

Linking Surface Pollutant Regimes and Remote Sensing through Atmospheric Chemical Transformation and Transport

Megi Caushaj^{1*}, Florian Mandija², Dhurata Premti³
¹Department of Industrial Chemistry, Faculty of Sciences, University of Tirana, Tirana, Albania

²National Center for Scientific Research (CNRS), Pierre-Simon Laplace Institute (IPSL), University of Versailles-Saint-Quentin-en-Yvelines (UVSQ), Paris, France

³Environmental Engineering Department, Faculty of Civil Engineering, Polytechnic University of Tirana, Tirana, Albania

*Corresponding email: megi_caushaj.fshnutdr@unitir.edu.al

Received: 12.07.2026; Accepted: 24.09.2026; Available online: 28.09.2026; Published: 30.09.2026

Cite this article: Caushaj, M., Mandija, F., & Premti, D. (2026). Linking Surface Pollutant Regimes and Remote Sensing through Atmospheric Chemical Transformation and Transport. *Trends in Ecological and Indoor Environmental Engineering*, 4(3), 65–83.

Background: Atmospheric pollution at the surface reflects interacting effects of emissions, chemical transformation, secondary formation, and atmospheric transport. Ground-based measurements provide detailed local information, whereas remote sensing describes atmospheric conditions over broader spatial domains. However, differences between surface and column observations complicate their interpretation. A process-based framework is therefore needed to relate recurrent surface pollutant regimes to atmospheric-column variability. **Objectives:** This study aims to develop and evaluate an exploratory process-based framework linking surface pollutant regimes with column-integrated atmospheric observations through chemical transformation, secondary formation, vertical redistribution, and transport. **Methods:** Hourly surface observations from an urban-background monitoring station in Tirana for 2010–2012 were analysed using principal component analysis (PCA), non-negative matrix factorization (NMF), and k-means clustering. The dataset comprised 16,632 observations of 11 pollutant variables. PCA identified dominant covariance structures, NMF characterized non-negative pollutant co-variation patterns, and clustering identified recurrent multivariate regimes. Separately, ERA5/GIOVANNI data for 2024–2025 were used to examine ozone (O₃), temperature (T), and relative humidity (RH) relationships using Pearson and Spearman correlations, partial correlations, spatial analysis, and distributional assessment. NMF was used exclusively for exploratory pattern recognition and not quantitative source apportionment. **Results:** The first three PCA components explained 58.8% of total variance, accounting for 27.2%, 21.9%, and 9.8%, respectively. NMF and k-means identified three recurrent pollutant patterns interpreted as traffic/combustion, coarse-particle/resuspension, and photochemical/secondary regimes. In the ERA5/GIOVANNI analysis, O₃ showed a strong negative correlation with temperature ($r = -0.825$, $p < 0.001$), a positive correlation with RH ($r = 0.344$, $p < 0.001$), and temperature and RH were negatively correlated ($r = -0.610$, $p < 0.001$). These results indicate substantial coupling between ozone variability and meteorological conditions. However, the observed statistical relationships cannot independently demonstrate specific chemical mechanisms or establish direct correspondence between historical surface regimes and contemporary atmospheric-column observations. **Conclusion:** The proposed "Invisible Bridge" provides an exploratory framework for interpreting surface–column differences through atmospheric transformation and transport. Its current evidence is process-based rather than a direct validation of PMF–remote-sensing integration, providing a basis for future temporally matched and quantitatively integrated observations.

Keywords: air pollution; atmospheric chemistry; particulate matter; ozone; remote sensing; atmospheric reanalysis; secondary aerosol formation; pollution regimes; non-negative matrix factorization.

INTRODUCTION

Atmospheric chemistry governs the composition, transformation, and redistribution of pollutants from the surface layer to the upper troposphere. The atmosphere is continuously influenced by both anthropogenic and natural processes, with pollutants being either directly emitted or formed through chemical reactions involving primary pollutants and other atmospheric constituents (Finlayson-Pitts & Pitts, 2000; Gaffney & Marley, 2003; Matthias, 2004). Primary pollutants are emitted directly from their sources, whereas secondary pollutants are produced through atmospheric transformation reactions and can substantially influence air quality and human health (Saxena & Sonwani, 2019; Xiang et al., 2023). Particulate matter, CO, SO₂, NO₂, and O₃ are among the principal atmospheric pollutants considered in air-quality assessment (Seinfeld & Pandis, 2006). Recent work has further emphasized that atmospheric chemical transformations may regenerate reactive species and maintain interconnected reaction pathways, indicating that observed pollutant variability cannot always be explained solely by emission intensity or meteorological conditions (Mandija, 2013; Mandija, 2015; Price et al., 2025).

Atmospheric transformations operate on different temporal and chemical scales. During daytime, solar radiation and hydroxyl radicals (OH) initiate oxidation processes involving primary pollutants and contribute to the formation of O₃ and other secondary compounds (Monks, 2005; King, 2013; Ball, 2014). During night-time, the formation of nitrate radicals (NO₃) and

the NO₃–N₂O₅ equilibrium provide alternative pathways for atmospheric oxidation and pollutant processing (Brown et al., 2006; Brown et al., 2016). Water vapour also participates in oxidative chemistry through OH-radical formation and therefore can influence the degradation and variability of O₃ (Seinfeld & Pandis, 2006; Mandija et al., 2018; Mandija et al., 2025). These coupled chemical and meteorological processes can modify the composition and distribution of pollutants after their initial emission and can contribute to secondary particulate formation.

Source-apportionment approaches provide an important means of characterizing pollutant mixtures at the surface. Positive matrix factorization (PMF) is a widely used receptor-model framework for identifying and quantifying source-related chemical patterns in particulate matter (Paatero & Tapper, 1994; Paatero, 2004; Song et al., 2007). However, PMF is inherently based on measurements collected at the receptor and therefore provides limited information on the vertical and horizontal redistribution of pollutants originating outside the local measurement domain. In particular, surface observations may contain the combined effects of direct emissions, atmospheric transformation, secondary formation, and transport, while the transformation history and spatial origin of the contributing pollutants cannot be fully resolved from the surface record alone.

Remote sensing provides complementary information because satellite observations characterize atmospheric pollutants over

a much broader spatial domain and, for many products, integrate information through the atmospheric column. Aerosol products therefore do not directly measure near-surface $\text{PM}_{2.5}$ in the same manner as ground-based observations; instead, quantities such as aerosol optical depth represent particle-related light extinction integrated through the atmospheric column (Diao et al., 2019). Consequently, a direct equivalence between surface $\text{PM}_{2.5}$ measurements and satellite-derived column observations cannot be assumed. The relationship between the two depends on atmospheric transport, vertical distribution, chemical transformation, and secondary formation.

This creates an important interpretive gap between surface-based source characterization and column-integrated remote sensing. PMF can characterize chemical source patterns at the receptor, whereas remote sensing can describe the spatial distribution of atmospheric pollutants but does not by itself provide source apportionment. The two observational perspectives therefore require a process-based mechanism capable of explaining how pollutants emitted or observed near the surface are chemically transformed and redistributed before contributing to the atmospheric column. The present study considers atmospheric chemical transformation and regeneration as this connecting mechanism. In this framework, primary pollutants can undergo oxidation and subsequent transformations, with part of the transformed material contributing to secondary $\text{PM}_{2.5}$ formation. At the same time, atmospheric transport can redistribute both primary and chemically processed pollutants vertically and horizontally. Chemical transformations may therefore influence the relationship between surface pollutant composition and column-integrated atmospheric observations without implying that the original source-related co-variation is necessarily lost.

The central scientific question addressed in this study is whether a combination of atmospheric chemical relationships and recurrent pollutant regimes can provide an empirical basis for interpreting the connection between surface observations and remote-sensing observations. In particular, the study examines whether relationships among O_3 , temperature, and relative humidity can characterize atmospheric conditions relevant to chemical transformation and whether recurrent pollutant co-variation patterns identified in surface observations can provide a regime-based context for understanding surface–column differences. The study therefore treats the relationship between the two datasets as process-based rather than as direct temporal validation: the historical surface-monitoring observations characterize recurrent pollutant regimes, whereas the more recent ERA5/GIOVANNI data characterize contemporary atmospheric variability and O_3 –T–RH relationships.

Based on the atmospheric-chemistry framework described above, we hypothesize that (i) atmospheric pollutant variability is organized into distinct recurrent regimes reflecting different combinations of fresh emissions, particulate enhancement/resuspension, and chemically processed or secondary pollution; (ii) O_3 variability is associated with temperature and relative humidity in a manner consistent with the combined influence of atmospheric dynamics and chemical processing; and (iii) the chemical transformation and vertical redistribution of pollutants provide the process-level link that helps explain why surface pollutant regimes and column-integrated remote-sensing signals can diverge or converge under different atmospheric conditions. These hypotheses are evaluated using complementary exploratory multivariate analyses of surface observations and statistical analysis of ERA5/GIOVANNI atmospheric variables.

Accordingly, the aim of this study is not to perform a new quantitative PMF source-apportionment analysis, but to develop and empirically examine a process-based framework, the proposed "Invisible Bridge", for interpreting how atmospheric chemical transformation, secondary formation, and transport connect surface pollutant regimes with column-integrated remote-sensing observations. To achieve this aim, we (1) characterize recurrent pollutant co-variation patterns in the Tirana surface-monitoring dataset using PCA, exploratory NMF, and k-means clustering; (2) identify and characterize the principal atmospheric regimes represented by these patterns; and (3) examine O_3 , temperature, and relative humidity relationships in the ERA5/GIOVANNI dataset to assess their relevance to atmospheric chemical processing and the surface–column linkage. The resulting framework is intended to establish when and why surface pollutant patterns may be reflected differently in column-integrated observations, while providing a basis for subsequent quantitative integration of PMF/ME-2 and remote-sensing data.

MATERIALS AND METHODS

Methodological positioning of PMF and NMF

Positive matrix factorization (PMF) is considered in this study as the established receptor-modelling framework for quantitative source apportionment of particulate matter. However, no PMF or ME-2 source-apportionment model was fitted to the present dataset. Instead, non-negative matrix factorization (NMF) was employed as an exploratory multivariate method to identify recurrent non-negative patterns of pollutant co-variation.

The distinction between NMF and PMF is maintained throughout the analysis. NMF does not incorporate the measurement-error structure, uncertainty weighting, or source-profile constraints that are fundamental to quantitative PMF applications. Consequently, NMF factors are interpreted exclusively as statistical pollutant-regime patterns and are not reported as emission-source contributions, source profiles, or quantitative source apportionment.

NMF was used in conjunction with principal component analysis (PCA) and k-means clustering to identify recurrent structures in the surface-monitoring observations. Agreement among these independent statistical representations was used to characterize robust pollutant regimes. PMF/ME-2 is retained as a subsequent quantitative step that could test whether the exploratory regimes identified here correspond to chemically resolved source profiles when appropriate speciated aerosol measurements and uncertainty estimates are available.

Accordingly, the present analysis should be regarded as an exploratory regime-characterization study rather than a PMF source-apportionment study. The term "source" is therefore avoided when referring to NMF factors unless supported by independent chemical or contextual evidence.

Chemical background: relevant mechanisms

We have previously discussed the concept of pollutant regeneration and its potential importance for linking surface and satellite observations. In this section, this process is further illustrated through representative atmospheric chemical reactions. We selected volatile organic compounds (VOCs), specifically benzene (C_6H_6) and ethylene (C_2H_4), as representative pollutants measured or relevant to the atmospheric environment of Tirana, Albania. Benzene and ethylene are considered here as representative urban VOCs because of their association with traffic and industrial emissions. These compounds participate in complex photochemical and oxidative reactions and undergo

atmospheric transformation (Figure 1). Through these reaction pathways, they contribute to the formation and redistribution of secondary atmospheric pollutants, including ozone (O_3),

thereby illustrating chemical mechanisms that may contribute to the observed variability in atmospheric pollutant concentrations.

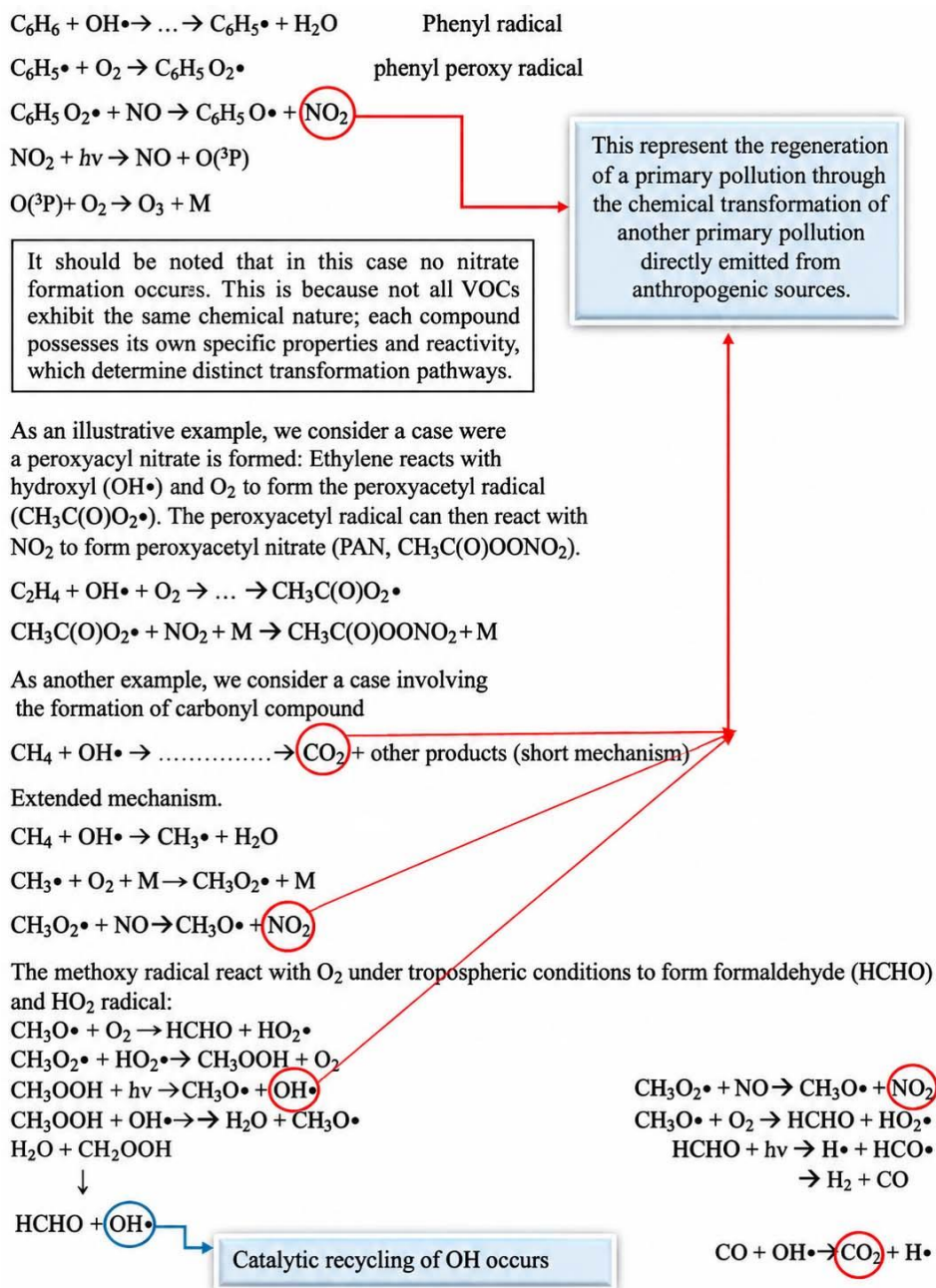


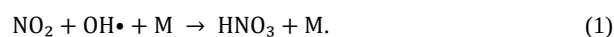
Figure 1. Conceptual scheme illustrating the chemical transformation and regeneration of reactive atmospheric species linking primary emissions to secondary pollutant formation

All reactions shown in Figure 1 are supported by existing literature, with references covering VOC oxidation pathways and both daytime and night-time chemistry of primary pollutants (Hallquist et al., 2009; Kleinman et al., 2002; Seinfeld & Pandis, 2006; Ball, 2014). Although the term "recycling" is commonly used in the atmospheric-chemistry literature for simplicity of expression (e.g., Lin et al., 2025), in this study the term is used with stricter conceptual precision. Specifically, we use "recycling" for cases in which reactive species participate temporarily in atmospheric reactions and are subsequently regenerated without implying reconstruction of the original primary pollutant. Accordingly, we use the term "catalytic recycling" to describe processes in which reactive radicals such as $OH \cdot$ participate in oxidation reactions and are

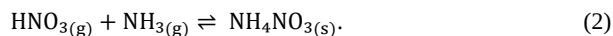
subsequently regenerated within the reaction cycle, allowing them to participate in further reactions without permanent removal from the system (Monks, 2005; Stone et al., 2012).

The benzene transformation pathways discussed here are consistent with mechanisms reported in previous studies (Albert & Davis, 2010; Yuan et al., 2016), while the formation of peroxyacetyl nitrate is described by Fischer et al. (2014).

Under daytime tropospheric conditions, NO_2 reacts with OH radicals to form nitric acid, representing an important homogeneous pathway for nitric acid formation (Finlayson-Pitts et al., 2002; Zhang et al., 2015):



Nitric acid subsequently participates in gas-particle partitioning and heterogeneous processes involving ammonia, amines, mineral dust, and sea-salt aerosol. Under suitable atmospheric conditions, ammonium nitrate (NH_4NO_3) can represent an important nitrate component of fine particulate matter (PM) (Zhang et al., 2015):



These pathways illustrate how nitrogen oxides participate in coupled atmospheric transformation cycles. Following oxidation, NO_2 -derived products may contribute to secondary aerosol formation or participate in subsequent atmospheric reaction pathways, depending on prevailing chemical and meteorological conditions. These processes can therefore influence both gaseous atmospheric chemistry and $\text{PM}_{2.5}$ composition.

Data and analysis

This study combines data analysis based on reanalysis datasets with a conceptual interpretation of atmospheric chemical processes. The goal is to establish a theoretical link between PMF and RS observations through the role of chemical transformations.

Ozone (O_3), air temperature (T), and relative humidity (RH) were obtained from the ERA5 reanalysis dataset (Hersbach et al., 2023). The data were accessed through the NASA GIOVANNI online portal (<https://giovanni.gsfc.nasa.gov/giovanni>). The study region covers Albania, with the city of Tirana selected as representative of a typical urban-industrial zone. This area is characterized by complex emission sources and significant weather variability, making it suitable for the present study.

The study uses two distinct observational components with different temporal coverage and complementary analytical roles that support the overall analysis. First, O_3 , T, and RH were obtained from the ERA5 reanalysis dataset through the NASA GIOVANNI platform. The ERA5/GIOVANNI analysis covers January 2024 to December 2025. These data were used to characterize contemporary spatial and temporal variability in atmospheric O_3 , T, and RH and to examine their statistical relationships in the atmospheric column and near-surface layers. For spatial consistency, the gridded data were aggregated as area-weighted means over Albania, with particular attention to the Tirana region.

Second, an independent historical surface-monitoring dataset from an automatic urban-background station in central Tirana was used to characterize pollutant co-variation and recurring atmospheric regimes. The surface-monitoring record covers 2010–2012 and contains hourly observations of particulate and gaseous pollutants. These data were analysed using PCA, NMF, and k-means clustering to identify representative traffic/combustion, coarse-particle/resuspension, and photochemical/secondary regimes.

The two datasets were therefore not treated as temporally coincident observations and were not used for direct year-to-year or point-to-point temporal validation or quantitative temporal comparison. Instead, they serve complementary purposes within the proposed conceptual framework: the historical surface observations provide an empirical characterization of pollutant regimes at ground level, whereas the more recent ERA5/GIOVANNI data provide an independent characterization of atmospheric variability and chemical–meteorological relationships relevant to the surface–column linkage. The use of different periods is an explicit feature of the case-study design rather than a data-merging procedure or an attempt at temporal alignment within the analysis.

Because the study is designed as a conceptual and semi-quantitative case study, the ERA5/GIOVANNI results should not be interpreted as climatological trends or as evidence of long-term changes in ozone behaviour. Similarly, the 2010–2012 surface-monitoring analysis is used to demonstrate the existence and structure of representative pollutant regimes rather than to establish their persistence over the 2024–2025 period. The integration of the two components is therefore process-based rather than temporally matched, with atmospheric chemical transformation, vertical redistribution, and transport providing the conceptual mechanism connecting surface pollutant regimes with remote-sensing and reanalysis observations.

The exploratory analysis of the historical surface-monitoring dataset used three complementary multivariate approaches. Principal component analysis (PCA) was used to summarize dominant orthogonal modes of variability and to generate scores for visualization and clustering. Non-negative matrix factorization (NMF) was used to identify non-negative co-varying pollutant patterns, while k-means clustering was applied to the first three PCA scores to identify representative atmospheric regimes. The detailed pre-processing, quality-control procedures, model configurations, stability assessment, and criteria used for selecting the NMF and k-means solutions are described in Section Statistical analysis of surface-monitoring data. The resulting NMF factors were treated as exploratory atmospheric regimes rather than quantitative source profiles.

The results downloaded from GIOVANNI (in NetCDF and CSV formats) were processed for the defined study area and time period. The pre-processing steps included:

- selection of pressure levels corresponding to the surface layer (temperature at 2 meters above the ground);
- temporal reindexing of all variables to hourly means;
- screening for missing (NaN) values and identification of abrupt variations (spikes);
- removal of anomalies and outliers to improve time-series consistency.

These steps ensured temporal synchronization among O_3 , T, and RH and removed potential numerical anomalies prior to correlation analyses.

Pairwise statistical relationships among the variables (O_3 –T and O_3 –RH) were assessed using Pearson correlation coefficients with two-tailed p-values. Spearman rank correlations were also calculated as a robustness check to account for nonparametric dependencies. Confidence intervals were calculated for the correlation estimates, and the Benjamini–Hochberg correction was applied to control for false positives when testing multiple correlations at different altitudes. Partial correlations were used to quantify the O_3 –T association while controlling for RH, and vice versa. Despite the linear assumptions underlying the correlation analyses, regression analyses were also applied to identify nonlinear dependencies in the data.

Data visualizations complemented the quantitative analyses. Scatterplots, distributional density plots, and the t-multimodality test were generated to examine the distribution of O_3 and its covariates. All graphical and numerical analyses were interpreted with emphasis on physical and chemical mechanisms rather than isolated statistical significance (p-values), ensuring the environmental relevance of the results.

LVS and PMF framework

In this framework, low-volume sampler (LVS) is referenced as the standard instrument commonly used for the collection of particulate matter samples in ambient air. The LVS operates at a typical flow rate of approximately $2.3 \text{ m}^3 \cdot \text{h}^{-1}$ and allows for the collection of atmospheric particles on filters suitable for

subsequent chemical analyses (Çaushaj et al., 2024). Although no direct sampling was performed in this work, the LVS is included in the methodological framework (see Figure 12) to illustrate the data acquisition process that precedes source apportionment analysis using the PMF model. The instrument's role in the workflow is to represent the initial phase of empirical data collection, ensuring that the PMF approach is conceptually linked to measurable field observations.

Air-quality monitoring data

Air-quality data for this study were obtained from the automatic urban-background monitoring station in central Tirana, operated by the National Environmental Agency (AKM) in collaboration with the Institute of Public Health. Continuous hourly measurements covering 2010–2012 were used. The station monitored CO, SO₂, NO_x (NO, NO₂), O₃, PM₁₀, benzene, and speciated VOCs (including ethylbenzene and o-xylene). Reported units were CO in mg·m⁻³; SO₂, NO_x, O₃, benzene, and PM₁₀ in µg·m⁻³; and VOCs (ethylbenzene and o-xylene) as mixing ratios (ppb).

The instruments conformed to European reference methods: PM₁₀ by EN 12341 (gravimetric or beta-attenuation sampler); SO₂ by UV-fluorescence (Teledyne/API Model 100E); NO_x by chemiluminescence (Model 200E); O₃ by UV-absorption photometer (EN 14625); CO by non-dispersive infrared (NDIR); and benzene/VOCs by gas chromatography (GC-based systems). Routine QA/QC and calibrations (zero/span checks and multi-point calibrations) followed manufacturer and EN protocols. The data were quality-checked and aggregated to hourly and daily averages for analysis.

Statistical analysis of surface-monitoring data

The historical surface-monitoring dataset from Tirana (2010–2012) was analysed using descriptive statistics and complementary multivariate techniques to characterize pollutant co-variation and recurrent atmospheric pollution regimes. The dataset comprised hourly observations of PM₁, PM_{2.5}, PM₁₀, NO_x, NO, NO₂, CO, O₃, SO₂, ethylbenzene, and o-xylene.

Data pre-processing and quality control

The surface-monitoring data were subjected to sequential quality control prior to statistical analysis. Missing, invalid, negative, and physically implausible observations were identified according to predefined quality-control criteria. Short data gaps were linearly interpolated to preserve local temporal continuity, whereas residual missing observations were replaced by the median value of the corresponding variable. Median imputation was applied only after interpolation and was used to minimize the introduction of artificial temporal patterns and extreme values.

The initial dataset comprised 16,632 contiguous hourly records spanning 20 May 2010 to 12 April 2012, corresponding to 182,952 individual pollutant measurements across the 11 species. Of these, 39,721 measurements (21.7%) were identified as missing or invalid, comprising 36,119 unreported values, 3,557 negative values, and 45 instrument fill values. Invalid measurements were set to missing and treated by linear interpolation of short gaps followed by median imputation of residual gaps. No hourly records were excluded, and all 16,632 records were retained for the primary analysis.

Variables were transformed and standardized where required before multivariate analysis to prevent differences in measurement scale from disproportionately influencing the results. PCA, NMF, and k-means clustering were performed using the same quality-controlled analytical dataset.

A sensitivity analysis was conducted by repeating the PCA, NMF, and k-means procedures using an alternative dataset in which residual missing observations were excluded rather than median-imputed. The resulting factor structures, explained variance, and cluster characteristics were compared with those obtained from the primary analysis.

Principal component analysis

Principal component analysis (PCA) was applied to the standardized pollutant matrix to identify the dominant orthogonal modes of variability and reduce the dimensionality of the dataset. Variables were standardized using z-score transformation to zero mean and unit standard deviation, and the PCA was based on the correlation matrix.

The number of components was evaluated using explained variance and the scree-plot criterion. The first three principal components were retained for interpretation and for subsequent k-means clustering. Component loadings were interpreted according to their magnitude, sign, and pollutant co-variation within the atmospheric context. PCA was used for dimensionality reduction and pattern recognition and was not considered a direct source-apportionment method.

Software: RStudio 2023.06.1+524 "Mountain Hydrangea" Release (547dcf861cac0253a8abb52c135e44e02ba407a1, 2023-07-07) for Windows; Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) RStudio/2023.06.1+524 Chrome/110.0.5481.208 Electron/23.3.0 Safari/537.36; NMF.

Non-negative matrix factorization

Non-negative matrix factorization (NMF) was used as an exploratory multivariate technique to identify non-negative pollutant co-variation patterns. NMF models with ranks from 2 to 6 were evaluated using multiplicative-update optimization and multiple random initializations.

For each factor rank, 50 independent runs were performed using random seeds 0–49. Convergence was defined as a relative change in the reconstruction objective of $< 1 \times 10^{-6}$ between successive iterations, with a maximum of 20,000 iterations per run. Reconstruction performance was quantified using the Frobenius norm, with normalized reconstruction error calculated as:

$$\|X - WH\|_F / \|X\|_F \quad (3)$$

Factor stability across independent runs was assessed after resolving permutation ambiguity using the Hungarian algorithm (linear sum assignment) applied to the pairwise cosine-similarity matrix between factor profiles from each pair of runs. Similarity between corresponding factor profiles was quantified using cosine similarity between L1-normalized factor profiles.

The final factor rank was selected based on the joint assessment of reconstruction error, factor stability, and interpretability. In the combined analysis, a three-factor solution was selected. NMF factors were interpreted as exploratory pollutant-regime patterns rather than quantitative emission-source profiles and were not used to derive source contributions or source-specific mass fractions.

NMF stability analysis

The stability of the selected three-factor solution was evaluated using 50 independent random initializations to assess the consistency of the resulting factor structure. Across the multi-start solutions, the mean pairwise similarity of matched factor profiles was 0.998 ± 0.005 , with a minimum similarity of 0.962 and a maximum of 1.000. The corresponding reconstruction

error was 0.0862 ± 0.0003 , and the coefficient of variation of the reconstruction error was 0.40%. Factor ordering and the dominant pollutant associations remained consistent across runs. These results indicate that the three-factor solution was stable with respect to random initialization and was not driven by a particular starting configuration.

K-means clustering and identification of atmospheric regimes

K-means clustering was applied to the scores of the first three principal components to classify observations according to their multivariate atmospheric characteristics. Candidate solutions with $k = 2-6$ clusters were evaluated using the elbow criterion based on within-cluster sum of squares (WCSS), together with silhouette scores and the physical interpretability of the resulting pollutant regimes.

The elbow analysis indicated a marked reduction in WCSS up to $k = 3$, followed by progressively smaller improvements for higher k values. The $k = 3$ solution provided an appropriate balance between within-cluster compactness, cluster separation, and physical interpretability and was therefore selected for the subsequent atmospheric-regime analysis. Cluster-specific pollutant concentrations were subsequently calculated to characterize the atmospheric regimes represented by the identified clusters.

The resulting clusters were interpreted as recurrent multivariate atmospheric states rather than independent emission-source categories.

Statistical integration of multivariate analyses

The robustness of the multivariate structure to the treatment of residual missing values was assessed by repeating PCA, NMF,

and k-means clustering after excluding residual missing observations instead of applying median imputation. The resulting component structure, factor profiles, and cluster characteristics were compared with those obtained from the primary analysis.

RESULTS

This section is divided into two major sections: the first one dealing with the remote sensing techniques on the data of ozone, temperature and relative humidity with respect to liquid, and the second one providing a case study of in-situ measurements in the same location.

Remote sensing analyses

The results presented in this section aim to demonstrate that RS technique provide a broader view of atmospheric pollutants, particularly gases, which PMF does not directly capture. Through the following analyses and interpretations of these correlations, we are also focused on the meteorological conditions, and how satellite-derived information can be linked to PMF results, enabling a more comprehensive interpretation of pollutant variability (Figure 2). The regeneration pathway shown in C_6H_6 and C_2H_4 scheme is important, as it illustrates how gases, after a series of chemical transformations can be regenerated and ultimately become part of $PM_{2.5}$.

For example: $NO_2 + OH \rightarrow HNO_3$, $HNO_3 + NH_3 \rightarrow NH_4NO_3$ which is one of the important components of $PM_{2.5}$.

A strong and statistically significant negative correlation was observed between ozone and temperature ($r = -0.825$, $p < 0.001$; Figure 2–3). This inverse relationship indicates that ozone concentrations decrease as temperature increases (Figure 3).

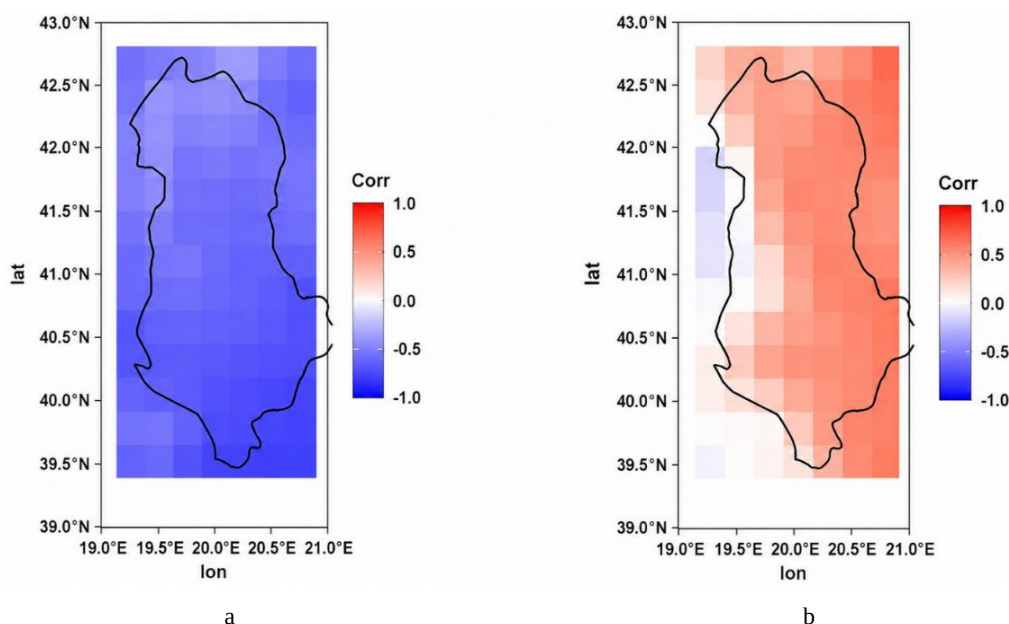


Figure 2. Spatial correlation maps of the study domain (Albania):
a – correlation between O_3 and RH; b – correlation between O_3 and temperature T

While surface-level ozone typically increases under high-temperature conditions due to enhanced photochemical production, the pressure-level data analysed here reveal a different mechanism. At mid- to upper-tropospheric levels, higher temperatures are often associated with warm, dry descending air masses that reduce ozone mixing ratios. Conversely, lower temperatures are associated with stratospheric intrusions or enhanced vertical mixing that transport ozone-rich air downward, resulting in higher ozone concentrations. Thus, this strong anticorrelation reflects the

influence of large-scale atmospheric dynamics and vertical transport processes (Kalabokas et al., 2020). The correlation between ozone and relative humidity was positive but weaker ($r = 0.344$, $p < 0.001$). This relationship suggests that humid air masses may be associated with higher ozone levels at these pressure levels, although the effect is less pronounced than that of temperature. A possible explanation is that convection and vertical uplift, often associated with moist air masses, facilitate the upward transport of ozone precursors from the boundary layer.

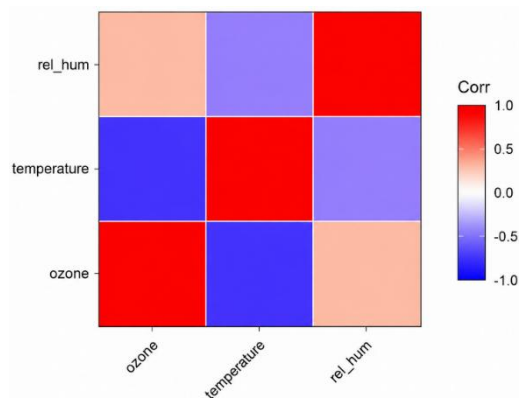


Figure 3. Correlation matrix among ozone, temperature, and relative humidity represented as a heatmap. Positive correlations are shown in red, while negative correlations are shown in blue, illustrating the strong negative relationship between O_3 and T and the inverse relationship between O_3 and RH

In addition, humid and cloudy conditions may reduce ozone photolysis rates, thereby increasing ozone lifetime. The relatively modest correlation indicates that humidity exerts a secondary influence among several additional processes, including convective mixing, dilution, and long-range transport.

Also, a moderate negative correlation was observed between temperature and relative humidity ($r = -0.610$, $p < 0.001$). This result is consistent with thermodynamic expectations, whereby warmer air masses over the region, especially during summer, tend to be drier under anticyclonic conditions, whereas colder air masses are generally moister and associated with frontal systems or cyclonic activity.

Up to this point, the phenomenon has been considered from a thermodynamic perspective; however, the chemical dimension, namely the influence of atmospheric chemistry, should also be considered. Multiple competing processes occur simultaneously in the atmosphere, meaning that the influence of chemical transformations cannot be excluded. It should be noted that the correlations were analysed pairwise: ozone–temperature (negative), ozone–humidity (positive), and temperature–humidity (moderate negative).

The aim, as developed throughout this research, is to investigate the chemical basis for the observed decrease in ozone concentration with increasing temperature and decreasing relative humidity, thereby linking these three variables (temperature, relative humidity, and ozone) within a unified correlation analysis. This analysis was conducted following an extensive review of the literature on atmospheric chemistry, indicating that a negative correlation may emerge when the two combined variables (temperature and RH) increase simultaneously, accompanied by a decrease in ozone. However, this relationship does not occur universally, as numerous chemical processes and pollutants coexist in the atmosphere. In the present analysis, idealized conditions involving only these three variables are assumed in order to better isolate their combined effect.

From a theoretical perspective and based on known chemical transformations, increasing temperature can contribute to ozone destruction through changes in water vapour concentration in the troposphere. An increase in relative humidity may enhance the production of hydroxyl (OH) radicals, which initiate chains of oxidative reactions contributing to ozone depletion.

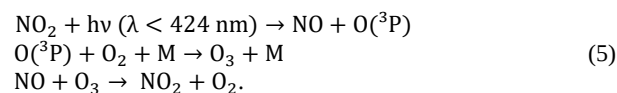
In many polluted regions, high surface ozone concentrations are strongly associated with temperature. This is due to the fact that during heatwave events, elevated temperatures coincide with atmospheric stagnation, enhanced photochemical activity, and

biogenic or biomass-burning emissions (Jacobson & Streets, 2009; Chen et al., 2009; Leibensperger et al., 2008; Murazaki & Hess, 2006; Nolte et al., 2008; Jacobson, 2008; Doherty et al., 2013; Wu et al., 2021). On the other hand, under warm climatic conditions, ozone concentrations may decrease, mainly due to elevated water vapour concentrations. This increase in humidity may favour ozone destruction, particularly in regions characterized by low NO_x concentrations (Fiore et al., 2012).

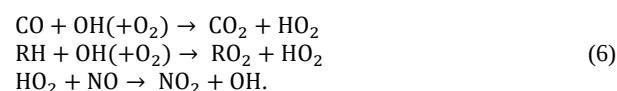
To explain these observed correlations, we propose the following chemical mechanisms, expressed by Equations 4–6. As temperature increases, the saturation vapour pressure increases, allowing more water vapour to be present before saturation is reached (Trenberth et al., 2005). At this stage, excited oxygen atoms O^* react with water vapour to generate hydroxyl radicals $OH^\bullet$, which play an important role in ozone removal. The key chemical reactions are expressed in Equation 1 (Sanderson, 2007).



When other chemical processes are considered, additional interactions modify this mechanism. For example, nitrogen oxides (NO_x) are known to enhance ozone formation in the troposphere at high concentrations, while low NO_x conditions lead to ozone depletion (Brown et al., 2006). The photolysis of NO_2 is the primary reaction driving tropospheric ozone formation, followed by regeneration cycle of NO_2 through its interaction with ozone, Equation 2:



Net ozone production occurs in the presence of ozone precursors, including carbon monoxide (CO), methane (CH_4), and volatile organic compounds (VOCs), which are oxidized by hydroxyl radicals (OH) to produce hydroperoxyl radicals (HO_2). These radicals act as intermediates, providing alternative pathways for the conversion of NO to NO_2 without consuming ozone (Jacob, 2000; Ivatt et al., 2022; Dewan & Lakhani, 2022).



Chain termination can occur either through self-reaction of radicals, forming more stable species (e.g., peroxides), or

through the reaction between hydroxyl radicals and NO_2 to form nitric acid (HNO_3). Because the formation of ozone is closely linked to the levels of NO_x , lower concentrations of NO_x lead to reduced ozone generation in the troposphere. As a result, the recovery of ozone that is depleted by other reactions is diminished. In warm and humid environments, any remaining ozone contributes to the creation of OH radicals, thereby enhancing the ozone destruction process rather than the production process (Dewan & Lakhani, 2022; Sanderson, 2007).

Atmospheric processes, such as lightning-generated NO_x emissions and stratosphere-troposphere exchange, may partially offset this reduction (Lamarque et al., 2011; Kawase et al., 2011). However, these factors have a localized or limited global impact and cannot fully compensate for the overall ozone loss caused by increased humidity (Kawase et al., 2011; Zeng et al., 2010). Lightning discharges produce NO_x by breaking down N_2 and O_2 at high temperatures, which then recombine into reactive nitrogen species. These NO_x compounds are vital for controlling the chemistry and concentration of tropospheric ozone (Fiore et al., 2012; Pyrgou et al., 2018).

In summary, high temperature (T) leads to increased relative humidity (RH). When it occurs alongside low NO_x concentrations, it results in a net decrease in tropospheric ozone. The Clausius-Clapeyron relationship shows that higher temperatures increase water vapor in the troposphere (Wallace & Hobbs, 2006). Atmospheric chemistry then explains the mechanisms behind ozone loss. This analysis highlights a paradox between atmospheric dynamics and atmospheric chemistry.

Although warm air associated with high temperatures has a greater capacity to hold water vapor (Wallace & Hobbs, 2006), relative humidity (RH) can decrease because the air's ability to hold moisture increases faster than the available supply of vapor. Therefore, RH depends not only on temperature, but also on the rate of surface evaporation that replenishes water vapor in the air (Byrne & O'Gorman, 2016).

The ozone density plots indicate a bimodal distribution, implying the existence of two separate atmospheric regimes (Figure 4). The regime associated with lower ozone levels likely pertains to summer conditions, which are influenced by significant photochemical degradation and the subsidence of dry air masses. Similar occurrences of summer ozone reduction due to photochemical processes and dry-air subsidence have been documented by Zanis et al. (2014). Conversely, the regime

exhibiting higher ozone levels may be linked to spring and autumn, times when stratospheric intrusions and the long-distance transport of ozone-rich air masses occur more frequently, aligning with earlier findings throughout the Mediterranean (Cristofanelli & Bonasoni, 2009; Viana et al., 2018).

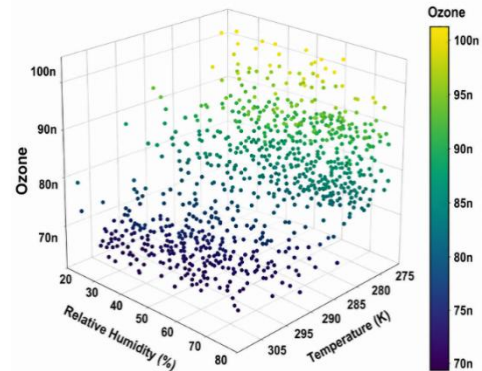


Figure 4. Three-dimensional scatter plot illustrates the relationship between ozone concentration (O_3), temperature and relative humidity

Overall, these results indicate that temperature is the dominant driver of ozone variability at high altitudes, whereas relative humidity exerts a secondary impact. The strong dynamical control observed here underscores the importance of accounting for vertical transport, seasonal circulation and stratosphere–troposphere exchange processes when interpreting ozone variability and its chemical behaviour. In the context of climate change the projected increase in heatwave frequency and atmospheric drying could further modulate ozone concentrations throughout the middle and upper troposphere with substantial implications for regional air quality and radiative forcing (Gaudel et al., 2018; Doherty et al., 2013).

Figure 5 shows correlations between ozone, air temperature, and relative humidity. Here, their distributions and regression analyses are presented. Above the diagonal, correlation coefficients are reported: a strong negative correlation between O_3 and temperature ($r = -0.825$, $p < 0.001$), a weak positive correlation between O_3 and RH ($r = 0.344$, $p < 0.001$), and a moderate negative correlation between temperature and RH ($r = -0.610$, $p < 0.001$).

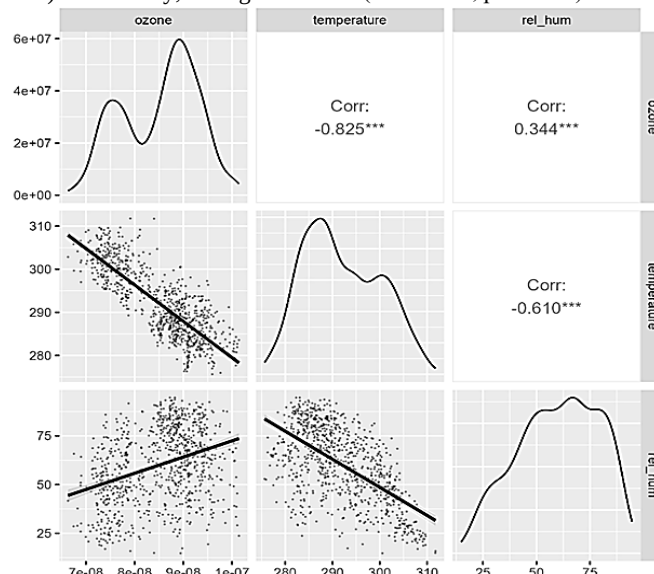


Figure 5. Pairwise correlation plot between ozone, temperature and relative humidity

In the scatterplots shown below the diagonal, the relationships between all pairs of variables are presented, along with a black fitted regression line to demonstrate the relationships indicated by the calculated Pearson correlations. On the diagonal, the density distribution of each variable is shown, with clear bimodal characteristics for ozone (indicating two seasonal regimes), a unimodal distribution of temperature, and a relatively wide range of relative humidity (50–70%). Both variables show consistent correlations with ozone and support the conclusion that temperature is the primary factor affecting ozone fluctuations, while relative humidity also has an impact but is less pronounced.

Overall, these results support the thesis that atmospheric chemical processes contribute to variations in air pollution, as discussed in relation to the connections established through processes occurring in atmospheric chemistry.

Night-time temperature displays a unimodal distribution centred around 284.7 K, indicating thermally stable conditions and limited vertical mixing across the lower troposphere (Figure 6a). Relative humidity values reach 70–90% across areas, consistent with the observed night-time mean of 78.53% enhanced moisture accumulation after sunset (Figure 6b).

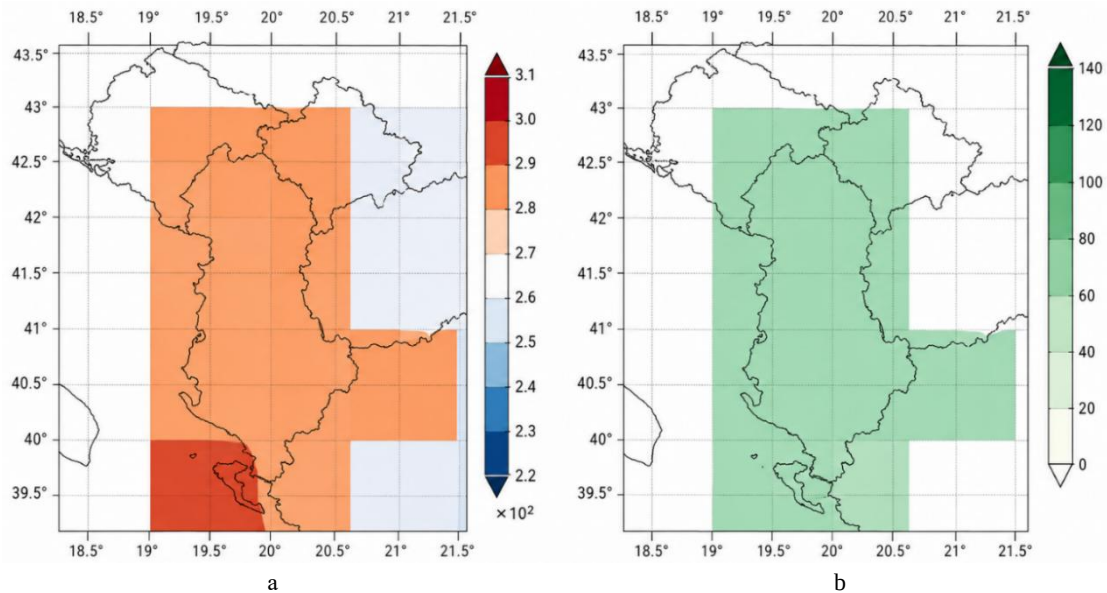
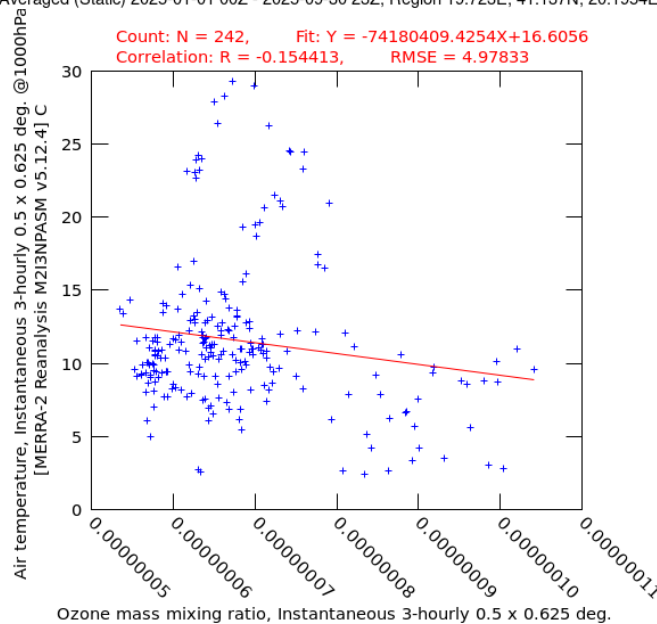


Figure 6. Night-time surface meteorological conditions over Albania: a – surface air temperature distribution (Giovanni, 2025); b – time-averaged surface relative humidity distribution

The correlation analysis shows a weak negative relationship between ozone and surface-air temperature ($R = -0.15$; Figure 7a), whereas the relationship between ozone and relative humidity is stronger and also negative ($R = -0.48$; Figure 7b). These results indicate that ozone variability is more strongly

associated with relative humidity than with temperature under the conditions represented in the dataset. The correlations are interpreted as statistical associations and do not, by themselves, establish a causal role of water vapor or hydroxyl-radical chemistry.

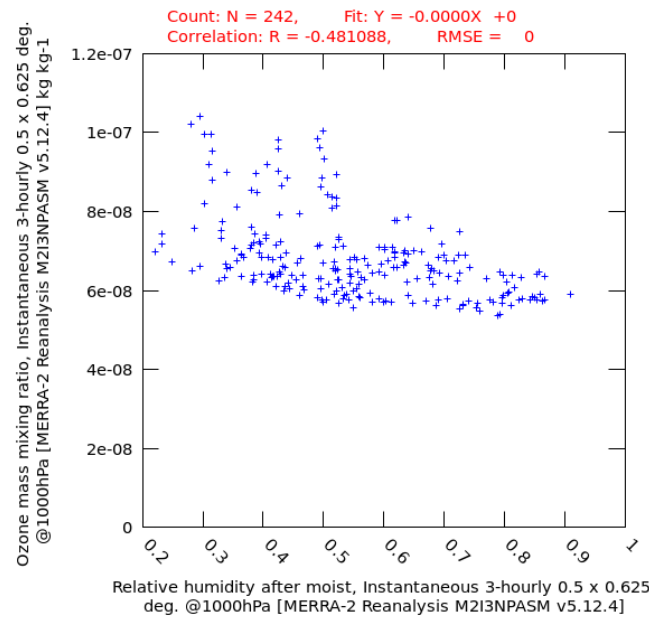
Scatter, Area Averaged (Static) 2025-01-01 00Z - 2025-09-30 23Z, Region 19.723E, 41.137N, 20.1954E, 41.5984N



- The user-selected region was defined by 19.723E, 41.137N, 20.1954E, 41.5984N. The data grid also limits the analyzable region to the this point: 20E, 41.5N. This analyzable region indicates the spatial limits of the subsetted granules that went into making this visualization result.

a

Scatter, Area Averaged (Static) 2025-01-01 00Z - 2025-09-30 23Z, Region 19.723E, 41.137N, 20.1954E, 41.5984N



- The user-selected region was defined by 19.723E, 41.137N, 20.1954E, 41.5984N. The data grid also limits the analyzable region to the this point: 20E, 41.5N. This analyzable region indicates the spatial limits of the subsetting granules that went into making this visualization result.

b

Figure 7. Ozone relationships with surface meteorological parameters over the Tirana region: a – scatter plot of ozone versus surface air temperature (T); b – scatter plot of ozone mass mixing ratio versus relative humidity at 1000 hPa (MERRA-2, 2025)

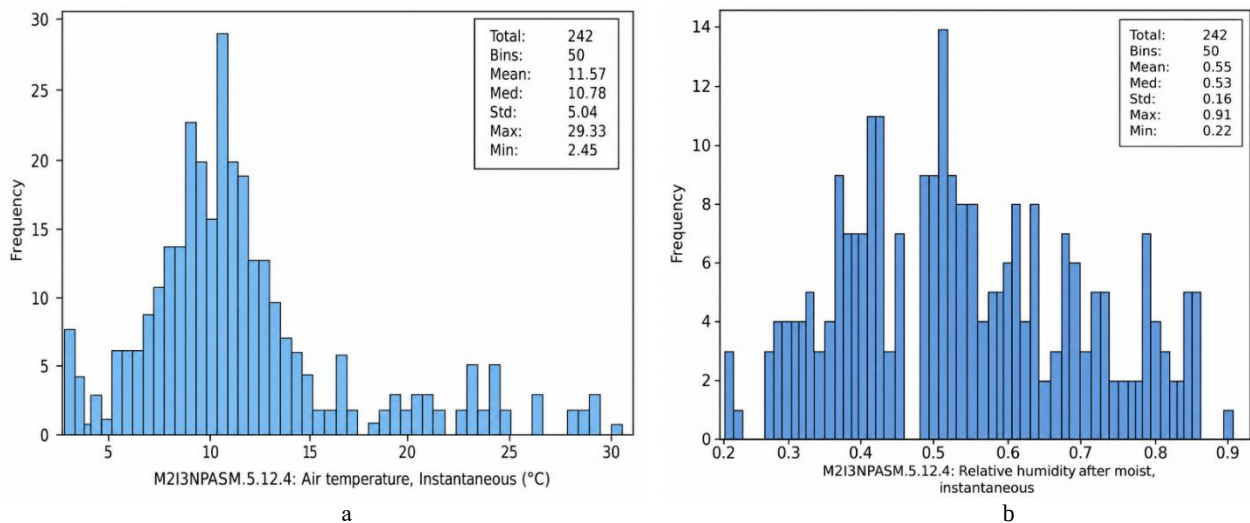


Figure 8. Histograms of surface meteorological parameters over the Tirana region: a – surface air temperature; b – relative humidity

Temperature distribution is unimodal, with a mean value of approximately 11 °C, reflecting stable surface cooling during night-time (Figure 8). Relative humidity exhibits a broad distribution (20–90%), confirming strong variability between daytime and night-time moisture regimes (Figure 9).

The bimodal ozone distribution supports the existence of two distinct atmospheric regimes: one driven by daytime photochemical production and another by night-time dynamical transport processes, reflecting contrasting atmospheric conditions.

Negative coefficients between ozone and temperature over the Tirana region dominate, mostly ($R \approx -0.4$), revealing spatial variability in ozone–temperature coupling influenced by topography and local circulation. Meanwhile, stronger negative correlations between ozone and relative humidity over Tirana indicate that high moisture content enhances ozone depletion through oxidation pathways involving OH radicals.

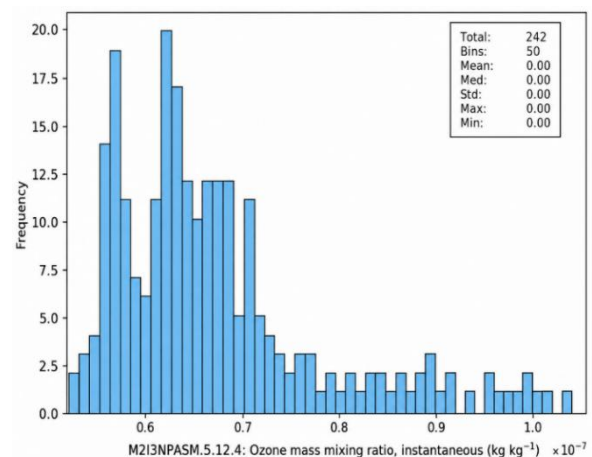


Figure 9. Histogram of ozone mass mixing ratio Tirana region

The strong negative correlation between ozone and temperature ($R = -0.82$) is further elucidated by the spatial and temporal patterns shown in the Giovanni analyses. Histogram results (Figures 8–9) indicate that night-time temperatures display a unimodal distribution centered around 284.69 K, while ozone concentration exhibits broader, multi-peaked distributions. Relative humidity also displays strong variation between daytime (mean 53.93%) and night-time (mean 78.53%) conditions. These distribution patterns, together with the time-series data, suggest that the ozone–temperature inverse correlation is not merely a statistical feature but reflects fundamental atmospheric processes. The bimodal ozone distribution particularly supports the existence of two distinct atmospheric regimes: one governed by photochemical production and another by dynamical transport mechanisms.

Multivariate analysis of surface-monitoring data

The multivariate analysis of the Tirana surface-monitoring dataset revealed three distinct and recurrent patterns of pollutant co-variation. PCA identified the dominant covariance structure, NMF resolved three characteristic non-negative pollutant patterns, and k-means clustering separated the

observations into three atmospheric regimes with distinct pollutant signatures.

Descriptive characteristics

The final analytical dataset comprised 16,632 hourly observations of 11 pollutant variables. Descriptive statistics are presented in Table 1.

PM₁, PM_{2.5} and PM₁₀ exhibited substantial variability across the monitoring record, indicating pronounced temporal changes in particulate pollution. O₃ also showed substantial variability, consistent with the presence of contrasting atmospheric conditions over the monitoring period.

The relationship between the fine-particle fractions was weak, with a correlation coefficient of $r = 0.035$ between PM₁ and PM_{2.5}. Thus, the two fractions showed limited linear co-variation within the monitoring dataset. This result alone does not establish differences in emission sources or formation mechanisms.

Cluster-specific mean concentrations are presented in Table 2 and demonstrate marked differences in the pollutant composition of the three identified regimes.

Table 1. Descriptive statistics of the 11 pollutant variables in the hourly surface-monitoring dataset, including mean, median, standard deviation (SD), minimum, and maximum values. Particulate matter (PM₁, PM_{2.5}, PM₁₀) and gaseous pollutants (NO_x, NO, NO₂, O₃, and SO₂) are reported in $\mu\text{g}\cdot\text{m}^{-3}$, CO in $\text{mg}\cdot\text{m}^{-3}$, and VOCs (ethylbenzene and o-xylene) in ppb

Species	Mean	Median	SD	Min	Max
PM ₁	22.6	17.485	18.0132	0.91	229.21
PM _{2.5}	31.4	26.99	15.922	8.37	102.98
PM ₁₀	50.751	40.775	38.658	1.24	429.71
NO _x	66.8	33.45	141.198	0	1057
NO	42.2	11.275	123.67	0	890.73
NO ₂	25	21.1	21.144	0	174.75
O ₃	18.1	15.88	12.802	0	83.22
CO	0.8	0.74	0.419	0	4.57
SO ₂	3.8	2.07	3.671	0	33.64
Ethylbenzene	0.2	0	0.718	0	17.46
o-Xylene	0.03	0	0.345	0	15.99

Table 2. Cluster centroids/cluster means ($k = 3$). Mean concentrations of the species for each cluster. The centroid table summarizes the three atmospheric states identified by clustering on PCA scores (e.g., traffic peak, photochemical haze, background). PM₁, PM_{2.5}, PM₁₀, NO_x, NO, NO₂, O₃, and SO₂ concentrations are expressed in $\mu\text{g}\cdot\text{m}^{-3}$, CO in $\text{mg}\cdot\text{m}^{-3}$, and VOCs (ethylbenzene and o-xylene) in ppb

Cluster	1	2	3
PM ₁	50.4	16.5	17.8
PM _{2.5}	18.1	18.8	18.1
PM ₁₀	107.4	38.9	41.5
NO _x	88.0	39.6	656.3
NO	53.7	19.9	611.1
NO ₂	34.3	19.5	73.7
O ₃	6.9	22.5	17.5
CO	1.4	0.7	0.7
SO ₂	5.1	4.2	0.6
Ethylbenzene	0.9	0.1	0.0
o-Xylene	0.1	0.0	0.0

Principal component analysis

The first three principal components explained 58.8% of the total variance, with PC1, PC2, and PC3 accounting for 27.2%, 21.9%, and 9.8%, respectively (Figure 10).

PC1 represented the dominant overall pollution gradient, with most pollutant variables showing substantial positive contributions to this component across the analysed dataset. PC2 provided a major contrast between particulate-dominated and gas-phase atmospheric conditions across the analysed environmental observations. PM₁, PM_{2.5}, and PM₁₀ were associated with one direction of PC2, whereas O₃ and combustion-related gaseous pollutants, including NO_x, NO₂, and CO, contributed to the opposing direction.

The PC1–PC2 biplot showed clear organization of observations into distinct regions of multivariate space (Figure 11). Traffic/combustion observations were associated with the NO_x-dominated direction, coarse-particle observations with the particulate loading structure, and photochemical/secondary observations with the O₃-related direction. The PCA structure therefore provided a consistent basis for distinguishing the major atmospheric pollution regimes identified by the subsequent clustering analysis.

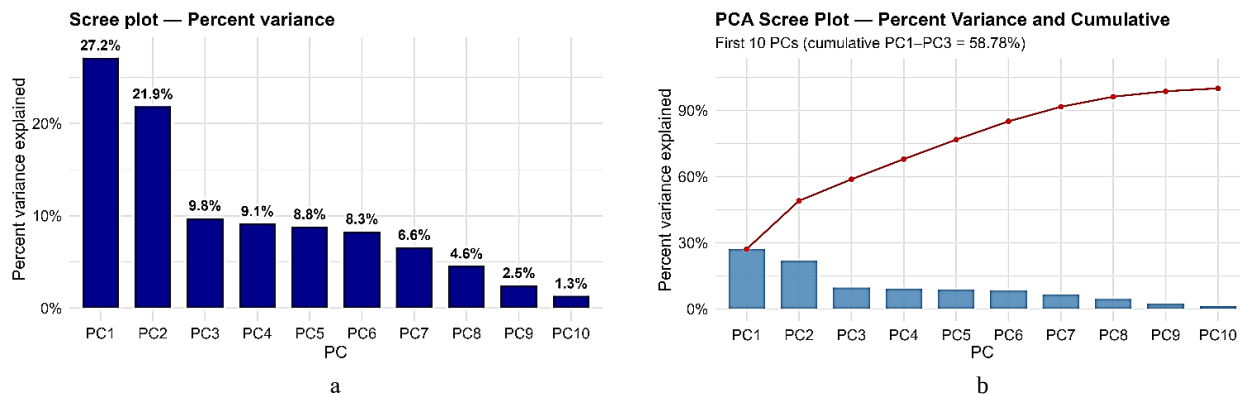


Figure 10. Principal component analysis (PCA) variance explained:

a – scree plot showing the explained variance ratio for PC1–PC11; b – contribution of each principal component to the total variance

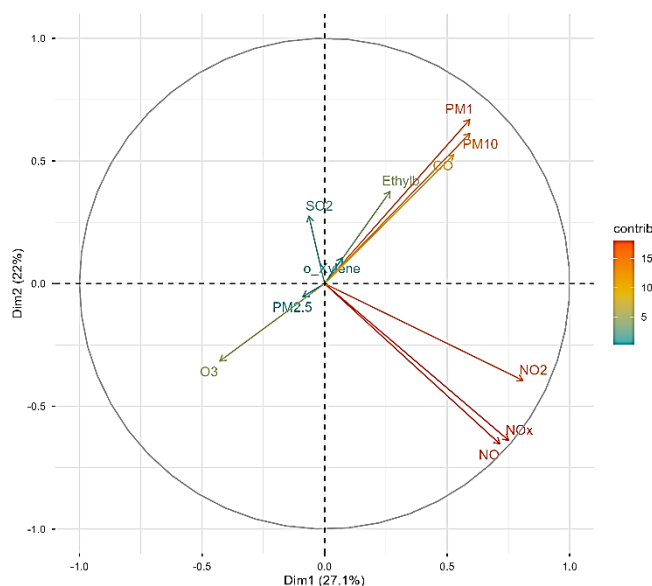


Figure 11. PCA biplot (PC1 vs PC2) with species loadings. Scatter of samples in PC1–PC2 space with superimposed loading vectors for species (biplot). Colour points by k-means cluster to visualize how clusters separate in PC space and how species drive that separation

NMF-derived pollutant patterns

NMF identified a three-factor solution with distinct pollutant signatures (Figure 12).

Factor 1 — Photochemical/secondary pattern.

Factor 1 was characterized by substantial contributions from O_3 , NO_x , $PM_{2.5}$ and NO_2 . The co-occurrence of these species defined an oxidant-rich pollutant pattern associated with chemically processed atmospheric conditions.

Factor 2 — Coarse-particle/resuspension pattern.

Factor 2 was dominated by PM_{10} and PM_1 , with comparatively smaller contributions from gaseous pollutants. This factor represented a particulate-enriched atmospheric state consistent with coarse-particle enhancement.

Factor 3 — Fresh combustion/traffic pattern.

Factor 3 was dominated by NO_x and NO , with additional contributions from CO , ethylbenzene, and *o*-xylene. The resulting pollutant combination was characteristic of a fresh combustion/traffic-related pattern.

The three NMF factors were consistent with the major pollutant groupings identified by PCA and clustering. The factors are interpreted as exploratory atmospheric patterns rather than quantitative emission-source profiles.

The multi-start stability analysis demonstrated that the dominant pollutant associations remained highly consistent across independent NMF initializations. The mean matched-factor similarity was 0.998 ± 0.005 , with a minimum of 0.962 and a maximum of 1.000. The normalized reconstruction error was 0.0862 ± 0.0003 , with a coefficient of variation of only 0.40%.

The three-factor NMF solution identified distinct pollutant regimes based on the normalized factor profiles (Figure 12). Factor 1 represents a photochemical/secondary regime dominated by O_3 , $PM_{2.5}$, and SO_2 , consistent with aged and oxidized air masses. Factor 2 corresponds to a coarse-particle/resuspension regime characterized by high PM contributions, consistent with mechanically generated or transported particles. Factor 3 represents a fresh combustion/traffic regime dominated by NO and NO_x , with minor contributions from CO and traffic-related VOCs, including ethylbenzene and *o*-xylene.

K-means clustering and atmospheric regimes

The clustering analysis identified three distinct atmospheric regimes (Figure 13). The WCSS decreased from 10,670 at $k = 2$ to 7,920 at $k = 3$, with progressively smaller reductions at $k = 4$ –6 (6,190, 5,480, and 4,840). The corresponding silhouette

scores were 0.632, 0.512, 0.335, 0.348, and 0.363. Together, these diagnostics supported retention of the three-cluster

solution. The three clusters exhibited distinct pollutant compositions (Table 2).

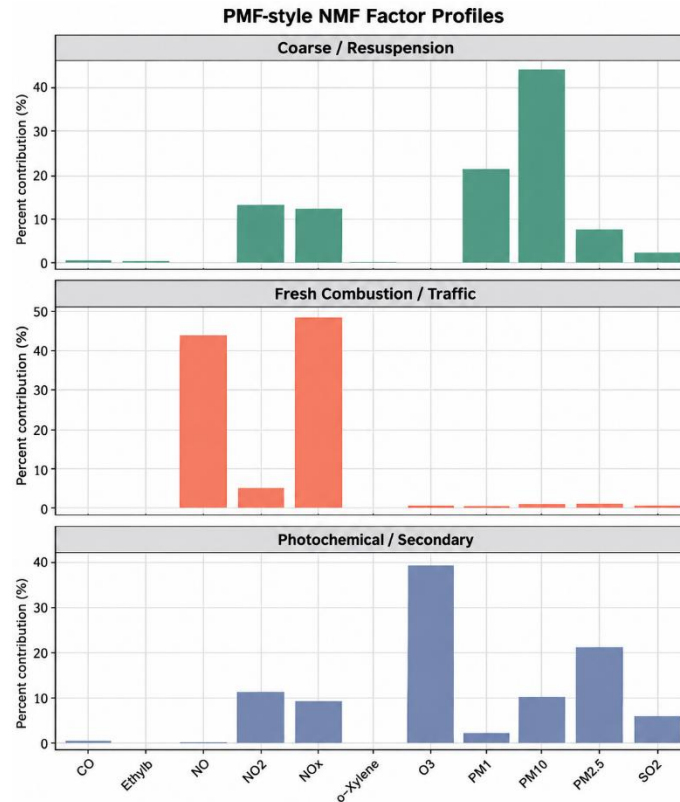


Figure 12. NMF pollutant regimes identified in the surface air-quality dataset

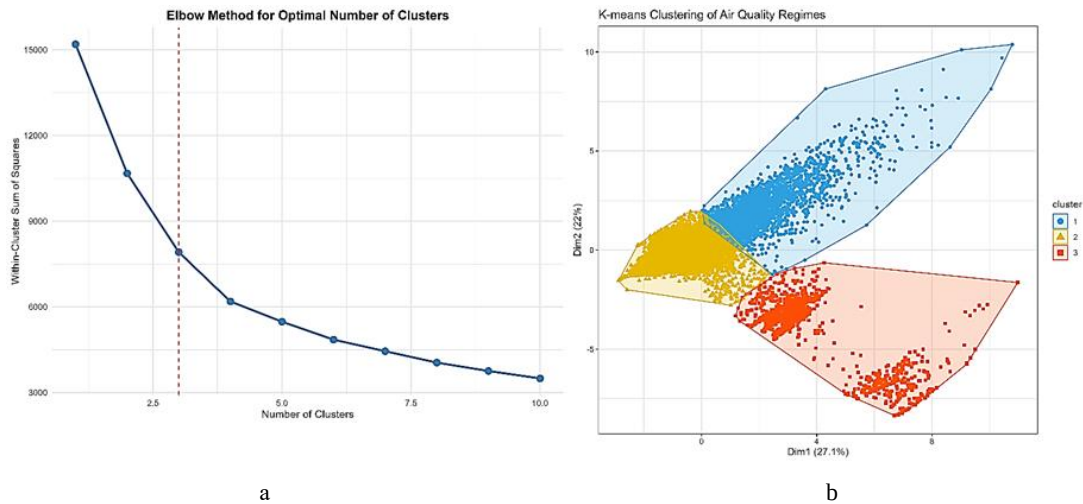


Figure 13. K-means clustering of atmospheric regimes based on principal component analysis (PCA): a – optimization of the number of clusters using the Elbow method; b – k-means clustering results shown as a PC1–PC2 scatter plot, with observations coloured by cluster

The traffic/combustion regime was characterized by elevated nitrogen oxides, particularly NO_x and NO, together with enhanced NO₂. The corresponding mean concentrations were NO_x = 656.3 $\mu\text{g}\cdot\text{m}^{-3}$, NO = 611.1 $\mu\text{g}\cdot\text{m}^{-3}$, and NO₂ = 73.7 $\mu\text{g}\cdot\text{m}^{-3}$. The coarse-PM regime was characterized by enhanced particulate concentrations, with mean PM₁₀ of approximately 107 $\mu\text{g}\cdot\text{m}^{-3}$ and PM₁ of approximately 50 $\mu\text{g}\cdot\text{m}^{-3}$. The photochemical/secondary regime was characterized by enhanced O₃ together with elevated PM_{2.5}, with a mean O₃ concentration of approximately 22.5 $\mu\text{g}\cdot\text{m}^{-3}$.

The convergence of the PCA, NMF, and clustering results provides consistent statistical evidence for three recurrent

atmospheric states: a fresh traffic/combustion state, a coarse-particle-dominated state, and a photochemical/secondary state.

The optimal number of clusters was determined using the Elbow method, which indicated three main clusters (Figure 13a). K-means clustering with $k = 3$ separated the observations into three distinct groups in the PC1–PC2 space (Figure 13b), representing the main atmospheric states identified in the dataset.

Sensitivity analysis

Repeating the multivariate analyses after excluding residual missing observations rather than applying median imputation

produced materially unchanged pollutant associations, factor structures, and three-regime clustering structure. The consistency of the results across the two missing-data treatments indicates that the principal multivariate patterns were not dependent on the selected imputation procedure.

DISCUSSION

Chemical mechanism validation

The empirical correlations observed in the Giovanni data provide quantitative support for the chemical mechanisms proposed. Specifically, negative ozone–temperature relationships support the mechanism where:

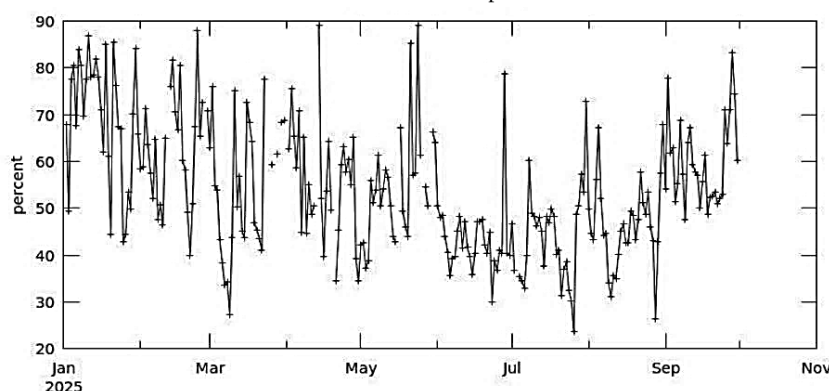
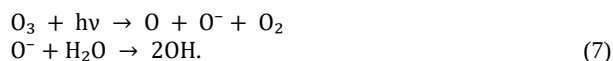


Figure 14. Surface Relative Humidity over Albania, January–September 2025

The observed day–night patterns in the correlation reflect the transition between daytime photochemical processes and night-time atmospheric chemistry:

- during the day, photolytic processes contribute to ozone formation, while OH radicals participate in ozone loss through oxidation reactions;
- during the night, ozone chemistry is influenced by the NO_3 – N_2O_5 equilibrium and the absence of sunlight-driven photolysis, which substantially alters the dominant reaction pathways affecting the ozone budget.

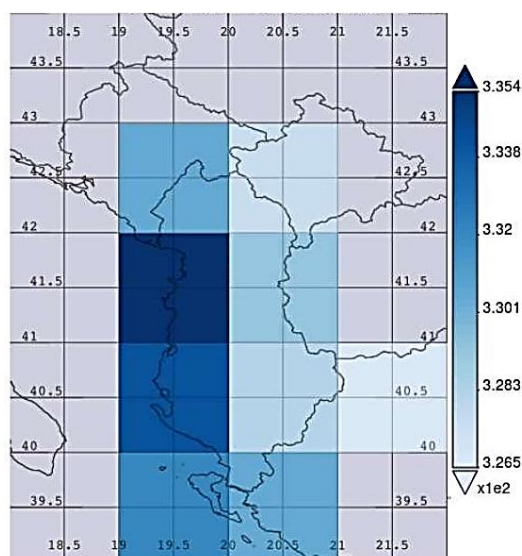


Figure 15. Spatial distribution of total column ozone (DU) over Albania for January–September 2025 based on AIRS data, representing a subset of the broader atmospheric analysis period (January 2024–December 2025)

For visualization, Figures 14 and 15 present a January–September 2025 subset of the broader ERA5/GIOVANNI analysis period (January 2024–December 2025). Figure 14 presents the temporal variation in surface relative humidity (RH) over Albania from January to September 2025. The time series illustrates the seasonal variability of atmospheric moisture conditions during the selected period. This interval represents a subset of the broader ERA5/GIOVANNI analysis period covering January 2024–December 2025 and was selected to provide a clear visualization of the seasonal relationship between T and RH.

An increase in temperature increases the atmosphere's water vapor holding capacity, consistent with the Clausius–Clapeyron relationship (Trenberth et al., 2005), and may consequently affect the production of OH radicals and the atmospheric ozone budget.

The data presented in Figure 14 clearly indicate seasonal variations in relative humidity (RH) from January to September 2025. RH values are generally lower during spring and summer and higher during winter and early autumn. The negative correlation between temperature and RH is consistent with the thermodynamic relationship between temperature and the atmosphere's water vapor holding capacity described by the Clausius–Clapeyron relationship (Trenberth et al., 2005). Thus, as atmospheric temperature increases, the saturation water vapor pressure also increases, allowing the atmosphere to hold a greater amount of water vapor.

Higher water vapor availability may facilitate OH radical production through the reaction of electronically excited oxygen atoms generated by ozone photolysis with water vapor (Seinfeld & Pandis, 2006). Consequently, the observed temporal variations are consistent with a potential link between temperature-dependent changes in RH and the atmospheric processes governing ozone formation and loss during the day, together with night-time pathways involving NO_3 – N_2O_5 chemistry.

Spatial correlation patterns indicate regional variations in these relationships, potentially reflecting differences in NO_x availability, VOC emissions, or meteorological conditions across the study domain.

Figure 15 illustrates substantial spatial variability in total column ozone across Albania. Higher ozone levels are generally observed in coastal and low-elevation areas, where stronger solar radiation and increased anthropogenic emissions of ozone precursors may enhance photochemical ozone production. In contrast, elevated mountainous areas generally exhibit lower ozone levels, which may be associated with differences in atmospheric transport, boundary-layer dynamics, and photochemical activity compared with lower-elevation regions. These spatial differences indicate considerable

regional variability in ozone distribution across the study area, reflecting the combined influence of photochemical processes, NO_x availability, atmospheric transport, topography, and meteorological conditions, consistent with the findings reported by Fiore et al. (2012) and Zanis et al. (2014).

Figure 16 provides a conceptual illustration of the linkage between Positive Matrix Factorization (PMF) and Remote Sensing (RS) in air-quality assessment. Satellite observations capture the column-integrated $\text{PM}_{2.5}$ burden, including both primary particles and secondary aerosols formed through chemical transformation and atmospheric transport. Gaseous precursors such as SO_2 , NO_x , NH_3 , and VOCs undergo

oxidation and subsequent gas-to-particle conversion, becoming incorporated into $\text{PM}_{2.5}$ during atmospheric processing and redistribution. In contrast, PMF relies on surface-based measurements collected by a low-volume sampler (LVS), which provides chemically resolved $\text{PM}_{2.5}$ data used to infer source contributions at ground level. While PMF attributes sources based on particle composition, it cannot resolve the atmospheric origin or transformation history of the gaseous precursors that contributed to the formation of those particles. Conversely, RS does not apportion sources but reflects the spatial distribution of the integrated atmospheric pollutant burden, allowing inference of regional transport patterns when combined with meteorological information.

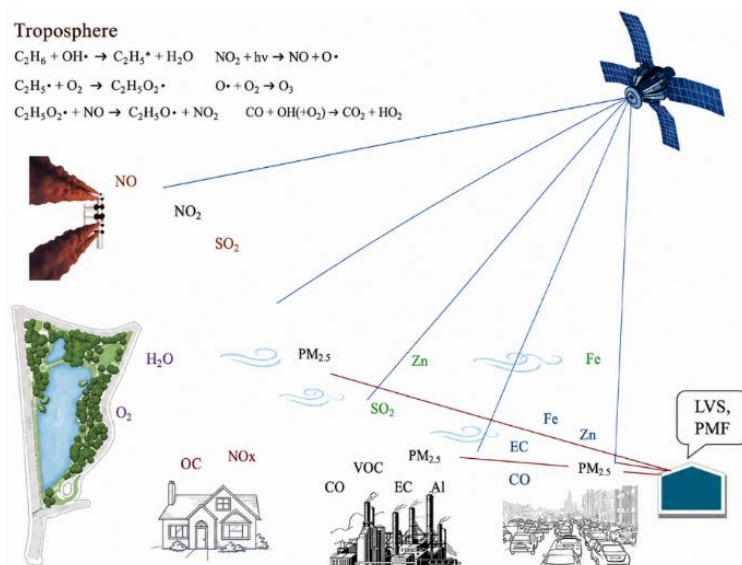


Figure 16. Conceptual framework integration PMF and remote sensing for pollutant regeneration in layers of atmosphere

Consequently, surface $\text{PM}_{2.5}$ represents the integrated outcome of multiple interacting processes, including local emissions, chemical transformation, secondary formation, and horizontal and vertical transport. This framework explains why PMF and RS offer different yet complementary perspectives on atmospheric pollution. Building on this mechanistic understanding, the following regime-based analysis (Section "Regime-based interpretation of PMF and the surface-column linkage") provides empirical evidence for this "Invisible Bridge" by demonstrating how distinct atmospheric states govern the divergence and convergence between surface-based pollutant patterns and column-integrated satellite observations.

Regime-based interpretation of PMF and the surface-column linkage

To further interpret the surface-column linkage within the proposed conceptual framework, regime-based statistical analysis was applied to identify recurrent atmospheric states and their associated pollutant characteristics.

Overall, the combined PCA, NMF, clustering, and correlation analyses reveal a coherent structure of atmospheric regimes characterized by fresh combustion emissions, coarse particulate episodes, and regionally aged photochemical conditions. These results demonstrate that $\text{PM}_{2.5}$ variability cannot be attributed solely to local emission sources but reflects the integrated effects of chemical transformation, secondary formation, and atmospheric transport. The clear separation of regimes, together with the limited coupling between fine-particle fractions, supports a process-based interpretation in which primary emissions undergo oxidation and secondary formation while being redistributed vertically and horizontally.

Although the weak PM_1 – $\text{PM}_{2.5}$ correlation cannot be considered definitive evidence of distinct sources, it may nevertheless provide conceptual insight into the complexity of fine-particle formation. This limited coupling suggests that PM_1 and $\text{PM}_{2.5}$ are not necessarily governed by identical dominant sources and may differ in chemical composition and formation pathways. From a process-based perspective, this observation can be interpreted as qualitative support for the multiple atmospheric regimes identified in this study, in which particulate matter may originate from different mechanisms, including regional transport, chemical transformation, and secondary formation. While not conclusive on its own, this interpretation is consistent with the broader framework developed here, emphasizing that $\text{PM}_{2.5}$ variability reflects the combined influence of emissions, atmospheric processing, and transport.

The exploratory NMF/PCA/clustering analysis provides an empirical characterization of recurrent pollutant regimes within the available surface monitoring dataset. These regimes provide contextual information for interpreting the surface-column relationship, but they do not constitute quantitative PMF source apportionment. This distinction is important because NMF identifies mathematical co-variation patterns without incorporating the measurement uncertainties required for quantitative PMF source apportionment.

The three identified regimes provide a framework for interpreting divergences between surface and column observations, providing an empirical basis for the "Invisible Bridge". The NMF-derived Traffic/Combustion regime is characterized by in-situ measurements but often produces only weak satellite column signals because of its limited vertical

extent and rapid dilution. By contrast, the Photochemical/Secondary regime (Factor 1) represents oxidized, vertically mixed products (secondary aerosols and O_3) that develop downwind and aloft. These species are more spatially extended and better represented in satellite AOD and tropospheric-column measurements, thus forming the bridge between surface emissions and column retrievals. The Coarse/Resuspension regime (Factor 2) occupies an intermediate role: transported or mechanically generated coarse particles can elevate AOD, but the surface–column relationship depends strongly on particle size distribution and plume height.

The NMF factors presented here are exploratory pollutant-regime descriptors and should not be interpreted as quantitative source apportionment or as substitutes for PMF. For quantitative source apportionment, receptor-based PMF/ME-2 should be applied to chemically speciated aerosol (OC/EC, inorganic ions, and trace metals), with bootstrap and dispersion diagnostics reported. Thus, combining NMF, PCA, and clustering with targeted PMF/ME-2 and satellite products provides a practical, evidence-based framework for bridging surface and column observations, improving source interpretation and clarifying when satellite products can be used as proxies for surface exposure.

Regime-based composites of satellite products show that periods dominated by the Photochemical/Secondary factor correspond to enhanced column-integrated signals, including tropospheric NO_2 , whereas Traffic/Combustion episodes are primarily expressed at the surface, with weaker column responses. This contrast provides empirical evidence for the proposed Invisible Bridge: primary emissions detected at the surface undergo chemical transformation and vertical redistribution, while their associated secondary products form an elevated pollutant layer that is preferentially captured by remote-sensing observations.

Implications for air-quality policy and satellite application

The conceptual framework proposed in this study has direct implications for both air-quality policy and the practical use of remote-sensing observations in environmental monitoring. By identifying chemical transformations as the mechanistic link between surface sources and atmospheric column observations, this work provides a scientific basis for interpreting satellite data in the context of regional pollution sources.

From a policy perspective, understanding how pollutants transform and redistribute within the troposphere can help environmental authorities better assess the effectiveness of emission-reduction strategies. Policies traditionally focus on primary emissions measured at the surface, but the framework presented here emphasizes that atmospheric chemistry and vertical and horizontal transport can significantly amplify or reduce pollutant concentrations observed from space. Incorporating these processes into regulatory assessments may support more accurate evaluations of air-quality trends, exceedances, and health-risk forecasts.

For satellite applications, this framework supports the use of RS datasets not only as descriptive tools but also as diagnostic indicators of atmospheric chemical processes. Patterns in O_3 – T – RH correlations, spatial gradients of ozone and humidity, and diurnal cycles can help identify photochemical hotspots, nocturnal chemistry regimes, and regions affected by long-range transport. This strengthens the operational relevance of satellite observations for applications such as early-warning systems, regional emissions analysis, and the design of targeted mitigation strategies.

An additional implication of this framework is that chemical regeneration processes can cause atmospheric pollutant concentrations to diverge significantly from surface-based estimates. As a result, the actual atmospheric burden may exceed what is inferred from emission inventories alone. Integrating chemical transformation processes into regulatory frameworks could therefore lead to more protective and scientifically grounded air-quality thresholds.

Overall, this conceptual linkage between PMF and RS opens new possibilities for integrating satellite data into air-quality management, encouraging the development of hybrid monitoring frameworks that combine chemical knowledge with multi-platform observations.

CONCLUSION

Chemical transformations of pollutants in the atmosphere regenerate primary pollutants, a portion of which can undergo further reactions and transform into $PM_{2.5}$ particles, thereby mediate the link between surface sources and satellite-observation variability. Positive matrix factorization (PMF) provides quantitative attribution of pollutant sources and composition near surface, while remote sensing (RS) offers spatially extensive observation of pollutant distribution across the troposphere. The integration of PMF and RS is thus essential to analyse both horizontal and vertical transport as well as chemical regeneration processes.

Correlation analyses conducted using Giovanni confirm that the O_3 – T relationship is spatially and seasonally variable, driven by both photochemical and thermodynamic mechanisms. Analysis of ERA5 reanalysis data for Albania, with particular attention to the Tirana region, over 2024–2025 revealed clear statistical relationships among ozone, temperature, and relative humidity. These relationships were interpreted as reflecting the combined influence of atmospheric dynamics, vertical transport, and chemical processing rather than as evidence of a long-term climatological trend. The historical surface-monitoring dataset from 2010–2012 was analysed independently to identify recurring pollutant regimes that provide empirical context for the interpretation of the contemporary atmospheric processes. These results indicate that higher temperatures increase the atmosphere's capacity to hold water vapor, which enhances OH radical production and accelerates ozone depletion, thereby explaining the strongly negative O_3 – T relationship. The Clausius-Clapeyron dependency explains the complicated interplay between atmospheric dynamics and chemical transformations.

The surface-monitoring and reanalysis components of this study have different temporal coverage and were intentionally used for complementary purposes. The historical surface-monitoring record (2010–2012) provides an empirical characterization of pollutant co-variation and atmospheric regimes, whereas the ERA5/GIOVANNI dataset (2024–2025) provides a contemporary characterization of O_3 , temperature, and relative-humidity variability. Because these datasets are not temporally coincident, the study does not claim direct temporal validation between surface and reanalysis observations. Instead, the proposed "Invisible Bridge" is evaluated at the process level, using atmospheric chemical transformation, secondary formation, vertical redistribution, and transport as the mechanisms linking surface pollutant regimes with column-integrated atmospheric observations.

This study establishes a conceptual framework, evidencing atmospheric chemical transformations as the critical link between ground-level source apportionment and remote sensing. While limited by its dataset size and its theoretical

nature, the work provides a solid scientific foundation for future developments by outlining a clear direction for quantitative integration of PMF, RS and complementary methods toward more comprehensive and advanced air quality assessment.

Combining PCA, NMF (PMF-style) and regime-based clustering yields a robust workflow that strengthens source apportionment and clarifies when satellite columns reflect surface pollution. Using a case study, we show that NMF identifies physically meaningful regimes that guide targeted PMF/ME-2 runs, PCA summarizes dominant variability for clustering, and regime-based composites with satellite products (tropospheric NO₂) quantify when column signals track surface concentrations. This integrated approach reduces factor mixing, increases the reproducibility of receptor solutions, and provides a practical framework to coordinate ground observations with remote sensing for urban air-quality assessment.

Acknowledgments

The author gratefully acknowledges the support and constructive discussions provided by colleagues during the development of this work. Their insights contributed to the interpretation and presentation of the results.

Author's statements

Contributions

Conceptualization: M.C.; Data curation: M.C., F.M.; Formal Analysis: M.C., F.M.; Investigation: M.C., F.M.; Methodology: M.C., F.M.; Project administration: F.M.; Software: M.C., F.M.; Supervision: F.M., D.P.; Visualization: M.C., F.M.; Writing – original draft: M.C., F.M.; Writing – review & editing: M.C., F.M., D.P.

Declaration of conflicting interest

The authors declare no competing interests.

Financial interests

The authors declare they have no financial interests.

Funding

Not applicable.

Data availability statement

The datasets used in this study are publicly available. Ozone (O₃), temperature (T), and relative humidity (RH) data were obtained from the ERA5 reanalysis dataset through the NASA Giovanni platform (<https://giovanni.gsfc.nasa.gov>). Surface air-quality data were obtained from the automatic monitoring station in Tirana, operated by the National Environmental Agency (AKM) and the Institute of Public Health.

AI Disclosure

Generative AI tools were used solely for language editing and proofreading. The authors take full responsibility for the manuscript's content, analysis, interpretations, and conclusions.

Ethical approval declarations

Not applicable.

Additional information

Publisher's note

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

The initial version of the research manuscript was published as a preprint: <https://doi.org/10.20944/preprints202607.1170.v1>.

REFERENCES

- Albert, D. R., & Davis, H. F. (2010). Collision complex lifetimes in the reaction C₆H₅+ O₂→ C₆H₅O+ O. *The Journal of Physical Chemistry Letters*, 1(7), 1107–1111. <https://doi.org/10.1021/jz100199z>.
- Ball, S. (2014, January). *Atmospheric chemistry at night* (ECG Environmental Brief No. 3). Environmental Chemistry Group, Royal Society of Chemistry.
- Brown, S. S., Dubé, W. P., Tham, Y. J., Zha, Q., Xue, L., Poon, S., ... & Wang, T. (2016). Nighttime chemistry at a high altitude site above Hong Kong. *Journal of Geophysical Research: Atmospheres*, 121(5), 2457–2475. <https://doi.org/10.1002/2015JD024566>.
- Brown, S. S., Ryerson, T. B., Wollny, A. G., Brock, C. A., Peltier, R., Sullivan, A. P., ... & Ravishankara, A. R. (2006). Variability in nocturnal nitrogen oxide processing and its role in regional air quality. *Science*, 311(5757), 67–70. <https://doi.org/10.1126/science.1120120>.
- Byrne, M. P., & O'gorman, P. A. (2016). Understanding decreases in land relative humidity with global warming: Conceptual model and GCM simulations. *Journal of Climate*, 29(24), 9045–9061. <https://doi.org/10.1175/JCLI-D-16-0351.1>.
- Çaushaj, M., Premti, D., Ylli, F., Pinguli, L., Manaj, H., Prifti, T., & Malollari, I. (2024). Air pollution analysis and continuous monitoring of particulate matter (PM_{2.5}) in an urban area. *Journal of Natural Sciences*, 35, 58–71. <https://jns.edu.al/wp-content/uploads/2024/09/4.-M-Caushaj-DH.Premti-et.al-Perf.pdf>.
- Chen, J., Avise, J., Lamb, B., Salathé, E., Mass, C., Guenther, A., ... & Larkin, N. (2009). The effects of global changes upon regional ozone pollution in the United States. *Atmospheric Chemistry and Physics*, 9(4), 1125–1141. <https://doi.org/10.5194/acp-9-1125-2009>.
- Cristofanelli, P., & Bonasoni, P. (2009). Background ozone in the southern Europe and Mediterranean area: influence of the transport processes. *Environmental Pollution*, 157(5), 1399–1406. <https://doi.org/10.1016/j.envpol.2008.09.017>.
- Dewan, S., & Lakhani, A. (2022). Tropospheric ozone and its natural precursors impacted by climatic changes in emission and dynamics. *Frontiers in Environmental Science*, 10, 1007942. <https://doi.org/10.3389/fenvs.2022.1007942>.
- Diao, M., Holloway, T., Choi, S., O'Neill, S. M., Al-Hamdan, M. Z., Van Donkelaar, A., ... & Vaidyanathan, A. (2019). Methods, availability, and applications of PM_{2.5} exposure estimates derived from ground measurements, satellite, and atmospheric models. *Journal of the Air & Waste Management Association*, 69(12), 1391–1414. <https://doi.org/10.1080/10962247.2019.1668498>.
- Doherty, R. M., Wild, O., Shindell, D. T., Zeng, G., MacKenzie, I. A., Collins, W. J., ... & Keating, T. J. (2013). Impacts of climate change on surface ozone and intercontinental ozone pollution: A multi-model study. *Journal of Geophysical Research: Atmospheres*, 118(9), 3744–3763. <https://doi.org/10.1002/jgrd.50266>.
- European Committee for Standardization. (2014). EN 12341:2014: *Ambient air—Standard gravimetric measurement method for the determination of the PM₁₀ or PM_{2.5} mass concentration of suspended particulate matter*.
- Finlayson-Pitts, B. J., & Pitts, J. N., Jr. (2000). *Chemistry of the upper and lower atmosphere: Theory, experiments, and applications*. Academic Press.
- Fiore, A. M., Naik, V., Spracklen, D. V., Steiner, A., Unger, N., Prather, M., ... & Zeng, G. (2012). Global air quality and climate. *Chemical Society Reviews*, 41(19), 6663–6683. <https://doi.org/10.1039/c2cs35095e>.
- Fischer, E. V., Jacob, D. J., Yantosca, R. M., Sulprizio, M. P., Millet, D. B., Mao, J., ... & Pandey Deolal, S. (2014). Atmospheric peroxyacetyl nitrate (PAN): a global budget and source attribution. *Atmospheric Chemistry and Physics*, 14(5), 2679–2698. <https://doi.org/10.5194/acp-14-2679-2014>.

- Gaffney, J. S., & Marley, N. A. (2003). Atmospheric chemistry and air pollution. *The Scientific World Journal*, 3(1), 199–234. <https://doi.org/10.1100/tsw.2003.18>.
- Gaudel, A., Cooper, O. R., Ancellet, G., Barret, B., Boynard, A., Burrows, J. P., ... & Ziemke, J. (2018). Tropospheric Ozone Assessment Report: Present-day distribution and trends of tropospheric ozone relevant to climate and global atmospheric chemistry model evaluation. *Elementa: Science of the Anthropocene*, 6, 39. <https://doi.org/10.1525/elementa.291>.
- Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., ... & Wildt, J. (2009). The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmospheric Chemistry and Physics*, 9(14), 5155–5236. <https://doi.org/10.5194/acp-9-5155-2009>.
- Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., ... Thépaut, J.-N. (2023). *ERA5 reanalysis data* [Data set]. European Centre for Medium-Range Weather Forecasts. <https://giovanni.gsfc.nasa.gov/>.
- Ivatt, P. D., Evans, M. J., & Lewis, A. C. (2022). Suppression of surface ozone by an aerosol-inhibited photochemical ozone regime. *Nature Geoscience*, 15, 536–540. <https://doi.org/10.1038/s41561-022-00972-9>.
- Jacob, D. J. (2000). Heterogeneous chemistry and tropospheric ozone. *Atmospheric Environment*, 34(12–14), 2131–2159. [https://doi.org/10.1016/S1352-2310\(99\)00462-8](https://doi.org/10.1016/S1352-2310(99)00462-8).
- Jacobson, M. Z. (2008). On the causal link between carbon dioxide and air pollution mortality. *Geophysical Research Letters*, 35(3), L03809. <https://doi.org/10.1029/2007GL031101>.
- Jacobson, M. Z., & Streets, D. G. (2009). Influence of future anthropogenic emissions on climate, natural emissions, and air quality. *Journal of Geophysical Research: Atmospheres*, 114(D8). <https://doi.org/10.1029/2008JD011476>.
- Kalabokas, P., Jensen, N. R., Roveri, M., Hjorth, J., Eremenko, M., Cuesta, J., ... & Beekmann, M. (2020). A study of the influence of tropospheric subsidence on spring and summer surface ozone concentrations at the JRC Ispra station in northern Italy. *Atmospheric Chemistry and Physics*, 20(4), 1861–1885. <https://doi.org/10.5194/acp-20-1861-2020>.
- Kawase, H., Nagashima, T., Sudo, K., & Nozawa, T. (2011). Future changes in tropospheric ozone under Representative Concentration Pathways (RCPs). *Geophysical Research Letters*, 38(5), L05801. <https://doi.org/10.1029/2010GL046402>.
- King, M. (2013, July). *Calculating photolysis rates and estimating photolysis lifetimes* (ECG Environmental Brief No. 1). ECG Bulletin, p. 21.
- Kleinman, L. I., Daum, P. H., Lee, Y. N., Nunnermacker, L. J., Springston, S. R., Weinstein-Lloyd, J., & Rudolph, J. (2002). Ozone production efficiency in an urban area. *Journal of Geophysical Research: Atmospheres*, 107(D23), ACH-23. <https://doi.org/10.1029/2002JD002529>.
- K-means clustering algorithm [Software]. *Statistical clustering method*.
- Lamarque, J. F., Kyle, G. P., Meinshausen, M., Riahi, K., Smith, S. J., van Vuuren, D. P., ... & Vitt, F. (2011). Global and regional evolution of short-lived radiatively-active gases and aerosols in the Representative Concentration Pathways. *Climatic Change*, 109(1), 191. <https://doi.org/10.1007/s10584-011-0155-0>.
- Leibensperger, E. M., Mickley, L. J., & Jacob, D. J. (2008). Sensitivity of US air quality to mid-latitude cyclone frequency and implications of 1980–2006 climate change. *Atmospheric Chemistry and Physics*, 8, 7075–7086. <https://doi.org/10.5194/acpd-8-12253-2008>.
- Lin, C., Wang, T., Li, M., Jiang, Y., Wang, Y., & Tan, Y. (2025). Shipping nitrogen oxides recycling via isoprene nitrate amplifies forest ozone. *Communications Earth & Environment*, 6(1), 421. <https://doi.org/10.1038/s43247-025-02413-y>.
- Mandija, F. (2013). Particulate matter concentrations over the region of north-west of Albania. *Journal of Materials and Environmental Science*, 4(6), 915–920.
- Mandija, F. (2015). Characteristics, spatial distributions and main aerosol sources around the area of Shkodra Lake. *International Journal of Continuing Engineering Education and Life Long Learning*, 25(3), 259–273.
- Mandija, F., Caushaj, M., Premti, D., Keckhut, P., Alraddawi, D., Frasher, N., & Vila, F. (2025). Midlatitude spatial and temporal correlations between aerosol properties and weather extremes. *International Journal of Environment and Pollution*, 75(2), 110–132.
- Mandija, F., Chavez Perez, V. M., Nieto, R., Sicard, M., Danylevsky, V., Añel, J. A., & Gimeno, L. (2018). The climatology of dust events over the European continent using data of the BSC DREAM8b model. *Atmospheric Research*, 209, 144–162. <https://doi.org/10.1016/j.atmosres.2018.03.006>.
- Matthias, A. D. (2004). *Monitoring near-surface air quality*. In *Environmental Monitoring and Characterization* (pp. 163–181). Academic Press. <https://doi.org/10.1016/B978-012064477-3/50012-6>.
- Monks, P. S. (2005). Gas-phase radical chemistry in the troposphere. *Chemical Society Reviews*, 34(5), 376–395. <https://doi.org/10.1039/B307982C>.
- Murazaki, K., & Hess, P. (2006). How does climate change contribute to surface ozone change over the United States?. *Journal of Geophysical Research: Atmospheres*, 111(D5). <https://doi.org/10.1029/2005JD005873>.
- NASA Goddard Earth Sciences Data and Information Services Center. (2025). *NASA Giovanni: Interactive visualization and analysis system* [Dataset]. <https://giovanni.gsfc.nasa.gov>.
- National Environmental Agency (AKM). (2024). *Air quality monitoring data for Tirana* [Dataset].
- Nolte, C. G., Gilliland, A. B., Hogrefe, C., & Mickley, L. J. (2008). Linking global to regional models to assess future climate impacts on surface ozone levels in the United States. *Journal of Geophysical Research: Atmospheres*, 113(D14). <https://doi.org/10.1029/2007JD008497>.
- Non-negative matrix factorization (NMF) [Software]. *Multivariate decomposition method*.
- Paatero, P. (2004). *User's guide for positive matrix factorization programs PMF2 and PMF3*, Part 1: Tutorial.
- Paatero, P., & Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics*, 5, 111–126.
- Price, P., Bortorff, B., Jenkins, J., Brune, W. H., & Stevens, P. S. (2025). Re-assessing hydroxyl radical chemistry in the atmosphere: Instrument interferences may explain previous measurement discrepancies. *Communications Earth & Environment*, 6(1), 325. <https://doi.org/10.1038/s43247-025-02308-y>.
- Principal component analysis (PCA) [Software]. *Statistical data analysis method*.
- Pyrgou, A., Hadjinicolaou, P., & Santamouris, M. (2018). Enhanced near-surface ozone under heatwave conditions in a Mediterranean island. *Scientific Reports*, 8(1), 9191. <https://doi.org/10.1038/s41598-018-27590-z>.
- Sanderson, M. (2007). Climate change, methane and ozone. *ECG Bulletin, January. Environmental Chemistry Group, Royal Society of Chemistry*. <https://www.envchemgroup.com/climate-change-methane-and-ozone.html>.
- Saxena, P., & Sonwani, S. (2019). *Primary criteria air pollutants: environmental health effects*. In *Criteria air pollutants and their impact on environmental health* (pp. 49–82). Singapore: Springer Singapore. https://doi.org/10.1007/978-981-13-9992-3_3.
- Seinfeld, J. H., & Pandis, S. N. (2006). *Atmospheric chemistry and physics*. (Note: standard textbook; edition not specified.)

- Song, Y., Tang, X., Xie, S., Zhang, Y., Wei, Y., Zhang, M., Zeng, L., & Lu, S. (2007). Source apportionment of PM_{2.5} in Beijing in 2004. *Journal of Hazardous Materials*, 146(1–2), 124–130. <https://doi.org/10.1016/j.jhazmat.2006.11.058>.
- Stone, D., Whalley, L. K., & Heard, D. E. (2012). Tropospheric OH and HO₂ radicals: Field measurements and model comparisons. *Chemical Society Reviews*, 41(19), 6348–6404. <https://doi.org/10.1039/C2CS35140D>.
- Trenberth, K. E., Fasullo, J., & Smith, L. (2005). Trends and variability in column-integrated atmospheric water vapor: Trenberth et al.: Trends and variability in column-integrated atmosphere water vapor. *Climate Dynamics*, 24(7), 741–758. <https://doi.org/10.1007/s00382-005-0017-4>.
- Viana, M., Padrosa, M., Querol, X., Alastuey, A., & García Dos Santos, S. (2018). *Ozone in southern Europe – Assessment and effectiveness of measures*. ETC/ACM Technical Paper 2017/3. European Topic Centre on Air Pollution and Climate Change Mitigation (ETC/ACM). Retrieved from https://www.eionet.europa.eu/etcs/etc-atni/products/etc-atni-reports/etcacm_tp_2017_3_ozone_s-europe/%40%40download/file/ETCACM_TP_2017_3_Ozone_S-Europe.pdf.
- Wallace, J. M., & Hobbs, P. V. (2006). *Atmospheric science: An introductory survey* (2nd ed., Chapter 3: Atmospheric Thermodynamics, pp. 79–90). Academic Press.
- Wu, C., Liu, L., Wang, G., Zhang, S., Li, G., Lv, S., Li, J., Wang, F., Meng, J., & Zeng, Y. (2021). Important contribution of N₂O₅ hydrolysis to the daytime nitrate in Xi'an, China during haze periods: Isotopic analysis and WRF-Chem model simulation. *Environmental Pollution*, 288, 117712. <https://doi.org/10.1016/j.envpol.2021.117712>.
- Xiang, W., Wang, W., Du, L., Zhao, B., Liu, X., Zhang, X., ... & Ge, M. (2023). Toxicological effects of secondary air pollutants. *Chemical Research in Chinese Universities*, 39(3), 326–341. <https://doi.org/10.1007/s40242-023-3050-0>.
- Yuan, B., Liggio, J., Wentzell, J., Li, S. M., Stark, H., Roberts, J. M., ... & de Gouw, J. A. (2016). Secondary formation of nitrated phenols: insights from observations during the Uintah Basin Winter Ozone Study (UBWOS) 2014. *Atmospheric Chemistry and Physics*, 16(4), 2139–2153. <https://doi.org/10.5194/acp-16-2139-2016>.
- Zanis, P., Hadjinicolaou, P., Pozzer, A., Tyrllis, E., Dafka, S., Mihalopoulos, N., & Lelieveld, J. (2014). Summertime free-tropospheric ozone pool over the eastern Mediterranean/Middle East. *Atmospheric Chemistry and Physics*, 14, 115–132. <https://doi.org/10.5194/acp-14-115-2014>.
- Zeng, G., Morgenstern, O., Braesicke, P., & Pyle, J. A. (2010). Impact of stratospheric ozone recovery on tropospheric ozone and its budget. *Geophysical Research Letters*, 37(9), L09805. <https://doi.org/10.1029/2010GL042812>.
- Zhang, R., Wang, G., Guo, S., Zamora, M. L., Ying, Q., Lin, Y., ... & Wang, Y. (2015). Formation of urban fine particulate matter. *Chemical reviews*, 115(10), 3803–3855. <https://doi.org/10.1021/acs.chemrev.5b00067>.