Supplementary Information to Milledge et al., 2021
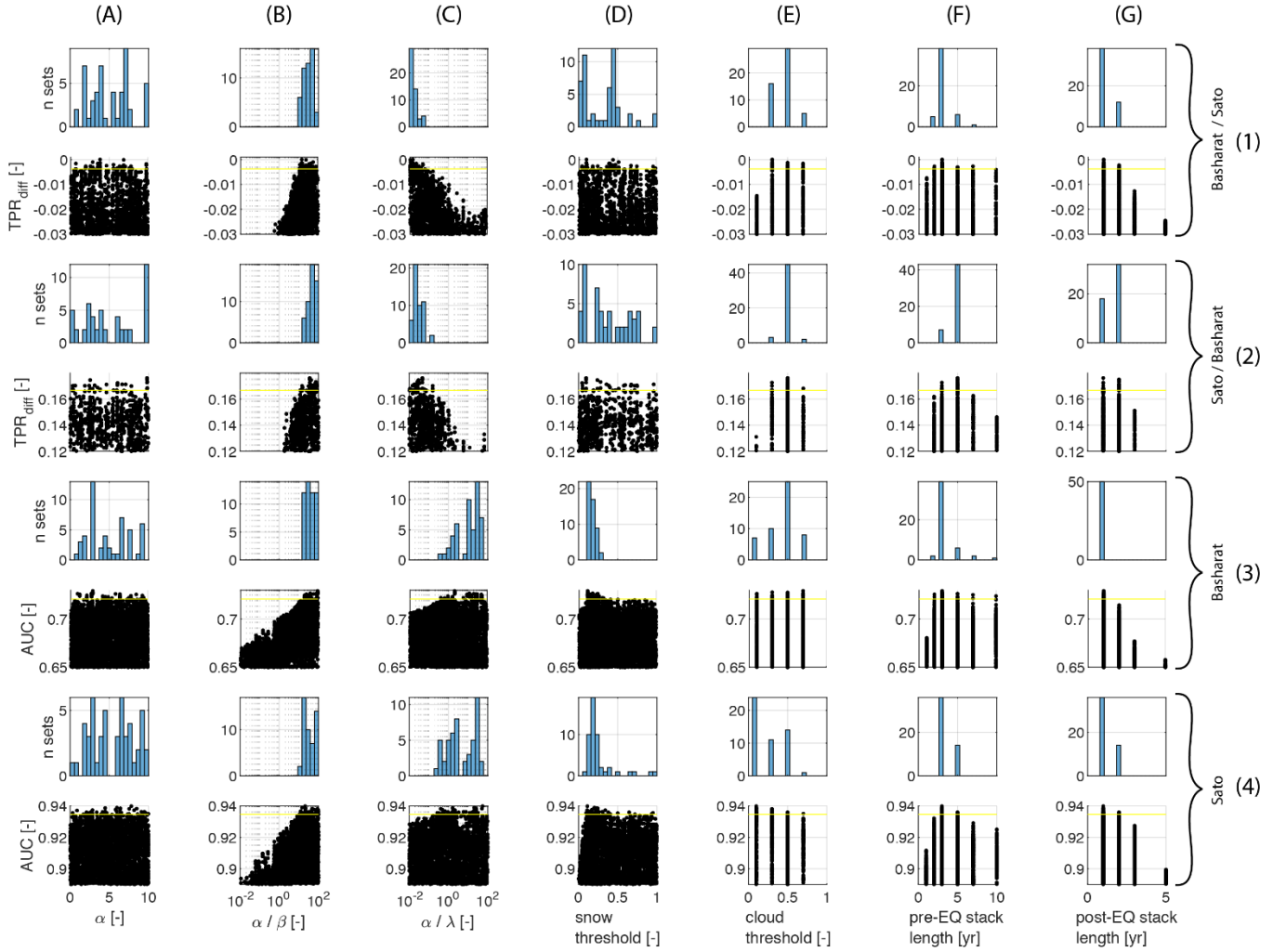
Figure S1:Dotty plots and posterior parameter distributions for Kashmir case study for the seven tuneable parameters associated with ALDI (columns) evaluated using the two of the test datasets (Sato et al., 2007; Basharat et al., 2016) and two performance metrics (rows): TPR_diff, the difference in TPR between ALDI and the competitor inventory at the FPR defined by the competitor inventory; and AUC, the area under the ROC curve, a more general indicator of classifier performance over the full range of FPRs. 'Basharat/Sato' refers to using Basharat as the check dataset and Sato as the competitor in row 1; 'Sato/Basharat' refers to the converse in row 2. Basharat is used as the check dataset in row 3, and Sato as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.
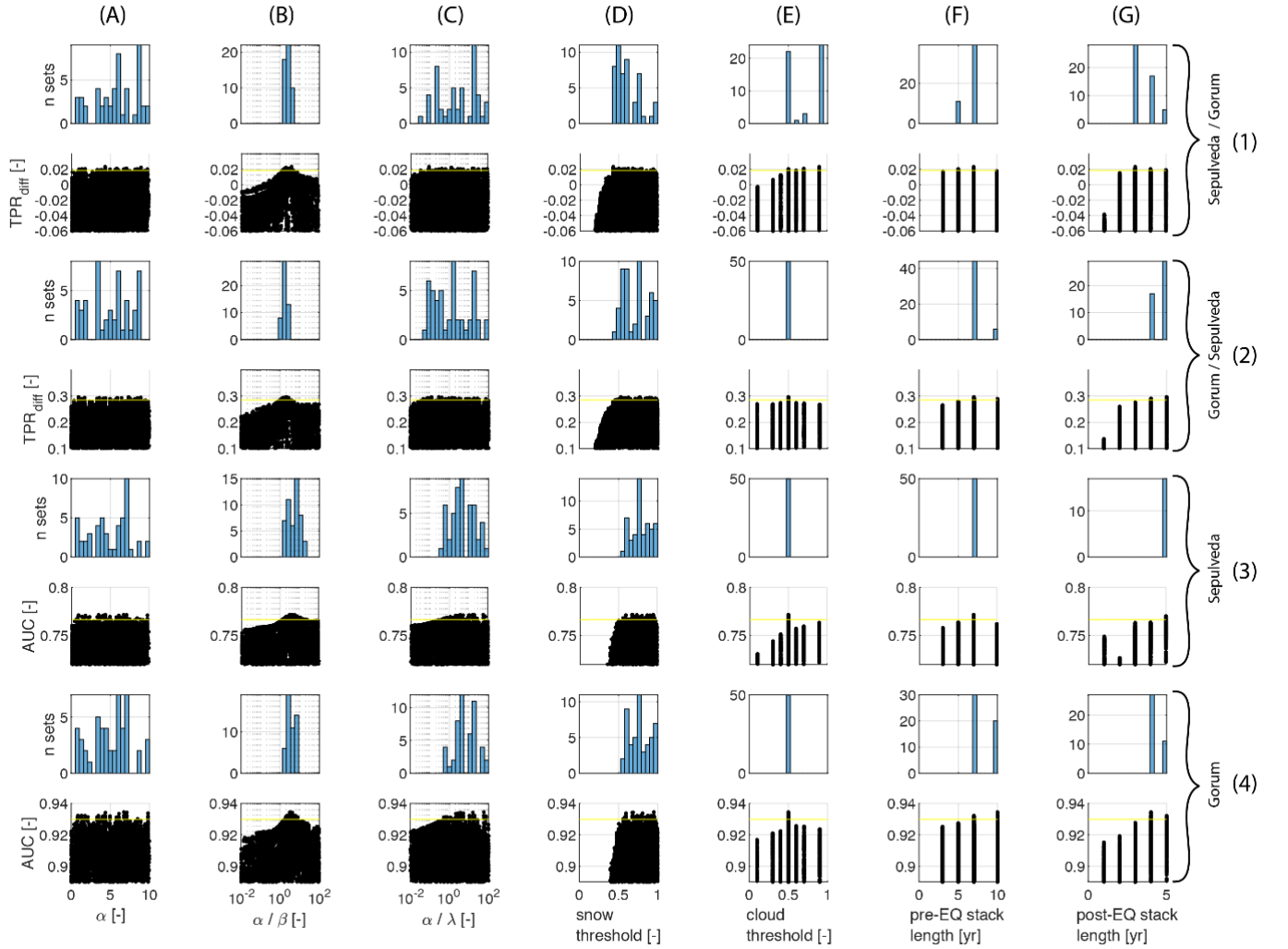
Figure S2: Dotty plots and posterior parameter distributions for Aisen case study for the seven ALDI parameters (columns) evaluated using the two of the test datasets (Sepulveda et al., 2010; Gorum et al., 2014) and two performance metrics (rows): TPR$_{diff}$ and AUC. 'Sepulveda/Gorum' refers to using Sepulveda as the check dataset and Gorum as the competitor in row 1; 'Gorum/Sepulveda' refers to the converse in row 2. Sepulveda is used as the check dataset in row 3, and Gorum as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.

15
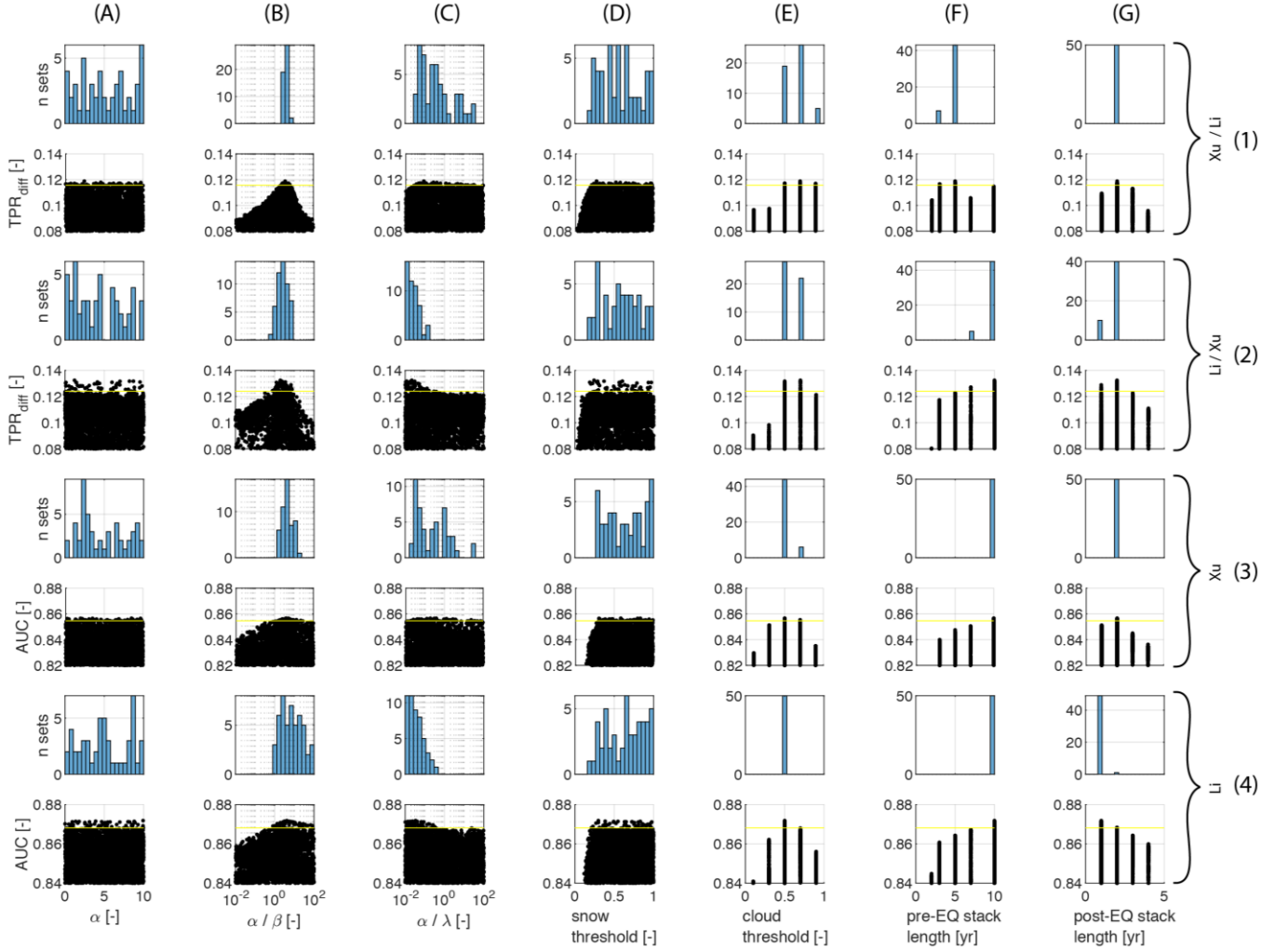
Figure S3: Dotty plots and posterior parameter distributions for Wenchuan case study for the seven ALDI parameters (columns) evaluated using the two of the test datasets (Li et al., 2014; Xu et al., 2014) and two performance metrics (rows): TPR_{diff} and AUC. 'Xu/Li' refers to using Xu as the check dataset and Li as the competitor in row 1; 'Li/Xu refers to the converse in row 2. Xu is used as the check dataset in row 3, and Li as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.
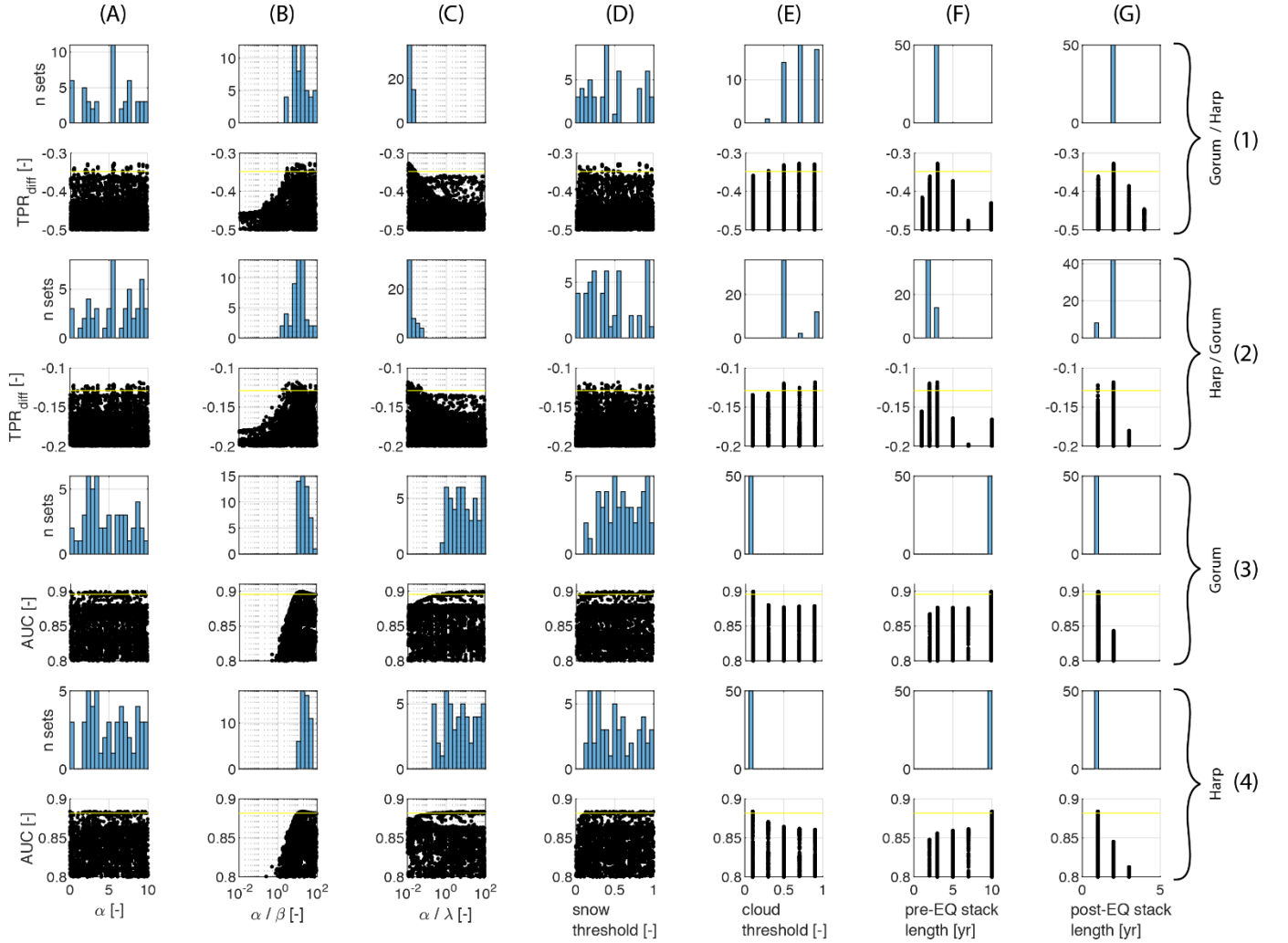
Figure S4: Dotty plots and posterior parameter distributions for Haiti case study for the seven ALDI parameters (columns) evaluated using the two of the test datasets (Gorum et al., 2013; Harp et al., 2016) and two performance metrics (rows): TPR$_{diff}$ and AUC. 'Gorum/Harp' refers to using Gorum as the check dataset and Harp as the competitor in row 1; 'Harp/Gorum' refers to the converse in row 2. Gorum is used as the check dataset in row 3, and Harp as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.
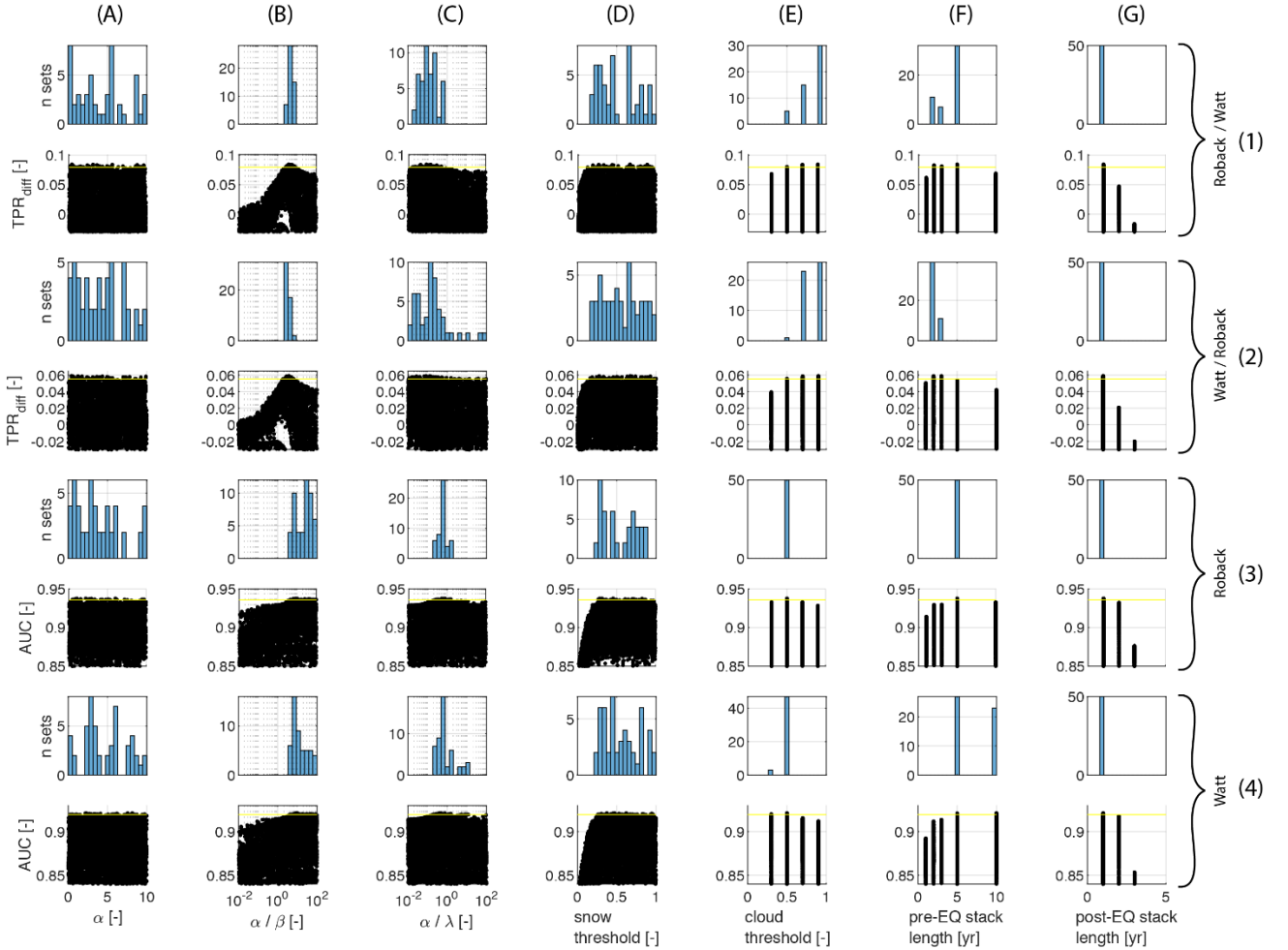
Figure S5: Dotty plots and posterior parameter distributions for Gorkha case study for the seven ALDI parameters (columns) evaluated using two of the test datasets (Roback et al., 2018; Watt, 2016) and two performance metrics (rows): TPR$_{diff}$ and AUC. 'Roback/Watt' refers to using Roback as the check dataset and Watt as the competitor in row 1; 'Watt/Roback' refers to the converse in row 2. Roback is used as the check dataset in row 3, and Watt as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.

40
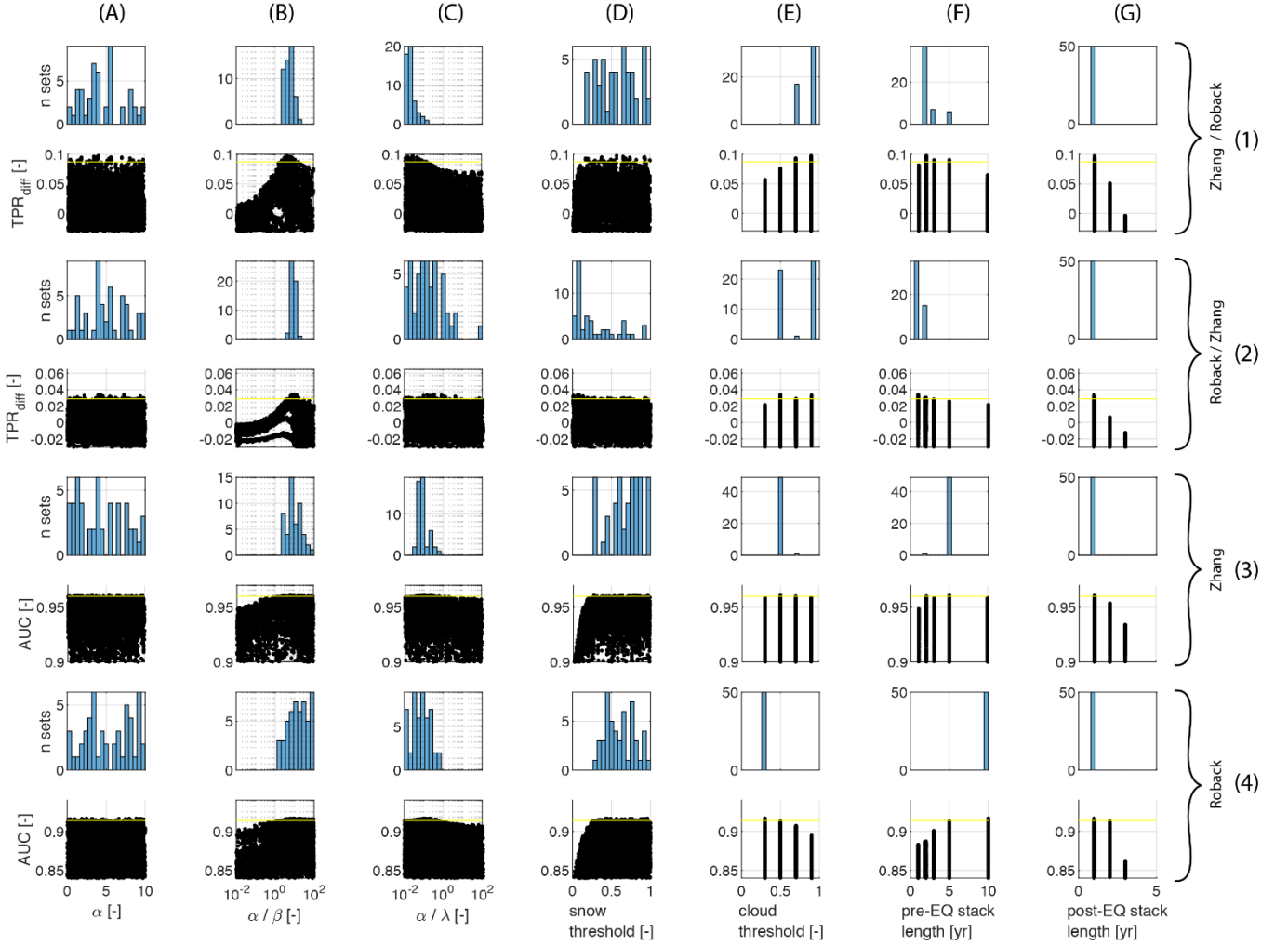
Figure S6: Dotty plots and posterior parameter distributions for Gorkha case study for the seven ALDI parameters (columns) evaluated using two of the test datasets (Zhang et al., 2016; Roback et al., 2018) and two performance metrics (rows): TPR_{diff} and AUC. 'Zhang/Roback' refers to using Zhang as the check dataset and Roback as the competitor in row 1; 'Roback/Zhang' refers to the converse in row 2. Zhang is used as the check dataset in row 3, and Roback as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.
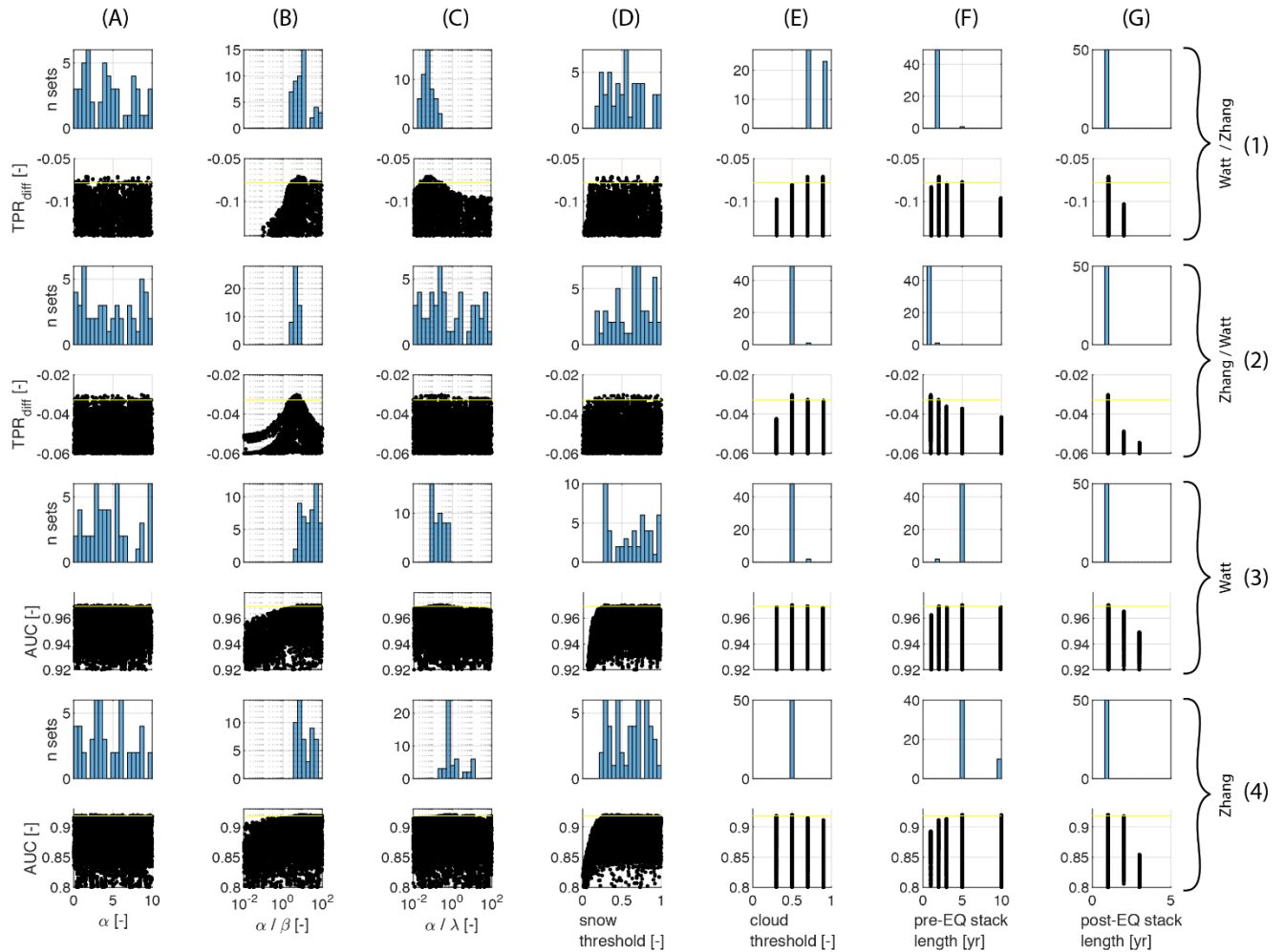
Figure S7: *Dotty plots and posterior parameter distributions for Gorkha case study for the seven ALDI parameters (columns) evaluated using two of the test datasets (Zhang et al., 2016; Watt, 2016) and two performance metrics (rows): TPR$_{diff}$ and AUC. 'Watt/Zhang' refers to using Watt as the check dataset and Zhang as the competitor in row 1; 'Zhang/Watt' refers to the converse in row 2. Watt is used as the check dataset in row 3, and Zhang as the check dataset in row 4. Points plotting above the yellow line are results for the best 100 parameters. In each case the parameter distributions are for the best 100 parameter sets evaluated using the same metric and datasets as the dotty plot below it.*

55