

RESEARCH

Open Access



Characterization of tropospheric ozone pollution, random forest trend prediction and analysis of influencing factors in South-western Europe

Jinyang Wang^{1,2}, Tianzhen Ju^{1,2*}, Bingnan Li³, Cheng Huang^{1,2}, Xuhui Xia^{1,2}, Jiaming Zhang^{1,2} and Chunxue Li^{1,2}

Abstract

Nowadays, environmental problems have gradually become the focus of world attention. In recent years, heat waves in many parts of Europe have increased ozone concentrations, fuelling ozone pollution. Therefore, this paper investigates the spatial and temporal distribution of tropospheric column ozone concentrations in South-western Europe, future trend changes, influencing factors, and potential source regions based on remotely sensed monitoring data from the OMI (Ozone Monitoring Instrument) from 2011 to 2021. The results show that the areas of high tropospheric column ozone concentrations are mainly concentrated in the northwest, Poland, and southeast coastal areas. At the same time, the monthly variation curve of column ozone concentration is bimodal. Trend change analyses indicate an upward trend in future column ozone concentrations in the southeastern part of the study area. The potential for increases also exists in parts of Germany, France, and Poland, which will need to be monitored. Random forest model projections found a slight decrease in column ozone concentrations in 2022 and 2023 of about 1–4 DU compared to tropospheric column ozone concentrations in 2021. The health risk assessment found that the number of all-cause premature deaths due to exposure to ozone was the highest in Germany. During the summer, when ozone pollution is high, the potential source area in the southeastern part of the study area is located at the border of the three countries, and synergistic management is recommended. In exploring the correlation between the influencing factors and ozone, it was found that there is a significant difference between the long-time and short-time series. In addition, the pathway analysis shows that the population size, distribution density, and forested area in southwestern Europe may be more sensitive to the production of tropospheric ozone.

Keywords O₃, Random forest forecasting, Hurst index, BenMAP-CE, Potential sources, OMI

*Correspondence:

Tianzhen Ju

jutianzhen@nwnu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Introduction

Tropospheric ozone is an essential secondary air pollutant and a key component of photochemical smog. Most studies in the European region show a strong relationship between high ozone concentrations and their precursors and meteorological parameters. In addition, the impact of motor vehicle emissions on air quality should be considered and is particularly pronounced during the year's warm season [17]. Tropospheric ozone has some negative impacts on human health, as well as on flora and fauna. The combination of drought and ozone stress in the European region, where summers are hot and dry, can lead to a 10–17% reduction in wheat yields [22]. Meanwhile, exposure to high levels of tropospheric ozone (O₃) can lead to phytotoxic effects, which may affect grape yield and quality [6]. It has also been found that short-term exposure to ozone (O₃) induces oxidative stress in human and mouse skin, leading to aberrant transcriptional expression of genes consistent with increased skin aging. Long-term exposure to high tropospheric ozone concentrations appears to cause skin aging (Kateryna et al. 2019).

In response to the refined effects of tropospheric ozone on human health, high ozone concentrations are quadratically related to daily mortality, an effect that is more pronounced in respiratory causes [7]. Toybi et al. examined long-term exposure to delicate particulate matter (PM_{2.5}), nitrogen dioxide (NO₂), and ozone (O₃) in relation to the incidence of hospitalization for acute myocardial infarction (AMI) and stroke. A positive correlation was found between long-term exposure to PM_{2.5}, O₃, and, to a lesser extent, NO₂, and hospitalization for AMI and stroke. It was found that for every 5 ppb increase in ozone, there was a 1.98 percent increase in mortality [34]. Tropospheric ozone is harmful to human health. In the Mediterranean region of Europe, exposure to lower ozone resulted in higher hospitalization rates and mortality in humans for cardiovascular diseases than respiratory diseases [33]. The human health effects of sub-exposure to ozone also differed after COVID-19. Estimation of the association between pollution exposure levels and non-accidental mortality using a Pine-Pose regression model found that changes in ozone concentration were associated with an increase in premature deaths during lockdown and quarantine [3]. Therefore, this paper must study the effects of ozone pollution on human health underexposure.

In recent years, extreme heat events have occurred in many parts of Europe due to heatwaves. The Mediterranean region has been affected by heatwaves that have resulted in severe hill fires in Greece, Turkey, and Italy. The high temperatures associated with heatwaves have exacerbated tropospheric ozone pollution over much of

Europe. The Copernicus Atmospheric Monitoring Service (CAMS) monitored "extremely high levels of surface ozone pollution" in western and southern Europe. In this paper, the Aura satellite-based Ozone Monitoring Instrument (OMI) tropospheric column ozone concentration data for the years 2011–2021 are selected. Spatial interpolation and spatial computation using ArcGIS were used to derive ozone's spatial and temporal distribution characteristics in southwestern Europe. Tropospheric ozone trends were analyzed using the ozone concentration trend (SLOPE) slope and the Hurst index. The Random Forest Model (RFM) was further used to project tropospheric ozone changes for 2022 and 2023. Meteorological mapping software (MeteoInfo Map) was used to analyze and explore the potential source areas of pollution in the heavily polluted Patras region in conjunction with the transport paths of air masses in the region. The impact assessment of human health effects of surface ozone was carried out using the BenMAP-CE model. Finally, the Pearson correlation analysis, through-center analysis, and elliptical center-of-gravity migration model were used to explore the natural and human factors. In order to provide a reference for the prevention and control of air pollution in the region.

Overview of the study area

South-western Europe is located in the western part of the Asian and European continent and the eastern coast of the Atlantic Ocean, most of which is located at latitudes of 35° to 60° north and belongs to the northern temperate zone. Broadly speaking, this region of South-western Europe can also be divided into four parts: Northern, Western, Central and Southern Europe, which covers an area of about 5 million square kilometers and includes more than 30 countries, about half of Europe. The western half of Europe, bordered by the Arctic Ocean in the north, the Atlantic Ocean in the west and the Mediterranean Sea in the south, is mostly located between 35° and 70° north latitude, with a gradual transition from maritime to continental climate from west to east. The climate varies widely from north to south, with a temperate maritime climate in the north and a Mediterranean climate in the south. The vast majority of the region is a plain with an east–west orientation. This allows westerly winds and cyclones to penetrate deep into the continent and extend into the Atlantic Ocean. South-western Europe receives year-round rainfall, with hot, rainy summers and mild, wet winters. The temperature difference between land and water is greater in winter, the pressure difference between land and water is correspondingly greater, and the winds are stronger.

As shown in Fig. 1, this region of study is part of South-western Europe, which contains 26 countries including the UK, France, Germany, Ireland, Poland, Serbia and Greece. As can be seen from the DMSP/OLS nighttime light data map, the average nighttime light intensity is higher in northwestern South-western Europe. South-western Europe is well developed in heavy industry, coal, steel, chemicals, and machinery and accounts for nearly 30% of the world's industrial production. The dominance of industry and energy use is one of the reasons for the high level of atmospheric pollution in the region. At the same time, heat waves and persistent high temperatures in recent years have led to a number of forest fires, and high temperatures can also exacerbate ozone pollution. According to the Copernicus Atmospheric Monitoring Service (CAMS), forest fires in Spain generated about 1.3 million tonnes of greenhouse gases, and in France, they generated about 344,000 tonnes of greenhouse gases, a record high.

Data sources

The tropospheric ozone and NO₂ data used in this study are from the OMI sensor, which is on board the EOS-Aura satellite launched by NASA (<https://www.earthdata.nasa.gov/>), a near-polar, sun-synchronous orbit scientific exploration satellite with an orbit altitude of 705 km, jointly developed by the Netherlands Agency for Aeronautics and Space and the Finnish Meteorological Service. OMI has a spatial resolution of 0.25° × 0.25°. OMI is used to observe backscattered radiation from the Earth's

atmosphere and the Earth's surface to obtain information [10] and can measure a wide range of data such as ozone, HCHO, NO₂, SO₂ column concentrations and profiles as well as aerosols, clouds, and surface UV radiation [45].

The meteorological data are derived from a dataset jointly provided by the National Centers for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) (<https://psl.noaa.gov/data/gridded/data.ncep.reanalysis.surface.html>), which contains data from 1948 to the present. The potential source study data in this paper is also based on meteorological data provided by the NCEP to explore the contribution of potential pollution sources in this study area. In the impact analysis of social factors, this paper uses data on population, PM_{2.5} pollution, greenhouse gas emissions, forested land area, and Other manufacturing (% of value added in manufacturing) from the World Development Bank database. (<https://databank.worldbank.org/source/world-development-indicators?l=en#>).

Because disease and population data are missing for individual countries, this study discusses the effects of ozone on human health for the 21 countries mentioned above. The ozone surface data used for the health benefits study were CAMS annual air quality (interim) reanalyses for the European region (CAMS European air quality reanalyses (Copernicus.eu)) with a spatial resolution of 0.1°, ~ 10 km. Disease data are based on the dataset provided by the World Health Organization (<https://platform.who.int/mortality/countries>).

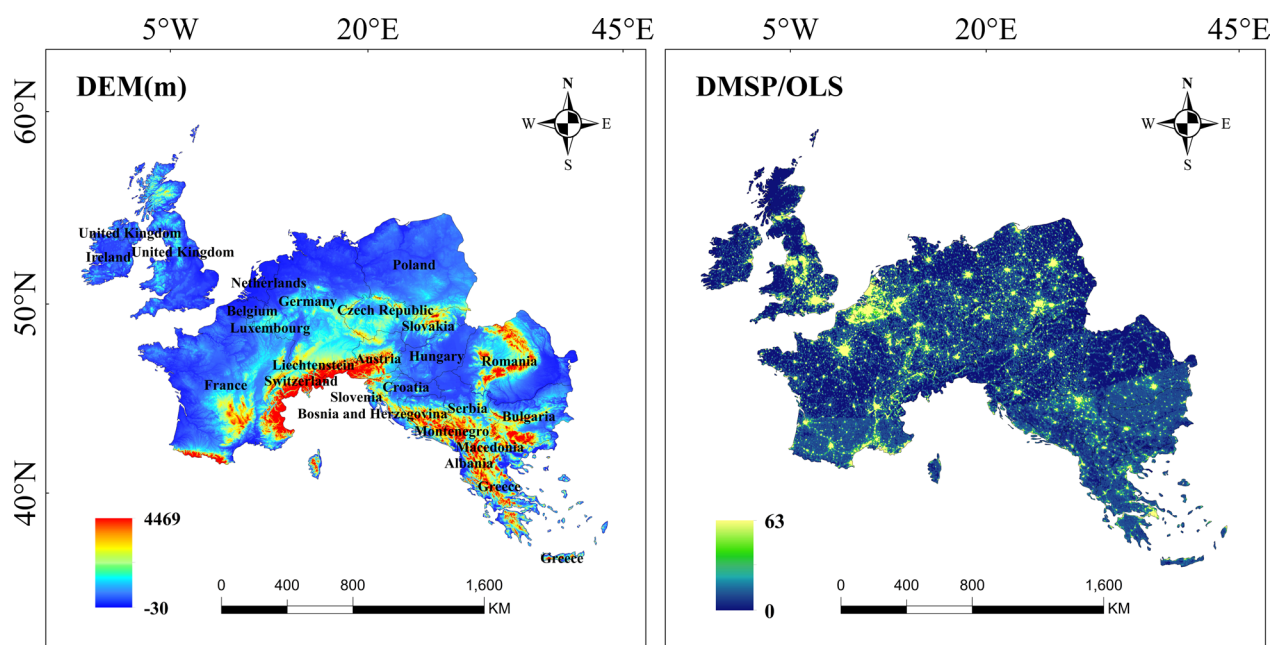


Fig. 1 Map of elevation and nighttime light data for the South-western European study area

Research methodology

After acquiring the daily column ozone concentration remote sensing data for the study area from 2010 to 2021 and obtaining the ozone data stored in the HDF-EOS strip (swath) data format, the data were batch processed using a Python program. Firstly, the Python software was used to extract latitude, longitude, column ozone concentration and cloudiness data. Secondly, the latitude and longitude of the data point coverage was extended by 0.5° to increase the confidence level of the pollutant data to maximize the area covered by the column ozone concentrations in the study area. Finally, the spatial and temporal distribution of column ozone concentrations over parts of South-western Europe was obtained in ArcGIS by grid calculation, kriging, averaging and mask extraction.

Global nighttime light data are obtained by extracting, projection correcting, and resampling the raw imagery to obtain raster data. The DN (Digital Number) value of the image element obtained after processing indicates the light intensity, and the grey scale value ranges from 0 to 63.

Pearson correlation analysis

In analyzing the spatial relationship between ozone relationship with natural factors (precipitation, temperature, NDVI), this paper adopts the Pearson correlation analysis, The correlation coefficient r reflects the positive or negative correlation between the variables and the strength of the correlation [11].

$$r_{xy} = \frac{\sum_{i=1}^n [(x_i - \bar{x})(y_i - \bar{y})]}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (1)$$

where x and y represent two variables, r_{xy} is the correlation coefficient between the x and y variables, and i is the year number.

One-dimensional linear regression

To study the spatial and temporal trends in column concentrations over 12 years in parts of South-western Europe, this paper uses a one-dimensional linear regression analysis to establish the relationship between column ozone concentrations and temporal variability, where the spatial variability characteristics of individual pixels at different times can be used to characterize the evolution of regional patterns for a given time series [24], calculated as Eq. (2):

$$\theta_{slope} = \frac{n \times \sum_{i=1}^n (i \times O_{3i}) - \sum_{i=1}^n i \sum_{i=1}^n O_{3i}}{n \times \sum_{i=1}^n i^2 - (\sum_{i=1}^n i)^2}, \quad (2)$$

where θ_{slope} is the slope of each column ozone concentration value, i.e., the coefficient of variation of the column ozone concentration for each pixel, i is the year for each year from 2010 to 2021, n is the number of years, and O_{3i} is the specific column ozone concentration value corresponding to year i . When $\theta_{slope} > 0$ indicates that the column ozone concentration also shows an increase with time, when $\theta_{slope} < 0$ then the opposite is true. The larger the absolute value of θ_{slope} , the greater the variation in column ozone concentration values will be [12].

Hurst index

This paper uses the Hurst index to predict future trends in annual column ozone concentrations in selected regions of South-western Europe, which can quantitatively characterize the long-range dependence or persistence of a time series and help to analyze future trends at a particular time or space. This paper uses Matlab software to perform pixel calculations to determine the Hurst index.

The calculation is based on the following principle: located at moments t_1, t_2, \dots, t_n , the response time series obtained at $\xi_1, \xi_2, \dots, \xi_n$, and for any positive integer $\tau \geq 1$, the time series are averaged as [40]:

$$\langle \xi \rangle_\tau = \frac{1}{\tau} \sum_{t=1}^{\tau} \xi(t) \quad \tau = 1, 2, 3, \dots, n. \quad (3)$$

The cumulative deviation is:

$$X(t, \tau) = \sum_{t=1}^{\tau} (\xi(t) - \langle \xi \rangle_\tau), \quad 1 \leq t \leq \tau. \quad (4)$$

The extreme differences are:

$$R(\tau) = \max_{1 \leq t \leq \tau} X(t, \tau) - \min_{1 \leq t \leq \tau} X(t, \tau) \quad \tau = 1, 2, 3, \dots, n. \quad (5)$$

The standard deviations are:

$$S(\tau) = \left[\frac{1}{\tau} \sum_{t=1}^{\tau} (\xi(t) - \langle \xi \rangle_\tau)^2 \right]^{\frac{1}{2}} \quad \tau = 1, 2, 3, \dots, n, \quad (6)$$

where the following relationship is satisfied between R , S , τ and H is the Hurst index:

$$\frac{R}{S} = \left(\frac{\tau}{2} \right)^H. \quad (7)$$

The range of values of the Hurst coefficient H corresponds to three cases, respectively [23]: (1) when $H=0.5$, it means that the indicators are completely independent and there is no interdependence, (2) when $0.5 < H < 1$, it means that the time series has persistence or long-range correlation, implying that the future ozone trend is consistent with the past and the closer the value of H is to 1, the stronger the persistence is; (3) when $0 < H < 0.5$, it means that the future overall trend is opposite to the past and is resistant to persistence and the closer the H value is to 0, the stronger the resistance to persistence.

Forest-based classification and regression

Forest-based classification and regression spatial statistics tools in ArcGIS Pro and MATLAB software were used to predict the distribution of ozone column concentrations in the study area for the years 2022 and 2023. A model is created, and predictions are generated based on an adaptation of the random forest model (RF model). Wherein the explanatory variables can be in the form of fields in the training element attribute table, raster datasets, and distance element neighborhood analysis values used for calculations used as additional variables. Compared with other "black box" machine learning algorithms, the RF algorithm is relatively simple and easy to operate, and the model is highly interpretable. At the same time, RF algorithms can effectively prevent variable covariance and avoid model overfitting problems [44].

BenMAP-CE model

The environmental benefit assessment model (BenMAP) is a geographic information system (GIS)-based model that estimates the health benefits associated with changes in air quality by creating population exposure surfaces in order to estimate changes in a range of health outcomes associated with air pollution (US EPA. BenMAP User's Manual). BenMAP analyzes the health effects of air pollution using a health impact function (HIF), which is shown below:

$$\Delta Y = (1 - e^{-\beta \Delta X}) \times Y_0 \times \text{POP}, \quad (8)$$

where ΔY is the estimated health effect of the change in pollutant concentration, Y_0 is the baseline incidence of health endpoints (i.e., mortality or morbidity), POP is the population affected by the change in air quality, ΔX is the change in air quality, and β is the coefficient of the relationship between the pollution concentration and the health effect (i.e., the coefficient of the exposure–response relationship) [18].

Table 1 Exposure–response coefficient of health effects

Health effect terminal	β (%)	95%CI (%)
All-cause mortality [5]	0.92	0.47 to 1.38
Cardiovascular mortality [5]	1.09	0.61 to 1.58
Respiratory mortality [5]	0.45	– 0.74 to 1.65

β indicates the percentage increase in population risk of death for each 1 $\mu\text{g}/\text{m}^3$ increase in O_3 concentration

The concentration threshold for surface ozone in this study was 57.24 $\mu\text{g}/\text{m}^3$ [35], and the exposure response factors are specified in Table 1.

Path coefficient

The pathway coefficient is a measure of the degree of correlation between variables. It removes the depletion of each factor in the middle and finally gets the contribution of each factor to the concentration of the column ozone. In the multivariate relationship, it is necessary to discuss the causal relationship between x_i and x_j and the relative importance to the variable y , which is formulated as follows (8) [19, 32]:

$$p_{ij} = b_i \frac{Z_{x_i}}{Z_y}, \quad (9)$$

where p_{ij} is the path coefficient, b_i is the partial regression coefficient, and Z_{x_i} and Z_y represent the standard deviation.

Potential source contribution analysis

Potential source contribution function (PSCF) is a method for identifying the likely location of pollution sources based on airflow trajectories using a conditional probability function as the basic principle, also known as residence time analysis [27, 31, 39]. A GIS program and meteorological mapping software (MeteoInfo) are used to create a rectangular grid (i, j) with a certain resolution that covers all the areas under study, based on the results of backward trajectory simulations, and thresholds are set for atmospheric pollutant concentrations. When one of the trajectories passes through a region where the pollutant concentration exceeds the set threshold, it is marked as a polluted trajectory. The number of nodes of the polluted trajectory passing through grid point (i, j), m_{ij} , is compared with the number of trajectory endpoints falling within that grid, n_{ij} , which is the result of the PSCF, corresponding to the calculation formula:

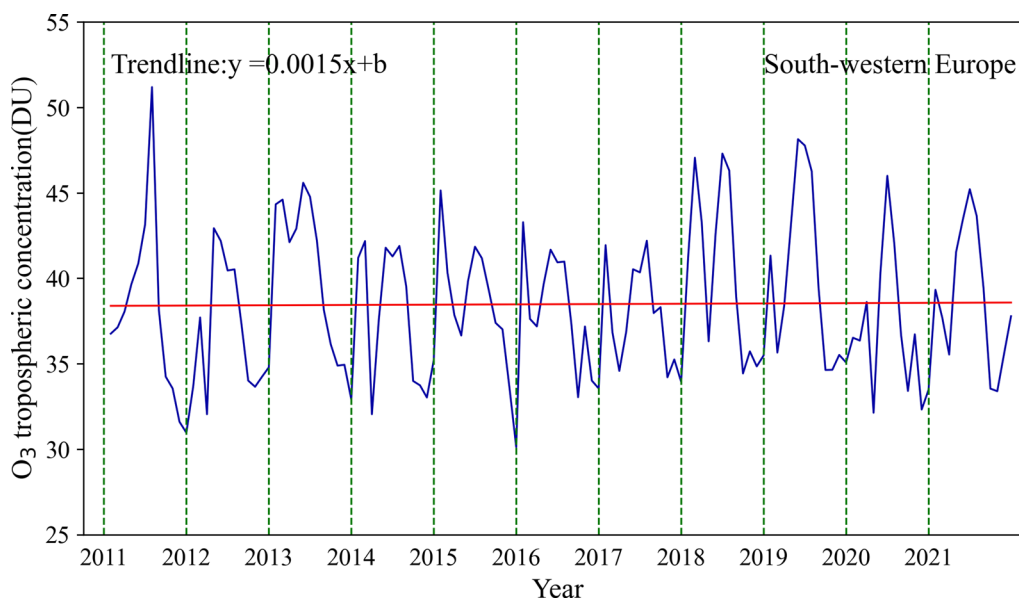


Fig. 2 Monthly trends in tropospheric column ozone concentrations in South-western Europe, 2011–2021

$$W_{ij} = \begin{cases} 1.0, & n_i > 3n \\ 0.70, & 1.5n < n_i \leq 3n \\ 0.42, & < n_i \leq 1.5n \\ 0.17, & n_i \leq n. \end{cases} \quad (10)$$

To reduce the uncertainty of the PSCF [13, 25], a weighting factor W_{ij} , when the n_{ij} of one of the grids is less than three times the average trajectory endpoints of each grid in the study area, is introduced to reduce the impact of the resulting error, calculated as:

$$WPSCF(i, j) = W(i, j) \times PSCF(i, j), \quad (11)$$

Weighting and calculation of the PSCF:

$$PSCF_{ij} = \frac{m_{ij}}{n_{ij}}. \quad (12)$$

Results and discussion

Monthly variation in column ozone concentrations

As can be seen in Fig. 2, the monthly tropospheric ozone concentrations in South-western Europe do not fluctuate significantly from 2011 to 2021 and show a "double peak" over the study period. Regarding the overall trend, the two peaks occurred in 2013 and 2018–2019, respectively. The change in concentrations from 2014 to 2017 is flat and low, with a rapid decline in 2020 and a slight increase in 2021. The overall trend over the 11a period is a slow increase with a slope of 0.0015, with monthly mean tropospheric column ozone concentrations ranging from 30 to 53 DU. Most of the high monthly mean values occurred from May to July each year, with low values

occurring in April and September to November each year. 11a The maximum value occurred in July 2011, with a column concentration value of 51.18 DU; the minimum was in December 2015, with a column concentration value of 30.19 DU.

Figure 3 shows the monthly mean tropospheric column ozone concentration values distribution in the study area. In this paper, the column ozone concentrations in the region are classified into seven classes, from low to high. Figure 3 shows that the spatial distribution of monthly mean column ozone concentrations is somewhat regular, with "high in the north and low in the south" from January to April and from October to December and high values in the south of the study area from May to September, contrary to the above mentioned periods.

The highest ozone concentration values were recorded in June, with levels 5–7 occupying 44.5% of the study area, while in July and August, the concentration range decreased by 13.8% and 38.9% of the study area, respectively. From the seasonal point of view, the concentration values were high in spring (March to May) and summer (June to August). However, the distribution of the areas of high concentrations in the two seasons differed. In summer, the high concentration areas are concentrated in the southern part of the study area, such as southern France, Hungary, Albania, Greece, eastern Bulgaria, and eastern Romania. High ozone values are found in spring in areas such as Poland and northern Germany. The lowest values are found in autumn, with a trend of shifting low values from northwest to southeast from September to November. The range of autumn values mainly floats

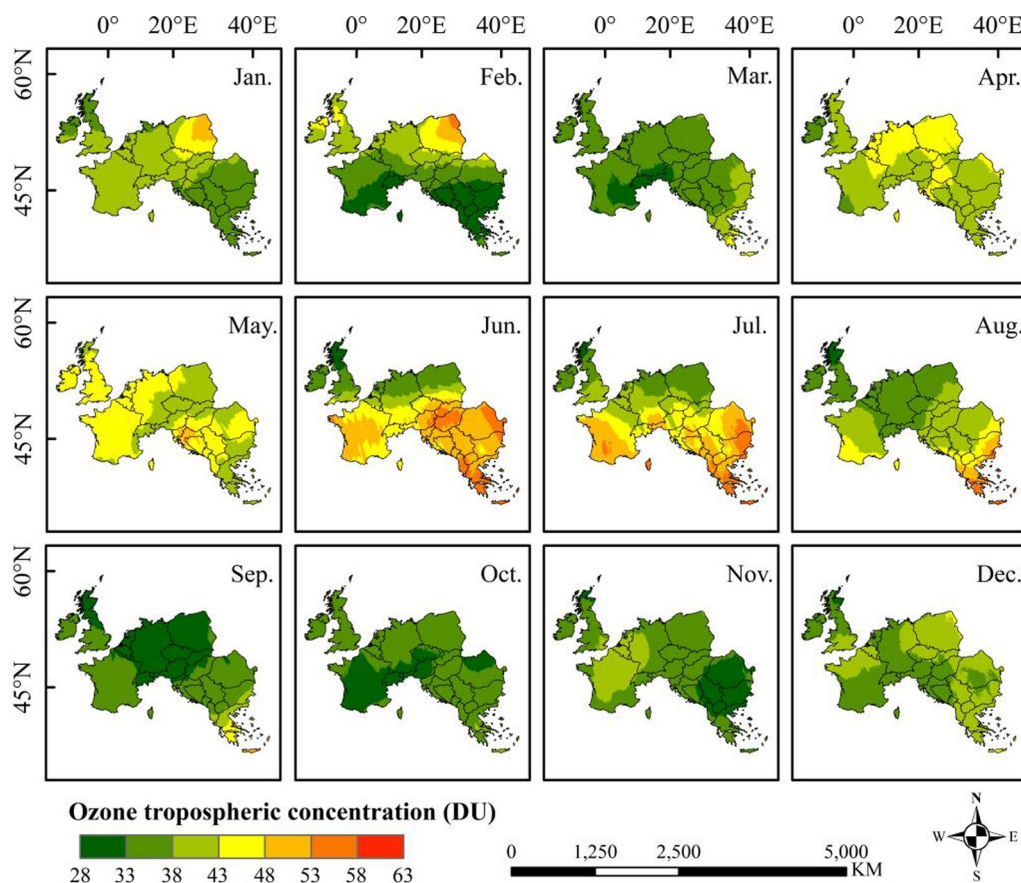


Fig. 3 Spatial distribution of monthly mean tropospheric ozone column concentrations in South-western Europe, 2011–2021

at lower levels, with 98%, 100%, and 100% of levels 1–3, respectively, and a seasonal average value of 34.2 DU.

Interannual variation in column ozone concentrations

Figure 4 shows the annual mean spatial distribution of column ozone concentrations in southwestern Europe from 2010 to 2021. Taking 4 DU as a class, the column ozone concentration values are divided into five classes with annual average values ranging from 33.01 DU to 50.28 DU. From the spatial distribution, the overall tropospheric column ozone concentration changes are relatively smooth from 2010 to 2021. Notably, ozone pollution in most countries in southwestern Europe was exacerbated by heatwaves from 2017 onwards, with significantly higher concentration values, which increased again in 2020 after a sudden decrease. The areas with high concentration values are mainly concentrated in the southeastern part of the study area, especially in Greece and Albania. Relatively high values were also observed in the southern United Kingdom, eastern Poland, and France. The generally low concentration values in central

Europe may be related to the region’s overall reduction of NOx precursors [46].

These trend changes and distributional features may be closely related to climate change in the study area, related policy measures, and the new crown epidemic. The amplification of the European constant wave train and the enhancement of the large-scale circulation background, such as the NAO (North Atlantic Oscillation) dipole, which has been affected by heat waves in recent years, have favored the establishment and enhancement of anticyclonic blocking anomalies in continental Europe, thus contributing to the generation of heat waves in Europe [39]. The Mediterranean basin, under the influence of subtropical high pressure, cloud-free conditions and high solar radiation intensity, is prone to photochemical solid degradation of volatile organic compounds (VOCs) in air masses transported from regions such as Central Europe, the Balkan Islands, and other regions (e.g., Asia and North America) [29], which in turn leads to ozone concentrations in the southwestern-southeastern and partly central parts of Europe that are relatively

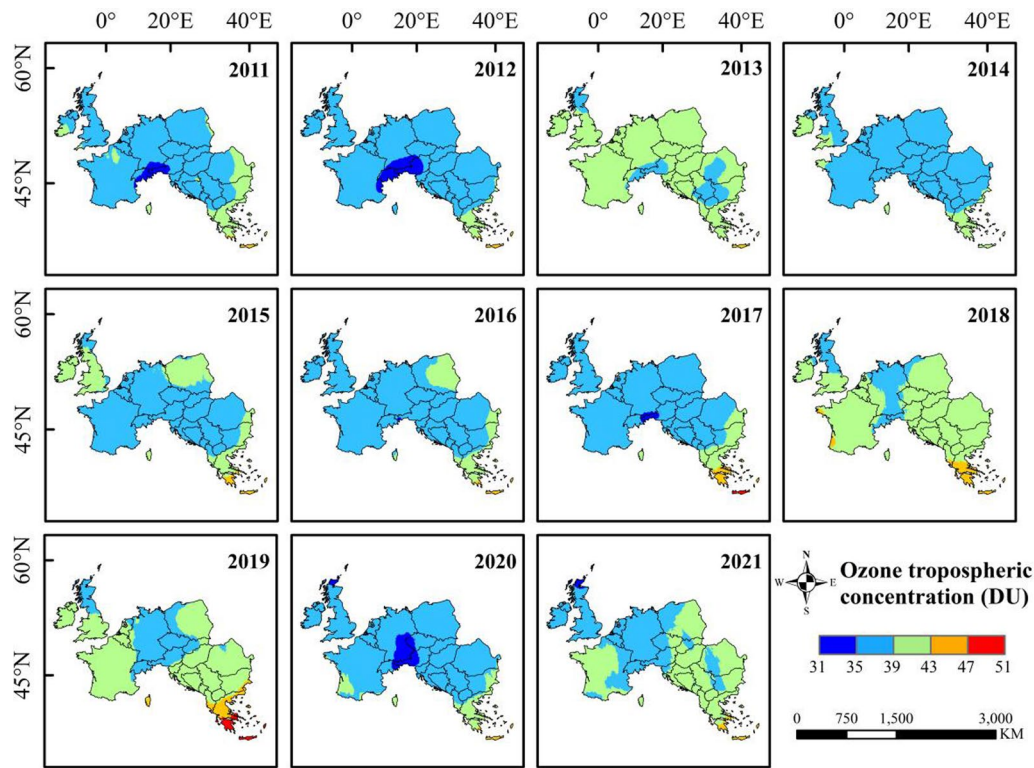


Fig. 4 Spatial distribution of annual mean tropospheric ozone column concentrations in South-western Europe, 2011–2021

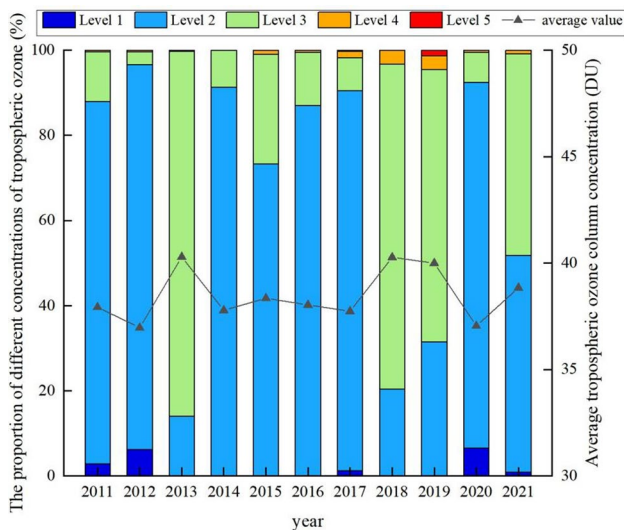


Fig. 5 Stacking of annual mean tropospheric column ozone concentrations in South-western Europe, 2011–2021

higher in south-eastern and parts of central South-western Europe compared to other regions.

As can be seen from Fig. 5, a plot of the percentage area of different classes of annual mean tropospheric ozone concentrations in southwest Europe, the column ozone

concentration values are higher in 2013, 2018, and 2019. The share of high concentration values in classes 4 and 5 (43–51 DU) reached its maximum in 2019. The remaining years are dominated by classes 1–2, which account for 51.85–94.77% of the overall study area. Due to the heatwave that swept through Europe in 2018, A study by the Met Office (Met Office) has revealed that the chances of this heatwave sweeping across the UK are thirty times greater due to emissions from human activity compared to previous years. South England has seen some of the longest-lasting warmth on record. The rest of Europe shows a similar increase in concentration values. After COVID-19, however, European ozone concentrations show a significant downward trend in most cases [26].

Trend change analysis

In order to provide a clearer picture of the year-to-year annual mean ozone trend in parts of SW Europe from 2011 to 2021, this study used Arcgis image spatial analysis to fit and calculate the slope of the ozone trend, as shown in Fig. 6a, where θ_{slope} denotes the slope of ozone change, $\theta_{\text{slope}} > 0$ indicates an upward trend and $\theta_{\text{slope}} < 0$ indicates a downward trend. As can be seen from the figure, the slope index of θ_{slope} in the southern part of the study area reaches a mean value of 0.026, with no significant

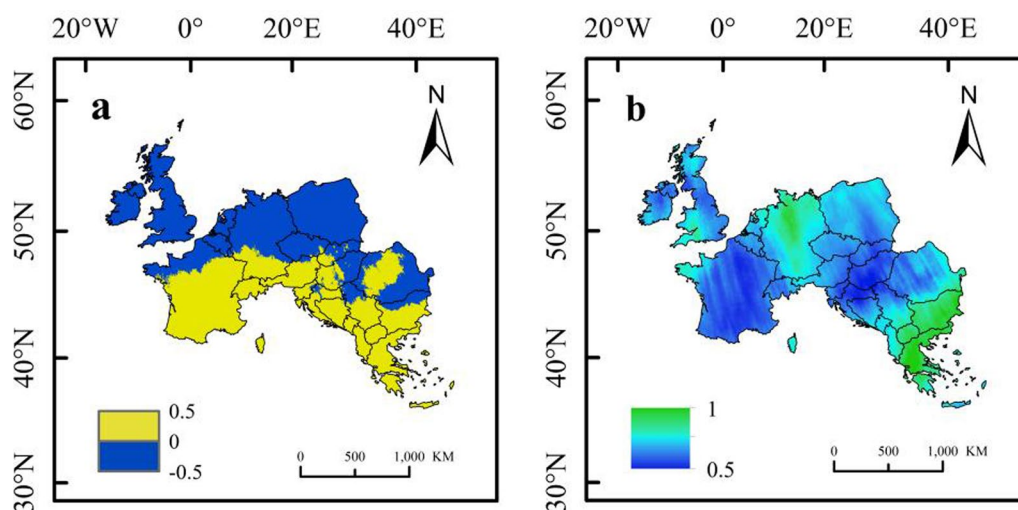


Fig. 6 Slope of ozone concentration change (a) and spatial distribution of Hurst index (b)

upward trend. On the contrary, the northern parts of the study area, such as the UK, northern Germany, northern Poland, and Romania, show a slow decreasing trend. This may be related to the EU's Seventh Environmental Action Programme [42]. According to the European Air Quality Report (2020) results, air pollutant emissions have been significantly reduced in most sectors after the project. In 2018, significant pollutants from road and non-road traffic still significantly declined despite the continued growth in passenger and freight traffic. At the same time, the decline in air pollutant emissions was closely linked to the stringency of outbreak control measures. Across Europe, countries such as Italy, France, Poland, Switzerland, Germany, the United Kingdom, and Belgium have seen significant declines in NO_x emissions, fuelling the downward ozone trend [42].

In this paper, the Hurst index is calculated pixel by pixel in space using MATLAB software according to the principle of R/S analysis. The Hurst index reflects the autocorrelation of the time series. As shown in Fig. 6b, the calculated values of the Hurst index ranged from 0.51 to 0.71. According to the calculations, the study area shows moderate persistence, which indicates a continuous upward trend of ozone concentration values in the southern part of the study area in the coming period. In contrast, the upward trend is more pronounced in countries such as Bulgaria, Macedonia and Greece. However, there is a continuing downward trend in ozone concentration values in northern Germany, the United Kingdom and north-west France. Changes in these trends may be related to local industrial, manufacturing and other sectors. Therefore, countries in the south-eastern part of the study area need to implement stricter air control

measures and strengthen monitoring to reduce the likelihood of air pollution events.

Random forest prediction of the characteristics of future ozone changes

The random forest regression model was used in Matlab to predict the month-by-month concentration values of column ozone concentration in 2021. In other studies, Xiong et al. [41] used a random forest (RF) model to predict and correct the maximum eight-hourly concentration of ozone in the Yangtze River Delta region of China. Asgari et al. [2] explored the development and application of parallel air quality prediction models based on machine learning algorithms such as random forests in Tehran, Iran.

Figure 7 shows the curve of predicted versus measured values for 2021, with an R^2 of 0.971 for the model. Figure 8 shows the spatial distribution of annual mean column tropospheric ozone concentrations measured in southwestern Europe in 2021 and the annual mean column ozone concentrations in 2021 predicted by Random Forest regression modeling. A comparison of these two concentration values shows that the maximum difference in annual mean concentration between measured and predicted values in the region is 4.6983 DU, and the difference in mean values is only 0.0036 DU. The overall spatial distribution pattern of column ozone concentrations maintains high consistency. Therefore, the model can predict the monthly average concentration values in 2022 and 2023 with relative accuracy.

Figure 9 shows the predicted monthly mean concentrations in South-western Europe in 2022 using the 2011–2021 monthly mean concentration model.

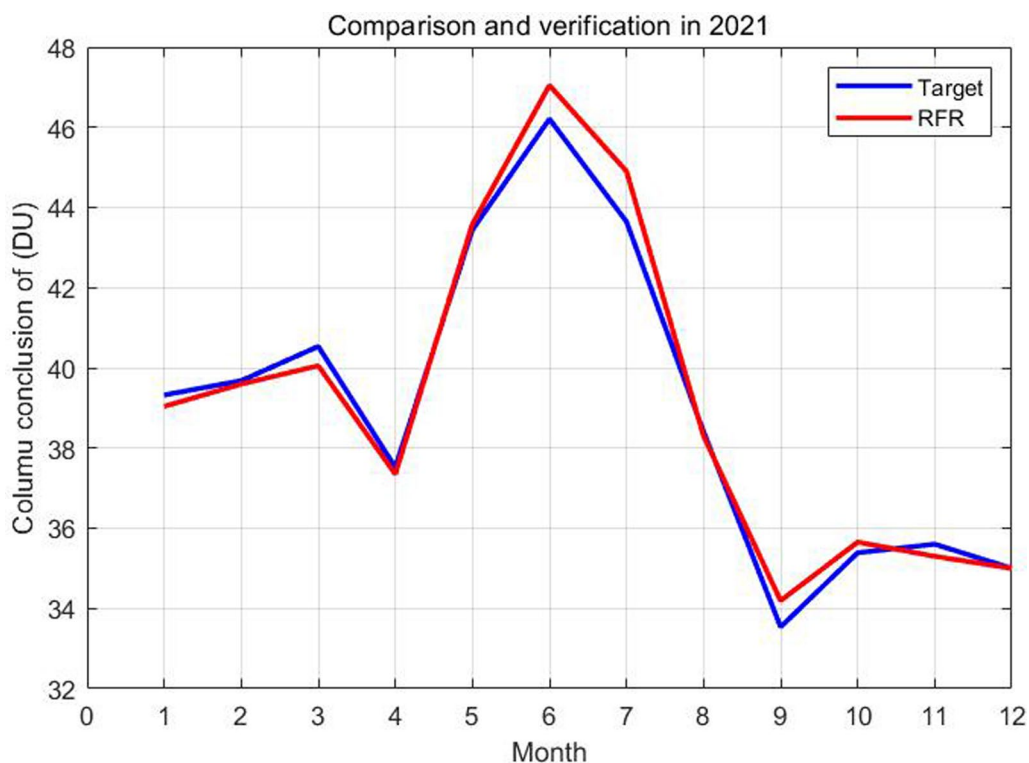


Fig. 7 Comparison of measured and predicted monthly mean tropospheric ozone concentration in South-western Europe in 2021

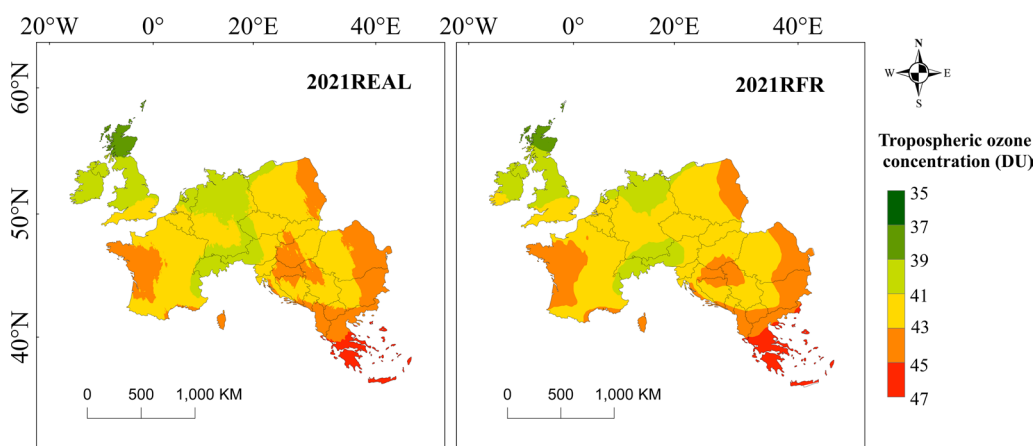


Fig. 8 The measured value of column tropospheric ozone concentration in 2021 and the predicted value of random forest

Similarly, Fig. 10 shows the predicted monthly mean ozone concentrations in South-western Europe in 2023 using the 2011–2021 monthly mean measurements and the 2022 monthly mean predictions. As can be seen in Figs. 9 and 10, the monthly mean values for 2022 and 2023 have two distinct peak months in a year. Higher column ozone values were recorded in February, March, and July. The column ozone concentration

values show a predominantly decreasing trend in the spring, rising again to a second peak around June but continuing to decline from early autumn onwards. The difference in column ozone concentration values between the two peaks is about 4 DU. The annual average concentration value for 2022 is about 38.43 DU, and the annual average for 2023 is about 37.89 DU.

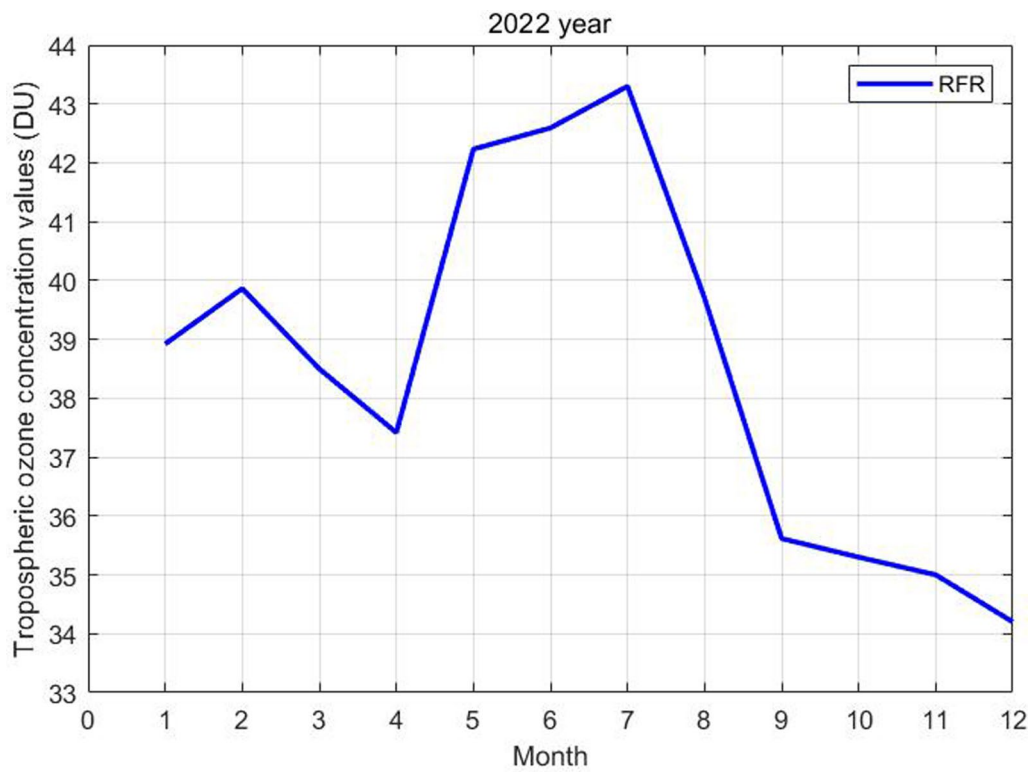


Fig. 9 Predicted monthly mean tropospheric ozone concentration in South-western Europe in 2022

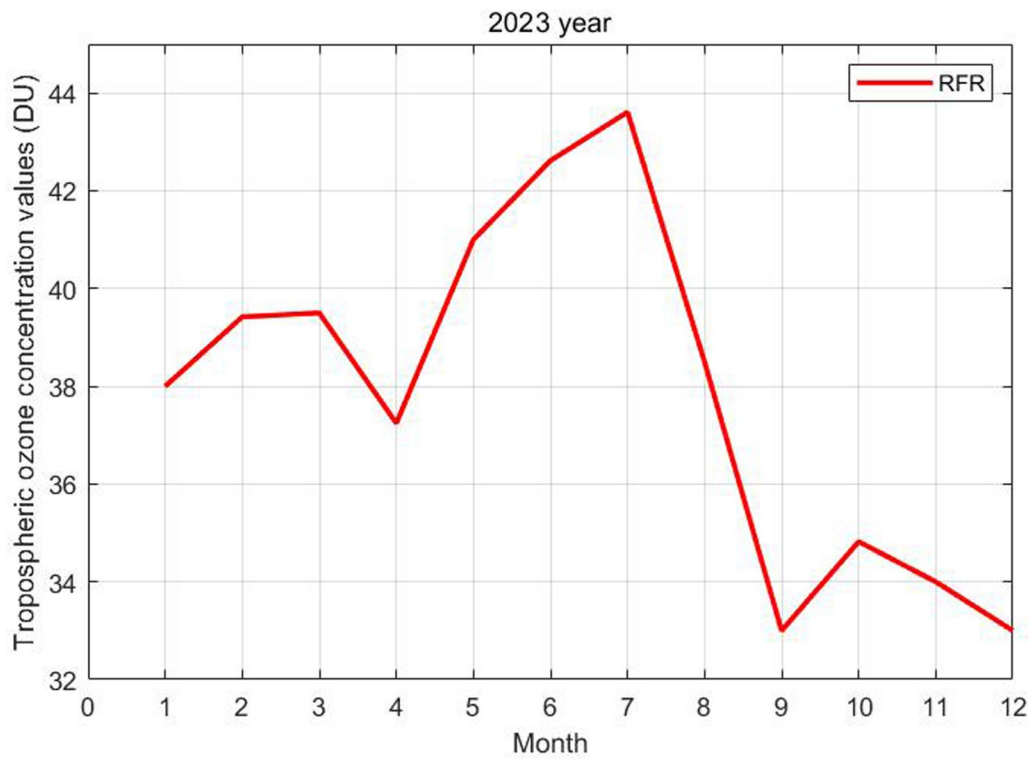


Fig. 10 Predicted monthly mean tropospheric ozone concentration in South-western Europe in 2023

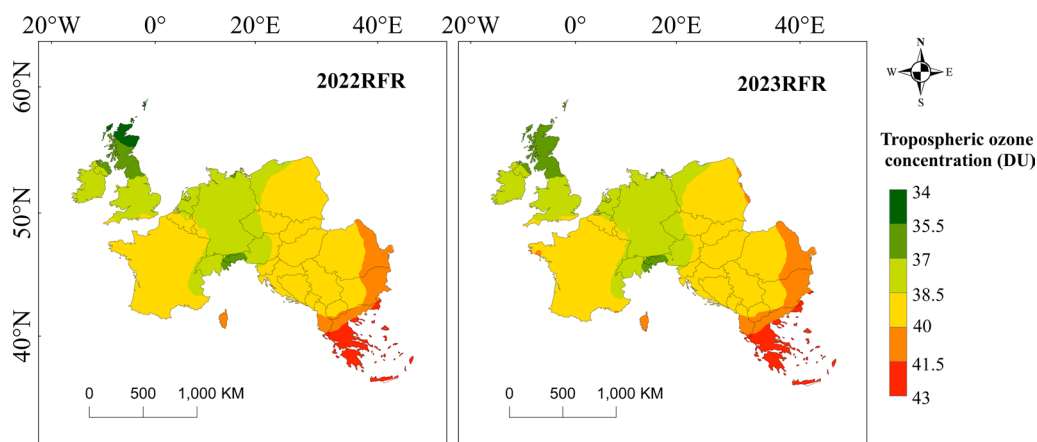


Fig. 11 South-western European annual mean for 2022 and 2023 predicted by Random Forest

As can be seen in Fig. 11, the predicted tropospheric column ozone concentrations for 2022 and 2023 are slightly lower compared to the 2021 concentrations, around 1–4 DU. However, the overall spatial distribution of the concentration values remains higher in the south-east than in the north-west. The spatial distribution of the column ozone concentrations in 2022 is slightly lower compared to the 2021 concentration values, around 3–6 DU. They are in the UK, western France, western Hungary, Croatia and north-western Serbia. Greece, western Bulgaria, and western Romania have higher ozone concentration values of about 40–43 DU, about 9.6% of the study area. Ma et al. [21] investigated the future trend changes and distribution of ozone concentrations in the Beijing–Tianjin–Hebei region using the Random Forest Prediction Model (RFPM). They concluded that ozone concentration values increase with increasing latitude. There are some discrepancies between the conclusions here and other studies. There are no significant differences between the 2023 predicted values compared to the 2022 predicted values. In terms of spatial distribution, there is an overall decrease of one order of magnitude in concentration values in the north of the UK compared to 2022.

From the above analysis, it can be seen that there is a clear "double peak" pattern in the monthly changes of concentration values in 2022 and 2023. Regarding annual average changes, the ozone concentration values from 2021 to 2023 show a decreasing trend yearly, with apparent spatial distribution characteristics. Concentration values in the southwestern and eastern regions are significantly higher than in other regions. Countries such as Bulgaria, Greece, Romania, Poland, and France need to pay more attention to the changes in ozone concentrations and take appropriate preventive, control, and management measures in due course.

Analysis of potential source contributions

In this paper, while studying the spatial distribution of tropospheric ozone concentration values, it was found that the distribution of high ozone concentration columns is concentrated in the southeastern coastal countries. Therefore, this study selected the industrial and commercial center of Patras (38°N, 21°E) in western Greece for potential source analysis. Daily mean column ozone concentrations were combined with daily 24-h backward trajectory routes, and the trajectory routes were divided into a $0.5^\circ \times 0.5^\circ$ grid. Potential source areas were mapped using MeteoInfo maps, and the images were optimized using Python to obtain the potential source areas of ozone for the four seasons in Patras. The larger the WPSCF value, the larger the area of the potential source area of O_3 , and the more significant the contribution to the O_3 concentration. To avoid the effect of the COVID-19, we selected the four seasons of O_3 in 2019 for the potential source contribution factor analysis.

As shown in Fig. 12, the main contribution of the potential source area centered on Patras contains the combined effect of its source area and the external source area module. Overall, there are seasonal differences between the four seasons of the potential source area. The contribution to ozone is higher in the northeast and southwest directions in summer and autumn, while the southern regions also contribute to the increase in ozone in autumn. Bulgaria, Macedonia, Serbia, and Italy contribute less to ozone. In the spring, the West contributes a significant influence on ozone. In winter, there is some influence on ozone concentrations in all directions except in the south-west. These areas include Albania, southern Bulgaria, and northern Greece. The regions with a more significant impact on ozone concentration values are located mainly in northern Greece. It is worth noting that during the summer months, the area of high impact

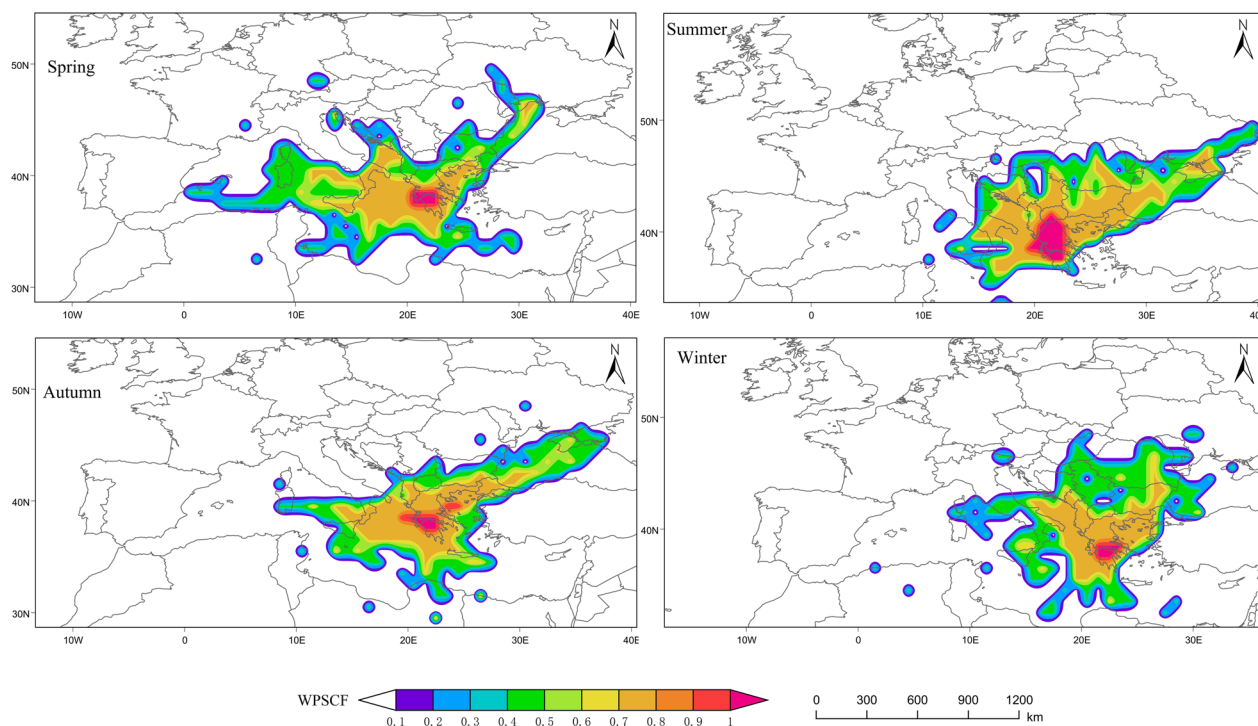


Fig. 12 Distribution of potential hazard sources in Patras in different seasons in 2019

on ozone is located at the border between southern Macedonia, south-eastern Serbia, and northern Greece. There is a high concentration of automotive and electrical industries, and some areas are transport hubs.

Border areas could consider developing joint national or interregional prevention, control, and management measures to reduce the risk of ozone pollution. In addition, countries in the region could synergistically reduce the generation of pollutants by introducing policies related to reducing carbon emissions, expanding green areas, promoting energy-saving retrofitting in sectors such as buildings and heavy industry and restructuring and optimizing the energy mix (development of non-petrochemical industries).

Number of premature deaths from diseases associated with ozone exposure and health assessments

In order to quantify the impacts of ozone pollution on human health, this study loaded annual averages of ozone surface concentrations in South-western Europe into BenMAP-CE. Firstly, the primary database for health effect assessment is constructed, including spatial grid setup, ozone surface data, loading of disease data and population spatial data, and construction of health effect function. The spatial distribution of grid-scale ozone concentrations was then modeled. The VNA spatial interpolation method calculated Inverse distance weights for the

defined spatial grids [14]. Finally, a model-based assessment of the number of premature deaths due to ozone pollution in the study area.

As shown in Fig. 13 and Table 2, the number of all-cause premature deaths in the United Kingdom, Austria, Poland, France, Switzerland, and Slovenia was higher in 2019 than in 2018, and Germany was the country with the highest number of all-cause premature deaths in both years. Bulgaria, Bosnia and Herzegovina, Germany, The Netherlands, Czech Republic, Romania, Luxembourg, Macedonia, Serbia, Slovakia, and Greece have all experienced a decrease in all-cause premature deaths. The rates of decrease were 39.62%, 42.62%, 15.6%, 48.82%, 54.18%, 64.16%, 67.97%, 30.09%, 23.76%, 9.06%, and 2.9%, respectively. There are more premature deaths from cardiovascular diseases than respiratory diseases in the study area. The highest number of premature deaths from cardiovascular diseases due to ozone pollution is in Germany, and the country with the most significant impact on respiratory diseases is France, with 11,049 premature deaths in 2019. The spatial distribution of premature deaths shows that the areas of high values leading to all-cause premature deaths, premature deaths from cardiovascular diseases, and premature deaths from respiratory diseases are predominantly located in the more developed countries in the northwest of the study area (UK, France, Germany). The remaining countries, such as

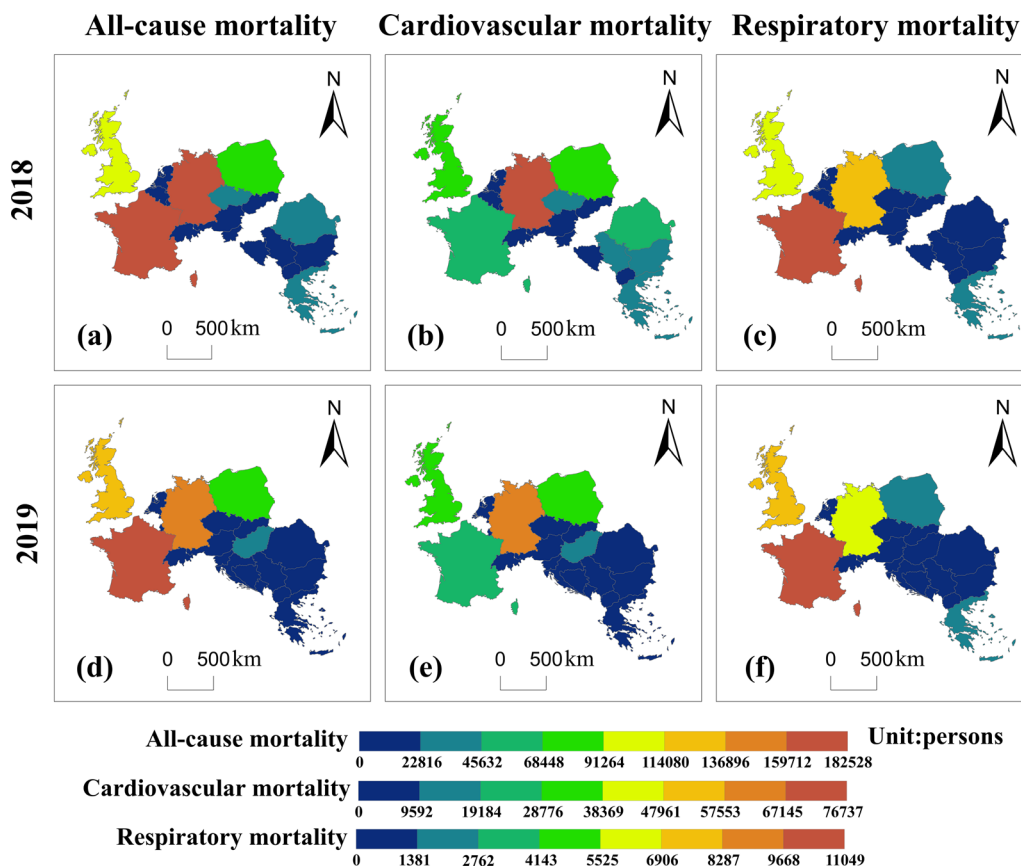


Fig. 13 The map for visualization of early death values for ozone pollution in the South-western Europe. **a** All-cause mortality in 2018. **b** Cardiovascular mortality in 2018. **c** Respiratory mortality in 2018. **d** All-cause mortality in 2019. **e** Cardiovascular mortality in 2019. **f** Respiratory mortality in 2019

Poland, Hungary, and Greece, are also affected to some extent. Lei et al. using the BenMAP-CE model to assess the health effects of ozone in the Qin-Jin region of China, found that the number of premature deaths due to cardiovascular diseases under ozone exposure was higher than the number of premature deaths due to respiratory diseases. This finding is consistent with the conclusions of this paper. Yang et al. (2019) used the BenMAP-CE model to analyze the US region’s health and economic impacts of O₃ and PM_{2.5}. It was found that with a concentration path at 4.5, an average annual reduction in O₃ could prevent about 1800 all-cause deaths in the future.

Correlation analysis between ozone and meteorological factors and NO₂

In this study, the correlations between two meteorological factors, precipitation and temperature, and the concentrations of nitrogen dioxide (NO_x is an ozone precursor, and nitrogen dioxide is used here as an example) and tropospheric column ozone concentrations were analyzed using Pearson’s correlation in order to

investigate their respective effects on atmospheric tropospheric column ozone concentrations.

Figure 14a shows the spatial correlation between precipitation and column ozone concentration. The correlation between column ozone concentration and precipitation in the study area was calculated using the Pearson correlation coefficient method. The results show that the correlation differs significantly from north to south, with a negative correlation in the northern part of the study area and a positive correlation in the southern part. A positive correlation was observed in about 45.8% of the study area, but it was weakened in the south-eastern part of Greece. There is a complex positive and negative correlation between the effects of water vapor content on tropospheric O₃. Precursor gas abundance is positively correlated with OH-controlled O₃ production and negatively correlated with OH-induced O₃ destruction [4]. The correlation between precipitation and ozone is small in northern France, southern Germany, northern Austria, Slovakia, and northern Romania.

Table 2 Values of premature deaths from ozone pollution in south-west Europe (by country)

Region	All-cause mortality		Cardiovascular mortality		Respiratory mortality	
	(in persons, 95%CI)		(in persons, 95%CI)		(in persons, 95%CI)	
	2018	2019	2018	2019	2018	2019
United Kingdom	104,769 (55,966, 150,254)	125,284 (67,640, 177,805)	29,831 (17,504, 41,234)	35,499 (21,064, 48,534)	6733 (4214, 13,739)	8066 (5084, 16,123)
Austria	11,035 (5831, 15,996)	20,181 (11,006, 28,358)	5024 (2915, 7024)	8997 (5396, 12,173)	399 (248, 832)	756 (480, 1482)
Bulgaria	16,768 (8918, 24,155)	10,124 (5296, 14,824)	13,461 (7862, 18,694)	7572 (4346, 10,698)	371 (232, 764)	205 (127, 436)
Belgium	12,471 (6557, 18,169)	–	3784 (2183, 5318)	–	633 (392, 1332)	–
Bosnia and Herzegovina	5828 (3099, 8395)	3344 (1746, 4906)	3387 (1978, 4704)	1682 (964, 2382)	170 (106, 350)	89 (55, 190)
Germany	182,528 (98,078, 260,258)	154,056 (82,145, 221,335)	76,737 (45,306, 105,427)	63,323 (37,085, 87,696)	7466 (4691, 15,062)	6199 (3875, 12,694)
Poland	76,453 (41,004, 109,213)	79,909 (42,985, 113,814)	36,066 (21,252, 49,647)	36,588 (21,627, 50,209)	1431 (898, 2898)	1391 (875, 2801)
France	176,202 (97,525, 244,104)	178,720 (98,994, 247,413)	20,678 (12,594, 27,568)	22,628 (13,792, 30,145)	10,857 (6955, 20,696)	11,049 (7081, 21,030)
Montenegro	–	182 (94, 270)	–	64 (36, 91)	–	6 (4, 13)
Netherlands	15,209 (7968, 22,238)	7784 (4028, 11,525)	4399 (2529, 6206)	2188 (1241, 3128)	714 (441, 1513)	352 (216, 765)
Czech Republic	27,730 (15,140, 38,926)	12,705 (6681, 18,508)	13,850 (8315, 18,720)	6280 (3624, 8826)	864 (548, 1689)	385 (239, 810)
Croatia	–	12,705 (6936, 17,836)	–	6224 (3737, 8413)	–	403 (256, 789)
Romania	32,216 (16,980, 46,824)	11,547 (5967, 17,119)	21,304 (12,324, 29,865)	7916 (4485, 11,332)	840 (521, 1759)	317 (194, 690)
Luxembourg	615 (327, 886)	197 (102, 292)	210 (122, 291)	64 (36, 92)	31 (19, 64)	10 (6, 22)
Macedonia	3081 (1640, 4435)	2154 (1130, 3144)	1950 (1140, 2705)	1058 (609, 1490)	91 (57, 187)	66 (41, 139)
Switzerland	8476 (4473, 12,303)	12,084 (6469, 17,293)	3046 (1764, 4264)	4131 (2429, 5697)	295 (183, 617)	436 (273, 886)
Serbia	16,526 (8809, 23,752)	12,600 (6646, 18,301)	9987 (5847, 13,836)	7095 (4107, 9938)	470 (294, 963)	357 (222, 747)
Slovakia	9295 (4966, 13,327)	8453 (4501, 12,162)	4601 (2700, 6358)	4199 (2455, 5824)	163 (102, 333)	134 (84, 275)
Slovenia	2934 (1556, 4240)	4290 (2317, 6087)	1316 (766, 1833)	1907 (1132, 2607)	67 (42, 139)	116 (73, 231)
Greece	33,071 (18,218, 46,024)	32,112 (18,019, 45,323)	14,115 (8554, 18,908)	13,518 (8167, 18,163)	1476 (943, 2840)	1500 (956, 2901)
Hungary	–	25,186 (13,545, 35,879)	–	14,367 (8491, 19,720)	–	1078 (678, 2170)

Figure 14b shows the spatial correlation between temperature and column ozone concentration. It can be seen from the figure that temperature is positively correlated with ozone concentration in most parts of southwestern Europe, with the positively correlated area accounting for 57.65% of the entire study area. The negative correlation is mainly concentrated in southern Germany, Poland, Macedonia, and western Austria.

Figure 14c shows the spatial correlation between NO₂, one of the ozone precursors, and ozone production. From the figure, it can be seen that all areas with higher ozone concentrations show a significant

positive correlation with NO₂, thus indicating that NO₂ concentration strongly influences ozone production in this study area. The south-eastern region includes the countries of Romania, Bulgaria, and North Macedonia, where the primary economic industries are petrochemicals, machinery, industry, manufacturing of hydraulic tools, metallurgy, and electrical equipment, respectively. Industrial pollutant emissions, harmful emissions from motor vehicles, and oil combustion all contribute to elevated NO_x concentrations, resulting in high levels of ozone pollution. Therefore, most areas with high ozone pollution

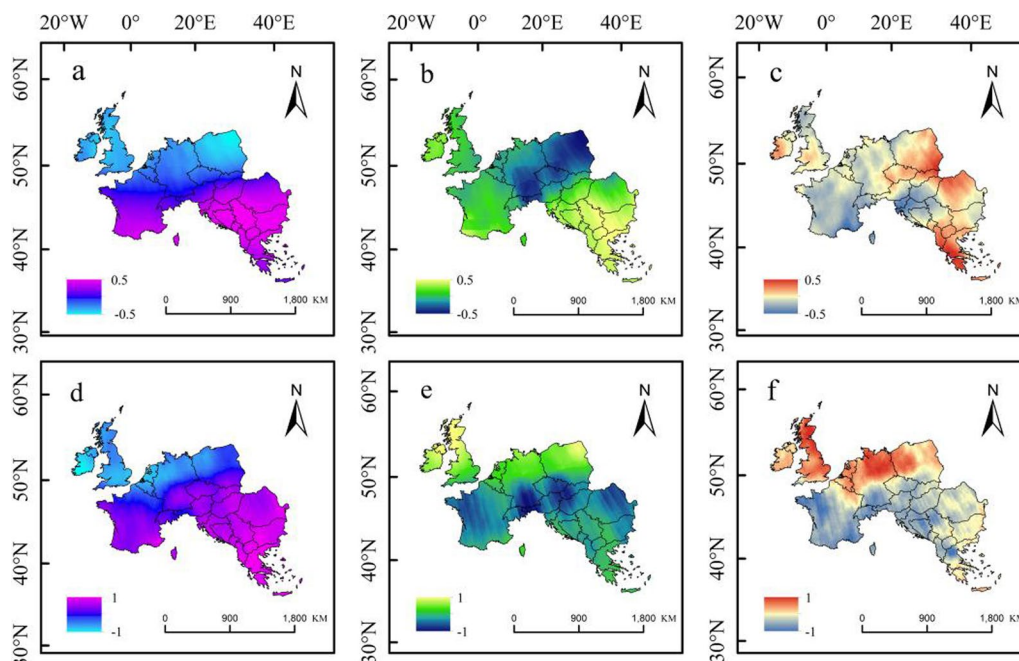


Fig. 14 Spatial correlation between column tropospheric ozone concentration and precipitation (a), atmospheric pressure (b) and NO_2 (c). Correlations between summertime tropospheric ozone and precipitable water (d), air temperature (e), and NO_2 (f), respectively

positively correlate with NO_2 . However, the relationship between ozone and the photochemical reaction concentrations of NO_x and VOC_5 is not simply linear but non-linear [1]. Nitrogen oxides and ozone react with each other under certain conditions, and each sector causes a change in the concentration of nitrogen oxides. This variation may cause a negative correlation between NO_2 and ozone. Hence, a negative correlation exists in eastern Bulgaria and southeastern Romania—the exact reasons for this need to be further explored. Marin et al. [20] reached similar conclusions in their study of the Bucharest region.

Tropospheric ozone concentration values in summer are the highest of the four seasons. Figure 14d–f correlates the three influencing factors with the summer ozone concentration values, respectively. The areas with high values of precipitable water–ozone correlation are distributed in the spring in western Bulgaria and the northern parts of Macedonia and Greece. In summer, the regions with high temperature–ozone correlations are concentrated in the northern part of the UK, northern Germany, north-eastern Poland, and south-eastern regions, in contrast to the year-round correlations. Nitrogen dioxide is a precursor of ozone, and the areas contributing to elevated ozone concentrations during the summer months are concentrated in the coastal areas of the UK, Germany, Poland, and the south-eastern part of the study area. This may be

due to high tropospheric ozone concentrations in the near-equatorial Atlantic in July due to the presence of the Azores anticyclone and its associated photochemistry and dynamics [16].

Impact of human activities on the concentration of ozone

The factors are scaled differently in multivariate analyses, and the variables are not independent. Therefore, exploring the influence of the study variable X_i on the dependent variable through the variable X_j is necessary. This paper uses the through-analysis method to investigate the direct and indirect effects of the influencing factors on the concentration of column ozone. Table 3 shows different anthropogenic factors' direct and indirect path coefficients on column ozone concentrations. Table 3 shows that the total coefficient of determination R^2 for anthropogenic ozone is 0.9996, and the residual path coefficient is 0.00601. The direct path coefficients for each variable from X to Y : $X_1-Y=0.3868$, $X_2-Y=1.1433$, $X_3-Y=-0.4496$, $X_4-Y=1.2827$, $X_5-Y=1.4749$, $X_6-Y=0.2878$, $X_7-Y=0.0442$. From the direct path coefficients, population density, total greenhouse gas emissions, and forest area significantly affect ozone. The indirect path coefficients from X_1 , X_2 , and X_6 to X_4-Y show that there is a contribution of population size and agglomeration to the production of greenhouse gases and a contribution of $\text{PM}_{2.5}$ emissions to the production of greenhouse gases. The indirect paths

Table 3 Path coefficient analysis between tropospheric ozone and human influencing factors

Item	Direct path coefficient	Indirect connection diameter coefficients						
		X_1-Y	X_2-Y	X_3-Y	X_4-Y	X_5-Y	X_6-Y	X_7-Y
Population, total (X_1)	0.3868	-	-0.3027	-0.1641	-0.5712	0.6490	-0.0340	0.0197
Population density (X_2)	1.1433	-0.0983	-	0.0040	-0.3903	-0.7191	0.1348	-0.0082
PM _{2.5} air pollution (X_3)	-0.4496	0.1388	-0.0100	-	0.4179	-0.3456	0.0086	0.0006
Total greenhouse gas emissions (kt of CO ₂ equivalent) (X_4)	1.2827	-0.1685	-0.3365	-0.1525	-	-0.3805	-0.0714	-0.0063
Forest area (% of land area) (X_5)	1.4749	0.1765	-0.5767	0.1151	-0.3508	-	-0.2049	0.0196
Urban population (X_6)	0.2878	-0.0454	0.5523	-0.0131	-0.3181	-0.9621	-	-0.0204
Other manufacturing (% of value added in manufacturing) (X_7)	0.0442	0.1767	-0.2002	-0.0061	-0.1874	0.6273	-0.1373	-

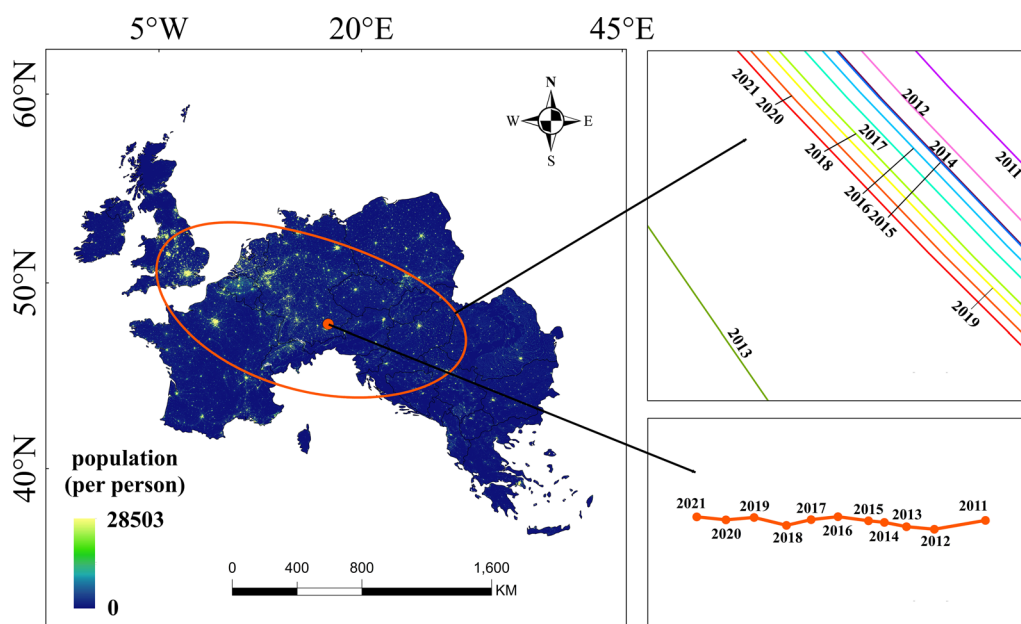


Fig. 15 Spatial distribution of population, standard deviation ellipse, and center-of-gravity migration trajectory in South-western Europe

of X_5 to X_3-Y show that forests contribute to the pollution of PM_{2.5} to some extent. Ozone concentrations were found to be positively correlated with visible foliar symptoms across Europe, and depending on the frequency of visible foliar symptoms, the risk of ozone to vegetation may be higher in the Alps and parts of continental Europe than in the Mediterranean region [9].

Overall, anthropogenic impacts on ozone are ranked in order of direct pathway coefficients: forested land area > greenhouse gas emissions > population density > PM_{2.5} air pollution > total population size > urban population density > other manufacturing (% of value added in manufacturing). Regarding indirect pathways, population and distribution density in south-west Europe may be more sensitive to ozone production, which is

related to other factors. Therefore, reducing population density, reducing forest plantations, coal, and oil energy consumption, and optimizing the energy mix can positively impact ozone pollution mitigation.

Based on the conclusions of the throughput analyses, it was found that population strongly influences ozone production. Therefore, this paper provides further analyses of population changes. As shown in Fig. 15, standard deviation ellipses and center-of-gravity migration trajectories were made in ArcGIS using point data on population distribution in South-western Europe from 2010 to 2021. The long half-axis of the ellipse indicates the direction of the main population distribution, and the short half-axis indicates the extent of the main population distribution; the more significant the difference between

the values of the long and short half-axes, the more pronounced the direction of the main population. The center-of-gravity point indicates the center of the whole data. As shown in Fig. 15, the center of gravity of the population has been broadly distributed over the 12 years in south-eastern Germany, in the Bavarian region. In particular, from 2010 to 2021, the population's center of gravity shows a regular migration from east to west year by year. At the same time, migration distances are short, with relatively long migration distances in 2011–2012. Except for 2013, the area of standard deviation ellipse is decreasing continuously in each year from 2010 to 2021. This situation indicates a robust spatial location clustering of population patches in the study area. This may be related to local solid manufacturing, electronic information, and the influence of the financial economy.

Conclusions

There is a clear pattern in the spatial distribution of column ozone concentrations in southwest Europe, with high-value areas mainly in the UK, Poland, France, and the southeast coast. After 2018, the heatwave exacerbated ozone pollution in most of the countries of southwest Europe, with elevated values in all areas except the central part of the study area. High-value areas were mainly in Greece, Albania, and parts of Bulgaria. Overall, the spatial distribution is characterized by significant but not large changes.

Regarding inter-annual variability, tropospheric column ozone concentrations peaked in 2013, 2018, and 2019 and were at lower levels in the remaining years. Seasonally, the seasonal mean of ozone column concentration showed summer (42.76 DU) > spring (40.5 DU) > winter (38.26 DU) > autumn (34.18 DU). Regarding monthly variation, the monthly mean values showed a "bimodal" pattern during the study period, ranging from 30 to 53 DU. Concentrations in the early part of the period were stable. Then, they declined rapidly in 2020, with high ozone values from March to May in areas such as Poland and northern Germany and high concentrations from June to August in the southern part of the study area.

The slope value for the southeastern part of southwest Europe is 0.026 with no significant increasing trend. The long-term correlation of the time-series trend Hurst index shows that the study area as a whole shows moderate persistence and that the southern part of the study area is likely to see an increase in ozone column concentrations in the future, and in particular, attention needs to be paid to the southeastern coastal areas for enhanced monitoring. In the Random Forest Forecasts (RFF), the monthly changes in the concentration values for the years 2022 and 2023 show a clear

"bimodal" pattern. Predicted column ozone concentrations in 2022 and 2023 show a slight decrease of about 1–4 DU compared to concentrations in 2021. Exposed ozone contributes to premature human deaths; premature deaths from cardiovascular diseases are generally higher than premature deaths from respiratory diseases.

The potential source area for columnar tropospheric ozone concentrations in the Patras region is influenced by national sources and neighboring countries. During the summer months, when ozone pollution is at its highest, this concentration value is jointly influenced by three countries, Greece, Macedonia, and Serbia, so that the establishment of joint national or regional prevention, control, and management measures in this area could be considered in order to maintain the health of the atmospheric environment. Based on the correlation analysis of the influencing factors, it can be seen that both the long-time and short-time series of the effect of precipitation on ozone concentration are negatively correlated in the northern part of the study area. The correlation between ozone concentration and temperature and NO₂ concentration differed significantly between summer and year-round. From the results of the thorough analysis, it can be learned that the population size and distribution density in southwestern Europe may be more sensitive to ozone production than other factors.

In addition, there are areas where further refinement is needed in this paper. In this paper, we use a random forest model to predict the tropospheric ozone concentration values for the next two years, and we use the ozone concentration value data from 2011 to 2020 for training to get the predicted values for 2021. Based on the column ozone concentration data and the annual average spatial distribution of the concentration values, the influence of the New Crown epidemic on the ozone concentration values is relatively small. Therefore, the year 2020 is included in this prediction model. This may explain the slight bias in the model predictions, and the model's accuracy may be further improved if the effect of the New Crown epidemic is excluded. In the health benefit analysis section, data for the study area must be supplemented due to missing data for some countries.

Author contributions

The author of this paper is JW, the first author, mainly responsible for data collection and processing, mapping, analysis and discussion, and paper writing. The responsible author is Professor TJ, who is mainly responsible for the topic selection and writing guidance. The co-author is BL, who is mainly responsible for the guidance and language polishing of the paper. The other co-authors, XX, CH and JZ were mainly responsible for formatting and paper fluency.

Funding

This work was supported by the Natural Science Foundation of Gansu Province (CN) (17YF1FA120) at the Key Laboratory of Resource Environment and

Sustainable Development of Oasis, Gansu Province. Environmental Science and Engineering Experiment (Practical Training) Teaching Center (202018).

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Code and materials availability

The air pollutant data used during this study are from the NASA, and other datasets analyzed are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

Ethics approval was not required for this research. This paper does not require consent to participate.

Consent for publication

The first author and the responsible author agree to publish the paper.

Competing interests

The authors have no relevant financial or non-financial interests to disclose.

Author details

¹College of Geography and Environmental Sciences, Northwest Normal University, No. 27, Qiujiawan, Anning District, Lanzhou City 730070, Gansu Province, China. ²The Key Laboratory of Resource Environment and Sustainable Development of Oasis, Lanzhou 730000, Gansu Province, China. ³Faculty of Atmospheric Remote Sensing, Shaanxi Normal University, Xi'an 710062, China.

Received: 6 November 2023 Accepted: 11 February 2024

Published online: 26 March 2024

References

- Archibald AT, Levine JG, Abraham NL et al (2011) Impacts of HOx regeneration and recycling in the oxidation of isoprene: consequences for the composition of past, present and future atmospheres. *Geophys Res Lett* 38(5). <https://doi.org/10.1029/2010gl046520>
- Asgari M, Yang W, Farnaghi M (2022) Spatiotemporal data partitioning for distributed random forest algorithm: air quality prediction using imbalanced big spatiotemporal data on spark distributed framework. *Environ Technol Innov* 27:102776
- Achebak H, Petetin H, Quijal-Zamorano M, Bowdalo D, Pérez García-Pando C, Ballester J (2021) Trade-offs between short-term mortality attributable to NO and O changes during the COVID-19 lockdown across major Spanish cities. *Environ Pollut* (Barking, Essex) 286:117220. <https://doi.org/10.1016/j.envpol.2021.117220>
- Ajayakumar RS, Nair PR, Girach IA, Sunilkumar SV, Muhsin M, Satheesh Chandran PR (2019) Dynamical nature of tropospheric ozone over a tropical location in Peninsular India: role of transport and water vapour. *Atmos Environ* 218:117018. <https://doi.org/10.1016/j.atmosenv.2019.117018>
- Bell ML, Dominici F, Samet JM (2005) A meta-analysis of time-series studies of ozone and mortality with comparison to the national morbidity, mortality, and air pollution study. *Epidemiology* 16:436–445
- Blanco-Ward D, Ribeiro AO, Paoletti E, Miranda AI (2021) Assessment of tropospheric ozone phytotoxic effects on the grapevine (*Vitis vinifera* L.): a review. *Atmos Environ* 244:117924. <https://doi.org/10.1016/j.atmosenv.2020.117924>
- Díaz J, Ortiz C, Falcón I, Salvador C, Linares C (2018) Short-term effect of tropospheric ozone on daily mortality in Spain. *Atmos Environ* 187:107–116. <https://doi.org/10.1016/j.atmosenv.2018.05.059>
- Fuks KB, Hüls A, Sugiri D, Altug H, Vierkötter A, Abramson MJ, Goebel J, Wagner G, Demuth I, Krutmann J, Schikowski T (2019) Tropospheric ozone and skin aging: results from two German cohort studies. *Environ Int* 124:139–144. <https://doi.org/10.1016/j.envint.2018.12.047>
- Ferretti M, Cailleret M, Haeni M, Trotsiuk V, Apuhtin V, Araminiene V, Buriánek V, Cecchini S, Dalstein-Richier L, Hůnová I, Jakovljević T, Kaoukis K, Neirync J, Nicolas M, Prescher A, Novotny R, Pavlendová H, Potočić N, Rupel M, Russ A, Stakėnas V, Verstraeten A, Vollenweider P, Žilindra D, Pitar D, Calatayud V, Gottardini E, Schaub M (2024) The fingerprint of tropospheric ozone on broadleaved forest vegetation in Europe. *Ecol Ind*. <https://doi.org/10.1016/j.ecolind.2023.111486>
- Gebhardt H, Zimmermann F, Matschullat J (2020) 1981–2020 winter ozone trends, Erzgebirge, Central Europe. *Chem Erde* 1:125738. <https://doi.org/10.1016/j.chemer.2020.125738>
- Huang R, Ju T, Dong H et al (2021) Analysis of atmospheric SO₂ in Sichuan-Chongqing region based on OMI data. *Environ Monit Assess* 193(12):849. <https://doi.org/10.1007/s10661-021-09638-2>
- Huang YY, Yang D, Feng L (2019) Spatial and temporal changes in vegetation cover and its drivers in Ningxia from 2000–2016. *J Ecol* 38(08):2515–2523. <https://doi.org/10.13292/j.1000-4890.201908.016>
- Hsu YK, Holsen TM, Hopke PK (2003) Comparison of hybrid receptor models to locate PCB sources in Chicago. *Atmos Environ* 37(4):545–562. [https://doi.org/10.1016/S1352-2310\(02\)00886-5](https://doi.org/10.1016/S1352-2310(02)00886-5)
- Jonson JE, Simpson D, Fagerli H, Solberg S (2006) Can we explain the trends in European ozone levels? *Atmos. Chem Phys* 6:51–66. <https://doi.org/10.5194/acp-6-51-2006>
- Juran S, Grace J, Urban O (2021) Temporal changes in ozone concentrations and their impact on vegetation. *Atmosphere* 12(1):82. <https://doi.org/10.3390/atmos12010082>
- Kulkarni PS, Bortoli D, Salgado R, Antón M, Costa MJ, Silva AM (2011) Tropospheric ozone variability over the Iberian Peninsula. *Atmos Environ* 45:174–182. <https://doi.org/10.1016/j.atmosenv.2010.09.029>
- Rozbicka K, Majewski G, Rogula-Kozłowska W, Rozbicki T (2020) Tropospheric ozone assessment in urban environment—Warsaw case study of selected heat waves. *J Atmos Solar Terr Phys* 209:105418. <https://doi.org/10.1016/j.jastp.2020.105418>
- Li SH (2018) Evaluation of health and economic benefits under PM_{2.5} and O₃ standards in Beijing-Tianjin-Hebei Region Based on BenMAP. Ph.D. Thesis, Tianjin Normal University, Tianjin, China.
- Shimin L, Tongqin W, Zhaoli L et al (2020) Correlation and pathway analysis of agronomic traits and yield of maize varieties tested in the Yellow and Huaihai Sea. *Anhui Agric Sci* 48(21):30–32. <https://doi.org/10.3969/j.issn.0517-6611.2020.21.009>
- Marin CA, Mureanu L, Radu C et al (2019) Wintertime variations of gaseous atmospheric constituents in Bucharest Peri-Urban Area. *Atmosphere* 10(8):478. <https://doi.org/10.3390/atmos10080478>
- Ma R, Ban J, Wang Q, Zhang Y, Yang Y, He MZ, Li S, Shi W (2021) Random forest model based fine scale spatiotemporal O₃ trends in the Beijing-Tianjin-Hebei region in China, 2010 to 2017. *Environ Pollut* 276:116635. <https://doi.org/10.1016/j.envpol.2021.116635>
- Nguyen TH, Cappelli G, Emberson L, Ignacio GF, Irimescu A, Francesco S, Fabrizio G, Booth N, Boldeanu G, Bermejo V, Bland S, Frei M, Ewert FA, Gaiser T (2024) Assessing the spatio-temporal tropospheric ozone and drought impacts on leaf growth and grain yield of wheat across Europe through crop modeling and remote sensing data. *Eur J Agron*. <https://doi.org/10.1016/j.eja.2023.127052>
- Pan YJ, Wang YL, Peng J, Shen H, Liu SQ (2012) Time series analysis of precipitation in the middle and lower reaches of Han River basin based on wavelet and R/S methods. *Geogr Res* 31(05):811–820. <https://doi.org/10.11821/yj2012050005>
- Peng S, Ju T, Liang Z et al (2022) Analysis of atmospheric ozone in Fenwei Plain based on remote sensing monitoring. *Environ Monit Assess* 194:412. <https://doi.org/10.1007/s10661-022-10082-z>
- Polissar AV, Hopke PK, Patero P et al (1999) The aerosol at Barrow, Alaska: long-term trends and source locations. *Atmos Environ* 33(16):2441–2458. [https://doi.org/10.1016/S1352-2310\(98\)00423-3](https://doi.org/10.1016/S1352-2310(98)00423-3)
- Pey J, Cerro JC (2022) Reasons for the observed tropospheric ozone weakening over south-western Europe during COVID-19: strict lockdown versus the new normal. *Sci Total Environ* 833:155162–155162. <https://doi.org/10.1016/j.scitotenv.2022.155162>
- Draxler RR, Hess GD (1998) An overview of the hysplit-4 modeling system for trajectories. *Aust Meteorol Mag* 47(4):295–308
- Rix M, Valks P, Hao N et al (2012) Volcanic SO₂, BrO and plume height estimations using GOME-2 satellite measurements during the eruption of

- Eyjaflajökull in May 2010. *J Geophys Res* 117:D00U19. <https://doi.org/10.1029/2011JD016718>
29. Sabolis A, Meskhidze N, Curci G, Palmer PI, Gantt B (2011) Interpreting elevated space-borne HCHO columns over the Mediterranean Sea using the OMI sensor. *Atmos Chem Phys* 11:12787–12798. <https://doi.org/10.5194/acp-11-12787-2011>
 30. Schulz G, Sanders T, Voynova YG, Bange HW, Dähnke K (2023) Seasonal variability of nitrous oxide concentrations and emissions in a temperate estuary. *Biogeosciences* 20:3229–3247. <https://doi.org/10.5194/bg-20-3229-2023>
 31. Shen ST, Yi MJ, Dai HX et al (2020) Analysis of PM_{2.5} transport path and potential source contribution in Bengbu city. *J Hefei Univ Technol (Nat Sci)* 43(4):543–551. <https://doi.org/10.3969/j.issn.1003-5060.2020.04.018>
 32. Xiaoyuan S, Zhongyuan Z, Shengwei Z et al (2016) Passage analysis of the influence of meteorological factors on snowmelt runoff in Xilin River. *China Soil Water Conserv Sci* 14(05):74–81. <https://doi.org/10.16843/j.sswc.2016.05.010>
 33. Sicard P, Khaniabadi YO, Perez S, Gualtieri M, De Marco A (2019) Effect of O₃, PM₁₀ and PM_{2.5} on cardiovascular and respiratory diseases in cities of France, Iran and Italy. *Environ Sci Pollut Res Int* 26(31):32645–32665. <https://doi.org/10.1007/s11356-019-06445-8>
 34. Schwartz JD, Yitshak-Sade M, Zanobetti A, Di Q, Requia WJ, Dominici F, Mittleman MA (2021) A self-controlled approach to survival analysis, with application to air pollution and mortality. *Environ Int* 157:106861. <https://doi.org/10.1016/j.envint.2021.106861>
 35. Turner M, Jerrett M, Pope C, Krewski D, Gapstur S, Diver W, Beckerman B, Marshall J, Su J, Crouse D et al (2015) Long-term ozone exposure and mortality in a large prospective study. *Am J Respir Crit Care Med* 193:1134–1142. <https://doi.org/10.1164/rccm.201508-1633OC>
 36. US EPA. BenMAP User's Manual [EB/OL]. Available online: <http://www2.epa.gov/benmap/manual-and-appendices-benmap-ce>. Accessed 17 Apr 2023.
 37. Yun X, Shen G, Shen H, Meng W, Chen Y, Xu H, Ren Y, Zhong Q, Du W, Ma J, Cheng H, Wang X, Liu J, Wang X, Li B, Hu J, Wan Y, Tao S (2020) Residential solid fuel emissions contribute significantly to air pollution and associated health impacts in China. *Sci Adv* 6(44):eaba7621. <https://doi.org/10.1126/sciadv.aba7621>
 38. Wang GC, Wang DQ, Chen ZL (2016) PM_{2.5} pollution characteristics and transport paths and potential source areas during severe pollution in Beijing in winter. *China Environ Sci* 36(7):1931–1937
 39. Wang QY, Luo DH, Wang JN (2016) Characteristics of summer heat waves in the European region and their connection with blocking circulation. *Clim Environ Res* 21(04):367–379. <https://doi.org/10.3878/j.issn.1006-9585.2015.15022>
 40. Wang XL, Hu BQ, Xia J (2002) R/S analysis method for hydrological time series trends and variation points. *J Wuhan Univ (Eng Edn)* 02:10–12. <https://doi.org/10.3969/j.issn.1671-8844.2002.02.003>
 41. Xiong K, Xie X, Mao J, Wang K, Huang L, Li J (2023) Improving the accuracy of O₃ prediction from a chemical transport model with a random forest model in the Yangtze River Delta region, China. *Environ Pollut* 319:120926. <https://doi.org/10.1016/j.envpol.2022.120926>
 42. Yao Y, Lan Y, Zhang HY, Zhang M (2021) Effectiveness, experience and inspiration of air pollution prevention and control in Europe. *Environ Sustain Dev* 46(06):176–180. <https://doi.org/10.19758/j.cnki.issn1673-288x.202106176>
 43. Yang P, Zhang Y, Wang K, Doraiswamy P, Cho SH (2019) Health impacts and cost-benefit analyses of surface O₃ and PM_{2.5} over the U.S. under future climate and emission scenarios. *Environ Res* 178:108687. <https://doi.org/10.1016/j.envres.108687>
 44. You JW, Zou B, Zhao XG et al (2019) Estimation of near-surface NO₂ concentration in China based on random forest model. *China Environ Sci* 39(03):969–979
 45. Zheng J, Yuan Y (2017) Insights from the latest European atmospheric environmental quality and its prevention and control policies. In: *Proceedings of the 2017 Annual Conference of the Chinese Society of Environmental Science on Science and Technology* pp. 665–667
 46. Zohdirad H, Jiang J, Aksoyoglu S, Namin MM, Ashrafi K, Prevot AS (2022) Investigating sources of surface ozone in central Europe during the hot summer in 2018: high temperatures, but not so high ozone. *Atmos Environ*. <https://doi.org/10.1016/j.atmosenv.2022.119099>

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.