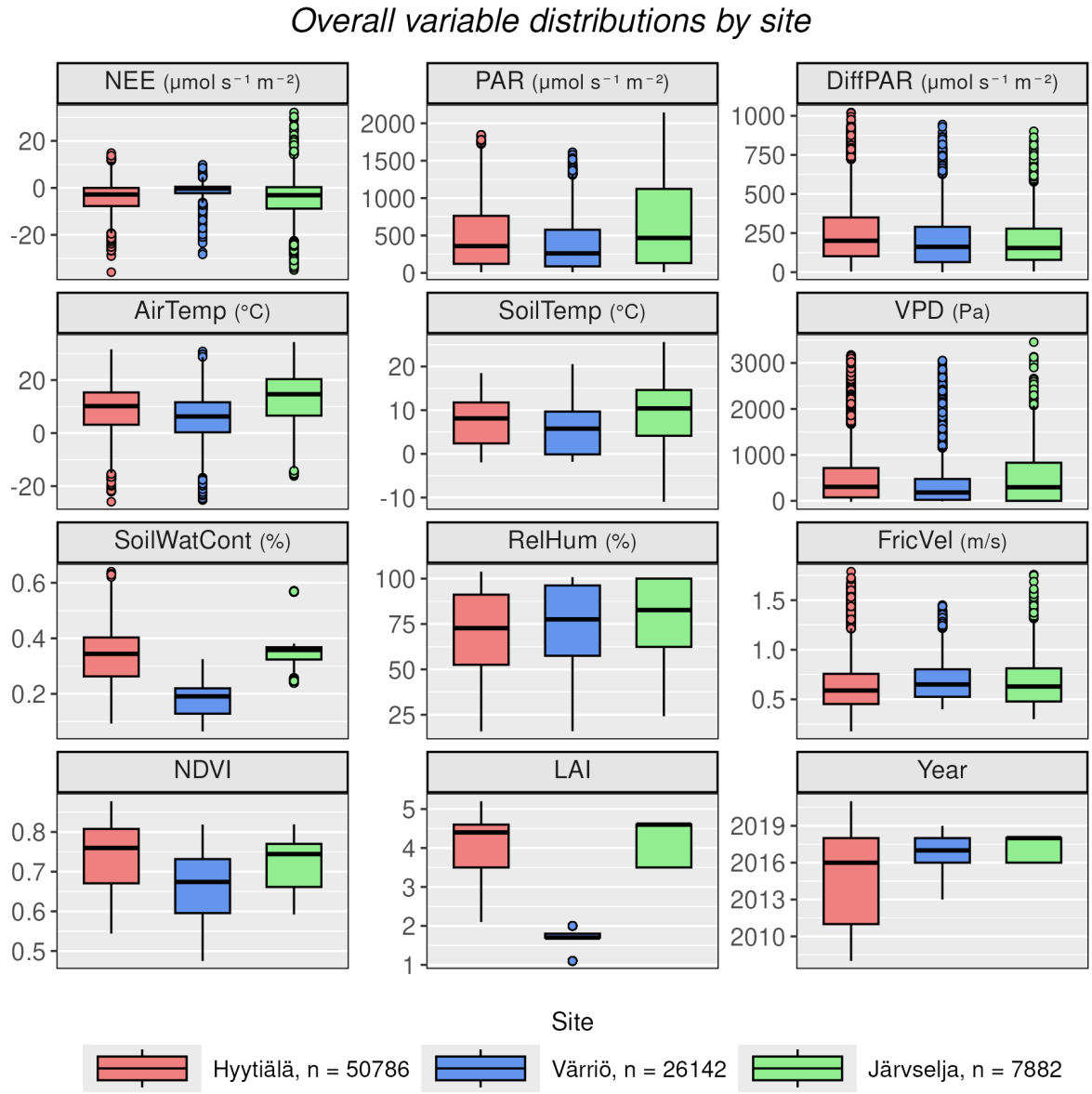

Figure A1: Boxplots showing the distribution of all input variables as well as NEE and the origin year across Värriö, Hyytiälä, and Järvselja.

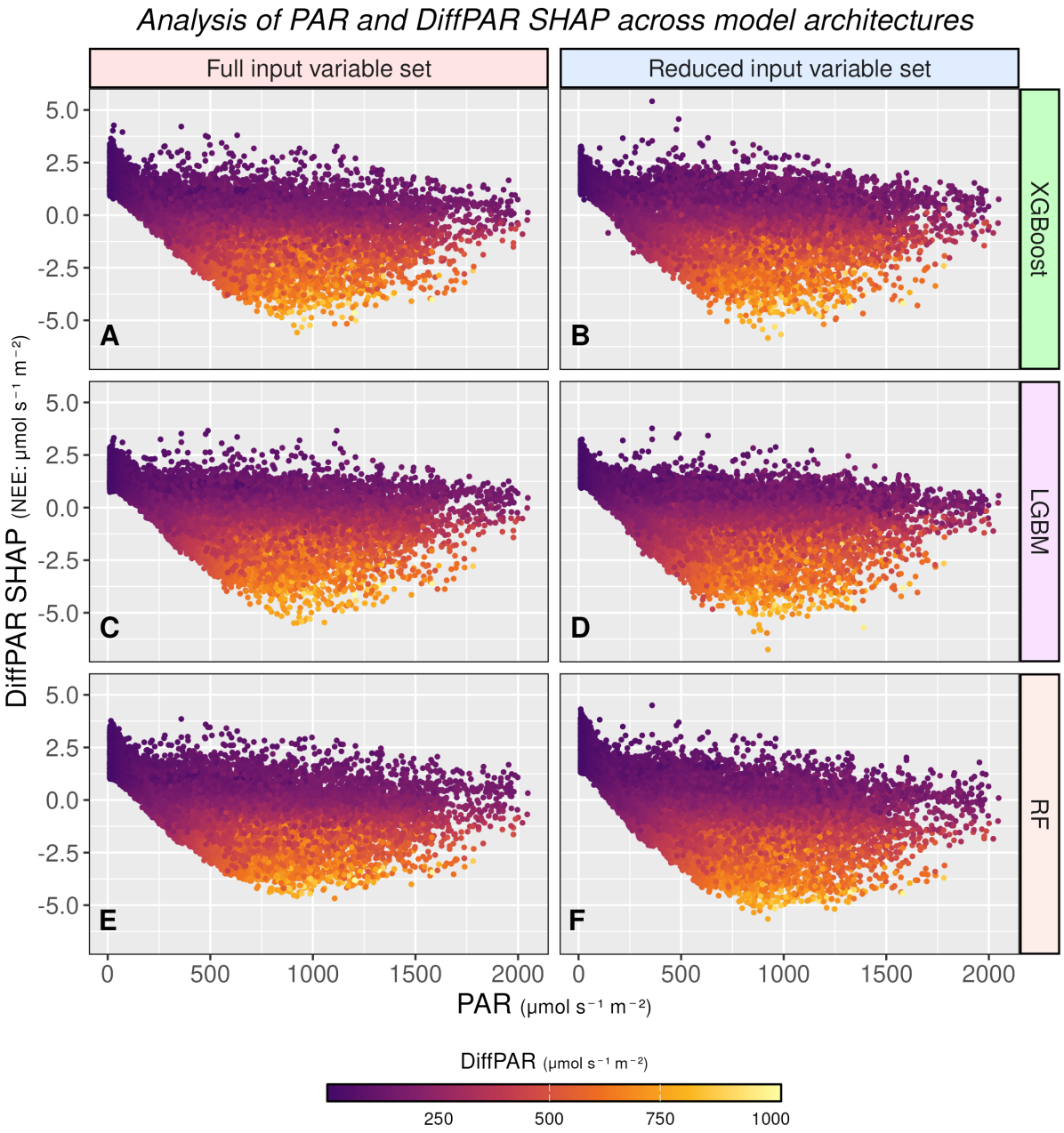

Figure A2: Scatterplots showing the relationship between PAR and DiffPAR SHAP across both input configurations, full and reduced, and all three ML models: XGBoost, Random Forest, and LGBM.

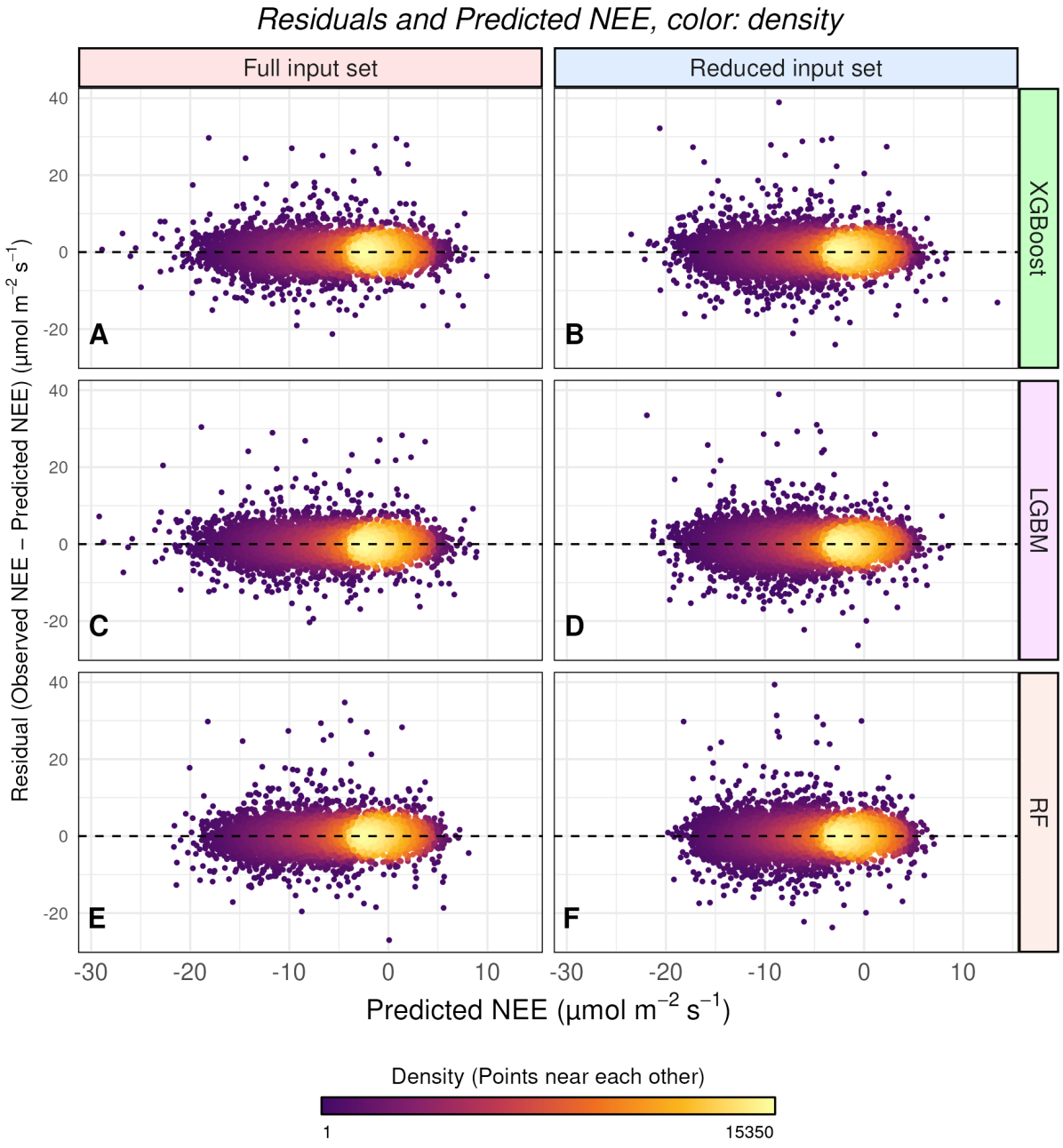

Figure A3: Residuals of the different ML models across both input configurations. A residual is the observed value minus the model-predicted value for a data point, i.e. the prediction error.

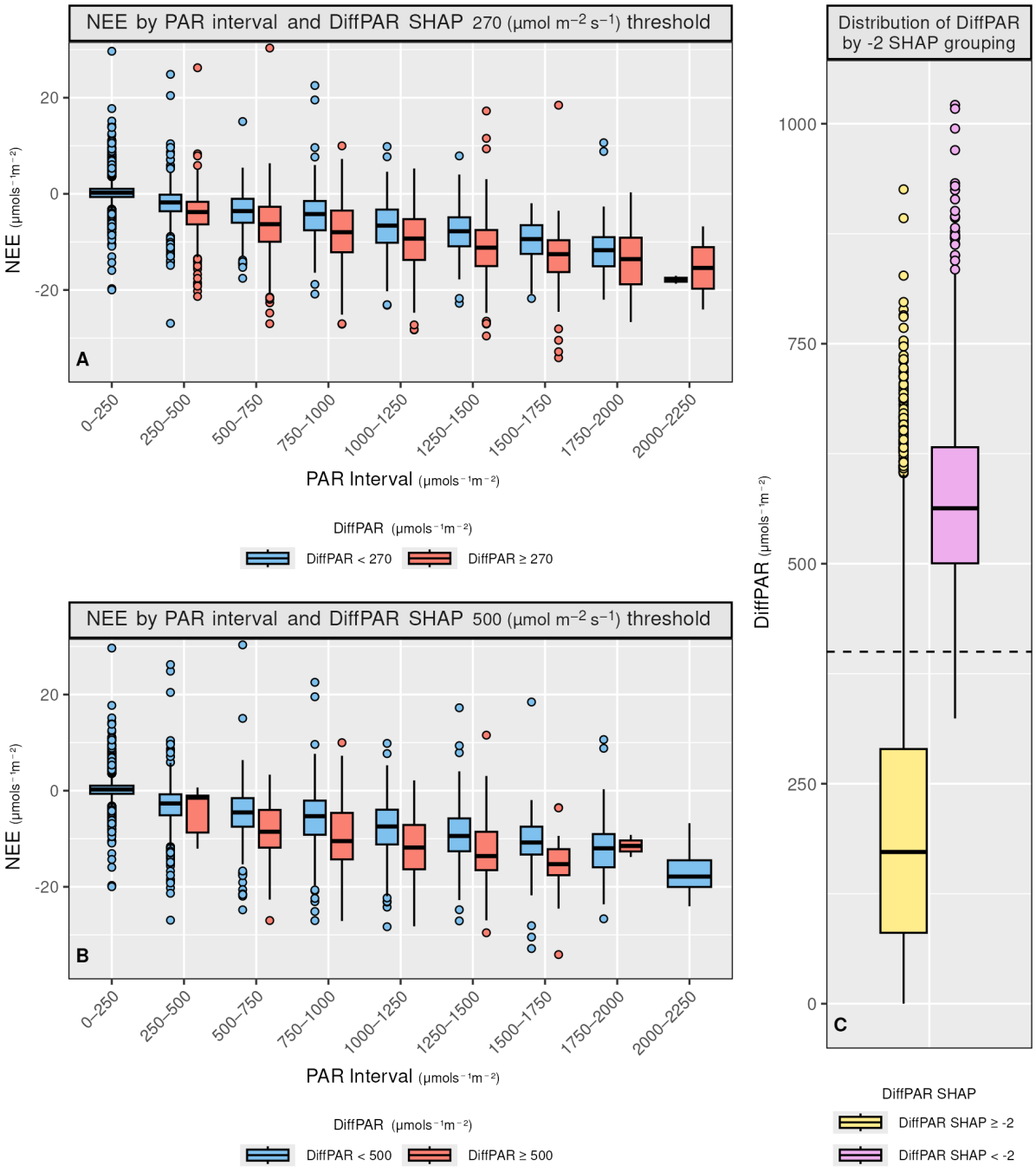

Figure A4: NEE stratified by PAR intervals using raw DiffPAR thresholds: A, 270 μmol m^{-2} s^{-1}; B, 500 μmol m^{-2} s^{-1}. DiffPAR stratified by SHAP-based grouping is shown in C.

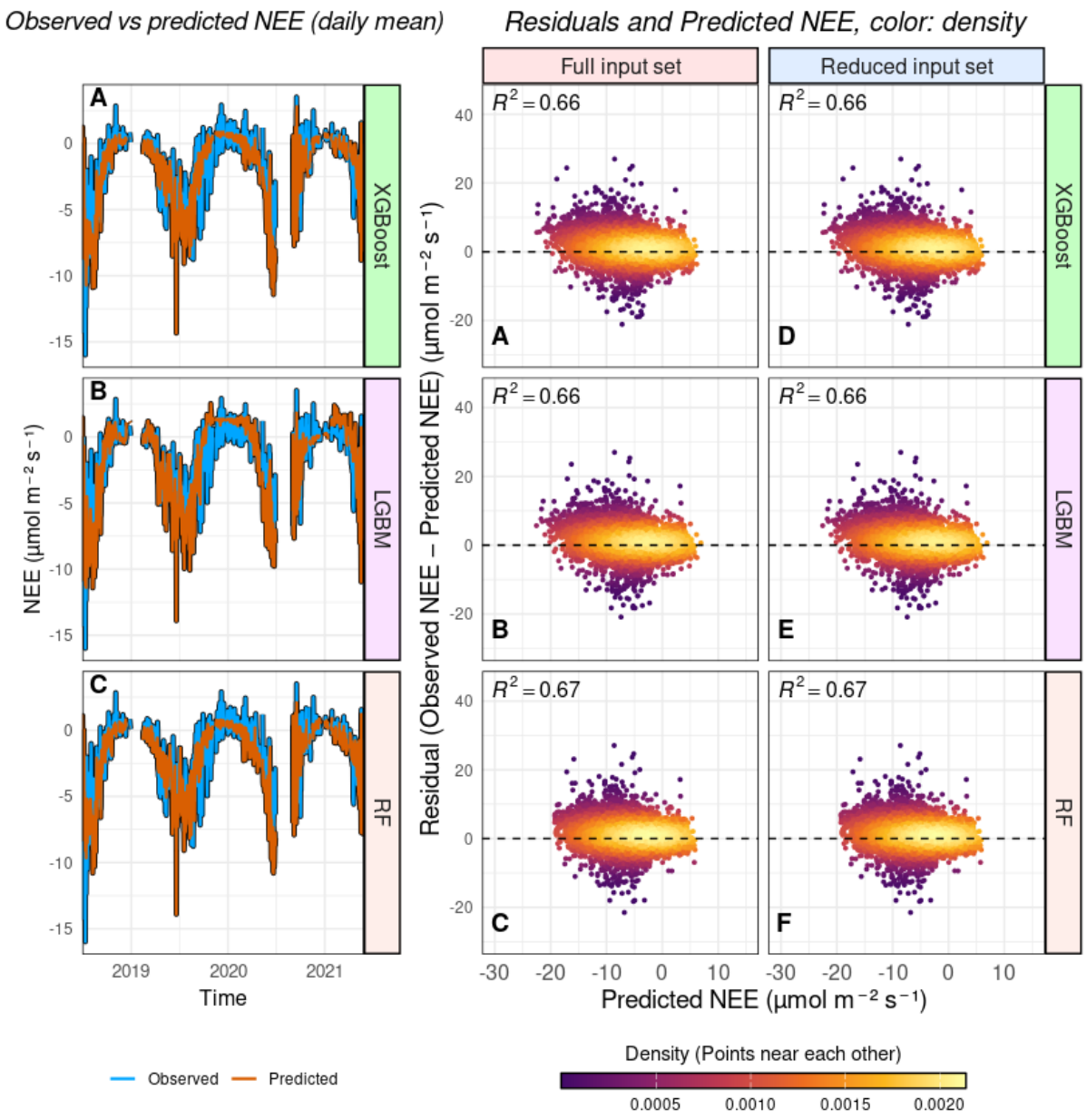

Figure A5: Observed versus predicted NEE, daily means, left, and residuals against predicted NEE, right, for the time-blocked holdout, where the earliest 75% of observations were used for training and the latest 25% for testing. Performance is lower than under the random split, but the main temporal structure is still captured.