|  |
| --- |
| **Figure 1** |
| *Violin plots of R2 for every model on ACT and ECLSK science datasets* |
|  |
| *Note: R2* over tentimes 10-CV (100 folds). Black dots indicate mean *R2.* Red, blue, yellow and green violins indicate GLMM tree, M-BART, most accurate PPD-based BA tree, and most accurate smearing-based BA tree respectively. |

|  |
| --- |
| **Figure 3** |
| *Effect of N on R2 and tree size for every dataset* |
|  |
| *Note: R2 (left) and tree size (right)* over tentimes 10-CV (100 folds) for all BA GLMM models on all seven datasets. *Blue dotted line marked by “GLMM” marks the median R2 or tree size of GLMM models. Red dotted line marked by “BART” marks the median R2 of BART model. Red dots connected by red solid line indicate median R2 or tree size.* |

|  |
| --- |
| **Figure 4** |
| *Effect of palt on R2 and tree size for every dataset* |
|  |
| *Note: R2 (left) and tree size (right)* over tentimes 10-CV (100 folds) for all BA GLMM models on all seven datasets. *Blue dotted line marked by “GLMM” marks the median R2 or tree size of GLMM models. Red dotted line marked by “BART” marks the median R2 of BART model. Red dots connected by red solid line indicate median R2 or tree size.* |

|  |
| --- |
| **Figure 2** |
| *Mean R2 of GLMM, M-BART, and most accurate BA GLMM models on ACT and ECLSK science datasets.* |
|  |
| *Note: 10 times 10-CV results in R2 per model. Bars over boxplots indicate results of Tukey tests. ns: not significant, \*: p < .05, \*\*: p < .01, \*\*: p < .01, \*\*\*: p < .001, \*\*\*\*: p < .0001.* |

|  |
| --- |
| **Table Appendix 1** |
| *Mean R2 and standard deviation for every model on every dataset* |
| *Note: M* is mean *R2* over tentimes 10-CV (100 folds). *SD* is standard deviation. |

|  |
| --- |
| **Table Appendix 2** |
| *Mean tree size and standard deviation for every model on every dataset* |
| *Note: M* is mean tree sizeover tentimes 10-CV (100 folds). *SD* is standard deviation. |

|  |
| --- |
| **Figure Appendix 3** |
| *Effect of palt on tree size for every model grouped by N* |
|  |
| *Note: Blue dotted line marked by “GLMM” marks the median tree size of GLMM model. Red dots connected by red solid line indicate median tree size.* |

|  |
| --- |
| **Table 1** |
| *R2 and Tree Size for ACT and ECLSK science datasets* |
|  |
| *Note.* For PPD- and Smearing-based BA trees, only the model with the highest R2 is displayed here. For the complete results see Appendix 1 and Appendix 2. |

|  |
| --- |
| **Table 2** |
| *R2 and Tree Size for ACT and ECLSK science datasets where generated data is Ngen = 10 times the size of the original dataset* |
|  |
| *Note.* For Smearing-based BA GLMM trees, only the model with the highest R2 is displayed here. For the complete results see Appendix 1 and Appendix 2. |