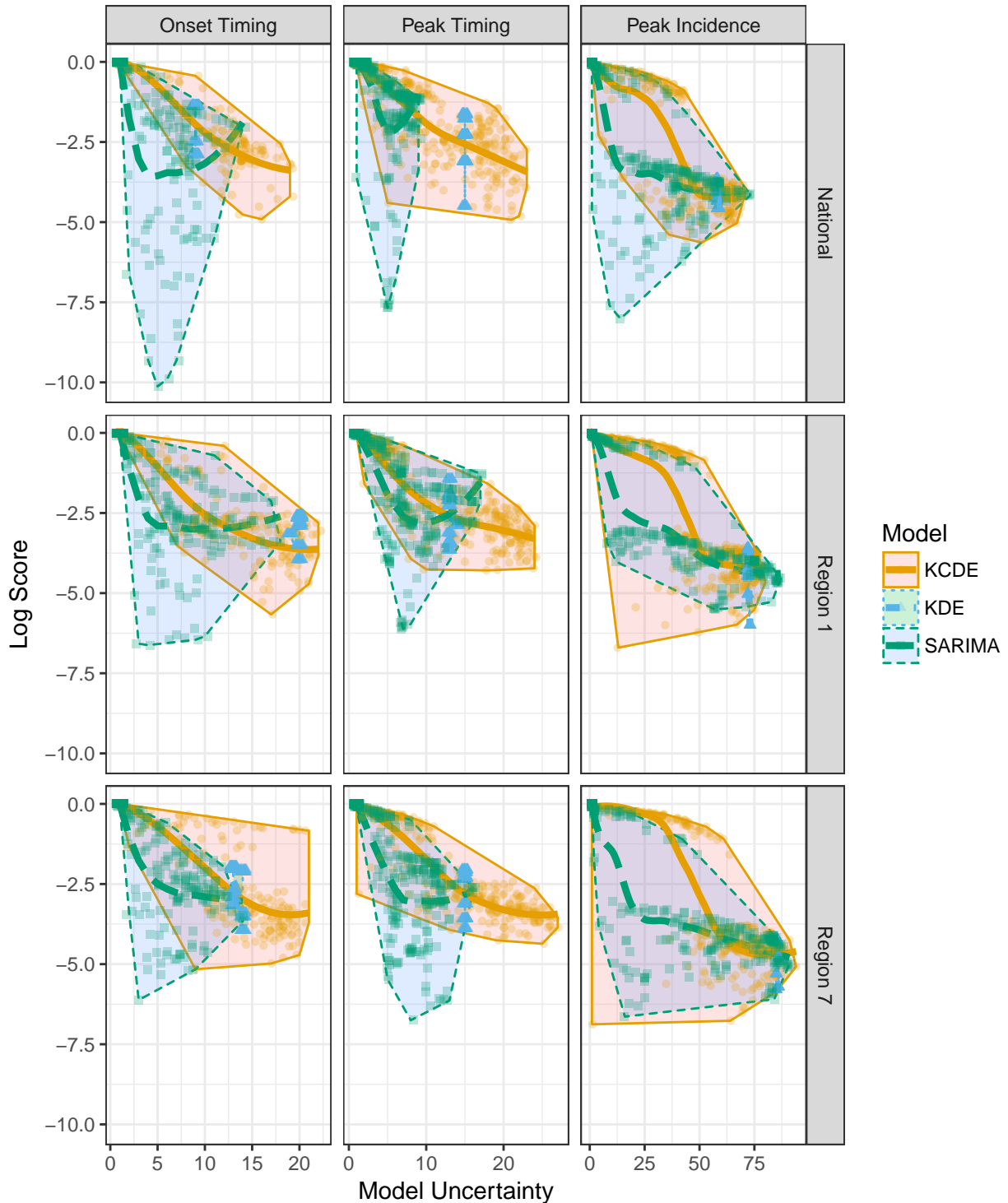


Log Scores vs. Model Uncertainty



S5 Fig. Log scores achieved by each component model vs. model uncertainty Model uncertainty is measured by the number of bins required to cover 90% of the predictive distribution. The plot summarizes results across all seasons in the training phase when all three component models produced predictions. The thick line is a smoothed estimate of mean log score at each value of model uncertainty; the shaded region indicates the convex hull of log scores achieved by each model; and the actual log scores achieved in each week are indicated with points. The KCDE and SARIMA models condition on all previously observed data within the current season, and generally have high certainty when the target event (season onset or season peak) has almost occurred or has already occurred.