ATMOSPHERIC ENVIRONMENT: X 16 (2022) 100195

Contents lists available at ScienceDirect

Atmospheric Environment: X

journal homepage: www.journals.elsevier.com/atmospheric-environment-x

Characterizing changes in extreme ozone levels under 2050s climate conditions: An extreme-value analysis in California

Bradley Wilson [a,*], Mariah Pope [a], Jeremy R. Porter [a,b], Edward Kearns [a], Evelyn Shu [a], Mark Bauer [a], Neil Freeman [a], Mike Amodeo [a], David Melecio-Vazquez [a], Ho Hsieh [a], Maggie Tarasovitch [a]

[a] First Street Foundation, 215 Plymouth St, Brooklyn, NY, 11201, United States
[b] Quantitative Methods in the Social Sciences, City University of New York's Graduate Center, 365 5th Ave., New York, NY, 10016, United States

ARTICLE INFO

Keywords:
Air pollution
Ozone
Climate penalty
Extreme-value theory
CMAQ

ABSTRACT

Ground-level ozone has a well established climate penalty associated with changing meteorological conditions under future climate change, but most existing research focuses exclusively on ozone-temperature relationships defined at specific locations. Using an extreme-value theory approach, we model relationships between temperature, humidity, and vapor pressure deficit and maximum daily 8 h ozone concentrations (MDA8) across California using a combination of ozone station network data and Community Multiscale Air Quality Modeling System (CMAQ) model output from 2008 to 2017. We use a spatial regression with time varying bias coefficients to fuse station observations and modeled output into a spatially explicit gridded ozone dataset at 12 km resolution, then fit independent Point Process models to every grid cell to estimate 5, 20, and 50 year ozone return-levels across California. We evaluate the impact of climate change on ozone return levels and episode days under RCP4.5 conditions using an ensemble of 18 downscaled CMIP5 climate models, finding changes of -6 to 8 ppb in effective return levels and −8 to 16 episode days, Our results further strengthen the evidence for anozone climate penalty in California and show the advantages of considering multiple climate variables when modeling ground-level ozone.

1. Introduction

Ground-level ozone (hereafter, ozone) pollution in the United States has improved over recent decades on average due to the reduction of anthropogenic emissions (US EPA, 2016). However, as of 2015, one in three Americans were still exposed to ozone levels that exceed National Ambient Air Quality Standards (NAAQS) (Nolte et al., 2018). Health implications associated with ozone exposure are well documented through known associations with both long and short-term adverse health effects, such as chest pain, coughing, asthma, respiratory diseases and infections, and premature deaths (Bell et al., 2007; Fann et al., 2015, 2021; Garcia-Menendez et al., 2015; Kinney, 2008; Orru et al., 2017; Silva et al., 2017; Tagaris et al., 2009). At current levels of anthropogenic emissions of ozone precursors, millions of Americans will continue to be exposed to harmful concentrations of ozone into the near future (Fann et al., 2016; Fann et al., 2021). Furthermore, research indicates that air quality gains associated with reduced anthropogenic emissions are likely offset by a 'climate penalty', or impacts associated with changing meteorological conditions and subsequent biogenic emissions that are conducive to ozone formation (Shi et al., 2019; Rasmussen et al., 2013; Silva et al., 2017; Stowell et al., 2017; Coates et al., 2016).

There are multiple pathways by which ozone levels shift under various climate scenarios. Ozone formation is driven by the interaction of precursor emissions, such as nitrogen oxides (NOx) and volatile organic compounds (VOCs), with sunlight, but is also affected by additional meteorological factors like temperature, humidity, wind patterns, vapor pressure deficit, mixing height, and cloud cover, among others. Among those factors, high temperatures, low humidity, and vapor pressure deficit are all well documented in regards to their promotion of increasing levels of ozone (Kavassalis and Murphy, 2017; Mahmud et al.,

* Corresponding author.
E-mail addresses: bradley@firststreet.org (B. Wilson), mariah@firststreet.org (M. Pope), jporter@gc.cuny.edu (J.R. Porter), ed@firststreet.org (E. Kearns), evelyn@firststreet.org (E. Shu), mark@firststreet.org (M. Bauer), neil@firststreet.org (N. Freeman), mamodeo@firststreet.org (M. Amodeo), dmeleciovazquez@firststreet.org (D. Melecio-Vazquez), ho@firststreet.org (H. Hsieh), maggie@firststreet.org (M. Tarasovitch).

https://doi.org/10.1016/j.aeaoa.2022.100195
Received 4 August 2022; Received in revised form 1 November 2022; Accepted 4 November 2022
Available online 11 November 2022

2008; Shen et al., 2016; Wells et al., 2021). Specifically, high temperatures increase the rate at which ozone is formed and can cause local vegetation to release more ozone precursors, such as isoprene (Nolte et al., 2018b), low levels of humidity can extend the lifetime of ozone (while high humidity promotes ozone depletion), and vapor pressure deficit directly modulates the uptake of ozone by trees through the opening and closing of stomata (Arnold et al., 2018; Kavassalis and Murphy, 2017). Additionally, ozone formation depends on the local VOC-to-NOx ratio. It is typically assumed that lowering NOx and/or VOC's will decrease local ozone concentrations; however, this is not always true due to the complex chemistry of the VOC-to-NOx ratio. For example, lowering NOx levels while VOC levels remain constant can actually increase peak ozone concentrations. While these factors create conditions in which ozone is able to develop and persist, shifts in natural or anthropogenic emission of ozone precursors (Fann et al., 2016, Kalashnikov et al. 2022) may also contribute to future ozone levels, but are not a primary focus of this study.

Meteorological parameters are used as key components within chemical transport and climate-chemistry models, which are frequently used to complement observations from air quality monitoring stations (Davis et al., 2011; Doherty et al., 2013; Fann et al., 2021; Nolte et al., 2021; Torres-Vazquez et al., 2022). Some models like the U.S. EPA Community Multiscale Air Quality Modeling System (CMAQ) can even be run in a 'coupled' mode with common meteorological models (e.g. Weather Research and Forecasting—WRF-CMAQ). Results from these studies using climate-chemistry models generally find the net effect of climate change on annual ozone levels in 2050 in the range of 1–5 ppb (e.g. Wu et al., 2008; Stowell et al., 2017; Bell et al., 2007), with some variation between regions. However, climate-chemistry models are somewhat less effective at capturing ozone variability and may underestimate local ozone maxima, with uncertainties arising from multiple sources including meteorological inputs and boundary condition specification, emissions inventories, and numerical representations of the underlying physical and chemical processes (Doherty et al., 2013; Parrish et al., 2014; Torres-Vazquez et al., 2022). Statistical modeling of ozone concentrations or 'ozone episodes' (days above high thresholds) are common alternative approaches and also integrate meteorological data to support predictions (Thompson et al., 2001). Past results show increases of up to tens of additional ozone episode days/year depending on region and climate scenario (Mahmud et al., 2008; Shen et al., 2016). However, statistical analyses are typically constrained to specific monitoring station locations where high quality time-series measurements are available. Understanding how to integrate different modeling approaches to characterize future ozone concentrations is an active area of research (e.g. Berrocal et al., 2012; Davis et al., 2011; Zhang et al., 2018; Turnock et al., 2020).

To these ends, in this study we build upon previous research in two areas. We use an extreme value theory (EVT) (Coles, 2001) model to investigate the 5-year, 20-year, and 50-year effective return levels of ozone concentrations and the number of days per year that ozone exceeds 71 ppb presently and in the 2050's. 71 ppb corresponds to an 'unhealthy for sensitive groups' day in the EPA Air Quality Index and is typically the level at which behavioral modifications start to be recommended. Compared to other EVT research (Gouldsbrough et al., 2022; Rieder et al., 2010; Shen et al., 2016), we evaluate a larger set of meteorological covariates to determine whether ozone tails have statistical relationships with more factors than high temperatures, and how climate-related shifts across multiple parameters affect future ozone predictions. Additionally, our modeling approach fuses estimated ozone concentration data from CMAQ with observations from monitoring stations. Combining multiple complementary data sources facilitates characterization of the ozone climate penalty with high spatial resolution.

2. Materials and methods

We focus our analysis on the state of California, which has the highest ozone nonattainment rate in the United States (US EPA, 2022). There are currently 37 out of 58 counties in California where at least a portion of the county exceeds 2015 National Ambient Air Quality Standards (NAAQS) for 8-h ozone. This equates to approximately 34.6 million people (93% of the population) per the 2010 population (US EPA, 2022). California also has a dense network of ozone monitoring stations and varying climate regimes, making it an ideal location to analyze climate-related effects on ozone. For this study, we source ozone concentration data from two sources: monitoring stations available from the EPA Air Quality System (AQS) data program and model results from CMAQ.

2.1. Developing gridded ozone concentration estimates

The EPA AQS data program collects air quality monitoring data for six common air pollutants at hundreds of stations across the United States. We focus on ozone observations calculated under the Ozone 8-h (MDA8) 2015 NAAQS. MDA8 values represent the highest 8-h mean ozone concentration measured over the course of a dayFor this study, we acquired daily MDA8 ozone concentration data at all AQS monitors in California for the years of 2008–2017 from the AQS data portal. We filtered the data range to the months of May–September; the months that constitute the 'ozone season', when ozone episodes are most frequent (Bloomer et al., 2009). We exclude monitoring stations that are missing more than 10% of data across the ten-year period, leaving 136 out of a total 226 stations. The average MDA8 ozone concentrations across all stations are shown in Fig. 1.

To achieve complete spatial coverage over California, we also sourceCMAQ model results for ozone concentrations from EQUATES (EPA's Air QUAlity TimE Series Project). CMAQ is a state-of-the-art modeling platform that combines current knowledge of atmospheric science and air quality modeling to track the movement of pollution over time and output volume-average pollutant concentrations across spatially continuous grid cells and vertical layers (US EPA, 2021). All EQUATES data were computed using CMAQ version 5.3.2 and WRF version 4.1.1. The CMAQ ozone daily average surface concentrations datasets are formatted to the same MDA8 standards for calculating 8-h average ozone concentrations on a grid with 12 km horizontal resolution over the surface layer of the contiguous US. To match the AQS data records, we downloaded daily CMAQ data between 2008 and 2017 for the May–September months and filtered it to the state of California using a US Census boundaries shapefile. The distribution of EPA stations overlaid on CMAQ grid cells is shown in Fig. 2.

Comparing CMAQ output to corresponding ozone measurements from AQS stations shows that ozone concentrations measured at AQS stations are typically higher than CMAQ model output (Fig. 3a). The mean difference among all stations was 7.3 ppb while the maximum difference and minimum difference were 20.3 and −4.7 ppb, respectively. These results are consistent with the EPA incremental evaluation of the CMAQ system (Appel et al., 2021), and motivation for data fusion efforts (e.g. Berrocal et al., 2012; Chen et al., 2014; Requia et al., 2020) that aim to combine station observations and climate-chemistry model output into a single consistent data layer. We address these differences by fusing both data sources into a single unified set of daily gridded ozone estimates using a geostatistical model that calibrates CMAQ modeled estimates to AQS station data. Our approach is conceptually similar to Berrocal et al. (2012) in using a spatial regression with time varying bias coefficients to relate model output to air quality observations. For each daily time step (t = 1:153), we estimated observed ozone concentrations using CMAQ concentrations of the corresponding grid cell and a spatially varying random field modeled with a Gaussian Random Field (GRF) Process (Equation (1)):

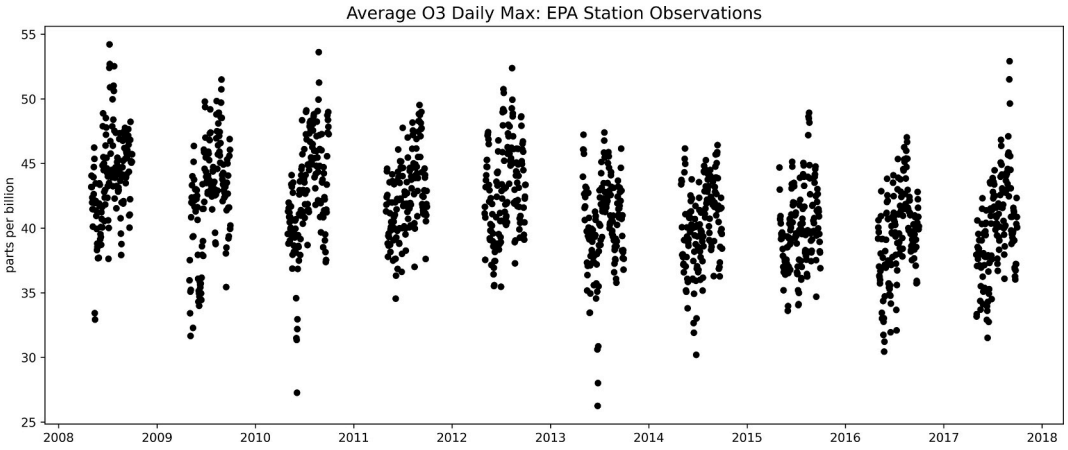

Fig. 1. Average MDA8 ozone concentrations across all 136 California stations between 2008 and 2017.

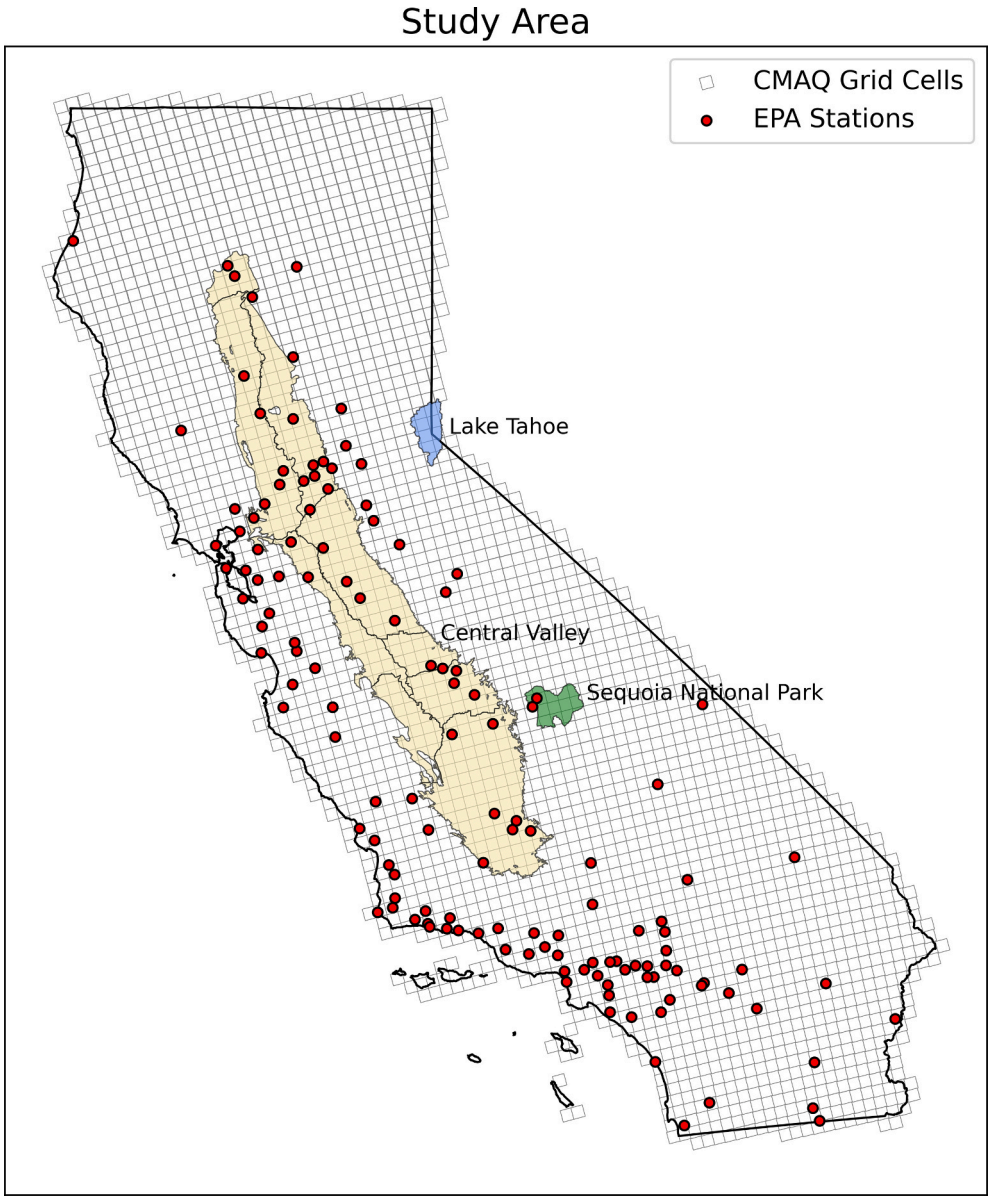

Fig. 2. Location of ozone monitoring stations and CMAQ grid cells across California including referenced geographical names.

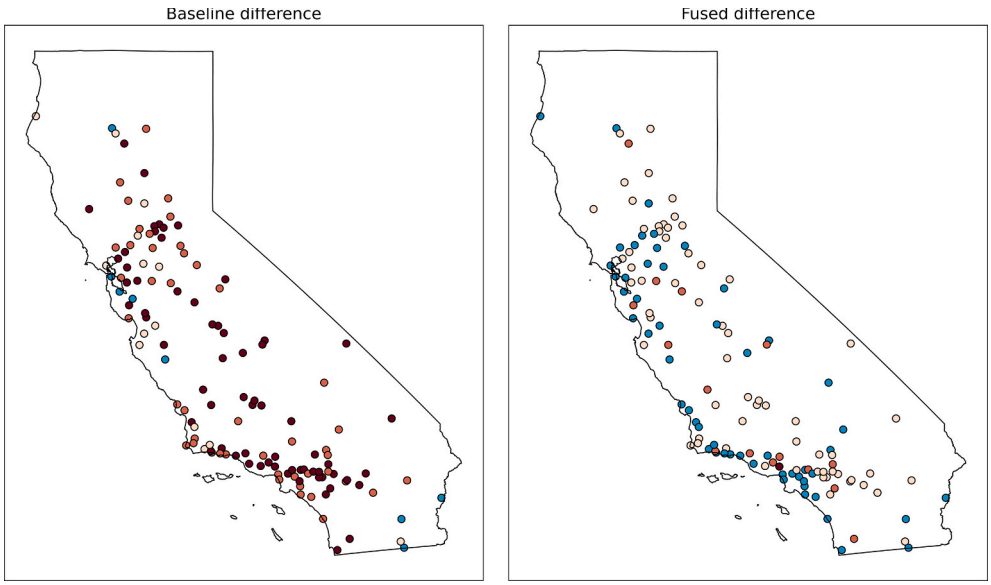

Fig. 3. Percent Difference between AQS station data and modeled output for the 2008–2017 time period, May-September.

$$Y_t(s) = B_{0(t)} + B_{1(t)}(x_t) + \varepsilon \quad (1)$$

Where $Y_t(s)$ is the observed concentration at AQS station, $B_{0(t)}$ is the value of the GRF, $B_{1(t)}(x_t)$ is the regression coefficient for CMAQ concentrations, and ε is the model error term—all varying with day (t).

We fit the model using the INLA-SPDE approach in R-INLA (Rue et al., 2009; Lindgren et al., 2011), using all CMAQ grid cells that contain AQS data stations. The SPDE approach uses a triangular mesh defined across the study region to model the GRF. To help inform estimation of spatial effect on the edges of the study area, we add observations from the closest AQS stations in neighboring states. After model fitting, we extract the spatial field estimates on the full CMAQ grid and use the coefficients from the fitted model to predict the ozone concentrations for every day in the complete time series. To assess model fit, we recalculated the mean differences between station observations and fused grid cell predictions and compared those to the differences calculated against the original CMAQ data (Fig. 3). We find the fused predictions are significantly closer to the AQS observations and less biased towards overprediction.

2.2. Covariate data

To match the daily time-series of MDA8 ozone concentrations, we acquired a consistent set of daily meteorological covariates for each CMAQ grid cell using Google Earth Engine. Although a small number of AQS ozone stations also record in-situ meteorological data and CMAQ uses WRF for meteorological inputs, we used an external source of gridded meteorological data to facilitate consistent comparisons across grid cells under both current and future conditions. For historical observations, we used daily covariate data from the University of Idaho Gridded Surface Meteorological Dataset (GridMET), which have a native spatial resolution of approximately 4 km (Abatzoglou, 2013). GridMET uses a climatically aided interpolation process to blend several gridded climate data and regional reanalysis datasets into a spatially and temporally complete set of surface meteorological variables (Abatzoglou, 2013). We directly downloaded maximum temperature and minimum humidity data and derived vapor pressure deficit using equations described in (Allen and Food and Agriculture Organization of the United Nations, 1998). All three covariates have established relationships with ozone formation. Daily observations were resampled onto the 12 km CMAQ grid using a weighted average of each covariate value.

To construct time series representative of future climate conditions, we first calculated change factors for each variable using data from 18 downscaled global climate models (GCMs) in the Multivariate Adaptive Constructed Analogs Applied to Global Climate Models (MACAv2: METDATA) (Abatzoglou and Brown, 2012) dataset. MACAv2 was validated across the Western US and has been shown to provide more accurate predictions in complex terrain compared to other downscaling methods (Abatzoglou and Brown, 2012). The MACAv2 dataset is also trained against GridMET data and evaluated on the same 4 km grid, which provides consistency in our covariate data sources. For this study, we used data from the Representative Concentration Pathway (RCP) 4.5 scenario to present 'middle of the road' estimates.

For each grid cell in the MACAv2 dataset, we calculated change factors for each variable containing the percentage difference in the mean ozone season temperature, minimum humidity, or vapor pressure deficit between the 2008–2017 and 2048–2057 time periods. Separate change factors were calculated for each of the 18 GCMs as well as an ensemble mean and 10th and 90th percentile values. We then used the change factors to scale the GridMET time series into multiple 2050s time series that account for the overall trend in each variable. Daily covariate values were scaled linearly by the change factor value for each respective variable, effectively creating mean-shifted versions of the present day time series. Because the EVT approach used in this study ends up averaging the output metrics across all yearly covariate values, the change factor approach is an efficient way to assess shifts in climate. The ensemble mean change factors for each of the three variables are shown in Fig. 4. Temperature and vapor pressure deficits are all positive and show the largest increases in the northeast and the smallest increases in the south. Humidity shows the opposite trend, with reductions in the northeast and increases in the south. Equivalent plots for the 10th and 90th percentiles are included in the supplementary information.

2.3. Extreme value theory model

To model extreme ozone levels across California, we used an EVT approach that focuses on modeling the tail behavior of ozone probability distributions (Coles, 2001). EVT has a long history of application in ozone-related research and has been shown to be an appropriate distribution for modeling ozone tail behavior (Rieder et al., 2010; Shen et al., 2016; Smith, 1989; Thompson et al., 2001). EVT approaches allow for flexible model specifications, including a non-stationary Point Process (PP) specification that allows for incorporating covariates on the model parameters. This is an important model consideration in light of

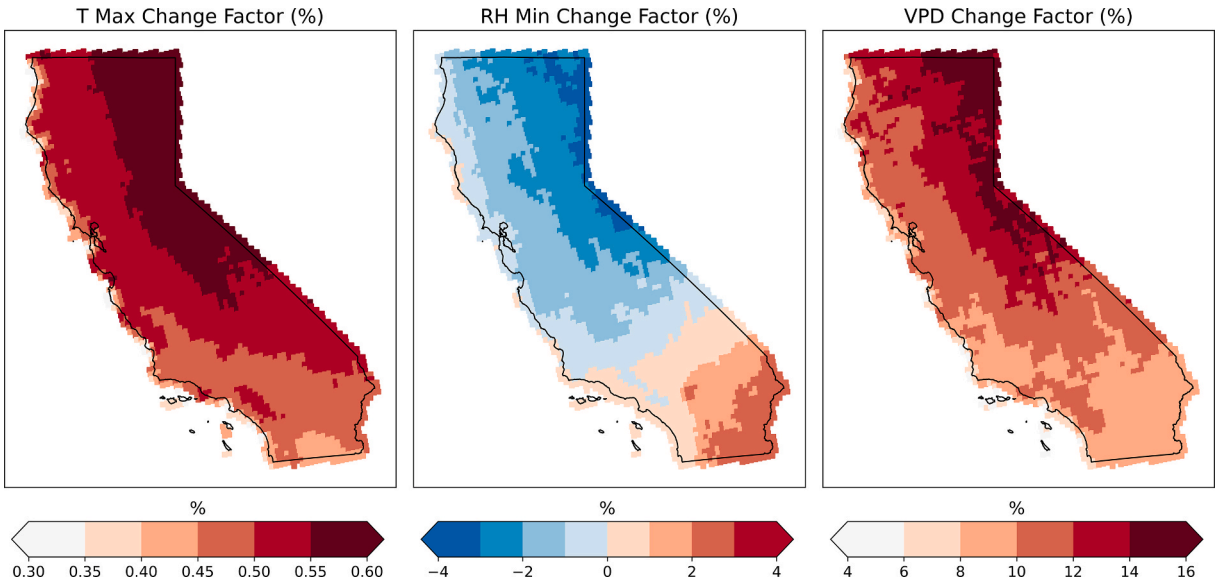

Fig. 4. Change factors (%) between current (2008–2017) and future (2048–2057) mean climate conditions for maximum temperature, minimum relative humidity, and vapor pressure deficit between May and September. .

previous research showing a strong non-stationarity between meteorological covariates and ozone concentrations (Wells et al., 2021).

Using the PP formulation from Gilleland and Katz (2016), a Poisson point process for intensity measure Λ on the set $A = (t_1, t_2) \times (x, \infty)$ is given by:

$$\Lambda(A) = (t_2 - t_1) \cdot \left[1 + \xi \frac{x - \mu}{\sigma}\right]^{\frac{-1}{\xi}} . I(1 + \xi \cdot (x - \mu) / \sigma \geq 0) \tag{2}$$

Where μ, σ, ξ are location, scale, and shape parameters, and I is an indicator function that is zero or one based on whether the specified condition is met. The time frame $[t_1, t_2]$ as [0,1] and parametrize the "season" length as 153 days, or the time period lasting from May 1st to September 31st. Conditional on the scale parameter being greater than zero, the Poisson rate parameter λ defines the frequency of exceeding a given threshold μ:

$$\lambda = \left[1 + \xi \frac{x - \mu}{\sigma}\right]^{\frac{-1}{\xi}} \tag{3}$$

We used a maximum likelihood approach to fit the model parameters. For a high threshold u, the PP log likelihood is given by:

$$l(\mu, \sigma, \xi; x_1, ..., x_n) = -k \ln \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{n} \left[1 + \frac{\xi}{\sigma}(x_i - u)\right]_{x_i > u} - n_y \left(1 + \frac{\xi}{\sigma}(u - \mu)\right)^{\frac{-1}{\xi}} \tag{4}$$

Where n_y is the number of years of data and $[.]_{x_{i>\mu}}$ is an indicator function. Following (Shen et al., 2016), we selected a threshold equal to the 90th percentile value of ozone concentrations at each station or grid cell, respectively. For a few grid cells with a 90th percentile value above 71 ppb, we fixed the threshold at 71 ppb to ensure consistent prediction for ozone episode days. To model non-stationary behavior given a covariate x, we specified linear predictors of the form $y(x) = y_0 + y_1(x)$ on the location (μ) and scale (σ) parameters, or both. We log transformed the scale parameter to ensure it remains positive valued.

We used a multi-stage process to identify the multiple candidate models for each grid cell, adapted from the generalized EVT workflow described in Towler et al. (2020). All model fitting was completed using the R package extRemes version 2.1-2 (Gilleland and Katz, 2016). Treating each grid cell independently, we first declustered the ozone time series using a runs based declustering algorithm (Coles, 2001) built into the extRemes package. This step helps satisfy the Poisson assumption for the Point Process model by replacing clusters of values above a threshold with a single value. We then iterated through three sets of candidate models, each nested within the previous set to facilitate goodness of fit comparisons between model options. The first model set included a baseline model with no covariates on any parameters and model with year as a predictor on the location parameter to test for stationarity in the ozone time series. The second and third sets of models tested the inclusion of temperature, vapor pressure deficit, and humidity as covariates on the location and scale parameters, respectively. At each stage, we used likelihood ratio tests to evaluate the significance of adding additional covariates to the model formulas. If the inclusion of covariates was significant (p<=.05), we proceeded to the next set of models with the chosen formulation. For choosing between meteorological covariates, we selected the covariate model with the lowest Akaike Information Criteria (AIC). To test model sensitivity, we also retained alternate model fits at each stage of the model fitting process. These included a baseline model with no covariates, a baseline model with only year as a covariate, a location parameter only model, and both location and location/scale models that forced the inclusion of the year covariate to further assess the temporal stationarity assumption. In total, this returned six candidate models for each grid cell.

Using the model candidates for each grid cell, we produced two primary metrics: effective return levels at 5-year, 20-year, and 50-year return periods, and the expected number of and the expected number of ozone 'episode days' (>71 ppb or 'unhealthy for sensitive groups' in the EPA Air Index). Both metrics were produced for current and 2050s climate conditions. Effective return levels are analogous to quantiles above for each grid cell, e.g. a 20-year effective return level corresponds to the ozone concentration that occurs with 5% probability above the threshold. We use the daily time series of current and future as inputs for both metrics and average outputs across all ten years to return expected annual values, with additional sensitivity tests on the future predictions using the 10th and 90th percentile change factors.. For future predictions in models with year as a covariate, we fixed the value of the year covariate to equal the time frame 2015–2017, the final three years in the dataset. To compare fits across model candidates, we compared the root mean squared error (RMSE) between model predicted effective return levels and corresponding quantiles from the fused predictions.

3. Results

Fig. 5 shows the RMSE across our model candidates. Models labeled with 'year' are the candidate models that included year as a covariate on the location parameter for all grid cells. We find significant reductions in RMSE (~5 ppb) in any model with meteorological covariate data and small differences between subsequent variations. Differences between the location and location/scale models are very small, but the candidate models with year as a covariate improve predictions for several specific outlier grid cells. The location/scale and year model shows the best overall performance, but only six grid cells pass the likelihood ratio test for inclusion of scale covariates (Fig. S9). Based on these findings, we select the location model with year and meteorological covariates as our preferred model that balances complexity and simplicity. The fitted meteorological covariates for each grid cell under this model are shown in Fig. 6. Results for the remainder of the model options are included in the supplementary information.

Fig. 7 shows ozone effective return levels in parts per billion at 5, 20, and 50 year return periods across California and their respective estimated changes under 2050s climate conditions. Results for the 10th and 90th percentile climate conditions are included in the supplementary information. Effective return levels range from 44-47 ppb to 109–164 ppb across California, with the lowest values along the central coastline and the highest values in the central and south central regions. These estimates are consistent with observational records for the areas that have historically seen high ozone exposures. We find the largest change in return levels (6–8 ppb) in the Central Valley and Lake Tahoe, with more modest increases of 0–4 ppb in the northern part of the state. Our results show small decreases 0–4 ppb in select regions, particularly along the coastline and across the southern portions of the state. These spatial patterns are preserved in the 10th and 90th percentile scenarios (Figs. S2 and S5). Examining the change factors (Fig. 4) in conjunction with the model covariates (Fig. 6) shows that these changes are driven by increasing minimum humidity levels in the southern regions and increasing vapor pressure deficit and temperature in the central and northern regions. The magnitude of predicted changes is generally not large enough to dramatically alter the air quality of regions with good to moderate ozone levels in current conditions, but these results do suggest that some areas in northern California without present-day concern are likely to experience some unhealthy days under 2050s climate conditions.

Fig. 8 shows the model predicted number of ozone episode days across California for current and 2050s conditions under an ensemble mean of CMIP5 climate projections. We find that a majority of California experiences less than ten ozone episode days per year. Higher values are geographically concentrated in the central and south-central regions, with some grid cells showing up to 15 ozone episode days per year. Model predictions under 2050s climate conditions have a greater effect in these areas with existing ozone pollution, with estimated increases of 5 to 16 additional annual ozone episode days. The climate effect is less pronounced throughout the rest of the state where smaller changes of one to four ozone episode days are expected. In total, these changes indicate a net climate penalty, with projected increases in ozone pollution outweighing the effects of localized reductions. However, changing climate conditions are not generally large enough to create ozone episode days in areas without existing ozone pollution..

4. Discussion

Although emission control strategies have been effective at reducing ozone concentrations across California, large portions of the state remain in non-attainment status under EPA standards and face compounding challenges under changing climate conditions that are conducive to ozone formation (Mahmud et al., 2008; Rasmussen et al., 2013). With growing evidence that ozone exposure contributes to a broad range of negative health effects (Bell et al., 2007; Fann et al., 2015, 2021; Garcia-Menendez et al., 2015; Kinney, 2008; Orru et al., 2017; Silva et al., 2017; Tagaris et al., 2009), especially in the short-term, understanding the potential burden of climate change on high levels of ozone pollution is an important area of research. Using an EVT modeling approach in combination with a fused data set of EPA station observations and CMAQ model output, we produced a spatially explicit characterization of extreme ozone concentrations under past and future climate conditions. Our results show increases in annual ozone episode days under 2050s climate conditions by an average of 5.4 episode-days (Fig. 8). Corresponding average changes in effective return level estimates range from −4 to 8 ppb, with a small net positive mean increase of 0.3 - 0.4 ppb across California ppb (Fig. 7). The number of additional episode days trends significantly higher n areas that currently experience high levels of ozone pollution, including much of Central California. These results are consistent with other research highlighting the emergence of a climate penalty of ozone in California (Shen et al., 2016).

Our methodological approach expands previous research in several important ways. While multiple meteorological variables including temperature, humidity, wind speed, and vapor pressure deficit have

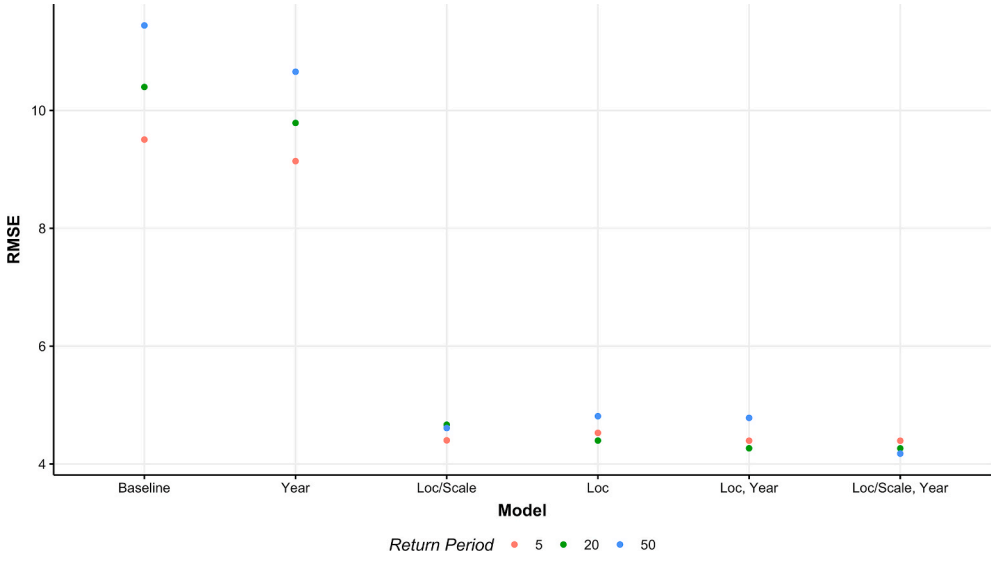

Fig. 5. RMSE by model and return period.

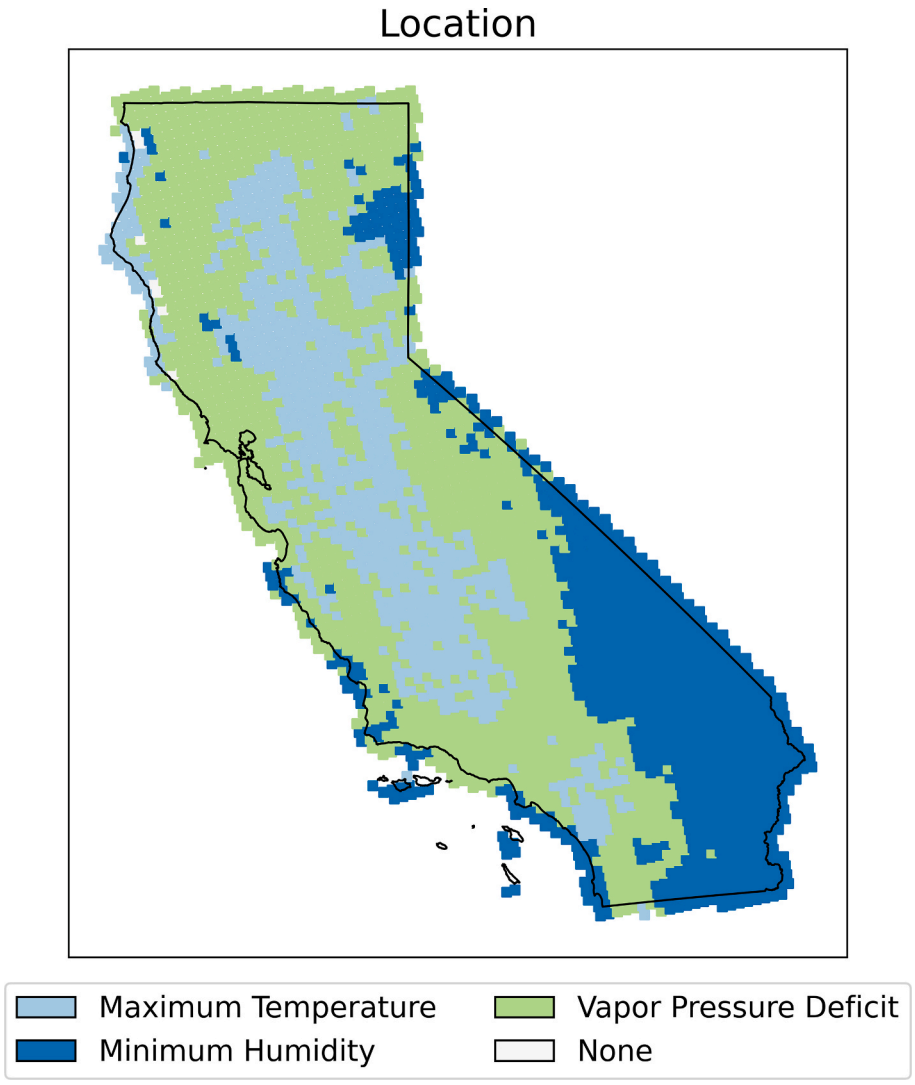

Fig. 6. Meteorological covariates on the location parameter selected for each grid cell.

been included as explanatory covariates in previous statistical models of ozone concentrations, extreme-value analyses have predominantly focused on extreme heat (Gouldsbrough et al., 2022; Shen et al., 2016). In this study, we evaluated an expanded covariate set, finding strong ozone tail dependence on temperature, humidity, and vapor pressure deficit at varying locations across California. These results were robust across multiple formulations of the EVT model, andare consistent with literature showing non-stationary covariate relationships across the U.S. (Wells et al., 2021). Our findings are also consistent with studies proposing the importance of vapor pressure deficit as an ozone predictor and an important meteorological dependency that helps regulate the ozone-humidity relationship (Kavassalis and Murphy, 2017). These findings highlight the importance of considering multiple meteorological covariates in analyses of ozone tail dependence, even in locations that have shown historically strong temperature-ozone relationships. By fitting a separate EVT model to each grid cell and station location, our approach allows for the highest degree of non-stationary covariate relationships. We found covariate relationships typically vary smoothly across space (Fig. 6), but there are instances where individual pixels are different from their spatial neighbors or stations are fit to different covariates than their containing grid cell. In future work, it is worth considering whether a regional modeling approach or another EVT implementation with spatial dependence would reduce the risk of overfitting individual locations.

Another novel feature of our analysis is the fusion of both EPA stations and CMAQ grid-based ozone estimates. To our knowledge, this is the first study that applies an EVT modeling approach to simulate extreme ozone concentrations under future climate conditions on a spatially explicit grid. Although CMAQ is a modeled data product and comes with more uncertainty, it provides consistent spatial coverage that allows for estimation in areas that lack ozone monitoring stations. Even in California, a state with one of the densest networks of ozone monitoring stations in the U.S., there exist significant gaps in coverage. We find that some of the highest predicted ozone return levels are located in these areas without station coverage (Fig. 7). These results reinforce the need to provide more complete representations of air pollution across space and time—a trend that is developing alongside new data assimilation techniques and low cost monitoring sensors (Berrocal et al., 2012; Davis et al., 2011; Requia et al., 2020; Skipper et al., 2021). Our model approach also has several limitations. The CMAQ grid resolution we used in this analysis is 12 km, which may not capture fine-scale spatial variation in meteorology or ozone levels. We applied a simple bias correction to bring the underlying CMAQ data more in line with the EPA station observations, but our procedure does not fully remove all differences between the two sources of information. . Implementing a more complex bias correction, potentially with other covariates, may further improve the accuracy of results. We also assumed that ozone precursor emissions and background ozone levels

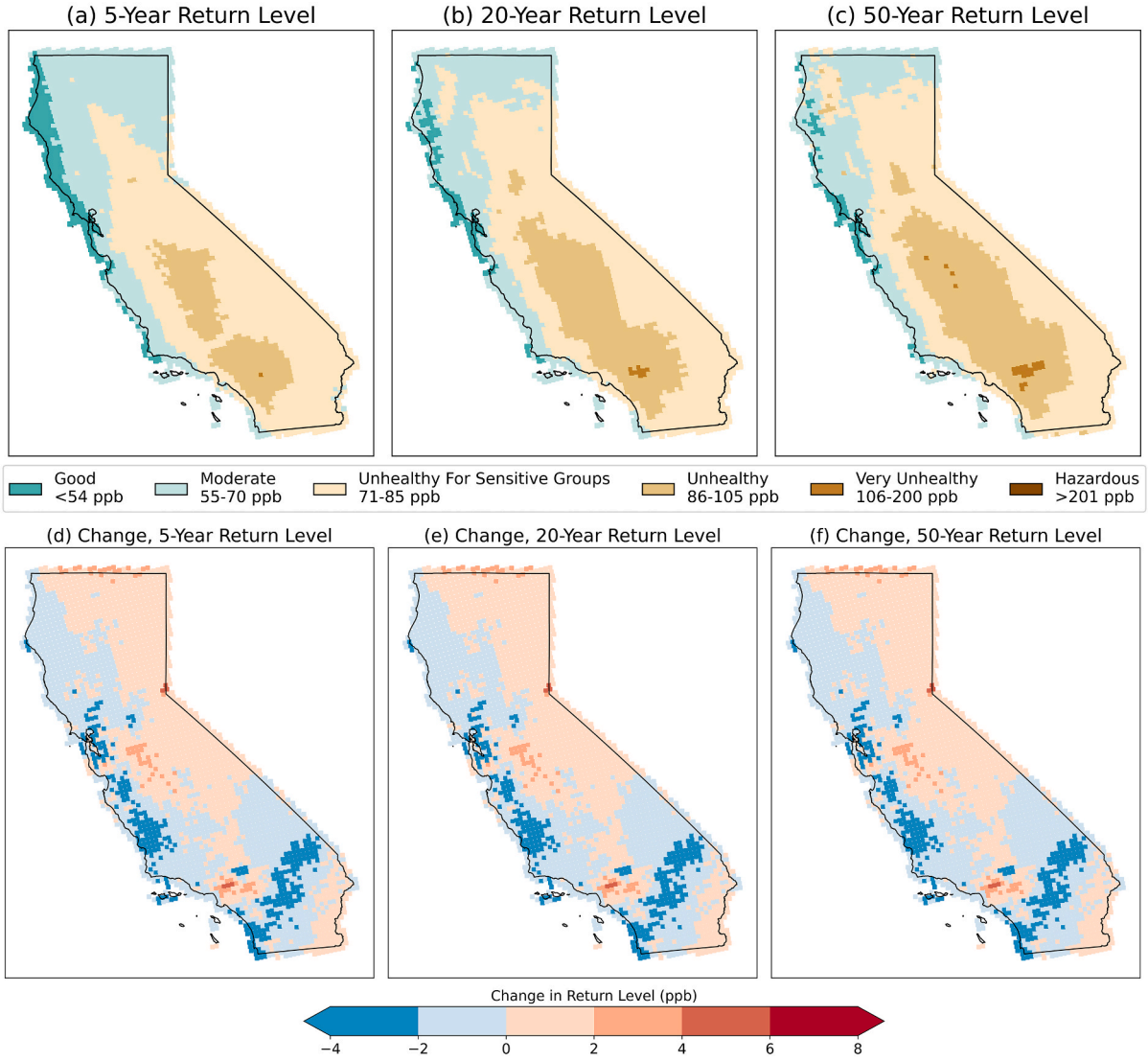

Fig. 7. Maps of effective return levels at return periods (5, 20, 50) and changes under 2050s climate conditions.

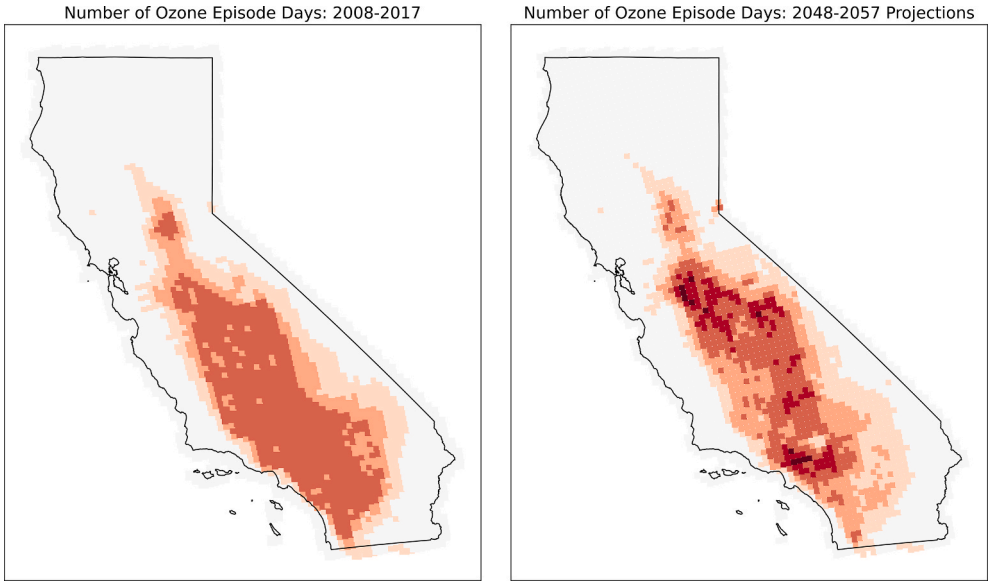

Fig. 8. Estimated number of ozone episode days under current and future climate conditions.

remained constant at 2008–2017 levels in this study. Background ozone levels are affected by a multitude of factors that could plausibly change under future climate conditions including air flow dynamics and turbulence levels, chemical reaction rates, and vertical mixing boundary layer (Ebi and McGregor, 2008). However, recent research points to California remaining a regional hotspot for background ozone concentrations, with frequent occurrences of stable anticyclonic conditions alongside high temperatures and low boundary layer ventilation (Gaudel et al., 2018; Pan and Faloona, 2022). These conditions are also increasingly present during wildfires, which have played a dominant role in ozone episode days in recent years (Pan and Faloona, 2022). Even without considering increased biomass burning during fires, California consistently shows increases in ozone concentrations across multiple climate scenarios and under both 2011 and projected 2040 emissions inventories (Nolte et al., 2021). Therefore, we expect the results of this research to be a conservative estimate of the total ozone climate penalty. As the multifaceted interactions between meteorological and anthropogenic factors are better understood into the future, the extreme-value theory approach presented in this paper could be revisited.

5. Conclusion

In order to better understand how projected 2050s climate conditions will affect extreme ozone levels across all of California, we trained a Point Process EVT model on a fused ozone surface produced from AQS and CMAQ ozone datasets from 2008 to 2017 and applied the results to 2050 climate conditions derived from downscaled CMIP5 GCMs. We found that a combination of daily temperature, humidity, and vapor pressure deficit covariates were able to effectively model observed ozone tail distributions between 2008 and 2017. Our results show net increases in ozone episode days and higher ozone return level concentrations across most of California under 2050s climate conditions, adding evidence to a larger body of literature suggesting that a climate penalty may offset planned emission reduction strategies. Continued research into characterizing the multiple pathways by which future climate conditions may be conducive to ozone formation is critically important for informing subsequent population health implications research—especially for sensitive populations that may ultimately be the ones most likely to be negatively affected by the ozone climate penalty.

Credit Author statement

Bradley Wilson, PhD: Conceptualization, methodology, software, validation, visualization, supervision, writing - original draft, writing - review and editing, **Mariah Pope:** Conceptualization, methodology, software, validation, visualization, supervision, **Jeremy R. Porter, PhD:** Conceptualization, methodology, writing - original draft, writing - review and editing, **Dr. Edward Kearns, PhD:** Conceptualization, writing - original draft, writing - review and editing, supervision, **Evelyn Shu:** Validation, writing - original draft, writing - review and editing, **Mark Bauer:** Validation, writing - original draft, writing - review and editing, **Neil Freeman:** Validation, writing - original draft, writing - review and editing, **Mike Amodeo:** Validation, writing - original draft, writing - review and editing, **David Melecio-Vazquez, PhD:** Validation, writing - original draft, writing - review and editing, **Ho Hsieh:** Validation, writing - original draft, writing - review and editing, **Maggie Tarasovitch:** Validation, writing - original draft, writing - review and editing, project administration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.aeaoa.2022.100195.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modeling. Int. J. Climatol. 33, 121–131. https://doi.org/10.1002/joc.3413.

Abatzoglou, J.T., Brown, T.J., 2012. A comparison of statistical downscaling methods suited for wildfire applications: statistical Downscaling for Wildfire Applications. Int. J. Climatol. 32, 772–780. https://doi.org/10.1002/joc.2312.

Allen, R.G., Food and Agriculture Organization of the United Nations (FAO), 1998. Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements. FAO irrigation and drainage paper. Food and Agriculture Organization of the United Nations, Rome.

Appel, K.Wyat, Bash, Jesse O., Fahey, Kathleen M., Foley, Kristen M., Gilliam, Robert C., Hogrefe, Christian, Hutzell, William T., Kang, Daiwen, Mathur, Rohit, Murphy, Benjamin N., Napelenok, Sergey L., Nolte, Christopher G., Pleim, Jonathan E., Pouliot, George A., Pye, Havala O.T., Ran, Limei, Roselle, Shawn J., Sarwar, Golam, Schwede, Donna B., Sidi, Rahim I., Spero, Tanya L., Wong, David C., 2021. The Community Multiscale Air Quality (CMAQ) model versions 5.3 and 5.3.1: system updates and evaluation. Geosci. Model Dev. 14, 2867–2897. https://doi.org/10.5194/gmd-14-2867-2021.

Arnold, S.R., Lombardozzi, D., Lamarque, J.F., Richardson, T., Emmons, L.K., Tilmes, S., Sitch, S.A., Folberth, G., Hollaway, M.J., Val Martin, M., 2018. Simulated global climate response to tropospheric ozone-induced changes in plant transpiration. Geophys. Res. Lett. 45, 13070–13079. https://doi.org/10.1029/2018GL079938.

Bell, M.L., Goldberg, R., Hogrefe, C., Kinney, P.L., Knowlton, K., Lynn, B., Rosenthal, J., Rosenzweig, C., Patz, J.A., 2007. Climate change, ambient ozone, and health in 50 US cities. Clim. Change 82, 61–76. https://doi.org/10.1007/s10584-006-9166-7.

Berrocal, V.J., Gelfand, A.E., Holland, D.M., 2012. Space-time data fusion under error in computer model output: an application to modeling air quality. Biometrics 68, 837–848. https://doi.org/10.1111/j.1541-0420.2011.01725.x.

Bloomer, Bryan J., Stehr, Jeffery W., Piety, Charles A., Salawitch, Ross J., Dickerson, Russell R., 2009. Observed relationships of ozone air pollution with temperature and emissions. Geophys. Res. Lett. 36 (9) https://doi.org/10.1029/2009GL037308.

Chen, G., Li, J., Ying, Q., Sherman, S., Perkins, N., Rajeshwari, S., Mendola, P., 2014. Evaluation of observation-fused regional air quality model results for population air pollution exposure estimation. Sci. Total Environ. 563–574, 0.

Coates, J., Mar, K.A., Ojha, N., Butler, T.M., 2016. The influence of temperature on ozone production under varying NOxconditions – a modelling study. Atmos. Chem. Phys. 16, 11601–11615. https://doi.org/10.5194/acp-16-11601-2016.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values, Springer Series in Statistics. Springer, London ; New York.

Davis, J., Cox, W., Reff, A., Dolwick, P., 2011. A comparison of CMAQ-based and observation-based statistical models relating ozone to meteorological parameters. Atmos. Environ. 45, 3481–3487. https://doi.org/10.1016/j.atmosenv.2010.12.060.

Doherty, R.M., Wild, O., Shindell, D.T., Zeng, G., MacKenzie, I.A., Collins, W.J., Fiore, A. M., Stevenson, D.S., Dentener, F.J., Schultz, M.G., Hess, P., Derwent, R.G., Keating, T.J., 2013. Impacts of climate change on surface ozone and intercontinental ozone pollution: a multi-model study. J. Geophys. Res. Atmospheres 118, 3744–3763. https://doi.org/10.1002/jgrd.50266.

Ebi, Kristie L., McGregor, Glenn, 2008. Climate Change, Tropospheric Ozone and Particulate Matter, and Health Impacts. Environ. Health Perspect. 116 (11), 1449–1455. https://doi.org/10.1289/ehp.11463.

Fann, N., Brennan, T., Dolwick, P., Gamble, J.L., Ilacqua, V., Kolb, L., Nolte, C.G., Spero, T.L., Ziska, L., 2016. Ch. 3: air quality impacts. The Impacts of Climate Change on Human Health in the United States: A Scientific Assessment. U.S. Global Change Research Program. https://doi.org/10.7930/J0GQ6VP6.

Fann, N., Nolte, C.G., Dolwick, P., Spero, T.L., Brown, A.C., Phillips, S., Anenberg, S., 2015. The geographic distribution and economic value of climate change-related ozone health impacts in the United States in 2030. J. Air Waste Manag. Assoc. 65, 570–580. https://doi.org/10.1080/10962247.2014.996270.

Fann, N.L., Nolte, C.G., Sarofim, M.C., Martinich, J., Nassikas, N.J., 2021. Associations between simulated future changes in climate, air quality, and human health. JAMA Netw. Open 4, e2032064. https://doi.org/10.1001/jamanetworkopen.2020.32064.

Garcia-Menendez, F., Saari, R.K., Monier, E., Selin, N.E., 2015. U.S. Air quality and health benefits from avoided climate change under greenhouse gas mitigation. Environ. Sci. Technol. 49, 7580–7588. https://doi.org/10.1021/acs.est.5b01324.

Gaudel, A., Cooper, O.R., Ancellet, G., Barret, B., Boynard, A., Burrows, J.P., Clerbaux, C., Coheur, P.-F., Cuesta, J., Cuevas, E., Doniki, S., Dufour, G., Ebojie, F., Foret, G., Garcia, O., Granados-Muñoz, M.J., Hannigan, J.W., Hase, F., Hassler, B., Huang, G., Hurtmans, D., Jaffe, D., Jones, N., Kalabokas, P., Kerridge, B., Kulawik, S., Latter, B., Leblanc, T., Le Flochmoën, E., Lin, W., Liu, J., Liu, X., Mahieu, E., McClure-Begley, A., Neu, J.L., Osman, M., Palm, M., Petetin, H., Petropavlovskikh, I., Querel, R., Rahpoe, N., Rozanov, A., Schultz, M.G., Schwab, J.,

Siddans, R., Smale, D., Steinbacher, M., Tanimoto, H., Tarasick, D.W., Thouret, V., Thompson, A.M., Trickl, T., Weatherhead, E., Wespes, C., Worden, H.M., Vigouroux, C., Xu, X., Zeng, G., Ziemke, J., 2018. Tropospheric Ozone Assessment Report: present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation. Elementa: Science of the Anthropocene 6, 39. https://doi.org/10.1525/elementa.291.

Gilleland, E., Katz, R.W., 2016. extRemes 2.0: an extreme value analysis package in *R*. J. Stat. Software 72. https://doi.org/10.18637/jss.v072.i08.

Gouldsbrough, L., Hossaini, R., Eastoe, E., Young, P.J., 2022. A temperature dependent extreme value analysis of UK surface ozone, 1980–2019. Atmos. Environ. 273, 118975 https://doi.org/10.1016/j.atmosenv.2022.118975.

Kavassalis, S.C., Murphy, J.G., 2017. Understanding ozone-meteorology correlations: a role for dry deposition: ozone-Meteorology Correlations: dry Dep. Geophys. Res. Lett. 44, 2922–2931. https://doi.org/10.1002/2016GL071791.

Kinney, P.L., 2008. Climate change, air quality, and human health. Am. J. Prev. Med. 35, 459–467. https://doi.org/10.1016/j.amepre.2008.08.025.

Lindgren, F., Rue, H., Lindstrom, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the SPDE approach (with discussion). J. Roy. Stat. Soc. B 73 (4), 423–498.

Mahmud, A., Tyree, M., Cayan, D., Motallebi, N., Kleeman, M.J., 2008. Statistical downscaling of climate change impacts on ozone concentrations in California. J. Geophys. Res. 113, D21103 https://doi.org/10.1029/2007JD009534.

Nolte, C.G., Dolwick, P., Fann, N., Horowitz, L.W., Naik, V., Pinder, R.W., Spero, T.L., Winner, D.A., Ziska, L.H., 2018. Chapter 13 : air quality. Impacts, risks, and adaptation in the United States: the fourth national climate assessment, volume II. U. S. Global change research program. https://doi.org/10.7930/NCA4.2018.CH13.

Nolte, C.G., Spero, T.L., Bowden, J.H., Sarofim, M.C., Martinich, J., Mallard, M.S., 2021. Regional temperature-ozone relationships across the U.S. under multiple climate and emissions scenarios. J. Air Waste Manag. Assoc. 71, 1251–1264. https://doi.org/10.1080/10962247.2021.1970048.

Orru, H., Ebi, K.L., Forsberg, B., 2017. The interplay of climate change and air pollution on health. Curr. Environ. Health Rep. 4, 504–513. https://doi.org/10.1007/s40572-017-0168-6.

Pan, K., Faloona, I., 2022. The impacts of wildfires on ozone production and boundary layer dynamics in California's Central Valley. Atmos. Chem. Phys. 22, 9681–9702. https://doi.org/10.5194/acp-22-9681-2022.

Parrish, D.D., Lamarque, J.-F., Naik, V., Horowitz, L., Shindell, D.T., Staehelin, J., Derwent, R., Cooper, O.R., Tanimoto, H., Volz-Thomas, A., Gilge, S., Scheel, H.-E., Steinbacher, M., Fröhlich, M., 2014. Long-term changes in lower tropospheric baseline ozone concentrations: comparing chemistry-climate models and observations at northern midlatitudes. J. Geophys. Res. Atmospheres 119, 5719–5736. https://doi.org/10.1002/2013JD021435.

Rasmussen, D.J., Hu, J., Mahmud, A., Kleeman, M.J., 2013. The ozone–climate penalty: past, present, and future. Environ. Sci. Technol. 47, 14258–14266. https://doi.org/10.1021/es403446m.

Requia, W.J., Di, Q., Silvern, R., Kelly, J.T., Koutrakis, P., Mickley, L.J., Sulprizio, M.P., Amini, H., Shi, L., Schwartz, J., 2020. An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States. Environ. Sci. Technol. 54, 11037–11047. https://doi.org/10.1021/acs.est.0c01791.

Rieder, H.E., Staehelin, J., Maeder, J.A., Peter, T., Ribatet, M., Davison, A.C., Stübi, R., Weihs, P., Holawe, F., 2010. Extreme events in total ozone over Arosa – Part 1: application of extreme value theory. Atmos. Chem. Phys. 10, 10021–10031. https://doi.org/10.5194/acp-10-10021-2010.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models using integrated nested Laplace approximations (with discussion). J. Roy. Stat. Soc. B 71 (2), 319–392.

Shen, L., Mickley, L.J., Gilleland, E., 2016. Impact of increasing heat waves on U.S. ozone episodes in the 2050s: results from a multimodel analysis using extreme value theory: predict Future Ozone Episodes. Geophys. Res. Lett. 43, 4017–4025. https://doi.org/10.1002/2016GL068432.

Shi, L., Liu, P., Zanobetti, A., Schwartz, J., 2019. Climate Penalty: climate-driven increases in ozone and PM2.5 levels and mortality. Environ. Epidemiol. 3, 365. https://doi.org/10.1097/01.EE9.0000610052.11501.34.

Silva, R.A., West, J.J., Lamarque, J.-F., Shindell, D.T., Collins, W.J., Faluvegi, G., Folberth, G.A., Horowitz, L.W., Nagashima, T., Naik, V., Rumbold, S.T., Sudo, K., Takemura, T., Bergmann, D., Cameron-Smith, P., Doherty, R.M., Josse, B., MacKenzie, I.A., Stevenson, D.S., Zeng, G., 2017. Future global mortality from changes in air pollution attributable to climate change. Nat. Clim. Change 7, 647–651. https://doi.org/10.1038/nclimate3354.

Skipper, T.N., Hu, Y., Odman, M.T., Henderson, B.H., Hogrefe, C., Mathur, R., Russell, A. G., 2021. Estimating US background ozone using data fusion. Environ. Sci. Technol. 55, 4504–4512. https://doi.org/10.1021/acs.est.0c08625.

Smith, R.L., 1989. Extreme value analysis of environmental time series: an application to trend detection in ground-level ozone. Stat. Sci. 4, 367–377.

Stowell, J.D., Kim, Y., Gao, Y., Fu, J.S., Chang, H.H., Liu, Y., 2017. The impact of climate change and emissions control on future ozone levels: implications for human health. Environ. Int. 108, 41–50. https://doi.org/10.1016/j.envint.2017.08.001.

Tagaris, E., Liao, K.-J., DeLucia, A.J., Deck, L., Amar, P., Russell, A.G., 2009. Potential impact of climate change on air pollution-related human health effects. Environ. Sci. Technol. 43, 4979–4988. https://doi.org/10.1021/es803650w.

Thompson, M.L., Reynolds, J., Cox, L.H., Guttorp, P., Sampson, P.D., 2001. A review of statistical methods for the meteorological adjustment of tropospheric ozone. Atmos. Environ. 14.

Torres-Vazquez, A., Pleim, J., Gilliam, R., Pouliot, G., 2022. Performance evaluation of the meteorology and air quality conditions from Multiscale WRF-CMAQ simulations for the long island sound tropospheric ozone study (LISTOS). J. Geophys. Res. Atmospheres 127, e2021JD035890. https://doi.org/10.1029/2021JD035890.

Towler, Erin, Llewellyn, Dagmar, Prein, Andreas, Gilleland, Eric, 2020. Extreme-value analysis for the characterization of extremes in water resources: A generalized workflow and case study on New Mexico monsoon precipitation. Weather Clim. Extrem. 29 https://doi.org/10.1016/j.wace.2020.100260.

Turnock, S.T., Allen, R.J., Andrews, M., Bauer, S.E., Deushi, M., Emmons, L., Good, P., Horowitz, L., John, J.G., Michou, M., Nabat, P., Naik, V., Neubauer, D., O'Connor, F. M., Olivié, D., Oshima, N., Schulz, M., Sellar, A., Shim, S., Takemura, T., Tilmes, S., Tsigaridis, K., Wu, T., Zhang, J., 2020. Historical and future changes in air pollutants from CMIP6 models. Atmos. Chem. Phys. 20, 14547–14579. https://doi.org/10.5194/acp-20-14547-2020.

US EPA, 2016. Ozone trends [WWW Document]. URL. https://www.epa.gov/air-trends/ozone-trends, 7.28.22.

US EPA, 2021. CMAQ. Zenodo. https://doi.org/10.5281/zenodo.5213949.

US EPA, 2022. Green book: California nonattainment/maintenance status for each county by year for all Criteria pollutants. https://www3.epa.gov/airquality/greenbook/jbtc.html, 7.13.22.

Wells, B., Dolwick, P., Eder, B., Evangelista, M., Foley, K., Mannshardt, E., Misenis, C., Weishampel, A., 2021. Improved estimation of trends in U.S. ozone concentrations adjusted for interannual variability in meteorological conditions. Atmos. Environ. 248, 118234 https://doi.org/10.1016/j.atmosenv.2021.118234.

Wu, S., Mickley, L.J., Leibensperger, E.M., Jacob, D.J., Rind, D., Streets, D.G., 2008. Effects of 2000 – 2050 global change on ozone air quality in the United States. J. Geophys. Res. 113, D06302 https://doi.org/10.1029/2007JD008917.

Zhang, J., Gao, Y., Luo, K., Leung, L.R., Zhang, Y., Wang, K., Fan, J., 2018. Impacts of compound extreme weather events on ozone in the present and future. Atmos. Chem. Phys. 18, 9861–9877. https://doi.org/10.5194/acp-18-9861-2018.