Earth's Future

RESEARCH ARTICLE
10.1029/2022EF003317

Key Points:

- Precursor emission changes are represented by the inclusion of ozone trend in the statistical downscaling models
- Statistical projections results are sensitive to the predictor types and Earth System Model output used
- Climate change scenarios have a major impact on future ozone, and greenhouse gas and air pollution reductions can mitigate health risks

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

E. Hertig,
elke.hertig@med.uni-augsburg.de

Citation:

Hertig, E., Jahn, S., & Kaspar-Ott, I. (2023). Future local ground-level ozone in the European area from statistical downscaling projections considering climate and emission changes. *Earth's Future*, *11*, e2022EF003317. https://doi.org/10.1029/2022EF003317

Received 2 NOV 2022
Accepted 23 JAN 2023

Author Contributions:

Conceptualization: Elke Hertig
Formal analysis: Sally Jahn, Irena Kaspar-Ott
Funding acquisition: Elke Hertig
Methodology: Elke Hertig
Project Administration: Elke Hertig
Supervision: Elke Hertig
Writing – original draft: Elke Hertig, Sally Jahn, Irena Kaspar-Ott
Writing – review & editing: Elke Hertig, Sally Jahn, Irena Kaspar-Ott

Future Local Ground-Level Ozone in the European Area From Statistical Downscaling Projections Considering Climate and Emission Changes

Elke Hertig[1], Sally Jahn[1], and Irena Kaspar-Ott[1]

[1]Regional Climate Change and Health, Faculty of Medicine, University of Augsburg, Universitätsstrasse 2, Augsburg, Germany

Abstract Ground-level ozone is a major air pollutant harmful to human health. In the scope of climate change, it is essential to provide high-quality local-scale assessments of the anticipated changes for public health and policy interventions. Assessments and projections of ground-level ozone usually rely on numerical modeling, but statistical approaches are also available. The present study enhances the validity of statistical downscaling by taking climate change as well as air pollution changes into account. Besides considering meteorological predictors such as air temperature, short-wave radiation, humidity, and wind, ozone trends from changes in precursor emissions were included in the statistical models. Meteorological and ozone predictor information extracted from reanalysis data for the observational period and output of seven Earth System Models (ESMs) for the projection periods were used, with three of them having interactive chemical modeling, while the other four used prescribed ozone changes. Ground-level ozone, more precisely daily maximum 8-hr running means (MDA8) as well as daily maximum 1-hr values (MDA1), at 798 measurement stations across the European area in the "ozone season" from April to September were assessed. Results depended strongly on whether only meteorological information or additional information about emission changes were considered. As a general picture under the consideration of climate and emission changes, decreasing ground-level ozone concentrations were projected under the moderate SSP2-4.5 scenario, while for the more pessimistic scenario SSP3-7.0 increasing ozone concentrations over Europe, especially at the end of the 21st century, were assessed.

Plain Language Summary Ground-level ozone is a gaseous air pollutant that is formed under sunlight from other air pollutants. Ground-level ozone is harmful to human health. There are concerns that under climate change, ground-level ozone will increase. To assess changes on local-scale various methods are available, which can roughly be divided into numerical models and statistical approaches. The present study developed statistical models further, to give a more realistic picture of future changes in ground-level ozone in the European area. Information about meteorological changes such as air temperature, atmospheric humidity, and wind as well as changes in air pollutants like nitrogen oxides was considered. During the 21st century decreasing ground-level ozone concentrations were projected under a moderate climate change scenario, while for a more pessimistic scenario increasing ozone concentrations over Europe, especially at the end of the century, were assessed. This result highlights the necessity to further reduce greenhouse gases and air pollution, bringing forth better protection of human health.

1. Introduction

Ground-level ozone (O_3) is a photochemical oxidant formed from precursors in a polluted atmosphere, mainly nitrogen oxides (NO_x) and non-methane volatile organic compounds (VOCs). O_3 at the surface is harmful to human health and the World Health Organization recently tightened the air quality guidelines based on new epidemiological evidence (World Health Organization, 2021).

While already many regions of the world frequently exceed the recommended levels, concerns have been raised that in the scope of ongoing climate change ground-level O_3 concentrations will further increase despite efforts for rigorous air pollution control. This "climate penalty" implies the need for stronger emission controls to achieve a given air quality standard (Wu et al., 2008). Processes that are discussed to contribute to elevated O_3 concentrations due to climate change comprise warming-induced biogenic VOC emission enhancements, faster chemistry kinetics, and faster peroxyacetyl nitrate decomposition (Lu et al., 2019). Porter and Heald (2019) found that temperature-dependent deposition and soil NO_x emissions also contribute to the "climate penalty".

Furthermore, enhanced stratosphere–troposphere exchange, changes in the large-scale atmospheric circulation and synoptic patterns, increased stagnancy, and changes in atmospheric humidity may lead to increases in surface O_3 (Lu et al., 2019; Porter & Heald, 2019). In general, models project increases in surface O_3, and most notably of extreme O_3 episodes, in polluted regions and decreases over rural areas and oceans (Schnell et al., 2016). This trend was found in Europe already in the observational period 2000–2015, with daily mean O_3 decreasing at rural sites, while in suburban and urban sites stable or increasing trends occurred (Boleti et al., 2020). This rural-urban difference is commonly attributed to the reduction of precursor emissions, which are most effective in rural regions with low NO_x concentrations. At low NO_x conditions, ozone production is limited by the supply of NO_x (NO_x-limited regime), whereas at high NO_x levels, ozone production linearly increases with VOCs concentration, but decreases with NO_x concentrations (VOC-limited regime) (Lu et al., 2019). For Europe, projections by Orru et al. (2019) indicate that the health burden due to O_3 increases due to climate change and the size of susceptible populations. However, because of reductions in O_3 precursor emissions, an overall decrease in the total health burden in 2050 was estimated. Extreme O_3 events are expected to increase in combination with rising heat events, in particular over mid-latitude regions like Central Europe or the Northeast of the United States (Jahn & Hertig, 2022; Shen et al., 2016). Thus, O_3 health risks could combine with increased health risks from heat exposure.

Projections of future tropospheric O_3 under climate change are commonly conducted using climate chemistry models coupled with general circulation models and regional climate models, respectively (a recent overview is given in Lu et al., 2019). Since these dynamical models are computationally expensive, the number of considered models and model runs is usually limited. For downscaling to regional to local scales, besides dynamical downscaling, various statistical analyses have been conducted, mostly focusing on the relationship between meteorological factors and synoptic conditions with ozone concentrations (e.g., Hertig, 2020; Hertig et al., 2020; Jahn & Hertig, 2021; Jahn & Hertig, 2022; Otero et al., 2016). Statistical analysis has the advantage of being computationally inexpensive, but it mostly comes with the limitation that it considers solely meteorological drivers, while the impact and changes of precursor emissions are not regarded.

Here, we used a statistical downscaling framework based on the Perfect Prognosis (PP) approach (for an overview of statistical downscaling methods see e.g., Gutiérrez et al., 2019) to assess future changes of local, ground-level O_3 in the European area. The PP models were established in the observational period using local, station-based O_3, as the target variable, more precisely daily maximum 8-hr running means (MDA8) as well as daily maximum 1-hr values (MDA1). Reanalysis-based variables served as predictors. To generate future projections, predictors simulated by the latest generation of Earth System Models (ESMs) from the Coupled Model Intercomparison Project Phase 6 were used.

In addition to the inclusion of information on meteorological changes, we factored changes in precursor emissions into the statistical downscaling models. This was done via the incorporation of O_3 time series derived from the ESM model output. These ESM O_3 time series reflect the larger-scale total changes of O_3, resulting from climate change and from changes of precursor emissions under different scenario assumptions. Thus, the precursor gases were not used as predictors in the statistical models. Yet the information about their change and impact on O_3 progression was included by directly taking ESM-modeled O_3. To assess the impact of different choices within the statistical downscaling, a sensitivity analysis was conducted, analyzing the impact of different predictor types (emission-related and meteorological, i.e., radiative and thermal, thermo-dynamic, and circulation-dynamic), different ESMs (with interactive computation of the atmospheric chemistry vs. prescribed O_3 changes), and different scenario assumptions (using the shared socioeconomic pathways SSP2-4.5 and SSP3-7.0). Section 2 presents the data and methods used in this study, while Section 3 elaborates first on the statistical downscaling models in the observational period, followed by the projection results for local MDA8 in the European area until the end of the 21st century, including the comparison of different predictor sets. Section 4 discusses the results further, while in Section 5 conclusions are drawn. Results for MDA1 were generally similar to MDA8 and can be found in the Supporting Information.

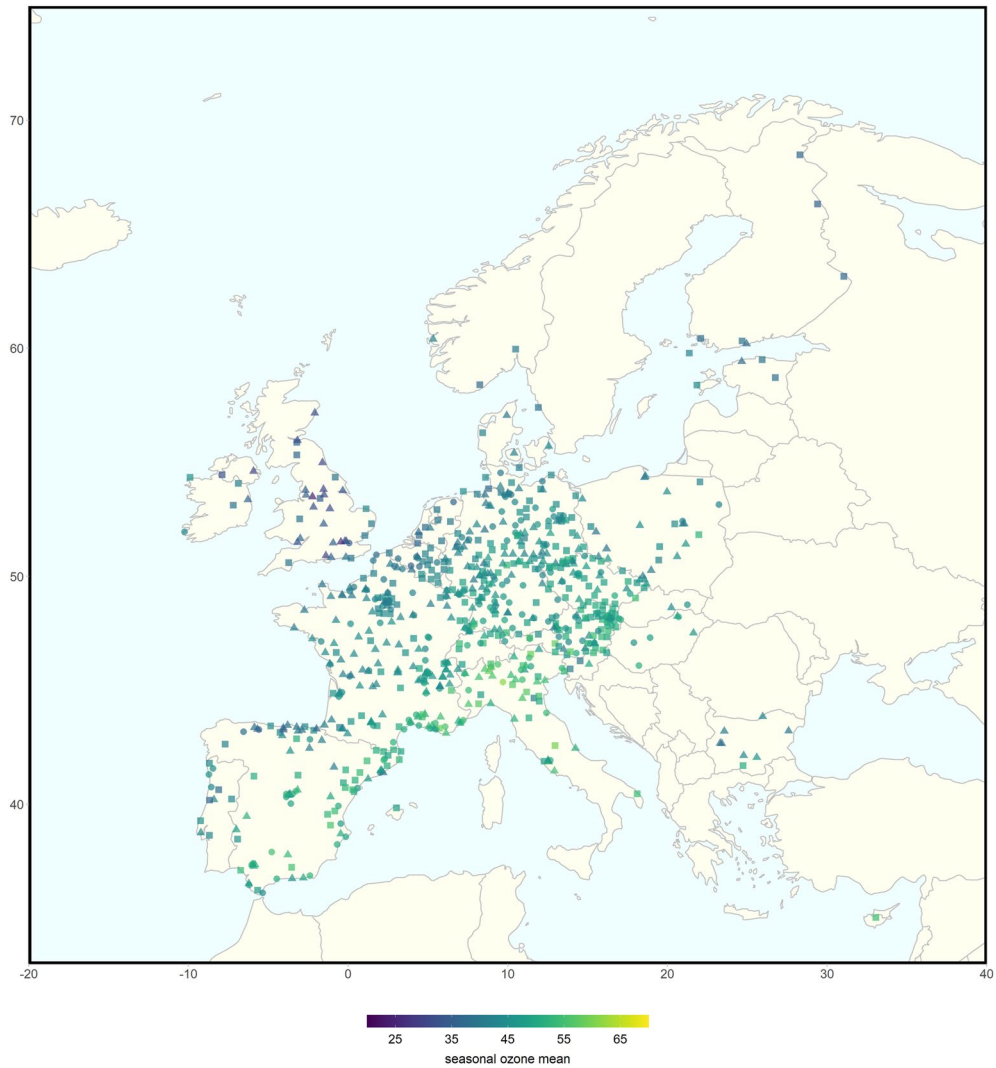

Figure 1. Location of all 798 ozone stations across Europe. Color indicates MDA8 concentrations as mean values [ppbv] in the season April to September for the period 2005–2019 and shape of points indicate the type of each station (triangle = urban, circle = suburban and rectangle = rural).

2. Data and Methods

2.1. Data

2.1.1. Station-Based Data

Hourly Air Quality eReporting O_3 pollution data from the European Environment Agency (EEA, 2017) were extracted. MDA8 as well as MDA1 were calculated and converted from µg/m^3 to ppbv. Stations with data from 2005 to 2019 and having more than 80% complete data per month were considered. To reduce the impact of direct formation and depletion processes next to emission sources of precursors (mainly NO_x), only urban, suburban, and rural background stations were incorporated in the analysis. Hence, no stations in traffic or industrial areas, located near major roads or industrial areas or sources, were included in the analysis. Consequently, we focus in our analysis on stations which are installed to monitor background concentration levels of ozone. These background levels are commonly considered to be representative for a given region and regarded as the average exposure of the general population. In summary, the initial selection led to an ozone station database with 798 background stations. The location of the stations, station type, and mean MDA8 concentrations for April-September 2005-2019 are illustrated in Figure 1. The distribution of the ozone stations per country is given in Table S1 in the Supplementary Material.

2.1.2. Reanalysis Data

The main meteorological drivers of O_3 in the European area comprise solar radiation, air temperature, atmospheric humidity, wind speed and direction, and air pressure (Hertig, 2020; Hertig et al., 2020; Jahn & Hertig, 2021; Otero et al., 2016). The meteorological variables enter the station-specific statistical downscaling models as predictors. Predictor data was retrieved from the ERA5 reanalysis data set of the European Center for Medium-Range Weather Forecasts (Hersbach and Dee, 2016). Data was downloaded for the whole European domain with a $1° \times 1°$ resolution. Variables were extracted at the 850 hPa level and preprocessed to get the following daily predictors for the period 2005 to 2019 (units in brackets): mean air temperature T850 (°C), specific humidity SH850 (g/kg), geopotential heights GH850 (m) as well as wind speed WS850 (m/s) and direction WD850(.), the two latter calculated from the zonal and meridional wind components UW850 (m/s) and VW850 (m/s). Furthermore, surface solar radiation downward SRD (W/m^2) was downloaded. Only variables which carry physically meaningful and relevant information to model MDA8 and MDA1 were included. Atmospheric changes due to climate change were covered by selecting circulation-dynamic, thermo-dynamic, thermal, and radiation-based predictors. The mean of the nine grid boxes covering the area over and around each ozone station location was used to define station-based daily predictor data.

For O_3, the Copernicus Atmosphere Monitoring Service (CAMS) reanalysis from the European Centre for Medium-Range Weather Forecasts (ECMWF) was used (Inness et al., 2019). The CAMS reanalysis contains information about atmospheric composition and is assimilated from satellite retrievals of total column CO, tropospheric column NO_2, aerosol optical depth, and total column, partial column, and profile O_3 retrievals. For our study, we used O_3 data at the 850 hPa level from 2005 to 2019 at a monthly resolution (O3month), spatially interpolated to $1° \times 1°$ grid boxes of latitude and longitude.

In preparation for our study, we analyzed two other tropospheric O_3 reanalysis products in addition to the CAMS reanalysis: the O_3 data from the ERA5 reanalysis and the TCR-2 reanalysis from Miyazaki et al. (2020), the latter showing good results in the O_3 evaluation study from Park et al. (2020) for East Asia. A correlation analysis with the O_3 station data showed that the CAMS reanalysis had the best agreement with the observational data over most of our study area. The ERA5 O_3 data could reflect the spatial patterns of O_3 distribution over Europe to a very limited extent.

2.1.3. Earth System Model (ESM) Data

To estimate tropospheric O_3 concentrations over Europe on a station basis for the mid (2041–2060) and late 21st century (2081–2100), all available data from the Coupled Model Intercomparison Project Phase 6 (CMIP6, Eyring et al., 2016) were used. Thus, the number of ESMs depended on the availability of data at the time of this publication. The atmospheric predictors (T850, SRD, GH850, UW850, VW850, SH850) were used at daily resolution. O_3 data at the 850 hPa level was only available at a monthly resolution. Data was extracted for the historical period and the two scenarios SSP2-4.5 and SSP3-7.0 (for a description of the SSPs see e.g., O'Neill et al., 2014). Seven ESMs could be used, which can be divided into two groups: models with O_3 prescribed from a data set (BCC-CSM2-MR, IPSL-CM6A-LR, MPI-ESM1-2-HR and MPI-ESM1-2-LR), and models which interactively compute the atmospheric chemistry by themselves (CESM2-WACCM, MRI-ESM2-0 and UKESM1-0-LL). Note that IPSL-CM6A-LR did not provide predictor data for 21 O_3 station locations in the alpine region.

For those ESMs that do not have an interactive chemistry model themselves, an O_3 dataset is provided as part of CMIP6. This CMIP6 O_3 dataset (Checa-Garcia, 2018) uses simulations of the CMAM and CESM-WACCM models that are part of the Chemistry-Climate Model Initiative (Eyring et al., 2013; Morgenstern et al., 2017). Thus, a dataset is available that contains a complete three-dimensional field for both the stratosphere and troposphere, spanning from pre-industrial times to the present, and to the end of the 21st century under the various SSP scenarios (Keeble et al., 2020; O'Neill et al., 2015). This is a key difference from the O_3 data used for CMIP5, which were based on stratospheric O_3 values from a combination of model and observational data between the 1970s and 2011, extended into the past and future under assumptions about changes in stratospheric chlorine and the 11-year solar cycle (Cionni et al., 2011). Note that for the models which do not model the O_3 concentrations interactively, the O_3 output is identical.

One of the three models with interactive chemical modeling is the CESM2-WACCM version of the Community Earth System Model version 2 (CESM2), which uses the Model for Ozone and Related chemical Tracers family of chemical mechanisms, covering the troposphere, stratosphere, mesosphere, and lower thermosphere

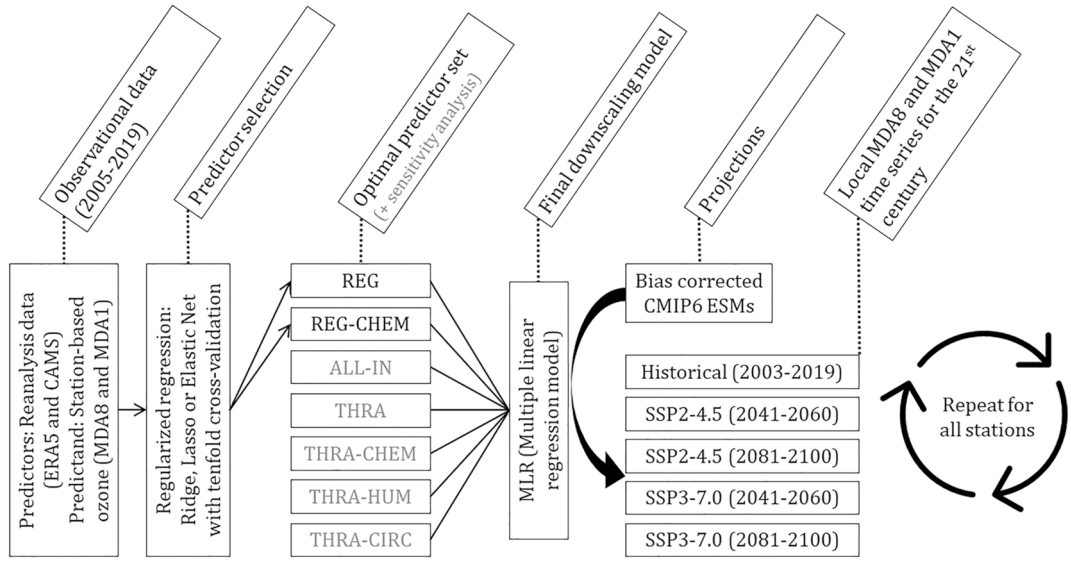

Figure 2. Schematic diagram of the downscaling methodology.

(Emmons et al., 2020; Gettelman et al., 2019). The MRI-ESM2-0 model includes an interactive chemistry model, called MRI-CCM2.1 module, which simulates the distribution and evolution of O_3 and other trace gases in the troposphere and middle atmosphere (Yukimoto et al., 2019). The third model in our study with interactive chemical modeling, UKESM1-0-LL, uses the UKCA model (United Kingdom Chemistry and Aerosol model) with unified stratospheric-tropospheric chemistry and close coupling between chemistry and aerosols (Sellar et al., 2019).

All ESM simulation output were re-gridded onto 1 × 1 spatial resolution to align with the reanalysis data. The units were converted accordingly, and WS850 and WD850 were calculated from UW850 and VW850.

2.2. Methods

Figure 2 provides a schematic diagram illustrating the methodology used in this study. Detailed information is given in the following subsections.

2.2.1. Statistical Downscaling Models and Sensitivity Analysis

Statistical downscaling models were built for the O_3 season from April to September, the time of the year when the highest O_3 concentrations occur over Europe. The models were used to establish station-specific relationships of meteorological and chemical predictors with MDA8 and MDA1. Daily time series of all selected and preprocessed meteorological variables as well as the monthly resolved O_3 data were used as initial predictors. Before entering any statistical model building process, each predictor was standardized by subtracting from the values the respective mean and then dividing it by the respective standard deviation. A station-specific modeling approach was chosen with an individualized predictor screening to generate site-specific optimums. Predictor selection was conducted using regularization with varying shrinkage methods (Hastie et al., 2009). Multiple Linear Regression (MLR), representing an effective tool to study the impact of predictors on the mean of the response variable, was used to model the relationship between all selected predictors and MDA8/MDA1.

A sensitivity analysis was conducted analyzing further varying settings of the statistical downscaling models, mainly differing in terms of the predictor variables used, to quantify more detailed the influence of specific predictors on the downscaling and projection results. A special focus was given to O3month as an emission-related predictor. All statistical downscaling models were evaluated by considering various model fit and performance metrics. Most of the data preparation and analysis including predictor selection, and hence regularized regression, model building as well as projections, were conducted using the R programming language (R version 4.1.0, IDE RStudio).

2.2.1.1. Predictor Selection and Main Models

Predictor selection was conducted using regularization with varying shrinkage methods to generate site-specific optimum predictor sets. Ridge, Lasso and Elastic Net regularized regression (Hastie et al., 2009) were considered. In this regard, ridge regression is considered best if all variables are useful as it shrinks all parameters but does not remove them. Lasso regression will be the best choice if the initial predictor set contains a lot of useless variables as these are excluded from the equation and simpler and easy to interpret results are generated. Elastic Net combines Lasso and Ridge regression penalty and thus combines the strengths of both options. It is the preferred option if, as in our case, there exist strong correlations between predictors. Elastic Net groups and shrinks the parameters associated with the correlated variables and leaves them in the equation or removes them all at once. The R package glmnet was applied for regularized regression and predictor selection. In an upstream tenfold cross-validation, the station-specific optimum value for alpha was defined. Lambda.1se instead of lambda.min was used as the former results in models with in general fewer parameters. This was preferred in our analysis since we aimed to identify the main drivers of O_3 at each site. As a result, MLR models based on site-specific predictor optimums were generated to establish the relationships between MDA8/MDA1 and their main drivers and to use these statistical downscaling models for station-based projections.

2.2.1.2. Model Fit and Performance

Fit and performance of the MLR statistical downscaling models were evaluated using two typical performance metrics: the adjusted coefficient of determination R^2 and the Root Mean Square Error RMSE. The fit and performance evaluation of the final station-based MLR models was embedded in its own tenfold cross-validation procedure. To judge overall performance, mean values were calculated for both metrics, and averaged for the 10 calibration (cal) and 10 validation (val) periods, respectively. Furthermore, adjusted R^2 was extracted from the final MLR models (R^2). Stations showing a sufficient model fit and performance were considered for further analysis and projections. In this regard, stations showing an R^2 value of at least 0.25 were selected.

2.2.1.3. Main Drivers

The final station-based statistical MLR downscaling models represent the results of the multistep predictor selection and model building approach that aimed to define the linkage between MDA8/MDA1 and their optimum predictor set. The impact of a predictor on the target variable can in general be interpreted in terms of the magnitude and the sign of the predictor's standardized regression coefficient. Thus, predictor variables can be ranked by importance to identify station-based main drivers. The most important drivers of MDA8/MDA1 were identified for each analyzed O_3 station location. This allows for the investigation of possible station-type- and region-specific differences and sensitivities under current and future climatic conditions.

2.2.1.4. Sensitivity Analysis

In addition to the station-specific optimal choice of predictors, different predictor configurations were tested to investigate the influence of specific predictor groups on downscaled MDA8/MDA1. The main model-building process and its framework described in the previous sections are used as a baseline and the following settings, differing mainly in terms of the predictors used, were evaluated.

1. Regularized (REG): Statistical models using the station-specific optimum predictor set from the regularized predictor selection described in Section 2.2.1.1.
2. Regularized and chemistry (REG-CHEM): Statistical models using the station-specific optimum predictor set from the predictor selection described in Section 2.2.1.1, with always setting by default O3month as a predictor to ensure the inclusion of the emission-related information content.
3. All in (ALL-IN): Statistical models with all predictor variables included (O3month, T850, SH850, GH850, SRD, WS850, and WD850). Thus, no regularization was used, and all predictor variables were kept.
4. Thermal, radiative (THRA): Statistical models using only T850 and SRD to assess the isolated role of these most important meteorological predictors for ground-level O_3.
5. Thermal, radiative and chemistry (THRA-CHEM): Statistical models including T850, SRD, and O3month to measure the combined impact of the essential meteorological predictors and of emission changes.
6. Thermal, radiative and thermo-dynamic (THRA-HUM): Statistical models using the predictors T850, SRD, and SH850 to assess the additional influence of atmospheric humidity.

7. Thermal, radiative and circulation-dynamic (THRA-CIRC): Statistical models using T850, SRD, GH850 as well as WS850 and WD850. This setting was used to identify the role of atmospheric circulation dynamics on O_3 and its change.

For the first two options, a feature selection based on regularized regression in accordance with the described approach for the original model-building process was conducted to generate station-specific optimum predictor sets. For all other options, station-based MLR models were directly generated with the respective predictor sets. Model fit and performance were evaluated for all settings. The relationships of the predictors with MDA8/MDA1 established in the observation period for each option were subsequently used to assess the response of MDA8/MDA1 to future changes in the respective predictor sets. Consequently, the projections of all varying settings from the sensitivity analysis could be compared.

2.2.2. Bias Correction of ESM Data

Prior to using the ESM data within the statistical downscaling models, they were bias corrected using a univariate quantile mapping method (using the R package MBC; Cannon, 2018). The atmospheric variables from the ERA5 reanalysis and the CAMS reanalysis for O_3 were used as the correction basis. Since the historical CMIP6 runs ended in 2014, they were extended to 2019 using data from scenario SSP3-7.0. Bias correction was performed for each grid point in the entire European study area. Like Cannon (2018) three consecutive months in a shifting window were used, and only the central month was kept. If less than 20% of the days per grid point and 3-month window were missing, the bias correction was calculated, otherwise, the grid point was flagged as a missing value. Bias correction was calculated separately for each scenario (SSP2-4.5 and SSP3-7.0), and for each time slice (2041–2060 and 2081–2100) using the reference period 2003–2019.

2.2.3. Statistical Projections

Statistical projections under climate and emission changes were generated by replacing reanalysis predictor data with bias-corrected ESM output in the MLR models. Time slice differences under the two selected scenarios SSP2-4.5 and SSP3-7.0 for the future periods 2041–2060 and 2081–2100 compared with the historical period 2003–2019 were used to illustrate the local O_3 changes under future climate change.

3. Results

Results are shown for MDA8 since most health-relevant guidelines and threshold values refer to this value. Results for MDA1 are similar and can be found in the Supplement. Note that differing numbers of ozone measuring stations were included in the various analyses, due to the different statistical downscaling performances of the stations under different predictand and predictor sets.

3.1. Statistical Downscaling Model Performance

The performance of the MLR models was evaluated using a tenfold cross-validation. To judge and compare the different predictor sets performance was calculated as mean values over the 10 calibration and 10 validation periods as well as for the final MLR models using the whole time series. Stations showing a final R^2 value of at least 0.25 were selected. This led to a reduction of the original 798 stations to 716 stations for the predictor set REG, and to 720 stations for both predictor sets ALL-IN and REG-CHEM. The criterion for exclusion of 0.25 led to 682 stations for the set THRA-CHEM, 671 stations for the set THRA-CIRC, 635 stations for the set THRA-HUM, and down to 623 stations for the predictor set THRA. Thus, with the predictor sets REG, REG-CHEM, and ALL-IN the vast majority of stations could be assessed, whereas the reductions of the number of assessable stations for the other predictor sets point to missing predictor information in these sets to adequately capture ozone variation in some regions.

The over all predictor combinations shared set of 623 stations was used for further comparison of the predictor performance in the observational period. Interestingly, all predictor sets yielded a comparable performance for the stations considered (Table 1). Thus, the predictors T850 and SRD (THRA) could already explain a large fraction of the daily MDA8 variability. The inclusion of other meteorological and chemical information did not substantially increase statistical performance. However, keep in mind that using the sets REG, REG-CHEM, and ALL-IN resulted in almost 100 stations more which could be assessed compared

Table 1
Statistical Downscaling Model Performance of the Different Predictor Sets Based on 623 Stations

Option	Predictors	R^2	R2.cal	R2.val	RMSE.cal	RMSE.val
1	REG	0.52	0.46	0.46	8.09	8.11
2	REG-CHEM	0.52	0.46	0.46	8.09	8.10
3	ALL-IN	0.52	0.48	0.47	8.01	8.03
4	THRA	0.46	0.40	0.40	8.55	8.55
5	THRA-CHEM	0.49	0.44	0.44	8.22	8.23
6	THRA-HUM	0.46	0.41	0.41	8.50	8.50
7	THRA-CIRC	0.48	0.43	0.42	8.37	8.38

Note. Given are the adjusted coefficient of determination R^2 and the Root Mean Square Error RMSE [in ppbv] as means over the 10 calibration (cal) and validation (val) periods as well as final R^2 for MDA8.

to the set THRA, indicating that THRA is a decisive predictor set only for some regions of Europe. Thus, stations with high model performance using THRA concentrated mainly over Central Europe, while for southern, western, and north-western European stations further predictors were needed.

For further analysis, we chose the predictor set REG-CHEM since it combines a good model performance, robust downscaling models derived from regularization and assures the inclusion of information on meteorology as well as emission changes in the projections. Thus, the results of the predictor set REG-CHEM for MDA8 from April to September at 720 measuring stations are discussed in the following.

3.2. Main Predictors

Specific predictors emerged as prominent factors governing daily MDA8 variability. Figure 3 shows the three most important predictors of the predictor set REG-CHEM for MDA8 in the O_3 season from April to September. In the majority of the statistical downscaling models O3month, SRD and T850 were selected. For most of the stations in Central and southern Europe, the most important predictor was either T850 or SRD, whereas for stations in northern Europe O3month was selected. For some maritime stations in southern Europe SH850 or WS850 constituted the most important predictor, indicating the role of meso-scale land-sea circulation systems (Figure 3, left). In Central Europe, the second most important predictor was still a thermal or radiative one. However, for many stations, O3month was selected. In southern Europe, SH850 emerged as an important driving factor (Figure 3, middle). O3month constituted the third most important predictor at most stations in Central Europe, but SH850 played a role for stations in eastern Europe as well as WD850 in western Europe (Figure 3, right), the latter pointing to the importance of inflow of polluted air masses.

Overall, for all the 720 stations O3month was selected as a predictor, for 715 stations SRD and 686 stations T850. All three predictors had strong positive regression coefficients in the MLR models. SH850 was included in the statistical downscaling models only for 339 stations, with a medium strong, negative coefficient. WS850 was selected for 613 stations, WD850 for 513 stations, and GH850 for 308 stations, all of them with relatively low, negative coefficients. In summary, O3month plays a decisive role in the statistical downscaling models to explain daily MDA8 variability, implying that the emission-related predictor information is important and enhances statistical downscaling model quality. The selection of SRD and T850 as prominent meteorological predictors reflects the photochemical, temperature-dependent build-up of O_3. SH850 only played an important role in some European regions, whereas the frequent selection of the negative association of WS850 with MDA8 points to in-situ O_3 formation under low-flow conditions.

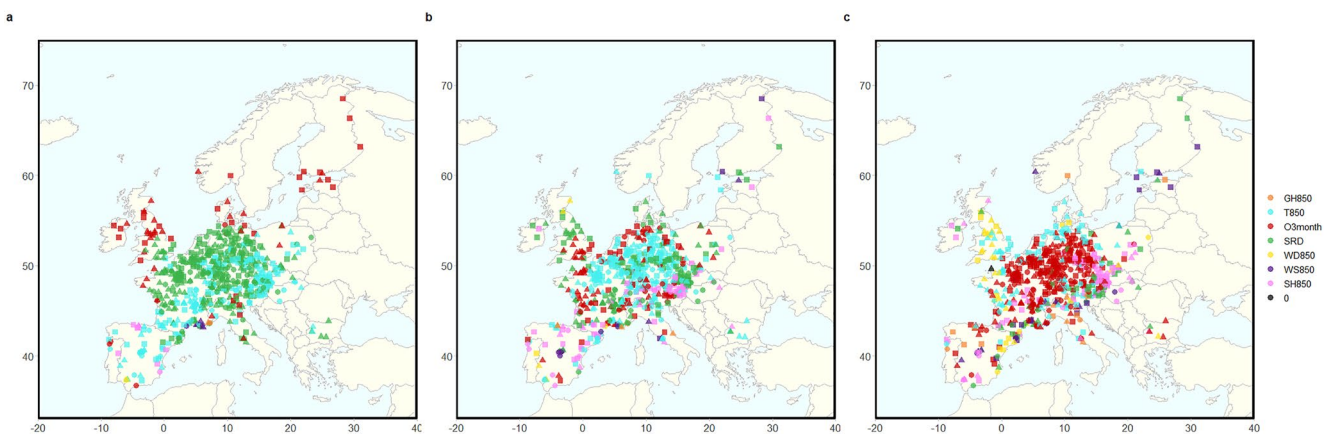

Figure 3. Predictors chosen in the multiple linear regression models using the predictor set REG-CHEM. From left to right: most important predictor, second most important predictor, third most important predictor for MDA8 (0 = no third predictor variable).

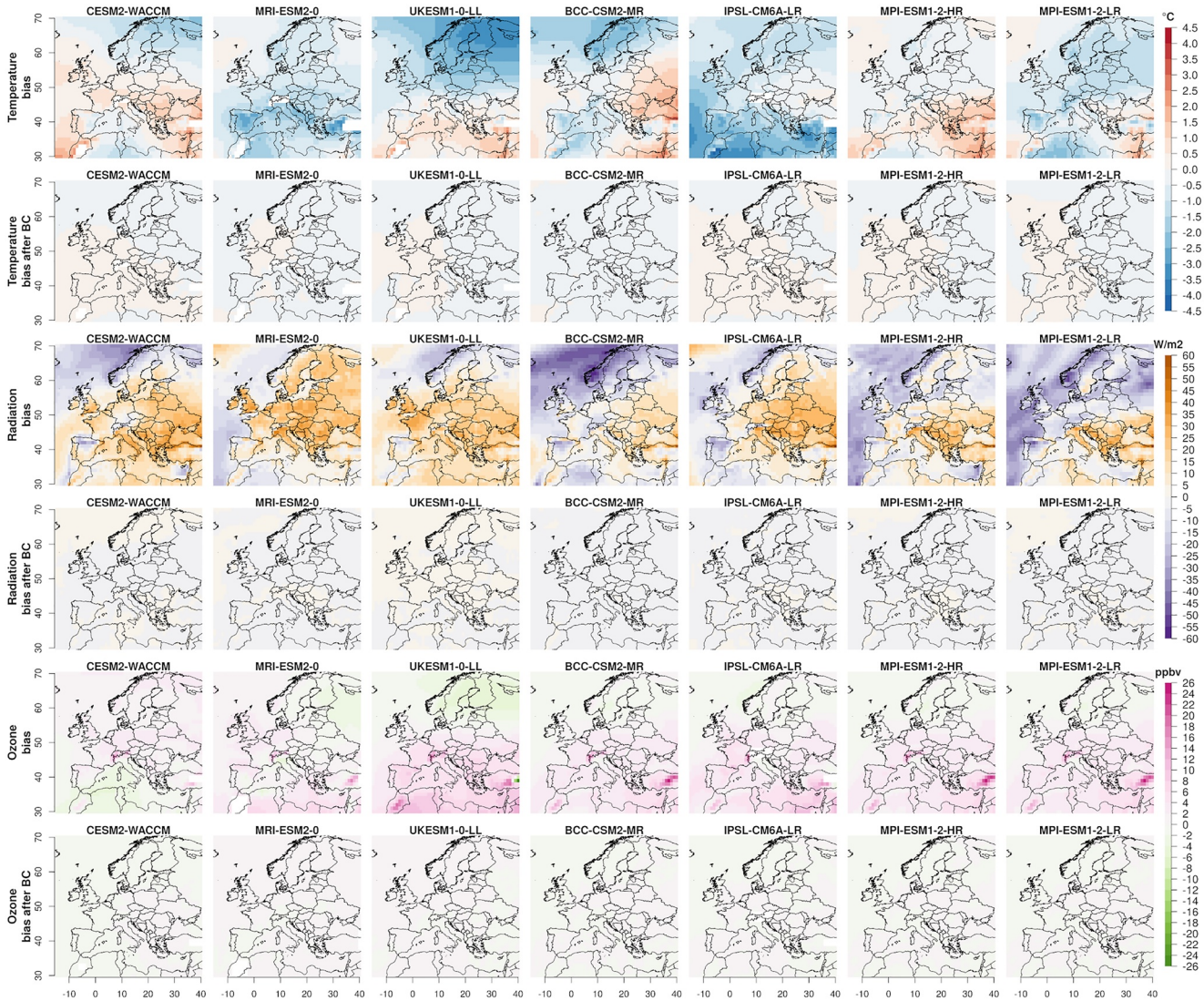

Figure 4. Biases before and after bias correction (BC) of the three most important predictors T850 [°C], SRD [W/m^2], and O3month at 850 hPa [ppbv] for the O_3 peak season April-September in the historical period 2003–2019. Bias correction reference for T850 and SRD: ERA5 reanalysis, for O3month: Copernicus Atmosphere Monitoring Service reanalysis.

3.3. Bias Correction

With the bias correction method QDM, the partly strong model deviations of the individual climate models could be satisfactorily adjusted. Figure 4 gives the raw and the adjusted output for the three most important predictors T850, SRD as well as O3month for the months of April to September in the historical period 2003–2019.

The raw output of the seven ESMs showed for T850 and SRD different amounts of biases, which were spatially different, compared to the observation-based ERA5 reanalysis. For T850 deviations of up to ±4.5°C were found, for SRD up to ±60 W/m^2 at individual locations.

O3month, shown at the bottom of Figure 4, had compared to the CAMS reanalysis an almost identical pattern for the models with no interactive chemistry modeling, with slightly too high O_3 values over Europe, strongest over the Alps and Turkey and small negative biases in the European North. The ESM UKESM1-0-LL with active chemistry modeling had the highest biases in the summertime, especially in the Mediterranean area.

The bias correction satisfactorily minimized model errors for all predictors. Model errors after bias correction were within ±0.35°C for T850, ±4 W/m^2 for SRD, and ±0.6 ppbv for O3month (see again Figure 4). The

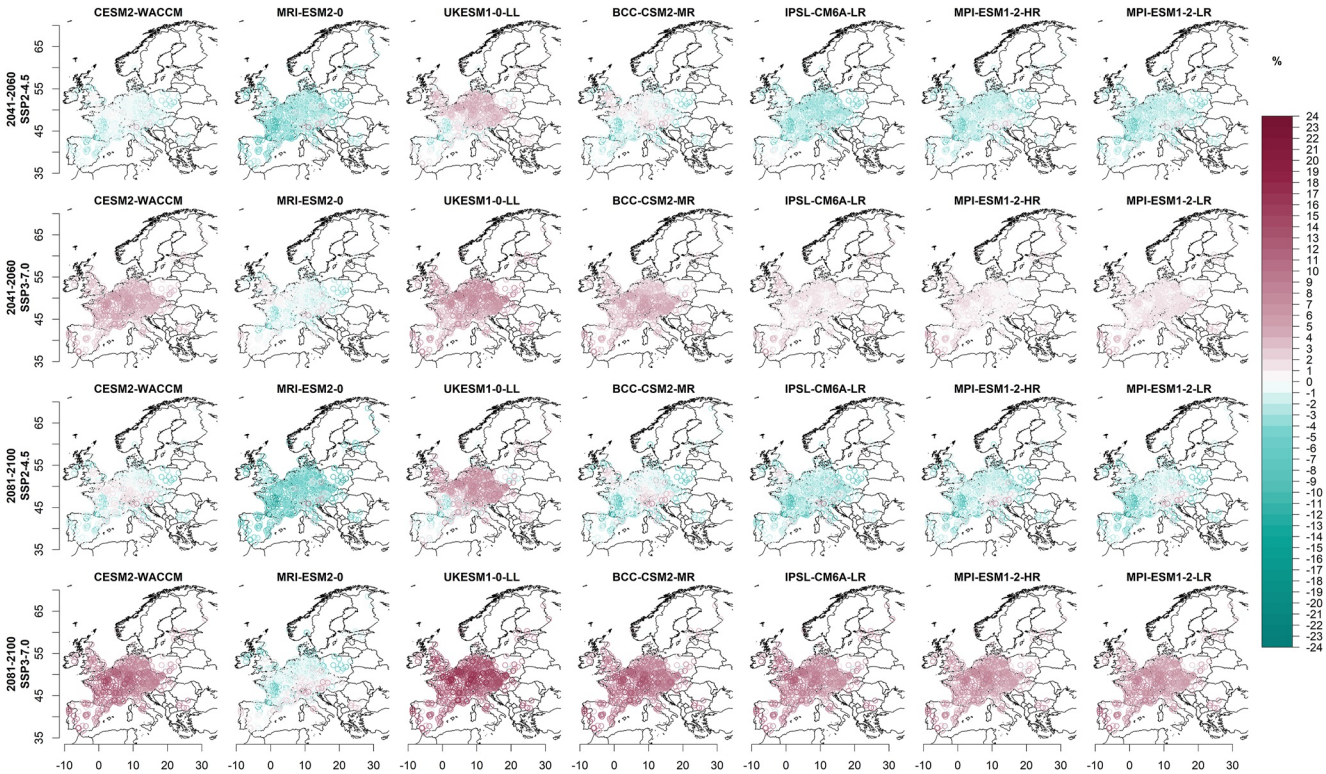

Figure 5. Change signals [%] of downscaled ground-level MDA8 concentrations at 720 O_3 measuring stations across Europe for all investigated CMIP6 earth system models. Shown are results for SSP2-4.5, SSP3-7.0, and the years 2041–2060 and 2081–2100.

remaining errors for the other predictors were also very small. We also examined whether the bias correction modified the change signals in future scenarios. It was found that the change signals were not affected by the bias correction (not shown here).

3.4. Statistical Projections Under Climate and Emission Changes

The results in Sections 3.4.1 and 3.4.2 are based on the statistical models using the station-specific optimum predictor set from the predictor selection described in Section 2.2.1.1. with always setting by default O3month as predictor (REG-CHEM). Results are given for MDA8 from April to September for the 720 stations which could be assessed using this predictor set. A discussion about the role of the other predictor sets on projected MDA8 is given in Section 3.4.3.

3.4.1. Individual ESM Results

Figure 5 shows the downscaled MDA8 changes for the single ESM ensemble members. The two top rows show the downscaled near future changes as average in the period 2041–2060, and the bottom rows illustrate the change signals at the end of the 21st century (2081–2100) for each station. The two scenarios SSP2-4.5 and SSP3-7.0 are shown one above the other in each case. A general trend can be seen from the figures of the individual models. SSP2-4.5 led to decreasing MDA8 concentrations for almost all ESMs, while the more pessimistic scenario SSP3-7.0 projected increasing concentrations over Europe, especially at the end of the 21st century. However, there were exceptions for the individual ESMs. For example, the MRI-ESM2-0 model consistently projected almost only negative trends even in the pessimistic scenario, while the UKESM1-0-LL model simulated predominantly increases in ground-level MDA8 concentrations even in the more optimistic scenario SSP2-4.5. At the end of the century, these were particularly strong.

To understand the reasons for the differing MDA8 projection results from the individual climate models, the change signals of the three most important predictors are examined below.

Figure 6 shows the projected T850 progression in the 21st century for the O_3 peak season from April to September. For each individual model of the CMIP6 ensemble, the change signals of the bias-corrected data

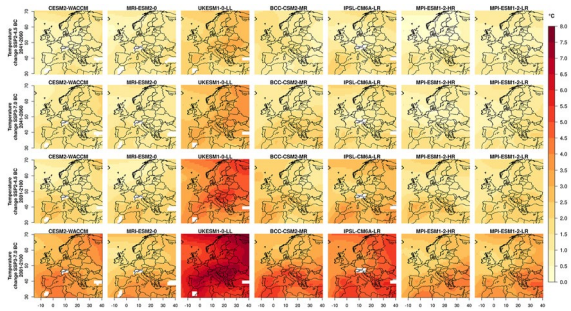

Figure 6. Change signals of bias corrected (BC) CMIP6 T850 [°C] for the O_3 peak season April-September for the years 2041–2060 (top rows) and 2081–2100 (bottom rows) compared to 2003–2019 and the two scenarios SSP2-4.5 and SSP3-7.0.

are given for the periods 2041–2060 and 2081–2100, and for the two scenarios SSP2-4.5 and SSP3-7.0. In the mid-21st century, the two scenarios were not yet very different from each other and projected a temperature increase of about 2°C, compared to the historical period 2003–2019 (shown in the top two rows of Figure 6). However, toward the end of the 21st century, especially scenario SSP3-7.0 yielded a further temperature increase over Europe. The differences between the individual climate models increased as well. In all projections, the UKESM1-0-LL model stood out with particularly strong warming trends, reaching up to 8°C in the more pessimistic scenario.

The change in SRD (Figure 7) shows a very heterogeneous picture of increasing and decreasing tendencies. With some exceptions, over the European land areas, an increase of radiation was projected and over the sea decreasing radiation sums. UKESM1-0-LL stood out, again, with a strongly increasing SRD north of the Mediterranean Sea at the end of the 21st century. The two very strong positive trends of T850 and SRD in the UKESM1-0-LL model induced the consistently increasing, partly very pronounced O_3 concentrations, as seen in Figure 5.

Figure 8 shows the projected changes in O3month from the CMIP6 ensemble. As expected, the four models with prescribed O_3 concentrations from the Chemistry-Climate Model Initiative O_3 data set yielded an almost identical O3month projection. The CESM2-WACCM model exhibited a very similar picture because it is part of the Chemistry-Climate Model Initiative and thus provided O_3 data for the other models without active chemistry modeling. It is noticeable that the two scenarios were fundamentally different from each other in their change signals. Scenario SSP2-4.5 led to decreasing O3month concentrations, while scenario SSP3-7.0 resulted in O3month increases. An exception was the MRI-ESM2-0 model with active chemistry, which also simulated predominantly decreasing O3month levels over Europe for scenario SSP3-7.0. This was reflected then in the results of the downscaled station MDA8 from Figure 5, where negative MDA8 concentrations occurred despite increasing T850 and SRD.

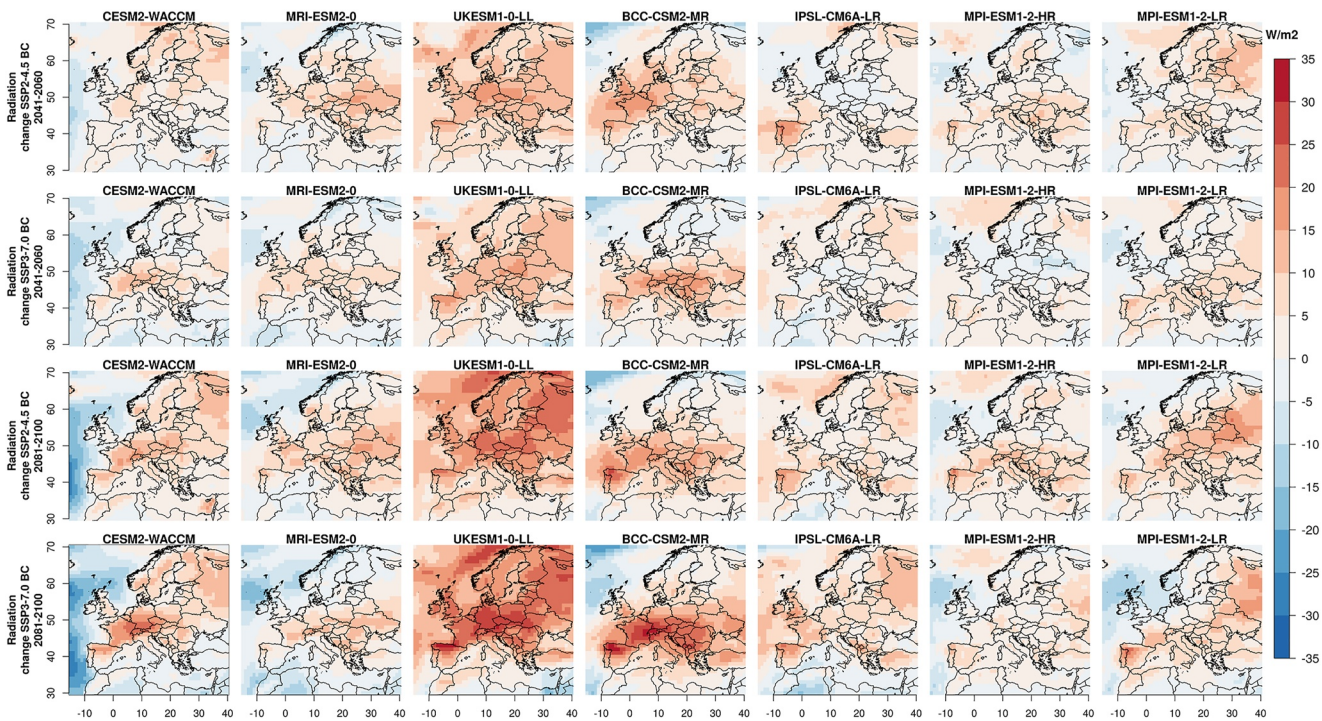

Figure 7. As Figure 6, but for SRD [W/m^2].

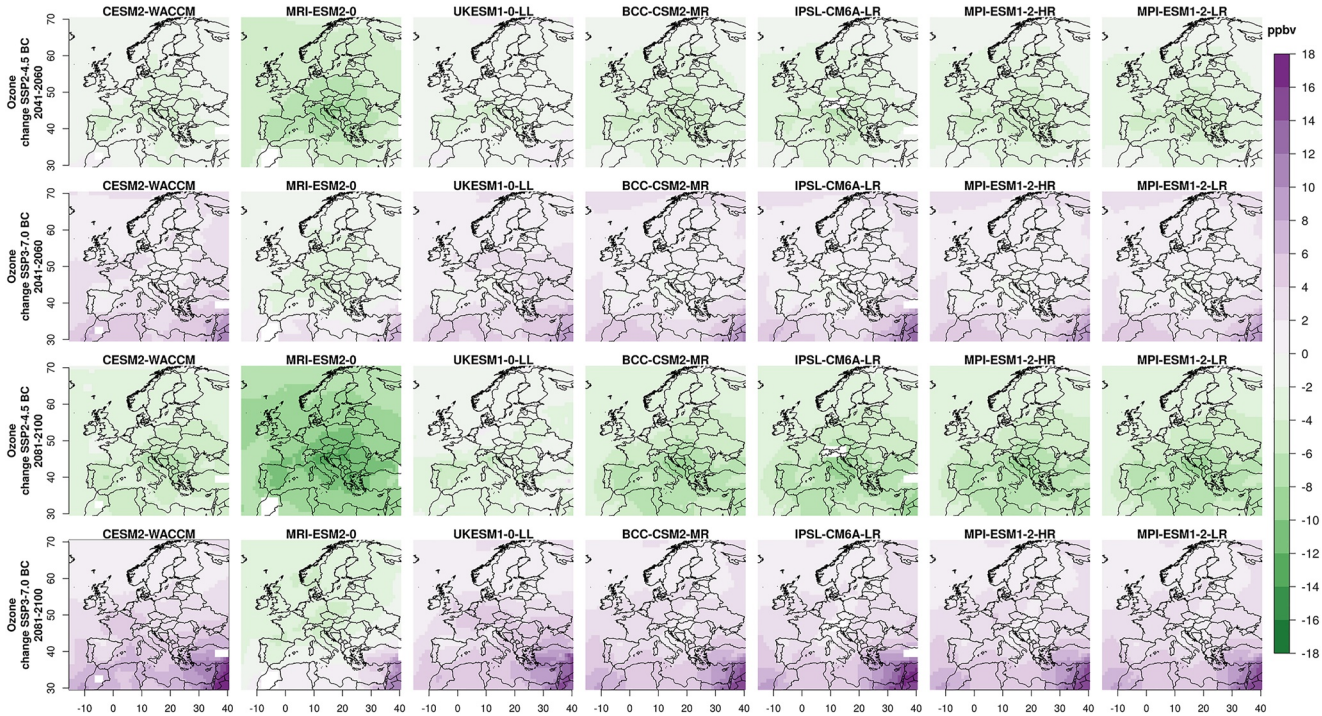

Figure 8. As Figure 6, but for O3month [ppbv].

The individual analysis of the climate models showed that a distinction between models with and without an active chemical component was not decisive for the downscaling results, since the other predictors could determine the future MDA8 concentrations, as shown for the model UKESM1-0-LL.

3.4.2. Multi-Model Mean

Figure 9 shows the projection results of the CMIP6 models as Multi-Model Mean (MMM). In a world that follows the SSP2 scenario, called "Middle of the Road", in which social, economic, and technological trends do not shift markedly from historical patterns and climate protection measures are adopted, the MDA8 concentrations over Europe were projected to decrease. The mean decrease over all stations was −1.75% at the mid-century and about −1.94% at the end of the century. Thus, most of the MDA8 reduction might already be achieved by mid-century, with a small additional decrease occurring in subsequent years. The negative trend was not simulated for the center of the United Kingdom and at some stations in the Alps. Some weak positive trends were observed here. The strongest decreases occurred in western France, where all CMIP6 models in our ensemble showed a consistent decline. However, if the world develops according to scenario SSP3, which is characterized by regional rivalries and a revival of nationalism, pushing global issues into the background, causing severe environmental degradation, and making mitigation and adaptation to climate change more difficult, projected ground-level MDA8 concentrations look very different. On average, the model ensemble projected a +1.82% increase in MDA8 concentrations across Europe by the mid-century, increasing considerably to +5.49% by 2100. By the end of the 21st century, increases of up to 16% were projected at some European stations. All regions were affected by strong increases, no clear regional differences could be identified.

3.4.3. Impact of Different Predictor Settings

The model performance in the observational period is just one indicator for the overall statistical downscaling model performance since it does not necessarily guarantee the suitability of a predictor set to assess future MDA8 changes. While a good observational model performance hints at the successful capture of the relationships under current climate and environmental conditions, for projections a predictor set has also to carry the relevant signals of climate and emission changes that impact future MDA8 progression.

As for the comparison of the model performance in Section 3.1, the common set of 623 stations was used, again. The comparison of the impact of different predictor settings on future MDA8 concentrations showed that there

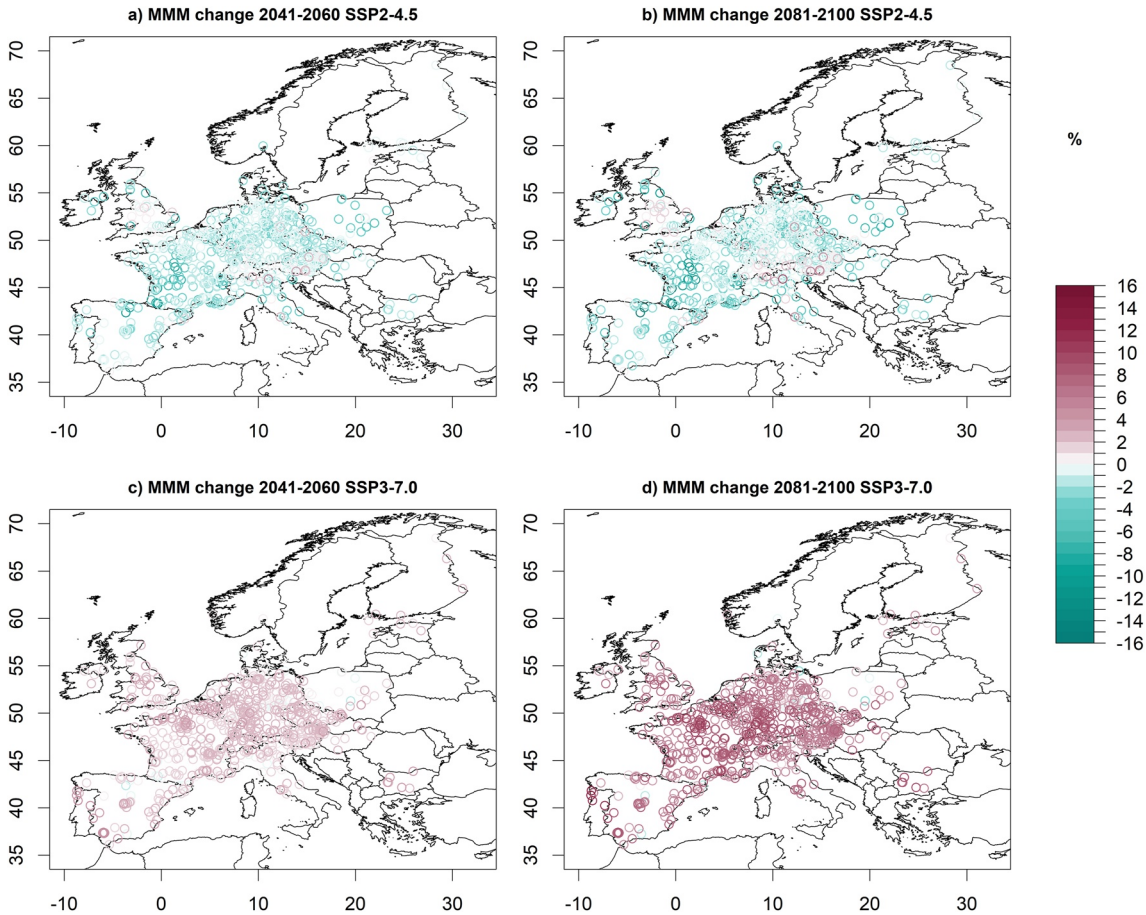

Figure 9. Multi-Model Mean (MMM) change [%] of downscaled ground-level MDA8 for 720 O_3 measuring stations over Europe. Shown are results for SSP2-4.5, SSP3-7.0 and the years 2041–2060 and 2081–2100 compared to the historical period 2003–2019.

Table 2
Multi-Model Mean (MMM) Change [%] of Downscaled Ground-Level MDA8 as Mean Over 623 O_3 Measuring Stations Over Europe Using Different Predictor Sets

Option	Predictors	SSP2-4.5 2041–2060 minus 2003–2019	SSP2-4.5 2081–2100 minus 2003–2019	SSP3-7.0 2041–2060 minus 2003–2019	SSP3-7.0 2081–2100 minus 2003–2019
1	REG	−1.76	−1.88	1.80	5.69
2	REG-CHEM	−1.75	−1.86	1.79	5.57
3	ALL-IN	−1.81	−1.92	1.81	5.64
4	THRA	1.64	2.78	1.64	3.81
5	THRA-CHEM	−1.69	−1.78	1.91	5.79
6	THRA-HUM	1.60	2.74	1.63	3.79
7	THRA-CIRC	1.54	2.62	1.53	3.63

Note. Shown are results for SSP2-4.5, SSP3-7.0 and the years 2041-2060 and 2081–2100 compared to the historical period 2003–2019.

are considerable differences when using only meteorological information compared to the inclusion of information on emission changes.

Table 2 shows the downscaled MDA8 changes as mean across the 623 measuring stations and the different ESMs. Using only meteorological predictors (THRA, THRA-HUM, THRA-CIRC) resulted in projected increases of MDA8 for both scenarios and time slices. At the middle of the century MDA8 increased by about 1.5% for both scenarios and at the end of the century increased by 2.6%–2.8% under SSP2-4.5, and by 3.6%–3.8% under SSP3-7.0. Despite the differing importance of the meteorological predictors across Europe to explain MDA8 variability (see Section 3.2) and their differing climate change signals, the increases were visible for almost all stations across the European area with no specific regional pattern (not shown). In contrast, the inclusion of information on emission changes (REG, REG-CHEM, ALL-IN, THRA-CHEM) modified the sign of change under SSP2-4.5 scenario assumptions, yielding decreases of about −1.7% to −1.9% during the 21st century. This points to the effectiveness of the emission reductions of the precursor gases which outweigh the O_3 increases caused by climate change. Under the more pessimistic SSP3-7.0 scenario, which does not include effective emission reductions, MDA8 was projected to increase. These increases were even stronger compared to the assessments using only meteorological information and amounted to 1.8%–1.9% at the

mid-century and 5.6%–5.8% at the end of the century, implying that there is a combined "climate change and emission penalty" on future O_3 concentrations. For all emission-including predictor sets the spatial pattern of MDA8 changes followed the patterns illustrated in Figure 9.

4. Discussion

A PP downscaling approach was adopted to assess local ground-level MDA8 and MDA1 concentrations in the European area. Despite many advancements in the current ESMs, there is still the need for downscaling to provide highly-resolved climate change information. In this regard, knowledge about local future ground-level O_3 concentrations is of particular interest to protect human health and to design corresponding mitigation and adaptation strategies. According to Liu et al. (2022) methods to obtain O_3 estimates can be divided into regional and global chemical transport models, statistical models, geostatistical data fusion, and machine learning models. While dynamical climate chemistry models are a useful tool to assess future O_3, they come with the drawback of being computationally expensive and thus the number of model runs is limited. Machine learning models have a very good performance, but they are generally characterized by low interpretability (Liu et al., 2022). Statistical models, which are very flexible, computationally inexpensive, and easy to interpret, have so far used meteorological information to assess future O_3 concentrations, neglecting information about the precursor emissions. In this study, this shortcoming was addressed by incorporating information about the precursor emissions into the statistical downscaling models. This was done not by using NO_x or VOCs directly as predictors, but via the inclusion of O_3 as a predictor, thus providing information about their final impact, including emission, transport, transformation, and deposition processes, as assessed in the reanalysis and climate model products.

The statistical downscaling models were calibrated and validated in the observational period using a regularization approach to carefully capture the predictors-predictand relationships. The selection of the predictors and their association with MDA8/MDA1 in the statistical models was in accordance with previous studies (e.g., Jahn & Hertig, 2021, 2022; Otero et al., 2016). Model performance was assessed using R^2 and RMSE. The overall assessable stations mean model performance amounted to about 50% explained variance (R^2 of approx. 0.5), which shows that local MDA8/MDA1 concentrations are not only governed by larger-scale factors, but also by local characteristics, which cannot be captured in such a downscaling approach. Furthermore, the model performance varied across the European area, pointing to a differing sensitivity of local MDA8/MDA1 variability to the larger-scale forcings, confirming the findings of Jahn and Hertig (2022) who showed that there exists a strong connection between ground-level O_3 to the larger-scale meteorology in Central Europe, whereas for northern and southern Europe other processes must be considered as well.

There has been large progress in chemistry-climate modeling leading to an advanced representation of O_3, but models still show partly large biases and considerable uncertainties exist (Young et al., 2018). The output of seven CMIP6 models was used in the current study, with three of them having interactive chemical modeling, while the other four used prescribed O_3 from the CMIP6 ozone dataset. However, statistically downscaled future MDA8/MDA1, assessed using the CMIP6 model output variables as predictors, could not be divided into two groups, interactive versus prescribed. Rather, the differences between the individual models across all models were dominant. The use of MRI-ESM2-0 output as predictors in the predictor set REG-CHEM yielded mainly negative MDA8/MDA1 trends, whereas UKESM1-0-LL output led to increases of MDA8/MDA1, with the other models providing assessment results in between. A general feature across almost all ESMs was decreasing MDA8/MDA1 concentrations under the SSP2-4.5 scenario, while the use of the more pessimistic scenario SSP3-7.0 led to increasing concentrations over Europe, especially at the end of the 21st century.

This general picture emerged when using meteorological factors as well as information about emission changes, while the pure meteorological predictor sets led to projected increases of MDA8/MDA1 in both scenarios. Regarding content this highlights the necessity to further reduce air pollution in the European area, bringing forth better protection of human health from direct emissions like NO_x as well as indirect pollutants such as O_3. From a methodological viewpoint, this points to the necessity to include information about emission changes in statistical projections to provide a more realistic assessment of future O_3 progression.

Multiple uncertainties exist within statistical downscaling assessments of ground-level O_3 under climate change which must be addressed properly. First, these uncertainties are related to the chosen downscaling method. Here, MLR with regularization and a thorough calibration/validation procedure was used which was appropriate for

the intended application and led to robust and reproducible results. Further, the choice of predictors affects the results and predictors were selected which are physically related to the predictand and contain information about climate change as well as changes of the precursor emissions. Climate change scenarios depend per se on the assumptions made within them and more than one scenario was regarded to illustrate a range of possible future evolution. Advancements in climate-chemistry modeling have led to more realistic assessments, but there are still many uncertainties. One way to attenuate these uncertainties was the use of multiple models spanning the range of possible future O_3 progression.

Recently, modern ESM ensemble learning methods for O_3 have been developed. These methods combine multiple algorithms and show high performance and stability. Thus, Requia et al. (2020) and Sun and Archibald (2021) developed multi-stage ensemble-learning frameworks based on random forest, gradient boosting, and neural network approaches, the latter also a 2-stage space-time Bayesian neural networks approach. While these methods greatly improve surface O_3 modeling, they are computationally relatively expensive, and they cannot treat the fusion of a small number of ESMs very well. So far, these methods have only been used with monthly O_3 data. Because the availability of daily data from the CMIP6 models is limited and yields a relatively small ensemble of models, complex model weighting or model fusion was not feasible in the present study. However, with our downscaling methodology, a weighting method of the individual ESMs would be possible as postprocessing. This was not done in this work, because the focus was rather on the sensitivity study of the predictors (particularly the influence of ozone as a predictor) and the influence of the different SSPs in the scenarios.

5. Conclusions

The provision of accurate estimates of local-scale ground-level O_3 under current conditions as well as under climate change is still a challenging task. Ground-level O_3 concentrations depend on the chemical and physical processes that govern the formation, transport, transformation, and deposition. These processes are controlled by multiple factors such as the characteristics and distribution of precursors, the larger-scale meteorology, and local-scale environmental conditions. Besides state-of-the-art chemical transport modeling, statistical approaches are useful tools to analyze and assess O_3. Under climate change, there are major concerns that O_3 concentrations will increase, despite efforts for better air pollution control. The results of the present study for local-scale ground-level ozone across the European area confirm this rating, with projected MDA8/MDA1 decreases under the SSP2-4.5 scenario, but increases under the scenario SSP3-7.0.

The present study contributes to the health-relevant topic of O_3 air pollution under climate and emission changes by assessing local-scale daily MDA8/MDA1 changes using the latest data and statistical methodology. In this regard, state-of-the-art ESM fusion methods (e.g., Liu et al., 2022; Sun & Archibald, 2021) are promising approaches to further improve the methodology presented in this study. Besides, for health impact assessments and public health and policy interventions, it is important to provide further information about changes in health-relevant thresholds such as the MDA8 value of 100 $\mu g/m^3$ for short-term O_3 exposure and of 60 $\mu g/m^3$ in the peak season (highest average in six consecutive months) for long-term exposure (World Health Organization, 2021). Furthermore, there is a need to refine exposure-response relationships from epidemiological studies to accurately estimate associated health burdens (Seltzer et al., 2018). Also, information about co-occurring heat stress as well as changes with respect to the susceptible population, as for instance considered in an assessment of future ozone and heat-related mortality by Orru et al. (2019), is useful information in this context. In this regard, the downscaling results of the present study provide a valuable basis for future studies.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

Air Quality eReporting data (EEA, 2017) are available at https://aqportal.discomap.eea.europa.eu/ with the raw data accessible at https://discomap.eea.europa.eu/map/fme/AirQualityExport.htm (since 2013) and https://discomap.eea.europa.eu/map/fme/AirQualityExportAirBase.htm (2004–2012). The authors acknowledge the European Centre for Medium-Range Weather Forecasts for provision of the ERA5 and CAMS data sets as

well as the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. ERA5 (Hersbach & Dee, 2016), CAMS (Inness et al., 2019), and CMIP6 (Eyring et al., 2016) data are available at https://cds.climate.copernicus.eu/, https://ads.atmosphere.copernicus.eu/cdsapp#!/search?type=dataset, and https://esgf-node.llnl.gov/projects/cmip6/, respectively. Most of the data preparation and analysis including regionalization, model building, and projections were conducted using the Python and R programming languages.

Acknowledgments
This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under project 408057478. Open access funding enabled and organized by Projekt DEAL.

References

Boleti, E., Hueglin, C., Grange, S. K., Prévôt, A. S., & Takahama, S. (2020). Temporal and spatial analysis of ozone concentrations in Europe based on timescale decomposition and a multi-clustering approach. *Atmospheric Chemistry and Physics*, *20*(14), 9051–9066. https://doi.org/10.5194/acp-20-9051-2020

Cannon, A. J. (2018). Multivariate quantile mapping bias correction: An N-dimensional probability density function transform for climate model simulations of multiple variables. *Climate Dynamics*, *50*(1), 31–49. https://doi.org/10.1007/s00382-017-3580-6

Checa-Garcia, R. (2018). CMIP6 ozone forcing dataset: Supporting information. *Zenodo*. https://doi.org/10.5281/zenodo.1135127

Cionni, I., Eyring, V., Lamarque, J. F., Randel, W. J., Stevenson, D. S., Wu, F., et al. (2011). Ozone database in support of CMIP5 simulations: Results and corresponding radiative forcing. *Atmospheric Chemistry and Physics*, *11*(21), 11267–11292. https://doi.org/10.5194/acp-11-11267-2011

Emmons, L. K., Schwantes, R. H., Orlando, J. J., Tyndall, G., Kinnison, D., Lamarque, J.-F., et al. (2020). The chemistry mechanism in the community Earth system model version 2 (CESM2). *Journal of Advances in Modeling Earth Systems*, *12*(4), e2019MS001882. https://doi.org/10.1029/2019MS001882

European Environment Agency (EEA). (2017). European air quality portal [Dataset]. Retrieved from https://aqportal.discomap.eea.europa.eu/

Eyring, V., Arblaster, J. M., Cionni, I., Sedláček, J., Perlwitz, J., Young, P. J., et al. (2013). Long-term ozone changes and associated climate impacts in CMIP5 simulations. *J. Geophys. Res.-Atmos.*, *118*(10), 5029–5060. https://doi.org/10.1002/jgrd.50316

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., & Taylor, K. E. (2016). Overview of the coupled model Intercomparison project phase 6 (CMIP6) experimental design and organization [Dataset]. Geoscientific Model Development, *9*(5), 1937–1958. https://doi.org/10.5194/gmd-9-1937-2016

Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., et al. (2019). The whole atmosphere community climate model version 6 (WACCM6). *Journal of Geophysical Research: Atmospheres*, *124*(23), 12380–12403. https://doi.org/10.1029/2019jd030943

Gutiérrez, J. M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., et al. (2019). An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. *International Journal of Climatology*, *39*(9), 3750–3785. https://doi.org/10.1002/joc.5462

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.

Hersbach, H., & Dee, D. (2016). ERA5 reanalysis is in production [Dataset]. Retrieved from https://www.ecmwf.int/en/newsletter/147/news/era5-reanalysis-production

Hertig, E. (2020). Health-relevant ground-level ozone and temperature events under future climate change using the example of Bavaria, Southern Germany. *Air Quality, Atmosphere & Health*, *13*(4), 435–446. https://doi.org/10.1007/s11869-020-00811-z

Hertig, E., Russo, A., & Trigo, R. M. (2020). Heat and ozone pollution waves in central and south Europe—Characteristics, weather types, and association with mortality. *Atmosphere*, *11*(12), 1271. https://doi.org/10.3390/atmos11121271

Inness, A., Ades, M., Agusti-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A. M., et al. (2019). The CAMS reanalysis of atmospheric composition [Dataset]. Atmospheric Chemistry and Physics, *19*(6), 3515–3556. https://doi.org/10.5194/acp-19-3515-2019

Jahn, S., & Hertig, E. (2021). Modeling and projecting health-relevant combined ozone and temperature events in present and future Central European climate. *Air Quality, Atmosphere & Health*, *14*(4), 563–580. https://doi.org/10.1007/s11869-020-00961-0

Jahn, S., & Hertig, E. (2022). Using clustering, statistical modeling, and climate change projections to analyze recent and future region-specific compound ozone and temperature burden over Europe. *GeoHealth*, *6*(4), e2021GH000561. https://doi.org/10.1029/2021GH000561

Keeble, J., Hassler, B., Banerjee, A., Checa-Garcia, R., Chiodo, G., Davis, S., et al. (2020). Evaluating stratospheric ozone and water vapor changes in CMIP6 models from 1850-2100.

Liu, X., Zhu, Y., Xue, L., Desai, A. R., & Wang, H. (2022). Cluster-enhanced ensemble learning for mapping global monthly surface ozone from 2003 to 2019. *Geophysical Research Letters*, *49*(11), e2022GL097947. https://doi.org/10.1029/2022gl097947

Lu, X., Zhang, L., & Shen, L. (2019). Meteorology and climate influences on tropospheric ozone: A review of natural sources, chemistry, and transport patterns. *Current Pollution Reports*, *5*(4), 238–260. https://doi.org/10.1007/s40726-019-00118-3

Miyazaki, K., Bowman, K., Sekiya, T., Eskes, H., Boersma, F., Worden, H., et al. (2020). Updated tropospheric chemistry reanalysis and emission estimates, TCR-2, for 2005–2018. *Earth System Science Data*, *12*(3), 2223–2259. https://doi.org/10.5194/essd-12-2223-2020

Morgenstern, O., Hegglin, M. I., Rozanov, E., O'Connor, F. M., Abraham, N. L., Akiyoshi, H., et al. (2017). Review of the global models used within phase 1 of the chemistry–climate model Initiative (CCMI). *Geoscientific Model Development*, *10*(2), 639–671. https://doi.org/10.5194/gmd-10-639-2017

O'Neill, B. C., Kriegler, E., Ebi, K. L., Kemp-Benedict, E., Riahi, K., Rothman, D. S., et al. (2015). The roads ahead: Narratives for shared socioeconomic pathways describing world futures in the 21st century. *Global Environment*, *42*, 169–180. https://doi.org/10.1016/j.gloenvcha.2015.01.004

O'Neill, B. C., Kriegler, E., Riahi, K., Ebi, K. L., Hallegatte, S., Carter, T. R., et al. (2014). A new scenario framework for climate change research: The concept of shared socioeconomic pathways. *Climatic Change*, *122*(3), 387–400. https://doi.org/10.1007/s10584-013-0905-2

Orru, H., Åström, C., Andersson, C., Tamm, T., Ebi, K. L., & Forsberg, B. (2019). Ozone and heat-related mortality in Europe in 2050 significantly affected by changes in climate, population and greenhouse gas emission. *Environmental Research Letters*, *14*(7), 074013. https://doi.org/10.1088/1748-9326/ab1cd9

Otero, N., Sillmann, J., Schnell, J. L., Rust, H. W., & Butler, T. (2016). Synoptic and meteorological drivers of extreme ozone concentrations over Europe. *Environmental Research Letters*, *11*(2), 024005. https://doi.org/10.1088/1748-9326/11/2/024005

Park, S., Son, S. W., Jung, M. I., Park, J., & Park, S. S. (2020). Evaluation of tropospheric ozone reanalyses with independent ozonesonde observations in East Asia. *Geoscience Letters*, *7*(1), 1–12. https://doi.org/10.1186/s40562-020-00161-9

Porter, W. C., & Heald, C. L. (2019). The mechanisms and meteorological drivers of the summertime ozone–temperature relationship. *Atmospheric Chemistry and Physics*, *19*(21), 13367–13381. https://doi.org/10.5194/acp-19-13367-2019

Requia, W. J., Di, Q., Silvern, R., Kelly, J. T., Koutrakis, P., Mickley, L. J., et al. (2020). An ensemble learning approach for estimating high spatiotemporal resolution of ground-level ozone in the contiguous United States. *Environmental Science & Technology*, *54*(18), 11037–11047. https://doi.org/10.1021/acs.est.0c01791

Schnell, J. L., Prather, M. J., Josse, B., Naik, V., Horowitz, L. W., Zeng, G., et al. (2016). Effect of climate change on surface ozone over North America, Europe, and East Asia. *Geophysical Research Letters*, *43*(7), 3509–3518. https://doi.org/10.1002/2016gl068060

Sellar, A. A., Jones, C. G., Mulcahy, J., Tang, Y., Yool, A., Wiltshire, A., et al. (2019). UKESM1: Description and evaluation of the UK Earth system model. *Journal of Advanced Modeling Earth System*.

Seltzer, K. M., Shindell, D. T., & Malley, C. S. (2018). Measurement-based assessment of health burdens from long-term ozone exposure in the United States, Europe, and China. *Environmental Research Letters*, *13*(10), 104018. https://doi.org/10.1088/1748-9326/aae29d

Shen, L., Mickley, L. J., & Gilleland, E. (2016). Impact of increasing heat waves on US ozone episodes in the 2050s: Results from a multimodel analysis using extreme value theory. *Geophysical Research Letters*, *43*(8), 4017–4025. https://doi.org/10.1002/2016gl068432

Sun, Z., & Archibald, A. T. (2021). Multi-stage ensemble-learning-based model fusion for surface ozone simulations: A focus on CMIP6 models. *Environmental Science and Ecotechnology*, *8*, 100124. https://doi.org/10.1016/j.ese.2021.100124

World Health Organization. (2021). *WHO global air quality guidelines: Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide* (p. 273). World Health Organization. Retrieved from https://apps.who.int/iris/handle/10665/345329

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. *Journal of Geophysical Research*, *113*(D6), D06302. https://doi.org/10.1029/2007jd008917

Young, P. J., Naik, V., Fiore, A. M., Gaudel, A., Guo, J., Lin, M. Y., et al. (2018). Tropospheric Ozone Assessment Report: Assessment of global-scale model performance for global and regional ozone distributions, variability, and trends. *Elementa: Science of the Anthropocene*, *6*. https://doi.org/10.1525/elementa.265

Yukimoto, S., Kawai, H., Koshiro, T., Oshima, N., Yoshida, K., Urakawa, S., et al. (2019). The meteorological research Institute Earth system model version 2.0, MRI-ESM2. 0: Description and basic evaluation of the physical component. *Journal of the Meteorological Society of Japan. Ser. II*, *97*(5), 931–965. https://doi.org/10.2151/jmsj.2019-051