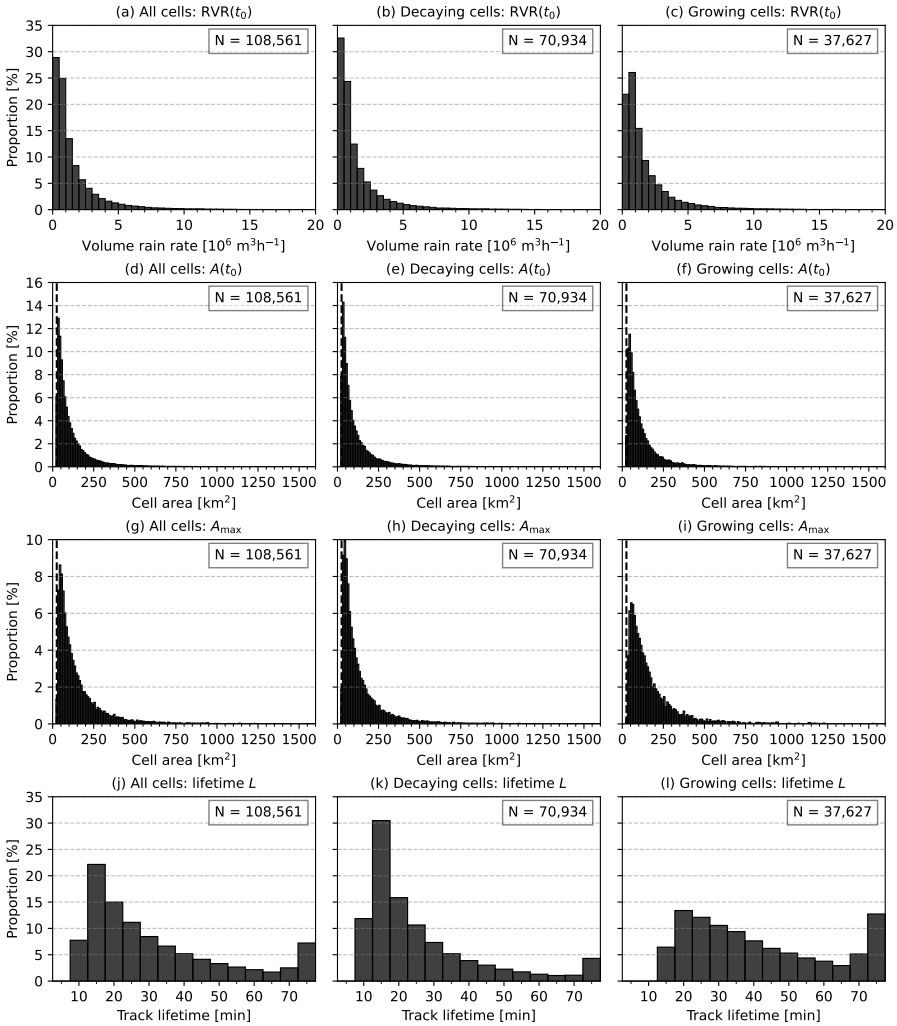

Figure S1. Same as Figure 4, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Histograms of cell and track feature values. The panels show (a-c) the volume rain rate at the nowcast creation time t_0; (d-f) cell area at t_0; (h-i) the maximum observed cell area; and (j-l) the observed cell track lifetime. Histograms are shown for all cells (panels a, d, g), decaying cells (panels b, e, h), and growing cells (panels c, f, i). The value in each panel indicates the number of cells in the histogram. In the cell area histograms (d-i), the vertical dashed line indicates the minimum cell area threshold of 25 km^2.

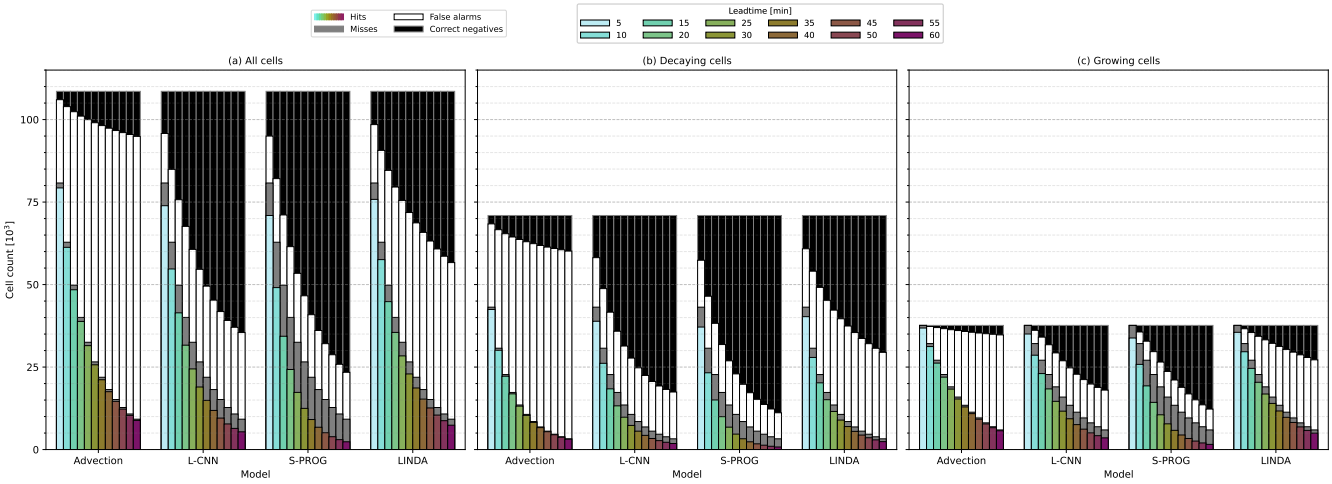

Figure S2. Same as Figure 6, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Number of convective cells used in the analysis by nowcast lead time for (a) all cells, (b) decaying cells, and (c) growing cells. Only the cells that are part of the tracks that existed at t_0 are considered. The coloured bars indicate the number of hits, i.e., cells that exist in both target observations and nowcast; the grey bars indicate misses, i.e., cells that exist in target observations but not in nowcast; the white bars indicate false alarms, i.e., cells that exist in nowcast but not in target observations; and the black bars indicate correct negatives, i.e., the number of cell tracks that existed in the input observations at t_0 and do not exist in target observations or nowcast at the given lead time.

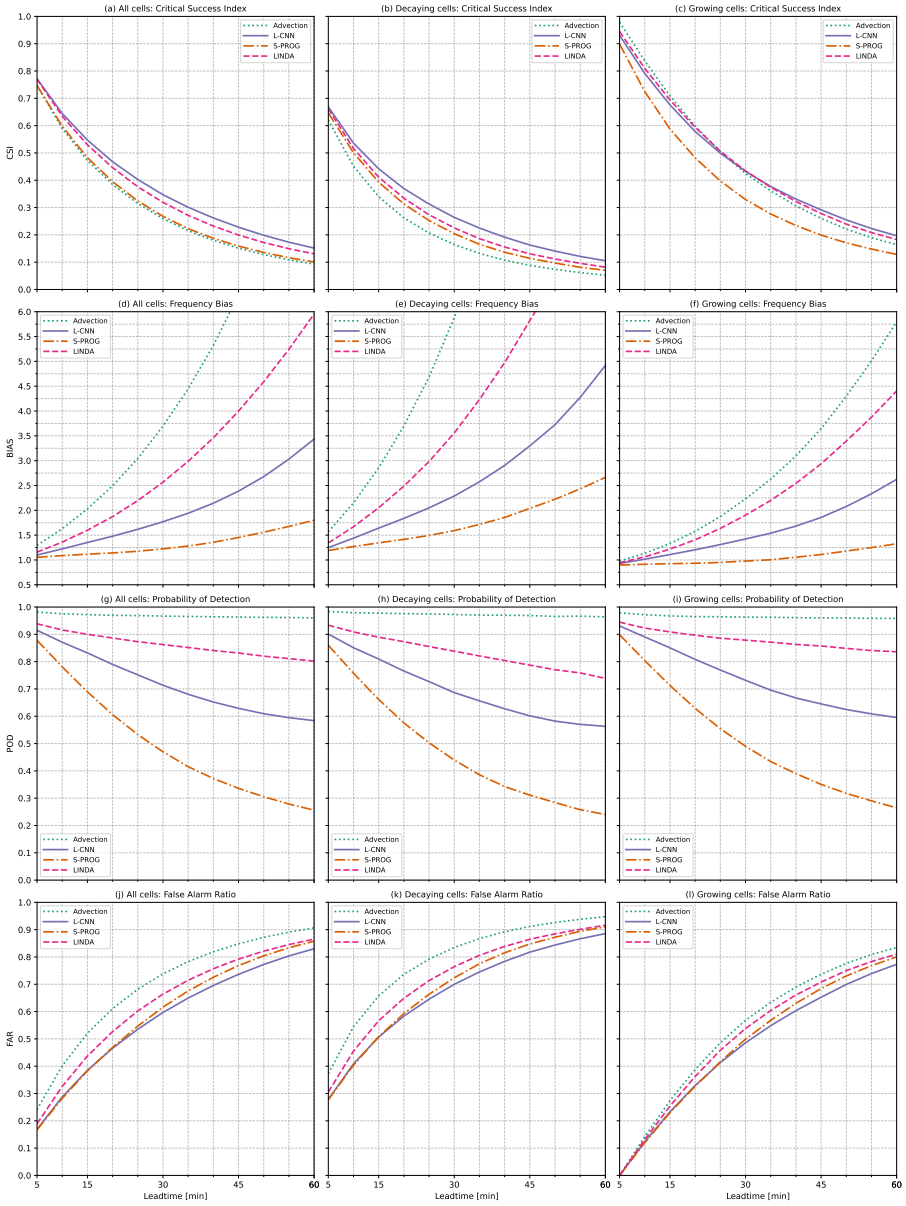

Figure S3. Same as Figure 7, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Contingency-based metrics of cell existence as a function of lead time, that is, whether a cell identified in the target observations was also identified in the nowcast. The panels show the Critical Success Index (CSI) for (a) all cell tracks, (b) decaying cell tracks, and (c) growing cell tracks; the Frequency Bias (BIAS) for (d) all cell tracks, (e) decaying cell tracks, and (f) growing cell tracks; the Probability of Detection (POD) for (g) all cell tracks, (h) decaying cell tracks, and (i) growing cell tracks; and the False Alarm Ratio (FAR) for (j) all cell tracks, (k) decaying cell tracks, and (l) growing cell tracks.

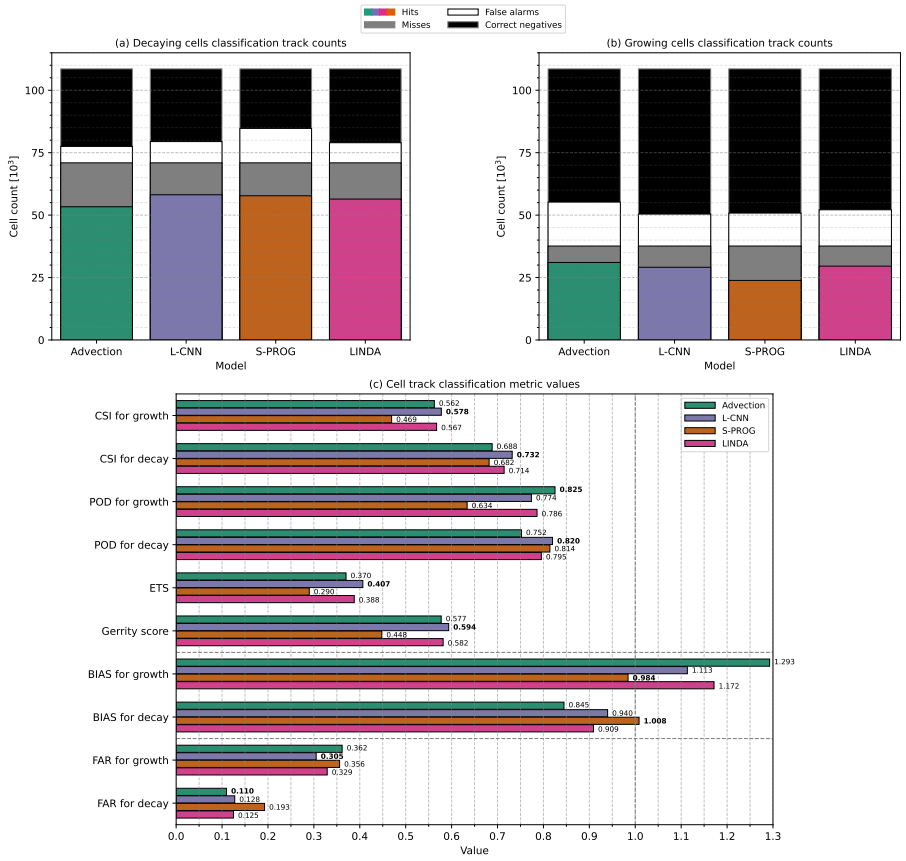

Figure S4. Same as Figure 8, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Number of hits (coloured bars), misses (grey bars), false alarms (white bars), and correct negatives (black bars) for the cell track classification into (a) decaying or (b) growing, and (c) contingency table-based metrics of the track classification into decaying or growing for the models. For the Critical Success Index (CSI), Probability of Detection (POD), False Alarm Ratio (FAR), and Frequency Bias (BIAS), the scores are calculated separately for growing and decaying cell tracks by changing the class that is considered the "true" class. For the Equitable Threat Score (ETS) the score is symmetric, and for the Gerrity score (GS), the multicategory version of the score is used; therefore, only one value is provided for both. The best model for each score is marked in the bolded value. For BIAS, the value closest to one, and for FAR, the lowest value are considered best, while for other scores the highest value is the best.

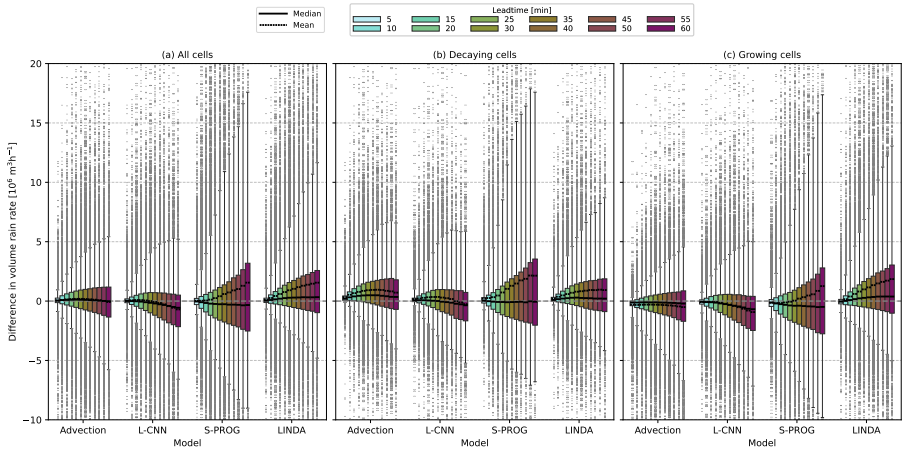

Figure S5. Same as Figure 9, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Box plots of differences between predicted and observed cell volume rain rates by nowcast lead time for (a) all cells, (b) decaying cells and (c) growing cells. The boxes show the 25th to 75th percentile range, and the whiskers the 5th to 95th percentile range. The solid line indicates the median and the dotted line the mean, and outliers are indicated by dots. A positive difference indicates overestimation of the volume rain rate by the model, and a negative difference underestimation.

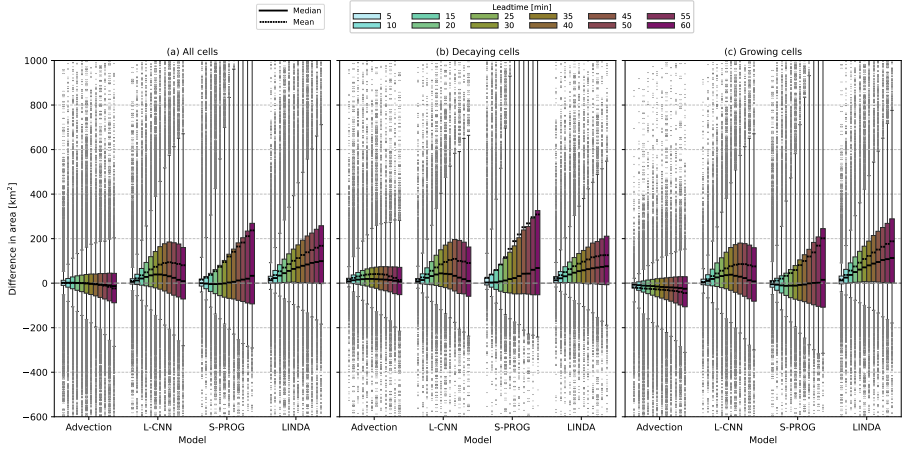

Figure S6. Same as Figure 10, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Box plots of differences between predicted and observed cell areas by nowcast lead time for (a) all cells, (b) decaying cells and (c) growing cells. The boxes show the 25th to 75th percentile range, and the whiskers the 5th to 95th percentile range. The solid line indicates the median and the dotted line the mean, and outliers are indicated by dots. A positive difference indicates overestimation of the cell area by the model, and a negative difference underestimation.

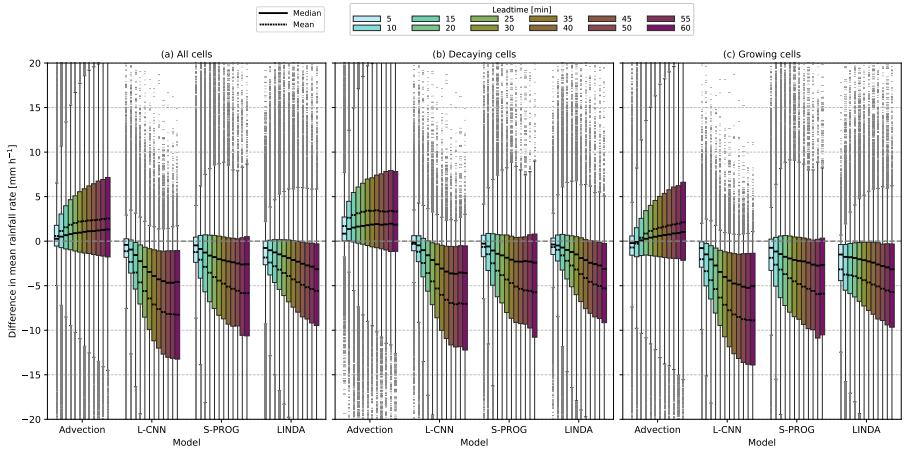

Figure S7. Same as Figure 11, but for a dataset where cell tracks with splits and merges have been removed from observed cell tracks. Box plots of differences between predicted and observed mean rain rate inside the cells by nowcast lead time for (a) all cells, (b) decaying cells and (c) growing cells. The boxes show the 25th to 75th percentile range, and the whiskers the 5th to 95th percentile range. The solid line indicates the median and the dotted line the mean, and outliers are indicated by dots. A positive difference indicates overestimation of the mean rain rate by the model, and a negative difference underestimation.