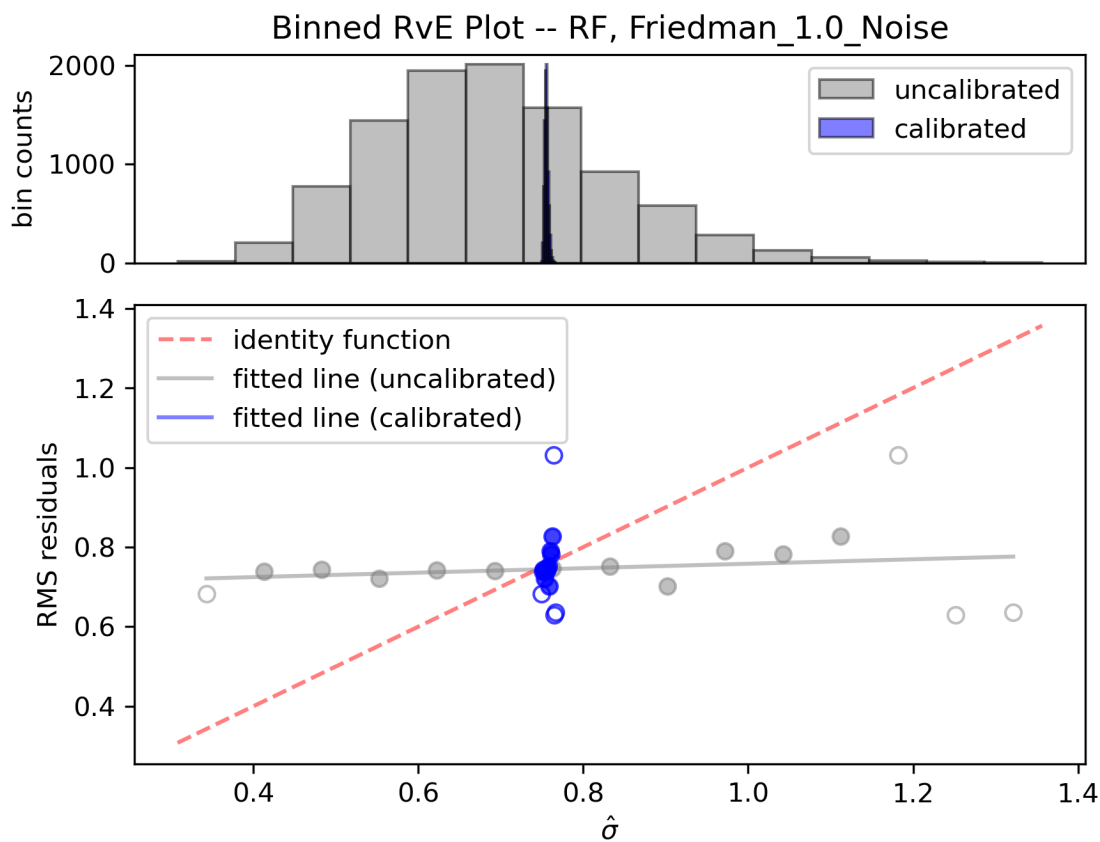
Supplementary Figure 53: Friedman with 0.5 sigma gaussian noise RvE



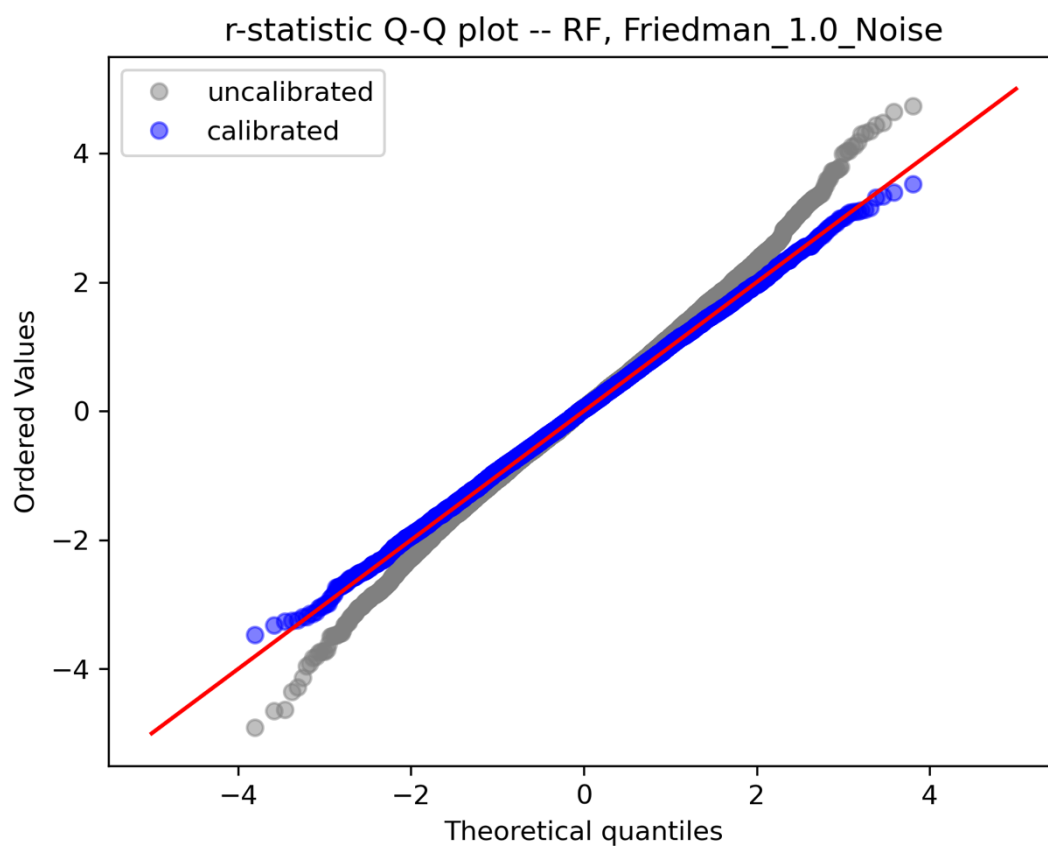
Supplementary Figure 54: Friedman with 0.5 sigma gaussian noise r-statistic Q-Q plot



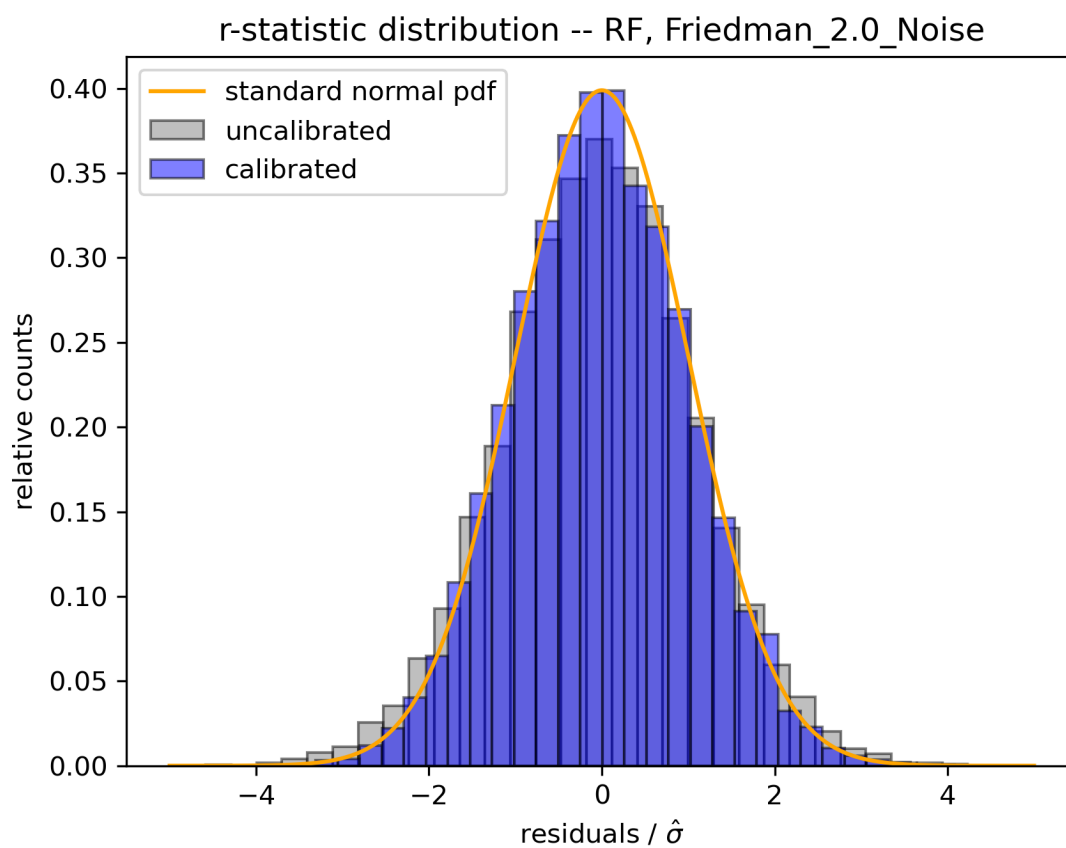
Supplementary Figure 55: Friedman with 1.0 sigma gaussian noise r-statistic



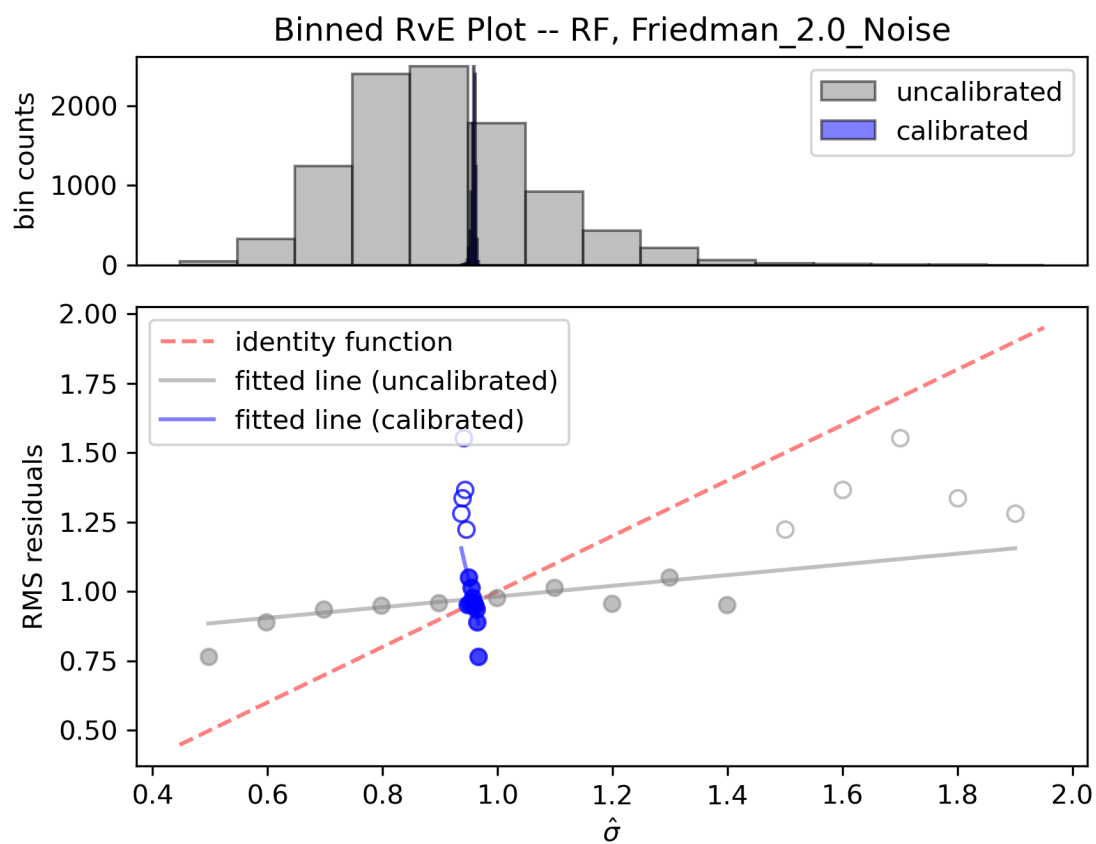
Supplementary Figure 56: Friedman with 1.0 sigma gaussian noise RvE



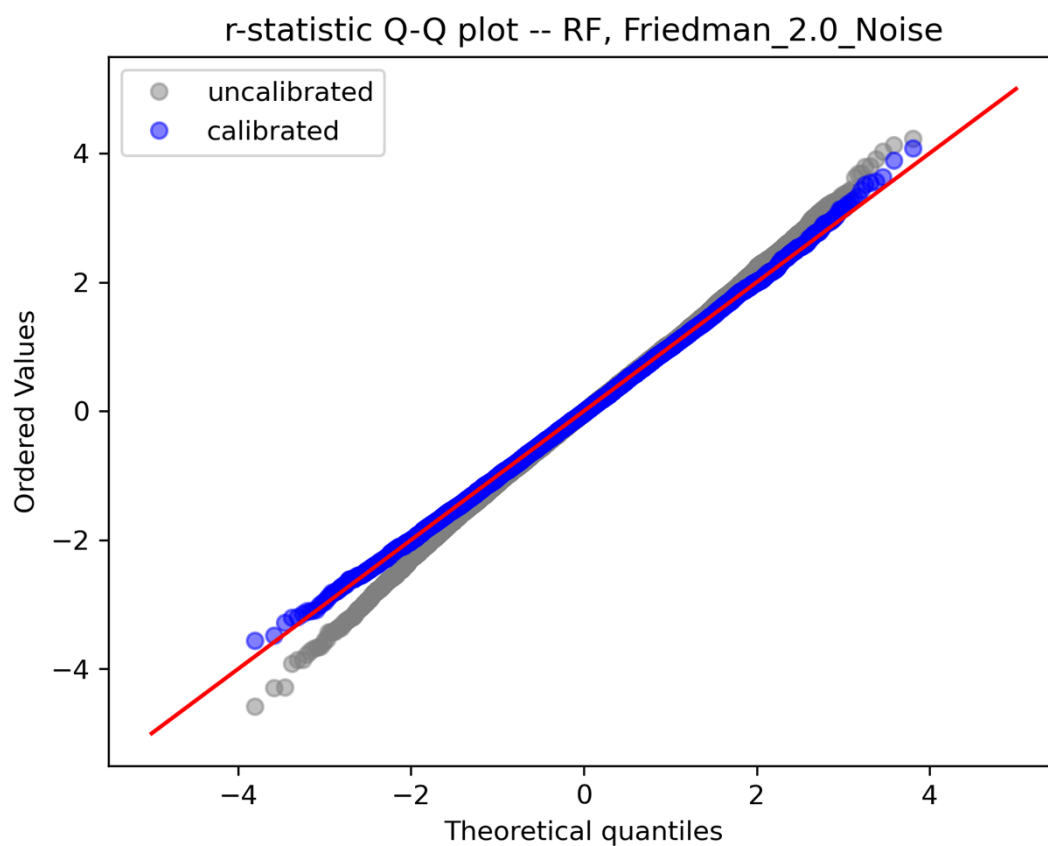
Supplementary Figure 57: Friedman with 1.0 sigma gaussian noise r-statistic Q-Q plot



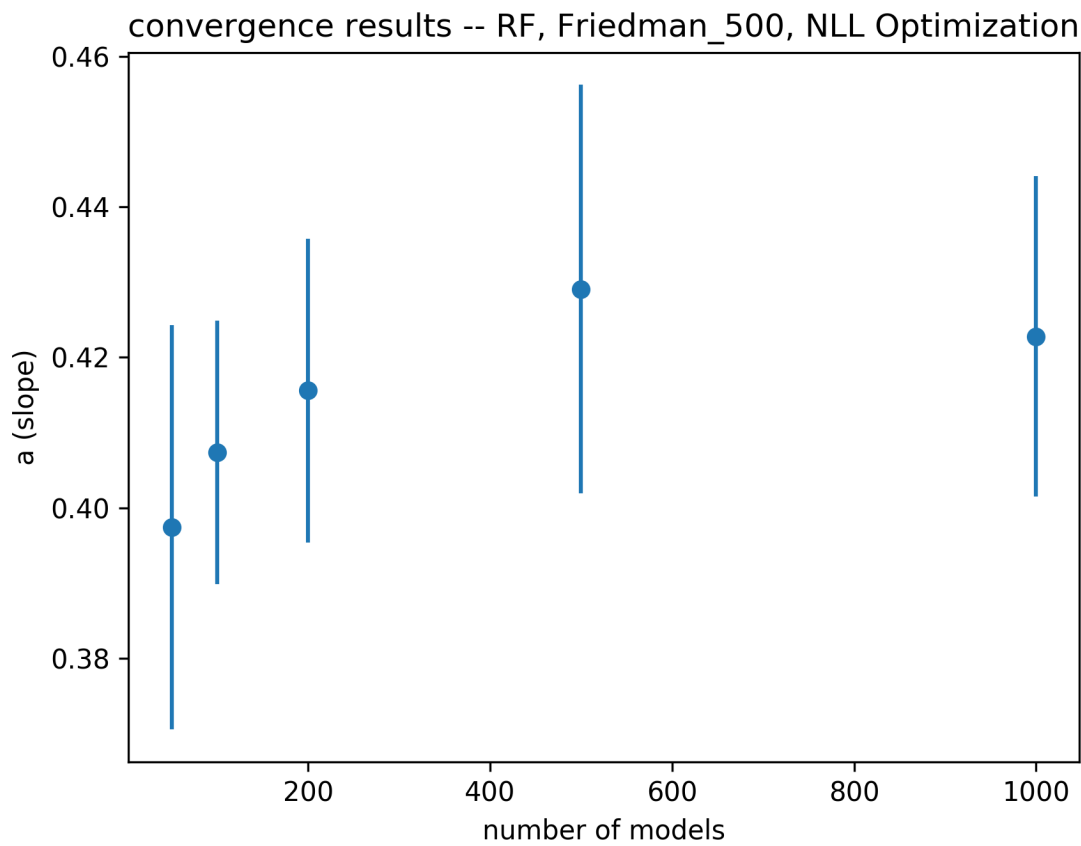
Supplementary Figure 58: Friedman with 2.0 sigma gaussian noise r-statistic



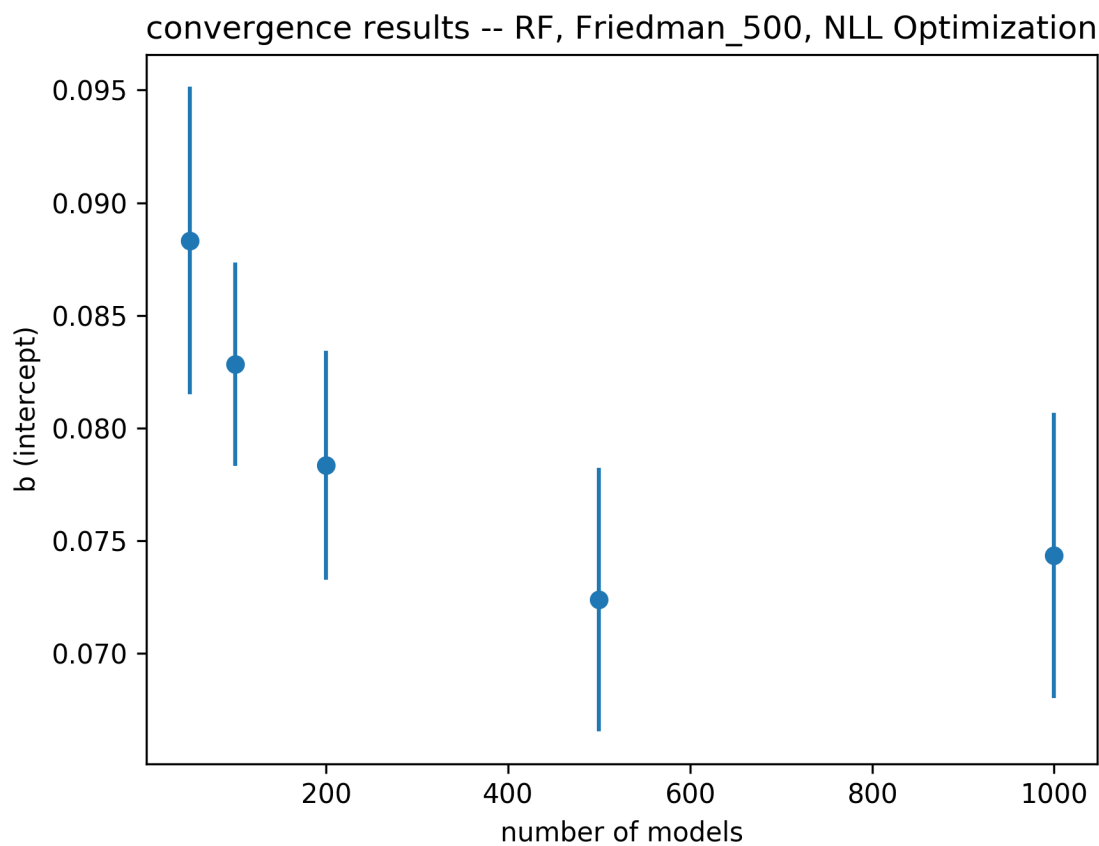
Supplementary Figure 59: Friedman with 2.0 sigma gaussian noise RvE



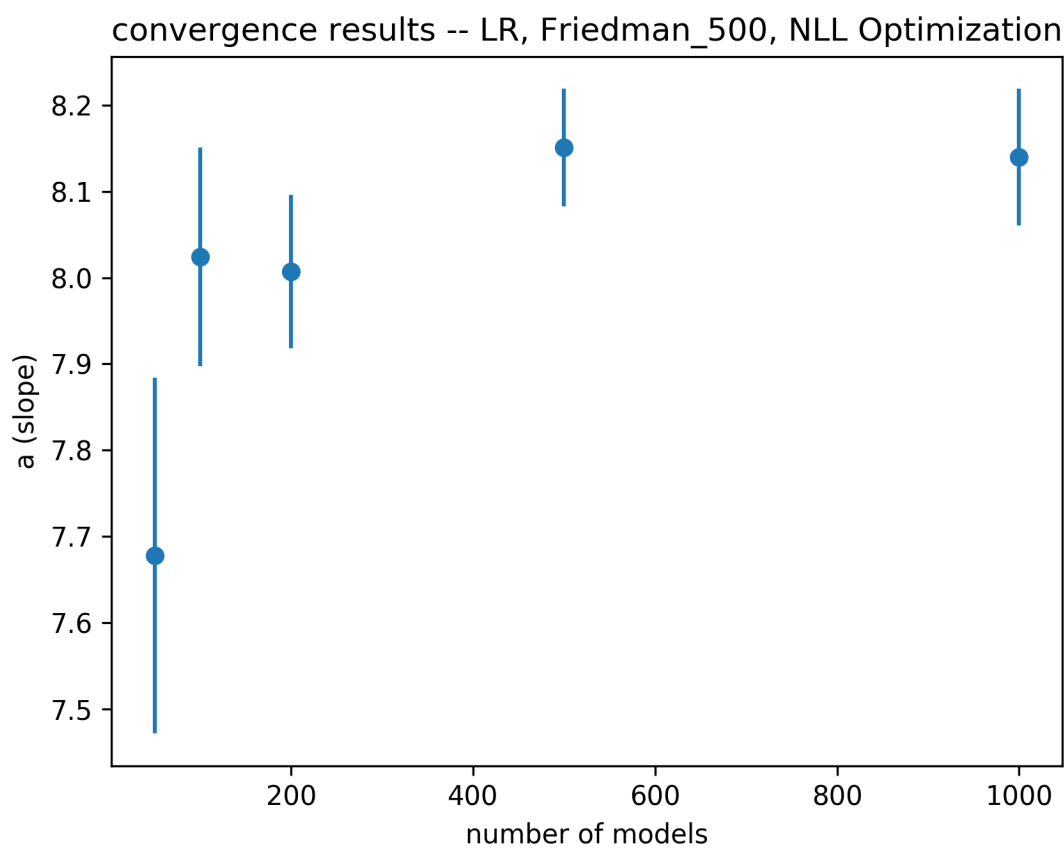
Supplementary Figure 60: Friedman with 2.0 sigma gaussian noise r-statistic Q-Q plot



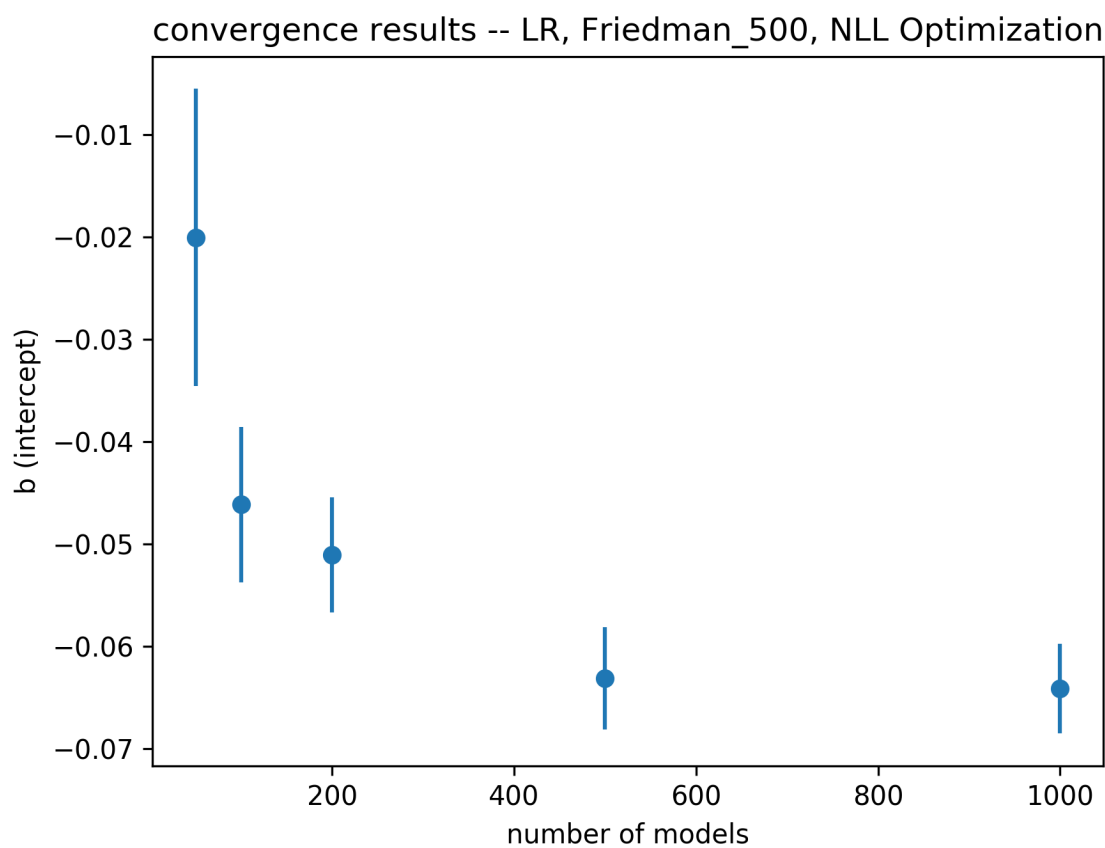
Supplementary Figure 61: Convergence of  $a$  for Friedman with random forest



Supplementary Figure 62: Convergence of  $b$  for Friedman with random forest

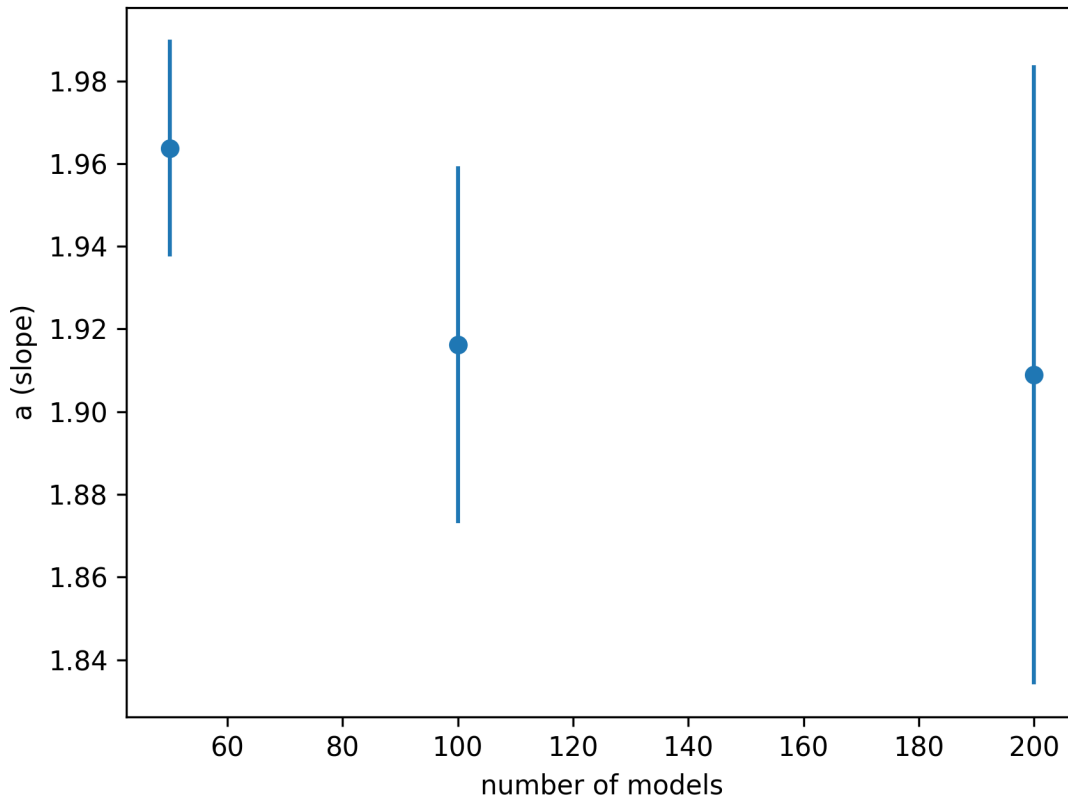


Supplementary Figure 63: Convergence of  $a$  for Friedman with linear ridge regression



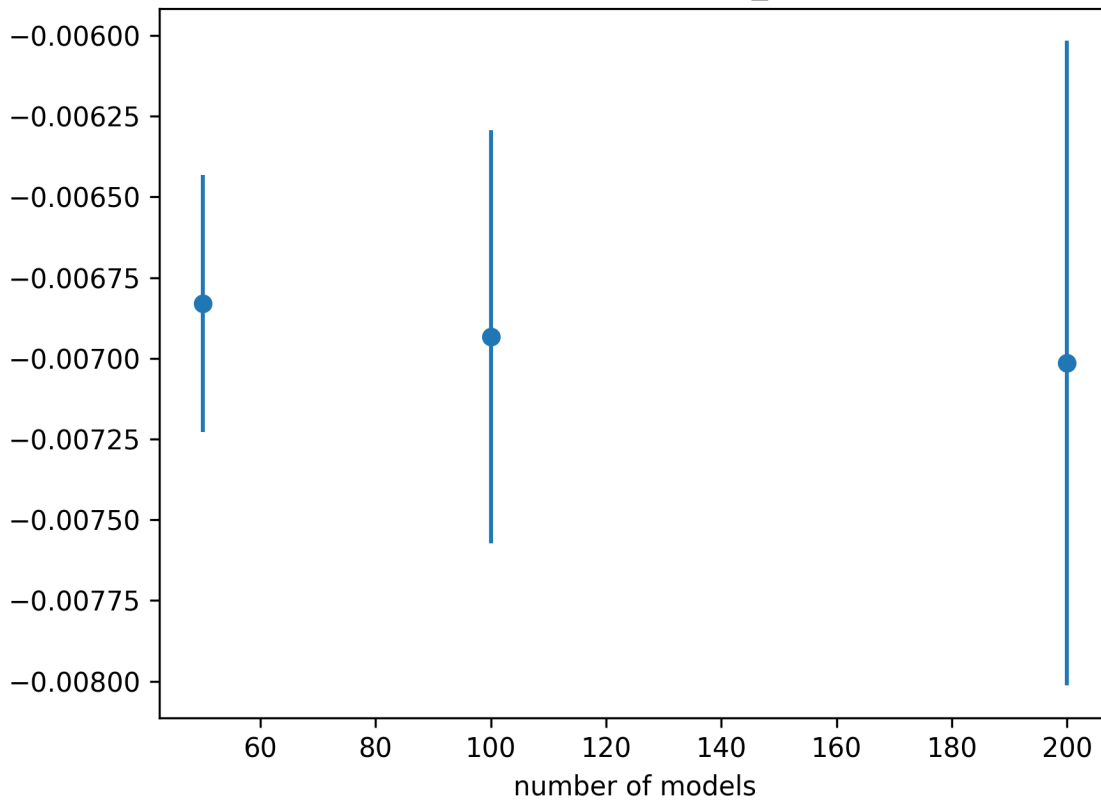
Supplementary Figure 64: Convergence of  $b$  for Friedman with linear ridge regression

convergence results -- GPR, Friedman\_500, NLL Optimization

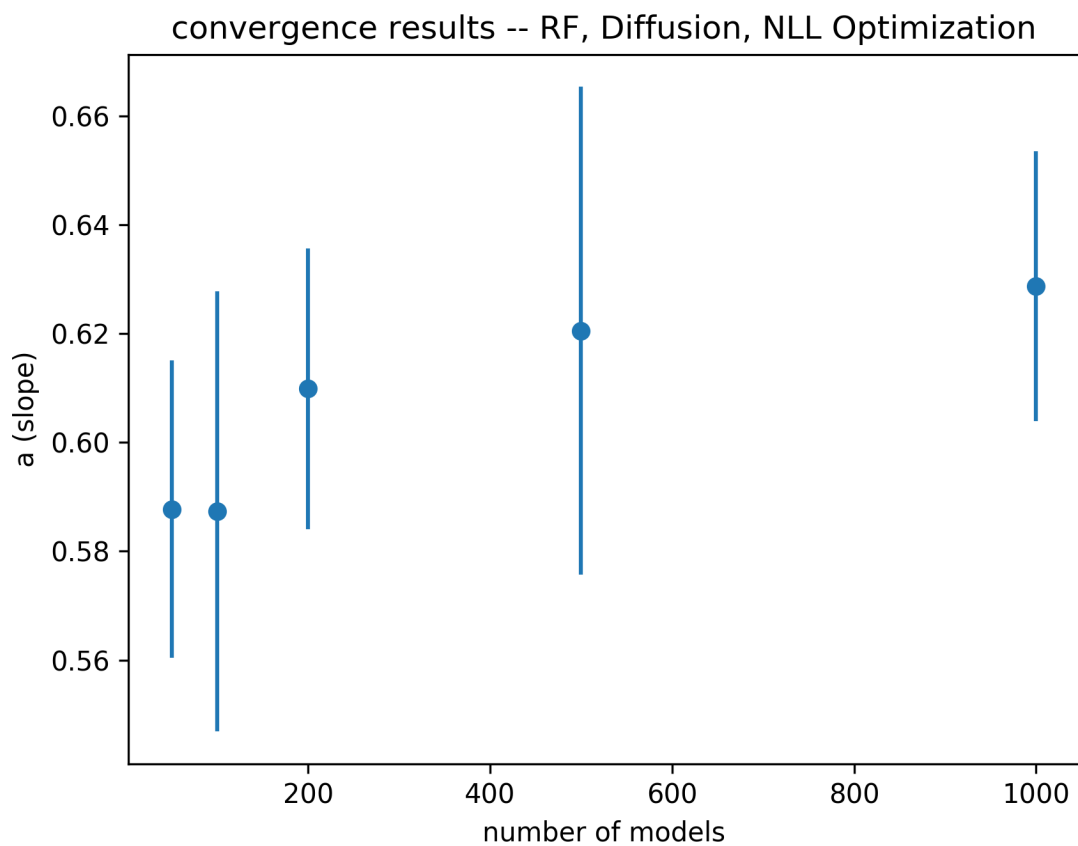


Supplementary Figure 65: Convergence of  $a$  for Friedman with GPR

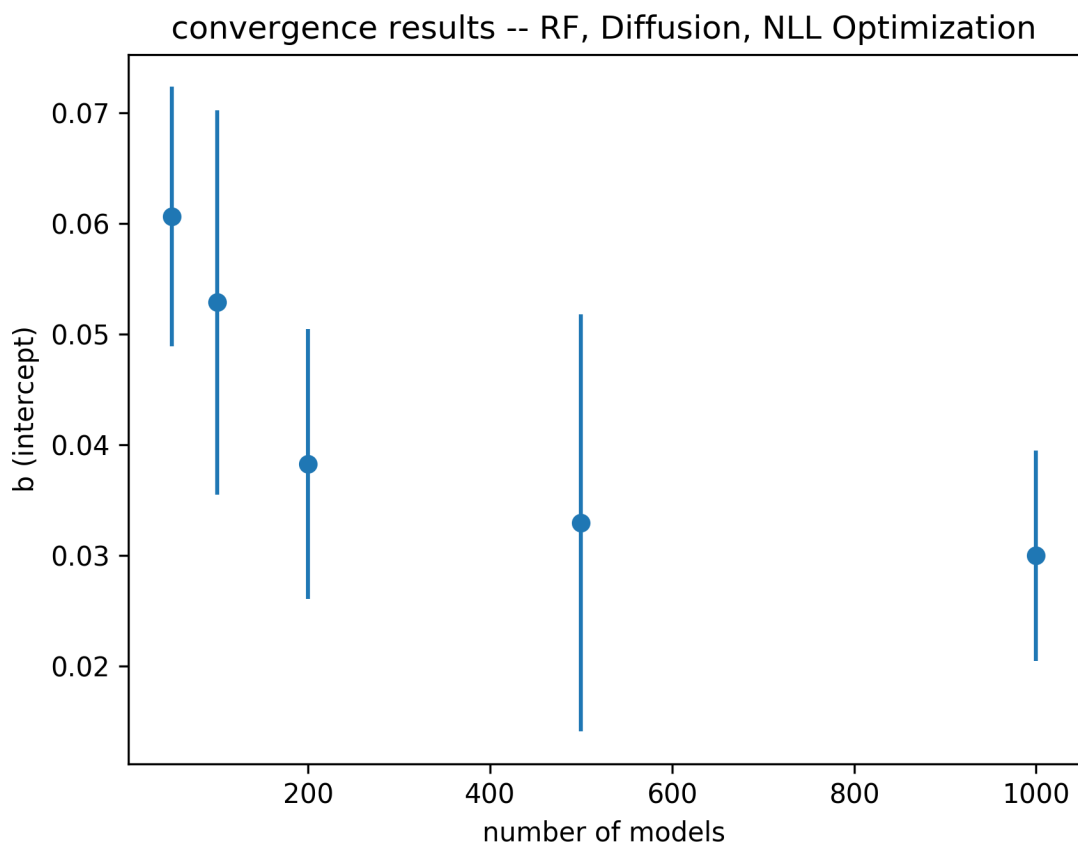
convergence results -- GPR, Friedman\_500, NLL Optimization



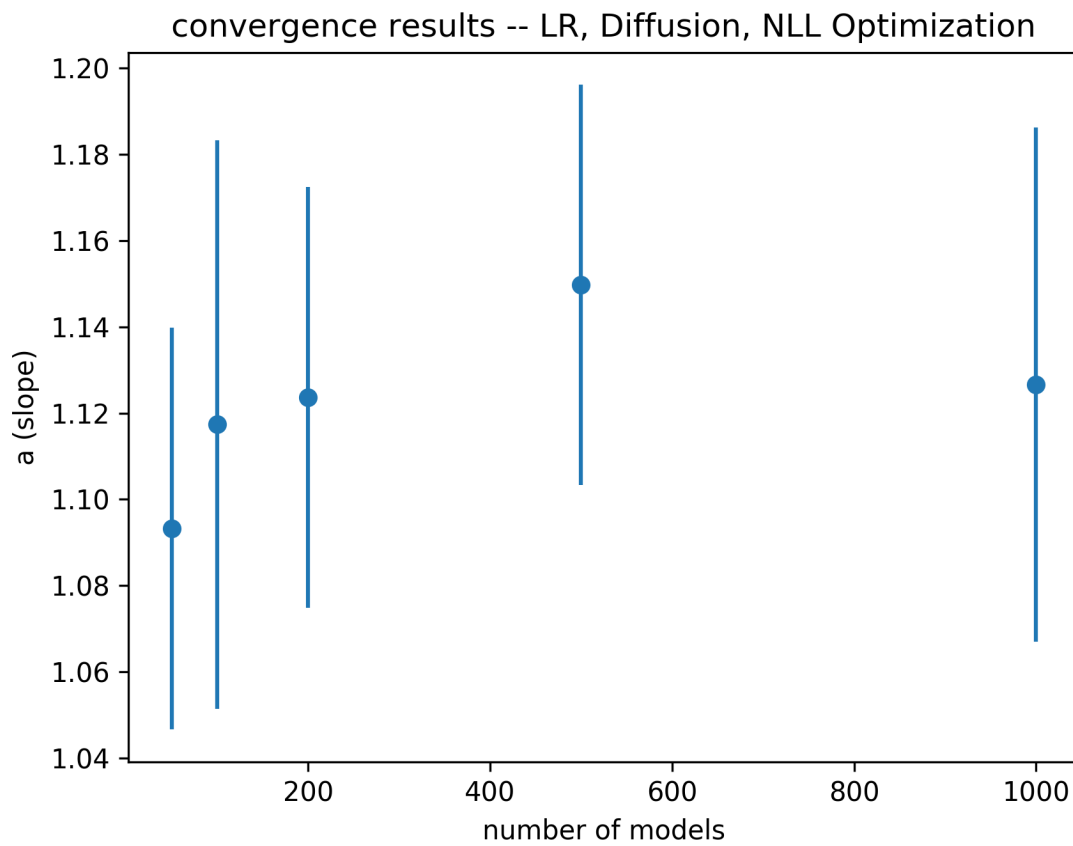
Supplementary Figure 66: Convergence of  $b$  for Friedman with GPR



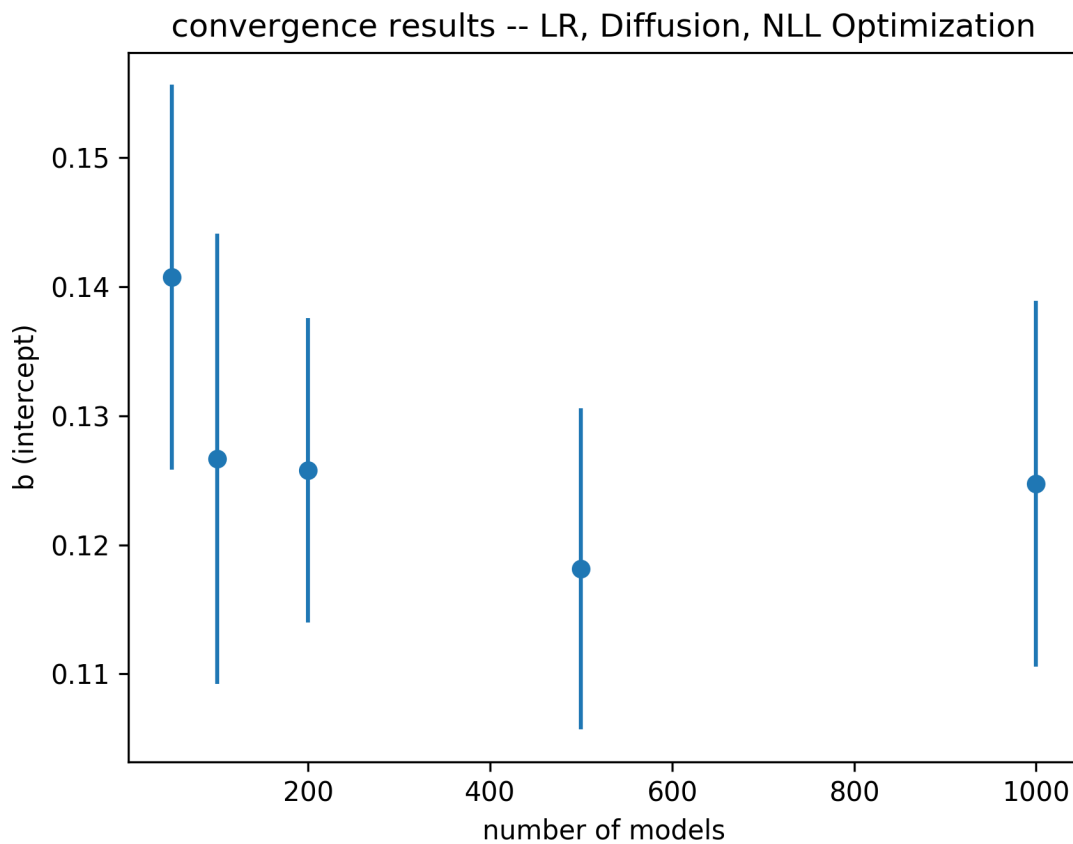
Supplementary Figure 67: Convergence of  $a$  for Diffusion with random forest



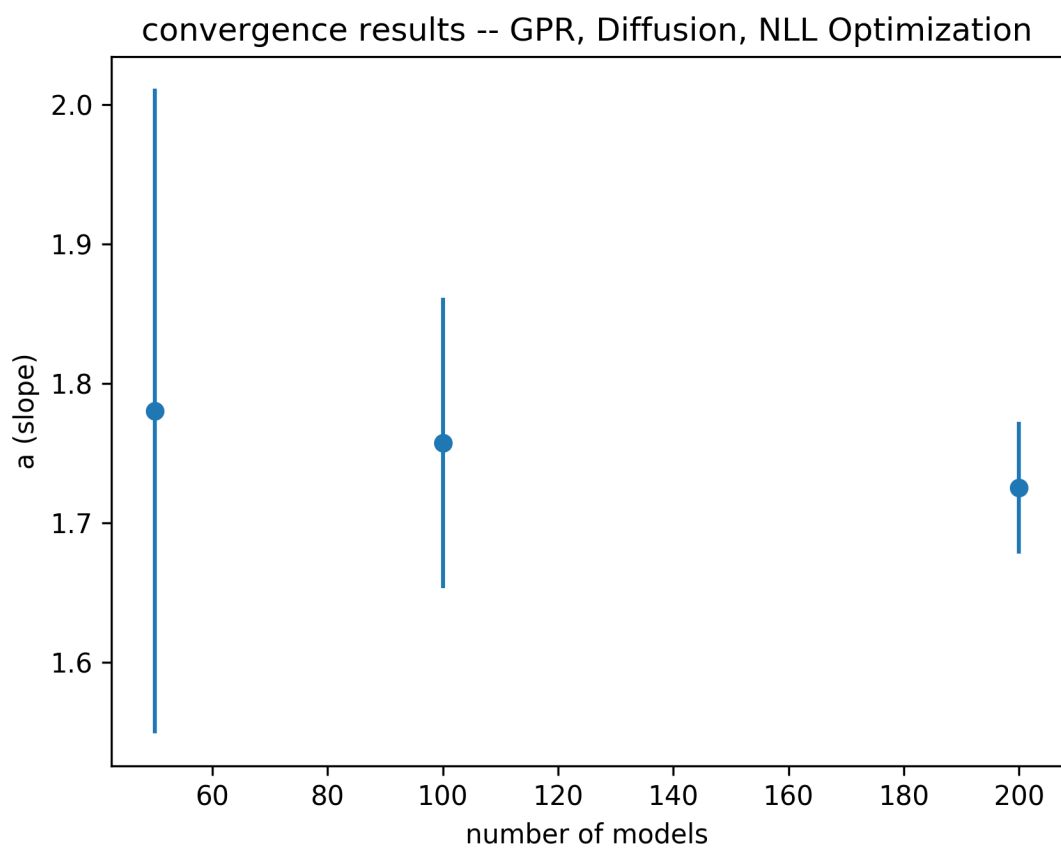
Supplementary Figure 68: Convergence of  $b$  for Diffusion with random forest



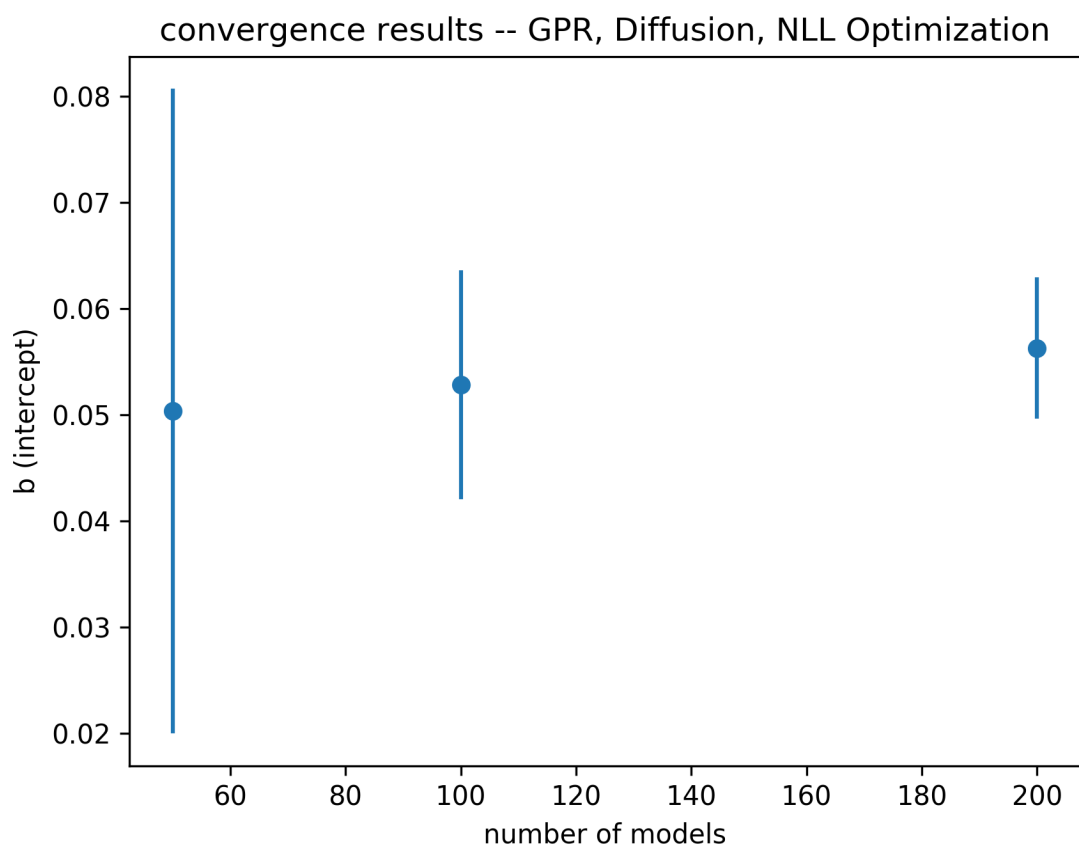
Supplementary Figure 69: Convergence of  $a$  for Diffusion with linear ridge regression



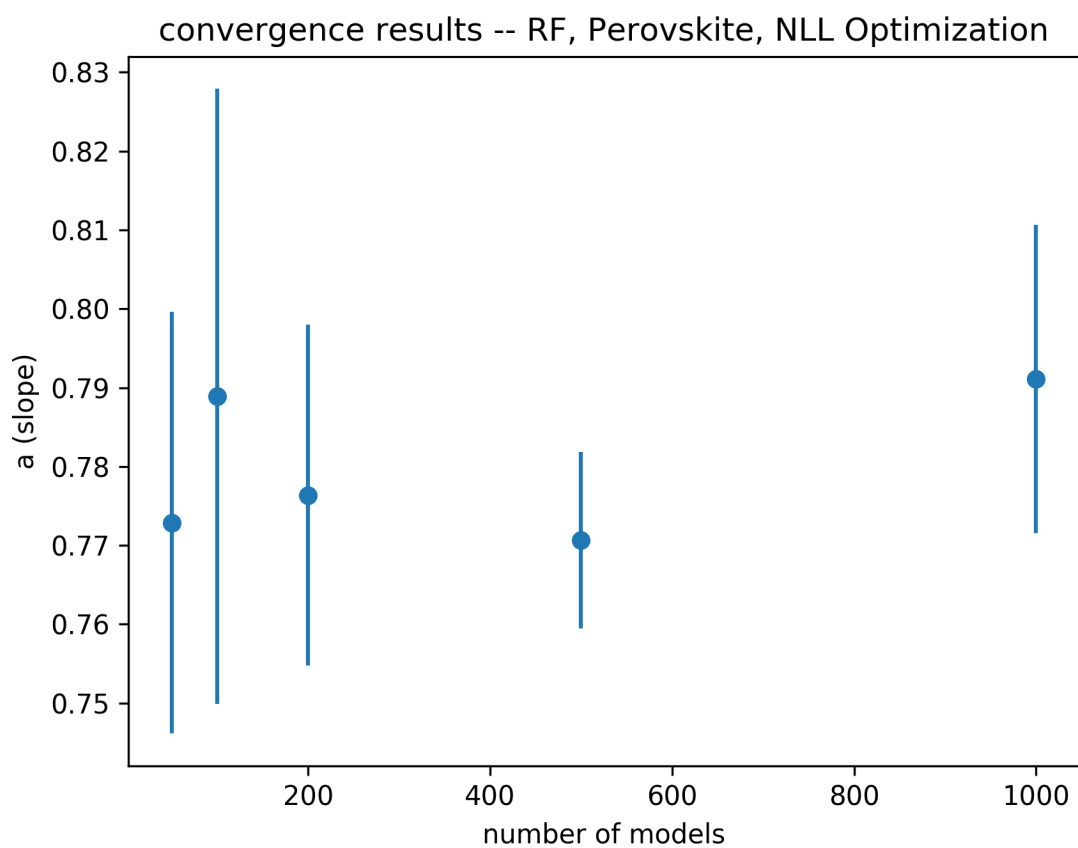
Supplementary Figure 70: Convergence of  $b$  for Diffusion with linear ridge regression



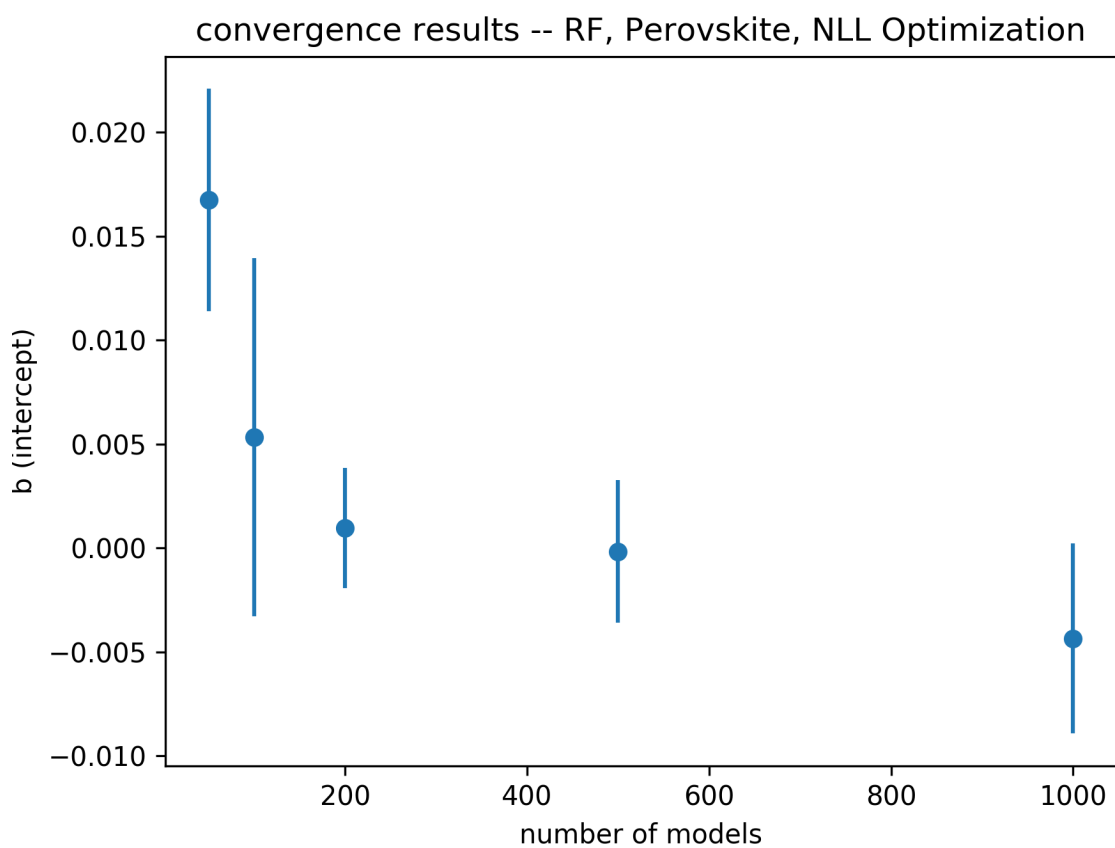
Supplementary Figure 71: Convergence of  $a$  for Diffusion with GPR



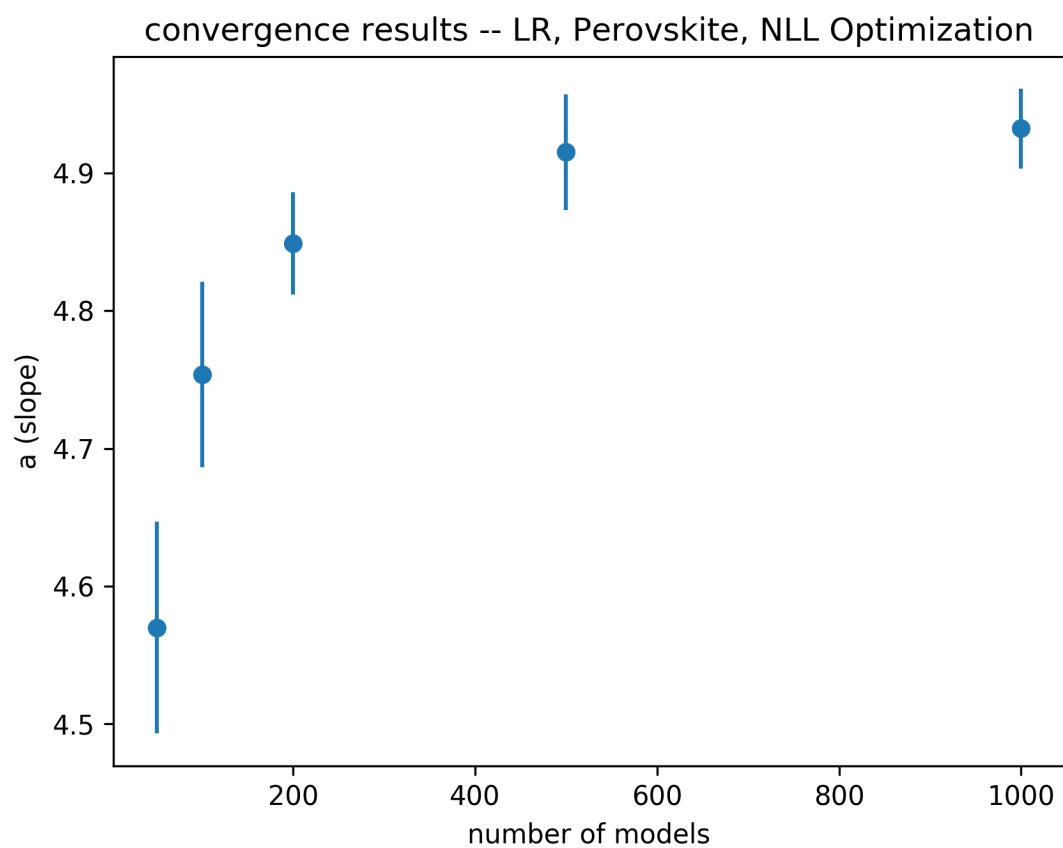
Supplementary Figure 72: Convergence of  $b$  for Diffusion with GPR



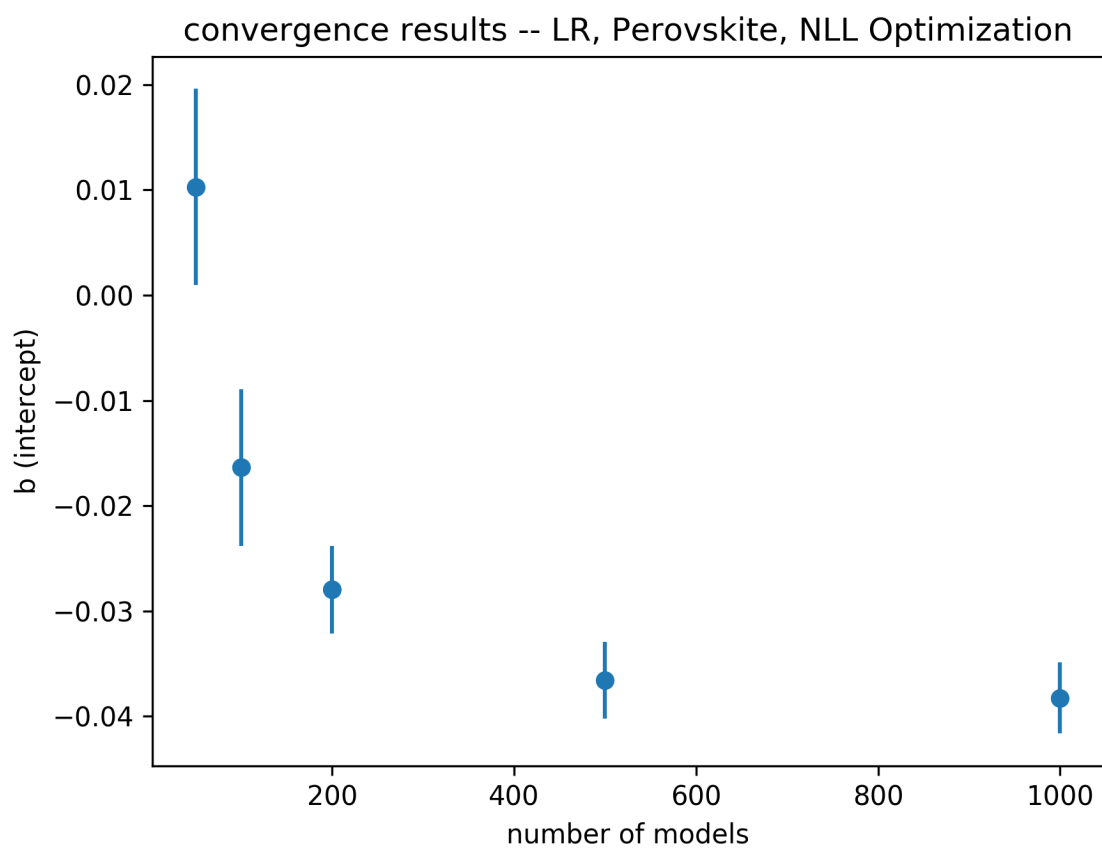
Supplementary Figure 73: Convergence of  $a$  for Perovskite with random forest



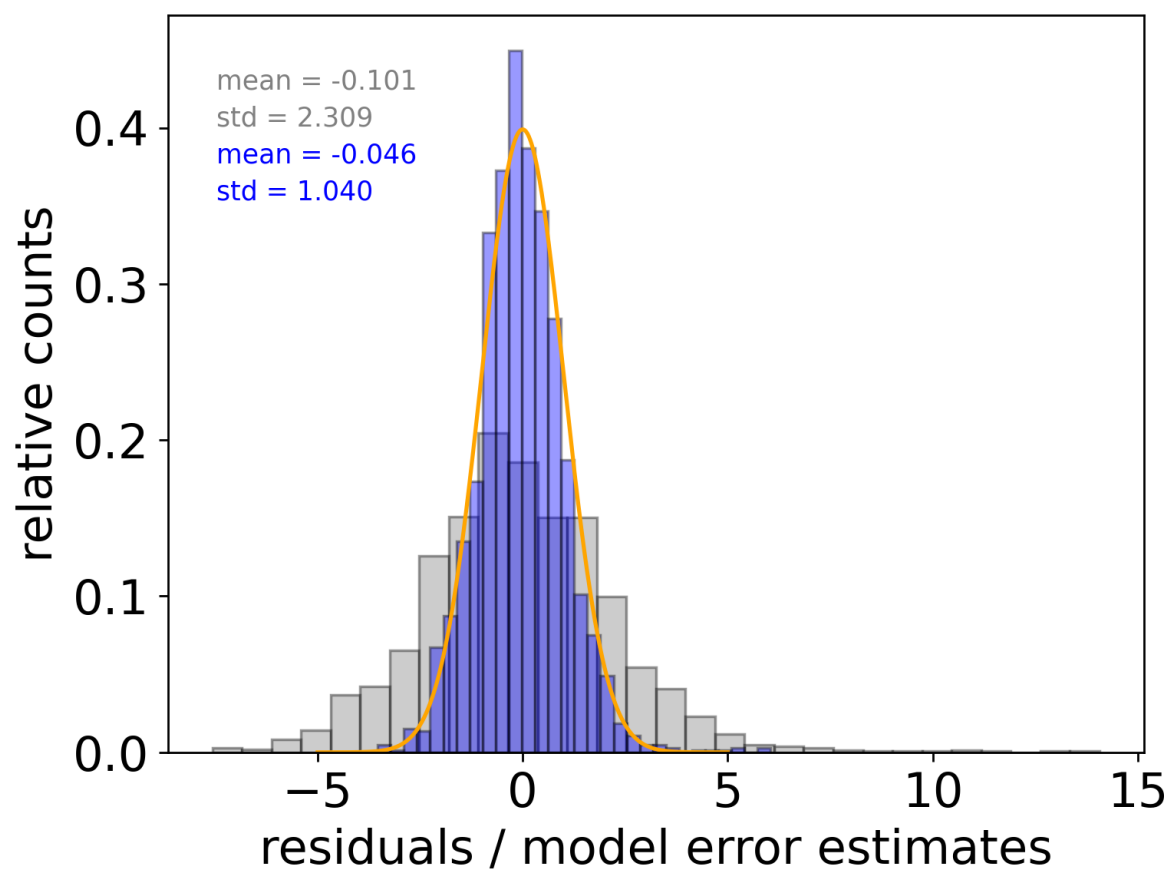
Supplementary Figure 74: Convergence of  $b$  for Perovskite with random forest



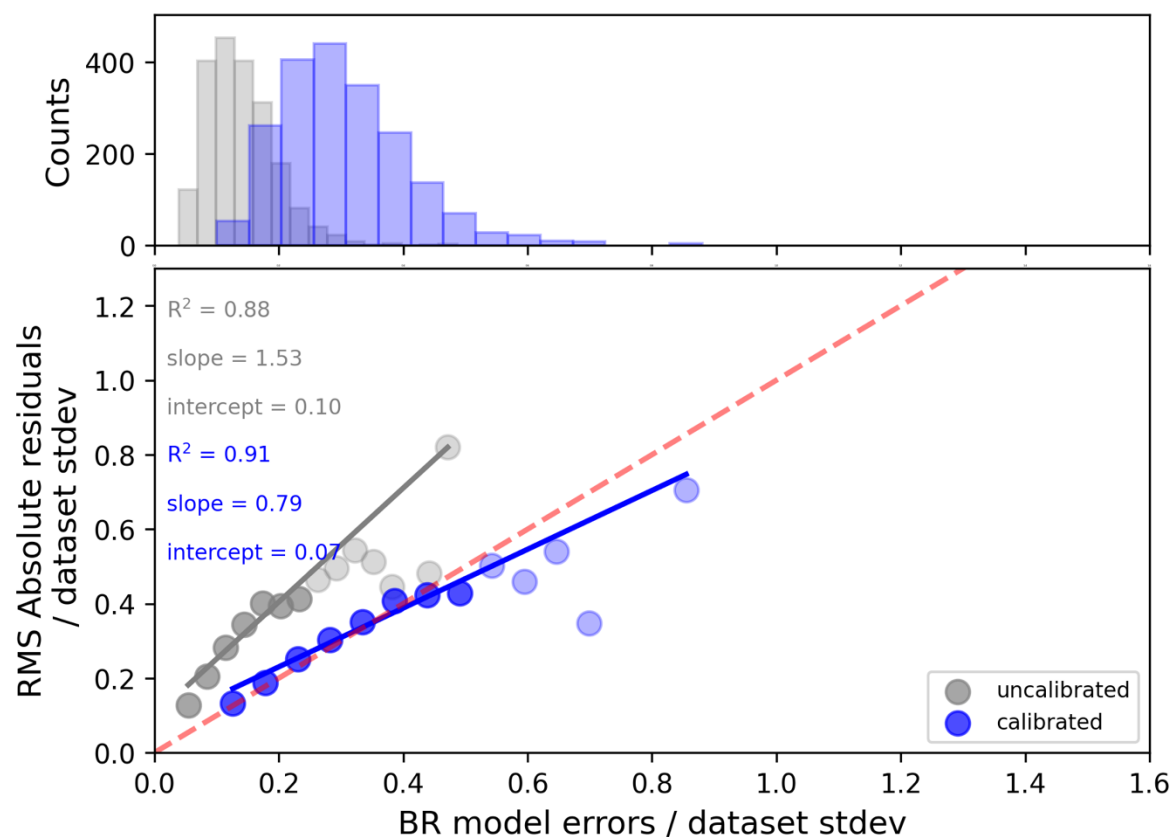
Supplementary Figure 75: Convergence of  $a$  for Perovskite with linear ridge regression



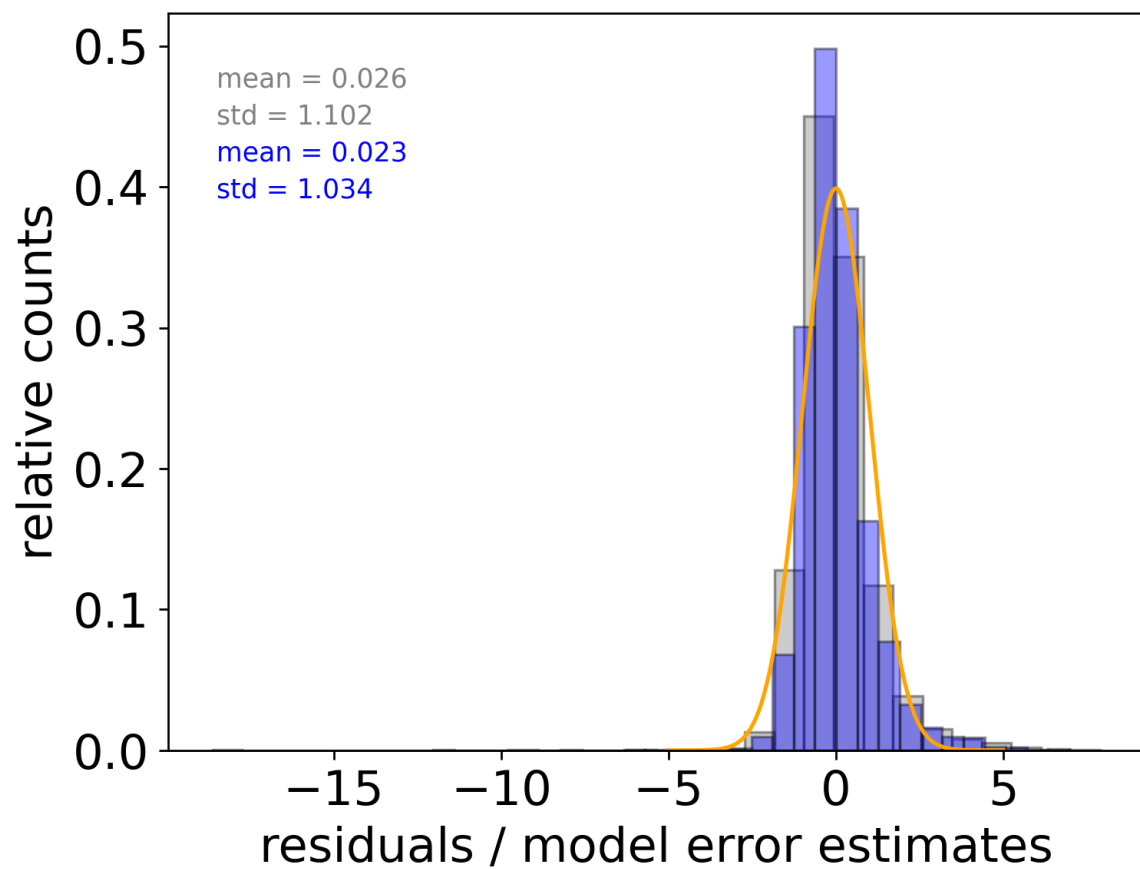
Supplementary Figure 76: Convergence of  $b$  for Perovskite with linear ridge regression



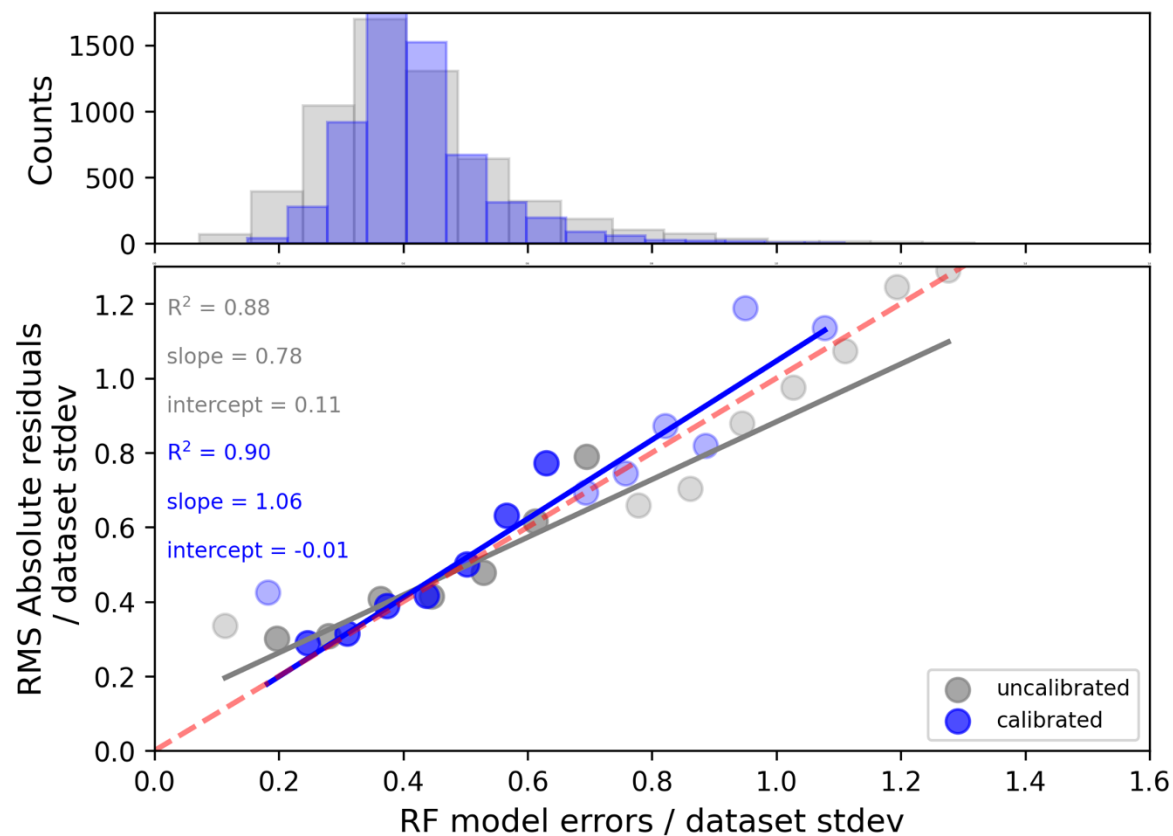
Supplementary Figure 77: Diffusion Neural Network (ensemble of 25) *r*-statistic



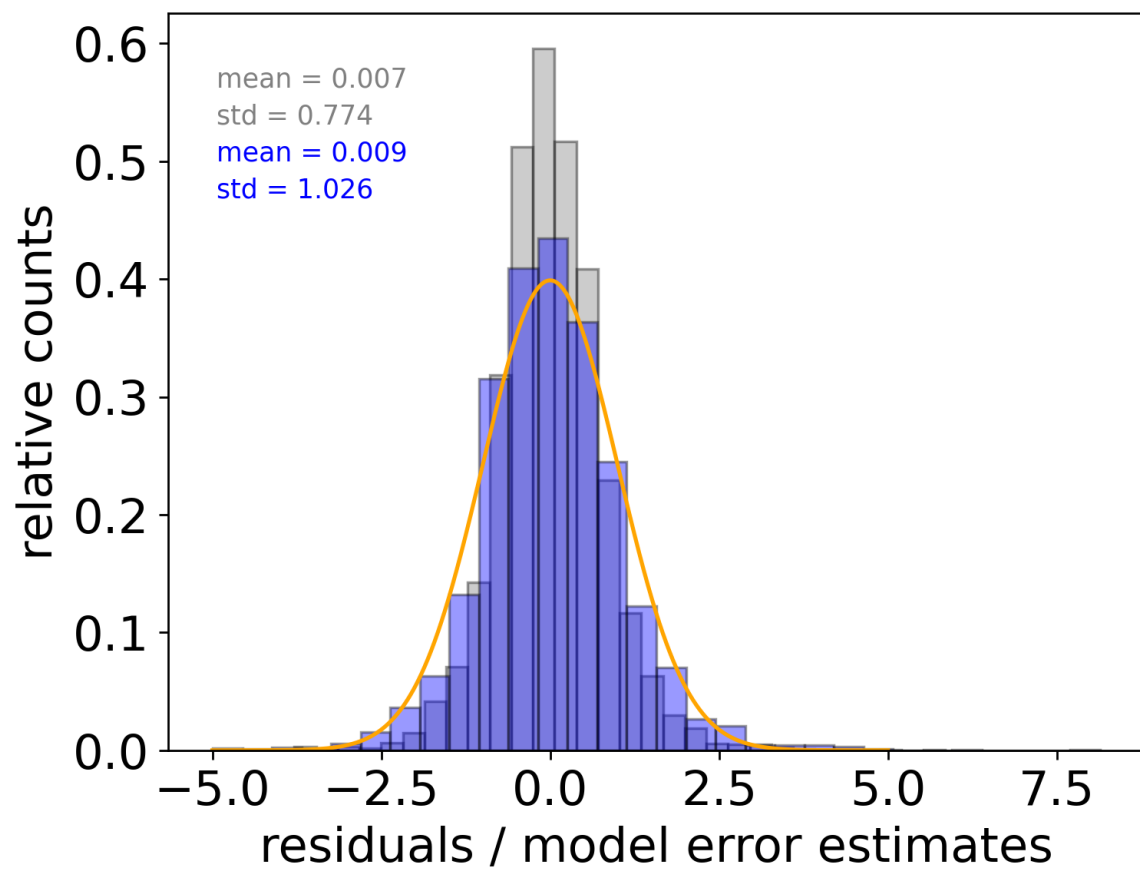
Supplementary Figure 78: Diffusion Neural Network (ensemble of 25) *RvE*



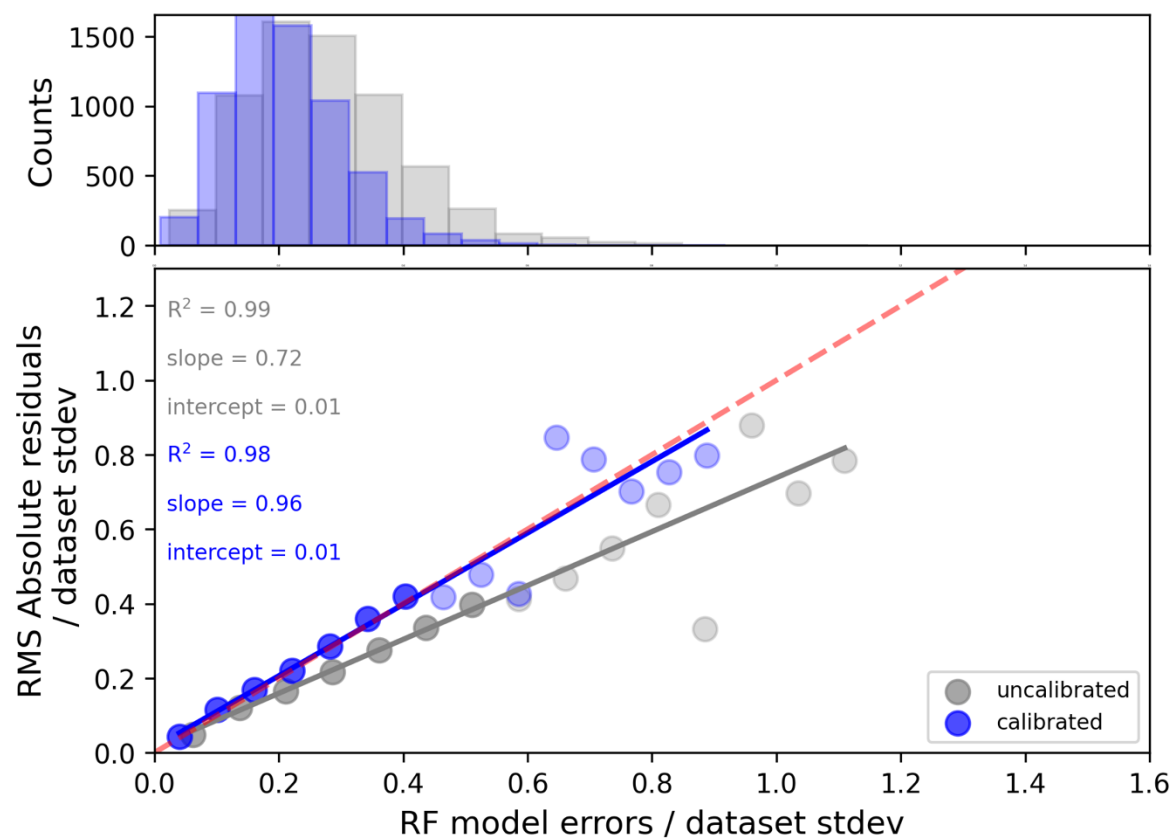
Supplementary Figure 79: Bulk modulus random forest  $r$ -statistic



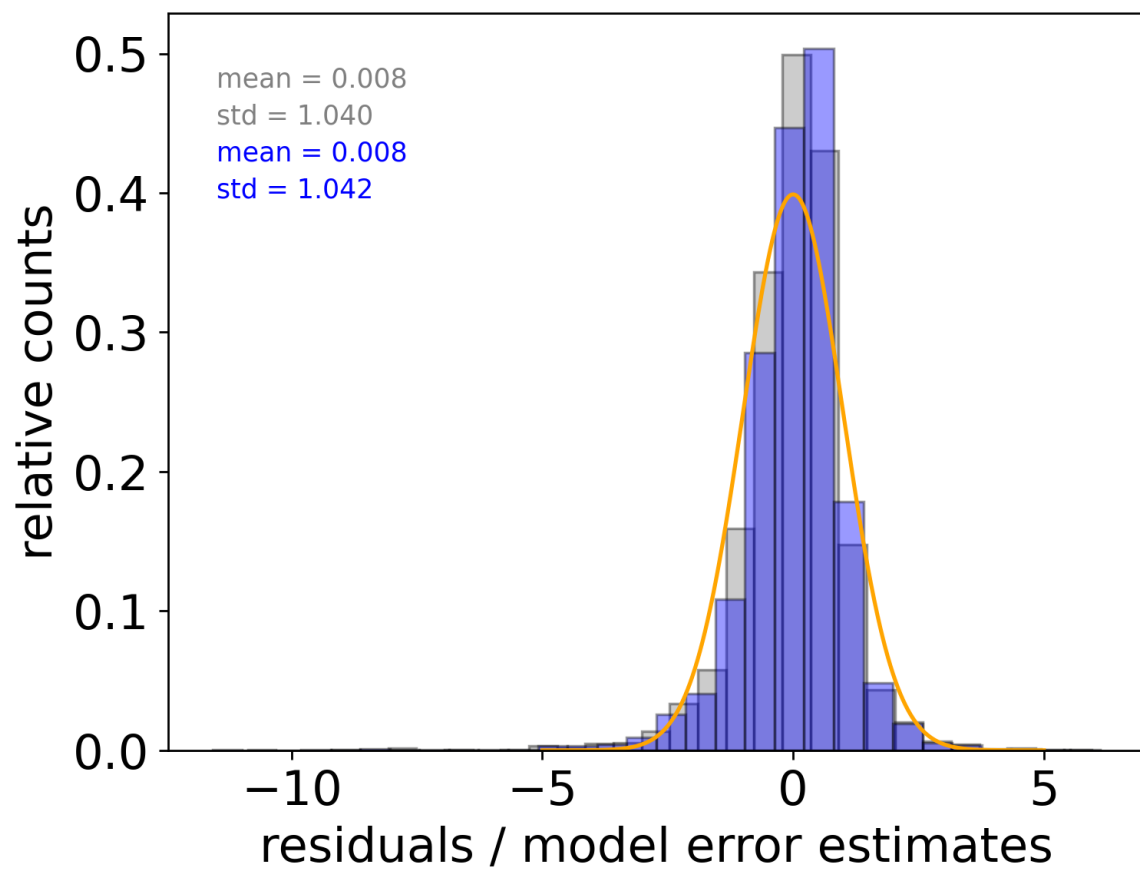
Supplementary Figure 80: Bulk modulus random forest  $R^2$



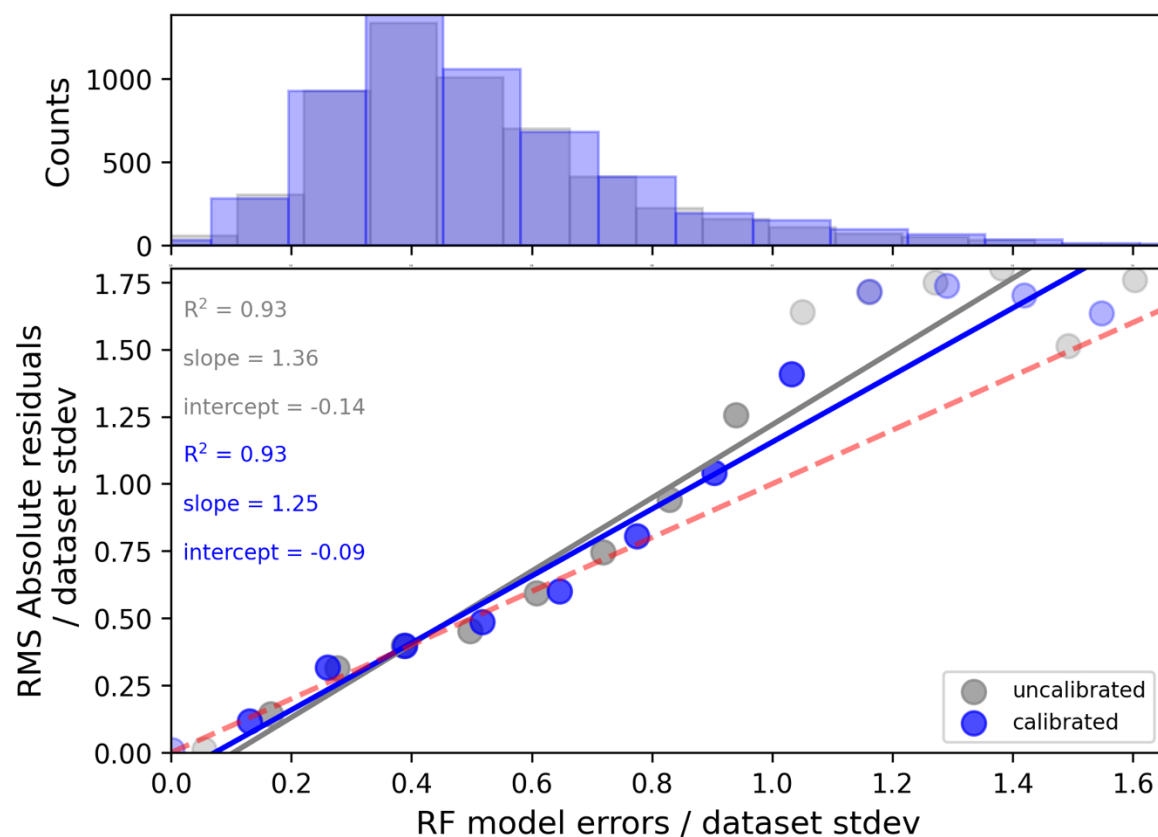
Supplementary Figure 81: Double perovskite gap random forest r-statistic



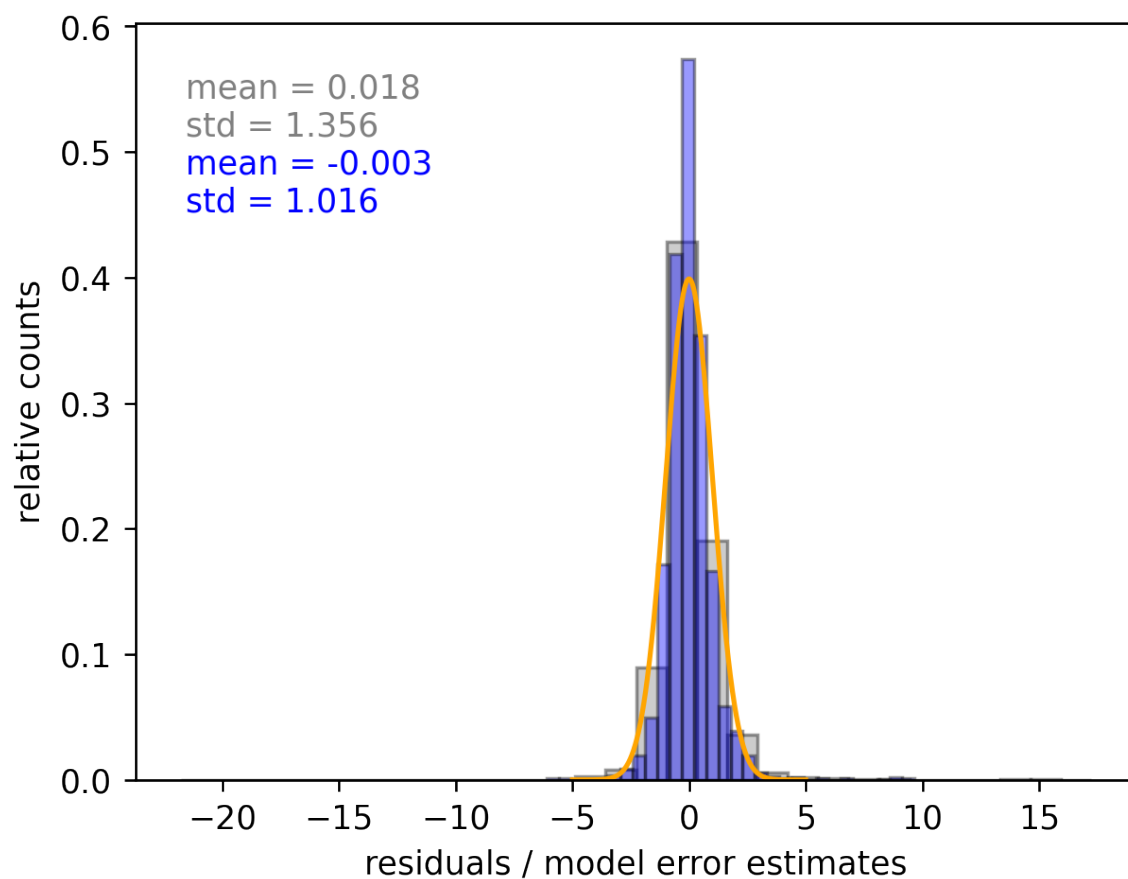
Supplementary Figure 82: Double perovskite gap random forest RvE



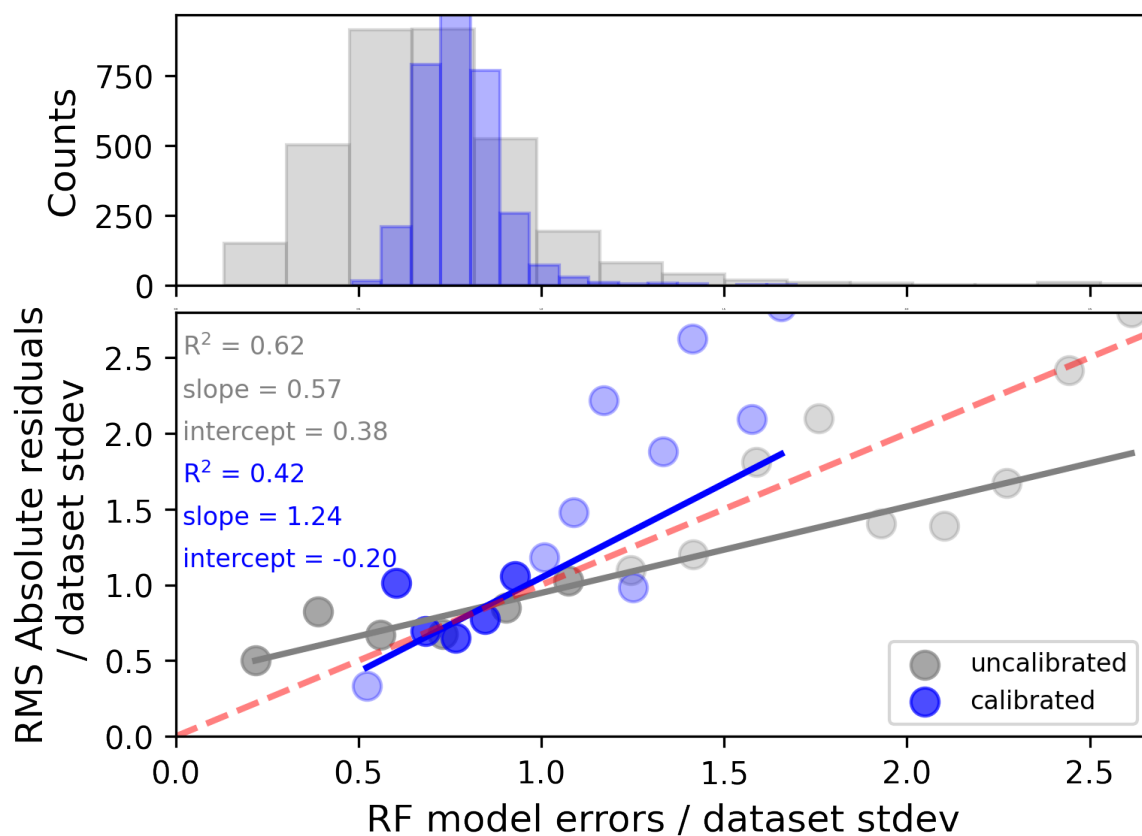
Supplementary Figure 83: Heusler random forest r-statistic



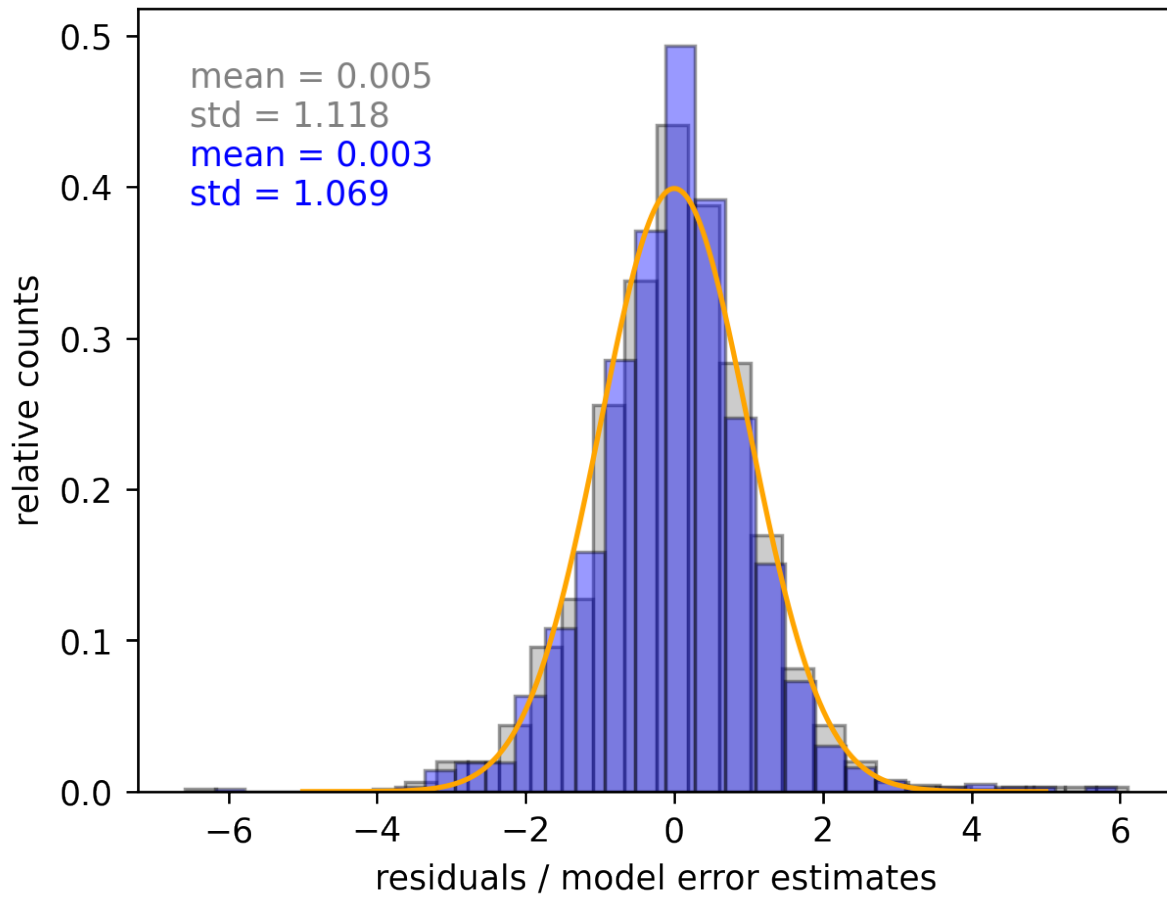
Supplementary Figure 84: Heusler random forest RvE



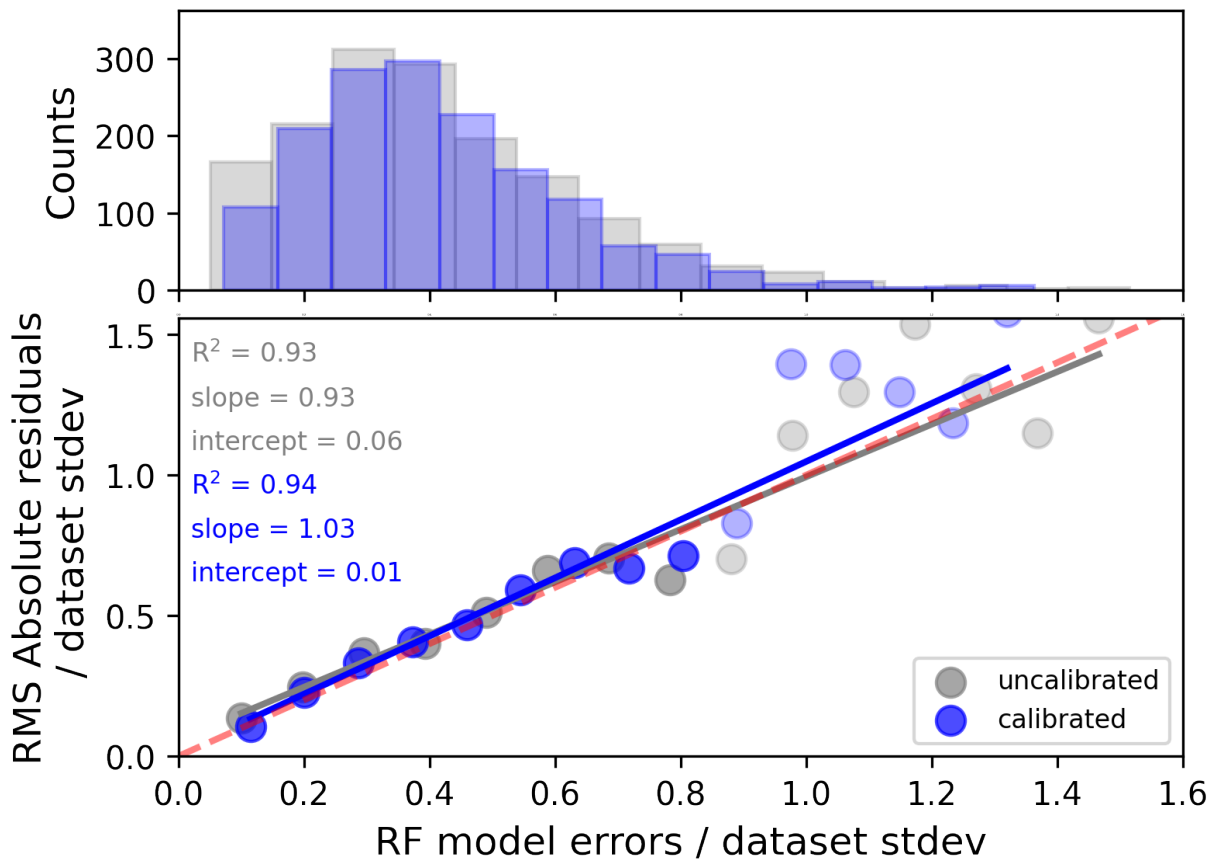
Supplementary Figure 85: Piezo random forest *r*-statistic



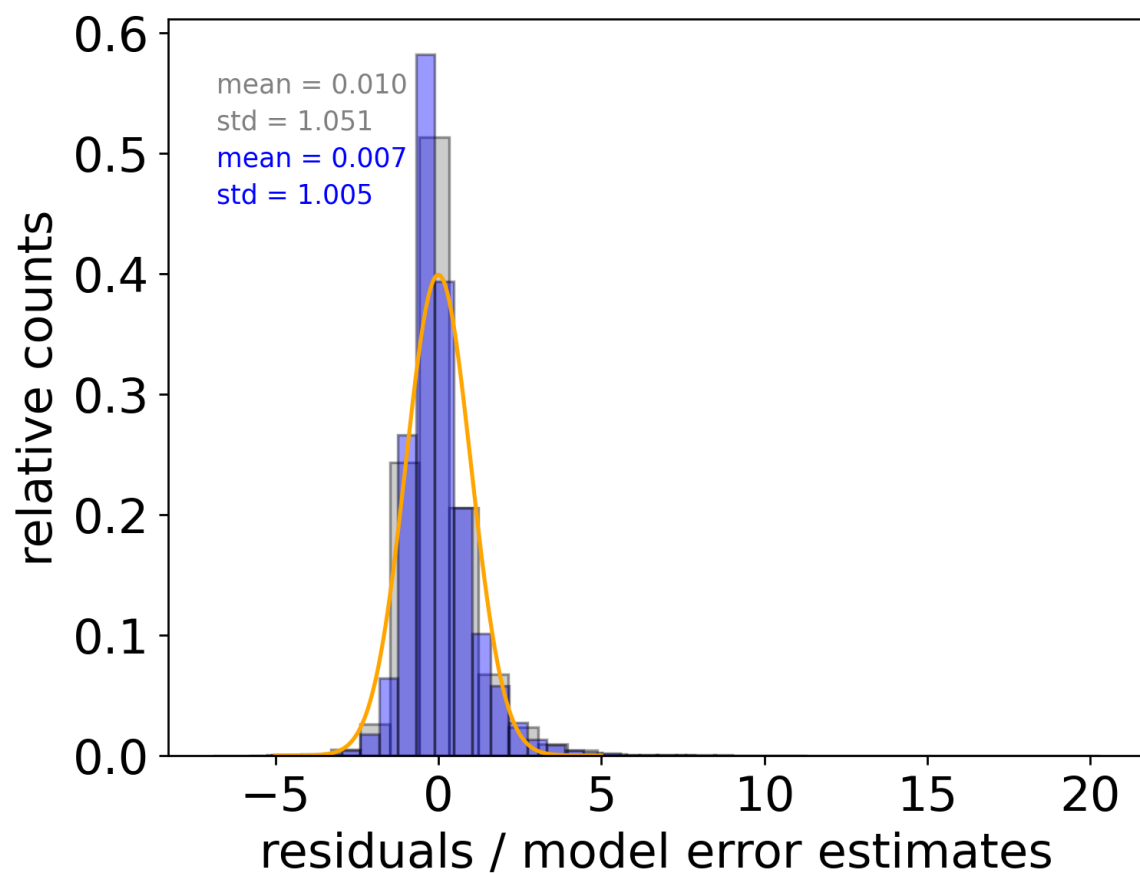
Supplementary Figure 86: Piezo random forest *R*<sub>vE</sub>



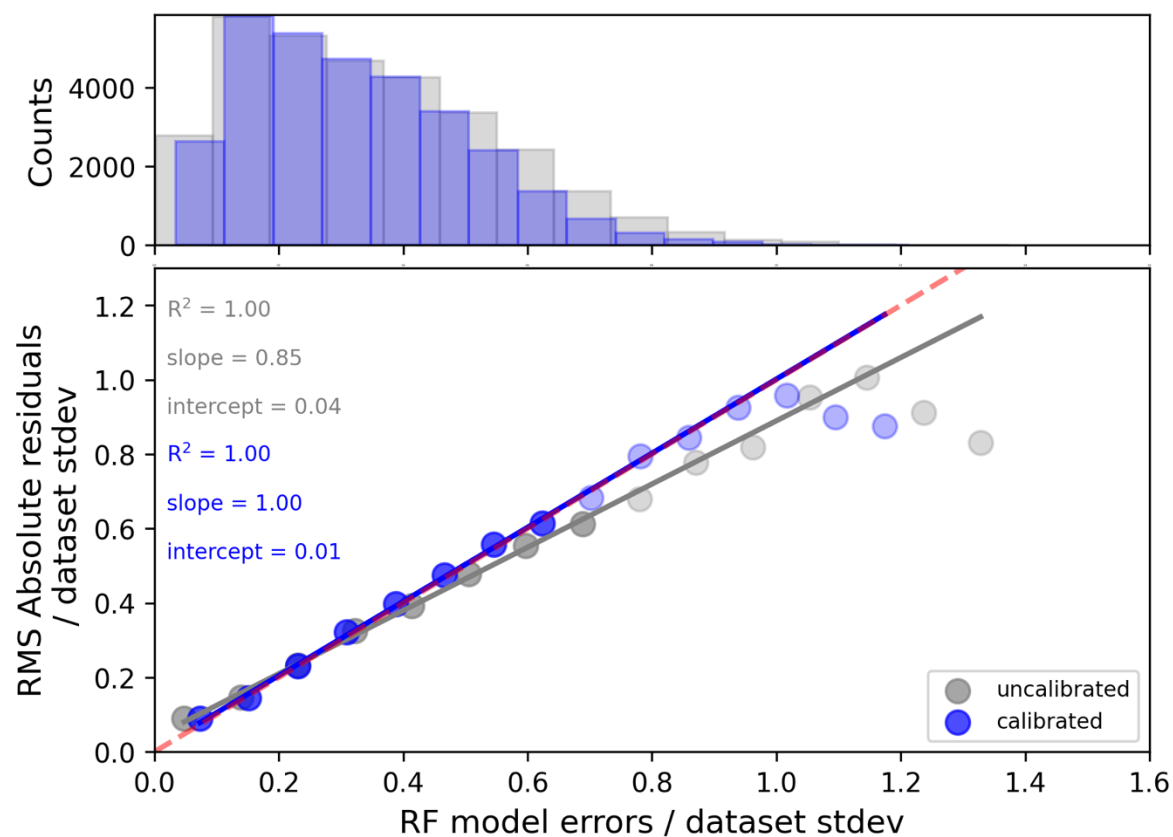
Supplementary Figure 87: Steel strength random forest  $r$ -statistic



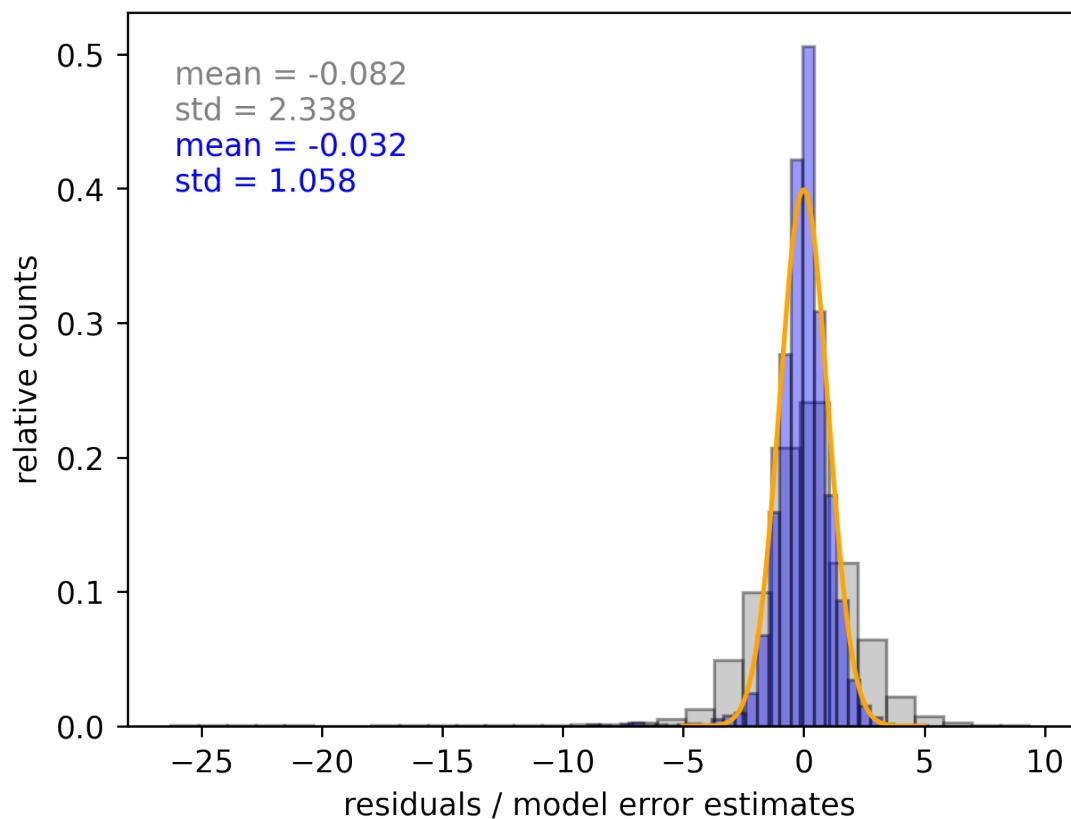
Supplementary Figure 88: Steel strength random forest  $RvE$



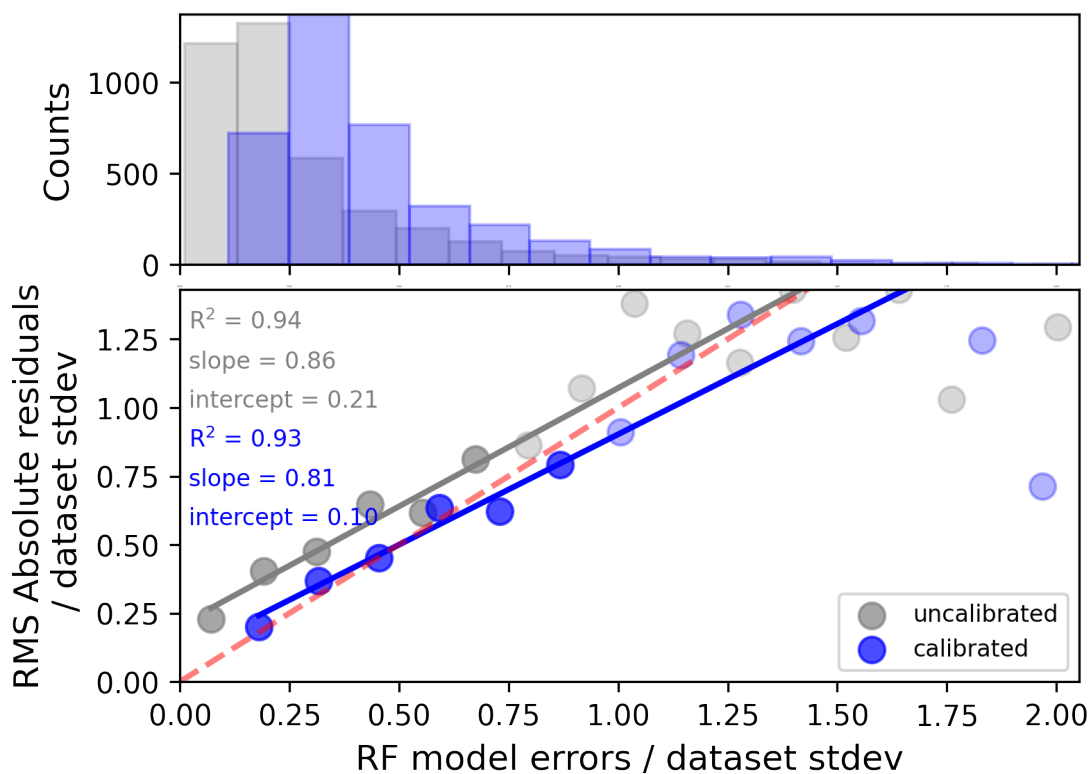
Supplementary Figure 89: Superconductor random forest r-statistic



Supplementary Figure 90: Superconductor random forest RvE



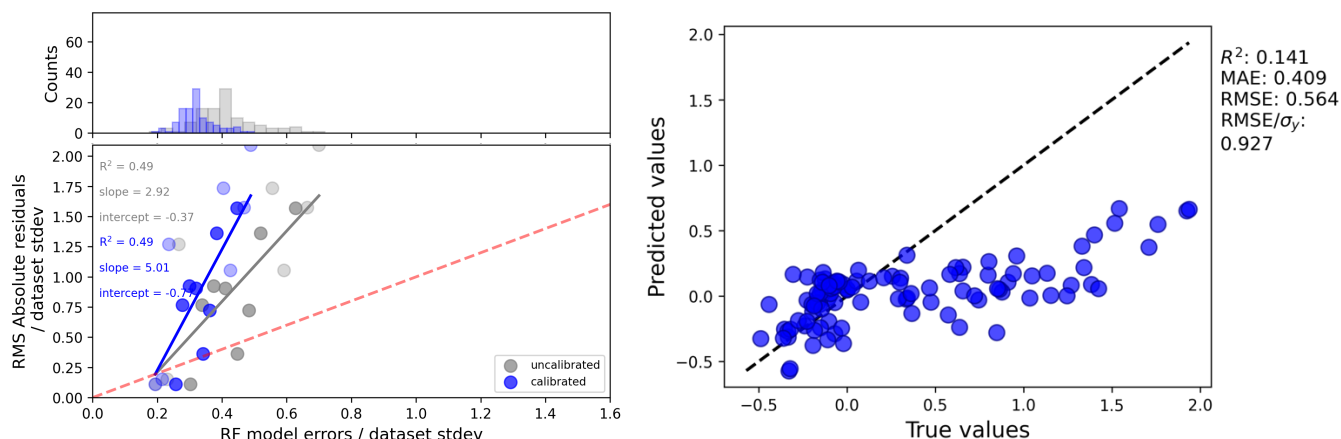
Supplementary Figure 91: Thermal conductivity random forest r-statistic



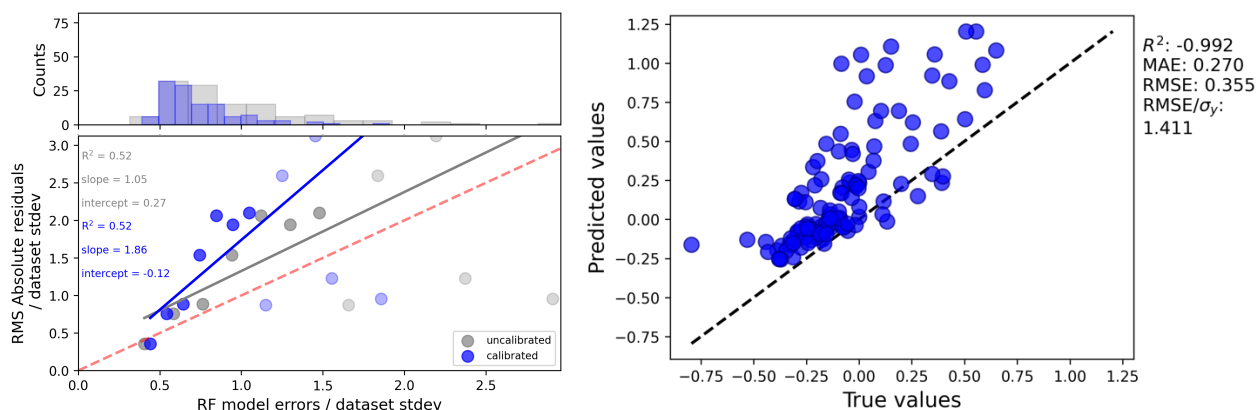
Supplementary Figure 92: Thermal conductivity random forest RvE

In Supplementary Figures 69-72, we apply our uncertainty quantification and recalibration method to the diffusion data with carefully chosen test sets that are expected to represent test data that is outside the domain of the trained model. Predictions on out of domain data are expected to result in poor model

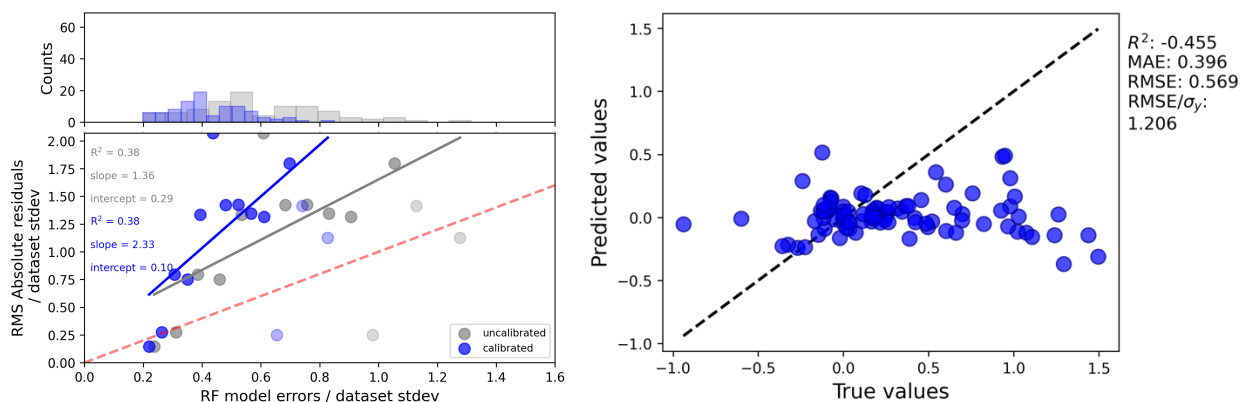
predictive ability and poor ability to accurately predict the model error bars. We use chemical intuition and domain knowledge to construct four different training-test set splits where the test set consists of materials we expect should be outside the training set. The approach is generally to create the test set by leaving out a specific type of chemistry for test. Test set 1 (Supplementary Figure 69): this is a “leave out non-transition metal host” test where the left out data are Al, Ca, and Mg hosts, which were chosen because they are the only host materials in the dataset that are not transition metals. Test set 2 (Supplementary Figure 70): this is a “leave out 3d transition metal host” test where the left out data are Fe, Cu and Ni hosts, which were chosen because they are the only host materials in the dataset that are transition metals belonging to the 3d row of the periodic table. Test set 3 (Supplementary Figure 71): this is a “leave out 4d/5d transition metal host” test where the left out data are Ag, Pd, Ir, Au and Pt hosts, which were chosen because they are the only host materials in the dataset that are transition metals belonging to the 4d or 5d row of the periodic table. Test set 4 (Supplementary Figure 72): this is a “leave out refractory metal host” test where the left out data are Mo, W and Zr hosts, which were chosen because they are the only host materials in the dataset that are refractory elements. In each case we show the RvE and parity plots, which demonstrate that the calibration does not work and that the model is highly inaccurate, respectively, both results consistent with the test data being outside the domain of the trained model.



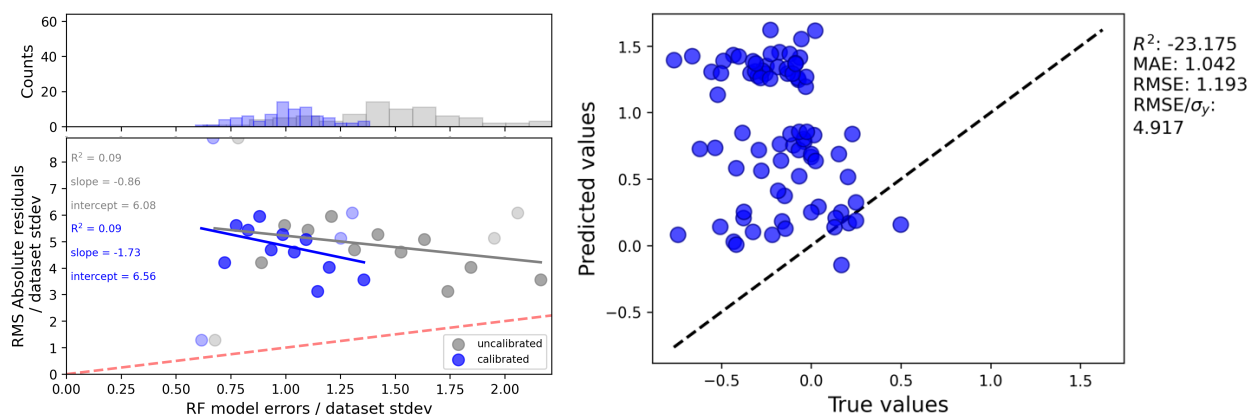
Supplementary Figure 93: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out non-transition metal host” domain test where the Al, Ca and Mg hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data



Supplementary Figure 94: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a “leave out 3d transition metal host” domain test where the Fe, Cu and Ni hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data.



Supplementary Figure 95: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a "leave out 4d/5d transition metal host" domain test where the Ag, Pd, Ir, Pt and Au hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data.



Supplementary Figure 96: Diffusion data random forest RvE (left-hand side) and parity plot (right-hand side) for a "leave out refractory metal host" domain test where the Mo, W and Zr hosts were left out and the remaining host data (except Pb) was used for training. Recalibration factors were obtained from 5 iterations of 5-fold CV (25 splits) on the training data.