medRxiv preprint doi: https://doi.org/10.1101/2024.08.30.24312858; this version posted August 31, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Supplemental Tables and Figures



Fig. S1. Map of California counties by region. Starred regions are considered "high incidence"; counties within these regions were modelled independently.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.30.24312858; this version posted August 31, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



Fig. S2. To generate 90% prediction intervals, we used a two-step bootstrapping process. In step 1, we sampled census tracts (n = 8,057) with replacement to generate 500 datasets, fit the models to each dataset (n = 1-500), and forecasted cases using each model. In step 2, we then resampled each estimate from a quasi-Poisson distribution with the mean set to each bootstrapped prediction and calculated the 5th and 95th percentile of the resulting distribution to obtain the 90% prediction interval. Figure created with BioRender.

medRxiv preprint doi: https://doi.org/10.1101/2024.08.30.24312858; this version posted August 31, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.

Table S1. Variables included in the five models included in the ensemble. Green cells were used as main effects and purple cells were used as interactions. Models 1-4 were generalized linear models (GLMs) and model 5 was a random forest (RF).

Variable	GLM 1	GLM 2	GLM 3	GLM 4	RF
Population (as an offset)					
Year (as a natural spline)					
Season (as a factor)					
Percent sand					
Impervious surface					
Elevation					
Total rainfall					
Lag 1 month					
Lag 3 month					
Lag 6 month					
Lag 9 month					
Lag 12 month					
Lag 15 month					
Lag 18 month					
Lag 21 month					
Lag 24 month					
Lag 27 month					
Lag 30 month					
Lag 33 month					
Lag 36 month					
Average Temperature					
Lag 1 month					
Lag 3 month					
Lag 6 month					
Lag 9 month					
Lag 12 month					
Lag 15 month					
Lag 18 month					
Lag 21 month					
Lag 24 month					
Lag 27 month					
Lag 30 month					
Lag 33 month					
Lag 36 month					
One year post drought (indicator)					
Two years post drought (indicator)					